上海交通大學

SHANGHAI JIAO TONG UNIVERSITY

学士学位论文

BACHELOR'S THESIS



论文题目:Argonaute蛋白热稳定性及动力学研究

学生姓名:	昝 冰	
学生学号:	517070910028	
专业:	物理学专业(国际班)	
指导教师:	洪 亮	
学院(系):	物理与天文学院	



上海交通大学 学位论文原创性声明

本人郑重声明:所呈交的学位论文《Argonaute蛋白热稳定性及动力学研究》,是本人在导师的指导下,独立进行研究工作所取得的成果。除文中已经注明引用的内容外,本论 文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡 献的个人和集体,均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本 人承担。

学位论文作者签名: 眷 *

日期: 2021 年 6 月 1 日



上海交通大学 学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定,同意学校保留并向国家 有关部门或机构送交论文的复印件和电子版,允许论文被查阅和借阅。本人授权上海交 通大学可以将本学位论文的全部或部分内容编入有关数据库进行检索,可以采用影印、 缩印或扫描等复制手段保存和汇编本学位论文。

保密口,在____年解密后适用本授权书。

本学位论文属于

不保密 ☑。

(请在以上方框内打"√")

学位论文作者签名: 眷 沐

指导教师签名: 沒亮

日期: 2021年6月1日

日期: 2021年6月1日



Argonaute 蛋白热稳定性及动力学研究

摘要

从原核生物中提取的 Argonaute 蛋白(pAgo)是一类核酸引导蛋白,由具有高度差异性 的蛋白组组成。pAgo 作为宿主的一类核酸内切酶,能够利用短片段干扰 RNA 或 DNA 作为 引物,对入侵 DNA 进行剪切从而使其失去活性。因此, p.4go 是一种能够被应用于基因剪辑 的理想蛋白酶。有趣的是,不同 pAgo 的结构非常相近,但是他们的生理温度却相差巨大。 在本次工作中,我们选择了三种具有不同生理温度的 pAgo,他们的生理温度范围分别为 37 °C,75 ℃和95 ℃,用来探究其热稳定性和体外基因剪切机理。我们发现,嗜热 p.4go 相较 于嗜常温 pAgo, 整体结构更为刚性并且在常温下解折叠吉布斯自由能更高。这是由于嗜热 pAgo 的氢键密度,盐桥密度和蛋白核心区域疏水基团比例高于嗜常温 pAgo,从而导致嗜热 pAgo 更为稳定。令人惊奇的是,结合圆二色谱和荧光染色实验,我们发现嗜热 pAgo 在其生 理温度发生了部分解折叠,也就是形成了熔融态。通过小角 X 射线散射发现,在 pAgo 各自 的生理温度,嗜热 p.Ago 表现出一种松散堆叠且柔性的结构,而嗜常温 p.Ago 为球状并且紧 凑的构象。此外,我们将解折叠试剂加入到嗜热 pAgo 中进而提高其解折叠比例和蛋白整体 柔性,发现解折叠比例提高后基因剪切活性提升。相应的酶活性实验分析表明,嗜热 pAgo 的熔融态能够提升目标基因的剪切效率和产物的释放, 而未对底物的亲和力产生影响。通过 我们的实验结果可以推测,嗜热 pAgo 在宿主体内可能通过熔融态进而实现其剪切入侵 DNA 的功能。并且, 嗜热 pAgo 在其生理温度为熔融态的这一结果, 能够为将嗜热 pAgo 改造为 具有常温活性的蛋白提供理论指导。

关键词: 原核生物 Ago 蛋白,基因剪辑,嗜热,嗜常温,熔融态,蛋白质工程



STUDY OF ARGONAUTE PROTEIN THERMOSTABILITY AND DYNAMICS

ABSTRACT

Prokaryotic Argonaute proteins (pAgos) constitute a highly diverse group of nucleic acid-guided proteins, which used as endonucleases in hosts to defense against the invading DNA by utilizing small interfering DNA guides. Therefore, pAgo is an ideally enzyme which can be applied to genome editing. Interestingly, whereas pAgos share the similar structures, the physiological temperatures (T_{phy}) are vastly diverse among them. Here, three pAgos with distinct T_{phy} ranged from 37 °C to 95 °C are chosen to investigate the thermostability and mechanism of genome-cleavage process in vitro. We explored that thermophilic pAgos are stiffer and have higher unfolding Gibbs free energy at room temperature than their mesophilic counterparts, which can be attributed to higher density of hydrogen bond and salt bridge and higher ratio of hydrophobic residues in thermophilic pAgos. Surprisingly, combining circular dichroism spectroscopy and fluorescence dye measurements, we explored the thermophilic pAgos are partially unfolding at their $T_{\rm phy}$. Interestingly, not only is ${\rm Mn}^{2+}$ necessary for the biochemical reaction but it also promotes the unfolding of pAgos. Further small-angle X-ray scattering identified the thermophilic pAgos present a loosely packed and flexible architecture at their T_{phy} but the mesophilic pAgo exhibits a globular compact conformation. Furthermore, urea-enhanced molten state in thermophilic pAgos which increases unfolding ratio and flexibility exhibits a higher activity of genome-cleavage. Complementary enzyme activity analysis revealed that the molten state promotes the efficiency of cleaving targets and releasing the products. Our findings suggested in vivo cleaving the invading DNA by thermophilic pAgos may be mediated via the molten state and which are fundamentally important for engineering the thermophilic pAgos to moderate temperature.

Key words: pAgo protein, genome-cleavage, thermophilic, mesophilic, molten state, protein

engineering



目 录

第一章 绪论	1
1.1 原核生物中 long-Argonaute 蛋白结构均	或构架1
1.1.1 原核 Argonaute 蛋白生物功能	1
1.1.2 DNA引导的DNA干扰路径	3
1.1.3 RNA 引导的 DNA 干扰路径	3
1.1.4 其他原核生物中 long-Argonaute	蛋白5
1.1.5 short-pAgos和PIWI-RE蛋白	5
1.2 基于原核生物 Argonaute 蛋白的基因	编辑功能6
1.2.1 long-Argonaute蛋白作为基因编	辑工具的优缺点6
1.2.2 Pyrococcus furiosus Argonaute 蛋	台的应用8
1.3 本章小结	12
1.3.1 本课题的提出	12
1.3.2 本课题的主要研究内容	13
第二章 热稳定性、动力学及生化反应研究方	法14
2.1 实验原理	14
2.1.1 荧光分光光度计	14
2.1.2 小角X射线散射	15
2.1.3 远紫外圆二色光谱仪	17
2.1.4 差示扫描量热仪	19
2.2 本章小结	23
第三章 Argonaute 蛋白热稳定性及基因剪切机	扎理研究24
3.1 Argonaute 蛋白热稳定性	24
3.2 Argonaute 蛋白在生理温度时构象	28
3.2.1 二级结构和三级结构表征	28
3.2.2 小角 X 射线散射表征	30
3.2.3 生物化学表征	32
3.3 本章小结	36
第四章 结论	38
参考文献	39
谢辞	44



第一章 绪论

Argonaute蛋白最初被发现于真核生物中,它们在RNA干扰 (RNAi)中具有关键作用^[1], ^[2]。作为RNA引导的静默复合物 (RISC)的功能核心,真核生物中的Argonaute蛋白 (eAgos) 使用短片段的单链RNA (ssRNA)分子作为引物靶向互补的目标RNA序列^[2-4]。目标RNA的翻 译过程可以通过结合或者剪切目标RNA的方式直接静默,或者通过结合目标RNA并募集额 外的静默蛋白间接静默^[5,6]。因此, eAgos可以在转录后调节基因表达,保护其宿主免受入侵 的RNA病毒的侵害,并通过降低转座子的迁移率来保持基因组的完整性^[2]。

随着同源原核生物中的Argonaute蛋白 (pAgos) 的发现,古菌和细菌中第一个全长 Argonaute结构被解析^[7,8]。此外,据报道,在~32%和~9%的分别测序的古细菌和细菌 基因组中,水平基因转移导致编码Argonaute蛋白的基因分布非常不均匀。由于古菌和细菌 缺乏RNA干扰途径^[9],因此原核生物的Argonaute蛋白的生理功能一直未被研究透彻。最近 有研究表明,原核生物的Argonaute蛋白具有独特的机制特征,并揭示了对其生理功能的意 义^[10-17]。有趣的是,已经有研究表明一些原核生物的Argonaute蛋白通过核酸引导剪切目标 DNA^[12-15]。鉴于与CRISPR相关的蛋白Cas9的具有核酸引导剪辑活性已经彻底改变了基因组 编辑领域,原核生物的Argonaute蛋白的变体可能能够应用于生物技术。例如癌症诊断和核 酸检测等。

1.1 原核生物中 long-Argonaute 蛋白结构域构架

基于Argonaute蛋白的结构域构架,原核生物的Argonaute蛋白可分为三类:分别为long pAgos,short pAgos和PIWI-RE蛋白(如图1-1a所示)。在这些类别中,只有long pAgos这一类蛋白在实验上被研究^[11-22]。虽然Argonaute蛋白的序列保守性很差并且功能差别性较大,但long pAgos的总体结构域架构与真核生物中Argonaute蛋白的结构域架构非常相似。eAgos和long pAgos均形成双裂叶支架,其中一个叶由氨基末端(N)和PIWI-Argonaute-Zwille (PAZ)结构域组成,而另一个叶由中间(MID)结构域和PIWI结构域组成^[7, 8, 22](如图1-1b所示)。MID结构域和PAZ结构域通常形成结合区域,能够促进引物脱氧核糖核酸或者引物核糖核酸的5'末端和3'末端结合在相应区域内^[22-24]。在目标核酸结合后,具有催化活性中心的Argonaute蛋白中的PIWI结构域(包含天冬氨酸-谷氨酸-天冬氨酸-X基序,其中X表示天冬氨酸、组氨酸或天门冬酰胺^[14])能够剪辑目标核酸。上述的活性中心位点氨基酸通常在eAgo和pAgo是不完整的,这可能跟结合目标核酸有关而跟剪切目标核酸无关^[11, 12, 25]。类似于eAgo中的N结构域,在一些pAgo中N结构域的功能被预测为能够像楔子一样能够加速剪切后目标核酸的解离^[26]。

1.1.1 原核 Argonaute 蛋白生物功能

2009年,一项生物信息学分析表明,编码Argonaute蛋白的基因和编码参与宿主防御的 蛋白质的基因存在于相同的操纵子中 (例如,核酸内切酶与限制性修饰系统相关联),这表 明原核生物的Argonaute蛋白可能在保护其宿主并防御入侵核酸方面具有重要作用,例如, 质粒入侵宿主细胞^[27]。事实上,这一假设至少已经由一些long pAgos所证实^[11-13,16]。值得注 意的是,虽然真核生物的Argonaute蛋白仅能通过引物RNA从而实现对入侵RNA的干扰,但 是一些原核生物的Argonaute蛋白能够与短片段的DNA引物结合,甚至在某些情况下,可以 直接靶向DNA^[12,13,18] (如表1-1所示)。这些结果表明,相较于真核生物的Argonaute蛋白,原



核生物的Argonaute蛋白功能更为强大,并且存在多样性的干扰机制。目前,Thermus thermophilus Argonaute蛋白(TtAgo)和Rhodobacter sphaeroides(RsAgo)中Argonaute蛋白的表征最为完善。虽然这两种pAgos均参与到宿主的防御体系中,但他们的防御机理并不完全相同。具体机理将在下一小节中阐述。



图1-1 Argonaute蛋白的演变和结构。(a) 原核生物中的Argonaute蛋白族谱和真核生物中的 Argonaute蛋白族谱。(b) 真核生物中Argonaute蛋白晶体结构,原核生物中的long-Argonaute 蛋白以及Argonaute蛋白-引物核酸-目标核酸的三元复合物晶体结构^[28]

表1-1 真核生物中Argonaute蛋白和原核生物中long-Argonaute蛋白对引物核酸和目标核酸的偏好性^[28]

eAgo and long pAgo proteins		Activity			
Host	Argonaute name	RNA-guided RNA interference	RNA-guided DNA interference	DNA-guided RNA interference	DNA-guided DNA interference
Eukaryotes	eAgo	+	-	-	-
Marinitoga piezophila*	MpAgo	+	+	-	-
Thermotoga profunda*	ТрАдо	-	+	-	-
Rhodobacter sphaeroides	RsAgo	-	+	-	-
Aquifex aeolicus	AaAgo	-	-	+	-
Natronobacterium gregoryi‡	NgAgo	-	-	unknown	unknown
Thermus thermophilus	TtAgo	-	-	+	+
Methanocaldococcus jannaschii	MjAgo	-	-	-	+
Pyrococcus furiosus	PfAgo	-	-	-	+

1.1.2 DNA 引导的 DNA 干扰路径

TtAgo参与宿主防御过程的第一个证据是在发现T.thermophilus菌株后获得的。由于嗜热菌菌株对外源DNA具有较高的自然吸收而被选择作为研究体系,其具有被转座子插入后破



坏的TtAgo基因^[12]。进一步的研究表明,在mRNA水平上,TtAgo并不调节参与宿主防御的其 他基因的表达。相反,*TtAgo*通过DNA引导的DNA干扰从而直接干扰在DNA水平上的入侵核 酸 (如图1-2a所示)。在体内, TtAgo结合短片段干扰DNA (siDNA) 引物 (长度为13-25个核苷 酸),从而用来结合和切割同源目标DNA。此外,在体内,*TtAgo*优先剪切入侵质粒DNA从 而产生siDNA,表明TtAgo可以区分入侵DNA和自身基因组DNA^[12]。最近的研究表明,TtAgo 通过其催化位点,利用一种称为 "chop" 的机制,能够直接结合质粒进行剪切,从而形成引 导DNA (如图1-2a所示)。在体外实验中,无引导DNA的TtAgo能够直接剪切具有一定程度不 稳定性双螺旋核酸。例如,含AT较高的的双链DNA (dsDNA) 或者具有局部碱基不匹配的双 链DNA会产生局部的解螺旋,从而产生短的双链DNA片段并且加载到TtAgo上。在TtAgo结 合短片段双链DNA分子后,与引导DNA互补的DNA单链(也就是passenger链)将会被释放 出去,从而使TtAgo仅保留一条单链DNA作为引物^[19]。当TtAgo-引导DNA复合物形成后,可 以结合并且剪切与引导DNA互补的单链目标DNA。此外,两个TtAgo-引导DNA复合物(引 导DNA序列为互补)可以使双链DNA发生断裂^[12](如图1-2a所示)。通过该机理可得知, TtAgo 通过采取降低细胞内质粒含量从而干扰入侵DNA。此外,TtAgo能够在体外剪切RNA,这一 结果说明TtAgo能够降低细胞内RNA含量从而降低入侵DNA在体内的转录^[16]。然而,TtAgo 能够降低RNA水平是直接由靶向RNA引起还是干扰DNA形成的间接结果仍未有定论。

类似于TtAgo,来自于古菌Pyrococcus furiosus(PfAgo)和Methanocaldococcus jannaschi (MjAgo)的Argonaute蛋白在体外均通过DNA引导对目标DNA进行剪切^[13, 15, 17]。此外, PfAgo已经被证明能够降低在P.furiosus中质粒的转化效率。虽然MjAgo的生理作用尚未被研 究,但是有工作已经表明,与TtAgo类似,MjAgo通过剪切更长的DNA底物从而产生 siDNA^[18]。有趣的是,源自M. jannaschii中染色体DNA并没有被MjAgo剪切^[18],这一结果证 明MjAgo能够区分入侵的DNA和自己的基因组DNA。尽管基因组DNA的色度化仅被证明可 以防止被MjAgo降解,但同时很容易推测其他基因包装蛋白在其他古细菌和细菌中具有相 同的功能。

