

参 考 文 献

- [1] Sharon, N.: *Scientific Am.*, **236** (6), 108, 1977.
- [2] Sharon, N. et al.: *Science*, **177**, 949, 1972.
- [3] Allen, A. K. et al.: *Biochem. J.*, **131**, 155, 1973.
- [4] Reeke, G. N.: *Advances in Expt. Med. and Biol.* (Eds Chowdhury, T. K. et al.), Plenum Press, Vol **35**, p. 13, 1974.
- [5] Lis, H. et al.: *Ann. Rev. Biochem.*, **42**, 451, 1973.
- [6] Sawyer, W. H.: *Advances in Expt. Med. and Biol.* (Eds Chowdhury, T. K. et al.), Plenum Press, Vol. **35**, p.71, 1974.
- [7] Edelman G. M. et al.: *Proc. Nat. Acad. Sci. U.S.A.*, **69**, 3580, 1972.
- [8] Hardman, K. D. et al.: *Biochemistry*, **11**, 4910, 1972.
- [9] Wright, S. C.: *J. Mol. Biol.*, **111**, 439, 1977.
- [10] Olsnes, S. et al.: *the Specificity and Action of Animal Bacterial and Plants Toxins* (Receptors and Recognition, Series B. Vol. 1) (Ed. Cuatrecasas), p.129, Chapman and Hall, 1977.
- [11] DeBray, H. et al.: *Protides Biol. Fluids* (Ed. Peeter, H.), Vol. **27**, p.451, Pergamon Press, 1981.
- [12] Kornfeld, S. et al.: *The Glycoconjugates II* (Eds. Horowitz, M. I. et al.), p.473, N. Y. Academic, 1978.
- [13] Brown, J. C. et al.: *Intern. Rev. Cytology*, **52**, 277, 1978.
- [14] Lis H. et al.: *The Antigen* (Ed. Michael, S.), Vol. **IV**, p. 429, Academic Press, 1977.
- [15] Stobo, J. D. et al.: *J. Immunol.*, **108**, 1, 1972.
- [16] Novogrodsky, A. et al.: *J. Immunol.*, **115**, 1243, 1975.
- [17] Novogrodsky, A. et al.: *Proc. Nat. Acad. Sci. U.S.A.*, **70**, 2515, 1973.
- [18] Bessler, W. et al.: *Biochem. Biophys. Res. Commun.*, **69**, 578, 1976.
- [19] Lis, H. et al.: *Advances in Experimental Medicine and Biology* (Eds. Chowdhury, T. K. et al.), Vol. **55**, p. 309 Plenum Press, 1974.
- [20] Olsens, S. et al.: *Protides Biol. Fluids* (Ed. Peeter, H.), Vol. **27**, p. 395, Pergamon Press, 1980.
- [21] Cuatrecasas, P.: *J. Biol. Chem.*, **248**, 3528, 1973.
- [22] 孙册等: «免疫技术在分子生物学和细胞学中的应用研究»交流讨论文札, 1980年9月第46页。
- [23] Fox, J. L.: *Chem. and En. News*, **56** (31), 25, 1978.
- [24] Reisner, Y. et al.: *Trends Biochem. Sci.*, **5**, 29, 1980.
- [25] Nakano, T. et al.: *Glycoconjugates Proc. 5th. Intern. Sym.* (Eds. Schauer, R. et al.), p. 452, 1979.
- [26] Edelman, G. M. et al.: *Methods Enzymol.* (Eds. Jakeby, W. B. et al.), Vol. **34**, p. 195, 1974.
- [27] Bavard, B. et al.: *Biochim. Biophys. Res. Commun.*, **77**, 489, 1977.
- [28] Skelly, C. R. et al.: *J. Gen. Virol.*, **44**, 679, 1979.
- [29] Latner, A. C. et al.: *Anal. Chem.*, **101**, 483, 1980.
- [30] 杨建中: «免疫技术在分子生物学和细胞生物学的应用研究»交流讨论会文札, 中国生化学会印, 1980年, 9月, 第8页。

[本文于1980年12月27日收到]

蛋白质晶体学——分子生物学的得力工具(一)

王 家 槐

(中国科学院生物物理研究所)

一、引 言

自从劳埃 (Von Laue) 在1912年成功地拍摄了第一张晶体的X射线衍射像后, 近七十年来, X射线晶体学一直是研究物质结构极其有用的工具。最初人们仅用X射线晶体结构分析方法测定一些晶胞内一个独立单位*, 只含几个原子的简单的矿物结构, 以后是十几个到上百个原子的有机化合物。五十年代后, 晶体学家成功地发展了一整套方法, 可以测定蛋白质分

