

重原子信息在直接法中的应用——含重原子晶体结构因子相位的统计估算

范海福 韩福森 贺存恒

(中国科学院物理研究所) (中国医学科学院)

1980年10月30日收到

提 要

重原子所提供的信息对测定含重原子晶体结构有重要的作用。如何利用重原子信息将影响到测定工作是否顺利乃至能否成功。本文提出一种利用重原子信息的新方法。这个方法能够更充分地利用重原子信息,并能使重原子法同直接法更有效地结合起来。

一、引 言

在单晶体结构分析发展的历史上,有几次重要的突破,如青霉素、维生素 B₁₂、肌红球蛋白等的单晶体结构分析,其对象都是含有重原子的晶体,而且其分析工作之所以获得成功都有赖于重原子所提供的信息。长期以来,含有重原子的晶体结构比不含重原子的晶体结构易于测定,已经成为一个公认的事实。直至六十年代中期还占主导地位的 Patterson 法,用于含重原子晶体比用于不含重原子的晶体,能够解决更为复杂得多的结构问题。六十年代中期以后,直接法在实际应用上一再取得突破。在解决不含重原子的晶体结构方面,其效能已大大超过了 Patterson 法。但与此同时却出现了一个怪现象:含重原子的晶体往往比不含重原子的晶体更难于用直接法求解。好像重原子提供的信息尽管大大有助于 Patterson 法,却偏偏对直接法起干扰作用。因此有人提出了一些从结构因子中扣除重原子贡献的办法,把含重原子的晶体“改造”成不含重原子的晶体,再用直接法求解^[1-3]。这类方法显然不可能很好地利用重原子的信息。到底是重原子的信息注定要同直接法作对,还是通常的直接法还不善于利用它? 我们认为问题属于后一种情况。我们相信:只要能够恰当地利用重原子的信息,对直接法来说,含重原子的晶体结构同样会比不含重原子的结构易于解决。如果这样,既然直接法在解决不含重原子的晶体结构方面远比 Patterson 法等其他方法有效,那末它在解决含重原子的晶体结构方面也将会有类似的表现。范海福等已就如何在直接法中利用重原子信息的问题发表过一系列文章^[4-7]。本文在以前工作的基础上提出了一种更为有效的利用重原子信息的方法。

二、含重原子晶体结构因子相位的统计估算

令 φ 表示结构因子的相位; φ_H 表示重原子贡献的相位。根据 Sim^[8] 的推导, φ —

$\varphi_H = \xi$ 的概率分布应为

$$P(\xi) = \exp[x \cos(\xi)] / 2\pi I_0(x), \quad (1)$$

其中

$$I_0(x) = \frac{1}{\pi} \int_0^\pi \exp[x \cos(\xi)] d\xi,$$

它是零阶第一类变型贝塞耳函数; $x = 2|F|/|F_H|/\Sigma f_L$; F ——结构因子; F_H ——重原子对结构因子的贡献; f_L ——轻原子的散射因子.

由(1)式可得 $|\varphi - \varphi_H|$ 的数学期望值为

$$\langle |\xi| \rangle = \frac{1}{\pi I_0(x)} \int_0^\pi \xi \exp[x \cos(\xi)] d\xi. \quad (2)$$

由(2)式可得结构因子相位的统计估算值

$$\varphi_E = \varphi_H + \text{sign}(\Delta\varphi) \langle |\xi| \rangle. \quad (3)$$

这里 $\text{sign}(\Delta\varphi)$ 代表 $\Delta\varphi = \varphi - \varphi_H$ 的正负号. 由(3)式所得的 φ_E 可用以代替 φ 来计算 E 图. 这样做能否成功, 取决于两个因素: 首先是由(3)式算出的 φ_E 是否同真正的结构因子相位 φ 相近; 其次是在实际应用时能否求出正确的 $\text{sign}(\Delta\varphi)$. 下面将分别探讨这两个问题.

