

直接法中的尝试法(I)

郑 启 泰

(中国科学院生物物理研究所)

1984年6月29日收到

提 要

应用直接法测定晶体结构,在归一化结构因子的计算、相位测定以及分子模型的建立中均需应用尝试原理。本文就归一化结构因子的计算提出了尝试分子式的计算方法,并成功地用于五个未知晶体结构测定,这对扩展直接法的使用效力是极为有益的。

一、引 言

从测定第一个晶体结构(氯化钠晶体)到目前对复杂的生物体系分子的晶体结构研究, X射线单晶结构分析学科已经渡过了七十余年的时光。围绕着结构分析的核心问题——结构因子的相位确定,数十年来,由于晶体学工作者的努力,迄今已经取得了解决这一问题的许多行之有效的方法,诸如尝试法、Fourier 综合法、Patterson 法、直接法、重原子法、同晶置换法、反常散射法、分子置换法等。同时,由于计算晶体学的发展(大型计算机的使用和各种单晶结构分析专用程序的建立)、衍射实验设备的完善,在“小分子”(其分子量常在一、二千以下,这不是一个绝对的概念也不是结构分析难易程度的标志)的结构分析中, X射线单晶结构分析已经发展为一种可供晶体学工作者使用的常规物理方法。其中,由于在直接法中采用了较为完善的数学模型,因而伴生有一套高效率的结构分析专用程序使之便于应用,所以它的使用更为普遍、更为有效,也便于非晶体学工作者的使用。当然,这并不是说晶体学中的相位问题已经全部解决,因为人们所期望的不仅在小分子中而且在较大分子的结构测定中也能具有当今直接法的应用效力。因此,即便到目前为止,直接法的“自动解析晶体结构”的能力已有很大的进展,但是在具体使用中尚有许多非数学形式的“技巧”问题,这就是在使用直接法测定晶体结构过程中时常会出现的“失败”情况。何以失败?一方面是由于以 Σ_2 关系式为基础的直接法,在进行相位迭代的过程中总是从部份已知相位的起始反射扩充到全部未知反射,有时在迭代的初始阶段,某些 Σ_2 关系式会有错误信息,如果它同时是一个概率很大的关系式,这势必导致一系列相位结果的失误,从而使结构测定失败,这是直接法的一个“先天不足”之处;另一方面,在直接法中还有另一个重要基础,即 $|E|$ 值的计算与选取。在直接法中是把归一化结构因子 $|E|$ 视为一个随机量处理,但直接法的这一假设并不与晶体结构的真实情况完全一致,只是一个近似性很好的假设。无数晶体结构的正确测定说明这一近似的可靠性、合理性。但是,直接法的使用还是经常遇到“失败”的情况,有时不单是在大的结构中出现,甚至一些

分子量很小的分子结构测定中也会出现周折,这些都涉及到 $|E|$ 值的计算问题.总之,这两方面因素的出现应归结为方法的不完备性而并非是晶体结构的“不可解”.影响 Σ_2 关系式的因素可归结为起始套反射的数目与选择,而 $|E|$ 值的计算则受归一化方式(如: Wilson, Debye, K 曲线等方式)、分子式确切与否的影响.时常出现这样的情况,貌似符合统计规律的 $|E|$ 值给不出正确结构解;而貌似奇异的 $|E|$ 值却给出正确的结构,这就是一种近似性的表现.因此,就出现了一个如何避免失败提高直接法应用效力的问题.不难看出,围绕 Σ_2 关系式与 $|E|$ 值两个因素,存在着一定的尝试性质,我们称之为尝试性原理.尝试法是结构分析的启蒙方法,早期的晶体结构测定都是以一定的尝试为基础的.不过在今天它却具有了新的内容.比如来自天然物的一些微量或极微量组分的结构测定,时常由于其量甚微而难以从其它物理、化学方法获得必要的甚至是起码的化学信息(如分子式).如何应用直接法测定它们的结构呢?这就提出了一个“纯粹”晶体学的结构问题——即不依附任何已知信息的一个结构测定问题.

针对此,我们在直接法使用的不同阶段采用了尝试分子式的计算、相位测定中的尝试法、模型建立中的尝试法等.结果表明,尝试性原理在直接法中的使用,大大地降低了“失败”例.直接法的使用已经从一种专门的数学方法而演变为“常规”的物理方法.因之,尝试法的应用就显得甚为迫切.多数结构的测定过程总是伴随着一定的尝试性质,如何适时地采用尝试法就成为提高直接法使用效力的一种需要了.

