

1

走近生命科学

有机大分子大厦的基石不过是 6 种元素 碳、氢、氧、氮、磷和硫。它们根据简单的‘游戏规则’(一些物理学化学定律)构建出丰富多彩的有机世界。雄踞这座分子大厦之顶的正是蛋白质分子和核酸分子 它们恰似生物化学和分子生物学的殿堂 人类最迫切地想要了解和解决的问题——衣食、健康、生态、环保 答案尽在其中。目前分子生物学在剧烈地改变整个人类社会，从提高平均寿命到攻克疑难病症，从高产作物的出现到全新物种的诞生。面对此景，谁会想到一个世纪前分子生物学的领地不过是一片荒原，又有谁能想象一个世纪后分子生物学的宫殿会是如此辉煌！

生命科学从 20 世纪 50 年代起进入分子水平。随后的 50 年 激动人心的研究成果不断出现，一位位成就卓著者走上诺贝尔奖领奖台。我们现在知道 DNA 是遗传物质，而蛋白质则是遗传信息的具体表现者。蛋白质和 DNA 之间关系的确定 为分子生物

学的飞速发展奠定了基础。

诺贝尔奖称得上当今科学领域中的最高奖项，但是科学贡献并不一定必须用诺贝尔奖来衡量。诺贝尔委员会的评奖原则中有一条即“奖项绝不能由三人以上分享”这保证了诺贝尔奖的专有性，但也设置了限定。事实上大量杰出的贡献者并没有获得诺贝尔奖。在核酸化学和蛋白质化学领域就有发明了连续式氨基酸序列测定的埃德曼 (Edman) 为沃森 (James Watson) 和克里克 Francis Crick 的 DNA 双螺旋模型提供了 X 射线衍射数据的女科学家富兰克林 (Rosalind Franklin) 发明了重组 DNA 技术的 H·博耶 (Herbert Boyer) 和科恩 (Stanley Cohen) 发现了癌基因 *ras* 的温伯格 (Weinberg) 合成出只含 L 型氨基酸的多肽的伯格曼 (Bergmann) 等等。更多的人则在科研竞争中落在了别人的后面，丧失了第一发现权。诺贝尔奖获得者宛如露出海平面的冰山，它只代表整个科技成果的一小部分。我们崇尚诺贝尔精神，追求科学上的创新性发现，但不可迷信诺贝尔奖和其获奖者。如果科学研究的最大目标成了诺贝尔奖本身，不能不说这是科学的悲哀。

因发现低密度脂蛋白 (LDL) 受体介导内吞而获得诺贝尔生理学医学奖的布朗 (Brown) 教授在复旦大学作讲演时，有学生问他：“怎样才能得诺贝尔奖？”他的回答是：“得奖不是一个科学家所追求的，有许多科学家虽然做出了巨大贡献却与奖无缘。对于一个科学家而言最重要的是他有没有解决自己最

想解决的问题。”分子生物学家们就是这样 各自钻研着自己感兴趣的问题 从各个角度共同努力 照亮了一个蛋白质和核酸的奇妙世界。

2

相识蛋白质

遭遇蛋白质

早期研究物质的方法和早期的化学理论一样简陋。一个半世纪前 人们连最简单分子的结构都搞不清楚 化学家们所能做到的只是把物质根据受热后的性质分为两类。一类 如金属、水、盐类 在受热再冷却后基本性质保持原样 另一类在受热后发生了质的变化。瑞典大化学家柏齐力乌斯 Berzelius 把后者称为有机物 因为它们大多直接或间接来自活的机体。不同的有机物在受热后的变化也不同 糖在受热后会变焦变黑；橄榄油在受热后会变为遇冷不凝的气体；还有一大类物质像蛋清受热后会由液态变为固态。1777年 法国化学家马凯 P. J. Macquer 把加热后能凝固的物质归为一类 称为“蛋白性物质”。

