

1.2 蛋白质晶体结构解析



测定蛋白质结构的意义

• 人体基因的主要功能是通过蛋白质来实现的，蛋白质扮演着构筑生命大厦的主要角色。人体中大约有10万种蛋白质。

蛋白质三维结构解析方法

X-射线晶体衍射法：85.3%（重点介绍）

核磁共振波谱：14.7%（重点介绍）

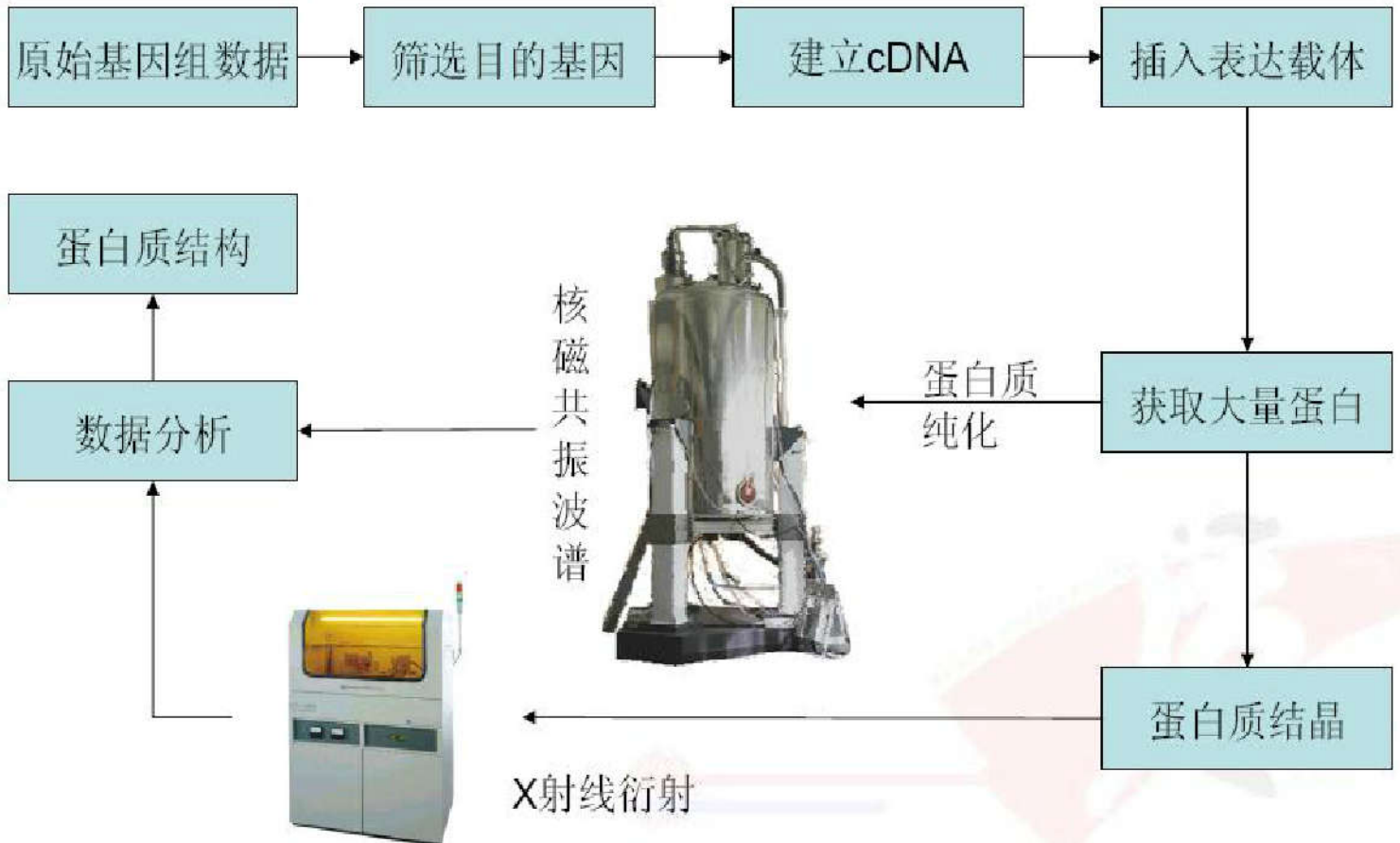
电镜三维重构、各种光谱技术、显微技术和计算机模拟

X-Ray晶体衍射目前仍然是蛋白质三维结构测定的主要方法。

优点：分辨率高，能精确确定生物大分子中各原子的坐标、键长、键角，给出生物大分子的分子结构和构型，确定活性中心的位置和结构

缺点：只能测定单晶，反映静态结构信息，无法测定溶液中的信息

蛋白质三维结构解析过程



登高必自卑，行远必自迩

(一) X-射线的发现



1895年11月8日，德国物理学家，50岁的伦琴在自己的实验室中偶然发现一种从阴极射线管中辐射出的新型射线，由于对管子发出的“东西”性质不确定，伦琴就把这种射线命名为“X射线”。



1901年第一届诺贝尔物理学奖评选时，29封推荐信中就有17封集中推荐他。伦琴最终获得了第一次诺贝尔物理学奖金。

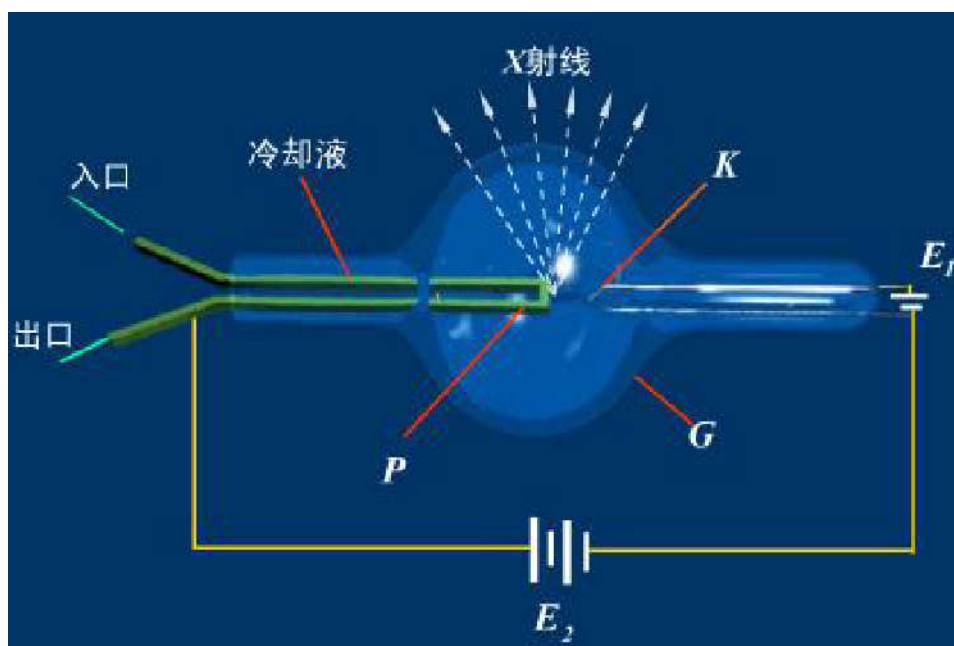


登高必自卑，行远必自迩

X射线本质



X射线是一种短波长($0.005\sim 10\text{nm}$)、高能量($2.5\times 10^5\sim 1.2\times 10^7\text{eV}$)的电磁波。它是原子内层电子在高速运动电子流冲击下，产生跃迁而发射的电磁辐射。



