

基于片段的药物设计研究进展

郭靖, 陆小云*

(暨南大学药学院, 广东 广州 510632)

[摘要] 基于片段的药物设计 (FBDD) 方法是一种通过筛选碎片库得到苗头片段, 然后采用基于结构的设计策略对苗头片段进行优化和改造获得先导化合物的研究。其研究过程分为 3 个部分: 1) 设计并构建片段库; 2) 片段库的筛选; 3) 利用基于结构的药物设计策略对苗头片段进行优化设计, 最终得到类药性的先导化合物或候选化合物。通过结合已经上市的 3 个抗肿瘤药物 (vemurafenib、venetoclax 和 erdafitinib) 和目前处于 II 期临床的 KRAS^{G12C} 抑制剂 AMG510 的研究为实例对 FBDD 的最新研究进展进行综述与展望。

[关键词] 药物发现; 片段筛选及优化; 基于结构的药物设计

[中图分类号] R914.4

[文献标志码] A

[文章编号] 1001-5094 (2020) 09-0698-12

Research Progress of Fragment-based Drug Design

GUO Jing, LU Xiaoyun

(School of Pharmacy, Jinan University, Guangzhou 510632, China)

[Abstract] Fragment-based drug design (FBDD) is a strategy that combines the fragments screening library and optimizes the fragments with structure-based drug design to obtain lead compounds. There are three basic steps for FBDD. First, design and establish a fragment library; second, use biophysical methods to screen active fragments comprehensively; finally, optimize the fragment assembly molecules to obtain drug-like lead compounds or candidates using structure-based drug design methods such as fragment linking, growth, and merging. In this article, we reviewed the basic principles and applications of FBDD and detailed the three approved anticancer drugs (vemurafenib, venetoclax and erdafitinib) and currently in clinical phase II KRAS^{G12C} inhibitor AMG510 to illustrate FBDD strategy.

[Key words] drug discovery; fragment screening and optimization; structure-based drug design

基于片段的药物设计 (fragment-based drug design, FBDD) 是一种通过筛选碎片库得到苗头片段, 随后基于结构的设计策略对片段进行优化和改造获得先导化合物的研究方法^[1]。FBDD 方法是为了克服传统高通量筛选的缺陷而逐步发展起来的新技术。高通量筛选由于盲目性大, 命中率低, 成本高, 对于部分药物靶点很难筛选得到理想的化合物。FBDD 方法基于药物靶标的活性位点是由多个子口袋组成, 首先筛选得到各个子口袋的特异性结合片段, 再将与靶蛋白各个子口袋特异结合的片段以合适的连接子连接起来, 组装成为高活性的化合物。FBDD 方法与传统的高通量筛选获得苗头化合物方法相比, 具有特异性高、类药性高、效率高等优势^[2]。

FBDD 作为发现苗头化合物的一种关键技术, 逐渐成为药物研发的主流方法。其研究过程可以分为 3 个步骤: 1) 片段库的设计构建 (片段通常符

合“3 规则”, 即: 相对分子质量 ≤ 300 ; 氢键供体数目 ≤ 3 ; 氢键受体数目 ≤ 3 ; 计算脂水分配系数 $cLogP \leq 3$; 旋转键数目 ≤ 3 ; 极性表面积 $\leq 60 \text{ \AA}^2$)^[3]; 2) 片段筛选 (使用多种生物物理技术方法筛选获得与靶标具有结合能力的苗头片段); 3) 片段优化 (基于靶标结构特征, 通过迭代优化策略将苗头片段优化成为先导化合物) (见图 1)^[4]。至今已有数十个采用基于片段药物设计方法的药物进入临床试验研究, 其中已有 3 个药物 (vemurafenib、venetoclax 和 erdafitinib) 被美国 FDA 批准上市。

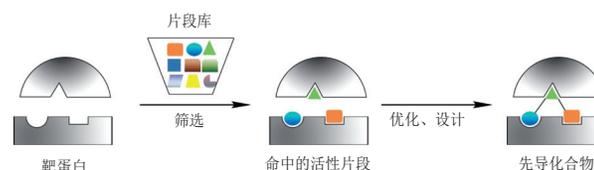


图 1 基于片段的药物设计流程示意图

Figure 1 Schematic diagram of fragment-based drug design process

1 FBDD 的研究方法

1.1 片段库的构建

FBDD 的第一个关键步骤是片段库的设计构建。

接受日期: 2020-07-15

* 通讯作者: 陆小云, 教授;

研究方向: 创新药物研发;

Tel: 020-85223259; E-mail: luxy2016@jnu.edu.cn

大多数片段库中的片段分子符合“3规则”。此外, 片段库的设计还包括评价化合物的纯度、稳定性和溶解性等方面。在片段初步筛选时, 片段的浓度一般都大于 $500 \mu\text{mol} \cdot \text{L}^{-1}$, 甚至更高^[3]。因此, 化合物在水溶液中必须有很好的溶解性。

在构建化合物库时, 还应避免包含“不良”化合物。“不良”化合物包括非特异性的结合片段、活性共价片段、泛测定干扰化合物 (pan-assay interference compounds, PAINS)、螯合剂及易聚集化合物等。具有这些特征的片段在筛选实验中, 即使对靶标没有特异性的亲和力, 依旧被作为是命中片段。通常有如下几种方法去除“不良”化合物: 1) 计算筛选可排除一些明显的“不良”化合物; 2) 通过表面等离子共振 (surface plasmon resonance, SPR) 实验中的中性传感器表面过滤掉一些“不良”化合物, 保证化合物的洁净度; 3) 在无靶蛋白结构信息时, 可通过观察配体的核磁共振 (nuclear magnetic resonance, NMR) 波谱去除自聚集化合物; 4) ALARM NMR (a La assay to detect reactive molecules by nuclear magnetic resonance) 方法鉴定泛测定干扰化合物等^[5-6]。

此外, 增加片段的复杂度和“3D特征”有助于识别诸如蛋白-蛋白相互作用等难成药性靶标的特异性抑制剂^[7]。在片段库构建时, 还需平衡片段库最大程度的多样化和包含一些复杂度相对较高的衍生物以提供初始的构效关系信息。一方面, 大量的片段蕴含着尽可能多的结合可能性, 可充分探索靶蛋白的结合位点; 另一方面, 获得命中片段初始的构效关系信息, 可为后续研究提供关键的指导信息。此外还有一些专门的片段库, 如天然产物片段库、氟原子片段库等直接用于筛选^[8-9]。

1.2 片段筛选

通常情况下, 命中片段与靶蛋白仅形成少量的高质量相互作用, 其结合亲和力一般在 $100 \sim 10\,000 \mu\text{mol} \cdot \text{L}^{-1}$ 的范围内。为了检测到筛选片段与蛋白之间较弱的结合作用, 必须使用高灵敏的筛选方法。与典型的高通量生物活性测试相比, 片段筛选时虽然片段浓度更高, 但是由于片段库中的片段往往具有良好的水溶性, 避免了因聚集而产生的假阳性分子。在众

多生物物理筛选方法中, NMR波谱法、表面等离子共振 SPR、X射线晶体学和热迁移实验 (thermal shift assay, TSA) 属于片段筛选的常用方法。

1.2.1 NMR 光谱法 NMR光谱法是片段筛选最常用的一种方法, 主要分为蛋白质观察法和配体观察法^[10-12]。在蛋白质观察法中, 片段结合信息是根据直接观察同位素标记的靶标蛋白化学位移的变化获得。配体滴定可提供二者相互作用的解离常数 (K_d), 并且共振归属可以揭示配体的结合位点。蛋白质观察 NMR方法具有很高的灵敏度, 被视为片段筛选的金标准。然而, 此方法需要大量同位素标记 [¹⁵N 和/或 (¹³C)] 蛋白, 且仅适用于相对分子质量低于 50 000 的蛋白质。与此对应的配体观察法, 仅需少量且无需同位素标记的蛋白质, 对蛋白质的相对分子质量大小亦无上限要求, 如饱和转移差 (saturation transfer difference, STD)^[13]、waterLOGSY (water-ligand observed via gradient spectroscopy)^[14] 和 ¹⁹F T2 实验法^[15]。

