von s unterscheiden, hat bei $|s|=\kappa$ (d.h. $p=0, y_M=0$) ein Minimum der Höhe $[(1-\kappa)/(1+\kappa)]^2$ und steigt nach den Grenzen $s=0$ ($|p|=\kappa$) hin auf den Wert 1 an. Zur Frage der praktischen Beobachtbarkeit von Fällen, wo das Maximum der d.p. auf der Seite positiver y liegt, kann z.B. auf die (222)-Interferenz von Kristallen des Zinkblende-Typs hingewiesen werden, bei Verwendung von Wellenlängen, die *zwischen* den Absorptionskanten beider Atomsorten liegen, sodass diese sehr verschieden stark absorbieren. Ihre zu (111) parallelen Ebenen liegen dann für (222) gerade um $d/2$ getrennt, die eine stärker *absorbierend*, die andere mit stärkerer *Real*-Streuung, womit die Voraussetzungen für eine zur üblichen inverse Kurven-Asymmetrie gegeben wären. Ein weiteres Beispiel, für welches ähnliche Verhältnisse zutreffen, gibt Cowley (1964) mit der (222)-Interferenz von Calcit, wo die Ca-Atome und die CO_3-Gruppen in um $d/2$ getrennten Ebenen liegen. Cowley bestätigt sogar, allerdings im Transmissions-Fall, experimentell für diese Interferenz die analog zu erwartende Inversion der Hell–Dunkel-Asymmetrie des R_0-Reflexes.

Die Verfasser haben zu danken der Direktion des AEG-Forschungsinstituts Frankfurt-Niederrad für die Möglichkeit zur Benutzung der dortigen Rechenanlage IBM 7040.

Literatur

BORRMANN, G. (1964). *Z. Kristallogr.* **120**, 143.
COLE, H. & STEMPLE, N. R. (1962). *J. appl. Phys.* **33**, 2227.
COWLEY, J. M. (1964). *Acta Cryst.* **17**, 33.
DE MARCO, J. J. & WEISS, R. J. (1965). *Acta Cryst.* **19**, 68.
KOHLER, M. (1933). *Ann. Phys. Lpz.* **18**, 265.
MILLER, F. (1935). *Phys. Rev.* **47**, 209.
PRINS, A. A. (1930). *Z. Phys.* **63**, 477.
RENNINGER, M. (1934). *Z. Kristallogr.* **89**, 344.
ZACHARIASEN, W. H. (1945). *Theory of X-ray Diffraction in Crystals.* New York: John Wiley.

Acta Cryst. (1967). **23**, 511

Bias, Feedback and Reliability in Isomorphous Phase Analysis

BY RICHARD E. DICKERSON, MARY L. KOPKA, JOAN C. VARNUM AND JON E. WEINZIERL
Gates and Crellin Laboratories of Chemistry, California Institute of Technology, Pasadena, California, U.S.A.*

(*Received* 2 *December* 1966)

The effect of assumed structural information about heavy atom sites on subsequent difference Fourier maps in the multiple isomorphous replacement phase analysis method is studied. Examples of good and bad derivatives of horse heart cytochrome C are used to illustrate the effect of the introduction of spurious derivatives into a phase analysis. Various means of discriminating between valid and invalid derivatives are compared, and suggested minimum standards for publication of a low resolution structure analysis are presented.

Introduction

In a conventional crystal structure analysis carried out at atomic resolution ($d_{\min}=\lambda/s_{\max}=\lambda/2 \sin\theta=0{\cdot}771$ Å for Cu $K\alpha$), the test of the final structure is the degree to which the structure factors calculated from the model agree with those actually observed. The course of analysis – the manner in which the model is obtained – is relatively unimportant. Electron density maps are ultimately more for display than for illumination, and their interpretation is seldom in doubt.

A much different situation prevails in macromolecular structure analysis using multiple isomorphous replacement methods at resolutions short of the atomic level. Here individual atomic positions cannot be found, and the electron density map itself is the sole end product. This map must then be interpreted in terms of a sensible structure, and it becomes vital that the investigator should know exactly how accurate his map is

One of the characteristics of Fourier series which makes Fourier refinement possible in small structures is the feedback property – the property that once an atom is put into the structure factor calculation and made to contribute to the phases (or signs), the atom persists in subsequent electron density maps. The degree and kind of persistence offers a clue as to whether it was put in correctly or not. In a certain sense the phases are *more* important than the amplitudes. Trial phases used with true amplitudes will, under favorable conditions, give a structure intermediate between trial and true structure. Trial phases with unit or random amplitudes will often give the trial model back again. If the trial model is so incorrect that the measured amplitudes are essentially random for this model, then the resulting electron density map can reproduce the false model and create a false impression of correctness. The saving factor is that observed and calculated amplitudes will disagree and the false model can be recognized for what it is. It is this last check which is lacking in low resolution protein or macromolecular structure analyses.

* Contribution No. 3458 from the Gates and Crellin Laboratories of Chemistry.

The correctness of the electron density map in a multiple isomorphous replacement phase analysis depends upon how well the heavy atoms have been located and characterized. Heavy atoms are usually located initially by difference Patterson methods of one type or another, and then characterized (extent of substitution or effective atomic number, and radial fall-off factor or empirical form factor) by methods such as Wilson-type plots or least-squares refinement. From this information, trial phases can be calculated, which can then be used in difference Fourier maps to improve the interpretation of the original heavy atom derivatives or to pull in multiple-site derivatives whose difference Patterson maps were too complex to interpret by themselves.

The danger in this process is that if a wrong interpretation of a heavy atom parameter is made and then if this interpretation enters into the phase analysis, subsequent rechecks involving those phases in difference Fourier maps may tend to confirm the wrong heavy atom. The investigator may carry on unwittingly with this wrong set of phases, produce an incorrect low resolution map of the molecule, and then spend fruitless hours trying to fit known polypeptide chain to what is only noise and error.

The purpose of this study was to see the effect of incorrect bias on a multiple isomorphous replacement analysis, to find out how serious the Fourier feedback problem was, and to try to find some criteria for assessing the reliability of assumed heavy atom parameters. If the investigator has made an error, how soon and in what manner will the subsequent analysis tell him so?

Experimental

The data used in this study were from crystals of the parent horse heart cytochrome C and from four heavy-atom derivatives, containing platinum, mercury, the combination of both at once, and palladium. Cytochrome C was extracted from whole frozen horse hearts by a method developed by Margoliash (1967) and crystallized in one to ten weeks from 95% saturated ammonium sulfate, 1*M* in sodium chloride. The platinum derivative, designated Pt30w, was prepared by diffusing $PtCl_4^{2-}$ into the pregrown crystals at a 7·5:1 mole ratio at *p*H 6 and photographed after 30 weeks. The mercury derivative, designated Hg14w, was prepared in a similar manner by diffusing in mersalyl (the sodium salt of salyrganic acid) at *p*H 6·8 and a 10:1 molar excess of heavy atom and was photographed after 14 weeks of aging. The double derivative was prepared using Pt and Hg at *p*H 6·8 at molar ratios of 1:1 and 3:1, respectively, and was photographed after seven weeks. Palladium as $PdCl_4^{2-}$ or $PdCl_2$ produced no intensity changes in ammonium sulfate, but was found to produce large and characteristic changes in crystals which had first been transferred from sulfate to 4·3 *M* mixed phosphate buffer (2:1 metal/protein mole ratio, *p*H 5·7, aged five weeks). The change of crystal medium itself altered only the innermost reflections, of fourth order or less. Observed structure factor data are shown in Table 1.

Cytochrome was found to crystallize in tetragonal space group $P4_1$ with cell dimensions $a=b=58{\cdot}5$, $c=42{\cdot}3$ Å. After Lorentz–polarization correction, parent and derivative data were placed on the same scale by a Wilson-type plot of $\ln(\langle F_P\rangle/\langle F_{PH}\rangle)$ *versus* s^2, where F_P and F_{PH} are observed parent protein and heavy-atom derivative structure amplitudes and $s=2\sin\theta$. Heavy-atom positions were found from difference Patterson maps using $(\Delta F)^2=(|F_{PH}|-|F_P|)^2$ as coefficients. Scale factors for scaling up the heavy-atom derivative data relative to the parent data were calculated by using an expression derived by Kraut (1961) based upon the assumption that in a centrosymmetric projection the origin peak should be the same height in a ΔI difference Patterson map as in a $(\Delta F)^2$ difference Patterson map:

$$K=\sum_{hk} F_P^2/\sum_{hk} F_P F_{PH}\,.$$

In practice the scaled-up difference Patterson maps were little improved over the level-scale maps ($\Sigma F_P=\Sigma F_{PH}$) other than in having a somewhat quieter background. In no case was any significant detail obscured by the use of level scale. Heavy atom structure factors for single sites were calculated from:

$$f_H=2Ae^{-Bs}\{\cos 2\pi(hx+ky)+\cos 2\pi(hy-kx)\}\,.$$

The extent of substitution, A, and the radial fall-off factor, B, were found by first assuming values of 1·00 and zero, respectively, and then making a Wilson-type plot of $\ln(\langle|\Delta F|\rangle/\langle|f_H|\rangle)$ *versus* s. This was found to give a better straight line plot than s^2, implying a sharper-than-Gaussian radial fall-off of heavy atom contribution near the origin of the diffraction pattern, and a flatter-than-Gaussian profile for the added heavy group. The assumption is physically reasonable at this resolution and the practical distinction is small. Unrefined heavy atom parameters are given in Table 2.

