

专论与综述

研究溶液中蛋白质三维结构的新方法 ——二维核磁共振技术

华庆新 王家槐

(中国科学院生物物理研究所, 北京)

近年来, 国际生物物理学会秘书长、瑞士苏黎世高等工业学院分子生物学与生物物理研究所 Wüthrich 教授及其同事, 发展了一整套以二维核磁共振(2D-NMR)技术测定溶液中蛋白质三维结构的新方法, 受到了国际学术界的高度重视^[1-4]。去年, Wüthrich 教授应我国生物物理学会之邀来华访问, 在中国科学院生物物理所报告了他的工作。

目前, 解析蛋白质三维结构的最有力方法, 当然仍推 X 射线晶体结构分析, 因为它能精确地确定生物大分子中各原子的座标, 确定各共价键的键长、键角。但它毕竟须在晶体条件下测定, 因而反映的主要还是静态结构。而生命现象的复杂多变的功能, 是在溶液态完成, 并与生物大分子在溶液态的多变的结构达到高度完善的统一。自七十年代以来, 随着核磁共振的一整套新技术的发展、完善, 核磁共振技术已能在溶液态或非晶态(如脂介质)中测定生物大分子三维结构及其动力学变化^[5], 因而引起学术界的重视。目前, X 射线晶体结构分析与高分辨率脉冲傅立叶核磁共振技术已成为测定蛋白质等生物高分子结构的两项极其重要的互补手段。

但是, 常规的核磁共振技术并不足以将分子量较大的蛋白质分子各吸收峰一一分辨开。目前世界上最高场强的核磁共振仪可达 600 兆赫。而一个几十万或几百万分子量的蛋白质, 拥有成千上万个质子, 其相邻质子间复杂偶合又情况各异, 即使在 600 兆赫的谱仪上仍不能将每个质子峰分开, 更难以逐个识别吸收峰了。

而近三、四年 2D-NMR 技术发展极快, 提供了一条全新的解析高分子量蛋白质的复杂图谱的途径。从 NMR 技术的发展看, 六十年代是连续波谱仪时代, 七十年代是脉冲傅立叶技术和双共振技术时代, 八十年代则可说是二维谱和自旋成像时代。Wüthrich 教授及其同事用 2D-NMR 作序列共振峰识别 (Sequential resonance assignment) 测定蛋白质完整的空间结构, 在碱性胰蛋白酶抑制剂 (BPTI)^[3], 胰高血糖素 (Glucagon)^[4] 和乳糖阻遏蛋白 (Lac-repressor)^[6] 上取得成功, 开辟了一条测定蛋白质溶液结构的崭新道路。目前此工作虽处于开创阶段, 但这一整套技术和思路已显示其蓬勃的生命力。

本文拟介绍 2D-NMR 的一般概念, Wüthrich 解析蛋白质三维结构的基本思路及某些实际应用。

(一)

2D-NMR 有别于常规一维核磁共振的主要点, 在于一维 NMR 只涉及一个频率变量, 是吸收峰强度对一个频率变量作图, 而二维谱则代表两个独立频率 ω_1 和 ω_2 的信号函数 $S(\omega_1, \omega_2)$, 是吸收峰强度对两个频率变量作图。二维图谱有两类表达方式, 图 1 是堆积图, 看起来像三维立体图。图 2 为等高线图, 与地形图上等高线类似。这两类图包含同样的信息, 立体的堆积图较直观, 但分析时往往主要用等高线图。有时一维 NMR 也可涉及时间、温度、浓度等参数

