

天然有机化学综述论文集

中国科学院有机化学研究所編輯

科学出版社

1966

单晶体X射线结构分析及其在研究天然有机物中的应用

范海福 林政炯 梁栋材

一、引言

单晶体X射线结构分析是测定物质内部结构的一种衍射分析方法。和其他测定物质结构的方法相比,单晶体衍射分析方法能以较少量的样品获得最全面的知识。它不仅能够提供分子结构的某些片段情况,而且可以给出分子在晶态下的全面而精细的立体图象。这个图象以百分之几Å的精确度描绘出原子在分子内部的立体分布以及分子之间在晶态下的相互联系。另一方面,这种方法也有它的缺点。首先,用目前以及最近将来可能有的实验技术,只能记录到所需实验数据的一半,而另一半则需用“猜”的办法来解决,因而工作量大、费时长*、需要较多的辅助知识;其次,它不分对象和场合,总是按照下面一套几乎一成不变的“程序”向研究者提供资料:晶体的对称性→晶胞参数、晶胞中分子数、分子量→原子在晶胞中的位置→分子的立体结构及分子间的连接情况→精确的键长、键角数据。这样就往往有一大部分的劳动要消耗在“制造”一些不受欢迎的“中间体”上。例如单单为了决定分子的绝对构型或者某一个化学键的长度,也必须先花费大量的时间去测定对称性、晶胞参数、晶胞中分子数、原子在晶胞中的位置等并非研究者所感兴趣的東西。已知的化学知识再丰富也只能使过程中的某些步骤简化,但却不能完全代替它们。最后,单晶体的衍射分析方法对样品还有较高的要求:样品必须是一颗完好的单晶体、具有合适的大小和形状**、对所用射线有足够的稳定性等等。所有这些缺点使单晶体的衍射分析方法远不如红外、紫外光谱以及其他一些物理方法那样广泛地应用于天然有机物的研究上。晶体的衍射分析方法,按射线的种类可分为X射线衍射、电子衍射、中子衍射三种,其中以X射线衍射的设备最完善、准确度最高、分析方法最成熟,因而应用也最广泛;如按样品的情况则又可分单晶体衍射和非单晶体衍射。所谓非单晶体,包括多晶体(均匀的或有一定织构的)及纤维样品。利用均匀的多晶样品(粉末样品)可以获得较准确的晶胞参数及衍射强度数据,但是粉末法对不同指标的衍射分辨本领很低、衍射图的解释又特别困难。其他非单晶样品也有类似的缺点。因此,在测定有机物的分子结构时,总是优先采用单晶体样品。本文以下的內容,除极个别的部分以外,都只限于单晶体。

X射线方法用于研究天然有机物自本世纪20年代始***,至今已有40多年。其效能

* 即使由电子计算机担任全部计算工作,要测定一个包含20—30个原子(氫原子除外)的分子结构,也需一年左右的时间。

** 通常希望样品尽量接近球状或圆柱状,直径在0.1—0.5mm之间。

*** W. H. Bragg 在1921年用X射线衍射方法研究萘和萘的晶体,开创了天然有机物的X射线分析工作,并首次用实验方法证实了有机物以分子状态存在于晶体中^[1]。

从验证化学结构式发展到独立测定分子结构(包括绝对构型);从只能给出半定量的数据发展到可以提供精确的键长、键角,其研究对象从简单的芳香族化合物^[1]发展到复杂的维生素 B₁₂^[2]乃至分子量达数万的蛋白质分子——肌红及血红球蛋白^[3]。目前已有大量的天然有机物用X射线衍射方法仔细地研究过,它们分布在天然有机化学的各个重要领域。X射线衍射分析的结果不但帮助精确测定天然有机物(特别是大分子)的分子量、帮助确定化学结构式、测定分子的立体结构,还为有机结构化学及生物化学的理论研究提供大量的键长、键角数据以作为建立和发展这些理论的实验基础。

尽管单晶体X射线结构分析在研究天然有机物方面已获得不少成果,但是它还远没有成为天然有机化学家所充分熟悉和乐于普遍采用的工具;它所接触到的天然有机物也还只是这个无边大洋中的一滴。究其原因,主要是由于上面提到过的一些缺陷。其次,晶体学工作者和化学工作者之间缺乏比较深入的相互了解和密切配合也妨碍了这种方法的迅速推广。因此,要使单晶体X射线结构分析这一工具更好地为天然有机化学的研究服务,除了迅速提高这一工具的效率外,增进晶体学工作者和化学工作者之间的相互了解和加强相互的协作是一个相当迫切的问题。本文的目的之一,就是希望在化学工作者了解单晶体X射线结构分析的方法和应用方面,起一点微小的作用。

二、方法略述

X射线单晶体结构分析的任务是根据晶体的衍射效应引出原子的立体分布。晶体的衍射效应通常用照相方法记录成所谓衍射图(图1);或用计数器记录成所谓衍射谱;原子

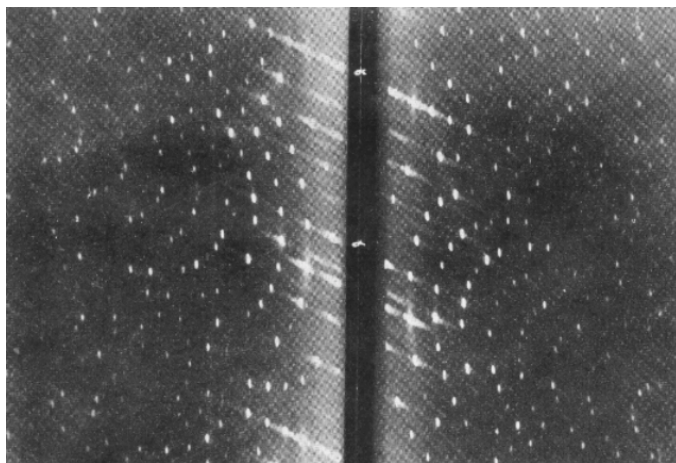
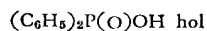


图 1



$\lambda_{\text{Cu},\text{Ni}}$ 滤波, 48 小时, 1962.9.20.

的立体分布则用所谓电子密度分布函数表示, 这个函数在每一个原子中心的位置都达到一个极大值。将这种函数绘成等位线图(图 2), 则每一个原子都相应于一个“高峯”, 峯顶的位置就是原子中心的位置, 从这些位置可以精确地画出分子的立体图象。X 射线单晶体结构分析的理论基础在于晶体的衍射结构振幅(一个表示晶体衍射效应的物理量)和晶体的电子密度分布函数之间有下列关系:

$$F(\mathbf{H}) = \Phi \rho(\mathbf{r}) = \sum_{j=1}^N f_j e^{2\pi i \mathbf{H} \cdot \mathbf{r}_j} \quad (1)$$

$$\rho(\mathbf{r}) = \Phi^{-1} F(\mathbf{H}) = \frac{1}{V} \sum_{\mathbf{H}} F(\mathbf{H}) e^{-2\pi i \mathbf{H} \cdot \mathbf{r}} \quad (2)$$

式中 $F(\mathbf{H})$ 是 $F(hkl)$ 的简写, 它代表衍射指标为 (hkl) 的结构振幅*; $\rho(\mathbf{r})$ 是 $\rho(xyz)$ 的

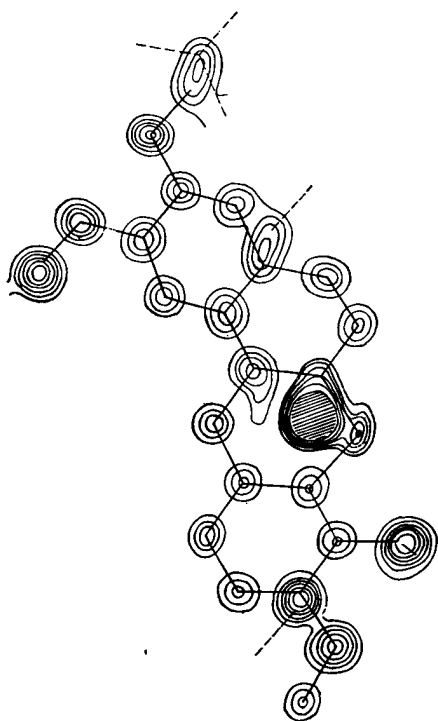


图 2 延胡索乙素碘氨酸盐的电子密度投影图

简写, 它代表晶体的电子密度分布函数; xyz 是晶胞坐标系中的分数坐标; f_j 是晶胞中第 j 个原子的散射因子(可以从理论计算求得); N 是晶胞中原子的数目; \mathbf{r}_j 是 x_j, y_j, z_j 的简写, 代表第 j 个原子在晶胞中的坐标; $\mathbf{H} \cdot \mathbf{r}_j = hx_j + ky_j + lz_j$; Φ 及 Φ^{-1} 分别代表傅立叶正变换及反变换; V 是晶胞的体积(可以从实验测得)。从式(2)可知, 只要由实验求得 $F(\mathbf{H})$ 就可以直接计算出 $\rho(\mathbf{r})$, 但是从实验不论是用照相方法还是用计数器方法都只能记录到 $F(\mathbf{H})$ 的绝对值而丢掉了它的相角或者正负号。这就给晶体结构分析带来了一个根本性的困难, 它要求我们从结构振幅的绝对值出发, 直接或者间接地找出它的相角或者正负号。看来这是一个不合情理的要求, 就象一付对联, 只告诉你上联而要你一字不差地猜出下联一样。但是如果知道了这付对联写的是什么内容以及写对联时有些什么背景, 再细致地研究了上联的笔调, 然后着手猜, 那么虽然不能绝对地一字不差, 但是也

不致于完全束手无策。单晶体的 X 射线结构分析也是如此。如果已知晶体化学成分、晶体的对称性、以及一些其他线索, 再细致地分析了结构振幅绝对值的变化情况, 就有可能近似地“猜”出它的“下联”——相角, 从而解决整个结构。单晶体 X 射线结构分析的整个过程就是以这个“猜下联”的问题, 亦即晶体学家通常所谓的“相角问题”, 作为中心环节的。

* $F(\mathbf{H})$ 在一般情况下是一个复数, 可以写成 $|F(\mathbf{H})| e^{i\alpha_{\mathbf{H}}}$, 其中 $|F(\mathbf{H})|$ 是 $F(\mathbf{H})$ 的绝对值; $\alpha_{\mathbf{H}}$ 是 $F(\mathbf{H})$ 的相角。当晶体有对称中心时, $\alpha_{\mathbf{H}}$ 只有 0° 和 180° 两种可能, 此时 $F(\mathbf{H})$ 可写成 $S_{\mathbf{H}} |F(\mathbf{H})|$, 其中 $S_{\mathbf{H}}$ 等于 $+1$ 或 -1 , 代表 $F(\mathbf{H})$ 的正负号。

在现阶段,相角问题的解决办法主要的有下面几种:

(一) 尝试法 这是一种比较原始的办法,其特点是连“猜”带“碰”。即先根据化学知识、晶体的对称性、以及其他一切可以利用的线索,“猜”出若干个可能的结构模型,然后将模型中的原子坐标 r_i 代入式(1)求算 $F(\mathbf{H})$, 碰碰看那一个模型算出的 $F(\mathbf{H})$ 能与实验求得的 $|F(\mathbf{H})|$ 吻合。如果都失败了,那就再“猜”一次、再“碰”一阵,直到找出合理的模型为止。一旦找到了基本上合理的结构模型,就可以开始所谓“逐步逼近”的工作:即利用模型上的原子坐标 r_i 计算出 $F(\mathbf{H})$, 取其相角或正负号,配以实验测得的 $|F(\mathbf{H})|$, 代入式(2)就可以算出近似的电子密度函数,由此又可以求出新的一套 r_i 。利用这套 r_i 计算出的结构振幅相角,将比由原始模型计算的更接近实际情况。循环重复上述步骤,就可以获得完全正确的结构振幅相角或正负号(如果不考虑实验误差的话)。以此代入式(2)就可以获得精确的原子坐标。在晶体结构分析发展的初期,尝试法曾经是一个主要方法。用这种方法不仅解决了一些已知化学结构的有机晶体,如六次甲基四胺^[4]、六卤代苯及六甲基苯^[5]、脲素及硫脲^[6]等,从而提供了第一批关于有机分子的立体化学数据,而且还成功地测定了一些未知化学结构的长链化合物的链长和羰基位置^[7], 开创了用X射线衍射方法测定化学结构式的工作。目前,尝试法虽然在许多方面都已被新的、更有效的方法所代替,但是在一些特殊的场合下,或者在分析过程的某些环节中却仍然占据着重要的位置。这是由于它本身在“猜”和“碰”的技巧上不断有所发展^[8],而且还由于其他方法往往需要它作为辅助工具。

(二) Patterson 法 尽管在1930年以前,利用X射线研究有机物已有一些成功的尝试,但是这一方法广泛地进入有机化学领域却是在1934年以后,这首先应归功于Patterson的发现。他在用X射线对天然有机物进行了多年的研究以后,感觉当时的X射线分析水平远远不能满足解决天然有机物结构的需要。用他自己的话来说,在当时,“我们所能做的一切就是测定一大堆空间羣”^{*}!为了改变这种情况,经过长期的探索,终于在1934年提出了一个所谓“原子间向量函数”(后来被称为Patterson函数)^[9],其定义如下:

$$P(\mathbf{u}) = \int_V \rho(\mathbf{r})\rho(\mathbf{r} + \mathbf{u})dV = \frac{1}{V} \sum_H |F(\mathbf{H})|^2 e^{-2\pi i \mathbf{H} \cdot \mathbf{u}} \quad (3)$$

其中 $P(\mathbf{u})$ 代表 Patterson 函数, \mathbf{u} 是它的自变量,其余符号与前相同。这个函数的形式和电子密度函数[式(2)]非常相似,只是以 $|F(\mathbf{H})|^2$ 代替了电子密度函数中的 $F(\mathbf{H})$ 。但是这样一来却使 Patterson 函数同电子密度函数有本质的区别:它可以毫不含糊地单靠实验数据求得。另一方面它同电子密度函数之间又存在以下的联系:在电子密度函数中每出现两个“高峯”(原子),相应地在 Patterson 函数上就会产生一个“高峯”,这个峯的大小正比于两原子序数的乘积,而其坐标和原点间的矢量则等于连接两原子的矢量(图3)。根据这一性质,可以用推理的办法由 Patterson 函数引出原子的坐标^[10]从而绕过了相角的问题。Patterson 方法特别适用于测定含“重原子”^{**}的晶体结构。因为通常“重原子”的 Patterson 峯数目较少而又特别突出,利用 Patterson 函数甚至在没有任何其他资料的情况下也很容易找出其坐标,一旦找到“重原子”的坐标,整个分析过程就可以大为简化。对于

* A. L. Patterson, Personal Reminiscences, in “Fifty Years of X-Ray Diffraction” ed. by P. P. Ewald (1962).