子的几百、几千个非氢原子在三维空间的排布。由于大型计算机以及自动衍射仪的出现, 七十年代大批蛋白质结构被测定, 达到原子分辨率水平 (优于 $2.5 \text{ \AA}$) 的约有一百个; 与此同时还测定了几个转移核糖核酸的晶体结构。从测定的精度看, 经过仔细修正的蛋白质结构, 一些定位较好的原子的位置误差已降到 $0.15 \text{ \AA}$ 左右^[1]。分子量高达几百万, 由几十万个非氢原

* 在晶体学里把晶胞内不能由晶体学对称元素相关联的部分叫独立单位。

子组成的病毒粒子的研究也已突破,其中西红柿矮小病毒、南方豆花叶病毒等晶体结构的研究,其分辨率已达到 $2.8 \text{ \AA}$ 左右^[2,3]。用大型同步辐射加速器作为强 X 光源,一些生物学动力过程,如肌肉收缩等也开始进入 X 射线晶体学的研究范围^[4]。

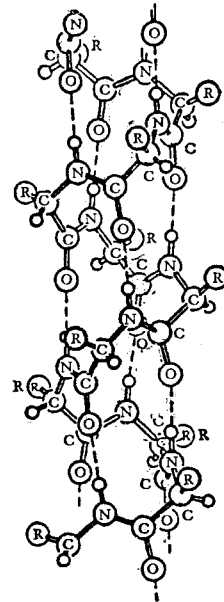
分子生物学一诞生就与 X 射线晶体学分不开。著名的蛋白质 α -螺旋和 β -折叠结构模型以及遗传分子 DNA 双螺旋模型均是借助于核苷酸和小肽段的单晶结构分析和纤维衍射图分析才得以建立。现在,一个新的学科分支——蛋白质晶体学(更广义称做生物晶体学)已经出现。它是分子生物学的得力工具。本文试图对目前用 X 射线晶体学方法研究生物学一些问题所取得的成果作一概述。

二、比较分子生物学

在经典生物学中,比较的方法是研究种系进化的有效办法。现代生物化学家在测定了大量取自不同种属的同一蛋白质的氨基酸排列顺序(称为一级结构)后,成功地用比较的方法使分类学在分子水平上发展起来^[5]。现在晶体学家在测定了大量的蛋白质三维结构后,也用比较方法对结构作研究。我们把这类在分子水平的比较研究称为比较分子生物学。

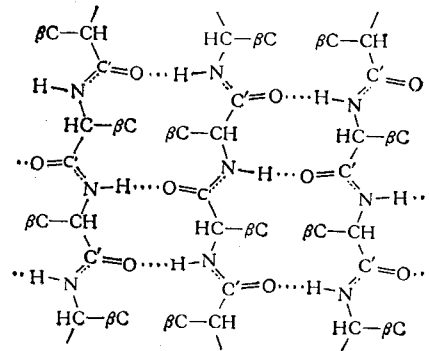
早期的比较研究首先揭示了蛋白质分子构象的一些共同特点:多肽链在折叠成三维结构时,相邻的一些氨基酸往往先形成一些 α -螺旋片段或 β -折叠结构(图 1),然后几个这样的片段又折叠成一个相对独立的结构区域;一条多肽链有时可折叠成一个以上的结构域(图 2)。由多条多肽链组成的蛋白分子,每条肽链独立折叠成一个亚基,然后几个亚基又按一定方式装配成球蛋白。分子是一个极其紧密的球状体。分子内部主要是疏水性的基团,外表则多是极性基团,极性基如出现在分子内部,通常互相配对形成氢键或静电键,有的参与形成活性中心区等,有一定生物学功能。

随着测定的结构越来越多,人们发现,初看如一团乱麻的蛋白质构象,实际上其折叠方式



(a)

a) α -螺旋结构模型



(b)

b) 反平行的 β 折叠层结构

图 1

[引自 Pauling, L. and Corey, R.B., *Proc. Nat. Acad. Sci.*, 37 729, 205, 1951)

可能是有限的^[6]。不仅不同种属的同一种蛋白质的构象基本相同,而且有时无论其一级结构还是生物学功能毫无相同之处的两种蛋白质,竟然有大体相同的折叠方式(借用数学术语,可称之为拓扑学相似性)*。典型例子是免疫球蛋白

