Acta Cryst. (1981). A**37**, 742–754

The Thermodynamic Approach to the Structure Analysis of Crystals

BY A. KHACHATURYAN, S. SEMENOVSKAYA* AND B. VAINSHTEIN

Institute of Crystallography of the Academy of Sciences of the USSR, Leninsky Prospekt 59, Moscow B-333, USSR

(*Received* 13 *October* 1980; *accepted* 26 *March* 1981)

Abstract

In a previous paper [Khachaturyan, Semenovskaya & Vainshtein (1979). *Sov. Phys. Crystallogr.* **24**, 519–524], the thermodynamic formulation of the procedure of crystal structure determination from the intensity array of X-ray reflections was proposed. This approach requires the introduction of analogues of such thermodynamic functions as the free energy, entropy, temperature and so on, and it is reduced to the solution of the specific 'kinetic equation'. It was shown that the equilibrium state of the system is described by the distribution of the scattering substance density which at low temperature corresponds to the true structure of the crystal. The problem of crystal structure determination is thus reduced to determination of the equilibrium distribution of the scattering substance. The Önsager equation formalism was applied to derive 'kinetic equations' describing the relaxation of the system to its equilibrium state. The test of the proposed procedure exemplified by the solution of the 'kinetic equation' to determine the crystal structure of L-proline, $C_5H_9NO_2$, with 32 atoms per unit cell is considered.

1. Introduction

In some cases the structure analysis proves to be easier if instead of the crystal under investigation we consider a model crystal consisting of homogeneous spheres of the same radius r_0 which replace the atoms of the real crystal. The transition to the model crystal composed of these spheres results in the following transformation of the intensity array (Khachaturyan, Semenovskaya & Vainshtein, 1979):

$$I_{\rm obs}(\mathbf{H}) \rightarrow I_{\rm exp}(\mathbf{H}) = \frac{I_{\rm obs}(\mathbf{H})}{|\tilde{f}(\mathbf{H})|^2} \cdot \left| \frac{v}{\Omega}\, \theta(H) \right|^2, \qquad (1)$$

where $I_{\rm obs}(\mathbf{H})$ are observed diffraction intensities in absolute units, $\mathbf{H}$ is the reciprocal-lattice vector, $\tilde{f}(\mathbf{H})$ is the average atomic factor of the crystal, Ω is the unit-cell volume, v is the volume of the sphere,

$$\theta(H) = 3 \cdot \frac{\sin 2\pi H r_0 - 2\pi H r_0 \cos 2\pi H r_0}{(2\pi H r_0)^3}. \qquad (2)$$

The transformation of the intensity array (1) is valid if all atoms composing the crystal have close atomic factors.

Let us divide the unit cell of the model crystal into a fine regular grid, N sites of which are equidistant and assume that the scattering substance can occupy only sites of the grid labelled by the vector $\mathbf{r}$.

This model allows the introduction of some 'elementary particles' of the scattering substance. The distribution of the scattering substance within a unit cell then may be described in terms of the distribution function of 'elementary particles' $C(\mathbf{r})$, which is either equal to unity or zero:

$$C(\mathbf{r}) = \begin{cases} 1 & \text{if the site } \mathbf{r} \text{ is occupied by a scattering substance 'particle'} \\ 0 & \text{otherwise.} \end{cases} \qquad (3)$$

For instance, the model crystal composed by spheres is described by the specific distribution function $C_0(\mathbf{r})$ which is equal to unity if $\mathbf{r}$ is within a sphere and zero if $\mathbf{r}$ is outside it (see Fig. 1).

The structure amplitudes corresponding to an arbitrary distribution of elementary 'particles' may be represented in the following form,

$$F(\mathbf{H}) \simeq \frac{1}{N} \sum_{\mathbf{r}} C(\mathbf{r})\, e^{-i2\pi \mathbf{H}\cdot\mathbf{r}}. \qquad (4)$$

One may readily see that intensities $|F(\mathbf{H})|^2$ related to the distribution $C_0(\mathbf{r})$ asymptotically tend to intensities $I_{\rm exp}(\mathbf{H})$ given by (1) as $N \rightarrow \infty$ [the accuracy of the representation (4) is greater the higher the value of N].

It is worth while emphasizing the analogy between the proposed model and the lattice gas model (Hill, 1956) and the binary substitutional solution model (Khachaturyan, 1963). In all these models the distribution function $C(\mathbf{r})$ describes two states: the presence

* Institute of Organo-Element Compounds, Academy of Science of the USSR, Moscow, USSR.

0567-7394/81/050742-13$01.00

of a particle at the site $\mathbf{r}$ if $C(\mathbf{r}) = 1$ and its absence if $C(\mathbf{r}) = 0$.

If we have an ensemble of distributions the mean structure amplitude over the ensemble can be introduced as a mean value of (4):

$$\langle F(\mathbf{H})\rangle = \frac{1}{N}\sum_{\mathbf{r}} n(\mathbf{r})\, e^{-i2\pi\mathbf{H}.\mathbf{r}}, \tag{5}$$

where $n(\mathbf{r}) = \langle C(\mathbf{r})\rangle$, $\langle\ \rangle$ is the symbol of averaging over the thermodynamic ensemble. This kinematic analogy with a lattice gas can be further extended to introduce thermodynamic functions. For example, Helmholtz free energy Φ can be obtained by the conventional procedure in terms of the partition sum (see, for instance, Landau & Lifchitz, 1958). To do it we should first choose a configurational Hamiltonian of the system. As a Hamiltonian $\mathscr{H}$ of the lattice gas consisted of 'particles' of scattering substance we can choose the generalized R index multiplied by the factor N:

$$\mathscr{H} = N\hat{R} = \frac{N}{4}\sum_{\mathbf{H}} a(\mathbf{H})\left[\left|\frac{1}{N}\sum_{\mathbf{r}} C(\mathbf{r})\, e^{-i2\pi\mathbf{H}.\mathbf{r}}\right|^2 - I_{\exp}(\mathbf{H})\right]^2, \tag{6}$$

where $I_{\exp}(\mathbf{H})$ is given by (1).

The Hamiltonian (6) assumes minimal value, equal to zero, for the distribution function $C(\mathbf{r}) = C_0(\mathbf{r})$ corresponding to the true structure for any set of positive quantities $a(\mathbf{H})$. Introduction of weight coefficients $a(\mathbf{H})$ into (6) makes the minimization procedure more flexible and precludes false structures corresponding to the relative minima. By choosing the Hamiltonian in the form (6) one can represent the partition sum in a form similar to that in the Ising problem:

$$Z = \sum_{\{C(\mathbf{r})\}} \exp\left(-\frac{N\hat{R}}{T}\right) = \sum_{C(\mathbf{r}_1)=0}^{C(\mathbf{r}_1)=1} \cdots \sum_{C(\mathbf{r}_N)=0}^{C(\mathbf{r}_N)=1} \exp\left\{-\frac{N}{4T}\sum_{\mathbf{H}} a(\mathbf{H}) \times\left[\left|\frac{1}{N}\sum_{\mathbf{r}} C(\mathbf{r})\, e^{-i2\pi\mathbf{H}.\mathbf{r}}\right|^2 - I_{\exp}(\mathbf{H})\right]^2\right\} \tag{7}$$

with the additional condition

$$\sum_{\mathbf{r}} C(\mathbf{r}) = \nu N, \tag{8}$$

where ν is the fraction of the unit cell filled by the scattering substance, T is the dimensionless temperature. In this case the free energy is

$$\Phi = -T\ln Z. \tag{9}$$

On the other hand, the free energy can be presented in the form

$$\Phi = NR - TS, \tag{10}$$

where

$$R = \langle\hat{R}\rangle = \sum_{\{C(\mathbf{r})\}} \hat{R}\,\frac{1}{Z}\exp\left(-\frac{N\hat{R}}{T}\right)$$

and NR is the analogue of the internal energy. The entropy S can be found by means of the Boltzmann equation:

$$S = -\ln\Gamma(R), \tag{11}$$

where $\Gamma(R)$ is the phase space volume corresponding to the given energy value R (the number of realizations of scattering substance distributions ensuring the same R index). As in ordinary statistical mechanics the temperature T introduced into (7) characterizes the degree of 'excitation' of the system (the increase of the 'energy' R). The 'excitation' results in the 'impairment' of the true distribution of scattering substance owing to its disordering. In this connection the ensemble of the considered 'excited' distributions is determined to ensure that the mean R index is equal to the given value R. The mean distribution over such an ensemble will deviate more from the true one the higher the temperature of the system is. The formal definition of the temperature is

$$\frac{1}{T} = \frac{dS}{dR}.$$

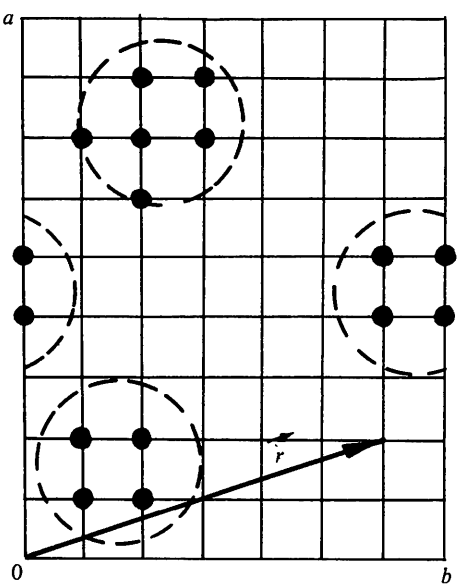

Fig. 1. The schematic drawing of the grid in a section of the model crystal. 'Particles' of the scattering substance occupy sites of the grid. Their segregations imitate atoms. ● is a 'particle' of the scattering substance where $C_0(\mathbf{r}) = 1$. The rest of the grid sites where $C_0(\mathbf{r}) = 0$ are vacant. Dashed circles show sections of spheres which are described by segregation of particles.

