

专论与综述

分子生物学的现状与展望

邹承鲁

(中国科学院生物物理研究所)

分子生物学是当前生物学发展的主流。自本世纪后半叶以来它是整个自然科学中发展最迅速的学科之一,不仅带动了生物学的全面发展,并且为农业、工业、医学以及国防方面的研究开辟了广阔的前景。本文仅就近年来的一些重要进展和动向作简略地介绍。

分子生物学大体可分为三大领域:一是蛋白质体系(包括酶),二是蛋白质-核酸体系(中心问题是分子遗传学),三是蛋白质-脂质体系(即生物膜)。现围绕这三大领域以及分子生物学在其他领域中的应用作些介绍。为了叙述方便,在生物大分子的标题下主要介绍蛋白质体系,但也附带谈到一些核酸方面的问题。

生物大分子

1. 生物大分子化学结构的测定

这是研究生物大分子结构与功能关系的基础。在蛋白质方面,已经测定化学结构的蛋白质(如果把不同种属的同一种蛋白质都计算在内),总数在一千种以上。其中最大的分子量已超过十万,如大肠杆菌的半乳糖苷酶^[1]和胶原蛋白。回想当初 Sanger 实验室,前后共五、六个人参加,历时八年,才完成了胰岛素这个仅 51 个氨基酸组成的最小的蛋白质化学结构的测定工作。二十多年来,蛋白质化学结构的分析技术有很大发展,如今借助于氨基酸自动分析仪和氨基酸顺序自动测定仪,顺利时一次可完成长达六、七十个氨基酸残基的顺序测定。但总的说来,毕竟还是比较麻烦的。

核酸化学结构的测定在七十年代初期还落后于蛋白质顺序的测定。那时已知化学结构的核酸为数有限;测定一个由二十个核苷酸组成

的核酸结构,一个熟练工作者需工作三年。但是近年来,核酸,特别是 DNA 化学结构的测定,有了飞跃的突破。一个熟练工作者,只需一天功夫,就可测定 250 个核苷酸组成的 DNA 链的顺序。所用方法主要有两种:一是英国发展起来的“正负法”,最近又改进为新的“末端终止法”,另一是美国发展的化学断裂法。迄今已知序列的最长的是 MS₂ 噬菌体 RNA,共 3569 个核苷酸^[2];DNA 顺序测定以 1977 年完成的 ϕ X-174 噬菌体 DNA(5386 个核苷酸)^[3] 为标志。近年来许多实验室竞相测定更大的 DNA,所含核苷酸数已激增至几万,如线粒体 DNA(约 17,000 核苷酸)的序列测定接近完成, λ 噬菌体 DNA(5 万核苷酸)片段顺序的测定已完成十分之九。可以预期今后几年将会有为数百万核苷酸的 DNA 顺序问世,甚至连过去不敢想象的数以亿计的核苷酸顺序的测定(如人的单个染色体 DNA)也有可能得以实现。

对比之下,蛋白质结构分析技术相形见绌了;已经测定的序列还未曾达到象核酸这样的长度。看来,今后蛋白质的顺序分析最好的方法是先搞到蛋白质的 mRNA,再反转录成 DNA,然后分析 DNA 的序列;同时在蛋白质的头尾选几部分作顺序分析,若能与 DNA 序列相互对应就可以推出蛋白质中的序列。现在已有人开始这样做了。例如,根据人干扰素基因 DNA 序列的测定和干扰素 N-末端部分序列的测定,推断出人干扰素的全部氨基酸排列顺序^[4]。

2. 生物大分子的高级结构

X 光晶体衍射技术的发展解决了生物大分子三度空间结构,使人们几乎可看到每个原子的位置。这方面的工作对分子生物学的发展作

出了重要的贡献。现在已有几十个蛋白质通过X光晶体衍射获得了高分辨率的三度空间的结构。核酸的空间结构测定还仅仅限于少数几个RNA。

不同种属执行同一生物功能的蛋白质，其化学结构有一定的相似性，而空间结构的相似性表现得更为突出。鱼线粒体和细菌的细胞色素C，在氨基酸序列中相差60多处，即差别达到百分之六十左右，但从X光晶体衍射的结果看，二者的空间结构却非常近似。

用X光晶体衍射测定比较大的蛋白质分子是一些激酶、脱氢和转换酶，如乳酸脱氢酶，甘油醛-3-磷酸脱氢酶和磷酸化酶等，分子量从14万—20万左右。由于脱氢酶和激酶都以核苷酸化合物为底物，它们在空间结构上有很多相似之处。脱氢酶空间结构大体可分为两部

分^[5]，这两部分都是较紧凑的结构，在蛋白质结晶学上称为结构域(Domain)，其中一部分与底物结合有关，另一部分则是与NAD⁺结合有关(图1)，这一部分结构域与许多激酶的空间结构有不少相似之处，这些相似之处可能是由于这两类酶都要和核苷酸结合所决定的。

对上述现象现在有两种解释。一种是“分化进化说”，即在很原始时脱氢酶和激酶是同一个酶，承担两个不同的任务；随着生物的进化，逐渐分化成两类酶，但仍保持着原始祖先的某些相似之处。另一种是“聚合进化说”即这两类酶原来是无关系的，但由于都要和核苷酸相结合，遂演变成了类似的结构。这两种学说不一定矛盾。即对整个生命世界，这两种情况也许都会发生，但对某一种酶，它只有一种可能性。看来在多数情况下分化进化的可能性比较大。

