

生物科学参考资料

第三集

科学出版社

生物科学参考资料

第 三 集

科 学 出 版 社

1 9 7 3

内 容 提 要

本集翻译了新近国外杂志上一些有关生物科学当前研究方向的论文。一共有14篇论文,包括理论性的综述;如第一篇以蛋白质的结构和历史来说明蛋白质的种间差异,提供一个十二亿年来的分子进化记录。第二篇以仿生学的观点来说明细胞依靠激素和神经纤维进行通信。又如“生物学个性的标志”一文,说明进行器官移植时,身体要有辨别外来的和自身的细胞的能力,每一个体的细胞具有独特的蛋白质标志。“丘脑下部的激素”一文说明垂体前叶接受发生于脑中的‘释放因子’的调节,并已分离出以人工合成了这类丘脑下部激素中的两种,证明其中一种可用来治疗不育症,另一方面,从这种人工合成的激素分子结构,打开一条节育的途径,制得其拮抗物质用作避孕药。“植物力能学的世界图式”是植物生态学方面的文章。“绘制人类染色体图”是探测一个基因和另一个基因的位置关系,同时在23对染色体中确定携带它的那一个染色体,进行此工作主要是应用统计学的技术。“干扰素的诱导”是介绍干扰素抵抗病毒感染的综述文章。“生物固氮的前景”和“土壤污染”二文对目前生产上的实际问题,也有所阐明。本集对国内生物学研究工作者有参考价值,对广大读者了解生物学科的动向和增长知识也有所帮助。

生 物 科 学 参 考 资 料

第 三 集

*

科 学 出 版 社 出 版

北京朝阳门内大街 137 号

中 国 科 学 院 印 刷 厂 印 刷

新华书店北京发行所发行 各地新华书店经售

*

1973 年 12 月 第 一 版 开本: 787 × 1092 1/18
1973 年 12 月 第 一 次 印 刷 印张: 10 2/9 插页: 1
印数: 0001—6,950 字数: 240,000

统一书号: 13031 · 153

本社书号: 273 · 13—6

定 价: 1.10 元

目 录

一种古老蛋白质的结构和历史	1
细胞的通信	18
生物学个性的标志	30
丘脑下部的激素	44
人体肌肉的收缩是如何调节的	56
分子生物学新动向	68
环一磷酸腺苷	76
生物力能学与肿瘤生长问题	87
植物力能学的世界图式	101
绘制人类染色体图	115
干扰素的诱导	129
生物固氮的前景	138
土壤污染	154
扫描电子显微镜	167

一种古老蛋白质的结构和历史

Richard E. Dickerson

“从酵母到人类氧化食物分子都需要一种细胞色素 *c*。此种蛋白质的种间差异提供一个十二亿年的分子进化记录。”

在十五亿到二十亿年前之间,在我们行星上的一些单细胞有机体发生了一个意义深远的变化,正是有助于多细胞有机体兴起的一个变化。这样一个用氧与食物分子化合的机构产生,使得从食物中可以比以前提取更多的能量。这个新的代谢机构的主要成分之一是细胞色素 *c*,它是一种蛋白质,今天其子孙后代能够在每一个有细胞核的生活细胞中发现。研究了从各种有机体提取的细胞色素 *c* 之后,有可能断定自从植物和动物分成两界以来,这种蛋白质的进化速度怎样,事实上也有可能给该变化提供十二亿年前这样一个近似年代。例如人类和黑猩猩的细胞色素 *c* 分子是完全相同的,有 104 个氨基酸连在一起,同样的顺序排列,卷曲成同样的三维结构。而人的细胞色素 *c* 与粗糙链孢霉 (*Neurospora crassa*) 的细胞色素 *c* 在 104 处中有 44 处相异,但是两种细胞色素 *c* 分子的三维结构实质上是相似的。我们现在能够解释为什么细胞色素 *c* 的 104 个氨基酸中很多是可变换的,也能解释为什么某些氨基酸完全不能改变,否则将破坏这种蛋白质的活性。

让我们尝试设想细胞色素 *c* 初次出现之前的地球。在这个行星上最早的有生命的有机体仅比一些食腐动物强一点,即从其周围的水中吸取富有能量的有机物(包括它们的邻居),放出能量低的分解产物。在普遍的无氧发酵过程中现在仍然存在这种生活方式的遗迹,如一个酵母菌由糖中提取能量而放出乙醇,或是一个运动员在非常激烈的运动时把葡萄糖转变成乳酸,引起肌肉痉挛。无氧发酵是所有生物的一种共同的生化遗产。

地球上只靠发酵过程作为能源能够支持多少生物的最高限度,决定于非生物方法合成高能化合物的速度。非生物方法指紫外线辐射,闪电,放射性或热等。当有些生物发展到能利用日光作为能源时,光合作用产生了,地球养活生物的能力大大增加。这是细菌和蓝藻的年代(见“The Oldest Fossils” Elso S. Barghoorn 著; *Scientific American*, 1971, 5 月)。

更先进形式的光合作用放出了一种腐蚀性有毒的气体到大气中,就是氧。某些细菌的反应是退缩到地球上的无氧角落,今天在那儿能够找到它们的子孙后裔。另一些细菌和蓝绿藻产生一些方法,用其自身的废物和氧化合,以中和气态氧。下一步

