

## 上海光源生物大分子晶体学用户在蛋白质特异性识别 DNA 的机理研究方面取得重要进展

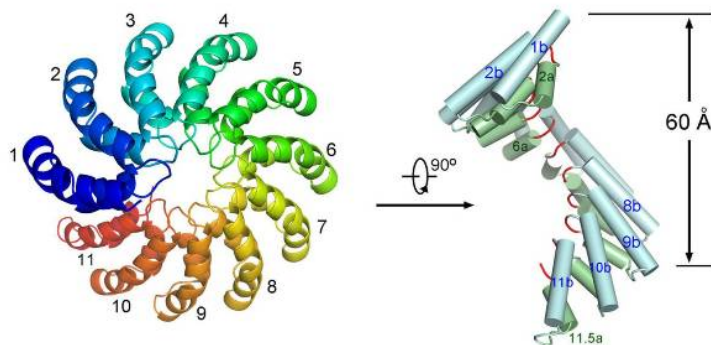
2012 年 1 月 5 日, 清华大学医学院颜宁教授研究组、生命学院施一公教授研究组以及美国普渡大学朱健康教授合作在《科学》在线发表论文, 报道转录激活因子样效应蛋白 (TALE) 特异识别 DNA 的分子机理。

TALE (Transcription Activator Like Effectors) 植物致病菌 *Xanthomonas* 通过 III 型分泌系统注入到宿主细胞内的一种蛋白质。TALE 蛋白的奇特之处在于它的 DNA 结合结构域——该 DNA 结合结构域不同于其他已知的 DNA 结合结构域, 它是由不同数量的重复单元组成的, 每一个重复单元特异识别一个 DNA 碱基对。大多数情况下每个重复单元由 34 个氨基酸组成, 这 34 个氨基酸中除了第 12、13 位的氨基酸变化较大之外, 其他氨基酸高度保守。这两个不保守的氨基酸被命名为 RVD (repeat variable diresidue)。每个重复序列中 12、13 位的氨基酸和被识别的核苷酸种类有特殊的一一对应关系。TALE 蛋白的特异 DNA 序列识别以及灵活的可组装性为它们在分子生物学中的应用提供了巨大的前景, 科学家们可以设计组装任意的 TALE 单元去识别目标 DNA 双螺旋序列。这一特性已经被用来构造切割特异双链 DNA 序列的 DNA 酶 TALEN (TALE nuclease), 成功用于在细胞基因组中引入定点突变、定点敲除等操作。理解 TALE 识别 DNA 的分子机制, 会极大地促进其在生命科学领域的应用。

TALE 蛋白到底是如何实现这种特殊的 DNA 识别方式呢? 为了回答这个有趣的问题, 颜宁、施一公、朱健康研究组合作, 选择了一

个经过改造的 TALE 蛋白 dHax3, 进行结构生物学和生物化学研究, 并利用上海光源 BL17U 线站成功解析了未结合 DNA 的 TALE 蛋白的晶体结构。结合进一步获得的 TALE 蛋白——DNA 复合物晶体结构, 清晰揭示了 TALE 蛋白特异识别 DNA 的机理。TALE 蛋白的重复单元组成 Helix-loop-helix 的结构围绕 DNA 呈右手螺旋状排列, RVD 这两个残基中只有第二位的氨基酸才与碱基特异识别。结构比较进一步展示了 TALE 蛋白类似于弹簧的伸展性。这些结构信息提供了 TALE 蛋白的改造基础, 极大地拓宽了 TALE 蛋白在生物技术应用上的前景。该论文 (Structural Basis for Sequence-Specific Recognition of DNA by TAL Effectors) 以 Highlight 的形式在 Science Express 发表。

能够及时利用上海光源收集到高质量的蛋白质晶体衍射数据, 为这项研究提供了巨大的便利, 保证了我国研究组在这一课题的激烈国际竞争中最终胜出。



TALE 蛋白质三维结构, 11 个 TAL 重复形成一个右手超螺旋图