X光晶体衍射也有它的缺点和不足之处。它测的是生物大分子晶态的结构，而绝大多数生物大分子在生物体内不是处于晶态，而是处于溶液状态或者是与膜紧密结合的。从晶态结构能否了解其生物功能？晶态结构与溶液结构及连在膜上时的结构是否一样？对此可以作一些回答。有一些蛋白质在晶态下仍表现一定的生物功能，如血红蛋白与氧的结合在晶态下一样可以进行。某些酶在晶态仍能表现其活性。经过这样的比较得出的一个大致的结论是：二者基本相同，但有时有细微差异。这个差异用酶来说明是最恰当的，因为可对晶态与溶液的酶活性进行比较。当然，由于测定晶态酶活性的方法和计算不同，得出的结果总是基于一定的假设之上，不那么可靠，但如果一个酶其晶态与溶液的活性在同一数量级，则可认为晶态结构基本上反应了溶液结构，反之，若二者有差别，像醇脱氢酶二者差一千倍，则两种状态可能会不大一样。

3. 低温酶学^[6]

X光晶体衍射的另一不足之处是它测的是静态的结构，而生物大分子的功能则存在于运动之中，因此，我们关心的是动态，希望能看到如何变化的过程。解决这一问题的手段之一是

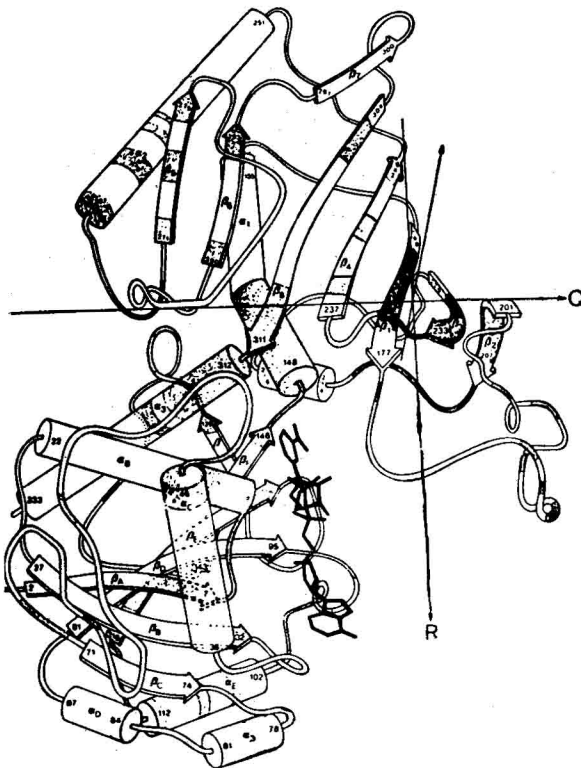


图1 甘油醛-3-磷酸脱氢酶示意图

上半是底物结合及催化结构域，下半是NAD结合结构域。圆筒代表 α -螺旋，箭头代表 β -折叠。黑线结构是NAD，灰色部分是已知5个种属的酶分子中氨基酸序列完全相同的部分。

将X光晶体衍射与低温生化结合起来。以酶催化的反应为例,这是非常快的过程。但温度每降低10℃,反应速度就可降低一半,如果把温度降至-50℃左右,反应速度就大大降低了,本来不稳定的中间物相对来说稳定了。如弹性蛋白酶的晶体在70%的甲醇溶液中冷至-55℃,然后用它的专一性底物苄氧羰基丙氨酸对硝基苯酯浸泡,可以得到酶的酰化中间物,在上述条件下可稳定二天,仍维持相当高的酰化比例,因此就有可能进行晶体衍射分析^[7]。从底物浸泡前后的差电子密度图(图2)可清楚地看出在活性部位丝氨酸羟基上的苄氧羰基丙氨酸的酰化物。这是人们第一次直接观察到一个酶和它的专一



图2 弹性蛋白酶的酰化中间物

A: 弹性蛋白酶晶体在70%甲醇溶液中在-55℃的X光衍射电子密度图,活性部位 Asp, His 和 Asp 的位置已在图上标明。

B: 与底物苄氧羰基丙氨酸酯浸泡后的电子密度图。与A相比在丝氨酸上新增的电子密度代表苄氧羰基丙氨酸。

性底物形成的,本来是不稳定的共价中间物。

4. 酶作用的过渡态

70年代以来酶学另一新进展是关于酶作用的过渡态的研究。过渡态的概念 Paccling 早在40年代就把它从化学动力学引入到生化领域,但当时并未引起重视。近十几年来由于积累了许多实验证据,才渐渐地受到注意^[8]。

