Acta Cryst. (1963). **16**, 317

Two Fourier functions for use in protein crystallography.

By L. K. Steinrauf, *Department of Chemistry, University of Illinois, Urbana, Illinois, U.S.A.*

(*Received* 2 *October* 1961 *and in revised form* 24 *August* 1962)

The difference Patterson and the Rossmann function (1960) are very practical starting points in the investigation of protein structure by the isomorphous replacement method. With these functions it has been possible in favorable cases to locate the heavy atoms substituted into the protein, which is necessary for the determination of the structure of the protein. The two functions to be proposed are intended to be used on more complicated heavy atom structures than are the above functions.

The first function is an electron density synthesis of the heavy atoms using as coefficients:

$$(|\mathbf{F}_{PH}| - |\mathbf{F}_P|) \exp (i\alpha_P) \qquad \text{I}$$

where $\mathbf{F}_P$ is the structure factor and α_P is the phase angle of the unsubstituted protein, $\mathbf{F}_{PH}$ is the structure factor of the protein with the heavy atom derivative, and $\mathbf{F}_H$ is the structure factor of the heavy atom derivative. For the noncentric case Function I is a very useful, although not an exact, representation of the electron density of the heavy atoms. It is very nearly the projection of $\mathbf{F}_H$ onto $\mathbf{F}_P$ (or the component of $\mathbf{F}_H$ that is parallel to $\mathbf{F}_P$) as may be seen in Fig. 1.

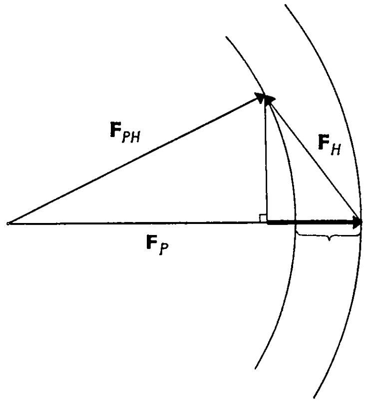
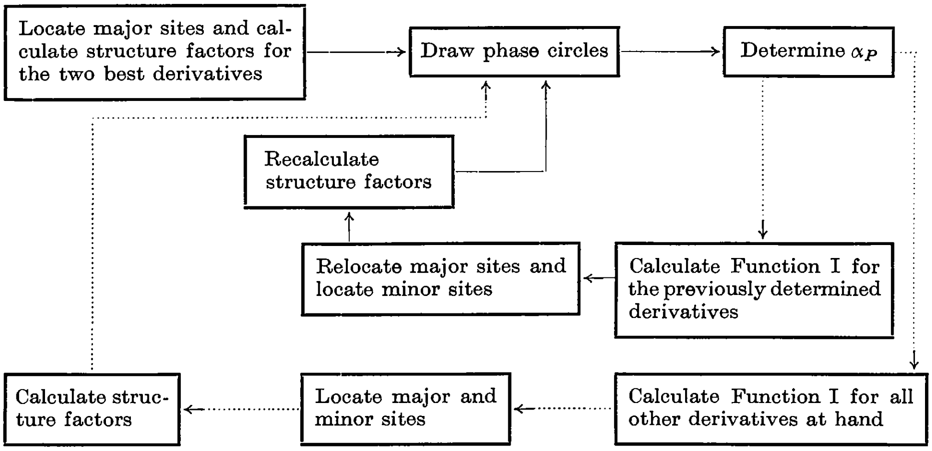


Fig. 1. The brace along $\mathbf{F}_P$ indicates the quantity $(|\mathbf{F}_{PH}| - |\mathbf{F}_P|) \exp (i\alpha_P)$, which will have the same phase angle and, with the conditions stated in the text, will be very nearly equal in length to the projection of $\mathbf{F}_H$, which is the heavily lined part of $\mathbf{F}_P$.

In practice this function has given accurate representation of heavy atoms as long as the average length of $\mathbf{F}_P$ is more than about four times the average length of $\mathbf{F}_H$.

Function I has been used to refine positional, vibrational and site-occupancy parameters of derivatives that are already rather well known. It has been used with some success in locating minor sites. The greatest use so far has been to locate the positions of derivatives which are too complicated to be located by difference Patterson or Rossmann functions; the addition of such derivatives to the phase circle determinations has been found to be well worth while. A flow chart is given below for the procedure used for triclinic lysozyme. The full lines follow the refinement cycle and the broken lines follow the path by which additional derivatives are added to the determination.