值得注意的是,对于*TtAgo和MjAgo*,*Argonaute*蛋白通过siDNA剪切目标DNA的这种经 典剪切机理比在没有引物核酸下对DNA进行chop更为有效^[18]。因此,虽然DNA的chop过程 可能在一定水平上提供对抗入侵核酸的先天免疫,但siDNA引导的DNA干扰能够有效去除 入侵者DNA。此外,*Argonaute*蛋白-siDNA复合物可能靶向同一入侵者DNA的多个拷贝体, 因此可能会在DNA复制前对遇到的入侵DNA进行干扰。

1.1.3 RNA 引导的 DNA 干扰路径

*RsAgo*包含不完整的 DEDX 四聚体催化中心,因此将其推测为催化无活性。编码 *RsAgo*的基因与编码假定的 DNA 核酸酶的基因聚集在一起,因此该核酸酶被预测有助于靶向核酸的降解^[26]。然而,尽管没有催化活性四聚体和可预测的核酸酶伴侣,在大肠杆菌中异源表达的 *RsAgo* 能够导致细胞内质粒的浓度降低并且同时质粒编码基因的表达也得到降低,但这一现象在 *RsAgo* 所在的原有宿主中并未发现^[11]。类似于 *TtAgo*, *RsAgo* 不影响参与宿主防御的内源性基因的表达,这表明 *RsAgo* 并非直接调解宿主防御机制^[11] (如图 1-2b 所示)。

在生物体内,*RsAgo*与短片段干扰 RNA (siRNA) 引物 (长度为14-19 nt 的核苷酸) 相关, 这表明它们来源于退化降解的 mRNA^[11]。此外,当 *RsAgo*在大肠杆菌中异源表达时,能够 结合 siRNA 引物(即在没有遗传共定位核酸酶的情况下),引物的生成并不完全依赖于遗传共 定位核酸酶或其他 *R. sphaeroides*-特异性宿主因子^[11]。siRNA 富含在核糖体 RNA 序列中。 此外,有趣的是,siRNA 也富含来自转座子编码或质粒编码的 mRNA 序列中^[11]。有科学家 提出 *RsAgo* 对 RNA 降解是通过采样方式从而产生引导 siRNA,但是如何实现这种 siRNA 的 增加目前尚不清楚。







b RsAgo-mediated RNA-guided DNA interference



图 1-2 原核生物 long-Argonaute 蛋白对核酸干扰机制。(a) TtAgo 介导 DNA 引物的 DNA 干 扰路径。(b) RsAgo 介导 RNA 引物的 DNA 干扰路径^[28]

此外, *RsAgo*-siRNA 复合物与 DNA 分子共纯化长度为 22-24 个核苷酸,富含于来自质 粒 DNA、基因编码转座子或噬菌体基因的外源序列^[11]。与 RNA 引物的互补性表明这些 DNA 分子是共纯化的靶链。这与 *RsAgo* 对 RNA 引物的亲和力高于对 DNA 引物的亲和力和 *RsAgo* 对目标 DNA 的亲和力高于对目标 RNA 的亲和力的结果是一致的^[20]。

对 RsAgo 结构的阐述揭示了氨基酸残基与 RNA 引物的 2'-羟基化基团的特异性相互作用。同时表明,与真核生物中的 Argonaute 蛋白、PfAgo 和 TtAgo 相反, RsAgo 的 PAZ 结构



域似乎可以缺失结合引物 3'端的结合口袋结构 (如图 1-3a, b 所示)。这些在结构上的偏差 可能增强 *RsAgo*-siRNA 复合物对目标 DNA 结合的稳定性,因为这些偏差可能使 siRNA 和 目标 DNA 之间发生完全碱基配对而不引起整体构象变化;在其他原核生物的 *Argonaute* 蛋 白中,这种构象变化与目标核酸的剪切相关^[29]。虽然通过 *RsAgo*-siRNA 复合物结合底物可 能足以通过抑制转录和/或复制进而静默入侵的 DNA,但也有研究表明目标 DNA 的结合也 可以导致未知核酸酶的聚集,从而在两端切割目标 DNA 序列,生成与 *RsAgo*-siRNA 复合物 共纯化中所观察到的 DNA 片段^[11]。潜在地,与 *RsAgo* 基因共定位的核酸酶发挥了这一作用; 然而,应该注意的是,在没有这种核酸酶的情况下^[11],这些短片段的 DNA 分子也可与 *RsAgo*-siRNA 复合物共纯化得到。这表明,在 *R. sphaeroides* 中,虽然这种核酸酶可能有助 于 *RsAgo* 引导的 DNA 干扰路径,但这种酶并不是必须的。

1.1.4 其他原核生物中 long-Argonaute 蛋白

其他已经被表征的 long-Argonaute 蛋白包括细菌 Archaeoglobus fulgidus (AfAgo)、 Aquifex aeolicus (AaAgo)、Marinitoga piezophila (MpAgo)和 Thermotoga profunda (TpAgo)。 至少在体外试验中,相比于单链 RNA, AfAgo 更偏好于单链 DNA 作为引物并且与目标 DNA 结合^[30]。然而,目前还未报道这种 pAgo 是否能够剪切目标 DNA。AaAgo 通过介导 DNA 引 导的 RNA 剪切,但是对 DNA 靶向性尚未被测试^[8]。编码 MpAgo 和 TpAgo 的基因簇与编码 CRISPR-Cas 酶的基因簇具有相似性,这表明了他们在功能上具有一定的相关性^[14]。相比之 下大多数其他的原核生物 Argonaute 蛋白通常使用 5'磷酸化的引物核酸, MpAgo 和 TpAgo 均使用 5'羟基化的短片段 RNA 作为引物从而剪切同源的单链目标 DNA^[14]。此外, MpAgo 可以使用 5'羟基化的核糖核酸作为引物切割目标单链 RNA^[14]。因此,虽然 MpAgo 和 TpAgo 的作用可能与其他 long-pAgos 的作用相似,但它们所使用的引物的产生和结合机制却并不 相同^[14]。

除了这些部分被表征的原核生物中的 *Argonaute* 蛋白,其他 long-*Argonaute* 蛋白仍然未 被发现。生物信息学分析表明,一些编码 long-*pAgos* 的基因与编码 Schlafen-like 的 ATP 酶 的基因相同,然而其他被预测出现在编码 Mrr、Sir2、Cas4-like 或 PLD 核酸酶的基因操纵子 中(如图 1-3b 所示)。对于这些与 p*Ago* 相关的酶的预测功能为:引物核酸的生成,双螺旋目 标 DNA 的解螺旋,和/或目标核酸的降解。

1.1.5 short-pAgos 和 PIWI-RE 蛋白

Short-pAgos 仅由一个中间结构域和一个 PIWI 结构域组成,其中催化四分体并不完整 ^[26](如图 1-3c 所示)。虽然它们缺乏 N 结构域和 PAZ 结构域,但是编码 short-pAgos 的基因通 常与编码未表征的 APAZ 结构域(类似于 PAZ 结构域) 的基因是相关联的^[26]。编码 APAZ 结 构域的基因通常与来自 Mrr、Sir2 或 TIR 家族的预测核酸酶结构域相融合^[26]。在其他 short-pAgo 变体中,预测的 APAZ-核酸酶结构域额外与 short-pAgo 融合^[26]。或者,Sir2-APAZ 结构域融合到 Schlafen-like ATP 酶结构域,或者融合到 Mrr 和 TIR 结构域中^[4]。有研究预测 short-pAgo 的功能是作为核酸引导的目标核酸识别平台,进而与相关的 APAZ-核酸酶在引物 生成和/或目标核酸的降解起到作用。

类似于 short-p*Agos*, PIWI-RE 蛋白同样缺乏 N 结构域和 PAZ 结构域, 但是 MID 结构 域和 PIWI 结构域都是保守的, 这表明它们为核酸引导的靶标识别提供了一个平台^[31] (如图 1-3d 所示)。与 short-p*Agos* 不同,一些 PIWI-RE 蛋白具有完整的催化四分体, 这表明它们 可以剪切目标核酸^[31]。此外, PIWI-RE 蛋白在 N 端含有唯一保守的"结构域 X", 这与 long-p*Agos* 中的 PAZ 结构域和 N 结构域并无关联^[31]。该结构域可能提供为螺旋酶、核酸酶 和 PIWI-RE 蛋白的共同出现提供一个相互作用的平台^[31]。大多数的 PIWI-RE 蛋白聚集在具 有 DinG-type 解旋酶和预测核酸酶的操纵子上。DinG-type 解旋酶主要被发现在噬菌体和质 粒的复制和转录过程中 RNA-DNA 双螺旋的解螺旋过程^[31]。目前,认为这些解旋酶参与到



使双链目标核酸解螺旋并且接触到 PIWI-RE-引导核酸复合物,和/或引导核酸-目标核酸的 解离,共同存在的这些核酸酶能够参与引导核酸的生成和/或目标核酸的剪切^[31]。



图 1-3 原核生物的 Argonaute 蛋白结构域构架。(a) long-pAgo 结构域构架。(b) long-pAgo 与其他核酸酶关联的结构域构架。(c) short-pAgo 与其他核酸酶关联的结构域构架。(d) PIWI-RE 与其他核酸酶关联的结构域构架^[28]

1.2 基于原核生物 Argonaute 蛋白的基因编辑功能

1.2.1 long-Argonaute蛋白作为基因编辑工具的优缺点

由于一些 long-p*Agos* 具有可编辑的并且能够靶向 DNA 的自然性质,因此很容易尝试 将 p*Agos* 应用于 DNA 编辑。接下来,我们将描述 p*Agos* 如何已在体外被应用于切割 DNA 靶标,并且讨论与已建立的基因组编辑工具相比下 p*Agos* 的优点和缺点。

*TtAgo*和*PfAgo*都可加载引导 siDNA,从而能够在体外特异性地线性化目标双链 DNA^[12, 13, 18, 19, 32]。由于目前所描述的 p*Ago*-引导核酸复合物只能够对单链目标核酸进行剪切,因此如果需要将双链 DNA 进行剪切,需将两个具有互补性单链 DNA 的 p*Agos*与目标双链 DNA 进行孵育,从而使双链 DNA 产生断裂。双链 DNA 的剪切机理已经通过 *PfAgo* 进行详细研究,并且提供了一种在质粒骨架或 DNA 插入^[32]中任何所需突出的部分产生双链 DNA 断裂的方法。无引导核酸的 p*Agos*也可以对双链 DNA 进行剪切,也就是 chop 机制。p*Agos*应该与 siDNA 充分接触达到饱和,从而防止对目标 DNA 脱靶^[18, 19]。考虑到 siDNA 引物的合成相对便宜,并且理论上可以设计成具有专门针对任何期望位置的单链 DNA,因此该技术可以潜在的替代限制性内切酶的位置。

CRISPR-effector 复合物,由 Cas9 或 Cas12a (也称为 Cpf1) 和可变的 RNA 引物组成,



Argonaute 蛋白热稳定性及动力学研究

是公认的可编程基因组编辑工具^[33, 34]。这些复合物可通过非同源末端连接或同源重组过程^[35],将可修复的基因组中产生双链 DNA 的断裂。除了引导 RNA 和目标 DNA 之间的互补性外,CRISPR-effector 复合物还需要目标 DNA 序列旁的前驱体相邻基序 (PAM),也就是PAM 识别区。PAM 识别区的存在限制了 CRISPR-effector 复合物可靶向的位点^[33, 34],因此 使 Cas 具有一定的局限性。

能够剪切单链 DNA 的 long-p*Agos* 是作为基因组编辑工具的 CRISPR-effector 复合物的 潜在替代工具。与 Cas9 和 Cas12a 不同, p*Ag*·os 不需要 PAM 识别区识别目标 DNA,因此, 较之 CRISPR-effector 复合物,它们能够靶向更多的序列,使其应用性更为宽广。此外,p*Agos* 通常使用的引物 DNA 序列长度较短,通常长度为 15-24 个核苷酸,明显短于典型的 CRISPR-effector 复合物所需引物 DNA 序列长度(Cas9 和 Cas12a 使用引物 RNA 序列长度 分别为 100 和 43 个核苷酸)。合成短片段的 DNA 引物比合成长片段的 RNA 引物更为便且 方便,并且可能实现高通量的引物生成,进而实现高通量基因组编辑筛选。此外, long-p*Agos* 的分子量大约为 75-85 kDa,大约是 Cas9 和 Cas12a 这两种酶分子量的一半。较小的分子量 可以使传递至宿主的传递效率更高。然而,与 Cas9 和 Cas12a 两种酶相反,产生双链 DNA 的断裂需要一对 p*Ago*-引导核酸复合物,也就是说, p*Ago* 不能直接剪辑双链核酸。

通过使用一对带有互补引导核酸的 long-pAgos 在体外产生双链 DNA 的断裂的可能性 最初仅描述了源自嗜热菌的 pAgos。在高温下,嗜热的 pAgo 活性非常高,双链 DNA 可能 在较高的温度下部分结构发生解螺旋,从而使 pAgo-引导核酸复合物能够剪切双链 DNA^{[12,13, ^{32]}。显然,这一特性使嗜热的 pAgos 不太适合作为嗜温生物的独立基因组编辑工具。解旋酶 活性的缺乏可能解释了为什么目前表征的 pAgos 无法在体外常温温度下有效地剪切双链 DNA。在体内,这些嗜热 pAgos 和其他嗜常温的 pAgos 可能依赖于双链 DNA 局部的解螺 旋 (例如 RNA 转录或 DNA 复制)从而实现有效的双链 DNA 靶向等这类自然过程。在细菌 的细胞中,已经有实验表明 pAgos 能够有效地干扰双链 DNA 质粒的传播与复制^[12,16]。然而, 嗜常温的 pAgos 靶向双链 DNA 需要进一步研究。至少,已经发现一些编码 pAgos 的基因与 编码预测的解旋酶的基因共存或融合,这些基因可能能够使 pAgos 靶向双链 DNA^[4,26]。}

对靶向 DNA 的嗜温型 pAgo 的检索导致了 Natronobacterium gregoryi Argonaute 蛋白 (NgAgo)^[36]的表征。在一项研究中,报道称 NgAgo 在 37 摄氏度时显示出 DNA 引导的 dsDNA 剪切,并且有可能在人类细胞中使用 NgAgo 进行基因组编辑^[36]。然而,这项工作却存在严重的争议,因为许多实验室无法复现其所描述的结果^[37-39]。最近的一项研究表明,NgAgo 可被编程进而下调斑马鱼中的特定基因^[40]。基因的敲除也存在于活性位点缺失的 NgAgo,因此有人认为基因敲除是 NgAgo-siDNA 复合物结合 DNA 的结果,从而引起该复合物诱导转录沉默^[40]。然而,另一项研究的数据表明,与早期发现相反,NgAgo 通过 DNA 引导从而对 RNA 剪切,因此在转录后调节基因表达。这一路径与真核细胞中 RNA 干扰路径类似^[41]。然而,这些数据需要进一步验证。虽然 RNA 干扰路径是一种有效的系统,可以敲除许多生物体中令人感兴趣的基因,但是一些生物体内缺乏功能性 RNA 干扰系统。该系统可以被"劫持"以进行基因编辑,或者由于 RNA 干扰路径介导的基因敲除所产生的脱靶效应而产生毒性^[42]。因此,NgAgo-siDNA 复合物在这些生物体中可潜在地作为 RNA 干扰的替代物发挥作用。

若将 pAgo-siDNA 复合物递送到细胞中,进而使用 pAgos 作为基因组编辑工具或目标 基因敲除工具,则会存在另一个潜在的限制。CRISPR-effector 复合物可以通过蛋白质转染 传递到细胞中。由于该复合物也可以在体内进行有效地组装,Cas9 或 Cas12a 和/或 RNA 引 物可以在病毒介导的转染或质粒转导后在细胞中表达^[33]。目前,至少一些 pAgos 在没有 siDNA 的情况下能够通过"chop"机制对 DNA 进行剪切,这些 pAgos 可能通过非特异性 DNA 降解引起毒性。此外,由于 siDNA 不能在体内表达,体外组装的 pAgo-siDNA 复合物的转



