

生物信息学及分子模拟在蛋白质降解靶向嵌合体 (PROTAC) 中的研究

朱 婕^{1#}, 聂锦宜^{1#}, 刘小益¹, 梁中洁^{1,2*}

(1. 苏州大学苏州医学院, 江苏 苏州 215123; 2. 江苏省精准诊疗药物创制工程研究中心, 江苏 苏州 215123)

摘要: 蛋白质降解靶向嵌合体 (proteolysis targeting chimera, PROTAC) 是一种利用泛素-蛋白酶体系统 (ubiquitin proteasome system, UPS) 对靶标蛋白进行降解的药物开发技术。与传统小分子药物通过“占位驱动”作用模式调控蛋白质活性的方式不同, PROTAC 通过形成稳定的靶蛋白-PROTAC-E3 泛素连接酶三元复合物, 并利用泛素-蛋白酶体系统降解靶标蛋白而发挥作用。然而, 目前只有少数 E3 泛素连接酶被用到 PROTAC 分子设计中, PROTAC 可靶向的靶标蛋白空间也有待进一步拓展; 另一方面, 靶蛋白-PROTAC-E3 泛素连接酶的三元复合物晶体结构解析难度较大, 计算模拟方法为构建三元复合物模型提供了有效的工具。基于此, 本文介绍了生物信息学在拓展 E3 泛素连接酶及扩大靶标蛋白空间方面的研究进展, 总结了计算模拟在 PROTAC 三元复合物结构模拟中的方法研究, 并探讨了 PROTAC 技术的发展趋势。

关键词: 蛋白质降解靶向嵌合体; E3 泛素连接酶; 靶标空间; 三元复合物结构模拟; 蛋白质翻译后修饰; 异双功能分子

中图分类号: R914

文献标识码: A

文章编号: 0513-4870(2024)06-1546-16

Research on bioinformatics and molecular simulation in proteolysis targeting chimeras (PROTAC)

ZHU Jie^{1#}, NIE Jin-yi^{1#}, LIU Xiao-yi¹, LIANG Zhong-jie^{1,2*}

(1. Suzhou Medical College of Soochow University, Suzhou 215123, China; 2. Jiangsu Province Engineering Research Center of Precision Diagnostics and Therapeutics Development, Suzhou 215123, China)

Abstract: Proteolysis targeting chimera (PROTAC) is a drug discovery strategy using ubiquitin proteasome system (UPS) to degrade the target protein. Unlike traditional small molecule drugs utilizing occupancy-driven pharmacology as the mode of action (MOA) to regulate protein activity, PROTACs function through forming stable target protein-PROTAC-E3 ubiquitin ligase ternary complex and use ubiquitin proteasome system to degrade the target protein. However, only a few E3 ubiquitin ligases have been used in PROTAC drug design now, and the space of target proteins that PROTAC can target needs to be further expanded. On the other hand, the complicated system of ternary crystal structures is difficult to capture and identify, computational simulation provides modeling of PROTAC-mediated ternary complex formation with effective approaches. In view of this, this review describes the recent progress of bioinformatics on expanding the landscape of E3 ubiquitin ligases and target proteins, and summarizes the methods of computation simulation in modeling PROTAC ternary complex. Finally, the trend of development about PROTAC is prospected.

收稿日期: 2023-11-17; 修回日期: 2024-03-21.

基金项目: 国家自然科学基金项目 (22377089); 江苏省高等学校重点学科建设项目 (PAPD); 大学生创新创业训练计划项目 (202210285179Y).

#共同第一作者.

*通讯作者 Tel: 86-512-65880103, E-mail: zjliang@suda.edu.cn

DOI: 10.16438/j.0513-4870.2023-1302

Key words: proteolysis targeting chimera; E3 ubiquitin ligase; target landscape; ternary complex computational simulation; protein post-translational modification; heterobifunctional molecule

1 PROTAC的原理及其作用特点

蛋白质降解靶向嵌合体 (proteolysis targeting chimeras, PROTAC) 的概念最早由 Crews 和 Deshaies 实验室于 2001 年提出, 该团队通过 PROTAC 技术成功实现了对靶蛋白甲硫氨酰氨肽酶 2 (methionine aminopeptidase-2, MetAp-2) 的降解^[1]。此后, 越来越多的 PROTAC 进入临床研究。其中靶向降解雄激素受体 (androgen receptor, AR) 的 PROTAC 进入 II 期临床, 靶向雌激素受体 (estrogen receptor, ER) 的 PROTAC 即将进入 III 期临床, 分别用于前列腺癌和乳腺癌治疗^[2]。目前为止, PROTAC 技术已被用于降解与各种疾病相关的不同靶蛋白, 在癌症、免疫疾病、神经退行性疾病及心血管疾病中显示出了良好的临床疗效^[3]。

PROTAC 是异型双功能分子, 一方面靶向靶标蛋白 (protein of interest, POI), 另一方面招募泛素连接酶 (ubiquitin ligases, E3), 实现对靶标蛋白的靶向降解。PROTAC 分子由三个部分组成: 靶蛋白结合的配体、E3 泛素连接酶结合的配体以及连接两个配体的连接子。通过诱导 E3 连接酶和靶蛋白接近形成三元复合物, 利用泛素-蛋白酶体系统 (ubiquitin proteasome system, UPS) 对靶蛋白进行降解^[4] (图 1A)。在真核细胞中, 泛素-蛋白酶体系统是通过降解缺陷和受损蛋白质来维持蛋白质稳态的主要机制。在该途径中, 蛋白质经由三步被蛋白酶体识别并进行靶向降解, 该过程涉及三种酶: 泛素激活酶 (ubiquitin-activating enzyme, E1)、泛素偶联酶 (ubiquitin-conjugating enzyme, E2) 及泛素连接酶 (ubiquitin ligase, E3)。首先, 活化的泛素 (Ub) 分子与 E1 泛素激活酶连接, 随后 E1 泛素激活酶与 E2 泛素偶联酶结合并将活化的泛素转移到 E2 泛素偶联酶上, 最后 E3 泛素连接酶催化泛素从 E2 泛素偶联酶转移到靶蛋白的赖氨酸侧链上^[5]。

相较于其他药物, PROTAC 具有靶向不可成药蛋白、作用时间长、可重复利用、克服耐药、高选择性和特异性及给药途径多等优势^[6] (图 1B), 具有非常广阔的应用前景。与传统药物需要长时间高强度地结合在活性口袋中的“占位驱动”作用模式不同, PROTAC 以“事件驱动”方式降解靶蛋白, 只需要产生结合作用就可能诱导靶蛋白降解, 大大提高可靶向蛋白的空间。即使 PROTAC 解离, 泛素化的靶蛋白会被继续降解, 靶蛋白仍需要较长时间才能恢复至发挥功能的水平。PROTAC 可以在下一个降解周期中被回收利用且活性

不变^[7], 避免了高剂量给药和相应的不良反应, 使得 PROTAC 以较低的浓度实现较高的活性。在临床前研究中已有初步数据表明这种低剂量、低频次用药的潜力。此外, 传统的小分子抑制剂仅能阻断蛋白活性, PROTAC 通过完全消除靶蛋白, 不仅可以阻断靶蛋白的催化功能, 还可以影响其他功能, 如蛋白间相互作用及组成大型蛋白复合物的能力等。在某种程度上, PROTAC 能够解决传统抑制剂引起的获得性耐药, 因为靶蛋白降解可能阻断细胞蛋白稳态的某些反馈机制^[6]。虽然有研究表明使用 PROTAC 治疗也可能导致耐药, 但这种耐药机制与小分子抑制剂不同, 是由于 E3 连接酶突变导致三元复合物的形成受限。这种耐药性可以通过改变 E3 连接酶来解决。

相较于单克隆抗体等大分子药物, PROTAC 可靶向大量细胞内靶标蛋白质, 具有较高的组织渗透性可实现口服用药。根据疾病和需求, 可开发出口服、注射及输注给药的 PROTAC 蛋白降解剂。诚然, 过去几十年 PROTAC 技术取得了很大的成功, 为药物开发提供了新的思路。但随之而来的是这项新技术面临的问题与挑战, 如 PROTAC 的药代动力学、靶蛋白位置受限、提高选择性和特异性及脱靶等问题^[6]。为此, 越来越多的研究针对这些问题开发了新的衍生技术, 例如使用抗体-PROTAC 偶联物 (antibody-PROTAC conjugates, APC) 将 PROTAC 偶联到肿瘤特异性抗体上^[8], 以肿瘤导向的方式选择性地靶向位于特定组织细胞的蛋白质, 实现特异性靶向降解; 还有三价 PROTAC (trivalent PROTAC), 通过提高分子的亲和力和协同性, 形成稳定的复合物更好地结合靶蛋白, 提高蛋白降解效率, 产生更有效的抗癌活性, 特别适用于获得耐药性的肿瘤^[9]; 此外, 针对 PROTAC 降解蛋白位置受限这一问题, 研究人员设计出了新的技术, 通过溶酶体降解途径降解膜蛋白及跨膜蛋白等, 可以靶向定位不同的靶蛋白^[10]。

尽管发展了很多新技术来弥补 PROTAC 本身的不足, 但这项技术仍存在一些不可忽视的挑战 (图 1B)。首先, PROTAC 的 ADMET 性质需要优化。PROTAC 本身不符合类药五规则, 其分子量在 700~1 200 Da 之间, 溶解度、透膜能力与口服生物利用度较差^[6]。其次, PROTAC 降解靶蛋白过程中有脱靶的可能性, 这种潜在毒性比传统药物更大, 且缺少有效的检测手段。尽管目前开发出的衍生技术如光化学靶向嵌合体 (PHOTAC)、半导体聚合物纳米-PROTAC (SPNpro) 及

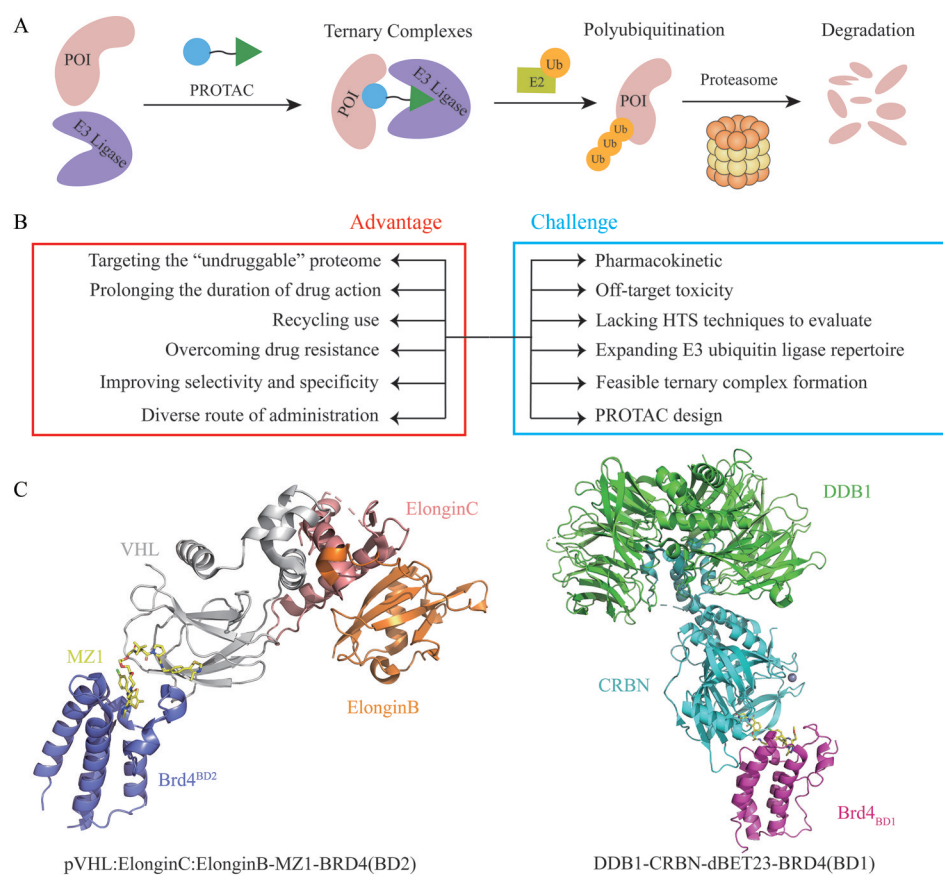


Figure 1 Proteolysis targeting chimeras (PROTAC). A: The mechanism of PROTAC; B: Advantage and challenge of PROTAC; C: Crystal structures of Brd4^{BD2}: MZ1: VHL-ElonginC-ElonginB complex (PDB ID: 5T35, left) and DDB1 Δ B-CRBN-dBET23-BRD4^{BD1} complex (PDB ID: 6BN7, right). POI: Protein of interest

