

mRNA 三级结构回转半径对蛋白质折叠速率的影响

田 旭, 李瑞芳, 彭诗雅, 高 姍, 邓俊超, 黄 娜, 付恒瑞

(内蒙古师范大学 物理与电子信息学院, 内蒙古 呼和浩特 010022)

摘 要: 回转半径(R_g)作为体现 mRNA 三级结构空间紧凑性的重要参量,其值越小对应的结构越紧密。构建含 105 个已知蛋白质折叠速率的蛋白质数据集,用 3dRNA 和 3DNA 软件计算了各蛋白质所对应的 mRNA 三级结构,从计算结果中整理出回转半径值,分析了蛋白质折叠速率与 mRNA 三级结构回转半径的相关性。研究结果显示, mRNA 三级结构 R_g 与蛋白质折叠速率呈极显著负相关($R = -0.426, P = 6.01 \times 10^{-6}$)。将 105 个蛋白质按二级结构及折叠类型分类,发现在各类蛋白质中 mRNA 三级结构 R_g 值与蛋白质折叠速率的相关程度存在差异。研究表明, mRNA 三级结构的回转半径对蛋白质折叠速率具有重要影响。

关键词: RNA 三级结构; 蛋白质折叠速率; 回转半径

中图分类号: Q61 **文献标志码:** A **文章编号:** 1001-8735(2026)01-0010-08

DOI: 10.3969/j.issn.1001-8735.2026.01.002

蛋白质是生物体内关键的生物大分子,由氨基酸按特定序列组成。蛋白质的折叠过程涉及多种非共价键相互作用,包括氢键和范德华力等^[1],其从变性状态到特定的生物学天然构象的转变,是一个复杂的动力学和热力学过程^[2-4]。在自然状态下,蛋白质通常能快速、可靠地从未折叠状态转变为具有特定三维结构的天然构象,这一结构对蛋白质功能有至关重要的影响。然而,错误的折叠可能导致无活性蛋白质的形成或淀粉样纤维的聚集,从而引发阿尔茨海默综合征、帕金森氏症、神经退行性蛋白病等蛋白质折叠疾病^[5-9]。蛋白质折叠过程是所有生物体系中最基本和重要的过程之一,也是生命科学中复杂而深奥的基础理论问题之一。因此,深入研究蛋白质的折叠过程对于理解其生物学功能和疾病机制具有重要意义。

蛋白质折叠过程遵循自由能减小的规律,即在一定时间内,通过过渡态系统(多个可能的路径)达到其自由能的全局最小值,形成热力学最稳定的天然构象^[10-12]。在蛋白质中,折叠速率存在很大差异,有些蛋白质在几微秒内完成折叠,而有些则需要几个小时。早期的研究发现,环境条件、蛋白质的各级结构和氨基酸序列等都对蛋白质折叠速率产生重要影响^[13-15]。随着蛋白质共翻译折叠理论的发展, mRNA 空间结构也被认为是影响蛋白质折叠速率的一个重要因素^[16]。

RNA 的折叠是一个层级式的过程,通过共轴螺旋堆积与三级接触的协同作用,最终形成紧凑的三维构象。目前, RNA 一级结构和二级结构的研究体系相对成熟,而对 RNA 三级结构的研究相对较少。在 RNA 的折叠过程中,通过双螺旋区的成核作用和二级结构单元间的“缩合”,最终形成具有生物学功能的三维构象。结合蛋白质的共翻译折叠过程,本研究小组分析了 mRNA 二级结构对蛋白质折叠速率的影响,发现 mRNA 二级结构对蛋白质折叠速率具有重要影响^[17-18]。值得注意的是, mRNA 的二级结构可通过氢键及其他相互作用折叠成 mRNA 三级结构,而 mRNA 的三级结构会直接影响到蛋白质的各级结构。因此,本研究假设, mRNA 的三级结构对蛋白质折叠速率也会有直接的影响。在 RNA 三级

收稿日期: 2025-05-12

基金项目: 内蒙古自治区自然科学基金资助项目“mRNA 对蛋白质折叠速率影响机制的研究”(2023MS03019); 呼和浩特市基础研究与应用基础研究计划资助项目“内含子间相互匹配在基因调控网络中的作用”(2025-规-基-58)。