导在理论上是唯一适合细胞递送的方法。即使 pAgos 可以在体内表达而不会引起毒性 (例 如,通过突变参与 DNA 剪切的残基),合成的 siDNA 的转染也可能干扰 RNA 干扰途径。之 后的研究需要进一步的探索可行的传递的方法。然而,至少对于 NgAgo,体外组装的 pAgo-siDNA 复合物的转染被证明能够适用于有效的细胞递送方式^[40]。在 2019 年,发现两种嗜常温的 pAgo 蛋白,分别为 CbAgo 和 LsAgo,这两种 pAgo 同样通过 DNA 引物剪切 DNA,但同时也发现了类似于 TtAgo 的对质粒的"chopping"机制。在之后的研究中,CbAgo 和 LsAgo 很有可能成为人体细胞内基因剪辑的有效工具。

1.2.2 Pyrococcus furiosus Argonaute 蛋白的应用

在本小节中,简要阐述 PfAgo 在基因检测中的具体原理及运用方法。如图 1-4a 所示, 首先将长度为 16 nt 的引导单链 DNA 和 PfAgo 孵育,然后将长度为 48 nt 的目标单链 DNA 加入到体系中。PfAgo 对目标单链 DNA5'端开始的第十个到第十一个核苷酸的磷酸二酯键 进行剪切。那么,长度为 48 nt 的单链 DNA 变为长度为 16 nt 的单链 DNA 和 32 nt 的单链 DNA。从图 1-4b 中的电泳胶图中可以看到,形成的长度为 16 nt 的单链 DNA,能够变为新 的引导 DNA。也就是说,在体系中存在两种引导 DNA。之后,将线性的序列长度为 2685 碱基对的长链 DNA 加入,具有两种引导 DNA 的 PfAgo 能够对这种超螺旋双链进行剪切。 从图 1-4c 中的电泳胶图中可以看到,如果在 PfAgo 中里面加入 PUC19 质粒和单链 DNA1, 而没有加入引导 DNA1,那么就不会对 PUC19 质粒进行剪切;如果同时将单链 DNA1和引 导 DNA1 加入,该体系能够对 PUC19 质粒进行剪切,降温后通过电泳跑胶可以看到剪切后 产物形成长度为 1815 碱基对的双链 DNA 和长度为 871 碱基对的双链 DNA。上述结果说明, PfAgo 能够实现两轮剪切功能,也就是能够剪切不同的底物。



图 1-4 PfAgo 实现两轮剪切的机理。(a) 单链 DNA 和 双链 DNA 的序列示意图。(b) 单链剪切产物电泳胶图。(c) 质粒剪切产物电泳胶图^[13]

由于质粒为超螺旋结构,而普通的 DNA 为双螺旋结构,因此 PfAgo 能否剪切双螺旋 DNA 仍然未知。如图 1-5a 所示,首先放入三种不同序列的引导 DNA 与 PfAgo 进行孵育从 而形成二元复合物,然后二元复合物会根据序列互补性结合到双螺旋 DNA 的特定位点从而 使 PfAgo 对目标位置进行剪切。剪切后,长度为 26 nt 的单链从双链 DNA 中被切下来形成 新的单链引导 DNA。PfAgo 和该单链引导 DNA 形成新的二元复合物后,能够剪切与该单链 引导 DNA 互补配对的单链 DNA。如图 1-5b 所示,第一列代表双链 DNA,第二列代表 3 个 长度为 16 nt 的单链 DNA,第三列代表的是合成的长度为 26 nt 的 DNA,第四列代表的是剪



Argonaute 蛋白热稳定性及动力学研究

切后产物,第五列代表的是剪切后的双链 DNA 和加入的长度为 16nt 的引导 DNA。从图 1-5b 中可以看到,双链 DNA 的确被 *PfAgo* 剪切。从图 1-5c 中可以看到,由从双链 DNA 中剪切 出来的单链 DNA 作为引导 DNA 后,可以继续剪切能够和新的引导 DNA 互补配对的单链 DNA。图 1-5d 表示了相应的剪切图示意。从以上实验结果能够表明,通过第一轮剪辑形成 新的短片段单链 DNA,可以继续和 *PfAgo* 蛋白结合形成二元复合物,进而直接引发第二轮 单链 DNA 的剪切。上述的所有反应,可以在一个试管中进行。这种方法能够极大的降低操 作复杂性。



图 1-5 引物数量对 PfAgo 剪切活性影响示意图。(a) 序列示意图。(b) 双螺旋剪切产物的电泳胶图。(c) 引物类型对剪切产物影响的电泳胶图。(d) 三种引物对剪切活性影响示意 图及电泳胶图^[13]

图 1-6 表明了基于 *PfAgo* 的基因检测技术。首先通过 PCR 或者 tHDA 对目标核酸区域 进行大幅度扩增。PCR 是通过高温使双螺旋 DNA 解螺旋变性,然后再设计特定的引物对目 标序列进行扩增; tHDA 是引进解旋酶后使双螺旋 DNA 解螺旋,然后在一定温度下进行扩 增。扩增后,目标双螺旋 DNA 含量大幅度提升。接着,在该体系中加入具有三种不同引物 的 *PfAgo*,能够将目标 DNA 片段进行剪切。将体系将至室温后,剪切下来的单链 DNA 能 够和带有荧光探针的单链 DNA 结合形成双螺旋 DNA,进而使荧光淬灭基团远离荧光基团, 从而发出荧光。因此,可以根据发光与否,检测是否有特定的核酸序列,从而能够对是否 有某疾病进行检测。





图 1-6 基于 PfAgo 的基因检测系统示意图[16]

从上述实验结果中可以看到,对于引导 DNA 序列的选择非常重要,其一方面是第二轮 基因剪切的引物,另一方面是基因检测中结合探针的部分。在该基因检测方法中,使用了 三种不同序列的引物。但是三种引物的设计较为麻烦,因此需要尝试只有一种或者两种不 同序列的引物 DNA,测试其检测活性如何。从图 1-7 中可以看到,如果只有一种引物,那 么剪切后的单链 DNA 无法从底物中掉下;如果有两种引物,剪切后的单链 DNA 很少从底 物中掉下来;而三种引物共存,则整体活性非常高。因此,在该基因检测中,三种引物的 存在必不可少。



图 1-7 引物数量对 PfAgo 剪切活性的影响[16]



在基因检测中,产生的单链 DNA 同样会对荧光产生影响。如图 1-8 和 1-9 所示,通过 设计不同的引物产生不同的单链 DNA,然后进行荧光检测。可以看到,荧光强度随着序列 的变化而变化。因此,需要探究该基因检测方法对该体系能否分辨单核苷酸突变的单链 DNA 序列。由于对于野生型的目标单链 DNA 和突变型的单链 DNA 所用的引物 DNA 都是同一 种,并且 *PfAgo* 只对特定位点进行剪切,因此不管目标单链 DNA 是否突变都能将单链 DNA 进行剪切。然而,荧光探针对剪切产物的序列是高度敏感的,所以可以通过对荧光探针的 序列设计,从而检测出是否存在突变和未突变的基因。该基因检测体系对于核苷酸的突变 具有非常高的敏感性,因此可以进行特定的突变 DNA 检测。检测系统的敏感性非常重要。 例如,对于两种 DNA,他们一个位点的突变都会引起肿瘤的产生。那么可以设计针对变异 体的荧光探针和针对野生型的进行探测进行检测。从图 1-9 中可以看到,当变异体的浓度仅 有 0.1%情况下,就能被检测出变异基因。







图 1-9 单位点突变检测灵敏性[16]

第 12 页 共 44 页



对于基于 *PfAgo* 检测系统更为潜在且强大的应用为在一个试管中,能够同时检测出多种目标 DNA 的存在。例如,如图 1-10 所示,有五种不同序列的 DNA,可通过设计十五种 不同序列的引物 DNA,使 *PfAgo* 能够对特定双链 DNA 的特定位点进行剪切,最后通过特定的荧光探针去结合剪切下来的单链 DNA,从而测出不同的荧光信号,进而能够达到一步 检测出不同的目标 DNA。



图 1-10 多基因一管法检测技术[16]

1.3 本课题的提出和主要研究内容

1.3.1 本课题的提出

在原核生物和真核生物中, Argonaute 蛋白己被证明为短片段核酸引导的基因静默过程的关键参与者。在真核生物中, Argonaute 蛋白是 RNA 引导静默复合物(RNA-induced silencing complex)的核心,通过利用 microRNA(miRNA)或短片段的 5'-磷酸化干扰 RNA 分子(siRNA)作为引导核酸将 RISC 复合物引导至与引导核酸互补的目标 RNA。在细胞质转录后基因静默过程中, Argonaute 蛋白-RNA 引导复合物可直接剪切目标 RNA 或直接静默转录的基因元素。而具体采取其中哪种方式,通常取决于 Argonaute 蛋白的类型以及引导核酸和目标核酸的序列互补性。至此, Argonaute 蛋白可以通过调节细胞内的 RNA 水平来协调各种细胞过程。

发现同源的原核生物中的 Argonaute 蛋白之后,便确定了 Argonaute 蛋白第一个全长晶体结构。由于古细菌和细菌缺乏 RNA 干扰途径,原核生物中 Argonaute 蛋白的生理功能仍然未知。最近对许多已经表征晶体结构的 Argonaute 蛋白的研究,例如 Pyrococcus furiosus Argonaute 蛋白 (PfAgo), Thermus thermophilus Argonaute 蛋白 (TtAgo), Methanocaldococcus jannaschii Argonaute 蛋白 (MjAgo), Clostridium butyricum Argonaute 蛋白 (CbAgo) 和 Limnothrix rosea Argonaute 蛋白 (LsAgo) 揭示, Argonaute 蛋白使用 5'端磷酸化的 DNA 作为引物干扰目标 DNA 的增长与复制。这一过程称为 DNA 引导的 DNA 干扰。结构学的研究进一步确定, Argonaute 蛋白将引物单链 DNA 的5'端和3'端分别固定在 MID 结构域和 PAZ 结构域,从而结合到形成稳定的二元复合物。进而成为结合互补目标核酸的模板。通过预





调节引导链的构象,种子区域中的碱基对目标核酸具有较高的亲和力并且能够特异性识别。因此,目标核酸和引导核酸之间能够相互作用。与目标核酸的碱基互补配对后,引导核酸的3'端会从 PAZ 结构域中释放,并同时伴随具有剪切功能的复合物的构象发生较大变化。 PIWI 结构域对目标 DNA5'端开始的第10个碱基与第11个碱基之间的磷酸二酯键进行剪切。 *Argonaute* 蛋白的剪切活性位点在 PIWI 结构域,其具有高度保守的催化四元体(DEDX, X = Asp, Asn 或 His),这些位点是实现剪切功能的必要氨基酸。这种典型的剪切机理表明, 原核生物的 *Argonaute* 蛋白参与宿主防御是通过引导 DNA 干扰入侵 DNA 的转录与复制, 从而抵抗古菌和细菌中的入侵核酸。

虽然 Argonaute 蛋白的基因剪切机理以及蛋白的三级和四级结构非常相似,但由于原核 生物的自然环境不同, Argonaute 蛋白的最佳活性温度不同。例如, Pyrococcus furiosus 的 Argonaute 蛋白在 87 ℃ 至 99.9 ℃ 范围内最为活跃,而 Thermus thermophilus 的 Argonaute 蛋白和 Clostridium butyricumum 的 Argonaute 蛋白分别在 75 ℃ 和 37 ℃ 时表现出最高的活 性。此外,虽然通过晶体结构已经揭示了剪切机理,但是仍然缺乏剪切过程中对 Argonaute 蛋白原位的结构跟踪。此外,不同的 Argonaute 蛋白是否显示出不同的热稳定性仍然未知。 如果存在不同热稳定性,那么在各自的最佳活性温度下,剪切过程中蛋白的结构状态也是 未知的。通过对这些问题的研究,能够了解 Argonaute 蛋白热稳定性受到什么因素影响,并 且能够根据不同的热稳定性将其运用于不同的生化环境检测中。此外,通过对剪切过程中 蛋白结构的原位跟踪,了解在其实现功能时蛋白的结构状态,从而能够对蛋白质改造工程 具有一定的指导意义。

1.3.2 本课题的主要研究内容

Argonaute 蛋白作为一种功能强大的基因编辑工具,具有非常可观的应用前景。 Argonaute 蛋白为一类 PIWI 家族。其中,具有剪切功能的 Argonaute 蛋白为 long-Argonaute 蛋白。不同的 long-Argonaute 蛋白均具有四个结构域,分别为 N 结构域、PAZ 结构域、MID 结构域和 PIWI 结构域。虽然结构相近,但由于生命体的生活环境不同,从不同生命体中提 取的 Argonaute 蛋白的生理温度不同,其跨度约为 25 ℃ 至 100 ℃。因此,探究具有不同生 理温度的 Argonaute 蛋白的热稳定性具有重要意义。此外,通过蛋白以及蛋白和相应底物的 晶体结构已经揭示 Argonaute 蛋白在实现其基因剪辑功能时构象的变化。然而,在溶液中 Argonaute 蛋白的结构变化仍然未知。本课题选取具有不同生理温度的 Argonaute 蛋白,通 过 nanoDSC 研究蛋白热稳定性,并且通过分子动力学模拟分析控制蛋白热稳定性的主要因 素。此外,通过圆二色谱和荧光染色实验原位跟踪溶液中 Argonaute 蛋白结构变化,并且通 过小角 X 射线表征蛋白在溶液中的柔性和相应的粗粒化构象。此外,我们通过生物化学实 验,定量化分析影响蛋白剪辑基因活性的控制因素。

本次工作选取三种原核生物的 Argonaute 蛋白: CbAgo (生理温度: 37 ℃), TtAgo (生 理温度: 75 ℃), PfAgo (生理温度: 95 ℃)。通过 nanoDSC 表征, PfAgo 的最为稳定, 其 次为 TtAgo 和 CbAgo。通过蛋白晶体结构 B-factor 和分子动力学模拟分析,发现嗜热的 PfAgo 和 TtAgo 较之嗜常温的 CbAgo 更为刚性,这是由于嗜热 Argonaute 蛋白具有更高的氢键和盐 桥密度以及蛋白核心区域更高的疏水基团比例。通过圆二色谱和荧光染色实验表征,发现 嗜热的 Argonaute 蛋白在其生理温度已经发生部分解折叠,而嗜常温的 Argonaute 蛋白仍然 保持完全折叠态。之后,通过 SAXS 实验表明嗜热的 Argonaute 蛋白在其生理温度的熔融态 结构更为柔性、松散和动态。我们将解折叠试剂施加于嗜热的 Argonaute 蛋白后,发现其基 因剪辑活性得到提升。最后,通过米式方程,揭示熔融态主要影响蛋白对基因的剪切速率, 而对蛋白对底物的亲和力并未影响。



第二章 热稳定性、动力学及生化反应研究方法

在本章节中,将阐述此次工作中所采用的实验方法及原理。

2.1 实验原理

2.1.1 荧光分光光度计

荧光分光光度计(PCR)是用于扫描液相荧光标记物所发出的荧光信号的一种仪器。PCR 能提供包括激发光谱、发射光谱以及荧光强度、量子产率、荧光寿命、荧光偏振等多种物理 参数,因此可从多个角度反映了分子的成键和结构情况。

PCR 的具体原理为由高压汞灯或氙灯发出的紫外光和蓝紫光经滤光片照射到样品池中, 激发样品中的荧光物质发出荧光。荧光经过滤过和反射后, 被光电倍增管所接收, 并通过图像或者数字的形式显示出来。荧光分光光度计的测试原理如图 2-1 所示。



图 2-1 荧光分光光度计的测试原理示意图

物质产生荧光是因为处在基态的物质在吸收激发光的能量后,跃迁到激发态。这些处于 激发态的分子并不稳定,在其返回基态的过程中会将一部分的能量以光的形式放出,从而产 生荧光。不同物质由于分子结构不同,导致其具有不同的激发态能级分布。即不同物质有不 同的特征荧光激发和发射光谱。

