

研究蛋白质溶液构象 的一种新方法

林克椿*

生命科学的研究已经进入了分子时代,人们希望能应用分子结构及其变化的观点来说明生命的基本过程,包括遗传及其变异;药物、毒物对细胞的作用机理;免疫过程及其缺陷的分子过程;衰老、癌症以及多种疾病的预防与治疗等等。所有这些现象的研究都建立在对生物大分子(特别是蛋白质和核酸)结构的了解的基础之上,而当前遗传工程、蛋白质工程等高科技的发展无疑也必须首先知道生物分子原来的结构和起主要作用的部位,然后才有可能加以修饰以满足实践需要。因此,分子结构的研究是分子生物学的核心课题,而在分子结构的研究中,在它发展的各个不同阶段上,近代物理学的理论与实验技术都是起决定性作用的因素。

拿X射线衍射作为一个例子可以充分说明这一点:如果能把蛋白质(或核酸)做成有规则的晶体,那么它的结构就可以用这种方法加以测定,即用X射线照射、获得它的晶体衍射图像、利用强度和位相关系反过来得出蛋白质的结构模型。廿世纪中叶开始对脱氧核糖核酸(DNA)和血红蛋白结构的研究奠定了分子生物学发展的基础,而且随着分辨率的不断提高,获得的信息也越多。例如我国在胰岛素的研究中目前已能达到 1.2 \AA 的分辨率,不仅能测出全部原子的空间位置,而且能够“看到”其中的水分子。当前已经测定出三维空间结构的蛋白质已经有大约400种之多,但能达到高分辨率的还只有10%左右。这方面的工作还需要继续和不断改进。

然而认真地考虑一下关于结构研究的现状就会发现,X射线衍射并不是唯一的、完美无缺的方法。首先它需要生物分子能形成晶体,实际上,生物体内的生物分子并非以晶体形式存在,而是更接近于“浸泡”在溶液(水)之中的状态。其次它的结构并不是刚性的,而是有一定的可变性,特别是在环境条件改变、或与其它分子相互作用而发挥其功能作用之时。这就需要寻找

能够研究在溶液中的大分子构象及其变化的方法。人们又一次求助于物理学,从六十年代开始进入了结构研究的第二个阶段,应用了圆二色、红外和核磁共振等来研究蛋白质的溶液构象。其中八十年代初发展起来的二维核磁共振已成为最有力的手段之一。

什么是二维核磁共振?

要了解二维核磁共振,首先需要知道一维核磁共振(1D-NMR)。由于原子核有自旋运动,因而在外磁场作用下将发生能级分裂。当再加一个相当于二个分裂能级差的射频场时,原子核即产生共振吸收,吸收峰形与磁场强度(或射频频率 ω)的关系即为核磁共振NMR谱。一维谱在分析分子结构成份、测定活细胞中的代谢以及在核磁成像等临床应用中已经发挥了很大作用。

把NMR技术应用于蛋白质分子时常遇到很多困难。由于蛋白质分子是由20种氨基酸组成的多肽链,每一种氨基酸中又有许多H原子,因而所得到的1D-NMR谱中许多峰重叠在一起,很难分辨。对于测定一种蛋白质在溶液中的构象,首先要能测出每一种氨基酸及其在多肽链中的顺序,然后再了解整个长链的弯曲折叠状况。为了取得这些信息,从1975年开始应用了二维核磁共振技术,即在原来1D-NMR测定的同时增加另一个射频 ω_2 ,使所得到的谱展开在一个XY

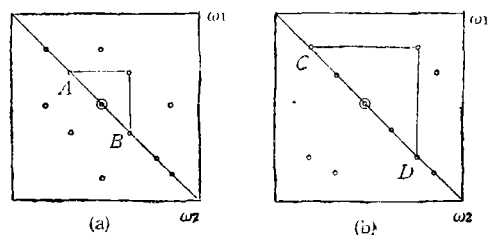


图 1

平面上,而不像1D-NMR那样只在X轴上。根据测谱时所用脉冲序列的不同,可以得到不同的2D-NMR谱,其中最典型的是二维相关谱(COSY)和二维核Overhauser谱(NOESY)。图1a,b分别代表一种假想的多肽的COSY谱和NOESY谱。它们的共同点是对角线上的点子相同,都和一维NMR谱完全一样,只是比一维谱分得更清楚一些而已。它们不同之处是在对角线以外的区域还有一些交叉峰,这些交叉峰在二种图上所给出的信息不同。在COSY谱中由交叉峰作平行于 ω_1 或 ω_2 轴的直线与对角线交于A、B二点,则这一交叉峰说明多肽链中同一个氨基酸残基上的二个H原子A、B之间有自旋偶合,通常能有偶合的这二个H原子距离不会超过三个键的间隔。而在