三、结构因子相位估算值同真实相位的比较

用一个典型的已知结构为例: 紫草乌碱乙碘酸盐的分子式为 $C_{34}O_{11}NH_{47} \cdot HI$. 晶体属单斜晶系; 空间群为 $P2_1$; 晶胞参数 $a = 12.58 \text{ \AA}$, $b = 14.38 \text{ \AA}$, $c = 11.00 \text{ \AA}$, $\beta = 114^\circ 36'$; 每晶胞含有两个分子. 在这个晶体中重原子的总散射贡献同轻原子的总散射贡献基本相当, 其重原子的原子序平方和同整个晶胞中全体原子的原子序平方和之比

$$\sum_1^{NH} Z_H^2 / \sum_1^N Z^2 = 0.581.$$

从这个比例看来, 这是一个典型的适合于用重原子法求解的结构. 但是, 实际上这个结构如用重原子法求解, 只能得到一个同时包含真实结构及其对映体的混杂图像. 使用 MULTAN 78 程序求解, 情况也是一样, 因为由 MULTAN 78 所得的结构因子相位退化为重原子本身的相位. 但是用(3)式计算, 却可以获得相当满意的结果(见表 1). 从表 1 可见, 在全部 85 个 $x \geq 3$, $|E| \geq 1.5$ 的衍射点中, φ_E 同真正相位 φ_R 之间的平均偏差是 8.5° , 最大的偏差是 48° , 相位偏差小于等于 15° 的衍射点占 91%. 当然, 要真正解出这个结构还必须设法求出(3)式中的 $\text{sign}(\Delta\varphi)$. 运用下节所述的办法, 这是不难做到的.

四、 $\text{sign}(\Delta\varphi)$ 的推引

1. 正负号组合法

这实际上是 MULTAN 法^[9]的一种变体. 先选出若干个衍射点作为起始套. 对每一

表 1

h	k	l	φ_E	φ_R	h	k	l	φ_E	φ_R
0	1	4	15	10	1	1	2	342	348
1	2	-5	347	353	1	2	6	346	355
1	2	8	168	163	1	3	-8	166	166
1	3	-1	343	332	1	4	-2	344	359
1	5	-6	192	194	1	6	-2	195	220
1	10	3	192	196	2	1	-3	164	176
2	3	-8	12	9	2	3	-1	157	158
2	4	-6	200	182	2	5	2	164	179
2	7	-5	195	205	2	7	-3	346	351
2	7	0	164	167	2	11	0	169	169
3	1	-2	18	52	3	2	-6	166	158
3	2	-1	193	193	3	3	-5	349	344
3	4	-6	12	6	3	4	-4	195	193
3	5	-5	193	195	3	6	-3	19	50
3	6	-1	195	210	4	1	-7	193	184
4	1	-2	163	161	4	3	-5	165	169
4	4	-8	346	355	4	4	-6	193	202
4	4	-3	344	357	4	5	-5	14	3
4	5	-4	16	11	4	5	-2	195	185
4	6	-6	12	28	4	6	-3	193	189
4	7	-7	350	327	4	8	-1	192	192
4	9	0	15	2	4	9	3	191	183
5	2	-8	11	8	5	2	-6	165	178
5	2	-5	192	193	5	2	-3	344	340
5	4	-6	345	352	5	4	-3	195	185
5	5	-2	14	17	5	5	1	193	193
5	6	-6	194	222	5	7	-7	168	158
5	11	1	10	11	6	1	-6	167	171
6	1	-1	195	199	6	2	-5	346	354
6	3	-6	13	22	6	3	-4	196	202
6	4	-8	348	352	6	5	-3	344	348
6	5	-1	167	176	6	7	1	164	155
6	9	-4	12	0	6	9	-1	191	183
6	10	0	349	356	7	1	-6	347	346
7	1	-5	14	25	7	2	-5	166	118
7	3	-6	191	190	7	3	-3	347	346
7	5	-6	349	355	7	5	0	13	20
7	6	-2	347	333	8	1	-8	12	2
8	2	-4	13	13	8	5	0	191	194
8	6	-6	167	177	8	7	-5	348	358
8	7	-3	193	187	8	8	-1	349	355
9	3	0	168	176	9	4	-6	191	206
9	8	-1	171	174					

个衍射点均考虑 $\text{sign}(\Delta\varphi)$ 等于 +1 及等于 -1 两种情况, 分别按 (3) 式求出两个可能的 φ_E 作为相应衍射点的两种起始相位。从各种可能的起始相位组合出发, 用带重原子校