接下来,简要介绍 PCR 在蛋白质解折叠中的应用。荧光染料在溶液中的荧光强度很低,因此 PCR 无法检测到荧光信号。对于自然状态下处于折叠态的蛋白,蛋白的疏水基团在蛋白的内部,蛋白的亲水基团在蛋白的表面。由于荧光染料只能和蛋白的疏水基团作用,因此荧光染料无法产生荧光信号。当温度上升示,氨基酸残基之间的相互作用力减弱,蛋白发生解折叠,疏水基团逐渐暴露于溶液中,从而使疏水基团与荧光分子结合,导致荧光信号增强。随着温度继续升高,暴露在溶液中的疏水基团越多,导致更多的荧光染料能够结合在蛋白的疏水基团上,荧光信号增强。因此,荧光强度正比于解折叠蛋白的比例。在蛋白完全解折叠后,荧光强度下降,这是因为解折叠后的蛋白开始发生聚集,疏水基团和荧光分子被包裹在聚集后的蛋白聚集体内部而导致荧光信号无法被检测到。因此,我们可以通过荧光 PCR 方法研究蛋白的解折叠过程和热稳定性。



2.1.2 小角 X 射线散射

X 射线是一种波长比可见光短得多(0.1-10 nm)的电磁波。虽然 X 射线中的光 子所携带的能量大、穿透能力强,但仍然会有部分 X 射线被材料所吸收和散射。根据 X 射线的吸收和散射能够得到不同的结构信息。SAXS 是 X 射线散射的一种。如图 2-2 所示,它是通过分析 X 射线在材料中传播时的弹性散射并以小角度(通常为 0.1°-10°, 因此以"小角度"命名)记录散射来实现的。根据可以记录清晰散射信号的角度范围, SAXS 能够提供 1 至 100 nm 之间尺寸的结构信息。记录的角度越小,探测到的物体 尺寸越大^[43]。



图 2-2 小角 X 射线散射原理示意图^[44]

X 射线光谱的原理为 X 射线与原子核外电子作用。物质由原子构成,而原子则由 致密的原子核和外层电子构成。如图 2-3 所示,当 X 射线穿过物质时,光子会被原子 的外层电子散射。通过放置在一定位置的探测器,能够收集被材料散射后的光子信号, 进而分析得出物质的结构信息。该方法的重点是光散射现象。其中,粒子、光子或光波 与其穿过的媒介物质中的粒子相互作用而射向不同方向的现象称为散射。



图 2-3 X 射线与物质相互作用示意图^[44]

入射 X 射线在电场的作用下,物质中原子的电子会在平衡位置发生振荡,从而成为一个辐射源。这个电子起到散射中心的作用,向周围辐射与入射波相同频率的电磁

第 16 页 共 44 页



Argonaute 蛋白热稳定性及动力学研究

波。这称为相干散射,也叫汤姆逊(Thomson)散射(如图 2-4 所示)。如果将入射光 子看做粒子,它与原子束缚电子相互碰撞,根据能量守恒和动量守恒原理,散射的 X 射线波长比入射 X 射线波长长,这叫康普顿(Compton)散射,由于散射波长与角度 相关,不能进行相干叠加,也叫非相干散射。在小角 X 射线散射研究中,仅讨论相干 散射。如前所述,散射信号考虑的主要是汤姆逊(Thomson)散射。



图 2-4 汤姆逊散射示意图以及推导

其中, I_0 是入射 X 射线的强度, 2θ 是观测角, I_e 是散射 X 射线的强度。

散射强度与质量的平方成反比,质子质量是电子的 1800 倍,所以在 X 射线散射 过程中,不用考虑原子核的散射效应。另外,在小角散射中,涉及的角度 θ 很小, cosθ 接近于 1,偏振因子约等于 1。实际上,可以将电子的散射强度作为一个常数,仅需 考虑不同电子产生的相干散射的叠加效应。

多个电子的汤姆逊散射会根据光程差进行叠加。如图 2-5 所示,对两个电子而言, r 表示它们之间的位置, S₀和 S 分别表示入射光和出射光的方向的单位矢量,在无限 远处(距离远大于波长),可以视出射光为平行光,因此两者之间的相位差为 q,q 的 模为 4πsin(θ)/λ,称为散射矢量,散射强度为它与电子位置的函数。如果将原子中电 子位置通过电子密度来表示,则是图 2-5 中最后所示的公式。F(q)为原子的散射因子。 如图 2-6 所示,有了 F(q),只需按照同样的方法把它们加起来就是分子的散射信号了。 实际测量过程中得到的是散射强度,即散射振幅的平方。对于溶液中的样品,散射信 号在所有方向平均,最终得到的结果为一条一维的散射曲线。

对于诸如蛋白质的生物大分子, SAXS 相对于晶体学的优势在于不需要结晶样品。 并且, SAXS 的特性允许研究这些分子的构象多样性^[44]。根据平均粒径,形状,分布 和表面体积比等参数, SAXS 可以确定颗粒系统的微米级或纳米级结构^[45-47]。其适用 范围可以是固体或液体,并且它们可以包含相同或另一材料的任意组合的固体,液体 或气体。SAXS 不仅可以研究颗粒,还可以研究有序体系(如薄片)和分形材料的结 构。该方法准确,无损,通常只需要最少的样品制备。SAXS 的应用范围非常广泛, 包括各种类型的胶体,金属,水泥,石油,聚合物,塑料,蛋白质,食品和药品。





The scattered intensity from the isolated molecule is then

$$I(\mathbf{q}) = \left| F(\mathbf{q}) \right|^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} f_i(q) f_j(q) \exp(i\mathbf{q} \cdot (\mathbf{r}_i - \mathbf{r}_j))$$

图 2-5 (a)和(b) 多个电子的汤姆逊散射示意图以及推导

2.1.3 远紫外圆二色光谱仪

远紫外圆二色光谱仪(far-UV CD spectroscopy)可以快速测定蛋白质的二级结构和折叠 特性。在蛋白质领域中,远紫外圆二色谱最广泛的用途是用于确定蛋白质是否折叠以及二级 结构的相对比例,以及探究氨基酸位点突变对蛋白的构象和稳定性影响。另外,它可以用于 研究蛋白质与蛋白质或蛋白质与配体之间的相互作用。far-UV CD 谱的优势在于能够快速测 试包含 20 μg 或更少蛋白质质量的生理缓冲液,并且给出定量的二级结构信息。但相比于 X 射线或 NMR,它无法提供残基的特定信息。

far-UV CD 谱的测试原理: 生物分子具有不同的右旋和左旋成分,它们能够表现出圆二 色性。不同的二级结构具有自身独特的圆二色谱特征曲线。圆二色性的定义为左旋和右旋圆 偏振光的不平等吸收。一光束具有与时间相关的电场和磁场。电磁辐射由彼此垂直、并沿传 播方向振荡的电场(*E*)和磁场(*B*)组成。当电场矢量仅在一个平面内振荡时会发生线性 偏振光,而当电场矢量的方向围绕其传播方向旋转而矢量保持恒定幅度时,则会发生圆偏振 光。该光束通过棱镜或滤光器极化后,其电场 *E* 将会在单个平面上呈震荡正弦曲线。从正 面看时,正弦波可以可视化为两个等长向量的结果,该向量描绘出圆圈,其中一个顺时针旋 转(*E*_R),另一个逆时针旋转(*E*_L)。在空间的单个点处,圆极化矢量将在波频率的一个周期 内绘制出一个圆。图 2-7 显示了在某一时刻某个位置的线性和圆偏振光的电场矢量,其中圆 极化矢量沿着传播方向(*k*)形成一个螺旋。



图 2-6 (a) 线性偏振光。(b) 圆偏振光

第 18 页 共 44 页



如图 2-7 所示,当不对称分子与光相互作用时,它们吸收左旋与右旋的圆偏振光的程度 不同(因此称为圆二色性),并且对于两个波也具有不同的折射率。这导致了光波平面旋转, *E*_R和 *E*_L向量的相加形成了椭圆向量,则形成椭圆偏振。CD 可以以 ΔE 为单位,不对称分子 对 *E*_R和 *E*_L的吸光度差异或椭圆度来表征。其中,椭圆度定义为正切角是切线的短轴与长轴 之比。



图 2-7 (a) 当 *E*_L和 *E*_R成分被同等吸收时,产生线性偏振光。(b) 当 *E*_L和 *E*_R成分被不 平等吸收时,产生椭圆偏振光

由于不同的蛋白质采用不同的构象,因此其生理功能具有高度特异性。也就是,蛋白质的结构确定蛋白质的功能。因此,对于蛋白质结构的测定对于理解蛋白质的功能至关重要。 不同的波长能够得到蛋白质不同的结构信息。如图 2-8 所示,远紫外 CD 光谱有两个吸收带: 在 190 nm 附近有很强的 pi 跃迁、在 210-220 nm 之间有较弱但较宽的 pi*跃迁。因此,在远 紫外(180-250 nm)区域能够探测蛋白质的 alpha helix 和 beta sheet,以及根据吸收峰的强弱 定量 alpha helix 和 beta sheet 的相对比例。



图 2-8 在远紫外 CD 光谱中观察到的(a) 肽键; (b) 相应的 UV 跃迁^[48]

上述两个吸收带产生了不同的特征带,可以对这些特征带进行反卷积,以估计在不同溶 液和环境条件下蛋白质的二级结构成分。如图 2-9 所示,虽然四种二级结构(alpha-helix、 beta-sheet、loop 和 random coil)具有各自的特征峰,但是由于大多数蛋白质均具有这四种 结构导致 CD 谱的信号曲线变得较为复杂。因此,通过 CD 谱分析二级结构,需要基于一组 具有已知二级结构成分的参考光谱曲线对实验数据进行拟合,从而通过回归分析估算样品光 谱中的各二级结构所占比例。







图 2-9 (a) 不同二级结构的圆二色光谱。(b) 不同蛋白质的圆二色光谱^[49]

CD 谱常用于蛋白质稳定性研究。温度的变化,pH 的变化等会影响生物分子的结构并 诱导其发生聚集。这种蛋白的聚集不仅能够引发生命体内各种疾病,而且影响生物药物的功效。通过 CD 谱能够能够定量且原位表征蛋白质折叠、解折叠以及聚集过程,并且能够探究 溶剂、pH 和配体对生物分子结构的影响。例如,如图 2-10 所示,在升温过程中监测目标吸 收峰的峰值随温度的变化,从而评估蛋白质的热稳定性。



图 2-10 (a) 在 222 nm 处获得的蛋白质熔融曲线。(b) 从 20°C 到 90°C 的 CD 光谱图进行了 测量,以评估溶菌酶二级结构的热稳定性^[49]

2.1.4 差示扫描量热仪

差示扫描量热法(Differential Scanning Calorimetry, DSC)属于量热学的一种实验方法。



基本原理是用温度程序(升/降/恒温)控制样品,观察样品端和参比端的热流功率差随温度 或时间的变化过程,从而得到样品在温度程序过程中的吸热、放热、比热变化等相关热效应 信息,并由此计算热效应的吸放热量(热焓)与特征温度(起始点,峰值,终止点等)。

DSC 不仅可以测定相变温度点,而且可以记录相变时的热量变化。DSC 曲线上的放热 峰和吸热峰分别代表材料放出的热量和的吸收热量。该法对试样产生的热效应能及时得到应 有的补偿,使得试样与参比物之间无温差、无热交换,试样升温速度始终跟随炉温线性升温, 保证了校正系数恒定,故而测量灵敏度和精度较高。

图 2-11 提供了差示扫描量热仪电池组件示意图。它由两个相同的池(坩埚)组成,其 中一个用于保存待研究的样品,另一个用于保存参考溶液(通常与制备样品溶液时使用的缓 冲液相同)。当温度以恒定的扫描速率升高或降低时,通过反馈机制,样品池与参考池(S 和 R)之间的温差(ΔT)保持等于零。为了保持ΔT = 0,必须连续监测温度变化对样品池 施加的差分功率。该差分功率与电池之间的热容差成正比,由此便构成了仪器测量的基本量。



图 2-11 (a) 差示扫描量热仪的电池组件示意图。(b) Netzsch DSC204F1 的结构示意图[50]

在通过扫描速率归一化之后,该差分电功率(通常以μW=μJ/s或以μcal/s给出)会 产生两个池之间的热容差异(单位为μJ/deg或μcal/deg)。理想情况下,如果两个电池相 同,则一次扫描就足以确定样品和参考溶液之间的热容量差异。然而,实际上,这两个单元 永远不会完美匹配,并且需要两次单独的扫描以减去和消除仪器效果。



图 2-12 热流型差示扫描量热仪的基本原理示意图^[50]

现代 DSC 装置包括样品坩埚、参比坩埚样品及传感器等,以 Netzsch DSC204F1 为例, 其结构示意图如图 2-11 (b)。两者之间保持热对称。开始测试后,温差信号是由腔体内两坩 间的一对热电偶(参比热电偶,样品热电偶)连续测量得到的,其基本原理示意图如图 2-12。 加热炉对样品坩埚进行热传导。我们需要把温差信号转化为热流信号给到传感器,故而利用



傅立叶热传导方程,由于热阻的存在,可知两端的加热热流差与温差信号成比例关系。所以, 只需通过热流校正,我们就能得到热流差信号。对该信号进行时间-温度作出连续图谱,即 可得到 DSC 图谱。

在我的课题中,差示扫描量热仪是用来测试蛋白未解折叠与解折叠态之间转变的能量与 其中热力学机制的。通过测量蛋白质或其他大分子的表观摩尔热容随温度的变化。随后对该 量的操纵产生了转变的完整的热力学特征。通常,可以从 DSC 获得三种不同类型的信息^[51], 即分子的热容;总体的热力学参数(温度引发的相变过程中的焓变、熵变、热容变化);以 及配分函数、附随的中间态及其热力学参数。





图 2-14 典型的量热扫描球状蛋白质的示意图。蛋白质溶液的热容(C_{p,p})显示了解折叠的特征峰。缓冲溶液的热容(C_{p,b})高于蛋白质溶液的热容。球状蛋白质原始状态的部分热容在 20 °C 时为 0.3 cal/g • K^[52]



图 2-15 热容变化示意图。虚线表示未折叠状态($C_{p,u}$)和自然状态($C_{p,n}$)的摩尔热容。这些值之间的差异定义了热容变化($\Delta C_p = C_{p,u} - C_{p,n}$)^[52]

在 DSC 数据分析中,需要使用统计热力学工具定义的最重要的量是平均过量焓<ΔH>。 该数量是在过渡期间填充的所有状态的焓贡献的总和。

第 22 页 共 44 页





图 2-16 从蛋白质的摩尔热容中减去自然状态的摩尔热容,即可得到< ΔC_p >。由两项构成: < $\Delta C_{p,tr}$ >和< $\Delta C_{p,bl}$ > $^{[52]}$

<ΔC_{p,tr}>是过渡过量热容量函数,在热容函数中定义特征过渡峰。<ΔC_{p,bl}>定义了基线中的"S形"偏移,通常与蛋白质解折叠或其他以ΔC_p呈正向变化为特征的转变有关。

$$\Delta H = \int_{T_0}^{T_f} < \Delta C_{p,tr} > dT$$
$$\Delta S = \int_{T_0}^{T_f} < \Delta C_{p,tr} > dlnT$$
图 2-17 由 DSC 曲线得出焓与熵

由图 2-17 中的公式,在对 DSC 曲线扣除极限后,通过积分,就能得到样品在解折叠前后的焓变与熵变。

2.2 本章小结

在本章节中,对小角 X 射线散射,DSC,圆二色谱以及荧光分光光度计的测试原理进行了详细地阐述。通过对测试原理的理解,能够更好的帮助理解实验设备对样品测试的物理量是什么,测得的是什么样的信息。

