

特约评述

DOI: 10.12211/2096-8280.2023-006

基于靶标结构的环肽分子计算设计

王凡灏², 来鲁华^{1, 2, 3}, 张长胜^{1, 2}

(¹ 北京大学前沿交叉学科研究院定量生物学中心, 北京 100871; ² 北京大学化学与分子工程学院, 北京分子科学国家研究中心, 北京 100871; ³ 北京大学-清华大学生命科学联合中心, 北京 100871)

摘要: 环肽在调控蛋白质-蛋白质相互作用方面具有独特的优势, 在新药研发领域受到了越来越多的关注。蛋白质相互作用界面一般较大而平坦, 相较于小分子化合物, 环肽分子更容易获得与这些靶标位点结合的高亲和力和高特异性。相较于线性多肽或蛋白质, 环肽结构一般具有更大的骨架刚性, 更难被酶降解, 从而在代谢上更稳定, 而且环肽更易于通过修饰改造增加跨膜活性, 从而结合细胞内的靶标蛋白。结构数据和结构建模方法是开发基于靶标结构计算设计环肽药物的基础。本文分析了蛋白质结构数据库中环肽与靶标蛋白结合情况, 介绍了目前环肽构象生成或结构预测的四类主要算法; 总结了基于靶标结构计算设计环肽分子的主要方法, 包括基于分子对接的虚拟筛选方法、借助于动力学模拟的设计方法、从头生成的设计方法以及具有跨膜活性的环肽设计方法; 并展望了数据驱动的机器学习方法在环肽设计领域中的可能应用以及未来环肽药物分子开发的可能方向。

关键词: 环肽; 多肽药物; 多肽设计; 构象生成

中图分类号: Q816 **文献标志码:** A

Target structure based computational design of cyclic peptides

WANG Fanhao², LAI Luhua^{1, 2, 3}, ZHANG Changsheng^{1, 2}

(¹Center for Quantitative Biology, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China;

²BNLMS, College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China; ³Peking-Tsinghua Center for Life Science, Peking University, Beijing 100871, China)

Abstract: Cyclic peptides (macrocycles) possess head-to-tail cyclic or partially cyclized substructures, which have received more and more attention in developing new drugs recently, since they have unique advantages in regulating protein-protein interactions (PPIs). Comparing to small-molecule compounds, it is easier to design cyclic peptide molecules that bind to target sites with high affinity and specificity, due to the broad and flat interfaces of PPIs and their large surfaces. Moreover, cyclic peptides are generally more rigid and difficult for digestion by proteases than their linear counterparts, making them more stable than linear peptides or proteins. Meanwhile, cyclic peptides are easier for modifications to increase transmembrane activity, targeting intracellular proteins through conformation adaptation or

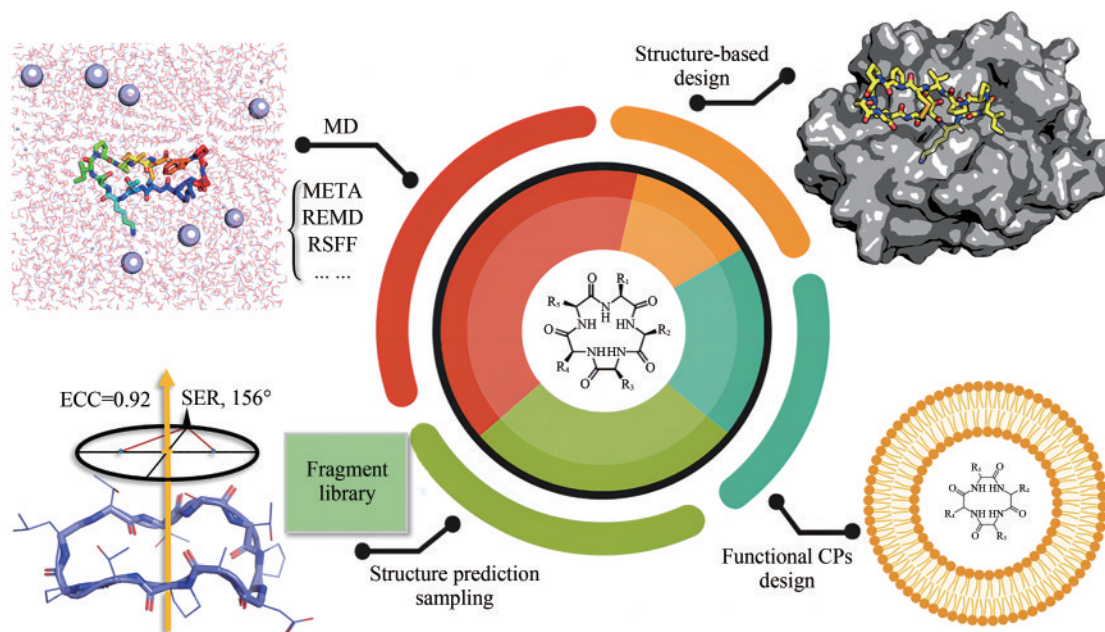
收稿日期: 2023-01-12 修回日期: 2023-02-23

基金项目: 国家自然科学基金 (21977007)

引用本文: 王凡灏, 来鲁华, 张长胜. 基于靶标结构的环肽分子计算设计[J]. 合成生物学, 2023, 4(3): 551-570

Citation: WANG Fanhao, LAI Luhua, ZHANG Changsheng. Target structure based computational design of cyclic peptides [J]. Synthetic Biology Journal, 2023, 4(3): 551-570

chemical modifications. 3D structure data and structure modeling technics are basis for designing structure based cyclic-peptide drugs. In this review, we assess the structures of cyclic peptides and target proteins available in protein structure database (PDB). Then, we review the algorithms of conformation generation or structure prediction for cyclic peptides, including homologous modeling, secondary structure prediction and optimization, backbone torsion sampling, and distance geometry method. We also summarize progress in target structure based computational design for cyclic peptides, including structure-based virtual screening, molecular dynamic simulation aided methods, *de novo* design algorithms, and the transmembrane cyclic peptide design. However, more generalized structure-based *de novo* design algorithms remains to be further explored, and methods to adopt unnatural amino acids or chemical modifications are also needs to be developed. It's worth noting that, with the increase of data for cyclic peptide 3D structures, the data-driven machine learning method may provide a more promising solution for improving the efficiency and effectiveness of structure based cyclic peptide *de novo* design and conformation generation to develop cyclic peptide drugs in the future.



Keywords: cyclic peptides; peptide drugs; peptide design; conformation generation

1 环肽药物

蛋白质-蛋白质相互作用在生物体系中发挥了关键作用, 很多相互作用体系, 如酶 / 抑制剂、细胞因子 / 细胞膜受体、病原体 / 受体、细胞信号转导蛋白质作用、分子机器组装体内部作用、蛋白质聚集等都被认为是潜在的重要靶标^[1-4]。通过调节蛋白质相互作用可以提高药物的专一性, 减少药物的毒副作用。蛋白质-蛋白质相互作用已经成为国内外药物研发机构竞相投入研发的药物靶点^[5-10]。

蛋白质-蛋白质相互作用界面一般较大而平坦,

较难进行高专一性结合的小分子化合物设计。蛋白质-蛋白质相互作用界面面积通常为 $1500 \sim 3000 \text{ \AA}^2$, 而小分子化合物总的溶剂可及表面积只有约 $150 \sim 500 \text{ \AA}^2$ 。相对于小分子, 多肽分子可以提供较大的结合界面, 并且具有易合成^[11]、结合专一性强、毒性小、免疫原性小等优势^[12-16]。使用天然多肽作为药物, 在调控分泌蛋白、细胞膜上受体等蛋白质功能方面已取得了很大成功^[17]。但由于天然多肽在很多情况下存在热稳定性差、易被蛋白酶降解、结合较弱的缺点, 导致这些多肽分子在体内的半衰期短、药效弱^[18-21]。所以, 改造这些多肽分子或全新设计稳定

的多肽是药物研发的重要目标^[22-23]。

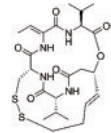
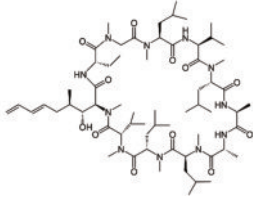
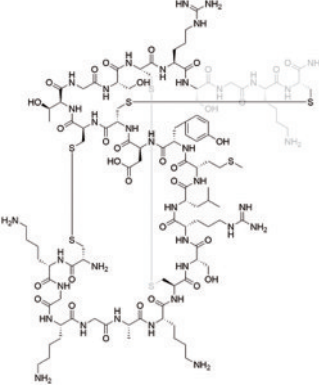
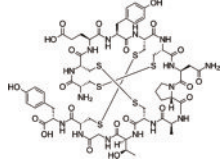
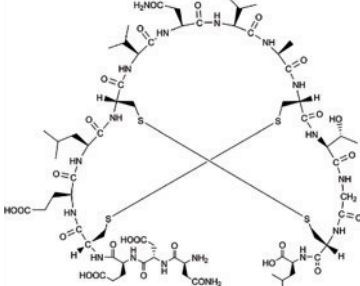
多肽分子的环化，如主链首尾形成酰胺键、残基侧链间形成二硫键或其他共价键，可以限制多肽的构象，是提高多肽药物稳定性的重要手段。环肽降低了肽链的柔性，因此还可以减少与靶蛋白结合的熵损失^[24-30]。在天然氨基酸环肽的基础上再引入*N*-甲基化氨基酸、D型氨基酸、类肽等非天然残基可以更进一步解决易被蛋白酶水解等问

题，并进一步拓宽其可设计化学空间^[31]。天然产物与人工设计的环肽分子越来越多地用作调节蛋白质-蛋白质相互作用，例如HDAC抑制剂^[32-34]、灰霉素^[35]、CXCR4拮抗剂^[36-37]、抑制HIF-1a/HIF-1b的cyclo-(CLLFVY)^[35, 38]等。

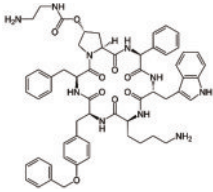
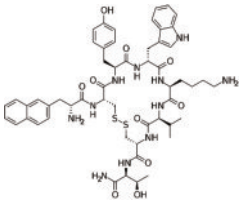
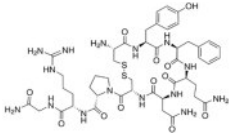
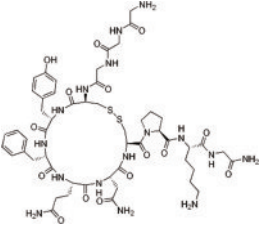
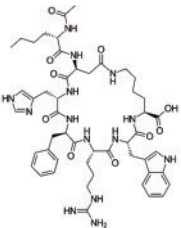
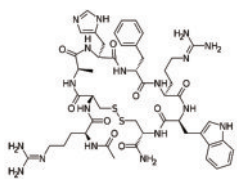
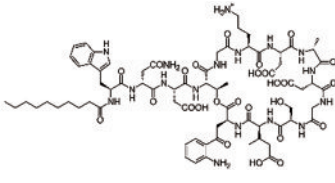
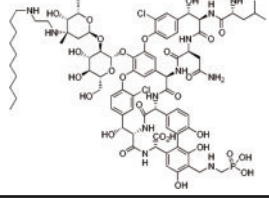
过去20年里经FDA批准上市的环肽药物有18种^[39-40]，包括激素类药物和靶向药，表1列出了这些药物的结构和功能。

表1 已获FDA批准的18种环肽类药物汇总表

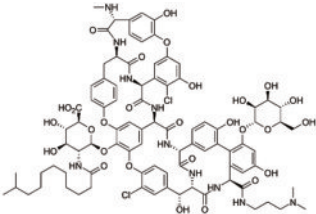
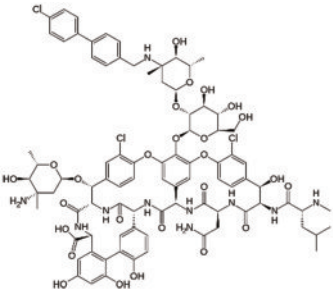
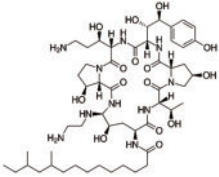
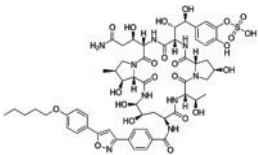
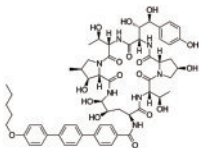
Table 1 Summary of 18 cyclic peptide drugs approved by FDA

