

孙晋武

(中国科学院微生物研究所, 北京)

一、概述

所谓“蛋白质工程”,就是人们通过对蛋白质结构和功能间规律的详细了解,按照人们预定的模式,人为地改变蛋白质结构,从而创造出有新奇性质蛋白质的过程。

“蛋白质工程”是近几年出现的一门新兴学科,它涉及各方面的知识和技能,包括对蛋白质一级结构和空间结构的全面了解;它们(如酶)的作用机制,即与作用物(如底物)的相互关系;X-光结晶学技术及计算机模拟蛋白质模型的方法;DNA 序列分析、人工合成寡

核苷酸以及基因工程方面的技能等等。采用基因工程的办法,原则上可以得到任何已知的蛋白质,而蛋白质工程技术则可按照人们的意愿改造它们,并创造出自然界未发现或没有的蛋白质,因此该学科的兴起,其意义是异常重大的。

所谓改造蛋白质,主要是对其中某些关键性的氨基酸作人为的更改,从而达到改变蛋白质性质的目的。这些性质包括增加酶的活性、提高蛋白质(或酶)的适应能力,改变酶的底物和反应专一性等。Wasserman 提出一套提高酶热稳定性的基因工程方案^[1],便能较好地帮助我们了解蛋白质工程的基本过程:

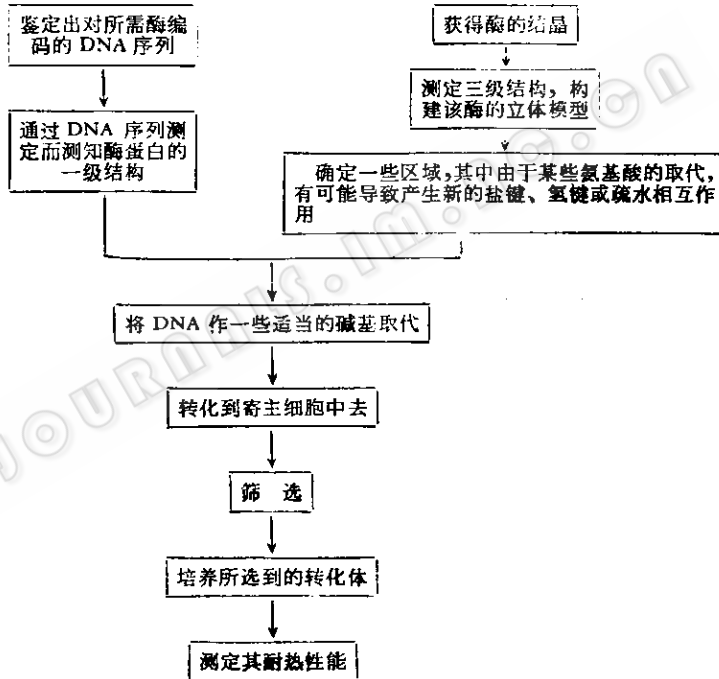


图1 提高酶热稳定性的基因工程方案^[1]

图中的“将 DNA 作一些适当的碱基取代”,即所谓的“位点定向诱变”(Site-Directed Mutagenesis),也有称之为“寡核苷酸定向诱变”的(Oligonucleotide-Directed Mutagenesis),这一方法的建立,为蛋白质工程提供了成功的可能。1982年,英国的 Winter 和 Fersht 等人采用这一手段,将嗜热脂肪芽孢杆菌(*Bacillus stearothermophilus*)中酪氨酸 tRNA 合成酶中与 ATP 结合部位的 Cys 35 变更为 Ser,并在大肠杆菌中获得引起突变的酶^[2],这一工作的成功使蛋白质工程迈出了重大的一步^[3]。与此同时或稍后,许多国家

的科学家用蛋白质工程创造出一系列新型蛋白质,并进行了大量理论与实践的探讨,使这一新兴学科呈现出巨大的生命力。1985年,英国皇家学会召开了专题讨论会,共有 14 篇论文谈到了在过去 3、4 年所发展起来的这种新方法在许多领域所取得的成绩,并预示了它的巨大应用前景^[4]。