Free-energy calculation in terms of the partition sum (7) is just as complex as that in the case of the Ising model of a lattice gas (binary alloy). However, as in the Ising model case, we can employ the mean-field (self-consistent field) approximation which is asymptotically correct in the particular case of the low temperatures (Suris, 1962), which is just the one we are interested in.

The free energy in the mean-field approximation can be written in the form*

$$\Phi = NR + T\sum_{\mathbf{r}} \{n(\mathbf{r}) \ln n(\mathbf{r}) + [1 - n(\mathbf{r})] \times \ln[1 - n(\mathbf{r})]\}, \tag{12}$$

where, according to the idea of the mean-field approximation, fluctuations are ignored,

$$R = \langle \hat{R} \rangle = \frac{1}{4}\sum_{\mathbf{H}} a(\mathbf{H}) \left\langle \left[\left| \frac{1}{N} \sum_{\mathbf{r}} C(\mathbf{r})\, e^{-i2\pi \mathbf{H}.\mathbf{r}} \right|^2 - I_{\exp}(\mathbf{H}) \right]^2 \right\rangle \tag{13}$$

$$\simeq \frac{1}{4}\sum_{\mathbf{H}} a(\mathbf{H}) [|\langle F(\mathbf{H}) \rangle|^2 - I_{\exp}(\mathbf{H})]^2$$

and $\langle F(\mathbf{H}) \rangle$ is defined by (5).

Equation (12) is the analogue of the corresponding equations for the free energy of a binary solution in the mean-field approximation (Khachaturyan, 1963, 1978), the first term of (12) is the analogue of the internal energy, while the second one is the analogue of the entropy term $-TS$ (see equation 10). The derivation of (12) is not given since it is identical to that for a binary solution (Khachaturyan, 1963).

Minimization of the free energy (12) with respect to $n(\mathbf{r})$ results in the equilibrium distribution of the scattering substance. As has been shown by Khachaturyan (1963), the function $n(\mathbf{r})$ describing the equilibrium distribution at $T = 0$ assumes only two values: 1 or 0. On the other hand, from (12) it follows that when $T \to 0$ the limit transition $\Phi \to NR$ occurs.

As follows from (6), the true structure ensures minimization of the R index with respect to the scattering substance distribution $C(\mathbf{r})$, which takes two values: 1 and 0.

Thus, when $T \to 0$ minimization of free energy (12) *is equivalent to the minimization of the R index in* (6) *for the desired variety of functions assuming the values either* 1 *or* 0. *In other words, when $T \to 0$ minimization of free energy* (12) *solves the problem of determination of the crystal structure for a given intensity array $I_{\exp}(\mathbf{H})$.*

* The Boltzmann constant K_B does not enter (9), (12), and (19) to (21) since all thermodynamic functions are expressed in dimensionless units.

It is worth while, however, to note that any crystal structure determination procedure based on the direct minimization of an R index with respect to atomic coordinates has no chance of being successful (Vand, Niggli & Pepinsky, 1960; Niggli, Vand & Pepinsky, 1960). In fact, the R index has a great number of the shallow relative minima in the phase space of atomic configuration. These minima arrest the minimization procedure before the absolute minimum is attained. This situation is close to that for real atomic systems. The configurational Hamiltonian of atomic systems has, as a rule, many shallow relative minima corresponding to metastable states. The system 'trapped' by a relative minimum at 0 K cannot escape it independently since the escape would require the expenditure of additional energy to overcome the energy barriers separating the minimum from the lower energetic states. The situation becomes more favourable at finite temperatures when the thermal fluctuations allow the system to escape from the metastable state. If the relative minimum is shallow it cannot be a 'trap' since the phase space volume of the minimum is small and, thus, the entropy of the state becomes small as well (according to the Boltzmann equation the entropy is proportional to the logarithm of the phase space volume). The drop of the entropy associated with the 'trapping' of the system by a shallow relative minimum would result in the increase of the Helmholtz free energy Φ since

$$\Phi = U - TS. \tag{14}$$

According to the second principle of thermodynamics any spontaneous process must be accompanied by a free-energy decrease. Thus, the above-described drop of the entropy and the corresponding free-energy increase which occurs when the system is 'trapped' by a shallow minimum state can make this state unstable. The latter conclusion is physically obvious. A shallow relative minimum cannot 'retain' the fluctuating system at finite temperatures. Continuing this line of reasoning we can conclude that elevating the temperature may result in the situation when all minima including the absolute one are not able to 'retain' the system. In other words, the free energy playing the part of the potential energy at finite temperatures is a much smoother function than the Hamiltonian.

The above formulated analogy between the R index and the configurational Hamiltonian is the basis of an idea which enables us to formulate the thermodynamic approach to the structure analysis of crystals. This approach implies the replacement of the R index by a 'smoother' function which is the formal analogue of the Helmholtz free energy. 'Smoothness' of this function may be controlled by temperature changes. If we have a 'smooth' function the real crystal structure can be found by means of conventional procedures which are variants of the gradient descent method.

2. Kinetic equations for density of scattering substance and for structure amplitudes

The mathematical analogy between the problem of the determination of crystal structure and the phase transition problem can be extended not only to thermodynamics but also to kinetics. In fact we can always imagine some kinetic process which involves 'particles' of scattering substance distributed over sites of the grid dividing the unit cell. Being spontaneous, this process results in a decrease in free energy (12) and, consequently, leads to the desired equilibrium state ensuring the absolute minimum of the free energy.

Roughly following the line suggested by Khachaturyan (1968) and Morris & Khachaturyan (1978), where the microscopic diffusion theory is developed, we can use the Önsager equation, which establishes a direct proportionality between the rate of relaxation parameters and thermodynamic driving forces – the variational derivatives of the free energy with respect to the relaxation parameters. In the case under consideration such an equation is

$$\frac{\mathrm{d}n(\mathbf{r},t)}{\mathrm{d}t} = -L\frac{\delta\Phi}{\delta n(\mathbf{r},t)}, \tag{15}$$

where t is the dimensionless 'time' and L is the kinetic coefficient. Since off-diagonal elements of the matrix of the kinetic coefficients $L(\mathbf{r}-\mathbf{r}')$ are neglected, (15) describes 'creation' and 'annihilation' of particles of the scattering substance at each site $\mathbf{r}$ of the grid rather than diffusional transitions between these sites. The kinetic process ceases on the attainment of the free-energy minimum when $\delta\Phi/\delta n(\mathbf{r}) = 0$. In this case, according to (15) $\delta n(\mathbf{r},t)/\mathrm{d}t = 0$, *i.e.* the density function $n(\mathbf{r},t)$ ceases to depend on 'time'.

Let us calculate the 'driving force' of the kinetic process $\delta\Phi/\delta n(\mathbf{r})$. The variation of (13) with respect to $n(\mathbf{r})$ yields

$$\frac{\delta R}{\delta n(\mathbf{r})} = \frac{1}{2}\sum_{\mathbf{H}} a(\mathbf{H})[|\langle F(\mathbf{H})\rangle|^2 - I_{\exp}(\mathbf{H})] \times\left[\langle F(\mathbf{H})\rangle\frac{\delta\langle F(\mathbf{H})^*\rangle}{\delta n(\mathbf{r})} + \langle F(\mathbf{H})^*\rangle\frac{\delta\langle F(\mathbf{H})\rangle}{\delta n(\mathbf{r})}\right]. \tag{16}$$

The variation derivative of (5) has the form

$$\frac{\delta\langle F(\mathbf{H})\rangle}{\delta n(\mathbf{r})} = \frac{1}{N}e^{-i2\pi\mathbf{H}.\mathbf{r}}. \tag{17}$$

Substituting (17) into (16), we obtain

$$\frac{\delta R}{\delta n(\mathbf{r})} = \frac{1}{N}\sum_{\mathbf{H}} a(\mathbf{H})[|\langle F(\mathbf{H})\rangle|^2 - I_{\exp}(\mathbf{H})]\langle F(\mathbf{H})\rangle \times e^{i2\pi\mathbf{H}.\mathbf{r}}. \tag{18}$$

It follows from (12) that

$$\frac{\delta\Phi}{\delta n(\mathbf{r})} = N\frac{\delta R}{\delta n(\mathbf{r})} + T\ln\frac{n(\mathbf{r})}{1-n(\mathbf{r})}. \tag{19}$$

Putting $L = 1$ into (15) (this results in the change of the 'time' scale only) and with (18) and (19), we may rewrite (15):

$$\frac{\mathrm{d}n(\mathbf{r},t)}{\mathrm{d}t} = -T\ln\frac{n(\mathbf{r},t)}{1-n(\mathbf{r},t)} - \sum_{\mathbf{H}} a(\mathbf{H})[|\langle F(\mathbf{H},t)\rangle|^2 - I_{\exp}(\mathbf{H})]\langle F(\mathbf{H},t)\rangle e^{i2\pi\mathbf{H}.\mathbf{r}}. \tag{20}$$

Representation of (20) in terms of structure amplitudes (the reciprocal-lattice representation) can be obtained by multiplication of (20) by the factor $(1/N)e^{-i2\pi\mathbf{H}.\mathbf{r}}$ and summation over $\mathbf{r}$.

$$\frac{\mathrm{d}\langle F(\mathbf{H},t)\rangle}{\mathrm{d}t} = -T\left\{\ln\frac{n(\mathbf{r},t)}{1-n(\mathbf{r},t)}\right\}_{\mathbf{H}} - a(\mathbf{H})[|\langle F(\mathbf{H},t)\rangle|^2 - I_{\exp}(\mathbf{H})]\langle F(\mathbf{H},t)\rangle, \tag{21}$$

where

$$\left\{X\right\}_{\mathbf{H}} = \frac{1}{N}\sum_{\mathbf{r}}(X)\,e^{-i2\pi\mathbf{H}.\mathbf{r}}$$

is the conversion operator of the Fourier transform. Equation (21) is the set of nonlinear differential equations for structure amplitudes $\langle F(\mathbf{H},t)\rangle$, the kernel of which includes the array of the normalized intensities $I_{\exp}(\mathbf{H})$.

