ATF3 同源二聚体选择性阻断突变体 —— 设计评估

目标：破坏 ATF3 同源二聚体，保留 ATF3 与其他 bZIP 蛋白的异源二聚体

1. 亮氨酸拉链 register（两种独立方法交叉验证）

ATF3 的拉链核心按 g-a-b-c-d-e-f 七肽重复排列。以保守的 Asn125（位于 a 位， 对应 GCN4 的 N16）为锚点，七肽编号 k = -2 ⋯ +2 :

k	g	a	b	c	d	e	f
-2	K110	T111	E112	C113	L114	Q115	K116
-1	E117	S118	E119	K120	L121	E122	S123
0	V124	N125	A126	E127	L128	K129	A130
+1	Q131	I132	E133	E134	L135	K136	N137
+2	E138	K139	Q140	H141	L142	I143	Y144

验证：

1. a 位 = T / S / N / I / K，与 Yu et al. 2024 文中对 ATF3 的描述完全一致 ("the ATF3 core a positions consist of T/S/K at a1, a2, and a5... only N and I at positions a3 and a4") 。
2. d 位全为 Leu（L114/121/128/135/142），完美 7 残基间隔。
3. 独立锚点校验：碱性区保守 Asn96 + 29 = 125，与 Leu 阶梯给出的锚点一致。

同一方法在 JUN 上给出 a = I/V/N/A/V，与文献 register 一致。

2. 决定配对特异性的是 g ↔ e′ 偶联

在平行二聚卷曲螺旋中，一条链 k 号七肽的 g 位与**对面链同一个** k 的 e′ 位 形成链间静电接触。这一点不是靠惯例假设，而是从 AlphaFold2 模型坐标测出来的： ATF3 g × 对面链 e′ 的 5×5 距离矩阵中，对角线元素在**每一行和每一列都是最小值** （2.9–7.4 Å，非对角线 8–40 Å） 。

野生型 ATF3 同源二聚体的 5 个偶联：

k	g · e′	相互作用
-2	K110 · Q115′	中性
-1	E117 · E122′	**排斥**
0	V124 · K129′	中性
+1	Q131 · K136′	中性
+2	E138 · I143′	中性

同源二聚体共 2 个 E–E 排斥（对称出现两次）、0 个吸引 —— 与论文 "the ATF3 target homodimer contains two electrostatic E–E repulsions" 完全吻合。 这为整套**打分函数**提供了外部校准（但不足以单独锁定 register —— 见 7.4） 。