作者简介: 田 旭(2000—),女,在读硕士研究生。

通信作者: 李瑞芳(1974—),女,教授,主要从事生物物理研究, E-mail: lirufang@imnu.edu.cn。

结构的形成过程中,回转半径作为表征空间构象动态变化的核心物理量,能够定量反映 mRNA 折叠过程中共轴螺旋堆积和远程相互作用的协同效率^[19]。

1 材料与方 法

1.1 数据集

本研究从整理的蛋白质折叠数据库中,经过去冗余处理,最终选取 105 个蛋白质作为研究样本。由于目前具有实验测定折叠速率数据的蛋白质数量有限,这些样本来自不同物种。这些蛋白质的折叠速率数据来源于文献[20-22]。在蛋白质折叠动力学研究中,折叠速率常数 k_f 的数值跨度极大(10^{-6} s^{-1} 到 10^3 s^{-1}),其原始数据分布严重偏斜,不满足统计分析的正态性要求。为此,研究领域内普遍采用其自然对数 $\ln k_f$ 进行分析。因此,论文选用 $\ln k_f$ 表述蛋白质折叠速率。

对这 105 个蛋白质进行了不同类型的分类:按照蛋白质的二级结构,可分为 24 个全 α 类蛋白质、35 个全 β 类蛋白质和 46 个混合类(α - β)蛋白质;按照蛋白质折叠类型,可分为 60 个二态蛋白质和 45 个多态蛋白质。

1.2 mRNA 三级结构的回转半径

首先,使用 RNAfold^[23] 软件计算每个蛋白质的 mRNA 二级结构,然后将每个蛋白质的 mRNA 序列和相应的二级结构输入 3dRNA^[24] 软件中。3dRNA 是一种基于 RNA 一级结构和二级结构预测 RNA 三维结构的软件,它通过基于物理的能量最小化和精修流程,旨在生成局部稳定的低能构象。其预测准确性已通过严格基准测试,被证明是当前最先进的方法之一,尤其对中等大小 RNA 具有竞争力^[24]。随后,将基于 3dRNA 预测出的结果导入 3DNA^[25] 中进行分析,从分析结果中整理得到每个蛋白质相应的 mRNA 三级结构的 R_g 值。

回转半径(R_g)具体的计算公式为

$$R_g = \sqrt{\frac{\sum_{i=1}^N m_i r_i^2}{\sum_{i=1}^N m_i}}, \quad (1)$$

其中: R_g 是 mRNA 三级结构的回转半径; N 是参与计算的粒子或基团总数量(如 mRNA 链中的核苷酸、磷酸基团等基本单元); m_i 是第 i 个粒子或基团的质量; r_i 是第 i 个粒子相对于系统质心的空间距离,其平方项反映该粒子对整体质量分布的离散贡献。 R_g 值表示通过加权平均质量分布的空间离散度,定量表征 mRNA 三级结构的紧凑性; R_g 越小,分子构象越紧密; R_g 越大,结构越松散。

R_g 是构象动态分布的综合统计量^[26],其数值反映的是结构变化的难易程度。该变化能力通过影响 mRNA 结构稳定性,最终调控蛋白质折叠速率。本文计算的是由 R_g 量化的结构动态特性,而非瞬时静态结构。

另外,由公式(1)可以发现, R_g 与质点组力学中的物理量——转动惯量紧密相连。转动惯量作为衡量物体转动惯性的关键物理量,其数值大小直接决定了物体运动状态改变的难易程度。具体而言,转动惯量越大,物体维持原有运动状态的能力越强,运动状态改变越难。

转动惯量的公式为

$$I = \sum_{i=1}^N m_i r_i^2, \quad (2)$$

转动惯量 I 与回转半径 R_g 的关系为

$$R_g = \sqrt{\frac{I}{\sum_{i=1}^N m_i}}。 \quad (3)$$

在生物大分子动力学中,分子转动惯量表征分子整体质量分布对旋转运动的惯性阻力。紧凑结构(低 R_g 值)对应的低转动惯量,表示分子更易通过旋转调整构象,从而加速折叠过程或功能运动。依据公式(1),蛋白质对应 mRNA 三级结构的 R_g 值取其总能量最小时的构象对应的数值。

