STUHRMANN, H. B. (1970*a*). *Acta Cryst.* **A26**, 297-306.
STUHRMANN, H. B. (1970*b*). *Z. Phys. Chem.* (*Frankfurt am Main*), **72**, 177-184, 185-198.
STUHRMANN, H. B. (1975). *Small-Angle Scattering of Proteins in Solution.* In *Neutron Scattering for the Analysis of Biological Structures. Brookhaven Symp. Biol.* **27**, IV-3-19.
STUHRMANN, H. B. & KIRSTE, R. G. (1965). *Z. Phys. Chem.* (*Frankfurt am Main*), **46**, 247-250.
STUHRMANN, H. B., KOCH, M. H. J., PARFAIT, R., HAAS, J., IBEL, K. & CRICHTON, R. R. (1977). *Proc. Natl Acad. Sci. USA*, **74**, 2316-2320.
TAUPIN, D. & LUZZATI, V. (1982). *J. Appl. Cryst.* **14**, 289-300.
WITZ, J. (1983). *Acta Cryst.* **A39**, 706-711.
YONATH, A. (1989). Personal communication.
YONATH, A., LEONARD, K. R. & WITTMANN, H. G. (1987). *Science*, **236**, 813-816.

Acta Cryst. (1991). A47, 744-748

E-Map Improvement in Direct Procedures

BY A. ALTOMARE, G. CASCARANO AND C. GIACOVAZZO

Istituto di Ricerca per lo Sviluppo delle Metodologie Cristallografiche, CNR, c/o Dipartimento Geomineralogico, Campus Universitario, 70124 Bari, Italy

AND D. VITERBO

Dipartimento di Chimica, Università della Calabria, 80030 Arcavacata di Rende, Cosenza, Italy

(*Received 11 October* 1990; *accepted 3 June* 1991)

Abstract

The quality of an E map is usually affected by systematic and/or random phase errors, by amplitude truncation effects in the series representation of the electron density, by the experimental uncertainty in the estimation of $|E|$ and by the intrinsic nature of the Fourier coefficients used (*i.e.* the E's). It is shown that simple supplemental calculations can improve the quality of an E map. Large molecular fragments can often be localized in the new map even when the original one is not easily interpretable.

1. Introduction

Usually a direct-methods procedure ends with one or more sets of approximated phases with which E maps are computed. The chemical significance of any trial solution is checked *via* atomic connectivity tables. Often one or more chemically sound fragments are well localized: the complete structure is then obtained by traditional least-squares and Fourier techniques. Sometimes a clear solution is not obtained: the map is uninterpretable because some atoms are occasionally missed or ghost peaks are present or the molecular geometry is distorted. But even in these cases a post mortem analysis of the structure may reveal the presence of correctly positioned atoms or fragments. Unfortunately, if such atoms or fragments were not *a priori* recognized in the E map, a procedure devoted to recovering the total from a partial structure (Karle, 1970; Beurskens, Prick, Doesburg, & Gould, 1979; Giacovazzo, 1983; Burla, Cascarano, Fares, Giacovazzo, Polidori & Spagna, 1989) would not be readily applicable. In this context it is of crucial importance to have a procedure which, from a traditional E map, is able to obtain a new more interpretable Fourier map.

The quality of a map depends on several factors among which the following three play a prominent role:

(*a*) *The phase errors.* In most cases these are unavoidable: large random errors can be tolerated without great loss of structural information in the E map while systematic errors have greater destructive effects (Silva & Viterbo, 1980).

(*b*) *Amplitude truncation effects* in the series representation of the electron density. Traditional direct methods do not phase reflections under the minimum threshold value $E_{\mathrm{Tr}} \simeq 1{\cdot}2$: in most of the practical applications E_{Tr} lies in the range 1·30-1·50. If phases are determined with sufficient accuracy the amplitude truncation effects are not really harmful (this is a necessary condition for the general success of direct methods). However, if this effect is associated with phase errors the final result is often destructive. A classical example is structures suffering from pseudotranslational symmetry: if no special action is undertaken the reflections actively used in the phasing process coincide with substructure reflections. Even when these reflections are accurately phased, and that it is not the rule, the information on the superstructure is completely lost in the E map.