相比之下 ATF3 : c-Jun 异源二聚体有 **4 个吸引配对**、**0 个排斥**，这解释了 ATF3 为何强烈偏好与 Jun 家族成员结合。

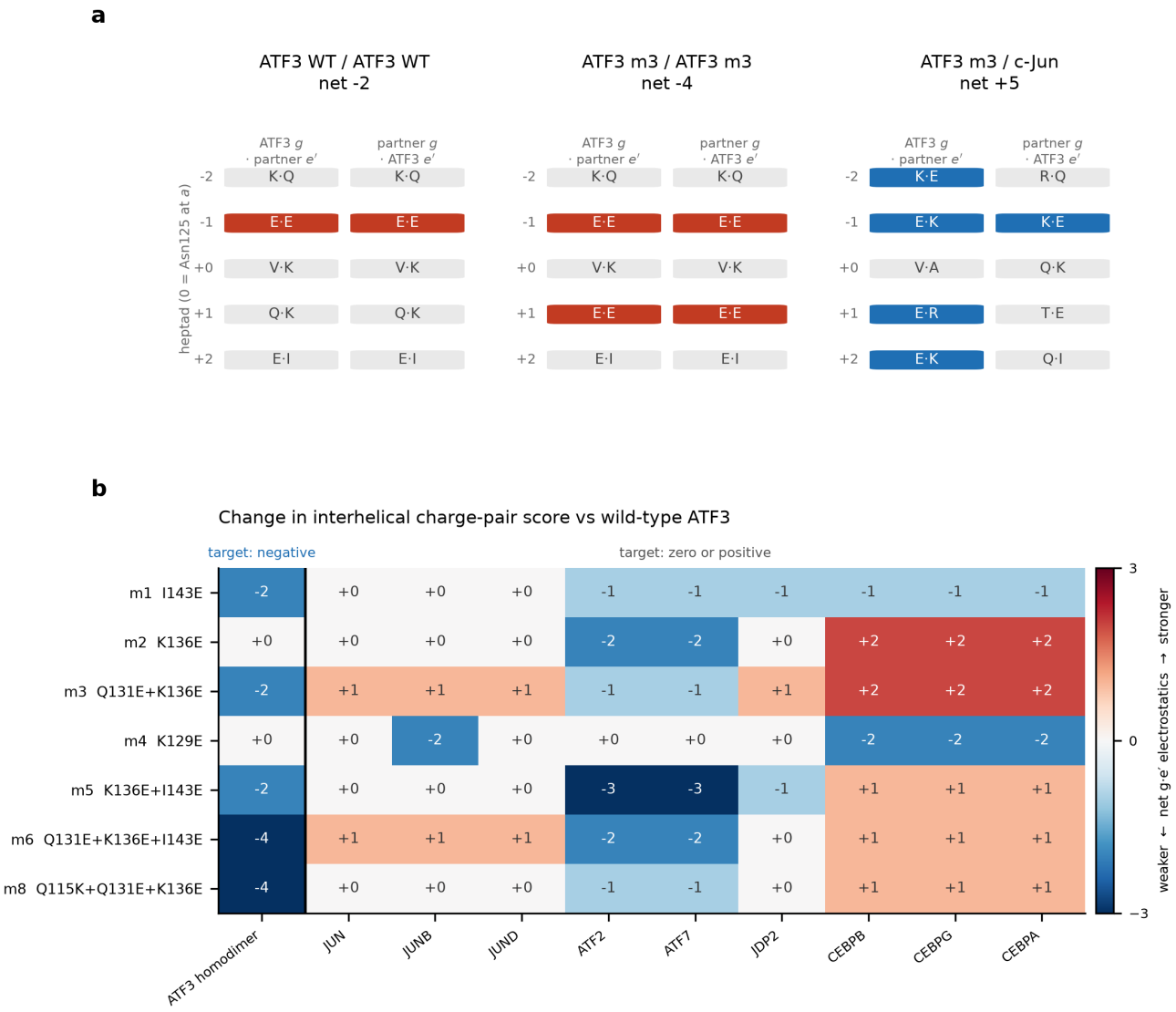


图 1 (a) g· e′ 偶联梯图。(b) 各突变体相对野生型的打分变化。

3. 核心设计原则

要在同源二聚体中制造排斥，必须让同一个 g· e′ 偶联的两端带同种电荷。

因为同源二聚体是对称的：只改 e 位时，它面对的仍是野生型的 g 位。 而在异源二聚体中，ATF3 的突变残基面对的是**伙伴蛋白未改变的**残基 —— 如果伙伴在该位带相反电荷，异源二聚体反而被加强。这正是同源/异源可被区分的 物理基础，也是 Yu et al. 用于筑选 ATF3W_aeg 的同一逻辑。

4. 五个突变体的评估

突变体	突变	位置	同源二聚体	Jun 家族	其他伙伴	结论
m3	Q131E + K136E	g(+1) + e(+1)	−2（新增 2 个 E131–E136′ 排斥）	+1（增强）	CEBP +2，JDP2 +1，ATF2/ATF7 −1	**五个中最优**（与 m1 在主判据上平手，靠次级判据胜出，见 7.3）
m1	I143E	e(+2)	−2（与已有 E138 形成 2 个排斥）	不变	ATF2/ATF7/JDP2/CEBP 各 −1	可行；主判据与 m3 平手，但损失 e 位 Ile 的疏水包装
m5	K136E + I143E	e(+1) + e(+2)	−2（与 m1 相同）	不变	ATF2/ATF7 −3	K136E 无额外贡献，只增加代价
m2	K136E	e(+1)	0（无变化）	不变	ATF2/ATF7 −2，CEBP +2	**达不到目的**（条件于 Gln 电中性，见 7.2）

突变体	突变	位置	同源二聚体	Jun 家族	其他伙伴	结论
m4	K129E	e(0)	0（无变化）	JUNB −2	CEBP −2	**达不到目的**（无条件成立，见 7.2）