Signs were determined first by inspection and then by the phase probability method of Blow & Crick (1959; Dickerson, Kendrew & Strandberg, 1961*a*). The program which was used alternates a phase determining cycle with one cycle of full-matrix least-squares refinement of heavy atom parameters against the fixed phases of the previous cycle, in a manner first used by Dickerson, Kendrew & Strandberg (1961*a*,*b*) with myoglobin, and subsequently programmed by Kraut, Sieker, High & Freer (1962) for chymotrypsinogen, Muirhead (1966) for haemoglobin, and Lipscomb, Coppola, Hartsuck, Ludwig, Muirhead, Searl & Steitz (1966) for carboxypeptidase. The program is summarized by Dickerson & Palmer (1967). However, except in the 'four-data' phase set used as a check, the parameters in this paper were deliberately *not* refined, and were those taken by inspection from the difference Patterson maps and Wilson plots.

Table 1. *Observed structure factors for parent cytochrome* C *and for heavy atom derivatives*

All data are on the same arbitrary scale. Missing data are represented by −1·00. *AS*=ammonium sulfate; 4·3=4·3 *M* mixed phosphate buffer; *N*=native cytochrome without heavy atoms.

h k	N in AS	Pt in AS	Hg in AS	Double in AS	N in 4.3	Pd in 4.3
1. 0.	23·60	−1·00	−1·00	−1·00	50·66	74·55
1. 1.	5·64	−1·00	23·53	−1·00	37·18	47·98
1. 2.	24·58	2·50	20·20	3·44	28·02	22·49
1. 3.	22·81	12·93	15·05	8·65	19·58	18·56
1. 4.	28·39	36·40	17·17	36·07	24·50	19·73
1. 5.	12·30	5·65	11·61	5·30	15·02	20·15
1. 6.	46·81	33·73	35·55	34·10	51·33	44·39
1. 7.	13·75	8·37	16·16	6·26	16·20	15·67
1. 8.	54·59	49·56	49·59	53·49	55·87	47·20
1. 9.	13·80	7·88	22·12	7·08	13·50	22·39
1.10.	30·66	32·49	34·24	34·62	32·62	26·23
1.11.	56·10	40·75	57·87	50·54	−1·00	59·15
1.12.	33·97	40·54	25·86	37·16	−1·00	29·84
1.13.	15·39	12·04	16·46	12·40	−1·00	24·47
1.14.	4·55	13·42	5·35	8·41	−1·00	6·46
1.15.	18·56	22·62	24·95	29·20	−1·00	17·83
2. 0.	2·27	30·00	20·91	15·21	13·10	25·39
2. 1.	2·45	2·39	4·24	−1·00	22·00	36·31
2. 2.	19·85	2·62	7·37	−1·00	36·03	35·99
2. 3.	30·41	38·34	34·64	38·48	27·29	45·42
2. 4.	141·55	125·55	144·13	180·31	163·83	154·04
2. 5.	16·75	4·76	17·27	5·45	6·46	9·90
2. 6.	46·45	40·86	40·50	43·11	46·24	56·46
2. 7.	16·97	6·88	16·26	7·59	16·80	15·40
2. 8.	10·34	7·43	12·32	6·81	13·21	7·22
2. 9.	48·58	41·10	35·75	39·63	49·51	46·84
2.10.	19·13	30·19	14·44	26·77	19·90	20·24
2.11.	19·01	20·15	23·84	22·49	−1·00	8·68
2.12.	20·08	7·76	23·53	11·59	−1·00	16·35
2.13.	54·42	55·81	46·86	62·09	−1·00	55·82
2.14.	52·90	57·40	53·93	63·89	−1·00	53·32
2.15.	21·20	25·02	24·14	28·65	−1·00	22·77
3. 0.	15·12	8·10	23·13	−1·00	10·83	13·76
3. 1.	28·50	26·51	34·24	28·31	17·38	17·39
3. 2.	15·03	36·25	18·58	38·73	14·49	14·98
3. 3.	62·27	66·04	62·82	71·48	60·94	77·40
3. 4.	14·23	31·42	15·45	32·80	2·97	27·05
3. 5.	30·26	14·08	34·24	19·28	22·87	17·28
3. 6.	11·76	22·54	15·35	21·89	15·66	30·33
3. 7.	43·49	25·12	33·23	25·18	40·49	41·52
3. 8.	8·16	8·98	5·15	6·86	4·20	8·88
3. 9.	13·87	27·02	8·08	19·54	16·64	13·22
3.10.	4·67	12·61	10·20	7·62	4·68	6·07
3.11.	6·04	9·00	5·55	7·98	−1·00	6·27
3.12.	40·49	28·75	33·53	31·13	−1·00	40·91
3.13.	32·05	27·26	31·92	29·86	−1·00	25·92
3.14.	53·73	44·14	58·78	55·66	−1·00	59·66
3.15.	8·37	9·74	19·90	32·35	−1·00	22·27
4. 0.	43·94	48·85	44·74	54·96	37·68	42·57
4. 1.	39·57	25·42	23·23	24·34	41·25	39·26
4. 2.	14·58	24·66	16·26	22·80	2·83	23·02
4. 3.	48·88	42·91	48·18	47·68	47·25	39·48
4. 4.	15·70	13·80	4·44	11·46	10·34	5·55
4. 5.	7·49	10·78	9·80	7·09	9·01	15·07
4. 6.	5·31	10·05	14·44	6·28	17·00	10·16
4. 7.	27·47	22·78	20·60	23·29	29·99	22·94
4. 8.	38·01	28·52	30·10	28·10	34·50	30·88
4. 9.	32·58	27·24	25·65	26·94	39·31	36·12
4.10.	5·71	14·90	9·09	7·83	4·90	9·78
4.11.	11·74	13·96	5·55	9·67	−1·00	7·58
4.12.	4·73	9·52	5·55	8·31	−1·00	10·30
4.13.	16·49	16·61	20·30	17·18	−1·00	20·09
4.14.	19·70	13·80	24·74	16·84	−1·00	16·76
4.15.	8·45	7·35	−1·00	12·79	−1·00	−1·00
5. 0.	76·63	44·71	58·38	52·35	78·71	68·11
5. 1.	58·91	54·83	54·03	63·66	65·93	56·70
5. 2.	67·00	38·34	74·13	50·09	71·91	64·44
5. 3.	32·26	32·97	25·86	32·94	36·18	28·21
5. 4.	8·45	16·81	5·86	11·08	4·39	5·79
5. 5.	37·78	46·57	44·14	47·34	40·38	43·19
5. 6.	30·87	36·45	27·67	37·98	28·50	28·75
5. 7.	17·54	28·71	8·99	19·66	16·96	19·41
5. 8.	35·81	29·50	27·98	28·58	31·11	21·06
5. 9.	47·96	28·86	46·36	35·28	52·83	42·62
5.10.	36·03	38·27	38·68	39·07	38·68	39·06
5.11.	22·67	29·37	19·80	25·50	−1·00	8·38
5.12.	6·02	12·61	11·41	8·41	−1·00	6·55
5.13.	28·48	26·24	28·28	29·72	−1·00	35·32
5.14.	20·36	16·93	11·51	18·19	−1·00	6·11
5.15.	39·75	−1·00	−1·00	−1·00	−1·00	−1·00
6. 0.	20·49	22·87	18·79	22·66	16·65	13·93
6. 1.	4·71	6·56	16·36	5·77	9·16	8·86
6. 2.	5·40	6·36	4·65	5·92	3·43	5·79
6. 3.	8·01	5·49	4·75	6·09	10·88	14·91
6. 4.	10·27	18·45	5·86	12·95	7·77	6·24
6. 5.	28·78	22·73	27·47	21·57	27·62	24·70
6. 6.	19·37	7·97	11·21	6·85	21·20	20·20
6. 7.	32·18	23·38	31·51	24·05	31·88	28·00
6. 8.	14·03	17·44	14·64	15·06	4·60	16·02
6. 9.	5·36	8·80	5·55	7·84	4·91	6·21
6.10.	7·85	7·59	9·90	8·05	6·58	6·29
6.11.	34·10	35·42	39·29	39·87	−1·00	44·92
6.12.	44·91	44·42	48·28	48·40	−1·00	43·89
6.13.	9·66	9·86	10·40	8·34	−1·00	11·75
6.14.	35·30	29·93	35·96	26·44	−1·00	30·57
7. 0.	21·16	35·18	16·36	29·96	16·76	26·00