与叶酸基团相连的 PROTAC (floate-PROTAC) 等^[11]来解决脱靶毒性的问题,但并不适用于所有情况。再者,目前尚无有效的高通量筛选技术用于快速、大量地评估 PROTAC 降解靶蛋白的能力,这降低了开发 PROTAC 的速度与成功率。

此外, E3 连接酶是降解靶蛋白选择性和特异性的关键因素之一,对于实现精准靶向蛋白质降解有重要意义。目前临床开发主要选择的是 CRBN (Cereblon) 和 VHL (Von-Hippel-Lindau) 两种 E3 连接酶。然而 E3 连接酶的突变会对降解效力产生影响,因此需要拓展 E3 连接酶的空间来提高 PROTAC 靶向蛋白的选择性和特异性^[12]。在 PROTAC 诱导靶蛋白降解的过程中,靶蛋白-PROTAC-E3 泛素连接酶三元复合物的形成是实现有效性和选择性降解靶蛋白的关键。对于 PROTAC 来说,靶蛋白结合配体和 E3 连接酶配体虽然无需像其他药物一样与蛋白质紧密结合,但两个配体与靶蛋白和 E3 连接酶有一定的亲和力。当 PROTAC 的浓度增加到一定值时,多余的 PROTAC 会分别与靶蛋白或 E3 连接酶形成二元复合物,这种现象称为钩效

应 (hook effect)^[12]。这对形成稳定的三元复合物是不利的,可能会导致严重不良反应,增大脱靶毒性。此外,靶蛋白与 E3 连接酶之间的相互作用也会影响三元复合物的形成,两者之间可以产生使三元复合物更加稳定的吸引作用,或是破坏三元复合物稳定性的排斥作用,在一定程度上影响 PROTAC 的降解效率,这就是所谓的协同性。协同性指 PROTAC 结合的三元复合物解离常数之比 (α),用于描述三元复合物的生成效率^[13]。三元复合物的第一个晶体结构 (BRD4-MZ1-VHL, PDB ID: 5T35) 说明 PROTAC 诱导的靶蛋白和 E3 连接酶之间的表面静电作用在稳定三元复合物方面具有正协同作用^[14]。较高的 α 被认为与较弱的钩效应相关,这是 PROTAC 开发设计的重要环节。尽管有研究表明,负协同性的三元复合物也可能导致降解,但整体而言协同性对于衡量 PROTAC 诱导的三元复合物的形成有重要意义,了解三元复合物的结构和稳定性有助于优化降解,改进 PROTAC 设计^[15]。然而目前仅有少数三元复合物的晶体结构得到解析,例如 BRD4 蛋白的不同结构域分别与 VHL 和 CRBN 作用

体系的三元复合物晶体结构(图1C)^[14,16]。在形成复合物过程中, PROTAC连接子的长度、组成和连接位置、配体蛋白作用以及两种蛋白质相互作用的方式均会导致不同的结合方式。由于目前的实验方法很难得到三元复合物的晶体结构,所以在缺失晶体结构的情况下很难对PROTAC进行合理改造。

综上, 拓展E3连接酶和靶标空间以及构建三元复合物对PROTAC开发和设计具有重要意义。所以, 本文首先介绍了生物信息学方法在拓展E3连接酶和可降解靶蛋白空间方面的研究, 尤其是针对PROTAC选择性和特异性的方面; 其次介绍了计算模拟方法包括蛋白对接、采样、聚类等方法以及目前构建的计算流程在三元复合物建模方面的研究; 最后介绍了一些其他与PROTAC一样基于邻近效应开发的新技术, 以及生物信息学和分子模拟技术未来的发展, 推动药物设计领域进一步发展。

2 生物信息学在拓展E3泛素连接酶和靶标空间中的研究

2.1 生物信息学在拓展E3泛素连接酶中的研究

细胞中的很多信号通路通过泛素-蛋白酶体途径选择性降解某些关键调控蛋白质, 这种对特定蛋白质高度特异的识别机制主要是由E3泛素连接酶决定的。E3泛素连接酶可以通过多种机制实现催化降解, 人类基因组中的600多个E3连接酶根据不同的催化机制主要分为三类(图2A)^[17]。最大的一类E3泛素连接酶是RING (really interesting new gene) 连接酶, 通过结合靶蛋白和E2-Ub将泛素转移到靶蛋白的赖氨酸残基上^[17]。第二类是HECT (homologous to E6-AP C-terminus) 连接酶, 通过结合E2-Ub将泛素转移到HECT结构域上, 催化结构域的半胱氨酸残基共价结合泛素并形成共价硫酯中间体, 最后将泛素转移到靶蛋白的赖氨酸残基上^[18]。第三类是RBR (RING-in-between-RING) 连接酶, 结合了RING和HECT连接酶的特征, 包含3个结构域: 与E2-Ub结合的RING1结构域、催化泛素转移的RING2结构域以及在两个RING结构域之间的IBR结构域。首先RING1识别E2-Ub, 然后泛素转移到RING2的催化半胱氨酸残基上形成硫酯中间体, 最终转移到靶蛋白的赖氨酸残基上^[19]。根据Ubihub数据库中收集的E3连接酶以系统发育树的形式展示了不同E3连接酶家族(图2B)^[20], 与上述三分类不同, 左边是单体形式的E3连接酶家族, 主要包括HECT、RBR、TRIM (tripartite motif)、TRAF (tumour necrosis factor receptor-associated factor) 等家族; 右边是依赖于多个亚基的复杂E3连接酶, 主要包括DCAF (DDB1-CUL4-associated factor)、SOCS (suppressor of cytokine

signaling)、BTB (broad-complex, tramtrack, and bric-a-brac)、FBXL (F-box with leucine rich amino acid repeats)、FBXW (F-box with WD 40 amino acid repeats)、FBXO (F-box only with uncharacterized domains) 等家族。

虽然E3连接酶有600多种, 但只有少数开发出相应的E3连接酶配体^[21]。E3连接酶配体是PROTAC降解效率和靶蛋白选择性及特异性的重要决定因素之一。目前用于设计PROTAC分子的E3连接酶仅限于VHL、CRBN、MDM2 (murine double minute 2) 及IAPs (inhibitor of apoptosis) 蛋白等E3连接酶。VHL、CRBN、MDM2及IAPs这四种E3连接酶都属于RING家族。目前进入临床的PROTAC分子大多是基于CRBN配体开发的, 少数是基于VHL配体开发的。严重依赖这两种E3连接酶使得PROTAC的临床开发面临挑战, 一旦发生突变, PROTAC耐药机制可能会迅速出现, 甚至可能会发生脱靶导致严重的不良反应^[22]。所以, 拓展可用于PROTAC的E3泛素连接酶是优化靶向蛋白质降解策略的必要途径。

拓展可用于PROTAC的E3连接酶(图2C) 首先可以寻找广泛适用的E3连接酶, 类似于CRBN和VHL, 可以与多数靶蛋白作用从而用于不同的疾病治疗。侯廷军教授课题组开发的PROTAC-DB数据库^[23]收集了目前PROTAC分子设计使用的E3连接酶。其次, 除了收集的E3连接酶, 研究人员使用如基于活性蛋白表达谱 (activity-based protein profiling, ABPP)^[24]的化学方法来扩展用于开发PROTAC的E3连接酶库, 开发了基于环指蛋白4 (ring finger protein 4, RNF4)、环指蛋白114 (ring finger protein 114, RNF114)、DCAF16 (DDB1-Cul4-associated factor 16) 及AhR (arylhydrocarbon receptor) 等E3连接酶配体的PROTAC^[25]。此外, 除了已经证明可以用于靶向蛋白质降解 (target protein degradation, TPD) 的E3连接酶之外, 部分E3连接酶没有可获取或者完整的晶体结构, 但是理论上也可用于开发PROTAC, 可以通过建模预测这些E3连接酶可能的结构模型。近年来, Google/DeepMind^[26]和RoseTTAFold^[27]等基于人工智能 (artificial intelligence, AI) 预测的三级结构模型质量极佳, 如DeepMind推出的AlphaFold数据库可以提供高质量的结构预测模型, 为许多迄今为止结构未知的靶蛋白和连接酶的药物发现提供了可能性^[21]。

尽管发现普遍适用的E3连接酶用于PROTAC分子设计是一种实用且具有价值的途径, 但是提高其选择性、特异性和避免不良反应是药物开发的主要目标。不同E3连接酶在降解方面的差异由几个因素决定, 包