* 拓扑学是几何学的一个分支,它研究连续可变几何图形的不变性。在拓扑学看来,各种形状(方或圆的)元把的杯子是等同的,但有把的杯子却是不等同的,不可互变的。

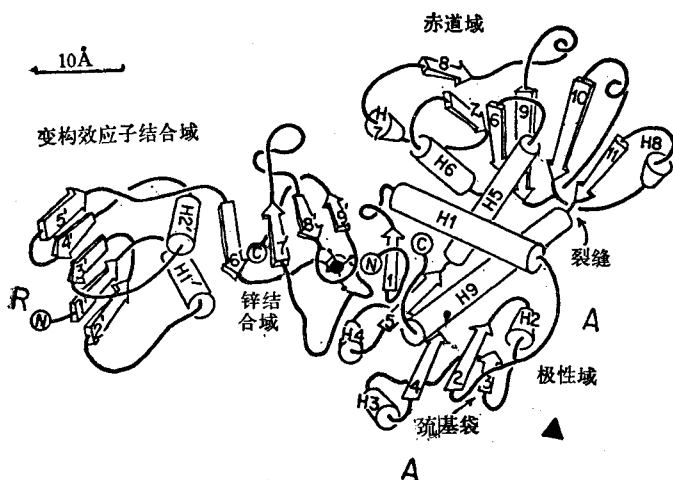


图 2

门冬氨酸转氨酶由 12 个亚基共 24 个结构域组成，其中 6 个催化亚基(c)，6 个调节亚基(R)。这里示出一个调节亚基的两个结构域——变构效应子结合域和锌结合域以及一个催化亚基的两个结构域——赤道域和极性域。图中箭头代表 β -结构，短圆柱代表 α 螺旋。▲ 号表示联系分子内三对亚基的三重旋转对称轴。[引自 Kantrowitz, E. R. et al.: *Trend in Biochem Sci.*, (1980) 6, 150]

白和过氧化物歧化酶^[8]。后者的功能是将体内可能存在的极其活泼的过氧化物自由基 O_2^- 转变为 H_2O_2 和 O_2 ，以免机体受其害。这两个蛋白质的一级结构，即使用电子计算机仔细地逐段比较，也找不出任何相似之处。在三维结构

性^[10,11]。

为了深入研究这种三维相似性，蛋白质晶体学家发展了一套定量的比较方法^[10,12]，其基本思想是：寻找一个旋转矩阵 $[c]$ 和一个平移向量 d ，它们能使两个相互比较的分子重叠在一起时，所有相应原子之间距离的平方和取极小。作这种定量比较时，一般只取主链原子，忽略掉不同的侧链基团，以突出比较肽链的走向。过氧化物歧化酶的一个亚基与免疫球蛋白的一个结构域经过这样的矩阵和向量的作用后重叠在一起时，两者相应原子间距离的均方差只有 2 Å 左右。

人们很快发现，三维结构的比较与一级结构的比较颇为不同。比较不同种属的同一蛋白质的一级结构时，人们可以追溯物种进化的过程。不同种属同一蛋白质氨基酸序列的差别主要是在某些位置上氨基酸残基的替换，而且这种替换与决定该蛋白质的遗传分子 DNA 上相应遗传密码子的变化，即核苷酸碱基的替换相关联。DNA 分子上发生一个核苷酸碱基的替换(突变)，反映到相应蛋白质分子上是一个氨基酸残基的替换(但并不都是一一对应关系)。在漫长的物种进化过程中，这种变化很多。表