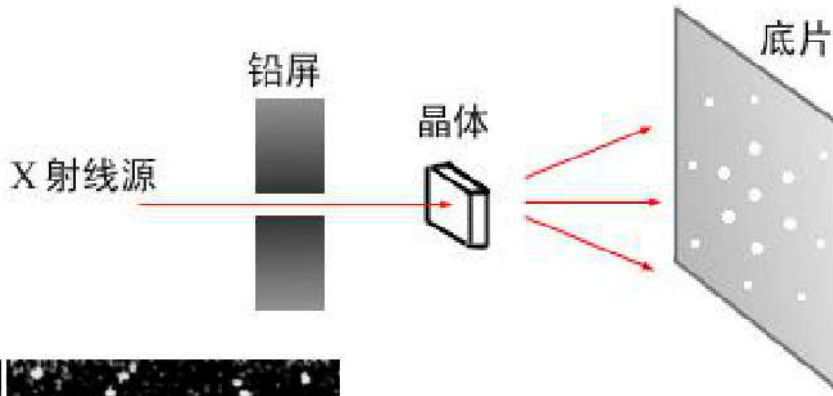
一般由高速电子撞击金属产生。如图所示，是一种产生X射线的真空管，**K**是发射电子的热阴极，**A**是由钨、钨或铜等金属制成的阳极。两极之间加有数万伏特的高电压，使电子流加速，向阳极**A**撞击而产生X射线。

登高必自卑，行远必自迩

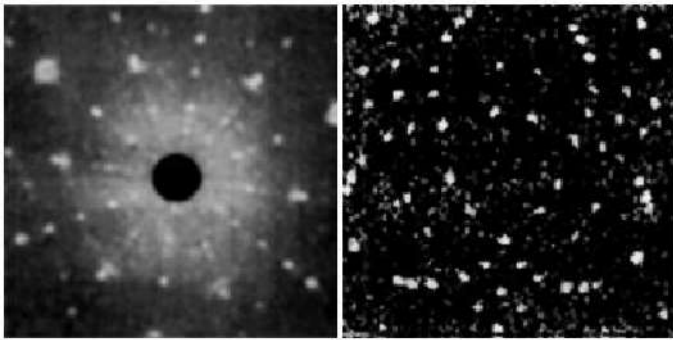
X射线衍射



1912年劳厄获得了X射线通过晶体后产生的衍射斑点图像（**劳厄衍射图**），证明了**X射线的波动性及其波长范围**。随后提出了表示原子排列周期与X射线波长间关系的著名的衍射方程（**劳厄方程**），并成功地解释了晶体衍射的实验结果。



1914年的诺贝尔物理学奖



(a)

(b)

劳厄(Laue)方程

$$a_0 (\cos\alpha_h - \cos\alpha_0) = h\lambda$$

$$b_0 (\cos\beta_k - \cos\beta_0) = k\lambda$$

$$c_0 (\cos\gamma_l - \cos\gamma_0) = l\lambda$$

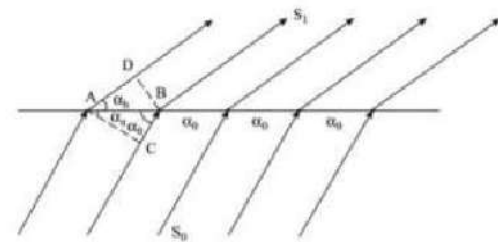


图 1.10 一行原子列对 X 射线的衍射

因在用X射线研究晶体结构方面所作出的杰出贡献，亨利·布拉格（William Henry Bragg）和劳伦斯·布拉格（William Lawrence Bragg）父子分享了1915年的诺贝尔物理学奖。

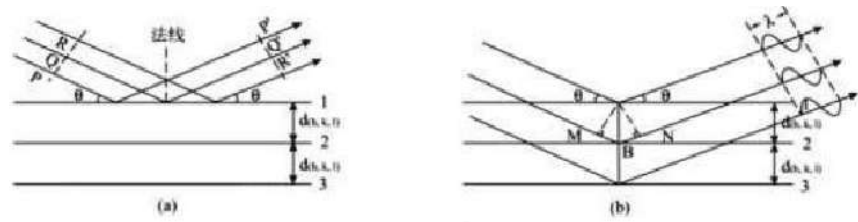
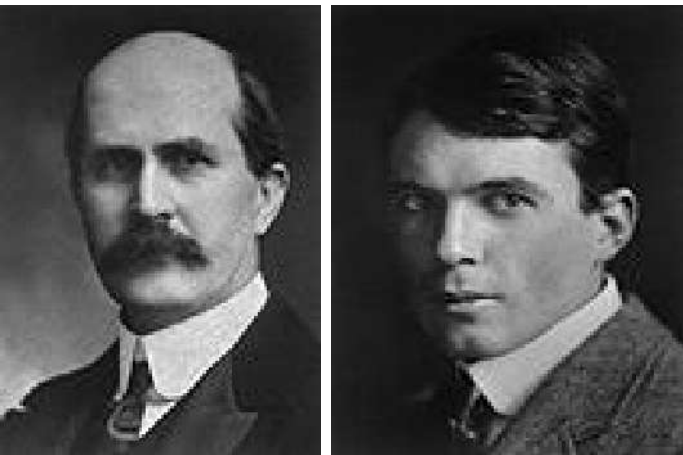


图 1.11 布拉格方程的推引

$$2 d \sin \theta = \lambda$$

$$(d = d / n)$$

X 射线衍射用于蛋白质结构的测定



1954年伯纳尔 (Bernal) 获得第一张胃蛋白酶晶体 X 衍射图片。



X-射线晶体结构分析基本原理

X射线衍射分析所依赖的基本原理是**X射线衍射现象**。

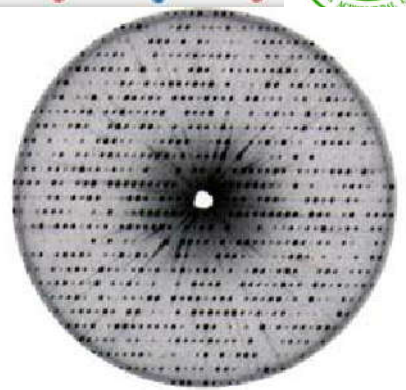
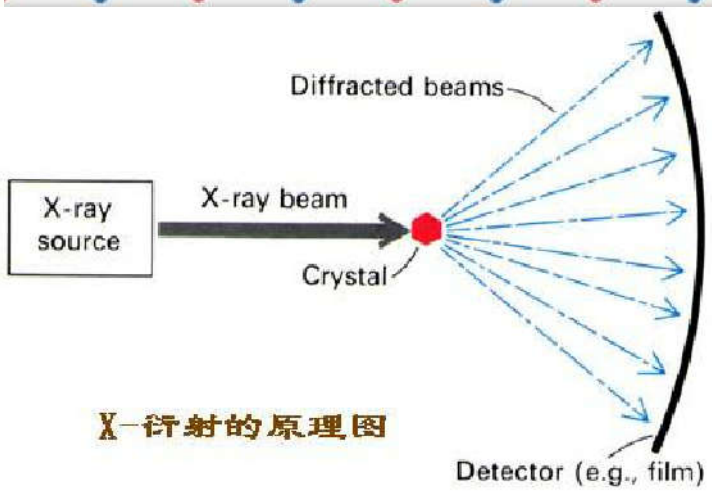
X射线衍射现象利用**X**射线的波长和晶体中原子的大小及原子间距同数量级的特性来分析晶体结构。