二、尝试分子式的计算

应用直接法测定晶体结构,在确定了晶体的对称群后,下一个环节将是如何计算归一化结构因子 $|E(hkl)|$ 值.通常 $|E|$ 值由下式确定:

$$|E(hkl)|^2 = \frac{F_0^2(hkl)}{\epsilon \sum_{j=1}^N f_j^2}, \quad (1)$$

式中 $F_0^2(hkl)$ 可以从实验强度数据经过统一、还原后得到; ϵ 称为消光因子,它依空间群及衍射指标类型而异; f_i 为第 i 个原子的散射因子.显见, $\sum_{j=1}^N f_j^2$ 的计算遍及全晶胞的 N 个原子,也即将涉及到待定结构的分子式信息.但是,在众多的天然有机分子结构测定中,有时因样品微量,在提供衍射实验之前无法得到有关分子式信息,或至多根据植物生源、分离提取过程做出某种假设,那么,如欲使用一个“纯晶体”去测定晶体结构,首要的一个问题便是如何设法得到一个试用的分子式.

在对称性测定中,我们可以获得晶体的对称群(唯一或几种可能)、晶胞参数($a, b, c, \alpha, \beta, \gamma$)与晶胞体积 V .在晶体密度(D_0)与分子量(M)之间存在着如下关系式:

$$D_0 = \frac{M \times z}{0.6023 \times V} = D_0(M), \quad (2)$$

式中 z 为晶胞内分子数,由衍射数据可以取得 z, V 的准确数值.因此,一旦获得了晶体密度值便可计算出分子量 M 的值(这曾是X射线晶体学发展的早期阶段的一种应用).一般

说来,我们可以使用悬浮法等一系列方法测定晶体的密度值,但是如果没有足够的样品,完成这些实验是存在一定困难的。对于天然有机分子,能不能根据有关衍射信息“给出”一个估算晶体密度呢?假如可能,下一步的问题就是如何从式(2)按所得的分子量去寻找试用分子式。兹分别叙述如下。

1. 估算晶体密度

对于天然有机分子,其组成元素多为碳、氮、氧、氢。有机分子之间符合一定的堆积规律^[1],晶体密度的大小标志着分子排列的致密程度。较密集的堆积(如分子内,分子间存在着氢键联系时)意味着分子的热振动效应的减少,而这又将直接反映在衍射图谱上具有较高的分辨率(当然晶体维度要符合衍射实验的要求);反之,衍射图谱的较低分辨率意味着分子具有较大的热振动效应,也即分子的排列较为松散。前者相应于较大的密度值,后者相应于较小的密度值。

我们统计了五十个有机分子的晶体密度值^[2]与衍射分辨率的关系。为方便起见,我

我们以衍射数据中可观察反射所占的比例做为分辨率的一种表示(记为 P)。由图 1 可知,尽管有机分子的密度范围常为 $0.9\text{--}1.9\text{g}\cdot\text{cm}^{-3}$ 之间(也有例外者),但图 1 表明,多数有机分子其密度值处于 $1.10\text{--}1.35\text{g}\cdot\text{cm}^{-3}$ 之间。对于宏观缺陷不甚明显的晶体,就可根据图 1 得到一个估算晶体密度值。我们提供的天然有机分子的可几密度值为 $1.30\text{g}\cdot\text{cm}^{-3}$ 。诚然,这一估算值带有一定的误差,但我们以下的试验表明,这并不影响获得正确的分子结构。

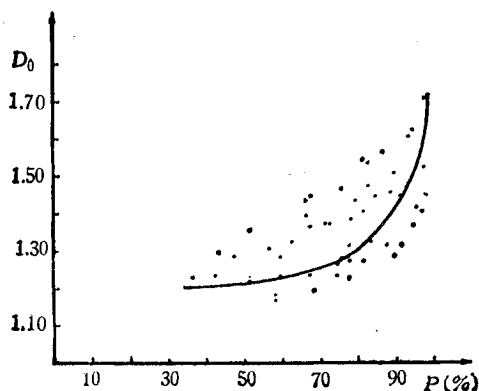


图 1

2. 尝试分子式的计算

由于天然有机分子中的元素多为 C, N, O, H, 由植物生源、提取过程的化学反应,多少可以得到一点化学信息。在直接法中,通常是把 C, N, O 原子视为“等同原子”处理的,在结构测定的早期阶段(获得正确的 E 图)是不予区别的。因此,一旦获得了估算晶体密度就可得到估算分子量 M , 下一步的问题是给出其中等同原子(常视为 C 原子)与 H 原子的数目。这一关系可以表示为