1.2.2 表面等离子共振 SPR是另一种广泛使用的片段筛选技术。SPR通过赖氨酸残基直接固定或使用生物素标记蛋白通过链霉亲和素-生物素相互作用将目标蛋白固定在传感器芯片上。片段流过芯片表面, 与蛋白质结合的片段增加传感器表面附近的质量, 改变界面的折射率, SPR仪器检测该变化。在SPR实验之前, 需要开发固定靶标的方法并测试固定靶标蛋白的功能性是否正常。SPR方法仅需要微量的蛋白质, 就可以用于数十个甚至数百个配体的筛选^[16-17]。SPR是一种定量方法, 当配体的溶解度高于 K_d 时, 可以从剂量反应曲线得出 K_d 。SPR的定量性质使其成为研究苗头片段类似物并确定其构效关系和结合动力学的理想方法。此外, SPR生物传感器和仪器的最新进展使该技术能够检测相对分子质量低于 100 的蛋白与小分子的结合, 并大大减少了由于非特异性结合导致的假阳性分子数目。

1.2.3 X射线晶体学 X射线晶体学通常被视为确定结构的方法, 但它也可用于片段筛选^[18]。在FBDD中使用X射线晶体学有时可识别出多个结合位点。晶体筛选的优势是可以直接获得苗头片段与靶蛋白的晶体结构, 这有助于片段的后续修饰。该方法可

能会产生结合力很弱的化合物,以至于用结晶学以外的任何方法都无法检测到它们的亲和力,命中的苗头片段只有经过修饰后,才可被其他技术方法检测,这强调了晶体学作为筛选方法的灵敏性^[19]。晶体学筛选需要强大的结晶系统,并且要求片段和蛋白具有较好的溶解性,对于一些溶解性差的膜蛋白则不适用。此外,X射线晶体学筛选尽管可以有效避免假阳性,但是有时会导致假阴性。

1.2.4 热迁移实验 TSA中小分子化合物结合靶蛋白,形成热力学稳定性增加的复合物。每种靶标蛋白都有自身的热熔曲线,随着温度的升高,蛋白会发生降解。当靶标蛋白结合小分子以后,稳定性增加,未降解蛋白的量会提高。在实验过程中根据蛋白熔解温度(T_m)的变化判断片段对蛋白的作用强弱^[20-21]。差示扫描荧光法(differential scanning fluorimetry, DSF)是TSA方法中的一种,每天可以筛选数百甚至上千个化合物。在DSF实验中,蛋白质的热解折叠可通过添加荧光染料来监控,该荧光染料与蛋白质变性时暴露出的疏水斑块结合^[22]。片段的结合亲和力和蛋白质解链的热力学性质共同决定了配体存在下蛋白质的解链温度,因此该方法筛选中,假阳性和假阴性很常见。DSF一般作为NMR筛选之前的一次富集筛选。

1.2.5 其他方法 片段筛选的方法还包括微量热泳动、弱亲和色谱、质谱技术及虚拟筛选等。微量热泳动技术是一种简单、快速、精确定量生物分子相互作用的方法。它可以测量分子在微观温度梯度场中的运动,检测蛋白水化层、电荷的改变^[23]。对于弱亲和色谱,蛋白质以类似于SPR的方式固定,待测试化合物流过阵列。该方法不是记录折射率的变化,而是测量与固定蛋白相互作用引起的化合物保留时间的差异^[24]。天然质谱法是片段筛选方法的重要补充^[25]。最简单的形式是片段/靶蛋白混合物通过软电喷雾电离进行离子化,根据靶蛋白分子离子质量的增加判断结合的片段,可以在多组分复杂体系中分析片段,片段/靶蛋白对的气相丰度可用作相对结合亲和力的粗略度量。分子对接可用于预测片段的结合模式及与蛋白的相互作用。分子对接应用于片段库的虚拟筛选,不仅需产生片段正确的结合模式,

还必须评估在溶剂环境中片段与蛋白结合时自由能的变化。分子对接的构象搜索算法和打分函数需要考虑片段的构象熵损失、去溶剂化效应、水相互作用及蛋白的构象变化等多种因素^[26]。因此大部分情况下,分子对接并不是片段库筛选的首选。当然随着对接方法和打分函数的改进和发展,分子对接在片段库筛选中的应用也将越来越多。

鉴于存在多种被广泛验证和应用的片段筛选方法,合理选择筛选方法是面临的首要难题。每种独立的方法都有假阳性和假阴性的风险,明智的策略是确定最佳组合和筛选顺序。一种策略是结合2种或多种筛选技术,并仅跟进公共的命中片段。这种严格的方法降低了公共命中片段假阳性的风险,但是增加了假阴性的风险。实际上,一些文献表明,用不同方法确定的苗头片段重叠度非常低^[27]。这可能是由于不同方法“筛选模式”和“验证模式”之间的差异导致的。另一种策略是选择一种已被证明对特定靶蛋白有效的筛选方法作为主要方法,依赖此筛选方法,直到化合物的亲和力达到其他方法可检测到为止。

1.3 片段优化

经过验证的命中片段活性一般在毫摩尔级别,往往与蛋白结合位点处氨基酸残基形成氢键、疏水相互作用等,且即使片段发展成为较大的分子,片段与蛋白的结合模式依然保守。这为片段的后续优化提供了一个高质量的起始结构。通常,X射线晶体结构数据为后续的结构改造提供了良好的基础。如果没有可用的晶体结构数据,则需综合利用分子对接和生化实验验证等方法确定片段的结合位点和作用模式。在获得活性片段与靶蛋白的结合信息后,利用基于结构的药物设计策略,兼顾化合物的合成可行性和类药性等对片段进行结构优化。片段优化主要有片段连接、片段生长及片段融合3种策略。

1.3.1 片段连接 在不同位点(最好是相邻位点)结合的非重叠片段通常可以连接在一起生成新的化学结构,这种称为片段连接的策略是片段修饰最常用的方法。通过连接片段衍生的化合物应具有比单个片段结合能之和更低的结合自由能。片段与靶蛋白以特定构象结合,理想的连接链将使连接后的化合

物中各片段保持与之前相同的作用模式, 且连接链与蛋白质可以形成额外、有利的相互作用, 并增加很少的熵损失。刚性连接链可降低结合过程中自由度和熵的损失。柔性连接链则赋予衍生物更多的可旋转键, 导致键合和溶液相中总体构象空间增加,

从而降低了对特定靶标的选择性。例如, 将作用于凝血酶 S1 口袋的对氯苯基四唑(片段 a)与作用于邻近口袋中的氨基醇(片段 b)连接得到先导化合物 **1**, IC_{50} 为 $1.4 \text{ nmol} \cdot \text{L}^{-1}$ (见图 2)^[28]。

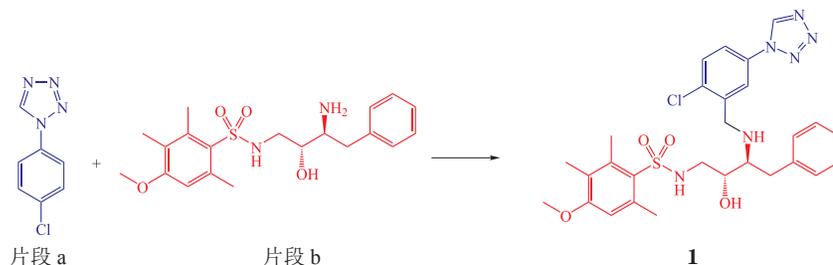


图 2 片段 a 与 b 连接得到先导化合物 **1** 的过程

Figure 2 The process of fragments a and b are connected to obtain lead compound **1**