- **m2 失效的原因**：e136 对面是 Q131（不带电）。E136 与 Q131 之间没有静电耦合，

对同源二聚体零影响。

- **m4 失效的原因**：e129 对面是 V124（疏水、不带电）。同理零影响，而且削弱了

与 JUNB（g(0)=E）和 CEBP（g(0)=R/K）的结合。

- **m3 成功的原因**：Q131E 与 K136E 同时改了 k=+1 偶联的两端 → 同源二聚体中

E131· E136′ 对称出现两次。而在异源二聚体里，ATF3-g131E 面对 Jun 的 R、CEBP 的 S， ATF3-e136E 面对 Jun 的 T、CEBP 的 R —— 所以与 Jun 和 CEBP 的结合反而更好。

5. 若需更强的效果：两个改进设计

在 10 个 e/g 位上穷举 E/K/Q 替换（单、双、三、四突变，共 6319 个方案）， 按"最大化同源二聚体排斥 / 最小化任一关键伙伴的损失"筛选：

设计	突变	同源二聚体	Jun 家族	最差伙伴	疏水位损失
m8（推荐）	Q115K + Q131E + K136E	−4（共 6 个排斥）	不变	−1（ATF2/ATF7）	无
m6	Q131E + K136E + I143E	−4（共 6 个排斥）	+1	−2（ATF2/ATF7）	1（I143）

m8 的第三个排斥来自 Q115K：e115 对面是已有的 K110 → K110· K115′ 排斥 ×2。 它不触碰任何疏水位，且 Jun 家族完全不受影响。

6. AlphaFold2-Multimer 结果

用 ColabFold 1.6.3 + AF2-Multimer v3 跑了 16 个复合物（WT + m1–m5 + m6/m8， 各做同源二聚体和 ATF3:c-Jun 异源二聚体），构建为 register 对齐的拉链肽段 （ATF3 101–152 / c-Jun 267–318，与论文所用合成肽范围相当）。

异源二聚体 —— 可用且结论明确：

- 全部 8 个模型：pLDDT 87.3–88.1，pTM 0.73，ipTM 0.72–0.73，recycle 0 即收敛。
- 所有变体之间 ipTM 极差仅 0.010。
- 设计的 g· e′ 配对在所有变体中都保持完好（两个方向分别为 2.7–7.5 Å 和 3.4–8.8 Å，

变体间差异 ⩽ 1 Å）。

- **→ 这五个突变（以及 m6/m8）都不会在结构上破坏 ATF3:c-Jun 界面。**

这是目标的"后半句"得到的直接支持。

同源二聚体 —— 不可用于排序，但本身是一条信息：

- ipTM 0.15–0.48，全部低于 0.5 的常规界面阈值。
- 关键问题：单次运行内 recycle 之间的 ipTM 波动达 0.19–0.25

（例如 m3：0.48 → 0.25 → 0.46），与变体之间的差异同量级。

- 8 个模型中有 **6 个被建成反平行**（链轴夹角余弦 ≈ −0.97），而 bZIP 是

专性平行二聚体 —— 即拓扑都是错的。

- **→ 不能用同源二聚体的 ipTM 给突变体排序**，我没有这样做。

有意义的是两者的**对比**：AF2 对 ATF3:c-Jun 立即收敛到高置信度，对 ATF3:ATF3 则给不出稳定解。这与论文的实验观察一致 —— ATF3 肽单独存在时无法自组装 （CD 只见上基线，SEC 只见单体峰）。

（注：加上碱性区的完整 bZIP 构建体 ATF3 86–155 确实给出平行、register 正确的 同源二聚体模型，pLDDT 81 / ipTM 0.41，第 1 节的 register 验证即基于此模型。 碱性区的共进化信号是把拓扑钉住的关键。）

为什么 AF2 本就不该用来排序这些突变体：AF2-Multimer 的置信度主要由 MSA 的共进化信号驱动，对单点电荷取代基本不敏感 —— 它会把任何 bZIP 拉链都判成平行 卷曲螺旋，不管 e/g 位带什么电荷。它在这里的恰当用途是**验证 register 和测量 实际几何距离**（本报告第 2 节），而不是当亲和力预测器。真正的定量依据是 g· e′ 电荷配对分析，它能精确复现论文独立报告的"2 个 E–E 排斥"。