$$Q \approx \frac{\sum_{\text{C,N,O}}}{\sum_{\text{H}}}, \quad (3)$$

式中 $\sum_{\text{C,N,O}}$ 表示等同原子数目, $\sum_{\text{H}}$ 表示氢原子数目, Q 称为原子比例系数。我们统计了五十个有机分子的结构数据^[2], 其实验条件为: θ 角范围 $\leq 68^\circ$, $\text{CuK}\alpha$ 辐射, 可观察点标准为 $I \geq 3\sigma(I)$, 晶体大小 $(d) \approx 2/\mu$ (μ 为线性吸收系数), 实验在常温下进行。图 2 所示为 Q 值分布图, 其可几值落于 0.8 与 1.3 两个范围; 图 3 所示为相应的晶体密度 D ,

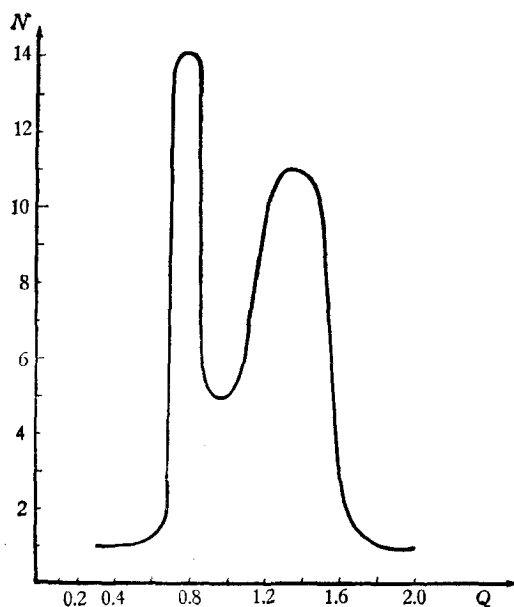


图 2

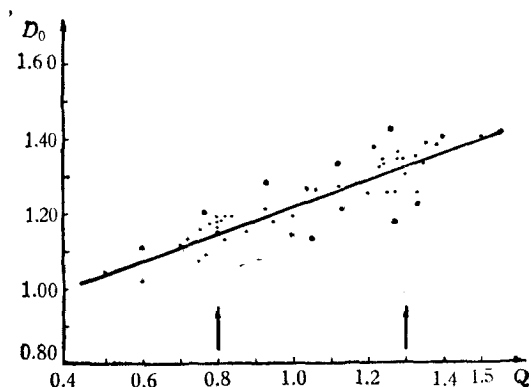
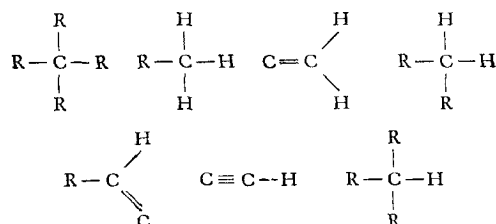


图 3

与 Q 值的关系。由图 3 可见, Q 的两个可几值落在 1.15 与 $1.30 \text{ g} \cdot \text{cm}^{-3}$ 的晶体密度值域, 对比图 1 可知, 在晶体衍射分辨率 P , 晶体密度值 D_0 与原子比例系数 Q 之间存在着一致的结论。就一般有机分子而言, Q 值的升高相应着较大的晶体密度值 (图 3), 反之亦然。因此, 由衍射谱图的分辨率可以得到晶体的估算密度值、原子比例系数, 由此就能给出一个尝试分子式。

从有机分子的结构特征看, Q 值大者表示该分子内等同原子数目多于氢原子数目; Q 值小者表示该分子内等同原子数目少于氢原子数目, 而氢原子数目的多少则是与分子中原子的种类、原子间的成键方式有关的。比如, 在碳原子中, 与氢原子的成键形式可有如下几种:



其中 R 可以是碳原子也可以是其它原子或取代基。通常的有机分子中碳、氢原子数目相近, 如果 Q 值高了, 表明该分子中或具有一定数目的双键甚至三键联结方式, 或含有较多的氧原子。

在氮原子中呈以下几种形式:

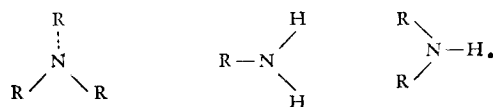


表 1

h	k	l	E^*	E	h	k	l	E^*	E
1	8	5	6.135	6.199	4	5	-5	2.911	2.914
2	0	-7	5.260	5.253	2	2	-3	2.867	2.866
1	-3	1	5.210	5.267	1	-5	-3	2.860	2.859
0	-11	-4	4.745	4.736	2	11	-3	2.805	2.798
2	5	6	4.716	4.716	1	3	-7	2.725	2.714
2	0	-6	4.561	4.550	4	6	-5	2.717	2.722
3	8	-2	4.241	4.231	4	7	3	2.712	2.718
3	8	-1	4.210	4.199	2	7	5	2.699	2.694
2	14	-3	4.209	4.212	1	3	8	2.635	2.638
2	12	-2	4.204	4.198	5	1	1	2.609	2.611
3	11	-2	4.200	4.198	3	-1	-6	2.608	2.605
2	-7	2	4.035	4.026	4	-7	-4	2.607	2.610
1	-4	1	3.982	4.022	0	-14	-3	2.598	2.597
2	3	-7	3.973	3.966	5	1	0	2.539	2.545
0	7	-8	3.962	3.966	1	17	0	2.555	2.539
1	3	-8	3.954	3.952	3	-4	-5	2.511	2.505
1	4	-7	3.948	3.939	0	5	-8	2.510	2.511
2	12	-3	3.948	3.943	1	-16	-3	2.489	2.492
0	-12	-4	3.898	3.893	0	-7	4	2.487	2.482
3	-3	-6	3.842	3.838	4	-4	-5	2.439	2.442
3	1	7	3.807	3.812	1	9	0	2.434	2.429
1	4	-8	3.798	3.797	3	9	-2	2.366	2.362
3	-4	-6	3.795	3.792	2	-5	2	2.344	2.342
2	4	6	3.754	3.751	2	14	2	2.366	2.340
4	5	0	3.717	3.716	2	3	2	2.333	2.332
3	-3	-5	3.716	3.702	4	4	-1	2.331	2.329
0	2	4	3.672	3.678	4	4	0	2.308	2.307
1	15	-3	3.665	3.659	3	9	-1	2.308	2.300
3	9	-6	3.557	3.562	3	10	3	2.278	2.279
1	-15	-3	3.534	3.530	4	-5	-5	2.270	2.271
1	16	-4	3.505	3.511	0	6	7	2.224	2.222
0	7	-9	3.488	3.497	3	-11	3	2.217	2.213
4	7	-1	3.444	3.444	1	10	4	2.212	2.207
4	5	-1	3.422	3.418	2	1	-6	2.207	2.202
4	-7	-5	3.411	3.416	2	-8	2	2.169	2.163
2	14	1	3.244	3.246	0	-12	1	2.164	2.158
1	6	-8	3.203	3.204	2	-1	-7	2.151	2.147
3	0	-6	3.187	3.183	1	-2	5	2.143	2.136
1	15	-4	3.087	3.089	2	-7	-2	2.140	2.135
3	10	2	3.068	3.068	5	3	0	2.136	2.142
1	6	1	3.060	3.068	1	7	-8	2.133	2.134
2	11	-2	3.028	3.022	1	15	1	2.131	2.127
4	8	-1	2.940	2.943	1	-10	-6	2.129	2.128
1	1	9	2.126	2.129	2	13	1	1.897	1.895
0	-8	-6	2.102	2.100	4	-6	-4	1.895	1.897
3	-10	3	2.095	2.092	2	-16	-2	1.890	1.894

表 1 (续)