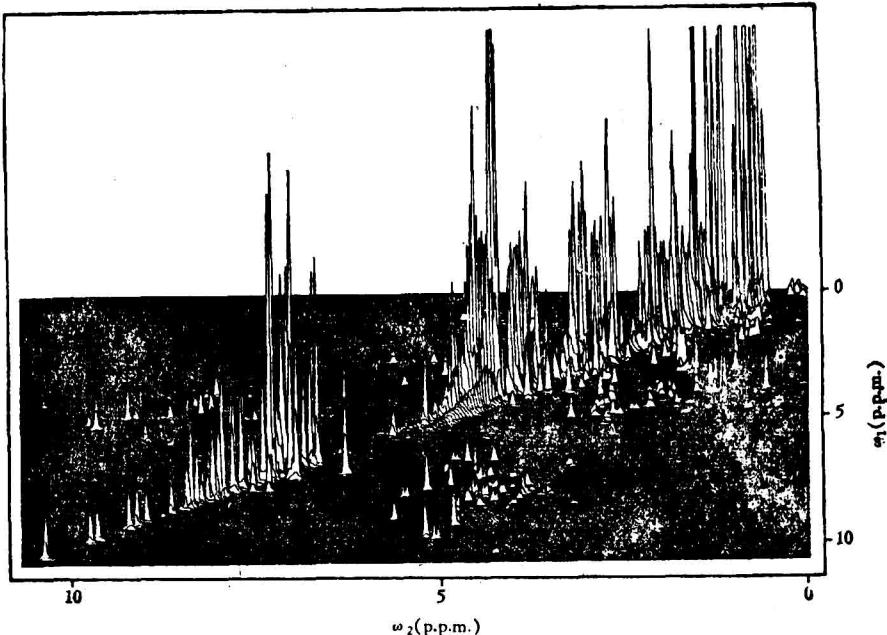


图 1 20 nM BPTI 的 500MHz¹H COSY 谱堆积图^[1]
pH4.6, 80°C, 于 90%H₂O 和 10%D₂O 混合溶剂中

作二维图谱,但只涉及一个频率变量,不能认为是 2D-NMR,而只代表一维 NMR 的二维堆积图。

2D-NMR 可分为二维分解图谱 (2D-resolved spectroscopy)、二维相关图谱 (2D-correlated spectroscopy 简称 COSY)、二维自旋回波相关谱 (2D-spinecho-correlated spectroscopy 简称 SECSY)、二维 NOE 谱 (2D-Nuclear Overhauser Enhancement spectroscopy 简称 NOESY) 等。

二维分解谱与双相层析原理相似。在第一个溶剂层析后,对不能完全分开的样品,可在垂直于第一个溶剂的展层方向上,作第二个溶剂层析。NMR 谱含有化学位移 δ 和偶合常数 J 等类信息,偶合亦有同核偶合与异核偶合。一维 NMR 分不开这些相互偶合的峰,可再加另一与原频率无关的频率变量。由于谱线强度是 J 和 δ 的函数,谱图即可以吸收峰强度对两个频率变量作图,得吸收峰的三维表达。这种分解谱可简化复杂的 NMR 图谱,使之便于分析,而又不减少其信息。

二维相关谱(包括 COSY 和 SECSY 谱)则可增加谱图的信息成份,其相应峰数和峰的复

杂性也都增加。由 2D-COSY 可得到蛋白质的¹H-¹H 偶合关系,确定氨基酸残基的质子自旋体系。一般 2D-COSY 谱可由两个非选择性 90° 脉冲序列得到^[7-9]即 $(90^\circ - t_1 - 90^\circ - t_2)_n$ 序列。

第一个 90° 脉冲使磁化强度矢量横倒, 经过发展期 t_1 , 不同的磁化矢量组分各以其特有的进动频率进动。第二个 90° 脉冲引起属于同一 J-偶合体系的磁化强度矢量的转换。在第二个 90° 脉冲后立即记录下作为 t_2 函数的自由感应衰减信号 FID。实验以一组等距离的 t_1 值测定。即对一个 t_1 值, 记录下作为 t_2 函数的 FID, 再换另一 t_1 值, 得到另一 FID, 等等。为得到足够大小的信噪比, 每个 t_1 值测定都要重复 n 次作累加。在每次记录的末尾, 经过 1—1.5 秒的弛豫延迟, 使自旋体系确实恢复平衡。在数据采集后 $S(t_1, t_2)$ 作二维 FT 变换, 将时间域的函数变为频率域的函数 $S(\omega_1 \omega_2)$ 。

生物学工作者更关心的是这些二维谱所提供的结构信息。以图 1 即 BPTI 的 COSY 堆积图为例,其纵、横轴均代表化学位移。自右至左,自上至下增大。在 $\omega_1 = \omega_2$ 的从右上至左