1.3.2 片段生长 片段生长指原子被添加到片段中, 增加与蛋白形成的相互作用。片段与蛋白复合物的晶体结构是片段成功生长的关键影响因素。蛋白与现有抑制剂的相互作用对于片段的生长同样具有重要的指导意义。从片段生长到先导化合物的过程中, 配体效率监控是一种方便且客观的评估片段增加的相对分子质量是否有效的方法。一般认为最终化合物的配体效率小于 0.3, 分子体积过大会损害药代动力学性质^[29]。尽管通常优选添加疏水性官能团以增加活性, 但是化

合物脂水分配系数过高也会导致如溶解度差, 易聚集及生物利用度低等结果。因此, 应首选考虑在片段的特定位置生长极性官能团, 如氢键供体或受体, 使其与蛋白质形成氢键相互作用, 然后在后续的生长步骤中增加疏水性片段。例如, 在多肽脯氨酰顺反异构酶 (peptidyl-prolyl cis/trans isomerase 1, Pin1) 抑制剂发现过程中, 在片段 c 的基础上增加疏水片段甲基、苄基和氢键供受体酰胺片段等得到先导化合物 **2**, 活性提高了约 90 倍 (见图 3)^[30]。

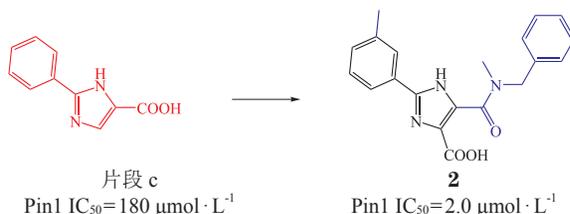


图 3 片段 c 生长为先导化合物 **2** 的过程

Figure 3 The process of fragment c grows into lead compound **2**

1.3.3 片段融合 片段融合是指利用与蛋白质有结晶结构的其他片段和已知配体的结构信息, 将重叠分子的结构部分并入片段中, 形成大的杂合分子。不同系列化合物活性变化的趋势可用于确定与蛋白结合的重要的片段和相互作用, 这些信息有助于将不

同系列中的重要片段杂化形成新的衍生物^[31]。例如, 发现乙酰胆碱结合蛋白抑制剂时, 基于片段 d ($IC_{50}=0.032 \text{ } \mu\text{mol} \cdot \text{L}^{-1}$) 与片段 e ($IC_{50}=20 \text{ } \mu\text{mol} \cdot \text{L}^{-1}$) 分别与靶蛋白形成共结晶结构, 将二者叠合后片段进行融合得到先导化合物 **3** ($IC_{50}=0.32 \text{ } \mu\text{mol} \cdot \text{L}^{-1}$) (见图 4)^[32]。

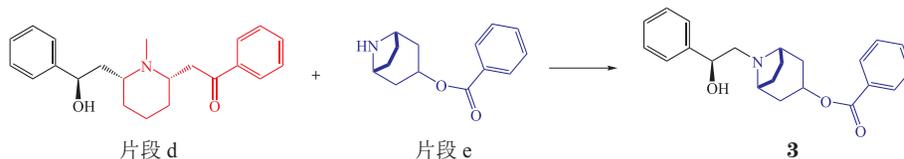


图 4 片段 d 与 e 融合得到先导化合物 **3** 的过程

Figure 4 The process of fragments d and e are fused to obtain lead compound **3**

2 FBDD 技术的药物实例介绍

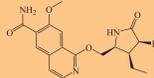
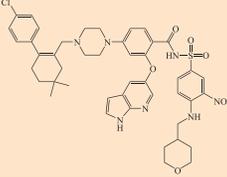
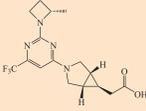
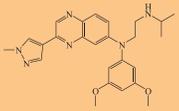
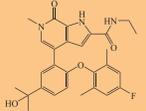
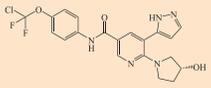
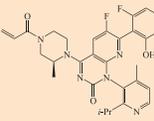
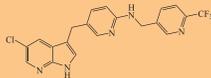
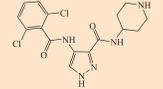
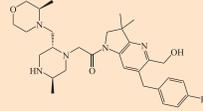
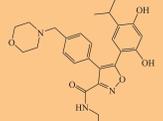
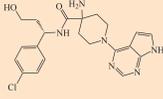
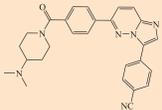
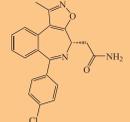
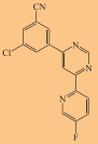
2.1 处于临床研究阶段和已上市药物

自 1996 年雅培公司报道了识别结合在靶蛋白活性口袋的低相对分子质量的片段技术——核磁共振构效关系研究法 (structure-activity relationships by nuclear magnetic resonance, SAR by NMR) 以来, FBDD 得到了广泛的应用。同时基于结构生物学、

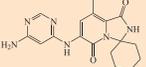
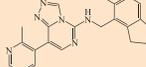
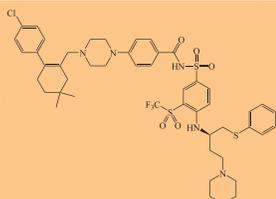
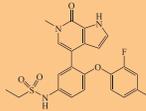
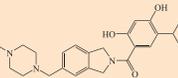
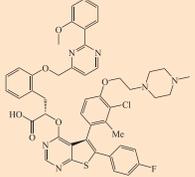
核磁共振技术和计算机技术的发展, FBDD 方法大大加快了苗头化合物的发现效率, 在苗头化合物优化成为先导化合物 / 候选化合物方面产生了不少成功的应用。据文献报道, 目前基于片段的药物发现方法获得的候选药物中有 3 个药物 (vemurafenib、venetoclax 和 erdafitinib) 已经被美国 FDA 获批上市, 另有约 21 个药物处于临床研究阶段 (见表 1)。

表 1 基于片段的药物发现方法获得的候选药物概览

Table 1 Overview of drug candidates obtained by fragment-based drug discovery methods

名称	作用靶点	结构式	研发状态	名称	作用靶点	结构式	研发状态
vemurafenib	B-Raf ^{V600E}		上市	PF-06650833	IRAK4		II 期
venetoclax	Bcl-2		上市	PF-06835919	KHK		II 期
erdafitinib	FGFR1-4		上市	ABBV-744	BET		I 期
ABL001	Bcr-Abl		III 期	AMG 510	KRAS ^{G12C}		I 期
PLX3397	CSF1R, KIT		III 期	ASTX029	ERK1/2		I 期
AT7519	CDK1, 2, 4, 5, 9		II 期	ASTX660	XIAP, cIAP1		I 期
AUY-922	HSP90		II 期	AZD5991	MCL1		I 期
AZD5363	AKT		II 期	ETC-206	MNK1/2		I 期
CPI-0610	BET		II 期	HTL14242	mGlu5 NAM		I 期

续表 1

名称	作用靶点	结构式	研发状态	名称	作用靶点	结构式	研发状态
eFT508	MNK1/2		II 期	MAK683	PRC2 EED		I 期
navitoclax (ABT263)	Bcl-2/Bcl-xL		II 期	mivebresib	BRD2-4		I 期
onalespib	HSP90		II 期	S64315	MCL1		I 期