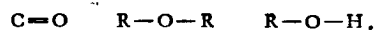
h	k	l	E^*	E	h	k	l	E^*	E
3	13	3	2.086	2.077	2	5	-8	1.888	1.891
1	-1	5	2.086	2.076	2	5	2	1.884	1.880
1	3	-3	2.079	2.085	0	6	-9	1.882	1.887
2	-6	2	2.079	2.075	1	1	-7	1.881	1.873
1	6	-4	2.075	2.070	4	-8	-5	1.871	1.876
2	-12	-6	2.073	2.078	1	-3	-3	1.861	1.866
1	-14	-3	2.070	2.068	1	-17	-2	1.846	1.849
0	0	-4	2.056	2.062	0	-14	-4	1.844	1.845
1	5	-4	2.039	2.035	1	16	-3	1.844	1.845
2	10	-7	2.038	2.042	0	5	-3	1.840	1.845
2	0	-2	2.038	2.043	1	-9	-7	1.837	1.836
0	-7	-7	2.023	2.022	3	2	-7	1.830	1.831
3	4	6	2.016	2.020	1	9	1	1.828	1.820
4	-3	-5	2.016	2.017	0	1	-5	1.813	1.809
2	-2	-6	2.012	2.007	2	15	-3	1.813	1.815
3	11	2	2.008	2.009	3	-8	3	1.808	1.804
4	3	-5	1.992	1.993	1	9	5	1.802	1.796
3	8	-5	1.987	1.986	0	-8	-5	1.792	1.787
1	18	0	1.982	1.988	4	-2	4	1.790	1.792
2	1	-7	1.980	1.976	2	3	-8	1.775	1.777
2	7	1	1.970	1.964	1	-10	-5	1.772	1.768
1	-13	-3	1.967	1.960	3	-1	-2	1.771	1.765
3	10	-2	1.966	1.963	0	-3	1	1.734	1.764
2	-5	6	1.964	1.962	1	0	-4	1.732	1.733
3	6	-6	1.962	1.962	1	10	0	1.731	1.726
2	7	6	1.958	1.960	1	4	-3	1.728	1.733
2	2	-2	1.958	1.962	3	0	-5	1.727	1.722
1	-6	-3	1.957	1.955	1	-6	2	1.725	1.729
4	5	-4	1.954	1.955	0	4	8	1.724	1.725
1	-6	-8	1.944	1.946	1	14	5	1.713	1.717
0	3	9	1.940	1.944	0	-9	0	1.707	1.707
0	-3	2	1.937	1.963	1	-2	-6	1.701	1.695
1	5	7	1.935	1.931	3	-2	-2	1.695	1.690
2	-9	-2	1.926	1.920	2	5	1	1.693	1.690
1	11	5	1.911	1.908	0	-12	0	1.692	1.687
0	5	8	1.909	1.911	1	13	-3	1.686	1.680
1	17	1	1.906	1.907	2	8	2	1.679	1.674
1	11	4	1.904	1.901	2	-2	0	1.672	1.679
1	2	-2	1.671	1.689	0	1	-9	1.497	1.499
1	-2	-4	1.669	1.670	1	0	-6	1.495	1.491
1	0	-3	1.650	1.661	2	10	6	1.493	1.498
3	0	7	1.641	1.652	1	5	6	1.478	1.475
4	-2	-6	1.646	1.650	2	10	-6	1.474	1.475
0	-10	0	1.631	1.628	2	12	2	1.459	1.459
2	-6	-7	1.629	1.629	1	-2	-5	1.457	1.453
1	4	2	1.627	1.634	3	-10	-1	1.457	1.453

表 1 (续)

h	k	l	E^*	E	h	k	l	E^*	E
0	5	3	1.626	1.628	2	-9	-5	1.456	1.453
0	-8	-1	1.616	1.618	0	-12	-3	1.453	1.450
1	-8	-7	1.596	1.596	1	-6	-6	1.451	1.447
1	8	1	1.589	1.587	0	-8	8	1.443	1.445
3	-6	-5	1.585	1.582	1	14	4	1.440	1.441
1	12	6	1.582	1.585	3	5	2	1.439	1.435
4	-5	-1	1.575	1.572	1	-3	-8	1.438	1.438
1	2	-7	1.574	1.570	2	10	5	1.435	1.436
1	-4	6	1.574	1.569	1	-2	-8	1.431	1.430
4	10	-2	1.566	1.570	2	2	-6	1.429	1.425
2	-1	-2	1.565	1.569	3	-3	-4	1.420	1.416
3	8	-6	1.563	1.565	1	1	-1	1.420	1.440
3	3	-4	1.563	1.558	3	8	0	1.416	1.412
0	-18	0	1.559	1.563	0	-9	5	1.416	1.411
0	-10	5	1.555	1.551	0	-9	4	1.411	1.407
2	7	-5	1.551	1.546	2	8	1	1.411	1.407
1	7	5	1.549	1.543	1	7	8	1.411	1.414
0	-12	-5	1.549	1.549	0	-10	-7	1.410	1.413
3	14	-2	1.539	1.543	5	1	-3	1.409	1.411
3	-4	-1	1.538	1.533	2	3	-3	1.406	1.404
1	13	-7	1.537	1.539	0	5	7	1.404	1.402
2	-14	-5	1.534	1.538	2	-3	-3	1.404	1.404
2	2	8	1.531	1.536	3	11	3	1.400	1.401
1	-10	5	1.531	1.527	2	-6	-5	1.399	1.395
1	-4	-3	1.526	1.529	1	13	-2	1.391	1.388
2	13	2	1.520	1.521	0	3	3	1.380	1.388
2	4	2	1.518	1.516	4	-3	-4	1.378	1.378
2	-2	-1	1.513	1.518	3	7	-6	1.373	1.374
1	12	1	1.511	1.506	0	1	-7	1.368	1.364
4	6	3	1.511	1.514	0	-16	4	1.365	1.367
1	-8	-6	1.507	1.503	2	-10	-7	1.363	1.366
1	0	-1	1.505	1.530					
2	-13	-4	1.500	1.501					
1	5	-3	1.499	1.499					