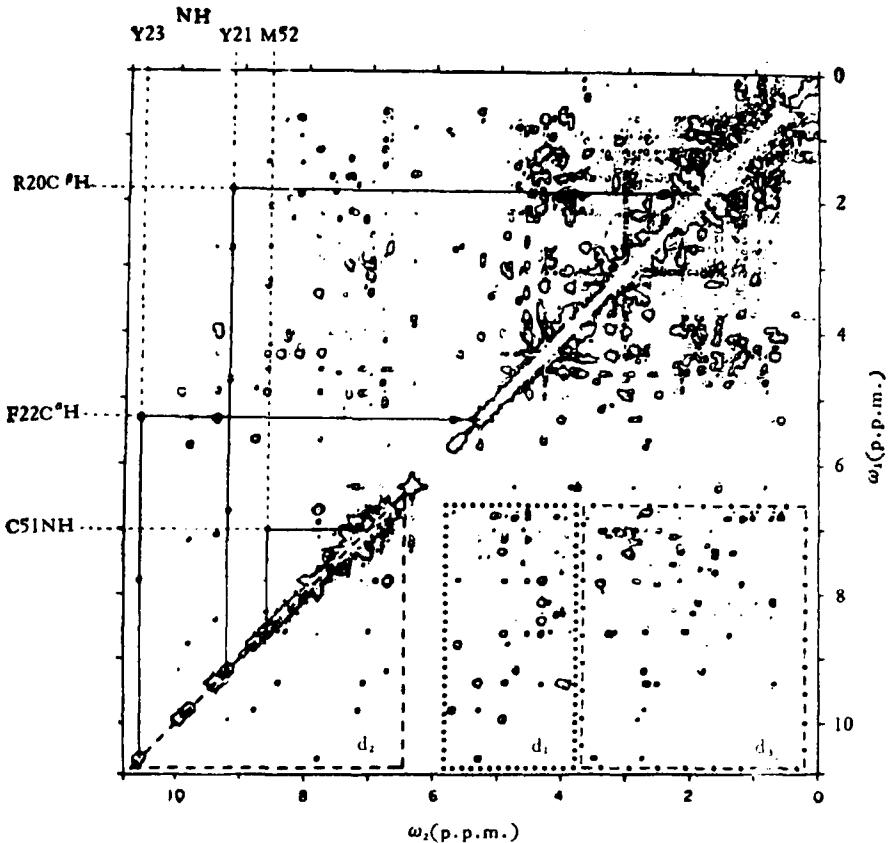


图 2 20 mM BPTI 的 500MHz ^1H - ^1H NOESY 谱等高线图^[3]

pD4.6, 36°C, D₂O 中

下的对角线上的峰，称对角线峰，实际它就是一维谱峰。最右上角是甲基峰，然后是次甲基峰和 α 碳原子上的质子峰 C^aH, 6—8ppm 区间为芳香环质子峰，左下端是酰胺质子峰。在重水 D₂O 中经一定时间的重氢交换，-NH 质子为氘取代。此酰胺质子峰不可见。在普通水 H₂O 中，则有此峰，且有 C^aH-NH 的 ^1H - ^1H 偶合。整个 COSY 谱以对角线为对称轴，其右下三角区与左上三角区完全等同，包含同样多的信息。而 $\omega_1 \neq \omega_2$ 的对角线以外的峰称为交叉峰，代表一维谱中两相互偶合的峰。即由一交叉峰作水平线，与对角线相交的对角线峰，与由此交叉峰作垂线与对角线相交的对角线峰偶合。分处 2D-COSY 谱不同区域的交叉峰，代表不同类型偶合，既可是 C^aH-C^bH 的偶合，也可是 NH-C_aH 偶合，C^bH-C^cH 偶合，以至芳环质子间的偶合。由此可见 2D-COSY 谱比一维谱峰数多，提供的信息多。它能提供多肽链上间隔三

个共价键的 ^1H - ^1H 偶合的完整图象。在 COSY 谱中特别重要的是 C^aH-NH 酰胺键上的质子偶合，它对确定蛋白质分子的二级结构特别重要。COSY 谱还可测氨基酸的自旋体系^[4]，这对蛋白质的谱线识别很有用。SECSY 谱与 COSY 谱在技术实现上略有不同，但所包含信息一样。

2D-NOESY 谱是由三个非选择性 90° 脉冲序列而得^[3,10]，即 $(90^\circ - \tau_1 - 90^\circ - \tau_m - 90^\circ - \tau_2)$ 序列。

在第一个 90° 脉冲后，经过发展期，各磁化强度矢量成份各以其特有的进动频率进动。在固定的 τ_m 混合期中，交叉弛豫导致空间上相邻的质子之间磁化矢量的交换。在第三个 90° 脉冲后，FID 作为 τ_2 的函数立即被记录。与 COSY 谱一样，NOESY 实验亦要以一组等长距离的 τ_1 作多次实验，每次都是累加至信号足够大。采样后作二维 FT 变换。图 2 中对角线峰

也就是一维谱峰，而对角线以外的峰则为 NOE 交叉峰，即代表不超过 5 \AA 距离的，在空间上（不通过共价键）接近的两质子之间有 NOE 联系。由 NOE 交叉峰作水平线和垂线，与对角线相交的两个对角线峰之间有 NOE 联系，其间距不超过 5 \AA 。NOE 强度是与两核间距的六次方成反比。图 2 中 NOE 交叉峰有三类。6.5—10 ppm 的 d_2 区，代表不同酰胺质子之间有 NOE 联系；3.8—5.8 ppm 的 d_1 区，代表酰胺质子与 $\text{C}^\alpha\text{-H}$ 质子之间有 NOE 联系；而在下端的 d_3 区，代表酰胺质子与 $\text{C}^\beta\text{-H}$ 质子之间有 NOE 联系。