7. 1.	25·33	30·06	16·06	28·30	26·11	35·56
7. 2.	8·44	24·65	13·84	24·62	4·74	5·13
7. 3.	5·83	5·86	4·95	6·46	3·82	5·86
7. 4.	22·95	35·02	21·31	34·33	25·63	33·19
7. 5.	13·72	6·41	12·02	7·00	13·89	13·79
7. 6.	4·60	8·04	7·27	7·24	7·45	5·87
7. 7.	4·68	10·69	5·45	7·51	4·58	11·74
7. 8.	33·48	34·66	38·18	36·03	33·36	38·73
7. 9.	22·19	18·15	17·78	13·21	27·13	22·20
7.10.	23·36	24·45	23·43	23·67	25·29	20·45
7.11.	16·43	19·24	12·32	13·01	−1·00	14·96
7.12.	36·38	32·77	32·12	32·99	−1·00	21·87
7.13.	10·01	14·91	16·16	8·16	−1·00	17·28
7.14.	5·29	−1·00	−1·00	−1·00	−1·00	−1·00
8. 0.	35·73	23·70	41·01	27·37	35·17	32·66
8. 1.	16·86	20·72	15·76	17·70	19·50	8·43
8. 2.	23·11	10·96	9·29	6·81	22·78	27·77
8. 3.	8·95	15·86	8·58	8·34	4·20	13·68
8. 4.	27·17	15·32	32·12	18·73	19·64	19·18
8. 5.	26·97	30·64	21·71	30·78	28·74	34·79
8. 6.	11·26	6·98	11·82	7·53	12·02	6·03
8. 7.	55·27	60·85	61·71	64·55	62·55	68·44
8. 8.	14·08	18·90	5·55	16·35	15·52	8·70
8. 9.	21·30	26·23	19·69	22·36	18·80	12·77
8.10.	31·15	27·89	32·02	32·35	35·94	27·95
8.11.	16·06	18·26	17·17	17·80	−1·00	13·58
8.12.	10·62	8·02	6·56	8·21	−1·00	11·89
8.13.	50·67	44·33	44·74	39·96	−1·00	−1·00
9. 0.	4·51	11·36	6·36	7·07	4·27	5·65
9. 1.	21·43	29·13	33·53	34·98	23·36	25·27
9. 2.	37·04	24·38	34·64	27·33	40·89	33·45
9. 3.	19·85	13·82	29·19	12·93	21·01	13·43
9. 4.	14·64	20·94	11·51	13·47	12·88	21·90
9. 5.	36·62	27·30	32·32	27·83	35·86	28·62
9. 6.	42·54	40·51	50·30	45·50	45·85	47·68
9. 7.	4·75	7·51	5·55	7·98	5·05	5·74
9. 8.	14·31	16·84	21·31	12·88	13·28	21·04
9. 9.	17·82	17·80	13·03	8·33	17·05	9·67
9.10.	18·47	18·28	17·88	14·91	16·96	21·24
9.11.	6·09	9·81	9·09	8·31	−1·00	6·37
9.12.	4·19	7·96	4·95	8·07	−1·00	6·13
9.13.	−1·00	−1·00	−1·00	−1·00	−1·00	−1·00
10. 0.	10·31	9·82	5·45	12·13	4·60	6·66
10. 1.	4·68	8·42	6·16	7·54	4·61	6·04
10. 2.	8·31	26·03	6·97	16·82	8·04	15·44
10. 3.	4·67	7·09	5·45	7·62	4·68	6·07
10. 4.	39·17	19·29	29·49	23·36	43·23	32·92
10. 5.	13·22	17·99	18·68	18·45	8·57	12·38
10. 6.	29·08	21·80	25·65	13·31	31·15	17·60
10. 7.	8·19	9·33	21·71	8·19	9·79	14·75
10. 8.	45·26	38·05	48·28	40·85	55·86	47·53
10. 9.	12·28	20·07	12·52	14·60	8·88	11·89
10.10.	28·81	21·87	27·77	23·90	35·21	25·88
10.11.	21·90	20·96	25·96	18·02	−1·00	33·50
10.12.	49·59	−1·00	−1·00	22·13	−1·00	−1·00
11. 0.	20·74	20·94	20·40	15·06	−1·00	19·11
11. 1.	15·04	24·75	5·55	20·24	−1·00	14·05
11. 2.	9·42	12·58	12·83	7·93	−1·00	6·99
11. 3.	15·63	22·25	24·24	26·56	−1·00	22·09
11. 4.	4·75	11·33	5·55	8·06	−1·00	6·30
11. 5.	31·73	21·78	31·92	21·35	−1·00	27·50
11. 6.	54·51	48·47	53·83	53·71	−1·00	64·62
11. 7.	35·00	38·62	27·37	36·89	−1·00	26·22
11. 8.	8·33	8·15	5·45	8·43	−1·00	6·51
11. 9.	17·32	18·21	22·83	22·05	−1·00	21·89
11.10.	5·73	11·92	4·95	8·03	−1·00	6·11
11.11.	22·20	−1·00	−1·00	15·00	−1·00	−1·00
12. 0.	10·86	23·24	9·59	20·40	−1·00	10·51
12. 1.	4·74	9·26	5·55	8·14	−1·00	15·67
12. 2.	21·03	35·68	23·53	35·53	−1·00	26·25
12. 3.	28·08	21·92	25·15	24·69	−1·00	30·66
12. 4.	27·72	9·52	29·39	11·78	−1·00	23·86
12. 5.	27·52	29·46	33·23	30·73	−1·00	41·15
12. 6.	29·90	19·44	31·92	18·31	−1·00	22·99
12. 7.	4·55	12·09	5·35	8·37	−1·00	6·44
12. 8.	36·24	30·85	25·75	26·41	−1·00	26·81
12. 9.	31·73	29·07	31·11	30·48	−1·00	30·31
12.10.	23·39	−1·00	−1·00	7·43	−1·00	−1·00
13. 0.	7·74	12·90	8·58	15·13	−1·00	15·72
13. 1.	4·73	12·04	5·55	8·42	−1·00	7·21
13. 2.	33·19	28·64	24·44	21·82	−1·00	35·47
13. 3.	19·85	9·69	20·00	11·91	−1·00	16·81
13. 4.	29·97	35·47	21·92	37·01	−1·00	30·58
13. 5.	66·06	55·99	63·93	62·28	−1·00	61·31
13. 6.	22·97	15·35	18·48	15·24	−1·00	22·50
13. 7.	12·07	8·01	5·05	8·16	−1·00	6·21
13. 8.	4·02	13·35	4·75	7·83	−1·00	5·94
13. 9.	−1·00	−1·00	−1·00	−1·00	−1·00	−1·00
14. 0.	27·93	33·72	23·63	33·92	−1·00	33·37
14. 1.	16·64	20·99	18·28	19·84	−1·00	11·13
14. 2.	25·64	12·72	22·02	13·86	−1·00	17·85
14. 3.	4·46	8·14	15·25	8·34	−1·00	8·06
14. 4.	4·37	13·16	5·15	8·26	−1·00	6·30
14. 5.	4·19	7·90	4·95	8·19	−1·00	6·11
14. 6.	23·33	7·73	17·98	12·31	−1·00	21·79
14. 7.	17·33	−1·00	−1·00	−1·00	−1·00	−1·00
15. 0.	12·32	19·32	9·59	18·78	−1·00	12·26
15. 1.	5·64	11·71	11·92	−1·00	−1·00	12·38
15. 2.	9·06	14·72	4·75	16·23	−1·00	14·91
15. 3.	12·05	9·74	10·50	11·17	−1·00	−1·00
15. 4.	22·97	13·54	34·24	30·13	−1·00	−1·00
15. 5.	21·50	−1·00	−1·00	−1·00	−1·00	−1·00

Table 2. *Unrefined heavy atom parameters for cytochrome derivatives*

Derivative	*x*	*y*	*B*	*A*
Pt30w	0·220	0·200	7·46	8·15
Hg14w	0·020	0·400	8·63	5·80
Pd	0·150	0·030	7·75	4·45
	0·090	0·300	7·75	4·45
Erronium	0·280	0·280	7·58	5·90

Of the two general classes of derivative tests, real space tests (Patterson, Fourier and convolution methods) and reciprocal space tests (figures of merit, *R* factors and sign or phase agreement), this paper is concerned chiefly with the former. Reciprocal space criteria will be mentioned in passing where relevant. The most widely used such criterion is the 'figure of merit', *m*, an attenuation factor ranging between 0 and 1 which is both the weighting factor to the *F*'s or ΔF's which insures the smallest r.m.s. error in electron density throughout a map (Blow & Crick, 1959), and the mean value of the cosine of the error in phase angle for the reflection in question (Dickerson *et al.*, 1961*b*).