19世纪的有机化学家在分析蛋白性物质的时候，发现这些化合物比其他有机物分子复杂得多。普通的有机物分子都能写出一个确定的化学式，如

酒精的化学式是 C_2H_5OH ，烷烃的化学式是 C_nH_{2n+2} 而对蛋白性物质却不能总结出一个通用的化学式。19世纪30年代荷兰生理学家米勒 G. J. Müller)用当时最先进的元素分析法对血清、蛋清、蚕丝等进行了分析。他发现这些物质的元素组成很相近都能用 $C_{40}H_{62}N_{10}O_{12}$ 来表示，他认为这是蛋白性物质所共有的。由于能用这个分子式来表示的物质在生物体中无处不在，米勒认为它们在生物体中是最重要的因此于1838年将这类物质命名为 *proteid*。与此同时瑞典化学家柏齐力乌斯用了另一个词 *protein* 它们的原意都是“最重要的第一位的”源自希腊语。这就是我们所熟悉的“蛋白质”(protein)。

只要有生命的地方就能找到蛋白质，它是生命体中含量最丰富、功能最重要的大分子从高等动植物到低等微生物乃至病毒都以蛋白质为主要成分。从这一点来说米勒和柏齐力乌斯的命名非常准确。但米勒的化学式从蛋白质化学来说是错误的，他想弄清蛋白质的完整分子式的努力注定要失败。因为生物体中的蛋白质大多分子量巨大而且种类繁多，不可能用一个化学式来代表。好在当时已有人开始从另一个角度去进行研究弄清了蛋白质的组成最终阐明蛋白质的结构打下了基础。

由氨基酸串起的“项链”

对大分子化学结构的最早研究可追溯到 19 世

纪。1812年 俄国化学家基希霍夫 (G. S. Kirchhoff) 发现, 加酸煮沸的淀粉生成一种与葡萄糖性质相同的糖。1819年 法国化学家布拉科诺 (H. Braconnot) 通过煮沸植物产品 (其中含纤维素) 也得到了葡萄糖。布拉科诺决定用同样的方法处理明胶 (一种蛋白性物质) 结果得到了一种有甜味的结晶体 布拉科诺最初怀疑这种结晶体是糖, 后来发现它是一种含氮的化合物。现在人们把布拉科诺所分离出的这种化合物叫做甘氨酸。

随后 化学家们开始利用浓酸、浓碱来消化蛋白质 (一般称这种操作为蛋白质的水解) 分离得到其他一些氨基酸。目前已经鉴定出来的, 存在于大多数蛋白质中的氨基酸共有 20 种。最后一种即苏氨酸是 1935 年美国化学家罗斯 (W. C. Rose) 发现的。一切证据都表明蛋白质是由氨基酸组成的大分子多聚物。这些氨基酸又是怎样连接在一起, 构成蛋白质的? 1902 年, 德国的有机化学家费歇尔 (Hermann Emil Fischer) 和奥夫梅斯泰 (Hofmeister) 同时提出了肽键理论。费歇尔出生在一个富裕的商人家庭, 他的父亲希望他能够进入家族企业, 而他的爱好却是自然科学。值得庆幸的是他的父亲不是那种独断专行的家长 在培养了他一段时间后 父亲失望了 觉得自己的儿子太笨了 不能做一名商人 最好还是去当学生。于是费歇尔如愿以偿地进入波恩大学攻读化学。天资、勤奋和命运的完美结合使得费歇尔在科学道路上一路顺风 :29 岁成为埃朗根大学的化学