根据这个理论,酶所以具有很高的催化效率就是由于它与过渡态中间物紧密结合,从而降低了底物形成它的过渡态时所需克服的能障(图3)。可见研究酶与过渡态中间物的结合方

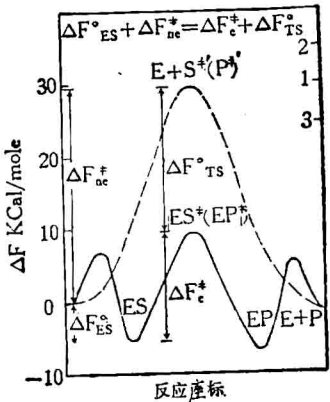


图3 酶催化反应的过渡态

图中虚线代表非酶反应,实线代表酶反应;S*是反应物(底物)的过渡态中间物;ES、EP分别代表酶与底物及产物的络合物;ES*是酶与过渡态中间物的络合物。

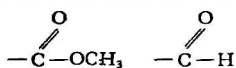
式对于了解酶作用的机理非常重要。但是ES*的半寿期很短,约在10⁻¹⁰—10⁻¹²秒的范围内,要研究它是很不容易的。目前的做法是寻找过渡态中间物的类似物,由于这些物质在结构上与S*类似,它们与酶结合常常比底物还要紧密,研究它们与酶的结合方式对于说明酶的作用机理也是很有帮助的。现在已经有六十多个酶发现了过渡态中间物的类似物。表1中是几个例子,从中可见,这些类似物通常都是酶的抑制剂。它们与酶的结合比底物还要牢固3—5个数量级,即Ki/Km约为1000至10万,说明它们在结构上比底物更接近于在反应过程中生成的过渡态中间物(真正的过渡态中间物与酶的结合应该比底物要牢固10⁷—10¹⁷倍)。这些抑制剂本身的结构,往往已经能对酶反应的机

表 1 一些酶的过渡态中间物的类似物

系统号	酶	底物	类似物	Ki/Km
1.1.3.2	乳酸氧化酶	乳酸	草酸	10^{-3}
1.1.1.27	乳酸脱氢酶	乳酸	草酸	10^{-4}
2.7.4.3	腺苷酸激酶	ATP	$A_{p5}A^*$	10^{-5}
3.4.4.10	木瓜蛋白酶	$Ac-Phe-Gly-CH_2^+$	$Ac-Phe-Gly-Cinal^*$	10^{-3}
4.1.2.7	醛缩酶	磷酸丙糖	磷酸乙醇酸	10^{-4}

* $A_{p5}A$, 见正文

* 底物和类似物羧基末端的结构分别为

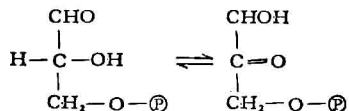


理提供重要线索。例如腺苷酸激酶催化的反应是:

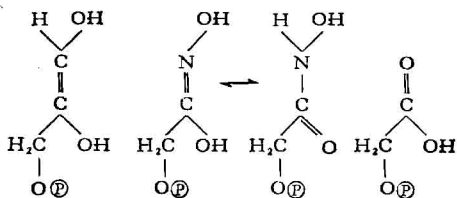


后来发现,两个腺苷分子在各自的核糖 5' 位上以五聚磷酸连接的分子 ($A_{p5}A$) 和相应以四聚磷酸连接的分子 ($A_{p4}A$) 对腺苷酸激酶的抑制能力极为不同;前者极强,后者很弱。这一事实表明在酶的活性部位同时存在结合 ATP 和 AMP 的部位,两者大体上是线性排列。处于过渡态时,ATP 分子的末端磷酸已略为离开它原来的位置,而趋近于 AMP 分子的磷酸基团。此时 ATP 和 AMP 两个分子的腺苷的距离大约相当于五聚磷酸的长度。

磷酸丙糖异构酶催化的反应是 D-甘油醛-3-磷酸和磷酸二羟丙酮的互变。



它的过渡态中间物可能具有烯醇类物质的结构。这一结构和磷酸乙醇酸及其羟肟酸的结构相比较时,可以看到其相似性。这些类似物不仅



是酶的强烈抑制剂,并且引人注意的是,丙糖磷酸异构酶的晶体浸泡在类似物的溶液后,引起酶的晶体形状明显改变,使得棒状晶体比浸

泡前缩短了百分之五左右。这种差别肉眼都可以觉察,但是在用底物浸泡时则观察不到。由此推断,过渡态中间物与酶的结合非常牢固,而且使酶分子发生明显的结构变化。

酶的过渡态的研究不仅对了解酶的催化机理有重要意义,而且也为药物设计指出了一条新的途径。过去设计药物总以为药物和底物的结构越象越好,而现在应从与酶的过渡态中间物的结构类似去考虑,才能获得更佳的效果。

5. 记忆性酶——单体酶的别构现象

酶的调节性质的研究已受到越来越多的重视。别构作用过去都认为是寡聚体蛋白才具有的功能,但近年来发现有些酶明明是一个单体,也具有类似于别构的现象。这种酶可称为滞后酶 (hysteretic) 或记忆性酶 (mnemonic)。如肝脏的葡萄糖激酶是一个单体的酶,分子量 4 万左右,其反应速度相对底物浓度作图所得曲线明显地呈 S 形。为何有此现象? 有一种解释是^[2]: 游离的酶在开始时以两种构象状态存在 (图 4, 分别以圆圈和菱形代表), 这两种构象的互变是一个相对缓慢的过程, 而在平衡时酶主要是以圆圈所代表的构象形式存在。当配基 (底物或效应剂) 与上述任一种构象状态的酶结合后, 都能将酶诱导变成相同的第三种构象状态, 以方形代表 ($\square$)。但 $\square$ 不能直接释放出产物, 必须经过 $\diamond$ 才能放出产物。所谓记忆性酶就是指酶在释放产物之后仍“记得”它在结合产物之前的构象是菱形的, 不再回到圆圈形了。由于开始反应时大部分酶是以圆圈形存在的, 而圆圈