A second widely used criterion is the Kraut *R* factor (Kraut *et al.*, 1962), defined by:

$$R_K \equiv \frac{\sum_{hkl} |\varepsilon_{hkl}|}{\sum_{hkl} |F_{PH}|}$$

where ε_{hkl} is the lack-of-closure error of the phase triangle of the derivative in question using the protein phase angle determined by all the derivatives together:

$$|\varepsilon_{hkl}| \equiv \big||\mathbf{F}_{PH}| - |\mathbf{F}_P + \mathbf{f}_H|\big| .$$

A third criterion is the 'centric *R* factor' of Cullis, Muirhead, Perutz, Rossmann & North (1961):

$$R_C \equiv \frac{\sum_{hkl} \big||\Delta F| - |f_H|\big|}{\sum_{hkl} |\Delta F|}$$

where:

$$|\Delta F| = \big||F_{PH}| - |F_P|\big| .$$

In the centrosymmetric zone, R_K and R_C differ only by a normalizing denominator:

$$R_K = R_C \cdot \frac{\sum_{hkl} |\Delta F|}{\sum_{hkl} |F_{PH}|}.$$

Kraut R_K values were calculated in the work to follow. These can be converted roughly to R_C's by multiplying by 3·3 for Pt, 5·0 for Hg, and 3·6 for the double derivative.

The 'four-data' check sign set was the result of the combination of the Pt30w data, another 1:1 Pt/cytochrome mole ratio set photographed at six weeks (Pt6w), and Hg photographed at one week (Hglw) and again at fourteen weeks (Hg14w). Scale factors, x and y, coordinates and degrees of substitution, A, were refined until no further significant changes occurred. The duplicate heavy atoms refined to the same sites to within 0·05 Å but differed in degree of substitution. Figures of merit and Kraut R_K values are listed in Table 4. The 'five-data' set of Table 4 consists of the above four and the double derivative.

Difference Fourier maps were calculated with individual terms both unweighted and weighted with the figure of merit for each reflection. No practical advantage was seen in the weighted maps in terms of added information or detail; moreover these maps had the disadvantage of having peak heights reduced by a factor of around 0·6 (the mean figure of merit). All maps shown in this paper are unweighted.

A complete series of cross difference maps was calculated, using parent protein signs determined from one derivative and ΔF values from a different derivative. In this context, the expression 'mercury signs' is to be interpreted as meaning the signs of the parent protein obtained by using the mercury derivative data alone, and not the signs of the mercury atom contribution itself. Crossover terms, where F_{PH} and F_P have opposite signs and where the heavy atom contribution

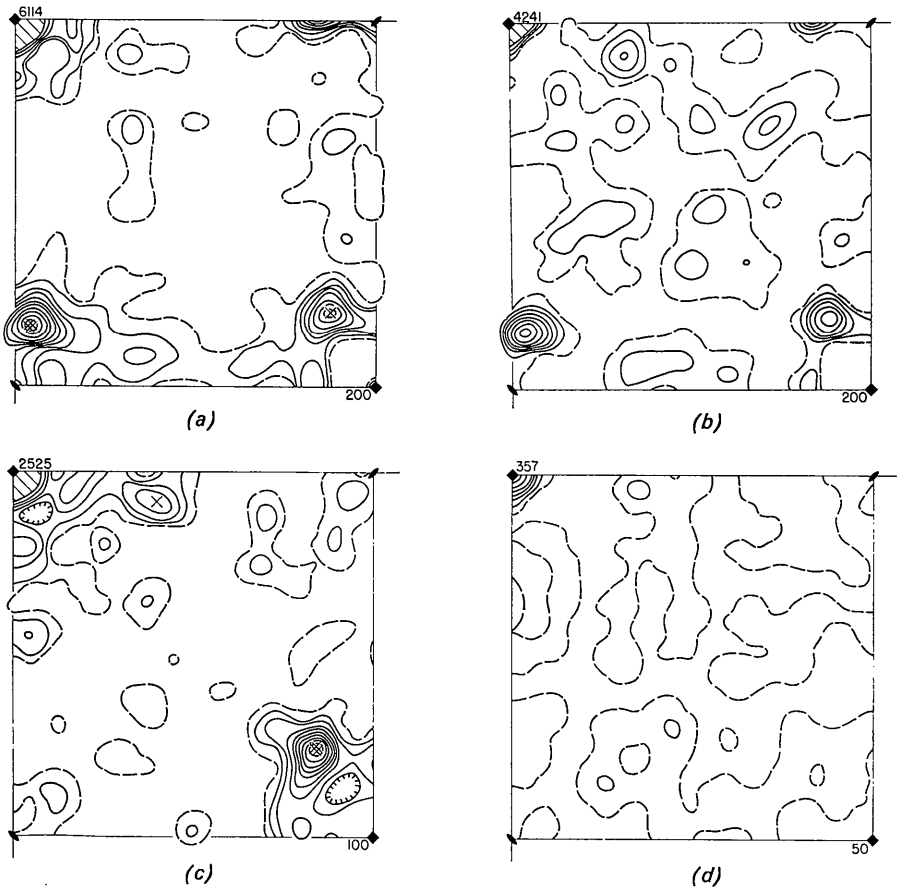


Fig. 1. $(\Delta F)^2$ difference Patterson maps of horse heart cytochrome C in $hk0$ projection at 3 Å resolution. All maps are to the same arbitrary scale. Contour intervals are marked in the lower right corner and the height of the origin peak in the upper left. The zero contour is dashed. Coordinates u (horizontal) and v (vertical) run from 0 to $\frac{1}{2}$, with the origin in the upper left corner. (*a*) $PtCl_4^{2-}$ derivative (Pt30w). Interpreted as arising from a single Pt site per molecule. The expected single peak is marked by × and the double peak by #. (*b*) $PtCl_4^{2-}$ map as in (*a*) but with the innermost 20 of the 180 reflections omitted. Note the essential similarity to (*a*). (*c*) Mersalyl derivative (Hg14w). Single site, peaks indicated as in (*a*). (*d*) 'Null Patterson' map comparing two unrelated sets of parent cytochrome data to give some idea of the experimental noise level.

is equal in magnitude to the *sum* of the magnitudes of F_{PH} and F_P, were rare and unimportant, and were neglected in the difference maps.

Valid derivatives: platinum and mercury

Fig. 1 shows the $(\Delta F)^2$ difference Patterson projections down the c axis for Pt30w, Hg14w and for one set of parent cytochrome data against another independent parent set. This last map, Fig. 1(*d*), provides an estimate of the amount of background noise to be expected from errors in intensity data, and suggests that features over 50 to 100 in the other maps are probably to be ascribed to genuine structural changes between parent and derivative and not to experimental data error. The double peak at $(x+y,y-x)$ is plain in both Pt and Hg maps, as is the single peak at $(2x,2y)$. Initial mean figures of merit (Table 4) from phase determination were around 0·4 for each of the derivatives taken alone, and 0·46 for the combination.

ΔF difference Fourier maps of three types for Pt and Hg are shown in Figs. 2 and 3: self-sign maps using signs determined by the derivative in question [Figs. 2(*a*) and 3(*a*)], cross-sign maps using signs determined by the *other* derivative [Figs. 2(*b*) and 3(*b*)], and check maps using refined four-data signs [Figs. 2(*c*) and 3(*c*)]. The self-sign maps give the best reproductions of the assumed peaks, of course, but prove nothing. Feedback of the initial assumptions in such cases is at its strongest, and any peak chosen at random for sign determination would have reappeared on the map and given the appearance of being 'confirmed' (see Fig. 9(*a*), for example). The real confirmations of the two derivatives are their cross-sign maps. There is no reason for a peak to show up where it does in Fig. 2(*b*), for example, unless: (*a*) the ΔF's are meaningful because of the correctness of the Pt derivative, and (*b*) the signs are meaningful because of the correctness of the Hg derivative. The self-sign Pt map, because of its bias in favor of a single-site Pt model, actually obscures information. The improvement in R_K for Pt30w in Table 4 from 17·5 to 13·4% is mainly a result of the addition of a half-weight secondary site immediately to the right of the principal site [refined position at × in Fig. 2(*a*)]. The asymmetry of the Pt peak is most pronounced in the cross-sign map, and the secondary site is less marked relative to the primary site in just that map whose signs are most heavily influenced by a one-site model.

The maps obtained using the two-derivative sign set and the refined four-data sign set (Figs. 2(*c*) and 3(*c*), and Table 3, *A*) offer no proof of derivative correctness beyond that of the cross-sign maps in spite of their improved appearance. This improved appearance arises from the introduction into the sign analysis of just those sites whose validity is being tested, plus the further advantage of least-squares fitting of the assumed models to the data. It is no surprise that they look superficially better. It is true that the least-squares program will often refine the effective substitution number of an incorrect site to zero (see erronium trials, below, and also Dickerson & Palmer, 1967). But it cannot add an omitted site, and the use of refined full-derivative signs clouds the proof by interrelating what would otherwise be two totally independent demonstrations of correctness.