此外,对同一样品进行多种实验方法测试,不仅能够获得多种样品的结构和动力学信息,也能够使实验结果之间相互验证和相辅相成,使得结果更为可靠。



第三章 Argonaute 蛋白热稳定性及基因剪切机理研究

基因剪辑技术,例如分子诊断和基因改造,对于生命科学、生物安全、医学治疗和环境跟踪非常重要^[53-57]。近期,CRISPR-Cas系统引起了学术界极大的注意,这种系统能够使科学家对核酸(DNA和RNA)的特定位点进行改变,从而导致相应物理性状发生改变。此外,通过该技术能够实现医疗分子诊断^[56-60]。类似于CRISPR-Cas系统,Argonaute蛋白是一种核酸引导内切酶。目前,许多科学家计划将该蛋白酶应用于一系列生物科技^[28, 61-63]。在CRISPR-Cas系统中,Cas核酸酶需要PAM识别区才能对特定序列进行剪辑^[64],而Argonaute蛋白并不需要该区域即可对核酸进行编辑,从而导致其具有更好的通用性和高效性。此外,Argonaute蛋白能够使用较之单链RNA更为稳定的单链DNA作为引物,而Cas核酸酶只能使用单链RNA作为引物。此外,较之于单链RNA的合成,单链DNA的合成成本更为低廉。在合适的生化条件下,Argonaute蛋白能够以极高的效率剪切和短片段干扰DNA(siDNA)或短片段干扰RNA(siRNA)相互补配对的单链DNA和单链RNA^[13, 20, 61, 65-68]。与CRISPR-Cas系统对比,Argonaute蛋白具有显著提高非侵入性检测的现实应用可能性和临床应用可能性的巨大潜力。例如液体活组织检查,特别是躯体内变异基因的早期诊断和多通路高效且精确体外的核酸检测等^[62, 63, 69, 70]。

Argonaute 蛋白是一个非常大的 PIWI 家族。不同的 Argonaute 蛋白具有不同的生理活性 温度、以及对引物核酸、目标核酸、核酸序列的偏好性等^[65]。例如, TtAgo 能够结合最短的 引物核酸序列长度为9 nt, CbAgo 为 14 nt, PfAgo 为 15 nt; TpAgo, ToAgo, TcAgo, FpAgo, CbAgo 各自的生理温度分别为 85 °C, 90 °C, 65 °C, 95 °C, 37 °C。Argonaute 蛋白根据生物 来源可分为真核生物中的 Argonaute 蛋白(eAgo)和原核生物中的 Argonaute 蛋白(pAgo) ^[4]。eAgo 主要利用单链 RNA 作为引物进而剪辑目标单链 RNA^[9]; 而 pAgo 主要利用单链 DNA 作为引物进而剪辑目标单链 DNA^[4]。Argonaute 蛋白剪辑的特定位点为从引物核酸 3'端开始 的第 10 和 11 位之间的磷酸二酯键^[13, 70]。因此,通过设计特定序列的引物核酸,能够对目 标核酸进行特定位点剪切,从而实现基因编辑。目前,对于 Argonaute 蛋白基因剪辑机理主 要通过晶体结构进行研究,而 Argonaute 蛋白热稳定性的研究和其在溶液中的核酸剪辑机理 的研究仍然未曾报道。因此,该方面的研究能够对 Argonaute 蛋白在未来应用于不同生物医 药和医疗诊断方面具有重要的意义。在本次工作中,我们选取了三种不同的 pAgo,分别为 生理温度为 95 °C 的 PfAgo,生理温度为 75 °C 的 TtAgo 和生理温度 37 °C 的 CbAgo^[13, 61]。 通过对这三种 pAgo 的研究能够使相应的研究结果具有一定的普适性。

3.1 Argonaute 蛋白热稳定性

首先,我们对 *PfAgo*, *TtAgo* 和 *CbAgo* 的结构进行探究。对于长 p*Ago* 家族,也就是具 有剪辑功能的 p*Ago*,其典型的整体结构为双叶型架构^[8,69]。如图 3-1a 所示, p*Ago* 具有四个 结构域,分别为 MID 结构域(紫色)、PIWI 结构域(蓝色)、N 结构域(绿色)和 PAZ 结 构域(红色)。不同的结构域具有不同的功能: MID 结构域为结合引物核酸的 3'端, PIWI 结构域对目标核酸进行定点剪切, N 结构域解离引物核酸和目标核酸复合物, PAZ 结构域结 合引物核酸 5'端^[17,29,71]。图 3-1b 为 *PfAgo*, *TtAgo* 和 *CbAgo* 三种蛋白晶体结构的位置重合, 可以看到,三种具有不同生理温度的 p*Ago* 具有相似的四级结构和三级结构,但是其 alpha-helix 和 beta-sheet,也就是二级结构却并不相同(如表 3-1 所示)。其中,嗜热的 *PfAgo*



和 *TtAgo* 较之 *CbAgo* 具有更高的 alpha-helix 和 beta-sheet 含量。通过 nanoDSC 测试所得到 的熔融曲线表明, *PfAgo*, *TtAgo* 和 *CbAgo* 的熔点分别为 46, 75 和 100 ℃ (如图 3-2 所示)。 并且,在室温下, *PfAgo* 和 *TtAgo* 具有更高的解折叠自由能(如图 3-1c 和图 3-3 所示)。这 些结果表明嗜热 pAgo 较之于嗜常温 pAgo 更为稳定。



图 3-1 PfAgo, TtAgo 和 CbAgo 的结构分析。(a) pAgo 的结构域示意图。PAZ 结构域、 MID 结构域、N 结构域和 PIWI 结构域分别着色为红色、黄色、绿色和蓝色。(b) 左图: PfAgo, TtAgo 和 CbAgo 中结构域的氨基酸序列。右图: PfAgo (绿色), TtAgo (蓝色) 和 CbAgo (黄色) 晶体结构的位置重合。(c) PfAgo, TtAgo 和 CbAgo 在室温下的解折叠自由能 变化。(d) PfAgo, TtAgo 和 CbAgo 晶体结构的 B-factor 对比。(e) PfAgo, TtAgo 和 CbAgo 的氢键密度和盐桥密度

p <i>Ago</i> s	Alpha-helix (%)	Beta-sheet (%)	Turn and loops (%)
PfAgo	31	29	40
TtAgo	29	31	40
CbAgo	29	25	46

表 3-1 PfAgo, TtAgo 和 CbAgo 二级结构相对含量(质量分数)





图 3-2 PfAgo, TtAgo 和 CbAgo 的熔融曲线。测试方法为 nanoDSC

p <i>Ago</i> s	Hydrophobic residue (%)
PfAgo	83
TtAgo	82
CbAgo	76

表 3-2. PfAgo, TtAgo 和 CbAgo 核心区域疏水氨基酸基团比例

为了比较蛋白整体结构的动力学,我们将 PfAgo, TtAgo 和 CbAgo 三种蛋白晶体结构的 B-factors 进行对比。如图 3-1d 所示,嗜热 pAgo 较之嗜常温的 CbAgo 更为刚性。通过分子 动力学(MD)模拟蛋白在水溶液中的情况发现相同的结果。这一结果表明,在溶液中,CbAgo 较之 PfAgo 和 TtAgo 更为柔性(如图 3-4a-c 所示)。此外,通过 MD 模拟我们发现,嗜热 pAgo 的氢键密度和盐桥密度更高(如图 3-1e 所示)。并且由于嗜热 pAgo 的核心处具有更高比例 的疏水氨基酸残基,导致其具有较强的非极性范德瓦尔斯力(如表 3-2 所示)。基于上述结 果分析与比较,可以得知,在嗜热 pAgo 蛋白中较强的残基间相互吸引作用能够增加焓变, 分子链的刚性能够降低熵变,从而导致更高的解折叠自由能变化。也就是,嗜热 pAgo 具有 更高的热稳定性。PfAgo, TtAgo 和 CbAgo 采用的是类似于核糖核酸酶(RNase H)的催化 活性中心四聚体,分别为天冬氨酸-谷氨酸-天冬氨酸-组氨酸^[13],天冬氨酸-谷氨酸-天冬氨酸 -天冬氨酸^[29]和天冬氨酸-谷氨酸-天冬氨酸-组氨酸^[13],天冬氨酸-谷氨酸-天冬氨酸 -天冬氨酸^[29]和天冬氨酸-谷氨酸-天冬氨酸-石糖同温度下表现出不同的剪切活性。 例如,CbAgo 能在室温下剪切底物而 TtAgo 和 PfAgo 无法剪切。我们通过对活性中心位点周 围的氢键和盐桥进行分析,发现 PfAgo 和 TtAgo 的氢键和盐桥数多于 CbAgo(如表 3-3 所示)。 并且,通过 MD 分析发现,嗜热的 PfAgo 和 TtAgo 的结构域较之 CbAgo 更为紧实(如图 3-4d





所示)。

因此,可以推断,较高的氢键数和盐桥数提高了嗜热蛋白整体的稳定性和活性中心的稳定性,从而使其能够在高温下实现功能。然而,由于 pAgo 的在实现其生化功能时,需要进行构象变化,则这种刚性并且紧凑的结构可能导致嗜热 pAgo 在常温下难以发生构象改变从而对底物的没有剪切活性。

p <i>Ago</i> s	Hydrogen bond	Salt bridge
PfAgo	8	4
TtAgo	6	4
CbAgo	5	2

表 3-3. PfAgo, TtAgo 和 CbAgo 催化中心周围氢键和盐桥数量



图 3-3 PfAgo, TtAgo 和 CbAgo 的解折叠自由能改变随不同解折叠试剂尿素浓度的变化





图 3-4 通过 MD 模拟分析 pAgo 结构。(a) MD 模拟示意图。(b) 室温下溶液中 PfAgo 和 CbAgo 的 B-factor 对比。(c) 室温下溶液中 TtAgo 和 CbAgo 的 B-factor 对比。(d) 室温下溶液中 PfAgo, TtAgo 和 CbAgo 的 R_g

3.2 Argonaute 蛋白在生理温度时构象

3.2.1 二级结构和三级结构表征

虽然传统的目标核酸剪辑机理已经通过具有单链 DNA(ssDNA)作为引导核酸和单链 DNA 作为目标核酸的 pAgo 三元复合物晶体结构所揭示,但是在 pAgo 的生理温度下原位监测溶液中蛋白结构的变化并未报道。在本小结中,我们利用远紫外圆二色谱(far-UV CD spectra)和荧光染料法(fluorescence dye method)分别用于记录在不同温度下无配体的 pAgo 和三元复合物(蛋白质-引导 DNA-目标 DNA)的二级结构和三级结构演变。对于蛋白二级结构的解折叠曲线是通过测量样品在 222 nm 波长处的椭圆率;对于蛋白三级结构的解折叠曲线是通过测量样品在 350 nm 波长处的荧光强度。引人注目的是,图 3-5a 表明, PfAgo 和 TtAgo 三级结构起始解折叠的温度(Ton)明显低于他们的生理温度(Tphy)。与之形成鲜明对比的是,CbAgo 的起始解折叠温度高于其生理温度(如图 3-5a 所示),并且这一结果与大多数典型的球状蛋白相似。也就是,蛋白实现其生化功能必须保持稳定的结构。此外,如图 3-5b 所示,三种 pAgo 的二级结构解折叠曲线也显示了相似的结果。因此,上述实验结果表明,嗜热的 PfAgo 和 TtAgo 在其生理温度时,蛋白的二级结构已经部分解折叠,并且三级结构也发生了一定的破坏。而对嗜常温的 CbAgo 在其生理温度时,处于完整的结构折叠态。

为了进一步证明在体外基因剪辑过程中 pAgo 的状态,我们在室温和各自的生理温度下,



将 pAgo 与 ssDNA 引导核酸和 ssDNA 目标核酸进行孵育从而形成三元复合物(图 3-6 中显示了引导核酸和目标核酸的核苷酸序列)。在生理温度下,与引导核酸和目标核酸结合的 CbAgo,其二级结构是完整的,与非活性温度下的 CbAgo 的二级结构相同(图 3-5c)。相比 之下, PfAgo 和 TtAgo 的二级结构在其各自的生理温度下已经发生了熔融(图 3-5d-e)。毫无 疑问, PfAgo 和 TtAgo 中的熔融状态,即保留了部分天然蛋白的结构,是嗜热 pAgos 本身的 固有性质,而非由引导核酸和目标核酸引发的构象变化。Mn²⁺是 pAgo 实现其功能的必要离 子。通过晶体结构发现,Mn²⁺主要结合在活性中心四聚体处,但是 Mn²⁺对于蛋白的稳定性 是未知的。我们通过对加入 Mn²⁺之后蛋白的解折叠过程进行研究。我们发现,如图 3-7 所示,Mn²⁺能够使 pAgo 的结构发生部分的解折叠。也就是说,Mn²⁺在一定程度上使蛋白的结构更为松散进而增加蛋白的柔性。这种蛋白柔性的提升可能能够增加 pAgo 在基因剪切过程 中构象的变化。



图 3-5 PfAgo, TtAgo 和 CbAgo 三级结构和二级结构的热熔融曲线。(a) PfAgo, TtAgo 和 CbAgo 三级结构的解折叠曲线。(b) PfAgo, TtAgo 和 CbAgo 二级结构解折叠曲线。(c) CbAgo-引导 核酸-目标核酸三元复合物远紫外圆二色谱。(d) TtAgo-引导核酸-目标核酸三元复合物远紫 外圆二色谱。(e) PfAgo-引导核酸-目标核酸三元复合物远紫外圆二色谱。图中的箭头和虚线 分别代表蛋白的起始解折叠温度和生理温度

5' C C T A C G C G C G C G G C T C C A A C T A C C A A 3' 3' G C G G T G G T C G A G G T T T -P 5'

图 3-6 引导单链 DNA 和目标单链 DNA 的核苷酸序列

蛋白的熔融态是一种由熵驱动的热力学平衡状态,这种状态与蛋白的天然状态和变性 状态明显不同。蛋白熔融态的结构致密性较低,通常具有一些类似于天然蛋白的二级结构, 但更重要的是其具有动态的三级结构^[72]。目前,在载脂蛋白 A-1^[73],细胞色素 c^[74]等蛋白中 发现了这种蛋白的熔融态^[75-79]。蛋白的熔融态是"半刚性"的,其保留了蛋白天然折叠模式的 主要特征。与典型蛋白核心区域较为刚性相反,蛋白的熔融态其内部更容易发生运动。例



Argonaute 蛋白热稳定性及动力学研究

如,包括许多内部侧链的旋转异构化和蛋白环的柔性^[80]。蛋白熔融态的特性对于活细胞中 蛋白必须根据各种不同条件进行构象调节而言似乎是理想的。例如在细胞质中,细胞膜附 近或内部,细胞器中等,蛋白需要发生构象变化或者解折叠,从而实现其生理功能^[77]。构 象变化对于 p.Ago 基因剪辑功能的实现至关重要。例如 PAZ 结构域在结合引导 DNA 的 3' 端后会发生空间重排^[17,22],引导 DNA 和目标 DNA 形成 DNA 双螺旋后 MID 结构域会释放 引导 DNA 的 3'端^[29],在对目标 DNA 剪切过程中的催化残基的定向排列以及剪切完毕后的 产物的释放等^[29]。与嗜热 pAgo 相比,嗜常温的 CbAgo 具有相对柔性的结构,并且可以有 效地通过自身调节从而实现其功能所需各个阶段所需的构象。相反, PfAgo 和 TtAgo 结构的 刚性会抑制构象的重排,并且由于较多的盐桥数和氢键数使其在较低温度下能够将催化活 性构象锁定在一定的位置从而导致无法调节构象实现相应功能。在生理温度下, 嗜热 pAgo 的熔融态保留了一定数量的天然结构,但却具有动态和灵活的构象集合,使其能够进行快 速的构象变化,从而达到实现生化功能的最佳构象。诱导契合模型(induced-fit model)被 视为通过局部底物引导的结构重排将一种紧密构象的集合转换为另一种独特的集合。也就 是,在不存在配体的情况下蛋白质仅采用一种构象,而通过配体结合后能够诱导成为具有 功能活性的构象。相比之下,嗜热 pAgo 的熔融态具有很高的柔性和多样化的平衡前异构体, 包括非活性的构象和具有活性的构象。在结合底物后能够使平衡向有利于活性构象的方向 移动从而实现相应的生化功能。



