

# 通过人工智能设计的 Talin 靶向肽结合物 调控细胞黏附斑功能

Modulating cell focal adhesion function by AI designed peptide  
binders for talin

一级学科： 物理学

研究方向： 生物物理

作者姓名：

指导教师：

答辩日期	20 年 月 日		
答辩委员会	姓名	职称	工作单位
主席			
委员			

天津大学福州国际联合学院

二〇二五 年 六 月

## 摘要

本文围绕细胞迁移与黏着斑功能调控，设计了人工智能（AI）辅助设计的 Talin 蛋白调控肽结合物（binder），并验证其与 talin 的结合力。Talin 作为连接整合素与肌动蛋白细胞骨架的核心力敏蛋白，其结构域 R3 对机械力高度敏感，构象变化直接影响黏着斑的组装与稳定性。为实现对 Talin-R3 结构域构象状态的特异性调控，我们借助 AlphaFold2、RFdiffusion 与 ProteinMPNN 等 AI 工具，设计了可稳定结合 Talin-R3 折叠态的肽结合物。通过荧光偏振实验、磁镊单分子力谱测量、TIRF 显微成像及细胞迁移实验，验证了 binder 对 Talin 结构及细胞黏着功能的调控效应。为应对 binder 大规模设计带来的验证挑战，我们构建了基于 SMMS（Single-Molecule Mechanical Selection）的高通量 PPI 检测平台，能够实现同时对上百种候选 binder 的快速、定量筛选，并具备高灵敏度、低成本与可扩展性等优势。本研究提出了研究靶向力敏蛋白功能的新方案，为 AI 驱动蛋白设计在细胞力学调控中的应用提供了新方法，有望在抗肿瘤转移、组织工程与药物开发等领域发挥广泛作用。

**关键词：**黏着斑；Talin 蛋白；肽结合物；人工智能蛋白设计；SMMS 高通量筛选平台

## 目录

第 1 章 绪论.....	1
1.1 引言.....	1
1.2 细胞迁移与黏着斑.....	2
1.2.1 细胞迁移的机制.....	2
1.2.2 黏着斑的组成、结构和作用机制.....	3
1.2.3 调节黏着斑功能的研究策略概述.....	4
1.3 作为力传感器的 Talin.....	6
1.3.1 Talin 的结构和机制.....	6
1.3.2 Talin 介导的黏附斑中的力转导.....	7
1.3.3 设计针对 Talin 蛋白的肽结合物.....	8
1.4 人工智能蛋白质工程.....	10
1.4.1 AI 设计 Talin 肽结合物.....	10
1.4.2 分子动力学模拟.....	11
1.5 SMMS pull down 平台.....	12
1.5.1 高通量蛋白-蛋白相互作用检测平台.....	12
1.5.2 AI 设计的结合分子的高通量验证平台.....	13
第 2 章 实验方法及原理.....	15
2.1 细胞及细菌实验.....	15
2.1.1 分子克隆与蛋白表达.....	15
2.1.2 荧光偏振实验.....	16
2.1.3 细胞培养与转染.....	18
2.1.4 细胞迁移实验.....	18
2.2 细胞黏附斑的荧光成像分析.....	19
2.2.1 免疫染色实验.....	19
2.2.2 全内反射荧光显微镜 (TIRF) 成像.....	19
2.3 单分子磁镊实验.....	20
2.3.1 磁镊实验的装置与原理.....	21
2.3.2 力的定量标定.....	22
2.3.3 磁镊实验中的典型测量内容.....	23
2.4 传统 pull-down 实验与蛋白质印迹 (Western Blot) 分析.....	23
2.5 基于 SMMS 的 Pull-down 实验用于蛋白-蛋白相互作用 (PPI) 检测.....	25
第 3 章 验证 AI 设计的肽结合物功能.....	27

3.1 分子动力学 (MD) 模拟验证结合物 .....	27
3.2 细胞体外验证结合物有效性 .....	28
3.2.1 Binder 以中等亲和力与 Talin R3 结合 .....	28
3.2.2 Talin R3 的拉伸曲线分析 .....	29
3.3 细胞体内验证结合物有效性 .....	31
3.3.1 引入 Binder 减少细胞黏附斑尺寸 .....	32
3.3.2 引入 Binder 后细胞迁移速度下降 .....	33
3.3.3 体内 Pull-down 实验 .....	34
第 4 章 开发高通量结合物筛选平台 .....	35
4.1 基于 SMMS 的高灵敏度蛋白 - 蛋白相互作用定量分析 .....	35
4.1.1 探索蛋白 - 蛋白相互作用的最优检测环境 .....	35
4.1.2 检测常见标签的蛋白 - 蛋白相互作用 (PPI) .....	39
4.1.3 构建用于验证 AI 设计 Binder 的高通量筛选平台 .....	41
4.2 结果讨论 .....	45
4.2.1 小结 .....	45
4.2.2 局限性与展望 .....	47
4.2.3 未来潜力 .....	48
第 5 章 总结 .....	49
参考文献 .....	50
发表论文和参加科研情况说明 .....	50
致谢 .....	56

## 第1章 绪论

### 1.1 引言

作为从简单的单细胞生物（如变形虫）到复杂的多细胞生物（如哺乳动物）中均普遍存在的生物学行为<sup>[1]</sup>，细胞迁移（cell migration）在其中至关重要。在胚胎发育过程中，细胞迁移在原肠胚形成、神经系统建构等关键形态发生阶段反复出现<sup>[2-3]</sup>。在成年个体中，迁移亦在多种生理与病理过程中发挥重要作用，例如在炎症反应中，白细胞定向迁移至损伤区域以发挥吞噬与免疫功能<sup>[4-5]</sup>；成纤维细胞与血管内皮细胞的迁移则是伤口愈合不可或缺的环节<sup>[6-7]</sup>。此外，在肿瘤转移过程中，癌细胞通过迁移实现从原发灶脱离、进入循环系统并在远端组织定植的过程<sup>[8]</sup>。近年来，细胞迁移在组织工程、再生医学等领域的研究与应用愈发受到重视，例如细胞在生物材料支架上的定植能力直接影响构建组织的结构与功能<sup>[9]</sup>。正因细胞迁移在生命过程中扮演着如此多样且关键的角色，过去几十年来相关调控机制的研究不断深入，逐渐形成包括单细胞迁移与集体迁移<sup>[10]</sup>、三维迁移机制<sup>[11]</sup>、细胞骨架动力学<sup>[12]</sup>、迁移形态学<sup>[13]</sup>以及黏着斑调控等多个重要子方向。

