

虚拟筛选

Virtual Screening, VS

主讲人：邹毅
Email: zouyi@cpu.edu.cn



1

- 虚拟筛选 (VS) 在药物发现中的意义
- 虚拟筛选的基本流程
- 案例



2

2

第一节 Part One

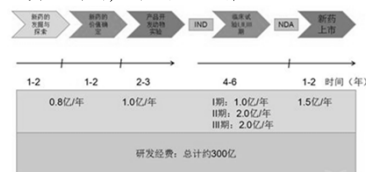
虚拟筛选在药物发现中的意义 The significance of virtual screening in drug discovery

3

3

引言

据统计数据显示, 成功研制一种新药需要约10—12年的时间, 耗资10亿—15亿美元。计算机辅助药物设计技术能通过干涉新药发现流程, 大量降低研发成本和时间, 增加社会经济效益。



资料来源: 药研机构和药企竞争力评价数据库系统研究组 撰制, 毛军文, 主编: 中国医药 2006(07)



4

4

筛选途径的药物发现方法

- **实验筛选:** 是使用生物物理、生物化学等方法鉴别小分子和蛋白质的相互作用, 从而筛选出对蛋白质功能有调节作用的先导化合物。基于大规模化合物库的实验筛选也称为高通量筛选 (HTS)。
- **虚拟筛选:** 是指在进行生物活性筛选之前, 根据预先设定的条件, 在计算机上对化合物分子进行预筛选, 以识别出最可能与靶标结合的小分子, 从而大大降低实际筛选化合物数目, 同时提高先导化合物的发现效率。



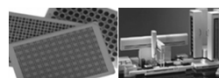
5

5

高通量筛选VS虚拟筛选

高通量筛选:

体外活性测试 (大量测试)
命中率: 0.01-0.001% 假阴性较高
需要先有大量化合物



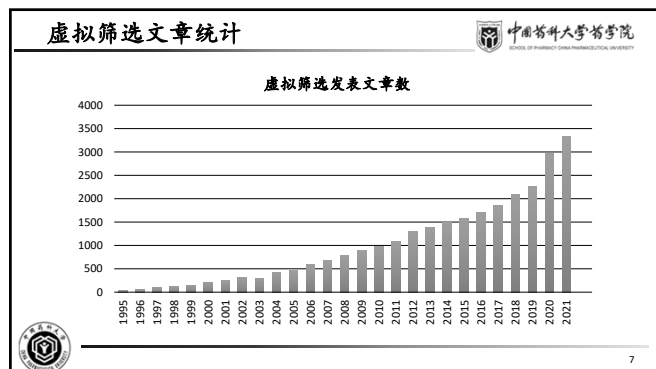
虚拟筛选:

计算机模拟分析
命中率: 2-24%
仅需测试少量化合物



6

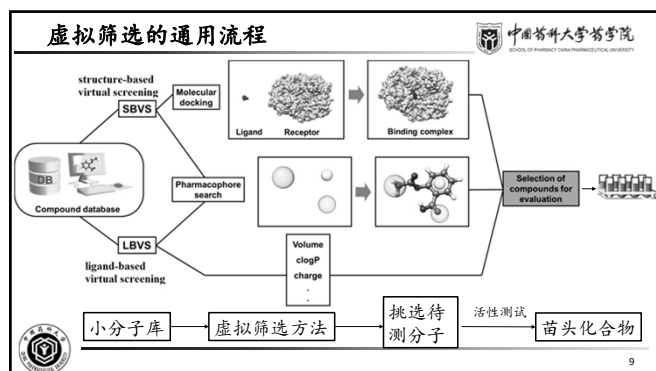
6



7



8



9

虚拟筛选化合物数据库

常用数据库:

ZINC、PubChem、DrugBank、ChEMBL、ChemDB、HMDB、BindingDB、SMPDB。

常用化合物库介绍:

- 1) 筛选化合物库分类: 多样性化合物库、上市药物分子库、已知活性库、靶点化合物库、天然产物库、片段库等等;
- 2) 常用化合物库品牌: Chemdiv、Enamine、Lifechemicals、Specs、Chembridge、Maybridge、Microsource、Vitas-M、Interbioscreen等

10

虚拟筛选化合物数据库

3) 各品牌化合物的优劣势介绍:

- a. ChemDiv: 140万库存分子, 多样性被认为是商业库中最好的, 每年新增20万以上的分子, 每年新增的量就和Specs全部的分子数一样多了, 而且有很多虚拟筛选的针对各靶点的子库, 被大量的高通量筛选的客户选用。
- b. Enamine: 100万以上, 乌克兰的公司, 多样性也不错, 但价格要高于chemdiv;
- c. Lifechemicals: 60多万种分子, 库存量充足, 如部分产品库存不足可很快完成合成及补充工作, 进入中国市场较晚, 价格中等;
- d. Specs: 20多万种分子, 价格最便宜, 最早进入中国市场, 已经被各大研究单位无数次的反复筛选过。化合物更新较慢, 溶解性, 多样性都相对比较差;
- e. Chembridge: 有70万种小分子, 将化合物分成1, 2, 3或4类, 价格都不一样, 是依赖于化合物合成的难易来分的。相对价格较贵

11

化合物数据库的预处理

1、类药性分析

目的: 排除理化参数不佳的化合物, 降低筛选成本。保留具有一定类药性(drug-likeness)特点的化合物

参数	最小值	最大值
logP	-2	5
分子量	200	500
氢键给体	0	5
氢键受体	0	10
摩尔折射率	40	130
旋转键数目	0	10
重原子数目	20	70
极性表面积 (TPSA)	0 Å²	120 Å²
电荷	-2	2