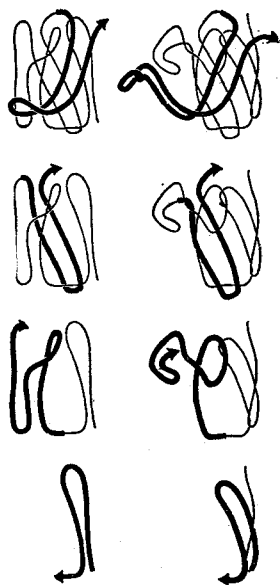


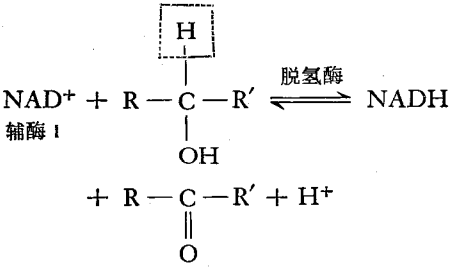
图 3 过氧化物歧化酶

(左)和免疫球蛋白可变结构域(右)肽链折叠的比较。自上而下，黑粗线表示肽链从氮端起的走向。两者惊人地相似^[8]。

1 是不同种属胰岛素一级结构的比较, 大约只有 40% 的氨基酸残基是已知各种属共有的。不同种属血红蛋白肽链 150 个左右氨基酸残基中只有 9 个是共同的。但当我们观察不同种属同一蛋白质的三维结构时, 却发现它们基本上是一样的。比如取自八目鳗(一种已知有胰岛素的最原始的脊椎动物)的胰岛素的三维结构与猪胰岛素三维结构的差别甚至不比同一晶体两个猪胰岛素分子三维结构之间的差别大, 然而这两个种属的胰岛素的一级结构相差甚远^[43]。三维结构模型的解释很快使人们明白了为什么有些氨基酸残基非常重要, 在进化过程中从不更换, 它们或者是对维持三维结构是必须的, 或者是活性部位的组成部分。有人比较了九个取自不同种属的血红蛋白的原子结构^[44], 它们之间一级结构差别最大的有 84% 氨基酸残基的排列是不同的。这些作者发现, 对三维结构比较重要的氨基酸残基有 59 个, 都在螺旋段的界面处或螺旋段与血红素基团的界面处, 且半数在分子内部。有些位于分子内部的氨基酸残基虽然可以变化, 但仍保持非极性性质, 仅是侧链基大小不同。这种变化只是导致各螺旋段的堆积几何有所不同。螺旋段间相对位置和定向可相差 7 Å 和 30°, 但变来变去, 整个分子却始终保持在血红素基团附近有一个几何形状相似的口袋, 允许氧分子可逆地进入分子内与血红素基相结合。这正是血红蛋白生物学功能所必不可少的, 这是自然选择的结果。总之, 在进化过程中, 三维结构比一维结构要保守得多。蛋白质三维结构之间的比较不能提供更多关于同种蛋白质的种系进化的信息。然而人们发现, 它也许能提供关于蛋白质分子本身进化过程的信息。

从分子进化角度看三维结构的相似性可以归纳为两种进化方式: 发散与收敛^[45]。发散型是指最早有一个原始的祖先分子, 在进化过程中衍生出一系列功能相关, 结构相似的不同蛋白质分子, 是“同根生枝”。血红蛋白与肌红蛋白无论在一级结构或生物功能方面都属同族, 其三维结构也相似(这是一类主要由 α-螺旋组

成的折叠形式), 这是明显属于发散型的一例。NAD-脱氢酶类也属于发散型。它们基本上都由两个结构域组成, 其中一个结构域结合辅酶 I (烟酰胺腺嘌呤二核苷酸, 简作 NAD), 故称 NAD-结合域, 另一个是底物结合域或叫催化域。尽管 NAD-域在不同脱氢酶内位置可能相异, 但它们的三维折叠都是几个 β-α-β 结构片段。目前已知三维结构的这类脱氢酶已有若干个。它们共同的功能是, 把底物上的一个质子和两个电子转移到电子受体辅酶 I 上, 从底物上下来的另一个质子则为蛋白质活性区的一个组氨酸侧链所接受。这种脱氢化过程首先是由



于这类蛋白质共有的 NAD-结合域里存在一个疏水的口袋, NAD⁺ 分子的腺嘌呤环正好适于躺在那个口袋内, 从而将它的另一端——接受电子和氢的烟酰胺环伸入蛋白质分子催化域的活性区。不同脱氢酶之差别是活性区构造不同, 以适于不同类型的脱氢反应。至于收敛型进化是指几乎毫不相干的分子在进化过程中共同利用一种三维结构安排, 是“殊途同归”。这里又有两种情况。一种是一旦一个空间排布非常合理的活性中心出现后, 它会被广泛利用于不同的场合。如碳酸酐酶、肝醇脱氢酶、羧肽酶和嗜热菌蛋白酶都有一个以 Zn 为中心的活性区, 围绕 Zn 的一组功能团的空间排布极其相似^[46]。另一种是从热力学角度来看, 能量较低的肽链折叠方式也会广泛被采纳为蛋白质三维结构的基本构架。如有人发现^[47], 在许多蛋白质内有 β-α-β 结构段, 其中两股 β-链是平行的, 而 α-螺旋则躺在与 β-链局部区域平行的位置(图 4)。因为通常把 α-螺旋和 β-折叠层称为蛋白质的二级结构, 所以他们把这种结构段叫做超二级结构。β-桶式的结构也相当普遍地