(*c*) *The Fourier coefficients* used for calculating the map. It is a traditional practice to use E coefficients at the conclusion of a phasing process: they produce a peakyness effect in the map which makes its interpretation easier in terms of atoms. On the other hand, false details, ripples of heavy atoms *etc.* are also produced which in some cases make the correct interpretation difficult. Such undesirable behaviour is generated by the intrinsic nature of the E coefficients [*i.e.* real atoms are replaced by unreal point atoms; $\langle|E|^2\rangle$ is never vanishing for any $(\sin\theta)/\lambda$ value] and by their imperfect estimation from experimental $|F|$ values.

An effective process offering the possibility of reducing phase errors by handling fragments produced by E maps is the *TRITAN* procedure (Refaat & Woolfson, 1988). A further technique able to improve the information provided by standard E maps has been described by Sheldrick (1982). The phase determination was divided into two stages: (a) cycles of tangent refinement are performed among a reduced number (say 200) of phases: the best (on the basis of early figures of merit) trials are then subjected to one cycle of tangent expansion up to say 400 phases. The main purpose of this step is to increase the number of trials without paying too much in terms of computer time; (b) the phases of the selected trials are submitted to partial structure recycling based on E values and on special least-squares procedures.

In the present paper we are mostly interested in obtaining maps in which the effects of amplitude truncation and those arising from the use of E coefficients are strongly reduced. In particular we will show that:

(1) direct phase expansion can be extended to $|E|$ values much smaller than threshold values normally used in standard direct procedures (or suggested in Sheldrick's paper). Our tests prove that there is some penalty to pay in terms of phase accuracy, but it is largely compensated by smaller amplitude truncation effects;

(2) the information contained in a standard E map can be improved without least-squares techniques. Recycling leads to an improved F (rather than E) map. Such a process can be applied by all those direct-methods programs not integrated in complete crystallographic packages.

2. The procedure

It is well known that highly regular structures show in E maps pronounced translational symmetries additional to that really present. Regular long chains, fused rings *etc.* are joined in an E map to translated spurious maxima whose intensity is often comparable with that of true atoms. Bürgi & Dunitz (1971) suggest extending the phasing process to smaller E values as a means of reducing the level of false details. Mo, Hjortås & Svinning (1973) warned against such a practice: the active use of small E values can introduce errors in the phasing process which, systematically propagated, can cause the failure of the phasing process. In some cases it may be advantageous to break off the phasing process at an early stage and to calculate E maps with a small number of terms rather than to involve an unreliable large number of phases.

Bürgi & Dunitz's suggestion may be harmonized with Mo, Hjortås & Svinning's warning on the condition that large phase errors associated with small E's have no effect on the phase determination of the largest E's. Such a condition is satisfied by the expansion procedure here described. The scenario is that where *SIR*88 (Burla, Camalli, Cascarano, Giacovazzo, Polidori, Spagna & Viterbo, 1989) is run in order to phase a number NLAR of reflections with values larger than E_{Tr}. The NTOP (in default NTOP = 10) sets of phases with the highest combined figure of merit CFOM are saved.

The new adopted expansion procedure is performed through the following main steps:

(a) A number NEXP of reflexions, immediately following the NLAR reflexions in a list sorted in decreasing order of E, is selected. We have fixed NEXP to $0{\cdot}8\times$NLAR, with NEXP ≤ 300.

(b) All triplets (called here ψ_E or psi-E triplets) relating one of the NEXP reflexions with two NLAR reflexions are set up. To make the procedure faster triplets are estimated in accordance with Cochran's formula.

(c) For the first of the NTOP trial solutions (that with the highest value of CFOM) the phases of the NEXP reflexions are determined by means of a weighted tangent formula (Burla, Cascarano, Giacovazzo, Nunzi & Polidori 1987).

(d) A weighted E map is computed with the expanded list of NLAR + NEXP reflexions and then automatically interpreted. Each reflexion is given the weight $D_1(\alpha)$ where α is the reliability parameter associated with each phase and $D_1(x) = I_1(x)/I_0(x)$. I_i is the modified Bessel function of order i.