It is of interest to note that the determination of the structure amplitudes proves to be possible only due to the first term in (21) since it ensures 'coupling' of structure amplitudes related to different reciprocal-lattice vectors. The 'coupling' may be visualized if the logarithm term in (21) is represented as the power-series expanison in $\Delta n = n(\mathbf{r}) - \nu$, and each term of the power series is expressed in terms of the structure amplitudes, *i.e.*

$$\begin{aligned}\left\{\ln\frac{n(\mathbf{r})}{1-n(\mathbf{r})}\right\}_{\mathbf{H}} &= \delta_{\mathbf{H},0}\ln\frac{\nu}{1-\nu} + \sum_{m=1}^{\infty}\frac{1}{m}\left[\left(\frac{1}{1-\nu}\right)^m - \left(-\frac{1}{\nu}\right)^m\right]\{\Delta n^m(\mathbf{r})\}_{\mathbf{H}} \\ &= \delta_{\mathbf{H},0}\ln\frac{\nu}{1-\nu} + \sum_{m=1}^{\infty}\frac{1}{m}\left[\left(\frac{1}{1-\nu}\right)^m - \left(-\frac{1}{\nu}\right)^m\right] \\ &\quad\times \sum_{\substack{(\mathbf{H}_1,\mathbf{H}_2,\ldots,\mathbf{H}_m)\\ \mathbf{H}_1+\mathbf{H}_2+\ldots+\mathbf{H}_m=\mathbf{H}}}\langle F(\mathbf{H}_1)\rangle \times\langle F(\mathbf{H}_2)\rangle\ldots\langle F(\mathbf{H}_m)\rangle,\end{aligned} \tag{22}$$

where

$$\Delta n(\mathbf{r}) = \sum_{\mathbf{H}\neq 0}{}' \langle F(\mathbf{H})\rangle e^{i2\pi\mathbf{H}.\mathbf{r}},$$

$$\delta_{\mathbf{H},0} = \begin{cases} 1 & \text{if } \mathbf{H} = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Priming means that the term corresponding to $\mathbf{H} = 0$ is omitted.

Equation (22) shows that kinetic equation (21) [and, hence, (20)] automatically extends the conventional approach in the structure analysis based on the consideration of triplets, quadruplets, quintets, *etc.* of reflections.

The second term of (21) by itself results in just independent relaxation of each structure amplitude and thus cannot lead to a change in phases.

The solution of (20) or (21) at $T \to 0$ and $t \to \infty$ gives the equilibrium densities and structure amplitudes and, thus, solves the problem of the crystal structure determination.

It is worth while to note that (20) is the analogue of the microscopic equation of the spinodal decomposition described and analysed by Khachaturyan (1968) and Morris & Khachaturyan (1978).

The time evolution of the densities $n(\mathbf{r},t)$ provides the free-energy decrease. Since the free energy depends implicitly on the 'time' through $n(\mathbf{r},t)$ we have

$$\frac{d\Phi}{dt} = \sum_{\mathbf{r}} \frac{\delta\Phi}{\delta n(\mathbf{r},t)} \cdot \frac{dn(\mathbf{r},t)}{dt},$$

or from (15) for $L = 1$

$$\frac{d\Phi}{dt} = \dot{\Phi} = -\sum_{\mathbf{r}} \left(\frac{dn(\mathbf{r},t)}{dt}\right)^2. \tag{23}$$

It follows from (23) that the first derivative of the free energy is always negative, *i.e.* kinetic equation (20) does provide the monotonous decrease of the free energy up to the attainment of the equilibrium state.

3. Determination of the L-proline structure

The above-formulated thermodynamic concept in the phase determination problem (see also Khachaturyan, Semenovskaya & Vainshtein, 1979) is applied below to the L-proline structure determination.

The structure of L-proline $C_4H_7(NH)COOH$ was solved by Kayushina & Vainshtein (1966). L-Proline is orthorhombic, space group $P2_12_12_1$ with $a = 11{\cdot}55$, $b = 9{\cdot}02$, $c = 5{\cdot}20$ Å. A unit cell contains four molecules. In the case of the space group $P2_12_12_1$ the density $n(\mathbf{r}) = n(x,y,z)$ is

$$n(x,y,z) = \sum_{\{hkl\}} p(hkl)\left\{A(hkl)\cos 2\pi\left(hx - \frac{h-k}{4}\right)\right.$$
$$\times \cos 2\pi\left(ky - \frac{k-l}{4}\right)\cos 2\pi\left(lz - \frac{l-h}{4}\right)$$
$$- B(hkl)\sin 2\pi\left(hx - \frac{h-k}{4}\right)$$
$$\left.\times \sin 2\pi\left(ky - \frac{k-l}{4}\right)\sin 2\pi\left(lz - \frac{l-h}{4}\right)\right\}, \tag{24}$$

where $x = r_x/a$, $y = r_y/b$, $z = r_z/c$ are dimensionless coordinates of the unit-cell point $\mathbf{r} = (r_x,r_y,r_z)$; h,k,l are indices of the reciprocal-lattice vector $\mathbf{H} = (h,k,l)$; $p(hkl)$ is the multiplicity factor of the reflection $\mathbf{H}$; $A(hkl)$ and $B(hkl)$ are real and imaginary parts of the structure amplitude:

$$\langle F(hkl)\rangle = A(hkl) + iB(hkl). \tag{25}$$

Summation in (24) is taken over all types of reflection $\{hkl\}$. Amplitudes $A(hkl)$ and $B(hkl)$ are given by the relations

$$A(hkl) = \frac{1}{N}\sum_{x,y,z} n(x,y,z)\cos 2\pi\left(hx - \frac{h-k}{4}\right)$$
$$\times \cos 2\pi\left(ky - \frac{k-l}{4}\right)\cos 2\pi\left(lz - \frac{l-h}{4}\right) \tag{26}$$

$$B(hkl) = -\frac{1}{N}\sum_{x,y,z} n(x,y,z)\sin 2\pi\left(hx - \frac{h-k}{4}\right)$$
$$\times \sin 2\pi\left(ky - \frac{k-l}{4}\right)\sin 2\pi\left(lz - \frac{l-h}{4}\right),$$

where N is the total number of the sites of the grid dividing the irreducible part of the unit cell described by the range

$$0 \le x < \tfrac{1}{2};\quad 0 \le y < \tfrac{1}{2};\quad 0 \le z < 1.$$

Summation in (26) is taken over all N sites of the grid. Taking into account (24) and (26) in (20) we can rewrite (20) in the form

$$\frac{dn(xyz,t)}{dt} = -T \ln \frac{n(xyz,t)}{1-n(xyz,t)} - \sum_{\{hkl\}} p(hkl)\, a(hkl) [A^2(hkl,t) + B^2(hkl,t) - I_{exp}(hkl)] \times \left[A(hkl,t) \cos 2\pi \left(hx - \frac{h-k}{4} \right) \times \cos 2\pi \left(ky - \frac{k-l}{4} \right) \cos 2\pi \left(lz - \frac{l-z}{4} \right) - B(hkl,t) \sin 2\pi \left(hx - \frac{h-k}{4} \right) \times \sin 2\pi \left(ky - \frac{k-l}{4} \right) \sin 2\pi \left(lz - \frac{l-z}{4} \right) \right]. \quad (27)$$

Equation (27) is the set of N differential equations for N unknowns $n(xyz,t)$. It describes the 'time' evolution of densities at each point $\mathbf{r} = (x,y,z)$ of the unit cell. These unknowns describe N degrees of freedom of the system. In the reciprocal-lattice representation of (27) the degrees of freedom will be described by the same number N of independent variables – structure amplitudes $\langle F(hkl,t) \rangle$. Thus, the solution of (27) will depend, in general, on N structure amplitudes.

This conclusion remains true even in the case when the kernel of (27) includes a reduced number of observed intensities. The latter occurs because of 'coupling' in (27) of structure amplitudes referred to various reflections. Since the number of structure amplitudes describing the solution of (27) is always equal to N, the proposed procedure does not result in the wave termination effect.