图 3-7 加入 Mn²⁺后 PfAgo, TtAgo 和 CbAgo 的热解折叠曲线。(a) 加入 Mn²⁺前后 PfAgo, TtAgo 和 CbAgo 的三级结构解折叠曲线。(b) 加入 Mn²⁺前后 PfAgo, TtAgo 和 CbAgo 的二 级结构解折叠曲线。图中的箭头代表蛋白的起始解折叠温度

与现代酶中的底物和催化特异性相比,在进化上,蛋白的熔融态可能具有显著的优势。 这是由于蛋白熔融态结构上的灵活性能够使其在古生物的细胞中发挥多种功能^[74, 81, 82],从 而为具有不同功能的蛋白的演化提供模版。根据 pAgo 高度保守的催化位点,催化效率可能 并不是天然蛋白质结构演变的驱动力。取而代之的是,功能的特异性,对不同温度或者 pH 值的内在稳定性,分子特异性识别或不适当的细胞运输等,都将有利于现代酶结构的明确 化。考虑到它们可能的传承,很容易推测蛋白熔融态可能是现在高度进化酶的起点。因此, 根据这一特性,在蛋白的设计或者从头设计中,可以根据蛋白的熔融态来设计出能够适应 不同环境的特定蛋白。

3.2.2 小角 X 射线散射表征

为了进一步探究 PfAgo, TtAgo 和 CbAgo 的微观结构如何随温度变化,我们将原位小角



X射线衍射对在不同温度下且在溶液中的 PfAgo, TtAgo 和 CbAgo 进行表征。从图 3-7 中可 以看到, 室温下溶液中三种 pAgo 的结构与晶体结构类似。在图 3-8a, 3-8b 和 3-8c 分别显 示了 PfAgo, TtAgo 和 CbAgo 的 Kratky 分析图。 Kratky 分析图(I(q) • q²和 q 的函数关系) 用于定性识别无序状态的蛋白并将其与球状蛋白区分^[83]。对于球状蛋白,其在低 q 区域具 有清晰的峰值^[83-85],并且曲线峰型呈现钟形。相反,对于理想的高斯链,曲线在高 q 值区 域为平台,这种曲线形状通常在未折叠的蛋白质中观察到,例如,内在无序蛋白^[83-85]。如 图 3-8a 所示, CbAgo 在其生理温度的 Kratky 曲线与室温下的相同,这表明嗜常温 pAgo 呈 球状并且为折叠的结构。这一结构,与 CbAgo 的晶体结构相似。也就是说, CbAgo 在生理 温度下其溶液中的构象与其晶体结构相同。将 CbAgo 升温至解折叠温度时,可以看到高 q 值区域平台上升,说明 CbAgo 中产生了高斯链。图 3-8b 和图 3-8c 分别表示的是 PfAgo 的 Krathy 曲线和 TtAgo 的 Krathy 曲线。在其各自生理温度下的 Kratky 曲线与其在室温下的 Kratky 曲线相比,代表结构特征的峰值明显降低,表明蛋白发生了熔融。然而,高q区域 的重合表明蛋白中并未有高斯链的形成。这些结果进一步表明,与 CbAgo 解折叠不同,嗜 热 pAgos 在其各自的生理温度下已经发生部分解折叠,也就是出现了熔融态,但这种熔融 态并不产生新的高斯链。这一结果与圆二色谱和荧光染料测试的结果一致。为了探究嗜热 pAgo 的熔融态是否增加了蛋白整体的柔性。我们将熔融态的 Porod-Debye 曲线与室温状态 下的 Porod-Debye 曲线进行对比。如图 3-8h-i 所示,正如所预期的, PfAgo 和 TtAgo 的熔融 态较之它们在室温下的构象,其整体柔性增加。因此,可以推测,这种柔性的蛋白状态能 够显著增加各种功能构象之间的快速变化,从而使蛋白的催化效率得到提升。为了得到可 视化的嗜热 p.Ago 熔融态构象,我们通过 SAXS 数据对嗜热的 PfAgo 和 TtAgo,嗜常温的 CbAgo 进行粗粒化模型模拟,相应的模型如图 3-8d-f 所示。从图中可以看到,虽然 PfAgo 和 TtAgo 在其生理温度下已经发生融化,但它们仍然呈现球状构象。这一结果说明嗜热 pAgo 的熔融 态需要维持球状构象才能实现其生化功能。如图 3-8g 所示, CbAgo 在其生理温度下的回转 半径(Rg)与其晶体结构相似,约为2.9 nm,但 PfAgo 和 TtAgo 熔融态的回转半径明显大于 其晶体结构,约为3.8 nm。这一结果表明,在它们各自的生理温度下,嗜热 p.Ago 的构象更 加延伸且松散,但仍然保持呈球状构象。



图 3-7 室温下溶液中 pAgos 的 SAXS 曲线。(a) 溶液中 PfAgo 的 SAXS 曲线与 MD 模拟中 PfAgo 的 SAXS 曲线对比。(b) 溶液中 TtAgo 的 SAXS 曲线与 MD 模拟中 TtAgo 的 SAXS 曲线对比。(c) 溶液中 TtAgo 的 SAXS 曲线与 MD 模拟中 TtAgo 的 SAXS 曲线对比





3.2.3 生物化学表征

在上述的结果中,我们发现, PfAgo和 TtAgo在其生理温度时发生部分解折叠。那么, 当适当增加解折叠比例后,嗜热 pAgo对底物的剪切活性是否能够得到提升。因此,在本小 节中,我们通过解折叠试剂尿素提高蛋白整体的解折叠比例,进而探究其对目标 DNA 剪切 活性。传统的核酸剪切活性检测是通过反应体系在反应一定时间后,通过电泳跑胶方法测 定产物的产率,该方法的缺点是无法实时跟踪产物变化,并且实验复杂,效率较低。因此, 我们设计了一种新的核酸剪切活性检测系统,该系统能够实时跟踪并且定量检测蛋白的剪 切活性。其中,荧光信号的强度与目标单链 DNA 被剪切的数量成正比,因此通过对荧光信 号的测定能够定量化剪切产物的数量。我们设计了特定序列的引物单链 DNA 和目标单链 DNA。其中,目标单链 DNA 的 3'端带有淬灭基团(BHQ1),在其 5'端带有荧光基团(6-FAM) (如图 3-9a 中的上图所示)。当通过蛋白剪切目标单链 DNA 后,淬灭基团和荧光基团会相 互分开,从而由探测器检测到荧光信号(如图 3-9a 中的下图所示)。因此,通过记录荧光强 度随时间的变化可进而监测蛋白剪切目标单链 DNA 的活性。如图 3-9b 和 3-9c 所示,我们 将 PfAgo和 TtAgo 与一定浓度的解折叠试剂尿素进行孵育,发现 PfAgo和 TtAgo 的解折叠比 例增加。此外,如图 3-9d 和 3-9e 所示,解折叠比例增加后蛋白的整体柔性得到了提升。有 趣的是,与相同温度下的无尿素体系相比,由尿素增加的 PfAgo 和 TtAgo 中熔融态促进了蛋



白对目标单链 DNA 的剪切活性(如图 3-9f 和 3-9g 所示)。然而,在更高的温度的下,尿素 并不能增加蛋白的剪切活性,这是因为过度的解折叠导致蛋白的结构发生过多的破坏,从 而导致活性降低。此外,如图 3-10 所示,根据荧光强度随时间变化,我们发现蛋白的单位 时间剪切产物量和最终剪切产物的产率都得到显著提高。为了排除尿素对目标单链 DNA 降 解的影响,我们将引导单链 DNA,目标单链 DNA 和尿素进行孵育,并将孵育后体系通过 电泳迁移率变动(EMSA)进行分析。如图 3-11 所示,发现尿素并未对单链 DNA 进行降解。 因此,这些结果可以表明,嗜热 pAgo 中的蛋白的熔融态对基因剪切活性具有重要意义。并 且,适当提高嗜热 pAgo 的解折叠比例,能够增加蛋白对底物的剪切活性。



图 3-9 不同温度下尿素对于 PfAgo 和 TtAgo 剪切活性的影响。(a) 上图:合成的长度为 16 nt 的引导单链 DNA (蓝色) 和合成的长度为 28 nt 的目标单链 DNA (红色),在目标单链 DNA 的 5'端和 3'端分辨标记 6-FAM 荧光基团和 BHQ1 淬灭基团。下图:通过荧光强度检测蛋 白剪切活性示意图。(b) 添加尿素前后 PfAgo 解折叠曲线。(c) 添加尿素前后 TtAgo 解折叠

曲线。(d) 添加尿素前后 *PfAgo* 的 Porod-Debye 曲线。(e) 添加尿素前后 *TtAgo* 的 Porod-Debye 曲线。(f) 不同温度下通过荧光信号强度定量 *PfAgo* 的剪切活性。(f) 不同温度下通过荧光信号强度定量 *TtAgo* 的剪切活性。在图(b)和(c)中灰色曲线表明尿素并不影响 荧光信号强度





图 3-11 将尿素,序列长度为 16 nt 的引物 DNA 和序列长度为 28 nt 的目标 DNA 孵育后, 通过 16%变性聚丙烯酰胺凝胶分析

从上述的结果中可以看到, 嗜热 pAgo 的熔融态的确能够提升剪切活性, 但熔融态如何 影响基因剪辑过程却仍然未知。在本次工作中, 我们通过 Michealis-Menten 动力学模型(米 式方程)定量化蛋白的催化效率:

v=k_{cat} [E]·[S] /(K_M + [S]) (3-1) 其中,v为反应速率,[E]为蛋白浓度,[S]为底物浓度,k_{cat}是催化常数,描述单个酶单位时 间转化底物分子为产物分子的数目。催化常数越大,催化效率越高。K_M表示底物对酶的活

第 34 页 共 44 页



性位点的亲和力。*K*_M越小,亲和力越大。酶的整体催化效率取决于 *k*_{cat} / *K*_M的比值,数值 越大,催化效率越高。我们将 p*Ago* 和引物单链 DNA 复合物与一系列不同浓度的目标单链 DNA 孵育,并且通过加入或不加入一定浓度的尿素从而探究解折叠比例对活性的影响。与 蛋白的熔融态可促进酶活性的结果一致,在不同底物浓度下,具有一定浓度尿素的 *PfAgo* 和 *TtAgo* 比没有尿素的剪切活性更高(如图 3-12 所示)。与嗜热 p*Ago* 在没有尿素的情况下 相比,解折叠比例提高后的嗜热 p*Ago* 具有更低的能耗和更高的催化活性(如表 3-4 中的 *k*_{cat}/*K*_m 比值所示)。值得注意的是,在有尿素存在下 p*Ago* 的 *k*_{cat} 是没有尿素的 p*Ago* 的 *k*_{cat} 的约 2.5 倍,表明催化活性大大提升。

通过阿伦尼乌斯方程可计算得到,在有或没有尿素的情况下,在 95 ℃ 下 PfAgo 的催化 动力学过程的能垒差和在 70 °C 下 TtAgo 的催化动力学过程的能垒差分别为 2.8 kJ/mol 和 2.7 kJ/mol。也就是说,解折叠比例的增加能够降低催化反应的能垒。与之相反,具有尿素的嗜 热pAgo和没有尿素的嗜热pAgo对目标底物显示出相似的亲和力(如图表3-4中的KM所示)。 为了促进引物单链 DNA 和目标单链 DNA 碱基互补配对, 引物 DNA 的种子区域将预先调 整其构象,从而减少熵损失,以较低的能耗和目标单链 DNA 配对。此外,引导单链 DNA 会将其 3'末端从 PAZ 结构域中释放到 N 结构域。然而,具有不同解折叠比例的 p.4go 表现 出相似的亲和力,表明熔融态增加的蛋白柔性并不会影响蛋白质的整体结构排列。kcat的较 大差异表明,尿素增加的蛋白熔融态(即更为动态的构象)可能有助于蛋白局部结构重排。 例如,催化位点的重排以及产物的释放等,从而显着提高了产物的转化比例。虽然,通常 认为催化残基的精确定位是有效催化的先决条件,但通过我们的发现表明,随着生化反应 的进行,增加嗜热 p.Ago 的催化活性中心的柔性能够增加对底物的催化效率。目前,绝大多 数的从头设计的蛋白都是熔融态,却无生化活性,并且蛋白质工程科学家已经做出了巨大 努力将这些分子转化为天然蛋白结构^[86, 87]。传统的假设是由结构决定生物功能。然而,我 们的结果表明, 嗜热 pAgos 中的熔融状态可以增强催化活性, 这不同于蛋白折叠或解折叠 过程中形成的中间状态。在这种情况下,这项工作表明熔融态对于某些嗜热蛋白实现高效 的催化活性非常重要。并且,通过通过对氨基酸序列的某些位点进行突变,增加蛋白整体 结构的柔性,进而使嗜热蛋白能够在低温下能够有效工作。

Protein	<i>T</i> (°C)	$K_{\rm M}(\mu{ m M})$	k_{cat} (min ⁻¹)	$k_{\text{cat}}/K_{\text{M}} (\mu \text{M}^{-1} \cdot \min^{-1})$
PfAgo	95	0.82 ± 0.29	7.56 ± 0.73	2.51
PfAgo + 1.28M Urea		1.18 ± 0.47	18.6 ± 1.35	5.00
TtAgo	70	1.6 ± 0.28	0.74 ± 0.05	0.64
<i>TtAgo</i> + 0.64M Urea		1.53 ± 0.53	1.91 ± 0.25	1.42

表 3-4. 添加尿素前后 PfAgo 和 TtAgo 的酶活性分析





图 3-12 Michealis-Menten 动力学分析。(a) *PfAgo* 在不同底物浓度下的剪切活性与米式方程 拟合。(b) *PfAgo* 与 1.28 M 尿素孵育后在不同底物浓度下的剪切活性与米式方程拟合。(c) *TtAgo* 在不同底物浓度下的剪切活性与米式方程拟合。(b) *TtAgo* 与 0.64 M 尿素孵育后在不 同底物浓度下的剪切活性与米式方程拟合

3.3 本章小节

在本章节中,我们分析比较了 PfAgo, TtAgo 和 CbAgo 的热稳定性及基因剪辑机理。通 过对比蛋白晶体结构的 B-factor 和 Porod-Debye 曲线发现,较之于嗜常温的 CbAgo, 嗜热的 TtAgo 和 PfAgo 更为刚性。通过分子动力学模拟计算表明,刚性的结构是由于在室温下嗜热 pAgo 具有更高的氢键密度、盐桥密度和更高比例的核心区疏水基团。通过圆二色谱和荧光 染色实验分别对二级结构和三级结构表征发现, CbAgo 在其生理温度下的构象与其在室温 相同,而 TtAgo 和 PfAgo 发生了部分解折叠,并且嗜热 pAgo 的熔融态是嗜热 pAgo 的本身 固有性质。之后,通过 SAXS 实验表明,熔融态的确提升了蛋白的柔性,并且表现出更为 伸展的、松散的球状结构。为了进一步验证熔融态对于嗜热 pAgo 极为重要,我们通过生化 反应实验证明提高蛋白解折叠比例后,剪切活性提升,并通过米式方程表明蛋白的熔融态 主要影响催化速率,而并未影响蛋白对底物的亲和力。至此,我们提出一种物理图像:对 于嗜常温的 pAgo,由于其较为柔性的结构,使其能够在较低的温度下调整构象从而实现其 特定的生物功能;而对于嗜热的 pAgo,由于其结构较为刚性而导致在低于其生理温度下非 常难以发生构象的调整。只有在高温下,部分结构被打破后,也就是熔融态,才能使其能 够快速的在各个构象之间发生变化,从而能够对底物进行高效的剪辑,但这种熔融态不能 过分伸展,仍需保持原有球状的构象。该物理图像的提出,为蛋白质工程提出一种新的改