名称	结构	功能简介
罗米地辛 (Romidepsin)		靶向组蛋白去乙酰化酶，具有抗癌活性，用于治疗T细胞淋巴瘤
伏环孢素 (Voclosporin)		靶向钙调神经磷酸酶，用于治疗狼疮性肾炎
齐考诺肽 (Ziconotide)		靶向G蛋白偶联受体，N型钙通道的有效和选择性阻滞剂
利那洛肽 (Linaclotide)		鸟苷酸环化酶 C(GC-C) 激动剂
普卡那肽 (Plecanatide)		鸟苷酸环化酶-C(GC-C)激动剂

续表

名称	结构	功能简介
帕西瑞肽 (Pasireotide)		生长抑素类似物,靶向生长抑素受体,可以提高生长抑素受体的激动剂活性
兰瑞肽 (Lanreotide)		生长素抑制剂的第二代类似物,靶向生长抑素受体,活性更高
加压素 (Vasopressin)		一种抗利尿激素,靶向血管加压素受体
特利加压素 (Terlipressin)		一种人工合成的长效抗利尿激素,靶向血管加压素受体
布美兰肽 (Bremelanotide)		黑素皮质素受体的激动剂,主要作用于 MC3 和 MC4 受体
赛美拉肽 (Setmelanotide)		黑素皮质素受体的激动剂,主要作用于 MC4 受体
达托霉素 (Daptomycin)		一种环状脂肽抗生素,破坏革兰氏阳性菌的细胞壁
特拉万星 (Telavancin)		万古霉素的半合成衍生物,抑制革兰氏阳性菌的细胞壁合成

续表

名称	结构	功能简介
达巴万星 (Dalbavancin)		第二代脂糖肽抗生素药物,比万古霉素更有效,也针对细胞壁的生物合成
奥利万星 (Oritavancin)		由氯霉素衍生的万古霉素的合成类似物,抑制细菌细胞壁的合成
卡泊芬净 (Caspofungin)		靶向1,3- β -葡聚糖合成酶,对侵袭性念珠菌病和其他形式的系统性真菌病有临床疗效
米卡芬净 (Micafungin)		靶向1,3- β -葡聚糖合成酶,对侵袭性念珠菌病和其他形式的系统性真菌病有临床疗效
阿尼芬净 (Anidulafungin)		靶向1,3- β -葡聚糖合成酶,对侵袭性念珠菌病和其他形式的系统性真菌病有临床疗效

这些分子主要来源于从自然界中分离的环肽,经改造优化提高了效价、药代动力学和药代动力学特性。例如通过将芳香萘和D-色氨酸加入生长抑素类药物兰瑞肽(Lanreotide, 结构见表1),可以稳定其超分子堆积的纳米管,有助于延长其半衰期、抑制激素水平和活性的能力。从技术角度来看,环肽药物的开发得益于快速发展的化学合成与修饰方法和现代分子生物学技术,环肽的分子结构不仅可以进行灵活修饰,一级序列也可以

快速突变和优化,以获得更高的产量与生物活性。例如由环孢素衍生而来的伏环孢素(Voclosporin, 结构见表1)就是利用化学修饰,将E-MePmt1分子基团取代环孢素A(CsA)中第一个残基Bmt1^[41-42],以优化其对靶标的结合能力,并提升了代谢稳定性^[43-45]。但是单纯通过大规模实验分离与筛选的方法获得可结合靶标蛋白环肽分子的效率低、成本高,所以计算设计方法将极大地帮助环肽药物的开发^[46-48]。

2 环肽的结构与环肽-靶标蛋白复合物结构

环肽的三维结构和环肽与靶标蛋白的复合物结构为理解环肽性质和功能机制提供了重要依据。已有很多工作收集了自由环肽的实验结构,例如,Beaufays等^[49]的研究中收集了38个PDB数据库中的环肽结构作为测试集,长度5~30个残基,通过侧链或主链之间的共价键连接成环;黄胜友实验

室^[50]在多肽构象生成方法MODPEP2.0的研究中,收集了蛋白质结构数据库(PDB)中采取二硫键连接成环的环肽结构。我们对PDB中的环肽配体与靶标蛋白的高分辨(≤ 2.5 Å)复合物结构进行了统计分析,截至2022年9月,PDB中包含了经过过去冗余的88个环肽配体与靶标蛋白的高分辨复合物结构,这些结构中的环肽配体长度为5~21个残基,靶标蛋白在相互作用界面上被埋藏的面积大部分在300~700 Å²之间(见表2)。

表2 PDB数据库中的环肽-靶标蛋白质复合物结构数据表

Table 2 Non-redundant cyclic peptide-target protein complex structures in the PDB database

PDB 编号	环肽配体序列	肽链 长度	环的 大小	环肽功能	靶蛋白界面 埋藏面积/Å ²	环化类型
1e4w	<u>SHFNEYE</u>	8	8	TGF- α 抗原表位类似物,与 tAb2 抗体结合	381	酰胺键
1sfi	<u>GRCTKSIPPICFPD</u>	14	14	胰蛋白酶抑制剂	612	酰胺键
2axi	<u>PFE[6cw]LDWEFp</u>	10	10	P53-HDM2 配体	510	酰胺键
3av9	<u>SAKIDNLD</u>	8	8	HIV 整合酶重组抑制剂	397	酰胺键
3ava	<u>ALKIDNLD</u>	8	8	HIV 整合酶重组抑制剂	395	酰胺键
3avb	<u>SLKIDNLD</u>	8	8	HIV 整合酶重组抑制剂	401	酰胺键
3avf	<u>DLKIDNLD</u>	8	8	HIV 整合酶重组抑制剂	418	酰胺键
3avg	<u>ADKIDNLD</u>	8	8	HIV 整合酶重组抑制剂	385	酰胺键
3avh	<u>ARKIDNLD</u>	8	8	HIV 整合酶重组抑制剂	388	酰胺键
3avi	<u>SLKIDNMD</u>	8	8	HIV 整合酶重组抑制剂	432	酰胺键
3avj	<u>ALKIDNMD</u>	8	8	HIV 整合酶重组抑制剂	437	酰胺键
3avk	<u>SLKIDNED</u>	8	8	HIV 整合酶重组抑制剂	419	酰胺键
3avm	<u>SRKIDNLD</u>	8	8	HIV 整合酶重组抑制剂	402	酰胺键
3avn	<u>SHKIDNLD</u>	8	8	HIV 整合酶重组抑制剂	394	酰胺键
3wne	<u>PKIDNG</u>	6	6	HIV 整合酶天然配体	378	酰胺键
3wng	<u>PKIDNp</u>	6	6	HIV-1 整合酶抑制剂	360	酰胺键
3wnh	<u>PK[nle]DNv</u>	6	6	HIV-1 整合酶抑制剂	367	酰胺键
3zgc	<u>GDEETGE</u>	7	7	红细胞核因子配体	343	酰胺键
4k1e	<u>GFCQRSIPPICFPD</u>	14	14	胰蛋白酶抑制剂	605	酰胺键
4kel	<u>GFCQRSIPPICFPN</u>	14	14	胰蛋白酶抑制剂	609	酰胺键
4y1d	<u>K[nle]DNv[45w]</u>	6	6	HIV-1 整合酶抑制剂	412	酰胺键
5n99	<u>NQpWQ</u>	5	5	链霉素和素配体	371	酰胺键
5xn3	<u>RGDINNNV</u>	8	8	SPSB2-iNOS 相互作用抑制剂	270	酰胺键
6jwm	<u>RGDINN</u>	7	7	SPSB2-iNOS 相互作用抑制剂	281	酰胺键
6jwn	<u>RGDINNVE</u>	9	9	SPSB2-iNOS 相互作用抑制剂	294	酰胺键
7k2e	<u>GDEETGE</u>	7	7	人源 KEAP1 蛋白抑制剂	360	酰胺键
7k2g	<u>GDEEAGE</u>	7	7	人源 KEAP1 蛋白抑制剂	403	酰胺键
7k2h	<u>GDPETGE</u>	7	7	人源 KEAP1 蛋白抑制剂	343	酰胺键
7k2m	<u>GEPETGE</u>	7	7	人源 KEAP1 蛋白抑制剂	320	酰胺键
6xbe	<u>rrLcPVPE</u>	8	8	NDM-1 金属- β -内酰胺酶抑制剂	379	酰胺键
6xbf	<u>rrLcPIPE</u>	8	8	NDM-2 金属- β -内酰胺酶抑制剂	452	酰胺键
1hqq	<u>RCCHPQCGAVEEC</u>	13	11	链霉素和素配体	413	二硫键
1jbu	<u>EEWEVLCWTWETCER</u>	15	7	凝血因子 VII 抑制剂	908	二硫键
1smf	<u>CTKSIPPEC</u>	9	9	胰蛋白酶抑制剂	414	二硫键
1sle	<u>CHPQGPPC</u>	8	8	链霉素和素配体	324	二硫键
1vpp	<u>RGWVEICAADDYGRCLTE</u>	18	9	生长素抑制剂	556	二硫键
1vwb	<u>CHPQFC</u>	6	6	链霉素和素配体	306	二硫键
2ck0	<u>CKEWLSTAPCG</u>	11	10	血管紧张素 II 抗体结合肽	504	二硫键
2nwn	<u>CSWRGLENHRMC</u>	12	12	丝氨酸蛋白酶抑制剂	597	二硫键
3g5v	<u>CGADSYEMEEDGVRKC</u>	16	16	EGFR 肽段	603	二硫键

