2.5. ELECTRON DIFFRACTION AND ELECTRON MICROSCOPY IN STRUCTURE DETERMINATION
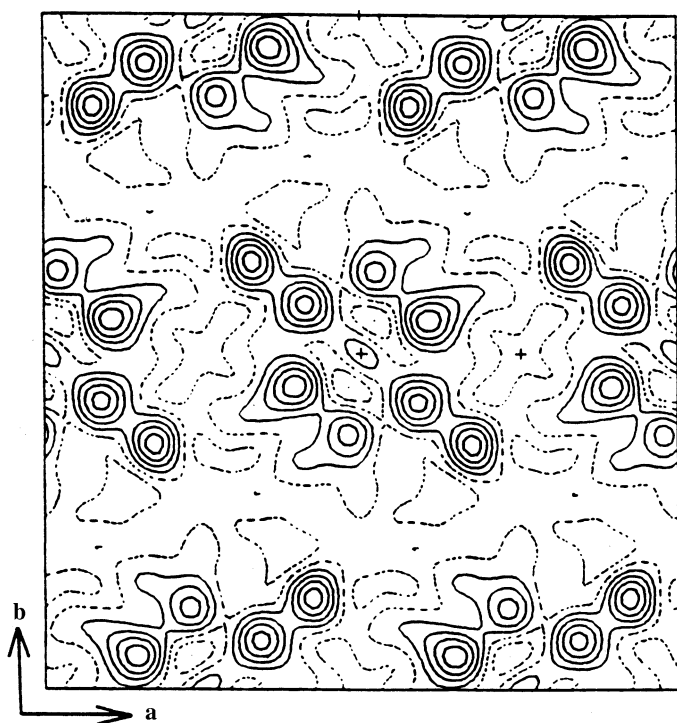

Fig. 2.5.7.1. Potential map for diketopiperazine ([001] projection) after a direct phase determination with texture electron-diffraction intensity data obtained originally by Vainshtein (1955).

microscope image (Dorset *et al.*, 1990, 1993). Most recently, such data have been used to determine the layer packing of a phospholipid binary solid solution (Dorset, 1994*d*).

An *ab initio* direct phase analysis was carried out with zonal electron-diffraction data from copper perchlorophthalocyanine. Using intensities from a *ca* 100 Å thick sample collected at 1.2 MeV, the best map from a phase set with symbolic unknowns retrieves the positions of all the heavy atoms, equivalent to the results of the best images (Uyeda *et al.*, 1978–1979). Using these positions to calculate an initial phase set, the positions of the remaining light C, N atoms were found by Fourier refinement so that the final bond distances and angles were in good agreement with those from X-ray structures of similar compounds (Dorset *et al.*, 1991). A similar analysis has been carried out for the perbromo analogue (Dorset *et al.*, 1992). Although dynamical scattering and secondary scattering significantly perturb the observed intensity data, the total molecular structure can be visualized after a Fourier refinement. Most recently, a three-dimensional structure determination was reported for C_{60} buckminsterfullerene based on symbolic addition with results most in accord with a rotationally disordered molecular packing (Dorset & McCourt, 1994*b*).

2.5.7.4. *The tangent formula*

Given a triple phase relationship

$$\phi_{\mathbf{h}} \simeq \phi_{\mathbf{k}} + \phi_{\mathbf{h-k}},$$

where **h**, **k** and **h** − **k** form a vector sum, it is often possible to find a more reliable estimate of $\phi_{\mathbf{h}}$ when all the possible vectorial contributions to it within the observed data set $\mathbf{k_r}$ are considered as an average, *viz:*

$$\phi_{\mathbf{h}} \simeq \langle \phi_{\mathbf{k}} + \phi_{\mathbf{h-k}} \rangle_{\mathbf{k_r}}.$$

For actual phase determination, this can be formalized as follows. After calculating normalized structure-factor magnitudes $|E_{\mathbf{h}}|$ from

the observed $|F_{\mathbf{h}}|$ to generate all possible phase triples within a reasonably high $A_{\mathbf{h}}$ threshold, new phase values can be estimated after origin definition by use of the tangent formula (Karle & Hauptman, 1956):

$$\tan \phi_{\mathbf{h}} = \frac{\sum_{\mathbf{k_r}} W_{\mathbf{h}} |E_{\mathbf{k}}| |E_{\mathbf{h-k}}| \sin(\phi_{\mathbf{k}} + \phi_{\mathbf{h-k}})}{\sum_{\mathbf{k_r}} W_{\mathbf{h}} |E_{\mathbf{k}}| |E_{\mathbf{h-k}}| \cos(\phi_{\mathbf{k}} + \phi_{\mathbf{h-k}})}.$$

