

Acta Cryst. (1977). **A33**, 520–521

Une nouvelle approche du problème de la phase. Par OTTO DIDEBERG, *Laboratoire de Cristallographie, Institut de Physique, Université de Liège au Sart Tilman, B-4000 Liège, Belgique*

(Reçu le 14 septembre 1976, accepté le 13 décembre 1976)

From the Patterson function U_i vectors can be found, with $P(U_i) \sim 0$. This information with $g(x) \geq 0$ can be used to derive a structure-factor equation of the type: $\sum_{\mathbf{h}} F_{\mathbf{h}} F_{\mathbf{k}-\mathbf{h}} \exp(-2i\pi\mathbf{h} \cdot \mathbf{U}_i) \equiv 0$. For a simple model, the Sayre model, we compare the phase information deduced from probability formulae, Sayre's equations and the new relations.

Introduction

Les méthodes directes, utilisées pour résoudre le problème de la phase, ont permis de déterminer ces dernières années un très grand nombre de structures cristallines. Leur mise en oeuvre ne nécessite aucune information sur la stéréochimie de la molécule. Cependant, dans certains cas, ces méthodes ne conduisent pas à la solution. Parmi ces cas, on a pu montrer que l'introduction d'informations relatives à la stéréochimie de la molécule peut permettre de résoudre le problème de la phase. Nous avons les méthodes proposées par Thiessen & Busing (1974) et surtout celle de Main (1975).

Nous avons pensé qu'il serait utile d'introduire non pas une information relative à la stéréochimie de la molécule, qui n'est pas toujours connue, même partiellement, mais celle contenue dans la fonction de Patterson. Cette fonction contient implicitement une telle information.

Principe

Nous savons que la fonction de Patterson est définie par l'expression

$$P(\mathbf{U}) = \int_{V_c} g(\mathbf{x})g(\mathbf{x} + \mathbf{U})d\mathbf{x}. \quad (1)$$

Si pour certaines valeurs de U_i on a $P(U_i) = 0$, comme $g(\mathbf{x}) \geq 0$, nous devons nécessairement avoir

$$g(\mathbf{x}) \cdot g(\mathbf{x} + \mathbf{U}_i) \equiv 0 \quad \text{quel que soit } \mathbf{x}.$$

Appliquons le théorème de convolution,

$$\sum_{\mathbf{h}} F_{\mathbf{h}} F_{\mathbf{k}-\mathbf{h}} \exp(-2i\pi\mathbf{h} \cdot \mathbf{U}_i) \equiv 0 \quad \text{quel que soit } \mathbf{k}. \quad (2)$$

La fonction de Patterson est une fonction paire, $P(U_i) = 0$ entraîne $P(-U_i) = 0$, et en prenant la partie réelle et imaginaire, (2) devient

$$\sum_{\mathbf{h}} |F_{\mathbf{h}} F_{\mathbf{k}-\mathbf{h}}| \cos(\phi_{\mathbf{h}} + \phi_{\mathbf{k}-\mathbf{h}}) \cos 2\pi\mathbf{h} \cdot \mathbf{U}_i \equiv 0 \quad (2a)$$

$$\sum_{\mathbf{h}} |F_{\mathbf{h}} F_{\mathbf{k}-\mathbf{h}}| \cos(\phi_{\mathbf{h}} + \phi_{\mathbf{k}-\mathbf{h}}) \sin 2\pi\mathbf{h} \cdot \mathbf{U}_i \equiv 0 \quad (2b)$$

$$\sum_{\mathbf{h}} |F_{\mathbf{h}} F_{\mathbf{k}-\mathbf{h}}| \sin(\phi_{\mathbf{h}} + \phi_{\mathbf{k}-\mathbf{h}}) \cos 2\pi\mathbf{h} \cdot \mathbf{U}_i \equiv 0 \quad (2c)$$

$$\sum_{\mathbf{h}} |F_{\mathbf{h}} F_{\mathbf{k}-\mathbf{h}}| \sin(\phi_{\mathbf{h}} + \phi_{\mathbf{k}-\mathbf{h}}) \sin 2\pi\mathbf{h} \cdot \mathbf{U}_i \equiv 0. \quad (2d)$$

Nous obtenons ainsi une série de relations entre phases. Voyons à présent les problèmes soulevés, lors de l'utilisation de ces relations.

La relation (2) est exacte si $P(U_i) = 0$. Or le calcul de $P(U)$, à partir des $|F_{\mathbf{h}}|^2$ montre que la fonction a très peu de zéro. Ceci résulte de la dispersion des électrons dans l'atome et de leurs facteurs de température. Il serait donc logique de calculer $P(U)$ à l'aide des $|E_{\mathbf{h}}|^2$ et de remplacer dans l'expression (2) $F_{\mathbf{h}}$ par $E_{\mathbf{h}}$. Malheureusement, à présent, nous ren-

controns une autre difficulté: la sommation devrait porter sur une infinité de termes.