续表

PDB 编号	环肽配体序列	肽链长度	环的大小	环肽功能	靶蛋白界面埋藏面积/A ²	环化类型
3p72	<u>C</u> TERMALHNLC	11	11	血小板糖蛋白 1b 抑制剂	521	二硫键
3wnf	CKIDNC	6	6	HIV-1 整合酶抑制剂	432	二硫键
4ib5	G <u>C</u> RLYGFKIHG <u>C</u> G	13	11	CK2beta 拮抗剂	456	二硫键
4m1d	<u>C</u> RIHIGPGRAFY <u>T</u> C	14	14	HIV-1 gp120 蛋白 V3 域类似物	588	二硫键
4ou3	<u>C</u> NGR <u>C</u> G	6	5	猪氨肽酶 N 抑制剂	366	二硫键
5dje	DCAWHLGELVW <u>C</u> T	13	11	抗体结合肽	610	二硫键
5co5	GCCSHPACAGNNQHIC	16	15	芋螺毒素	692	二硫键
5eoc	<u>C</u> QLINTNGSWHIC	13	13	丙型肝炎病毒 E2 表位 I 类似物	468	二硫键
5h5q	<u>C</u> RVDLQGWRR <u>C</u> RR	13	11	人源 GPX4 抑制剂	505	二硫键
5th2	<u>C</u> QFDQSTRRL <u>K</u> C	12	12	L5Q 中间位变体类似物	709	二硫键
5vb9	<u>C</u> WVLEYDMFGALH <u>C</u> R	15	14	IL-17 抗体抑制剂	571	二硫键
5wxr	GAC <u>S</u> ARGLENHAAC	14	12	尿激酶型纤溶酶原激活物抑制剂	526	二硫键
5xco	RRRR <u>C</u> PLYISYDPV <u>C</u> RRRR	19	11	K-Ras(G12D 突变体)抑制剂	650	二硫键
6e5m	<u>C</u> TKSIPP <u>Q</u> C	9	9	β-胰蛋白酶抑制剂	434	二硫键
1bm2	[slz][pY]VNV <u>P</u>	6	6	GBR2-sh2 结构域高活性配体	336	其他环化
1bzh	DAD <u>E</u> [flr]L[<u>a</u> ea]	7	7	酪氨酸磷酸酶抑制剂	398	其他环化
1vw1	HPQGPPC([l <u>e</u> a])K	9	8	链霉素和素配体	300	其他环化
4zjx	<u>C</u> [dab]RWT <u>K</u> CL	8	7	肉毒杆菌神经毒素(血清型 A)抑制剂	604	其他环化
5nes	<u>c</u> WwKkKkKwW <u>c</u> [8vh]	12	12	靶向铜绿假单胞菌糖蛋白的抗菌肽	259	其他环化
5ney	<u>C</u> wWkKkKkWw <u>C</u> [oxe]	12	12	靶向铜绿假单胞菌糖蛋白的抗菌肽	271	其他环化
5nf0	<u>c</u> WwKkKkKwW <u>c</u> [oxe]	12	12	靶向铜绿假单胞菌糖蛋白的抗菌肽	283	其他环化
6b67	<u>M</u> [pS]I[pY]VA[48v]	7	7	PPM1A 活性调节剂	528	其他环化
6dn6	INN <u>N</u> [<u>a</u> bu]	5	5	iNOS-SPSB 蛋白-蛋白相互作用抑制剂	234	其他环化
6dn7	WDINN <u>N</u> [<u>b</u> al]	7	7	iNOS-SPSB 蛋白-蛋白相互作用抑制剂	254	其他环化
6dn8	[gzi]VDINN <u>N</u> [cy3]	8	8	iNOS-SPSB 蛋白-蛋白相互作用抑制剂	258	其他环化
6nnv	F[mea][9kk][sar]DV[mea]Y[sar]WYL <u>C</u> K	14	13	PD-L1 抑制剂	510	其他环化
6u4a	WWI <u>P</u> [aly]VK[aly]G <u>C</u>	11	11	BRD2-BD1 抑制剂	587	其他环化
6u74	WKT <u>I</u> [aly]G[aly]TWRT[aly]Q <u>C</u>	14	14	BRD2-BD1 抑制剂	582	其他环化
6u8m	W[aly]KAILPG[aly]ILKTLHIC	17	17	BRD2-BD1 抑制剂	504	其他环化
6wgn	GyFVNFRNFR <u>T</u> FR <u>C</u> G	15	14	K-Ras(G12D 突变体)抑制剂	641	其他环化
6xci	cP[lwi][hyp]Ep <u>k</u> [nle]	8	8	NDM-3 金属-β-内酰胺酶抑制剂	457	其他环化
6xib	[b74]CKG[<u>f</u> tr][<u>f</u> tr]DHY[3wx] <u>C</u> A	12	11	PCSK9 抑制剂	546	其他环化
6xic	[b74]CKa[<u>f</u> tr][<u>f</u> tr]DHY[3wx][<u>d</u> hl]	11	11	PCSK9 抑制剂	518	其他环化
6xid	[v7m]CKa[<u>f</u> tr][<u>f</u> tr]PTY[3wx] <u>C</u> A	12	11	PCSK9 抑制剂	537	其他环化
6xie	[b74]CAa[apd][<u>f</u> tr]QT[0a1][3wx] <u>C</u>	11	11	PCSK9 抑制剂	496	其他环化
6xif	[v7s]GAa[apd][v7p][gnc][olt][0a1][3wx] <u>G</u>	11	11	PCSK9 抑制剂	528	其他环化
6xs5	[48v]yIIDTPLGVFLSSLKR	17	17	人源 Vps29 抑制剂, 结构稳定剂	529	其他环化
6xs7	[48v]yTTIYTWTPLGTFPRIR	17	17	人源 Vps30 抑制剂, 结构稳定剂	654	其他环化
6xs8	[48v]yGYDPLGLKYFA	13	13	人源 Vps31 抑制剂, 结构稳定剂	380	其他环化
6xsa	[48v]YLPTITGVGHLWHPL	15	15	人源 Vps32 抑制剂, 结构稳定剂	623	其他环化
6yw1	[48v]yVWLTDTWVLSRT	14	14	促进 HIF 脯氨酰羟化酶 2 结晶的环肽配体	671	其他环化
7bph	yFESVYAIWGT <u>L</u> C <u>G</u>	14	13	GNAS 抑制剂	520	其他环化
7k2k	[<u>b</u> al]DEET <u>G</u> E	7	7	人源 KEAP1 蛋白抑制剂	349	其他环化
7k2l	[<u>b</u> al]NPET <u>G</u> E	7	7	人源 KEAP1 蛋白抑制剂	330	其他环化
7k2o	[<u>a</u> bu]DPET <u>G</u> E	7	7	人源 KEAP1 蛋白抑制剂	346	其他环化
7k2p	[<u>d</u> av]DPET <u>G</u> E	7	7	人源 KEAP1 蛋白抑制剂	344	其他环化
7k2r	[<u>b</u> 3a]DPET <u>G</u> E	7	7	人源 KEAP1 蛋白抑制剂	350	其他环化
7rov	A[060]PLYISYDPV <u>C</u> RA	14	12	K-Ras(G12D 突变体)抑制剂	564	其他环化

注：下划线标注的残基为共价连接的残基，大写代表天然 L 型氨基酸残基，小写字母代表 D 型氨基酸残基，括号内为化学修饰的天然残基、除 D 型氨基酸外的非天然残基、环化分子基团，其中 [pY]、[pS] 分别代表磷酸化酪氨酸、磷酸化丝氨酸

这些结构中，31 个环肽配体通过酰胺键环化，23 个环肽配体通过二硫键环化，41 个环肽配体中含 D 型氨基酸或其他非天然氨基酸，或者是引入

了新的基团实现合环，例如图 1(c) 的环肽是由线性肽两末端的半胱氨酸通过与 α,α'-二氯-间二甲苯反应形成的合环^[51]。其中不含 D 型氨基酸环肽的主链

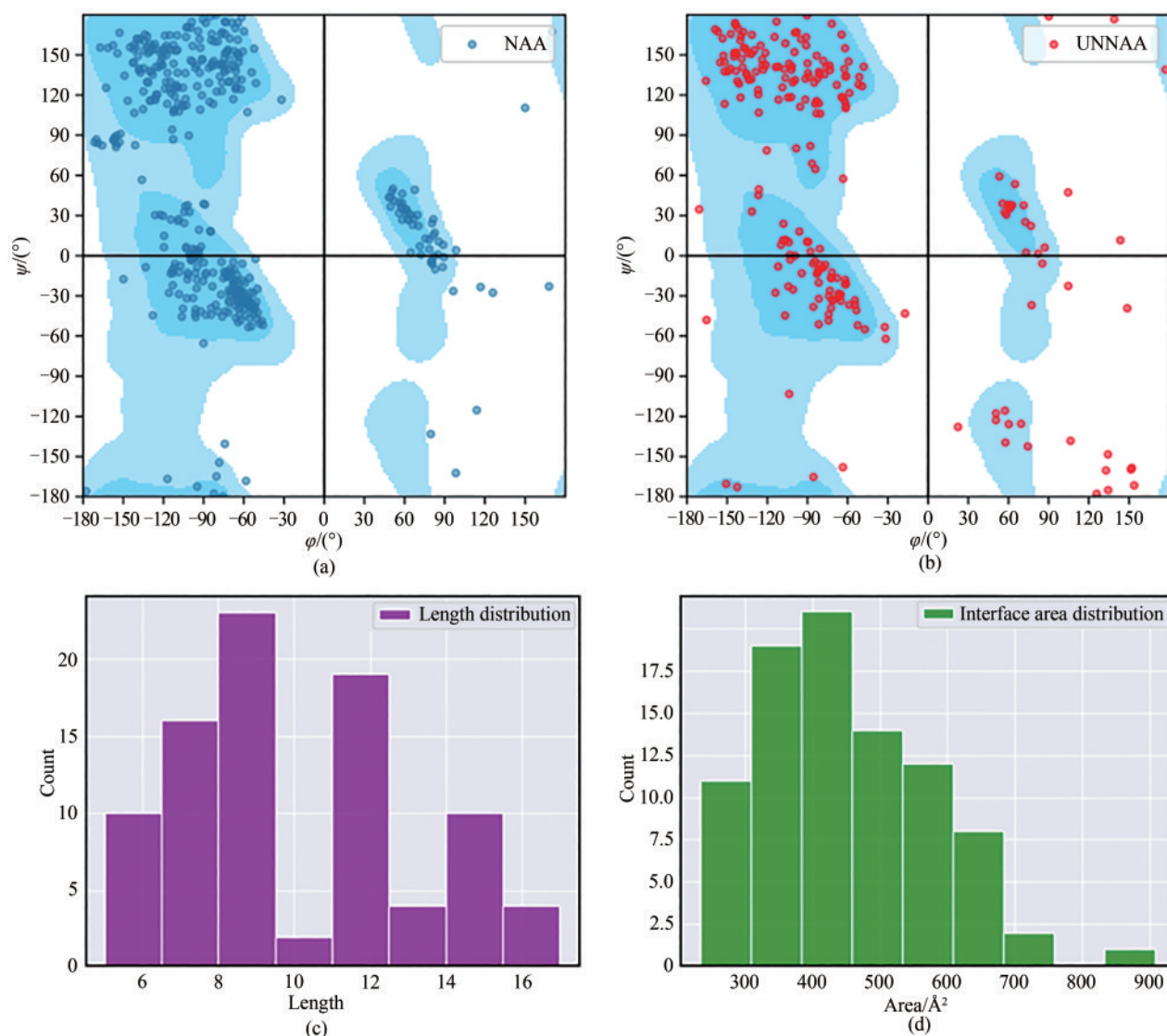


图1 PDB中环肽-靶标复合物数据集（见表2）中环肽配体的参数统计图

（a）、（b）中蓝色背景分布为天然蛋白质体系的氨基酸残基扭转角 ψ/ϕ 的分布。

（a）仅含天然氨基酸残基的环肽配体主链扭转角分布图（ ψ/ϕ ）；（b）存在非天然氨基酸残基的环肽配体主链扭转角分布图（ ψ/ϕ ）；

（c）数据集中所有环肽配体环序列长度分布图；（d）数据集中所有环肽配体与靶标之间界面面积分布图

Fig. 1 Parameters of cyclic peptide ligands with the cyclic peptide-target complex data set (see Table 1) in PDB

Distribution of torsion angles of the main chain of cyclic peptide ligands containing natural (a) and non-natural (b) amino acid residues (ψ/ϕ), in which blue cloud highlights the distribution of the torsion angle ψ/ϕ of amino acid residues in natural proteins; Length distribution of the loop sequences of all cyclic peptide ligands in the data set (c); Distribution of the interface area between all cyclic peptide ligands and targets (d)

二面角 ϕ 和 ψ 分布与天然蛋白的Ramachandran统计分布相近，如[图1(a)]所示，这些二面角大都在Ramachandran统计允许的范围之内。Riniker团队^[52]对PDB数据库和CSD数据库中的大环化合物（包括环肽）的统计结果表明，环上可旋转单键数 ≤ 8 时，才会导致环上的二面角分布受环应力的影响而偏离相应线性分子的分布范围^[52]，而环肽分子的环内一般含有 ≥ 5 个残基，可旋转单键数一般 ≥ 10 。

含D型氨基酸的L/D混合环肽的主链二面角 ψ 和 ϕ 的分布[见图1(b)]趋近于甘氨酸的非手性Ramachandran统计分布。

此外，有些环肽结构中存在着一些较为刚性的局部结构以稳定分子骨架，从而减少与靶标蛋白结合过程中的熵损失，增强环肽配体与靶标蛋白的结合^[26-27, 30, 53-54]，如图2中的 β 发卡末端环化形成的环肽配体结构[(a)以侧链二硫键合环，(b)以