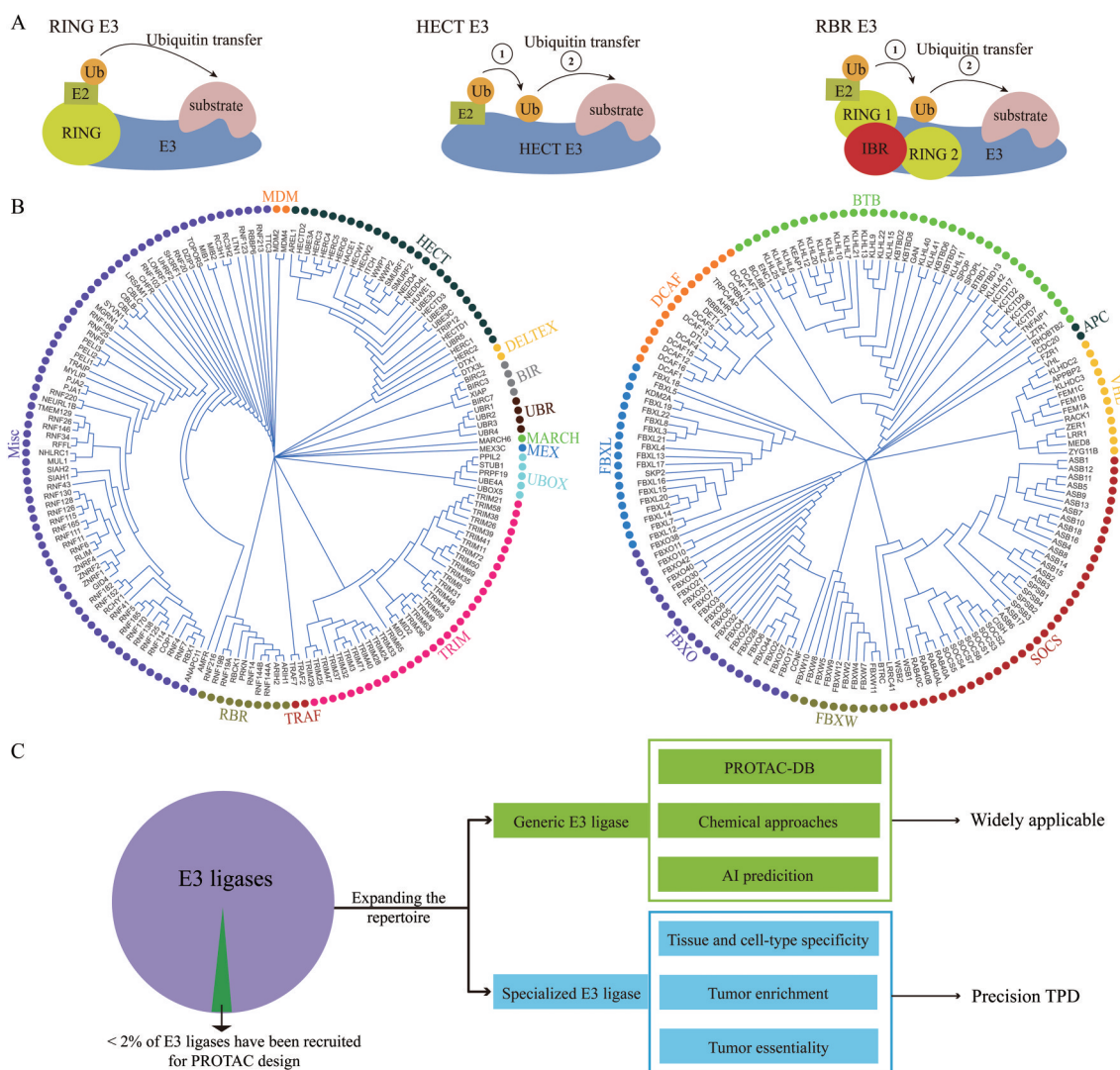


Figure 2 Expanding the arsenal of E3 ubiquitin ligases. A: Classes of E3 ubiquitin ligases; B: Phylogenetic trees of E3 ubiquitin ligases; C: Expanding the ligase landscape. TPD: Target protein degradation

括连接酶与靶蛋白之间的形状互补程度、连接酶和靶蛋白之间形成稳定三元复合物的能力、连接酶和靶蛋白的亚细胞定位差异以及连接酶和靶蛋白的细胞类型特异性表达谱^[21]。其中, E3连接酶的特征表达谱对于发现新的E3连接酶具有重要作用。鉴定具有独特表达谱的E3连接酶, 可以提高PROTAC的选择性并降低细胞毒性, 实现精确靶向蛋白质降解(图2C)。目前, GTEx数据库(genotype-tissue expression)、TCGA数据库(The Cancer Genome Atlas)及HPA数据库(Human Protein Atlas)等公开可用的数据库^[28]可以用于分析E3连接酶的表达来鉴定具有独特表达谱的E3连接酶^[12]。目前为止, 已经确定了几种具有组织特异性表达的E3连接酶, 如在骨骼肌中富集的KLHL40(Kelch-like family member 40)和KLHL41(Kelch-like family member 41)^[29]以及在中枢神经中富集的RNF182(ring

finger protein 182)^[30]等。此外, 还有一些E3连接酶在某些组织或细胞类型中表现出低表达的反向特异性。Bcl-XL(B-cell lymphoma extra large)是一种抗凋亡蛋白, 但是针对这个靶蛋白的药物可能会导致血小板水平降低。VHL在血小板中的表达较低, 通过使用PROTAC募集VHL连接酶, 可以靶向特定的肿瘤类型, 而不会出现血小板减少的不良反应^[31]。

除了分析E3连接酶的组织特异性, 研究其肿瘤特异性(tumour specificity)或肿瘤富集程度(tumor enrichment)有助于实现精准靶蛋白降解(图2C)。通常, E3连接酶的富集与肿瘤对该连接酶表达的依赖相一致。可以通过使用DepMap^[32](Cancer Dependency Map)开发的CERES算法(computational correction of copy-number effect in CRISPR-Cas9 essentiality screens)^[33]分析多个肿瘤细胞系的E3连接酶的依赖性

评分, 衡量 E3 连接酶对特定肿瘤细胞生长和存活的重要性^[34]。肿瘤细胞对基于重要的 E3 连接酶的 PROTAC 产生抗性的能力较低, 但基于这类 E3 连接酶开发的 PROTAC 的毒性还有待测试。此外, 癌症-睾丸抗原 (cancer-testis antigens, CTA) 也是 E3 连接酶的一种, CTA 在正常睾丸中表达受限, 但在多种癌症中高度过表达。黑色素瘤抗原 (melanoma antigen, MAGE) 家族^[35]通常被称为 MAGE-RING 连接酶 (MRLs), 参与组成作为 RING E3 连接酶的底物招募蛋白, 并以多亚基复合物的形式起作用^[36]。虽然不是所有的 MAGE E3 连接酶都是肿瘤特异性的, 但这种组织、疾病特异的 E3 连接酶是可以用于拓展 E3 连接酶空间研究的一个领域^[21]。

PROTAC 通过形成稳定的三元复合物发挥作用, 因此低结合亲和力的配体也可以发挥有效功能。由于钩效应取决于 E3 连接酶和靶蛋白的表达水平, 因此 PROTAC 在不同细胞类型和组织之间的降解效率会有所不同^[28]。所以特异的 E3 连接酶对于形成稳定的三元复合物并降低药物毒性具有重要意义, 可以靶向更具挑战性的靶蛋白。

2.2 生物信息学在拓展可降解靶标空间中的研究

PROTAC 技术已成功用于降解多种靶蛋白, 包括激酶、核受体、表观遗传蛋白及错误折叠蛋白等。越来越多的靶蛋白被证实“PROTACtable”, 即该蛋白可用于设计 PROTAC 进行靶向降解。进入临床的 PROTAC 靶蛋白包括雄激素受体 (androgen receptor, AR)、雌激素受体 (estrogen receptor, ER)、白细胞介素-1 受体相关激酶 4 (interleukin-1 receptor-associated kinase 4, IRAK4)、信号转导和转录激活因子 3 (signal transducer and activator of transcription 3, STAT3) 等^[6]。PROTAC-DB 数据库涵盖了 3 270 个 PROTAC 分子、365 个靶蛋白配体、82 个 E3 连接酶配体、1 501 个连接子及 664 个三元复合物模型的信息。PROTAC 不需要结合在靶蛋白具有生物学功能的活性位点, 扩大了靶

蛋白的空间, 被认为难以成药的靶标也可以通过 PROTAC 实现靶向降解。

据统计, 超过 80% 的人类蛋白质是“不可成药”的靶蛋白, 了解这些蛋白是否可以被 PROTAC 靶向降解 (PROTACtability) 是关键。Schneider 等^[37]从组学层面首次对潜在的药物靶标进行了系统全面的评估, 通过整合来自各种公开可用数据源的信息对可能作为药靶的蛋白进行分析, 收集了六个方面的信息: 临床开发 (clinical)、文献报道 (literature)、泛素化 (ubiquitination)、蛋白质半衰期 (turnover)、小分子配体 (small-molecule binder) 和定位 (location) (图 3)。

根据六个方面的信息系统地分析靶标蛋白, 并将靶标蛋白分为不同的类。第一方面, 检索 ClinicalTrials.gov 网站查看蛋白质是否是进入临床阶段的 PROTAC 药物靶点, 并将其分为三个类别: ① 进入 IV 期临床 PROTAC 的靶蛋白; ② 进入 II 期或者 III 期临床 PROTAC 的靶蛋白; ③ 进入 I 期临床 PROTAC 的靶蛋白。第二方面, 通过检索 PubMed 数据库查看靶蛋白是否已有文献报道的 PROTAC 方面的研究。第三方面, 通过收集蛋白质上的泛素化位点信息将靶蛋白分为两类: ① 具有 UniProt 数据库中提供的泛素化位点信息; ② 具有 PhosphoSitePlus 数据库、mUbiSiDa 数据库和基于蛋白质组学实验方法^[38]提供的泛素化位点。第四方面, 收集靶蛋白的半衰期数据, 半衰期非常短的蛋白质可能会限制 PROTAC 的降解效率。Mathieson 等^[39]使用基于细胞培养条件下稳定同位素标记 (stable isotope labeling with amino acids in cell culture, SILAC) 的蛋白质组学来确定各种细胞类型中蛋白质的半衰期范围。第五方面, 检索靶蛋白是否具有小分子配体。靶蛋白的小分子配体只需要适当的结合亲和力, 即使是瞬时的相互作用也足以发挥 PROTAC 降解功能。在 ChEMBL 数据库中报道的基于靶标的活性测定中, 小分子配体测得的活性至少为 $10 \mu\text{mol}\cdot\text{L}^{-1}$ 。第六方面, 通过检索 UniProt 数据库和 GO 数据库得到蛋白的位置数据, 为每个靶蛋白分

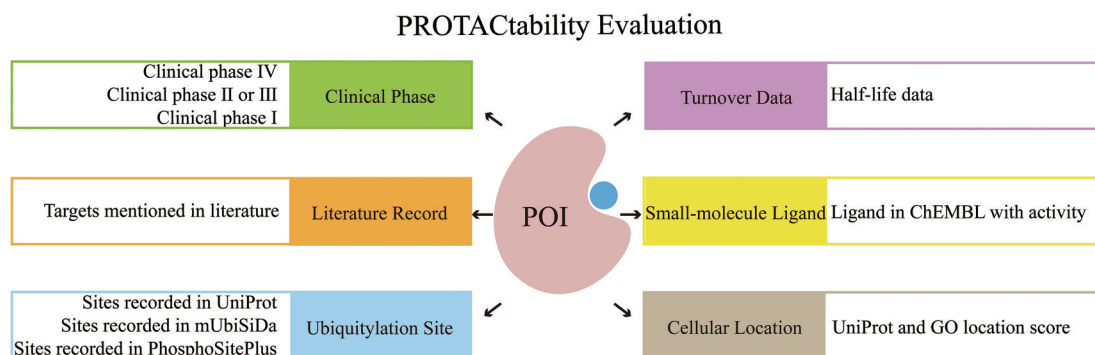


Figure 3 Expanding the PROTACtable target proteins

配一个位置分数。位于细胞质或细胞核中的蛋白质被认为是位于有利的位置,膜蛋白被认为是位于灰色的位置,其余被认为位于不利的位置,然后进一步根据置信度分类,得到位置分数。

根据以上针对靶蛋白的六类信息可以进一步将靶标简化为4个 PROTACTability 类别:临床优先级(存在临床数据)、文献优先级(有文献报道)、具有可被降解的机会(在具有泛素化位点、具有半衰期数据和有报道的小分子配体三者中至少满足一个,并且位置得分好)及不完整证据(能被降解的证据不足)。前三个类别的靶蛋白被认为是“PROTACTable”的靶蛋白,可以评估并量化潜在的可被降解的靶标蛋白质^[37]。