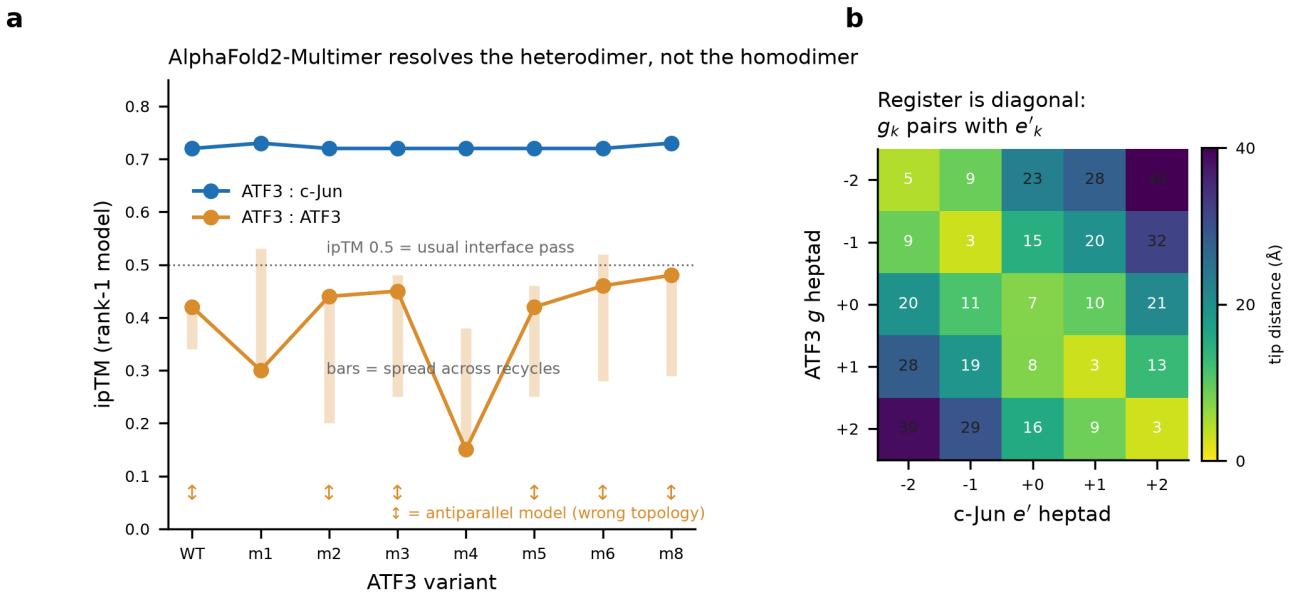


图 2 (a) 16 个复合物的 ipTM。(b) register 对角线距离矩阵。

7. 验证与敏感性分析

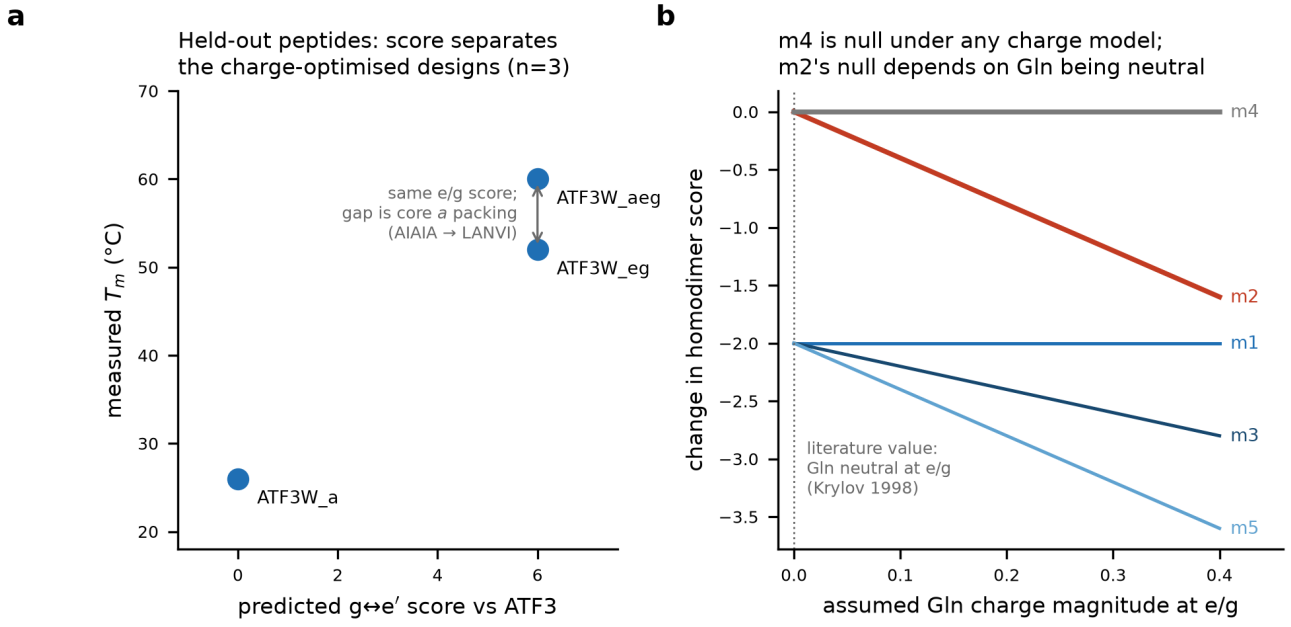


图 3 (a) 留出肽验证。(b) Gln 电中性假设的敏感性。

7.1 留出验证（论文中未用于构建打分的数据）

论文自己筛出的三条肽（Materials & Methods 中给出序列）构成一组留出测试集。 用同一套打分、同一套 register 规则（anchor 取 a3，与 ATF3 对齐）直接计算：

肽	同源二聚体排斥对	其中 E–E / K–K	与 ATF3 的吸引对	实测 Tm
ATF3W_a	0	0 / 0	0	26 ° C

肽	同源二聚体排斥对	其中 E–E / K–K	与 ATF3 的吸引对	实测 Tm
ATF3W_eg	10	6 / 4	6	52 ° C
ATF3W_aeg	10	6 / 4	6	60 ° C

论文独立报告的 ATF3W_aeg 数值是"ten unfavorable（6 E–E and 4 K–K）"和 "six favorable E–K" —— **两项均逐位复现，无任何调参。**

实测 Tm 的排序也一致：e/g 全为 Q 的 ATF3W_a（打分 0）对应 26 ° C，两条电荷优化 肽（打分 +6）对应 52/60 ° C。打分**无法**区分 eg 与 aeg —— 这是正确的，因为两者 差别在核心 a 位（AIAIA vs LANVI），而 e/g 打分不建模核心包装。论文也把 52→60 ° C 的增益归因于核心。该盲区由此被定位并界定，而非被掩盖。