B-Raf: mouse sarcoma virus oncogene homologous gene B (小鼠肉瘤病毒致癌基因同源基因 B); Bcl-2: B-cell lymphoma-2 (B 淋巴细胞瘤 -2); FGFR: fibroblast growth factor receptor (成纤维细胞生长因子受体); *Bcr-Abl*: breakpoint cluster region-abl oncogene (断点簇区域 -abl 癌基因); CSF1R: colony stimulation factor-1 receptor (集落刺激因子 1 受体); KIT: kit proto-oncogene protein (原癌基因蛋白质 kit); CDK: cyclin-dependent kinase (周期蛋白依赖性激酶); HSP90: heat shock protein 90 (热休克蛋白 90); AKT: protein kinase B (蛋白激酶 B); BET: bromodomain and extraterminal domain (溴结构域和末端外结构域); MNK: mitogen-activated protein kinase-interacting kinase (丝裂原活化蛋白激酶相互作用激酶); Bcl-xL: B-cell lymphoma-extra large (特大型 B 细胞淋巴瘤); IRAK4: interleukin-1 receptor-associated kinase 4 (白细胞介素 -1 受体相关激酶 4); KHK: ketohexokinase (果糖激酶); *KRAS*: K-ras2 Kirsten rat sarcoma viral oncogene homolog (Ki-ras2 Kirsten 大鼠肉瘤病毒致癌基因同源基因); ERK: extracellular regulated protein kinases (细胞外调节蛋白激酶); XIAP: X-linked inhibitor of apoptosis (X 连锁凋亡抑制蛋白); cIAP1: cellular inhibitor of apoptosis protein 1 (细胞凋亡抑制蛋白 1); *MCL1*: myeloid cell leukemia 1 (髓细胞白血病基因 1); mGlu5 NAM: (NAM-type 代谢型谷氨酸受体); PRC2 EED: polycomb repressive complex 2 embryonic ectoderm development (多梳抑制型复合物 2 胚胎外胚层发育蛋白); BRD: bromodomain-containing protein 4 (溴结构域蛋白 4)

2.2 药物研究案例分析

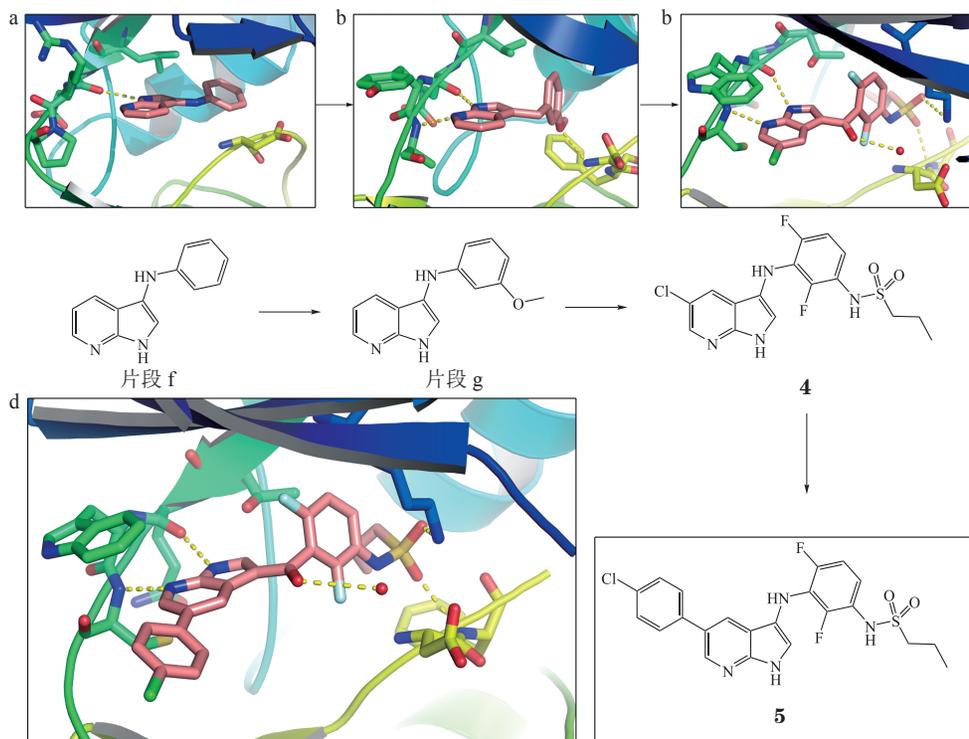
2.2.1 Vemurafenib *BRAF* 是人类重要的原癌基因之一。

BRAF^{V600E} 突变主要发生于黑色素瘤、结肠癌和甲状腺癌中。该突变导致下游 MEK (mitogen-activated extracellular signal-regulated kinase) -ERK 信号通路持续激活, 从而促进细胞的增殖和肿瘤的侵袭转移^[33]。Plexxikon 公司研究人员建立了数量约 20 000 个、相对分子质量在 150~350 的片段库用于选择性的 *BRAF*^{V600E} 抑制剂的研究。在 200 $\mu\text{mol} \cdot \text{L}^{-1}$ 浓度下, 基于该片段库针对多种具有典型性结构的激酶家族成员进行筛选。其中 238 个化合物在 200 $\mu\text{mol} \cdot \text{L}^{-1}$ 浓度下对 Pim-1 (proviral integrations of moloney virus 1)、p38 和 CSK (C terminal src-family kinase) 3 种激酶的抑制率大于 30%。随后对这 238 个化合物和上述 3 种激酶进行共结晶结构解析。如图 5a 所示, 片段 f (3-苯胺基-7-氮杂吡啶) 与 Pim-1 共结晶结构显示, 7-氮杂吡啶片段可与 Pim-1 铰链区形成氢键相互作用。但该片段对 Pim-1 的结合力较弱 ($\text{IC}_{50} \approx 200 \mu\text{mol} \cdot \text{L}^{-1}$)。

进一步对不同激酶的结构分析, 在保持与铰链区具有双氢键结合的基础上, 针对片段 f 的 2、3、4、5、6 位进行取代优化, 从而提高该类骨架抑制剂的活性与选择性。其中片段 g [3-(3-甲氧基苄基)-7-氮杂吡啶] 对 FGFR1 的抑制 IC_{50} 为 1.9 $\mu\text{mol} \cdot \text{L}^{-1}$ 。片段 g 与 FGFR1 的共晶结构显示, 7-氮杂吡啶片段可与 FGFR1 的铰链区形成双氢键作用, 甲氧基与关键氨基酸形成氢键作用, 从而限制化合物的结合构象 (见图 5b)。基于以上共结晶分析, 进一步在片段 g 的氮杂吡啶片段 3、4 和 5 位进行优化, 得到代表化合物 PLX4720 (4)。PLX4720 优先结合在 *BRAF*^{V600E} 和 *BRAF*^{WT} 的活性构象中, 对 *BRAF*^{V600E} 和 *BRAF*^{WT} 的 IC_{50} 分别为 0.013 和 0.16 $\mu\text{mol} \cdot \text{L}^{-1}$ ^[34]。然后将 PLX4720 的 5 位氯原子用 4-氯苯基取代, 最终得到 PLX4032 (vemurafenib, 5)^[35-36]。Vemurafenib 的结合模式与 PLX4720 相似, 结合在 *BRAF*^{V600E} 活性构象口袋 (见图 5d)。利用生物活性技术筛选活性片段, 随后采用基于片段生长的 FBDD 技术成功研发

获得 vemurafenib 结合模式见图 5d, 从最初的片段筛选到 vemurafenib 被美国 FDA 批准上市仅仅用了 6 年时间。2011 年, 美国 FDA 批准 vemurafenib 用于治

疗 $BRAF^{V600E}$ 突变的恶性黑色素瘤。Vemurafenib 也是首个成功的经 FBDD 技术开发并被批准上市的药物。



a: 片段 f 与 Pim-1 共结晶结构; b: 片段 g 与 FGFR1 的共结晶结构; c: PLX4720 与 $BRAF^{V600E}$ 的共结晶结构; d: vemurafenib 与 $BRAF^{V600E}$ 共结晶结构