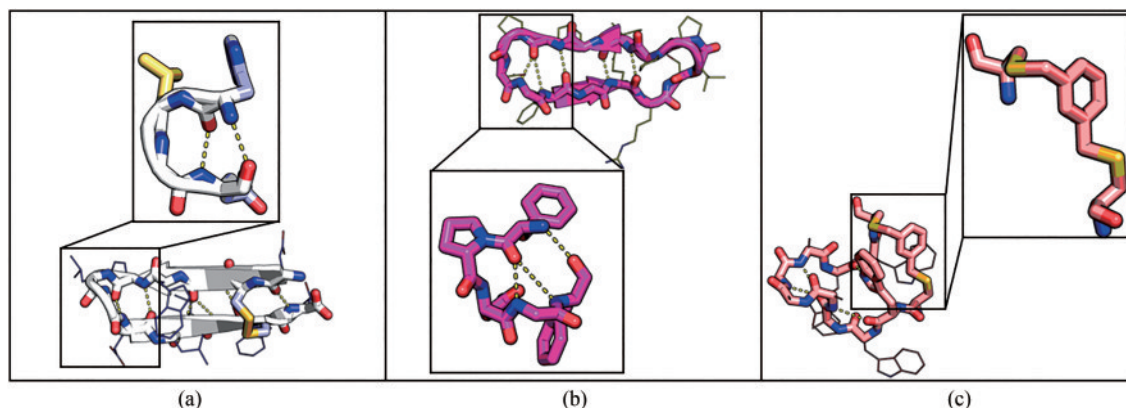


图2 PDB编号为5DJC (a)、4K1E (b) 和5NES (c) 的复合物结构环肽配体结构示意图

Fig. 2 Cyclic peptide ligand structures for the complexes in PDB 5DJC (a), 4K1E (b) and 5NES (c)

首尾酰胺键合环], β 片层之间的氢键和 β 转角处形成的1,3、1,4氢键相互作用, 稳定了环肽的构象, 增强了骨架的刚性。脯氨酸也可以限制环肽主链的运动, 因此在刚性环肽骨架的从头设计工作中引入脯氨酸也是稳定目标骨架结构的重要策略^[55-56]。

分子动力学模拟 (MD simulation) 是研究蛋白质和多肽结构动态变化的重要手段。基于分子动力学模拟的环肽药物分子的构象研究推动了研究者们对环肽构象的深入理解与经验积累, 为实现大规模环肽分子从头设计提供了理论依据^[57]。由于蛋白质中的环结构 (loop) 与环肽结构较为相似, 通用的分子动力学模拟力场参数根据蛋白质 loop 结构进行重新拟合调整后可以更好地用于环肽分子的结构优化与构象采样^[58-63]。Geng等^[64]就曾以晶体结构为基准, 比较了四种多肽力场 Amber99sb-ildn^[65]、OPLS-AA/L、RSFF1^[66]和RSFF2^[67]预测全反式 (酰胺键) 环肽结构的能力。其中基于蛋白质环形结构 (loop) 库进行参数化的RSFF1和RSFF2力场得到的预测结果与晶体结构匹配最佳^[68]。同其他柔性体系一样, 增强采样的方法可以加速环肽构象采样的过程, 提升采样的效率。其中最为广泛使用的是副本交换分子动力学 (REMD)^[69-71]和元动力学 (META)^[72-73]。

3 环肽分子结构建模 (结构预测/构象生成)

环肽分子结构建模的目标是生成环肽的低能

量的合理构象, 对于较刚性的环肽, 如环化的 β 发夹, 就是预测出准确的折叠结构。从结构数据库中收集环肽结构建立数据集来检验由环肽结构建模算法的表现。计算生成的环肽构象是环肽与靶标蛋白对接筛选的前提。经检验, 环肽分子的结构采样方法以及打分评价方法, 可直接用于大规模的基于靶标结构的环肽分子的改造、生成、评价^[74]。

3.1 基于模板结构的比较模建算法

黄胜友实验室^[50]开发的MODPEP2.0 (<http://huanglab.phys.hust.edu.cn/software/modpep2>) 可以快速产生二硫键环化的环肽构象。他们从PDB数据库中收集了3~30个氨基酸长度的二硫键环肽结构, 通过聚类选择后得到的不同长度的二硫键环肽结构集合作为目标环肽构象生成的模板库。MODPEP2.0选取与目标肽序列相似性高的, 并且结构分辨率高的结构做模板, 选取概率如下:

$$P_i = \exp(w_s \times \frac{s_i}{s_{\max}} + w_r \times \frac{r_i}{r_{\max}})$$

式中, s_i 是基于目标序列和第 i 个结构的序列相似性; r_i 是第 i 个结构的分辨率, 除以相应的最大值 (s_{\max}, r_{\max}) 是为了对不同的打分项进行归一化; w 为两项的权重系数。之后在选定的模板主链上应用侧链构象转子库逐个残基安装侧链, 得到目标二硫键环肽的构象, 并通过选取不同的模板生成多样的构象^[50]。对于环以外序列的结构, 算法用MODPEP的早期版本通过片段组装方法完成^[75]。该团队构建了测试方法对MODPEP2.0生成

构象的准确性进行评价,当生成10种或100种构象时,与测试数据集中结构最接近的C α RMSD平均值分别为2.20 Å、1.66 Å^[50, 75]。

张阳实验室^[76]发展的I-TASSER是基于结构采样和优化作蛋白质结构预测的代表性方法之一,加入环化所需的共价距离约束就可用于环肽结构预测^[76],提供在线计算服务(<http://zhang.bioinformatics.ku.edu/I-TASSER>)。该算法通过多种序列比对方法识别与连续片段相似的PDB模板结构,将模板的主链作为这个片段的主链结构,并通过基于模板片段副本交换的蒙特卡洛采样方法探索整个蛋白质或多肽的结构空间,获得能量较低的结构模型。环肽中的环化共价键可作为几何约束加入能量函数。在Peplook的环肽结构测试集中,I-TASSER表现优异^[49, 76]。

3.2 基于主链结构预测和分子力场优化的方法

PEPstrMOD算法(<https://webs.iitd.edu.in/raghava/pepstrmod/>)从序列预测所得二级结构(α 螺旋、 β 片、转角、无规等)类型构建多肽的初始主链构象,使用Amber力场对结构进行能量最小化和动力学模拟优化^[77]。对于环肽,在分子拓扑结构中加入相关共价键连。PEPstrMOD将末端修饰、D-氨基酸、非天然氨基酸、翻译后修饰等残基类型的力场参数加入Amber力场,使之可以预测含有非天然氨基酸残基的多肽结构^[77-79]。PEPstrMOD算法在来自于PDB数据库的ModPep、ModPep16、CyclicPep三个数据集上进行了检验,预测得到的主链原子的均方根偏差为3.81~4.05 Å之间。

PEP-FOLD算法(<https://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms:PEP-FOLD3>)也是先从序列预测各残基处的主链构象,然后利用力场做优化。不同的是PEP-FOLD将主链二级结构扩展为27种类型,基于预测的所有位置的这27种类型的概率分布,使用贪婪算法构建粗粒化的三维模型,使用蒙特卡洛方法作粗粒化力场优化后再构建全原子模型。同样地,在分子拓扑结构中加入相关共价键连用于环肽的结构预测^[80-81]。PEP-FOLD的最新版本改进了粗粒化力场的范德华项和

用于主链结构预测的片段库,提高了多肽结构的预测性能^[80-83]。

3.3 基于主链扭转角采样的环肽结构预测算法

环肽局部连续片段的构象运动是两端都受约束的运动,与力学中的运动学闭环(kinematic closure, KIC)类似,例如对于机器人,在给定肩部和指尖固定位置的情况下,确定机器人手臂内部关节的可能位置与扭转角度范围。Rosetta(<https://www.rosettacommons.org/software>)中的genKIC是运动学闭环算法在蛋白质环区结构预测、蛋白质主链结构扰动中的成功应用。genKIC算法选择一肽段上第一个、中间某一个和最后一个残基的C α 原子固定为转动中心,然后对非转动中心的残基用Ramachandran概率对 ψ/ϕ 随机采样并对N-C α -C键角采样,得到的肽链在固定残基处断开,之后运动学闭环算法找到使肽链重新闭合的3个转动中心残基的 ψ/ϕ 值,从而获得肽段的新构象^[84]。使用Rosetta势能函数和蒙特卡洛采样实现对构象空间的探索,获得蛋白质环区或环肽的低能构象^[85-86]。

Peplook(http://www.biosiris.com/en/Online_order/PepLook/PepLook_order.html)是一种对蛋白质或多肽的主链二面角进行玻尔兹曼随机采样来预测结构的算法^[49, 87]。Peplook从64对 ψ/ϕ 角度中对每一个氨基酸随机取值^[88-90],根据这些主链二面角值构建结构,并计算体系能量,在环肽结构预测中体系能量加入首尾二硫键或酰胺键原子的距离相关项。每一轮生成1万个结构,前一轮得到的体系能量和主链二面角取值的关系决定下一轮中各 ψ/ϕ 角度对被取到的概率,采样100~500轮得到低能量构象^[49]。在38个环肽测试集上,Peplook建模最佳结构的主链原子的均方根偏差的平均值为3.8 Å^[49, 91]。

3.4 基于距离几何的构象生成算法

构象生成器ETKDG是小分子化合物构象生成最常用的方法之一。ETKDG通过分子的拓扑结构和晶体结构中的二面角数据,生成分子的距离边界矩阵,根据该矩阵限定的原子间距离的范围随

机产生一个距离矩阵, 然后由距离矩阵产生三维结构, 之后优化构象^[92]。最近该方法改进了对大环化合物分子和环肽的生成。mETKDG使用椭圆几何约束来限制环肽整体环系骨架, 并且加入可调整的库仑相互作用作为酰胺原子之间的奖励式方法来模拟跨环分子内氢键^[52]。mETKDG的环肽建模性能在由PDB、CSD筛选的环肽分子测试集上得到了验证, mETKDG的重构主链原子的均方根偏差的平均值为1.23 Å。目前mETKDG已经被写入常用的RDkit (<http://www.rdkit.org/>) Python工具包中^[52, 92]。

通过基于分子对接的虚拟筛选, 可以缩小实验测试的分子范围, 从而更高效地发现新的结合靶标的环肽分子。分子动力学模拟是详细研究和改进已知环肽分子的重要方法。近年来, 以Rosetta为代表的从头生成算法大大拓展了设计环肽分子的化学空间, 为针对特定靶标结构设计全新的环肽药物分子提供了解决方案。对于细胞内的靶标蛋白, 需要在设计中考虑跨膜活性。因此本节将介绍这4个方面的计算设计方法(图3), 表3总结了这些算法的主要思路和适用目的。