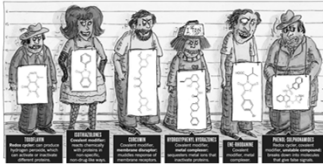
12

化合物数据库的预处理

中国药科大学药学院
SCHOOL OF PHARMACY CHINA PHARMACEUTICAL UNIVERSITY

2、假阳性分子排除

目的：一些化合物易于与生物大分子发生化学反应,在基于受体、酶或细胞检测实验中总是表现为阳性,而实际上为假阳性

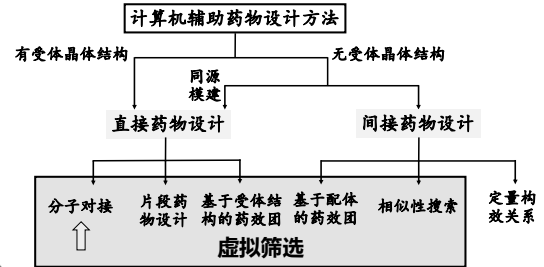


13

13

虚拟筛选的不同方法

中国药科大学药学院
SCHOOL OF PHARMACY CHINA PHARMACEUTICAL UNIVERSITY

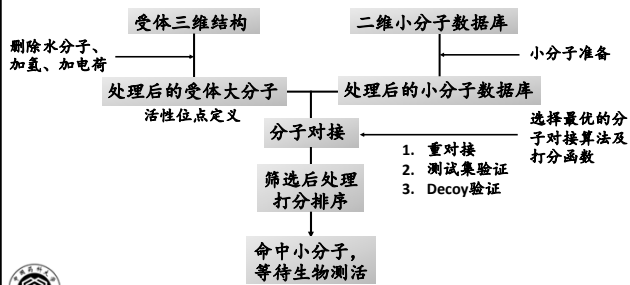


14

14

基于分子对接的虚拟筛选步骤

中国药科大学药学院
SCHOOL OF PHARMACY CHINA PHARMACEUTICAL UNIVERSITY



15

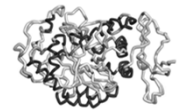
15

基于分子对接的虚拟筛选步骤

中国药科大学药学院
SCHOOL OF PHARMACY CHINA PHARMACEUTICAL UNIVERSITY

蛋白结构的获取

- 经实验得到蛋白晶体结构
- ✓ X-ray晶体衍射
- ✓ NMR核磁共振
- ✓ 冷冻电镜
- 同源建模
- 基于人工智能的蛋白结构预测



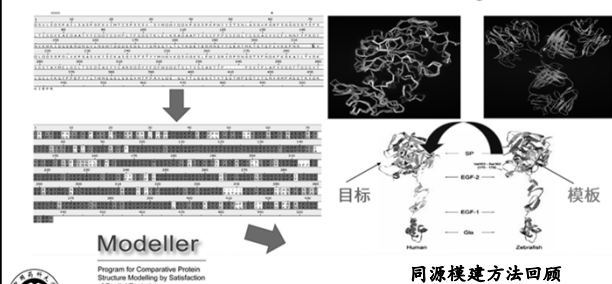
特点：
结果准确，但投入大、周期长、
风险大

16

16

基于分子对接的虚拟筛选步骤

中国药科大学药学院
SCHOOL OF PHARMACY CHINA PHARMACEUTICAL UNIVERSITY



17

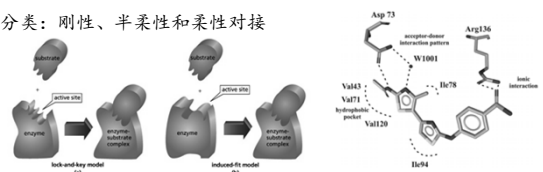
17

基于分子对接的虚拟筛选步骤

中国药科大学药学院
SCHOOL OF PHARMACY CHINA PHARMACEUTICAL UNIVERSITY

分子对接

- 根据几何匹配和能量匹配，寻找两个分子之间的最佳匹配（相互作用）模式。
- 分类：刚性、半柔性性和柔性对接



18

18

基于分子对接的虚拟筛选步骤



分子对接的核心步骤

搜索算法 (搜索配体与受体的可能结合模式)

- 配体的可能构象
- 配体在活性位点的可能取向

打分函数 (对配体和受体的结合模式进行打分寻找最合理的结合)

- 力场打分函数
 - 经验打分函数
 - 知识打分函数
- 多种打分综合评价：一致性打分



19

基于分子对接的虚拟筛选步骤



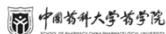
对接方法的选择

项目	类别	软件
打分功能	力场	DOCK
	经验	LUDI、FlexX、Autodock
搜索算法	半经验	GOLD
	随机 (GA)	GOLD、Autodock
	随机 (MC)	ICM、Glide
	随机 (其他)	Tabu
配体柔性	刚性配体	全部软件
	半柔性配体	FlexX、Dock、AutoDock、Glide、Surflex
	全柔性配体	GOLD
	刚性受体	全部软件
受体柔性	结合位部分柔性	GOLD、SLIDE、DOCK、AutoDock、FlexX
	结合部分全柔性	ICM



3

基于分子对接的虚拟筛选步骤



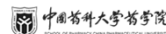
评判对接算法和打分函数精度的标准

- (1) “对接能力”，其被用于评判一个小分子的预测结合构象与实验结合构象间的差异，通常用RMSD值的大小来判断；当两者间的 $RMSD < 2 \text{ \AA}$ 时，一般认为是一个成功的对接；
- (2) “筛选能力”，其测试了一个打分函数对抑制剂和非抑制剂的区分能力；通常使用ROC 曲线来评判；
- (3) “排序能力”，其用于评测打分函数重现小分子结合自由能的能力；通常用预测结合自由能和实验值间的 Pearson 和 Spearman 相关性来判断。

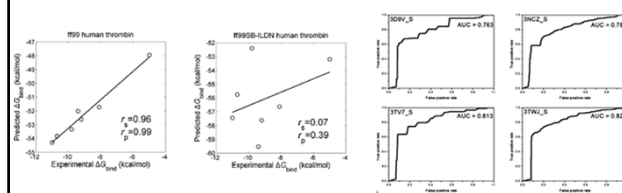


21

基于分子对接的虚拟筛选步骤



评判打分函数精度的标准



相关性分析

ROC相关性分析



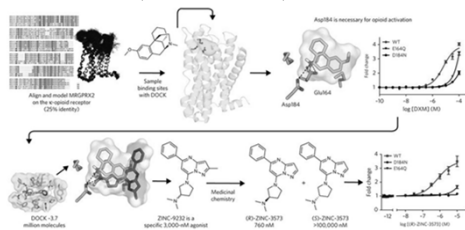
J. Chem. Inf. Model. 2015, Vol. 55, Issue 4, Pages 861-871

3

基于分子对接的虚拟筛选步骤



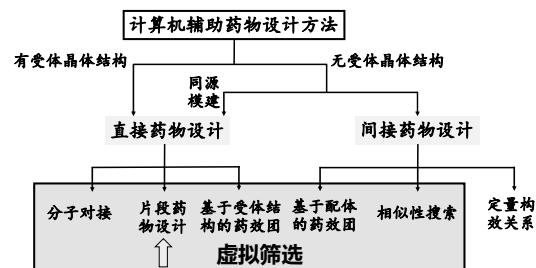
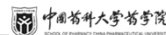
作者通过同源模建构建了MRGPRX2蛋白结构,在此基础上通过分子对接成功找到了一个选择性的激动剂,并得到实验验证,可作为药物开发的先导化合物。