Afin de mieux apprécier ces divers facteurs et de vérifier la validité des expressions (2a)–(2d), nous avons utilisé un modèle idéal, celui de Sayre (1952).

Calcul du modèle de Sayre

Nous prenons une maille unidimensionnelle et centrée. La périodicité est de 20 Å, avec huit atomes par maille ($x_i = \pm 0,113, \pm 0,234, \pm 0,361, \pm 0,438$). Avec la radiation du cuivre ($\lambda = 1,54$ Å) pour un angle $2\theta = 135^\circ$, l'indice maximum des réflexions observées est de 24.

Nous définissons la densité électronique $g(x)$ par

$$g(x) = \delta \sum_{i=1}^N \exp[-\pi\delta^2 a^2 (x - x_i)^2].$$

Elle représente N courbes de Gauss de largeur $1/\delta$ centrées sur x_i . Nous avons alors

$$F_{\delta}(h) = \sum_{i=1}^N \exp\left(-\frac{\pi}{a^2\delta^2} h^2\right) \exp(2i\pi h x_i)$$

et

$$E(h) = \frac{1}{\sqrt{N}} \sum_{i=1}^N \exp(2i\pi h x_i) = \frac{F_{\infty}(h)}{\sqrt{N}}.$$

Nous pouvons calculer $E(h)$ et $F_1(h)$, $F_2(h)$, $F_5(h)$ les trois derniers facteurs de structure correspondent à un δ valant respectivement 1, 2 et 5 Å⁻¹. Ces diverses valeurs sont présentées dans le Tableau 1. Nous avons également consigné dans ce tableau les valeurs du facteur de structure de la fonction $g^2(x)$ pour $\delta = 2$. Ils se calculent à l'aide de l'expression

$$F_{\delta}^2(h) = \frac{\delta}{\sqrt{2}} \sum \exp\left(-\frac{\pi h^2}{2a^2\delta^2}\right) \exp(2i\pi h x_i).$$

Nous pouvons à partir des facteurs de structure calculer

$$P_{\delta}(U) = \frac{2}{a} \sum_{h=0}^{24} |F_{\delta}(h)|^2 \cos 2\pi h U.$$

La comparaison des valeurs obtenues par cette relation, avec celles tirées de l'expression (1) conduit aux remarques suivantes: les minima de la fonction $P_{\delta}(U)$ se déplacent suivant la valeur de δ et sont rarement des zéros. De plus $P_{\infty}(U_i) \sim 0$ avec $U_i = 0,2750$, alors qu'il y a un maximum de la fonction (1) pour $U_i = 0,2780$. Pour le calcul des expressions (2a) et (2b), nous choisissons un certain nombre de vecteurs U_i , soit m . Afin de pouvoir apprécier la valeur des $P(U_i)$ ainsi sélectionnés, nous définissons un paramètre:

$$\text{RES} = \left(\sum_{i=1}^m \left| \sum_{h=0}^{24} |F_{\delta}(h)|^2 \cos 2\pi h U_i \right| \right) / m.$$

Le calcul du premier membre des expressions (2a) et (2b), pour $k = 1, 24$, permet d'estimer la fiabilité des expressions en fonction de δ et m . Dans chaque cas nous avons 48 relations

Tableau 1. *Facteurs de structure* ($\times 10^3$) *pour le modèle considéré*

h	$E(h)$	$ F_1(h) $	$ F_2(h) $	$ F_3(h) $	$ F_2^2(h) $
0	2828	8000	8000	8000	11313
1	-501	1406	1414	1417	2002
2	-207	568	582	586	825
3	-249	656	691	702	986
4	-679	1695	1862	1919	2674
5	192	445	516	537	749
6	-1026	2187	2703	2869	3960
7	-327	630	841	911	1247
8	905	1549	2258	-2509	3434
9	1924	2881	4642	5306	7108
10	-955	1232	2220	2618	3463
11	387	424	864	-1055	1376
12	-560	511	1194	1514	1945
13	-503	377	1020	1348	1704
14	414	251	797	1101	1366
15	-2194	1060	3989	5781	7036
16	1047	396	1791	2731	3257
17	1116	326	1790	2883	3362
18	679	151	1016	1734	1975
19	-102	17	142	258	286
20	-169	21	217	421	455
21	-278	25	331	684	721
22	-85	5	93	207	211
23	-695	31	696	1665	1654
24	-1764	54	1610	4163	4008

Tableau 2. *Validité des expressions*

δ	m	RES	Cas douteux
1	17	21,37	0
2	35	8,97	0
2	10	1,83	2
5	27	1,59	37
∞	18	1,17	43

de la forme $\sum_{i=1}^m \left| \sum_{h=0}^{24} F_h F_{k-h} \cos(\phi_h + \phi_{k-h}) \cos 2\pi h U_i \right| = 0$ et

nous considérons comme douteuses celles dont le premier membre est en module plus grand que deux fois RES. Les résultats les plus significatifs sont repris dans le Tableau 2.