7.2 对 Gln 电中性假设的敏感性 —— 一处需要更正

打分把 Gln 在 e/g 位视为电中性。扫过"Gln 等效电荷量 wQ"后：

wQ	m1	m2	m3	m4	m5
0.0	−2.00	0.00	−2.00	0.00	−2.00
0.2	−2.00	−0.80	−2.40	0.00	−2.80
0.4	−2.00	−1.60	−2.80	0.00	−3.60

- **m4 的"零影响"是无条件的**：e129 对面是 V124，在任何电荷模型下都不带电。
- **m2 的"零影响"是有条件的**：e136 对面是 Q131。若 Gln 带有可观的部分电荷，

K136E 会变成真实的去稳定突变（wQ=0.4 时达 −1.6，接近 m1）。

把 Gln 视为 e/g 位的电中性填充物是该领域的标准做法，也正是本文所依赖的 （文中引 Krylov, Barchi & Vinson 1998 说明 Q "neither favored nor disfavored for many opposing residues"，并据此在 Library a 中把 e/g 全固定为 Q）。所以这个假设 有外部支持，但**m2 的判定应表述为条件结论，m4 则不需要条件。**

7.3 排序稳健性

对 6 个打分参数（Gln 电荷 0–0.4、His 电荷 0–1、Arg 相对强度 0.7–1.3、 排斥/吸引比 0.3–1.5、g–e′ 取向不对称性 0.7–1.4、末端七肽权重 0.3–1.0） 随机取 5000 组：