(e) A peak search program automatically finds peak maxima which are arranged in decreasing order of the intensity INT. The standard molecule-recognition program of *SIR*88 produces molecular fragments and labels (in a chemical sense) peaks. In this step labelling is not used: the heaviest atomic species is associated with all the peaks. With the ith peak an occupancy factor

$$\mathrm{oc}(i) = \mathrm{occ}(i)\times W_1(i)\times w_2(i)$$

is associated, where occ(i) is the crystallographic occupancy, $w_1(i)\leq 1$ is a weight depending on the ratio INT(i)/INT(1) and $w_2(i)$ is 1 for peaks belonging to a molecular fragment, 0·667 for isolated peaks and 0 for peaks too close to higher-intensity peaks.

(f) A structure-factor calculation is made. An observed Fourier map is then computed with coefficients

$$W|F_{\mathrm{obs}}| \exp[B(\sin^2\theta)/\lambda^2] \exp(i\varphi_{\mathrm{calc}})$$

where B is the average isotropic thermal factor calculated in the normalization process. $|F_{\mathrm{obs}}| \exp[B(\sin^2\theta)/\lambda^2]$ instead of $|F_{\mathrm{obs}}|$ is used in order to improve peakyness (it is the expected value of $|F_{\mathrm{obs}}|$ in the absence of thermal motion). W is Sim's (1960) weight, calculated on the (arbitrary) assumption that 2/3 of the structure is well localized.

(*g*) Steps (*e*) and (*f*) are repeated twice. Then peak-search and molecule-recognition programs are again run and peaks relabelled.

3. Experimental

The success of the procedure relies on the expectation that phase expansion to smaller E's allows the reduction of amplitude termination errors without a remarkable loss of reliability in phase estimates. In order to check such an assumption we have applied the procedure to the crystal structures quoted in Table 1. In Table 2 for each structure the following parameters are given:

NRIF	total number of reflexions
NLAR	number of reflexions with largest E chosen by *SIR*88 for solving the crystal structure
EMIN	minimum E value for NLAR reflexions
$N\Sigma_2$	number of triplets among the NLAR reflexions
$E\psi$MIN	minimum E value for NEXP reflexions
$N\psi_E$	number of ψ_E triplets
$G\psi_E$	minimum value of the reliability parameter (according to Cochran) for ψ_E triplets
$\langle\Delta\varphi\rangle_{\Sigma_2}$	average phase error for NLAR reflexions. In order to simulate a realistic distribution of the error $\Delta\varphi$ is obtained from true phases after one cycle of tangent refinement on the NLAR reflexions
$\langle\Delta\varphi\rangle_{\psi_E}$	average phase error for NEXP reflexions after phase expansion from NLAR reflexions.

Some observations may be made about Table 2:

(1) Our experience shows that a large number of ψ_E triplets is a necessary condition for a reliable phase expansion. In our tests the minimum threshold value for the reliability parameter of a ψ_E triplet is 0·6. This guarantees for all the structures a sufficiently large number of ψ_E triplets.

(2) The phase expansion to NEXP reflexions is usually quite reliable. Obviously $\langle\Delta\varphi\rangle_{\psi_E}$ is always larger than $\langle\Delta\varphi\rangle_{\Sigma_2}$: often their difference is quite negligible but even when it is very large the map calculated with NLAR + NEXP reflexions is remarkably better (see later) than the original one.

The application of *SIR*88 to the test examples quoted in Table 1 very often provides the complete or nearly complete crystal structure. We only give here a short report about some examples where standard *SIR*88 provides unsatisfactory partial structures.

Table 1. *Code name, space group and crystallochemical data for test structures*

Structure code*	Space group	Molecular formula	Z
APAPA	$P4_12_12$	$C_{30}H_{37}N_{15}O_{16}P_2 . 6H_2O$	8
AZET	$Pca2_1$	$C_{21}H_{16}Cl\,N\,O$	8
BED	$I4$	$C_{26}H_{26}N_4O_4$	8
BOBBY	$P2_13$	$Na^+Ca^{2+}N(CH_2CO_2)_3^{3-}$	4
CEPHAL	$C2$	$C_{18}H_{21}NO_4$	8
DIOLE	$I\bar{4}2d$	$C_{10}H_{18}O_2$	16
GOLD	Cc	$C_{28}H_{16}$	8
GRA4	$P\bar{1}$	$C_{30}H_{22}N_2O_4$	2
INOS	$P2_1/n$	$C_6H_{12}O_6 . H_2O$	8
LOGAN	$P2_12_12_1$	$C_{17}H_{26}O_{10}$	4
NEWQB	$P\bar{1}$	$C_{24}H_{20}N_2O_5$	4
NO55	$Fdd2$	$C_{20}H_{24}N_4$	16
PGE2	$P1$	$C_{20}H_{32}O_5$	1
QUINOL	$R\bar{3}$	$C_6H_6O_2$	54
RIFOLO†	$P2_1$	$C_{39}H_{49}N\,O_{13}$	2
TUR10	$P6_322$	$C_{15}H_{24}O_2$	12

