

同步辐射中的蛋白质结构和功能研究

董宇辉

(中国科学院高能物理研究所 100049)

20 世纪后半叶发展最迅速的一个学科是结构生物学，这是一门通过蛋白质结构来诠释蛋白质功能的学科，DNA 双螺旋结构和血红蛋白以及肌红蛋白的结构确定标志着这个学科的建立。蛋白质是生命活动主要的执行者，只有形成了正确的结构，蛋白质才能具有正确的功能。因此蛋白质的结构研究，不仅可以让我们得以深入了解生物体内诸多生命过程的细节，更有助于重大疾病的预防和治疗、新型高效药物和疫苗的研发等。正是由于蛋白质结构研究在基础科学和医药、疫苗研发中的重要作用，结构生物学获得了蓬勃的发展。

目前获得蛋白质结构的方法主要有 3 种：X 射线晶体学、核磁共振和冷冻电镜。

核磁共振是一种应用非常广泛的技术，在化学、物理及生物学的多个领域中都有广泛的应用。如果在磁场中有一个核磁矩不为零的原子核，那么这个原子核的自旋能级将会发生塞曼分裂。在这种条件下，这个原子核就能够共振吸收某一特定频率的射频辐射。而这个核磁共振的频率随着周围化学环境的不同也略有变化，因此通过核磁共振方法确定出共振峰的位置后，就能够计算出相应的原子核周围的化学环境，即原子之间的键长、键角、二面角等立体化学参数。这样我们就获得了搭建整个蛋白质结构的片段信息，从而搭建出整个分子的三维结构。这个方法的优点在于可以研究溶液状态下的结构以及动力学过程，但是由于谱仪的分辨能力有限，分子量很大（即分子量大于 4 万道尔顿，相当于 360 个氨基酸残基）的生物大分子所产生的共振峰由于数目过多难以分辨，导致结构无法解析。同时为了产生核磁共振信号，还必须把没有核磁矩的原子核替换成为有磁矩的同位素，例如把 ^{14}N 替换为 ^{15}N ， ^{12}C 替换为 ^{13}C 等。虽然这些同位素替换的技术已经相当成熟，但是制备这些同位素置换的

样品仍是一个昂贵和费时的的工作。

冷冻电子显微学在研究分子量巨大的体系方面具有优势。这种方法可以使用二维晶体，也可以使用单颗粒重建技术来获得三维结构。然而二维晶体的生长，遇到的困难和生长三维晶体类似，也是一个依靠经验尝试的过程。利用单颗粒重建技术，需要拍摄很大数量的生物大分子图像，才能够得到足够的信息确定结构的细节。理论上用单颗粒重建技术得到分辨率为 1 \AA 的结构需要处理数以百万计的分子图像，目前还没有完成如此多的图像处理工作的可能性。最近随着新型探测器的发展，在图像质量上得到了大幅度的改善，因此冷冻电镜单颗粒重建获得的结构也接近了原子分辨率，但是由于电子与蛋白质的相互作用截面很大，多重散射效应使得蛋白质内部结构的分辨率不如表面部分。

确定蛋白质分子的三维结构最常用和最准确的手段是 X 射线晶体学。蛋白质可以在适当的条件下长出晶体，因此衍射技术可以用到蛋白质的结构解析上来。其实利用晶体学衍射解析结构的历史非常悠久，自劳厄第一次观察到晶体的衍射现象以来，人们就意识到可以利用衍射来解决物质的结构问题。布拉格父子第一次演示了利用衍射解析结构的例子，因此结构分析逐步成为了晶体学最重要的作用之一。晶体学早年的工作集中在小分子晶体的结构解析上，一直到二十世纪五六十年代，肯德鲁 (Kendrew) 和佩鲁茨 (Perutz) 的工作解析了肌红蛋白和血红蛋白的晶体结构，结构生物学这个今天的“明星”学科诞生了。今天，利用 X 射线晶体学解析蛋白质结构这门学科也被称为“蛋白质晶体学”，属于 X 射线晶体学的一个分支。之所以在劳厄和布拉格父子的工作完成几十年之后才得到第一个蛋白质的结构，就是因为蛋白质的结构远远要比小分子复杂，解析蛋白质的晶体结构要困难得多。

很幸运,随着 X 射线晶体学的各种理论,以及实验设备的蓬勃发展,现在 X 射线晶体学已经成为确定蛋白质结构的主力技术。到目前为止,在蛋白质数据库(Protein Data Bank, PDB)中已经收录了超过 10 万个蛋白质结构(表 1),这些结构中有 9 万多个是用 X 射线晶体学解析出来的,由此可见 X 射线晶体学在确定蛋白质结构中的主导地位。

表 1 在蛋白质数据库中收录的结构数目统计(来自 PDB 网页 <http://www.pdb.org>)

实验方法	蛋白	核酸	蛋白/核酸复合物	其他	合计
X 射线晶体学	85247	1556	4537	4	91344
核磁共振	9277	1089	216	7	10589
电镜	576	67	190	0	833
综合方法	63	3	2	1	69
其他	157	4	6	13	180
合计	95320	2719	4951	25	103015

而在这些利用蛋白质晶体学解析的结构中,很大一部分是利用同步辐射的晶体学衍射手段测定的。根据 Biosync 数据库的统计,2013 年利用 X 射线晶体学一共解析了 9361 个蛋白质及蛋白质复合物的结构,其中 8538 个结构是利用同步辐射上的晶体学技术解析的。同步辐射是解决蛋白质等复杂晶体结构关键的光源,同步辐射提供了极高的 X 光光强,很好的准直性和方便可调的入射 X 光能量,这些特性对于生物大分子的结构测定非常有利。正是由于分子生物学技术、探测器技术、晶体学理论和同步辐射装置的发展,使得结构生物学得到了迅猛的发展,每年解析出来的结构数目急剧增加,对生物学各个领域都产生了深