在众多参与细胞迁移的分子机制中，黏着（附）斑（focal adhesions, FAs）是细胞感知并响应细胞外基质（ECM）机械信号的核心枢纽。它不仅作为物理锚点介导细胞附着，还整合细胞内外的机械与化学信号，从而驱动迁移行为，并在细胞增殖、分化等过程中发挥调控作用<sup>[14-15]</sup>。随着高分辨成像与力学测量技术的发展，研究人员发现黏着斑的结构特征（如大小、密度、分子组成）与细胞迁移的速度、方向性和群体协调性密切相关<sup>[16]</sup>。因此，调控黏着斑功能不仅有助于理解基础细胞行为，也在抗肿瘤转移、组织工程等应用领域具有重要前景。目前，调节黏着斑功能的研究手段不断丰富，包括信号通路干预、小分子抑制剂、基因编辑、光遗传学调控及蛋白-蛋白相互作用干预等方法（详见第1.2节）。然而，针对特定力敏蛋白，如Talin等，对其特异性、可控性调节手段的研究仍处于初步阶段，从而限制了在精细调控黏着斑结构与功能上的进一步突破。

近年来，人工智能（AI）驱动蛋白质工程迅速发展，特别是AlphaFold2、RF Diffusion、ProteinMPNN等工具的出现，为特异性结合蛋白或短肽（binder）的结构预测与功能设计提供了新路径<sup>[17-19]</sup>。这些方法有望以更高效的方式，开发

出针对组成黏着斑复合蛋白的关键力响应蛋白如 **Talin** 等的高亲和力调节分子，为实现对黏着斑的精准操控提供新手段。本章将围绕以下几个方面展开：首先回顾细胞迁移和黏着斑的基本机制与关键组成成分，并讨论当前调节黏着斑功能的研究策略（第 1.2 节）；随后，我们聚焦于 **Talin** 蛋白，分析其结构、功能与力学特性，并探讨通过设计结合 **Talin** 的短肽分子调控黏着斑功能的可能性（第 1.3 节）；然后介绍实现该目标所采用的 AI 辅助设计工具（第 1.4 节）；最后，为了推进 AI 设计结合物的方法广泛应用，我们提出了采用 **SMMS**（Single-molecule-mechanical-selection）来实现一个针对 AI 设计的蛋白的高通量筛选平台（第 1.5 节）。

## 1.2 细胞迁移与黏着斑

### 1.2.1 细胞迁移的机制

细胞迁移是一个由多个步骤协调进行的动态过程，包括极性建立、前缘伸展、黏附形成、细胞体推进与后缘解离<sup>[1]</sup>。迁移的起始阶段，细胞首先建立前后极性，并在前缘通过肌动蛋白聚合作用形成伪足（如板状伪足 **lamellipodia** 和丝状伪足 **filopodia**），由此推动细胞膜向外延展。随后，细胞通过整合素和黏着斑与细胞外基质形成锚定点，为后续力的传递提供物理基础<sup>[20]</sup>。细胞体推进主要依赖肌动蛋白-肌球蛋白（**actomyosin**）系统所产生的收缩力，由 **RhoA-ROCK** 通路调控的应力纤维结构提供牵引力，将胞体向前拉动<sup>[21]</sup>。与此同时，后缘黏着斑解离，使得细胞后部能够顺利脱附，保证迁移效率。整个过程中，小 GTP 酶家族（如 **Rac1**、**Cdc42** 和 **RhoA**）通过调控肌动蛋白动力学与黏附的形成与解离，确保迁移各阶段协调推进<sup>[22]</sup>。

不同类型的细胞可采用多种迁移模式，包括依赖强黏附的间充质迁移、快速穿行但黏附依赖较弱的变形虫式迁移，以及保持细胞间连接的集体迁移<sup>[23]</sup>。其中，间充质型迁移对黏着斑高度依赖，细胞通常呈拉伸状态，通过“黏附形成—力传递—黏附解离”的循环进行运动。黏着斑在前缘的组装提供了有效的牵引力，而其在后缘的解离保证了细胞的持续推进<sup>[16]</sup>。

另外，黏着斑不仅是细胞与 **ECM** 之间的物理连接点，更是信号传导平台，它能够整合机械力和生化信号，以调控迁移的方向性、速度与稳定性<sup>[24]</sup>。因此，解析黏着斑的结构与调控机制，对于深入理解细胞迁移过程具有重要意义。本小节之后将专门讨论黏着斑的组成、功能以及调节策略。