1.3 回归分析

对蛋白质作 R_g 值与蛋白质折叠速率的回归分析,记录回归结果中 Multiple R 和 F 的值。Multiple R 称相关系数,用 R 表示,公式为

$$R = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (4)$$

F 对应的是显著性统计量,一般情况下等于弃真概率 P 值,即样本为假的概率,本文通过 P 值检验分析结果的可靠性。

2 结果分析

2.1 回转半径与蛋白质折叠速率的相关性

通过对 105 个蛋白质样本的线性回归分析,发现 mRNA 三级结构 R_g 与蛋白质折叠速率呈极显著负相关($R = -0.426, P = 6.01 \times 10^{-6}$)。为直观展示这一趋势,绘制了 R_g 值与 $\ln k_f$ 的热图及散点图(图 1)。在热图中为了直观体现出 $\ln k_f$ 随 R_g 的变化关系,将 105 个蛋白质样本按 R_g 值升序排列并赋予连续编号 N_p ,再按排列顺序画出蛋白质折叠速率热图,这样就将蛋白质折叠速率与 R_g 值关联起来,可以直观地呈现出 $\ln k_f$ 随 R_g 的变化关系。由图 1 可知,随着 R_g 值的增大(即 mRNA 三级结构越松散),蛋白质折叠速率 $\ln k_f$ 呈下降的趋势,直观地呈现出显著的负相关性。

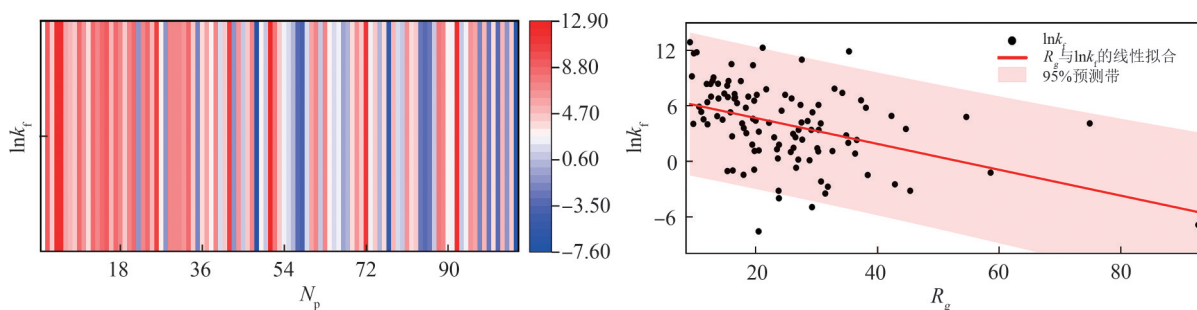


图 1 mRNA 三级结构 R_g 与蛋白质折叠速率的相关性热图和散点图

Fig. 1 The correlation of mRNA tertiary structure R_g with protein folding rates of heatmap and scatter plot

2.2 回转半径对不同二级结构类蛋白质折叠速率的影响

mRNA 三级结构 R_g 与不同二级结构类蛋白质折叠速率相关性结果见表 1。表 1 显示,相关系数 R 值对所有蛋白质及不同二级结构类蛋白质均为负数,说明 R_g 值越大,相应的蛋白质折叠速率越小。对于本文所选所有蛋白质,回转半径与蛋白质折叠速率呈极显著负相关($R = -0.426, P = 6.01 \times 10^{-6}$);对于全 α 类蛋白质,二者呈显著负相关($R = -0.554, P = 4.9 \times 10^{-3}$);对于混合类(α - β)蛋白质,二者呈极显著负相关($R = -0.508, P = 2.7 \times 10^{-4}$)。而对全 β 类蛋白质,二者无相关性($R = -0.174, P = 0.325$)。这一差异散点图(图 2)已直观体现:在全 α 类和混合类(α - β)蛋白质的散点图中,可以观察到明显的负相关趋势,数据点分布与各自的拟合线方向一致,全 α 类斜率更陡。而在全 β 类蛋白质的散点图中,数据点分布较为分散,没有形成明确的线性趋势,拟合线接近水平,说明了回转半径的大小对不同类蛋白质折叠速率的影响不同。依据二级结构组成将蛋白质样本划分为三类:全 α 类蛋白质、全 β 类蛋白质及混合类(α - β)蛋白质,对各类别样本的回转半径进行升序排列后实施连续编号(即 $N_p = 1$ 为组内最小 R_g 值,编号增大表示 R_g

表 1 mRNA 三级结构 R_g 对不同二级结构类蛋白质折叠速率的影响

Tab. 1 The effect of mRNA tertiary structure R_g on protein folding rates across secondary structure classes

类型	R	P
总体	-0.426	6.01×10^{-6}
α	-0.554	4.9×10^{-3}
β	-0.174	0.325
α - β	-0.508	2.7×10^{-4}