正因为如此,近年来一些国家十分关心蛋白质工

在写作过程中,张树政教授和姜书勤副教授对本文提出了宝贵意见并作了热情的指导,特此致谢。

程方面的工作。美国 Genex 公司于 1984 年花了 150 万美元建立起蛋白质工程实验室，成为这一领域中第一家实力雄厚的生物技术公司，并可获得从 Allied 公司那里得到为期五年的 1650 万美元的保证金^[1]。英国科学和工程研究委员会 (SERC) 最近宣布成立以 R. Freedman 为首的蛋白质工程研究委员会，负责发展和协调在英国现有的五、六个与这方面工作有关的研究中心，在今后十年中，将有 200 万英镑的资金用于这方面的课题^[2]。日本通产省将从 1986 年 4 月起执行一项六年计划，投资 300 亿日元致力于蛋白质工程，内容包括蛋白质的设计、酶和抗体的合成等等^[3]。

二、位点定向诱变

位点定向诱变是一种以人工合成的寡核苷酸在一已知 DNA 序列中某特定点上实现变更而在体外引起诱变的方法。Hutchison 等人于 1978 年报导了这一方法，他们当时是在单链噬菌体 ϕ X174 中进行的^[4]。后来有人根据其原理，对克隆到 pBR 322 中的基因进行了诱变^[5]。近年来随着单链噬菌体 M13 运载体系统的发展，由于它具有许多可被克隆的位点；转化产生的噬菌斑根据有无插入物而呈现不同的颜色而易于筛

选；这类运载体在寄主细胞中拷贝数多(约有 200 个)，而表达效率高；另外在这一系统可插入的外源 DNA 长度变化范围较大以及能快速简便地提得单链模板 DNA 等一系列优点，因而已被广泛地应用于位点定向诱变工作。下面简略介绍一下 Zoller 和 Smith 所述的方法(图 2)^[10]。

首先将计划诱变的外源 DNA 片段用常规方法克隆到 M13 运载体系统中去(具体方法可参见文献 [11] 和 [12])，根据转化细胞在含有 IPTG 和 Xgal 的平皿上所呈现的颜色，挑出含有插入了外源 DNA 的白色噬菌斑，由此提取单链 DNA 作为下一步反应的模板，一般有 1—2 μ g 这样的 DNA 就够做一次实验了。

人工合成一段寡核苷酸作为导致突变的诱变剂，其长度一般为 14—21 个核苷酸。该寡核苷酸与上述作为模板的被克隆的外源 DNA 上某一部分序列除对其中需要变异的核苷酸作了相应的变更外，其余部分完全互补。为防止 DNA 多聚酶中外切核酸酶活性的作用，这种变更一般置于此寡核苷酸的中部(实际上只需在变更处之后有三个核苷酸起保护作用就够了)。通常一次实验只需 250 p mol 这样的寡核苷酸。

以此寡核苷酸为引物，与上述单链模板在一系列预备实验所确定的条件下(如 55 $^{\circ}$ C 处理 5 分钟，再冷至室温)进行退火，由此形成的一段双链 DNA，除了在变更处有一错配 (mismatch) 外，其它部位都能完全互补，然后在 dNTP 存在下，用 *E. coli* DNA 多聚酶 I Klenow 大片段将引物进行延伸，与此同时进行同位素标记，最后经 T₄ 连接酶将延伸物两头连接成闭环状双链 DNA (cc-DNA)。为了提高突变频率，必须将没有连接和延伸不完全的副产物除去而将 cc-DNA 富集纯化出来，可采用 Sephadex 凝胶过滤和碱性蔗糖密度梯度的方法，另外也可用 S₁ 核酸酶处理、将单链 DNA 结合到硝酸纤维素薄膜上、琼脂糖凝胶电泳、氯化铯密度梯度离心等方法来达到这一目的。经富集后的 cc-DNA 便可用于随后的对感受态 *E. coli* 细胞的转化。由于这样的 cc-DNA 是由野生型和突变型分子共同构成的杂合双链，由此转化得到的噬菌斑也有野生型和突变型之分，故必须进行筛选以得到经突变的重组体。筛选的方法很多：限制性内切酶位点筛选法是根据所要的突变体新生或破坏了某个内切酶切点，而在同位素标记后，经该内切酶切割呈现出与野生型不同的内切酶图谱而加以区分。另一种杂交筛选法是最简单有效的方法。它是将上述的人工合成的寡核苷酸进行同位素标记，用作为筛选突变体的探针。其原理是该寡核苷酸能与突变克隆 DNA 形成完美的配对，而比它与野生型 DNA 有错配存在显得更稳定一些，因此只需将突变型与野生型噬菌斑在低温下与探针杂交，然后逐渐提高温度进行洗涤，直至只有突变型噬菌斑可以杂交为止。纯化所得到的突变型噬菌体，