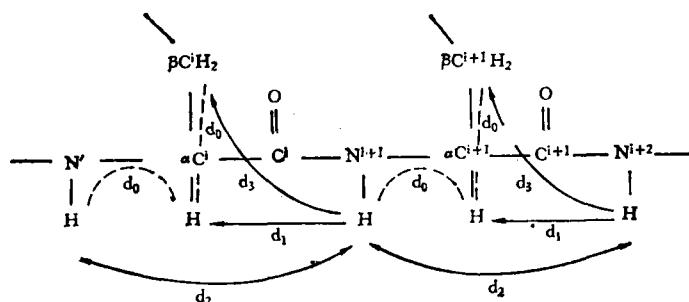


图 3 多肽链上残基内及相邻残基间的 ^1H - ^1H 偶合及 NOE 联系图

d_0 、 d_1 、 d_2 、 d_3 ，详见正文

图 3 为 COSY 或 SECSY 谱及 NOESY 谱所反映的偶合和 NOE 联系。 d_0 为 COSY、SECSY 反映的相隔不超过三个共价键的质子间的偶合，即只反映同一残基内的 ^1H - ^1H 偶合。而分属不同残基的质子起码相隔四个共价键，其偶合太弱。图 3 中 d_1 、 d_2 、 d_3 即为 NOESY 反映的多肽链上相距小于 5 \AA 的不同质子间，通过空间而不是通过共价键的 NOE 联系。注意，这里没有画出相距更远或两条肽链之间的 d_1 、 d_2 、 d_3 。实际上由于蛋白质卷曲，使序列间隔很远的残基间会有 NOE 联系。而正基于这点，Wüthrich 等人即通过 COSY、NOESY 谱的结合应用，沿肽链顺序，一个一个地识别出各共振峰，并分析推测其溶液三维结构。

(二)

从 2D-NMR 测定溶液中蛋白质三维结构

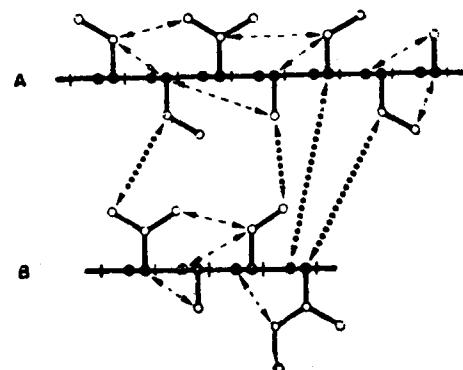


图 4 以 ^1H NMR 测定蛋白质空间结构的策略图示^[11]
A, B 为两肽段，短竖线分开相邻残基；●——肽链骨架上的 $\text{C}^\alpha\text{-H}$ 和 NH 基团；○——氨基酸侧链上的 CH 基团。余见正文

的要点，亦即 Wüthrich 方法的总体策略如图 4 所示。

首先按多肽链上氨基酸的排列顺序把每个氨基酸残基上的各质子（ CH 和 NH ）的共振峰识别出。然后，由于多肽链在卷曲过程中某些相邻残基可形成 α 螺旋和 β 折迭等二级结构，使这些

相邻残基的酰胺质子之间或酰胺质子与 $\text{C}^\alpha\text{-H}$ 质子之间处于较短的距离范围内（如图 4 中 A 肽段或 B 肽段内以 \longleftrightarrow 相连者）。当这些二级结构单位进一步卷曲，可导致一级结构上相去甚远的肽段相互靠拢，形成三维结构，从而使图 4 中 A 肽段与 B 肽段之间一些质子（以 \longleftrightarrow 相连者）的空间距离靠拢。当所有这些近程或远程的残基上质子间距落人 5 \AA 范围内，它们的 ^1H - ^1H NOE 联系均可在 NOESY 谱上得到反映。亦即对 NOESY 谱的仔细分析，可提供不同残基上各质子间的距离制约信息，它们反映了蛋白质不同部位二级结构和三维卷曲的状况。由于蛋白质分子的三维结构既可用其每个组成原子的整套座标参数表达，也可以用各原子间的距离制约条件描述，两者可以互相换算^[11]。这样，我们就可从 2D-NMR 获得蛋白质三维结构的粗模型。再借助于能量修正等计算技术^[12]，此模型还可进一步修正、改善。下面对如何按氨基酸排列顺序识别共振峰及测定二级