当**X**射线入射到样品晶体分子上时，分子上的每个原子使**X**射线发生散射，这些散射波之间相互叠加形成衍射图形。

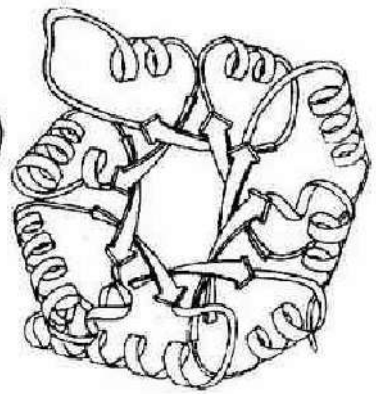
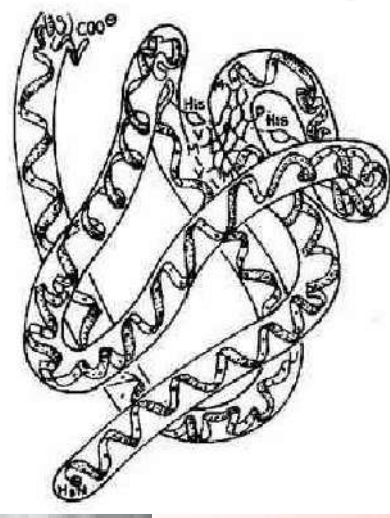
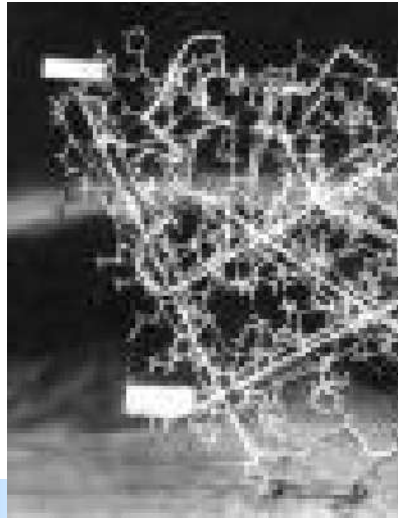
衍射图形能给出样品内部结构的许多资料，如原子间的**距离**、**键角**，分子的**立体结构**、**绝对构型**、原子和分子的**堆积**、有序或无序的**排列**以及非计量的程度等。

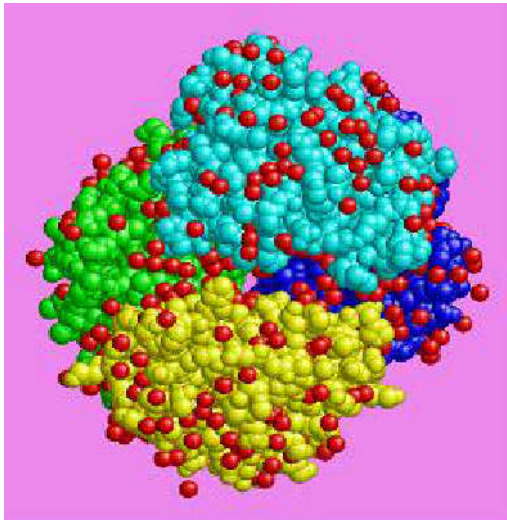
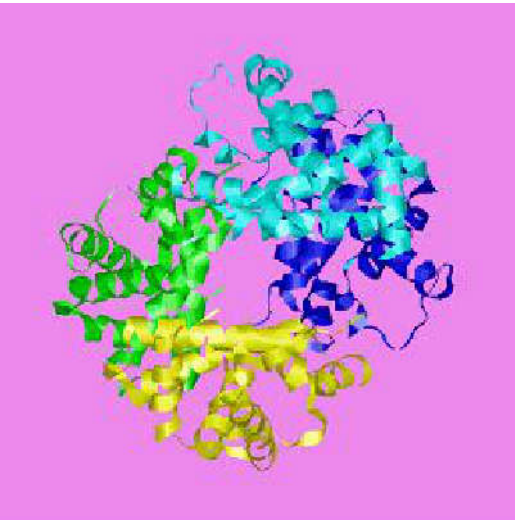
登高必自卑，行远必自迩

1957年肯特罗 (Kendrew) 完成肌红蛋白的0.6 nm分辨率的蛋白质晶体结构



肌红蛋白结晶X-光衍射图





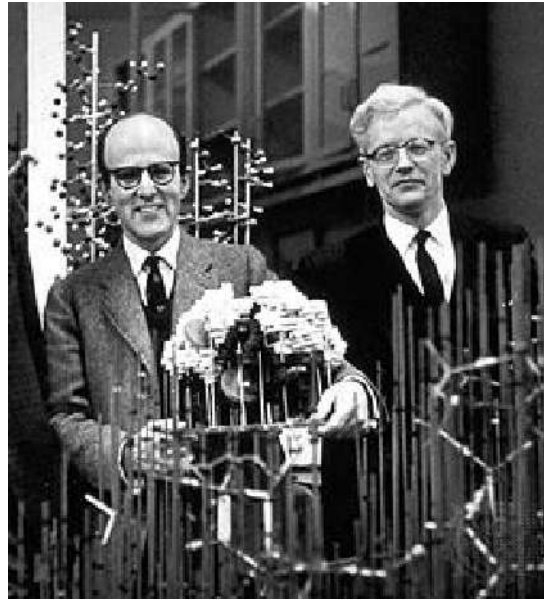
<http://hpdb.hbu.edu.cn/structure/proteinstructure10.asp>



<http://www.hsgq.pudong.edu.sh.cn/webinfo/jiaoyanzujianshe/huaxue/zhuye/minrentang/1962.htm> (1962)



由于测定出蛋白质的精细结构, 两位英国科学家M. F. 佩鲁茨和J. C. 肯德鲁获得1962年的诺贝尔化学奖。



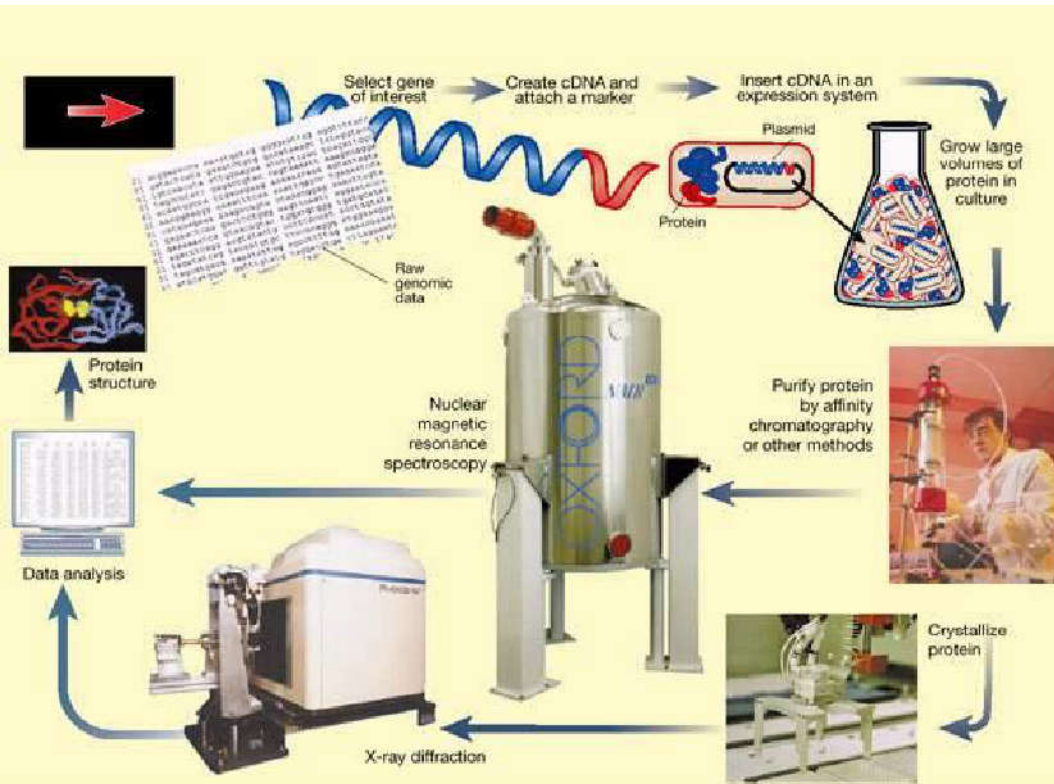
www.britannica.com/nobel/micro/46132.html