Nat Chem Biol 2017 Vol. 13 Issue 5 Pages 529-536

23

虚拟筛选的不同方法

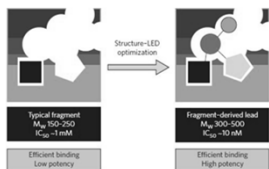


3

基于片段设计的虚拟筛选步骤

中国药科大学药学院
SCHOOL OF PHARMACY CHINA PHARMACEUTICAL UNIVERSITY

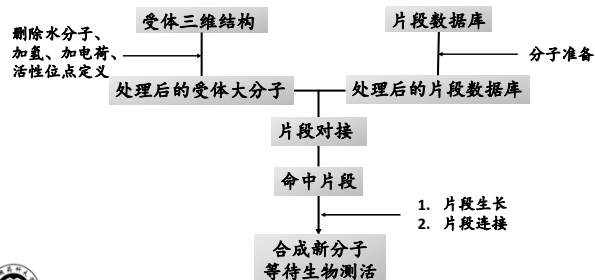
基于片段的药物发现方法的主要内容是设计并建立由片段分子组成的化合物库，并对库中的片段进行生物活性的筛选从而找到苗头片段分子。再利用X射线晶体学、核磁共振等技术对这些片段与靶蛋白的结合模式与结合强度进行分析,根据这些结构信息对片段分子进行结构优化得到先导化合物。



25

25

基于片段设计的虚拟筛选步骤

中国药科大学药学院
SCHOOL OF PHARMACY CHINA PHARMACEUTICAL UNIVERSITY

26

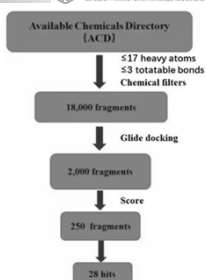
26

基于片段设计的虚拟筛选步骤

中国药科大学药学院
SCHOOL OF PHARMACY CHINA PHARMACEUTICAL UNIVERSITY

研究案例

凝血因子VIIa是一种新的抗血栓形成靶标，与S1口结合的大多数因子VIIa抑制剂包含疏水和阳离子基团，导致不好的膜渗透性和口服吸收。为了发现中性S1靶向抑制剂并改善药代动力学特征，Cheney等研究人员进行了基于片段的虚拟筛选研究。



27

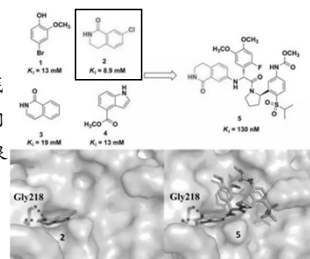
27

基于片段设计的虚拟筛选步骤

中国药科大学药学院
SCHOOL OF PHARMACY CHINA PHARMACEUTICAL UNIVERSITY

研究案例

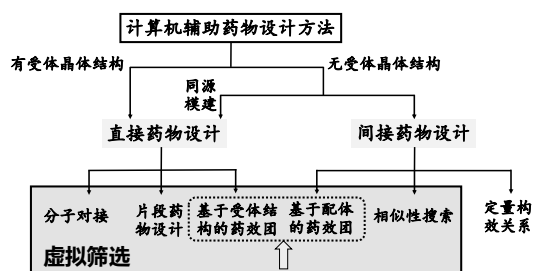
代表性命中片段 ($K_i = 8.9 \text{ mM}$) 与凝血因子VIIa S1口袋结合并与氨基酸残基Gly218形成氢键。基于命中片段内酰胺的结合模式，发现了许多具有良好抑制效能和改善渗透性的类似物。如化合物5是细胞可渗透的，并且具有 130 nM 的 K_i 值。