是利用这些废物的氧化所放出的能量(如果你要烧你的废品,你不妨用它所生的火取暖),这是氧化过程(或呼吸)的开始,是地球上对生命有用的能量增加供应上的第二个大突破。

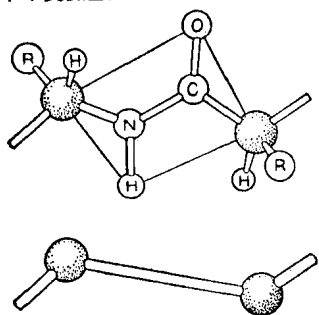
当一个酵母细胞把糖氧化成二氧化碳和水,而不是停止在乙醇,就使它从每克燃料多得 19 倍的能量。运动员的肌肉中的乳酸和氧化合后,痉挛就消失了,他由他的葡萄糖中得到相当大的能量偿还。使有用能量的供应增大这样显著时的任何代谢上的改进,都会在生命的发展上有一个变革性的效果。现在我们相信,细胞的分化和多细胞动植物的出现仅在有这样大的新供能条件下才能发生。

细菌和蓝绿藻是原核体(前核细胞),其遗传物质 DNA 不是局限在一个有组织的细胞核内,它们的呼吸和光合作用机构(如存在的话)同样是分散的。绿藻及所有高等动植物是真核体(有真核的细胞),其 DNA 在一个细胞核内,在叫做线粒体的细胞器内进行它们的呼吸。在真核体植物,光合作用在另一种叫做叶绿粒的细胞器中进行。线粒体是所有真核细胞的“发电厂”。线粒体的作用是分解来自食物的含能量多的分子,使它们和氧化合,利用产生的能量以合成 ATP 分子储存起来。所有真核细胞的线粒体的化学结构是相同的,好象一旦最适宜的化学机构建立起来,就不再改变了。

至少有 20 种专一的酶参与生物氧化,它们起初作为来自食物分子的电子或氢原子的受体,后来又作为供体。在生物氧化过程的最后部分,可以发现一系列的细胞色素分子(用附书的英文字母区别不同的细胞色素),它们都与一个含铁的血红素结合,血红蛋白中也有同样的血红素。电子传递通过一系列的细胞色素分子:从细胞色素 *b* 到细胞色素 *c₁*,从细胞色素 *c₁* 到细胞色素 *c*,到细胞色素 *a* 和 *a₃*,最后到氧原子上,在该处氧原子与氢离子化合生成水。这是一个渐进的过程,一小部分一小部分地释放能量而不是一下子放完。电子由细胞色素 *b* 转移到细胞色素 *c₁*,以及在细胞色素 *a₁*、

图 1 细胞色素 *c* 分子的骨架(彩色图)

每一种利用氧进行呼吸的生物,其细胞内部有一种这类蛋白质分子。此图以简单形式阐明 104 个氨基酸怎样连接成一个紧紧包围一正铁血红素基团(一个复合花束,中间有一个铁原子)的连续的长链。图中不同颜色表示在进化过程中 104 个氨基酸位置中的每一个位置能有多少变异。有些种属没有第 104 个氨基酸,脊椎动物以外的所有种属在链的起点还有额外的氨基酸,有多到 8 个的(见表 1)。在进化过程中最不易改变的(可能也是最重要的)氨基酸用红和橙色表示,有改变的氨基酸按其出现的氨基酸种类的数目,分别用黄绿,蓝绿,蓝,紫表示,必需的正铁血红素基团用玫瑰色表示。每一个氨基酸只画出它的 α -碳原子(带有一独特支链的原子)。左下方的小图上部说明两个氨基酸怎样连接成一个肽键,连到 α -碳的支链以标有“R”的圆球表示。小图下部是在大图中细胞色素 *c* 骨架所使用的图式。所有的肽键($-\text{CO}-\text{NH}-$)都省略了,只画了与正铁血红素接触的那些支链。35 个不变氨基酸的位置,都注明了氨基酸的缩写。



Ala	丙氨酸	Leu	亮氨酸
Asp	天门冬氨酸	Lys	赖氨酸
Asn	天门冬酰胺	Met	蛋氨酸
Arg	精氨酸	Phe	苯丙氨酸
Cys	半胱氨酸	Pro	脯氨酸
Gly	甘氨酸	Ser	丝氨酸
Glu	谷氨酸	Thr	苏氨酸
Gln	谷氨酰胺	Trp	色氨酸
His	组氨酸	Tyr	酪氨酸
Ile	异亮氨酸	Val	缬氨酸

a_2 和氧之间转移的时候,放出的能用来合成ATP,ATP是细胞代谢的一个通用的能原。

大多数细胞色素是与线粒体膜紧密连接的,但是其中之一,细胞色素 c 易溶于水介质中,并能分离出纯品。其他成分可以作为多酶复合体分离出来: b 和 c_1 ,作为一种细胞色素还原酶复合体, a 和 a_3 作为一种细胞色素氧化酶。还原酶把电子给细胞色素 c ,氧化酶再把电子接受过来。以下事实可以说明所有的真核体彼此是多么相似,即发现了任何种属的植物、动物或真核体微生物的细胞色素 c 在试管内都能和任何一种属的细胞色素氧化酶发生反应。蠕虫或灵长类,鲸或小麦在线粒体膜下都是完全一样的。

细胞色素 c 的进化

因为细胞色素 c 是这样古老的蛋白质,同时是如此之小,并且易于纯化,对进化过程有兴趣的蛋白质化学家非常注意它。已经测定了 40 种以上真核体生物的细胞色素 c 的全部氨基酸顺序。在图 1 和表 1 中对其中 38 种的顺序进行了比较。我们掌握细胞色素 c 分子进化的资料比任何其他蛋白质的都多。西北大学的 Margoliash 和加利福尼亚大学的 Smith 在落杉矶最先发现不同种生物的氨基酸顺序是有差异的,并且其差异的程度与两种生物在进化树中相距的远近是互相对应得相当好的。Margoliash,威斯康辛大学的 Fitch,还有其他研究人员对这些差异进行了精细的电子计算机计算和分析,推论出生物的精族系树,完全不用传统的解剖学论据。这种族系树与根据经典形态学所得的结果是一致的。显然对氨基酸顺序进行比较是研究进化过程的一种有力工具。