** 所谓重原子,通常是相对于 C, N, O 等原子而言,其序数一般都在 10 以上。

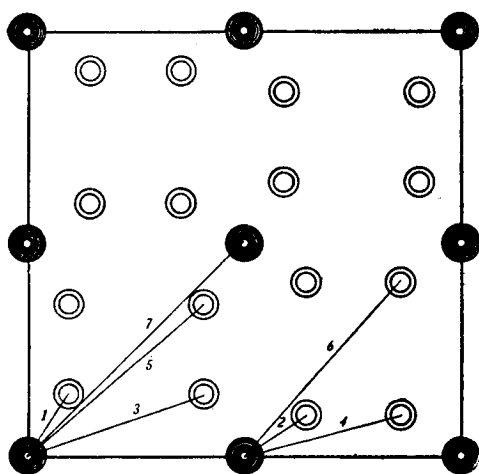


图 3a KH_2PO_4 的电子密度图

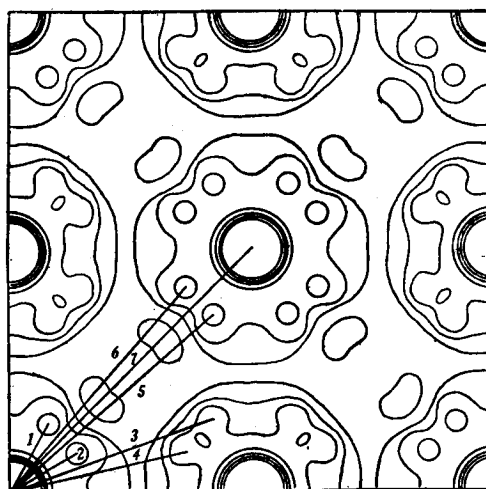


图 3b KH_2PO_4 的 Patterson 图

不含重原子的晶体结构,分析起来要困难一些。但是,在这种情况下,Patterson 方法也要比其他方法更有把握。因此到目前为止,大部分无重原子的晶体结构都是用它来解决的。

Patterson 法的出现为单晶体结构分析揭开了新的一页,把它从依附于化学假设的地位解放出来,使分析过程从“猜”和“碰”变成了逻辑推理。这个方法经过一系列的发展^[11-15]现已成为结构分析中最常用的一种方法,并且还是另外一些重要分析方法(重原子法、同晶型置换法等)所不可缺少的基础。

(三) 重原子法 当晶胞中含有比较“重”的原子时,由于它对 X 射线的散射能力远大于其他原子,因此可以近似地认为大部分结构振幅的相角或者正负号只由重原子决定。这样,只要知道重原子的位置就可以近似地推算结构振幅的相角或正负号,从而获得一个近似的电子密度分布函数。借此就能够求出其他原子的粗略位置,再利用逐步逼近就可以得到精确的结构。这就是所谓重原子法,其整个过程示意表如图 4。在这种方法中,所用“重原子”的“重量”(原子序)应当适中。太轻,不足以控制大部分结构振幅的相角;太重,则又“月明星稀”,所得电子密度图将只见重原子而难辨轻原子,它们都将淹没在重原子的“余辉”——由实验误差引起的散乱背景——之中。H. Lipson 和 W. Cochran^[16]根据衍射强度分布的统计规律认为,最理想的情况是使晶胞中重原子序数的平方和约略等于轻原子序数的平方和。此时约有四分之三的结构振幅其相角或正负号由重原子控制。

重原子法的原理早在 1929 年就已由 W. L. Bragg 提出^[17],但是它的大量应用以及理论上的充实^[18]却是在 Patterson 函数发现以后。因为在此以前还没有一种可靠的办法能够迅速而有效地从实验数据求出重原子的位置。由于一般的有机分子都很容易用置换、成盐、络合等方法加入卤素或者某些金属原子,这一方法和 Patterson 法相结合已成为用 X 射线研究有机晶体,特别是在测定化学结构及分子的构型、构象等方面,应用最广泛的工具^[19]。维生素 B_{12} 的晶体结构(目前已测定的最复杂结构*)^[2]就是用典型的重原

* 分辨率低于 1\AA 的球蛋白结构除外。



图 4a Patterson 图

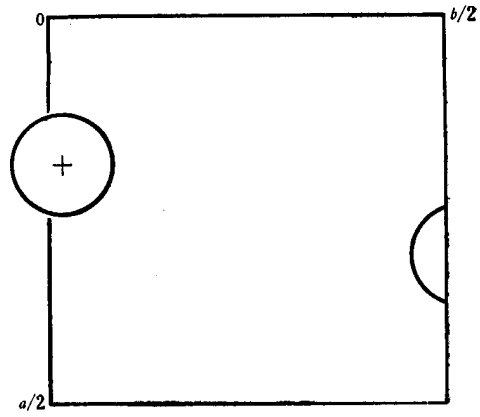


图 4b 重原子位置

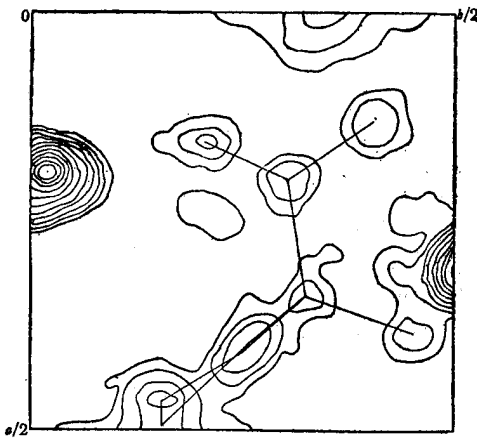


图 4c 重原子相角的电子密度图

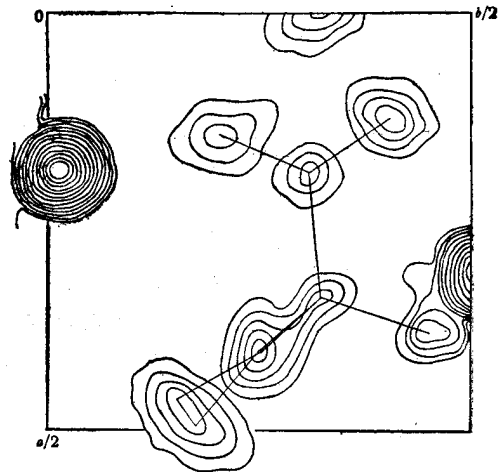


图 4d 逼近两次的电子密度图

子法测定的。重原子法也有美中不足之处。这就是重原子的存在使轻原子的位置不易测得非常准确。于是当需要获得高精度度的结果时,总尽量避免使用含重原子的晶体,因而也就不能利用重原子法。

(四) 同晶型置换法 在晶胞中某一个或者几个位置上增加、减少、或者调换一些原子而保持其对称性及晶胞参数不变,这称为同晶型置换。置换前和置换后的晶体称为“同晶型”的晶体。晶体在置换前后其衍射效应会发生改变,根据这个改变,如果知道发生置换的位置及置换原子的种类就可以从下面的推导求出结构振幅的相角或正负号。

设有两同晶型晶体,其一包含置换部分 Q_1 (由置换原子组成) 及非置换部分 P (由剩余的原子组成); 其二包含置换部分 Q_2 及非置换部分 P 。则这两个晶体的结构振幅分别为

$$F_1(\mathbf{H}) = F_P(\mathbf{H}) + F_{Q_1}(\mathbf{H})$$

$$F_2(\mathbf{H}) = F_P(\mathbf{H}) + F_{Q_2}(\mathbf{H})$$

由此得:

$$F_2(\mathbf{H}) - F_1(\mathbf{H}) = F_{Q_2}(\mathbf{H}) - F_{Q_1}(\mathbf{H}) \quad (4)$$

上式的 $F_{Q_2}(\mathbf{H})$ 及 $F_{Q_1}(\mathbf{H})$ 可根据置换位置及置换原子的种类算出, 而 $F_2(\mathbf{H})$ 及 $F_1(\mathbf{H})$ 的绝对值则可从实验求得. 由这些数据就可以解出 $F_2(\mathbf{H})$ 及 $F_1(\mathbf{H})$ 的相角. 这样解出的相角在一般情况下有两个答案 (见图 5)*, 而其中只有一个是正确的. 要判别真伪, 可以

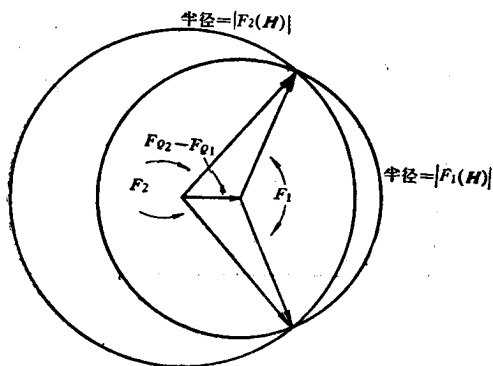


图 5

可以再引进一个同晶型晶体. 假设它由置换部分 Q_3^{**} 及非置换部分 P 组成, 则

$$F_3(\mathbf{H}) = F_P(\mathbf{H}) + F_{Q_3}(\mathbf{H})$$

$$F_3(\mathbf{H}) - F_1(\mathbf{H}) = F_{Q_3}(\mathbf{H}) - F_{Q_1}(\mathbf{H}) \quad (5)$$

此时, 对于 $F_1(\mathbf{H})$ 的相角, 式(4)和式(5)将有一个公共解, 这就是正确的解.

同晶型置换法的原理首先由 Cork 提出, 经过一系列的发展^[20,21], 现已成为解决复杂结构的最有效工具. 应用同晶型置换法需有两个条件. 第一个条件是必须有足够数目***、合于应用的同晶型晶体: 要求置换原子

同时又是“重原子”, 而且不同晶体中重原子的序数有较大的差别. 这对于有机化合物并不十分困难, 因为同种有机分子的不同卤代物、同种有机分子的不同卤氢酸盐、以及同种有机分子与不同金属盐类形成的络合物等, 往往容易形成同晶型晶体, 而卤素原子或者金属原子则既是置换原子又是重原子. 第二个条件是必须知道置换原子的位置. 由于一般置换原子的数目较少而且又是重原子, 这个问题不难用 Patterson 法解决. 一旦满足了上述条件, 同晶型置换法就能发挥出巨大的威力. 特别是在测定大分子的结构方面, 同晶型置换法远比重原子法有效. 例如要测定一个包含 1000 个轻原子的分子结构, 用重原子法就需要在晶胞中加入一个序数约为 200 的原子! 而用同晶型置换法则只要一个序数约 70 的原子就够了. 由于同晶型置换法需要较大的工作量, 而且对许多不很大的有机分子又不容易找到三个置换位置不尽相同的同晶型晶体, 因此它的应用远不如重原子法那样广泛. 但是对于复杂的分子, 特别是生物大分子结构的直接探索, 同晶型置换法却是目前最有效的方法. 利用这种方法已经成功地测定出肌红和血红球蛋白的分子结构^[3,21], 使人们第一次“看”到了生物大分子内部的立体结构情况.

(五) 反常散射法 一般认为原子散射因子与所用 X 射线的波长无关. 这就是正常散射的情况. 当原子的序数较大, 特别是当其吸收边与所用 X 射线波长接近时, 就可以观察到违反上述情况的现象. 此时的散射称为反常散射. 当原子发生反常散射时, 其原子散射因子可以表示如下^[22]:

$$f = f^{\circ} + \Delta f' + i\Delta f'' \quad (6)$$

其中 f° 代表正常散射因子, 而 $\Delta f'$ 和 $\Delta f''$ 分别代表反常校正值的实分量和虚分量, 它们都是入射 X 射线波长的函数, 其数值 (见参考文献^[23]) 可以从理论计算求得. 如果一个晶体结构由正常散射的原子 P 及反常散射原子 Q 所组成, 则其结构振幅可写成

* 只有在晶体具有对称中心时答案才是唯一的. 此时由于 $F_1(\mathbf{H})$, $F_2(\mathbf{H})$, $F_{Q_1}(\mathbf{H})$, $F_{Q_2}(\mathbf{H})$ 都变成实数, 图 5 的两个三角形就变成一根直线.