4.1 基于分子对接的虚拟筛选方法

4 基于靶标结构的环肽分子计算设计

基于靶标结构可以改造设计或全新设计与之结合的环肽分子, 在具有备选的环肽库的情况下,

基于大规模环肽数据库的虚拟筛选是最早出现的计算设计方法, 环肽结构库和对接及打分算法是这一方法的核心内容。

Duffy等^[95]于2011年便尝试过生成用于虚拟

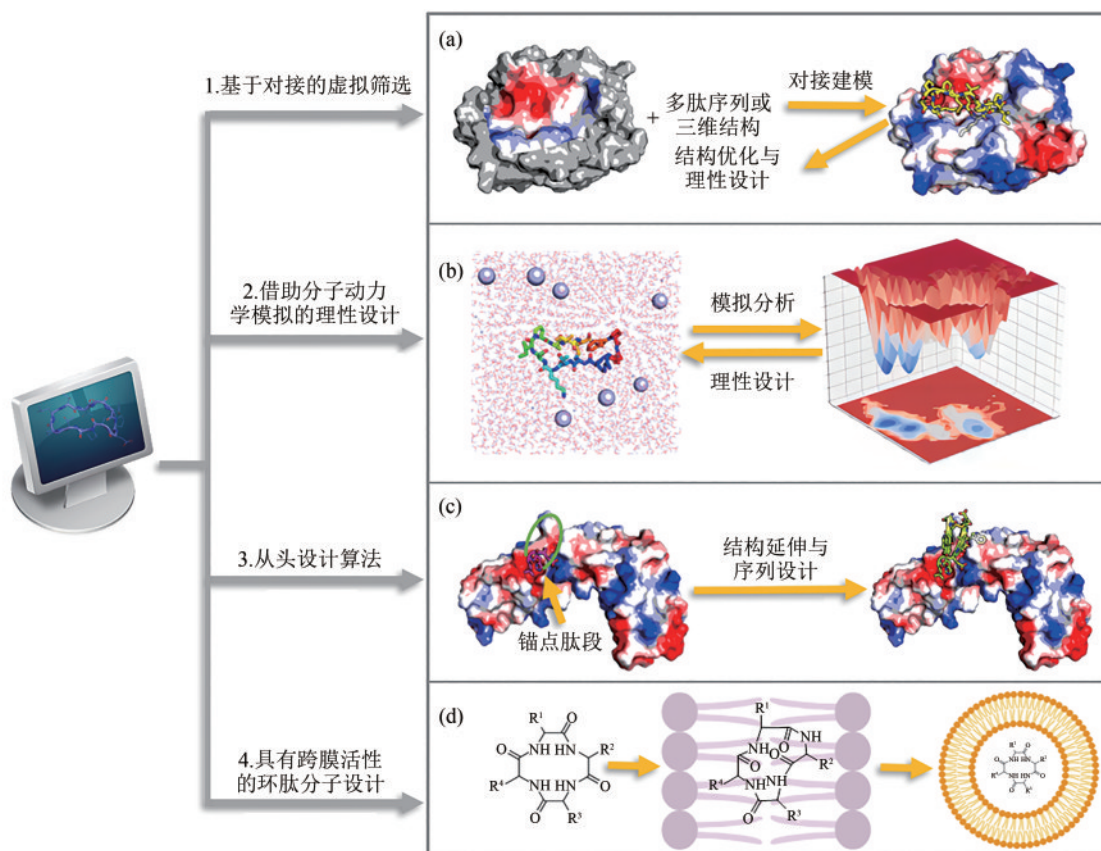


图3 基于靶标结构的环肽设计算法

(a) 基于分子对接的虚拟筛选算法; (b) 基于分子动力学模拟的理性设计算法; (c) 从头设计算法; (d) 跨膜环肽分子的设计算法

Fig. 3 Overview of computational methods for target structure based cyclic peptide design

(a) Virtual screening algorithms based on molecular docking; (b) Rational design algorithms based on molecular dynamics simulation;

(c) De novo design algorithms; (d) Design algorithms for transmembrane cyclic peptides

表3 基于靶标结构的环肽分子计算设计算法

Table 3 Structure based computational design algorithms of cyclic peptides

类型	代表算法	算法简介	算法目的
基于对接的虚拟筛选算法	AutoDock CrankPep ^[93] , HADDOCK ^[94] 等分子对接算法	将环肽结构库中的分子与靶标蛋白对接,基于对接打分富集可能结合靶标的环肽。对接打分是对环肽和靶标蛋白结合的粗略但高效的评估	从现有环肽、已知结构环肽或某一类环肽等环肽库中富集可能结合靶标蛋白的分子,为实验发现未知的结合环肽提供候选分子
借助分子动力学模拟的设计算法	REMD ^[69] , META ^[72] , BE-META ^[73] 等增强采样算法	借助于分子动力学模拟增强采样算法对环肽以及环肽与靶标蛋白的复合物结构进行采样。通过动力学采样对目标环肽和靶标蛋白的结合能进行精细的计算和比较	对已知结合的环肽分子进行分析和改造设计,估算结合自由能,研究各残基对结合的贡献,找到对环肽进行改进的可能方案
从头设计算法	Rosetta Anchor Extension ^[34]	以某个对结合能贡献较大的基团为起点,在靶标蛋白表面通过主链优化采样与序列设计生长出全新的环肽	不受限于已知结合环肽或是环肽结构库,针对靶标蛋白从头设计环肽配体,并能引入化学修饰与非天然氨基酸
具有跨膜活性的环肽分子设计	Rosetta ^[56]	引入主链酰胺键的N-甲基化修饰或降低侧链极性的非天然氨基酸	在设计中加入环肽跨膜活性的考量

筛选的环肽数据库,开发了CycloPs,该程序可以根据规则与用户预定义的约束,大量生成一维SMILES,通过RDkit建模为三维环肽分子结构,结合后续的虚拟筛选构成了完整的环肽药物的设计流程。但是环肽药物发现的虚拟筛选方法往往局限于数据库的规模与复合物对接建模的准确度。

Sanner实验室^[96-97]开发的AutoDock CrankPep(或简称ADCP)是由小分子-蛋白对接软件AutoDock改进而来的多肽(环肽)对接软件(<https://ccsb.scripps.edu/adcp>),可以对接长度在20个氨基酸残基以内的线性多肽与环肽结构。他们应用了基于多肽骨架的曲轴运动微扰方式,支持对具有环肽主链的构象进行高效采样,并在AutoDock基本的Monte Carlo采样过程中引入了主链环化势能项,使得程序能够从线性多肽序列出发的扩展线性肽开始对接计算,从而在没有初始环状构象可用的情况下实现对接与环化同时进行,最终得到环肽-蛋白复合物结构。最后,距离依赖性势能确保多肽环化。这种势能不同于之前的基于环化势能的方法,ADCP中性线性多肽能够在模拟过程中环化、断裂和重组^[96-97],中间状态的存在允许算法在对接过程中探索主链的各种构象组合并确定最佳二硫键或首尾环化位点。ADCP测试的平均 f_{nc} (对接得到的复合物构象对于天然复合物构象中的原位配体-受体相互作用对的召回率,在蛋白质复合物结构预测比赛CAPRI中, f_{nc} 高于0.3对应

中等对接精度, f_{nc} 超过0.5则认为对接精度较高)超过0.5^[98]。ADCP是环肽-受体复合物相互作用建模和结构预测的有效工具,可用于开发设计环肽药物,并在相当一段时间内被认为是该领域最先进的SOTA(state of the arts)算法^[93]。

与ADCP类似,HADDOCK算法的最新2.4版本(<https://wenmr.science.uu.nl/haddock2.4>)在对接过程中也应用了末端二硫键几何约束或者首尾氨基酸距离约束,使多肽结构在对接过程中逐步环化^[94]。在对接流程中,用于对接的环肽是用PyMOL的内置功能从多肽序列生成起始构象^[99],HADDOCK会利用环化约束指导线性多肽结构逐步收紧多肽首尾或二硫键之间的距离(通过HADDOCK)使线性肽强制环化。环化之后,HADDOCK会将生成的环肽结构分别与各自的受体结构进行对接,并实施9种对接方案以获得高质量的配体结构^[94]。HADDOCK在环肽-蛋白质复合物建模方面的表现与ADCP基本相当^[96][在完全未结合状态从头对接的情况下与ADCP的表现相当,在重建模(redock)性能评估时略优于ADCP(HADDOCK中的默认模式)]。由于HADDOCK还可以结合各种实验数据(例如关于肽构象的核磁共振谱图信息)来指导对接,因此基于已知结构的测试性能可被认为是下限,通过额外的实验数据输入可以进一步增强算法性能^[94]。环肽配体的对接建模为环肽药物分子的虚拟筛选与理性设计提供了高效可靠的工具。

4.2 借助于动力学模拟的设计方法

较为传统的药物分子改性与理性设计方法仍然适用于设计环肽配体，这类方法向环肽复合体系的迁移应用往往依赖于高效的环肽配体分子构象采样算法。

Razavi等^[79]使用计算筛选和REMD模拟相结合的算法，通过模拟LapD中的 β -发夹(hairpin)结构，成功设计了环肽配体。LapD是一种细菌蛋白，其与LapG的相互作用对生物膜的形成至关重要^[78-79, 100-101]。该团队首先使用隐式溶剂模型的REMD模拟用于快速筛选潜在的环肽模拟物^[69]，随后应用显式溶剂模型进一步进行模拟采样结合虚拟筛选方法得到4种最有利的的设计产物，最终成功设计出了类似LapD β -发夹结构的一系列环肽配体^[73]。若已知体系的快速自由度，则可以应用BE-META模拟方法以实现高效构象采样。例如，在基于isoDGR的 $\alpha\beta$ 3环肽拮抗剂的设计中便使用了BE-META结合多构象对接的方法^[102]，从而在其构象系综中发现最有利于结合的配体构象，并根据这些优势构象进一步修改isoDGR分子结构，衍生物的平衡系综更倾向于与受体产生强结合的状态^[69]。

4.3 从头生成设计方法

从头设计方法可以得到与已知肽段结构完全不同的全新结构环肽，需要同时考虑生成环肽结构的稳定性以及跟靶标蛋白的亲和力，在环肽结构采样的同时优化与靶标蛋白的计算结合能。

在不考虑靶标的自由环肽结构生成方面，Hosseinzadeh等^[55]利用基于Rosetta软件的genKIC算法生成L/D氨基酸混合的7~14个残基的环肽的主链骨架结构。用全甘氨酸的序列和甘氨酸、脯氨酸非手性的Ramachandran概率对 ψ/ϕ 随机采样，生成闭环的主链骨架，最终得到200多个预测可折叠成单个稳定结构的环肽，其中12个结构得到实验验证。此工作的算法流程支持加入D-氨基酸来拓宽构象空间并且能够形成双环（同时存在首尾酰胺键连接与二硫键连接）拓扑结构以增强大环刚性，这些结构几乎完全覆盖了已知环肽化合物可能的构象空间，并探索到了大量非天然的局部二级结

构，大大拓宽了理性环肽药物设计和虚拟筛选方法的可用起始结构空间^[55]。之后，Hosseinzadeh等^[34]在2021年提出了锚扩展(anchor extension)方法，将genKIC环肽结构设计方法用于直接在靶标蛋白表面设计环肽配体，是一次里程碑式的尝试。这种方法需要目标受体表面的三维结构，并利用已知有较强结合配体分子的对结合能贡献较大的基团作为锚，置于原位（同时删除配体的其余部分结构），锚即为环肽结构向外延伸的起始位点。锚扩展方法使用Rosetta软件中的广义运动学闭环方法(genKIC)在锚点周围构建环肽骨架，生成的合理大环骨架从锚点出发向外延伸进行主链优化采样与序列设计，逐步增强其与受体之间的相互作用^[84-86]。作者选定的测试体系是HDAC2和HDAC6分别与一种蓝藻毒素Largazole结合的复合物结构，保留其中的2-硫-2-氨基-7-磺酰基庚酸(SHA)的部分作为锚，来从头设计环肽配体^[103-104]，最终得到了多个 IC_{50} 为纳摩尔量级的产物(HDAC2与SHA的 IC_{50} 和与环肽配体最优 IC_{50} 分别为5.4 $\mu\text{mol/L}$ 、9.1 nmol/L，HDAC6则为0.6 $\mu\text{mol/L}$ 、5.4 nmol/L)。

锚扩展方法需要结合较强的“锚”残基结构作为起始。Delaunay等^[105]提出的Des3PI算法(https://github.com/des3pi/Public_Des3PI)能够生成这些“锚”残基，该团队利用20种氨基酸残基与靶标蛋白表面进行对接筛选，得到结合较强的位点，然后将对接在表面的氨基酸残基间通过甘氨酸简单连接成环得到环肽配体分子^[105]。这种方法普适性较强，但是由于主链结构柔性较大，生成的环肽配体刚性与结合稳定性弱于锚扩展工作中精细设计的环肽配体结构。

4.4 具有跨膜活性的环肽分子设计

药物分子的跨膜活性决定了其是否可以作用于细胞内的靶标^[39]。多数多肽药物的靶标为细胞膜上的受体，近年来在可以进入细胞内的多肽设计领域取得了重要进展。实验研究表明主链酰胺键的N-甲基化修饰以及降低侧链的极性均可提高环肽的跨膜活性^[39]。Sindhikara等^[106]通过动力学模拟的方法发现，环肽分子的主链往往在磷脂双