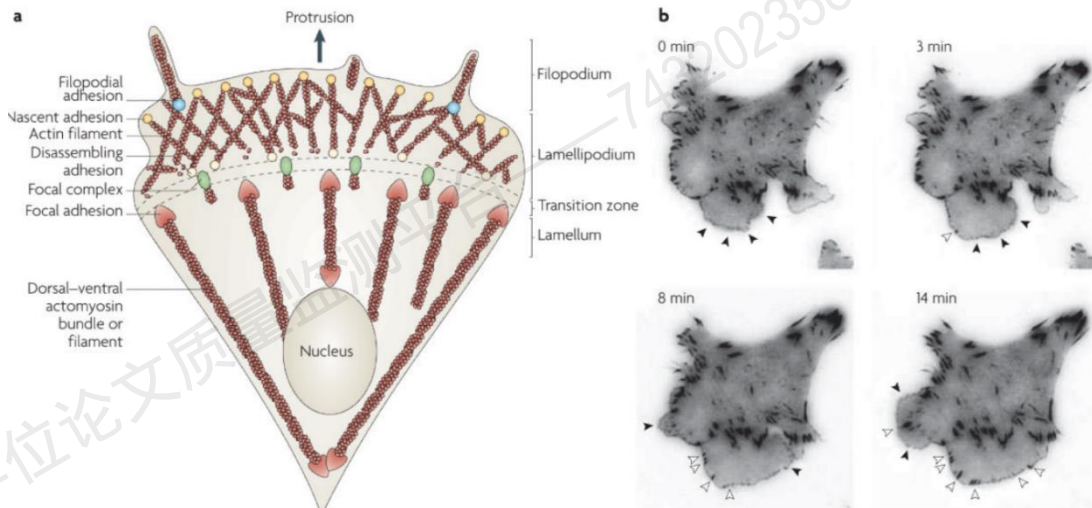


图 1.1 正在迁移的细胞的结构组成 (Adapted from Parsons et al., 2010)。(a) 黏附与细胞前缘突出结构 (丝状伪足 filopodia 和板状伪足 lamellipodia) 的形成密切相关。最初的黏着 (新生黏着) 形成于板状伪足中 (尽管黏着也可能与丝状伪足有关), 新生黏附的组装速率与突出结构的形成速率相关。新生黏附要么在板状伪足与板层 (lamellum, 即过渡区域) 交汇处解体, 要么延伸并成熟。黏着从新生阶段向黏着复合体 (focal complexes) 和黏着斑 (focal adhesions) 成熟的过程中, 伴随着肌动蛋白丝的束化与交联, 肌动蛋白-肌球蛋白 (actomyosin) 介导的收缩力可稳定黏附形成并增大其面积。(b) 表达 paxillin-mEGFP 的 CHO 细胞贴附在纤连蛋白 (fibronectin) 上的 TIRF (全内反射荧光) 显微镜图像。0、3、8 和 14 分钟的图像捕捉了新生黏附的周转与成熟过程。实心箭头标记新生黏附, 空心箭头则指示正在朝细胞中心延伸的成熟黏着结构<sup>[15]</sup>。

### 1.2.2 黏着斑的组成、结构和作用机制

黏着斑 (focal adhesions, FAs) 作为细胞与细胞外基质 (ECM) 之间的多蛋白连接结构, 是细胞黏附、力传导和信号整合的核心平台<sup>[14]</sup>。它通过多种细胞骨架和信号分子介导细胞的运动、分化、生长和存活。黏着斑的形成、结构与功能反映了细胞对周围环境物理与化学信号的感知和响应能力。

在组成上, 黏着斑是由上百种结构和信号蛋白组装而成的多分子复合物, 被称为“黏着复合体” (adhesome)<sup>[25]</sup>。核心蛋白包括整合素 (integrins)、Talin、Vinculin、Paxillin、FAK (focal adhesion kinase)、Kindlin、Zyxin 等 (图 1.2)。这些蛋白通过直接或间接方式连接细胞骨架与 ECM, 并参与力学调控和信号传导。整合素是最外层与 ECM 直接结合的跨膜受体, Talin 通过其 FERM 结构域激活整合素并连接 actin 细胞骨架<sup>[26]</sup>。Vinculin 可被张力激活, 与 Talin 和 actin 结合, 进一步增强黏着斑的机械稳定性<sup>[27]</sup>。在信号通路的激活中, FAK 和 Src 等激酶通过磷酸化作用调节 Paxillin 等蛋白的活性, 从而调控黏着斑的组装和解离<sup>[28]</sup>。

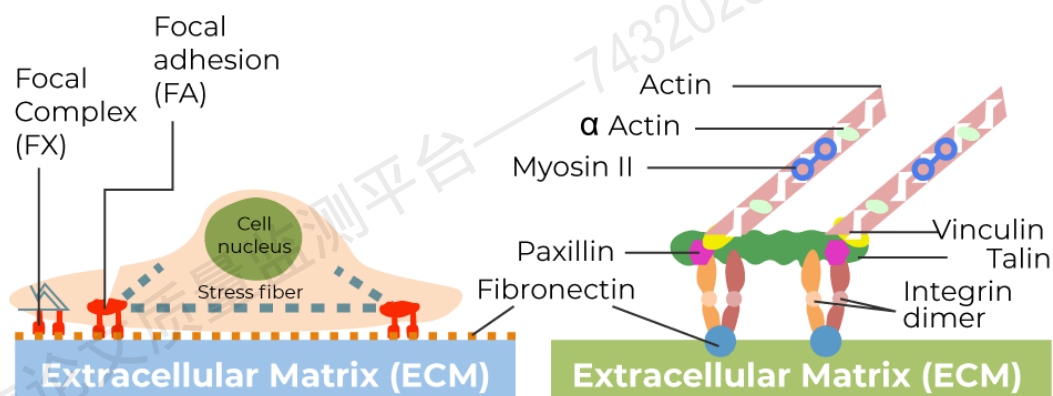


图 1.2 黏附斑的组成与结构示意图。左图显示黏附斑主要位于细胞与细胞外基质（ECM）之间的界面处。右图呈现了黏着斑组成成分以及它们之间的空间位置与结构。这些多蛋白复合体由一系列结构和信号蛋白组成，包括 paxillin、纤连蛋白（fibronectin）、整合素（integrin）、talin、vinculin 以及肌动蛋白（actin）等，它们共同介导细胞与 ECM 之间的黏附、力的传递以及力转导过程。

在结构层面，超高分辨成像显示黏着斑具有明显的垂直分层结构，包括底部的整合素层、中间的连接层（如 Talin、Vinculin）和上部的信号层（如 FAK、Paxillin）<sup>[29]</sup>，揭示了力与信号在空间上的传导路径。

功能上，黏着斑不仅传递细胞骨架产生的牵引力，还感知外部 ECM 的刚度、拓扑与化学属性。通过“外力-蛋白构象变化-信号传导”的路径，黏着斑能将力学输入转化为生化反应，从而调控基因表达、细胞生命行为与运动方向<sup>[30]</sup>。此外，黏着斑还具有动态性，其组装与解离受肌动蛋白牵引力、 $\text{Ca}^{2+}$ 信号、磷酸化事件等多因素调节<sup>[31]</sup>。由此可见，黏着斑不仅是静态的结构单元，而是一个动态响应的“力学-生化接口”。其核心蛋白的构象变化（如 Talin 在力加载下的解折叠）、相互作用强度的变化以及亚细胞定位的改变，均影响细胞的迁移能力与形态稳定性。