** Q_3 的位置不能和 Q_1 完全相同.

*** 对于非中心对称晶体, 需要三个同晶型晶体; 对于中心对称晶体, 则只需两个.

$$F(\mathbf{H}) = F_P(\mathbf{H}) + F'_Q(\mathbf{H}) + F''_Q(\mathbf{H}) \quad (7)$$

其中

$$F_P(\mathbf{H}) = \sum_P f_P e^{2\pi i \mathbf{H} \cdot \mathbf{r}_P};$$

$$F'_Q(\mathbf{H}) = \sum_Q (f'_Q + \Delta f'_Q) e^{2\pi i \mathbf{H} \cdot \mathbf{r}_Q};$$

$$F''_Q(\mathbf{H}) = \sum_Q i \Delta f''_Q e^{2\pi i \mathbf{H} \cdot \mathbf{r}_Q}.$$

由此可见 $F''_Q(\mathbf{H})$ 的相角永远超前于 $F'_Q(\mathbf{H})$ 的相角 90° 。因此当一个非中心对称的晶体有反常散射效应时 $|F(\mathbf{H})|$ 将不等于 $|F(-\mathbf{H})|$ (图6), 二者的大小关系由晶体中分子的绝对构型决定。此性质一方面可用于测定分子的绝对构型, 另一方面还可以用于测定非中心对称晶体的结构振幅相角:

由式(7):

$$F(\mathbf{H}) = F_P(\mathbf{H}) + F'_Q(\mathbf{H}) + F''_Q(\mathbf{H})$$

$$F(-\mathbf{H})^* = F_P(-\mathbf{H})^* + F'_Q(-\mathbf{H})^* + F''_Q(-\mathbf{H})^*$$

其中 $F(-\mathbf{H})^*$ 是 $F(-\mathbf{H})$ 的共轭复数。

从图6可见 $F_P(-\mathbf{H})^* = F_P(\mathbf{H})$; $F'_Q(-\mathbf{H})^* = F'_Q(\mathbf{H})$; $F''_Q(\mathbf{H})^* = -F''_Q(\mathbf{H})$ 。因此

$$F(\mathbf{H}) - F(-\mathbf{H})^* = 2F''_Q(\mathbf{H}) \quad (8)$$

只要知道 Q 原子的位置就可以算出式(8)右边的 $F''_Q(\mathbf{H})$ 。由于只有重原子才具有可觉察的反常散射效应, 因此 Q 原子必然是重原子, 而其位置很容易从 Patterson 函数求出。式(8)左边 $F(\mathbf{H})$ 及 $F(-\mathbf{H})^*$ 的绝对值则可以从实验求得。根据这些数据就可以解出 $F(\mathbf{H})$ 及 $F(-\mathbf{H})$ 的相角(图7)。图7和图5相似, 也包含两个可能的答案。要决定其取舍, 可以再引用另一种波长的X射线衍射数据, 或者使与同晶型置换法结合。至于中心对称的晶体, 利用反常散射效应, 原则上可以获得类似于同晶型置换法的结果。

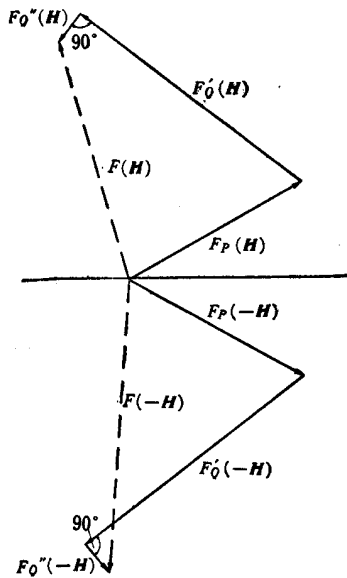


图 6

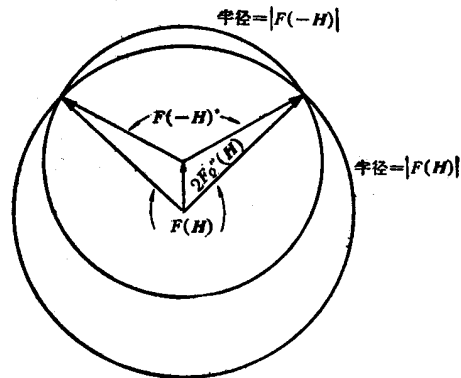


图 7

由于一般重原子的反常散射效应只占其总散射效应的5%，最大的也只有20%左右，不容易准确测量，这一方法目前多半局限于判别分子的绝对构型，或者作为同晶型置换法的补充。至于用作解决相角问题的独立手段，则尚有待进一步发展。关于反常散射法的更详细叙述可参看有关文献^[24]。

(六) 直接法：按照比较狭义的理解，所谓直接法，就是从结构振幅的绝对值“直接”地推导其相角的办法，而相对地，其他一切方法都称为“间接法”。直接法最早的尝试始于 Banerjee^[25]，而比较集中和迅速的发展则是在1947年 Harker 和 Kasper 的工作^[26]发表以后。直接法的特点是：原则上它可以不依赖于化学假设而纯粹从衍射强度直接地推出晶体结构；对于无重原子的晶体结构，有时候可以比其他方法更快地获得结果；整个过程很容易由电子计算机来操作，因此便于实现自动化。但是这类方法的局限性较大，它要求晶体具有对称中心、晶胞中原子数目较少*、晶胞中原子的序数大致相同、衍射强度的分布符合统计规律等等。到目前为止，总的说来，直接法的效率还远没有达到 Patterson、重原子等方法水平。关于直接法的详细情况可查有关文献^[25-27]。

目前，“相角问题”仍然是研究单晶体X射线结构分析中的关键性问题。围绕这个问题已经进行了而且还在进行着大量的研究。新近的成果及动向可以从最近的两个学术会议**中略见一斑。在这些工作中，最突出的趋势是使电子计算技术和各种解决相角问题的方法紧密地结合起来，以达到使晶体结构分析过程“自动化”的目的。

单晶体X射线结构分析的“程序”示于图8。它可分为三个阶段。第一阶段是准备阶段，它从准备晶体样品开始而以定出晶体的对称性及收集完衍射强度数据而告终；第二阶段是分析阶段，其任务是解决相角问题、求出电子密度，因而是最关键的阶段；第三阶段是精确修正阶段。它通常被看作第二阶段的一部分，但是其目的、任务和方法都与第二阶段迥异，而且由于计算技术和修正技巧的发展，结构分析的精确度大大提高，要求高精确度的工作也迅速增多，使得这一部分已经越来越有独立的必要。第三阶段的任务是将第二阶段所得的结构模型精确化，以便获得研究化学键的性质或者其他理论问题所需的精确键长、键角数据。所用方法最主要的有差值级数法^[28]和最小二乘法^[29]。这两种办法都依据同一原则：即认为最准确的原子坐标应该能使结构振幅的理论值最接近其实验值。因此系统地改变每一原子的坐标参数及温度参数(校正原子热振动效应的参数)，使得结构振幅理论值和实验值之差趋于极小，就可以获得实验误差所容许的最准确的原子坐标。通常用以表示结构振幅理论值和实验值之间偏差的因子称为偏离因子，其定义为：

$$R = \sum_H \left| |F_{\text{实验}}| - |F_{\text{理论}}| \right| / \sum_H |F_{\text{实验}}|$$

它在一定程度上代表着分析结果的准确度。但是准确度的更直接和更确切的表示则用原子坐标参数的标准误差。此两者之间有一定的联系^[30]。根据经验，一个正确的结构模型通常可使 $R < 30\%$ ，而当 R 小至 10% 时结果已相当准确，其时原子参数的标准误差约为 $\pm 0.02 \text{ \AA}$ 。

* 到目前为止，用直接法解出的最复杂结构含有36个独立原子。参看 I. L. Karle, H. Hauptman, J. Karle, and A. B. Wing, *Acta Cryst.* **11**, 257 (1958)。

** 1960: "Computing Methods and the Phase Problem in X-Ray Crystal Analysis" (已篇成专集出版), ed. by R. Pepinsky et al., Pergamon Press, Oxford (1961)。

1962: "Neue Fortschritte zur experimentelle und theoretischen Methodik der Kristallstrukturforschung". München. 见 *Physik. Verhandl.* **13**, 116-133 (1962)。

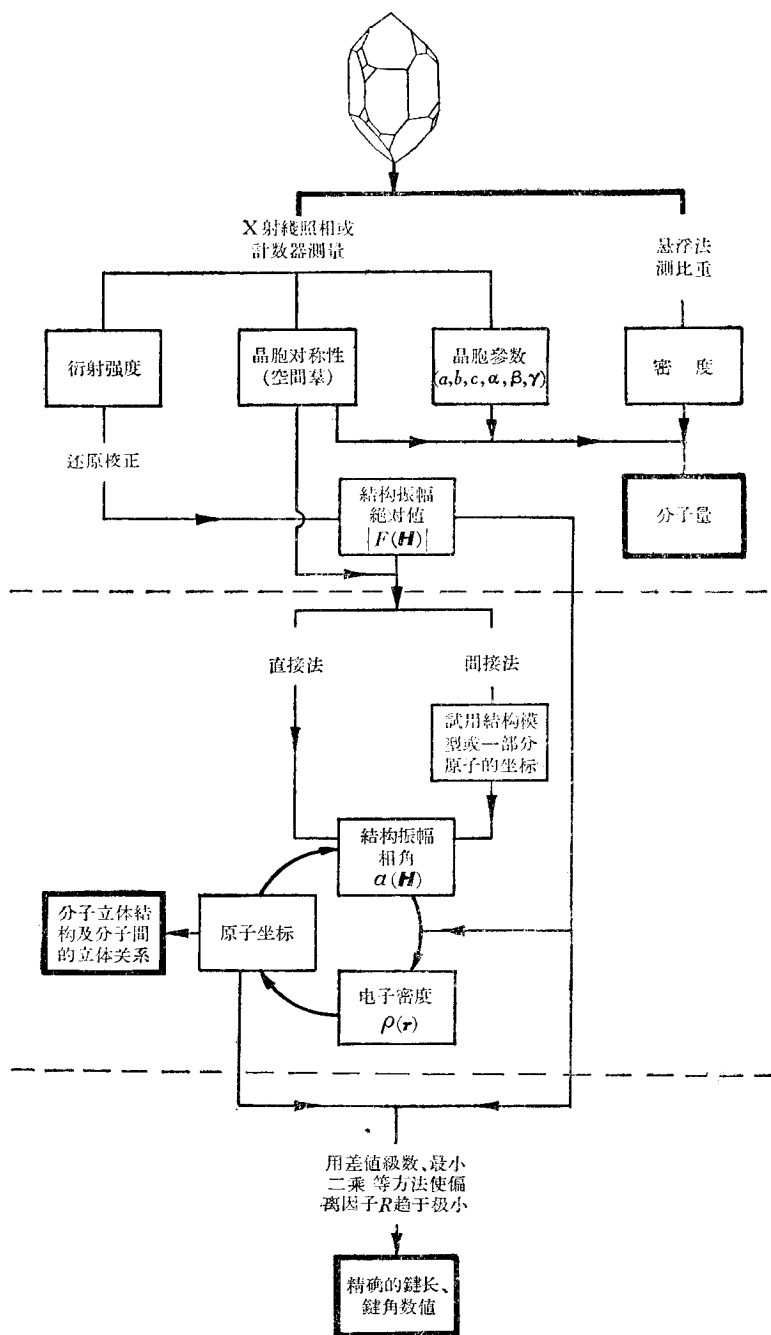


图 8

三、应用简介

从单晶体X射线结构分析三个阶段都可以产生有用的结果。例如第一阶段虽然主要是为后两阶段作准备，但是仍然可以通过所测得的晶胞参数提供分子量的数据。其根

据可由下式表示:

$$M = N\rho V/Z$$

式中 M 是分子量; N 是阿佛加得罗数(6.023×10^{23} 克分子 $^{-1}$); ρ 是晶体密度; V 是晶胞体积(可从晶胞参数算出); Z 是晶胞中分子数。因此只要测得晶胞参数和晶体的密度, 再知道晶胞中分子数或者近似的分子量, 就可以求出准确的分子量。如果不知道分子数和近似的分子量, 则只能求出分子量的整倍数 ZM 。此时若能测出同一物质在不同结晶状态下的各种 ZM 值, 则其最小公约数就可能是分子量的真实值。X射线方法一般不宜独立地测定分子量, 但是由于它的准确度相当高*, 而样品用量又极少**, 因而往往是其他方法的得力助手。其结果除可与一般物理化学方法的结果相互验证外, 有时还可以纠正它们的错误。例如由化学方法测得 Clerodin 的化学式原为 $C_{21}H_{30}O_6$, 但X射线测得的分子量表明, 若保持 C, H, O 的比例基本上不变, 其化学式应改为 $C_{24}H_{34}O_7$ 。这在后来的电子密度图中得到证实^[31]。在测定结晶态蛋白质的分子量方面, X射线方法有着重要的地位^[32]。它除了给出准确的分子量以外, 还可以顺便提供有关分子几何形状的一些线索。胰岛素的分子量原由一般物理化学方法测得为 30000—50000。D. Crowfoot^[33]用X射线方法测得 $ZM = 37600$ 。因该晶体的晶胞由三个等同部分组成, 故 Crowfoot 认为胰岛素的分子量或其整倍数应约等于 12000。其后, 关于胰岛素的分子量, 在 6000 和 12000 之间发生了争论。最近的X射线观察^[34]发现胰岛素在许多不同的结晶状态下都以一个不变的二聚体作为最小单位。此二聚体的两个等同部分由二重旋转轴联系, 每一部分的“分子量”为 5733, 而整个二聚体的分子量则二倍其数。

但是, 单晶体X射线结构分析的主要应用, 远不在于测定分子量, 而在于观察物质的分子结构及分子间在晶态下的作用方式。为解决这些问题, 绝大多数都需要利用第二、三阶的结果。目前, 在天然有机物领域中, 这种类型的应用按对象的特点、研究的目的、以及解决问题的途径, 大致可分为两个方面:

(一) 直接解决天然有机物分子的化学结构、相对构型、绝对构型、构象等问题 单晶体X射线结构分析可以从一个大致上正确的分子式(晶体学中通常称为化学式)开始, 对上述问题全面地加以回答; 也可以从更多的化学资料出发, 只对构型、构象中的个别问题作出结论。但是由于所有这些问题通常都必须等到结构分析的第二阶段基本结束才可以解决, 因此上述两种情况虽有难易之分, 但是所需时间的长短却无大差别。

利用X射线方法解决天然有机物的化学结构, 恐怕是这一方法对许多天然有机化学家最有吸引力的一种应用。尤其当研究对象的量非常少而其化学反应又比较特殊时更是如此。另一方面, 对晶体学家说来, 这又是一个异常艰巨但却兴味无穷的课题。这使它反过来成为促进单晶体X射线结构分析理论和技术发展的一个主要动力。下面是各个历史发展阶段中具有代表性的一些例子。