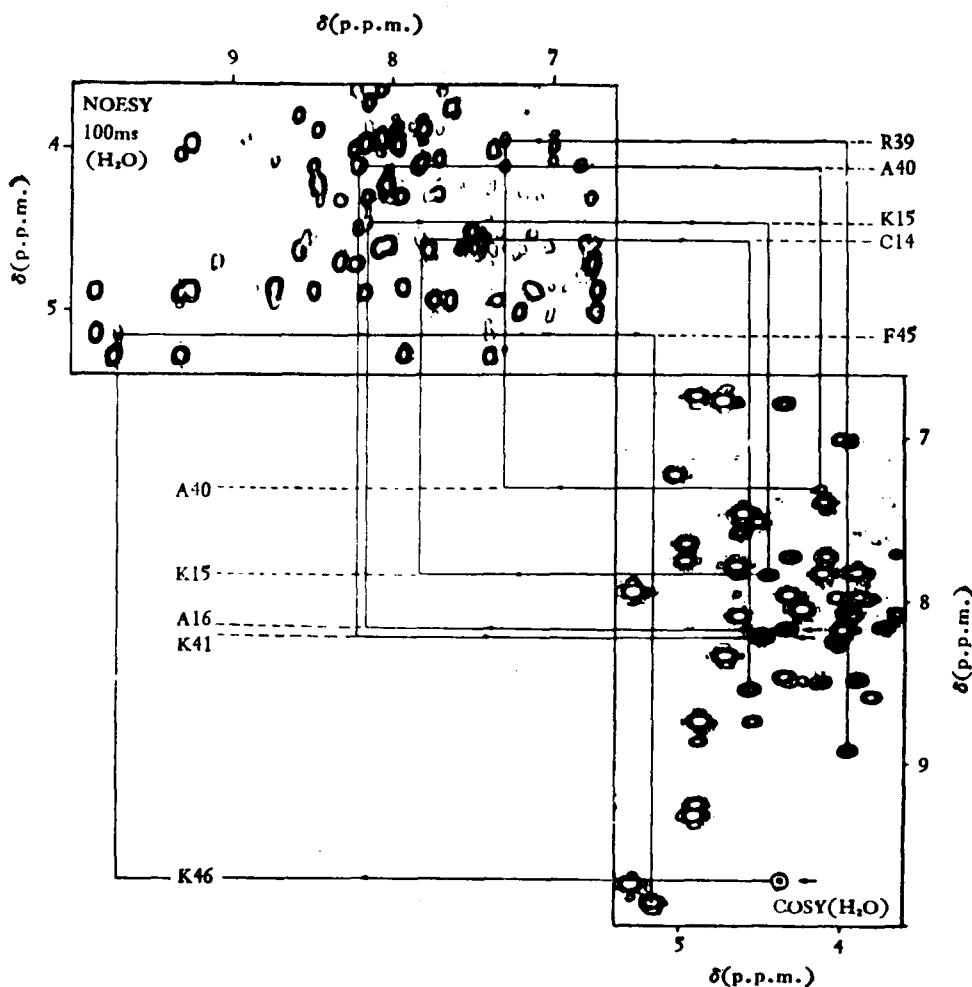


图 5 20mM BPTI 的 500MHz ^1H 谱 COSY-NOESY 结合图^[1]
左上为 NOESY 的 d_1 区, 右下为相应 COSY 区 (H_2O 为溶剂, 68°C)

结构分三步介绍如下。

第一步, 按氨基酸残基顺序识别共振峰, 这是整个测定最基础的一步。若在 COSY 谱上先认出与某残基相应的共振峰, 则可参照一级结构, 综合应用 COSY 谱与 NOESY 谱来识别各共振峰。由于 COSY 和 NOESY 谱都按对角线对称, 沿对角线裁开的半张图已包含全部信息, 而同样实验条件下 COSY 和 NOESY 的对角线峰相等, 都等于其一维谱峰。Wüthrich 巧妙地利用这点, 在对角线右下方放半张 COSY 谱, 在左上方放半张 NOESY 谱, 使分析工作大大简化。

图 5 为该结合图的一部分, 只包含 $\text{C}^{\alpha}\text{H}$ -

NH 交叉峰。在 COSY 谱上若已识别出 Lys-46 的 $\text{C}_{i-1}^{\alpha}\text{H}-\text{N}_i\text{H}$ 交叉峰 (图 5 右下方箭头所指处), 以此为起始点, 向左作水平线与对角线的交点应是 Lys-46 酰胺基上质子峰。由此作垂线向上, 在 NOESY 谱上碰到的唯一的峰即 $\text{C}_{i-1}^{\alpha}\text{H}-\text{N}_i\text{H}$ 交叉峰, 则必为 Lys-46 的 NH 与前一残基 Phe-45 的 $\text{C}^{\alpha}\text{H}$ 交叉峰、再沿此峰向右作水平线与对角线的交点为 Phe-45 $\text{C}^{\alpha}\text{H}$ 质子峰。由此交点作垂线向下, 在 COSY 谱上碰到的唯一的一个峰必为 Phe-45 的 $\text{C}^{\alpha}\text{H}-\text{NH}$ 交叉峰。由此, 我们已看到序列共振识别的一般方法。图 5 还显示如何由 41 位残基识别到 39 位; 从 16 位残基识别到 14 位的共振峰的路线。