28

28

虚拟筛选的不同方法

中国药科大学药学院
SCHOOL OF PHARMACY CHINA PHARMACEUTICAL UNIVERSITY

29

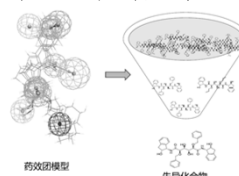
29

基于药效团模型的虚拟筛选步骤

中国药科大学药学院
SCHOOL OF PHARMACY CHINA PHARMACEUTICAL UNIVERSITY

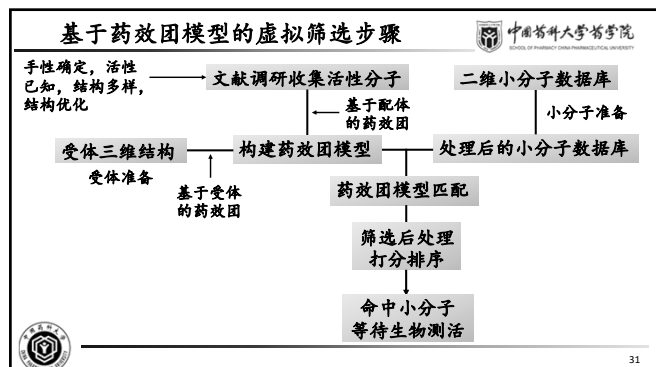
药效团模型

药物活性分子中对活性起着重要作用的“药效特征元素”及其空间排列形式。包括了结合特征、结构和特性约束的信息，以作为数据库检索的提问方式。

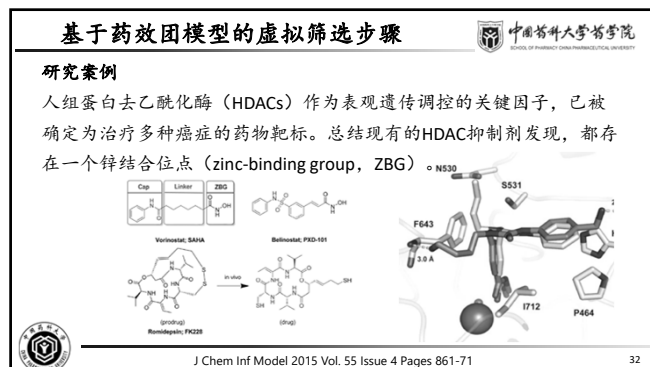


30

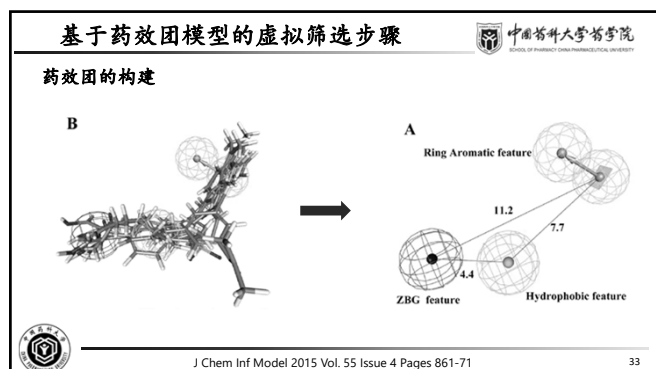
30



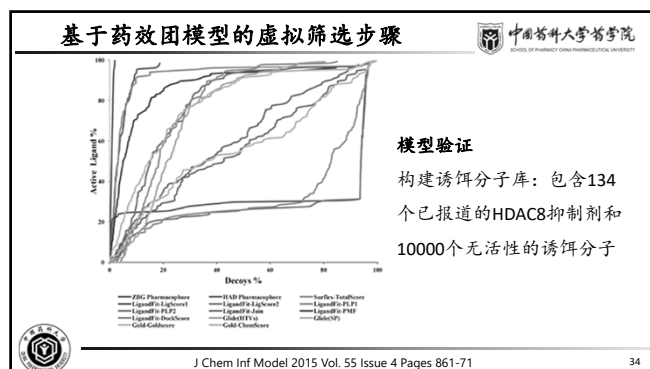
31



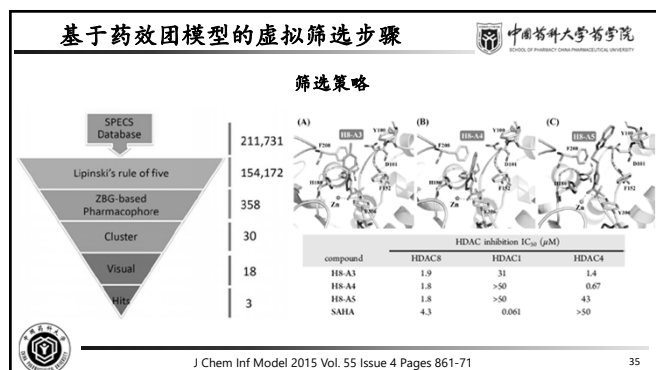
32



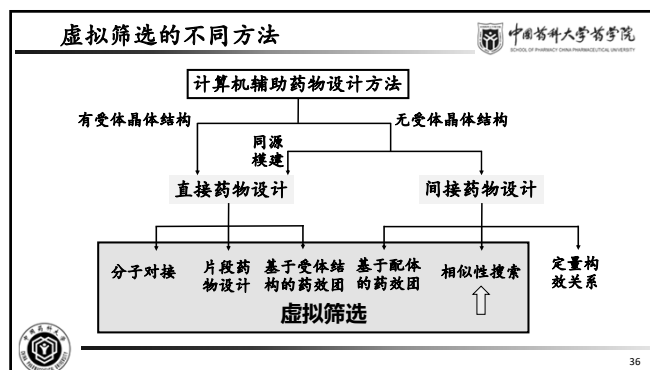
33



34



35



36

基于相似性搜索的虚拟筛选



相似性搜索

主要包括二维/三维指纹相似性搜索和形状相似性搜索。其基本假设都是“相似性原理”，即相似的分子倾向于具有相似的理化性质，并表现出相似的生物活性。

相似性搜索的过程

首先需要将已知的活性化合物和数据库中化合物进行相似性计算打分，得分越高的分子相似性越大；然后根据设定的相似性阈值，将满足相似性要求的分子从数据库中挑选出来。



37

37