### 1.2.3 调节黏着斑功能的研究策略概述

细胞黏着斑（focal adhesions, FAs）作为连接细胞骨架与胞外基质（ECM）的关键机械和信号界面，其功能调控对于理解细胞行为与开发抗肿瘤策略具有重要意义<sup>[32]</sup>。近年来，针对黏着斑结构与功能的调控方法不断拓展，主要包括信号通路干预、小分子药物、基因调控技术、光遗传学方法、以及蛋白-蛋白互作抑制剂等。



在通过信号通路干预方案中,研究表明可以利用 Rho 家族小 GTP 酶(如 RhoA、Rac1、Cdc42)来调控黏着斑形成和细胞骨架动态重塑<sup>[33]</sup>。例如,RhoA-ROCK 信号轴控制应力纤维形成和黏着斑成熟<sup>[21]</sup>; Rac1 激活可促进新生黏着斑形成,支持细胞迁移前突的扩张<sup>[34]</sup>。此外,FAK/Src 通路通过磷酸化 Paxillin 和其他黏着斑蛋白,调节其组装、解离与下游信号活化<sup>[15,28]</sup>。在小分子药物干预方案中,研究人员基于上述信号调控轴开发了多种小分子靶向药物,如 FAK 抑制剂(PF-573228,VS-6063)、ROCK 抑制剂(Y-27632)、Src 抑制剂(SU6656)等。这些药物可抑制黏着斑形成或促进其解组装,被广泛应用于抑制肿瘤细胞迁移和侵袭<sup>[35-38]</sup>。在基因编辑与光遗传调控方案中,RNA 干扰(siRNA)与 CRISPR-Cas9 系统被用于特异性敲低 Talin、Vinculin、Kindlin 等核心 FA 蛋白,从而研究其在力传导与信号整合中的功能<sup>[39-41]</sup>。近年来,光遗传学手段也被引入细胞黏附调控,如利用光敏蛋白控制 RhoA/FAK 激活状态,进而调节黏着斑组装与细胞极性<sup>[42-43]</sup>。蛋白-蛋白互作抑制与合成肽干预方案近年来发展为针对 FA 蛋白间相互作用界面的抑制。例如,设计短肽 peptides 或小分子干扰 Talin 与整合素、Talin 与 Vinculin 的结合界面,已被证明可显著影响 FA 稳定性与细胞迁移行为<sup>[44-46]</sup>。Stapled peptides、macrocytic peptides 等具备结构稳定性强、亲和力高的优势,已成为靶向 FA 关键蛋白的研究热点<sup>[47-49]</sup>。

以上的这些调节黏着斑功能的方案各有优缺点:信号通路调控常涉及 Rho GTP 酶、FAK/Src 等多个通路,具有靶点丰富、机制明确的特点,但由于作用链条较长,往往缺乏特异性,如果作为药物副作用较大。小分子抑制剂如 FAK 或 ROCK 抑制剂具有合成简便、细胞通透性好等优点,但在靶向性与稳定性方面仍存在诸多问题<sup>[37]</sup>。基因干预手段(如 RNAi、CRISPR)可敲除或沉默特定 FA 蛋白,从源头上改变其表达水平,但其不可逆性、不可控性及时效限制使其在功能精细调节方面存在不足。光遗传学方法可实现对 FA 相关通路的时空精准调控,但对系统构建要求高,技术门槛较高,难以实现临床转化<sup>[42]</sup>。蛋白-蛋白相互作用抑制策略通过干扰特定蛋白界面比如 Talin-Integrin,具有较高靶向性,但其分子设计复杂,通透性与稳定性仍需进一步优化<sup>[47]</sup>。

综上所述,以上策略在特定场景中取得了较好效果,但如何以高特异性、良好可控性、明确结构基础的方式,直接调控核心力敏蛋白(如 Talin),仍是该领域极具挑战的前沿问题。近年来,伴随 AI 驱动蛋白结构预测与设计工具的快速发展,一种新兴策略逐渐浮现:通过 AI 辅助设计、开发可结合 Talin 蛋白的合成肽分子(binder)以调节其功能状态。该策略在理论上具备结构明确、作用位点精准、结合亲和力可控、序列优化空间大等多重优势,有望成为实现高精度调控

黏着斑功能的有效工具。我们将在下节介绍该方法的研究关注对象，Talin 和它所主导的力传导通路。

### 1.3 作为力传感器的 Talin

#### 1.3.1 Talin 的结构和机制

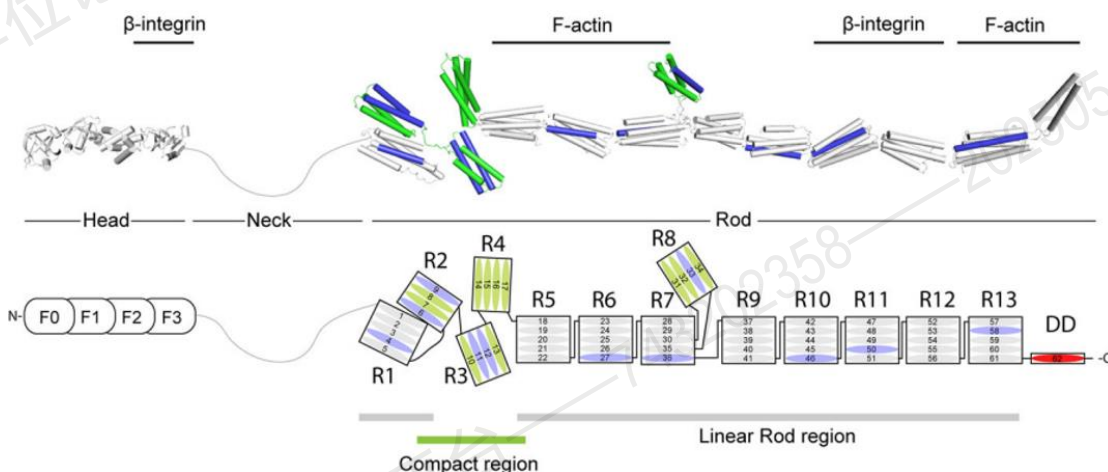


图 1.3 全长 Talin 的结构 (Adapted from Yan et al., 2015)。Talin 是一种大型多结构域蛋白，它由一个 N 端的“头部”，通过一个长且无结构的“颈部”区域（约 10 kDa）连接至一个大型的 C 端“杆状”区域（约 220 kDa）。其中，头部含有一个非典型的 FERM 结构域（约 50 kDa）。Talin 的杆状部分由 13 个螺旋束（R1–R13）组成，并以一个末端单螺旋结构结束，该结构支持 Talin 的二聚化。该杆状区域包含 11 个 vinculin 结合位点（VBS，图中以蓝色显示），这些位点全部埋藏在螺旋束内部<sup>[50]</sup>。