登高必自卑, 行远必自迩



蛋白质晶体结构的X射线衍射分析包含五个主要步骤：

- 1、样品制备
- 2、蛋白质结晶和晶体生长
- 3、衍射数据收集和处理
- 4、位相求解
- 5、模型建立和修正



- 克隆、表达、纯化
- 结晶
- 数据收集及处理
- 相角的测定
- 相角的改进（优化）
- 电子密度图的解释
- 修正
- 结构的描述和与功能关系的研究

(1) 样品制备



大量表达、分离和纯化目标蛋白
表达体系

1. 细菌系统（大肠杆菌E. coli等）
2. 酵母系统
3. 昆虫细胞
4. 哺乳动物细胞
5. 细胞外体系

(2) 蛋白质结晶和晶体生长



蛋白质结晶原理

与小分子结晶一样，蛋白质在溶液中处于过饱和状态时，分子间可以规则的方式堆积起来形成晶体析出。

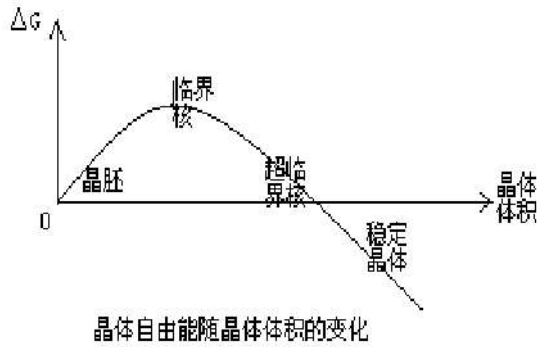
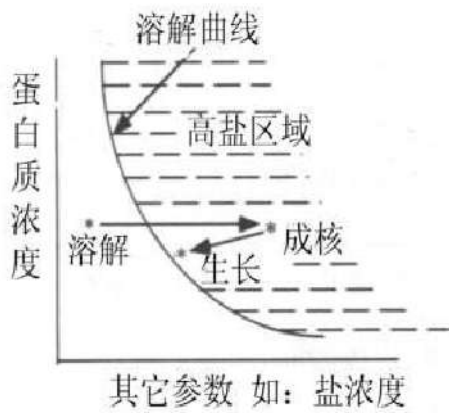
蛋白质结晶方法

整批结晶法

透析法

液液扩散法

气相扩散法



蛋白质在其溶液中的结晶其实是一种缓慢沉淀的过程。溶液中存在的作用力有疏水相互作用、静电力、氢键作用力等，当吸引力与排斥力达到平衡时溶液处于稳定状态。当吸引作用增大，或分散的或排斥的相互作用减小时，分子开始趋向与聚集状态。例如在液态环境中，如果生物大分子溶液很浓，没有足够的水维持其溶剂化作用，则分子可能聚集成非晶态沉淀，也可能结晶出来。

常用的使蛋白质沉淀的方法是加入沉淀剂，这种方法是通过降低水的流动性来增加蛋白质的有效浓度。常用沉淀剂：聚乙二醇（PEG），盐（如硫酸铵）。

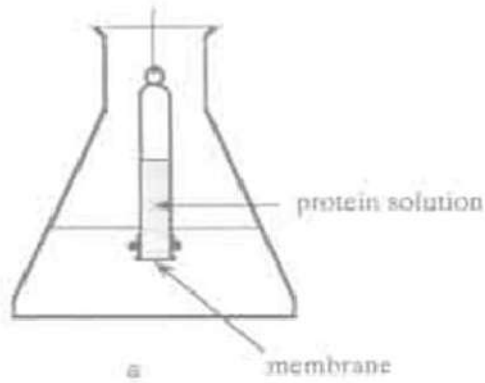
另外一种方法：减小蛋白质分子间的排斥力或是增大他们之间的吸引力。如加入有机溶剂，改变溶液的PH，温度等。



整批结晶法(batch crystallization)

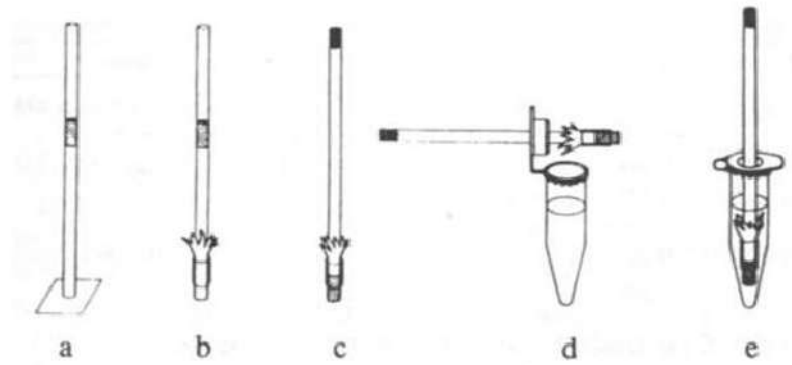
这是最古老也是最简单的蛋白质结晶的方法。该方法的原理是瞬间将沉淀剂加入蛋白质溶液中，使溶液突然达到高度饱和状态。如果运气好的话，晶体会逐渐从过饱和溶液中析出并且不需要其他步骤。

■ 透析法 (dialysis)



适用于对于大量蛋白质样品进行结晶。

■ 微量透析法 (liquid-liquid diffusion)



- 图a和图b中，蛋白质溶液保留在毛细管中。图c中，蛋白质溶液被甩到管底并与膜接触。图d和图e中，毛细管置于盛有渗析液的Eppendorf管中。

■ 液液扩散法



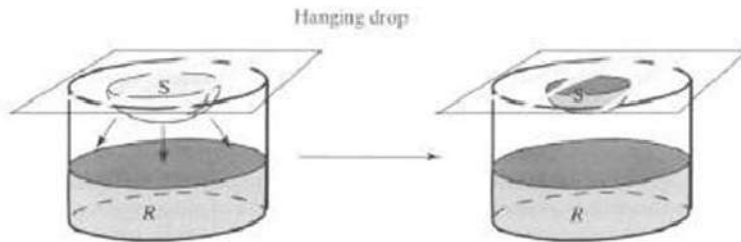
熔点毛细管中的液液扩散
图示为沉淀剂密度较大时

——这种方法是在小孔毛细管中将蛋白质溶液与含有沉淀剂的溶液相互覆盖，让其自己缓慢扩散的过程。

——蛋白质溶液与沉淀剂以1:1混合。因此沉淀剂浓度应为最终所需浓度的**2**倍。

气相扩散法

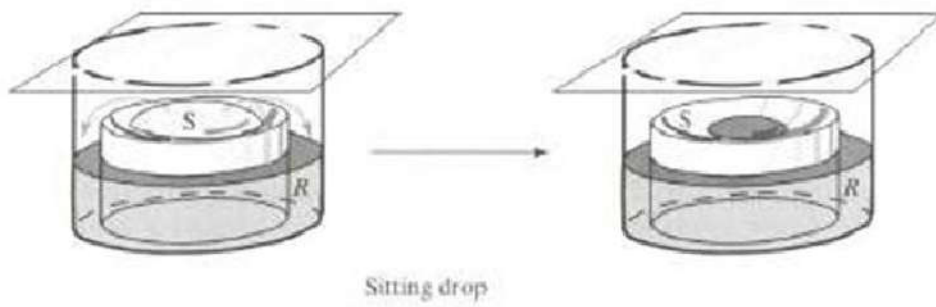
■ 悬滴法



- 使用带有空穴的盘子，蛋白质溶液的液滴悬挂在盖玻片的下方，盖玻片覆盖在密封有沉淀剂溶液的空穴上方。盖玻片的表面经过硅化处理以阻止液滴在玻璃表面的铺展。

- 这种方法是利用气相平衡的原理，使任何挥发性的组分在小液滴和大样品池间达到平衡，使蛋白质液滴中沉淀剂及蛋白质的浓度逐渐增加，达到过饱和状态，最终析出晶体
- 非常适用于只有少量样品，却要筛选大量条件的情况。