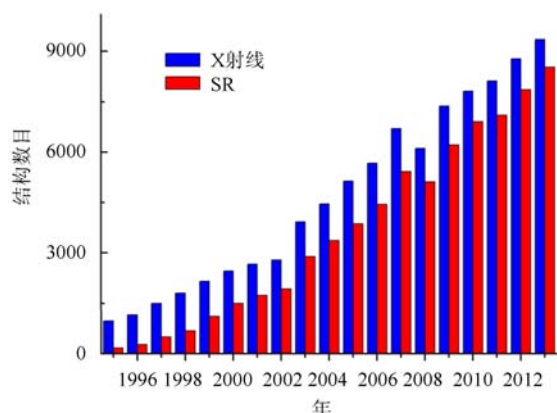


图 1 Biosync 统计的利用 X 射线晶体学解析的蛋白质结构数目。蓝色为利用 X 射线晶体学解析的蛋白质结构,红色为利用同步辐射上的 X 射线衍射解析的蛋白质结构

远的影响。现在生物学的研究者在浏览诸如 *Nature*、*Science*、*Cell* 这样的主流杂志时,经常能看到利用同步辐射解析结构的研究文章发表。

除数量之外,同步辐射对大分子结构解析更具意义的影响是使许多复杂结构的解析成为可能,如超大的大分子复合物和组装体等。迄今,基于同步辐射结构解析的研究工作已获得 5 项诺贝尔奖,分别是 1997 (ATP 合酶结构),2003 (离子通道结构),2006 (RNA 聚合酶结构),2009 (核糖体结构)和 2012 年 (G 蛋白偶联受体结构)的诺贝尔化学奖。因此,自肯德鲁和佩鲁茨创立起来的结构生物学,是目前受同步辐射推动作用最大的学科。

同步辐射之所以在蛋白质结构中如此重要,是由蛋白质结构解析的特点所决定的,特别是晶体结构解析中无法回避的相位问题。

1) 相位问题

利用晶体衍射解析结构,其基本原理可以描述如下:在 X 射线照射到单晶样品上时,会发生所谓的衍射现象。这个现象在一百年前就由劳厄观察到了。根据晶体学的理论,从晶体出来的衍射点,和晶体中的某一个晶面对应,表征晶体中一个晶面使用的是它的密勒指数 (hkl)。密勒指数的定义比较晦涩,详细的解释读者可以参考晶体学的教科书,在这里可以简单的介绍一下。

在晶体中,无数的原子周期性排列成规则的晶格,我们总可以在晶格中找出很多处于同一平面上的原子,这些原子所处的平面就叫“晶面”。我们还可以发现,这些晶面中很多是相互平行的,也是等价的,因为上面的原子排列完全一样。在晶体学中,这些等价的晶面就由密勒指数来标记。密勒指数的定义比较复杂,首先在这一系列平行而且等价的晶面中选出离原点距离最近的一个,这个面在晶体的三个晶轴上、并以晶格常数为单位测得的截距(例如在三个轴上的截距分别为 1, 3, 2),将三个轴上的截距计算出三个倒数 (1, 1/3, 1/2),然后约简为 3 个没有公约数的整数(都乘以 6,得 6, 2, 3),即将其化简成最简单的整数比 (6:2:3),就得到这组晶面(请记住这是一系列平行的晶面)的密勒指数 (623)。图 2 就

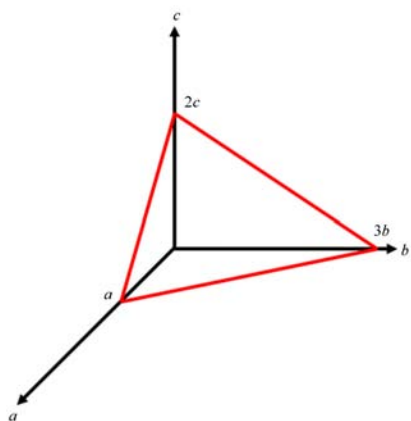


图2 密勒指数为(623)的晶面示意图。 a 、 b 、 c 分别代表晶体的三个晶轴，三条红色线段组成的晶面在这三个晶轴上的截距分别为 a 、 $3b$ 、 $2c$ ，以晶格常数(a 、 b 、 c)为单位，则截距为(1, 3, 2)，截距的倒数为(1, 1/3, 1/2)，约化成互质的整数就是(6, 2, 3)

是(623)这个晶面的示意图。

因此，X射线照射到晶体上产生一系列的衍射点，每个点都有一个名字，或者标记(hkl)。带有某个(hkl)标记的衍射点，它的“特性”可以用一个叫结构因子的参数来描述，这个

$$F(h, k, l) = \sum_j f_j \exp[2\pi i(hx_j + ky_j + lz_j)] \quad (1)$$

其中 $F(h, k, l)$ 是每个衍射点所对应的结构因子， f_j 是晶胞中第 j 个原子的散射因子，其位置在(x_j, y_j, z_j)。

另一种表达结构因子的方式是通过晶胞内的电子密度：

$$F(h, k, l) = \frac{1}{V} \int_V \rho(x, y, z) \exp[2\pi i(hx + ky + lz)] dx dy dz \quad (2)$$