分子遗传学

1. 左旋 DNA 双螺旋的发现

DNA 分子双螺旋模型的提出是分子生物学发展史上的一个里程碑,它打开了探索生物体遗传奥秘的大门,奠定了分子遗传学的基础。随着遗传密码的发现,“中心法则”的确立,以及基因工程的出现,分子遗传学以其一系列的重大成果有力地推动了分子生物学的发展。

最近关于左旋 DNA 的报道^[1],又一次引起了人们的重视。这是根据人工合成的胞嘧啶和鸟嘌呤相间排列的寡聚脱氧核苷酸得到的结果;在天然的 DNA 分子中,类似的序列也是肯定存在的。但是这种结构在天然 DNA 分子中究竟是广泛存在,还是在少数情况下的一种特例,它的生物功能又是什么,都有待于进一步研究。

根据 X 射线结构分析的结果,左旋 DNA 呈细长而伸展的状态,碱基对偏离轴心而靠近双螺旋的外侧。碱基的这种暴露状态使其更易受外界因素的影响。由于磷原子的连接线呈左旋锯齿状,因此称为 Z-DNA,以区别通常的 B-DNA。

2. 遗传密码的例外

遗传密码最初主要是根据以大肠杆菌为实验材料所取得的结果测定出来的。近几年来,由于蛋白质和核酸顺序分析的进展,遗传密码也就可以通过比较蛋白质的顺序, RNA 的顺序和 DNA 的顺序来加以验证。一直到不久前,还没有发现有一处与遗传密码表有出入的,这说明遗传密码表在生物界是广泛适用的。但是最近在测定线粒体 DNA 全部顺序的过程中,将其序列与由线粒体 DNA 编码合成的蛋白质序列相对照,发现通用的遗传密码表出现了三处例外(表 2 中加括号处), CUA 应是 Leu,但是在线粒体中却是 Thr; AUA 由 Ile 变为 Met; UGA 由终止信号变为 Trp。这一发现最近已得到不同实验室的工作者证实。

这是一个意外的重要发现。过去认为线粒体在某些方面类似于原核生物,并以此做为真

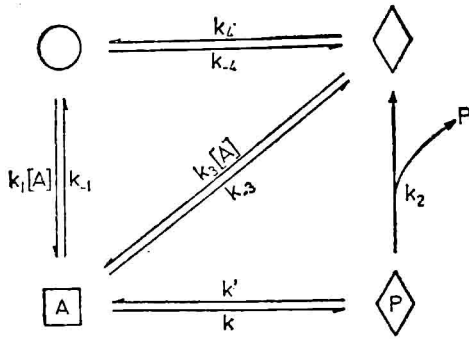


图 4 滞后酶构象变化示意图(说明见正文)

变成菱形的构象变化是一个比较慢的过程,随着反应的进行,酶不再经历这一慢的过程了,这是反应开始时有一段滞后期的原因。这也是以反应速度相对底物浓度作图时得到 S 形曲线的原因。

这种慢构象转变,在蛋白质化学中不仅是有解释的,而且是有先例的。比如两个结构域之间相互关系的变化,肽链的比较大的卷曲的改变以及两个螺旋之间相对关系的变化都是慢过程,所以滞后现象不仅有实验证据,而且和构象研究的结果也是符合的。

6. 生物大分子的人工合成

我们国家首先合成了胰岛素,现在又在进行核酸的合成,在生物大分子的人工合成方面做出了不少成绩。但是无论是国际还是国内,在蛋白质的合成方面十几年来没有重大的突破,至今要合成一个蛋白质仍是非常艰巨的工作。在核酸的合成方面,经过 9 年努力, Khorana 和 24 位共同工作者报道了大肠杆菌酪氨酸转移核糖核酸前体基因包括起动和终止信息的全合成工作^[10]。他们用的方法是化学合成和酶促合成相结合,合成的量极少,以至要设计出非常灵敏的测活方法,才能测出转录活性,即利用 *E. Coli* 噬菌体的变种,灵敏度为 10^{-14} 克分子。

总的看来蛋白质和核酸的人工合成仍是一项十分艰难的任务。一个简单的细胞,如支原体就含有六、七百种蛋白质和相应的核酸。大肠杆菌有人估计约含一千种以上的蛋白质。合成一个最小的蛋白质和核酸尚且如此困难,因此要合成一个完整的细胞恐怕还是相当遥远的事。

表 2 遗传密码表

	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	终止	终止(Trp)	A
	Leu	Ser	终止	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu(Thr)	Pro	Glu	Arg	A
	Leu	Pro	Glu	Arg	G
A	ILe	Thr	Asn	Ser	U
	ILe	Thr	Asn	Ser	C
	ILe(Met)	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