Talin 是生物系统中一种高度保守、约 270 kDa 的细胞骨架连接蛋白。作为整合素活化与黏着斑组装的关键桥梁分子，Talin 在细胞迁移、形态维持与力信号转导中发挥核心作用<sup>[51]</sup>。哺乳动物细胞中主要表达 Talin-1 和 Talin-2 两个同源亚型，Talin-1 广泛存在于迁移性细胞和肿瘤细胞（如 HeLa 细胞）中，在力学感应和细胞运动中尤其关键<sup>[52]</sup>。

如图 1.3 所示，Talin 的主要结构域由 N 端的 FERM 结构域（HEAD）（F0–F3）、C 端的 Rod 结构域（R1–R13）和一个二聚化结构域（DD）组成，N 端与 C 端借由一段不规则的链接区（NECK）连接<sup>[50]</sup>。其中，FERM 结构域通过 F3 亚结构域直接结合整合素  $\beta$  亚基胞质尾部，从而激活整合素构象变化，诱导其与 ECM 的高亲和结合状态<sup>[53]</sup>。此外，FERM 结构域还能与磷脂膜上的 PIP2 结合，协助 Talin 定位至质膜内侧<sup>[54]</sup>。Talin 的 Rod 结构域由一系列螺旋束组成，它们在

未受力时呈折叠闭合状态，多个亚结构域中包含 Vinculin 结合位点（VBSs）和 Actin 结合位点（ABS2/ABS3）<sup>[55]</sup>。在受到机械拉伸时，Rod 区域特定模体（如 R3、R9）发生部分或完全展开，暴露出这些隐蔽的相互作用位点，从而招募 Vinculin、Actin 等下游因子<sup>[56-57]</sup>。

Talin 激活过程遵循“自抑制 - 构象释放 - 活化组装”的多级调控模型。在细胞静息状态下，Talin 的 Rod 结构域与 FERM 域之间存在稳定的“自抑制”构象，阻止其与整合素和 Actin 等结合<sup>[58]</sup>。而在 PIP2 聚集、Rap1-GTP 上调或外部牵引力刺激下，Talin 构象被解锁，其 FERM 结构域转为开放状态，与整合素胞质尾结合并激活整合素，同时 Rod 区在拉伸下暴露 VBSs 并招募 Vinculin<sup>[59-60]</sup>。这一“力控构象开关”机制使 Talin 具备了力学调节与信号整合的双重能力。

### 1.3.2 Talin 介导的黏附斑中的力转导

Talin 作为黏着斑中的核心桥梁分子，其独特之处在于兼具机械连接功能与力学感应能力。Talin 被 PIP2 聚集、Rap1-GTP 上调或外部牵引力刺激下等因素激活后，它的自抑制构象被解锁，从而 FERM 结构和活化态的整合素尾部结合，将整合素从低亲和力状态转变为高亲和力构象，使其牢固结合 ECM<sup>[58]</sup>。同时，Talin 与质膜上的 PIP2 结合，协助其定位至黏附点。随着细胞骨架张力的增加，Talin 的杆状区折叠的螺旋束结构被逐步展开，暴露出多个隐藏的 Vinculin 结合位点（VBSs）<sup>[57]</sup>，从而招募大量 Vinculin 与之结合，进一步增强整合素-Talin 复合物与肌动蛋白细胞骨架之间的连接，并促进肌动蛋白的聚合<sup>[61-62]</sup>，从而增强黏附斑的稳定性，并强化细胞与 ECM 之间的黏附<sup>[63]</sup>。这种构象驱动的力响应过程也使 Talin 成为一个“机械信号开关”：其受力状态不仅决定了下游蛋白的招募，还反过来调节黏着斑的生长速度与方向性<sup>[64-65]</sup>。

单分子力谱实验与张力传感器研究表明，不同 Rod 亚结构域的展开力阈值不同（如 R3 与 R9 较易解折叠，R1 和 R11 较稳定），暗示 Talin 可通过分段展开方式调节不同功能位点的开放顺序，构建出一种“力控调控序列”<sup>[66-67]</sup>。此外，其构象变化还可能改变其与其他调控蛋白（如 RIAM、KANK、DLC1 等）的结合状态，进一步扩展其调控网络<sup>[68]</sup>。