另一个结果最初看来可能是惊人的。细胞色素 c 仍在很慢的进化,其速率如在一地质学时期内平均计算是对各种生物近似恒定的。十年前加利福尼亚理工学院的 Pauling 和 Zuckerkandl 最先作出血红蛋白的分子进化的这种分析。如对血红蛋白和细胞色素 c 进行比较,可发现细胞色素 c 的变化更慢得多。这是为什么呢?蛋白质链的合成是来自 DNA 的指令,突变的发生是在 DNA 内。是不是造血红蛋白的 DNA 比造细胞色素 c 的 DNA 更常发生突变呢?没有理由这样想。所以解释一定在于自然选择或筛选过程,以试验是否突变分子能够起其应起的作用。

在讨论经过考验的形成一个有效的细胞色素 c 的各种“化学式”之前,还要简短描述一下蛋白质的结构。所有蛋白质分子都是由首尾相连的氨基酸构成的。20 种氨基酸的每一种一端都有一个羧基($-\text{COOH}$),另一端有一个氨基($-\text{NH}_2$)。把一个

图 2 带支链的细胞色素 c 分子

通过 X-射线结晶衍射测定的马心细胞色素 c 在氧化状态下的结构。 α -碳原子标明了数字,连接 α -碳原子的肽键($-\text{CO}-\text{NH}-$)只以实键表示。为了清晰,去掉了分子“背后”的三个支链:亮氨酸 35,苯丙氨酸 36 和亮氨酸 98。从三维结构中能够看出围绕着正铁血红素的分子内的支链是疏水的(红和桔黄),亲水支链(蓝或紫)的氨基酸在外部,在外面它们通常是与水接触的。一个重要例外是苯丙氨酸 82 的疏水支链,它位于正铁血红素左侧的分子表面。亲水支链之间的区域,在异亮氨酸 81 之上,是一个空穴,它显然是对溶剂分子开放的。当细胞色素 c 被氧化的时候,在这个空穴之上的赖氨酸 13,与一个大的氧化酶复合体相互作用。

表1 38种生物中细胞色素c的组成

(没有另一种蛋白质在这么多生物中作了这样详尽的分析,按氨基酸的化学性质分成类别,分别作出下列的记号)

疏水,芳香环: F. 苯丙氨酸; W. 色氨酸; Y. 酪氨酸。疏水,非芳香: I. 异亮氨酸; L. 亮氨酸; M. 蛋氨酸; V. 缬氨酸。亲水,碱性: H. 组氨酸; K. 赖氨酸; X. 甲基化赖氨酸。双性的,疏水很弱,极性不荷电: A. 丙氨酸; B. 天门冬酰胺或天门冬氨酸; C. 胱氨酸; N. 天门冬氨酸; P. 脯氨酸; Q. 谷氨酰胺; S. 丝氨酸; T. 苏氨酸; Z. 谷酰胺或谷氨酸。无支链,一个氢原子: G. 甘氨酸。链的起点有“a”,表示有一个甲基(CH₃)与氨基末端接连;链的起有“h”,表示没有甲基。

	-9	-5	-1	1	5	10	15	20	25	30	35	40	45							
人、猩猩																				
恒河猴																				
马																				
牛、羊、猪																				
狗																				
兔																				
抹香鲸																				
大灰袋鼠																				
鸡、火鸡																				
鹌鹑																				
北京鸭																				
龟																				
蛇																				
牛蛙																				
金枪鱼																				
小蜜鱼																				
野生蓖麻蚕																				
菸草天蛾																				
金蝇																				
果蝇																				
面包酵母																				
假丝酵母																				
链孢霉																				
小麦胚																				
荞麦种子																				
向日葵种子																				
绿豆																				
花椰菜																				
西葫芦																				
芝麻种子																				
芝麻子																				
棉花种子																				
蓖麻子																				
苘麻种子																				
氨基酸种类的数目	1	3	5	5	1	3	3	4	4	2	3	2	1	3	3	6	1	2	3	1

表 1 (续)

哺乳动物	人、猩猩、恒河猴、马、驴、猪、牛、羊、狗、兔、香鼠、抹香鲸、大灰袋鼠	50	55	60	65	70	75	80	85	90	95	100	105
其他脊椎动物	鸡、火鸡、鸭、鹅、北京鸭、电鳗、牛蛙、金枪鱼、小鲨鱼	50	55	60	65	70	75	80	85	90	95	100	105
昆虫	野生蓖麻蚕、菸草大蛾、金蝇、果蝇	50	55	60	65	70	75	80	85	90	95	100	105
低等植物	面包酵母、假丝酵母、链孢霉	50	55	60	65	70	75	80	85	90	95	100	105
高等植物	小麦胚、荞麦种子、向日葵种子、花椰菜、西葫芦、西葫芦、芝麻种子、芝麻子、槐花种子、槐花子	50	55	60	65	70	75	80	85	90	95	100	105
氨基酸种类的数目	25	31	37	45	52	58	65	72	80	88	96	104	112

氨基酸的羧基和另一个氨基酸的氨基连接起来,一定要去掉一个分子的水,产生一个肽键($-\text{CO}-\text{NH}-$)。因为一个氨基酸只有一部分(当然是有决定意义的部分)进入一个蛋白质链,化学家把它叫做“残基”。因此常说一个甘氨酸残基或一个苯丙氨酸残基在蛋白质链上某一个部位。

邻近肽键的碳称为 α -碳,它是很重要的,因为每一个氨基酸在该部位都有一个独特的支链。支链也可能就只有一个氢原子(如甘氨酸就这样),或者可以有很多原子,如包括有一个六碳的“芳香环”(如苯丙氨酸、色氨酸和酪氨酸)。