3 计算模拟方法在 PROTAC 三元复合物结构模型中的研究

3.1 构建 PROTAC 三元复合物结构模型的计算分析方法

“靶蛋白-PROTAC-E3 连接酶”三元复合物的形成是 PROTAC 诱导靶蛋白降解的关键,只有当 PROTAC

与靶蛋白和 E3 连接酶同时结合形成稳定的三元复合物时,才能有效且特异性地降解靶蛋白。然而,复杂的三元复合物晶体结构通常难以鉴定,近几十年来,生物物理方法为描述三元复合物的形成提供了有力的工具。为了更好地理解 PROTAC 三元复合物形成的结构基础,分子模拟可用于研究 PROTAC 对靶蛋白的选择性,并对降解活性提供分子水平的解释^[12]。目前用于构建 PROTAC 三元复合物的计算方法主要包括结构生成和分析(图4)^[13],结构生成主要包括蛋白质-蛋白质对接(protein-protein docking)、PROTAC 构象采样(PROTAC sampling)和分子动力学模拟(molecular dynamics);分析主要包括聚类(clustering)、MM/GBSA(molecular mechanics/generalized born surface area)和正则模式分析(normal mode analysis)等(表1)^[13,40]。

3.1.1 蛋白质-蛋白质对接方法 为了获得合理的三元复合物模型,首先需要获取蛋白质的结构信息,确定其是否有可获取的晶体结构以及是否结合配体。在确定了使用的蛋白结构后,对接靶蛋白和 E3 连接酶得到

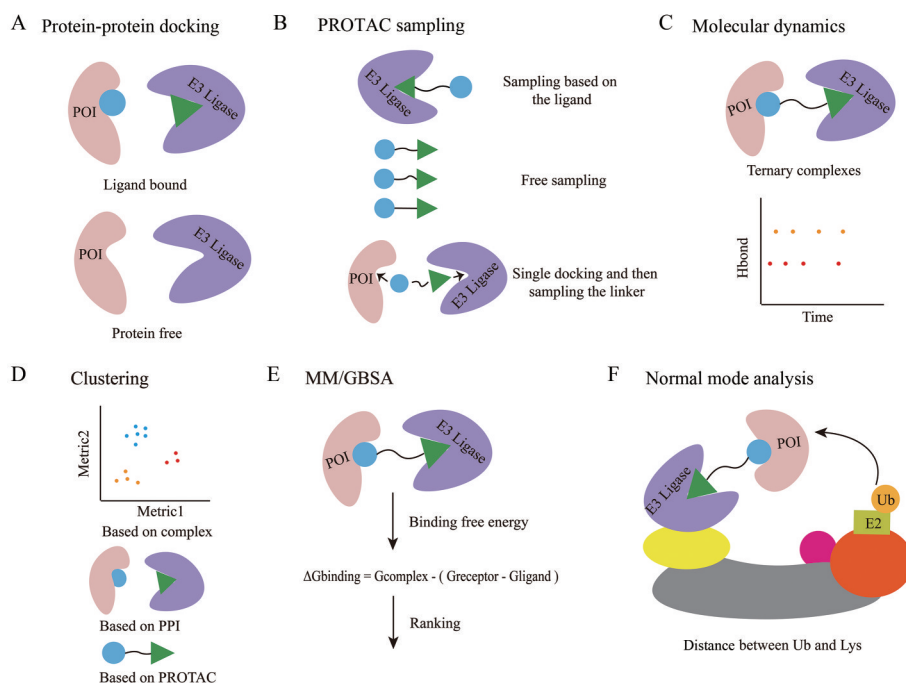


Figure 4 Computational methods applied in target protein degradation (TPD). A: Protein-protein docking; B: PROTAC conformational sampling; C: Molecular dynamics simulation; D: Clustering; E: MM/GBSA; F: Normal mode analysis

Table 1 Characteristics of the key computational methods applied in TPD

Model generation technique		Analysis technique	
Protein-protein docking	Accurate prediction of binary protein-protein complexes	Clustering	Categorize the conformational space to get representative poses
PROTAC conformational sampling	Sample PROTAC conformations restrained to the protein-protein complex	Estimating energies with MM/GBSA	Improve the ranking accuracy of ternary complexes and estimate the stability and cooperativity
Molecular dynamics (MD)	Obtain the conformational ensemble of ternary complex	Normal mode analysis	Describe the collective motion of the protein complexes

蛋白质相互作用的结构模型, 然后在此基础上生成 PROTAC 结合的三元复合物 (图 4A)。但这些蛋白相互作用结构模型并非都是合理的, 需要对其进行打分评估, 然后选择最佳的结构进行优化。Schiedel 等^[41]最先使用蛋白对接软件 HADDOCK^[42]用于 PROTAC 诱导靶蛋白降解的研究中。后来, 越来越多的蛋白质-蛋白质对接工具被开发用于得到合理的蛋白对接构象。近年来, 基于蒙特卡洛 (Monte Carlo) 算法的 RosettaDock^[43]已成为 TPD 中使用最广泛的蛋白质-蛋白质对接工具之一。Nowak 等^[16]使用 RosettaDock 对 CRBN 和溴结构域 BRD4 的结合域 1 (BD1) 进行对接产生了 20 000 个对接结构, 并在 200 个能量最低的构象中鉴定出了一种与复合物晶体结构非常相似的构象。并通过计算前 200 个构象中的 BRD4 配体 JQ1 和 CRBN 配体来那度胺的溶剂暴露原子对之间的最短距离设计了 PROTAC。其他常用的蛋白对接工具包括 ZDOCK^[44]、HDOCK^[45]、ClusPro^[46]、PatchDock^[47] 以及 LightDock^[48]等。

3.1.2 PROTAC 构象采样方法 蛋白质-蛋白质对接得到的是三元复合物中靶蛋白和 E3 连接酶之间可能的结合构象, 而实际上三元复合物通过 PROTAC 诱导形成。PROTAC 到底以何种构象结合到两个蛋白中需要对 PROTAC 在对应的蛋白对接构象中进行采样, 得到适合 PROTAC 结合的蛋白质-蛋白质结构。PROTAC 构象采样的方法包括: ① 可以在蛋白质本身结合的配体的基础上对 PROTAC 进行采样; ② 可以对整个 PROTAC 分子直接进行采样; ③ 可以分别对接蛋白和配体, 通过生成连接子得到完整的 PROTAC 分子 (图 4B)。Zaidman 等^[49]和 Bai 等^[50]都在蛋白质-蛋白质作用的基础上生成连接子的构象来得到 PROTAC 构象集。目前主要采用单独采样来获取 PROTAC 构象, 包括单独采样连接子和根据化学信息对整个 PROTAC 进行采样。PROTAC 对接可以在蛋白质结合位点周围产生几何形状并对其进行评分, 也可以使用对接评分函数来评估 PROTAC 构象的结合能。蛋白质-蛋白质对接与 PROTAC 构象采样方法的结合可预测潜在三元复合物的结构模型^[51]。目前常用的构象生成软件是 RDKit, 这是一种基于 Python 的开源工具。常用的软件还有 OMEGA^[52]和 MOE 等。

3.1.3 分子动力学模拟方法 蛋白质是动态的, 这一特性在预测 PROTAC 三元复合物时特别重要。分子对接只是从打分函数和构象空间采样的层面得到合理的结合方式确定其是否可以结合, 但无法判断 PROTAC 与靶蛋白和 E3 连接酶是否能稳定结合。此外, 分子对接无法捕获蛋白质-配体诱导的拟合效应或 apo (单

体)-holo (复合物) (结合前后) 状态构象变化^[53]。分子动力学 (molecular dynamics, MD) 模拟基于原子间相互作用的物理学的一般模型, 预测蛋白质或其他分子系统中每个原子如何随时间运动。MD 以飞秒时间分辨率揭示分子中所有原子的位置, 为重要的分子过程提供原子水平的解释^[54]。重要的是, 这种模拟还可以预测生物分子在原子水平上对突变、翻译后修饰、质子化或配体结合等扰动的反应机制^[55]。通过进行 MD 模拟观察 PROTAC 三元复合物的运动情况可以评估复合物的稳定性^[13]。Weerakoon 等^[56]使用 MD 模拟研究了 MZ1 和 dBET6 的构象变化, 发现 MD 结果产生的构象系综与核磁共振技术测量结果较为一致。常用的分子动力学模拟程序有 LAMMPS、AMBER、CHARMM、Tinker、NAMD、GROMACS 和 OpenMM 等。

3.1.4 聚类方法 蛋白质-蛋白质对接、PROTAC 构象采样和分子动力学模拟会产生广泛的靶蛋白-PROTAC-E3 连接酶三元复合物的构象空间, 需要对该构象空间进行进一步分析。在这种情况下, 聚类可以对 PROTAC 三元复合物的构象集合进行分组得到不同的聚类簇。同一聚类簇内 PROTAC 三元复合物构象相似, 不同聚类簇间构象差异较大, 可以选择不同聚类簇中的代表性构象开展进一步分析。这种聚类分组差异取决于使用的算法, 常用的聚类算法包括 K 均值聚类 (K-means clustering)、高斯混合模型 (Gaussian mixture module) 的期望最大化聚类、基于密度的噪声应用空间聚类、凝聚聚类和层次聚类 (hierarchical clustering) 等。可以根据目的选择不同的聚类方法, 一些计算建模管道在蛋白质-蛋白质对接方法后集成了聚类分析, 也有建模工具对三元复合物模型进行聚类从而得到较合理的复合物模型。Weerakoon 等^[56]在 MD 模拟结果的基础上进行了聚类, 并基于聚类绘制了网络分析三元复合物在不同状态之间的转变。因此, 聚类可以对构象空间进行分类从而从模拟中获得稳定的代表性构象。

3.1.5 MM/GBSA 方法 虽然聚类方法能够提供蛋白质构象空间的分组信息, 但是其他指标如能量对于评估三元复合物稳定性更加适用。在评估配体-受体相互作用的不同方法中, 结合自由能的计算起着重要作用。计算自由能的方法包括热力学积分 (thermodynamic integration, TI)、自由能微扰 (free energy perturbation, FEP)、分子力学/泊松-玻尔兹曼 (广义玻恩) 表面积 (molecular mechanics/Poisson Boltzmann (Generalized Born) Surface Area, MM/PB(GB)SA) 及线性相互作用能 (linear interaction energy, LIE)。其中, MM/GBSA 计算是估计小分子配体与生物大分子结合自由能的常用方法。通常基于 MD 模拟结果, 将结合自由能拆分成

分子力学项和溶剂化能分别计算。Liao等^[57]开发了一种基于MD的计算流程,将蛋白质-蛋白质对接与配体对接相结合,生成起始的三元复合物构象候选物。针对这些结构采用MD模拟得到稳定的三元复合物构象,对构象集合进行聚类并使用MM/GBSA进行能量打分,对得到的三元复合物构象进行排名。Li等^[40]提出了采用MD对PROsettaC产生的三元复合物进行重新排序的新策略。通过MM/GBSA计算评估PROTAC诱导的三元复合物的稳定性和协同性。由于蛋白质-蛋白质对接、PROTAC构象采样及MD模拟产生了大量构象,MM/GBSA通过计算能量衡量三元复合物组分的结合程度,评估三元复合物模型的合理性。