- m3 在 78% 的参数组下排第一，其余 22% 由 m1 夺冠（无第三者）。
- **m1 与 m3 在主判据上是平手**：两者的 (dhomo, min_dhet) 同为 (−2, −1)。

区分它们的是次级判据 —— m3 的平均伙伴效应为 +0.89 而 m1 为 −0.67， 且 m3 **不触碰疏水位**（m1 牺牲 I143 的包装）。**"m3 优于 m1"这一结论来自次级判据，而非主判据。**

- 伙伴面板留一法（依次剔除 9 个伙伴中的任一个）：m3 在全部 10 种面板下均排第一，

m1 均排第二。结论不依赖任何单个伙伴。

7.4 register 可识别性 —— 第二处需要更正

此前我把"复现论文的 2 个 E–E 排斥"当作 register 的校准依据，这**说法过强**。 实测：register 平移 ±1 个七肽后，同源二聚体打分仍为 −2、E–E 对仍为 2 （因为唯一的带电排斥对 E117· E122′ 落在平移后的窗口内部）。

平移	同源二聚体打分	E–E 对	a 位签名
−1	−2	2	RTSNI
0	−2	2	TSNIK
+1	−2	2	SNIKL

所以 E–E 计数校准的是**打分函数**，而非 register。真正锁定 register 的是三项： **a 位签名** TSNIK（唯一与论文一致的平移）、**碱性区 Asn96 + 29 = 125** 的独立 偏移校验、以及**模型坐标中 g×e′ 距离矩阵的对角线双最小**。这三项本报告都做了， 结论不变，但依据应如此表述。

7.5 未评估的内容

- **无统计模型，因此没有残差诊断或分布假设可检验。** 电荷配对打分是确定性的

组合计数，不是拟合估计量；"假设检验/残差诊断"在此是范畴错误（先前版本误列）。
可检验的是打分方案的物理假设，即 7.1–7.4。

- MM-GBSA **ΔΔG 系列**：流程已搭好并验证（WT 全长 bZIP 同源二聚体

E_int = −106 kcal/mol，两次独立运行重复性 ~1 kcal/mol），但未对全部变体运行。
原因：这些突变每条链净电荷改变 2 个单位，单构象 GB 相互作用能会被净电荷的 去溶剂化项主导，而非 g· e′ 配对几何。本机无 GPU，每个结构约 70 分钟亦不现实。

- **核心 a/d 包装未建模**。7.1 已量化这个盲区的大小：它正是 ATF3W_eg 与

ATF3W_aeg 之间那 8 ° C。本设计的所有突变都在 e/g 位，不改变核心，故不影响结论；
但若将来考虑核心位突变，这套打分不适用。

- FOS、FOSL1、NFE2L2、MAFG 未纳入伙伴面板：这些蛋白在 a 位没有保守 Asn

（FOS 的 a 位为 T/T/K/I/K），无法用同一锚点可靠转移 register。FOS 本身也不是 ATF3 的主要配对伙伴。

8. 建议的实验验证

1. 优先构建 m3；若需更强的同源二聚体破坏，用 m8（Q115K+Q131E+K136E）。

不建议做 m2 和 m4 —— 按配对分析它们对同源二聚体零影响。

2. 同源二聚体检测：拉链肽段的 CD 热变性（Tm）或 SEC，参照论文的做法。
3. 异源二聚体保留度：与 c-Jun、ATF2、CEBPß 分别做共免疫沉淀或 CD/ITC。

注意 **ATF2/ATF7 是所有方案的共同代价**（−1 至 −3），应单独验证。

4. 注意两点：

- 全长 ATF3 结合 CRE DNA 时，二聚体会被 DNA 协同稳定。6 个电荷排斥在溶液中

足以破坏二聚体，但在 DNA 上仍可能形成同源二聚体，需用 EMSA 确认。

- 同源二聚化被阻断后，突变体会更多地进入异源二聚体池，异源二聚体的占比

预期上升 —— 这对您的目标是有利的，但解读表型时需考虑。

5. 所有突变均位于拉链（≥115），碱性区（86–110）未被触碰，DNA 结合面完整。