20种氨基酸根据它们支链的特点能分成三类(见图2):5种是亲水的,放入水溶液中,它们就带上一个正电荷或是一个负电荷。5种中有3种是碱性的(精氨酸、组氨酸和赖氨酸),另两种是酸性的(天门冬氨酸和谷氨酸)。7种不溶于水,因此称为疏水的,它们是上面提到的支链上有环的3种氨基酸,加上亮氨酸、异亮氨酸、蛋氨酸和缬氨酸,其余8种氨基酸对水的反应是双性的,有丙氨酸、天门冬酰胺、半胱氨酸、谷酰胺、甘氨酸、脯氨酸、丝氨酸和苏氨酸。

现在让我们看一看有多少有成效的细胞色素 c 化学式因种属不同而有差异。人和马的细胞色素 c 分子的104个氨基酸中有12处不同。高等脊椎动物——哺乳类、鸟类和爬行类——与鱼类的细胞色素 c 平均有19处氨基酸不同。脊椎动物和昆虫的细胞色素 c 平均有27处氨基酸不同;还有,昆虫和植物的细胞色素 c 分子在链的开始处比脊椎动物的多几个氨基酸残基。两个细胞色素 c 之间相异最大的是人和链孢霉之间的差异;它们有40%以上位置的氨基酸不一样。具有这样大的差异的两个分子怎么能执行同样的化学机能呢?

当我们注意这些变化在什么部位的时候开始找到一个答案。氨基酸顺序的某些部分,如图2所示,是不变的。在已知种属的细胞色素 c 中,104个氨基酸位置中有35个是完全不变的,其中包括一个从残基70—80的长顺序。35个位置为15种氨基酸占据着,在图1中以红色显示。另有23个位置只为两种不同但很近似的氨基酸中的一个所占据。在这23个位置上有18套可变换的氨基酸对。在图1中以橙色显示。有17个位置,自然选择显然只接受三种氨基酸组成的一套氨基酸,在图1中以黄绿显示。

在X-射线结构分析之前,由化学顺序研究中已知,在可替换的位置上能够替代的氨基酸几乎总是有同样的化学特性。一般说来,必须都是亲水的,或疏水的,或者是对水中性的。这种替换叫做保守的替换,因为它们保存了蛋白质链那一部分的总的化学性质。

链上只有几个位置能够允许较大的改变。例如残基89能是酸性的(天门冬氨酸或谷氨酸),碱性的(赖氨酸),极性但不荷电的(丝氨酸、苏氨酸、天门冬酰胺和谷酰胺)弱疏水的(丙氨酸)或无支链的(甘氨酸)。分子中这一位置表现严格禁用的一种支链是一种大的疏水的支链。这样的“无关重要区”是很少的,所以总的说来细胞色素 c 是一个进化上保守的分子。♥

我们没有理由去想细胞色素 c 的基因比血红蛋白的基因发生突变更慢,或是顺

序上不变的、保守的和变化较大的区域反映细胞色素 *c* 基因内的突变率有任何差异。突变大概是随机的, 在这些种间对比中我们所看到的是经过生存能力的严格考验之后留下来的分子。不变部位不改变的原因显然是由于这些部位的任何突变都是致死的, 因而被淘汰了。别处只要保存分子的主要化学性质能耐受保守的变化。根本的改变大体指出与蛋白质起作用无关紧要的部分。

这是单从顺序对比尽我们可能作的解释, 按照分子不同部位的主要特性和非主要特性对变异性作的解释, 是合理的, 但是科学常常是被合理而又不正确的假说所折磨。要前进一点, 需要知道氨基酸顺序怎样卷曲成为一个起作用的分子。简而言之, 我们需要蛋白质的三维结构。

分子的三维结构

作者与 Margoliash 共同研究, 在 1963 年开始作马心的细胞色素 *c* 的 X-射线结晶学分析。当细胞色素 *c* 在线粒体内转移电子时, 它在一种氧化型(正铁细胞色素)和一种还原型(亚铁细胞色素)之间互变, 正铁血红素中的铁原子交替地在 +3 或 +2 的氧化状态。我们决定用氧化型开始分析, 这主要是一个策略上的决定, 因为如果我们要辨认电子转移过程, 两种氧化状态毕竟都是需要的。

在 X-射线结晶学, 把一束 X-射线对准一个要研究的物质的纯晶体, 记录下光束由各种角度射击样品时产生的衍射图样。进入样品 X-射线本身由于结晶内电荷的分布而偏斜各种角度。设计了高度复杂的电子计算机程序, 由数以万计的 X-射线衍射数据推算出电荷的三维分布。由这种分布又能推断出蛋白质分子中氨基酸支链的分布。

五年前我们得到氧化型马心细胞色素 *c* 的最早的低分辨图, 三年前得到高分辨图。这些图已用于构造蛋白质的精细三维模型。也能把三维空间座标输送给计算机, 得到简单的可以用投影放大法观察的球、棒式的图, 使我们能看到三维的蛋白质卷曲着的链(参见图 5、图 6)。就在一年前, 我们计算了还原型细胞色素 *c* 第一个高分辨图。现在我们正在改进这个模型, 并对两种氧化状态进行比较。

当我们解决了最早的高分辨结构的时候, 我们注意到细胞色素 *c* 的氨基酸顺序的几种显著特点。我们知道整个进化中最顽固保守的位置是为三种特殊类型的残基占据: 荷正电的(碱性的)赖氨酸残基; 三个疏水的芳香族的残基——苯丙氨酸、色氨酸和酪氨酸的残基; 四个疏水的非芳香族的残基——亮氨酸、异亮氨酸、蛋氨酸和缬氨酸的残基。借助于图 2 和图 3, 能够找出这些部位。图 1 和图 2 中颜色, 用红和橙色表示疏水的残基, 用绿、黄、蓝和紫表示中性残基和碱性及酸性亲水残基。