Les expressions se vérifient bien, si deux conditions sont remplies: d'une part, $P(U_i) \sim 0$ et d'autre part, les effets de limitation de série sont faibles.

Comparaison des méthodes

Le problème suivant concerne l'utilisation pratique de telles relations. Nous pouvons notamment les utiliser lors de l'examen des solutions proposées, par exemple, dans les méthodes de multisoluation. Cependant dans le cas que nous avons examiné, il est surprenant de constater le parallélisme étroit entre les relations trouvées et celles actuellement utilisées. Prenons une expression du type (2a), on a

$$\sum_h |F_h F_{k-h}| \cos(\phi_h + \phi_{k-h}) \cos 2\pi h U_i \equiv 0$$

pour m vecteurs U_i

$$\sum_{i=1}^m \sum_h |F_h F_{k-h}| \cos(\phi_h + \phi_{k-h}) \cos 2\pi h U_i \equiv 0$$

qui dans le cas centrosymétrique peut se mettre sous la forme

$$\sum_h a_h S_h S_{k-h} = 0.$$

En groupant les $|a_h|$ élevés, nous pouvons construire des inégalités du type

$$\left| \sum_h a_h S_h S_{k-h} \right| < \sum_{h'} |a_{h'}| + \varepsilon;$$

ε est l'erreur admise, ici $\varepsilon = 2RES$.

Reprenons quelques inégalités, ainsi construites dans l'exemple traité.

(a) Relations du type Σ_1 : pour $k = 18$ nous obtenons

$$|-10,73S_{18} + 16,84S_9^2| < 20,21$$

à mettre en rapport avec les relations Σ_1 de Hauptman et Karle

$$S(E_{2h}) \sim S(|E_h|^2 - 1).$$

(b) Relations du type Σ_2 : par exemple si $k = 24$

$$|21,96S_{24} - 29,20S_9S_{15}| < 13,66$$

comme l'a montré Cochran, nous pouvons associer une probabilité $P_+ = \frac{1}{2} + \frac{1}{2} \text{tgh}(N^{-1/2}|E_h E_{h'} E_{h-h}|)$ à la relation $S_9 S_{15} S_{24} = +1$, on a $P_+ = 0,995$.

(c) Quartets: nous retrouvons ainsi des relations entre plans sous la forme de quartets. Nous pouvons associer une probabilité

$$P_+ = \frac{1}{2} + \frac{1}{2} \text{tgh}[N^{-1}|E_h E_k E_l E_{-h-k-l}|(E_{h+k}^2 + E_{h+l}^2 + E_{k+l}^2 - 2)]$$

à l'expression $S_h S_k S_l S_{-h-k-l} = 1$ (Giacovazzo, 1975).

Dans notre modèle, pour $k = 19$, on a:

$$|15,34S_9S_{10} + 6,04S_4S_{15}| < 11,11$$

d'où $S_9S_{10}S_4S_{15} = -1$.

Si nous calculons P_+ , on a $P_+ = 0,35$ soit une probabilité de 0,65 d'avoir un quartet négatif.

(d) Bien d'autres relations entre phases peuvent être trouvées, par exemple pour $k = 16$

$$|1,92S_8^2 + 2,22S_7S_9 - 5,42S_{16}| < 8,22$$

qui exclut la solution $S_7S_9 = +1$ et $S_{16} = -1$.

Conclusion

Les équations de Sayre permettent de trouver le même type de relation, ainsi pour $k = 18$, si l'on utilise les valeurs consignées dans le Tableau 1

$$|21,55S_9^2 - 23,24S_{18}| < 45,47.$$

Cependant la différence essentielle entre les deux méthodes réside dans le fait que pour utiliser les relations de Sayre on doit connaître le second membre $|F_2^2(h)|$. Cette condition malheureusement n'est pas réalisée en pratique. Comparés aux méthodes probabilistes les nouvelles relations sont incontestablement plus lourdes à utiliser pour déterminer la valeur des phases. Ainsi son utilisation, soit pour affiner la valeur de celle-ci, soit pour tester la fiabilité d'une solution proposée par une méthode de multisoluation paraît beaucoup plus prometteuse.

Nous remercions le Professeur J. Toussaint pour l'aide qu'il nous accorde dans nos travaux.

Références

GIACOVAZZO, C. (1975). *Acta Cryst.* A 31, 252-259.
 MAIN, P. (1975). *Acta Cryst.* A 31, S10.
 SAYRE, D. (1952). *Acta Cryst.* 5, 60-65.
 THIESSEN, W. E. & BUSING, W. R. (1974). *Acta Cryst.* A 30, 814-821.