* For the sake of brevity complete references for known structures are not given. The reader is referred to magnetic tapes distributed by the crystallographic groups in Göttingen and York.

† Cerrini, Lamba, Burla, Polidori, & Nunzi (1988).

RIFOLO

48 out of 53 non-H atoms are located by standard *SIR*88. The complete structure is provided by application of our procedure.

BED

For this structure the average phase error for NEXP reflexions was markedly larger than for NLAR reflexions (40° against 19°). However, our process proved effective: 33 atoms out of 34 are located after application of our procedure while only 27 are correctly located by *SIR*88.

Direct comparison of the molecular fragments provided by the standard molecule-recognition program of *SIR*88 with the molecular fragments obtained at the end of our procedure is very informative. Two examples are shown.

AZET

In Fig. 1(*a*) the *SIR*88 output is shown. It is rather difficult to recognize in it the two symmetry-independent molecules contained in the asymmetric unit. In Fig. 1(*b*) the complete crystal structure is shown. The three solid circles denote the atoms not found after the application of our procedure.

APAPA

In Fig. 2(*a*) the *SIR*88 output is shown: 44 non-H atoms (out of 63) are correctly located. Of the 53 peaks connected in the fragment, 44 are correctly located (the crossed positions correspond to ghosts). In Fig. 2(*b*) the complete crystal structure is shown. Only three atoms (corresponding to the solid circles) are not found.

Table 2. *Fit parameters for test compounds*

	NRIF	NLAR	EMIN	$N\Sigma_2$	$E\psi$MIN	$N\psi_E$	$G\psi_E$	$\langle\Delta\varphi\rangle_{\Sigma_2}$	$\langle\Delta\varphi\rangle_{\psi_E}$
APAPA	3241	426	1·44	2858	1·22	2356	0·600	17 (11)	32 (31)
AZET	1910	342	1·31	3003	1·03	8000	0·700	28 (25)	41 (43)
BED	3570	286	1·55	2750	1·33	2757	1·000	19 (12)	40 (25)
BOBBY	318	68	1·21	1867	0·890	2364	0·759	11 (12)	17 (20)
CEPHAL	2770	334	1·48	3115	1·22	4874	0·842	10 (10)	16 (14)
DIOLE	786	182	1·2	2731	0·96	2054	0·600	9 (9)	24 (23)
GOLD	3891	374	1·52	3262	1·31	8000	0·700	16 (16)	20 (18)
GRA4	4211	394	1·60	3774	1·15	6941	0·600	0 (0)	3 (3)
INOS	2934	304	1·61	2341	1·28	3286	0·600	0 (0)	1 (0)
LOGAN	2428	258	1·54	2733	1·27	3377	1·000	8 (8)	16 (13)
NEWQB	3673	473	1·54	2347	1·24	1856	0·600	0 (0)	5 (1)
NO55	1660	246	1·34	2688	1·12	5609	0·600	15 (11)	29 (22)
PGE2	1560	300	1·23	2594	0·96	4308	1·000	24 (20)	44 (44)
QUINOL	2844	296	1·60	2848	1·23	3709	0·600	0 (0)	2 (0)
RIFOLO	5429	362	1·72	2280	1·51	5529	1·000	10 (10)	21 (18)
TUR10	1340	219	1·35	2704	1·06	2509	0·835	10 (9)	30 (25)

Concluding remarks

A procedure for improving standard *E* maps is described. Amplitude truncation effects in the series representation of the electron density are strongly reduced by phasing a supplemental number of reflexions. *F* instead of *E* coefficients are used in the Fourier series and are combined with useful criteria for modifying the electron density. The final molecular models proved to be far better than the original ones.