从该分子的氨基酸顺序的化学分析中得知碱性的和疏水的基团倾向于沿着链群集。例如在 22—27, 38 和 39, 53—55, 86—91 这些区域找到碱性的残基。在 9—11, 32—37, 80—85, 94—98 这些区域找到疏水的残基。14 和 17 位置(半胱氨酸)的

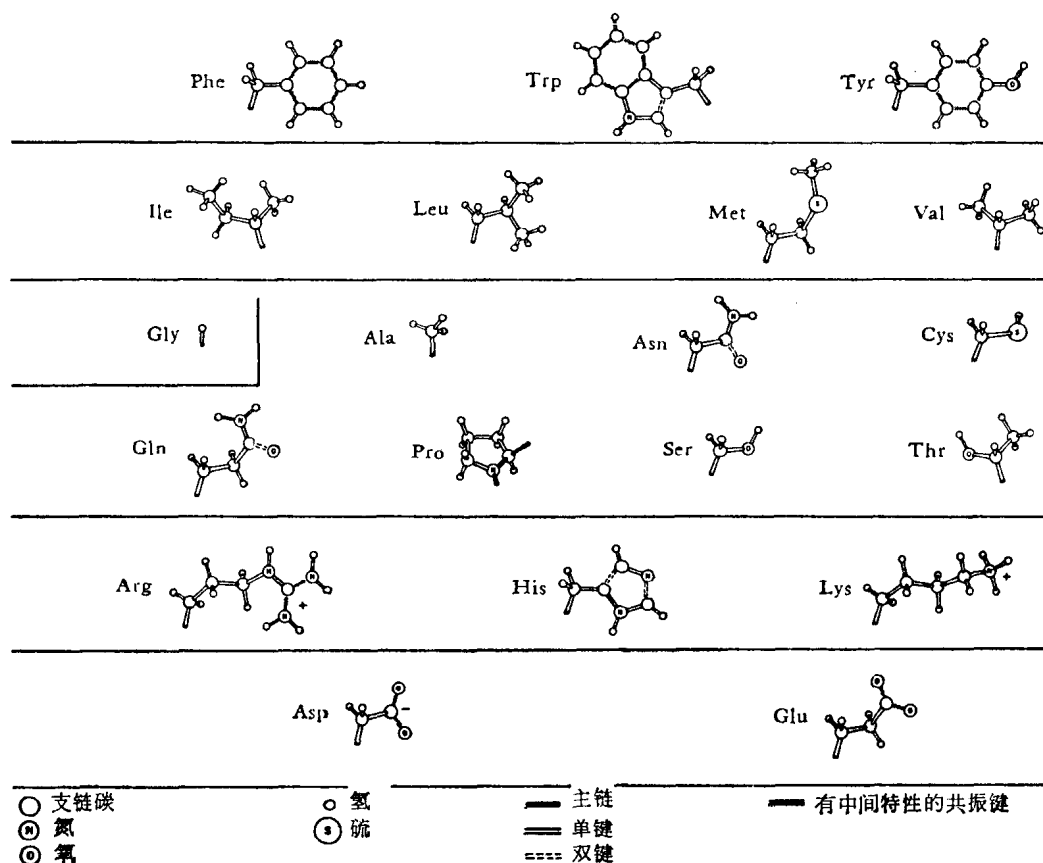


图3 20种氨基酸支链的结构
 (图中的各种氨基酸英文缩写见图1的注解)

残基, 18位的残基(组氨酸)是不变的, 这是能够理解的, 因为他们形成连到正铁血红素基团上的键。还不怎么能理解的是从70—80的长区段同样是不变的。在有结构上的证据之前, 曾猜测蛋氨酸(在80处)可能在血红素(一侧是组氨酸18)的另一侧连到铁原子上, 但是单从化学证据来看, 不可能十分肯定。

从化学分析中也知道马的细胞色素c有12个甘氨酸, 这些甘氨酸或是不变的或是保存在绝大多数的种属中。也还知道在8个苯丙氨酸或酪氨酸(支链上有芳香环)中有7个或是在所有种属中不变, 或是只能以另一个替代。残基36, 苯丙氨酸或酪氨酸在三种生物中以异亮氨酸替代, 它的支链虽然是非芳香族的, 但至少是和它替代的支链一样大, 并也是疏水的。

所有这些类似点和保守性在X-射线分析之前就知道了, 但是不能用结构来解释。设想每一个残基是由自然选择把它放在所居的部位, 每一个残基包含关于细胞色素分子功能部分的重要信息。自然选择不作用于一个氨基酸顺序, 而是作用于这个卷曲的并与其他生物分子结合而起作用的分子。有顺序而没有卷曲的指令, 好象是有另件的清单而没有整个机器的蓝图一样。

细胞色素 c 和进化

现在有了细胞色素 c 的蓝图, 让我们更仔细地看看图 1。为了使图式简单, 不画上支链, 只保留与正铁血红素基团相连接的那些支链。此外, 沿着主链只描画出 α -碳原子(当有支链时, 从此分支伸出)。连接 α -碳的肽键($-\text{CO}-\text{NH}-$)只以直线简单表示。所以该图是细胞色素分子的一个简化的卷曲图形。

我们看见扁平的正铁血红素基团(一个对称的碳原子和氮原子的花束, 其中心有一个铁原子)居于一个裂隙内, 只有一边向外界暴露。如果血红素直接参与电子出入分子内外, 转移就可能沿着这个边缘发生。如图所示半胱氨酸 14 和 17, 组氨酸 18 从右侧使正铁血红素就位, 另一个从左侧束缚血红素的基团就是蛋氨酸 80。