其中 V 为晶胞的体积， $\rho(x, y, z)$ 为晶胞中的电子密度分布，积分区域为整个晶胞的体积。

可以证明，这两个定义是一致的。X射线发生衍射时，主要是电子在发生作用，因此，X射线衍射探测到的，或者感知到的，是晶体样品中的电子密度。电子密度的分布和原子位置是一致的：有原子存在的地方，电子密度一定很大，反之亦然。

需要指出的是，结构因子是一个复数，既有实部，又有虚部。从另一个角度来讲，结构因子包括两部分：

(1) 振幅，即结构因子的绝对值；(2) 相位，表征结构因子和实轴之间的夹角。如图3所示，如果将结构因子的实部和虚部在一个直角坐标系中表示出来，那么我们会得到二维空间的一个矢量，在实轴（相当于X轴）的投影就是它的实部，在虚轴（相当于Y轴）

的投影就是虚部。那么这个矢量的长度就是振幅（就是这个复数的绝对值），与实轴的夹角就是相位。

结构因子其实就是电子密度的傅里叶变换，如果把结构因子进行反变换，我们就得到电子密度

$$\rho(x, y, z) = \sum_{h, k, l} F(h, k, l) \exp[-2\pi i(hx + ky + lz)] \quad (3)$$

计算出电子密度，我们就可以通过电子密度确定原子位置了。

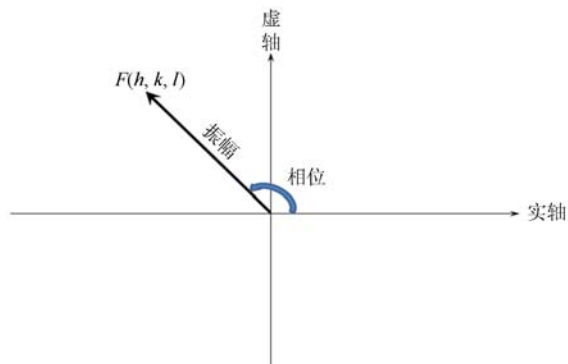


图3 结构因子是一个复数，绝对值为振幅，与实轴的夹角为相位

在晶体学的衍射实验中，我们利用各种各样的探测器来探测衍射点的强度，这些衍射点的强度正比于结构因子振幅的平方。所以通过测量衍射点的强度，我们就可以知道结构因子的绝对值。但是，另一个参数，相位却是不能直接测量出来的。从另一个角度来看，相位这个参数，其实不是唯一的，你可以把晶胞中的原点位置挪动一下，里面分子的结构不动，得到另外一套结构因子。这套新的结构因子里面，绝对值不变，但是相位改变了。于是大家会问，相位这种测量不到，又没有明确数据的玩艺，难道还会对结构解析有重大影响吗？很遗憾，答案是很令人沮丧的：决定结构的重要数据，不是可以测量得到的振幅，反而是有点“虚无缥缈”的相位，准确地说，是相位之间的关系。

下面是一个有趣的例子，经常用来说明相位的重要性。假设用一只鸭子的图像进行傅里叶变换，得到了鸭子的振幅和相位。当然，如果你用这个结构因子去做逆傅里叶变换，得到的自然是鸭子的图像（图4、图5、图6图片来自 <http://www.ysbl.york.ac.uk/~cowtan/fourier/magic.html>）。

你又把猫的图像如法炮制，得到猫的振幅和相位，逆傅里叶变换得到的自然是猫的样子。

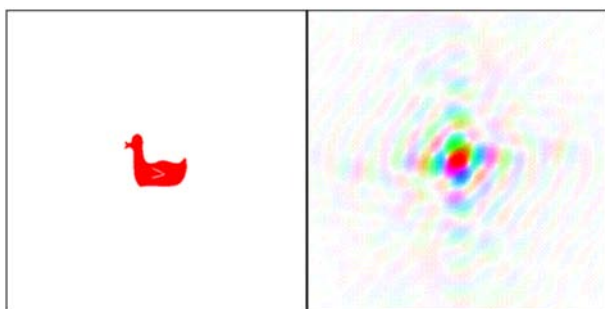


图4 鸭子的傅里叶变换

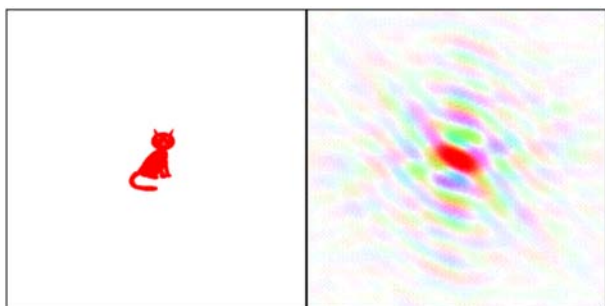


图5 猫的傅里叶变换

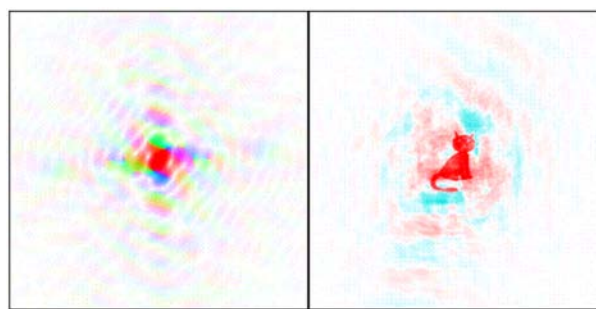
接下来的事情有点出乎大家的预料了。假如你把鸭子的振幅和猫的相位结合在一起，也能得到一套结构因子，把这套结构因子拿去作逆傅里叶变换，得到的结果却是一个接近猫的结构；反之，把猫的振幅和鸭子的相位结合在一起，结果出来的是鸭子。决定结构的，不是振幅，而是相位！