E^* 系由尝试分子计算而得; E 为正确分子式计算所得。

在氧原子中则呈如下形式:



诚然, $|E|$ 值的计算是个关系甚大的问题^[3], 不过正如我们的试验所证明的: 倘若分子中的原子分布接近“随机状态”, $|E|$ 值的计算方法选择的依赖性较小; 反之, 一旦分子出现了“择优”取向而偏离其随机假设时, 尽管给定正确的分子式仍然无法取得正确的结构^[4].

3. 误差分析

使用尝试分子式对归一化结构因子的影响如何呢? 由(1)式知,按 Wilson 统计, $|E|$ 值通常依 $\sum_i f_i^2$ 而异. 假定以 $\sum_i f_i^2$ 代表由尝试分子式所得的值,以 $\sum_i f_i^2$ 表示正确分子式的值,那么

$$\begin{aligned}\sum_i f_i^2 &= (N + \delta N)f_C^2 + (M + \delta M)f_N^2 + (P + \delta P)f_O^2 + (L + \delta L)f_H^2 \\ &= \sum_i f_i^2 + (\delta N f_C^2 + \delta M f_N^2 + \delta P f_O^2 + \delta L f_H^2) \\ &= \sum_i f_i^2 + \delta \sum_i f_i^2,\end{aligned}\quad (4)$$

式中 N, M, P, L 分别表示分子中的 C, N, O, H 原子的正确数目; $\delta N, \delta M, \delta P, \delta L$ 分别表示由尝试分子式引进的误差. 由(4)式可知,项 $\delta \sum_i f_i^2$ 引进一个系统误差,它对 $|E|$ 值计算的影响是一致分布的.

我们曾以例 1 为例分别用正确分子式(E)与尝试分子式(E^*)计算了 $|E|$ 值(表 1), 获得十分满意的结果. 这一例证说明尝试分子式的可靠性.

在把衍射分析方法引向一个常规物理方法的过程中, 尝试分子式解决了使用直接法的一个关键问题, 这就增强了衍射分析方法应用的独立性, 提高了它的自动解析程度.

三、应 用

我们曾对四个未知晶体结构测定使用了尝试分子式的计算, 结果是十分成功的.

1. Y1005 晶体结构测定

Y1005 晶体呈透明棕色片状. 其空间群为 $P\bar{1}$, 晶胞参数 $a = 4.763(1)\text{\AA}$, $b = 17.838(2)\text{\AA}$, $c = 8.774(1)\text{\AA}$, $\alpha = 93.55(1)^\circ$, $\beta = 96.15(1)^\circ$, $\gamma = 93.94(1)^\circ$, $V = 712.85(6)\text{\AA}^3$, 晶胞内分子数 $z = 2$. 选取 $I \geq 3\sigma(I)$ 为可观察点标准, 采用 $\text{CuK}\alpha$ 辐射, 用 P2_1 四圆衍射仪收集强度数据, 独立衍射点 $N = 1923$. 可观察点 $N_{ob} = 1815$, 由此得到 $P = 0.94$. 由图 1 得到估算晶体密度值为 $1.35\text{g} \cdot \text{cm}^{-3}$, 由图 3 得到相应的 Q 值为 1.4 左右. 按该样品的植物生源曾认为分子中含有 N 原子, 故尝试分子式给定为 $\text{C}_{15}\text{N}_2\text{O}_4\text{H}_{15}$, 由此计算 $|E|$ 值(表 1 E^* 一列). 应用 MULTAN-78 程序测定其分子结构, 在 E 图上获得廿个非氢原子位置, 其中包含了一个错误原子(图 4, 标以 $\times$ 者). 经 Fourier 综合与最小二乘法修正获得正确的分子结构, 其分子式为 $\text{C}_{16}\text{O}_7\text{H}_{16}$, 计算晶体密度值为 $1.49\text{g} \cdot \text{cm}^{-3}$.