Equation (1) with $v = 4\pi r_0^3/3$, $\Omega = abc$ and $f(\mathbf{H}) = \bar{f}(\mathbf{H})\, e^{-BH^2}$ was applied to normalize the theoretical intensity array reported by Kayushina & Vainshtein (1966). The value $r_0 = 0{\cdot}65$ Å is the radius of spheres imitating C, N and O atoms, $B = 0{\cdot}975$ Å^2 is the value determining the Debye–Waller factor,

$$\bar{f}(\mathbf{H}) = 3 - 1{\cdot}071H + 3{\cdot}55\, e^{-H^2/0{\cdot}138} \quad (28)$$

is the average atomic factor attributed to spheres, $H = 2 \sin \theta/\lambda$ in Å^{-1}.

Since the importance of weak diffraction maxima as 'information carriers' about the crystal structure is less than that of strong ones (Vainshtein & Kayushina, 1966), we shall use the array consisting of strong reflections only. Normalized intensities $I_{exp}(hkl)$ of the 100 strongest reflections regrouped in the order of decrease of their magnitudes are listed in Table 1.

Numerical solution of (27) was carried out by iterations

$$n(xyz,t_m) = n(xyz,t_{m-1}) + \left(\frac{dn(xyz,t)}{dt} \right)_{t=t_{m-1}} \Delta t_{m-1}, \quad (29)$$

where t_m is the 'time' corresponding to the mth iteration cycle, Δt_m is the 'time' step, $dn(xyz,t)/dt$ is given by the right side of (27). The calculation was carried out for $N = 10 \times 10 \times 10 = 1000$ sites of the grid with coordinates

$$x = \tfrac{1}{20}(j_1 - \tfrac{1}{2}), \quad y = \tfrac{1}{20}(j_2 - \tfrac{1}{2}), \quad z = \tfrac{1}{10}(j_3 - \tfrac{1}{2}), \quad (30)$$

where j_1, j_2, j_3 are integers within the ranges $1 \le j_1 \le 10$, $1 \le j_2 \le 10$, $1 \le j_3 \le 10$. These sites cover the irreducible part of the unit cell. The nearest distance between sites corresponds to the resolution of the order of $a/20 \sim b/20 \sim c/10 \sim 0{\cdot}5$ Å.

The 'time' step Δt_{m-1} in (29) was estimated proceeding from the condition of the monotonous decrease of the free energy $\Phi(t)$:

$$\Phi(t) = \frac{N}{4} \sum_{\{hkl\}} p(hkl)\, a(hkl) [A^2(hkl,t) + B^2(hkl,t) - I_{exp}(hkl)]^2 + T \sum_{x,y,z} \{ n(xyz,t) \ln n(xyz,t) + [1 - n(xyz,t)] \ln [1 - n(xyz,t)] \}. \quad (31)$$

It is equal to

$$\Delta t_{m-1} = \begin{cases} -\dfrac{\dot{\Phi}}{\ddot{\Phi}} + \left[\left(\dfrac{\dot{\Phi}}{\ddot{\Phi}} \right)^2 - \dfrac{2D}{\ddot{\Phi}} \right]^{1/2} \\ -\dfrac{\dot{\Phi}}{\ddot{\Phi}} \quad \text{if} \left[\left(\dfrac{\dot{\Phi}}{\ddot{\Phi}} \right)^2 - \dfrac{2D}{\ddot{\Phi}} \right] < 0 \\ 200 \quad \text{if } \Delta t_{m-1} \ge 200, \end{cases} \quad (32)$$

where

$$\dot{\Phi} = \frac{d\Phi}{dt} = -\sum_{x,y,z} \left(\frac{dn(xyz,t)}{dt} \right)^2_{t=t_{m-1}}$$

$$(\ddot{\Phi})_{t_{m-1}} = [(\dot{\Phi})_{t_{m-1}} - (\dot{\Phi})_{t_{m-2}}]/\Delta t_{m-2}$$

$$\frac{50D}{N} = \tfrac{1}{4} \sum_{\{hkl\}} p(hkl)\, a(hkl)\, I_{exp}(hkl) - T[v \ln v + (1-v) \ln (1-v)] \quad (33)$$

$$v = \frac{32v}{\Omega} = \frac{32(4\pi r_0^3/3)}{abc} = 0{\cdot}068. \quad (34)$$

Table 1. *Calculated values of intensities* $|\langle F(hkl,\infty)\rangle|^2$ *and phases* φ_{hkl} *for the* 100 *strongest reflections of the normalized intensities array for* L-*proline after the six-stage crystal structure determination procedure* (*see Table* 3)

The minimum temperature is $T = 10^{-7}$. The final R_L value is 0·21.

$h\,k\,l$	I_{exp}	$\|\langle F\rangle\|^2$	φ_{true}	φ	$h\,k\,l$	I_{exp}	$\|\langle F\rangle\|^2$	φ_{true}	φ
0 0 0	4·60 − 03	5·28 − 03	0·00 + 00	0·00 + 00	0 0 2	1·62 − 03	1·58 − 03	0·00 + 00	0·00 + 00
0 2 0	6·40 − 04	6·06 − 04	3·14 + 00	3·14 + 00	2 1 1	5·19 − 04	4·90 − 04	1·43 + 00	1·50 + 00
2 1 0	4·94 − 04	4·63 − 04	3·14 + 00	3·14 + 00	0 4 0	4·76 − 04	4·24 − 04	3·14 + 00	3·14 + 00
0 4 2	3·92 − 04	3·53 − 04	−3·14 + 00	3·14 + 00	4 0 0	2·74 − 04	2·44 − 04	0·00 + 00	0·00 + 00
5 1 0	2·56 − 04	1·76 − 04	1·57 + 00	1·57 + 00	4 2 0	2·56 − 04	1·89 − 04	3·14 + 00	3·14 + 00
4 1 0	2·53 − 04	1·86 − 04	3·14 + 00	3·14 + 00	2 0 1	2·31 − 04	1·93 − 04	−1·57 + 00	−1·57 + 00
7 2 0	2·12 − 04	1·96 − 04	−1·57 + 00	−1·57 + 00	2 0 0	1·99 − 04	2·30 − 04	0·00 + 00	0·00 + 00
0 3 2	1·94 − 04	1·68 − 04	−1·57 + 00	−1·57 + 00	3 4 0	1·86 − 04	1·36 − 04	1·57 + 00	1·57 + 00
2 1 2	1·59 − 04	1·47 − 04	−2·89 + 00	−3·06 + 00	7 2 1	1·56 − 04	1·41 − 04	−3·12 + 00	−3·01 + 00
3 3 0	1·49 − 04	8·31 − 05	−1·57 + 00	−1·57 + 00	3 2 1	1·47 − 04	1·25 − 04	−2·35 + 00	−2·59 + 00
2 0 2	1·42 − 04	9·64 − 05	0·00 + 00	0·00 + 00	5 2 0	1·40 − 04	1·09 − 04	−1·57 + 00	−1·57 + 00
5 5 0	1·17 − 04	1·11 − 04	−1·57 + 00	−1·57 + 00	0 1 1	1·15 − 04	7·71 − 05	−1·57 + 00	−1·57 + 00
1 0 3	1·04 − 04	8·37 − 05	1·57 + 00	1·57 + 00	6 1 1	1·03 − 04	9·24 − 05	1·57 + 00	1·61 + 00
1 0 1	1·02 − 04	8·56 − 05	1·57 + 00	1·57 + 00	5 1 2	9·91 − 05	9·61 − 05	2·12 + 00	1·81 + 00
3 4 1	9·83 − 05	3·33 − 05	1·29 + 00	6·20 − 01	0 1 2	9·58 − 05	7·83 − 05	1·57 + 00	1·57 + 00
3 3 2	9·16 − 05	4·16 − 05	−1·68 + 00	−1·92 + 00	2 2 2	8·58 − 05	5·85 − 05	−1·61 + 00	−1·93 + 00
2 5 1	8·42 − 05	9·10 − 05	−1·47 + 00	−1·52 + 00	4 3 1	8·41 − 05	5·99 − 05	1·78 + 00	1·66 + 00
0 2 1	8·32 − 05	4·95 − 05	0·00 + 00	0·00 + 00	0 2 2	8·20 − 05	6·88 − 05	3·14 + 00	3·14 + 00
6 0 0	7·91 − 05	7·55 − 05	3·14 + 00	3·14 + 00	3 1 0	7·89 − 05	7·66 − 05	1·57 + 00	1·57 + 00
0 4 1	7·77 − 05	4·46 − 05	−3·14 + 00	3·14 + 00	5 4 1	7·62 − 05	6·98 − 05	1·84 − 01	3·39 − 02
0 6 0	7·09 − 05	6·77 − 05	0·00 + 00	0·00 + 00	3 1 2	6·92 − 05	3·98 − 05	2·03 + 00	1·66 + 00
4 3 0	6·77 − 05	3·73 − 05	0·00 + 00	0·00 + 00	5 2 2	6·76 − 05	5·43 − 05	−1·95 + 00	−1·92 + 00
6 3 0	6·67 − 05	3·94 − 05	0·00 + 00	0·00 + 00	7 2 2	6·52 − 05	5·04 − 05	−1·22 + 00	−1·47 + 00
0 2 3	6·47 − 05	6·10 − 05	0·00 + 00	0·00 + 00	6 2 0	6·34 − 05	3·26 − 05	0·00 + 00	0·00 + 00
2 4 2	6·30 − 05	2·08 − 05	2·90 + 00	2·22 + 00	5 3 2	6·21 − 05	5·09 − 05	−9·71 − 01	−1·20 + 00
5 1 1	6·13 − 05	3·71 − 05	−1·91 + 00	−1·57 + 00	1 3 0	6·00 − 05	3·98 − 05	−1·57 + 00	−1·57 + 00
4 1 1	5·71 − 05	2·11 − 05	1·72 + 00	1·72 + 00	2 2 0	5·63 − 05	3·52 − 05	3·14 + 00	3·14 + 00
1 7 0	5·61 − 05	4·95 − 05	1·57 + 00	1·57 + 00	3 2 0	5·60 − 05	4·23 − 06	−1·57 + 00	−1·57 + 00
5 3 1	5·56 − 05	3·06 − 05	9·60 − 01	1·17 + 00	1 2 1	5·47 − 05	4·79 − 05	−2·03 + 00	−2·23 + 00
5 2 1	5·41 − 05	1·45 − 05	−2·96 + 00	−3·11 + 00	1 6 0	5·23 − 05	2·18 − 05	−1·57 + 00	−1·57 + 00
2 3 1	5·08 − 05	1·98 − 05	2·30 + 00	2·00 + 00	2 3 0	5·02 − 05	4·74 − 05	−3·14 + 00	3·14 + 00
3 2 3	4·98 − 05	2·20 − 05	−2·09 + 00	−2·21 + 00	6 1 0	4·95 − 05	5·37 − 05	3·14 + 00	3·14 + 00
5 3 0	4·94 − 05	5·47 − 05	−1·57 + 00	−1·57 + 00	2 3 3	4·85 − 05	3·23 − 05	2·08 + 00	2·16 + 00
0 3 4	4·79 − 05	3·16 − 05	−1·57 + 00	−1·57 + 00	1 3 2	4·31 − 05	3·20 − 05	−1·12 + 00	−1·43 + 00
1 1 2	4·23 − 05	9·14 − 06	2·71 + 00	1·98 + 00	3 5 1	4·18 − 05	4·44 − 05	−1·01 − 01	2·51 − 02
3 1 1	4·17 − 05	3·43 − 05	−2·30 + 00	−2·44 + 00	9 1 1	3·97 − 05	1·89 − 05	−2·69 + 00	−2·63 + 00
6 3 2	3·92 − 05	3·28 − 05	1·19 + 00	1·07 + 00	3 0 1	3·88 − 05	3·18 − 05	1·57 + 00	1·57 + 00
2 4 1	3·86 − 05	5·38 − 05	1·50 + 00	1·49 + 00	9 3 1	3·76 − 05	3·64 − 05	1·78 − 01	3·24 − 01
2 4 0	3·76 − 05	8·63 − 06	−3·14 + 00	3·14 + 00	0 1 4	3·74 − 05	1·62 − 05	1·57 + 00	1·57 + 00
3 7 0	3·60 − 05	7·64 − 06	1·57 + 00	1·57 + 00	2 1 3	3·58 − 05	4·57 − 05	6·48 − 01	1·29 + 00
0 4 3	3·51 − 05	5·11 − 05	3·14 + 00	3·14 + 00	0 6 2	3·48 − 05	3·41 − 05	0·00 + 00	0·00 + 00
5 1 3	3·43 − 05	2·46 − 05	−1·64 + 00	−1·36 + 00	2 3 2	3·34 − 05	3·25 − 05	2·43 + 00	2·81 + 00
2 5 0	3·30 − 05	3·07 − 05	0·00 + 00	0·00 + 00	4 4 1	3·24 − 05	2·64 − 05	−2·68 + 00	−2·79 + 00
2 5 2	3·21 − 05	1·47 − 05	−5·71 − 01	−1·84 − 01	2 2 3	3·05 − 05	9·44 − 06	2·95 + 00	3·10 + 00
0 4 4	2·97 − 05	3·97 − 05	−3·14 + 00	3·14 + 00	5 7 0	2·91 − 05	7·30 − 06	1·57 + 00	1·57 + 00
2 0 3	2·84 − 05	3·70 − 05	−1·57 + 00	−1·57 + 00	3 0 3	2·75 − 05	1·46 − 05	1·57 + 00	1·57 + 00
3 1 3	2·55 − 05	2·77 − 05	−2·00 + 00	−2·01 + 00	6 6 0	2·51 − 05	2·15 − 05	−3·14 + 00	3·14 + 00
1 2 3	2·46 − 05	2·07 − 05	−1·51 + 00	−1·50 + 00	4 0 2	2·46 − 05	2·71 − 05	0·00 + 00	0·00 + 00
3 4 2	2·44 − 05	1·87 − 05	1·75 + 00	1·63 + 00	2 2 4	2·41 − 05	6·74 − 06	−1·36 + 00	−1·58 + 00
3 2 2	2·28 − 05	1·46 − 06	−7·73 − 01	−3·59 − 01	8 5 0	2·27 − 05	4·29 − 06	−3·14 + 00	3·14 + 00