第 36 页 共 44 页



Argonaute 蛋白热稳定性及动力学研究

造方式:可通过将某些特定位点的氨基酸突变或者将柔性链引入蛋白中,使其在低于生理 温度下具有柔性的结构,进而实现低温下的剪切活性。



第四章 结论

在本次工作中,通过利用 nanoDSC,圆二色谱,荧光染色法和分子动力学模拟等方法 原位研究溶液中嗜热 pAgo (PfAgo 和 TtAgo) 和嗜常温 pAgo (CbAgo) 的结构和动力学。并 且通过酶活性实验和小角 X 射线散射研究 pAgo 的 DNA 剪辑功能对温度和蛋白柔性的依赖 性。

对于本次研究的系统,主要有以下发现:

- 1. PfAgo, TtAgo 和 CbAgo 的晶体结构非常相近,并且催化活性中心高度保守。PfAgo 具有最高的熔点和解折叠吉布斯自由能变化,其次为 TtAgo 和 CbAgo。
- 2. 通过晶体结构 B-factor 和分子动力学模拟,发现嗜热 pAgo 较于嗜常温 pAgo 更为刚性,并且嗜热 pAgo 的结构域更为紧实。此外,较之嗜常温的 CbAgo,嗜热的 PfAgo 和 TtAgo 具有更高的氢键密度、盐桥密度和蛋白核心区域疏水氨基酸残基比例,从而表现出较低的柔性和较高的热稳定性。此外,嗜热 pAgo 的活性中心位点的氢键数和盐桥数更高,从而使其催化中心更为稳定。刚性的结构可能降低了嗜热 pAgo 在常温下的活性,这是由于 pAgo 在实现剪切功能时需要较大的构象变化。
- Mn²⁺不仅是 pAgo 实现其生化功能的必要离子,同时能够增加蛋白的解折叠比例, 从而使蛋白更为柔性。
- 我们从实验上首先发现, PfAgo 和 TtAgo 在其生理温度下表现出一种熔融的、松散的、动态的、伸展的但保持球状的构象,这是嗜热 pAgo 的本质。即使在蛋白-引物-目标物三元复合物条件下嗜热 pAgo 仍然为熔融态。这一结果打破了传统"结构决定功能"的观点。
- 5. 通过小角 X 射线散射发现, 嗜热 pAgo 的熔融态, 较之其晶体结构, 具有更大的回转半径, 并且提高了蛋白的整体柔性。
- 6. DNA 剪辑生化实验和小角 X 射线散射证明,将嗜热 pAgo 的解折叠部分提升后, 蛋白的柔性增加,并且熔融态确实能够提升基因剪切效率,包括反应速率和最终剪 切产物含量。通过米式方程分析,嗜热 pAgo 的熔融态能够增加剪切活性并且帮助 产物的释放,但并未影响 pAgo 对底物的亲和力。

通过本研究中的结果可以推测, 嗜热 pAgo 代表了另一种嗜热蛋白, 其主要的生理功能, 也就是基因剪切, 可能通过熔融态实现。此外, 我们的发现证明一种高度动态且非原始的 构象集合可与高效的催化活性有效兼容。因此, 酶改造工程师能够将其对蛋白设计或从头 设计的注意力从结构精确活性点设计转移到创造合适的动力学网络框架的设计。虽然我们 已经发现嗜热 pAgo 在其生理温度时为熔融态并且提供了粗粒化的构象, 但是我们并未给出 具体的原子分辨的构象。在之后的实验中, 可利用核磁共振技术对熔融态结构进行解析, 从而给出原子分辨的蛋白结构。



参考文献

- [1] BOHMERT K, CAMUS I, BELLINI C, et al. AGO1 defines a novel locus of Arabidopsis controlling leaf development [J]. The EMBO journal, 1998, 17(1): 170-80.
- [2] KETTING R F. The Many Faces of RNAi [J]. Developmental Cell, 2011, 20(2): 148-61.
- [3] HAMMOND S M, BERNSTEIN E, BEACH D, et al. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells [J]. Nature, 2000, 404(6775): 293-6.
- [4] SWARTS D C, MAKAROVA K, WANG Y, et al. The evolutionary journey of Argonaute proteins [J]. 2014, 21(9): 743.
- [5] HUTVAGNER G, SIMARD M J. Argonaute proteins: key players in RNA silencing [J]. Nature Reviews Molecular Cell Biology, 2008, 9(1): 22-32.
- [6] HANNON G J. RNA interference [J]. Nature, 2002, 418(6894): 244-51.
- [7] SONG J-J, SMITH S K, HANNON G J, et al. Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity [J]. Science, 2004, 305(5689): 1434.
- [8] YUAN Y-R, PEI Y, MA J-B, et al. Crystal Structure of A. aeolicus Argonaute, a Site-Specific DNA-Guided Endoribonuclease, Provides Insights into RISC-Mediated mRNA Cleavage [J]. Molecular Cell, 2005, 19(3): 405-19.
- [9] SHABALINA S A, KOONIN E V. Origins and evolution of eukaryotic RNA interference [J]. Trends in Ecology & Evolution, 2008, 23(10): 578-87.
- [10] WANG B, LI S, QI H H, et al. Distinct passenger strand and mRNA cleavage activities of human Argonaute proteins [J]. Nat Struct Mol Biol, 2009, 16(12): 1259-66.
- [11] OLOVNIKOV I, CHAN K, SACHIDANANDAM R, et al. Bacterial argonaute samples the transcriptome to identify foreign DNA [J]. Molecular cell, 2013, 51(5): 594-605.
- [12] SWARTS D C, JORE M M, WESTRA E R, et al. DNA-guided DNA interference by a prokaryotic Argonaute [J]. Nature, 2014, 507(7491): 258-61.
- [13] SWARTS D C, HEGGE J W, HINOJO I, et al. Argonaute of the archaeon Pyrococcus furiosus is a DNA-guided nuclease that targets cognate DNA [J]. Nucleic Acids Research, 2015, 43(10): 5120-9.
- [14] KAYA E, DOXZEN K W, KNOLL K R, et al. A bacterial Argonaute with noncanonical guide RNA specificity [J]. Proceedings of the National Academy of Sciences, 2016, 113(15): 4057-62.
- [15] ZANDER A, HOLZMEISTER P, KLOSE D, et al. Single-molecule FRET supports the two-state model of Argonaute action [J]. RNA biology, 2014, 11(1): 45-56.
- [16] SWARTS D C, KOEHORST J J, WESTRA E R, et al. Effects of Argonaute on Gene Expression in Thermus thermophilus [J]. PLOS ONE, 2015, 10(4): e0124880.
- [17] WILLKOMM S, OELLIG C A, ZANDER A, et al. Structural and mechanistic insights into an archaeal DNA-guided Argonaute protein [J]. Nat Microbiol, 2017, 2: 17035.
- [18] ZANDER A, WILLKOMM S, OFER S, et al. Guide-independent DNA cleavage by archaeal Argonaute from Methanocaldococcus jannaschii [J]. Nature Microbiology, 2017, 2(6): 17034.
- [19] SWARTS D C, SZCZEPANIAK M, SHENG G, et al. Autonomous Generation and Loading of DNA Guides by Bacterial Argonaute [J]. Molecular Cell, 2017, 65(6): 985-98.e6.
- [20] MIYOSHI T, ITO K, MURAKAMI R, et al. Structural basis for the recognition of guide RNA



and target DNA heteroduplex by Argonaute [J]. Nat Commun, 2016, 7: 11846.

上海交通大学 GHANGHAI JIAO TONG UNIVERSITY

- [21] WANG Y, JURANEK S, LI H, et al. Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex [J]. nature, 2008, 456(7224): 921-6.
- [22] WANG Y, SHENG G, JURANEK S, et al. Structure of the guide-strand-containing argonaute silencing complex [J]. Nature, 2008, 456(7219): 209-13.
- [23] MA J-B, YUAN Y-R, MEISTER G, et al. Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein [J]. Nature, 2005, 434(7033): 666-70.
- [24] SONG J-J, LIU J, TOLIA N H, et al. The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes [J]. Nature Structural & Molecular Biology, 2003, 10(12): 1026-32.
- [25] LIU J, CARMELL M A, RIVAS F V, et al. Argonaute2 is the catalytic engine of mammalian RNAi [J]. Science, 2004, 305(5689): 1437-41.
- [26] MAKAROVA K S, WOLF Y I, VAN DER OOST J, et al. Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements [J]. Biology direct, 2009, 4(1): 1-15.
- [27] KWAK P B, TOMARI Y. The N domain of Argonaute drives duplex unwinding during RISC assembly [J]. Nature structural & molecular biology, 2012, 19(2): 145.
- [28] HEGGE J W, SWARTS D C, VAN DER OOST J J N R M. Prokaryotic Argonaute proteins: novel genome-editing tools? [J]. 2018, 16(1): 5-11.
- [29] SHENG G, ZHAO H, WANG J, et al. Structure-based cleavage mechanism of Thermus thermophilus Argonaute DNA guide strand-mediated DNA target cleavage [J]. Proc Natl Acad Sci U S A, 2014, 111(2): 652-7.
- [30] PARKER J S, ROE S M, BARFORD D. Structural insights into mRNA recognition from a PIWI domain–siRNA guide complex [J]. Nature, 2005, 434(7033): 663-6.
- [31] BURROUGHS A M, IYER L M, ARAVIND L. Two novel PIWI families: roles in inter-genomic conflicts in bacteria and Mediator-dependent modulation of transcription in eukaryotes [J]. Biology direct, 2013, 8(1): 1-15.
- [32] ENGHIAD B, ZHAO H. Programmable DNA-guided artificial restriction enzymes [J]. ACS synthetic biology, 2017, 6(5): 752-7.
- [33] HSU P D, SCOTT D A, WEINSTEIN J A, et al. DNA targeting specificity of RNA-guided Cas9 nucleases [J]. Nature biotechnology, 2013, 31(9): 827-32.
- [34] ZETSCHE B, GOOTENBERG J S, ABUDAYYEH O O, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system [J]. Cell, 2015, 163(3): 759-71.
- [35] SANDER J D, JOUNG J K. CRISPR-Cas systems for editing, regulating and targeting genomes [J]. Nature biotechnology, 2014, 32(4): 347-55.
- [36] GAO F, SHEN X Z, JIANG F, et al. DNA-guided genome editing using the Natronobacterium gregoryi Argonaute [J]. Nature biotechnology, 2016, 34(7): 768-73.
- [37] CYRANOSKI D. Replications, ridicule and a recluse: the controversy over NgAgo gene-editing intensifies [J]. Nature News, 2016, 536(7615): 136.
- [38] LEE S H, TURCHIANO G, ATA H, et al. Failure to detect DNA-guided genome editing using Natronobacterium gregoryi Argonaute [J]. Nature biotechnology, 2017, 35(1): 17-8.
- [39] JAVIDI-PARSIJANI P, NIU G, DAVIS M, et al. No evidence of genome editing activity from Natronobacterium gregoryi Argonaute (NgAgo) in human cells [J]. PloS one, 2017, 12(5): e0177444.



- [40] QI J, DONG Z, SHI Y, et al. NgAgo-based fabp11a gene knockdown causes eye developmental defects in zebrafish [J]. Cell research, 2016, 26(12): 1349-52.
- [41] YE S, BAE T, KIM K, et al. DNA-dependent RNA cleavage by the Natronobacterium gregoryi Argonaute [J]. BioRxiv, 2017: 101923.
- [42] QI L S, LARSON M H, GILBERT L A, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression [J]. Cell, 2013, 152(5): 1173-83.
- [43] KRATKY O. A survey [M]. Small angle X-ray scattering. 1982.
- [44] BURGER V M, ARENAS D J, STULTZ C M. A structure-free method for quantifying conformational flexibility in proteins [J]. Scientific reports, 2016, 6(1): 1-9.
- [45] PEDERSEN J S. Analysis of small-angle scattering data from colloids and polymer solutions: modeling and least-squares fitting [J]. Advances in colloid and interface science, 1997, 70: 171-210.
- [46] PEDERSEN J S. Form factors of block copolymer micelles with spherical, ellipsoidal and cylindrical cores [J]. Journal of applied crystallography, 2000, 33(3): 637-40.
- [47] PEDERSEN J S. Determination of size distribution from small-angle scattering data for systems with effective hard-sphere interactions [J]. Journal of applied crystallography, 1994, 27(4): 595-608.
- [48] LADOKHIN A S, FERNÁNDEZ-VIDAL M, WHITE S H. CD spectroscopy of peptides and proteins bound to large unilamellar vesicles [J]. The Journal of membrane biology, 2010, 236(3): 247-53.
- [49] GREENFIELD N J. Using circular dichroism spectra to estimate protein secondary structure[J]. Nature protocols, 2006, 1(6): 2876.
- [50] CHOMA C T. Characterizing Protein Stability by DSC [J]. Calorimetry Sciences Corporation, 2006.
- [51] FREIRE E. Differential scanning calorimetry [J]. Protein stability and folding, 1995: 191-218.
- [52] GARBETT N, DELEEUW L, CHAIRES J B. High-throughput DSC: A Comparison of the TA Instruments Nano DSC Autosampler System[™] with the GE Healthcare VP-Capillary DSC[™] [J].
- [53] WU L, QU X. Cancer biomarker detection: recent achievements and challenges [J]. Chemical Society Reviews, 2015, 44(10): 2963-97.
- [54] CHOI J R, HU J, TANG R, et al. An integrated paper-based sample-to-answer biosensor for nucleic acid testing at the point of care [J]. Lab on a Chip, 2016, 16(3): 611-21.
- [55] MUMFORD R A, MACARTHUR R, BOONHAM N. The role and challenges of new diagnostic technology in plant biosecurity [J]. Food Security, 2016, 8(1): 103-9.
- [56] GOOTENBERG J S, ABUDAYYEH O O, LEE J W, et al. Nucleic acid detection with CRISPR-Cas13a/C2c2 [J]. 2017, 356(6336): 438-42.
- [57] KOMOR A C, BADRAN A H, LIU D R J C. CRISPR-based technologies for the manipulation of eukaryotic genomes [J]. 2017, 168(1-2): 20-36.
- [58] BARRANGOU R, DOUDNA J A J N B. Applications of CRISPR technologies in research and beyond [J]. 2016, 34(9): 933-41.
- [59] WU W Y, LEBBINK J H, KANAAR R, et al. Genome editing by natural and engineered CRISPR-associated nucleases [J]. 2018, 14(7): 642-51.
- [60] LI Y, LI S, WANG J, et al. CRISPR/Cas Systems towards Next-Generation Biosensing [J]. Trends Biotechnol, 2019, 37(7): 730-43.