这就是晶体学中最核心的问题：相位问题。我们利用衍射实验测量得到的信号，缺少了确定结构最关键的相位信息。如何找到这个相位信息，成为晶体学发展的主线之一。

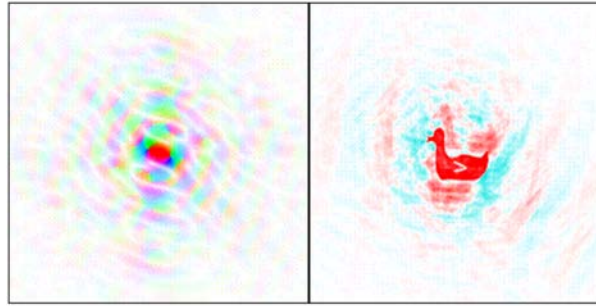
2) 同晶置换法

同晶置换法是蛋白质晶体学发展中里程碑式的成就之一。当年肯德鲁和佩鲁茨在解析肌红蛋白和血红蛋白结构的时候，为了解决相位问题，提出了这个划时代的方法，开拓了结构生物学这一学科。

这个方法的关键在于，假如我们在生长出蛋白质单晶的基础上，通过某种方式（比如将晶体浸泡在含有重金属离子的溶液中），把重金属离子引入到蛋白质的晶体内部，而不引起蛋白质结构的变化，那么我们就获得了一个“同晶置换”的晶体。因为重原子不导致晶体结构的变化，所以加入重原子之后，晶体的晶型没有改变，这就是所谓的“同晶型”。很显然，这是一个近似的假设，重原子的加入，或多或少会改



鸭子振幅+猫相位



猫振幅+鸭子相位

图6 交换振幅和相位的结果：决定结构的是相位而不是振幅

变蛋白质的结构，但是还是有可能获得结构变化不大的同晶置换晶体的。

由于重金属的加入，因此原来的晶体和同晶置换的晶体之间的结构因子有下列关系：

$$F_{PH}(h, k, l) = F_P(h, k, l) + F_H(h, k, l) \quad (4)$$

其中 $F_P(h, k, l)$ 为原来晶体的结构因子， $F_{PH}(h, k, l)$ 为加入重原子后的同晶置换晶体的结构因子， $F_H(h, k, l)$ 为纯粹由重原子组成的晶体（这是一个虚拟的晶体，即把原来的蛋白质全部去掉，只留下重原子，由于晶体的周期性仍然存在，这些重原子还是组成一个虚拟的晶体）的结构因子。

重原子的位置可以通过其他方法（例如解析小分子结构中很成熟的帕特森法）来确定。这样我们就有了 $F_P(h, k, l)$ 和 $F_{PH}(h, k, l)$ 这两个复数的结构因子的振幅，以及 $F_H(h, k, l)$ 这个完整的结构因子。注意 $F_H(h, k, l)$ 这个值，它是通过重原子的位置，利用公式 (1) 计算出来的。我们不仅知道其振幅，还知道其相位。于是我们就能够找出 $F_P(h, k, l)$ 的相位来，具体的做法可以参考图 7：

不过，半径为 F_P 和 F_{PH} 的两个圆有两个交点，因此这个方法只给出两个可能的解。我们可以再次制备另外一个同晶置换晶体，就能够确定出唯一的解来。

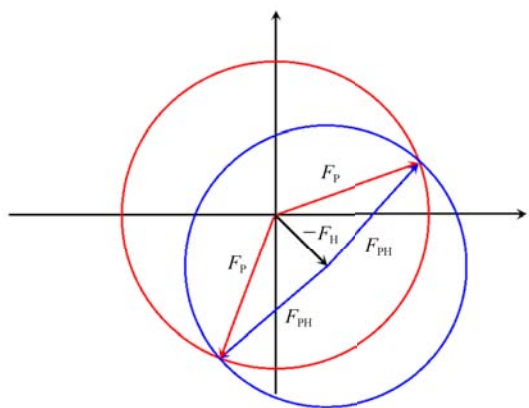


图7 同晶置换的示意图

或者采用其他方法，例如直接法，判断哪个解是真实的。这方面的细节，读者可以参考蛋白质晶体学的教科书。

同晶置换法是一种间接测量衍射相位的方法，它的出现，在晶体学发展上具有划时代的意义，因此肯德鲁和佩鲁茨获得了 1962 年的诺贝尔化学奖。

3) 反常散射

同晶置换法能够得到应有的前提是要制备出同晶置换的晶体来。这个任务可不容易完成。首先，将蛋白质长成晶体就不容易，需要长时间的尝试。大家也知道，重金属离子的毒性是相当大的，往往会造成蛋白质的变性，即结构破坏，要想把重金属离子渗入到蛋白质晶体的晶格中又不改变蛋白的结构，难度可想而知。因此，在同步辐射得到广泛应用以前的相当长时间内，虽然已经有了可行的结构解析技术，解析一个蛋白质结构也是非常困难的事情。