3.1.6 正则模式分析方法 正则模式分析(normal mode analysis, NMA)广泛用于预测蛋白质中的振动模式,而蛋白质的振动模式通常与生物学功能相关。低频率的运动模式对应于蛋白质的全局运动,而高频运动模式对应于蛋白质的局部构象改变。研究表明,这些低频运动模式对应于蛋白质中与功能相关的全局运动,并且构象转变遵循其中一种或几种正则模式^[58]。NMA通常与粗粒化模型结合起来用于简化计算^[59],如通过使用各向异性网络模型(anisotropic network model, ANM)将蛋白质结构简化为节点和边组成的弹性网络。网络中的每个节点表示一个氨基酸残基,边表示它们之间的作用。通过这种简化的弹性网络分析正则模式可以用来研究蛋白质的运动,获得残基全局运动方向的信息^[60]。ANM通过从最低频率的正则模式中提取信息,识别蛋白质中柔性较强的部分及其运动方向,预测蛋白质的构象改变。通过跟踪正则模式振幅随频率的变化,ANM可以对蛋白质的动态行为进行量化并跟踪大的构象变化。在研究PROTAC发挥作用的过程,ANM方法通过计算E2-Ub与靶蛋白上可及的赖氨酸残基之间的距离,分析这些数据与实验降解的相关性^[13]。

在构建靶蛋白-PROTAC-E3连接酶三元复合物的模拟计算中,这些建模和分析方法往往交叉使用。Drummond等^[61]提出了构建靶蛋白-PROTAC-E3连接酶三元复合物的四种方法包括:①在两种结合配体蛋白的基础上生成连接子作为PROTAC得到的三元复合物作为起始构象,然后对整个三元复合物进行采样;②独立取样PROTAC构象,然后添加到刚性蛋白质上;③在其中一种蛋白质的基础上对PROTAC进行采样,然后添加第二种蛋白质;④对PROTAC构象进行采样,但通过蛋白质-蛋白质对接添加可能的E3连接酶-靶蛋白作用结构。研究经验表明^[16,49,50],发现靶蛋白和E3连接酶之间有利的相互作用对于研究三元复

合物更有意义^[15]。对蛋白质-蛋白质和PROTAC或连接子构象单独采样是目前构建三元复合物模型的主要方法。

在三元复合物研究中,Schiedel等^[41]通过使用HADDOCK^[42]对接得到靶蛋白Sirt2与E3连接酶CRBN的结合构象,然后使用打分函数、范德华力和静电相互作用进一步进行评估复合物结构模型;对得到的构象进行聚类,选择得分最高的3个簇进一步分析;使用分子对接程序GOLD将PROTAC对接到Sirt2-CRBN复合物中得到对接构象,使用MM/GBSA对得到的构象分析,选择低能构象进行氢键分析。对接结果表明PROTAC与Sirt2-CRBN的结合模式与解析的结合小分子的Sirt2和CRBN晶体结构基本一致,得到了合理的Sirt2-PROTAC-CRBN三元复合物模型。Crews等^[62]采用分子对接得到p38 α -PROTAC-VHL三元复合物模型,通过120 ns的MD模拟并进行层次聚类得到代表性构象;然后进一步分析PROTAC的构象以及VHL与p38 α 之间的蛋白质-蛋白质相互作用研究哪些相互作用对于形成稳定的三元复合物是有利的。在Crews实验室的另一项研究^[63]中,在靶向降解p38 α 的研究基础上研究p38 MAPK家族另一个亚型p38 δ 的降解;同样通过分子对接得到相应的三元复合物模型并进行短时间MD模拟,通过分析两个三元复合物的结构差异来解释PROTAC分子针对不同p38亚型的选择性降解,指导PROTAC分子的设计和优化。

研究靶蛋白-PROTAC-E3连接酶三元复合物的实验方法主要包括表征三元复合物形成的生物物理方法和三元复合物结构测定方法。生物物理方法主要包括时间分辨荧光能量转移(time-resolved fluorescence energy transfer, TR-FRET)、均相光激化学发光免疫分析技术(amplified luminescent proximity homogeneous assay linked immunosorbent assay, AlphaLISA)、荧光偏振分析(fluorescence polarization, FP)、等温滴定量热法(isothermal titration calorimetry, ITC)和表面等离子体共振技术(surface plasmon resonance, SPR)等。通过TR-FRET、AlphaLISA、FP、ITC及SPR等体外的生物物理方法可以表征三元复合物形成过程中的亲和力、热力学和动力学等特征信息;在活细胞检测中,纳米生物发光能量共振转移(nanobioluminescent resonance energy transfer, NanoBRET)和基于荧光素酶的二亚单位系统(NanoLuc binary technology, NanoBiT)等技术可以对活细胞中三元复合物的形成进行动态监测;计算模拟上对三元复合物进行分子动力学模拟并结合MM/GBSA进行结合能的计算,能够预测PROTAC的结合能和形成三元复合物的协同性,并能够提高三元

复合物结构模型预测的精度。三元复合物结构测定的方法主要包括X射线晶体衍射技术(X-ray)、核磁共振技术(nuclear magnetic resonance, NMR)和电子显微镜(electron microscope, EM)等实验方法,计算方法流程中蛋白质-蛋白质对接、PROTAC构象采样、分子动力学模拟等步骤可以获得三元复合物结构的构象空间,通过对已知晶体结构的三元复合物进行结构建模,可以评估计算方法对三元复合物结构建模的合理性^[13]。

3.2 PROTAC三元复合物结构模型的计算流程

目前用于PROTAC三元复合物研究的计算方法主要有四种,一是基于对接得到三元复合物模型,如Drummond等^[61]提出的四种方法、Zaidman等^[49]提出的PRosettaC及侯廷军等^[51]提出的PROTAC-Model等;二是结合MD模拟研究三元复合物,如Li等^[40]提出的一种基于MD对PRosettaC产生的三元复合物进行重新排序的策略;三是基于泛素化模型研究三元复合物,如Bai等^[64]模拟了CRL4A连接酶复合物构象,根据泛素与靶蛋白上赖氨酸的距离将三元复合物分为活性(productive)和非活性(unproductive)复合物;四是

基于机器学习预测三元复合物,如利用机器学习实现对PROTAC小分子靶向降解能力的预测的Deep-PROTACs^[65]。其中,PRosettaC^[49]和PROTAC-Model^[51]整合了RosettaDock^[43],可在线预测三元复合物模型。

3.2.1 PRosettaC Zaidman等^[49]开发了一种基于分子建模软件Rosetta^[43]构建PROTAC介导的三元复合物的工具PRosettaC(图5A)。首先,使用RDKit随机生成PROTAC构象并给出蛋白结合配体两端之间的距离阈值得到距离分布;然后在这个距离阈值的限制上使用PatchDock^[47]软件对蛋白质对接得到构象空间进行过滤,可以大大减少对接的构象;然后对这些构象使用RosettaDock^[43]软件进行局部对接得到50个高分辨率的模型,在这些蛋白对接模型的基础上构建完整的PROTAC;得到PROTAC三元复合物模型并使用Rosetta对结构进行分析,过滤能量高的复合物选择得分最高的200个模型;最后对这些三元复合物进行聚类,得到的簇代表构象被认为是最接近晶体结构的复合物模型。该方法通过使用PROTAC的距离分布来限制蛋白质-蛋白质对接的搜索空间,然后使用能量优化及聚类方法来限制配体构象的搜索空间。

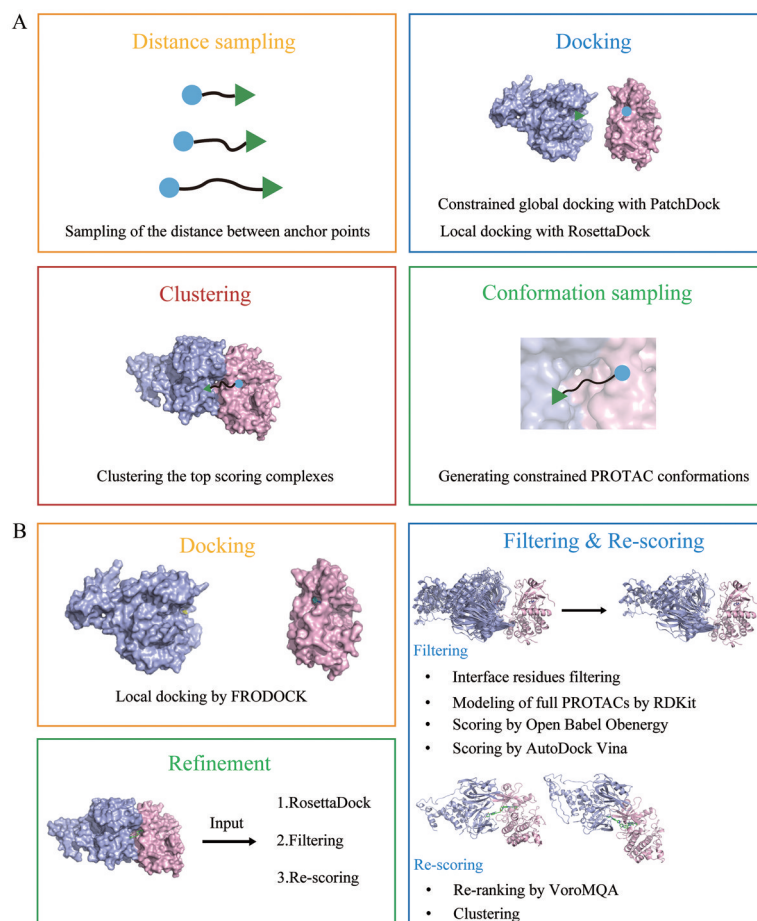


Figure 5 Modeling of PROTAC mediated ternary complexes. A: PRosettaC protocol; B: PROTAC-Model protocol

3.2.2 PROTAC-Model 侯廷军等提出整合了 FRODOCK^[66]、RosettaDock^[43]以及几种过滤和评分方法的 PROTAC-Model (图 5B)^[51]。使用 FRODOCK 软件^[66]进行局部蛋白质-蛋白质对接, 然后对生成的蛋白对接构象进行筛选, 首先根据界面残基数对接构象进行过滤; 然后使用 RDKit 对 PROTAC 构象进行建模, 去除不合理的对接模型; 利用 Open Babel Obenergy (Obenergy)^[67]工具计算了 PROTAC 构象的能量, 并去除了具有不利原子碰撞的模型; 使用 AutoDock Vina (Vina)^[68]评估 PROTAC 与蛋白质-蛋白质复合物的结合模式, 保留 Vina 能量评分小于 0 kcal·mol⁻¹ 的复合物; 筛选出的模型进一步由 VoromQA 方法^[69]重新排名; 最后, 采用聚类算法对这些模型进行聚类。在筛选构象过程中, 也可以选择 FRODOCK 提供的每个簇中的最佳模型输入 RosettaDock^[43]进行优化。然后以类似的方式对生成的模型进行筛选、重新排序和聚类。该计算流程将预测的具有良好降解能力的三元复合物模型整合到了更新的 PROTAC-DB 数据库中。