The only ΔF difference maps which are really worth anything in proving out derivatives, in summary, are the cross-sign maps in which the sign set used has no dependence upon the parameters of the derivative whose validity is being tested. Self-sign maps are utterly worthless except as journal illustrations.

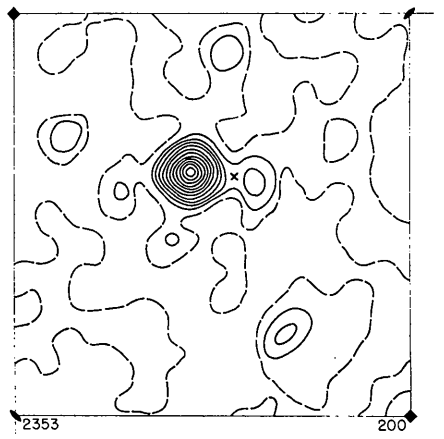
An example of an incorrect interpretation: $PdCl_4^{2-}$

Fig. 4(*a*) shows the difference Patterson map between crystals in 4·3 *M* mixed phosphate buffer and similar crystals with $PdCl_4^{2-}$ added. A study of the effect of medium on $|\Delta F_{hk0}|$ in ammonium sulfate and in 4·3 and 5·0 *M* phosphate buffer showed that, although several of the low-order terms changed in intensity, only the 100 and 110 reflections actually changed sign. This change of sign was accounted for in all the ΔF maps to follow. The 'interpretation' of the Pd map will illustrate some of the potential hazards of the process.

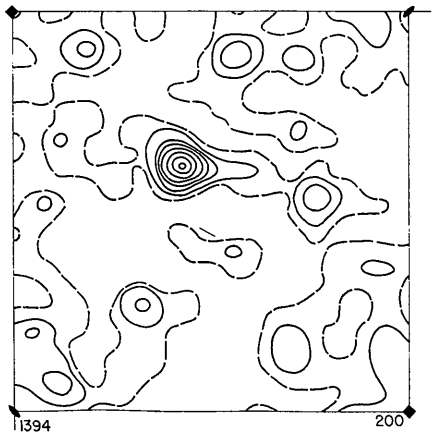
Although this Pd map is noisier than either of the Pt or Hg maps, it can be fitted by a two-site model. These two sites, marked *A* and *B* in Fig. 5, account for the peaks in Fig. 4(*a*) which are marked by ×, #, or #. Only the two satellite peaks below and to the upper right of the very large peak near the origin are unexplained. If the data are cut off at 6 Å resolution, then these peaks merge and the interpretation is even more convincing. At 4 Å the peak heights are very nearly correct; the highest peak is interpreted as the near-superposition of two double peaks, the next two as near-superpositions of a double and a single, and the next two as double sites. With due allowance for origin diffraction ripple, experimental error particularly in low-order reflections, and a certain degree of non-isomorphism, it is easy to be persuaded that the interpretation is right.

The Pd self-sign difference Fourier map [Fig. 6(*a*)] supports this conclusion and even suggests a minor third site, *C*. With two exceptions, all ten of the new *C–C*, *C–A* and *C–B* vectors required do show up on or near features of the Patterson map. Nevertheless, this interpretation is wrong from start to finish, and it is important to see what kind of warning signs reveal this.

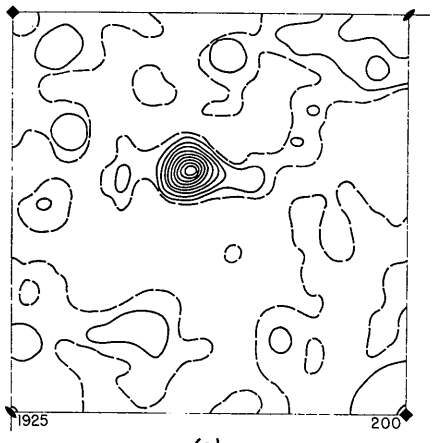
The first sign of trouble is the low figure of merit for the Pd derivative taken alone, 0·3 as compared with 0·4 for Pt or Hg. Admittedly, the absolute value of the figure of merit is a shaky measure of correctness. Experiments with cytochrome C and with triclinic hen egg-white lysozyme suggest that the mean figure of merit rises sharply with (*a*) the number of derivatives, (*b*) the number of sites per derivative, (*c*) the number



(a)

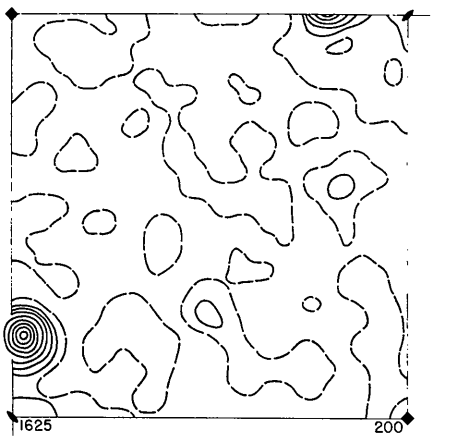


(b)

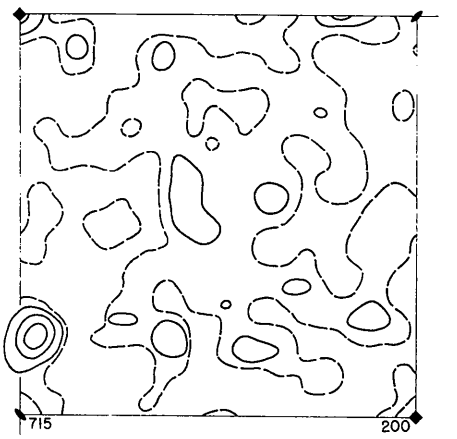


(c)

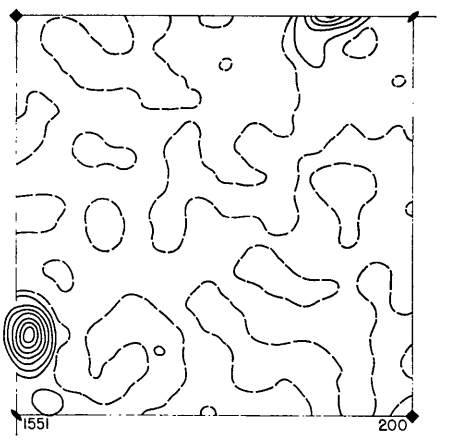
Fig. 2. ΔF Difference electron density projections using platinum ΔF's and the sign sets indicated. These maps and all ΔF maps to follow are on the same arbitrary scale. The contour interval in each map is given in the lower right corner and the height of the main peak in the lower left. Zero contours are dashed. Coordinates x (horizontal) and y (vertical) run from 0 to $\frac{1}{2}$, with the origin in the upper left corner. All maps are unweighted. (*a*) Pt signs. (*b*) Hg signs. (*c*) 4-Data signs.



(a)



(b)



(c)

Fig. 3. ΔF difference electron density projections using mercury ΔF's and the sign sets indicated. Same conventions and scale as Fig. 2. (*a*) Hg signs. (*b*) Pt signs. (*c*) 4-Data signs.

of parameters adjusted by least squares per site (such as anisotropic temperature factors), and (*d*) the extent to which the r.m.s. errors in the derivatives are underestimated. As an example, the mean figure of merit for a three-dimensional phase analysis at 6 Å of triclinic hen egg-white lysozyme using three derivatives with 5, 6, and 6 sites was 0·83 (Dickerson & Steinrauf, unpublished), yet the analysis is now believed to have been at least partially in error. In a variant of this particular test, when the same lysozyme data were used, with the same number of sites in the three derivatives, but with atomic coordinates chosen from a table of random numbers, the mean figure of merit was 0·69. But in spite of the problems in interpreting absolute values of *m*, the *relative* values and changes in mean

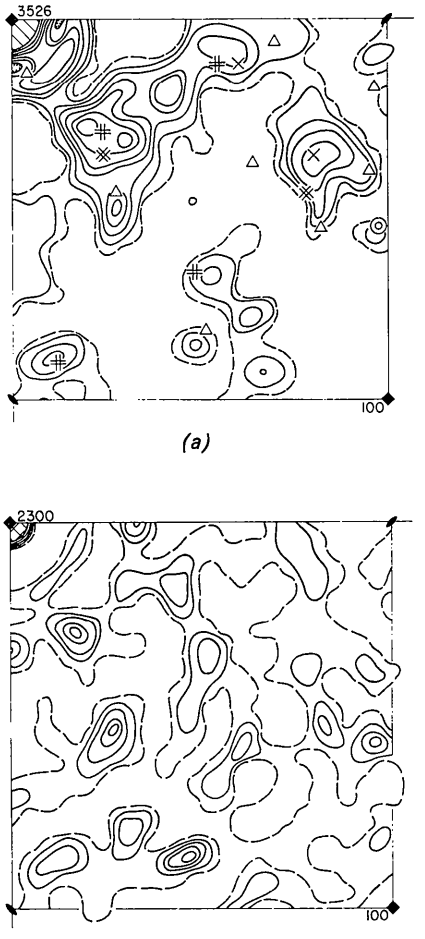


Fig. 4. $(\Delta F)^2$ Difference Patterson maps of the $PdCl_4^{2-}$ derivative in 4·3 *M* phosphate buffer against parent crystals in 4·3 *M* phosphate buffer. Same conventions and scale as Fig. 1. (*a*) Full 4 Å data. Expected peaks are marked as follows:

× # Single and double peaks from sites *A* and *B* taken individually as shown in Fig. 5.
≠ Cross vector peaks between sites *A* and *B*.
○ ◎ Single and double peaks from site *C* of Fig. 5.
△ Cross vector peaks between site *C* and site *A* or *B*.