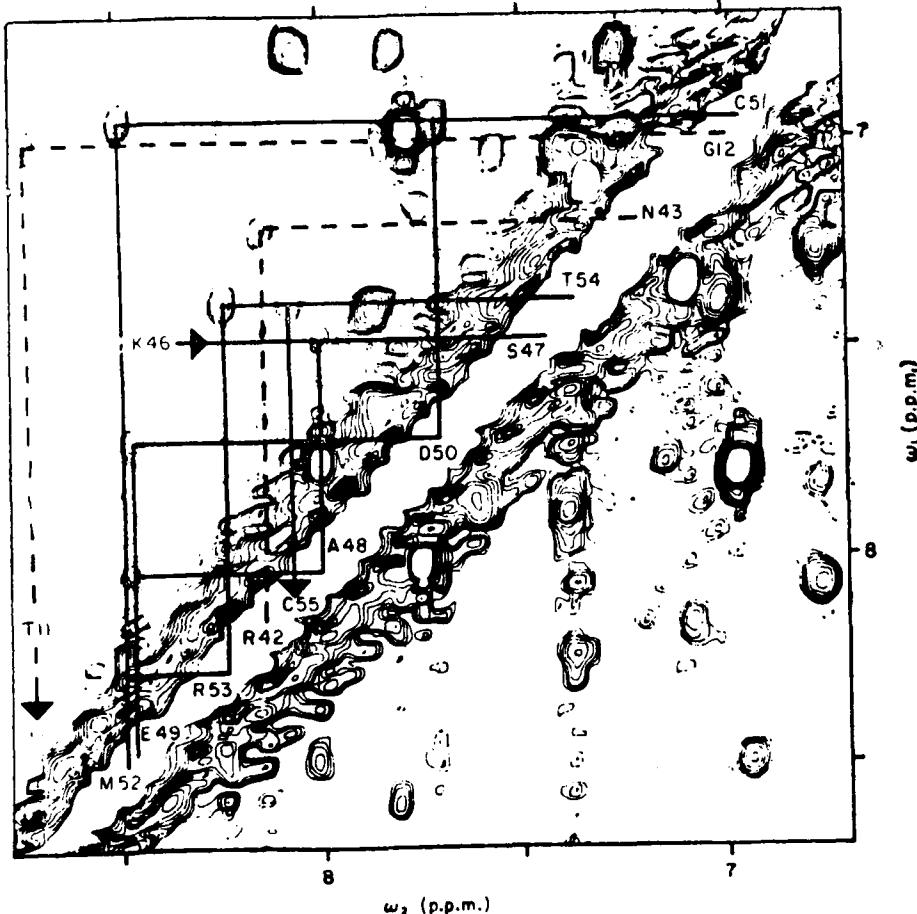


图 6 BPTI 500MHz¹H NDESY 谱的 d_2 区^[3]

其余条件同图 5

实际上由于分辨率等各种原因，单用 $C^{\alpha}H-NH$ 交叉峰区不能解决全部问题，有时要用 NOESY 谱中 $C^{\beta}H-NH$ 交叉峰区，与 COSY 谱中同样区域相配合，有时可单独使用 NOESY 谱中 $N_iH-N_{i+1}H$ 交叉峰区。

图 6 即为用 $NH-NH$ 区识别 BPTI 中 46—55 位残基的共振峰。这里，从对角线上一个 NH 的质子峰出发，会碰到两个交叉峰，它们相应于与前、后两个残基的 NH 的 NOE 联系。假定已识别出其中一个峰，就可能由此起始点出发，沿一定方向按顺序识别下去。图 6 从 LYS-46 出发，与对角线相交的 NH 峰是 Ser-47（因已知 Phe-45 的 NH 不在此处），其间通过的交叉峰只能是 Ser-47 与 Ala-48 之间的 $NH-NH$ 峰。由此峰作垂线与对角线的交