分子层内会将极性的羰基氧和氮上的氢原子通过环肽分子的构象变化(酰胺键顺反异构转变)收容到环内,这大大降低了分子的极性,使之与磷脂细胞膜融合,穿过细胞膜后又可以被极性较大的细胞内环境诱导,使环肽分子采取相对较为极性的构象,易于与极性的蛋白靶标结合,形成较为稳定的复合物^[18, 107-108]。因此此类环肽分子具有随环境极性不同转变构象的特性,这些特性决定了此类环肽具有作为可跨膜药物分子的潜力^[109-110]。Baker实验室^[56]综合了这些特性进行了具有优良膜渗透性的环肽设计并且取得了成功。该团队改进了genKIC算法,在采样的过程中,允许整个化学基团(如氮甲基)被放置,这就可以在环肽结构中引入非天然氨基酸修饰,并且Rosetta的残基类型被重新整理为残基和补丁,这些补丁可以通过Rosetta的能量函数来对已有的化学实体做化学修饰^[111]。研究者们利用优化调整后的Rosetta软件设计了一系列长度为6~12的环肽结构,这些结构均为化学修饰或支持构象转变的刚性环肽结构,随后进行了实验合成与结构解析,实验测出的结构与设计结构叠合较好(BB-RMSD<1.5 Å),最后测试了新结构的膜通透性和口服生物利用效率,得到了较好的效果^[56]。此工作在环肽药物设计做了进一步的探索,从化学修饰角度着手,大量设计结构稳定且具有跨膜活性潜在环肽药物分子,为环肽药物分子的虚拟筛选工作进一步提供了可靠基础。

5 总结与展望

目前环肽作为调控蛋白质-蛋白质相互作用的重要研究对象,已经获得了越来越多的关注,有关环肽分子的结构、动态构象、与靶标蛋白质相互作用的实验数据也逐渐增多。基于这些实验数据,研究者通过分子动力学模拟、分子构象采样等方法研究了环肽的构象空间,开发了基于比较模建、基于主链扭转角采样、基于主链二级结构预测和分子力场优化和基于距离几何的环肽分子构象生成或结构预测算法。蛋白质中环区(loop)与环肽的结构和动态性质相似,这些结构数据也是进行环肽构象研究和结构建模方法测试

的重要数据来源。以这些研究为基础,人们已经开始探索基于分子对接的虚拟筛选方法、借助于分子动力学模拟的方法和基于结构采样的从头生成方法来针对目标的靶标结构进行环肽分子的计算设计,近些年对环肽类化合物的生物膜透过性的研究也取得了进展。表4总结了这些环肽分子结构建模和基于靶标结构进行设计的方法。基于靶标结构进行环肽设计的关键是能够同时准确高效地评估环肽分子结构构象的稳定性、环肽配体与蛋白质靶标结合强度。更加高效准确的环肽序列和构象采样算法、环肽与靶标蛋白质结合能的评估方法准确,将提高环肽分子计算设计的效率和成功率,从而推动环肽药物与功能化环肽分子的开发。另外,对已有的环肽药物的性质和功能机制进行研究和学习,以及将计算设计方法和噬菌体展示、酵母展示等高通量实验筛选技术相结合,是促进环肽药物开发的途径^[112]。

近年来以深度学习为代表的机器学习技术快速发展,在图像识别、语音识别以及自然语言处理等领域取得突破。在蛋白质结构研究领域,以AlphaFold2为代表的基于深度学习技术的端对端方法展现出令人惊叹的由蛋白质序列预测折叠结构的效率和精度,显著超越了基于经验势能打分和采样的方法。结构设计是结构预测的反问题,机器学习技术正在推动蛋白质或多肽设计从基于物理势能函数的阶段到数据驱动的新阶段。例如目前基于深度扩散模型的RFdiffusion^[113]、Chroma^[114]等工作,已经实现了多种场景下的蛋白质(多肽)结构从头生成,基于靶标结构的小分子配体也可以通过深度学习模型从头生成,比如LiGANN^[115]、DeepLigBuilder^[116]等。这些生成模型的成功依赖于大量的已知三维蛋白质与复合物结构数据,基于机器学习的环肽结构生成算法或许可以借鉴蛋白质或多肽设计的相关算法,并通过引入蛋白质环区结构的数据以弥补环肽结构数据少的不足。另外,环肽设计领域未来的挑战是更加多样化的功能环肽的从头设计,首先需要进一步拓宽环肽的可设计化学空间^[117-121],比如引入非天然氨基酸、N-甲基化修饰与更丰富的环化方式等等^[122-124]。在设计更多可用环肽骨架的同时能够引入非天然的化学修饰改性,并对修饰后的分子进行有效模拟采样与结构优化,实现对环肽

表 4 本文所涉及的算法总体简介
Table 4 A brief introduction of algorithms included

算法名称	目标任务	测试体系或数据集	测试集规模	环肽序列长度	算法性能
MODPEP 2.0 ^[50]	仅 SS-环肽结构建模	非冗余 SSCP 库	193	约 30	前 100 构象与前 10 构象的平均 C _α RMSD 为 2.20 Å 和 1.66 Å, 优于 mETKDG 与 PEP-FOLD
Peplook ^[87]	环肽结构建模	环肽建模困难数据集	38	5~30	平均 BB-RMSD 为 3.8 Å, 优于 PEP-FOLD
PEPstrMOD ^[77]	环肽结构建模	CyclicPep	34	10~30	平均 C _α -RMSD 为 4.06 Å, 优于 PEP-FOLD
PEP-FOLD ^[80, 82-83]	环肽结构建模	环肽建模困难数据集	34	10~30	平均 RMS 为 2.75 Å
I-TASSER ^[76]	环肽结构建模	环肽建模困难数据集	35	10~30	平均 BB-RMSD 为 2.5 Å
mETKDG ^[52]	环肽结构构象生成	mc-PEP-set	2	10, 11	最优 BB-RMSD 分布均值分别为 0.60 Å, 1.27 Å
AutoDock CrankPep ^[96]	环肽配体对接	PDB 数据库中非冗余 环肽-蛋白复合物	CNCP: 18; SSCP: 20	CNCP: 6~14; SSCP: 6~20	对于 CNCP,SSCP 的最优平均 fnc 分别为 0.86 和 0.70
HADDOCK 2.4 ^[94]	环肽配体对接	PDB 数据库中非冗余 环肽-蛋白复合物	CNCP: 18; SSCP: 12	CNCP: 6~14; SSCP: 6~14	CNCP 的平均 RMSD 为 1.5~2.0 Å(略优于 AutoDock CrankPep),SSCP 的最优平均 RMSD 为 3.0 Å
DES3PI ^[105]	环肽配体从头设计	靶标蛋白: Ras, Mcl-1, Aβ protofibril	—	—	生成环肽配体的 AutoDock CrankPep 对接打分可以达到 -9~-12 kcal/mol
Rosetta	环肽单体理性 从头设计 ^[55]	—	—	7~14	生成了 200 种环肽结构, 其中部分进行了结构解析, BB-RMSD<1.6 Å
	环肽配体 从头设计 ^[34]	靶标蛋白: HDAC2, HDAC6	—	7~10	设计环肽配体实验测得最优 IC ₅₀ 为: HDAC2 9.1 nmol/L; HDAC6 5.4 nmol/L
	跨膜活性环肽 从头设计 ^[56]	—	—	6~12	生成了具有跨膜活性的环肽结构, 与实验解析结构的 BB-RMSD<1.5 Å, 具有跨膜活性与较好的口服利用率

注: SSCP 代表环化方式为侧链二硫键成环的环肽分子; CNCP 代表通过酰胺键主链首尾相连的环肽分子。

分子多样化设计与功能调控。此外, 自然界中的环肽分子不只是单环结构, 通过提取天然产物或基于较为传统的环肽合成方法即可产生双环肽、三环肽等较为复杂的主链环系拓扑结构, 这些额外的环状结构可以进一步提供空间约束以增强分子刚性, 但同时也大大增加了构象空间复杂性^[125-126]。然而目前对于环肽体系的计算研究大多停留于单环结构。这些需求与挑战为环肽设计领域未来的研究提供了契机。

参 考 文 献

[1] IVANOV A A, KHURI F R, FU H. Targeting protein-protein interactions as an anticancer strategy[J]. Trends in Pharmaceutical Sciences, 2013, 34(7): 393-400.

[2] SCOTT D E, BAYLY A R, ABELL C, et al. Small molecules, big targets: drug discovery faces the protein-protein interaction challenge[J]. Nature Reviews Drug Discovery, 2016, 15(8): 533-550.

[3] JONES S, THORNTON J M. Principles of protein-protein interactions[J]. Proceedings of the National Academy of Sciences of the United States of America, 1996, 93(1): 13-20.

[4] NOOREN I M A, THORNTON J M. Diversity of protein-protein interactions[J]. The EMBO Journal, 2003, 22(14): 3486-3492.

[5] REYNÈS C, HOST H, CAMPROUX A C, et al. Designing focused chemical libraries enriched in protein-protein interaction inhibitors using machine-learning methods[J]. PLoS Computational Biology, 2010, 6(3): e1000695.

[6] ROGNAN D. Rational design of protein-protein interaction inhibitors[J]. MedChemComm, 2015, 6(1): 51-60.

[7] MORELLI X, BOURGEAS R, ROCHE P. Chemical and structural lessons from recent successes in protein-protein interaction inhibition (2P2I)[J]. Current Opinion in Chemical Biology, 2011, 15(4): 475-481.

[8] SPERANDIO O, REYNÈS C H, CAMPROUX A C, et al. Rationalizing the chemical space of protein-protein interaction in-