值升高), 构建三个独立数据集: 全 α 类蛋白质(24 个样本, N_p : 1-24)、全 β 类蛋白质(35 个样本, N_p : 1-35)、混合类(α - β)蛋白质(46 个样本, N_p : 1-46)。基于上述分类数据, 进一步生成每一类蛋白质折叠速率随 R_g 变化关系的热图(图 2), 定量揭示了 mRNA 三级结构回转半径对蛋白质折叠速率的调控效应存在影响。

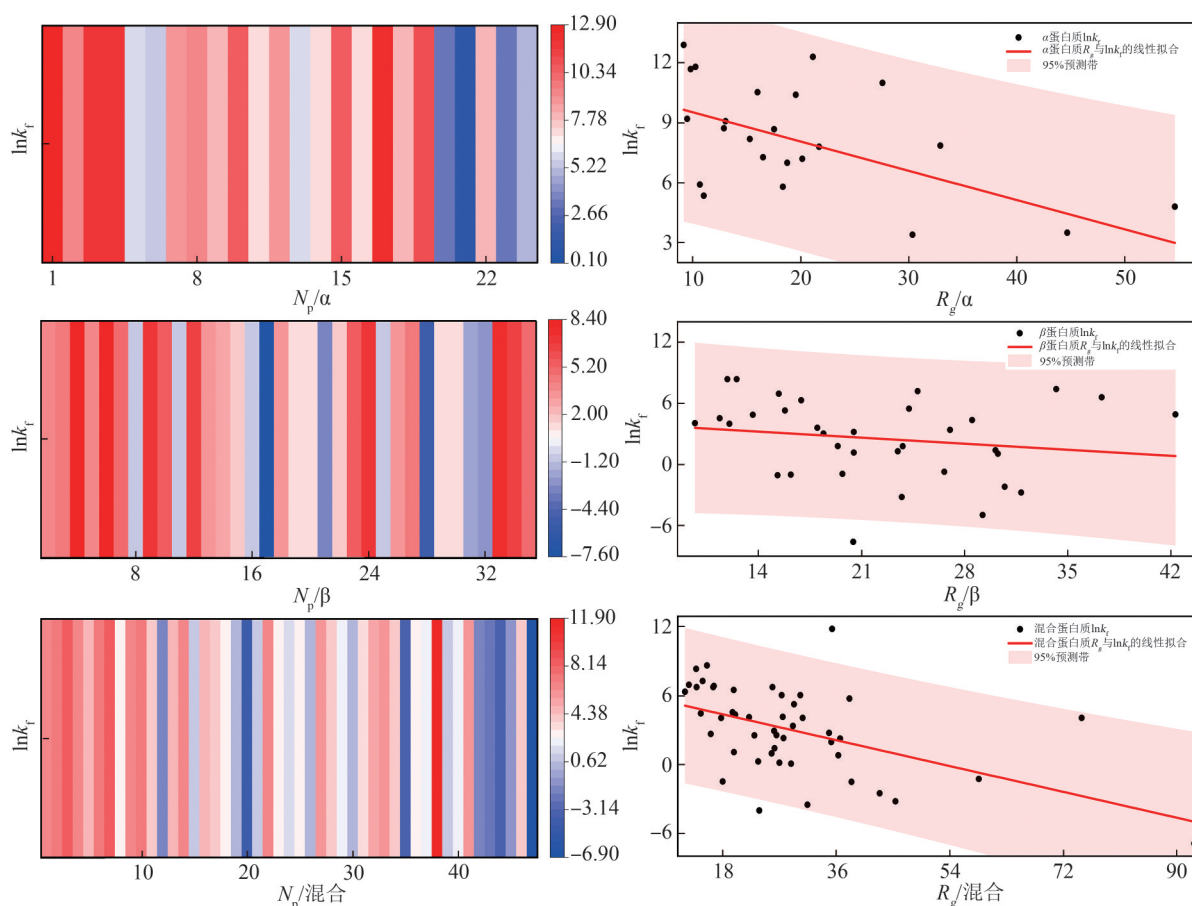


图 2 mRNA 三级结构 R_g 与不同二级结构类型蛋白质折叠速率的相关性热图和散点图

Fig. 2 The correlation between mRNA tertiary structural R_g and protein folding rates across secondary structure classes of heatmap and scatter plot