核生物的共生起源学说的依据。但是,现已证明线粒体遗传密码和真核及原核生物的不同,这对于我们探讨生物体早期的进化,是一个有价值的线索。

3. 染色体的结构与功能

真核细胞基因表达的调节控制是分子遗传学当前面临的重要问题之一。

动物的不同器官例如大脑、肌肉、肝脏,或是植物的花、叶、茎等,分别由不同类型的细胞组成。这些不同类型的细胞却是由同一个受精卵分化出来的,并都含有发育成为一个完整个体所需的全部遗传信息。近年来,从一个植物体细胞培养成为一个完整的植株已经成功。在动物方面也有不同类型的体细胞都含有相同的基因的实验根据。这些体细胞都携带着相同的遗传信息,为什么又各不相同呢?这是因为基因表达的全过程是在严格的调节和控制之下进行的。

染色体的结构大体上像串珠^[12],但线不从珠子中心穿过,而是绕在它外面。珠子由 8 个蛋白质分子构成。DNA 的双螺旋就绕在珠子的外面,两个珠子之间有一段游离的 DNA 结构,把这样的结构用一定的 DNA 酶处理,两个珠子之间的 DNA 被去掉,得到的东西暂名为

核小体。用 X 光衍射和电镜观察可以看到它的粗结构是一个扁圆体,长 110 Å,厚 57 Å,蛋白质在中心,有 140—200 个残基对的 DNA 双螺旋绕在碱性蛋白质(组蛋白)的外面。扁圆体好象一个半张开的蚌,中间有一点是空的。这样的结构比较稳定。从这一结构还看不出蛋白质的调控作用。组蛋白的作用看来是维系 DNA 稳定的结构。另一类种类很多的蛋白质是酸性蛋白和中性蛋白,统称非组蛋白。现在看来主要是这些蛋白质对调控起作用。但由于非组蛋白种类多,含量低,变化快,所以分离很困难。近年来技术上有所突破,发展了两向凝胶电泳^[3]:一向是凝胶等电聚焦,根据蛋白质的等电点来分离,另一向是十二烷基硫酸钠-凝胶电泳,根据蛋白质的分子大小来分离。采用这一技术,大肠杆菌细胞破碎后在电泳图上可分辨出 1000 种蛋白质的点子。虽然分出的量很少,但用荧光标记抗体法测定,相当灵敏,最低量可测出 1000 个分子。此法加以改进后,一次两向凝胶电泳,在理论上可分出 7000 个点。离体培养的 Hela 细胞,发生一个单一的基因突变后,仅影响到一个蛋白质的改变,在两向电泳中就可看出差别。这一方法为分析基因表达的调节控制反映在细胞内所合成的蛋白质的种类和数量的差异,提供了极为灵敏的手段。

4. 真核基因的插入顺序

真核生物的多数基因都是不连续的,存在有插入序列。这些插入序列在基因表达过程中的加工与剪接是近二、三年来最受关注的问题之一。

以胰岛素为例,胰岛素分子由 A、B 两条肽链组成,但实际上,生物体首先合成的是一个称为前胰岛素原的分子(即在 B 链之前有一段 24 肽称为前肽,B 链和 A 链之间有一段 35 肽,称为 C 肽),它需要经过两步加工,切去前肽和 C 肽,才形成胰岛素分子,把人胰岛素基因与胰岛素结构相比较,可看到这个基因。在 C 肽编码区的中间有 786 个碱基的插入序列^[14](图 5)。

为什么真核细胞的基因存在插入序列?它起什么作用?现在流行的说法认为,它的存在可

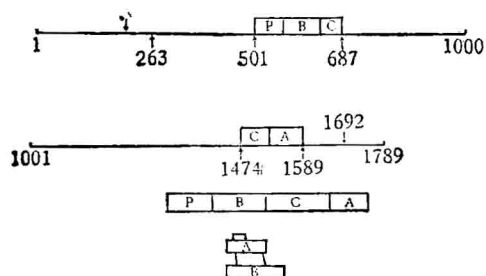


图5 人胰岛素基因简图

图中 P、B、C、A 分别为前肽、B 链、C 链和 A 链,其长度分别为 24、30、35、21 个氨基酸;* 所标位置的 DNA 序列为 TATAAAG,是 RNA 多聚酶结合位点,转录从核苷酸顺序号 263 处开始,在 688 至 1473 之间有一段 786 个核苷酸的插入序列,在 1692 处是信使 RNA 和多聚 A 的连接点,值得注意的是 C 链插入段,插在 C 链第 9 个氨基酸缬氨酸密码 GTG 的 GT 之间。

使蛋白质的合成分段进行,然后把几段连起来成为一个完整的蛋白质。这种分段合成方式,有助于加速真核细胞的进化速度。当基因发生一个变异后,它只影响一个肽段而不影响其它肽段,因此,在不同肽段各自发生单点变异后,这些肽段之间的不同排列组合就可以形成许多不同的产物。真核细胞进化很快,品种又繁多,看来是与插入序列的存在有关的。一些蛋白质插入序列常常分布在两个结构域中间。比如前面提到的激酶与脱氢酶,这两类酶分子中连接核苷酸的那部分结构域很相似,而连接底物的那部分结构域则很不一样,这样插入序列的存在可以使不同的结构域独立发生变异和进化,然后通过它们排列组合,增加了变异和进化的可能性。

不仅蛋白质的合成是分段进行,需要加工的,甚至连 tRNA 这样一种比较小的分子也是需要经过加工的。插入顺序的现象在真核细胞中是普遍存在的。但去年又发现了例外,这个例外恰恰又是线粒体^[5]。线粒体 DNA 编码的一些蛋白质,从目前已知情况来看,都是没有插入序列的。从这点看起来,线粒体又有点像原核细胞了。这个事实似乎是支持共生或寄生的学说的。