点为 Ala-48 的 NH 质子峰，可按此法顺序识别下去。

第二步，找出起始点，对氨基酸残基侧链共振峰作识别。欲作序列共振识别，必先找出起始点，这主要靠 COSY 谱。前已指出，COSY 谱中还包含 $C^{\alpha}H-C^{\beta}H$ ， $C^{\beta}H-C^{\gamma}H$ 及芳环质子间的偶合峰，可用作识别各氨基酸侧链的自旋体系。反言之，若由 COSY 谱测知某残基的自旋体系，则也提供了这是何残基的重要信息。为此，可将样品溶于重水中，使酰胺质子充分为氘取代，则主链酰胺键质子峰消失，使图谱更简化。在这种 COSY 谱中，Gly、Ala、Val、Ile、Thr 等有特征的谱线。若某个蛋白质中只含有一个 Ala，则借助于侧链自旋体系的识别，可在 COSY 谱上找出该 Ala 的共振峰作为序列识别。

别的起始点。另外，亦可以蛋白质中某些特征的小肽段作起始点。在一维 NMR 测定中，用蛋白质化学修饰或作氨基酸取代，或由不同种属的同一蛋白的 NMR 谱线比较，都可识别出相当一些残基的共振峰，作 2D-NMR 图谱解析的起始点。

第三步，二级结构测定。在氨基酸残基序列共振识别的基础上，可进一步识别具规则二级结构的肽段。这里有两个方法。一是样品于重水后立即记录 COSY 谱，使不位于二级结构区的酰胺质子迅速为氘取代，而参与二级结构中氢键作用的酰胺质子还来不及与重氢交换，仍在 COSY 谱上显示。此时，凡 COSY 谱中能观察到酰胺质子峰的肽段极有可能是二级结构区。二是基于二级结构区的构象特点。 β 结构是一种伸展的肽链，主链质子之间相距一般在 4 Å 以上，唯一的例外是 $C_i^aH - N_{i+1}H$ 的间距（即 d_1 ）为 2.2 Å。 α 螺旋是一种扭折的肽链，相邻酰胺质子间距为 2.8 Å，而 $C_i^aH - N_{i+3}H$ 之间间距，即 $d_1(i, i+3)$ 也只有 3.4 Å。因此，若在 NOESY 谱上观察到连续的强的 d_1 联系，则极有可能是 β 折迭；若观察到一段强的 d_2 和 $d_1(i, i+3)$ 联系区，则很可能是 α 螺旋。其具体分析方法可参看文献^[6]。

(三)

下面以乳糖阻遏蛋白与 DNA 结合区的探测为例，说明 Wüthrich 方法如何成功地用于测定溶液中蛋白质三维结构，从而对一些生物学问题提供了重要信息。

近年来，用 X 射线晶体学方法已测定了三个基因调控蛋白的三维结构。它们是 λ 噬菌体的一对阻遏蛋白——Cro 蛋白和 CI 蛋白的 DNA 结合域以及大肠杆菌中异化作用基因激活蛋白（Catabolite gene activator protein，即 CAP）。这三个蛋白的一个共同特点是，其与 DNA 结合的部位都是一个螺旋—转折—螺旋的“V”字型结构单位，尤其是 Cro 和 CI 蛋白的第二段 α 螺旋正好可嵌入 B-DNA 的深沟中，使 DNA 与蛋白之间得以进一步发生特异性识

别作用^[13]。根据一级结构相似性的研究，人们推测，这种螺旋—转折—螺旋单位很可能在其他许多与 DNA 有专一性作用的基因调控蛋白包括乳糖阻遏蛋白中普遍存在^[14]。这是近年来在核酸与蛋白相互作用研究中的一个重要发现。问题是要从实验上判定这种普遍性究竟是否存在。很多年来，试图得到乳糖阻遏蛋白晶体的努力都失败了，晶体学工作无法开展。而 Wüthrich 及其同事用 2D-NMR 在溶液态确实测得乳糖阻遏蛋白的 DNA 结合域有两段相邻的 α 螺旋^[6]，其位置与根据一级结构比较研究所作的预测一致。他们还从 NOE 联系探测到这两段相邻的 α 螺旋也是通过中间的转折肽形成“V”字形结构单位。这是对蛋白质与核酸相互作用研究的一个重要贡献。

Wüthrich 还对 Glucagon 作了 2D-NMR 测定^[4]，观察到这个只有 29 个残基的小肽与十二烷基磷酸胆碱（DPC）混合形成微泡时，其 19—29 位残基取 α 螺旋构象，而分子 N 端则为较伸展的构象，从而探测到在活性状态下的构象。因篇幅所限，不再详述。

总之，2D-NMR 与 X 射线结构分析互为补充，二者是目前观察蛋白质三维结构的最主要方法。而 2D-NMR 随着方法的日臻完善，它对测定蛋白质溶液构象的极端重要性将更趋明显。