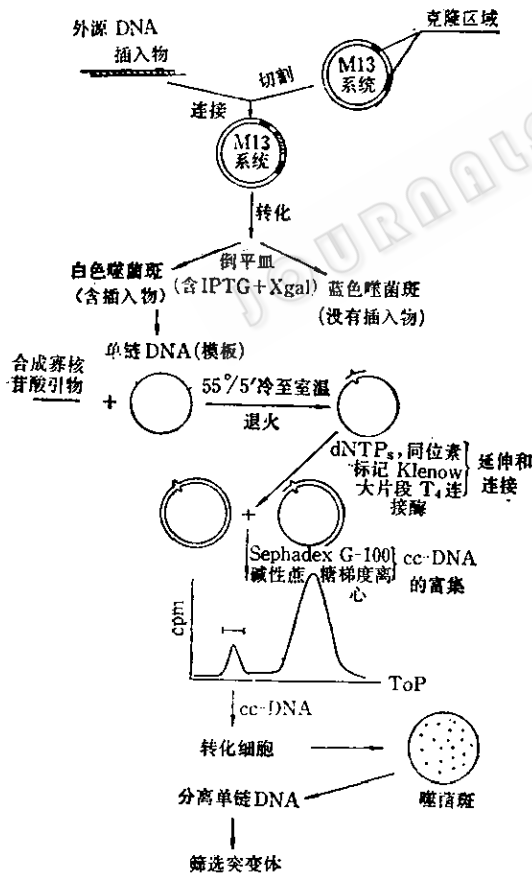


图 2 位点定向诱变示意图

提纯其 DNA, 为了证实所需位点确实实现了突变, 一般还需对此 DNA 进行序列分析。

三、几个实例

1. 在 *T.* 溶菌酶中引入二硫键以提高热稳定性; Perry 和 Weitzel^[13] 以 *T.* 溶菌酶作为研究材料。这是一种胞壁质酶, 它的生物学功能是使感染细胞的细胞壁破裂以利于裂解释放出复制的噬菌体。该蛋白的 X-光晶体结构已经清楚。它具有二个不配对的半胱氨酸 (Cys 54 和 Cys 97), 热稳定性较差。通过理论计算认为: 蛋白质构象稳定性程度有赖于其一级结构中远离氨基酸之间交联数目的增加。他们从 *T.* 噬菌体的晶体结构中寻找在一级结构中相距甚远、但在空间位置上靠得很近 (在 5.5 Å 范围内) 的具有侧链的 β 碳原子的氨基酸, 根据计算机模拟分析来构建一些二硫键模型, 最终选定以半胱氨酸来取代野生型 *T.* 溶菌酶分子中 3 号位的异亮氨酸 (Ile 3 \rightarrow Cys), 让取代后的半胱氨酸与原分子中 97 位半胱氨酸构成一对二硫键, 从而产生一 *T.* 溶菌酶突变型。这种由 Ile 3 \rightarrow Cys 的变更, 便是借助于位点定向诱变的方法, 以人工合成的寡核苷酸 pAGAATTATGAATTGTTTGAATGTTA 为引物, 借助于 M13 mp10 而完成的。突变基因在 *lacII* 启动子的作用下, 在 *E. coli* 中得以表达, 从而获得了 (Ile 3 \rightarrow Cys) *T.* 溶菌酶, 该酶已被提纯。

经活性测定, 突变酶与野生型酶在活力上没有差别, 但在热稳定性上却有明显的改进。由图 3 可见这两种酶在 67°C 下保温, 野生型酶的初半衰期为 11 分钟, 而 (Ile 3 \rightarrow Cys) *T.* 溶菌酶, 初半衰期为 28 分钟。尤为突出的是延长保温时间, 突变酶的活力不会降到起始活力的 50% 以下, 而野生型酶在 180 分钟后活力仅剩 0.2% (图 3)。