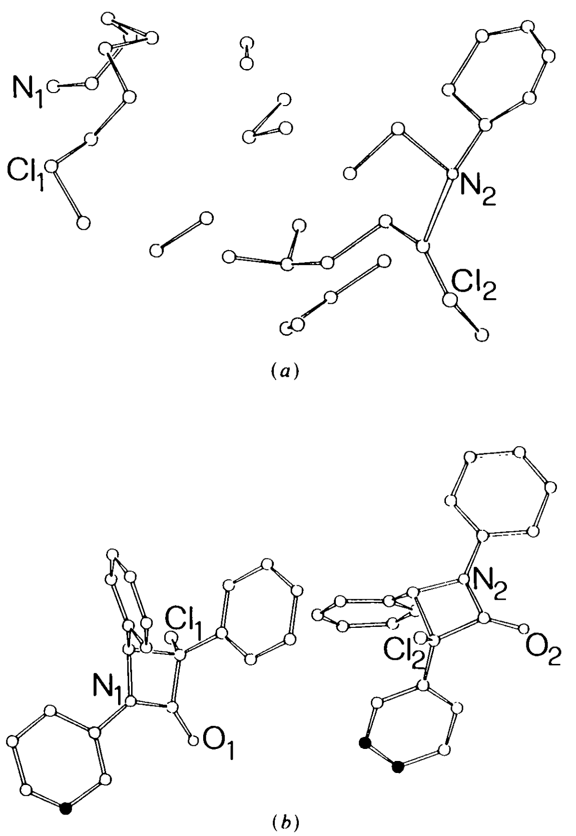

Fig. 1. (*a*) AZET: molecular fragment provided by standard *SIR*88. (*b*) AZET: symmetry-independent molecules. Solid circles denote atoms not located after application of our procedure.

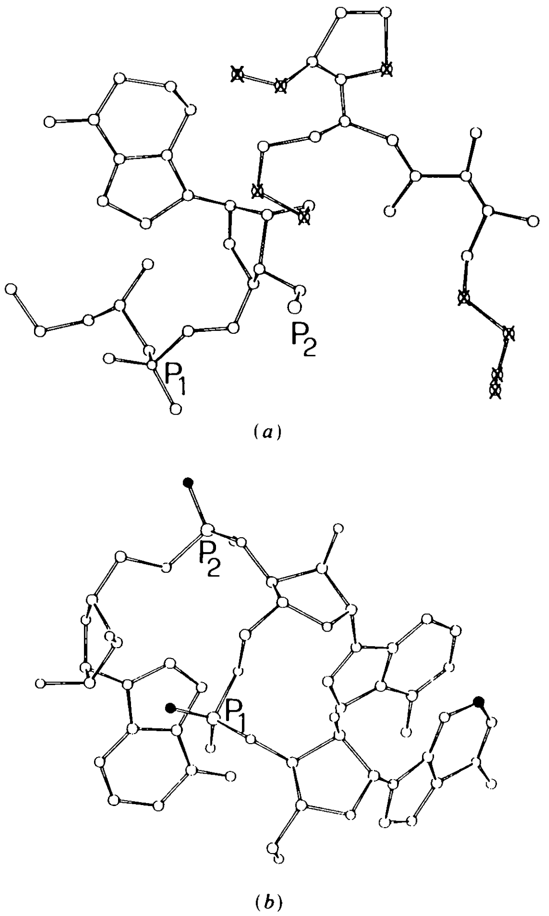

Fig. 2. (*a*) APAPA: molecular fragments provided by standard *SIR*88. Crossed circles correspond to ghosts. (*b*) APAPA: crystal structure. Solid circles denote atoms not located by our procedure.

References

BEURSKENS, P. T., PRICK, A. J., DOESBURG, H. M. & GOULD, R. O. (1979). *Acta Cryst.* **A35**, 765-772.

BÜRGI, H. B. & DUNITZ, J. D. (1971). *Acta Cryst.* **A27**, 117-119.

BURLA, M. C., CAMALLI, M., CASCARANO, G., GIACOVAZZO, C., POLIDORI, G., SPAGNA, R. & VITERBO, D. (1989). *J. Appl. Cryst.* **22**, 389-393.