3.3 构建 PROTAC-靶标-E3-E2-泛素复合物结构模型

研究表明三元复合物的结合亲和力并不总与靶蛋白降解相关, 尽管靶蛋白-PROTAC-E3 连接酶三元复合物的形成是实现靶向降解的关键, 但三元复合物的形成并不总会导致 TPD, 这表明有其他因素可以驱动降解^[70]。靶蛋白泛素化是 PROTAC 诱导的 TPD 过程中三元复合物形成后的重要步骤。PROTAC 参与三酶级联反应, E2/Ub 结合 E3 使 Ub 转移到靶蛋白表面的赖氨酸残基上。有研究提出并非每个三元复合体构象都能被认为是可以诱导靶标泛素化的“活性”构象, 这可能取决于靶标上可接近的赖氨酸到 E2 上的 Ub 的方向或距离。Bai 等^[64]提出了一种基于结构的计算方法, 通过将 E2-E3 连接酶复合物结构与三元复合物叠合建模, 使用基于 CRBN 的 Cullin-Ring 泛素连接酶 4A (CRL4A) 复合物, 构建了三元复合物结合 CRL4A 的泛素化模型 (target/PROTAC/CRBN/DDB1/CUL4A/Rbx1/NEDD8/E2/Ub), 这些蛋白间的相互作用使 Ub 转移到靶蛋白上 (图 6)^[64]。该方法旨在从结构模型角度研究基于 CRBN 的 PROTAC 系统中泛素化三元复合物形成与赖氨酸邻近性之间的关系。

4 PROTAC 技术展望

4.1 PTM 介导的蛋白质稳定性技术

泛素化修饰是普遍存在的蛋白质翻译后修饰 (protein translational modifications, PTMs) 的一种重要类型。作为蛋白质调控功能的重要方式, PTM 能够调控蛋白质的稳定性^[71]。泛素化通过泛素-蛋白酶体系

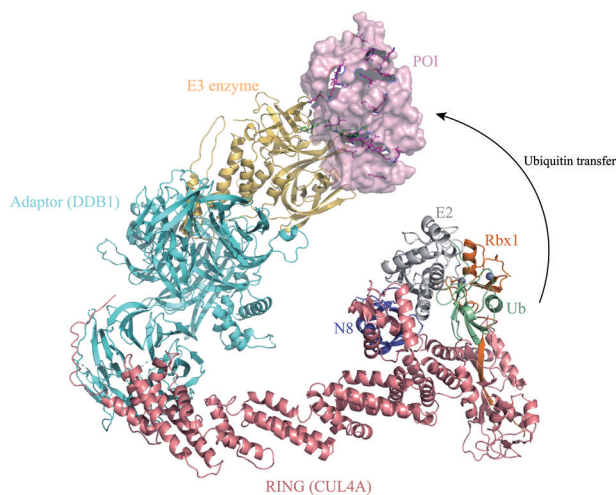


Figure 6 CRL4A ligase complex model

统参与调节蛋白质的稳定性, 对体内 80% 的蛋白质进行降解。此外, 几种常见的 PTM 也会参与调节蛋白质稳定性的途径, 如甲基化、磷酸化及乙酰化等, 这些修饰由相应的修饰书写酶 (writer enzyme) 和擦除酶 (eraser enzyme) 控制底物蛋白质的修饰水平。目前检测到的 PTM 数量远远超过已有功能报道的 PTM 位点数目, 因此解释功能未知的 PTM 位点是当前研究的一大挑战^[72]。随着精准医疗的快速发展, 修饰组学的研究也逐渐深入, 如磷酸化蛋白质组学通过对磷酸化的蛋白质进行全面分析, 可以对磷酸化修饰进行定性、定位、定量及功能分析^[73]。所以在研究 PTM 驱动的蛋白质稳定性调控中, 可以借助功能 PTM 蛋白质组学来系统地评估 PTM 对蛋白质稳定性的影响。

新型蛋白质组学技术的应用有望大大加快识别和阐明调控蛋白质稳定性的 PTM 位点。目前主要使用基于高通量质谱 (mass spectrometry, MS) 的定量蛋白质组学和细胞培养条件下稳定同位素标记技术 (SILAC), 研究哪些 PTM 会对蛋白质的稳定性产生影响, 从而对功能 PTM 位点进行快速筛选。另一方面, 通过干扰 PTM 书写酶或擦除酶的功能可以系统地测量细胞受到扰动后蛋白质稳定性的变化。将 SILAC 技术和串联质谱标签 (tandem mass tag, TMT) 结合, 这种技术称为多重蛋白质组动力学分析 (multiplexed proteome dynamics profiling, mPDP), 可以精确评估和比较多种扰动的影响, 揭示在不同扰动条件下蛋白质的稳定性变化。蛋白质的稳定性可以受单独 PTM 位点调控, 也可以受多种 PTM 共同调控。单个 PTM 作为 PTM 调控网络中的一部分, 这种 PTM 网络是通过蛋白质-蛋白质相互作用形成的, 一个 PTM 可以通过网络相互作用作为下一个 PTM 的启动子^[74]。这些 PTM 参与调控蛋白质稳定性的方式揭示了联合使用

擦除酶和书写酶的抑制剂来靶向 PTM 修饰的蛋白质可以成为一种潜在的 药物设计策略^[75]。通过靶向参与共同调控蛋白质稳定性的 PTM 来拓展潜在的 药物靶标空间。

4.2 PTM 介导的邻近诱导双分子技术 (CIP)

近年来, TPD 这一研究领域迅速发展, 随着 PROTAC 进入临床研究, 药物开发进入异型双功能药物 (heterobifunctional drugs) 阶段。PROTAC 设计理念成为异型双功能分子设计策略的先驱。这一策略利用邻近效应诱导治疗 (proximity-inducing therapeutic), 通过合理设计的连接子连接招募分子与靶标结合分子, 使靶蛋白与招募分子募集来的功能蛋白质靠近, 从而起到理想的生物学效应^[76]。根据募集的功能蛋白质的分类可以扩展异型双功能分子的类别, 异型双功能分子可以通过招募其上游功能蛋白质作为效应子 (如 E3 泛素连接酶、内体/溶酶体、核糖核酸酶 L、蛋白磷酸酶/激酶、乙酰转移酶等) 来调节靶蛋白的功能, 使效应子和靶蛋白靠近以加速它们的相互作用。通过这种化学诱导接近 (chemically induced proximity, CIP) 技术, 越来越多的异型双功能分子被设计用于靶向不可成药的蛋白^[10]。在靶向蛋白质降解中, 除了泛素-蛋白酶体系统, 溶酶体途径也是一种重要的蛋白质降解途径, 包括自噬-溶酶体途径和内吞-溶酶体途径。其中, 自噬靶向嵌合体 (autophagy-targeting chimera, AUTAC) 及自噬小体绑定化合物 (autophagosome-tethering compounds, ATTEC) 等技术通过自噬-溶酶体途径降解靶蛋白; 溶酶体靶向嵌合体 (lysosome targeting chimera, LYTAC) 技术通过内吞-溶酶体途径降解靶蛋白 (图 7A)^[77]。这些方法可以选择性地识别蛋白质/细胞器并将其转运到溶酶体中进行降解。自噬系统在溶酶体介导的细胞内物质降解中起主要作用, 使受损的细胞器、细胞内碎片等底物降解。内吞系统主要针对细胞外和膜相关的靶蛋白。基于溶酶体的靶蛋白降解可以克服蛋白酶体介导的降解的局限性。然而, 对这些方法的理解和大规模应用仍处于起步阶段。

除了靶向蛋白质降解外, 针对 PTM 的异型双功能分子也可以通过调节靶蛋白上的 PTM 来改变蛋白质的功能。这种异型双功能分子招募调节 PTM 的功能蛋白质使其与靶蛋白靠近发挥作用, 如磷酸化诱导嵌合小分子 (phosphorylation-inducing chimera, PHICS) (图 7B)、蛋白磷酸酶募集嵌合体 (phosphatase-recruiting chimera, PHORC/phosphorylation targeting chimeras, PhosTAC) (图 7B) 及乙酰转移酶募集嵌合体 (acetylation-tagging-system, AceTAG)^[76]。这些异型双功能分子通过改变靶蛋白的 PTM 状态而不是其表达水平来

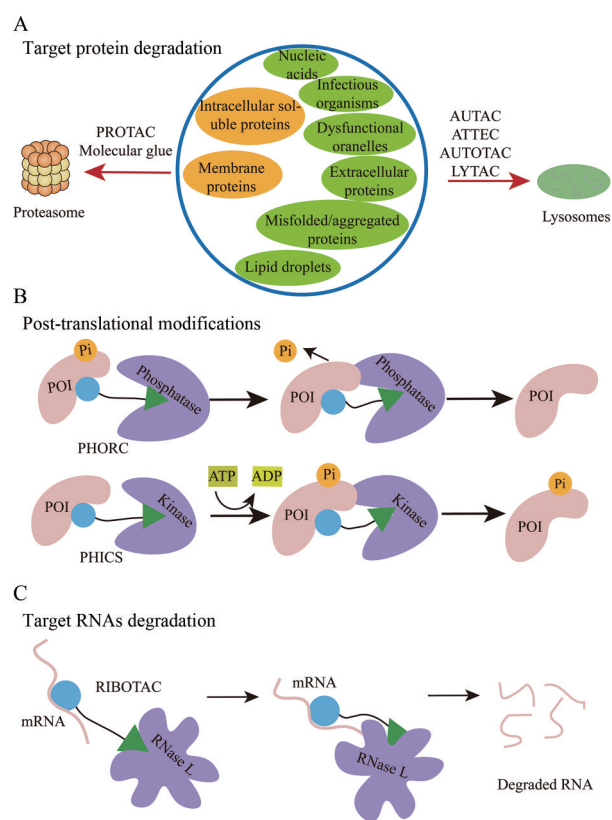


Figure 7 Different bifunctional molecules based on chemically induced proximity (CIP). A: Target protein degradation; B: Mechanism of PHORC and PHICS; C: Target RNAs degradation

精确调节其功能。此外, 除了直接靶向蛋白质的 PTM, 目前也有策略来靶向翻译前机制, 例如靶向 RNA 降解、RNA 干扰和基因编辑等方法^[71]。核糖核酸酶嵌合体 (RIBonuclease TArgeting Chimera, RIBOTAC) 是一类降解 RNA 的异型双功能分子技术^[78], 由 RNA 靶向配体、招募核糖核酸酶 (RNase L) 配体部分以及连接子组成 (图 7C)。RIBOTAC 将内源性核糖核酸酶 L 募集到特定 RNA 上并激活核糖核酸酶, 通过诱导邻近降解靶 RNA^[79]。然而, 这种方法最显著的局限性是其低细胞渗透率^[11]。此外, 高选择性的 RNA 小分子配体设计难度大, 易脱靶, 并且 RIBOTACs 不适用于在细胞核中发挥正常功能的 RNAs, 因为其主要在细胞质中起作用^[80]。

异型双功能分子处于小分子药物设计新趋势的最前沿, 通过将不同的效应子募集到靶蛋白用于调控不同的生物学过程。然而, 关于异型双功能分子的一些关键问题仍然需要解决, 包括钩效应、可控性、稳定性及脱靶效应等。通过开发招募更多内源效应子的异双功能分子, 可以极大地丰富靶向调节机制。在未来几年, 优化异双功能分子以克服其成药性是一项重大的研究挑战^[76]。