The first 'time' step was chosen to be $\Delta t_0 = 100$. The density cut-off

$$n(xyz,t) = \begin{cases} 10^{-12} & \text{if } n(xyz,t) < 10^{-12} \\ (1 - 10^{-6}) & \text{if } n(xyz,t) > (1 - 10^{-6}) \\ n(xyz,t) & \text{if } 10^{-12} \le n(xyz,t) \le (1 - 10^{-6}) \end{cases} \quad (35)$$

was used.

As mentioned in § 1 the main shortcoming of the crystal structure determination is associated with relative side minima of the free energy which arrest the relaxation process. This shortcoming can be overcome if the following three conditions are satisfied.

1. The relaxation occurs at the high temperature to suppress the metastable states (to eliminate the relative side minima of the free energy). This temperature,

however, should always be below the critical temperature of the 'order–disorder' transition (Khachaturyan, Semenovskaya & Vainshtein, 1979)

$$T_0 = \nu(1-\nu)\max[a(\mathbf{H})\,I_{\exp}(\mathbf{H})] \qquad (36)$$

[compare with equation (2.62) of Khachaturyan (1978)], where the symbol max[ ] implies the maximum value of the function within the brackets.

2. The initial distribution $n(xyz,0)$ should be close to the state of the absolute minimum corresponding to the real structure of the crystal. Since the real structure of the crystal is unknown we started the crystal structure determination from the initial distribution $n(xyz,0)$ produced by the random-number generator.

3. The R index (13) and (27) depends on a small number of strong intensities $I_{\exp}(hkl)$ [small number of non-zero weight coefficients $a(hkl)$].

A decrease in the number of reflections taken into account in the R index [see (13)] yields a decrease in the number of relative minima of the R index. It improves our chances to attain the absolute minimum of the free energy (31).

In the case under consideration, condition 3 was satisfied because (27) depends on only four normalized intensities. One of these four reflections is 000. Its structure amplitude by definition is equal to ν.*

Three other reflections were selected to meet the following characteristics:

(i) Maximal intensities $I_{\exp}(hkl)$.

(ii) Phases of the selected structure amplitudes may be changed independently by the value π if the appropriate shift of the origin in the unit cell is made.

(iii) Good 'couplings' of the three selected structure amplitudes with all others.

Condition (i) is taken since it is easier to get the coincidence of the calculated and observed amplitudes for weak reflections than for strong ones (Vainshtein & Kayushina, 1966). Thus, strong reflections are more valuable. The selection of a reflection triplet satisfying condition (ii) results in the situation when the phases of the reflections can be considered to be known with reasonable accuracy. Finally, condition (iii) enables us to determine the rest of the structure amplitudes proceeding from the selected triplet. 'Coupling' of amplitudes corresponding to various reciprocal-lattice vectors occurs owing to the nonlinear logarithm term in (27). It provides the interrelation between all amplitudes and their phases and eventually is the sole reason why the phase determination proves to be possible. The simplest way to select a triplet of reflections providing the best 'coupling' is to carry out several iterations (29). The triplet which yields the maximal number of non-zero amplitudes after iteration should be considered as the best one.

The two best triplets for L-proline which possess the three above-mentioned characteristics are

$$211,\ 720,\ 321 \qquad (37a)$$

and

$$211,\ 210,\ 720. \qquad (37b)$$

Regardless of the fact that the kernel of (27) depends on as little as four normalized intensities, the high-temperature solution of (27) at $t \to \infty$ proceeding from the random density distribution $n(xyz,0)$ gives the rough structure of the crystal. The solution was obtained with about 200 successive iterations (29) with the weight factors

$$a(hkl) = \begin{cases} 3 \text{ for } 000 \\ 3,\ 6 \text{ and } 9 \text{ for } 211,\ 720 \\ \quad \text{and } 321, \text{ respectively} \\ 0 \text{ for other reflections} \end{cases} \qquad (38)$$

at the temperature $T = 9 \times 10^{-5}$ close to the critical temperature. The critical temperature T_0 is determined by (36):

$$\begin{aligned} T_0 &= \nu(1-\nu)\max[a(hkl)\,I_{\exp}(hkl)] \\ &= \nu(1-\nu)\,a(211)\,I_{\exp}(211) \\ &= 0{\cdot}068(1-0{\cdot}068)\times 3 \times 5{\cdot}1866 \times 10^{-4} \\ &= 9{\cdot}86 \times 10^{-5}. \end{aligned}$$

The details of the calculation procedure are described at the beginning of this section. The calculation results obtained for the triplet 211, 720, 321 are listed in Table 2. The columns of Table 2 are the reflection numbers, Miller indices (hkl), the array of the normalized intensities $I_{\exp}(hkl)$, the phases $\varphi_{\text{rand}}(hkl)$ corresponding to the random initial distribution $n(xyz,0)$, calculated phases $\varphi_{99}(hkl)$ and $\varphi_{197}(hkl)$ after 99 and 197 iteration cycles and the true phases φ_{true}, respectively.