从早期的用 X-射线对蛋白质的研究已知氨基酸顺序常常卷曲成螺旋构型称为 α -螺旋; 有时许多氨基酸呈现一种波浪状或皱纹构型称为 β -片。细胞色素 c 没有 β -片, 只有两个 α -螺旋的区段, 由残基 1—11 和残基 89—101 形成。该蛋白链的绝大部分紧紧地缠绕在正铁血红素的周围, 为在其他蛋白质中显著的 α -和 β -构型只留下很小的空间。

正如人们能利用细胞色素 c 来研究进化, 人们也能利用进化来研究细胞色素 c。前已提到, 图 1 指出每一位置可以接受氨基酸种类的变异量。细胞色素分子中机能的很重要的位置, 在种属间没有差异或很少, 而在种属间变化很大的位置大概是对细胞色素 c 这个能生存分子不怎么重要。

正铁血红素裂隙是不变的, 说明强大的选择压力要保持正铁血红素基团环境在整个进化过程中的恒定。不变的残基 70—80 卷曲而成分子的左侧以及正铁血红素所居的空穴。分子的右侧的变化是中等的, 只能接受 1、2 或 3 种不同的氨基酸。分子的背侧是多变的, 对于残基 58 和 60, 以及 α -螺旋背侧的另 4 个残基, 他们中每一个位置在各种属中可为 6 种以上的氨基酸占据。这些是强有力的线索, 指出分子中是否为电子转移或者为两个大的分子复合体(还原酶和氧化酶)相互作用有重要作用的部分。

分子是怎样卷曲的

如果我们现在再看图 2(图中显示马心细胞色素 c 的所有支链), 就能够理解很多进化上的保守性(和前面一样, 肽键仍然仅画直线, 它们的原子位置是了解的, 但是与本文不是特别有关的)。此图中根据所接受氨基酸的特性选用各种颜色把不同的位置分类(亲水的, 疏水的或双性的); 所用的颜色是: 红色表示疏水的、带芳香环的氨基酸, 橘黄色表示疏水的、非芳香的氨基酸, 蓝色表示亲水的、碱性的氨基酸, 紫色表示亲水的、酸性的氨基酸, 绿色表示双性的氨基酸(弱疏水的、极性但不荷电的), 黄色表示无支链的(有氢原子)甘氨酸。

非极性疏水的一类大多在分子的内部,远离外面的水环境;荷电的(酸性或碱性的)总是在外面。这种安排是一个卷曲蛋白质的“油滴”模型的好样品。按照这种模型,当在细胞内部合成一个氨基酸链的时候,疏水的(或“油的”)支链尽可能远离水环境、群集在分子中心的自然趋势,有助于链的合适卷曲。还能作一个更有力的叙述:如果蛋白质分子执行功能需要多肽链的某些部分向内卷曲,那么自然选择就会促进在该点上保留疏水支链,于是合适的卷曲就形成了。一个荷电的(或是亲水的)支链能被推进蛋白质分子的内部,但是要付出相当大的能量代价。在大多数情况,在某一部分有一个荷电的基团有助于保证链在该点上居于卷曲的分子的外面(现在只知道一两种情况有在催化机理中起某种作用的荷电基团存在于蛋白质内部)。

我们现在能够知道疏水支链的进化保守性的道理,明白亲水残基赖氨酸具有保守性的原因之一:它们有助于使分子适当地卷曲。根本性的改变如在支链上阻止适当卷曲,则是致死的。没有卷曲,就没有细胞色素;没有细胞色素,就没有呼吸;没有

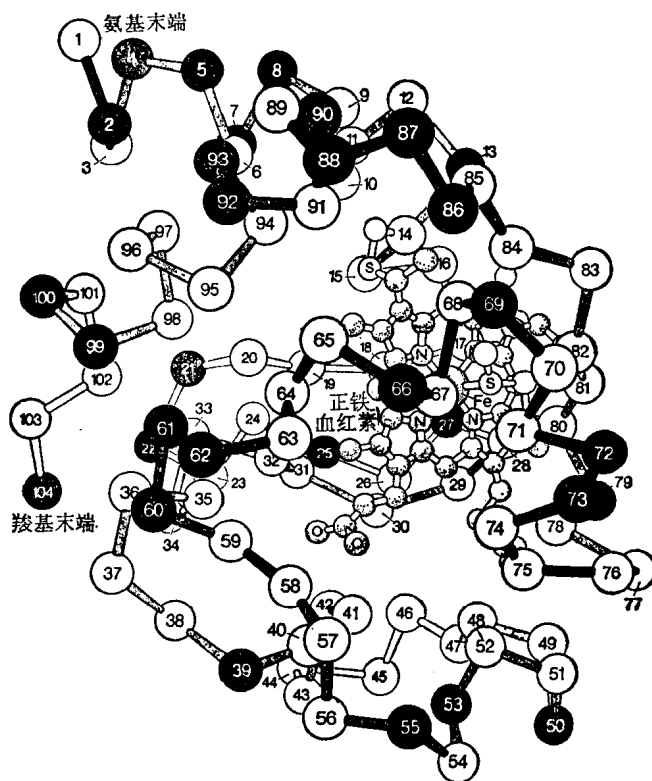


图4 马心细胞色素c背面上的电荷分布图

19个亲水的赖氨酸(带正电荷5, 7, 8, 13, 22, 25, 27, 39, 53, 55, 60, 72, 73, 79, 86, 87, 88, 99, 100)的大多数分布在分子的两侧。12个荷负电的(酸性的)支链(2, 4, 21, 50, 61, 62, 66, 69, 90, 92, 93, 104)中有9个群集在分子的上部中心的一个区带内。虽然酸性支链的具体位置有变化,但整个进化中保持了这个区带的荷负电的特性。从小麦到人没有一种生物的细胞色素c在这个区带内有少于六个酸性的氨基酸,没有一种生物在此分子的其他地方有多于五个酸性的氨基酸。在同一种生物没有发现上述的最大值和最小值。很象是这些荷电区带参与细胞色素c和其他大分子的结合。