* 作者系北京医科大学教授

NOESY 谱中用同一方法所得到的对角线上二点C、D, 则说明不在同一残基上的二个H原子由于距离接近(一般不超过 5 Å)也有偶合(称为核 Overhauser 增强效应 NOE), 所以多肽链很可能由于弯曲或折叠才使得它们如此接近. 这就是用 2D-NMR 分析结构的出发点. 注意实际的二维谱应该是在 XY 平面上向 z 方向升高的许多峰, 称为堆积图. 通常都只画出 XY 截面上代表不同强度与高度的大小不等的等高线图以便于分析.

多肽链氨基酸序列的确定

确定蛋白质(或多肽)结构的第一步是用 2D-NMR

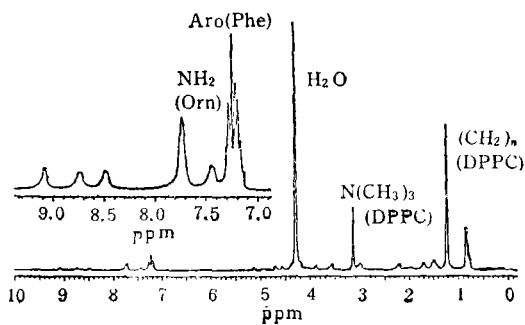


图 2

谱辨认出每个氨基酸残基的 H 峰, 以及它们之间的顺序关系, 这就需要充分利用前面所提到的 COSY、NOESY 谱以及两种谱的结合. 如果从生物化学分析已经知道这种多肽的结构关系, 将大大简化这一过程.

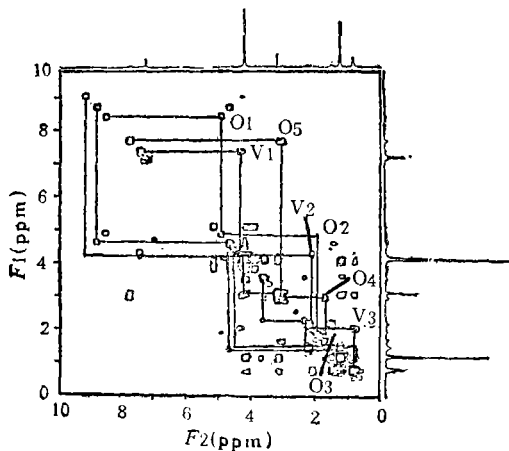


图 3

这里用我们做过的一项工作——短杆菌肽 S 与脂相互作用时构象的改变——作为例子加以简要说明. 短杆菌肽 S 是由苯丙氨酸、亮氨酸、鸟氨酸、缬氨酸、脯氨酸组成的二条平行而反向的 β 折叠链在脯氨酸与苯丙氨酸间互相连接形成的一种环状十肽. 它在脂中的一维谱见图 2, 左上角是放大的几个属于氨基酸残基峰,

这几个峰清晰可见, 取其中之一作为分析的起点, 例如 7.67 ppm 处的鸟氨酸 NH₂ 峰, 在图 3 的 2D-COSY 谱中对角线的第 4 个点即为此峰, 作平行线可遇到交叉点 O5, 从这一点作垂线与对角线上交点即为鸟氨酸的 C^αH 峰, 再从这点找到 O4, 从而找到 C^γH、依此类推, 从 O3、O2、O1 可找到 C^β、C^α 和 NH, 故鸟氨酸的全部 H 峰都可确定.