■ 坐滴法



- 如果蛋白质溶液表面张力低，在悬滴法中蛋白质溶液在盖玻片上铺展而不能形成液滴，此时可用坐滴法。

杂质、晶核及其他因素如温度、溶液PH值及振动等都会影响结晶过程。

为使蛋白质结晶，要求：

- 蛋白质纯度足够高，不含其他化合物。
- 蛋白质分子表面性质相同，聚集形式均一。
- 蛋白质溶解在合适的溶剂中。
- 溶液达到过饱和，使蛋白质形成小的聚集体，作为晶体生长的晶核。晶核形成，晶体才开始生长。
- 溶液的过饱和度要降至最低水平，避免太多的晶核产生。
- 晶体生长尽量缓慢以便在结构中达到最大的有序度，方法：改变温度，蛋白质浓度等。

物理因素：温度、重力、压力、震动、时间、电场磁场、介质的电解质性质和粘度、均相或非均相成核等

化学因素：pH值、沉淀剂类型和浓度、添加剂、离子种类、离子强度、过饱和度、氧化还原环境、蛋白质浓度等

生化因素：蛋白质纯度、配合体、抑制剂、化学修饰、遗传修饰、蛋白质的聚集状态、蛋白质水解、蛋白质自身的对称性、蛋白质的稳定性和等电点等

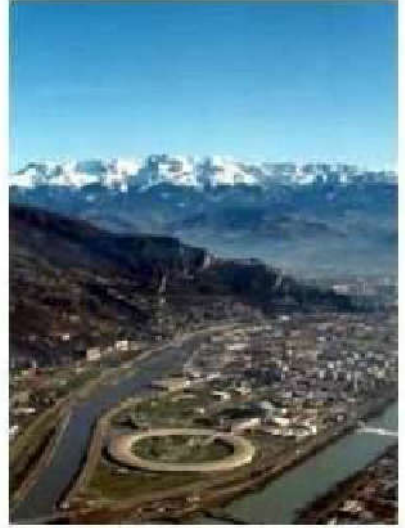


3、衍射数据收集和处理

第三代同步辐射光源的应用使得用**20-40 um**大小的晶体解析高分辨率结构已经成为现实

目前世界上比较著名的同步辐射工作站有多个：

APS (USA) ; ESRF (France) ; SPring-8 (Japan)



艺术家画的座落于格勒诺布尔[法国东南部城市]Grenoble的欧洲同步辐射装置。
周长是844.39米。

同步辐射器是一个庞大而且昂贵的装置。环的直径有10到几百米的大小。

4、位相求解



相位测定是X射线晶体结构分析的核心问题。由于相位不能由实验术探测到，因而只能通过一些间接的方法推导出来。

(1) 分子置换法(MR)

分子置换法就是把已知结构的蛋白质分子放到待测蛋白质晶体的晶胞中建立起初始结构模型并借助此模型计算待测蛋白质晶体各个衍射点的相角的方法。

(2) 多对同晶型置换法(MIR)

在蛋白质晶体中引入**散射能力强的重金属原子**如Pb和Hg等作为标志原子，制备出重原子的衍生物，然后求出这些重原子在晶胞中的坐标，根据坐标计算出重原子散射波在各个衍射点的相角，最后推测出蛋白质分子在各个衍射中的位相。

(3) 多波长反常散射法(MAD)

利用**同步辐射波长连续可变**的特点，使用一个重原子衍生物作为母体，用一个晶体就可以收集到重原子反常散射吸收边两侧的多套数据，并解出结构。

5、模型建立和修正



晶体学R因子一般要求达到**0.2**以下

键长偏差大约为**0.015Å**

键角偏差约为**3°**

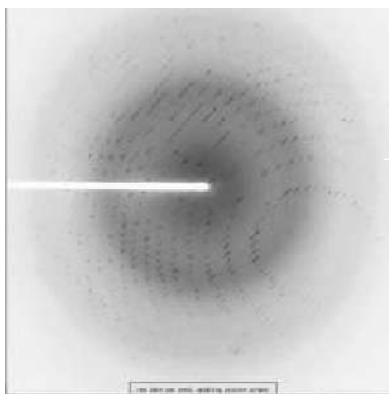
二面角构象分布要求除了甘氨酸的二面角构象是随机的外，其他残基的二面角构象分布受到立体化学的限制