Since Function I is an approximation for noncentric reflections, peak heights calculated by its use will be expected to be low. The average value of a projection is the average value of the cosine, which is $2/\pi$ or 0·63. Therefore, the height of a peak using Function I should be about 60% of the height of the same peak of a synthesis using the structure factors of the heavy atoms, if they were known. This has been checked for an artificial heavy atom structure for which the calculated structure factors were projected onto random 'protein' structure factors obtained from a table of random numbers. The resulting projections together with the phases of the 'protein' were used in a synthesis of Function I. The peak heights so obtained averaged 58% as high as those obtained from a synthesis using the original structure factors of the heavy atoms; the background was considerably higher but the positions of the peaks differed by only insignificant amounts. It might be pointed out that Function I is similar to that usually used to calculate hydrogen positions for noncentrosymmetric structures.



Therefore, the hydrogen peaks in a difference Fourier leaving out hydrogen atoms should be only about 60% as high in a noncentric projection as in a centric projection.

The second function to be proposed is a Patterson synthesis using as coefficients:

$$(|\mathbf{F}_{PH_1}|-|\mathbf{F}_P|)^2+(|\mathbf{F}_{PH_2}|-|\mathbf{F}_P|)^2-(|\mathbf{F}_{PH_1}|-|\mathbf{F}_{PH_2}|)^2 . \qquad \mathrm{II}'$$

The Rossmann function gives the self-vectors of the heavy atoms of derivative one and the self-vectors of the heavy atoms of derivative two as positive peaks and the cross-vectors between one and two as negative peaks. In practice there may be many sites in each derivative, in which case the negative peaks may be obscured by the positive peaks. In Function II the self-vectors have been removed from the Rossmann function and the cross-vectors remain as positive peaks. This function does not involve any additional complications in scaling over the original Rossmann function, although the background will be higher.

Function II may be simplified to give:

$$(|\mathbf{F}_{PH_1}|-|\mathbf{F}_P|)(|\mathbf{F}_{PH_2}|-|\mathbf{F}_P|) . \qquad \mathrm{II}''*$$

It is easy to show that Function II″ is a cross-correlation between derivative one and derivative two, just as the Patterson is a self-correlation.

It should be pointed out that the algebraic sign of the coefficients of Functions I and II can, and frequently will, be negative. This sign must be included in the synthesis. The uses of these two functions and the results of a study of the heavy atom chemistry of triclinic lysozyme will be submitted for publication in the near future.

Reference

Rossmann, M. B. (1960). *Acta Cryst.* **13**, 221.

---

* This algebra was kindly shown by Dr R. E. Dickerson.

Acta Cryst. (1963). **16**, 318

Refinement of the crystal structure of 6-amido-3-pyridazone.

By Paul Cucka, *American-Standard, Research Division, Monroe and Progress Streets, Union, New Jersey, U.S.A.*

(*Received* 2 *October* 1962)

The structure of 6-amido-3-pyridazone was determined some years ago (Cucka & Small, 1954; I in the following), but the refinement was left incomplete because of the lack of computing facilities. The present article reports the results of a refinement of the structure carried out on an IBM 7090 computer using the least squares program written by Martin, Busing & Levy (1962).

The unit cell and intensity data were the same as used in I. Weights were assigned to the structure factors according to the scheme

$$\begin{array}{ll} |F_o| \le 4|F_{\min}| & \sqrt{w}=|F_o|/4|F_{\min}| \\ |F_o| \ge 4|F_{\min}| & \sqrt{w}=4|F_{\min}|/|F_o| \end{array}$$

$F_{\min}$ being the smallest observed F; zero weights were assigned to unobserved reflections and to the very strong $\bar{2}01$ reflection. Trial parameters were taken from the previous work (I), including estimates for the hydrogen atom positional parameters. Scattering factors for carbon, nitrogen and oxygen were taken from Berghuis *et al.*