The surprising characteristic of the proposed technique is that even use of the array consisting of only four reflections yields the true phases of the 23 strongest reflections (compare columns 6 and 7). This takes place although at $t \to \infty$ the linear R index

$$R_L = \sum_{\{hkl\}} p(hkl)\big|[\,|\langle F(hkl)\rangle| - |F_{\exp}(hkl)|\,]\big| \times \left[\sum_{\{hkl\}} p(hkl)|F_{\exp}(hkl)|\right]^{-1} \qquad (39)$$

* The zeroth term of (27) corresponding to $\mathbf{H} = (hkl) = 0$ is $a(0)[(1/N\sum_{\mathbf{r}} n(\mathbf{r}))^2 - \nu^2]\,\nu$. It ensures the conservation condition of 'particles'. The latter enables us to avoid the complication associated with the limitation (8). We, in fact, consider the model crystal as a system with a variable number of particles.

Table 2. *Calculated phases of the* 48 *strongest reflections for* L-*proline after* 99 *and* 197 *iteration cycles at* $T = 9 \times 10^{-5}$ ($T_0 = 9{\cdot}83 \times 10^{-5}$)

Calculation begins with the triplet 211, 720, 320 and random density $n(\mathbf{r}, 0)$. The R_L value is 0·97 after 197 iteration cycles.

N	$h\,k\,l$	I_{exp} ($\times 10^4$)	φ_{rand}	φ_{99}	φ_{197}	φ_{true}
1	0 0 0	46·000	0	0	0	0
2	0 0 2	16·198	π	0	0	0
3	0 2 0	6·401	π	π	π	π
4	2 1 1	5·186	1·578	1·468	1·470	1·428
5	2 1 0	4·940	0	π	π	π
6	0 4 0	4·757	0	π	π	π
7	0 4 2	3·915	π	π	π	π
8	4 0 0	2·736	0	0	0	0
9	5 1 0	2·561	$-\pi/2$	$\pi/2$	$\pi/2$	$\pi/2$
10	4 2 0	2·555	0	π	π	π
11	4 1 0	2·532	π	π	π	π
12	2 0 1	2·311	$\pi/2$	$\pi/2$	$-\pi/2$	$-\pi/2$
13	7 2 0	2·116	$-\pi/2$	$-\pi/2$	$-\pi/2$	$-\pi/2$
14	2 0 0	1·990	π	π	0	0
15	0 3 2	1·940	$\pi/2$	$\pi/2$	$-\pi/2$	$-\pi/2$
16	3 4 0	1·860	$-\pi/2$	$-\pi/2$	$\pi/2$	$\pi/2$
17	2 1 2	1·587	0·496	1·726	−3·038	−2·889
18	7 2 1	1·555	−2·685	−3·044	−3·156	−3·122
19	3 3 0	1·494	$\pi/2$	$\pi/2$	$-\pi/2$	$-\pi/2$
20	3 2 1	1·474	−3·222	−3·187	−3·185	−2·35
21	2 0 2	1·417	0	0	0	0
22	5 2 0	1·397	$-\pi/2$	$-\pi/2$	$-\pi/2$	$-\pi/2$
23	5 5 0	1·169	$\pi/2$	$\pi/2$	$-\pi/2$	$-\pi/2$
24	0 1 1	1·145	$-\pi/2$	$-\pi/2$	$\pi/2$	$-\pi/2$
25	1 0 3	1·037	$\pi/2$	$\pi/2$	$-\pi/2$	$\pi/2$
26	6 1 1	1·029	−1·389	−1·138	1·537	1·565
27	1 0 1	1·023	$-\pi/2$	$-\pi/2$	$-\pi/2$	$\pi/2$
28	5 1 2	0·991	−0·842	1·622	1·628	2·116
29	3 4 1	0·983	0·156	0·056	−0·014	1·293
30	0 1 2	0·957	$\pi/2$	$\pi/2$	$\pi/2$	$\pi/2$
31	3 3 2	0·916	−1·384	−1·550	−1·601	−1·679
32	2 2 2	0·857	1·513	−0·867	−2·753	−1·610
33	2 5 1	0·842	−1·057	−1·700	−1·607	−1·471
34	4 3 1	0·840	1·244	−1·732	−1·635	1·775
35	0 2 1	0·832	π	π	π	0
36	0 2 2	0·819	π	π	π	π
37	6 0 0	0·791	π	π	π	π
38	3 1 0	0·789	$\pi/2$	$\pi/2$	$\pi/2$	$\pi/2$
39	0 4 1	0·777	0	0	π	π
40	5 4 1	0·761	2·561	2·476	−0·019	0·184
41	0 6 0	0·708	0	0	π	0
42	3 1 2	0·691	−1·226	−1·322	1·768	2·027
43	4 3 0	0·677	π	0	π	0
44	5 2 2	0·675	−1·094	−1·830	−1·595	−1·948
45	6 3 0	0·666	0	0	π	0
46	7 2 2	0·652	2·008	−2·195	−1·569	−1·221
47	0 2 3	0·647	0	0	π	0
48	6 2 0	0·634	π	0	π	0

is equal to 0·97 only and the densities are limited by the range $0{\cdot}04 \leq n(xyz,\infty) \leq 0{\cdot}14$. The maximal discrepancy of phases is 48° for 321.*

* Similar results were obtained for the triplet (37*b*). The analysis of 20 other various triplets demonstrated that the ten strongest reflections always have true calculated phases regardless of the fact that the 'coupling' of the triplets under investigation with other reflections is not good enough [condition (iii) is not satisfied].

Table 2 shows that the phases of weak reflections are not improved during the high-temperature relaxation. An improvement may be achieved if the temperature is lowered, the addditional reflections are included in the input intensity array and the new relaxation process starts from the initial distribution $n(\mathbf{r},0)$ designed making use in (24) of the phases of the strongest reflections.

The successive solutions of (27) carried out with the gradual extension of the input intensity array and decrease of the temperature are, actually, the stages of the refinement process. As the temperature decreases, the system approaches the location of the absolute minimum of the R index, the function $n(\mathbf{r},\infty)$ reaching the value 1 inside a sphere, simulating an atom, and 0 outside it. It should, however, be mentioned that the cut-off procedure (35) does not allow one to get to the theoretical limit $R_L \to 0$.

The calculations showed that in the opposite case when the temperature T exceeds the critical temperature T_0 heterogeneities dissipate and at $t \to \infty$ the transition to the disordered state $n(\mathbf{r},\infty) = v$ occurs. The latter is in complete agreement with the theoretical conclusions.

Below we shall give a brief account of the structure determination procedure for L-proline. The procedure consisted of several stages. Each ith stage is the solution of (27) at the temperature T_i when the kernel of (27) depends on the intensities of the M_i strongest reflections from Table 1. The calculation in each ith stage begins with the designing of the initial distribution $n_i(xyz,0)$ at $t = 0$. To do this the phases of the M_i strongest reflections calculated from solution of (27) at $t \to \infty$ in the foregoing $(i-1)$th stage were used in (24). Analogously, the phases of M_{i+1} strongest reflections calculated from the densities $n_i(xyz,\infty)$ were used in the $(i+1)$th stage and so on.

The value M_i in all stages (except the first stage when $M_1 = 4$) is the number of strongest reflections of Table 1, selected for the calculations at the ith stage. The design of the initial distribution $n_i(\mathbf{r},0)$ was always carried out on the basis of (24), where

$$\begin{cases} A_i(hkl,0) = \dfrac{1}{K_i}\sqrt{I_{exp}(hkl)}\cos\varphi_{i-1}(hkl) \\ B_i(hkl,0) = \dfrac{1}{K_i}\sqrt{I_{exp}(hkl)}\sin\varphi_{i-1}(hkl) \\ A_i(000,0) = v \\ B_i(000,0) = 0. \end{cases} \tag{40}$$

$\varphi_{i-1}(hkl)$ is the phase of the reflection hkl calculated in the $(i-1)$th stage [the phases $\varphi_0(211)$, $\varphi_0(720)$, $\varphi_0(321)$ for the first stage are presented in the column for φ_{197} of Table 2], K_i is the attenuation factor. The numerical values of the coefficients K_i are listed in Tables 3 and 4.

Table 3. *The calculation scheme for the six-stage procedure of the crystal structure determination for* L-*proline*

A number M_i is the input intensity array in the ith stage of calculation, T_i is the temperature, K_i is the attenuation factor in equation (40). The R_L^i index is calculated for 100 reflections.

i (number of the stage)	1	2	3	4	5	6
M	4 →	15 →	25 →	35 →	60 →	90 →
T	10^{-5}	10^{-5}	10^{-5}	10^{-5}	10^{-6}	10^{-7}
K	20	20	20	20	10	5
R_L	1·06	0·93	0·85	0·82	0·32	0·21

Table 4. *The calculation scheme for the three-stage procedure of the crystal structure determination for* L-*proline*

The R_L value is calculated for the 100 strongest reflections [equation (39)].

i (number of the stage)	1	2	3
M	4 →	15 →	25 →
T	10^{-5}	10^{-5}	10^{-7}
K	20	20	20
R_L	1·06	0·93	0·45

All amplitudes with the exception of those referred to the M_i reflections of the input array were put to zero.