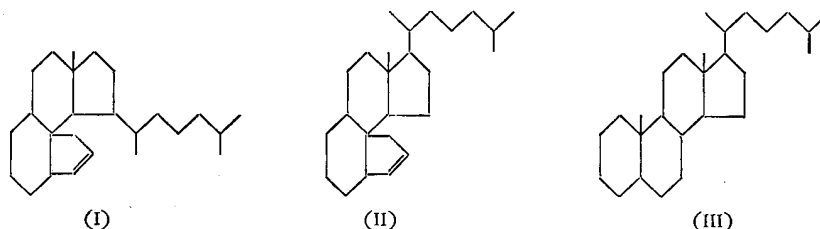
Shearer(1925)^[7]研究了一系列长链化合物, 发现了碳链在晶胞中取向的规律以及链长和晶胞参数之间的关系, 成功地测出一系列长链化合物的链长。另外, 他还根据衍射强度数据用尝试法测定出羰基的位置以及它在碳链中所占的长度。这一工作从结构的角度来看虽不完整, 但是它证明了X射线方法尽管在技术状态非常原始的阶段仍然能够为测

* 测定一般物质分子量的误差约为 0.1%; 测定蛋白质等大分子则约为 1%。

** 约 0.1 毫克即可。

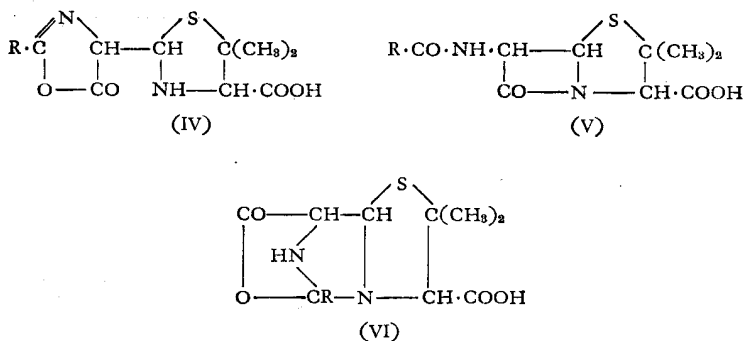
定化学结构作出贡献,从而迈开了利用X射线研究未知结构有机物的第一步。

另一个早期的例子是1932年 Bernal 对麦角甾醇的研究。当时一般认为麦角甾醇的结构式为(I)。Bernal^[35]根据X射线测得的晶胞参数,结合对晶体的光性研究,指出其分子大小应为 $5.0 \times 7.2 \times 17-20 \text{ \AA}$ 。这显然与(I)矛盾,因而他建议其结构式应改为(II)。随后 Rosenheim 及 King^[36]重新探讨了麦角甾醇的结构并获得了正确的结构式(III)。这



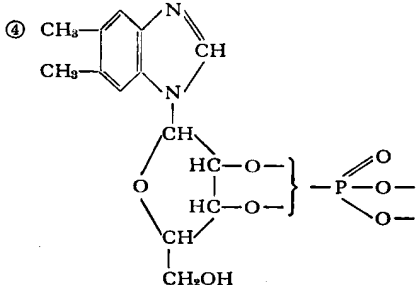
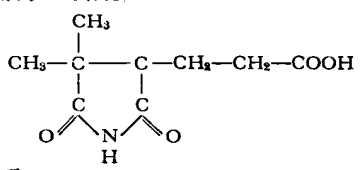
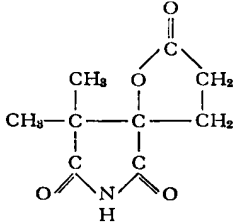
是只利用属于结构分析第一阶段的结果而对分子形状、大小及结构式的判断作出重要贡献的一个出色例子。这一方法后来还常用于估计晶态蛋白质分子的大小和形状。

自1934年 Patterson 法创立起至第二次世界大战结束约十年间,单晶体X射线结构分析在理论上和技术上都有迅速的发展。青霉素的X射线结构分析则可以说是这种发展的集中表现。这一工作始于1943年而终于1945年。整个过程都和化学方法密切配合。X射线方法不仅以最直观的方式确证了青霉素的化学结构和测定了它的分子构型,而且还帮助解决了化学降解和合成过程中提出来的一系列鉴定问题。从单晶体X射线结构分析的角度看,这一工作不仅利用了几乎全部当时已有的成就,而且在分析过程中还创造了有价值的新方法。在分析方案上,这一工作采用了几种晶型不尽相同的晶体同时分析、相互对比的办法;在求解结构模型时,采用了 Patterson 法、同晶型置换法、重原子法和尝试法;在运用尝试法中 Bunn 发展了光学模拟技术,创造了所谓“蝇目法”。另外他还提出了所谓“误差综合法”,用以判别试用结构模型中原子位置的真伪。在计算傅立叶级数中还采用了当时的最新成就——穿孔卡片计算机^[38],进行了当时罕见的三维电子密度级数修正工作。青霉素的X射线结构测定是从它的苯基钠盐、苯基钾盐和苯基铷盐入手的。在刚开始分析时,化学方法给出了三种可能的结构式(IV),(V)和(VI)。X射线分析结果最后确证其结构为(V),并由电子密度图(图9)获得其构型(图10)。



维生素 B₁₂ 的晶体结构分析^[39]可以作为五十年代的代表。它开始于1948年而基本完成于1955年。其分析过程及结果示于表1。虽然这一工作在八、九年以前即已基本结

表 1

年份	化学及光谱结果	X 射线结果
1948	—	开始干晶体 X 射线照相工作
1949 1951	<p>I. 测出分子式:</p> <p>① $C_{61-64}H_{83-92}N_{14}O_{18-20}PCo$. (1949)</p> <p>② $C_{63}H_{84}N_{14}O_{14}PCo$. (1951)</p> <p>II. 发现有下列碎片:</p> <p>① $-C\equiv N$</p> <p>② $-NHCH_2CHOHCH_3$ (1 或 2 个)</p> <p>③ $-NH_2$ (5—6 个)</p> <p>④ </p> <p>⑤ $CoC_{40-46}H_{89-98}O_{4-12}N_{4-5}$</p> <p>III. 对碎片的连接方式及含钴碎片的内容有如下估计:</p> <p>① 类核甙酸部分可能一方面由苯并咪唑中的 N_9, 另一方面则由磷酸根通过丙醇胺与 Co 原子相连.</p> <p>② 含 Co 碎片中可能有一个类似卟啉的平面基团.</p>	<p>I. 从干晶体的三维 Patterson 分析:</p> <p>① 发现了 Co 和 $-C\equiv N$ 在晶胞中的位置.</p> <p>② 发现在钴原子周围有一个垂直于 $Co-C\equiv N$ 连线的平面基团, 其大小和形状类似于卟啉.</p> <p>③ 发现在 $Co-C\equiv N$ 连线的反方向有一条链状的痕迹, 估计是核甙酸与钴原子的“接头”.</p> <p>II. 利用钴原子相角求出三维电子密度, 经分析后得以下结果:</p> <p>① 再次看到与“1”相似的情况.</p> <p>② 发现了苯并咪唑及磷原子的位置, 并肯定了呋喃核糖上磷酸的取代位置是 3' (图 11).</p> <p>III. 对湿晶体作了三维 Patterson 和三维电子密度图分析, 结果与干晶体相似.</p>
1952 1954	<p>制成游离的含钴碎片——一个六羧酸, 并测得其分子式为 $C_{47}H_{60}O_{16}N_6CoCl$.</p>	<p>I. 从 $B_{12}SeCN$ 晶体的三维 Patterson 分析: 证实了干晶体中 $-C\equiv N$ 的位置.</p> <p>II. 从 $B_{12}SeCN$ 晶体的三维电子密度分析: 发现了钴原子周围的平面基团是一个特殊的五环系统(图 12).</p> <p>III. 从六羧酸 $C_{47}H_{60}O_{16}N_6CoCl$ 晶体的三维电子密度分析:</p> <p>① 证实了由“II”发现的五环系统(图 13).</p> <p>② 找出了五环系统的大部分侧链(图 14).</p>
1955 1956	<p>游离出琥珀酰亚胺</p> <p></p> <p>及</p> <p></p>	<p>I. 根据新的化学资料(见左), 并对比四种晶体(主要是六羧酸晶体及 B_{12} 干晶体)的分析结果最后获得整个分子的化学结构(VII)及分子构型(图 15—17).</p> <p>II. 经过逐步电子密度逼近、差值函数等方法修正后获得整个晶体结构全貌——包括结晶水位置、分子间连接情况, 甚至还找到部分氢原子的位置(图 18).</p> <p>III. 用 X 射线反常散射方法测定出分子的绝对构型.</p>