X射线衍射实验和结构计算过程

Fourier变换与Fourier反变换
实空间

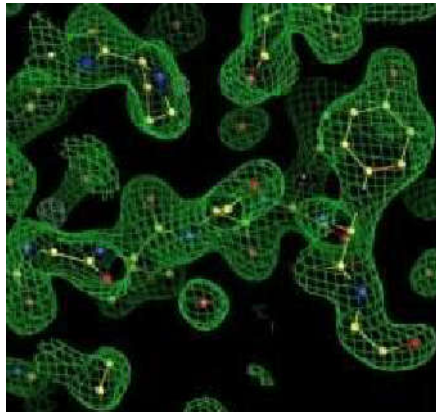
倒易空间

衍射图片

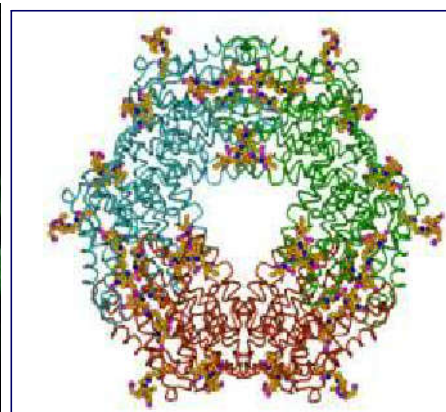


相角求解

电子密度

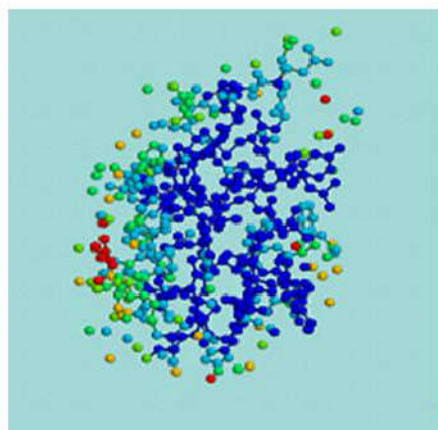


三维结构



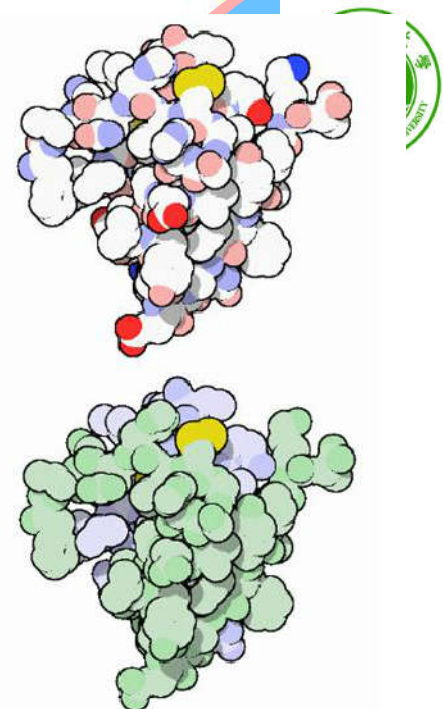
胰岛素

- 于1971年和1972年分别得到分辨率为2.5埃和1.8埃的猪胰岛素晶体测定，这是中国阐明的第一个蛋白质的三维结构。

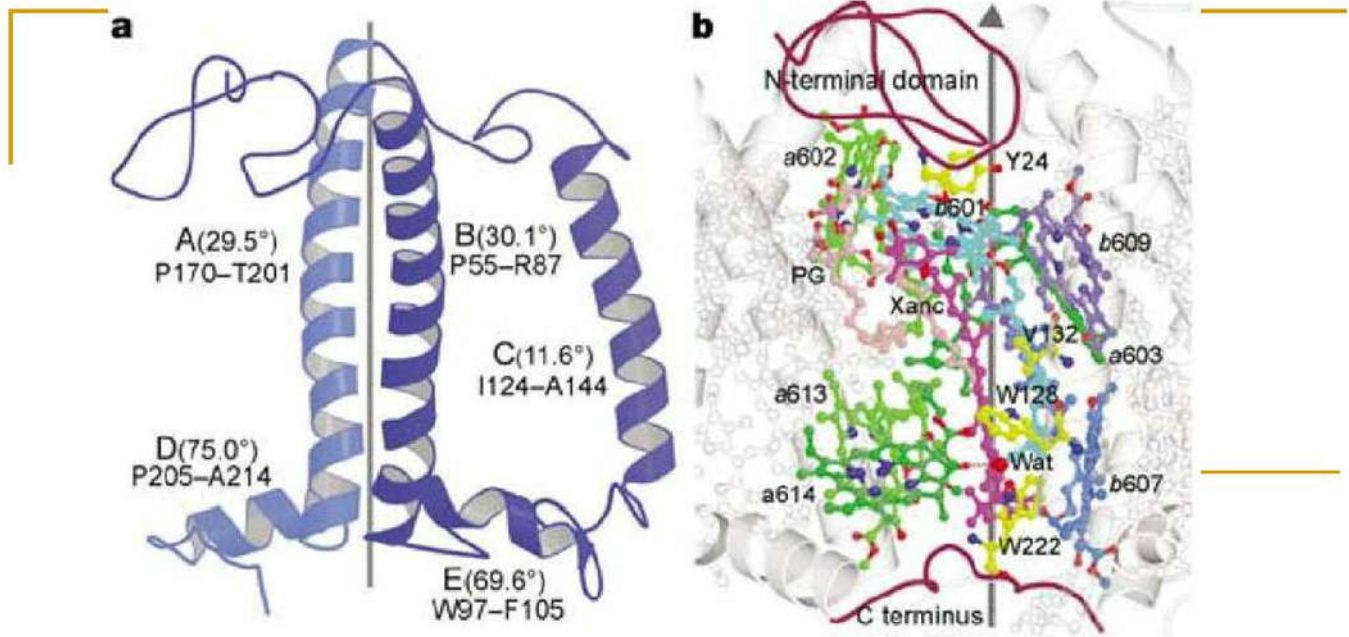


“中国蛋白质晶体学研究水平和世界发达国家一样高！胰岛素晶体最好的电子密度图在北京，不在牛津。”
——多萝西·霍奇金，1972

猪胰岛素(蛋白质编号4ins)的两条小链



菠菜捕光蛋白LHC II (膜、疏水, 2.72Å) Nature, 2004, Mar.

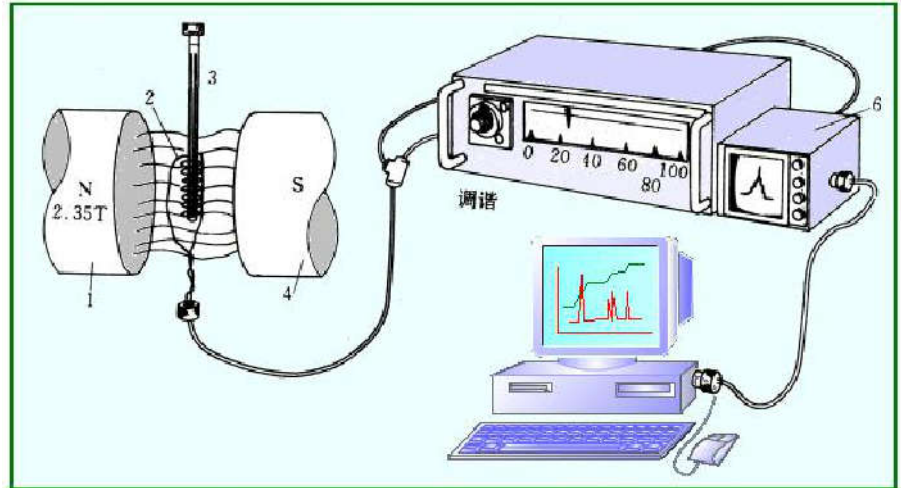


LHC-II是绿色植物中含量最丰富的主要捕光复合物，这种复杂的分子体系镶嵌在生物膜中，具有很强的疏水性，难以分离和结晶。对其晶体结构的测定是国际公认的高难课题，也是一个国家结构生物学研究水平的重要标志。03年利用北京同步辐射实验室生物大分子晶体学线站获得了该晶体的高分辨率衍射数据，最终获得2.72Å分辨率的晶体结构。

(二) 核磁共振测定生物大分子方法



核磁共振成像原则上也是吸收光谱。核磁共振(Nuclear Magnetic Resonance), 就是处于某个**静磁场**中的**自旋核系统**受到相应频率的射频磁场作用时, **共振吸收**某一特定频率的射频辐射的物理过程。核磁共振波谱是测量原子核对射频辐射(约4~600MHz)的吸收, 这种吸收只有在高磁场中才能产生。

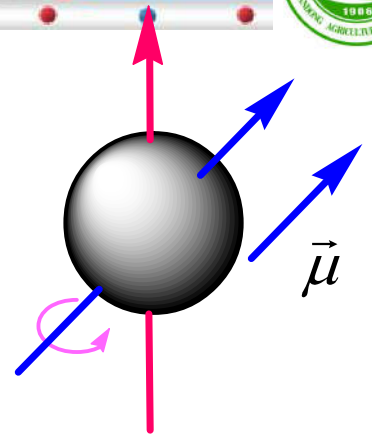


核磁共振波谱仪

1) 一些原子核像电子一样存在自旋现象，因而有自旋角动量：

$$P = \frac{h}{2\pi} [I(I+1)]^{1/2}$$

I 为自旋量子数



(2) 由于原子核是具有一定质量的带正电的粒子，故在自旋时会产生核磁矩： $\mu = \gamma P$

γ 磁旋比，即核磁矩与自旋角动量的比值，不同的核具有不同的磁旋比，它是磁核一个特征（固定）值。

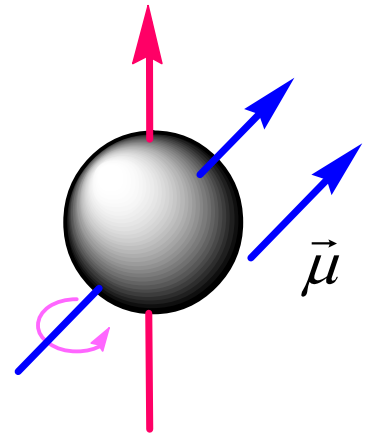
(3) μ 与 P 方向平行。

若原子核存在自旋，产生核磁矩，这些核的行为很象磁棒，在外加磁场下，核磁体可以有 $(2I+1)$ 种取向。

只有自旋量子数 (I) 不为零的核都具有磁矩

质量数 (a)	原子序数 (Z)	自旋量子 (I)	例子
奇数	奇或偶	$\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$	$I = \frac{1}{2}, {}^1H_1, {}^{13}C_6, {}^{19}F_9, {}^{15}N_7$ $I = \frac{3}{2}, {}^{11}B_5, {}^{35}Cl_{17}, I = \frac{5}{2}, {}^{17}O_8$
偶数	偶数	0	${}^{12}C_6, {}^{16}O_8, {}^{32}S_{16}$
偶数	奇数	1, 2, 3, \dots	$I = 1, {}^2H_1, {}^{14}N_7, I = 3, {}^{10}B_5$