Table 1. *Atomic coordinates*

Atom	x/a	$\sigma(x/a)$	y/b	$\sigma(y/b)$	z/c	$\sigma(z/c)$
C_1	0·2943	0·0009	0·4218	0·0005	0·5398	0·0010
C_2	0·2031	0·0011	0·4712	0·0006	0·3618	0·0011
C_3	0·1351	0·0010	0·4154	0·0005	0·1952	0·0011
C_4	0·1479	0·0009	0·3073	0·0005	0·2008	0·0010
C_5	0·0647	0·0008	0·2443	0·0005	0·0196	0·0009
N_1	0·2311	0·0007	0·2578	0·0004	0·3579	0·0007
N_2	0·3021	0·0007	0·3165	0·0004	0·5212	0·0008
N_3	0·0733	0·0007	0·1435	0·0004	0·0346	0·0009
O_1	0·3611	0·0007	0·4643	0·0003	0·7047	0·0007
O_2	0·9904	0·0006	0·2900	0·0004	0·8645	0·0007
H_1	0·196	0·012	0·543	0·008	0·356	0·015
H_2	0·068	0·012	0·436	0·008	0·096	0·016
H_3	0·364	0·013	0·289	0·008	0·613	0·016
H_4	0·385	0·013	0·616	0·007	0·842	0·016
H_5	0·513	0·012	0·401	0·007	0·907	0·015

Table 2. *Bond lengths*

$C_1{-}C_2$	1·4270 ± 0·0095 Å	$C_2{-}H_1$	0·932 ± 0·105 Å
$C_1{-}O_1$	1·2435 ± 0·0077	$C_3{-}H_2$	0·822 ± 0·093
$C_1{-}N_2$	1·3694 ± 0·0084	$N_2{-}H_3$	0·803 ± 0·099
$C_2{-}C_3$	1·3397 ± 0·0097	$O_1{-}H_4$	2·138 ± 0·096
$C_3{-}C_4$	1·4037 ± 0·0091	$O_1{-}H_5$	1·901 ± 0·093
$C_4{-}C_5$	1·5015 ± 0·0084	$N_3{-}H_4'$	0·882 ± 0·096
$C_4{-}N_1$	1·3049 ± 0·0078	$N_3{-}H_5'$	1·024 ± 0·094
$C_5{-}N_3$	1·3100 ± 0·0084		
$C_5{-}O_2$	1·2374 ± 0·0071		
$N_1{-}N_2$	1·3539 ± 0·0068		

Table 3. *Bond angles*

∠ $O_1{-}C_1{-}C_2$	127·8 ± 1·3°	∠ $C_3{-}C_4{-}N_1$	122·7 ± 1·2°
$O_1{-}C_1{-}N_2$	119·3 ± 1·0	$C_3{-}C_4{-}C_5$	119·7 ± 1·1
$C_2{-}C_1{-}N_2$	114·0 ± 1·0	$C_5{-}C_4{-}N_1$	117·6 ± 1·0
$C_1{-}C_2{-}H_1$	120·4 ± 13·5	$C_4{-}C_5{-}N_3$	118·1 ± 1·0
$C_1{-}C_2{-}C_3$	120·4 ± 1·2	$C_4{-}C_5{-}O_2$	118·5 ± 1·0
$H_1{-}C_2{-}C_3$	119·1 ± 13·7	$O_2{-}C_5{-}N_3$	123·4 ± 1·2
$C_2{-}C_3{-}H_2$	127·4 ± 15·6	$N_1{-}N_2{-}C_1$	126·8 ± 1·2
$C_2{-}C_3{-}C_4$	119·6 ± 1·2	$N_1{-}N_2{-}H_3$	118·9 ± 14·1
$H_2{-}C_3{-}C_4$	112·0 ± 12·1	$C_1{-}N_2{-}H_3$	113·7 ± 13·1
$C_4{-}N_1{-}N_2$	116·2 ± 1·0		

Table 4. *Deviations of atoms from molecular plane*

Atom	Deviation from plane	Estimated error
C_1	0·0035 Å	0·0082 Å
C_2	0·0022	0·0110
C_3	0·0269	0·0104
C_4	0·0001	0·0095
C_5	−0·0347	0·0067
N_1	0·0138	0·0062
N_2	0·0218	0·0072
N_3*	−0·1112	0·0058
O_1	−0·0347	0·0088
O_2*	−0·0021	0·0061

* Not included in determination of molecular plane.