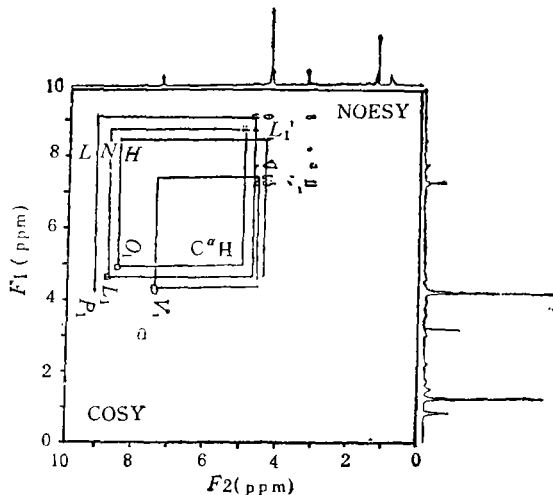


图 4

怎样从鸟氨酸的峰来寻找相邻残基的峰? 由于相邻残基间的 H 往往超过了三个键, 只有当多肽链有折叠时, 它们之间才有偶合, 这就需要利用 NOESY 谱了. 实际上无论是 COSY 还是 NOESY 谱, 相对于对角线来说, 左右两侧是完全相同的, 所以为了分析方便起见, 可以把这两张图沿对角线剪开, 然后把 COSY 和 NOESY 谱沿对角线重新组合成一张半为 COSY 半为 NOESY 的图, 其简化后的图见图 4. 在此图上从 O1 看到鸟氨酸的 C^αH (在对角线上), 然后作与 ω₁ 平行的垂线将遇到 NOESY 谱中的交叉峰 L', 从 L' 作平行于 ω₂ 的线将对角线相交, 这就是与鸟氨酸相邻的亮氨酸的 NH. 从这点出发, 再次利用 COSY 可得到亮氨酸的各个 H 峰. 用类似的方法反复利用 COSY/NOESY 结合图即可找出多肽中全部氨基酸残基的所有 H 峰.

从 NOE 效应计算质子间的距离

前面已经了解到 NOESY 谱中的交叉峰代表虽不在同一残基内, 但空间距离小于 5 Å 的两个质子之间有偶合. 这种偶合程度的大小和二质子间距离的六次方成反比, 所以 NOESY 谱中每个交叉峰的强度(大小)实际上反映了与其对应的两个质子之间的距离. 当然, 从某一具体的 NOESY 谱中所能得到的只是这些距离之间的相对关系, 因为它和实验条件有关. 因此还需要利用其它方法(例如作一个自旋回波谱, 从谱

线分裂程度了解距离的绝对值)加以校正,这样就能得到所有具有 NOE 效应的质子之间距离的绝对值了。

NOE 是一种空间效应,即二个 ^1H 核之间并非通过键的作用互相偶连,而是通过空间距离接近而产生偶连。这一点非常重要,它表明由氨基酸残基连接而成的多肽链并不是完全伸展的线性长链,而是有弯曲、盘旋或者其它力或键(例如二硫键)的形成,使整个多肽(或蛋白)缩合成球状或其它比较紧密的结构,以致于不但相邻近的残基之间、而且距离很远的二个残基之间也可以产生偶连,从而出现 NOESY 交叉峰。而要把一个蛋白质中所有的质子间距离都测定出来也不是一件很容易的事,只有在提高 2D-NMR 分辨率与有效性之后才能做到,这方面已经取得了很大进步,目前还在发展之中。

用距离几何计算法定蛋白结构

从理论上说,如果每二个质子之间的距离都为已知,则这个整体结构就可以计算出来了,这种算法称为距离几何算法(distance geometry algorithms)。此算法在数学研究中早已建立,只是把它和蛋白质的结构测定结合起来还是七十年代以后发展起来的。原因是对于蛋白质来说情况要复杂得多,从 NOESY 交叉峰得到的二个质子,可能在多肽键的骨架上,也可能在某一氨基酸残基的侧链上。除了质子间距离这一参数之外还需要考虑很多因素。例如二个残基以肽键相连结时,由于共轭效应,肽键必须在一个平面之内;侧链中的基团可以围绕某一个键作旋转运动,但由于空间位阻关系,并不是任意转角都是被允许的;即使在能被允许的转角范围内,还有一个总体能量一般应取最小值即最稳定状态的问题等等。