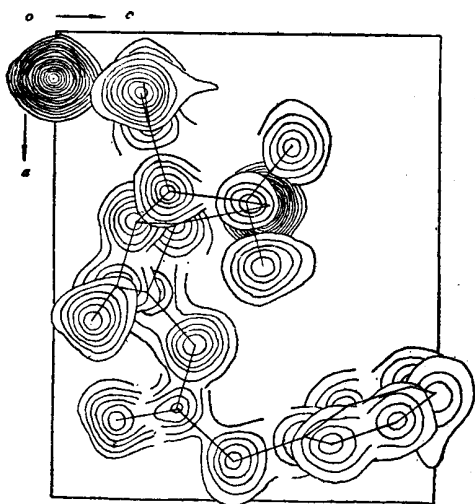


图 9

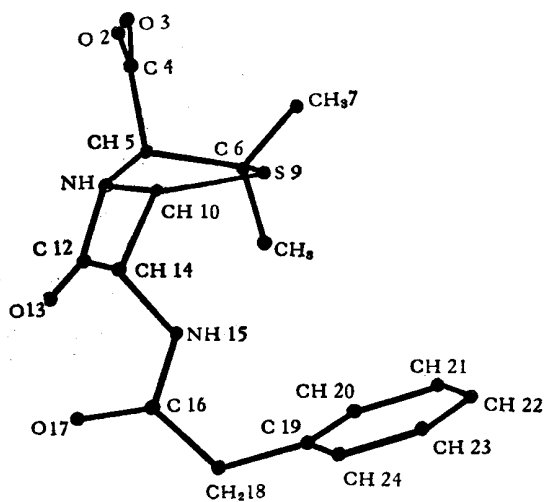


图 10

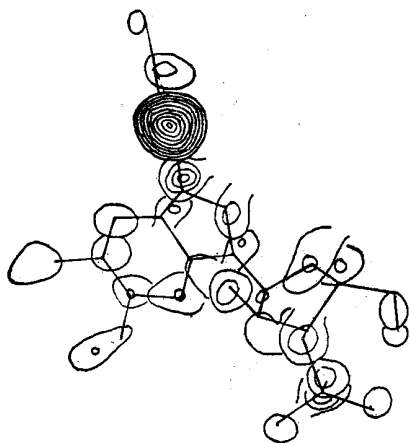


图 11



图 12

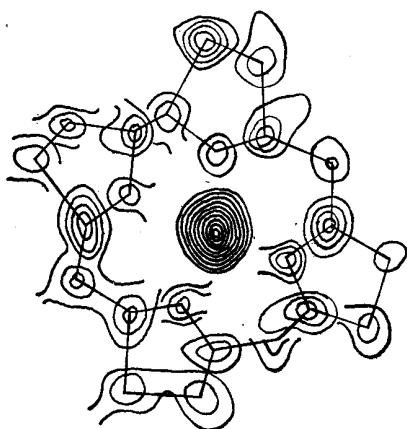


图 13

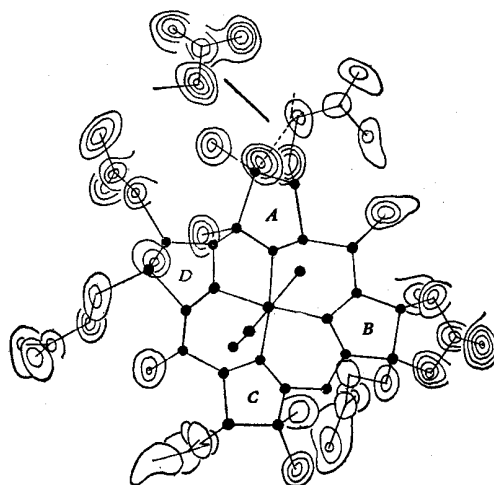


图 14

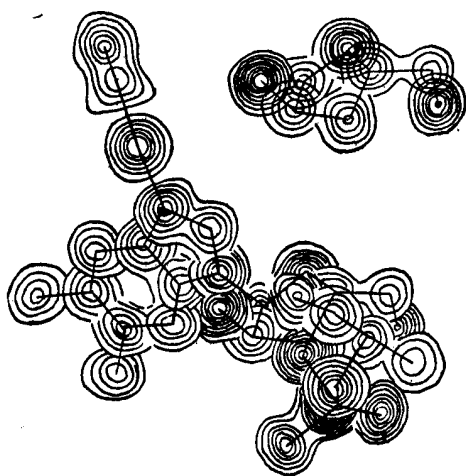


图 15

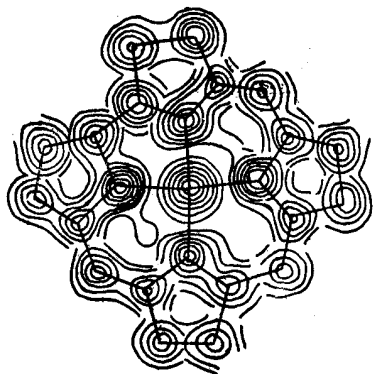


图 16

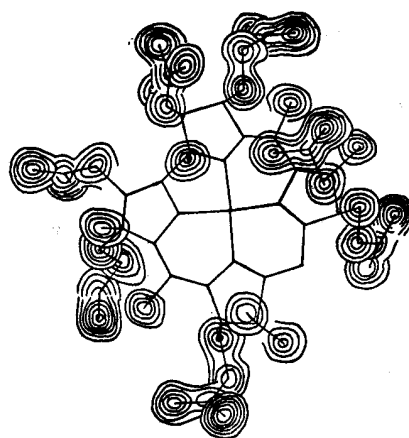


图 17

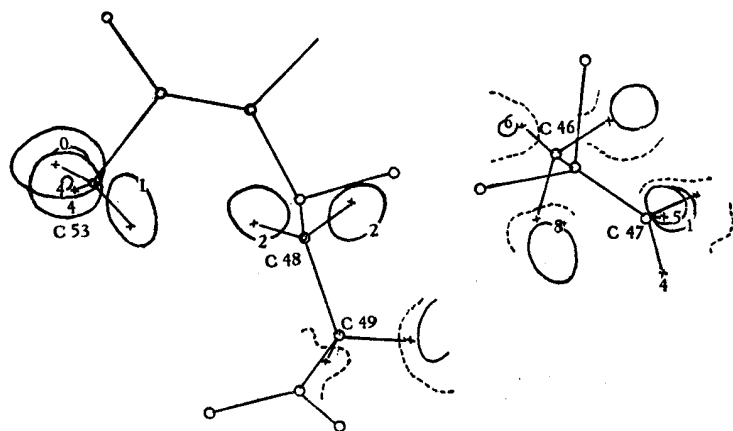
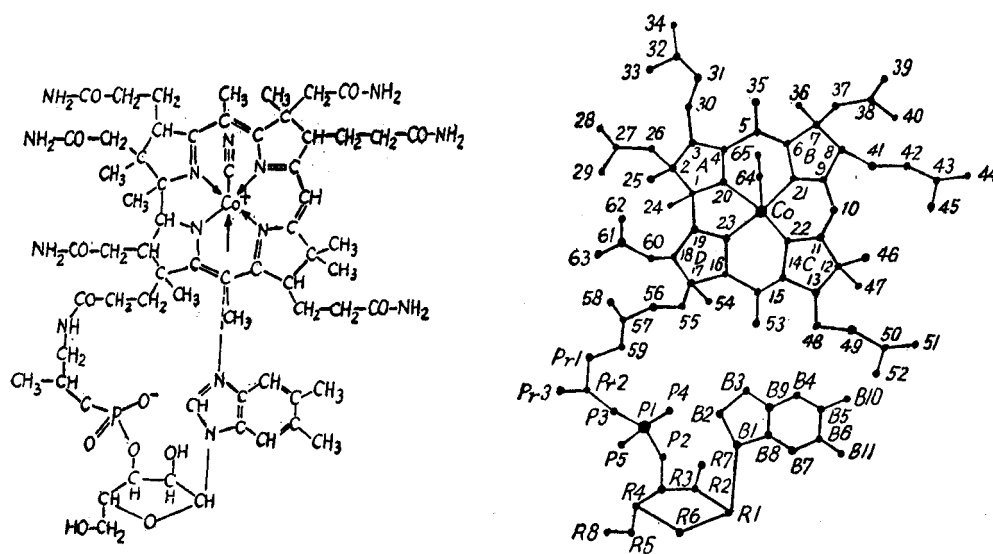


图 18

束,但是直到今天它仍然是单晶体X射线结构分析最高水平的标志之一。在这个工作中,就整体而言,化学方法与X射线方法紧密合作、互相验证、互相补充。而在过程中的某些关键性环节,则X射线又发挥了最大限度的作用。例如在解决 B_{12} 分子中含钴碎片的结构时,在几乎完全没有化学结构资料的情况下,X射线方法独立地找出了其中68个原子(不包括氢原子)的位置。这在今天仍然是仅有的例子。维生素 B_{12} 的分析方案与青霉素相仿,采用了几种晶体相互对比的办法,以保证结果的可靠和减少分析的困难。在寻找结构模型的过程中,重原子法的运用技巧达到了新的高峰,所积累的资料^[2,39]为有关重原子法的理论^[40]提供了一个宝贵的实验补充,对今后复杂的非中心对称晶体结构分析工作有重要的价值。这一晶体结构的彻底解决,除了在化学和医学上有一定意义外,还为复杂分子的结构化学积累了有价值的资料。

近四、五年来,由于理论和技术方面的进一步发展,特别是由于电子计算机的广泛应用,X射线方法在测定分子结构方面,应用的深度和广度都迅速地提高。最突出的成就是在晶态蛋白质分子的研究方面初步地解决了肌红和血红球蛋白晶体的结构。



(VII)

肌红球蛋白的分子量为17,000,由一条含153个氨基酸的肽链和一个血红素辅基组成。有关这个晶体结构的分析过程,已经发表过一系列的文章^[40]。它主要是依靠所谓“多对同晶型置换法”^[21]解决的。由于结构异常复杂,衍射数据的收集和处理都分阶段进行。起初只收集了相当于6 Å分辨率的数据,从其电子密度图获得了肽链的卷曲情况(图19);继之收集了2 Å分辨率的数据,其电子密度图已经能够显出血红素辅基以及其他一些原子基团的轮廓(图20 a, b);最近,由于利用了单晶体衍射的自动计数器记录技术^[41]和新型号的快速电子计算机(如IBM 7090),分辨率已提高

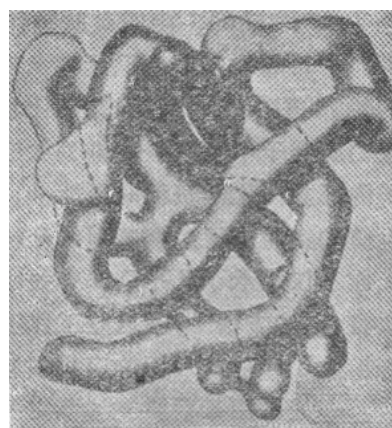


图 19

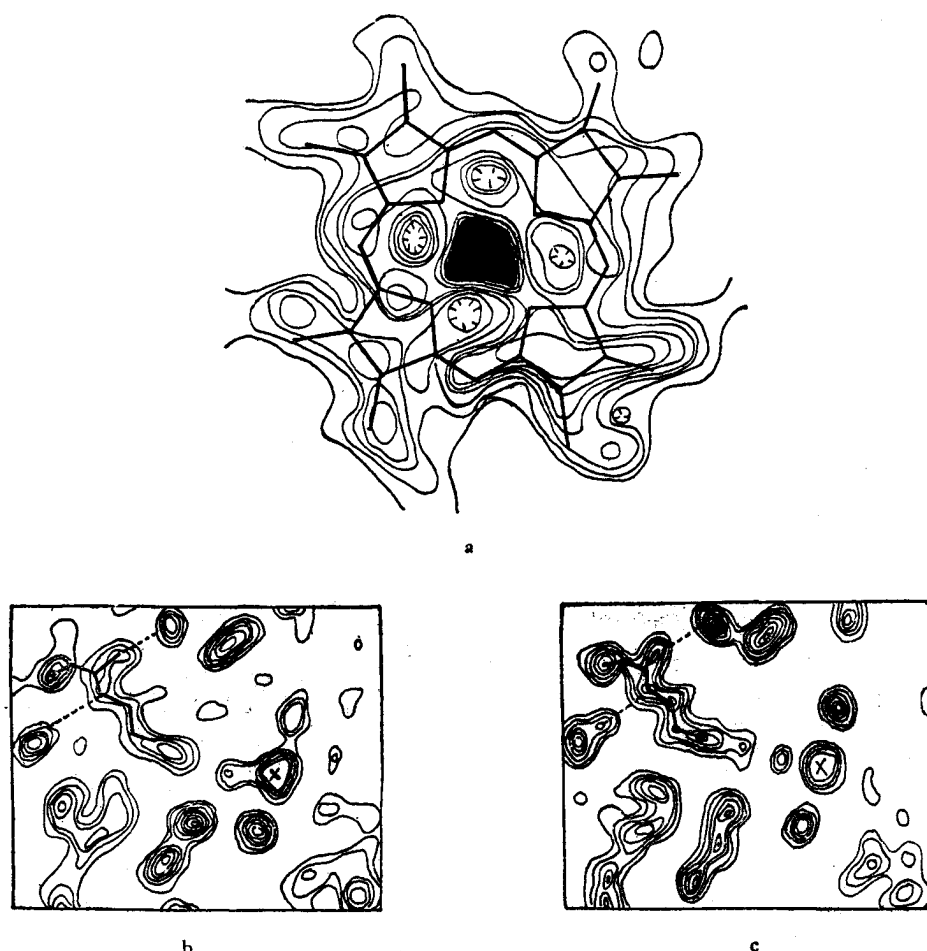


图 20

到 1.4 \AA ，其电子密度图适足以分辨相邻的成键原子（图 20c）。利用上述资料，结合化学线索，已经能够确切地推断出约 120 个氨基酸的位置，剩下约 30 个氨基酸的位置也大都有了可能的解释。预计不久的将来可望确切地求出绝大部分原子的位置。血红蛋白的分子量为 65000，由四条肽链和四个血红素辅基组成。其分析过程及方法与肌红蛋白相似^[42]。目前的分辨率只达到 5.5 \AA ，从分析结果可以看到其肽链的卷曲情况及血红素辅基的轮廓（图 21），并指出其氧化状态和还原状态的结构有显著的差别。这两项工作为测定蛋白质分子中氨基酸的排列次序开辟了新的途径，并且证实了从实验直接获得蛋白质分子内部立体化学资料的可能性，从而为在分子的水平上研究某些生物化学过程提供了有价值的实验方法。就 X 射线分析而言，肌红、血红蛋白以及其他一些晶态蛋白质分子的工作不仅继承了几乎全部有关小分子分析方法的成就，而且还发展出一套带有“晶态蛋白质特色”的方法^[43]，使这些工作不仅在对象上，而且也在方法上逐渐形成一个完整的体系，既扩大了 X 射线分析的应用领域，又丰富了 X 射线分析方法的内容。

在分子量较小的天然有机物领域，近年来进行了大量直接服务于化学研究的工作。

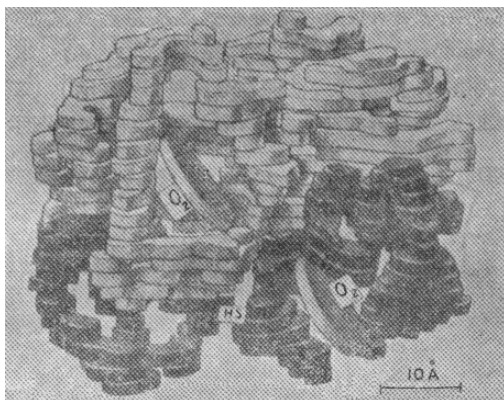
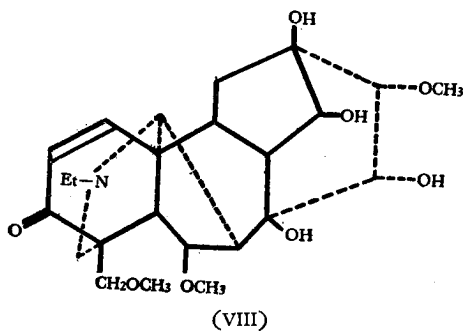


图 21

比较主要的研究对象有生物碱、萜类和甾族等。国外从事这方面工作的主要单位及研究者有英国 Glasgow 大学的 Robertson 和 Sim (Robertson 的主要合作者) 及 Robertson 领导下的实验室、Cardiff 大学 Wilson 的实验室、以及 Oxford 大学的 Hodgkin 及其实验室、澳大利亚的 Mathieson 和 Fridrichsons (Mathieson 的合作者)、加拿大的 Maria Przybylska、此外还有美国的 Pepinsky 以及 Lipscomb 等。他们所采用的分析方法大都是三维电子密度逐步逼近(重原子法)或三维 Patterson 函数分析,在修正结构参数时,普遍地采用了三维最小二乘法,计算工作的绝大部分都由电子数字计算机担任。在实验技术方面,Mathieson 还广泛地采用了低温(摄氏负一百多度)照相技术以提高分辨率和准确度。关于 Robertson 的实验室近年来(截至 1962 年底)在生物碱、萜类等方面的工作,Sim 及 Robertson 在 1963 年英国化学学会举办的年会*中已作报导。关于近年来 X 射线方法对上述各类天然有机物研究的概况可参看有关该类化合物研究进展的评论**。下面只从 X 射线分析的角度举一些难度较高的例子。它们大都是由于量少、化学行为反常、结构特殊或其他原因,不易用化学方法测定的。这些结构的解决在化学上都有一定的意义。

Demethanol aconinone (VIII) 的结构是分别用化学方法和 X 射线方法解决的^[41]。在



* Anniversary Meeting of the Chemical Society at Cardiff on March 28th, 1963. 并可参看 J. M. Robertson, *Proc. Chem. Soc.* 229 (1963).