The weight factors $a(hkl)$ used in (27) in the first stage were

$$a(hkl)=\begin{cases}1 \text{ for } \mathbf{H}=000\\ 5,\ 10 \text{ and } 20 \text{ for } 211,\ 720\\ \quad \text{and } 321 \text{ respectively}^*\\ 0 \text{ for other reflections.}\end{cases} \tag{41}$$

In the following stages

$$a(hkl)=\begin{cases}1 \text{ for } \mathbf{H}=000\\ 5 \text{ for first } M_i \text{ reflections}\\ \quad \text{—in Table 1}\\ 0 \text{ for all others.}\end{cases} \tag{42}$$

The equilibrium states were attained after about 100 iteration cycles of (29) in all stages when the free energy did not change any more. Determination of the L-proline structure consisted of six stages. The characteristics of these stages, *viz* the number of reflections M_i of the input array, the temperature T_i and the attenuation factor K_i are presented in Table 3. Each arrow designates the relaxation process [the solution of (27) from the initial densities $n(\mathbf{r},0)$ to the equilibrium densities $n(\mathbf{r},\infty)$] at given T_i and M_i. The numerical values in the last line of Table 3 describe the linear R_L^i index (39) of the equilibrium distribution calculated over the 100 strongest reflections during the ith stage. Table 1 gives the calculated intensities and phases of these 100 strongest reflections after 150 iteration cycles of the final stage of the crystal structure determination. The R_L value for these 100 reflections is 0·211. All calculated phases were found to be true values (compare the columns φ_{true} and φ in Table 1). The maximal deviation of phases is 41·8° for the 112 reflection. The root-mean-square deviation calculated for the 100 reflections is 10·6°. Fig. 2(*a*) displays the numerical values of the final density function $n_6(xyz,\infty)$ in ten sections normal to the z axis and separated by the distance $\Delta z = 0{\cdot}1$ ($0{\cdot}1c = 0{\cdot}52$ Å). The distances between the nearest sites of the net (30) in the xy plane are 0·05 and correspond to $0{\cdot}05a = 0{\cdot}58$ Å and $0{\cdot}05b = 0{\cdot}451$ Å, respectively. Fig. 2(*a*) shows the irreducible part of the unit cell. The same sections of the model crystal with the true positions of atoms substituted for spheres are displayed in Fig. 2(*b*). The area of the sections of spheres in Fig. 2(*a*) is larger than that in Fig. 2(*b*). This effect was caused by the temperature fluctuations which are inherent to the proposed thermodynamic approach. In fact, the calculations carried out at various temperatures showed that the area of the sphere sections decrease if the temperature is lowered.

Comparison of Figs. 2(*a*) and 2(*b*) shows that the L-proline structure is determined correctly. Small displacements of atomic positions with respect to the true ones are observed for carboxyl group atoms C(1), O(1) and O(2) only. These displacements, however, do not exceed the distance between the nearest sites of the grid and, therefore, are within the calculation accuracy. The other reason for these displacements may be associated with the approximations of the model, *viz* replacements of C, N and O atoms by spheres and ignoring the H-atom contribution to the scattering amplitude.

The densities $n_6(\mathbf{r},\infty)$ displayed by Fig. 2(*a*) were obtained proceeding from the triplet 211, 720, 321. The same results for $n_6(\mathbf{r},\infty)$ were obtained proceeding from the triplet 211, 210, 720.

Fig. 3 shows five sections of the function $n_3(\mathbf{r},\infty)$ calculated in the three-stage procedure with the 25 reflections included in the input array on the final stage (see Table 4). These sections were selected to reveal the places where the difference with the corresponding sections in Fig. 2(*a*) can be observed. Regardless of the large R_L value ($R_L = 0{\cdot}45$), the calculated structure is still close to the true one. Some difference is observed for the C(1) atom in the carboxyl group. Fig. 3 shows that the C(1) atom is located in the section $z = 0{\cdot}85$ instead of the section $z = 0{\cdot}15$. Decrease of density $n_3(\mathbf{r},\infty)$ in the position of the C(4) atom and a slight

* The coefficients $a(hkl)$ for the triplet (37*b*) were 5, 5 and 10, respectively.

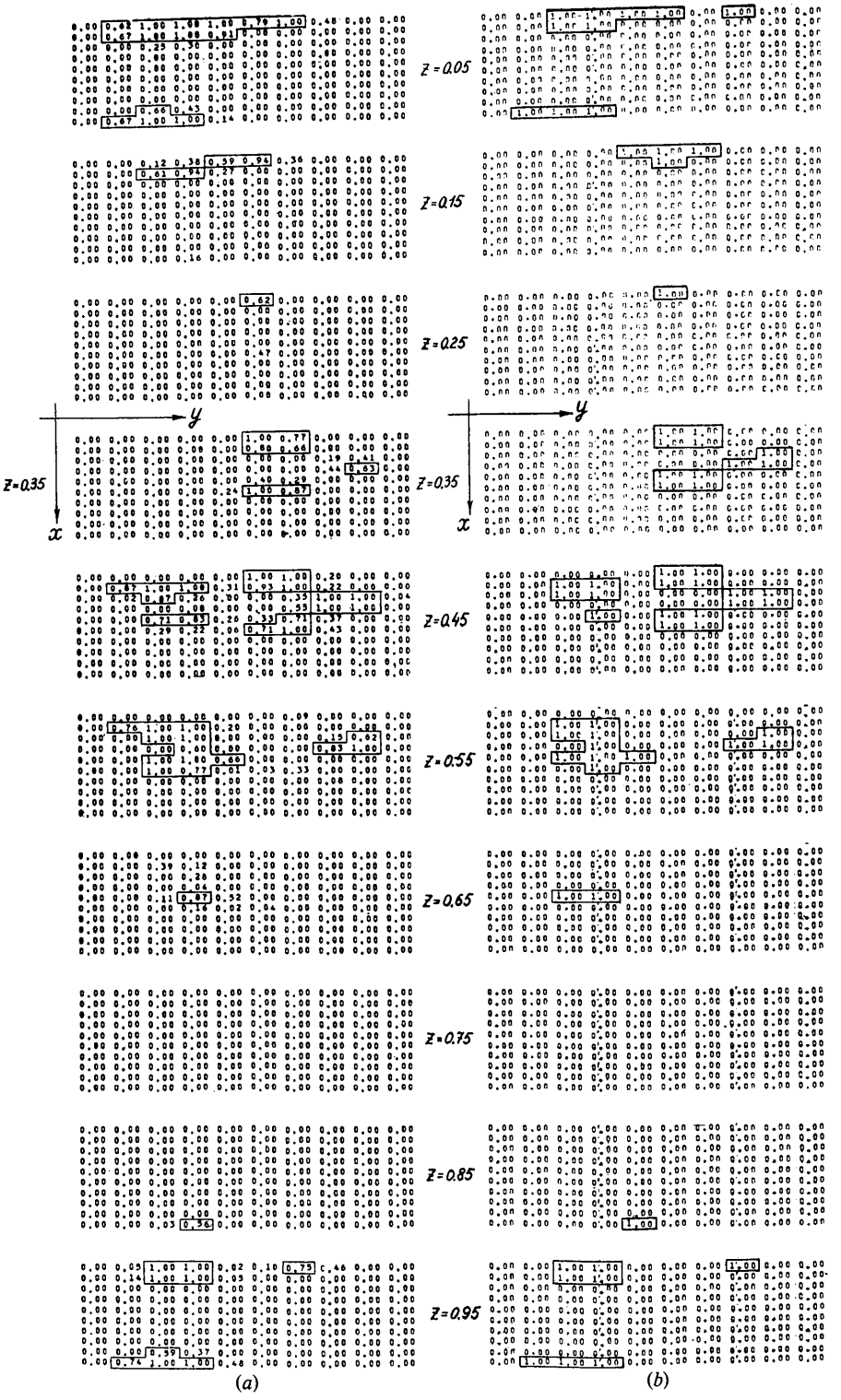

Fig. 2. (*a*) Densities $n_6(xyz,\infty)$ at all $N = 1000$ sites of the grid in the irreducible part of the unit cell of L-proline ($\frac{1}{2}a \times \frac{1}{2}b \times c$) calculated in the six-stage scheme illustrated in Table 3. Each section contains 10×10 sites in the xy plane. The numbers characterize the z coordinate of the section. The solid contours envelop the regions where $n_6(xyz,\infty) \geq 0{\cdot}56$. (*b*) Densities corresponding to the true distribution function for the model crystal assuming the values 1 inside the sphere and zero outside it. Centres of a sphere with the radius $r_0 = 0{\cdot}65$ Å are located in the true positions of C, N, and O atoms of L-proline. The solid contours envelop regions where densities are equal to unity.

shift of the neighbouring C(2) and C(3) atoms of the pyrrolidine ring C(2)C(3)C(5)N by a unit net distance is also observed. The obtained position of the C(1) atom is just a mirror image of its true position with respect to the C(2)C(3)C(5)N ring plane. In fact, this change is equivalent to the realization of the *trans* configuration of the C(4) atom and the carboxyl group C(1)O(1)O(2) with respect to the plane of the pirrolidine ring.* Since the C(4) atom is separated by 0·6 Å from the plane of the pyrrolidine ring (Kayushina & Vainshtein, 1966) the difference between the structure amplitudes of the obtained *trans* configuration and the true *cis* configuration is not considerable and the selection between these two configurations can be made in the later stages of the refining procedure.