图 5 基于 FBDD 技术发现 vemurafenib 的过程

Figure 5 The process of discovering vemurafenib based on FBDD technology

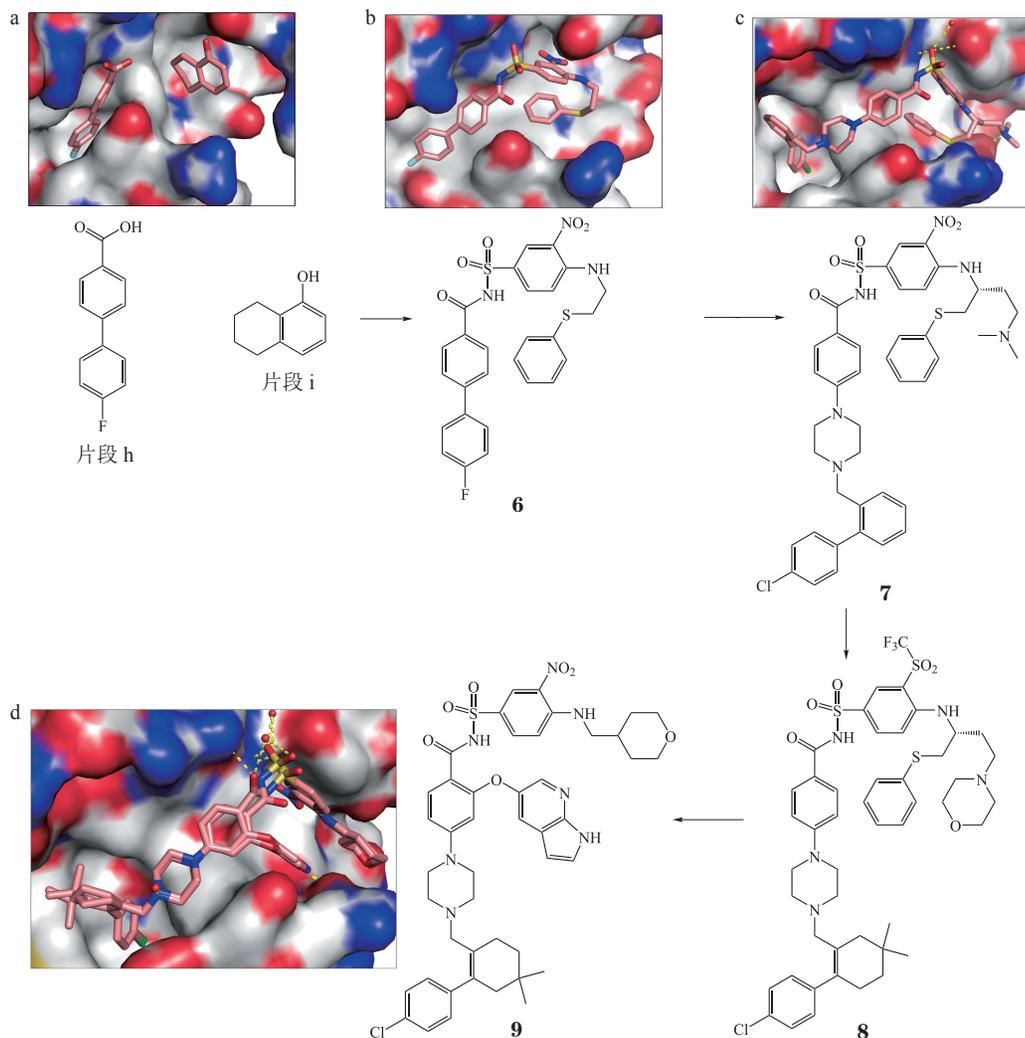
2.2.2 Venetoclax 内源性细胞凋亡主要通过 Bcl-2 蛋白家族调控。Bcl-2 蛋白家族又可分为 2 类: 抗凋亡蛋白, 比如 Bcl-2、Mcl-1 和 Bcl-XL; 促凋亡蛋白, 比如 Bax (Bcl-2 associated X protein)、Bak (Bcl-2 homologous antagonist/killer) 和 Bim (Bcl-2 interacting mediator of cell death)。在肿瘤细胞中, Bcl-2 等抗凋亡蛋白会过量表达, 与促凋亡蛋白 Bak 或者 Bax 结合形成二聚体, 使 Bax 或者 Bak 无法发挥促凋亡的功能, 从而避免细胞凋亡的发生。研究发现, 与 Bcl-2 等抗凋亡家族蛋白结合的结构域是 BH3。因此, 设计 BH3 模拟物是目前靶向 Bcl-2 家族蛋白的主要策略^[37]。

2005 年, 雅培公司运用“SAR by NMR”方法筛选小分子片段库, 识别与 Bcl-XL 上 BH3 结合的疏水性凹槽有相互作用的小分子片段^[38]。初步的筛选结果表明, 片段 h[4'-氟-4-羧酸二苯基, $K_d = (300 \pm 30) \mu\text{mol} \cdot \text{L}^{-1}$] 和片段 i[5, 6, 7, 8-四氢萘-1-醇,

$K_d = (4300 \pm 1600) \mu\text{mol} \cdot \text{L}^{-1}$] 分别结合到疏水性凹槽的相邻位点 (见图 6a)。片段 h 的羧基与 Arg139 相互作用, 而 4'-氟苯基占据由 Tyr101、Leu108、Val126 和 Phe146 形成的疏水口袋。片段 i 则占据了原本 BH3 多肽链上 Ile85 结合的空腔。设想将 2 个片段连接将会得到活性更高的分子, 在考察不同的连接方式和片段后, 采用磺酰胺连接链得到先导化合物 6, 其对 Bcl-XL 亲和性大幅提高 [$K_i = (0.036 \pm 0.0016) \mu\text{mol} \cdot \text{L}^{-1}$] (见图 6b)。随后采用基于结构的药物设计策略, 增加分子与靶标结合的疏水作用, 获得活性更强的化合物 ABT-737 (7)^[39], 图 6c 显示化合物 ABT-737 与 Bcl-XL 的结合模式。然而由于 ABT-737 的口服生物利用度较差, 进一步通过增加分子的柔性和亲水基团的优化策略, 获得候选分子 ABT-263 (navitoclax, 8), 该分子目前处于 II 期临床研究^[40]。随后在 ABT-263

的基础上, 雅培公司进一步优化旨在提高分子的选择性, 最终获得候选分子 ABT-199 (venetoclax, **9**)^[41], 其与 Bcl-XL 的结合模式见图 6d。2016 年,

美国 FDA 批准 ABT-199 用于治疗慢性淋巴细胞白血病。



a: 片段 h 与 i 分别与 Bcl-XL 上结合 BH3 疏水性凹槽结合示意图; b: 化合物 **6** 与 Bcl-XL 的结合模式; c: venetoclax 与 Bcl-XL 的结合模式; d: ABT-199 与 Bcl-XL 的结合模式

图 6 基于 FBDD 技术发现 venetoclax 的过程

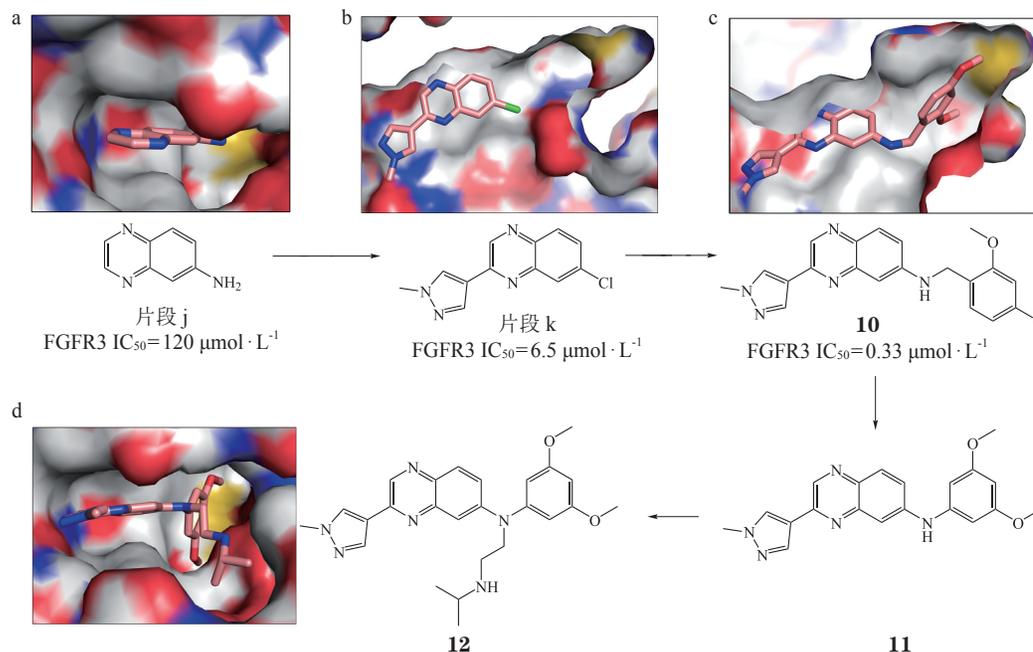
Figure 6 The process of discovering venetoclax based on FBDD technology