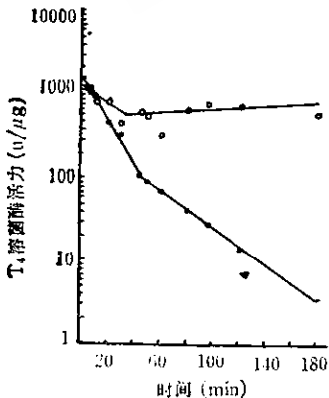


图 3 野生型和突变型 *T.* 溶菌酶的热失活曲线^[13]。

●野生型 *T.* 溶菌酶。○突变型 (Ile 3 \rightarrow Cys) *T.* 溶菌酶

2. 在酵母中合成具抗氧化功能的人 α_1 -抗胰蛋白酶: 白细胞弹性蛋白酶会造成肺弹性的丧失而引起肺

气肿, 另外它也是成人呼吸窘迫综合症的成因之一, α_1 -抗胰蛋白酶能抑制弹性蛋白酶而减少其危害。它是一种由 394 个氨基酸构成的糖基化了的血清蛋白质。而位于 α_1 -抗胰蛋白酶活性中心的第 358 位残基甲硫氨酸被氧化, 就会明显降低它对弹性蛋白酶的抑制作用。在体内由于巨噬细胞释放出氧自由基而造成 α_1 -抗胰蛋白酶的失活。因此在上述疾病中, α_1 -抗胰蛋白酶的氧化失活, 可能在由弹性蛋白酶造成肺损伤方面起着重要的作用。治疗肺气肿需要大剂量的 α_1 -抗胰蛋白酶, 如有一种抗氧化的 α_1 -抗胰蛋白酶, 就能减少应用剂量。美国 Rosenberg 等人^[14] 采用 M13 位点定向诱变方法, 以缬氨酸来单一取代处于酶活性部位的第 358 位甲硫氨酸, 利用酵母酸性磷酸酶 (PH05) 的启动子和 pC1/1 酵母穿梭质粒, 最后转化酵母株 AB103, 得到有活性的、非糖基化的、具有抗氧化功能的人 α_1 -抗胰蛋白酶。

在具体工作中, 他们是将编码 238—394 氨基酸的 cDNA 片段克隆到 M13 mp10 运载体中, 分离出单链噬菌体 DNA 作为位点定向诱变的模板, 以合成的 'GACCTCGGGGGGATAGACACGGGTATGGC' 作为诱变的引物。在该寡核苷酸中下面划有短线的部位以 C 代替了野生型的 T, 从而使新合成的互补 DNA 能为缬氨酸 (GTG) 而非甲硫氨酸 (ATG) 编码。以此引物为探针, 就可从转化体中鉴别出正的噬菌斑, 用 DNA 序列分析法证实了密码子的更改。

新构成的 α_1 -抗胰蛋白酶 (Val) 在抑制弹性蛋白酶活力方面比野生型 α_1 -抗胰蛋白酶 (Met) 具明显的抗氧化功能。N-氯琥珀酰亚胺 (N-Chlorosuccinimide, 简称 NCS) 是一种氧化剂, 能将易感的甲硫氨酸转变成相应的亚砷。实验表明, 对弹性蛋白酶的抑制作用与人 α_1 -抗胰蛋白酶含量成正比, 但 NCS 可氧化 α_1 -抗胰蛋白酶, 使其丧失对弹性蛋白酶的抑制效应。酵母产生的野生型 α_1 -抗胰蛋白酶的抑制效应同外加入 α_1 -抗胰蛋白酶情况一样, 但诱变的 α_1 -抗胰蛋白酶 (Met \rightarrow Val) 则不受 NCS 氧化, 因此它对弹性蛋白酶的抑制作用并不受氧化剂 NCS 所左右。

3. 酪氨酰-tRNA 合成酶的改造: 这是一个经典的例子, 自 1982 年开始, 英国的 Fersht 和 Winter 等人用蛋白质工程手段对嗜热脂肪芽孢杆菌的酪氨酰-tRNA 合成酶进行了一系列系统的氨基酸更改, 并测定了这些更改对酶的内部结构及它的催化动力学方面的影响, 取得了令人信服的结果。

他们先是采用位点定向诱变的方法将该酶活性部位第 35 位的半胱氨酸更改为丝氨酸 (Cys 35 \rightarrow Ser), 并在大肠杆菌中获高效表达, 酶蛋白占全部可溶性蛋白的 50%^[15]。它的催化性质与野生型酶没有区别, 但催化能力却比野生型酶差, 说明在活性部位结构的变化会导致催化能力的变更。

其后,他们采用同样手段,根据立体结构提供的信息,又作了一系列在不同部位变更氨基酸的工作^[11-17]。如将 Thr 51 变成 Ala 或 Pro,都能降低酶对 ATP 的 K_m 值而提高酶活,其中 (Thr51→Pro) 的酶活力提高达 25 倍。最近他们根据对该酶分子中可能与底物之间形成的 11 个氢键中的八个作了氨基酸变更,并对其分子中的氢键结合情况和酶的生物学专一性进行了分析,认为生物学专一性在某种程度上是由氢键结合所引起的,并在极大程度上受带电残基所左右^[18]。