教授 40 岁接任柏林大学化学教授一职（只有最杰出的科学家才能得到的席位），50 岁作为糖化学之父荣获 1902 年诺贝尔化学奖。如果说这位学者有一丝遗憾 那就是他没能从事自己最喜欢的物理学。费歇尔在化学上有过人天赋 他善于研究物质间的本质联系 并能想出办法合成它们 这是在今天也会令科学家们瞩目的成就。更令人敬佩的是 费歇尔没有在荣誉和成就面前止步，1899 年开始他投身于蛋白质化学 并做出了巨大的贡献。20 世纪初 费歇尔用滚热的沸腾强酸处理蛋白质 所得产物中含有氨基酸和大小不一的片段 他断定蛋白质分子包含了由氨基酸组成的长链。像研究糖和嘌呤一样，费歇尔首先寻求能有效分离氨基酸的办法 由此发现了一种新的氨基酸——脯氨酸。随后他试图从已得到的有光学活性的氨基酸合成蛋白质。费歇尔尝试把一个 L 型氨基酸的氨基和另一个 L 型氨基酸的羧基连起来 使几个氨基酸连接成链。1901 年他首先完成了这种缩合工作（两个氨基酸相连的过程中会去掉一个分子的水 称为缩合作用）把一个甘氨酸和另一个甘氨酸连接了起来 这是最简单可行的缩合。1907 年费歇尔合成了一个由 18 个氨基酸构成的链 由 15 个甘氨酸和 3 个亮氨酸组成。不过这个分子没有明显的蛋白质特性 费歇尔的解释是链还不够长。他把所合成的链称为肽 并把羧基和氨基之间通过缩水形成的键称为肽键。

但怎样证明天然蛋白质是由这样的肽键结构连

接起来的呢？费歇尔的关门弟子伯格曼设计了一种方法，可把各种不同的 L 型氨基酸连在一起组成短肽。美国的生物化学家弗吕东 J. S. Fruton 用伯格曼的方法制备了一些肽，这些肽可被消化液分解成小的片段。当时已经证明消化液只能打开一种分子键，所以这就说明在合成的短肽中，氨基酸之间的键一定和真正蛋白质中连接氨基酸的键一样。这个证据消除了人们对费歇尔关于蛋白质的肽键结构学说的怀疑。

体外合成蛋白质

人工合成蛋白质是当时所有研究蛋白质化学的科学家们的梦想。只有这样做才能表明人们对蛋白质结构的认识是正确的；同时人工合成蛋白质的意义十分重大。曾经风行一时的活力论（认为只有活的机体才能制造有机物的观点）已在 1828 年被德国化学家沃勒（Wohler）的实验所否定。他在实验室中加热氰酸铵，当时普遍认为它是无机物，得到了尿素（尿的成分之一）。1845 年德国化学家科尔比（A. W. H. Kolbe）第一次由元素制成有机物醋酸，彻底粉碎了活力论。从理论上说，亦属有机物的蛋白质也应当能够被合成出来。蛋白质在生命中的地位使得合成它要比合成其他有机分子更有意义，无论是在理论上，还是在实践上。

在合成蛋白质的道路上，科学家们一丝不苟，循

序渐进 从最简单的开始做起。20 世纪早期 科学家所能合成的肽很小，而且在性质上和蛋白质大相径庭。1916 年瑞士化学家阿布德哈尔登 (E. Abderhalden) 制成了含 19 个氨基酸的肽 这个记录保持了 30 年。费歇尔和伯格曼的方法一次只能加上一个氨基酸 随着肽链的延长 反应步骤的增多 产率也越来越低，所以他们的方法不适于合成长的肽链。而当时的化学家已经知道蛋白质的分子量十分巨大 这样一个肽不过是蛋白质分子的一个极小片段。

如何合成长的肽链呢？以色列化学家卡乔斯基 (E. Katchalski) 和哈佛的化学家伍德沃德 (Woodward) 同时报道他们用连锁反应成功地聚合出多肽来。这些多肽链一般由一种氨基酸组成，也可以用两种氨基酸的混合物制成含两种氨基酸的多肽链。这些成品还不能算是真正的蛋白质，因为自然界中的蛋白质往往是由 20 种氨基酸以特定顺序连起来的多肽链。