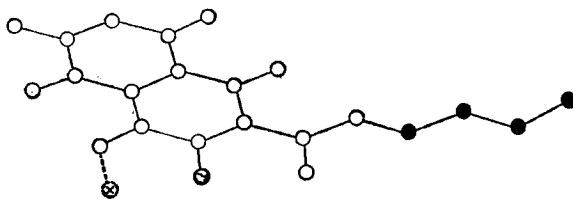


图 4

2. Eu16 晶体结构测定

空间群为 $P2_1$, 晶胞参数 $a = 8.669(2) \text{ \AA}$, $b = 10.771(5) \text{ \AA}$, $c = 10.436(2) \text{ \AA}$, $\beta = 104.21(2)^\circ$, $V = 944.58 \text{ \AA}^3$, $z = 2$. $N = 1361$, $N_{\text{obs}} = 1020$, $P = 0.75$. 由图 1 得到估算晶体密度值为 $1.15 \text{ g} \cdot \text{cm}^{-3}$ 左右, $Q \approx 0.8$, 给定尝试分子式为 $\text{C}_{20}\text{O}_2\text{H}_{26}$, 由此计算了 $|E|$ 值. 应用直接法测定了分子结构, 正确分子式为 $\text{C}_{20}\text{O}_2\text{H}_{32}$, 计算晶体密度值为 $1.08 \text{ g} \cdot \text{cm}^{-3}$. (见图 5).

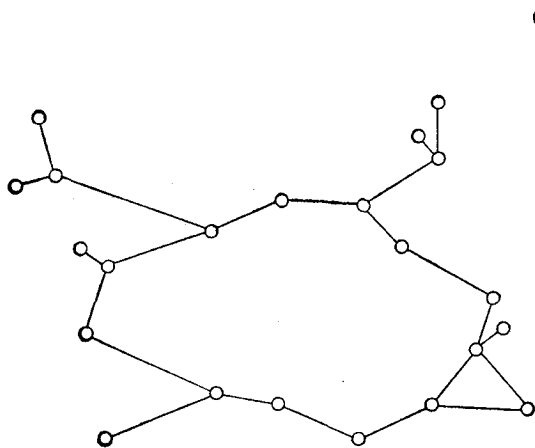


图 5

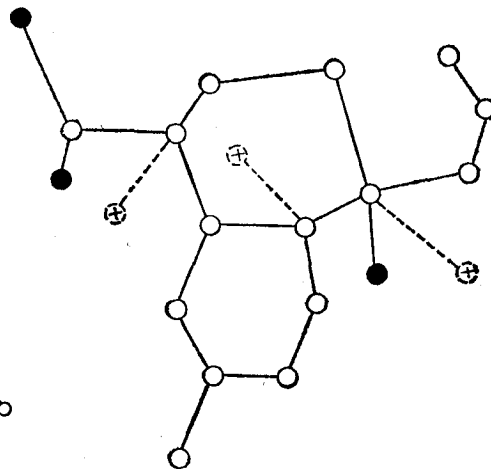


图 6

3. FOR4 晶体结构测定

空间群为 $P2_12_12_1$, 晶胞参数 $a = 6.296(3) \text{ \AA}$, $b = 9.379(3) \text{ \AA}$, $c = 26.659 \text{ \AA}$, $V = 1574.22 \text{ \AA}^3$, $z = 4$. $N = 1498$, $N_{\text{obs}} = 878$, $P = 0.59$. 由图 1 得到估算晶体密度为 $1.10 \text{ g} \cdot \text{cm}^{-3}$, $Q \approx 0.70$, 由此给出尝试分子式为 $\text{C}_{22}\text{H}_{30}$. 应用直接法测定分子结构, (图 6), 分子式确定为 $\text{C}_{16}\text{ONH}_{27}$, 计算晶体密度值为 $1.05 \text{ g} \cdot \text{cm}^{-3}$.

4. GFP 晶体结构测定

空间群为 $C2$, 晶胞参数 $a = 12.940(12) \text{ \AA}$, $b = 6.253(8) \text{ \AA}$, $c = 24.010(24) \text{ \AA}$,