呼吸,就没有生命。进化上原因和效果这样十分清楚是很少有的。

关于赖氨酸的功能还知道得更多。赖氨酸不仅是在外面,它们还群集在分子表面的两个荷正电的区域,被另一荷负电的区域所隔开。电荷的这种分隔在任何的其他蛋白质结构中还没有发现过,我们将会看到,它的存在,也许是因为细胞色素 *c* 和两个分子复合体(还原酶和氧化酶)相互作用,而不是象酶那样和小的底物分子发生作用。电荷的排列被认为是大分子彼此辨认过程的一部分。

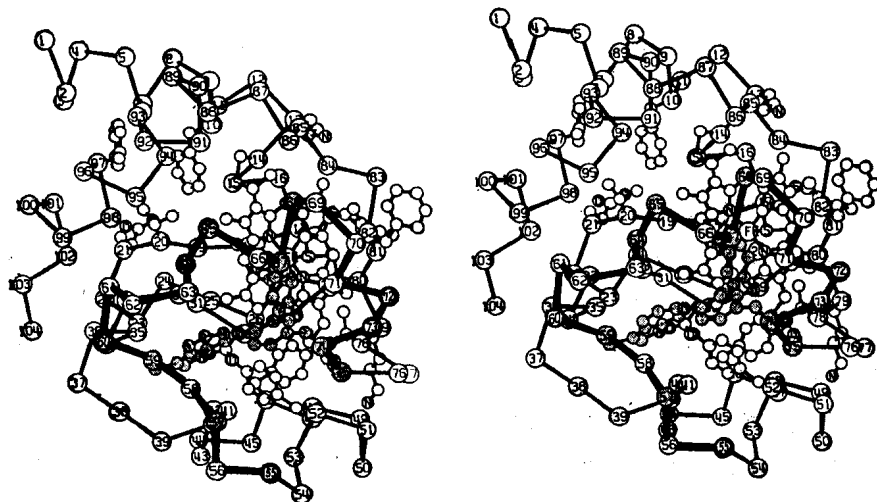


图5 氧化细胞色素 *c* 分子左侧的体视对

为了清晰,只画了少数关键的支链。在 55—75 的主链限定成一个环“左槽”,为强疏水的支链填充,其 α -碳有小斑点。其中三个支链有芳香环:色氨酸 59,酪氨酸 67 和酪氨酸 74,也用浅颜色。环绕左槽亲水的荷正电的支链用一粗黑线。

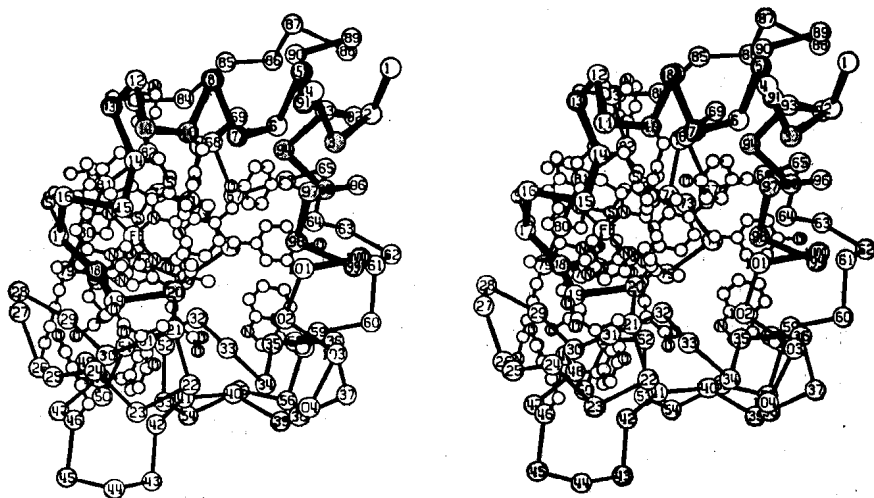


图6 氧化细胞色素 *c* 右侧的体视对

表明两个形成 α -螺旋的顺序: 1—11 的顺序和 89—101 的顺序。两个 α -螺旋和 12—20 的链形成右槽的轮廓。右槽象左槽一样排列有疏水的支链。

19个赖氨酸的大部分在分子的左侧和右侧(见图4)。在图5和图6的体视对中能分别观察分子的左侧和右侧。在左侧8个赖氨酸围绕一个,为疏水基团紧紧填充的链环(从55—75),包括不变的酪氨酸74,色氨酸59和酪氨酸67在最内部。虽然我们仍不了解电子转移机理,但已提出三个不变残基的芳香族环,当细胞色素 c 被还原时能给电子提供一个向内的途径。另外8个赖氨酸是在右侧,在大到足以支持另一个大分子的疏水支链的槽内。

这个右侧槽为两个 α -螺旋所束缚,也被第一个 α -螺旋由残基12—20的延续所束缚。在该槽内有两个大的芳香的支链:苯丙氨酸10(不变)和酪氨酸97(也可能是苯丙氨酸,但不会是别的)。总之,在右侧是一个排列有疏水基团的槽(包含两个芳香环),外面包围有一个正电荷的外环。正如我们实验室有人看到这个模型时所说的,它象空间船的一个入坞环。