2.3 回转半径对不同折叠类蛋白质折叠速率的影响

mRNA 三级结构的回转半径与不同折叠类蛋白质折叠速率相关性分析结果见表 2。由表 2 可知, 不论是二态类蛋白质还是多态类蛋白质, 它们的折叠速率与相应的 mRNA 三级结构回转半径均呈显著负相关。由散点图(图 3)可知, 两类蛋白质的散点图整体趋势都显示为负相关, 数据点随 R_g 增大向向下分布的趋势。这一结果表明, 对于不同折叠类蛋白质, 它们的折叠速率均随回转半径的增大而减小。本文依据折叠动力学类型(二态与多态)对蛋白质样本进行分类筛选, 并将每类样本的回转半径升序排列后进行连续编号($N_p=1$ 为组内最小 R_g), 其中二态类蛋白质的 N_p 取值范围为 1~60, 多态类蛋白质的 N_p 取值范围为 1~45, 然后基于样本编号及对应数据, 构建蛋白质折叠速率随 R_g 变化关系的热图(图 3), 可直观反映折叠速率的相对变化趋势。

表 2 mRNA 三级结构 R_g 对不同折叠类蛋白质折叠速率的影响

Tab. 2 The effect of mRNA tertiary structure R_g on protein folding rates across kinetic folding classes

类型	R	P
总体	-0.426	6.01×10^{-6}
二态	-0.071	3.92×10^{-2}
多态	-0.120	1.95×10^{-2}

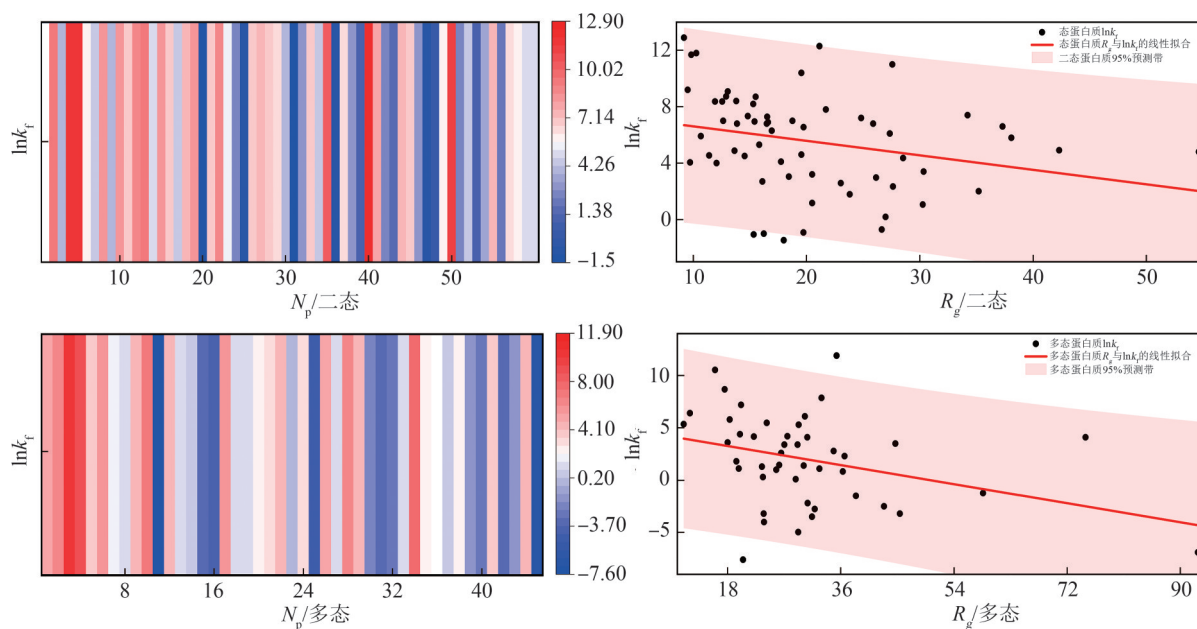


图 3 mRNA 三级结构 R_g 与不同折叠类蛋白质折叠速率的相关性热图和散点图

Fig. 3 The correlation between mRNA tertiary structural R_g and protein folding rates across distinct folding classes of heatmap and scatter plot