4.3 生物信息学及计算模拟技术在 PROTAC 发展中的应用

在人类基因组的 600 多种 E3 连接酶中, 只有不到 2% 的 E3 连接酶参与靶向蛋白降解的研究。已有关于靶蛋白的系统分析用来评估其 PROTAC stability^[37], 然而 E3 连接酶空间尚未被系统地定量表征, 招募更多潜在的 E3 连接酶对于实现精确靶向蛋白质降解有重要意义。除了 2.1 部分阐述的根据 E3 连接酶表达谱、结构可用性和功能必需性来招募更多潜在的 E3 连接酶, 近期, Liu 等^[22]为了更全面地分析 E3 连接酶, 通过整合 Ge 等^[81]、Ubihub^[20]和 UbiBrowser2.0^[82]的数据获取了全面的 E3 连接酶列表, 进一步提出还可以根据 E3 连接酶的化学配体、蛋白质-蛋白质相互作用 (PPI)、PPI 界面分析和细胞位置来拓展 E3 连接酶空间。通过数据库检索收集 E3 连接酶配体数据, 从中收集至少与一类配体结合的 E3 连接酶; 并通过基于深度学习的虚拟筛选模型预测 E3 与药物之间的相互作用, 揭示潜在的 E3 连接酶; 根据四种类型来源收集 E3 连接酶和靶蛋白的互作构建其 PPI 的综合图谱, 根据全面的 E3 酶和靶蛋白的 PPI 信息扩展 E3 连接酶的空间; 界面残基突变可能会破坏 E3 连接酶和靶蛋白的相互作用, 通过收集 PPI 界面信息来分析突变是否位于界面处, 这有助于快速识别针对特定靶蛋白可能有降解活性的 E3 连接酶; 通过数据库检索确定 E3 连接酶的细胞位置, 其中细胞质和细胞核被认为是设计 PROTAC 的有利位置, 通过注释细胞位置可以发现更多潜在的新型 E3 连接酶。尽管目前的研究有局限性, 招募更多潜在的 E3 连接酶对于扩大靶蛋白空间、提高选择性降解及指导 PROTAC 设计等方面有重要意义。

在 PROTAC 三元复合物结构模型构建方面, 虽然近年来已经报道了几种三元复合物的建模方法, 但三元复合物的结合并不能通过单独计算 PROTAC 与每种蛋白质之间的相互作用能来描述^[40]。其中, PROTAC 介导的三元复合物的结合稳定性和协同作用是决定 PROTAC 降解效率的关键因素。钩效应的出现不利于三元复合物稳定存在, 较高的协同性被认为与弱的钩效应相关, 因此协同性的预测对 PROTAC 的设计与优化至关重要。Li 等^[40]通过分子动力学模拟得到 PROTAC 三元复合物的低能结构, 结合 MM/GBSA 计算结合能并预测协同性; 然后在 PROsettaC 的基础上对三元复合物构象进行新的打分排序, 提高了三元复合物构象预测的准确度。

另一方面, 随着人工智能方法的蓬勃发展, 越来越多的研究开始结合深度学习方法来预测三元复合物的结构模型, 对 PROTAC 的降解效率进行预测并指

导 PROTAC 设计。Zheng 等^[83]提出了一种深度生成模型 (PROTAC-RL) 对给定的靶蛋白和 E3 配体开展 PROTAC 设计并进行采样, 通过强化学习生成具有最佳药代动力学的 PROTAC, 并使用 PROsettaC^[49]和分子动力学模拟开展构象的聚类和筛选分析, 这种深度学习和分子模拟的结合可以促进 PROTAC 的合理设计和优化。由于当前三元复合物建模方法的局限性, Li 等^[65]规避了三元复合物的建模过程, 从确定的蛋白-配体结构中提取出五个部分, 即靶蛋白口袋、E3 连接酶口袋、靶蛋白配体、E3 连接酶配体以及连接子, 并使用五个模块分别提取相应特征; 提出了一个以图神经网络为基础的深度学习模型 DeepPROTACs, 来预测设计出的 PROTACs 对于靶蛋白的降解功效。

近期研究^[15]提出了一种生成 PROTAC 三元复合物结构模型的新方法。考虑蛋白构象和 linker 本身几何形状的合理性, 将基于配体生成的 linker 一分为二分别连接到 E3 连接酶-配体复合物和靶蛋白-配体复合物上; 然后基于傅里叶快速采样算法生成蛋白质-蛋白质复合物构象, 得到靶蛋白和 E3 连接酶的低能构象和潜在合理的 linker, 并通过能量最小化、聚类和排名进行优化; 通过对已发表的三元复合物结构进行测试, 该模型可以高精度地预测三元复合物结构; 并且根据 Bai 等^[50]的方法预测三元复合物的协同作用和降解效率。该方法将蛋白结构纳入深度学习的模型中, 同时得到三元复合物结构及 PROTAC 的降解效率。将生物信息学、计算模拟与 AI 结合的方法在预测三元复合物的形成、降解效率预测以及指导 PROTAC 设计方面具有巨大的应用前景。

5 小结

PROTAC 技术利用泛素-蛋白酶体系统对靶蛋白进行泛素化降解, 与其他小分子药物和大分子药物相比, PROTAC 展现出了独特的优势。与此同时, 该技术同样面临着诸多挑战, 尤其在提高选择性和特异性方面, 很大程度上取决于 E3 泛素连接酶的选择。目前真正可使用的 E3 连接酶很少, 因此拓展 E3 泛素连接酶空间对合理设计 PROTAC 以提高选择性非常重要。此外, PROTAC 无需与蛋白质紧密结合, 实现降解的关键是形成稳定的三元复合物, 高效的选择性降解一定程度上依赖于 PROTAC 诱导的低能结合构象蛋白质间接接触的可塑性。因此, 从结构上研究三元复合物的形成对于理解选择性降解具有重要意义。然而目前报道的三元复合物的晶体结构很少, 而分子模拟有助于了解 PROTAC 如何诱导靶蛋白和 E3 连接酶作用形成三元复合物, 进而指导 PROTAC 设计并深化对 PROTAC 构效关系的理解。随着生物信息学和计算模拟方法的

快速发展, 针对三元复合物晶体结构数量少且建模方法准确度有限的情况下, 许多研究通过结合深度学习及结构模拟开展三元复合物模型的构建, 并对 PROTAC 的降解效率进行预测, 对指导 PROTAC 的设计起到了积极的推动作用。

经过多年的研究与发展, PROTAC 领域已衍生出了许多基于化学诱导接近理论的新技术, 基于 CIP 理论的药物设计领域已经进入了异型双功能分子阶段。总体而言, PROTAC 在药物发现方面的未来前景是巨大的, 并将为扩展可成药蛋白质组开启新的可能。

作者贡献: 朱婕负责文章撰写和绘制插图; 聂锦宜负责文字修改; 刘小益负责部分文献检索及文字修改; 梁中洁负责文章整体设计、修改、审查与指导。

利益冲突: 本文不存在利益冲突。

References

- [1] Sakamoto KM, Kim KB, Kumagai A, et al. Protacs: chimeric molecules that target proteins to the Skp1-Cullin-F box complex for ubiquitination and degradation [J]. *Proc Natl Acad Sci U S A*, 2001, 98: 8554-8559.
- [2] Sakamoto KM, Kim KB, Verma R, et al. Development of PROTACs to target cancer-promoting proteins for ubiquitination and degradation [J]. *Mol Cell Proteomics*, 2003, 2: 1350-1358.
- [3] Yao T, Xiao H, Wang H, et al. Recent advances in PROTACs for drug targeted protein research [J]. *Int J Mol Sci*, 2022, 23: 10328.
- [4] Sheng XY, Wu SH, Li BL, et al. Advances in the optimization of the linker in proteolysis-targeting chimeras (PROTAC) [J]. *Acta Pharm Sin (药学报)*, 2021, 56: 445-455.
- [5] Xie BH, Hu ZY, Ning WT, et al. The research progress of PROTACs for breast cancer treatment [J]. *Acta Pharm Sin (药学报)*, 2020, 55: 2053-2061.
- [6] Zeng S, Huang W, Zheng X, et al. Proteolysis targeting chimera (PROTAC) in drug discovery paradigm: recent progress and future challenges [J]. *Eur J Med Chem*, 2021, 210: 112981.
- [7] Xu YR, Zhang QS, Wu JY, et al. Recent progress in targeting degradation of FAK based on PROTAC [J]. *Acta Pharm Sin (药学报)*, 2021, 56: 1571-1579.
- [8] Maneiro MA, Forte N, Shchepinova MM, et al. Antibody-PROTAC conjugates enable HER2-dependent targeted protein degradation of BRD4 [J]. *ACS Chem Biol*, 2020, 15: 1306-1312.
- [9] Imaide S, Riching KM, Makukhin N, et al. Trivalent PROTACs enhance protein degradation *via* combined avidity and cooperativity [J]. *Nat Chem Biol*, 2021, 17: 1157-1167.
- [10] Li K, Crews CM. PROTACs: past, present and future [J]. *Chem Soc Rev*, 2022, 51: 5214-5236.
- [11] Wang C, Zheng C, Wang H, et al. The state of the art of PROTAC technologies for drug discovery [J]. *Eur J Med Chem*, 2022, 235: 114290.
- [12] Guenette RG, Yang SW, Min J, et al. Target and tissue selectivity of PROTAC degraders [J]. *Chem Soc Rev*, 2022, 51: 5740-5756.
- [13] Ward JA, Perez-Lopez C, Mayor-Ruiz C. Biophysical and computational approaches to study ternary complexes: a 'cooperative relationship' to rationalize targeted protein degradation [J]. *Chembiochem*, 2023, 24: e202300163.
- [14] Gadd MS, Testa A, Lucas X, et al. Structural basis of PROTAC cooperative recognition for selective protein degradation [J]. *Nat Chem Biol*, 2017, 13: 514-521.
- [15] Ignatov M, Jindal A, Kotelnikov S, et al. High accuracy prediction of PROTAC complex structures [J]. *J Am Chem Soc*, 2023, 145: 7123-7135.
- [16] Nowak RP, Deangelo SL, Buckley D, et al. Plasticity in binding confers selectivity in ligand-induced protein degradation [J]. *Nat Chem Biol*, 2018, 14: 706-714.
- [17] Ramachandran S, Ciulli A. Building ubiquitination machineries: E3 ligase multi-subunit assembly and substrate targeting by PROTACs and molecular glues [J]. *Curr Opin Struct Biol*, 2021, 67: 110-119.
- [18] Bernassola F, Chillemi G, Melino G. HECT-type E3 ubiquitin ligases in cancer [J]. *Trends Biochem Sci*, 2019, 44: 1057-1075.
- [19] Wang P, Dai X, Jiang W, et al. RBR E3 ubiquitin ligases in tumorigenesis [J]. *Semin Cancer Biol*, 2020, 67: 131-144.
- [20] Liu L, Damerell DR, Koukouflis L, et al. UbiHub: a data hub for the explorers of ubiquitination pathways [J]. *Bioinformatics*, 2019, 35: 2882-2884.
- [21] Bekes M, Langley DR, Crews CM. PROTAC targeted protein degraders: the past is prologue [J]. *Nat Rev Drug Discov*, 2022, 21: 181-200.
- [22] Liu Y, Yang J, Wang T, et al. Expanding PROTACtable genome universe of E3 ligases [J]. *Nat Commun*, 2023, 14: 6509.
- [23] Weng G, Shen C, Cao D, et al. PROTAC-DB: an online database of PROTACs [J]. *Nucleic Acids Res*, 2021, 49: D1381-D1387.
- [24] Ward CC, Kleinman JI, Brittain SM, et al. Covalent ligand screening uncovers a RNF4 E3 ligase recruiter for targeted protein degradation applications [J]. *ACS Chem Biol*, 2019, 14: 2430-2440.
- [25] Ishida T, Ciulli A. E3 ligase ligands for PROTACs: how they were found and how to discover new ones [J]. *SLAS Discov*, 2021, 26: 484-502.
- [26] Tunyasuvunakool K, Adler J, Wu Z, et al. Highly accurate protein structure prediction for the human proteome [J]. *Nature*, 2021, 596: 590-596.
- [27] Baek M, Dimaio F, Anishchenko I, et al. Accurate prediction of protein structures and interactions using a three-track neural network [J]. *Science*, 2021, 373: 871-876.
- [28] Kannt A, Dikic I. Expanding the arsenal of E3 ubiquitin ligases for proximity-induced protein degradation [J]. *Cell Chem Biol*, 2021, 28: 1014-1031.