(1) 在很强的外磁场中，某些磁性原子核可以分裂成两个或更多的量子化能级。

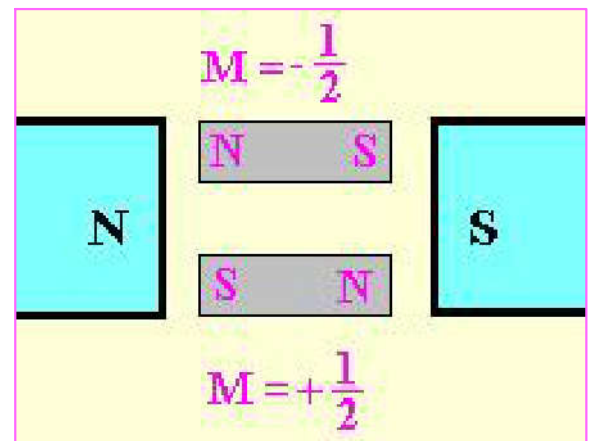
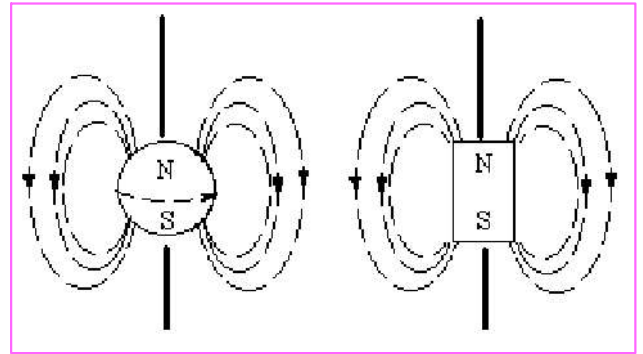


(2) 用一个能量恰好等于分裂后相邻能级差的电磁波照射，该核就可以吸收此频率的波，发生能级跃迁，从而产生 **NMR** 吸收。

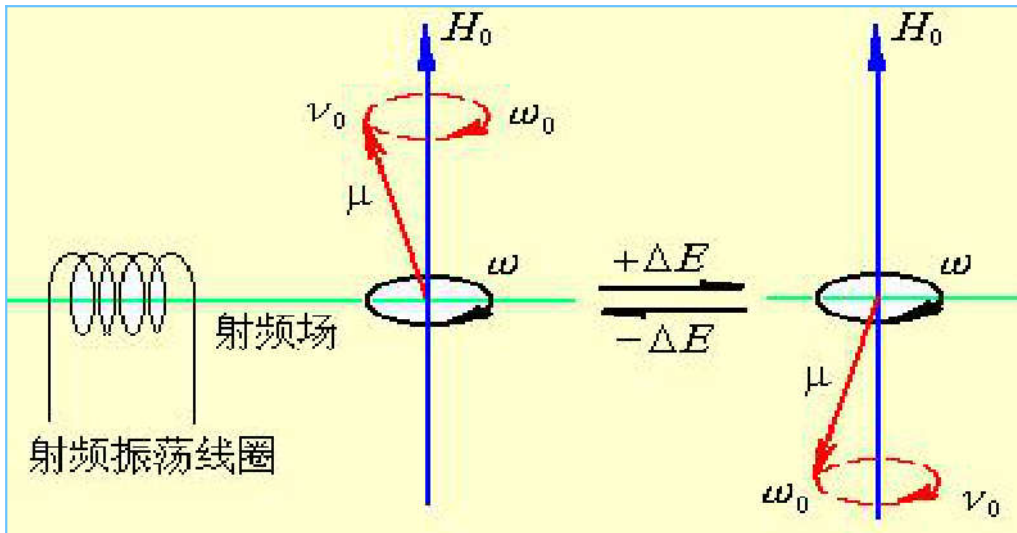
氢核 ($I=1/2$)，两种取向（两个能级）：

(1) 与外磁场平行，能量低，磁量子数 $m=+1/2$;

(2) 与外磁场相反，能量高，磁量子数 $m=-1/2$;



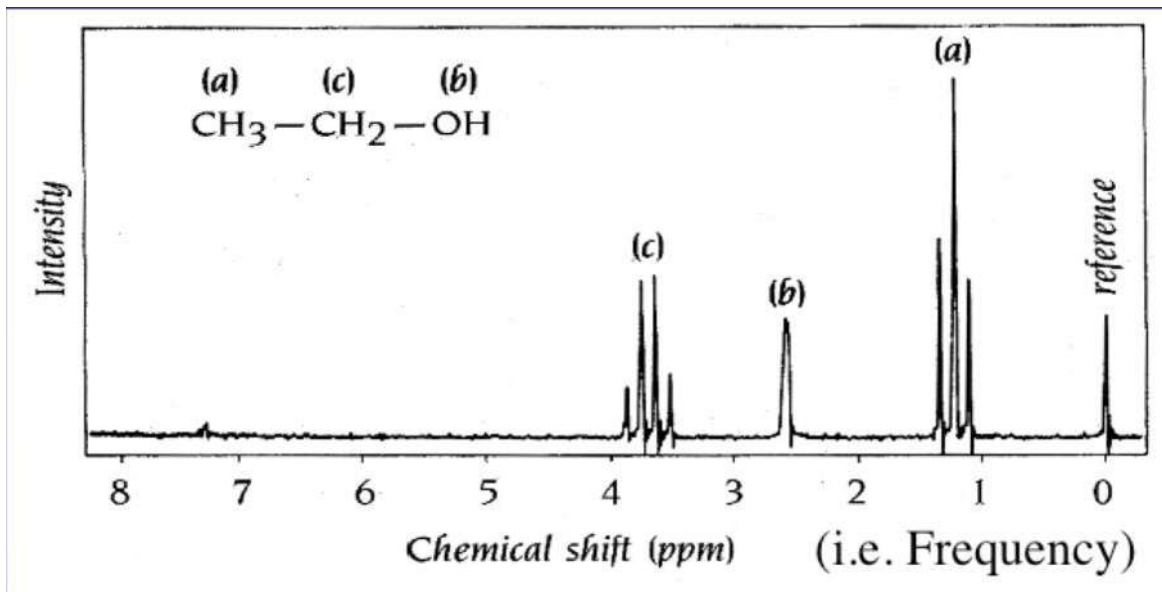
- (1) 核有自旋(磁性核)
- (2) 外磁场,能级裂分;
- (3) 照射频率与外磁场的比值 $\nu_0 / B_0 = \gamma / (2\pi)$



不同的原子处于不同的化学环境，即处于不同的电磁环境，因此表现出不同的共振频率：

$$\nu = \gamma (B_0 \pm B) / 2 \pi$$

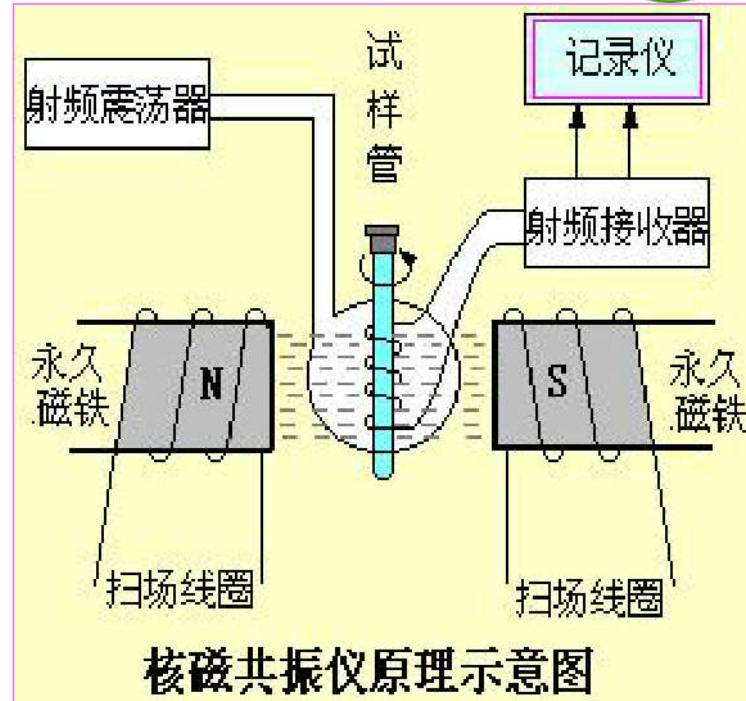
$$\omega = 2 \pi \nu = \gamma (B_0 \pm B)$$