(*b*) Same as (*a*), but with the innermost 20 reflections omitted. Compare the destruction of features of (*a*) with the analogous Pt maps [Fig. 1(*a*) and 1(*b*)].

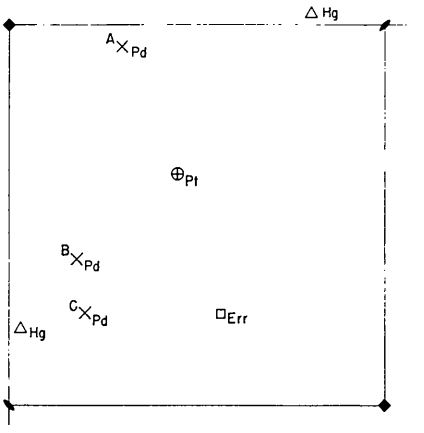


Fig. 5. Heavy atom sites. Same axis conventions as for ΔF maps. Sites are: ⊕ Pt; △ Hg; × Pd (3 sites); □ Erronium.

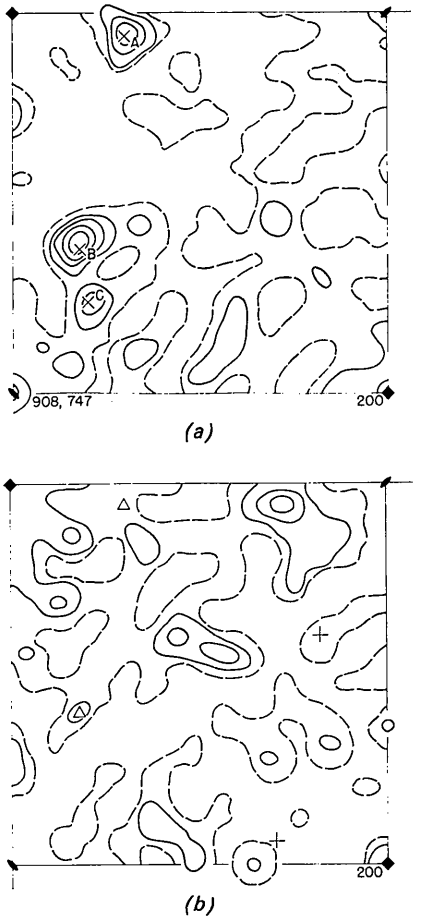


Fig. 6. ΔF Difference electron density projections using Pd ΔF's and the sign sets indicated. Same conventions as Fig. 2. (*a*) Pd signs. (*b*) 4-Data signs. The two principal sites are marked by △, and their symmetry-related alternates by +.

figure of merit during the course of analysis are useful in differentiating between good and bad derivatives under controlled conditions.

The clearest indication of trouble comes from the ΔF maps. The Pd difference map with four-data signs [Fig. 6(*b*)] fails to reproduce the Pd sites, either as originally chosen or with a shift of origin by $(\frac{1}{2},\frac{1}{2},0)$. Instead, 'ghost' Pt and Hg peaks appear, reflecting the fact that, although the sign set is a reasonably correct parent protein set, it has been biased in favor of the Pt and Hg sites used to obtain it. In the corresponding map using Pt signs alone, the Pt ghost peak becomes stronger at the expense of the Hg, and with Hg signs the reverse is true. The Pt and Hg difference maps using Pd signs (Fig. 7) fail to reproduce the Pt and Hg sites, and the Pt map appears to have ghost Pd peaks as well.

In addition to these direct space tests, a test in reciprocal space of agreement between $hk0$ signs determined by Pd and those from Pt, Hg, or any combination of them, demonstrates the incompatibility of the Pd set. A symmetry-permissible shift of origin in Pd by $(\frac{1}{2},\frac{1}{2},0)$, resulting in the reversal of sign of all reflections with $h+k$ odd, is of no help, and the conclusion must be drawn that either the Pd or the Pt and Hg taken together are wrong.

The Pd trials illustrate one of the pitfalls of interpreting difference Patterson maps, namely the ability to fit anything with enough concentrated effort. With two assumed sites per molecule, in space group $P4_1$, there will be six double peaks and two single peaks in the quadrant shown. With three sites per molecule, there will be three single peaks and fifteen double peaks. With such a great number of peaks available and with the low order 'checkerboarding' effect discussed below, some kind of a fit or another is assured.

At low resolution, and especially in projection, errors in a handful of low-order reflections can introduce serious error ripples into the difference Patterson map. If only one or two terms are involved, the erring reflections can be identified, but if several are acting in concert, it may not be apparent that low-order error fringes are the cause of the observed peaks. Errors in

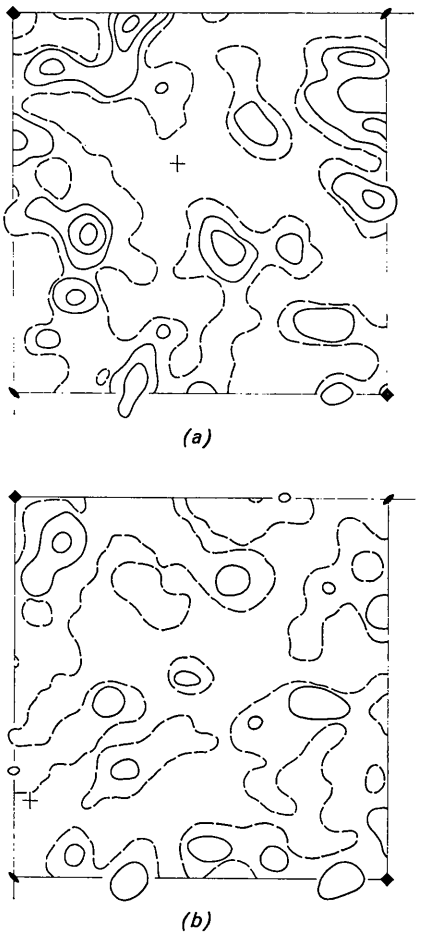


Fig. 7. (*a*) Pt and (*b*) Hg ΔF difference maps with Pd signs. Expected heavy atom sites are marked with a +.

Table 3. *Peak heights in ΔF difference electron density maps*

A. Pt30w difference maps

Signs:	Pt	Hg	Pt+Hg	4-Data	Pd	Err
Pt peak heights:	2353	1394	2067	1925	(−700)†	(0)†
Heights with Err*:	2139	1231	2134	—	—	—

B. Hg14w difference maps

Signs:	Pt	Hg	Pt+Hg	4-Data	Pd	Err
Hg peak heights:	715	1625	1553	1551	(−50)†	(142)†
Heights with Err*:	919	1425	1423	—	—	—

C. Erronium difference maps

Signs:	Err	Err+Pt	Err+Hg	Err, Pt, Hg	4-Data
Err peak heights:	1234	700	677	(345)†	(−16)†

D. Double derivative difference maps

	Unweighted maps			Weighted with figure of merit		
Signs:	Pt	4-Data	Hg	Pt	4-Data	Hg
Pt peak heights:	1711	1642	1232	1144	1367	643
Hg peak heights:	283	722	803	167	603	562
Pt/Hg ratio:	6·05	2·28	1·54	6·85	2·27	1·14

* Peak heights obtained from sign sets after erronium is added to the analysis.
† No longer the highest feature on the difference map.

somewhat higher order terms in tetragonal space groups lead to the grid pattern of fringes which Kraut calls 'checkerboarding'.

If it proves possible to fit an eight-atom model – two per molecule – to $8\times7=56$ peaks in a three-dimensional difference Patterson map, the temptation to accept it as valid is strong. But the danger lies in the fact that the nodes or intersections of a set of low-order error fringes *themselves* will form a vector set. In the presence of such features, one will automatically be able to find four, or six, or eight atomic sites which will explain peaks which are really nothing more than low-order fringe nodes. However, an atomic site set obtained in such a way will itself tend to show the regularity of the error nodes, with sites falling in parallelograms or on a small number of parallel lines. This may be the proper explanation of the difference Patterson map of the HgI_4^{2-} derivative of triclinic lysozyme shown in Fig. 10 of Dickerson (1964).