2.2.3 Erdaftinib FGFRs 是一类典型的受体酪氨酸激酶, 其家族包括 FGFR1、FGFR2、FGFR3 和 FGFR4 共 4 种受体。FGFRs 基因的融合、重排、易位和扩增与多种肿瘤的发生和发展密切相关^[42]。杨森制药公司研究人员对 Astex 专有数据库开展基于片段的筛选, 得到铰链区结合片段 j 和 k, 进一步利用片段 j 和 k 与 FGFR3 的共晶结构展开虚拟筛选, 得到苗头化合物 **10** (见图 7c)。化合物 **10** 是在片段 k 的 7 位引入 2,4-二甲氧基苄胺片段。随后, 又

用 3,5-二甲氧基苯胺片段替换苄胺片段, 伸向守门残基后面的疏水性口袋, 得到化合物 **11**, 该化合物对 FGFR3 的抑制活性达 $0.012 \mu\text{mol} \cdot \text{L}^{-1}$ 。进一步优化该类化合物对 FGFR 家族成员的抑制活性和药代动力学性质, 用含有碱性氨基的烷基链取代化合物 **11** 中苯胺上的氢原子, 从而使侧链末端的碱性氮原子与 DFG 基序上的 Asp641 产生作用, 最终得到 JNJ-42756493 (erdaftinib, **12**)^[43], 结合模式见图 7d。Erdaftinib 对 FGFR1、FGFR2、FGFR3

和 FGFR4 各亚型的 IC_{50} 分别为 0.0 012、0.0 025、0.003 和 0.0 057 $\mu\text{mol} \cdot \text{L}^{-1}$ [44]。2019 年美国 FDA 批准 erdafitinib 用于接受化疗后病情发展并且肿瘤中

存在特定 *FGFR* 基因突变的局部晚期或转移性尿路上皮癌患者的治疗。



a: 片段 j 与 FGFR3 的共结晶结构; b: 片段 k 与 FGFR3 的共结晶结构; c: 化合物 10 与 FGFR3 的共结晶结构; d: erdafitinib 与 BRAF^{V600E} 共结晶结构

图 7 基于 FBDD 技术发现 erdafitinib 的过程

Figure 7 The process of discovering erdafitinib based on FBDD technology

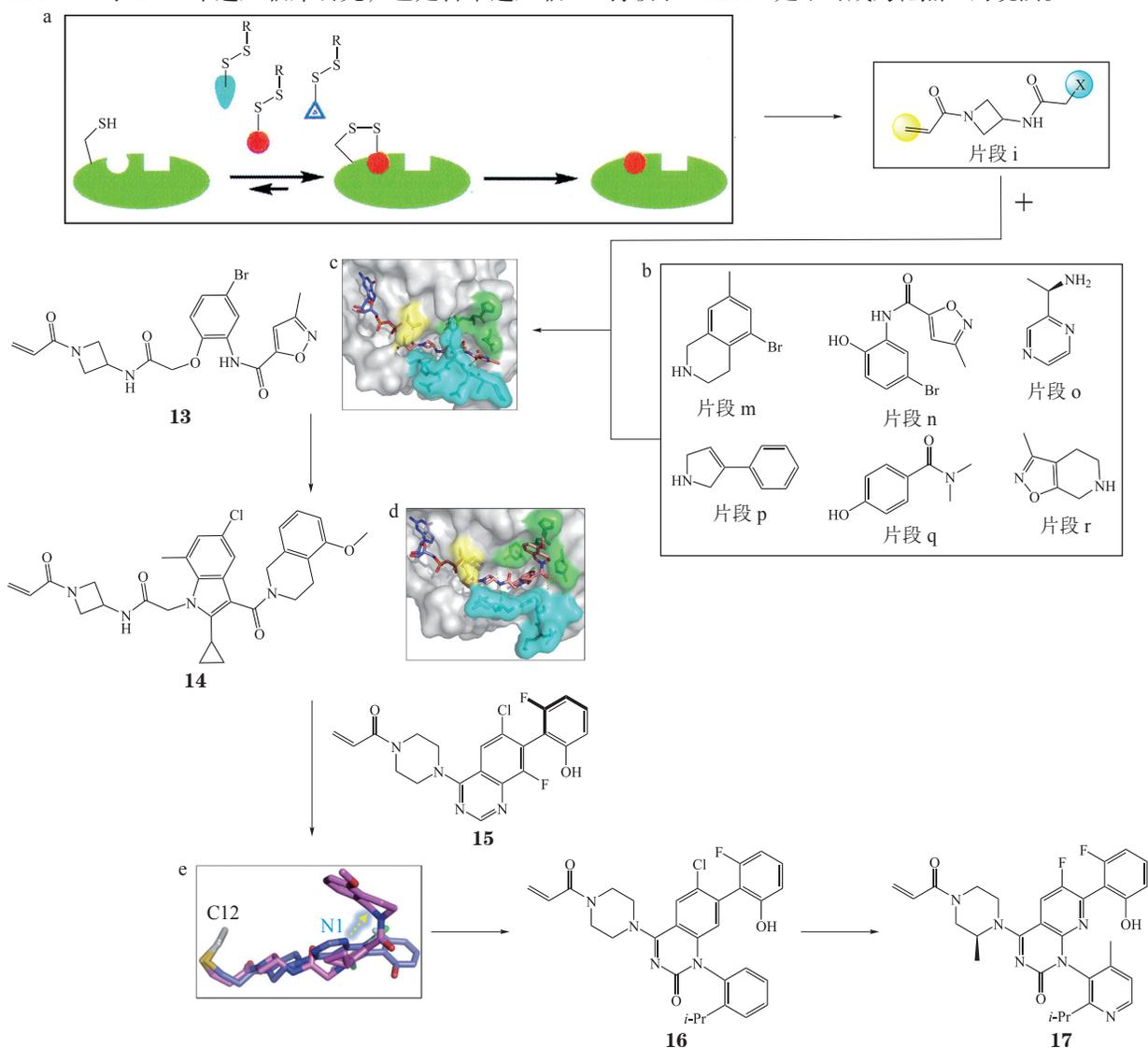
2.2.4 AMG 510 *RAS* 基因家族目前已知的成员包括 *KRAS*、*NRAS* 和 *HRAS*，其中 *KRAS* 突变最为常见，大约占 85%。*KRAS*^{G12C} 突变在非小细胞肺癌中约占 13%，在结肠直肠癌中约占 3%~5%，在其他实体瘤中约占 1%~2% [45]。近年来，针对 *KRAS*^{G12C} 抑制剂的开发是该领域的研究热点之一，尤其是基于变构的 switch II 的抑制剂研究。为了获得高质量的苗头化合物，安进公司研究人员基于 Carmot 公司开发的“Tethering”方法针对 *KRAS*^{G12C} 进行基于 FBDD 的药物设计 [46]。“Tethering”方法适合于含有反应性氨基酸残基 Cys 的靶标蛋白，化合物库中小分子结构中都包含二硫键。片段与靶标蛋白共孵育后，利用电喷雾质谱判断可以“修饰”靶标蛋白的片段（见图 8a）。经过“Tethering”方法筛选，成功获得丙烯酰胺类不可逆结合的“bait”片段 i。随后基于 3 300 个片段分子库建立新化合物库筛选，结合质谱技术确定了结合较好的苗头化合物 13。化合物 13 与 *KRAS*^{G12C} 的晶体结构显示其结合在 P2 口袋，且