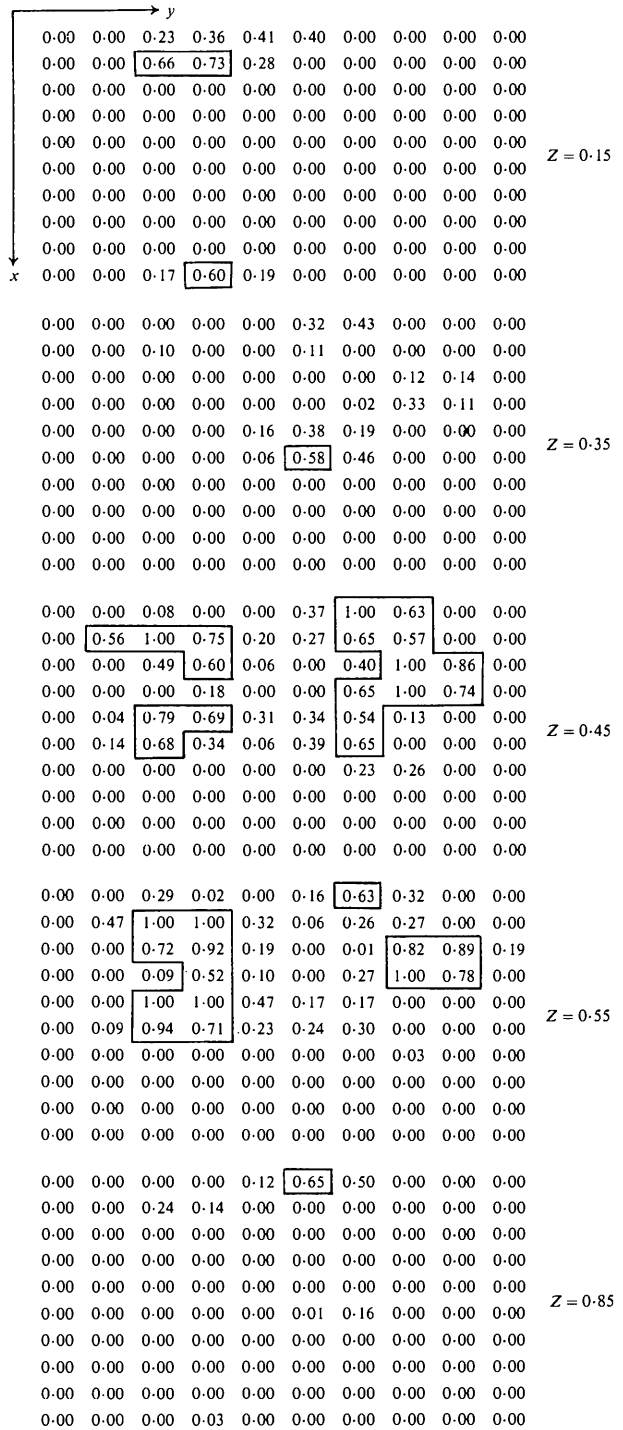

Fig. 3. Densities $n_3(xyz,\infty)$ at five sections of the irreducible part of the unit cell after the three-stage procedure (see Table 4). Each section contains 10 × 10 sites in the xy plane. The numbers characterize the z coordinate of the section. The solid contours envelop regions where $n_3(xyz,\infty) \geq 0{\cdot}52$.

Therefore, the intensity array including even the 25 strongest reflections practically seems to be sufficient for the structure determination. Reduction of the intensity array to the 15 strongest reflections and carrying out the two-stage procedure of the structure determination at $T = 10^{-7}$ yields, in general, the same structure as that presented in Fig. 3. The resolution, however, becomes slightly worse [the C(2) atom is not resolved and $R_L = 0{\cdot}597$]. The comparison of the calculations with various numbers of the strongest reflections taken into consideration (90, 25 and 15) demonstrates that the structure may be solved even if from 15 up to 25 reflections are used. It gives a large time-saving effect. The three-stage crystal structure determination with a 25-reflection array takes about 15 min of computer time (BESM-6). The six-stage calculations with the 90-reflection array take about 1·2 h.

The above considered calculations were carried out for the theoretical intensities reported by Kayushina & Vainshtein (1966). Observed intensities were determined by these authors by the photo method with the visual estimation of intensities. According to modern standards the accuracy of these measurements cannot be considered to be sufficient. However, we undertook similar computer calculations with the observed intensity array (Kayushina & Vainshtein, 1966). The calculations result in the same structure as that obtained for the theoretical intensity array. The calculated densities, however, correspond to the superposition of the *trans* and *cis* configurations of the true L-proline molecules.

The theoretical analysis and many concrete computer calculations which were carried out in the course of this study enable us to believe that the application of the proposed method to more complex compounds with a larger number of atoms per unit cell would not result in other principal difficulties than increase in the computation time. We will test this conclusion by carrying out a crystal structure determination for an organic structure with 54 atoms per irreducible part of the unit cell.

* Some compounds, including proline, form both *trans* and *cis* configurations (Kayushina & Vainshtein, 1966).

In conclusion, it is of importance to note that the proposed method can be readily extended to the cases when the form factors of atoms composing the crystal differ considerably. The improvement of the method would just require the introduction of several density functions for each type of atom and additional computer time.

References

HILL, T. L. (1956). *Statistical Mechanics*. New York: McGraw-Hill.

KAYUSHINA, R. L. & VAINSHTEIN, B. K. (1966). *Sov. Phys. Crystallogr.* **10**, 698–706.

KHACHATURYAN, A. G. (1963). *Sov. Phys. Solid State*, **5**, 16–24.

KHACHATURYAN, A. G. (1968). *Sov. Phys. Solid State*, **9**, 2040–2044.

KHACHATURYAN, A. G. (1978). *Ordering in Substitutional and Interstitial Solid Solutions. Progress in Materials Science*, Vol. 22, edited by B. CHALMERS, J. W. CHRISTIAN & T. MASSALSKY, pp. 1–150. Oxford: Pergamon Press.

KHACHATURYAN, A. G., SEMENOVSKAYA, S. V. & VAINSHTEIN, B. K. (1979). *Sov. Phys. Crystallogr.* **24**, 519–524.

LANDAU, L. D. & LIFCHITZ, E. M. (1958). *Statistical Physics*. London: Addison-Wesley.

MORRIS, J. W. & KHACHATURYAN, A. G. (1978). 107th AIME Ann. Meeting, Denver.

NIGGLI, A., VAND, V. & PEPINSKY, R. (1960). *Acta Cryst.* **13**, 1002.

SURIS, R. A. (1962). *Sov. Phys. Solid State*, **4**, 850–858.

VAINSHTEIN, B. K. & KAYUSHINA, R. L. (1966). *Kristallografiya*, **11**, 526–535.

VAND, V., NIGGLI, A. & PEPINSKY, R. (1960). *Acta Cryst.* **13**, 1001.

Acta Cryst. (1981). A**37**, 754–762

A New Type of Satellite in Plagioclases

BY H. JAGODZINSKI AND B. PENZKOFER*

Institut für Kristallographie und Mineralogie der Universität, Theresienstrasse 41, 8000 *München* 2, *Federal Republic of Germany*

(*Received* 2 *March* 1981; *accepted* 1 *April* 1981)

Abstract

A new type of satellite has been observed in plagioclases $Ca_xNa_{1-x}[Al_{1+x}Si_{3-x}O_8]$ with $0{\cdot}5 \le x \le 1$. Their intensities are such that they could be observed only by applying a special focusing technique. They appear in the neighbourhood of some strong reflections with $h + k = 2n$, $l = 2n$. It is shown that the new satellites are most probably caused by a three-dimensional array of domains of two structures, differing only by small displacements of the atoms. Since the translation lattices of the two structures have the same geometry it is concluded that the two structures are twins of an acentric plagioclase correlated with a centre of symmetry. The size of the domains is 80 Å approximately. The diffraction of such submicroscopically intergrown twins is calculated for lamellae and blocks and compared qualitatively with the experimental results. The approximate periodicity of the domains could be destroyed by very long exposure to X-rays.

* Present address: Deutsches Patentamt, Zweibrückenstrasse 12, 8000 München 22, Federal Republic of Germany.

0567-7394/81/050754-09$01.00

Introduction

It is well known that plagioclases $Ca_xNa_{1-x}[Al_{1+x}Si_{3-x}O_8]$ exhibit a continuous range of miscibility at high temperatures and tend to unmix at low ones. Since two diffusion processes govern the dynamics of this exsolution the various intermediate stages of unmixing are characterized by complicated structures, forming very complex arrays of domains which will be described here only briefly. With the generally adopted notation for indices with respect to the structure of pure anorthite with $a = 8{\cdot}18$, $b = 12{\cdot}88$, $c = 14{\cdot}17$ Å, $\alpha = 93°\,10'$, $\beta = 115°\,51'$, $\gamma = 91°\,13'$, we call

a reflections	$h + k = 2n$, $l = 2n$,
b reflections	$h + k = 2n + 1$, $l = 2n + 1$,
c reflections	$h + k = 2n$, $l = 2n + 1$,
d reflections	$h + k = 2n + 1$, $l = 2n$.

In the composition range $0{\cdot}5 \le x \le 0{\cdot}8$ the following types of satellites have been described by Bown & Gay (1958) while supersatellites have been found by Jagodzinski & Korekawa (1965): *e* satellites