生 物 膜

从整个分子生物学的领域看来,有关生物

膜的研究是近年来发展最为迅速的领域之一。如果以在荷兰出版的。在国际上有影响的生物化学与生物物理杂志每年发表的论文数作为指标,有关生物膜的论文已经超过核酸、蛋白质、酶等分支领域而跃居首位。

在生物体内,除整个细胞包以外周膜外,细胞内部也有广泛的膜结构,如细胞器膜,内织网膜等。细胞间的相互识别和交通,细胞与外周环境的相互作用,细胞内部新陈代谢得以有条不紊的进行,都和膜有密切关系。现在知道,膜主要由磷脂和蛋白质构成。功能旺盛的膜,如线粒体膜,蛋白质含量高于脂质。反之,如神经鞘,则脂质含量高于蛋白质。对于膜的结构,现在一般认为所谓流动镶嵌模型(见图 6)能较好地反映多数实验结果。

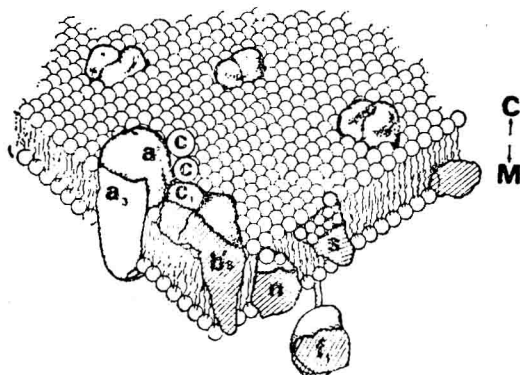


图6 线粒体膜一角的示意图

图中 C 和 M 分别代表线粒体膜的外侧及内侧。a、b、c 等都是和氧化磷酸化有关的酶系。它们的不对称分布是化学渗透学说的依据之一。

1. 蛋白质在生物合成中的前导肽段

除了和膜紧密结合的蛋白质以外,还有不少蛋白质在合成后,或者是输送到细胞外,或者输送到某一细胞器内以发挥其正常功能的。前者如胰岛素等激素蛋白,后者如线粒体、溶酶体等内部的多种酶,都必需在合成后穿透细胞膜以达到其最终目的地。现在知道这些蛋白质在生物合成过程中首先合成的是处于 N-末端的一段起引导作用的肽段。这个肽段通常大约含有二、三十氨基酸,其中疏水性氨基酸含量较丰,因此易于穿透膜中间的疏水区。在整个分子穿过膜以后,这一肽段通常经蛋白水解酶水解除去。前面提到的前胰岛素原,在 N-末端的一

段含有 24 个氨基酸的肽段 (图 5), 就是它的前导肽段。

2. 化学渗透学说

近年来, 生物能的研究受到了很大重视。一个人为维持生命每天平均需要的能量如果按二千大卡计算, 则全世界四十亿人口, 总消耗量相当于 4×10^{11} 瓦。如果把衣食住行所需的能量都加在一起, 大约要乘上十七、八倍 (各国消耗量不同; 此处按平均数估算) 即 7×10^{12} 瓦, 或 2.1×10^{20} 焦耳。这个数字看起来很大, 其实仅相当于地球从太阳接受的能量的二万分之一。当然, 地球表面大部分面积是海洋, 它所接受的能量很大一部分我们不能利用。但仅以照在陆地上通过光合作用所固定的碳的能量来算, 则也有 3.2×10^{21} 焦耳, 比现在的总消耗量还大十多倍。这就是生物能力学为什么受到重视的原因。

生物体的氧化磷酸化或光合磷酸化其基本过程是非常相似的, 对于这两种能量转化过程的机制, 在五十年代和六十年代先后提出了化学中间物学说和化学渗透学说。经过多年的争论, 虽然还不能完全定论, 但化学渗透学说显然已占优势。提出这一学说的 Mitchell 也因此获得了 1978 年度的诺贝尔化学奖。化学渗透学说的基本点认为^[16], 在氧化磷酸化过程中, 通过呼吸链酶系的作用, 将底物分子上的质子从膜的内侧传至外侧, 从而造成了质子在膜两侧分布的不均衡, 形成了质子梯度差。这个梯度差就是能量的来源, 通过腺三磷酶的逆反应, 把质子从膜的外侧再输回内侧, 来消除这一质子梯度差, 同时就合成了腺三磷。

3. 膜表面的受体物质

生物膜的另一重要功能是细胞膜内外的信息传递。现在知道在细胞膜表面, 广泛存在着—类称为受体的蛋白质, 而激素等物质对于细胞新陈代谢的调节作用, 就是首先通过和受体分子的特异性结合而实现的。许多药物的作用也是通过与细胞膜表面的受体的结合而实现的。例如吗啡就是通过和某些脑细胞表面的受体结合, 调节这些细胞的新陈代谢, 从而发挥其镇痛作用的。近年来通过对受体的研究, 不仅