相当部分的蛋白质内含有 Mn、Fe、Cu、Zn 等金属元素，即便蛋白质中没有金属元素，也可以通过分子生物学和生物化学的方法，把蛋白质中的部分 S 元素换成同一家族的 Se 元素，这种方法现在已经发展得非常成熟。人们注意到，在某些特定的波长下，这些金属元素对于 X 射线的散射行为出现异常。在本文的公式 (1) 中，在通常情况下，每个原子的散射因子 f 是一个实数，但是这些金属离子在某些特定波长上，其散射因子变成了复数，可以表示为

$$f = f_0 + f' + if'' \quad (5)$$

其中 f_0 为无异常行为时的散射因子， f' 和 f'' 是发生异常时散射因子修正项的实部和虚部。这种现象就是所

谓的“反常散射”。其原因通过量子力学很容易理解。金属离子中的电子是按照不同的能级排列的，入射 X 射线的能量如果接近两个电子能级的能量差时，就会发生共振，使得吸收大幅度加强，导致散射行为的强烈改变。以 Cu 离子为例，其中 f' 、 f'' 等修正项随着入射 X 射线能量的变化如图 8（引自 <http://skuld.bmsc.washington.edu/scatter/>）。

因此，如果利用散射因子变化剧烈的几个波长的 X 射线分别来进行衍射实验，那么蛋白质中其他的原子，由于 X 射线波长变化不大，而且没有反常散射出现，这些原子的散射因子改变很小，可以忽略不计。但是那些发生了反常散射的离子，其散射因子有很大的变化，相当于换成了另一种原子。这样，我们通过

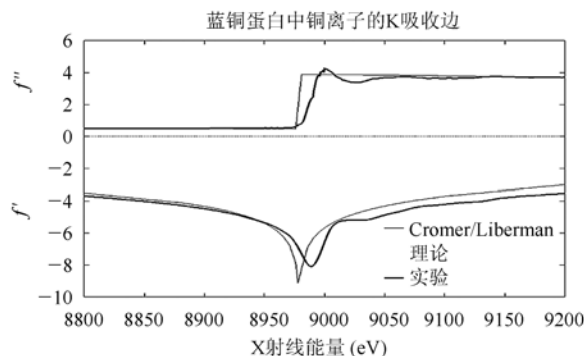


图8 反常散射系数随入射 X 射线能量的变化

改变入射 X 射线的能量，就达到了替换金属离子的效果，换句话说，我们实现了完全同晶（晶体没有改变，而是 X 射线的波长变了）的同晶置换。所以，通过利用反常散射，可以用同晶置换类似的方法来解析相位。

利用反常散射来解析蛋白质结构的想法由来已久，直到 1988 年左右，亨德里克森 (W. Hendrickson) 等人才将理论体系发展完善，并用反常散射实现了蛋白结构的解析。其原因也不难理解，实现这种实验，需要灵活改变 X 射线的能量，这就需要一种全新的光源。同步辐射正好是能够满足这种需求的光源。因此，同步辐射的出现，使得反常散射得以应用，大大提升了蛋白质晶体学解析结构的能力。2009 年莱马克里斯南 (Venkatraman Ramakrishnan)、施泰茨 (Thomas A. Steitz) 和尤纳斯 (Ada E. Yonath) 三位科学家，在核糖体结构和功能方面做出了杰出的贡献而获得诺贝尔化学奖。在获奖的公告材料中，明确指出了同步辐射

的贡献:

……还有其他的一些技术进展使得核糖体晶体结构解析成为可能,其中包括 CCD 面探测器,精确和自动的 X 射线衍射分析和波长可调的同步辐射的应用,可以利用反常散射来确定相位……

4) 同步辐射

前面已经提到同步辐射对于蛋白质结构解析的重要性。究竟什么是同步辐射呢?同步辐射其实就是相对论性带电粒子作曲线运动放出的电磁辐射,通俗地说,就是让运动速度接近光速的电子拐弯,发出的“光”(光也是一种电磁辐射)。这种光亮度极高,而且波长分布很广,从红外一直到硬 X 射线都有极高的亮度,而且是具有偏振特性的。而对于结构生物学者来说,需要的是一种性能优异的 X 射线,而同步辐射装置提供的 X 射线是结构生物学研究梦寐以求的光源。

同步辐射被人们所认识和发现经历了一段很长的时间。早在 1898 年,莱昂纳德(A. Lienard)通过理论预言,相对论性带电粒子沿半径为 R 作圆周运动时能发出电磁辐射,并给出了瞬时辐射功率的表达式。

从这以后不断的有理论上的预言,包括 1947 年我国的朱洪元先生给出了高速带电粒子在磁场中的辐射谱和角分布。到 20 世纪 40 年代,在粒子加速器的建造中,每一次加速器能量的提高都重新激起对这个问题的注意,原因也很清楚,这种辐射使得需要加速的粒子丢失能量,阻碍了粒子能量的进一步提升。

实验上探测该种辐射始于 1946 年,布卢伊特(J. P. Blewett)和鲍德温(G. C. Baldwin)在美国纽约斯克内克塔迪的通用电气实验室 100 MeV 电子回旋加速器(betatron)上进行电磁辐射的探测。布卢伊特坚信电子的电磁辐射有一个很宽的频谱,故鲍德温与同事用镀银的面包圈形天线在电子回旋加速器的真空盒外部和内部均作了接受基波和高次谐波的实验,结果什么也没有测到。这是由于不透辐射的天线镀层完全屏蔽了辐射。这样使同步辐射的发现推迟了一年。

1947 年 4 月,艾尔德(F. R. Elder)等人在美国通用电气实验室的 70 MeV 电子同步加速器上首次观察到了电子的电磁辐射,因此命名为同步辐射(同步加速器(synchrotron)的辐射(radiation))。假如一