正项的正切式^[6]进行相位扩展和修正, 选出品质因子最高的一套或几套计算 E 图。

2. 分量关系式法

若重原子的分布是中心对称的, 由 (2) 式可得结构因子实分量及虚分量绝对值的估算值:

$$\alpha = |E| \cos(\langle |\xi| \rangle) \text{sign}(F_H), \quad (4)$$

$$|\beta| = |E| \sin(\langle |\xi| \rangle). \quad (5)$$

用全部已知的 α , 加上一个未知正负号的 β 组成一个大起始套。令这个 β 的符号为正, 以确定对映体, 然后利用分量关系式^[4]

$$\beta_K = \frac{\theta}{V} \sum_{K'} \alpha_{K'} \beta_{K-K'}, \quad (6)$$

就可以求出其余 β 的正负号^[9]。

上述两种方法各有长处。正负号组合法可以利用现有的 MULTAN 程序略加修改即可使用, 不受重原子分布的对称性所限制, 执行起来较为简便。后一种方法可以利用大起始套的优势, 对映体的选定较有把握, 在克服相位退化^[10]方面也更为有效。

五、讨 论

本文提出的方法同扣除重原子贡献的方法相比有显著的优点。从测定一个非中心对称的含重原子晶体结构的过程, 可以看到这两种方法的明显差别。使用扣除重原子贡献的办法^[5], 在扣除重原子后所剩下的是一个完全未知的由轻原子组成的结构。这样一个“人造结构”其结构因子绝对值包含了统计估算的误差, 而且其结构因子的相位需要从头开始推引。重原子所提供的信息在此丝毫不起作用。事实上, 这样的方法等于把一个结构拆成互不相关的两个结构来测定, 其中一个只含重原子, 另一个只含轻原子。使用本文的方法, 一旦测定出重原子位置以后, 推引相位的问题就转化为推引正负号的问题, 即推引 $\Delta\varphi$ 的正负号或推引 β 的正负号。这相当于把非中心对称的结构问题转化为简单得多的中心对称的结构问题。此外, 由于充分利用了重原子所提供的信息, 在推引正负号时, 所用起始套可以比通常的起始套大得多。

范海福等以前曾提出过一些将重原子信息同直接法结合的方法。主要有, 用直接法消除由重原子引起的多解问题^[5], 以及利用重原子信息限制起始相位的排列方式^[7]等。同前一种方法相比, 本文的方法无须先找出轻原子的多解坐标, 这就省去了一道颇为麻烦的手续。同后一种方法相比, 当重原子所占比重较大时, 本文的方法可以获得较为准确的起始相位, 可以缩短相位扩展和修正的时间。但如重原子所占比重较小时, 本文的方法就可能不如受限相位排列法有效。

参 考 文 献

- [1] W. Hoppe and J. Gassmann, *Acta Cryst.* **B24**(1968), 97.
- [2] Th. E. M. Van den Hark, P. Prick and P. T. Beurskens, *Acta Cryst.*, **A32**(1976), 816.
- [3] S. E. Hull and M. J. Jrwin, *Acta Cryst.*, **A34**(1978), 863.

- [4] 范海福, 物理学报, 21(1965), 1114.
[5] 范海福、郑启泰, 物理学报, 27(1978), 169.
[6] 韩福森、范海福、古元新, 物理学报, 30(1981), 111.
[7] 范海福、韩福森, 本刊本期.
[8] G. A. Sim, *Acta Cryst.*, 12(1959), 813.
[9] G. Germain, P. Main and M. M. Woolfson, *Acta Cryst.*, A27(1971), 368.
[10] 范海福、千金子, 物理学报, 30(1981), 594.

ON THE APPLICATION OF HEAVY ATOM INFORMATION
IN DIRECT METHODS
—PHASE ESTIMATION OF THE STRUCTURE FACTORS IN
CRYSTALS CONTAINING HEAVY ATOMS

FAN HAI-FU HAN FU-SEN
(*Institute of Physics, Academia Sinica*)

HE CUN-HANG
(*Chinese Academy of Medical Science*)

ABSTRACT

Information drawn from the heavy atom plays an important role in the determination of crystal structures containing heavy atoms. To what extent would the determination proceed successfully depends on how the heavy atom information should be used. In this paper a new procedure is proposed to make full use of the heavy atom information and make the combination between heavy atom method and direct method more efficient.