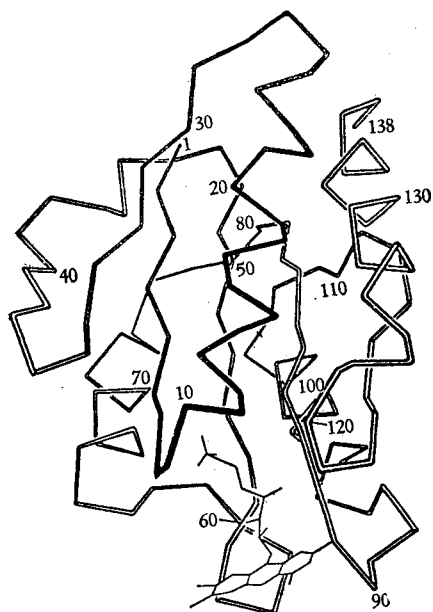


图4 β - α - β 型超二级结构

图中示出黄素蛋白的肽链走向,其中从氨基端第一个氨基酸残基到第35位那一段用黑粗线画出的肽段清楚地显示这种超二级结构的排布 [引自 Burnet, R. M. et al.: J. Biol. Chem., 249, 4383, 1974.]

解开了长期使分子遗传学家迷惑不解的一些问题的症结,而且立即与蛋白质三维结构的结构域概念联系到一起了。免疫球蛋白是目前研究最深入的例子之一。一个免疫球蛋白由两条重链和两条轻链组成(图5)。每条轻链有一个可

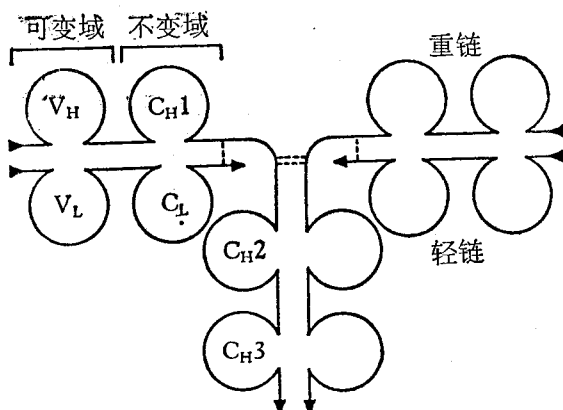


图5 免疫球蛋白分子示意图

整个分子由两条重链和两条轻链组成。轻链有一个可变域 V_L 和一个不变域 C_L 。重链有一个可变域 V_H 和三个不变域 C_{H1} , C_{H2} 和 C_{H3} [19]

存在于各类蛋白质,包括一些球状植物病毒的外壳蛋白中^[2,3]。当然,目前区分两个有一定相似性的蛋白分子在进化关系上究竟是收敛型还是发散型还很没有把握。但经验告诉我们^[8],发散型的、遗传上同源分子间往往有下列特点:多肽链内常常是整段肽链的缺失或插入(有时甚至是很长的一段),尤其是在多肽链的首尾两端,多肽链内的硫硫键或者减少、或者增加,却从来不交换其联接方式;小的缺失或插入易出现在连结二级结构的“环区”,二级结构单位可长、可短、可增加、可消失或略有位移,但不交换其相对位置,也即联接二级结构单位的三维拓扑学是不变的,相关的蛋白质常常有不同的亚基数,即使它们有相同的亚基数,它们也可能采用不同的亚基堆积方式。

在蛋白质晶体学家积累了这些三维结构比较资料的同时,分子遗传学家在1977—1978年间报道了一个惊人的发现^[17]:病毒和动物体的基因是分片分散在DNA分子上,而不是整个基因完整地联结在一起*。这个发现不仅很快

变的结构域 V_L 和一个不变的结构域 C_L ,每条重链则有一个可变的结构域 V_H 和三个不变的结构域 C_{H1} , C_{H2} , C_{H3} 。一个动物可以有上千种 V_H 和上千种 V_L 。这就提供了上百万种不同的 V_H 和 V_L 的组合,用于结合各种各样的异体入侵物——抗原。事实上每个结构域本身由一百个左右氨基酸残基组成,分别折叠成一个一个紧密而稳定的球状单位,彼此间极其相似。一如前述,现已证明^[18,19],每个这样的结构域是由一个基因段复制出来的,这个基因段被称为微基因,它们分散在DNA分子的不同地方。当DNA分子在细胞核内转录成RNA时,最初仍保留原状,后来这个RNA分子为一个拼接酶切去“无用”部分,然后把有用部分拼接成信息核糖核酸mRNA,跑出细胞核,到达蛋白质合成地点,转译产生免疫球蛋白(图6)。这种功能能使机体很快对外界异物的入侵发生反应,迅速大量产生能对抗该异物的免疫分子品