4. 其它例子: Villafranca 等人于 1983 年报道^[19],他们通过这种方法得到了大肠杆菌二氢叶酸还原酶的三种突变酶: Asp27→Asn, Gly95→Ala 和 Pro39→Cys。二氢叶酸还原酶是一种只有 159 个氨基酸的单链蛋白质,结构比较简单。为了更好地解释该酶的作用机制,他们对位于该酶空间结构上不同部位的这三个氨基酸作了变更,得到了突变酶。通过与野生型酶活力的比较和动力学分析,认为 Asp27 在催化功能上起重要作用。如在 Gly 95 位对保守的 Cis 肽键作局部结构的扰动,就会导致活力的丧失。用 Cys 代替 39 位的 Pro,就可能在新的 Cys39 与原来 85 位的 Cys 之间构成二硫键,而形成一全新的酶。

Brake 等人为了研究在酿酒酵母中成熟的外源蛋白分泌机制,对与此分泌有关的酵母外激素 α -因子作了一些工作^[20]。 α -因子是一个只有 13 个氨基酸残基的肽,它在被合成时先产生一个具有 165 个氨基酸的前体,其中含 83 个氨基酸构成的前导物 (Leader) 及 4 个 α -因子,每个前面都有短的间隔肽 (Spacer peptide),作者用位点定向诱变的方法将间隔肽、特别是前导物区域作了些变更,看它们对外源蛋白(成熟的人表皮生长因子: hEGF) 分泌能力的影响。通过比较认为,此前导物和间隔肽看来包含有将成熟蛋白断裂下来并分泌出去的必需信号。

此外美国 Cetus 公司的美籍华人麦富智将 β -干扰素中的两个半胱氨酸变成丝氨酸,从而使这种干扰素能在低温中保存半年还不降低活力,为干扰素的临床应用铺平了道路。

上面所列数例,从不同角度阐明了采用“蛋白质工程”技术的目的、方法和所得结果。正如前面所述,“蛋白质工程”作为一门新兴学科,目前还只处于起步阶段。由于它能按照人们的意志构造出新型蛋白质,通过它了解更多的蛋白质结构与功能的关系以及生命的奥秘,创造出更多有益于人类的蛋白质新产品,随着工作的进展,它对人类的贡献将是无法估价的。

参 考 文 献

- [1] Wasserman, B. P.: *Food Technology*, 38(2): 78—89, 1984.
- [2] Winter, G. et al.: *Nature*, 299(21): 756—758, 1982.
- [3] Ulmer, K. M.: *Science*, 219: 666—671, 1983.
- [4] *Biotech News*, 4(6), 1985.
- [5] Pramik, M. J., *GEN*, 4(3), 1984.
- [6] *Bio/Tech*, 3(3): 201, 1985.
- [7] *Bio/Tech*, 3(8): 682, 1985.
- [8] Hutchison, C. A. et al.: *J. Biol. Chem.*, 253(18): 6551—6560, 1978.
- [9] Wallace, R. B. et al.: *Nucl. Acids Res.*, 9: 3647, 1981.
- [10] Zoller, M. J. and M. Smith: *Nucl. Acids Res.*, 10(20): 6487—6500, 1982.
- [11] Messing, J.: *Methods in Enzymology*, Vol. 101, p. 20—77, Eds. by R. Wu et al., Academic Press, New York, 1983.
- [12] 孙晋武: *微生物学通报*, 10(5): 229—233, 1983.
- [13] Perry, L. J. and R. Wetzel: *Science*, 226: 555—557, 1984.
- [14] Rosenberg, S. et al.: *Nature*, 312(1): 77—80, 1984.
- [15] Wilkinson, A. J. et al.: *Biochemistry*, 22: 3581—3586, 1983.
- [16] Wilkinson, A. J. et al.: *Nature*, 307(12): 187—188, 1984.
- [17] Carter, P. J. et al.: *Cell*, 38: 835—840, 1984.
- [18] Fersht, A. R. et al.: *Nature*, 314(21): 235—238, 1985.
- [19] Villafranca, J. E. et al.: *Science*, 222: 782—788, 1983.
- [20] Brake, A. J. et al.: *Proc. Natl. Acad. Sci. U. S. A.*, 81: 4642—4646, 1984.

(本文于1986年1月27日收到)