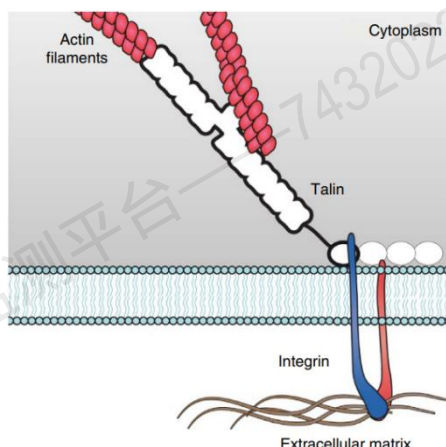


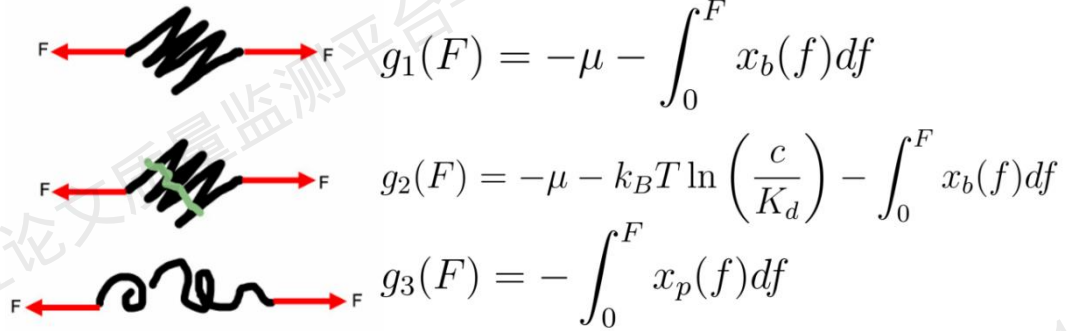
图 1.4 Talin 力传导结构示意图 (Adapted from Yao et al., 2016)。图中被标注为蓝色的细胞膜将细胞内 (上) 与细胞外基质 (下) 分隔。Talin 的头部域通过连接跨膜蛋白整合素和外界的 ECM 连接, 另一端杆状域 (白色) 连接到肌动蛋白细胞骨架 (红色)。该 Talin-Integrin-Actin 结构构成 FA 的力传导通路<sup>[64]</sup>。

Talin 构象变化与其功能状态存在显著对应关系。静息状态下, Talin 处于“自抑制折叠”构象, 仅少量与整合素结合, 黏着斑处于初期组装阶段; 而在强牵引力作用下, Talin 逐步解折叠, 暴露出多个功能位点, 黏着斑向成熟状态发展, 招募 Vinculin、Actin 等形成稳定结构<sup>[56]</sup>。若外部力降低或收缩力解除, Talin 重新折叠, VBSs 遮蔽, 黏着斑随之解体。这种构象-功能耦合机制使 Talin 成为连接细胞骨架与黏附界面之间的动态调控核心。Talin 所调控的力转导在包括心脏、肺、肝脏和肾脏等多个器官中起关键作用, 而其功能失调已被证明与多种疾病相关<sup>[30]</sup>。

### 1.3.3 设计针对 Talin 蛋白的肽结合物

目前, 针对如 Talin 等力传导蛋白的研究方法主要依赖于基因扰动技术, 如基因敲除、敲低和突变技术。然而, 考虑到 Talin 的力敏感特性, 迫切需要一种无需基因修改的机制来调节其结构域的力学稳定性。因此本课题提出了设计结合 Talin 结构域的调控肽分子结合物。在选取设计结合物对应的 Talin 的杆状结构域中, 基于对机械刺激的高敏感性, R3 被选为首要研究对象。与其他杆状结构相比, R3 在较低的力 (约 5 pN) 下便可快速发生展开和重新折叠, 意味着它是最先响应生理小力的结构域。此外, R3 含有两个 VBSs, 是最早招募 vinculin 的区域, 从而启动黏附加强过程<sup>[64]</sup>。那么, 设计一种专门与 R3 折叠态结合的肽结合物 (binder), 就能显著延长其处于折叠构象的时间, 使 vinculin 难以被激活, 从而导致下游的一系列力传导机制大大削弱, 比如减少了 vinculin 的招募, 削弱与

肌动蛋白的关联，进而降低黏附斑强度和尺寸。此策略在如降低肿瘤的迁移与转移能力等多个重要医学领域具有极高应用价值的潜力。



$$g_1(F) = -\mu - \int_0^F x_b(f)df$$

$$g_2(F) = -\mu - k_B T \ln\left(\frac{c}{K_d}\right) - \int_0^F x_b(f)df$$

$$g_3(F) = - \int_0^F x_p(f)df$$

图 1.5 结合物存在时 Talin R3 的三种状态。图中黑色为 Talin，绿色为 binder。从上到下三种状态分别是：未结合的折叠态、结合后的折叠态和展开态。对应的公式是这些状态在力  $F$  作用下的自由能。

下面简单介绍肽结合物 binder 将 Talin R3 锁定在其折叠状态的理论依据：当 binder 与杆状结构域相互作用时，存在三种可能的状态：未结合的折叠态、结合后的折叠态和展开态（如图 1.5 所示）。每种状态的自由能如图 1.5 右侧公式所示。在未结合的折叠态下，受力  $F$  时，其自由能为：

$$g_1(F) = -\mu - \int_0^F x_b(f)df \quad (1.1)$$

其中， $x_b(f)$ 是在力  $F$  下折叠的杆状结构域的伸长量， $\mu$  是杆状结构域的折叠能量。在结合后的折叠状态下，受力  $F$  时，其自由能为：

$$g_2(F) = -\mu - k_B T \ln\left(\frac{c}{K_d}\right) - \int_0^F x_b(f)df \quad (1.2)$$

其中， $K_d$ 是 binder 与杆状结构域结合的解离常数， $c$  是 binder 的浓度， $k_B$ 是玻尔兹曼常数， $T$ 是温度。对于展开态，自由能为：

$$g_3(F) = - \int_0^F x_p(f)df \quad (1.3)$$

其中， $x_p(f)$ 是在力  $F$  下展开的杆状结构域的伸长量。由于存在临界力  $F_c$ 使得杆状结构域在展开和折叠状态之间具有相等的概率，我们可以得到以下关系：

$$\int_0^{F_c} (x_p(f) - x_b(f))df = \mu + k_B T \ln\left(1 + \frac{c}{K_d}\right) \quad (1.4)$$

从式 1.4 中我们可以推断出，如果结合亲和力较强，即  $K_d$  值较低，临界力  $F_c$  将会更高。如果我们为 Talin 杆状结构域设计了一个强结合的 binder，我们可以显著提高展开杆状结构域所需的临界力。

## 1.4 人工智能蛋白质工程

近年来,人工智能(AI)在蛋白质结构预测和功能设计领域取得了突破性进展,极大推动了蛋白质工程的革新。尤其在 AlphaFold2 和 RoseTTAFold 等结构预测算法成功之后,蛋白质的三维结构获得了前所未有的解析速度和精度,使得蛋白功能预测与靶向设计成为可能<sup>[18,69]</sup>。同时,以扩散模型(diffusion models)和图神经网络(GNN)为基础的蛋白设计工具,如 RFdiffusion、ProteinMPNN 和 ESMfold 等,也逐渐在靶点结合肽、纳米抗体、人工酶等方向展示出强大的生成能力<sup>[19,70-71]</sup>。