参 考 文 献

- [1] Wüthrich, K., Wider, G., Wagner, G., and Braun, W.: *J. Mol. Biol.*, **155**, 311—319, 1982.
- [2] Billeter, M., Braun, W., and Wüthrich, K.: *ibid.*, **155**, 321—346, 1982.
- [3] Wagner, G., and Wüthrich, K.: *ibid.*, **155**, 347—366, 1982.
- [4] Wider, G., Lee, K-H, and Wüthrich.: *ibid.*, **155**, 367—388, 1982.
- [5] 华庆新：《生物化学与生物物理进展》，1981，第三期。
- [6] Zuiderweg, E. R. P. Kaptein, R., and Wüthrich, K.: *Proc. Natl. Acad. Sci.*, **80**, 5837, 1983.
- [7] Aue, W. P., Bartholdi, E. and Ernst, R. R.: *J. Chem. Phys.*, **64**, 2229—2246, 1976.
- [8] Anil Kumar, Wagner, G., Ernst, R. R. and Wüthrich, K: *Biochem. Biophys. Res. Commun.*, **96**, 1156—1163, 1980.
- [9] Wagner, G., Anil Kumar and Wüthrich, K: *Eur. J. Biochem.*, **114**, 375—384, 1981.

- [10] Anil Kumar, Ernst, R. R. and Wüthrich, K.: *Biochem. Biophys. Res. Commun.*, 95, 1—6, 1980.
- [11] Braun, W., Bösch, C., Brown, L. R., Gö, N. and Wüthrich, K.: *Biochim. Biophys. Acta*, 667, 377—396, 1981.
- [12] Levitt, M.: *J. Mol. Biol.*, 82, 393, 1974.
- [13] 王家槐:《生物化学与生物物理进展》,1984年,1期,1页。
- [14] Matthews, B. W. et al.: *Proc. Nat. Acad. Sci. U. S. A.* 79, 1428, 1982.

(本文于1982年8月7日收到)

自由基和癌

赵保路

(中国科学院生物物理研究所)

张建中

(中国科技大学研究生院)

自由基和癌的关系是很复杂的，并且一直是一个有争论的问题。但综观对这个问题研究的历史和现状，可以肯定地说，自由基和癌的关系是非常密切的。要研究癌就不能忽视自由基在致癌和癌的发展乃至治疗中的作用。研究自由基最有效的方法是电子自旋共振(ESR)波谱技术。

一、癌发展过程中自由基的变化

人们早就观察到在肿瘤生长过程中，组织里自由基的含量发生变化。有的认为癌组织中自由基含量比正常组织的大，而另一部分作者则发现组织癌变后自由基含量减少。如Kolomijisova等人观察到^[1]鼠肝肿瘤中自由基的

含量比正常组织减少大约三分之二，在肉瘤-45中亦观察到类似的结果^[1]。但Petyaev的结果^[2]则表明，在一些鼠白血病和肉瘤m-1中自由基含量明显增加。为了解决这一矛盾，很多人研究了肿瘤发展过程中自由基含量变化的动力学规律。

Saprin等人首先对鼠白血病-La发展过程中自由基含量的变化进行了动力学研究^[3](见图1)。

由图可知，开始自由基含量逐渐增加，大约第四天达到最大值。这个现象出现在白血病其它症状如脾增大、血象改变等之前。自由基变化动力学曲线的最高点和脾重量开始迅速增大的时间是一致的。在第五到六天，自由基浓度开始下降。在动物死亡时，即第七、八天，自由基含量低于正常值。

Emanuel等人亦得到类似的结果。例如，对160个患骨髓性白血病和170个患慢性淋巴性白血病人的自由基含量测定发现，它们比正常的高2到4倍^[4]。对人胃癌的研究指出，在第一期自由基含量为正常胃粘膜的3倍，从第二期开始，胃癌中自由基含量下降，在第四期明显低于正常值^[5]。

在研究了肿瘤生长的动力学规律之后，他们提出，移植肿瘤生长的早期阶段可以用一指数关系近似表示^[6]:

$$\phi = \phi_0 e^{\phi t}$$

式中 ϕ 是肿瘤在时间 t 时的体积， ϕ_0 为肿

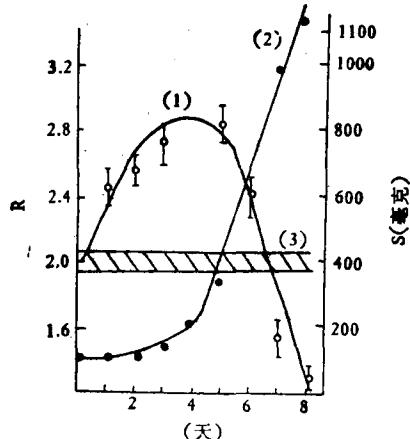


图 1

(1) 白血病鼠脾中自由基含量的变化；(2) 脾重量的变化；(3) 对照

R: ESR 信号相对大小 S: 脾重量