核磁共振波谱仪



1. 永久磁铁：提供外磁场，要求稳定性好，均匀，不均匀性小于六千万分之一。扫场线圈。
2. 射频振荡器：线圈垂直于外磁场，发射一定频率的电磁辐射信号。60MHz或100MHz。
3. 射频信号接受器（检测器）：当质子的进动频率与辐射频率相匹配时，发生能级跃迁，吸收能量，在感应线圈中产生毫伏级信号。
4. 样品管：外径5mm的玻璃管，测量过程中旋转，磁场作用均匀。



核磁共振的步骤

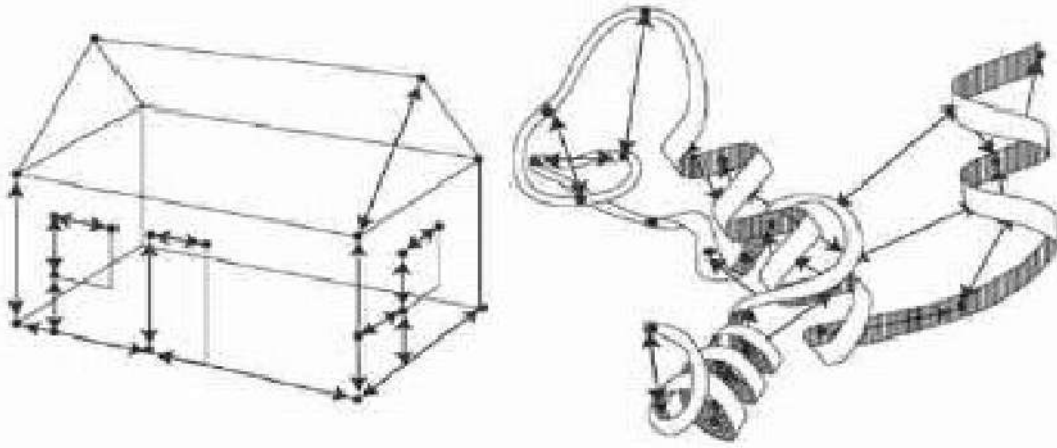


登高必自卑，行远必自迩

二维核磁共振(2DNMR)



NMR可获得分子内各核的化学环境、核间的耦合关系、空间构象等信息，但是当分子较大的时候，由于裂分谱线间的重叠，因此在测定了同一种核的一维谱之后，需要了解两种或三种不同核之间的联系关系，如C-H，N-H等就需要用到2DNMR，它是2个频率变量的函数，吸收峰对2个频率变量作图。



瑞士科学家库尔特·维特里希则发明了“利用核磁共振技术测定溶液中生物大分子三维结构法”。这种方法的优点是可对溶液中的蛋白质进行分析,进而可对活细胞中的蛋白质进行分析,能获得“活”蛋白质的结构,其意义非常重大。

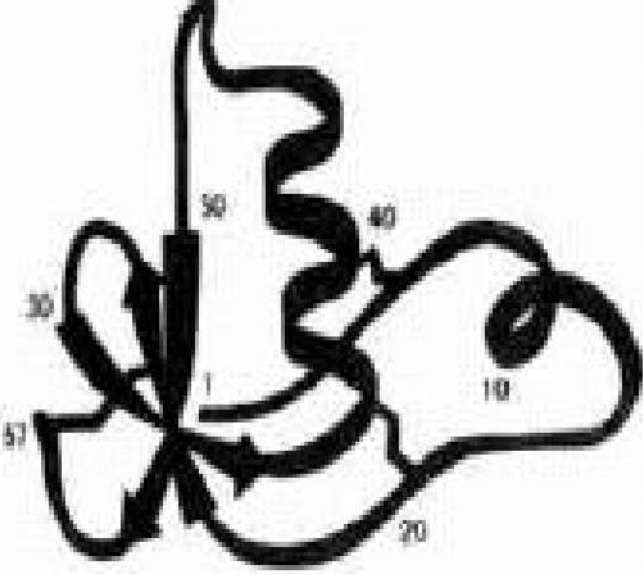
这种方法的原理可以用测绘房屋的结构来比喻首先选定一座房屋的所有拐角作为测量对象,然后测量所有相邻拐角间的距离和方位,据此就可以推知房屋的结构。维特里希选择生物大分子中的质子(氢原子核)作为测量对象,连续测定所有相邻的2个质子之间的距离和方位,这些数据经计算机处理后就可形成生物大分子的三维结构图。

❖核磁共振本身不能展示样体的内部结构。要得到内部的图像，就要将不同梯度的磁场加以结合，即改变穿过样本的磁场强度。这样就有无数二维的图像，彼此重叠后就得到样本内部空间的三维图像

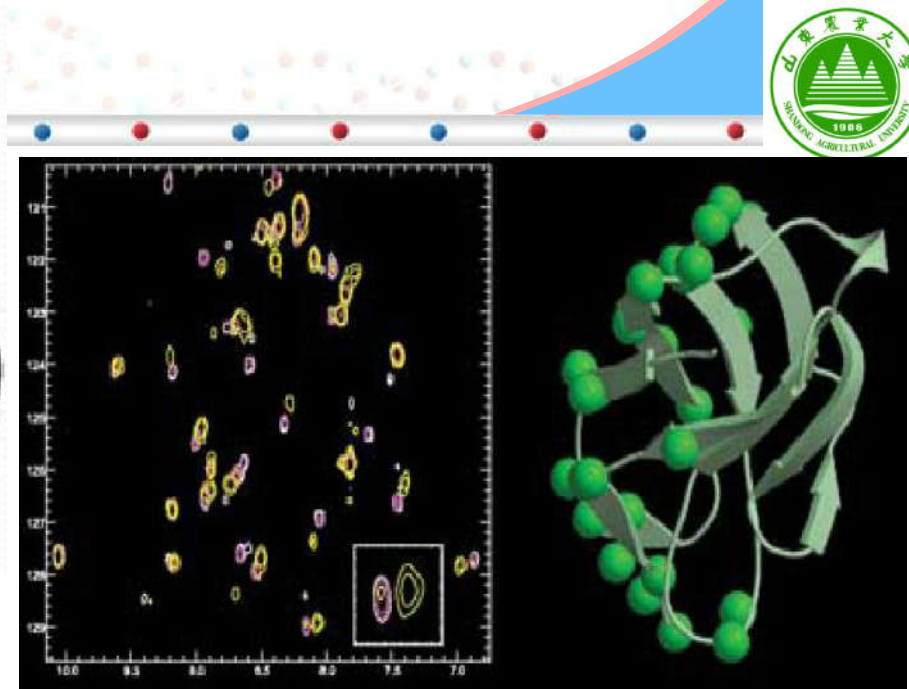
❖可以把一维谱中的重叠峰在二维或三维方向展开，便于NMR谱的解析。还能间接检测到在普通NMR方法中得不到的多量子跃迁

早期核磁共振主要用于对核结构和性质的研究，如测量核磁矩、电四极距、及核自旋等

后来广泛应用于分子组成和结构分析，生物组织与活体组织分析，病理分析、医疗诊断、产品无损监测等方面



1985年, 维特里希等人公布了第一次利用NMR 法测定的溶液中蛋白质——蛋白酶抑制剂IIA (proteinase inhibitor IIA) 的结构



NMR图谱得到的蛋白质三维结构
from: 厦门大学生命科学学院