- [61] HEGGE J W, SWARTS D C, CHANDRADOSS S D, et al. DNA-guided DNA cleavage at moderate temperatures by Clostridium butyricum Argonaute [J]. Nucleic Acids Research, 2019, 47(11): 5809-21.
- [62] SONG J, HEGGE J W, MAUK M G, et al. Highly specific enrichment of rare nucleic acid fractions using Thermus thermophilus argonaute with applications in cancer diagnostics [J]. Nucleic Acids Res, 2020, 48(4): e19.
- [63] HE R, WANG L, WANG F, et al. Pyrococcus furiosus Argonaute-mediated nucleic acid detection [J]. 2019, 55(88): 13219-22.
- [64] WANG J, LI J, ZHAO H, et al. Structural and mechanistic basis of PAM-dependent spacer acquisition in CRISPR-Cas systems [J]. 2015, 163(4): 840-53.
- [65] SWARTS D C, MAKAROVA K, WANG Y, et al. The evolutionary journey of Argonaute proteins [J]. Nat Struct Mol Biol, 2014, 21(9): 743-53.
- [66] RIVAS F V, TOLIA N H, SONG J-J, et al. Purified Argonaute2 and an siRNA form recombinant human RISC [J]. 2005, 12(4): 340-9.
- [67] KUZMENKO A, YUDIN D, RYAZANSKY S, et al. Programmable DNA cleavage by Ago nucleases from mesophilic bacteria Clostridium butyricum and Limnothrix rosea [J]. 2019, 47(11): 5822-36.
- [68] KUZMENKO A, OGUIENKO A, ESYUNINA D, et al. DNA targeting and interference by a bacterial Argonaute nuclease [J]. Nature, 2020, 587(7835): 632-7.
- [69] HEGGE J W, SWARTS D C, VAN DER OOST J. Prokaryotic Argonaute proteins: novel genome-editing tools? [J]. Nat Rev Microbiol, 2018, 16(1): 5-11.
- [70] ENGHIAD B, ZHAO H J A S B. Programmable DNA-guided artificial restriction enzymes [J]. 2017, 6(5): 752-7.
- [71] WEE LIANG M, FLORES-JASSO C F, SALOMON WILLIAM E, et al. Argonaute Divides Its RNA Guide into Domains with Distinct Functions and RNA-Binding Properties [J]. Cell, 2012, 151(5): 1055-67.
- [72] PANDE V S, ROKHSAR D S. Is the molten globule a third phase of proteins? [J]. Proceedings of the National Academy of Sciences, 1998, 95(4): 1490.
- [73] GURSKY O, ATKINSON D. Thermal unfolding of human high-density apolipoprotein A-1: implications for a lipid-free molten globular state [J]. Proceedings of the National Academy of Sciences, 1996, 93(7): 2991.
- [74] JAMES L C, TAWFIK D S. Conformational diversity and protein evolution a 60-year-old hypothesis revisited [J]. Trends in Biochemical Sciences, 2003, 28(7): 361-8.
- [75] VAMVACA K, VÖGELI B, KAST P, et al. An enzymatic molten globule: Efficient coupling of folding and catalysis [J]. Proceedings of the National Academy of Sciences of the United States of America, 2004, 101(35): 12860.
- [76] DYSON H J, WRIGHT P E. Coupling of folding and binding for unstructured proteins [J]. Current Opinion in Structural Biology, 2002, 12(1): 54-60.
- [77] ARRIGONI C, ROHAIM A, SHAYA D, et al. Unfolding of a Temperature-Sensitive Domain Controls Voltage-Gated Channel Activation [J]. Cell, 2016, 164(5): 922-36.
- [78] PERVUSHIN K, VAMVACA K, VÖGELI B, et al. Structure and dynamics of a molten globular enzyme [J]. Nature Structural & Molecular Biology, 2007, 14(12): 1202-6.
- [79] BOSE H S, WHITTAL R M, BALDWIN M A, et al. The active form of the steroidogenic acute regulatory protein, StAR, appears to be a molten globule [J]. Proceedings of the



National Academy of Sciences, 1999, 96(13): 7250.

- [80] BALDWIN R L, ROSE G D. Molten globules, entropy-driven conformational change and protein folding [J]. Current Opinion in Structural Biology, 2013, 23(1): 4-10.
- [81] HONAKER M T, ACCHIONE M, ZHANG W, et al. Enzymatic Detoxication, Conformational Selection, and the Role of Molten Globule Active Sites* [J]. Journal of Biological Chemistry, 2013, 288(25): 18599-611.
- [82] DEGRADO W F. Catalytic molten globules [J]. Nature, 1993, 365(6446): 488-9.
- [83] KIKHNEY A G, SVERGUN D I. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins [J]. FEBS Lett, 2015, 589(19 Pt A): 2570-7.
- [84] BERNADO P, SVERGUN D I. Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering [J]. Mol Biosyst, 2012, 8(1): 151-67.
- [85] SCHROER M A, PAULUS M, JEWORREK C, et al. High-pressure SAXS study of folded and unfolded ensembles of proteins [J]. Biophys J, 2010, 99(10): 3430-7.
- [86] BETZ S F, RALEIGH D P, DEGRADO W F. De novo protein design: from molten globules to native-like states: Current opinion in structural biology 1993, 3:601–610 [J]. Current Opinion in Structural Biology, 1993, 3(4): 601-10.
- [87] JÄCKEL C, KAST P, HILVERT D. Protein Design by Directed Evolution [J]. Annual Review of Biophysics, 2008, 37(1): 153-73.



谢辞

四年前,我来到上海交通大学物理与天文学院。时光荏苒,转眼间我已成为一名即将 步入博士阶段的本科毕业生。借本次毕业论文,我将诚挚地感谢在毕业设计工作中帮助我的 老师和同学。

首先,我要感谢我的毕业设计导师——洪亮老师。他有着海纳百川、有容乃大的精神。 是他教我热力学统计物理,引发我对热力学的兴趣,为我今后的研究打下一定基础;是他给 我一个良好的实验环境,让我能够尽情使用实验室的资源,使用经费购买所需药品试剂;是 他愿意付出心血指点我实验、倾听我并不完美的汇报;是他在我备考期间,愿意在百忙之中 抽出时间为我分析形势。感谢短短一年来在洪老师组里的学习,我认识到了物理思维的重要 性,了解到作为一名研究生应该怎样去发展课题,也学习到了合作的技巧与重要性。今后, 尽管我不再属于他的组,但是我仍憧憬在 Ago 蛋白及其相关课题上能够和洪老师组开展更多 合作,我将带着学到的该课题相关知识与理解,加入到其中,作出更大贡献。

其次,我要感谢实验室的黄娟师姐。这个课题由蛋白样品制备开始,如果没有黄娟师 姐倾囊相授纯化技术,课题将无从开始,我将无法毕业。另外我还要感谢郑力荣师兄,是他 在我实验出错时耐心指导;是他在我思路不清时拨云见日;是他将自己五年来的科研经验毫 无保留地传授给我;是他在我懈怠时以亲身经历激励鼓舞;是他在繁忙的工作中抽时间帮我 修改开题报告和毕业论文以及答辩 PPT;是他愿意花时间听我预演开题答辩和毕业答辩。还 要感谢我的陆慧师兄,是他愿意在生科院繁重的任务下拨冗来实验室指导我的实验操作。

接着,我要感谢上海交通大学分析测试中心的刘永佳老师。是她帮助我完成 nano-DSC 实验。更要感谢上海交通大学,能给我就读于此、毕业于此的机会。

最后,还要感谢我的父母,没有他们的关爱我将无法无忧无虑地学习。

四年的本科生涯即将结束,我将要进入博士生涯。再次感谢在毕业设计中帮助我的老师、同学和父母。我将在博士阶段继续努力,争取早日在生物物理领域做出贡献,回报你们。



STUDY OF ARGONAUTE PROTEIN THERMOSTABILITY AND DYNAMICS

Structural studies were initiated to gain insight into the properties of PfAgo, TtAgo and *CbAgo.* The overall two-lobed molecular architecture is typical for the family of long pAgos. pAgos share a common bilobal molecular architecture, with the C-terminal lobe consisting of the MID (colored in violet) and PIWI (P element wimpy testis, colored in blue) domains and the N-terminal lobe containing the N-terminal (colored in green) and PAZ (colored in red) domains. The superposition of the folding topology of those pAgos obtained from the crystal structures indicates that the pAgos with distinct T_{phy} share the similar quaternary and tertiary structure but vary in alpha-helix and beta-sheet in terms of secondary structure. However, the thermal melting curves obtained from nanoDSC show the melting temperature (T_m) of PfAgo, TtAgo and CbAgo are 46, 75 and 100 °C, respectively, and higher unfolding free energy change in thermophilic pAgos, manifest that thermophilic pAgos are more stable than their thermophilic counterpart. To compare the overall dynamics, the B-factors derived from crystal structures of PfAgo, TtAgo and CbAgo show that thermophilic pAgos are globally stiffer than CbAgo. Furthermore, Porod-Debye plots of pAgos obtained from small angle X-ray (SAXS) reflect that CbAgo is more flexible in solution than *PfAgo* and *TtAgo*, which is consistent with the results of crystallography B-factors. By performing molecular dynamics (MD) simulations, it revealed thermophilic pAgos present higher density of hydrogen bond and salt bridge and stronger nonpolar van der Waals interactions due to the higher ratio of internal hydrophobic residues. Based on such comparisons, these stronger attractive interactions between residues and rigidity of molecular chains in thermophilic pAgos contribute to increase in enthalpy change and decrease in entropy change, respectively, thus resulting higher free energy change, i.e., higher thermostability in thermophilic pAgos. PIWI domain of PfAgo, TtAgo and CbAgo adopt an RNase H-like fold with the catalytic tetrad Asp-Glu-Asp-His, Asp-Glu-Asp-Asp and Asp-Glu-Asp-Asp, respectively, which are highly conserved but exhibit different cleavage activity at the same temperature. We estimated the hydrogen bond and salt bridge around the catalytic sites and found that the numbers of hydrogen bond and salt bridge around the catalytic sites of PfAgo and TtAgo are higher than them of CbAgo. Therefore, it can be demonstrated that the hydrogen bond and salt bridge enhance the stability of overall structure and catalytic sites in thermophilic pAgos and, thus enable them to work at high temperature but may decrease the enzyme activity at moderate temperature corresponding the resulting rigid structure.

Although the canonical target cleavage mechanism has been revealed by the crystal structure of pAgo ternary complex with guide ssDNA and target ssDNA, the in-situ monitoring the evolution of pAgo structure at T_{phy} has not been reported. Here, far-UV CD spectra and fluorescence dye method are applied to record the secondary and tertiary structure of ligand-free pAgos and ternary complexes (protein-guide-target) at different temperature, respectively. Unfolding profiles for the secondary and tertiary structure in each protein was measured by the ellipticity at a wavelength of 222 nm and the fluorescence intensity at a wavelength of 350 nm,



respectively. Strikingly, the onset temperature of unfolding (T_{on}) of tertiary structure in *PfAgo* and *TtAgo* is obviously below T_{phy} . In sharp contrast, T_{on} of *CbAgo* is higher than its T_{phy} and is similar to the typical globular proteins. Moreover, melting of the secondary structure in three p*Agos* presents the similar results. Therefore, it can be demonstrated that secondary structure is partially unfolded and tertiary structure is disrupted in thermophilic p*Agos* at T_{phy} , while their mesophilic counterpart *CbAgo* functions at a folded state. To further identify the state of p*Agos* under genome cleavage process *in vitro*, we characterized the p*Agos* in complex with ssDNA guide and complementary ssDNA target at moderate temperature and T_{phy} . At T_{phy} , the secondary structure of *CbAgo* bound with guide DNA and target DNA is intact as inferred from the one at non-activity temperature. In contrast, the secondary structure of *PfAgo* and *TtAgo* appear largely melted at their respective T_{phy} . Undoubtedly, the molten state, i.e., retained part of native structure, in *PfAgo* and *TtAgo* during cleaving is intrinsic nature of thermophilic p*Agos* themselves.

To gain further insights into the evolution of microscopic structure as a function of temperature, we conducted in-situ SAXS on *PfAgo*, *TtAgo* and *CbAgo*. Kratky plots ($I(q) \cdot q^2$ as a function of q) are used to qualitatively identify disordered states and distinguish them from globular particles. Globular proteins confer a bell-shaped Kratky plot with a well-defined maximum. Conversely, an ideal Gaussian chain presents a plateau at large q values which normally observed experimentally in unfolded proteins. The Kratky profile of CbAgo at its T_{phy} is the same as that at moderate temperature, which indicates that the mesophilic pAgo present a globular and folded structure. As for *PfAgo* and *TtAgo*, the Kratky profiles of them show an explicit decrease in the structural features compared with the profiles at moderate temperature. These results further manifest that thermophilic pAgos are partially unfolded and which are analogous to the results from CD spectra and fluorescence dye measurement. The corresponding Porod-Debye plots reflect the flexibility of proteins. As expected, the molten state in *PfAgo* and TtAgo increases the globular flexibility compared with them at moderate temperature and thus may enhances the rapid conformation change among various functional states. The resulting reconstructions of the ab initio calculations obtained from SAXS measurements present the coarse-grained models of mesophilic and thermophilic pAgos. Although PfAgo and TtAgo are melted, they still present a globular conformation. Whereas radius of gyration (R_g) of CbAgo at its T_{phy} is similar to that of crystal structure, Rg of PfAgo and TtAgo are larger than their crystal structure. These results reflect that the conformations of thermophilic pAgos are more extended and loosely packed but globular at their respective T_{phy} .

Inspired by the molten state in the thermophilic pAgos at T_{phy} , we therefore investigated the cleavage of target ssDNA by *PfAgo* and *TtAgo* that increasing the unfolding ratio by unfolding-reagent urea. We designed a nuclei acid cleavage activity detection system in which the intensity of fluorescence signal is directly proportional to the activity of target ssDNA cleavage. The target ssDNA is designed with a quencher (BHQ1) at 3'-end and a fluorescence group (6-FAM) at 5'-end. The fluorescence signal can be detected when the quencher was separated from the fluorescence group after cleaving the target ssDNA by pAgo. Thus, the activity of cleaving target ssDNA can be monitored by recording the fluorescence intensity as a function of time. The details of cleavage efficiency determined by fluorescence intensity are described in Methods. We incubated the *PfAgo* and *TtAgo* with a certain concentration of urea and discovered that the unfolding ratio in those two pAgos are increased. Interestingly, the cleavage activity of target ssDNA is promoted by urea-enhanced molten state in *PfAgo* and *TtAgo* compared with the

Argonaute 蛋白热稳定性及动力学研究



urea-free systems at the same temperature. Moreover, we found that both of cleaving rate and yield of cleaving product are increased according to the fluorescence intensity. In order to exclude the effect of urea on the degradation of target ssDNA, electrophoretic mobility shift assay (EMSA) was applied to analyze the production of incubation of urea, guide ssDNA and target ssDNA and found that the ssDNA are not degraded. Thus, we can demonstrate that the molten state in the thermophilic p*Agos* is significant on high activity in genome-cleavage.

The intriguing question arises as how the molten state in thermophilic pAgos affect genome-cleavage process. Here, we quantified catalytic efficiency by Michealis-Menten kinetics model. The catalytic efficiency of the enzyme is determined by the ratio of k_{cat}/K_{M} . We incubated the pAgo-guide complex with a series concentration of target and with or without urea. Consistent with the observation that molten state can promote the enzyme activity, PfAgo and TtAgo with a certain concentration of urea is more efficient than them without urea at different concentrations of target. The enhanced unfolding ratio of thermophilic pAgos with urea compared with thermophilic urea has an amazingly small energetic cost and higher catalytic efficiency. Notably, k_{cat} of pAgos with urea for target is approximately 2.5 times higher than that of pAgos. By applying Arrhenius equation, the difference of energy barrier for this kinetic process with or without urea is ~ 2.8 kJ/mol for PfAgo at 95 °C and ~ 2.7 kJ/mol for TtAgo at 70 °C. In contrast, pAgos and pAgos with urea show similar affinities of target. To facilitate base pairing with a complementary target, the seed region of the guide will prearrange to decrease entropy penalty and the guide will release its 3'-end from PAZ domain to N domain. However, the similar affinity between pAgos with different unfolding ratio indicates that the increased flexibility does not affect the global ordering of protein. The large differences k_{cat} suggest that urea-enhanced molten state, i.e., more dynamical conformation, may makes contributions to local restructuring, such as rearrangement of catalytic sites and product release, and thus significantly enhances the yield of production. Although precise positioning of catalytic residues is generally considered to be a prerequisite for efficient catalysis, our findings suggest that there is little penalty associated with locking the catalytically active conformation of thermophilic pAgos in place as the reaction proceeds. Many de novo designed proteins have turned out to be molten globules and protein engineers have made considerable effort to convert these molecules into native-like proteins. The conventional assumption has been that the biological function is determined by structure. However, our results show that the molten state in thermophilic pAgos can enhance the catalytic activity, which different from the intermediate state formed during the folding or unfolding of enzyme. In this case, there have been hints in this work that molten state is necessary for some thermophilic proteins to achieve high enzyme activity.