当科学家们解析出天然蛋白质的氨基酸顺序后 (详见第三章) 合成蛋白质的道路上出现了第一缕曙光。接下来的问题是如何按照解析出的氨基酸顺序合成蛋白质。因为在合成一个具有特定氨基酸顺序的多肽链的过程中 每连接一个氨基酸残基 都要经过几个步骤，要想得到一个足够长的多肽就必须每步都有高产率。提高产率的方法有三 (1) 减少其他讨厌的副反应。接肽反应的主角应当只是氨基酸残基的氨基和待连上去的氨基酸的羧基，它们缩去

一个水分子形成肽键，完成肽链的延长。而 N 末端（蛋白质有两个端点，一端是个氨基 称为蛋白质的 N 末端 另一端是个羧基 称为蛋白质的 C 末端。）氨基酸残基的自由氨基、C 末端残基的自由羧基、侧链上的一些活泼基团 都能和接肽的试剂发生作用。所以接肽以前必须首先把这些基团加以封闭或保护，以免和接肽试剂发生作用生成不需要的肽键或其他键。作为保护基 必须符合这样的条件 在接肽时能起保护作用，而在接肽反应后又容易除去。应用最广泛的氨基保护基是苄氧甲酰基，羧基保护基通常用烷基。(2) 增加反应基团羧基和氨基的活性。羧基的活化最早采用的是酰氯法，这个反应条件激烈，已不大采用。现在通常使用比较温和的羧基活化法如叠氮化合物法、活化酯法。氨基的活化比较简单 不需要化学修饰 只需在接肽时以有机碱保证氨基的自由状态即可。(3) 提高缩水剂的效率。目前最有效的缩水剂是 N, N'-二环己基碳二亚胺 (DCCI)。

在实验室里最先合成的多肽是催产素。催产素和加压素是脑垂体后叶分泌的由 8 个氨基酸组成的激素 催产素在体内能促进子宫收缩和刺激产乳 加压素能提高血压，调控肾的功能。1933 年从垂体后叶提纯的催产素的高含硫量引起了美国化学家迪维尼奥 Du Vigneaud 教授的极大兴趣。因为他一直以来都在从事硫代谢的研究，特别是对含硫的半胱氨酸和甲硫氨酸做了许多重要的研究工作，他认为催

产素的高含硫量肯定和它的生理活性有关。进一步提纯这两种激素后，迪维尼奥发现它们的结构类似于蛋白质，只是分子量小得多。为了与真正蛋白质区分，这样的小肽链被命名为多肽。在测定了催产素中氨基酸的顺序后，含硫的分子找到了 2 个半胱氨酸。迪维尼奥还解析出催产素和加压素的结构，它们的肽链在一点上连接成环 就像“6”或“9”的形状。环由 5 个氨基酸组成 尾巴是另外的 3 个氨基酸。2 个半胱氨酸的—SH 基连成—S—S— 构成了环的一部分。分子的结构知道了，检验其正确性的办法就是合成它，这是更加困难的工作。整个合成工作分为两步：(1) 一个一个把氨基酸连起来，保证 2 个硫原子在正确的位置，一个在链尾，一个接近中部；(2) 把两个硫原子用键连起来，使肽链成环。最激动人心的时刻到了 经检验 人工合成的催产素的化学性质和生理性质与天然催产素完全相同。1955 年，迪维尼奥被授予诺贝尔化学奖。正如诺贝尔奖委员会在致迪维尼奥教授的信中所说：“您以杰出的知识和技术合成了一个复杂的、具有重要功能的生物分子，但更重要的是这个合成背后的意义。它证明了具有重要生理功能的物质可以用我们熟悉的氨基酸和我们熟悉的化学定律合成。它同时表明氨基酸在肽链中的顺序至关重要，这个结果对阐明真正蛋白质的秘密可能非常重要。”