年前,人们在回旋加速器上观察到电磁辐射,估计这种辐射该叫“回旋辐射”(betatron radiation)了。

该电子同步加速器是由波洛克(Herbert C. Pollock)领导建造的。当时在建造过程中加速器没有用辐射屏蔽完全包住,由于波洛克担心发生电击穿,采用了透明的真空盒,且在辐射屏蔽的角上安装了一个镜子,以观察是否发生电火花。实验时,技术员哈伯(Floy Haber)冒险的瞥了一下镜子,发现了强烈的弧光。经研究,这就是大家以前一直讨论的由相对论电子产生的电磁辐射。

鲍德温 1975 年在 *Physics Today* 上发表一篇文章,说到他与同步辐射的发现失之交臂的情况。认为如果当年他作某些细致的设计和自觉的违反辐射安全规则,他将发现者,同步辐射将以电子回旋辐射命名。

在同步辐射发现后约十年,1956 年汤博林(D. H. Tombouljian)和哈特曼(P. L. Hartman)在康奈尔大学的 320 MeV 同步加速器上经过两周的测量,给出了 $80 \sim 300 \text{ \AA}$ 范围内的同步辐射的空间分布和角分布,并指出同步辐射光源的软 X 光谱有其独特的优点。他们用掠入射摄谱仪拍照了 Be 箔和 Al 箔在 K 和 L 吸收边附近的透射谱。该同步辐射的先驱工作引起了进一步开展同步辐射应用研究的极大兴趣。

1961 年,美国国家标准局在 180 MeV 的电子同步加速器上建造了第一个用于紫外辐射的同步辐射装置,其实是附加在加速器真空盒切线方向的一个真空装置。在该装置上马登(R. P. Madden)和寇丁(K. Coding)测量了惰性气体波长在 $80 \sim 600 \text{ \AA}$ 的吸收谱,在光致电离的连续区发现了许多以前没有观察到的共振结构。该工作的成功进一步在世界范围内引起了将同步加速器作为紫外和软 X 光源的兴趣。接着欧洲、美国、苏联和日本都相继建立同步辐射的研究机构。这些同步辐射装置都是寄生在已有的加速器上,与高能物理研究处于共生状态。直到 1969 年威斯康星(Wisconsin)的 Tantalus 同步加速器才专门用于同步辐射研究。20 世纪 70 年代后专用的同步辐射光源渐渐多起来。目前已发展到约有 22 个国家,50 个实验室,80 台储存环和同步加速器用于同步辐射研究。

同步辐射的发展历经三代。第一代同步辐射装置

基本上是和高能物理、粒子物理的加速器装置共用的，由于没有专门考虑同步辐射实验的需求，加速器设计也没有进行相应的优化，性能虽然比实验室的 X 射线光源要好，但是还不能满足很多实验的需求。现在在北京的北京同步辐射装置（Beijing Synchrotron Radiation Facility）就属于第一代同步辐射装置。第二代装置就是专门设计的光源了，其代表性的装置有位于美国的 NSLS、德国的 Doris-III、日本的 Photon Factory、英国的 Daresbury。而第三代装置则进行了优化，大量使用各种提高亮度的插入件，因此其亮度很高，成为目前同步辐射的主力装置，代表性的装置有欧洲的 ESRF、美国的 APS、日本的 SPring-8、英国的 Diamond 等。

1981 年美国的第二代同步辐射装置 NSLS 上建成了世界上第一条用于测定蛋白质等生物大分子晶体结构的线站，很快世界各地的同步辐射装置上纷纷建成了蛋白质晶体学线站，目前这些线站的数目约有 150。这些性能优异的实验设施，为结构生物学的发展提供了有力的支撑。

5) X 射线自由电子激光和新一代的同步辐射

同步辐射装置提供的 X 射线具有很多优异的性能，正好能够满足结构生物学中解析蛋白质的需求。蛋白质主要由 C、N、O、S 等原子组成，即使形成了晶体，其衍射能力很弱，必须要高亮度的 X 射线才能产生足够的衍射信号。正好同步辐射装置发出的 X 射线，其亮度比常规的 X 光机要高出十几个数量级，所以使用同步辐射 X 射线能够在很短的时间内获得高质量的衍射数据，对于结构解析是非常方便的。蛋白质晶体的质量很难和小分子晶体相比，由于蛋白质分子本身的柔性，生长出来的晶体缺陷很多，这样需要准直性很好的光源才能获得明锐的衍射点，也只有同步辐射发出的 X 射线，具有极好的准直性，质量不高的晶体也能得到很好的衍射数据。同时同步辐射的准直性使得我们可以把 X 射线聚焦到一个很小的范围内，所以，尺寸很小的晶体也能收集衍射数据，进而解析结构。目前在同步辐射上，尺寸小到几个微米的蛋白质晶体也能够用来解析结构，由于对样品尺寸的要求大幅度降低，解析结构的难度也就相应降低了。最重

要的一点是，同步辐射提供了反常散射这种有力的方法，对于解决蛋白质晶体衍射的相位问题非常有效，这点在前面已经详细描述过了。

可以看到，同步辐射的优势在于灵活可调的 X 射线波长，以及极高的强度和微小的光斑。因此也不难理解，未来同步辐射的发展也就集中在如何提高 X 射线的强度和减小聚焦光斑，这意味着更小的晶体也能用于结构解析，毕竟生长蛋白质晶体是一个困难的任務。我们可以看到，同步辐射光源发展了三代，其强度逐步增加，光斑尺寸进一步减小。