The simplest way to see if the inner reflections are producing fake detail is to leave them out and see what happens to the map. The Pt and Pd difference Patterson maps of Figs. 1(*b*) and 4(*b*) differ from those of Figs. 1(*a*) and 4(*a*) only in having the innermost 20 of the 180 reflections removed – all those within the 11·5 Å limit. The features of the Pd map are now obliterated, leaving nothing but noise. The Pt map is virtually unaffected. As the null Patterson of Fig. 1(*d*) indicates that the features of the Pd map are higher than the experimental noise level, the peaks of Fig. 4(*b*) most probably arise from extensive non-specific or at least multi-specific binding to the protein, and the derivative is useless.

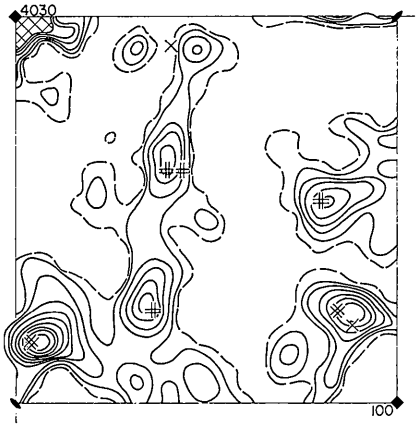


Fig. 8. $(\Delta F)_2{}^2$ Patterson map of the double derivative with Pt and Hg diffused in simultaneously. Pt–Pt and Hg–Hg single and double vectors are indicated by × and #, Pt–Hg double weight cross vectors by #.

A valuable control: the double derivative

The best check on the validity of two individual derivatives is the simultaneous introduction of both sites into the same crystal. Now, in addition to the self peaks, there must appear the *cross* peaks in the difference Patterson map. It is difficult to see how a fortuitous combination of error ripples of the type discussed in the previous section could produce just the required peaks in the individual maps, and then these peaks plus just the right extra peaks in the double derivative Patterson map. Fig. 8 shows such a difference Patterson map for the simultaneous Pt/Hg double derivative. The cross peaks (#) show up exactly where they are predicted to be, and the assumed model explains all of the highest features of the map.

An example of the greater sensitivity of the weighted maps to feedback errors is provided by the Pt/Hg peak height ratios in difference Fourier maps of the double derivative using sign sets from various sources (Table 3, *D*). When the sign set has been obtained from a single derivative corresponding to one of the two sites, that site is emphasized in the difference map at the expense of the other, leading to peak ratios in the unweighted maps ranging between 6·05 and 1·54. The final refined value was 2·8, close to that produced by the four-data sign set. The weighted maps, although producing the same nearly correct value with the four-data signs, exaggerate the emphasis on the sign-determining site when single-derivative signs are used.

Table 4. *Reciprocal space test of derivative quality*

		R_k					
Derivatives	*m*	Pt6w	Pt30w	Hg1w	Hg14w	Double	Other
Pt30w, unrefined	0·370		17·0				
Hg14w, unrefined	0·396				10·6		
Pt+Hg, unrefined	0·459		19·0		12·4		
Four-data, refined	0·575	15·5	17·5	11·3	11·1		
Double derivative, unrefined	0·328					17·0	
Five-data, refined, with secondary Pt site	0·641	12·6	13·4	11·1	10·7	13·8	
Palladium unrefined	0·291						15·3
Erronium, unrefined	0·272						17·4
Err+Pt	0·360		19·2				22·9
Err+Hg	0·361				12·3		23·3
Err+Pt+Hg, unrefined	0·430		19·9		12·9		26·5
Err+Pt+Hg, refined	0·485		16·9		10·8		21·0

The effect of a wrong derivative on subsequent analysis: erronium

The checks discussed above should ordinarily be sufficient to exclude a wrong interpretation. But what would be the effect of adding a totally erroneous derivative to the sign analysis? How serious would the feedback problem be, and at what point would the mistake be discovered?

To answer this question, such a site was selected and was used with the Pd data with the four innermost salt-sensitive reflections discarded. The site selected (Fig. 5) was chosen so as to be removed from any real site, not to be on any obvious low-order fringes which included a real site, and not to be compatible with the Pd difference Patterson map. This wrong site with its now pseudo-random ΔF data will be referred to as 'erronium', Err. Structure factor calculations, Wilson plots for A and B, and sign determinations were carried out for Err just as for either of the true derivatives, and the parameters obtained are given in Table 2. The Wilson plot was more ragged than usual, but not disturbingly so.

The self-sign map [Fig. 9(*a*)] is reassuring; the peak is almost as prominent as is the Hg peak in its own self-sign map. But the addition of either Pt or Hg to the sign analysis halves the Err peak, and the addition of both good derivatives wipes it out altogether [Fig. 9(*b*)–(*d*)]. In contrast, going from either a Pt or a Hg self-sign map to the two-derivative Pt+Hg map causes only a 5–10% drop in peak height. The map with Err (or Pd) ΔF's and the four-data signs [Fig. 6(*b*)] shows no trace of the Err peak.

The Pt and Hg cross-sign maps using Err signs show no trace of the expected peaks (Fig. 10). But the addition of Pt to Err in the sign analysis brings in the resultant Pt peak in the difference map almost as strongly as with Pt signs alone. In fact, each of the Pt maps with signs of Pt+Err, Hg+Err and Pt+Hg+Err is nearly as good as the corresponding map without Err (Table 3, *A*, *B*). Exactly the same behavior is observed with Hg ΔF maps and the corresponding sign sets. The Err contribution is so meaningless that to a good approximation the sign analysis behaves as if the Err is simply not there. It only serves as a perturbing source of random error which lowers the quality of the maps and raises background noise. The relatively small drop in mean figure of merit upon addition of erronium (Table 4) corroborates the unsystematic nature of its contribution.

The refinement program provides the most dramatic evidence of the incorrectness of erronium. Before re-

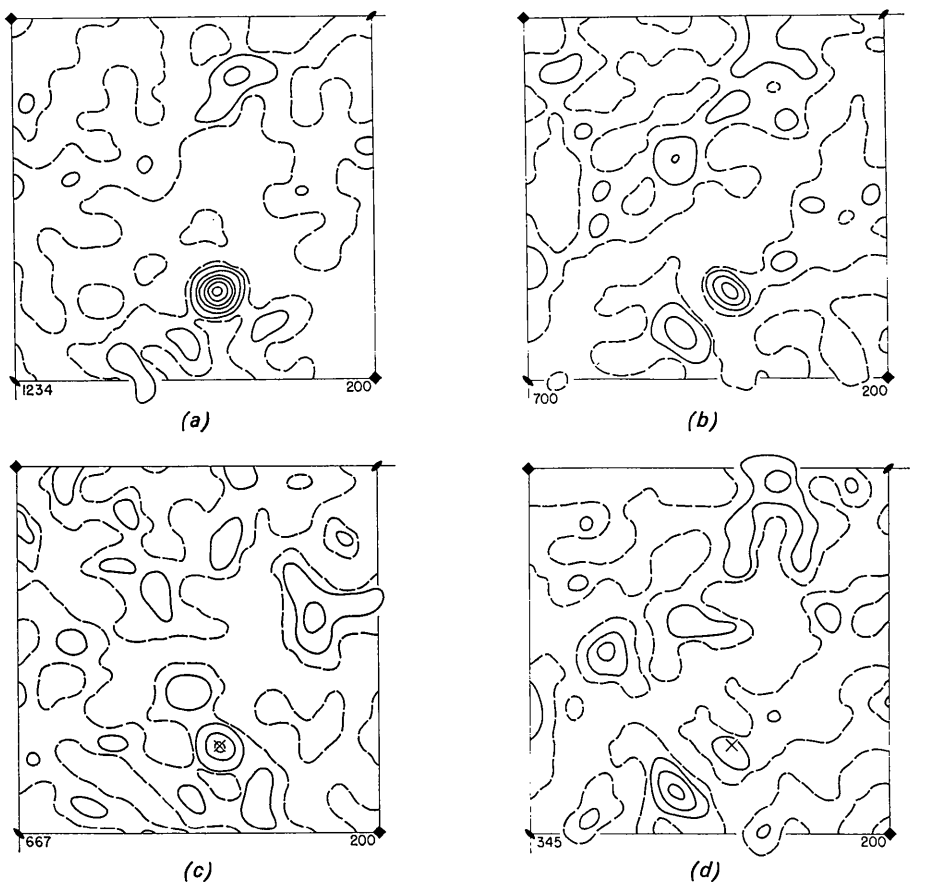


Fig. 9. Erronium ΔF maps with signs indicated. (*a*) Err signs. (*b*) Err+Pt signs. (*c*) Err+Hg signs. (*d*) Err+Pt+Hg signs.

finement, the four-data set gave a mean figure of merit of 0·52, and ten cycles of refinement brought this up to 0·58. The Err+Pt+Hg set before refinement gave a value of 0·43, and ten cycles brought this up to 0·48. However, in the course of this refinement the occupancy number of the erronium site fell drastically as shown in Fig. 11.