### 1.4.1 AI 设计 Talin 肽结合物

AI 驱动蛋白设计流程一般包括:目标蛋白结构建模、功能位点识别、初始序列生成、结合界面优化和候选结构的物理-化学打分<sup>[17-19]</sup>。在本研究中,我们首先从 UniProt 数据库中获取 Talin R3 的氨基酸序列,并将其输入 AlphaFold2。AlphaFold2 会预测 Talin R3 结构域的三维结构。然后,R3 的三维结构的 PDB 文件将会作为结合对象输入给 RF Diffusion,随着扩散过程的进行,与 Talin R3 结合位点互补的高置信度的结合物骨架 backbone 得以生成。该过程确保所设计的蛋白 binder 能够良好地匹配 Talin 的分子表面,最大化相互作用潜力。随后,将 RF Diffusion 所生成的主链 backbone 输入到 ProteinMPNN,它将对结合界面进行序列优化,生成多种具有高结合亲和力和结构稳定性的候选肽序列。最后,使用 AlphaFold 对设计的结合物与 Talin R3 结构域的结合进行了测试。每个设计的结合物与 Talin R3 序列分别输入进行分析。计算了结合界面处的 PAE 值(iPAE),并将其作为评估结合物质量的标准。根据评分选出了排名靠前的序列进行后续的实验验证。

综上所述,这些 AI 工具协同工作,实现了对 Talin 靶向肽类 binder 的高效优化设计,接下来可通过分子动力学模拟和实验结合分析进一步验证其效果。这种 AI 辅助的工作流程加速了基于肽的治疗手段与力生物学研究的发展,提供了一种精准设计分子相互作用的强大手段。



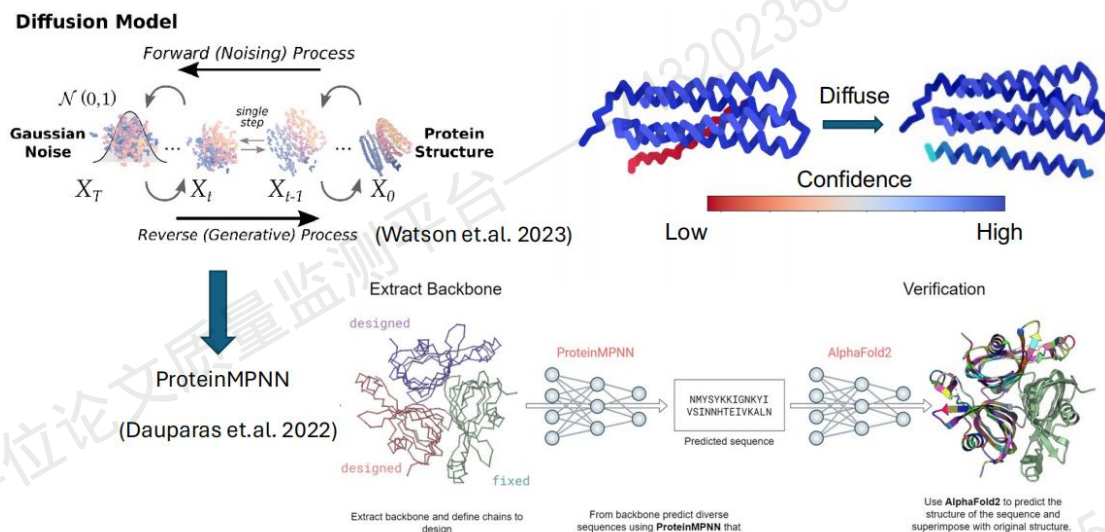


图 1.6: Talin R3 binder 设计流程 (Adapted from Dauparas et al., 2022, Watson et al., 2023)。以 AlphaFold2 预测的 Talin R3 结构作为 RF Diffusion 的输入，RF Diffusion 通过一个基于噪声的正向过程和一个生成式的反向过程生成全新的蛋白质主链结构。主链随后通过 diffusion 模型进行基于置信度的优化，以提高其与 Talin R3 的互补性。ProteinMPNN 用于为该主链设计氨基酸序列，使其能够准确折叠成目标结构，并确保高度稳定性和特异性。最终结构将通过 AlphaFold2 进行验证<sup>[17-19]</sup>。

### 1.4.2 分子动力学模拟

分子动力学(molecular dynamics, MD)模拟是一种基于经典力场的计算方法，广泛应用于蛋白质构象变化、稳定性分析及分子间相互作用的研究中。通过追踪原子随时间的运动轨迹，MD 能在原子水平揭示蛋白质在不同环境条件下的构象变化与动态行为，为结构预测和功能分析提供了重要补充<sup>[72]</sup>。

本研究中，我们拟利用 GROMACS 平台对 Talin-R3 与设计结合肽的复合结构进行 100 ns 经典 MD 模拟。首先，MD 可通过自由能计算(如 MM-PBSA/GBSA 或 FEP)来估算结合亲和力，从而量化 binder 与 Talin R3 之间的相互作用强度。此外，MD 还可通过分析模拟过程中形成的氢键、疏水相互作用以及盐桥等，识别关键的结合位点。模拟还可以评估 binder-Talin R3 复合物的稳定性，常用的评估参数包括 RMSD (均方根偏差) 及关键相互作用在时间上的持续性<sup>[73-74]</sup>。通过上述详细信息，MD 模拟可以辅助进一步筛选具有良好结构稳定性与结合能力的结合物的候选序列以用作后续的实验验证。

## 1.5 SMMS pull down 平台

在表达了肽结合物蛋白后，我们需对其结合 Talin 的能力做实验验证。更一般地，我们需要一个能够高通量测试蛋白-蛋白相互作用（PPI）的平台。AI 可以在一次运行中生成数十到数百个候选 binder，这对传统的 PPI 验证方法如 pull-down 实验构成了巨大挑战。这是因为传统的 PPI 检测方法耗时较长：在蛋白表达之后，pull-down 实验通常还需进行过夜孵育、洗脱、电泳、膜转印和免疫印迹（Western blot）等多个步骤，实验时间数以天计。而且这些方法的通量有限，通常一次最多只能检测十种蛋白，成为当前验证大规模 AI 设计的 binder 库的瓶颈。