这种说法不完全是儿戏的。从化学工作中知道细胞色素 c 和细胞色素氧化酶复合体之间的吸引主要是静电的,包括在氧化酶上的荷负电的基团和在细胞色素 c 上的荷正电的碱基。左侧或右侧群集的赖氨酸一定是与这个连接有关。Kazuo Okunuki(奥贯胜男)已指出,当只有一个正电荷,赖氨酸13,被一个庞大的芳香族的化学基团所封闭,细胞色素与其氧化酶的反应能力减少一半。化学上封闭的赖氨酸13意指物理上封闭血红素裂隙的上部。赖氨酸13更靠近于右侧的正电荷团(相对左侧而言),所以很可能正铁血红素裂隙和右槽一起是“辨认”氧化酶复合体的分子表面的部分。

左侧的荷正电区带和背后的荷负电区域的作用是什么呢?荷正电区带,由于有它的三个芳香环,可能是连接还原酶的地方,我们实质上还不知道这种连接的化学性质。荷负电区域可能是“垃圾堆”,分子表面不重要的部分,在该处有充分的负电荷可防止过度的总的正电荷。6个最易变的氨基酸部位在该处的这一事实支持上述的概念。

另一方面,同样可能负电荷的聚集有某种机能。酸性的氨基酸在各种生物都是保存下来了,虽然在某种微妙的方式下在早期的顺序比较时被忽略了。即使带负电荷的个别残基因种属不同而有差别,选择压力还是保持了分子表面这一区带的负电性。因为有几段蛋白质链弯向和弯出这个酸性区域,如果我们只看伸展出来的顺序,负电荷的保持不是立即显而易见的。这很好地说明了分子进化的原理即自然选择作用于卷曲的有机能的蛋白质,不是单独作用于蛋白质的氨基酸顺序。

如我们细心地观察甘氨酸残基的部位,我们就能了解为什么这么多残基是进化上不变的。血红素基团是如此之大,以致104个氨基酸好容易才能够缠绕着它。有很多地方,一个链距正铁血红素或另外的链太近,以致不适合有支链。正是在这些点我们发现带有单个氢原子作为支链的甘氨酸。

保守性的最后一种类型,芳香环链的保守性,更难解释。酪氨酸和苯丙氨酸常常在卷曲的细胞色素 c 分子中邻近成对:残基10和97在右槽,46和48在正铁血红素裂隙下面,67和74与色氨酸一道在疏水的左槽。只有残基36(可能是酪氨酸,苯丙氨

酸或异亮氨酸),在分子的背后似乎只有一个填空的作用,它是一个“油砖”。

左槽的三个芳香环在还原作用中可能与电子转移有关。右槽的两个环也可能用于电子转移,或者只能有助于把疏水的窄口局限在槽的中间。分子底部酪氨酸 48 用氢键连到正铁血红素的一个丙酸支链帮助支持血红素就位。在金枪鱼和鲤的细胞色素 *c*, 残基 46 是酪氨酸,其电子密度图指出该残基也以一个氢键连接血红素的另一个丙酸基。这两个酪氨酸与半胱氨酸 14 和 17 一起帮助固定血红素的位置,在血红蛋白或肌球蛋白没有看到这种情况。

苯丙氨酸 82 是个谜。从不是酪氨酸或其他残基,把其油性的支链伸出到正铁血红素裂隙左侧上面的水环境中,在该处它没有可以看到的作用。为了使它在该处一定要付出能量代价。为什么如此一个大的疏水基团要在分子的外部呢?为什么经过整个进化过程它是绝对不变的呢?单独看氧化的分子不可能说,但是在本题的末尾我们看看近来揭示出来的还原细胞色素的结构时,我们立刻会看到解答。

现在能够大体上解释细胞色素 *c* 在真核体生物整个历史中表现进化保守性的结构上原因。细胞色素 *c* 在结构上作了分析的蛋白质中是独特的,它的表面上有分隔的电荷区。前面讨论这些区的作用是推测的,可能是很错误的。我们能确信的是这些区域在分子的执行功能上有作用。在十亿年以上的分子进化中,所有种属单独靠偶然的会或共同的祖先,不能与成对的和暴露的芳香基一起维持这些正电区和负电区。这个保守的顺序在叫我们“看!”现在我们应该足够聪明知道去寻找什么?

蛋白质进化的速率

了解了上述的背景,我们就有了回到以前提出的一个问题的准备,什么东西决定不同蛋白质的进化速率?我们先作一个图,纵轴是进化上分枝点两边的两种生物之间例如鱼类和爬行类之间或爬行类和哺乳类之间氨基酸顺序的平均差异。横轴是两个族系分离距今的时间,按地质学记载断定的。如果为细胞色素 *c* 画这样一个图,可以发现所有的分枝点都落在一直线附近,说明进化上变化的平均速率是恒定的(图 7)。

怎么能是这样?在这长久的时期内,生物的外部形态变换多样成为今天的棉花、面包霉、果蝇、响尾蛇和黑猩猩,细胞色素 *c* 怎么能以如此恒定的速率在变化呢?这说明用蛋白质作为研究进化的工具有根本的优越性。自然选择最终是对所有生物群体有影响,唯一的成功标准是群体生存,生殖和留下新一代的能力。我们研究活的机体愈深入到分子水平,它们变得愈相似,把一个蛤蜊与一匹马相区分的形态学差异愈不那么重要。一种化学机构能适合很多种不同的生物。相反地,自然选择能够影响到的某种生物的一个外部变化,常常不是一个酶分子的效应,而是一整套代谢途径的效应。

细胞色素 *c* 内的始终一致的变化速率只意味着呼吸小体——线粒体——的调整是这样的好,线粒体是这样的不受自然选择的影响,以致于选择压力在分子水平上经