简单的短肽激素合成成功后，人们把目光瞄向了真正的蛋白质。1955 年 英国剑桥大学的桑格

(Frederick Sanger) 用色素标记和部分水解的方法确定了第一个蛋白质——牛胰岛素分子的氨基酸连接顺序(即一级结构)。此后不久,中国的几位有机化学家和生物化学家在北京讨论了胰岛素人工合成的可能性。当时国外某些实验室也试图合成胰岛素,但在第一步就失败了,他们不能把拆开的胰岛素的两条链再连起来,于是认为合成胰岛素是不可能的,至少暂时不可能。1959年,在国家科委的组织领导下,由北京大学化学系、中国科学院生物化学研究所和上海有机化学研究所的钮经义、王应睐、龚岳亭、邹承鲁、杜雨苍、季爱雪、邢其毅、汪猷和徐杰诚等共同组成一支研究队伍,开始了胰岛素的合成研究。实验采取了三步走的策略:首先成功地将天然胰岛素的A、B两条链拆开,再重新连接而得到了重合成的天然胰岛素结晶,为下一步的人工合成摸索了条件,确定了路线。随后人工合成B链和A链,并分别与天然的A链和B链连接而得到半合成的胰岛素;最后将人工合成的A链和B链连接而得到全合成的结晶胰岛素。经过数年的共同努力,第一个用人工合成方法得到的活性蛋白质——结晶牛胰岛素,终于在1965年降生在中国大地上。经鉴定,人工合成的胰岛素,它的结构、生物活性、物理化学性质、结晶形状,都和天然的牛胰岛素完全一样。结果公布之后,立即引起世界科学界的极大关注,它标志着中国科学家在蛋白质和多肽合成化学领域已经处于世界领先地位,为我们这个伟大的文明古国赢得了新

的荣誉。当年，诺贝尔委员会主席亲自来到中国考察授奖事宜。谁知考察之后他为难了：按照诺贝尔奖的有关条例 授奖对象最多不能超过 3 人 而参加人工合成牛胰岛素研究的人员，仅骨干就有 20 多位，不符合授奖规章。因此中国的科学家们与诺贝尔奖失之交臂 而 1982 年国家特发给在胰岛素合成工作中做出贡献的科学家们的国家自然科学一等奖也就具有了和诺贝尔奖同等的分量。一位来中科院参观的外国专家这样评论：“虽然中国发射了原子弹 但毕竟落在了美国、苏联的后面 而在合成蛋白质上中国人走在了世界的最前头！”

有机化学法合成长的肽链遇到的困难之一就是新氨基酸连到肽链上后如何提纯新肽链，以供下次连接反应所用。这个问题不解决，合成肽链的成本就不会降低到可商品化的程度。1959 年梅里菲尔德 (R. Bruce Merrifield) 的创新解决了这个难题。梅里菲尔德 1921 年 7 月 15 日出生在得克萨斯州沃思堡，但在加利福尼亚度过了他的青少年时光。在蒙特贝罗中学青年梅里菲尔德显示出对化学的浓厚兴趣，同时他还参加了学校的天文学俱乐部和学校的年度科学竞赛。1949 年从加州大学洛杉矶分校研究生毕业后 他离开加州 来到了纽约的洛克菲勒医学研究所 (即后来的洛克菲勒大学) 成为伍利 (D. W. Woolley) 博士的助手。他们的工作是研究肽生长因子，这就要求他们在实验室中合成肽链。也正是这个需要把梅里菲尔德引上了 1984 年诺贝尔化



梅里菲尔德

学奖领奖台。为了轻易地把产物和其他成分分开 梅里菲尔德把待合成的肽链的第一个氨基酸连接到聚苯乙烯树胶小颗粒上 这些小颗粒在所用的溶液中不能溶解 只要过滤就能把它们与反应后剩余的氨基酸和其他杂质分离。在合成肽

链的循环中 每次加一种氨基酸进行反应 然后除去剩余的未连接氨基酸 再加入下一种氨基酸 依次下去 直至得到所要的肽链。这样大大地提高了合成肽链的速率和产率。梅里菲尔德等利用这种自动化的多肽合成仪于 1969 年完成了牛胰核糖核酸酶的人工合成 这是第一个人工合成的酶。梅里菲尔德的这个发明的最大好处是可以使人工合成多肽自动化 把研究人员从繁琐的合成程序中解放了出来。