为了解决这一难题，在完成对 AI 设计的 binder 的初步实验验证之后，我们进一步尝试将单分子力学筛选（SMMS, Single-Molecule Mechanical Selection）技术开发为一种高通量、定量的平台用于研究 PPI、验证 AI 设计的 binder，并测试了一些其他的蛋白的相互作用网络。SMMS 技术通过施加精确可控的机械力，仅保留高亲和力、特异性的相互作用，同时有效排除弱或非特异性相互作用<sup>[75]</sup>。这种内建的特异性筛选机制，结合其可同时并行筛选数百个蛋白对的潜力，有望克服传统方法的关键限制。

### 1.5.1 高通量蛋白-蛋白相互作用检测平台

SMMS 技术基于利用经过校准的机械力来剥离通过待测分析物连接在孔板表面的超顺磁性微珠。分析物可以是蛋白质，此时微珠和孔板表面分别预涂被两种能够结合该蛋白不同位点的特异性传感分子（Sensor），如抗体等。该检测方式可用于识别目标分析物的存在，在生物化学研究中具有重要意义，尤其在液体疾病生物标志物检测方面<sup>[75]</sup>。对于这类应用，SMMS 展现出多个优势，包括依赖机械门控的高特异性、单颗微珠级别的高灵敏度、检测时间快（约 30 分钟）、操作简便以及成本低廉。在本研究中，我们将扩展 SMMS 的应用，使其能够用于蛋白-蛋白相互作用（PPI）的检测。与现有主要用于检测单一蛋白不同，此处的分析物为两个蛋白形成的复合物。我们将在两个相互作用蛋白的末端分别融合不同标签，从而使微珠和孔板表面能够分别包被能够特异性识别这些标签的传感分子。PPI 的检测不仅对于理解生物功能至关重要，也有助于发现疾病生物标志物，并推动靶向结合分子的设计。



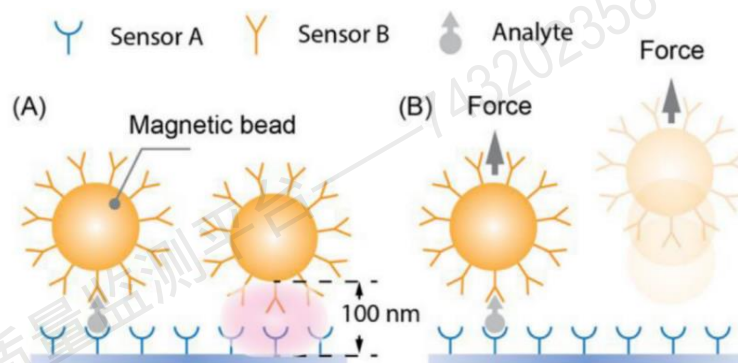


图 1.7 SMMS 原理示意图 (Adapted from Zhou et al., 2023)。 (A) 实验中，底板和磁珠表面分别修饰探针分 Sensor A 和 Sensor B，可分别识别分析物 (Analyte) 中的 Tag A 和 Tag B； (B) 当磁珠与底物通过 Analyte 介导形成连接后，施加磁力，磁力会将分析物种没有发生 PPI 的非特异性信号都去除。最终保留下来的微珠密度作为检测信号，通过普通光学显微镜即可进行定量分析<sup>[75]</sup>。

### 1.5.2 AI 设计的结合分子的高通量验证平台

在本工作中，我们将优化 SMMS 技术，实现对 AI 预测的蛋白结合分子进行高通量、定量验证。我们将建立 SMMS 实验流程，能够在 1–2 小时内验证超过 100 个 AI 设计的结合分子，实时对候选分子进行亲和力排名。SMMS 的性能将与 ELISA、Western blot 和 Co-IP 等传统方法进行对比评估，并建立自动化数据处理流程，将 SMMS 结果直接整合至 AI 驱动的结合分子设计平台中。

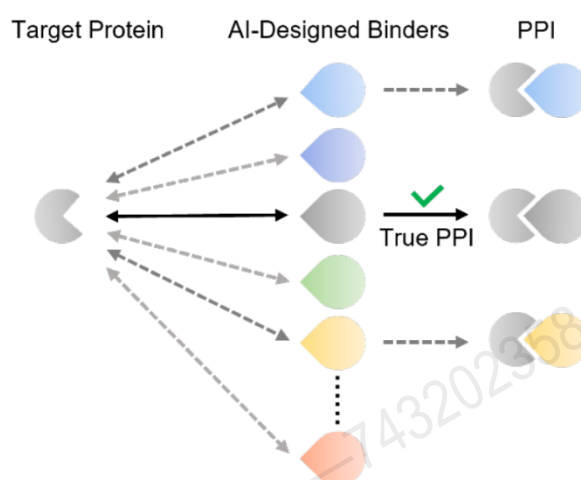


图 1.8: 一个由 AI 设计生成的大型结合分子库，其中只有一小部分是真正能与目标结合的有效分子。

蛋白设计的目标是识别具有期望结构和功能的新序列，其中一个重要应用是设计能够以高亲和力结合特定目标蛋白的结合分子。目前先进的 AI 蛋白设计方法通常需要从大量 AI 预测的候选分子中进行广泛的实验验证，以确保设计成功

<sup>[17-19]</sup>。在实际操作中，只有极少数 AI 生成的候选分子是真正有效的结合分子（见图 1.8）。大量候选分子迫切需要高通量筛选方法。而现有的常规方法如 Western blot 和 ELISA 通量低，扩展性差；酵母表面展示等高通量方法则依赖诸如流式细胞仪（FACS）等专用设备，限制了其普及性。我们将对 SMMS 技术进行优化，该技术具有高特异性、高灵敏度、低成本和快速处理等优势，用于建立一个高通量的结合分子筛选平台。