** 例如关于生物碱方面可参看: 高怡生,“近年来生物碱化学的进展”(见本集)。

X 射线分析中所引用的化学资料仅有一个分子式以及关于 N 原子和 $-C_2H_5$ 相连的线索。

Aspidospermine 的结构, 从化学方法探知有四种可能 (IX), (X), (XI), (XII)。X 射线方法^[45]从其碘甲基取代物求得其结构为 (XIII)。其电子密度图示于图 22。

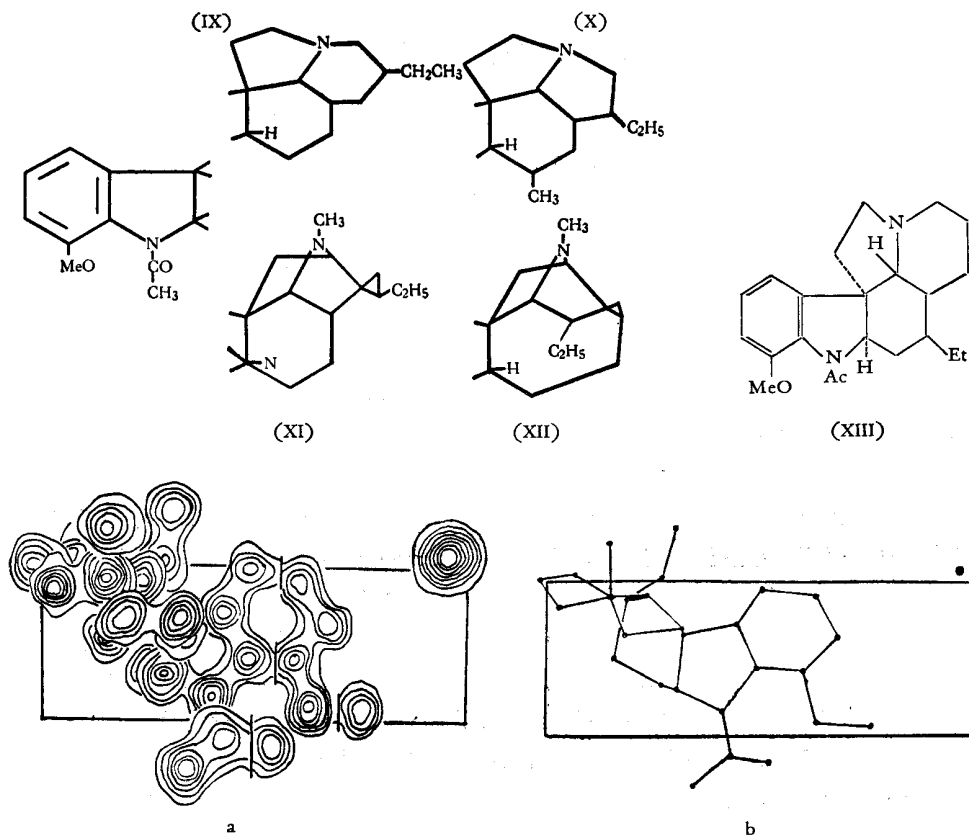
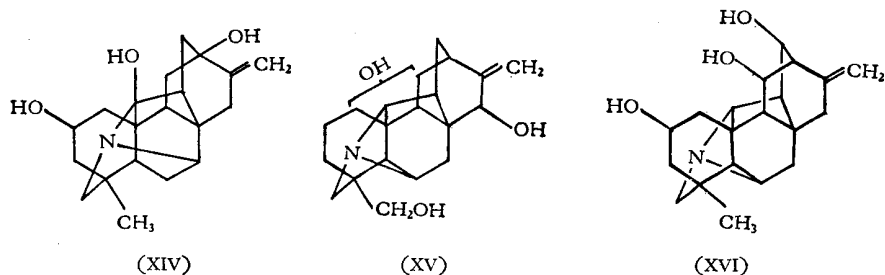
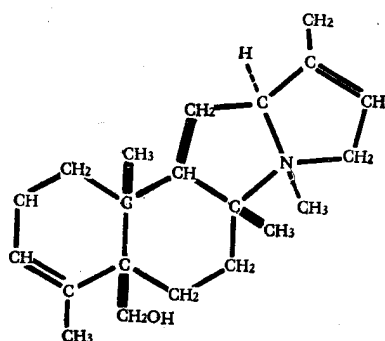


图 22

Hetisine 的结构, 由不同作者用化学方法得出两种结论 (XIV) 和 (XV), 但都不能完全解释该化合物的性质。X 射线方法最后获得正确的结构, 并测出其绝对构型如 (XVI)^[46]。

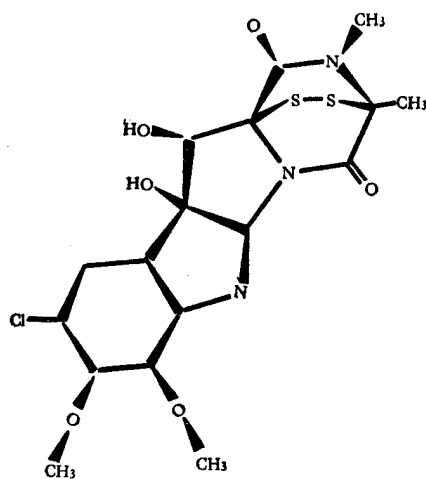


关于 thelepogine 的化学线索只有一个不很确切的分子式 $C_{20}H_{33}ON$ 。X 射线分析的结果^[47]将分子式改正为 $C_{20}H_{31}ON$, 并测得化学结构和绝对构型 (XVII)。



(XVII)

Sporidesimin 的化学资料也只有一个分子式 $C_{18}H_{20-22}O_6N_3S_2Cl$ ，因有重排及反常化学反应，不易用化学方法测定。X射线方法结合核磁共振研究测得其相对构型如 (XVIII)^[48]。



(XVIII)

Cedrelone 已知分子式及部分结构，用 X 射线方法^[50]通过其碘乙酸盐求得化学结构及相对构型为 (XIX)。电子密度图见图 23

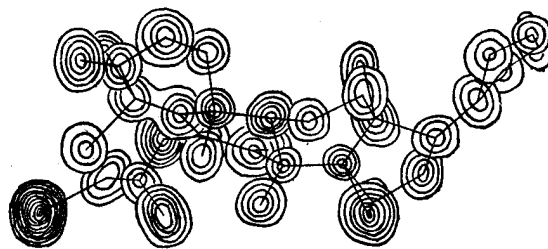
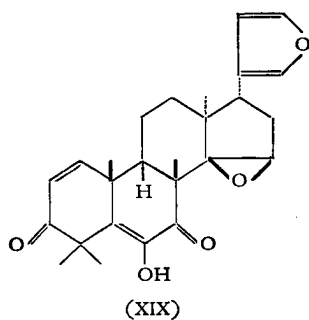
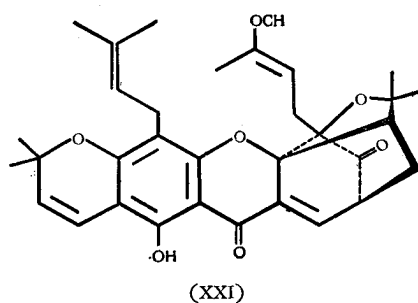
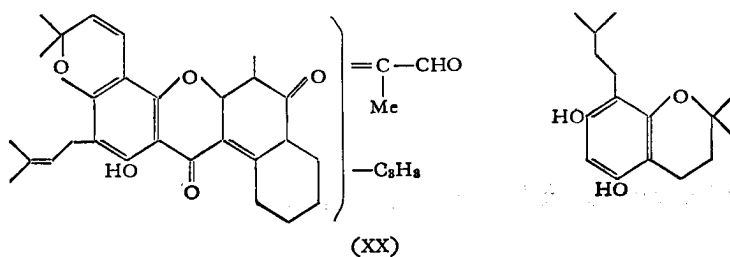


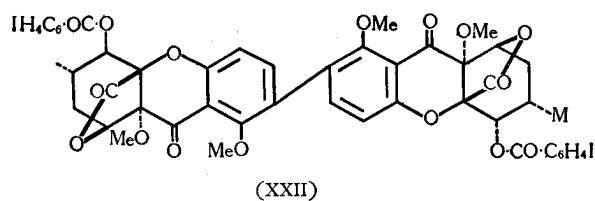
图 23



Morellin 的部分结构由化学方法探知如 (XX), 经 X 射线方法确定为 (XXI)^[51].



Ergoflavin 的分子式已知为 $C_{30}H_{26}O_{14}$, 用化学方法探知分子中有两个酚的羟基、两个醇的羟基、两个羰基和两个 γ -内酯环。其结构用 X 射线方法从其碘代苯甲酸盐测得, 并利用 X 射线反常散射法确定其绝对构型如 (XXII)^[52]。相应的电子密度图见图 24。



(二) 通过简单物质的精密测定以获得有关大分子结构的立体化学资料及经验规律
天然产物中的高分子化合物, 如橡胶、纤维素、特别是在生命问题上有重大意义的蛋白质和核酸等, 它们在一定程度上是由简单而相仿的单体聚合而成的长链分子, 其结构既复杂而结晶性能又差。除球蛋白能长成良好的单晶体外, 一般高分子物质在天然状态或经过适当处理后, 只能获得纤维状 (其内部结构仅在伸长方向上有某种程度的周期性)

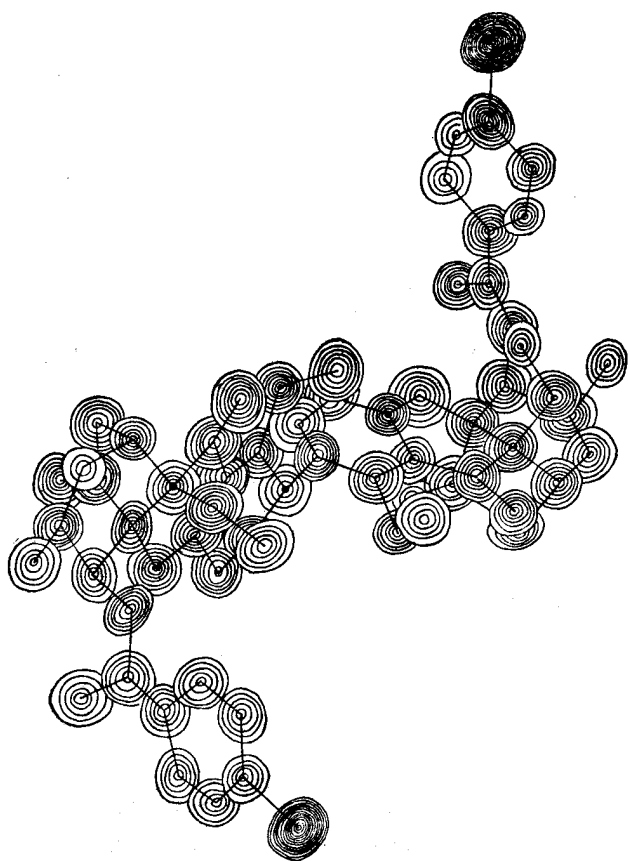


图 24

或粉末状的样品,因此衍射效应并不明显。要解决这一类化合物的结构问题,除了大分子自身的衍射数据外,还要利用其他物理、化学、生物学等方面的资料。其中,由单晶体X射线结构分析获得的有关大分子碎片(单体、二聚体及相应的简单物质)的结构资料尤为重要。积累有关长链分子中单体及其相互连接方式的结构参数,并探讨分子间相互作用的因素,可有助于推引或修正大分子的结构模型。

这一类型的工作,其对象的化学结构多半是已知的,分析的目的主要是获得原子精确的空间分布情况。为了保证高准确度的要求,通常都尽可能不用含重原子的晶体。因此三维 Patterson 函数成了解决这类结构的最主要工具。另外,由于在这类工作中,已知的化学资料较多,尝试法往往也能发挥重要的作用。为了获得高度准确的结果,一般的修正技术已嫌不足,目前普遍用带有各向异性温度校正的最小二乘法 and 差值级数法相结合。这些计算工作都必须在电子数字计算机中进行。随着电子计算技术的发展,近年来已逐渐采用“全矩陣式”的最小二乘电子计算程序以代替较老的“对角线式”程序,这就更进一步提高了结果的准确度。目前,在许多准确的分析工作中,X射线方法已经可以直接获得氢原子的确切位置(参看图 25)。

在有关蛋白质及核酸的简单物质中,系统地进行过精密X射线分析的有:氨基酸、肽、酰胺、嘧啶碱、嘌呤碱、核甙及核甙酸等。由它们所积累的资料,对了解蛋白质和核酸

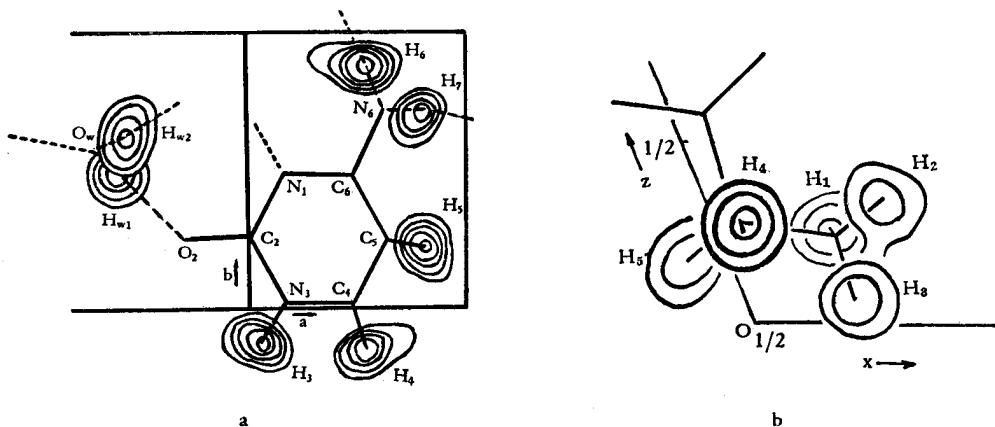


图 25

的结构起着重要的作用。

对于有关核酸的简单物质，单晶体 X 射线结构分析所获得的早期重要结果有 S. Furberg 1950 关于核甙的分析^[53]，他指出碱基并不与核糖呋喃环平行，而是近似地相互垂直，从而证明了以往 Astbury 关于核酸结构模型的假定是不正确的。

Watson-Crick 的双螺旋脱氧核糖核酸结构模型是根据其钠盐纤维的衍射数据及立体化学资料推引出来的^[54]。为了检验和修正核酸的结构模型，除了必须进一步改良样品的结晶性能，以获得更大量的衍射数据外，系统地、精密地分析各种核酸“碎片”的结构很有必要。从这些分析，可以获得核酸中碱基的互变异构现象、碱基配对的方式、核糖呋喃环的构型、以及磷酸、核糖与碱基三者的相互连接方式等重要资料。最近几年，这方面工作进展相当迅速，它们都达到很高的准确度。键长的最大可能误差（按标准误差的三倍计算）低至 0.01—0.02 Å 左右。有关这方面的几个重要工作列于表 2。通过这些分析，已经获得了一些有意义的资料。例如，嘌呤与嘧啶碱呈平面构型，其互变异构式以酮式及氨基式较稳定；Watson-Crick 模型中的鸟嘌呤与胞嘧啶碱基对应该形成三个而不是两个氢键；核糖呋喃环上 C₂' 或 C₃' 应偏离其余原子所组成的平面约 0.5 Å；碱基与戊糖环的相对取向是反型的^[60,62]，等等。这些资料已用于脱氧核糖核酸模型的修正中^[63]。