BURLA, M. C., CASCARANO, G., FARES, V., GIACOVAZZO, C., POLIDORI, G. & SPAGNA, R. (1989). *Acta Cryst.* **A45**, 781-786.

BURLA, M. C., CASCARANO, G., GIACOVAZZO, C., NUNZI, A. & POLIDORI, G. (1987). *Acta Cryst.* **A43**, 370-374.

CERRINI, S., LAMBA, D., BURLA, M. C., POLIDORI, P. & NUNZI, A. (1988). *Acta Cryst.* **C44**, 489-495.

GIACOVAZZO, C. (1983). *Acta Cryst.* **A38**, 685-692.

KARLE, I. (1970). In *Crystallographic Computing*, edited by F. R. AHMED, pp. 155-164. Copenhagen: Munksgaard.

MO, F., HJORTÅS, I. & SVINNING, T. (1973). *Acta Cryst.* **A29**, 358-362.

REFAAT, L. S. & WOOLFSON, M. M. (1988). *Acta Cryst.* **A44**, 349-353.

SHELDRICK, G. M. (1982). In *Computational Crystallography*, edited by D. SAYRE, pp. 506-514. Oxford: Clarendon Press.

SILVA, A. M. & VITERBO, D. (1980). *Acta Cryst.* **A36**, 1065-1070.

SIM, G. (1960). *Acta Cryst.* **9**, 804-810.

Acta Cryst. (1991). A**47**, 748-753

N-Dimensional Diamond, Sodalite and Rare Sphere Packings

BY M. O'KEEFFE

Department of Chemistry, Arizona State University, Tempe, AZ 85287, *USA*

(*Received* 15 *March* 1991; *accepted* 5 *June* 1991)

Abstract

The diamond, lonsdaleite and sodalite structures are generalized to N dimensions and described. Ring counts and coordination sequences for the $N+1$-connected nets are given for $N \leq 6$. Simple analytic expressions are given for coordination sequences for diamonds, sodalites and the primitive hypercubic lattice. Replacing the vertices of diamonds and sodalites by regular simplices produces rare (open) stable sphere packings; general expressions for the density of these are given.

Introduction

Recent developments in the theory of quasicrystals and incommensurate structures have led to an increased interest in crystallography in more than three dimensions (*e.g.* Janssen, 1986); this is an area that has also been of enormous stimulation to pure and applied mathematics although there the emphasis has been mainly on dense lattices and sphere packings (Conway & Sloane, 1988).

In three-dimensional crystal chemistry the four-connected nets are of particular importance being, among many other things, the basis of some elemental structures and of the framework silicates and hydrates. The diamond structure has a special place among these as it is the only such net with all vertices (atoms), edges (bonds) and angles equivalent. Such a net I term *regular* (it is noted that graph theorists usually employ this term in a much less restrictive sense). In this paper some properties of the N-dimensional analogs of diamond and its simplest polytype (in three dimensions, lonsdaleite) are described. Another important net is that of sodalite; it has all vertices and edges equivalent (*quasiregular*). Generalizations of this net are also described. Some reasons why the properties of these structures might be of interest are given below after some necessary basic definitions.

Recent discussions of nets (Stixrude & Bukowinski, 1990; O'Keeffe, 1991*a*) have focused particular attention on coordination sequences (Brunner, 1979; Meier & Moeck, 1979) and ring statistics (Marians & Hobbs, 1990). A coordination sequence consists of the numbers, n_k, of kth neighbors of a vertex, a kth neighbor being one for which the shortest path to the reference vertex consists of k edges. One can define (O'Keeffe, 1991*a*) a dimensionless local topological density

$$\rho_k = \sum_{i=1}^{k} n_i/k^3. \qquad (1)$$

The limit as $k \rightarrow \infty$ is called the global topological density ρ_∞. The generalization to N dimensions is obvious. The number of neighbors n_{Nk} for an N-dimensional structure can sometimes be fit to a polynomial

$$n_{Nk} = \sum_{i=0}^{N-1} a_{Ni}k^i. \qquad (2)$$

In that event $\rho_\infty = a_{N,N-1}/N$. In this work the coefficients have been empirically determined from a count of a large number of neighbors, the number of shells counted being at least twice the number of coefficients.

0108-7673/91/060748-06$03.00