为了获得更高的亮度，近年来出现了一种全新的光源：X 射线自由电子激光（XFEL）。自由电子激光也是利用相对论性的电子来产生 X 射线，不过利用了 X 射线光子和电子之间的相互作用，对电子束进行调制，使得电子的运动状态一致，于是就产生了与我们常见的激光一样的“受激辐射”（而同步辐射光源中，电子发射 X 射线是一个自发辐射的过程），我们得到的就是 X 射线波段的激光。正是由于这种光源的受激辐射特性，因此其变成了一种脉冲宽度非常窄（10 ~ 100 fs 量级甚至更小，而同步辐射能做到 ps 量级）的光源，峰值亮度要比同步辐射高十几个数量级，而且是完全相干的光源。由于这种全相干的特性，X 射线自由电子激光的一个重要科学目标，就是实现“不需要晶体的结构解析”，这样可以让结构生物学完全摆脱晶体生长这个瓶颈。

X 射线自由电子激光和同步辐射是相互补充的两种不同类型的光源。同步辐射研究的是静态的结构，或者比较慢的动态结构；而自由电子激光则能提供在 fs 量级的动态结构。这两种光源，一动一静，正好是科学研究都必须的两个方面。

2009 年，美国 SLAC 国家实验室建成了世界上第一台 X 射线自由电子激光装置 LCLS，几年后，日本也建成了 SACLA 这个 X 射线自由电子激光装置。目前针对 X 射线自由电子激光的方法学研究方兴未艾，也出现了一大批激动人心的结果。2000 年，瑞典乌普萨拉大学的哈伊杜（Janos Hajdu）研究组报道了分子动力学模拟的结果，表明溶菌酶分子在 XFEL 的高强度脉冲作用下，其结构可以在 2 fs 的时间内保持

不变, 5 fs 的时间内结构的误差还可以忍受, 超过 10 fs 后, 分子结构将会发生很大的变化。这个结果说明, 只要 XFEL 的脉冲足够短, 就可能在分子结构被破坏之前获得衍射信号, 由此提出了“diffraction-before-destruction”的概念。

2006 年, 德国 DESY 的查普曼 (Henry Chapman) 在当时的一台软 X 射线自由电子激光装置 FLASH 上对“diffraction-before-destruction”进行了实验验证。FLASH 虽然不能提供硬 X 射线的自由电子激光, 但是其软 X 射线激光的强度也完全足以破坏样品。在第一个脉冲的时间内, 查普曼成功获取了一个非晶体样品的衍射信号, 并将正确的结构重建出来。当然第二个脉冲获得的信号就完全不同了, 因为样品已经被上一个脉冲完全破坏了。

2011 年, 查普曼在美国建成的 LCLS 自由电子激光装置上成功获得了光合反应中心 I (PSI) 晶体的低分辨衍射数据。通过喷流装置, 他们将尺寸在 0.2 ~ 2 μm 的 PSI 晶体输送到自由电子激光光束中, 利用康奈尔大学研发的快速读出像素探测器 (pnCCD) 获得了衍射数据。这些低分辨的衍射数据和在同步辐射上获得的同样晶体的数据完全一致, 证明了利用 XFEL 可以获得准确的晶体衍射数据。2012 年, 查普曼研究组利用 LCLS 的 XFEL 获得了溶菌酶晶体的高分辨衍射数据, 并利用分子置换法解析出了晶体结构。这次实验使用的晶体尺寸很均一, 为 $1 \times 1 \times 3 \mu\text{m}^3$, 采用康奈尔大学研发的 CSPAD 探测器, 获得了 1.8 \AA 分辨率的溶菌酶晶体在常温下的结构, 和在同步辐射 (瑞士光源) 上获得的结构完全一致。这个实验表明, XFEL 可以获得高分辨的衍射数据, 而且能够解析出晶体结构。

溶菌酶毕竟是一种结构已知的蛋白质。在 2012 年, 查普曼研究组还报道了来自非洲锥虫 (*Trypanosoma brucei*) 的 cysteine protease cathepsin B 的高分辨结构。这是该酶在酶原状态的结构, 利用活性状态的结构作为模板, 采用分子置换法得到。结构清楚地显示了酶原状态未切除的肽段及糖基化的残基。自此, 利用 XFEL 获得高分辨的衍射数据, 并利用分子置换法解析结构的可行性完全得到验证, 但是对于全新结构的

解析还需要进一步验证。

2013 年, 德国施利希廷 (Ilme Schlichting) 研究组报道了利用单波长反常散射解析结构的结果。该实验利用浸泡的方式, 在溶菌酶晶体中引入重元素 Gd, 利用 Gd 的反常散射解析相位, 获得了结构。至此, XFEL 解析全新结构的可行性也得到验证。还有一个值得一提的工作是, 2013 年查普曼报道了利用 XFEL 得到的 G 蛋白偶联受体 (GPCR) 的高分辨结构。该实验证明了膜蛋白结构也可以在 XFEL 上获得, 对于结构生物学研究者是很有吸引力的。

XFEL 最吸引人的一点就是不需要晶体进行结构解析。由于不存在周期对称性, 但是在物理原理上对单颗粒的散射实验和晶体衍射非常类似 (一样是对样品的电子密度进行傅里叶变换, 一样需要解决相位问题), 这种实验经常被称为相干衍射成像。2011 年, 哈伊杜 (Janos Hajdu) 研究组报道了用 LCLS 的 XFEL 获得 mimivirus 病毒颗粒的衍射数据, 并重构了二维结构。虽然该实验还存在很多问题, 得到的二维结构可靠性也不高, 至今也没有报道三维结构, 但是这个结果表明了利用非晶体获取结构是可行的。