Thus, in this example of a bad derivative, the bad derivative difference map was damaged by the introduction of only one good derivative and was wiped out by two; the bad sign set could not pull in the good derivatives in difference maps, and the bad site refined down towards zero occupancy when permitted to do so.

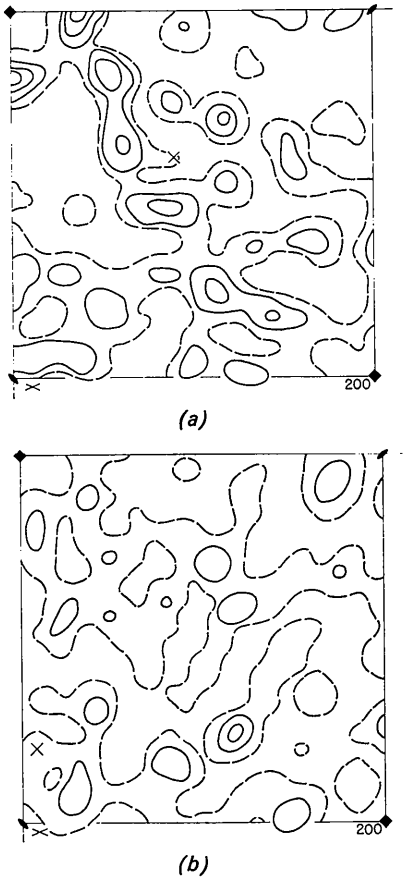


Fig. 10. (*a*) Pt and (*b*) Hg ΔF difference maps with Err signs. Expected heavy atom sites are marked with a ×.

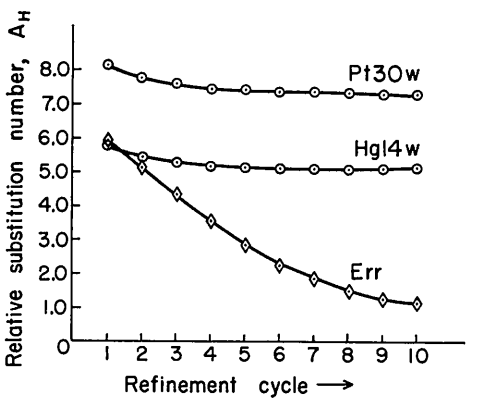


Fig. 11. Effective atomic numbers or site occupancy numbers of Pt, Hg, and Err during phase refinement.

Conclusions

These trials seem to illustrate some fairly general principles:

(1) In interpreting difference Patterson maps, there is considerable danger of deception by false detail from errors in low-order terms, particularly at 6 Å resolution. If many sites are required, if vector peaks from these sites tend to overlap in multiple peaks, if the sites themselves tend to fall on a common plane or to form parallelograms or anything like them, and especially if the Patterson map is sensitive to the removal of any selection of inner reflections, then the interpretation should be viewed with mistrust.

(2) Difference Patterson maps of multiple-site derivatives (3 or more per asymmetric unit in $P4_1$, for example) probably will not be interpretable in isolation and will have to be brought in with the aid of other derivatives. But this roundabout interpretation must then be shown to be compatible with the original difference Patterson map.

(3) The surest proof of the validity of a derivative by difference Fourier methods is a cross-sign map using signs or phases which are absolutely untainted by information about the derivative in question.

(4) A completely wrong derivative will have less effect on the phase analysis than might have been expected, especially in the presence of other very good derivatives. Careful use of cross difference Fourier maps will reveal the presence of a spurious derivative.

(5) Trends in the refinement of the substitution numbers will often point out false derivatives or false secondary sites. Although the absolute values of the mean figure of merit can be deceptive, relative changes in m as derivatives are combined *are* informative.

(6) The feedback problem need not be severe. Ghost peaks, while sometimes visible, would seldom be confused with real peaks unless one were overenthusiastic about choosing 'subsidiary sites'. If the minor sites of one derivative turn out to coincide with the principal sites of other derivatives, then there are grounds for skepticism.

An investigator who publishes a low-resolution protein structure analysis is asking the reader to take an unusually large amount on faith. The results, once obtained, do not necessarily make any obvious chemical sense, the usual way of casually judging a structure analysis. As a compensation, therefore, the path of analysis should be crystal-clear.

A reasonable minimum standard of publication might include the following:

(*a*) Complete F data for parent and derivatives. At low resolution, this amounts to only 1000–1500 reflections per derivative, a not excessive number even by

conventional standards. Since these data are the justification for everything else, they should be easily accessible to the reader.

(*b*) The $(\Delta F)^2$ difference Patterson map for *each* derivative used, whether it was interpretable by itself or not. The map should have marked on it the locations of the vectors expected from the heavy atom sites as finally adopted.

(*c*) A summary of the manner in which the heavy atom sites were deciphered, and in particular how those derivatives whose difference Patterson maps were uninterpretable in isolation were pulled into the analysis.

(*d*) ΔF difference maps for each derivative using phases or signs obtained from other unrelated derivatives. The parameters of these derivatives should be as they were before any refinement in combination with the derivative in question.

(*e*) Mean figures of merit and other refinement criteria such as the Kraut R factor for each derivative separately (or for pairs of derivatives in three dimensions), and for the final combination of all derivatives before and after refinement.

(*f*) Commentary on any unusual features of the refinement, such as the previously mentioned wiping out of the erronium site, which would permit one to judge the derivatives.

The opportunities for self-deception in a low resolution analysis are limitless. If the fundamental difference Patterson maps are interpretable, then their publication should be a matter of record. If some are *not* interpretable, and if the derivatives are used in the phase analysis, then publication of the maps becomes a matter of obligation. Enough supplementary information should then be provided to convince the average crystallographer that the derivatives are valid *in spite of* the uninterpretability of the difference Patterson maps. In view of the difficulties of interpreting the final structure, the onlooker may legitimately ask, 'If you don't know where you are going, how do you know when you are there?' The only answer is that the course of analysis must be so transparent and so obvious that *any* end product, no matter how unexpected in appearance, will be accepted.

We would like to thank Drs Jon Bordner and David Eisenberg for their help in the calculations involving the double derivative, and Miss Lillian Casler for the preparation of all of the figures. The authors are indebted to the United States Public Health Service for their support in the form of research grant GM 12121, under which this work was carried out. One of the authors (J.W.) is also the holder of a National Institutes of Health predoctoral traineeship.

References

BLOW, D. M. & CRICK, F. H. C. (1959). *Acta Cryst.* **12**, 794.

CULLIS, A. F., MUIRHEAD, H., PERUTZ, M. F., ROSSMANN, M. G. & NORTH, A. C. T. (1961). *Proc. Roy. Soc.* A, 265, 15.

DICKERSON, R. E. (1964). In *The Proteins*, Vol. II, p. 623. New York: Academic Press.

DICKERSON, R. E., KENDREW, J. C. & STRANDBERG, B. E. (1961*a*). *Acta Cryst.* **14**, 1188.

DICKERSON, R. E., KENDREW, J. C. & STRANDBERG, B. E. (1961*b*). In *Computing Methods and the Phase Problem in X-ray Crystal Analysis*, p. 236. New York: Pergamon Press.

DICKERSON, R. E. & PALMER, R. A. (1967). *Acta Cryst.* Submitted for publication.

KRAUT, J. (1961). Private communication.

KRAUT, J., SIEKER, L. C., HIGH, D. F. & FREER, S. T. (1962). *Proc. Nat. Acad. Sci. Wash.* **48**, 1417.

LIPSCOMB, W. N., COPPOLA, J. C., HARTSUCK, J. A., LUDWIG, M. L., MUIRHEAD, H., SEARL, J. & STEITZ, T. A. (1966). *J. Mol. Biol.* **19**, 423.

MARGOLIASH, E. (1967). In *Methods in Enzymology*. (Edited by S.P. Colowick and N.O. Kaplan). In the press.

MUIRHEAD, H. (1966). Private communication.

Acta Cryst. (1967). **23**, 522

The Effect of Errors in the Intensities on the Phase Angles Determined by the Isomorphous Replacement Method

BY V. AMIRTHALINGAM

Bhabha Atomic Research Centre, Chemistry Division, Bombay-74, India

(*Received* 11 *October* 1966)

Expressions are derived to estimate the error in phase angle α determined by the isomorphous replacement method due to the systematic and random errors in the intensities. When α is small, the error in α is fairly large whereas when α approaches 90°, the error is small and reasonably constant. This is compared with the case of phases obtained by the anomalous dispersion method.

The evaluation of phases by the isomorphous replacement method requires the intensities in absolute values. However, in practice there are always some errors: systematic ones like scale factor, absorption *etc.*, and random errors which occur during the estimation of the intensities. But it is possible to examine the problem