发现了一些比吗啡更有效的药物, 还发现在高等动物脑中天然存在着—类比吗啡有更强镇痛作用的多肽物质。由于这类物质是脑中原来就有的, 因此这一类的研究有可能发展出既有镇痛作用而又不会像吗啡那样“成瘾”的药物来。

细胞膜的表面性质还对细胞的分裂繁殖有重要的调节作用。例如在创伤愈合过程中, 正常细胞的繁殖在细胞与细胞间发生接触后即行停止, 不会无限制地繁殖下去。这称为细胞分裂的接触抑制。但是对于肿瘤细胞, 这种接触抑制却不起作用, 因而造成肿瘤细胞的无限制的恶性繁殖。近年来对细胞表面受体物质研究结果表明, 在细胞癌变过程中, 表面受体物质的分布有明显变化。对细胞表面性质的进一步研究, 将会对肿瘤防治提供有价值的线索。

分子生物学在其他领域的渗透和应用

分子生物学的影响已经渗入到基础和应用生物学的每一个分支领域, 从而全面地带动了生物学的发展。分子生理学、分子细胞学、分子病理学、分子药理学等学科已为广大生物学家所熟悉。这里仅就个别重要的和尚未广泛为人注意的应用, 举一些例子。

1. 进化论

进化论的研究是经典生物学领域受到分子生物学影响的一个例子。过去主要依靠不同种属的生物体在形态、解剖方面的比较, 来决定它的亲缘关系。对于形态结构比较简单的如微生物, 它们的亲缘关系就较难确定。近年来随着蛋白质和核酸化学结构测定方法的进展, 对于在不同种属生物体中起相同作用的蛋白质或核

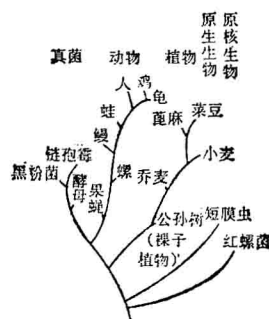


图 7 按细胞色素 C 氨基酸顺序决定的系统进化树

酸的化学结构进行了比较。例如对将近一百个生物种属(包括动物、植物、真菌、细菌等)的细胞色素c的氨基酸序列进行了测定,发现亲缘关系越近,在序列上的差异越小。因此根据它们在结构上差异的程度,可以断定它们在亲缘关系上的远近,并且可以由此得到反映生物进化的系统进化树(图7)。它与按经典方法得到的进化树基本符合^[17]。采用分子生物学的方法来研究分类和进化有这样一些好处:首先,生物大分子的结构反映了生命活动更为本质的方面;其次,根据结构上差异的程度可对亲缘关系给出一个定量的,因而也是更准确的概念;第三,对于在形态结构上非常简单的微生物,只有用这种方法才能得出可靠的结果。例如现在已经根据核糖体15S RNA的核苷酸序列,系统地研究了原核生物的进化^[18],不仅对它们之间的亲缘关系,第一次有了比较全面的了解,并且对原核生物的进一步分化进化,真核生物的起源,都提供了宝贵的资料。值得在此提出的是过去曾经认为是处于原核细胞和病毒之间的支原体,现已证明是梭菌一属的近亲。

2 生态学

生态学方面的一些问题也已经开始从分子水平进行研究,从而形成了分子生态学。例如澳大利亚的一种昆虫,虽在气候条件类似的南美地区也不能繁殖。现已查明,这种昆虫需利用其所食植物中的一种甾体,经过三步酶反应合成其变态所必需的一种激素,而南美地区的同科但不同属的植物中不含此种甾体。于是昆虫无法合成此种激素,因而无法繁殖。又如,某些田鼠越冬能力的强弱,和它们体内糖代谢体系中某些关键性的酶(磷酸葡萄糖变位酶)的催化效率强弱有密切关系^[19]。高效率的糖代谢体系,保证了体内贮存食物的充分利用。

3 老年学

生物体为什么会衰老?有一种看法认为与转录、转译等过程中的错误有关。例如曾经对老年与幼年的大白鼠的醛缩酶进行了比较,发现老年的酶虽然其免疫性质没什么变化,但催化活性却明显地降低了。这可能是由于老年动

物体内转录和转译发生了误差,合成出的酶分子在某些关键部位的氨基酸序列有错误,因此影响了酶的催化活性。

在生物体内,复制、转录和转译都是会出现错误的,不过几率很小。例如DNA的复制在体内出现错误的几率大约是 10^{-9} ,在体外大约是 10^{-5} ,这是由于机体内有一套纠正错误的校正机制。

在DNA分子内腺嘌呤(A)应该是与胸腺嘧啶(T)配对的,但是在腺嘌呤1位氮原子和6位碳原子之间有一个双键,它可形成类似烯醇式的结构,(图8下方左边腺嘌呤结构式)这种形式是非常不稳定的,存在的比例是极