表 2

化合物名称	使用数据	准确度	文献
胸腺嘧啶水合物 (Thymine monohydrate)	三维衍射强度	0.01 Å	[55]
胞嘧啶-5-醋酸 (Cytosine-5-acetic acid)	三维衍射强度	0.018 Å	[56]
胞嘧啶水合物 (Cytosine monohydrate)	三维衍射强度	0.012 Å	[57]
1-甲基胸腺嘧啶 (1-Methylthymine)	三维衍射强度	0.01 Å	[58]
腺嘌呤盐酸盐 (Adenine hydrochloride)	三维衍射强度	0.03 Å	[59]
腺嘌呤核甙-5'-磷酸 (Adenosine-5'-phosphate)	三维衍射强度	0.04 Å	[60]
胸腺嘧啶核甙磷酸钙盐 (Calcium thymidylate)	三维衍射强度	0.07 Å	[61]

最近有两个直接显示碱基配对的可能方式的结构分析工作颇有意义。一个是 9-乙基鸟嘌呤和 1-甲基胞嘧啶的 1:1 络合物^[61]。分析结果证明这一对碱基正是以 Watson-Crick 模型中所假定的方式结合起来(图 26 a, b)。另一个是 1-甲基胸腺嘧啶和 9-甲基腺

嘌呤的 1:1 络合物^[65]。这一对的结合方式却与 Watson-Crick 模型不一致(图 26 c, d)。

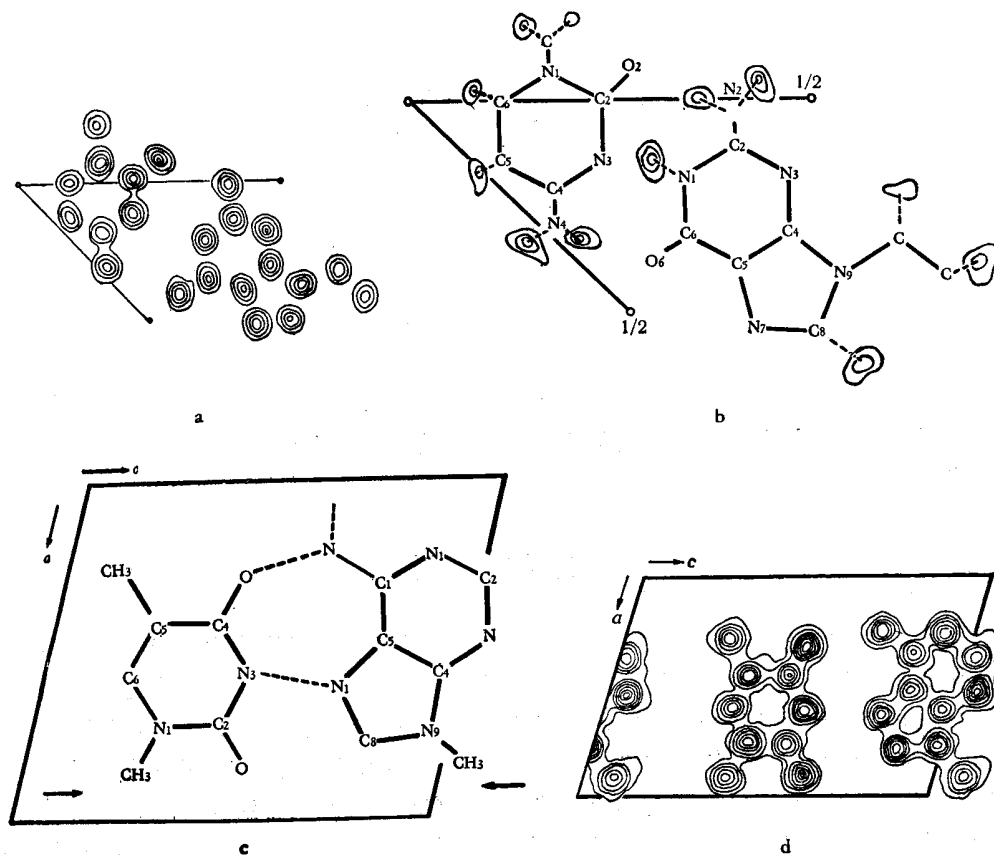


图 26

在解决多肽链构型问题,亦即蛋白质的二级结构问题中,相应小分子的结构分析工作所起的作用更为突出。在这方面,美国麻省理工大学 Pauling 领导下的实验室作出了特别重要的贡献。到目前,近二十种主要的氨基酸和十余种简单肽已被分析^[66]。其中有半数以上经用三维强度数据修正过,所测得原子坐标的最大可能误差(以标准误差的三倍计算)一般都在 0.02—0.03 Å 左右。从 1940 年开始,随着实验资料越来越丰富和精确,为多肽链提出的标准键长、键角数据不断得到修正^[67]。1953 年 Pauling 与 Corey 所提出的标准键长、键角数据已分别精确到 0.03 Å 和 4°(图 27)。

基于上述实验工作, Pauling 等总结出一系列多肽链构型的基本原则^[68],其中主要的有两个:(1)肽键键长平均为 1.32 Å,具有一定双键成分,因此整个酰胺基团应取平面构型。此外,酰胺基团上亚胺基与羰基应为反型分布(图 27);(2)酰胺基团之间应形成尽可能多的 N—H···O=C 氢键,此氢键的长度与方向有一定的限制。据此, Pauling 等于 1951 年推引出符合上述结构原则及精确键长、键角数据的 α-螺旋构型和折迭层构型^[69],许多纤维蛋白的衍射分析及球蛋白的单晶衍射分析都证明了这些构象的存在。最近完成的甘氨酸-苯丙氨酸-甘氨酸三肽的分析^[70],已直接证明了折迭层构象确是能量上稳定的构型。

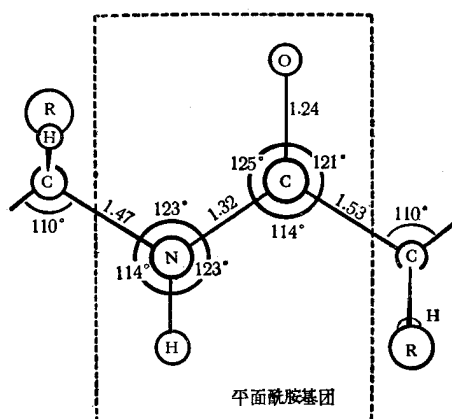


图 27

可以预期,更大量、更深入地分析与大分子有关的简单物质的晶体结构,将会为大分子本身的结构提供更多、更有意义的资料。

四、结 语

随着天然有机物结构研究工作的进一步发展,单晶体X射线结构分析将会获得日益广泛和深入的应用。相应地,对这一工具也将提出越来越高的要求。“高速度”、“高难度”和“高精度”将是单晶体X射线结构分析的主要奋斗目标,而其中以速度问题最为重要。因为这是单晶体X射线结构分析能否被广泛采用的关键,而且提高分析对象的难度和分析结果的精确度也必须以高速度作为基础才会具有现实的意义。速度问题的最彻底解决办法是使整个过程“全盘自动化”。它包括实验数据的自动记录和自动分析,前一问题目前已经解决^[4],剩下的只是分析过程自动化问题。但这是一个复杂和困难得多的问题。解决这个问题的出路有三:(1)用电子计算机来执行现有的分析过程,包括自动利用晶体化学知识以及物理、化学线索,自动运用各种解决相角问题的方法,自动分析并修正各级近似的电子密度图,等等;(2)改造现有的分析过程,使它进一步系统化和数学化,然后由电子计算机来执行;(3)创造新的易于由电子计算机自动执行的方法。这三方面是相互联系、相互补充的。目前在第一、二两方面条件较成熟。尤其是有关重原子法的自动化^[71]、Patterson图的自动分析^[72]、晶体化学知识的自动利用^[73]等尝试,都是很有希望和有实用意义的工作。要彻底解决自动化的问题需要晶体学家、数学家以至化学家的密切配合、共同努力。虽然目前离“全盘自动化”的目标还有很远的距离,但是就现有的技术水平以及今后的发展,可以认为:要做到对于中等复杂程度的晶体,在调整好样品和仪器以后若干小时内就可以从电子计算机的输出中读得有关晶体及分子结构的各种数据,这并不纯粹是一个幻想。另一方面,为了解决物质的结构,单晶体X射线结构分析方法现在不是,将来也不会是唯一的和孤立的方法。它的发展过程并不是排斥其他方法的过程,而是最大限度地联合其他方法的过程。特别是在进一步探索更复杂的大分子结构方面,X射线方法更需要和其他物理方法、化学方法、以至生物学方法密切地结合。这将为各方面的工作者提供广阔的合作园地。

[致謝] 在本文写作过程中,蒙吳乾章先生、黃芬、黃文英等同志提出宝贵的意见和建议,特别是吳乾章先生给予多方关怀和鼓励,作者特表深切的感谢。

参 考 文 献

- [1] W. H. Bragg, *Proc. Phys. Soc.*, **34**, 33 (1921); **35**, 167 (1922); *Z. Krist.* **66**, 22 (1928).
- [2] D. C. Hodgkin, J. Kamper, J. Lindsey, M. MacKay, J. Pickworth, J. H. Robertson, C. B. Shoemaker, J. G. White, R. J. Prosen, K. N. Trueblood, *Proc. Roy. Soc.*, **A242**, 228 (1957); **A266**, 494 (1962).
- [3] J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. C. Phillips, V. C. Shore, *Nature*, **185**, 422 (1960); M. F. Perutz, M. G. Rossmann, A. F. Cullis, H. Muirhead, G. Will, A. C. T North, 同上, **185**, 416 (1960).
- [4] R. G. Dickinson, A. L. Raymond, *J. Amer. Chem. Soc.*, **45**, 22 (1923); G. W. Gonell, H. Mark *Z. physik. Chem.*, **107**, 181 (1923).
- [5] S. B. Hendricks, C. Bilicke, *J. Amer. Chem. Soc.*, **48**, 3007 (1926); R. G. Dickinson, C. Bilicke, 同上, **50**, 764 (1928); K. Lonsdale, *Nature*, **122**, 810 (1928); *Proc. Roy. Soc.*, **A123**, 494 (1929); *Trans. Faraday Soc.*, **25**, 352 (1929).
- [6] K. Becker, W. Janke, *Z. physik. Chem.*, **99**, 242, 267 (1921); H. Mark, K. Weissenberg, *Z. Physik*, **16**, 1 (1923); S. B. Hendricks, *J. Amer. Chem. Soc.*, **50**, 2455 (1928); R. W. G. Wyckoff, *Z. Krist.*, **75**, 529 (1930); **81**, 102 (1932); R. W. G. Wyckoff, R. B. Corey, 同上, **89**, 462 (1934).
- [7] G. Shearer, *Proc. Roy. Soc.*, **A108**, 655 (1925); R. Robinson, *Nature*, **116**, 45 (1925).
- [8] W. L. Bragg, *Z. Krist.*, **70**, 475 (1929); W. L. Bragg, H. Lipson, 同上, **95**, 323 (1936); W. L. Bragg, *Nature*, **154**, 69 (1944); A. W. Hanson, H. Lipson, C. A. Taylor, *Proc. Roy. Soc.*, **A218**, 371 (1953); H. Lipson, C. A. Taylor, "Fourier Transforms and X-Ray Diffraction" G. Bell and Sons, London 1958; C. W. Bunn, "Chemical Crystallography" Oxford Univ. Press 1961, pp. 270—300.
- [9] A. L. Patterson, *Phys. Rev.*, **46**, 372 (1934).
- [10] A. L. Patterson, *Z. Krist.*, **90**, 517 (1935).
- [11] D. Harker, *J. Chem. Phys.*, **4**, 381 (1936); M. J. Buerger, *J. Appl. Phys.*, **17**, 579 (1946); I. Langmair, D. Wrinch, *Nature*, **142**, 581 (1938); D. M. Wrinch, 同上, **142**, 955 (1938); *Phil. Mag.*, **27**, 98, 490 (1939); **39**, 692 (1948).
- [12] M. J. Buerger, *Acta Cryst.*, **3**, 87,243 (1950); **4**, 531 (1951); *Proc. Nat. Acad. Sci.*, **36**, 376 (1950); **39**, 669, 674, 678 (1953).
- [13] J. Clastre, R. Gay, *Compt. Rend.*, **230**, 1876 (1950); J. Garrido, 同上, **230**, 1878 (1950); C. A. Beevers, J. H. Robertson, *Acta Cryst.*, **3**, 164 (1950); J. H. Robertson, 同上, **4**, 63 (1951); J. Donohue, K. N. Trueblood, 同上, **5**, 414 (1952); D. McLachlan, Jr., Bull. No. 50, Utah Eng. Exp. Sta. **41**, No. 6, 40 (1950); *Proc. Nat. Acad. Sci.*, **37**, 115 (1951).
- [14] D. McLachlan, Jr., *Z. Krist.*, **115**, 305 (1961); A. D. Mighell, R. A. Jacobson, *Acta Cryst.*, **16**, 443 (1963).
- [15] M. J. Buerger, "Vector Space and its application in crystal structure investigation" John Wiley and Sons, New York 1959; В. И. СИМОНОВ, *Кристаллография*, **4**, 302 (1959); J. Fridrichsons, A. McL. Mathieson, *Acta Cryst.*, **15**, 1065 (1962).
- [16] H. Lipson, W. Cochran, "The Determination of Crystal Structures" Chapter 7, 207, G. Bell and Sons, London 1953.
- [17] W. L. Bragg, *Proc. Roy. Soc.*, **A123**, 537 (1929).
- [18] V. Luzzati, *Acta Cryst.*, **6**, 142 (1953); M. M. Woolfson, 同上, **9**, 804 (1956); G. A. Sim, 同上, **10**, 177, 536 (1957); **12**, 813 (1959).
- [19] A. McL. Mathieson, *Rev. Pure Appl. Chem.*, **5**, 113 (1955); J. M. Robertson, *Proc. Chem. Soc.*, 229 (1963).
- [20] J. M. Cork, *Phil. Mag.*, **4**, 688 (1927). J. M. Robertson, *J. Chem. Soc.*, 615 (1935); 1195 (1936); J. M. Robertson, J. Woodward, 同上, 219 (1937); 36 (1940); J. M. Bijvoet, E. H. Wiebenga, *Naturwiss.*, **32**, 45 (1944); J. M. Bijvoet, *Proc. Koninkl. Ned. Abad. Wetenschap.*, (B) **52**, 313 (1949); D. Harker, *Acta Cryst.*, **9**, 1 (1956); M. F. Perutz, 同上, **9**, 867 (1956); F. H. C. Crick, B. S. Magdoff, 同上, **9**, 901 (1956); A. Hargreaves, 同上, **10**, 196 (1957); G. Kartha, 同上, **14**, 680 (1961).
- [21] D. W. Green, V. M. Ingram, M. F. Perutz, *Proc. Roy. Soc.*, **A225**, 287 (1954); J. Bodo, H. M. Dintzis, J. C. Kendrew, H. W. Wyckoff, 同上, **A253**, 70 (1959); R. E. Dickerson, J. C. Kendrew,