基于相似性搜索的虚拟筛选



研究案例1

雌激素受体（ERs）是ER阳性乳腺癌和骨质疏松症的有效靶标。亚型选择性配体的数量的增加有助于ERa和ERb之间的差异的研究。

研究方法：

1. 基于分子相似性使用参考化合物对SPCS数据库进行相似性搜索，相似系数设定为0.45-0.99；所有命中化合物进行基于对接的虚拟筛选，最终获得具有合适结合构象的化合物；
2. 购买纯度合适的命中化合物；
3. 运用不同生物技术对化合物进行活性评估。



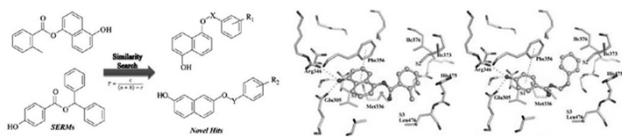
38

38

基于相似性搜索的虚拟筛选



研究案例1



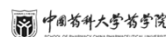
通过相似搜索和分子对接的组合策略的虚拟筛选，发现了21种化合物作为雌激素受体的有效配体，其中两个具有高的抑制率（>70%）。



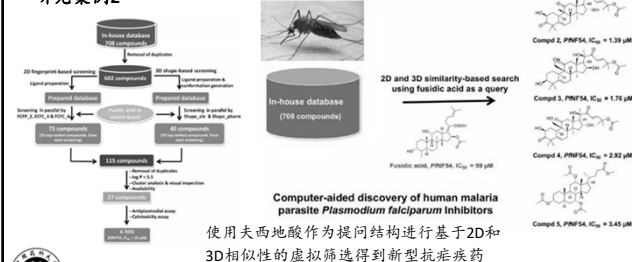
39

39

基于相似性搜索的虚拟筛选



研究案例2



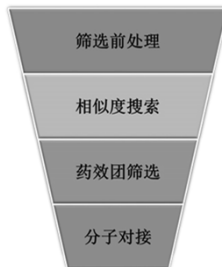
40

40

不同虚拟筛选方法的联合应用



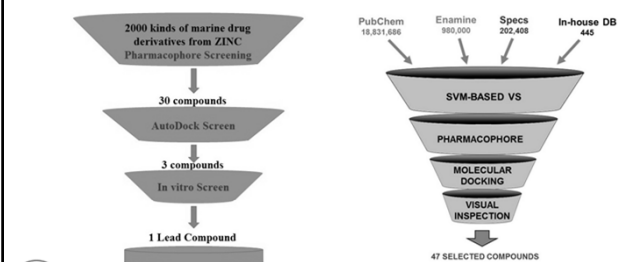
SBDD以锁匙模型为基础，寻找可以与靶标结合的配体分子；而SBDD以相似性原理为基础，寻找与已知抑制剂具有潜在相同作用方式的配体分子。SBDD结果准确，可直接预测结合能，缺点是需受体结构，速度慢；LBDD速度快，无需蛋白结构，但是有时由于缺乏具体结合方式的指导，结果外推时常常误差很大。这两种方法、理论是独立发展，各有其适用条件和利弊的。所以一个虚拟筛选项目通常会整合多种CADD技术。