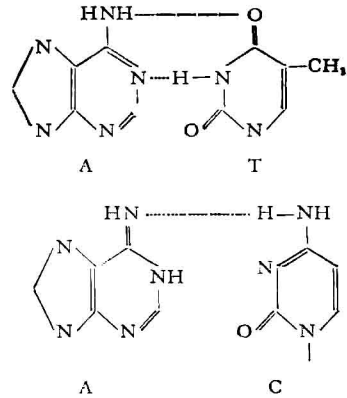


图8 腺嘌呤的异构化及碱基配对

低的,但如处于这种形式的结构就不能与T配对,反而比较容易与胞嘧啶C配对(图8)。如果A的这种异构现象正好在复制或转录的时候发生,就会出现错误。这样的产物由于碱基配对出现了问题,该处的螺旋就不紧,从而机体内的酶系可以对此进行校正和修复。即内切酶识别和切开错误配对的地方;外切酶再把错误的一段核苷酸切去;露出的空档由聚合酶修复,最后通过联结酶重新接好。生物体到了老年的阶段很可能是由于这一套校正和修复机制运转不正常,因此才造成合成产物不正常的结果。应该指出,分子老年学还是一门新学科,许多问题还有待于进一步研究。看来蛋白质生物合成迟缓,合成产物错误是比较普遍的现象,但出现错误的原因还有不同看法^[20]。

4. 基因工程

这几年有好几个蛋白质都通过基因工程的方法在实验室里搞出来了,其中包括胰岛素、干扰素、生长激素等。现在国外有好几家公司在建厂,据估计到八十年代末期,仅美国一个国家基因工程产物的总产值将达到五亿美元。其中主要的兴趣是在干扰素上,希望能用它来治疗癌症和病毒病。干扰素有种族专一性,动物的干扰素不能代替人的,而人的干扰素如果从白细胞中提取,则产量有限;当然也可以搞体外细胞培养,但也不如用基因工程的方法优越。由于现在仅仅开始拿到一定量的干扰素来进行初步的临床试验,所以干扰素能否对癌症及病毒病有良好的疗效,还不能最后下断语;若是确有疗效,则将是分子生物学对人类的又一重大贡献。

胰岛素的基因工程如果把整个胰岛素基因通过载体送到大肠杆菌中去合成的话,合成出来的将是前胰岛素原,如何切去前肽和C肽将遇到一系列问题。因此现在的做法是人工合成A链的基因和B链的基因,分别把它们送到大肠杆菌中去,得到的A链和B链通过重组,合成胰岛素。值得提出的是,这一工作中A、B链重组的条件是我国首创的。

附带提一下,对于基因工程,我的理解既然称为工程,就该体现工程的特点,按照设计的蓝图,进行施工,拿出预期的成品,具体来说,就是把一个生物体中携带的某一特定的遗传信息,通过一定的方法转移到另一生物体中,使之获得前者的遗传特征,就叫基因工程。如果事先并无明确的目的,盲目地引入一些遗传物质,搞出什么就算什么,恐怕是不能称之为工程的。

5. 肿瘤

近年来戒烟的人逐渐多起来了。我这里再来举一个这方面的例子,促进一下更多的同志

下决心戒烟。

已经知道吸烟或者是空气中的多环芳烃类物质与肺癌有关系。实际上多环芳烃本身并不致癌,但它进入体内后,在芳烃羟化酶的作用下,在芳烃环上加上了一个羟基或是环氧基,就变成了一个强烈的致癌物。芳烃羟化酶每人体内都有,但含量高低可以相去甚远。因此肺癌发病不仅和吸入芳烃有关,也和体内羟化酶含量有关。但即使对于羟化酶含量较低的同志,大量吸烟,仍然会极大地增加了生肺癌的危险性。由于大城市空气污染,不吸烟的人也会吸入一定量的芳烃,对于羟化酶高的人也可能导致肺癌。

本文经施建平同志帮助整理,特此致谢。

参 考 文 献

- [1] Fowler, A. V. et al.: *J. Biol. Chem.*, **253**, 5521, 1979.
- [2] Fiers, W. et al.: *Nature*, **260**, 500, 1976.
- [3] Sanger, F. et al.: *ibid.*, **265**, 685, 1977.
- [4] Houghton, M. et al.: *ibid.*, **285**, 536, 1980.
- [5] Walker, J. E. et al.: *Eur. J. Biochem.*, **108**, 549, 1980.
- [6] Douzou, P.: *Quart. Rev. Biophys.*, **12**, 521, 1979.
- [7] Makinen, M. W. et al.: *Ann. Rev. Biophys. Bioeng.*, **6**, 301, 1977.
- [8] Gaudour, R. D. et al.: eds. *Transition States of Biochemical Processes*, Plenum Press, New York, 1978.
- [9] Richard, J. et al.: *Eur. J. Biochem.*, **49**, 195, 1974.
- [10] Khorana, H. G.: *Science*, **203**, 614, 1979.
- [11] A. H. J. Wang. et al.: *Nature*, **282**, 680, 1979.
- [12] Laskey, R. A. et al.: *ibid.*, **286**, 763, 1980.
- [13] Garrels, J. I.: *TIBS*, **5**, 281, 1980.
- [14] Bell, G. I. et al.: *Nature*, **284**, 26, 1980.
- [15] Rasti, E. et al.: *Cell*, **18**, 501, 1979.
- [16] Mitchell, P.: *Science*, **206**, 1148, 1979.
- [17] Fitch, W. M.: *J. Mol. Evolution*, **8**, 13, 1976.
- [18] Fox, G. E. et al.: *Science*, **209**, 457, 1980.
- [19] Brown, A. J. L.: *Nature*, **269**, 803, 1977.
- [20] Hoeman, G.: *Intern. J. Biochem.*, **12**, 515, 1980.

[本文于1980年12月24日收到]