* 美国威斯康辛大学一个生化小组最近在植物的基因里也发现了相同的现象。

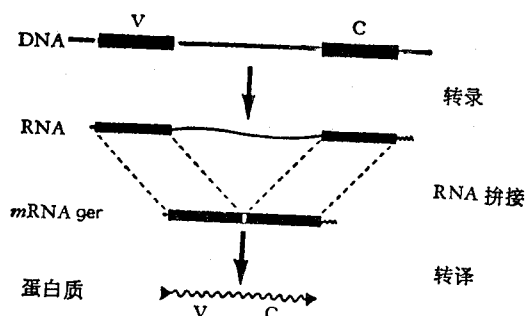


图6 免疫球蛋白轻链的产生过程示意图

V_L 和 C_L 两个结构域的微基因在 DNA 分子分离的区域^[19]

种,以防异物为害机体。事实上,每个细胞只产生一种免疫分子,其差别就是可变区基因片有突变,特定的异物结合到相应细胞表面的受体上,会刺激该种细胞大量分裂并且产生有效的免疫分子。当然,并非所有的基因段都正好相应于一个蛋白质的结构域。但基因分片与蛋白质结构域之间的联系使我们对于分子进化的认识大大深化。由此可以理解发散型和收敛型无非是在进化过程中不同的分子利用了一个相同

的微基因或甚至更小的基因片段,再加上别的基因片段,拼接而成。每个基因段的遗传和变异也因而有一定的独立性,并且一个新的分子的进化成型并不要求先破坏一个老的分子。这使整个生物界对自然环境有了更大的适应能力。

总结上述,作者倾向于认为:蛋白质分子在进化过程中,大量涉及的是一个氨基酸残基替换的量变,它们起因于 DNA 分子中一个核苷酸碱基的替换。自然选择则只保留那些稍加改善或至少不损害分子基本功能的替换。这就是同一蛋白质分子的种属差异现象。而这种量变积累到一定程度后,会发生质变,有可能出现新的基因片段组合方式(包括成段的插入和缺失)。如果这种组合更有助于机体对环境的适应,那么一个新的分子品种就会诞生。随着越来越多的三维结构的测定以及蛋白质分子折叠过程研究的深入,联系分子生物学其它分支的进展,比较分子生物学一定会提供越来越多的信息(未完待续)。

核磁共振及其在分子生物学中的应用(二)

华庆新

(中国科学院生物物理研究所)

三、NMR 在分子生物学中的应用

最早从事生物样品 NMR 研究的是 Jarde-tzky^[10]。1957 年他用 2—3M 的 Gly、Ala 等氨基酸作 ^1H 谱,观察共振峰随 pH 的变化。第一个蛋白质的 NMR,也是在 1957 年由 M. Saunders^[11] 等人测定了 RNase 的 ^1H 谱。以后, NMR 在分子生物学中得到了广泛应用。

1. 作分析研究 (1) 鉴定生物分子成分。最简单的方法是将图谱与已知化合物分子的谱线进行比较。一般由 δ 、J 强度,即可大体确定分子成分。研究完整组织时,由于 NMR 的非破坏性,已受到广泛重视。如在人、蛙、蟾蜍的

腓肠肌,新发现了一种磷的代谢成分——甘油-3-磷酸胆碱^[4]。其含量有种属差异,与生理状况亦有关。这为进一步研究提供了线索。目前要推测未知化合物的分子结构,必须仔细分析图谱。这在有机化学中用得较多,而在生物学中还不常用。1965 年有人^[12]曾预言 NMR 可用作短链多肽(如三肽)的序列测定。实际上,它比常规定序,并没有太多优点。但最近又有人用 Gd^{3+} 谱线加宽探针,作了小肽序列测定,也有人作了多糖糖链的序列测定^[13]。

(2) 确定生物分子的物理、化学状态 只要在所测 pH 范围内,样品有可解离的基团, NMR 即可测定分子的解离态,测定样品的 pH。