3 结论与讨论

本文分析了 mRNA 三级结构中回转半径对蛋白质折叠速率的影响, 以期对蛋白质折叠机制的阐释提供新的研究思路。总体来看, 基于 105 个蛋白质的线性回归分析表明, mRNA 三级结构回转半径 R_g 与折叠速率呈显著负相关 ($R = -0.426, P = 6.01 \times 10^{-6}$)。从动力学角度看, 在相同质量条件下, R_g 值越小, 描述物体转动惯性的转动惯量值越低, 更易改变运动状态。从热力学角度看, 低 R_g 值更易形成紧凑的 mRNA 结构, 通过空间限制效应将新生肽链约束于有限空间内, 提高折叠相关残基 (如疏水或带电基团) 的局部浓度。这一机制与受限环境中分子碰撞概率升高的物理原理一致, 从而加速疏水核心或氢键网络的形成^[27]。此外, 协同折叠机制通过 mRNA 的共轴螺旋堆积或远程相互作用, 形成动态结构框架, 协调核糖体移动与肽链折叠的时空同步性^[28]。

对于不同二级结构类蛋白质, 回转半径 (R_g) 与蛋白质折叠速率的关联不同。对 α 类蛋白质, 二者呈显著负相关, 此现象表明 α 类蛋白质折叠进程对 mRNA 三级结构存在较高依赖性。较大的 R_g 值意味着更为松散的 mRNA 空间布局, 转动惯量越高, 更难改变运动状态。在混合类 (α - β) 蛋白质层面, R_g 与蛋白质折叠速率呈极显著负相关。这一结果表明混合类 (α - β) 蛋白质的折叠进程对 mRNA 三级结构存在高度依赖性, 这可能是 mRNA 结构动态性与氨基酸序列的协同作用显著影响共翻译折叠路径导致的^[26]。在结果中未发现 β 类蛋白质的折叠速率与 R_g 的相关性, 原因在于 β 类蛋白质的折叠起始依赖于局部氢键网络, 形成高度协同的 β -折叠片层结构^[29], 这种折叠模式依赖局部相互作用的精确排列, 而非全局构象调整, 因此可能对 mRNA 三级结构的扰动相对不敏感, 形成相对自主稳定的折叠模式, 一定程度上隔绝了 mRNA 三级结构波动干扰, 展现出蛋白质折叠机制的精妙选择性与自主性, 也提示不同二级结构类蛋白质在进化中发展出各异折叠调控策略以适配功能需求。

对于不同折叠类型的蛋白质, 二态蛋白质和多态蛋白质的折叠速率虽均与 mRNA 三级结构回转半径呈显著负相关, 但相较于多态蛋白质, 二态蛋白质的相关性更弱 (R 绝对值小)。这一现象可能与二态折叠的动力学特性有关: 二态折叠通常发生在较小的、拓扑结构相对简单的蛋白质中, 其折叠过程高度协同, 近似一个“全有或全无”的相变过程, 主要受限于跨越单一、高能垒过渡态的速率^[30-31]。这种

高度协同的折叠机制,可能降低其折叠速率对 mRNA 结构提供的空间限制、协同运动等外部因素的敏感性,使其折叠速率更多由蛋白质自身的氨基酸序列及内在折叠途径决定^[32]。相比之下,多态折叠通常发生在分子更大、结构更复杂的蛋白质中,包含中间态或亚稳态,其折叠路径可能包含多个步骤和构象探索过程^[33]。这一差异表明,mRNA 三级结构对蛋白质折叠速率的调控效力与折叠机制的复杂性密切相关,在多态折叠中,mRNA 结构可能通过协调多步构象转变或降低中间态能垒来更显著地影响折叠动力学。

为进一步验证三种不同二级结构类蛋白质 mRNA 三级结构的回转半径对蛋白质折叠速率的影响,将数据库中的蛋白质进一步细分,对每一类蛋白质进行 mRNA 三级结构的回转半径与蛋白质折叠速率的相关性分析(表 3)。从表 3 可知,二态 α 类蛋白质的负相关($R=-0.267, P=4.05 \times 10^{-2}$)趋势与全 α 类结论一致;而二态 β 类蛋白质($R=2.45 \times 10^{-5}, P=0.982$)与多态 β 类蛋白质($R=-0.042, P=0.545$)中均未发现显著相关性,这与“ β 类蛋白质折叠对 mRNA 三级结构不敏感”的整体论断相符。但对于多态 α 类蛋白质则无相关性,原因可能是因样本量较小($n=8$),导致其存在的相关性未能被充分检测。结果还表明 mRNA 三级结构 R_g 与二态混合类(α - β)蛋白质折叠速率没有相关性,推测在二态混合类(α - β)蛋白质中, β 型结构可能对折叠过程具有更强的调控作用;mRNA 三级结构 R_g 与多态混合类(α - β)蛋白质的折叠速率呈显著负相关,这可能反映其 α 螺旋结构域在多态混合类(α - β)蛋白质折叠动力学中的主导作用。