目前固相合成多肽已应用于医药工业 用来合成一些有重要医学意义的短肽。这些合成品和天然产品有同样的生理效果 有时甚至更好。人工合成的催产素没有混杂的加压素 因此比提取的天然产品还要好。但人类手中的技术和生物体比起来还是相形见绌 利用自动合成多肽的仪器合成一个九肽 (含有 9 个氨基酸残基 需要 27 小时 平均每个肽键需要 3 小时, 而在一个细菌细胞内合成一个含 100

个氨基酸残基的蛋白质只需要 5 秒。

第一个被分离纯化的蛋白质

人工合成胰岛素的巨大成功是站在那些分离、解析出胰岛素结构的科学家肩上取得的。只有得到了很纯的胰岛素分子 分析出它的氨基酸顺序 才能根据这张自然界的蓝图合成胰岛素。虽然是和其他千千万万个分子共同完成生命功能，但历史把胰岛素推上了一个令其他蛋白质分子们羡慕的地位。每本生物化学教科书都会详细报道胰岛素的氨基酸组成、结构和功能 3 位科学家因它获得诺贝尔奖 他们是：发现胰岛素的班廷 (Frederick Banting) 和麦克劳德 (Macleod)，以及测定牛胰岛素氨基酸顺序的桑格。而准备合成蛋白质的中国科学家也选择了胰岛素。如果说胰岛素是当时的明星，那么把它领上舞台的是两位加拿大科学家班廷和麦克劳德。

虽然同是加拿大人，班廷和麦克劳德有着完全不同的医学背景。班廷是一位技术一般的军医，没有研究经验，更不知道糖尿病和胰腺功能的关系。而麦克劳德是多伦多大学某实验室的负责人，在糖代谢方面享有国际声誉。不过在发现胰岛素中起主要作用的却是班廷。早在 1889 年，明科夫斯基 (Minkowski) 和冯·梅林 (von Mering) 发现完全切除胰腺会导致糖尿病 把糖尿病和胰腺联系到了一起 当时所知胰腺的功能是分泌消化液)。经过仔细而又

繁琐的解剖实验 朗格汉斯 Langerhans 和迪亚马雷 (Diamare) 断定是没有导管的胰岛细胞分泌了抗糖尿病因子。为纪念发现者 这种胰岛细胞被冠名为朗格汉斯细胞。谢弗 (Sharpey Schafer) 把这个可能的抗糖尿病因子命名为胰岛素 (insulin) 但在糖尿病病人身上施用胰腺抽提物或胰腺却没有取得成功。当时估计可能这种抗糖尿病因子被胰腺分泌的消化液破坏了。1908 年齐尔泽 (Zuelzer) 针对 6 例糖尿病病人试用胰腺的乙醇抽提物 取得良好的效果, 一个重病病人的尿中没有了糖。他制备的抽提物在明科夫斯基诊所中由福施巴赫 (Forschbach) 试用 却没有取得同样良好的效果。于是这个小组放弃了这项研究。1920 年, 班廷被一篇关于胰腺细胞功能的文章所吸引, 从此走入了研究糖尿病和胰岛素的大门。这篇文章是巴依 (Moses Baron) 写的 他指出结扎通向胆囊的胰管后, 胰腺的葡萄状腺细胞萎缩, 而没有导管的朗格汉斯细胞却没有较大损伤。班廷当即想到如果结扎胰管会破坏分泌胰蛋白酶的葡萄状腺细胞 那么提纯抗糖尿病因子就不会被破坏 那时的胰腺抽提物可能会起到抗糖尿病的作用。班廷向麦克劳德提出了他的构想 但麦克劳德的反应却很冷淡, 因为班廷的主意不流畅, 且有些地方不合生理学原理。但班廷坚定不移 要求尝试。几个月后 麦克劳德看到了班廷方案的价值 同意资助他的研究 并给班廷两个学生助理和 10 条狗。大学生贝斯特 (Charles Best) 是两个助理之一 他负责测量血中和