- hibitors[J]. *Drug Discovery Today*, 2010, 15(5/6): 220-229.
- [9] LAGASSÉ H A, ALEXAKI A, SIMHADRI V L, et al. Recent advances in (therapeutic protein) drug development[J]. *F1000Research*, 2017, 6: 113.
- [10] CHENG S S, QU Y Q, WU J, et al. Inhibition of the CDK9-cyclin T1 protein-protein interaction as a new approach against triple-negative breast cancer[J]. *Acta Pharmaceutica Sinica B*, 2022, 12(3): 1390-1405.
- [11] NICOLAOU K C, BODDY C N, BRÄSE S, et al. Chemistry, biology, and medicine of the glycopeptide antibiotics[J]. *Angewandte Chemie International Edition*, 1999, 38(15): 2096-2152.
- [12] BODANSZKY M. Principles of Peptide Synthesis[M]. Heidelberg: Springer, 2012.
- [13] MERRIFIELD R B. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide[J]. *Journal of the American Chemical Society*, 1963, 85(14): 2149-2154.
- [14] BEHRENDT R, WHITE P, OFFER J. Advances in Fmoc solid-phase peptide synthesis[J]. *Journal of Peptide Science*, 2016, 22(1): 4-27.
- [15] AMBLARD M, FEHRENTZ J A, MARTINEZ J, et al. Methods and protocols of modern solid phase peptide synthesis[J]. *Molecular Biotechnology*, 2006, 33(3): 239-254.
- [16] CHAN W C, WHITE P D. Fmoc solid phase peptide synthesis: a practical approach[M]. New York: Oxford University Press, 2000.
- [17] LEADER B, BACA Q J, GOLAN D E. Protein therapeutics: a summary and pharmacological classification[J]. *Nature Reviews Drug Discovery*, 2008, 7(1): 21-39.
- [18] SANTOS G B, GANESAN A, EMERY F S. Oral administration of peptide-based drugs: beyond Lipinski's rule[J]. *ChemMedChem*, 2016, 11(20): 2245-2251.
- [19] DANHO W, SWISTOK J, KHAN W, et al. Opportunities and challenges of developing peptide drugs in the pharmaceutical industry[C]// The Proceedings of the 20th American Peptide Symposium-Peptides for youth: Advances in Experimental Medicine and Biology, 2009, 611: 467-469.
- [20] COOPER B M, IEGRE J, O'DONOVAN D H, et al. Peptides as a platform for targeted therapeutics for cancer: peptide-drug conjugates (PDCs)[J]. *Chemical Society Reviews*, 2021, 50(3): 1480-1494.
- [21] WANG J, YADAV V, SMART A L, et al. Stability of peptide drugs in the colon[J]. *European Journal of Pharmaceutical Sciences*, 2015, 78: 31-36.
- [22] VALEUR E, GUÉRET S M, ADIHOU H, et al. New modalities for challenging targets in drug discovery[J]. *Angewandte Chemie International Edition*, 2017, 56(35): 10294-10323.
- [23] BRUZZONI-GIOVANELLI H, ALEZRA V, WOLFF N, et al. Interfering peptides targeting protein-protein interactions: the next generation of drugs?[J]. *Drug Discovery Today*, 2018, 23(2): 272-285.
- [24] JENSEN K K, ANDREATTA M, MARCATILI P, et al. Improved methods for predicting peptide binding affinity to MHC class II molecules[J]. *Immunology*, 2018, 154(3): 394-406.
- [25] BEUMING T, FARID R, SHERMAN W. High-energy water sites determine peptide binding affinity and specificity of PDZ domains[J]. *Protein Science*, 2009, 18(8): 1609-1619.
- [26] ZHOU H X, GILSON M K. Theory of free energy and entropy in noncovalent binding[J]. *Chemical Reviews*, 2009, 109(9): 4092-4107.
- [27] MURPHY K P, XIE D, THOMPSON K S, et al. Entropy in biological binding processes: estimation of translational entropy loss[J]. *Proteins: Structure, Function, and Bioinformatics*, 1994, 18(1): 63-67.
- [28] VERTERAMO M L, STENSTRÖM O, IGNJATOVIĆ M M, et al. Interplay between conformational entropy and solvation entropy in protein-ligand binding[J]. *Journal of the American Chemical Society*, 2019, 141(5): 2012-2026.
- [29] EKBERG V, RYDE U. On the use of interaction entropy and related methods to estimate binding entropies[J]. *Journal of Chemical Theory and Computation*, 2021, 17(8): 5379-5391.
- [30] PAGE M I. Entropy, binding energy, and enzymic catalysis[J]. *Angewandte Chemie International Edition*, 1977, 16(7): 449-459.
- [31] HUIGENS III R W, MORRISON K C, HICKLIN R W, et al. A ring-distortion strategy to construct stereochemically complex and structurally diverse compounds from natural products[J]. *Nature Chemistry*, 2013, 5(3): 195-202.
- [32] YOON S, EOM G H. HDAC and HDAC inhibitor: from cancer to cardiovascular diseases[J]. *Chonnam Medical Journal*, 2016, 52(1): 1-11.
- [33] WEST A C, JOHNSTONE R W. New and emerging HDAC inhibitors for cancer treatment[J]. *The Journal of Clinical Investigation*, 2014, 124(1): 30-39.
- [34] HOSSEINZADEH P, WATSON P R, CRAVEN T W, et al. Anchor extension: a structure-guided approach to design cyclic peptides targeting enzyme active sites[J]. *Nature Communications*, 2021, 12(1): 3384.
- [35] KLING A, LUKAT P, ALMEIDA D V, et al. Targeting DnaN for tuberculosis therapy using novel griselimycins[J]. *Science*, 2015, 348(6239): 1106-1112.
- [36] DEMMER O, FRANK A O, HAGN F, et al. A conformationally frozen peptoid boosts CXCR4 affinity and anti-HIV activity[J]. *Angewandte Chemie International Edition*, 2012, 51(32): 8110-8113.
- [37] XIONG J P, STEHLE T, ZHANG R G, et al. Crystal structure of the extracellular segment of integrin $\alpha V \beta 3$ in complex with an Arg-Gly-Asp ligand[J]. *Science*, 2002, 296(5565): 151-155.
- [38] MIRANDA E, NORDGREN I K, MALE A L, et al. A cyclic

- peptide inhibitor of HIF-1 heterodimerization that inhibits hypoxia signaling in cancer cells[J]. *Journal of the American Chemical Society*, 2013, 135(28): 10418-10425.
- [39] BUCKTON L K, RAHIMI M N, MCALPINE S R. Cyclic peptides as drugs for intracellular targets: the next frontier in peptide therapeutic development[J]. *Chemistry*, 2021, 27(5): 1487-1513.
- [40] WANG Y, CHEETHAM A G, ANGACIAN G, et al. Peptide-drug conjugates as effective prodrug strategies for targeted delivery[J]. *Advanced Drug Delivery Reviews*, 2017, 110/111: 112-126.
- [41] CROXTALL J D, SCOTT L J. Lanreotide autogel®[J]. *Drugs*, 2008, 68(5): 711-723.
- [42] CAPLIN M E, PAVEL M, ĆWIKŁA J B, et al. Lanreotide in metastatic enteropancreatic neuroendocrine tumors[J]. *The New England Journal of Medicine*, 2014, 371(3): 224-233.
- [43] ANGLADE E, ASPESLET L J, WEISS S L. A new agent for the treatment of noninfectious uveitis: rationale and design of three LUMINATE (Lux Uveitis Multicenter Investigation of a New Approach to Treatment) trials of steroid-sparing voclosporin[J]. *Clinical Ophthalmology*, 2008, 2(4): 693-702.
- [44] KUGLSTATTER A, MUELLER F, KUSZNIR E, et al. Structural basis for the cyclophilin A binding affinity and immunosuppressive potency of E-ISA247 (voclosporin)[J]. *Acta Crystallographica Section D, Biological Crystallography*, 2011, 67 (Pt 2): 119-123.
- [45] HEO Y A. Voclosporin: first approval[J]. *Drugs*, 2021, 81(5): 605-610.
- [46] BECK J G, CHATTERJEE J, LAUFER B, et al. Intestinal permeability of cyclic peptides: common key backbone motifs identified[J]. *Journal of the American Chemical Society*, 2012, 134(29): 12125-12133.
- [47] MEIRELES L M C, MUSTATA G. Discovery of modulators of protein-protein interactions: current approaches and limitations[J]. *Current Topics in Medicinal Chemistry*, 2011, 11(3): 248-257.
- [48] WANG L, LI L, FU W T, et al. Optimization and bioevaluation of Cdc37-derived peptides: an insight into Hsp90-Cdc37 protein-protein interaction modulators[J]. *Bioorganic & Medicinal Chemistry*, 2017, 25(1): 233-240.
- [49] BEAUFAYS J, LINS L, THOMAS A, et al. *In silico* predictions of 3D structures of linear and cyclic peptides with natural and non-proteinogenic residues[J]. *Journal of Peptide Science*, 2012, 18(1): 17-24.
- [50] TAO H Y, WU Q L, ZHAO X J, et al. Efficient 3D conformer generation of cyclic peptides formed by a disulfide bond[J]. *Journal of Cheminformatics*, 2022, 14(1): 26.
- [51] RILEY B T, ILYICHOVA O, COSTA M G S, et al. Direct and indirect mechanisms of KLK4 inhibition revealed by structure and dynamics[J]. *Scientific Reports*, 2016, 6: 35385.
- [52] WANG S Z, WITEK J, LANDRUM G A, et al. Improving conformer generation for small rings and macrocycles based on distance geometry and experimental torsional-angle preferences[J]. *Journal of Chemical Information and Modeling*, 2020, 60(4): 2044-2058.
- [53] ANDREWS M D, DACK K N, DE GROOT M J, et al. Discovery of an oral, rule of 5 compliant, interleukin 17A protein-protein interaction modulator for the potential treatment of psoriasis and other inflammatory diseases[J]. *Journal of Medicinal Chemistry*, 2022, 65(13): 8828-8842.
- [54] SOUROJON M C, MOCHLY-ROSEN D. Peptide modulators of protein-protein interactions in intracellular signaling[J]. *Nature Biotechnology*, 1998, 16(10): 919-924.
- [55] HOSSEINZADEH P, BHARDWAJ G, MULLIGAN V K, et al. Comprehensive computational design of ordered peptide macrocycles[J]. *Science*, 2017, 358(6369): 1461-1466.
- [56] BHARDWAJ G, O'CONNOR J, RETTIE S, et al. Accurate *de novo* design of membrane-traversing macrocycles[J]. *Cell*, 2022, 185(19): 3520-3532.e26.
- [57] MCHUGH S M, ROGERS J R, YU H T, et al. Insights into how cyclic peptides switch conformations[J]. *Journal of Chemical Theory and Computation*, 2016, 12(5): 2480-2488.
- [58] GAVENONIS J, SHENEMAN B A, SIEGERT T R, et al. Comprehensive analysis of loops at protein-protein interfaces for macrocycle design[J]. *Nature Chemical Biology*, 2014, 10(9): 716-722.
- [59] DUFFY F J, O'DONOVAN D, DEVOCHELLE M, et al. Virtual screening using combinatorial cyclic peptide libraries reveals protein interfaces readily targetable by cyclic peptides[J]. *Journal of Chemical Information and Modeling*, 2015, 55(3): 600-613.
- [60] DUFFY F, MAHESHWARI N, BUCHETE N V, et al. Computational opportunities and challenges in finding cyclic peptide modulators of protein-protein interactions[M]. *Methods in Molecular Biology: Cyclic Peptide Design*. New York: Springer New York, 2019: 73-95.
- [61] CHE Y. Design of cyclic peptides as protein recognition motifs [M]//*Methods in Molecular Biology: Cyclic Peptide Design*. New York: Springer New York, 2019: 97-106.
- [62] SANTINI B L, ZACHARIAS M. Rapid *in silico* design of potential cyclic peptide binders targeting protein-protein interfaces[J]. *Frontiers in Chemistry*, 2020, 8: 573259.
- [63] SCOTT C P, ABEL-SANTOS E, WALL M, et al. Production of cyclic peptides and proteins *in vivo*[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1999, 96(24): 13638-13643.
- [64] GENG H, JIANG F, WU Y D. Accurate structure prediction and conformational analysis of cyclic peptides with residue-specific force fields[J]. *The Journal of Physical Chemistry Let-*