由于 X 射线自由电子激光极高的亮度和相干性, 使得解析蛋白质结构需要的晶体尺寸可以变得很小, 这将大幅度降低结构解析的难度。X 射线自由电子激光还可能实现不需要晶体的结构解析, 这对于结构生物学的影响将更加深远。

新一代的光源在结构生物学以及其他学科中潜在的重大推动作用, 使得世界各国纷纷提出了建设计划。在 X 射线自由电子激光的建设方面, 继美国的 LCLS、日本的 SACLA 建成之后, 美国将建设新的装置 LCLS-II, 欧洲决定在德国建设一台更强的 X 射线自由电子激光装置: 欧洲 X 射线自由电子激光 (European XFEL, 简称 EXFEL), 瑞士也计划建设自己的自由电子激光装置 SwissFEL, 意大利、韩国也提出了建设计划。在中国, 上海光源建成之后, 科学院正在开展软 X 射线自由电子激光装置的研制, 计划在未来建设硬 X 射线的自由电子激光装置。

同时, 新型的同步辐射装置也在建设, 这些新的同步辐射装置, 将能提供更高的亮度, 为科学研究提

供更多的机会。目前美国的 NSLS-II, 瑞典的 MAX-IV, 巴西的 Sirius 正在建设, 美国的 APS, 欧洲的 ESRF 和日本的 SPring-8 等已经运行了一段时间的光源也提出了新的改造计划。中国科学院也在规划建设新的同步辐射装置: 高能同步辐射光源 (High Energy Photon Source, 简称 HEPS)。这些新的同步辐射光源, 将比现有的同步辐射光源亮度上提高 2 个数量级, 为结构生物学提供了更好的研究前景。我们可以利用更小的晶体即可解析结构, 大大降低了结构解析的难度。需要指出的是, 同步辐射和 X 射线自由电子激光装置结合起来, 一个研究静态结构, 一个研究动态结构, 不仅为结构生物学提供了全新的研究天地, 更重要的是, 能够看到蛋白质在发挥功能时的“分子电影”, 对于通过结构阐释功能, 其意义更加重大。

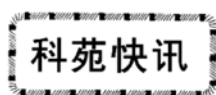
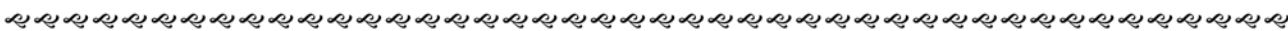
总结和展望

在晶体学一百年的发展历程中, 基于蛋白质晶体学的结构生物学研究为人类理解生命过程, 开发新药

起到了极其重要的作用。蛋白质晶体学, 或者结构生物学可以说是最近几十年来最为活跃的“明星学科”, 这一点从 *Nature*、*Science* 这些大家关注的科学杂志上, 频繁出现用蛋白质晶体学获得结构的研究文章可见一斑。

在蛋白质晶体学解析结构的过程中, 同步辐射提供了重要的手段。高亮度、微焦点、能量可调, 为解析蛋白质结构提供了不可或缺的技术手段。正是得益于同步辐射的大规模应用, 才使得结构生物学得到蓬勃的发展。

科学的需求永远是没有止境的。X 射线自由电子激光的出现, 新一代同步辐射光源的酝酿, 无不体现着科学前沿研究对实验装置和实验技术的进一步需求。我们相信, 同步辐射装置, 以及未来将要兴起的 X 射线自由电子激光和新一代光源, 必定能为蛋白质结构研究以及其他学科的前沿研究注入新的活力。



中国物理 C 发表粒子数据组
的《粒子物理综述》
施 郁

(复旦大学物理学系 200433)

《粒子物理综述》(Review of Particle Physics) 是粒子物理国际合作组织、依托于劳伦斯伯克利实验室的“粒子数据组 (Particle Data Group, 简写 PDG)”所作, 自 1978 年以来每两年 (偶数年) 发表一次印刷版, 每年更新一次网络版。它对粒子物理, 以及天体物理宇宙学中的相关领域, 作全面的总结。粒子数据组的唯一工作就是写《粒子物理综述》。2014 年 PDG 有来自 24 个国家的 140 个单位的 206 个作者。

《粒子物理综述》的前身可追溯到 1957 年著名物理学家盖尔曼 (M. Gell-Mann) 和罗森费尔德 (A. H. Rosenfeld) 在 *Annual Review of Nuclear Science* 发表的 *Hyperons and Heavy Mesons (Systematics and Decay)*。此后经历过好几个不同的名字。以前都发表于国际上主要的期刊, 如 *Nuclear Physics*, *Review of Modern Physics*, *Physics Letter B*, *Physical Review*

D, *European Physical Journal C*, *Journal of Physics G*.

2014 年的《粒子物理综述》有厚厚的 1675 页。分 3 部分: 粒子物理总结表 (Particle Physics Summary Tables)、

综述、表与图 (Reviews, Tables and Plots) 和粒子列表 (Particle Listings)。它包含很多综述文章、表和数据。

令人兴奋的是它今年 (2014 年) 9 月发表在中国物理学会旗下的中国物理 C 第 38 卷第 9 期, 封面上, *Chinese Physics C* 和 *Chinese Physical Society* 赫然在目。希望这一事件标志着中国物理期刊走向世界。