本研究揭示了 mRNA 三级结构的回转半径对蛋白质折叠速率的调控作用。统计分析结果显示,mRNA 三级结构 R_g 值与蛋白质折叠速率呈显著负相关,表明较低 R_g 值对应的紧凑 mRNA 三级结构,可通过增强共翻译折叠过程中的空间限制效应和共轴堆积效应,动态协调核糖体运动状态,从而提升蛋白质折叠效率。进一步研究发现,mRNA 三级结构的 R_g 值与其转动惯量呈正相关,这一特性可能通过降低构象调整的惯性以促进折叠进程。此外,不同二级结构类蛋白质中,mRNA 三级结构回转半径对蛋白质折叠速率的影响存在明显差异。

参考文献:

- [1] DILL K A, MACCALLUM J L. The protein-folding problem, 50 years on[J]. Science, 2012, 338(6110): 1042-1046.
- [2] SHAKHNOVICH E. Protein folding thermodynamics and dynamics: Where physics, chemistry, and biology meet[J]. Chemical Reviews, 2006, 106(5): 1559-1588.
- [3] PIANA S, LINDORFF-LARSEN K, SHAW D E. Protein folding kinetics and thermodynamics from atomistic simulation [J]. Proceedings of the National Academy of Sciences of the United States of America, 2012, 109(44): 17845-17850.
- [4] ZHAO Q Y. Nature of protein dynamics and thermodynamics[J]. Reviews in Theoretical Science, 2013, 1(1): 83-101.
- [5] HARTL F U. Protein misfolding diseases[J]. Annual Review of Biochemistry, 2017, 86: 21-26.
- [6] CHITI F, DOBSON C M. Protein misfolding, amyloid formation, and human disease: A summary of progress over the last decade[J]. Annual Review of Biochemistry, 2017, 86: 27-68.
- [7] SWEENEY P, PARK H, BAUMANN M, et al. Protein misfolding in neurodegenerative diseases: Implications and strategies[J]. Translational Neurodegeneration, 2017, 6: 6.
- [8] JIANG Y, DEANE C M, MORRIS G M, et al. It is theoretically possible to avoid misfolding into non-covalent lasso entanglements using small molecule drugs[J]. PLoS Computational Biology, 2024, 20(3): e1011901.
- [9] ZHAO P, WANG C, SUN S H, et al. Tracing genetic diversity captures the molecular basis of misfolding disease[J].

表 3 不同类蛋白质 mRNA 三级结构 R_g 对蛋白质折叠速率的影响

Tab. 3 The effect of mRNA tertiary structure R_g on protein folding rate across different protein classes

类型	R	P
总体	-0.426	6.01×10^{-6}
二态 α	-0.267	4.05×10^{-2}
二态 β	2.45×10^{-5}	0.982
二态 α - β	-0.135	0.101
多态 α	-0.281	0.177
多态 β	-0.042	0.545
多态 α - β	-0.167	0.038