The reliability of the phase estimate depends on the variance $V(\phi_{\mathbf{h}})$, which is directly related to the magnitude of $\alpha_{\mathbf{h}}$, *i.e.*

$$\alpha_{\mathbf{h}}^2 = \left[\sum_{\mathbf{k_r}} A_{\mathbf{h, k}} \cos(\phi_{\mathbf{k}} + \phi_{\mathbf{h-k}}) \right]^2 + \left[\sum_{\mathbf{k_r}} A_{\mathbf{h, k}} \sin(\phi_{\mathbf{k}} + \phi_{\mathbf{h-k}}) \right]^2;$$

$A_{\mathbf{h, k}}$ is identical to the A value defined in the previous section. In the initial stages of phase determination $\alpha_{\mathbf{h}}$ is replaced by an expectation value α_E until enough phases are available to permit its calculation.

The phase solutions indicated by the tangent formula can thus be ranked according to the phase variance and the determination of phases can be made symbolically from the most probable triple-product relationships. This procedure is equivalent to the one described above for the evaluation of phase-invariant sums by symbolic addition. This procedure may allow determination of a large enough basis phase set to produce an interpretable map.

An alternative procedure is to use an automatic version of the tangent formula in a multisolution process. This procedure is described in Chapter 2.2. After origin definition, enough algebraic unknowns are defined (two values if centrosymmetric and four values, cycling through phase quadrants, if noncentrosymmetric) to access as many of the unknown phases as possible. These are used to generate a number of trial phase sets and the likelihood of identifying the correct solution is based on the use of some figure of merit.

Multisolution approaches employing the tangent formula include *MULTAN* (Germain *et al.*, 1971), *QTAN* (Langs & DeTitta, 1975) and *RANTAN* (Yao, 1981). *RANTAN* is a version of *MULTAN* that allows for a larger initial random phase set (with suitable control of weights in the tangent formula). *QTAN* utilizes the $\alpha_{\mathbf{h}_{est}}$ definition, where

$$\alpha_{\mathbf{h}_{est}} = \left\{ \sum_{\mathbf{k}} A_{\mathbf{h, k}}^2 + 2 \sum_{\mathbf{k} \neq \mathbf{k'}} \sum A_{\mathbf{h, k}} A'_{\mathbf{h, k}} \frac{I_1(A_{\mathbf{h, k}}) I_1(A'_{\mathbf{h, k}})}{I_0(A_{\mathbf{h, k}}) I_0(A'_{\mathbf{h, k}})} \right\}^{1/2},$$

for evaluating the phase variance. (Here I_0, I_1 are modified Bessel functions.) After multiple solutions are generated, it is desirable to locate the structurally most relevant phase sets by some figure of merit. There are many that have been suggested (Chapter 2.2). The most useful figure of merit in *QTAN* has been the NQEST (De Titta *et al.*, 1975) estimate of negative quartet invariants (see Chapter 2.2). More recently, this has been superseded by the minimal function (Hauptman, 1993):

$$R(\phi) = \frac{\sum_{\mathbf{h, k}} A_{\mathbf{h, k}} (\cos \phi_{\mathbf{h, k}} - t_{\mathbf{h, k}})^2}{\sum_{\mathbf{h, k}} A_{\mathbf{h, k}}},$$

where $t_{\mathbf{h, k}} = I_1(A_{\mathbf{h, k}})/I_0(A_{\mathbf{h, k}})$ and $\phi_{\mathbf{h, k}} = \phi_{\mathbf{h}} + \phi_{\mathbf{k}} + \phi_{-\mathbf{h-k}}$.

In the first application (Dorset *et al.*, 1979) of multisolution phasing to electron-diffraction data (using the program *QTAN*), *n*-beam dynamical structure factors generated for cytosine and disodium 4-oxypyrimidine-2-sulfinate were used to assess the effect of increasing crystal thickness and electron accelerating voltage on the success of the structure determination. At 100 kV samples at least 80 Å thickness were usable for data collection and at 1000 kV this sample thickness limit could be pushed to 300 Å – or, perhaps, 610 Å if a partial structure were accepted for later Fourier

references