- B. E. Strandberg, *Acta Cryst.*, **14**, 1188 (1961); R. E. Dickerson, J. C. Kendrew, B. E. Strandberg, in "Computing Methods and the Phase Problem in X-Ray Crystal Analysis" ed. by R. Pepinsky et al., Pergamon Press, Oxford, 236, 1961.
- [22] R. W. James, "The Optical Principles of the Diffraction of X-Rays" G. Bell and Sons, London 149, 1950.
- [23] "The International Tables for X-Ray Crystallography" Published for the International Union of Crystallography, The Kynoch Press, Birmingham (1962). 卷III, Tables 3.3.2 A—C.
- [24] M. J. Buerger, "Crystal-structure Analysis", John Wiley and Sons, Chapter 20, New York, 1960; 仁田 勇, "X线结晶学", 丸善株式会社, 1960. (下) IV, 第7章; Y. Okaya and R. Pepinsky in "Computing Methods and the Phase Problem in X-Ray Crystal Analysis" ed. by R. Pepinsky et al., Pergamon Press, Oxford 273, 1961.
- [25] K. Banerjee, *Proc. Roy. Soc.*, **A141**, 188 (1933).
- [26] D. Harker, J. S. Kasper, *J. Chem. Phys.*, **15**, 882 (1947); *Acta Cryst.*, **1**, 70 (1948).
- [27] A. И. Китайгородский, "Теория структурного анализа", Глава IV, Изд. АН СССР, Москва, 1957; W. Cochran, *Acta Cryst.*, **11**, 579 (1958); M. J. Buerger, "Crystal-structure Analysis", Chapter 21, John Wiley and Sons, New York, 1960; M. M. Woolfson, "Direct Methods in Crystallography", Oxford Univ. Press 1961; A. И. Китайгородский, Б. К. Вайнштейн, И. М. Руманова, З. В. Звонкова, *Ж. структурн. химии*, **2**, 622 (1961).
- [28] A. D. Booth, *Nature*, **161**, 765 (1948); W. Cochran, *Acta Cryst.*, **4**, 408 (1951).
- [29] E. W. Hughes, *J. Amer. Chem. Soc.*, **63**, 1737 (1941).
- [30] V. Luzzati, *Acta Cryst.*, **5**, 802 (1952).
- [31] J. C. Paul, G. A. Sim, T. A. Hamor, J. Monteath Robertson, *J. Chem. Soc.*, 4133 (1962).
- [32] F. H. C. Crick in "Methods in Enzymology", 卷IV, 127, Acad. Press, New York, 1957.
- [33] D. Crowfoot, *Proc. Roy. Soc.*, **A164**, 580 (1938); D. Crowfoot D. Riley, *Nature*, **144**, 1011 (1939).
- [34] B. W. Low, J. R. Einstein, *Nature*, **186**, 470 (1960).
- [35] J. D. Bernal, 同上, **129**, 277 (1932).
- [36] O. Rosenheim, H. King, *Chem. and Ind.*, (London) **10**, 464, 954 (1932); *Nature*, **130**, 315 (1932).
- [37] D. Crowfoot, C. W. Bunn, B. W. Roger-Low, A. Turner-Jones in "The Chemistry of Penicilline" Chapter 11, Princeton Univ. Press, 1949.
- [38] L. J. Comrie, G. B. Hey, H. G. Hudson, J. R. Statist. Soc. *Suppl.*, **4**, 210 (1937); P. A. Shaffer, V. Schomaker L. Pauling, *J. Chem. Phys.* **14**, 648, 659 (1946); E. G. Cox, G. A. Jeffrey, *Acta Cryst.*, **2**, 341 (1949); J. Donohue, V. Schomaker, 同上, **2**, 344 (1949); M. D. Grens, J. S. Kasper, 同上, **2**, 347 (1949); E. G. Cox, L. Gross, G. A. Jeffrey, 同上, **2**, 351 (1949).
- [39] D. C. Hodgkin, J. Pickworth, J. H. Robertson, R. J. Prosen, R. A. Sparks, K. N. Trueblood, *Proc. Roy. Soc.*, **A251**, 306 (1959); J. G. White, 同上, **A266**, 440 (1962); D. C. Hodgkin F. R. S., J. Lindsey, R. A. Sparks, 同上, **A266**, 475, 494 (1962).
- [40] J. C. Kendrew, *Acta Cryst.*, **1**, 336 (1948); *Proc. Roy. Soc.* **A201**, 62 (1950); J. C. Kendrew, P. J. Pauling 同上, **A237**, 255 (1956); J. C. Kendrew R. G. Parrish, **A238**, 305 (1956); J. C. Kendrew, G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, D. C. Phillips, *Nature*, **181**, 662 (1958); M. M. Bluhm, G. Bodo, H. M. Dintzis J. C. Kendrew, *Proc. Roy. Soc.*, **A246**, 369 (1958); G. Bodo, H. M. Dintzis, J. C. Kendrew, H. W. Wyckoff, 同上, **A253**, 70 (1959); J. C. Kendrew, R. E. Dickerson, B. E. Strandberg, R. G. Hart, D. C. Phillips, V. C. Shore, *Nature* **185**, 422 (1960); R. E. Dickerson et al., *Acta Cryst.*, **14**, 1188 (1961); J. C. Kendrew, H. C. Watson, B. E. Strandberg, R. E. Dickerson, D. C. Phillips, V. C. Shore, *Nature*, **190**, 666 (1961); J. C. Kendrew, *Science*, **139**, 1259 (1963).
- [41] U. W. Arndt, D. C. Phillipst, W. A. Wooster, A. M. Wooster, *J. Sci. Instrum.*, **39**, 103 (1962); W. Arrington et al., *Physik. Verhandl.*, **13**, 127 (1962); W. Wooster, 同上, **13**, 127 (1962); J. P. Cowan et al., *Acta Cryst.*, **16**, 221 (1963); B. B. Cetlin et al., 同上, **16**, 943 (1963); W. A. Wooster, *J. Sci. Instrum.*, **40**, 14 (1963); M. H. Mueller, L. Heaton, S. S. Sidhu, *Rev. Sci. Instrum.*, **34**, 74 (1963).
- [42] J. Boyes-Watson et al., *Proc. Roy. Soc.*, **A191**, 83 (1947); M. F. Perutz, 同上, **A195**, 474 (1949); L. Bragg, M. F. Perutz, 同上, **A213**, 425 (1952); L. Bragg, F. R. S., E. R. Howells, M. F. Perutz, **A222**, 33 (1954); M. F. Perutz, 同上, **A225**, 264 (1954); D. W. Green, V. M. Ingram, M. F. Perutz, 同上, **A225**, 287 (1954); E. R. Howells, M. F. Perutz, 同上, **A225**, 308 (1954); L. Bragg, M. F. Perutz, 同上, **A225**, 315 (1954); D. M. Blow, 同上, **A247**, 302 (1958); A. F. Cullis et al., *Acta Cryst.*, **13**, 1053 (1960); M. F. Perutz, M. G. Rossmann, A. F. Cullis, H. Muirhead,

- G. Will, A. C. T. North, *Nature*, **185**, 416 (1960); A. F. Cullis et al., *Proc. Roy. Soc.*, **A265**, 15 (1961); **A265**, 161 (1962); M. F. Perutz, *Nature*, **194**, 914 (1962); H. Munhead, M. F. Perutz, 同上, **199**, 633 (1963); M. F. Perutz, *Science*, **140**, 863 (1963).
- [43] (a) J. C. Kendrew, *Progress in Biophys.*, **4**, 244 (1954); R. E. Dickerson, J. C. Kendrew, B. E. Strandberg, in "Computing Methods and the Phase Problem in X-Ray Crystal Analysis", 236, ed. by R. Pepinsky et al., Pergamon Press, Oxford, 1961; L. K. Steinrauf, *Acta Cryst.*, **16**, 317 (1963).
- [44] K. Wiesner, M. Götz, D. L. Simmons, L. R. Fowler, F. W. Boachelor, R. F. C. Brown, G. Büchi, *Tetrahedron Letters*, No. 2, **15** (1959); M. Przybylska, *Can. J. Chem.*, **1116**, 1843 (1959); *Acta Cryst.*, **14**, 428 (1961).
- [45] J. F. D. Mills, S. C. Nyburg, *J. Chem. Soc.*, 1458 (1960).
- [46] M. Przybylska, *Can. J. Chem.*, **40**, 566 (1962); *Acta Cryst.*, **16**, 871 (1963).
- [47] J. Fridrichsons A. McL. Mathieson, *Tetrahedron Letters*, No. 26, 18 (1960).
- [48] J. Fridrichsons A. McL. Mathieson, 同上, 1265 (1962).
- [50] I. J. Grant, J. A. Hamilton, T. A. Hamor, J. Monteath Robertson, G. A. Sim, *J. Chem. Soc.*, 2506 (1963).
- [51] G. Kartha, G. N. Ramachandran, H. B. Bhat, P. Madhavan Nair, V. K. V. Raghavan and K. Venkataraman, *Tetrahedron Letters*, No. 7, 459 (1963).
- [52] J. D. M. Asher, A. T. McPhail, J. Mantheath Robertson, J. V. Silverton, G. A. Sim, *Proc. Chem. Soc.*, 210 (1963).
- [53] S. Furberg, *Acta Cryst.*, **3**, 325 (1950).
- [54] F. H. C. Crick, J. D. Watson, *Proc. Roy. Soc.*, **A223**, 80 (1954).
- [55] R. Gerdil, *Acta Cryst.*, **14**, 333 (1961).
- [56] R. E. Marsh, R. Bierstedt, E. L. Eichhorn, 同上, **15**, 310 (1962).
- [57] G. A. Jeffrey, Y. Kinoshita, 同上, **16**, 20 (1963).
- [58] K. Hoogsteen, 同上, **16**, 28 (1963).
- [59] W. Cochran, 同上, **4**, 81 (1951).
- [60] J. Kraut, L. H. Jensen, 同上, **16**, 79 (1963).
- [61] K. N. Trueblood, P. Horn, V. Luzzati, 同上, **14**, 965 (1961).
- [62] L. Pauling, R. B. Corey, *Arch Biochem. Biophys.*, **65**, 164 (1956); M. Spencer, *Acta Cryst.* **12**, 59, 66 (1959).
- [63] R. Langridge et al. *J. Biophys. Biochem. Cytol.*, **3**, 767 (1957); *J. Mol. Biol.*, **2**, 38 (1960).
- [64] E. J. O'Brien, *J. Mol. Biol.*, **7**, 107 (1963).
- [65] K. Hoogsteen, *Acta Cryst.*, **12**, 822 (1959); **16**, 907 (1963).
- [66] A. Rich, D. W. Green, *Ann. Rev. Biochem.*, **30**, 93 (1961); J. C. Kendrew, M. F. Perutz, *Ann. Rev. Biochem.*, **26**, 327 (1957).
- [67] R. B. Corey, *Chem. Rev.*, **26**, 227 (1940); "Advances in Protein Chemistry", **4**, 385 (1948); R. B. Corey, J. Donohue, *J. Amer. Chem. Soc.*, **72**, 2899 (1950); R. B. Corey, L. Pauling, *Proc. Roy. Soc.*, **B141**, 10 (1953); T. Hahn, *Z. Krist.*, **109**, 438 (1957).
- [68] L. Pauling, R. B. Corey, *Fortschr. Chem. Org. Naturstoff*, **11**, 180 (1954); L. Pauling in "Symposium on protein structure" Wein-Springer-Verlag, 17, 1957.
- [69] L. Pauling, R. B. Corey, H. R. Branson, *Proc. Nat. Acad. Sci.*, **37**, 205 (1951); L. Pauling R. B. Corey, 同上, **37**, 729 (1951); **39**, 253 (1953).
- [70] R. E. Marsh, J. P. Glusker, *Acta Cryst.*, **14**, 1110 (1961).
- [71] L. I. Hodgson, J. S. Rollett, *Physik. Verhandl.*, **13**, 117 (1962).
- [72] R. A. Jacobson et al. in "Computing Methods and the Phase Problem in X-Ray Crystal Analysis", 125 ed, by R. Pepinsky et al., Pergamon Press, Oxford, 1961.
- [73] H. J. Milledge, *Proc. Roy. Soc.*, **267**, 566 (1962).