由于计算机技术在各领域中的应用及其不断发展,现在上述困难逐步得到了解决。例如关于一个简单分子的结构,把各种限制性因素作为基本要求输入,从图象上就可以清楚地看到各种被允许的结构状态。这种分子图象技术显然对于解决蛋白质溶液构象提供了良好的条件。此外计算机运算速度的提高也是一个很重要的因素。

有了上述条件,就可以把从 2D-NMR 所得到的全部距离数据,再加上各种限制因素输入计算机,由此得出溶液中蛋白质的构象模型。

从这里可以看出用 X 射线衍射法和 2D-NMR 法不同,X 射线衍射法通过原子的散射得出结构,由它测定每个原子的位置从而求得总的结构;而 2D-NMR 则是通过许多 ^1H 原子核之间的距离,利用各个距离的制约关系来求得总的结构。它们都能得出蛋白质的三维立体结构图象。由于 2D-NMR 法本身发展较晚,目前分子量在 10,000 以下的蛋白质结构已能测定,随着方法的改善,很快可以扩大到分子量为 20,000 的蛋白质,如果再结合其它技术(例如同位素标记)可达到

30,000 至 40,000。但它的突出优点是测定溶液中的蛋白质构象,这是 X 射线衍射无法比拟的。

2D-NMR 测定结构与分子动力学

不论是用 X 射线衍射法或 2D-NMR 法所得到的结构都是代表着蛋白质分子在某一特定条件下的时间平均位置。实际上,每个原子受到周围其它原子的作用,其位置不是固定的,它所受到的合力类似于弹性力,使原子不断振荡运动,因此从某一瞬间看,处于某一种状态,而在经过一段短时间(例如 10^{-12} 秒)之后则处于另一种状态。连续不断地取不同时间,将会得出不同的结构。这方面的研究是八十年代发展起来的又一个关于分子结构研究的第三个阶段,被称为分子动力学。分子的这种运动在晶体结构研究这第一阶段影响不太大,而在溶液构象研究的第二阶段显然是一个不能不考虑的因素。这首先表现在二个 ^1H 原子核之间距离的变化上,因此任意二个核之间的距离 $|r_i - r_j|$ 实际上在上限 U_{ij} 与下限 L_{ij} 之间变化,即

$$L_{ij} \leq |r_i - r_j| \leq U_{ij}$$

而 2D-NMR 法中由于应用了距离几何计算,所得到的构象

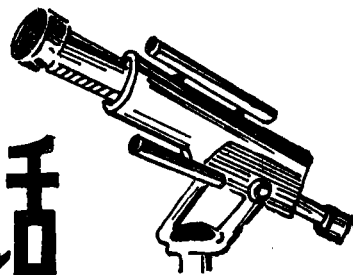
$$\{r_i = (x_i, y_i, z_i); i = 1, 2, \dots, N\}$$

显然和输入的各距离参数有关。这就决定了用 2D-NMR 测定结构的精确度,实际上得到的结构,其精确度是一对对溶液中距离几何关系的均方根距离(RM SD)的平均值。例如对于一种分子量较小的蛋白质 Tandamist 来说,多肽骨架的 RMSD 值为 0.85 \AA ,骨架以及蛋白质内部的氨基酸侧链为 1.04 \AA ,而对所有的重原子则为 1.53 \AA 。这些数值的不同说明,分子的不同区域其局部 RMSD 很不相同。内部侧链和骨架测定结果都很好,但在蛋白质的表面,则无序性增大。产生这一现象的原因既可能是在距离几何计算时所输入之值有不同的范围,也可能是在某些分子区域距离的限制性较少,也就是分子动力学的差别。这类研究不仅有助于了解结构本身不是一种刚性的而是具有一定活动程度的这一事实,而且进一步还将有助于了解分子的不同区域在蛋白质发挥功能时的构象变化。而这正是近代分子生物学要求解决但还没有彻底了解的问题,也是 2D-NMR 的进一步发展要以局部 RMSD、静态与动态无序性之间的关系作为重要方向的原因。

2D-NMR 的现状与展望

用 2D-NMR 测定溶液中蛋白质或核酸的构象是八十年代发展起来的一种新技术。第一个被测定的溶液中球蛋白结构是 1985 年 Williamson, Havel 和 Wüthrich 所报道的,可见历史还很短。但是它的出现却具有很重要的意义,因为它对生物分子结构的测定已经跨越了第一阶段即晶体结构的阶段,标志着人们对于更接近于活体状态下的分子已经开始有办法精确

关于小行星的谈话



· 卞 德 培 ·

编者按: 1989 年底之前, 传闻一颗小行星将撞击地球, 惹起许多人的恐慌。消息很快得到了澄清, 原来是场以讹传讹的虚惊。

甲: 那颗一时成了新闻人物的小行星, 究竟是怎么回事?