- ters, 2016, 7(10): 1805-1810.
- [65] LINDORFF-LARSEN K, PIANA S, PALMO K, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field[J]. *Proteins: Structure, Function, and Bioinformatics*, 2010, 78(8): 1950-1958.
- [66] JIANG F, ZHOU C Y, WU Y D. Residue-specific force field based on the protein coil library. RSFF1: modification of OPLS-AA/L[J]. *The Journal of Physical Chemistry B*, 2014, 118(25): 6983-6998.
- [67] ZHOU C Y, JIANG F, WU Y D. Residue-specific force field based on protein coil library. RSFF2: modification of AMBER ff99SB[J]. *The Journal of Physical Chemistry B*, 2015, 119(3): 1035-1047.
- [68] SLOUGH D P, YU H T, MCHUGH S M, et al. Toward accurately modeling *N*-methylated cyclic peptides[J]. *Physical Chemistry Chemical Physics*, 2017, 19(7): 5377-5388.
- [69] SUGITA Y, OKAMOTO Y. Replica-exchange molecular dynamics method for protein folding[J]. *Chemical Physics Letters*, 1999, 314(1/2): 141-151.
- [70] RHEE Y M, PANDE V S. Multiplexed-replica exchange molecular dynamics method for protein folding simulation[J]. *Biophysical Journal*, 2003, 84(2): 775-786.
- [71] MELI M, COLOMBO G. A Hamiltonian replica exchange molecular dynamics (MD) method for the study of folding, based on the analysis of the stabilization determinants of proteins[J]. *International Journal of Molecular Sciences*, 2013, 14(6): 12157-12169.
- [72] LAIO A, PARRINELLO M. Escaping free-energy minima[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2002, 99(20): 12562-12566.
- [73] PIANA S, LAIO A. A bias-exchange approach to protein folding[J]. *The Journal of Physical Chemistry B*, 2007, 111(17): 4553-4559.
- [74] TAYLOR R D, REY-CARRIZO M, PASSIOURA T, et al. Identification of nonstandard macrocyclic peptide ligands through display screening[J]. *Drug Discovery Today Technologies*, 2017, 26: 17-23.
- [75] YAN Y M, ZHANG D, HUANG S Y. Efficient conformational ensemble generation of protein-bound peptides[J]. *Journal of Cheminformatics*, 2017, 9(1): 59.
- [76] ZHANG Y. I-TASSER server for protein 3D structure prediction[J]. *BMC Bioinformatics*, 2008, 9: 40.
- [77] SINGH S, SINGH H, TUKNAIT A, et al. PEPstrMOD: structure prediction of peptides containing natural, non-natural and modified residues[J]. *Biology Direct*, 2015, 10(1): 73.
- [78] DAMAS J M, FILIPE L C S, CAMPOS S R R, et al. Predicting the thermodynamics and kinetics of helix formation in a cyclic peptide model[J]. *Journal of Chemical Theory and Computation*, 2013, 9(11): 5148-5157.
- [79] RAZAVI A M, WUEST W M, VOELZ V A. Computational screening and selection of cyclic peptide hairpin mimetics by molecular simulation and kinetic network models[J]. *Journal of Chemical Information and Modeling*, 2014, 54(5): 1425-1432.
- [80] THÉVENET P, SHEN Y M, MAUPETIT J, et al. PEP-FOLD: an updated *de novo* structure prediction server for both linear and disulfide bonded cyclic peptides[J]. *Nucleic Acids Research*, 2012, 40(W1): W288-W293.
- [81] BINETTE V, MOUSSEAU N, TUFFERY P. A generalized attraction-repulsion potential and revisited fragment library improves PEP-FOLD peptide structure prediction[J]. *Journal of Chemical Theory and Computation*, 2022, 18(4): 2720-2736.
- [82] ZHAO D X, LU K, LIU G B, et al. PEP-FOLD design, synthesis, and characteristics of finger-like polypeptides[J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2020, 224: 117401.
- [83] SHEN Y M, MAUPETIT J, DERREUMAUX P, et al. Improved PEP-FOLD approach for peptide and miniprotein structure prediction[J]. *Journal of Chemical Theory and Computation*, 2014, 10(10): 4745-4758.
- [84] MANDELL D J, COUTSIAS E A, KORTENME T. Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling[J]. *Nature Methods*, 2009, 6(8): 551-552.
- [85] COUTSIAS E A, SEOK C, WESTER M J, et al. Resultants and loop closure[J]. *International Journal of Quantum Chemistry*, 2006, 106(1): 176-189.
- [86] GO N, SCHERAGA H A. Ring closure and local conformational deformations of chain molecules[J]. *Macromolecules*, 1970, 3(2): 178-187.
- [87] THOMAS A, DESHAYES S, DECAFFMEYER M, et al. PepLook: an innovative in silico tool for determination of structure, polymorphism and stability of peptides[J]. *Advances in Experimental Medicine and Biology*, 2009, 611: 459-460.
- [88] CRAVEUR P, JOSEPH A P, ESQUE J, et al. Protein flexibility in the light of structural alphabets[J]. *Frontiers in Molecular Biosciences*, 2015, 2: 20.
- [89] JOSEPH A P, AGARWAL G, MAHAJAN S, et al. A short survey on protein blocks[J]. *Biophysical Reviews*, 2010, 2(3): 137-145.
- [90] ETCHEBEST C, BENROS C, HAZOUT S, et al. A structural alphabet for local protein structures: improved prediction methods[J]. *Proteins: Structure, Function, and Bioinformatics*, 2005, 59(4): 810-827.
- [91] THOMAS A, DESHAYES S, DECAFFMEYER M, et al. Prediction of peptide structure: how far are we?[J]. *Proteins: Structure, Function, and Bioinformatics*, 2006, 65(4): 889-897.
- [92] RINIKER S, LANDRUM G A. Better informed distance geometry: using what we know to improve conformation generation[J].

- Journal of Chemical Information and Modeling, 2015, 55(12): 2562-2574.
- [93] ZHANG Y Q, SANNER M F. Docking flexible cyclic peptides with *AutoDock CrankPep*[J]. Journal of Chemical Theory and Computation, 2019, 15(10): 5161-5168.
- [94] CHARITOU V, VAN KEULEN S C, BONVIN A M J J. Cyclization and docking protocol for cyclic peptide-protein modeling using HADDOCK_{2.4}[J]. Journal of Chemical Theory and Computation, 2022, 18(6): 4027-4040.
- [95] DUFFY F J, VERNIERE M, DEVOCELLE M, et al. CycloPs: generating virtual libraries of cyclized and constrained peptides including nonnatural amino acids[J]. Journal of Chemical Information and Modeling, 2011, 51(4): 829-836.
- [96] ZHANG Y Q, SANNER M F. AutoDock CrankPep: combining folding and docking to predict protein-peptide complexes[J]. Bioinformatics, 2019, 35(24): 5121-5127.
- [97] FORLI S, HUEY R, PIQUE M E, et al. Computational protein-ligand docking and virtual drug screening with the AutoDock suite[J]. Nature Protocols, 2016, 11(5): 905-919.
- [98] LENSINK M F, BRYLSBAERT G, MAURI T, et al. Prediction of protein assemblies, the next frontier: The CASP14-CAPRI experiment[J]. Proteins, 2021, 89(12): 1800-1823.
- [99] LASHKOV A A, TOLMACHEV I V, EISTRIKH-HELLER P A, et al. P₃FepRestr: Plugin to PyMOL molecular graphics system for calculating the free energy of ligand-receptor binding[J]. Crystallography Reports, 2021, 66(5): 861-865.
- [100] MOZAFFARI S, SALEHI D, MAHDIPOOR P, et al. Design and application of hybrid cyclic-linear peptide-doxorubicin conjugates as a strategy to overcome doxorubicin resistance and toxicity[J]. European Journal of Medicinal Chemistry, 2021, 226: 113836.
- [101] YEDVABNY E, NERENBERG P S, SO C, et al. Disordered structural ensembles of vasopressin and oxytocin and their mutants[J]. The Journal of Physical Chemistry B, 2015, 119(3): 896-905.
- [102] SPITALERI A, GHITTI M, MARI S, et al. Use of metadynamics in the design of isoDGR-based $\alpha\beta 3$ antagonists to fine-tune the conformational ensemble[J]. Angewandte Chemie International Edition, 2011, 50(8): 1832-1836.
- [103] KIM H J, BAE S C. Histone deacetylase inhibitors: molecular mechanisms of action and clinical trials as anti-cancer drugs[J]. American Journal of Translational Research, 2011, 3(2): 166-179.
- [104] RITO M, SILVA F, FELIX A. Histone Deacetylase (HDAC) inhibitors as potential anticancer drugs in Neuroendocrine Carcinomas (NEC) of cervix and endometrium[J]. Virchows Archiv, 2015, 467: S 15-S16.
- [105] DELAUNAY M, HA-DUONG T. Des3PI: a fragment-based approach to design cyclic peptides targeting protein-protein interactions[J]. Journal of Computer-Aided Molecular Design, 2022, 36(8): 605-621.
- [106] SINDHIKARA D, BORRELLI K. High throughput evaluation of macrocyclization strategies for conformer stabilization[J]. Scientific Reports, 2018, 8: 6585.
- [107] WHITTY A, ZHONG M Q, VIARENGO L, et al. Quantifying the chameleonic properties of macrocycles and other high-molecular-weight drugs[J]. Drug Discovery Today, 2016, 21(5): 712-717.
- [108] AHLBACH C L, LEXA K W, BOCKUS A T, et al. Beyond cyclosporine A: conformation-dependent passive membrane permeabilities of cyclic peptide natural products[J]. Future Medicinal Chemistry, 2015, 7(16): 2121-2130.
- [109] WANG C K, SWEDBERG J E, HARVEY P J, et al. Conformational flexibility is a determinant of permeability for cyclosporin[J]. The Journal of Physical Chemistry B, 2018, 122(8): 2261-2276.
- [110] ROSSI SEBASTIANO M, DOAK B C, BACKLUND M, et al. Impact of dynamically exposed polarity on permeability and solubility of chameleonic drugs beyond the rule of 5[J]. Journal of Medicinal Chemistry, 2018, 61(9): 4189-4202.
- [111] ALFORD R F, LEAVER-FAY A, JELIAZKOV J R, et al. The Rosetta all-atom energy function for macromolecular modeling and design[J]. Journal of Chemical Theory and Computation, 2017, 13(6): 3031-3048.
- [112] MUTTENTHALER M, KING G F, ADAMS D J, et al. Trends in peptide drug discovery[J]. Nature Reviews Drug Discovery, 2021, 20(4): 309-325.
- [113] WATSON J L, JUERGENS D, BENNETT N R, et al. Broadly applicable and accurate protein design by integrating structure prediction networks and diffusion generative models[EB/OL]. bioRxiv, 2022[2022-12-20]. <https://www.biorxiv.org/content/10.1101/2022.12.09.519842v1>.
- [114] INGRAHAM J, BARANOV M, COSTELLO Z, et al. Illuminating protein space with a programmable generative model [EB/OL]. bioRxiv, 2022[2022-12-20]. <https://www.biorxiv.org/content/10.1101/2022.12.01.518682v1>.
- [115] SKALIC M, SABBADIN D, SATTAROV B, et al. From target to drug: generative modeling for the multimodal structure-based ligand design[J]. Molecular Pharmaceutics, 2019, 16(10): 4282-4291.
- [116] LI Y B, PEI J F, LAI L H. Structure-based *de novo* drug design using 3D deep generative models[J]. Chemical Science, 2021, 12(41): 13664-13675.
- [117] SUBTELNY A O, HARTMAN M C T, SZOSTAK J W. Ribosomal synthesis of *N*-methyl peptides[J]. Journal of the American Chemical Society, 2008, 130(19): 6131-6136.
- [118] KAWAKAMI T, MURAKAMI H, SUGA H. Messenger RNA-programmed incorporation of multiple *N*-methyl-amino acids into linear and cyclic peptides[J]. Chemistry & Biology, 2008,

- 15(1): 32-42.
- [119] GOTO Y, MURAKAMI H, SUGA H. Initiating translation with D-amino acids[J]. RNA, 2008, 14(7): 1390-1398.
- [120] FUJINO T, GOTO Y, SUGA H, et al. Reevaluation of the D-amino acid compatibility with the elongation event in translation[J]. Journal of the American Chemical Society, 2013, 135(5): 1830-1837.
- [121] ACHENBACH J, JAHNZ M, BETHGE L, et al. Outwitting EF-Tu and the ribosome: translation with d-amino acids[J]. Nucleic Acids Research, 2015, 43(12): 5687-5698.
- [122] HUANG Y J, LIU T. Therapeutic applications of genetic code expansion[J]. Synthetic and Systems Biotechnology, 2018, 3(3): 150-158.
- [123] YOUNG D D, SCHULTZ P G. Playing with the molecules of life[J]. ACS Chemical Biology, 2018, 13(4): 854-870.
- [124] ARRANZ-GIBERT P, VANDERSCHUREN K, ISAACS F J. Next-generation genetic code expansion[J]. Current Opinion in Chemical Biology, 2018, 46: 203-211.
- [125] LOMBARDI A, MARASCO D, MAGLIO O, et al. Miniaturized metalloproteins: application to iron-sulfur proteins[J]. Proceedings of the National Academy of Sciences of the United States of America, 2000, 97(22): 11922-11927.

- [126] QUARTARARO J S, ESHELMAN M R, PERARO L, et al. A bicyclic peptide scaffold promotes phosphotyrosine mimicry and cellular uptake[J]. Bioorganic & Medicinal Chemistry, 2014, 22(22): 6387-6391.



通讯作者: 张长胜(1981—),男,副研究员。研究方向为蛋白质设计、计算结构生物学。

E-mail: changshengzhang@pku.edu.cn



第一作者: 王凡灏(1998—),男,博士研究生。研究方向为基于靶标结构的多肽药物分子设计。

E-mail: wangfh2020@stu.pku.edu.cn