丙烯酰胺与 Cys12 形成共价结合，酰胺基团的羰基与 Lys16 形成氢键。处于中间部位的酰胺连接基团与氨基酸 Ala59 和 Tyr96 形成一对氢键，异噁唑部分指向处于关闭构象的 switch II 口袋（见图 8c）。

为提高化合物的活性，进一步采用 Chemotype Evolution 的组合筛选方法（多轮分子库构建和筛选）获得活性更优的第 2 代苗头化合物 14。与化合物 13 的结合构象比较，化合物 14 的四氢异噁唑部分可占据由 His95/Tyr96/Gln99 组成的隐藏口袋，且 switch II 处于开放式构象（见图 8d）。化合物 14 在人胰腺癌细胞上表现较好的抑制活性（ $IC_{50}=0.067 \mu\text{mol} \cdot \text{L}^{-1}$ ）。然而，药代动力学性质研究显示由于清除率相对较高，导致化合物 14 的生物利用度较低。随后，安进公司将已报道 GDP-*KRAS*^{G12C} 抑制剂 ARS-1620 (15) 与先导化合物 14 进行结合构象叠合（见图 8e），推测在 ARS-1620 噁唑母核的 1 位引入取代苯环等取代基可占据由 His95/Tyr96/Gln99 构成的隐藏口袋，合成得到化合物 16 [47]。进一步基于结构和药

代动力学性质的优化, 最终得到 AMG 510 (17)。AMG510 于 2019 年进入临床研究, 也是首个进入临

床实验的 KRAS^{G12C} 抑制剂。AMG510 的成功研发也打破了“KRAS 是不可成药靶点”的说法。



a: 高通量质谱分析共价蛋白抑制剂复合物, 用于发现不可逆抑制剂; b: 筛选得到的与蛋白非共价结合片段; c: 苗头化合物 **13** 与 KRAS^{G12C} 共结晶结构; d: 化合物 **14** 与 KRAS^{G12C} 共结晶结构; e: 化合物 **14** 与 ARS-1620 分别与 KRAS^{G12C} 结合构象的叠合

图 8 基于 FBDD 技术发现 AMG 510 的过程

Figure 8 The process of discovering AMG 510 based on FBDD technology

3 结语与展望

FBDD 的研究过程包括高质量片段库的设计构建, 蛋白-片段相互作用的验证和基于结构药物设计指导下的苗头片段的优化等。通过 FBDD 方法发现的片段通常具有较高的配体效率, 在获得片段与靶蛋白的结构信息基础上, 利用配体结合效率为指导方向, 快速优化片段。近年来, 随着筛选技术的不断进步, 检测片段与靶蛋白之间弱结合作用的表

征变得越来越容易, 使得 FBDD 技术成为近年来发展最快的药物研究新方法之一。

尽管 FBDD 技术目前取得了巨大的成功, 但是仍存在以下几个方面需要改进: 首先是片段库的质量, 过去几十年, FBDD 技术强调了片段质量的重要性, 即要确保进行实验筛选的片段具有合适的物理化学特性, 例如溶解度和避免干扰读数 (如聚集等)。现在需要同时保持片段具有良好的理化性

质且分子量足够小以满足大面积化学空间的需求。其次, 提高用于检测片段结合筛选技术的灵敏度。FBDD 方法需要高灵敏地识别片段分子与靶蛋白的结合, 需要发展使用微量蛋白, 实现 $1\ 000\ \mu\text{mol}\cdot\text{L}^{-1}$ 亲和力下可靠地检测片段与蛋白的结合。最后, 如何高效地优化苗头片段到先导化合物也是一个很大

的挑战。在结构生物学和分子对接等技术指导下, 充分利用片段连接、生长和融合等方法提高先导化合物优化设计的成功率。总之, 基于片段的药物发现是一项融合药物化学、结构生物学、计算机辅助药物设计等多学科的药物研究新兴技术, 在未来新药研发方面将发挥重要的作用。

【参考文献】

- [1] Fattori D. Molecular recognition: the fragment approach in lead generation[J]. *Drug Discov Today*, 2004, 9(5): 229-238.
- [2] 程俊飞, 盛春泉, 董国强. 基于片段的药物发现 (FBDD) 研究进展[J]. *中国药物化学杂志*, 2018, 28(6): 504-507.
- [3] Congreve M, Carr R, Murray C, et al. A 'rule of three' for fragment-based lead discovery?[J]. *Drug Discov Today*, 2003, 8(19): 876-877.
- [4] Erlanson D A, Fesik S W, Hubbard R E, et al. Twenty years on: the impact of fragments on drug discovery[J]. *Nat Rev Drug Discov*, 2016, 15(9): 605-619.
- [5] Huth J R, Mendoza R, Olejniczak E T, et al. ALARM NMR: a rapid and robust experimental method to detect reactive false positives in biochemical screens[J]. *J Am Chem Soc*, 2005, 127(1): 217-224.
- [6] Baell J B, Holloway G A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays[J]. *J Med Chem*, 2010, 53(7): 2719-2740.
- [7] Morley A D, Pugliese A, Birchall K, et al. Fragment-based hit identification: thinking in 3D[J]. *Drug Discov Today*, 2013, 18(23/24): 1221-1227.
- [8] Over B, Wetzel S, Grütter C, et al. Natural-product-derived fragments for fragment-based ligand discovery[J]. *Nat Chem*, 2013, 5(1): 21-28.
- [9] Vulpetti A, Dalvit C. Design and generation of highly diverse fluorinated fragment libraries and their efficient screening with improved ^{19}F NMR methodology[J]. *ChemMedChem*, 2013, 8(12): 2057-2069.
- [10] Shuker S B, Hajduk P J, Meadows R P, et al. Discovering high-affinity ligands for proteins: SAR by NMR[J]. *Science*, 1996, 274(5292): 1531-1534.
- [11] Stockman B J, Dalvit C. NMR screening techniques in drug discovery and drug design[J]. *Prog Nucl Magn Reson Spectrosc*, 2002, 41(3/4): 187-231.
- [12] Lepre C A. Practical aspects of NMR-based fragment screening[J/OL]. *Methods Enzymol*, 2011, 493: 219-239[2020-07-15]. <https://doi.org/10.1016/B978-0-12-381274-2.00009-1>.
- [13] Haselhorst T, Lamerz A C, Itzstein M. Saturation transfer difference NMR spectroscopy as a technique to investigate protein-carbohydrate interactions in solution[J/OL]. *Methods Mol Biol*, 2009, 534: 375-386[2020-07-15]. https://link.springer.com/protocol/10.1007%2F978-1-59745-022-5_26.
- [14] Dalvit C, Fogliatto G, Stewart A, et al. WaterLOGSY as a method for primary NMR screening: practical aspects and range of applicability[J]. *J Biomol NMR*, 2001, 21(4): 349-359.
- [15] Dalvit C, Fagerness P E, Hadden D T A, et al. Fluorine-NMR experiments for high-throughput screening: theoretical aspects, practical considerations, and range of applicability[J]. *J Am Chem Soc*, 2003, 125(25): 7696-7703.
- [16] Danielson U H. Integrating surface plasmon resonance biosensor-based interaction kinetic analyses into the lead discovery and optimization process[J]. *Future Med Chem*, 2009, 1(18): 1399-1414.
- [17] Giannetti A M. From experimental design to validated hits: a comprehensive walk-through of fragment lead identification using surface plasmon resonance[J/OL]. *Method Enzymol*, 2011, 493: 169-218[2020-07-15]. <https://linkinghub.elsevier.com/retrieve/pii/B978-0-12-381274-2.00008-X>. Doi: 10.1016/B978-0-12-381274-2.00008-X.
- [18] Jhoti H, Cleasby A, Verdonk M, et al. Fragment-based screening using X-ray crystallography and NMR spectroscopy[J]. *Curr Opin Chem Biol*, 2007, 11(5): 485-493.
- [19] Davies T G, Wixted W E, Coyle J E, et al. Monoacidic inhibitors of the kelch-like ECH-associated protein 1: nuclear factor erythroid 2-related factor 2 (KEAP1:NRF2) protein-protein interaction with high cell potency identified by fragment-based discovery[J]. *J Med Chem*, 2016, 59(8): 3991-4006.
- [20] Lo M C, Aulabaugh A, Jin G X, et al. Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery[J]. *Anal Biochem*, 2004, 332(1): 153-159.
- [21] Vedadi M, Niesen F H, Allali-Hassani A, et al. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination[J]. *Proc Natl Acad Sci USA*, 2006, 103(43): 15835-15840.
- [22] Mashalidis E H, Śledź P, Lang S, et al. A three-stage biophysical screening cascade for fragment-based drug discovery[J]. *Nat Protoc*, 2013, 8(11): 2309-2324.
- [23] Jerabek-Willemsen M, Wienken C J, Braun D, et al. Molecular interaction studies using microscale thermophoresis[J]. *Assay Drug*