- Nature Communications, 2024, 15(1): 3333.
- [10] ONUCHIC J N, WOLYNES P G. Theory of protein folding[J]. Current Opinion in Structural Biology, 2004, 14(1): 70-75.
- [11] PANDEY P, ALEXOV E. Most monogenic disorders are caused by mutations altering protein folding free energy[J]. International Journal of Molecular Sciences, 2024, 25(4): 1963.
- [12] WOLYNES P G. Evolution, energy landscapes and the paradoxes of protein folding [J]. Biochimie, 2015, 119: 218-230.
- [13] DALTON B A, NETZ R R. pH modulates friction memory effects in protein folding [J]. Physical Review Letters, 2024, 133(18): 188401.
- [14] TSUBOYAMA K, DAUPARAS J, CHEN J, et al. Mega-scale experimental analysis of protein folding stability in biology and design[J]. Nature, 2023, 620(7973): 434-444.
- [15] CAMPOS L A, MUÑOZ V. Targeting the protein folding transition state by mutation: Large scale (un) folding rate accelerations without altering native stability[J]. Protein Science, 2024, 33(7): e5031.
- [16] NOLLER H F, HOFFARTH V, ZIMNIAK L. Unusual resistance of peptidyl transferase to protein extraction procedures[J]. Science, 1992, 256(5062): 1416-1419.
- [17] 李瑞芳, 于志芬, 黄俏. mRNA 的二级结构对蛋白质折叠速率的影响[J]. 生物物理学报, 2014, 30(7): 497-508.
- [18] 李瑞芳, 李宏, 郭春阳, 等. mRNA 环结构对蛋白质折叠速率的影响[J]. 生物化学与生物物理进展, 2018, 45(6): 672-680.
- [19] CHEN S J, DILL K A. RNA folding energy landscapes[J]. Proceedings of the National Academy of Sciences of the United States of America, 2000, 97(2): 646-651.
- [20] WAGAMAN A S, COBURN A, BRAND-THOMAS I, et al. A comprehensive database of verified experimental data on protein folding kinetics[J]. Protein Science, 2014, 23(12): 1808-1812.
- [21] MANAVALAN B, KUWAJIMA K, LEE J. PFDB: A standardized protein folding database with temperature correction [J]. Scientific Reports, 2019, 9: 1588.
- [22] GROMIHA M M. Multiple contact network is a key determinant to protein folding rates [J]. Journal of Chemical Information and Modeling, 2009, 49(4): 1130-1135.
- [23] ZHAO Y J, HUANG Y Y, GONG Z, et al. Automated and fast building of three-dimensional RNA structures [J]. Scientific Reports, 2012, 2: 734.
- [24] WANG J, WANG J, HUANG Y Z, et al. 3dRNA v2.0: An updated web server for RNA 3D structure prediction[J]. International Journal of Molecular Sciences, 2019, 20(17): 4116.
- [25] LU X J, OLSON W K. 3DNA: A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures[J]. Nucleic Acids Research, 2003, 31(17): 5108-5121.
- [26] MUSTOE A M, BUSAN S, RICE G M, et al. Pervasive regulatory functions of mRNA structure revealed by high-resolution SHAPE probing[J]. Cell, 2018, 173(1): 181-195. e18.
- [27] ELLIS R J. Macromolecular crowding: Obvious but underappreciated [J]. Trends in Biochemical Sciences, 2001, 26(10): 597-604.
- [28] FAURE G, OGURTSOV A Y, SHABALINA S A, et al. Role of mRNA structure in the control of protein folding[J]. Nucleic Acids Research, 2016, 44(22): 10898-10911.
- [29] DOBSON C M. Protein folding and misfolding[J]. Nature, 2003, 426(6968): 884-890.
- [30] JACKSON S E. How do small single-domain proteins fold?[J]. Folding and Design, 1998, 3(4): R81-R91.
- [31] DAGETT V, FERSHT A R. Is there a unifying mechanism for protein folding?[J]. Trends in Biochemical Sciences, 2003, 28(1): 18-25.
- [32] WOLYNES P G. Recent successes of the energy landscape theory of protein folding and function[J]. Quarterly Reviews of Biophysics, 2005, 38(4): 405-410.
- [33] BALDWIN R L, ROSE G D. Is protein folding hierarchic? II. Folding intermediates and transition states[J]. Trends in Biochemical Sciences, 1999, 24(2): 77-83.

Effect of mRNA Tertiary Structure Radius of Gyration on Protein Folding Rate

TIAN Xu, LI Ruifang, PENG Shiya, GAO Shan, DENG Junchao, HUANG Na, FU Hengrui

(College of Physics and Electronic Information, Inner Mongolia Normal University, Hohhot 010022, China)

Abstract: The radius of gyration (R_g) is an important parameter reflecting the spatial compactness of the mRNA tertiary structure, with smaller values indicating tighter structures. This paper constructed a protein dataset containing folding rates for 105 known proteins and calculated the corresponding mRNA tertiary structures using the 3dRNA and 3DNA software. From the computational results, it extracted the radius of gyration values and analyzed the correlation between protein folding rates and the radius of gyration of the mRNA tertiary structure. The results show a highly significant negative correlation between the mRNA tertiary structure R_g and protein folding rates ($R = -0.426$, $P = 6.01 \times 10^{-6}$). By classifying the 105 proteins based on secondary structure and folding type, differences are found in the correlation levels between the mRNA tertiary structure R_g and protein folding rates across various protein categories. This study indicates that the radius of gyration of the mRNA tertiary structure has a significant impact on protein folding rates.

Key words: RNA tertiary structure; protein folding rate; radius of gyration

【责任编辑 张颖娟】