41

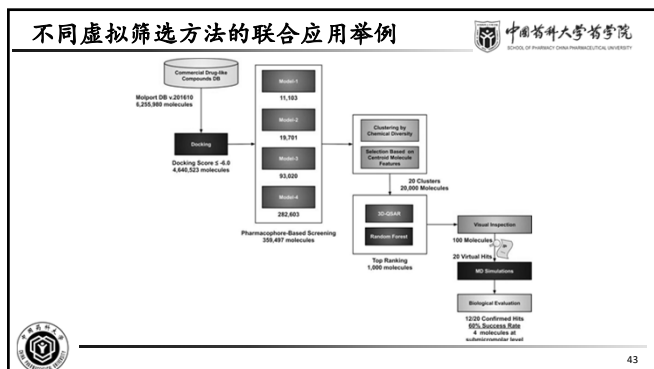
41

不同虚拟筛选方法的联合应用举例



42

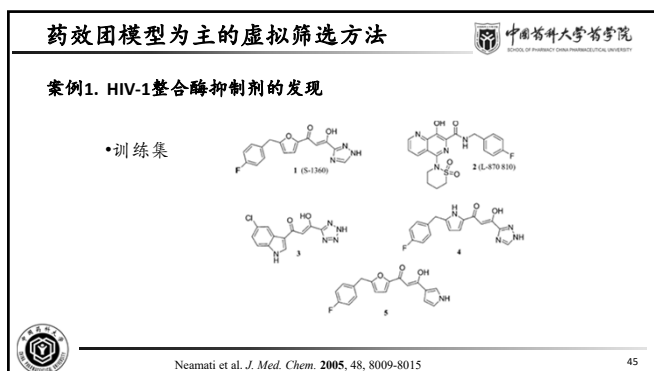
42



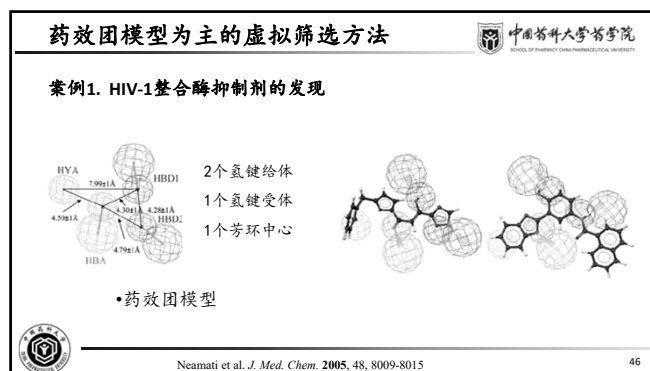
43



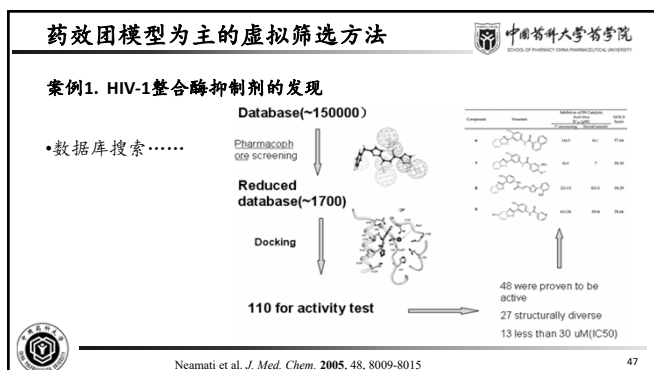
44



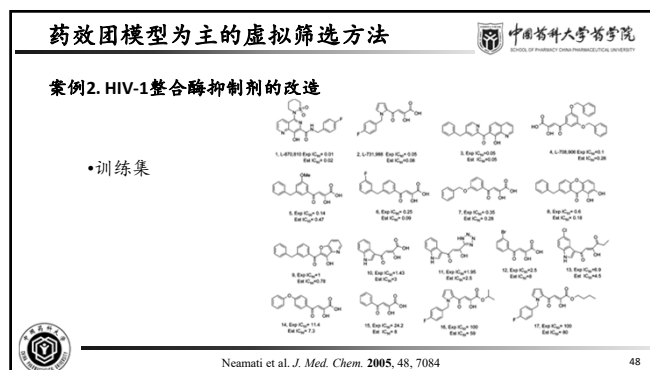
45



46



47

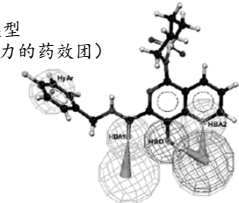


48

药效团模型为主的虚拟筛选方法

案例2. HIV-1整合酶抑制剂的改造

药效团模型
(具有活性预测能力的药效团)