- Dev Technol*, 2011, 9(4): 342-353.
- [24] Meiby E, Simmonite H, Strat L, *et al.* Fragment screening by weak affinity chromatography: comparison with established techniques for screening against HSP90[J]. *Anal Chem*, 2013, 85(14): 6756-6766.
- [25] Ohlson S, Duong M D. Fragment screening for drug leads by weak affinity chromatography (WAC-MS)[J/OL]. *Methods*, 2018, 146: 26-38[2020-07-15]. [https://linkinghub.elsevier.com/retrieve/pii/S1046-2023\(17\)30395-X](https://linkinghub.elsevier.com/retrieve/pii/S1046-2023(17)30395-X). Doi: 10.1016/j.ymeth.2018.01.011.
- [26] Law J M, Fung D Y, Zsoldos Z, *et al.* Validation of the SPROUT *de novo* design program[J]. *J MOL STRUC-THEOCHEM*, 2003, 666-667(1): 651-657.
- [27] Wielens J, Headey S J, Rhodes D I, *et al.* Parallel screening of low molecular weight fragment libraries: do differences in methodology affect hit identification?[J]. *J Biomol Screen*, 2013, 18(2): 147-159.
- [28] Howard N, Abell C, Blakemore W, *et al.* Application of fragment screening and fragment linking to the discovery of novel thrombin inhibitors[J]. *J Med Chem*, 2006, 49(4): 1346-1355.
- [29] Veber D F, Johnson S R, Cheng H Y, *et al.* Molecular properties that influence the oral bioavailability of drug candidates[J]. *J Med Chem*, 2002, 45(12): 2615-2623.
- [30] Potter A, Oldfield V, Nunns C, *et al.* Discovery of cell-active phenylimidazole Pin1 inhibitors by structure-guided fragment evolution[J]. *Bioorg Med Chem Lett*, 2010, 20(22): 6483-6488.
- [31] Erlanson D A, McDowell R S, O'Brien T. Fragment-based drug discovery[J]. *J Med Chem*, 2004, 47(14): 3463-3482.
- [32] Edink E, Rucktooa P, Retra K, *et al.* Fragment growing induces conformational changes in acetylcholine-binding protein: a structural and thermodynamic analysis[J]. *J Am Chem Soc*, 2011, 133(14): 5363-5371.
- [33] Davies H, Bignell G R, Cox C, *et al.* Mutations of the *BRAF* gene in human cancer[J]. *Nature*, 2002, 417(6892): 949-954.
- [34] Tsai J, Lee J T, Wang W R, *et al.* Discovery of a selective inhibitor of oncogenic B-Raf kinase with potent antimelanoma activity[J]. *Proc Natl Acad Sci USA*, 2008, 105(8): 3041-3046.
- [35] Yang H, Higgins B, Kolinsky K, *et al.* RG7204 (PLX4032), a selective BRAF^{V600E} inhibitor, displays potent antitumor activity in preclinical melanoma models[J]. *Cancer Res*, 2010, 70(13): 5518-5527.
- [36] Bollag G, Tsai J, Zhang J Z, *et al.* Vemurafenib: the first drug approved for *BRAF*-mutant cancer[J]. *Nat Rev Drug Discov*, 2012, 11(11): 873-886.
- [37] Roberts A W, Seymour J F, Brown J R, *et al.* Substantial susceptibility of chronic lymphocytic leukemia to BCL2 inhibition: results of a phase I study of navitoclax in patients with relapsed or refractory disease[J]. *J Clin Oncol*, 2012, 30(5): 488-496.
- [38] Oltersdorf T, Elmore S W, Shoemaker A R, *et al.* An inhibitor of Bcl-2 family proteins induces regression of solid tumours[J]. *Nature*, 2005, 435(7042): 677-681.
- [39] Bruncko M, Oost T K, Belli B A, *et al.* Studies leading to potent, dual inhibitors of Bcl-2 and Bcl-xL[J]. *J Med Chem*, 2007, 50(4): 641-662.
- [40] Park C M, Bruncko M, Adickes J, *et al.* Discovery of an orally bioavailable small molecule inhibitor of prosurvival B-Cell lymphoma 2 proteins[J]. *J Med Chem*, 2008, 51(21): 6902-6915.
- [41] Levenson J D, Phillips D C, Mitten M J, *et al.* Exploiting selective BCL-2 family inhibitors to dissect cell survival dependencies and define improved strategies for cancer therapy[J/OL]. *Sci Transl Med*, 2015, 7(279): 279ra40[2020-07-15]. <https://stm.sciencemag.org/content/7/279/279ra40.short>. Doi: 10.1126/scitranslmed.aaa4642.
- [42] Hallinan N, Finn S, Cuffe S, *et al.* Targeting the fibroblast growth factor receptor family in cancer[J/OL]. *Cancer Treat Rev*, 2016, 46: 51-62[2020-07-15]. <https://www.sciencedirect.com/science/article/pii/S0305737216300135?via%3Dihub>. Doi: 10.1016/j.ctrv.2016.03.015.
- [43] Murray C W, Newell D R, Angibaud P. A successful collaboration between academia, biotech and pharma led to discovery of erdafitinib, a selective FGFR inhibitor recently approved by the FDA[J/OL]. *MedChemComm*, 2019, 10: 1509-1511[2020-07-15]. <https://pubs.rsc.org/en/content/articlepdf/2019/md/c9md90044f>. Doi: 10.1039/C9MD90044F.
- [44] Perera T P S, Jovcheva E, Mevellec L, *et al.* Discovery and pharmacological characterization of JNJ-42756493 (erdafitinib), a functionally selective small-molecule FGFR family inhibitor[J]. *Mol Cancer Ther*, 2017, 16(6): 1010-1020.
- [45] Adjei A A. Blocking oncogenic ras signaling for cancer therapy[J]. *J Natl Cancer Inst*, 2001, 93(14): 1062-1074.
- [46] Erlanson D A, Braisted A C, Raphael D R, *et al.* Site-directed ligand discovery[J]. *Proc Natl Acad Sci USA*, 2000, 97(17): 9367-9372.
- [47] Lanman B A, Allen J R, Allen J G, *et al.* Discovery of a covalent inhibitor of KRASG12C (AMG 510) for the treatment of solid tumors[J]. *J Med Chem*, 2020, 63(1): 52-65.



[专家介绍] 陆小云: 博士, 研究员, 博士生导师, 国家优秀青年基金获得者, 863 青年科学家, 广东省杰出青年基金获得者等。近年来重点围绕肺结核和肺癌等重要肺部疾病的临床需求, 针对疾病密切相关的生物学靶标, 利用“基于结构的药物设计”为技术手段, 设计和合成全新的药物先导物, 为治疗肺部疾病的创新药物研发提供化学物质基础。以第一或通讯作者在 *Angew Chem Int Edit* 和 *J Med Chem* 等本专业期刊发表 SCI 论文 60 余篇; 参编英文专著 3 部; 申请发明专利 30 余项, 授权 12 项(美国 1 项), 部分已实现转移转化。先后主持 5 项国家级科研项目和多项省部级项目。