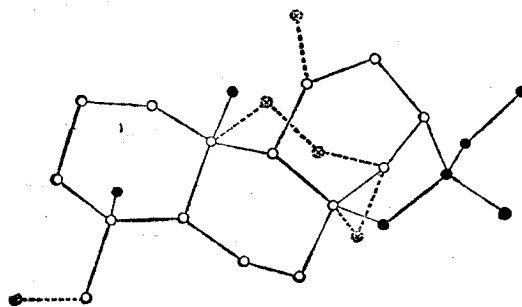


图 7

$\beta = 113.24(9)^\circ$, $V = 1785.18 \text{ \AA}^3$, $z = 4$. $N = 1317$, $N_{ob} = 1053$, $P = 0.80$. 由图 1 得到估算晶体密度值为 $1.20 \text{ g} \cdot \text{cm}^{-3}$, $Q \approx 0.80$, 给定尝试分子式为 $\text{C}_{21}\text{O}_2\text{H}_{28}$. 由此计算了 $|E|$ 值, 并用直接法测定其分子结构 (图 7). 正确的分子式应为 $\text{C}_{20}\text{O}_2\text{H}_{34} \cdot \text{H}_2\text{O}$, 计算晶体密度值为 $1.21 \text{ g} \cdot \text{cm}^{-3}$.

以上四个天然有机分子属于“纯晶体”结构测定例, 它表明了尝试分子式的应用效力.

5. 在一种极端的情况下, 若以碳原子的散射因子 (f_c) 取代氢原子的散射因子 (f_H) (见 (4) 式) 情况将如何呢?

我们选定未知化合物 $\text{C}_{15}\text{OH}_{20}$ 进行试验, 空间群为 $P2_1/n$, 晶胞参数 $a = 11.184(4) \text{ \AA}$, $b = 15.460(6) \text{ \AA}$, $c = 7.298(3) \text{ \AA}$, $\beta = 96.5(1)^\circ$. 应用直接法测定其结构. 按尝试分子式 $\text{C}_{20}\text{H}_{20}$ 计算, 由 E 图获得正确分子结构 (图 8(a)); 由尝试分子式 C_{15} 计算所得 E 图出现了较多的伪峰 (图 8(b)).

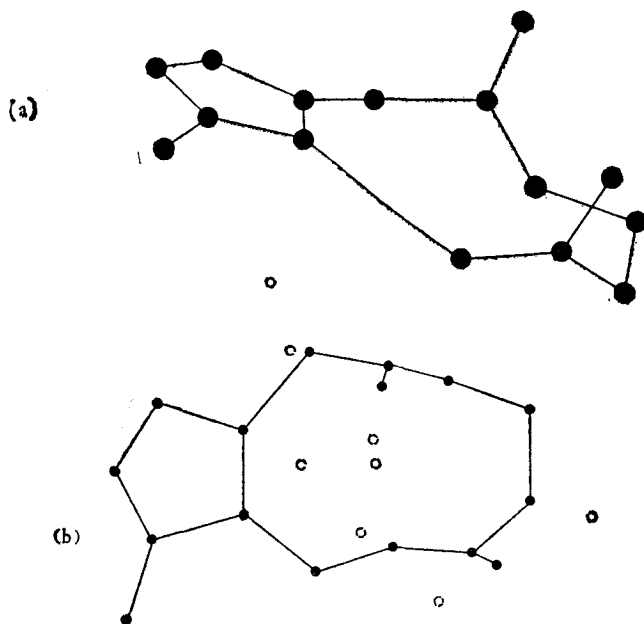


图 8

以碳原子取代氢原子, 其效果类似于提高“重原子”的贡献, 随着有机分子中氢原子数目的增加, 这一影响将表现在高衍射角度的衍射点的 E 值计算的偏离, 并将增加 E 图中的“噪声比”.

J. Clardy 教授提供 Y1005, Eu16, FOR4, GFP 四个晶体衍射数据, 谨致谢意.

参 考 文 献

- [1] A. I. Kitaigorodskii, *Organic Chemical Crystallography*, New York, (1961).
- [2] 数据采自 *Cryst. Struct. Comm.*, 8(1979); 9(1980); 10(1981).
- [3] Y. Subramanian, S. R. Hall, *Acta Cryst.*, A38(1982), 577; S. R. Hall, Y. Subramanian, *Acta Cryst.*, A38(1982), 590; 598.
- [4] 郑启泰, 物理学报, 待发表.

THE TRIAL AND ERROR PROCEDURES IN THE DIRECT METHOD (I)

ZHENG QI-TAI

(Institute of Biophysics, Academia Sinica)

ABSTRACT

Some trial and error procedures are often used in the process of structure determination (calculating normalized structure factor, determining phase and building molecular model) by the direct method. The present paper provides a procedure to calculate a trial molecular formula (TMF) on the basis of the diffraction data. Five unknown structures have been solved successfully by means of the TMF. This method is helpful in extending the application of the direct method.