乙: 那是颗不大的小行星, 直径约 220 米。1989 年 3 月 31 日, 两位美国天文学家霍特和托马斯, 在搜寻近地小行星时, 发现了它, 它的临时编号为 1989 FC。经过几天的观测, 它的轨道根数被确定下来。他们惊讶地发现, 8 天多之前, 它曾一度走到离地球只 69 万公里处, 还不到地、月间距离的两倍。天文距离动辄以亿万公里或若干光年来计量, 69 万公里可说是近在咫尺, 传说小行星要撞击地球实源于此。不过, 对于直径 12700 多公里的地球来说, 1989 FC 还远在地球直径五六十倍之外呢!

甲: 刚才说到小行星轨道根数, 指的是什么?

乙: 为了描述一个天体的运行, 以及它轨道的形

状、大小和在空间的位置等, 就需要掌握其轨道的一些参数, 叫做轨道根数。它们是: 轨道半长径 a , 偏心率 e , 倾角 i , 升交点黄经 Ω , 近日距 q , 近日点角距 ω , 周期 P 和过近日点时刻 T 等。1989 FC 的这些根数是:

a : 1.04 天文单位 e : 0.36 i : 5°
 q : 0.65447 天文单位 Ω : 180° ω : 255°
 T : 1989 年 1 月 13 日 P : 1.033 年

甲: 过去有过比 1989 FC 更接近地球的小行星吗?

乙: 还从来没有过。在此之前, 一直保持着离地球最近记录的小行星是“赫姆斯”, 1937 年 10 月 30 日, 它创记录地从距离地球 80 万公里的近处经过。它和 1989 FC 都属于阿波罗型小行星, 或者叫近地小行星。

测定了。当然, 它的精确度还在不断提高之中, 暂时还不能完全和 X 射线衍射相比拟, 因此可以说是生物分子晶体结构测定的一种重要补充。但是由于它具有能直接了解溶液中的结构这一突出优点, 以及作为和分子动力学研究的结合从而成为结构与功能相互关系的桥梁这一重要作用, 受到了世界各国科学家的重视。因此近几年来在国际上发展很迅速。精确度也在逐步提高, 对某些蛋白质或其局部, 溶液中 2D-NMR 测定已经可以和 2\AA 分辨率的 X 射线衍射的精确度相比。其结果是一方面通过二种方法所得结构相同或相近, 说明它们的等同性; 另一方面用 2D-NMR 反映不同区域的无序性不同, 特别是蛋白质表面和内部侧链之间的区别 (如多肽激素胰高血糖素和金属原子比较丰富的分子), 从而成为一种极有用的补充信息。除此以外, 2D-NMR 法还对用化学方法测定的氨基酸序列分析结果可以加以检验, 以纠正以往不正确的理解, 还

有一点值得提出的是 2D-NMR 分析的结果有时不一定要非常完整 (这涉及繁杂的计算过程) 就可以得到有用的信息。例如前面提到的短杆菌肽 S 在溶液中是一种环状平面型结构, 但和脂相互作用后可以发现鸟氨酸两侧的缬氨酸和亮氨酸之间也有 NOESY 交叉峰出现, 这说明在和脂作用后平面构型以鸟氨酸为中心二侧成分折叠起来插入到脂中的可能性, 也是短杆菌肽 S 和细胞膜作用后的一种重要构象变化。要知道本文中所谈到的都还是蛋白质在溶液中的构象, 但是人们进一步感到极大兴趣的是在细胞膜中它们的构象如何? 在与外界物质作用后又有什么变化? 这些问题都涉及到生命的重大基本课题, 例如分子识别和信息传递等。然而至今还没有找到合适的手段。也许 2D-NMR 的改进和发展是解决这类问题的有希望途径之一。