- [29] Ehrlich KC, Baribault C, Ehrlich M. Epigenetics of muscle- and brain-specific expression of KLHL family genes [J]. *Int J Mol Sci*, 2020, 21: 8394.
- [30] Liu QY, Lei JX, Sikorska M, et al. A novel brain-enriched E3 ubiquitin ligase RNF182 is up regulated in the brains of Alzheimer's patients and targets ATP6V0C for degradation [J]. *Mol Neurodegener*, 2008, 3: 4.
- [31] Khan S, Zhang X, Lv D, et al. A selective BCL-X(L) PROTAC degrader achieves safe and potent antitumor activity [J]. *Nat Med*, 2019, 25: 1938-1947.
- [32] Pacini C, Dempster JM, Boyle I, et al. Integrated cross-study datasets of genetic dependencies in cancer [J]. *Nat Commun*, 2021, 12: 1661.
- [33] Meyers RM, Bryan JG, Mcfarland JM, et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells [J]. *Nat Genet*, 2017, 49: 1779-1784.
- [34] Nieto-Jimenez C, Morafraila EC, Alonso-Moreno C, et al. Clinical considerations for the design of PROTACs in cancer [J]. *Mol Cancer*, 2022, 21: 67.
- [35] Weon JL, Potts PR. The MAGE protein family and cancer [J]. *Curr Opin Cell Biol*, 2015, 37: 1-8.
- [36] Feng Y, Gao J, Yang M. When MAGE meets RING: insights into biological functions of MAGE proteins [J]. *Protein Cell*, 2011, 2: 7-12.
- [37] Schneider M, Radoux CJ, Hercules A, et al. The PROTACtable genome [J]. *Nat Rev Drug Discov*, 2021, 20: 789-797.
- [38] Kim W, Bennett EJ, Huttlin EL, et al. Systematic and quantitative assessment of the ubiquitin-modified proteome [J]. *Mol Cell*, 2011, 44: 325-340.
- [39] Mathieson T, Franken H, Kosinski J, et al. Systematic analysis of protein turnover in primary cells [J]. *Nat Commun*, 2018, 9: 689.
- [40] Li W, Zhang J, Guo L, et al. Importance of three-body problems and protein-protein interactions in proteolysis-targeting chimera modeling: insights from molecular dynamics simulations [J]. *J Chem Inf Model*, 2022, 62: 523-532.
- [41] Schiedel M, Herp D, Hammelmann S, et al. Chemically induced degradation of sirtuin 2 (Sirt2) by a proteolysis targeting chimera (PROTAC) based on sirtuin rearranging ligands (SirReals) [J]. *J Med Chem*, 2018, 61: 482-491.
- [42] Dominguez C, Boelens R, Bonvin AM. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information [J]. *J Am Chem Soc*, 2003, 125: 1731-1737.
- [43] Gray JJ, Moughon S, Wang C, et al. Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations [J]. *J Mol Biol*, 2003, 331: 281-299.
- [44] Pierce BG, Wiehe K, Hwang H, et al. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers [J]. *Bioinformatics*, 2014, 30: 1771-1773.
- [45] Yan Y, Zhang D, Zhou P, et al. HDock: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy [J]. *Nucleic Acids Res*, 2017, 45: W365-W373.
- [46] Kozakov D, Hall DR, Xia B, et al. The ClusPro web server for protein-protein docking [J]. *Nat Protoc*, 2017, 12: 255-278.
- [47] Schneidman-Duhovny D, Inbar Y, Nussinov R, et al. PatchDock and SymmDock: servers for rigid and symmetric docking [J]. *Nucleic Acids Res*, 2005, 33: W363-W367.
- [48] Jimenez-Garcia B, Roel-Touris J, Romero-Durana M, et al. LightDock: a new multi-scale approach to protein-protein docking [J]. *Bioinformatics*, 2018, 34: 49-55.
- [49] Zaidman D, Prilusky J, London N. PROsettaC: Rosetta based modeling of PROTAC mediated ternary complexes [J]. *J Chem Inf Model*, 2020, 60: 4894-4903.
- [50] Bai N, Miller SA, Andrianov GV, et al. Rationalizing PROTAC-mediated ternary complex formation using Rosetta [J]. *J Chem Inf Model*, 2021, 61: 1368-1382.
- [51] Weng G, Li D, Kang Y, et al. Integrative modeling of PROTAC-mediated ternary complexes [J]. *J Med Chem*, 2021, 64: 16271-16281.
- [52] Hawkins PC, Skillman AG, Warren GL, et al. Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database [J]. *J Chem Inf Model*, 2010, 50: 572-584.
- [53] Saikia S, Bordoloi M. Molecular docking: challenges, advances and its use in drug discovery perspective [J]. *Curr Drug Targets*, 2019, 20: 501-521.
- [54] Salmaso V, Moro S. Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: an overview [J]. *Front Pharmacol*, 2018, 9: 923.
- [55] Hollingsworth SA, Dror RO. Molecular dynamics simulation for all [J]. *Neuron*, 2018, 99: 1129-1143.
- [56] Weerakoon D, Carbajo RJ, De Maria L, et al. Impact of PROTAC linker plasticity on the solution conformations and dissociation of the ternary complex [J]. *J Chem Inf Model*, 2022, 62: 340-349.
- [57] Liao J, Nie X, Unarta IC, et al. *In silico* modeling and scoring of PROTAC-mediated ternary complex poses [J]. *J Med Chem*, 2022, 65: 6116-6132.
- [58] Skjaerven L, Martinez A, Reuter N. Principal component and normal mode analysis of proteins; a quantitative comparison using the GroEL subunit [J]. *Proteins*, 2011, 79: 232-243.
- [59] Bauer JA, Pavlovic J, Bauerova-Hlinkova V. Normal mode analysis as a routine part of a structural investigation [J]. *Molecules*, 2019, 24: 3293.
- [60] Alexandrov V, Lehnert U, Echols N, et al. Normal modes for predicting protein motions: a comprehensive database assessment and associated web tool [J]. *Protein Sci*, 2005, 14: 633-643.
- [61] Drummond ML, Williams CI. *In silico* modeling of PROTAC-mediated ternary complexes: validation and application [J]. *J Chem Inf Model*, 2019, 59: 1634-1644.

- [62] Bondeson DP, Smith BE, Burslem GM, et al. Lessons in PROTAC design from selective degradation with a promiscuous warhead [J]. *Cell Chem Biol*, 2018, 25: 78-87.
- [63] Smith BE, Wang SL, Jaime-Figueroa S, et al. Differential PROTAC substrate specificity dictated by orientation of recruited E3 ligase [J]. *Nat Commun*, 2019, 10: 131.
- [64] Bai N, Ricking KM, Makaju A, et al. Modeling the CRL4A ligase complex to predict target protein ubiquitination induced by cereblon-recruiting PROTACs [J]. *J Biol Chem*, 2022, 298: 101653.
- [65] Li F, Hu Q, Zhang X, et al. DeepPROTACs is a deep learning-based targeted degradation predictor for PROTACs [J]. *Nat Commun*, 2022, 13: 7133.
- [66] Garzon JI, Lopez-Blanco JR, Pons C, et al. FRODOCK: a new approach for fast rotational protein-protein docking [J]. *Bioinformatics*, 2009, 25: 2544-2551.
- [67] O'Boyle NM, Banck M, James CA, et al. Open babel: an open chemical toolbox [J]. *J Cheminform*, 2011, 3: 33.
- [68] Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading [J]. *J Comput Chem*, 2010, 31: 455-461.
- [69] Olechnovic K, Venclovas C. VoroMQA: assessment of protein structure quality using interatomic contact areas [J]. *Proteins*, 2017, 85: 1131-1145.
- [70] Han B. A suite of mathematical solutions to describe ternary complex formation and their application to targeted protein degradation by heterobifunctional ligands [J]. *J Biol Chem*, 2020, 295: 15280-15291.
- [71] Amirian R, Azadi Badrbani M, Izadi Z, et al. Targeted protein modification as a paradigm shift in drug discovery [J]. *Eur J Med Chem*, 2023, 260: 115765.
- [72] Liang Z, Liu T, Li Q, et al. Deciphering the functional landscape of phosphosites with deep neural network [J]. *Cell Rep*, 2023, 42: 113048.
- [73] Zhu F, Yang S, Meng F, et al. Leveraging protein dynamics to identify functional phosphorylation sites using deep learning models [J]. *J Chem Inf Model*, 2022, 62: 3331-3345.
- [74] Zhu F, Deng L, Dai Y, et al. PPICT: an integrated deep neural network for predicting inter-protein PTM cross-talk [J]. *Brief Bioinform*, 2023, 24: bbad052.
- [75] Lee JM, Hammaren HM, Savitski MM, et al. Control of protein stability by post-translational modifications [J]. *Nat Commun*, 2023, 14: 201.
- [76] Hua L, Zhang Q, Zhu X, et al. Beyond proteolysis-targeting chimeric molecules: designing heterobifunctional molecules based on functional effectors [J]. *J Med Chem*, 2022, 65: 8091-8112.
- [77] Paudel RR, Lu D, Roy Chowdhury S, et al. Targeted protein degradation *via* lysosomes [J]. *Biochemistry*, 2023, 62: 564-579.
- [78] Borgelt L, Haacke N, Lampe P, et al. Small-molecule screening of ribonuclease L binders for RNA degradation [J]. *Biomed Pharmacother*, 2022, 154: 113589.
- [79] Meyer SM, Tanaka T, Zanon PRA, et al. DNA-encoded library screening to inform design of a ribonuclease targeting chimera (RiboTAC) [J]. *J Am Chem Soc*, 2022, 144: 21096-21102.
- [80] Dey SK, Jaffrey SR. RIBOTACs: small molecules target RNA for degradation [J]. *Cell Chem Biol*, 2019, 26: 1047-1049.
- [81] Ge Z, Leighton JS, Wang Y, et al. Integrated genomic analysis of the ubiquitin pathway across cancer types [J]. *Cell Rep*, 2018, 23: 213-226.
- [82] Wang X, Li Y, He M, et al. UbiBrowser 2.0: a comprehensive resource for proteome-wide known and predicted ubiquitin ligase/deubiquitinase-substrate interactions in eukaryotic species [J]. *Nucleic Acids Res*, 2022, 50: D719-D728.
- [83] Zheng SJ, Tan YH, Wang ZY, et al. Accelerated rational PROTAC design *via* deep learning and molecular simulations [J]. *Nat Mach Intell*, 2022, 4: 739-748.