2个氢键受体, 1个氢键给体, 1个芳环中心

Neamati et al. *J. Med. Chem.* 2005, 48, 7084

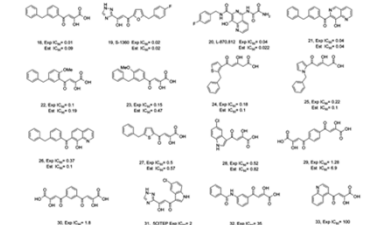
49

药效团模型为主的虚拟筛选方法

案例2. HIV-1整合酶抑制剂的改造

• 活性预测

药效团活性预测结果与最终的实验活性测定结果进行比较, 结果非常一致, 线性相关常数 r 达0.85。



Neamati et al. *J. Med. Chem.* 2005, 48, 7084

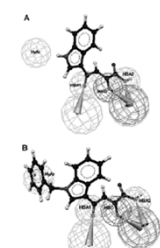
50

药效团模型为主的虚拟筛选方法

案例2. HIV-1整合酶抑制剂的改造

化合物改造

- 16个化合物中, 有一个化合物与药效团模型的匹配方式如图A所示, 该化合物只能与药效团中的三个特征相匹配, 芳香疏水中心没有匹配上。
- 依据这一信息, 研究者在该化合物上添加了一个疏水的芳环侧链, 然后进行了活性测试。
- 发现改造之后的化合物 (如图B) 所示, 与之相比, 活性有了很大的提高。

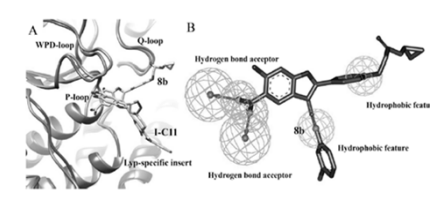


Neamati et al. *J. Med. Chem.* 2005, 48, 7084

51

药效团模型为主的虚拟筛选方法

案例3. 基于CBP模型发现新型Lyp抑制剂



药效团模型构建案例

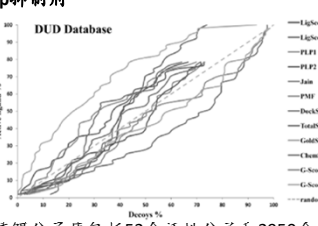
J Med Chem 2014 Vol. 57 Issue 22 Pages 9309-22

52

药效团模型为主的虚拟筛选方法

案例3. 基于CBP模型发现新型Lyp抑制剂

Docking Program	Score Function	EF (DUD Database)				Success Rate ^a (%)
		1%	2.5%	5%	10%	
LigScore1	3.152	1.124	1.124	1.124	1.124	76.9
LigScore2	3.152	0.878	0.878	0.878	0.878	
PLP1	3.152	0.878	0.878	0.878	0.878	
PLP2	3.152	0.878	0.878	0.878	0.878	
PLP3	3.152	0.878	0.878	0.878	0.878	
GoldScore	3.152	0.878	0.878	0.878	0.878	98.1
GoldScore2	3.152	0.878	0.878	0.878	0.878	
GoldScore3	3.152	0.878	0.878	0.878	0.878	
GoldScore4	3.152	0.878	0.878	0.878	0.878	
GoldScore5	3.152	0.878	0.878	0.878	0.878	



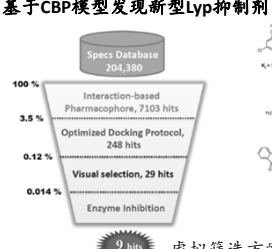
对接及其打分函数评价 (诱饵分子库包括52个活性分子和2950个无活性的诱饵分子)

J Med Chem 2014 Vol. 57 Issue 22 Pages 9309-22

53

药效团模型为主的虚拟筛选方法


案例3. 基于CBP模型发现新型Lyp抑制剂



虚拟筛选方案


J Med Chem 2014 Vol. 57 Issue 22 Pages 9309-22

54



中国药科大学药学院
SCHOOL OF PHARMACY CHINA PHARMACEUTICAL UNIVERSITY

谢谢大家



55