$$\mathbf{q}_k = \mathbf{J}\mathbf{p}_k = \left\{\frac{(1/V)\sum_{\mathbf{h}}[2a(\mathbf{h})\theta(\mathbf{h}) - b(\mathbf{h})]\exp(-2\pi i\mathbf{hx})}{p_k(\mathbf{x})}\right\} = \left[\frac{Q_k(\mathbf{x})}{p_k(\mathbf{x})}\right], \quad (15.1.5.26)$$

where the vector is partitioned as shown above, and

$$a(\mathbf{h}) = (V/N)\sum_{\mathbf{y}}\rho(\mathbf{y})p_k(\mathbf{y})\exp(2\pi i\mathbf{hy}), \quad (15.1.5.27)$$

$$b(\mathbf{h}) = (V/N)\sum_{\mathbf{y}} p_k(\mathbf{y})\exp(2\pi i\mathbf{hy}). \quad (15.1.5.28)$$

Similarly, vector $\mathbf{s}_k$ in equation (15.1.5.16) is obtained from

$$\mathbf{s}_k = \mathbf{J}^T\mathbf{q}_k = (2/V)\rho(\mathbf{x})\sum_{\mathbf{h}}\theta(\overline{\mathbf{h}})[2a(\mathbf{h})\theta(\mathbf{h}) - b(\mathbf{h})]\exp(-2\pi i\mathbf{hx}) - Q_k(\mathbf{x}) + p_k(\mathbf{x}), \quad (15.1.5.29)$$

where $Q_k(\mathbf{x})$ is defined in equation (15.1.5.26).

The remaining calculations in equations (15.1.5.12), (15.1.5.13), (15.1.5.15), (15.1.5.17) and (15.1.5.18) require either the inner product of a pair of vectors or a linear combination of vectors, both of which are very quick to calculate. Each iteration of the conjugate gradient requires four FFTs, as described in equations (15.1.5.26–15.1.5.29).

15.1.5.2.3. *The diagonal approximation*

The full-matrix solution to equation (15.1.5.4) requires a significant amount of computing, although it can be achieved using FFTs. The diagonal approximation to the normal matrix has been used as an alternative method of solution to the electron-density shift in equation (15.1.5.4) (Main, 1990*b*). As with the full-matrix calculation, it can be done entirely by FFTs and a linear combination of vectors.

The diagonal element of the normal matrix, $\mathbf{J}^T\mathbf{J}$, in equation (15.1.5.7) is

$$d_0(\mathbf{x}) = (4/N)\rho(\mathbf{x})\left[\rho(\mathbf{x})\sum_{\mathbf{h}}|\theta(\mathbf{h})|^2 - \sum_{\mathbf{h}}\theta(\mathbf{h})\right] + 2. \quad (15.1.5.30)$$

The right-hand side of equation (15.1.5.7), $-\mathbf{J}^T\varepsilon(\mathbf{x})$, is identical to the residual vector, $r_0(\mathbf{x})$, which can be calculated from equation (15.1.5.22). Therefore, the solution to the electron-density shift, $\delta\rho(\mathbf{x})$, can be calculated from

$$\delta\rho(\mathbf{x}) = r_0(\mathbf{x})/d_0(\mathbf{x}). \quad (15.1.5.31)$$

Compared with the full-matrix solution, all the calculations involved in between equations (15.1.5.12) and (15.1.5.18) and the subsequent iterations are spared in the diagonal approximation. This makes calculation by the diagonal approximation much faster than by the full-matrix method.

15.1.6. Example

To demonstrate the effect of different constraints on phase improvement, various density-modification techniques were applied to an MIR data set for which the refined structure coordinates are available. The test structure is 5-carboxymethyl-2-hydroxymuconate isomerase, solved by Wigley *et al.* (1989). MIR phases were available to 3.7 Å, with SIR information to 2.6 Å. Density modification was used to improve and extend phases to the limit of the data at 2.1 Å. The structure includes threefold noncrystallographic symmetry.

The MIR and density-modified phases are compared by plotting the mean of the cosine of the phase error, weighted by the figure of merit and structure-factor amplitude, as a function of resolution (Zhang *et al.*, 1997),

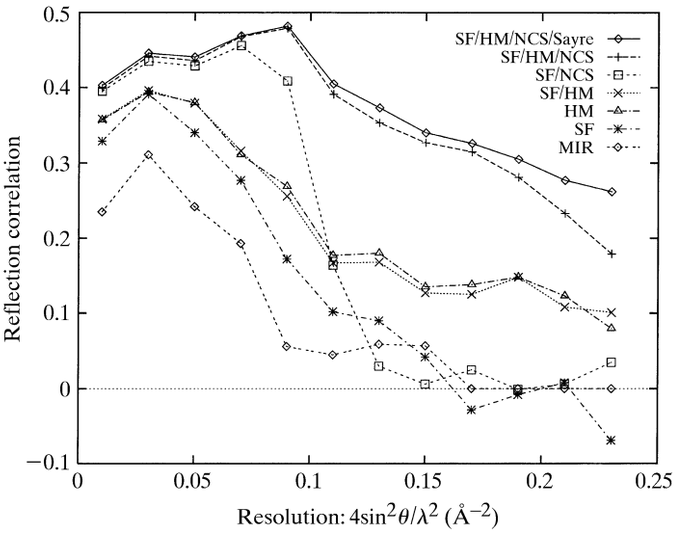

Fig. 15.1.6.1. Phase correlations after different combinations of density modifications.

$$C_f = \left\langle w|F|^2\cos(\varphi - \varphi_0)\right\rangle \Big/ \left(\left\langle w^2|F|^2\right\rangle\left\langle|F|^2\right\rangle\right)^{1/2}. \quad (15.1.6.1)$$

This phase correlation over all reflections is equivalent to map correlation. The results of density modification by various techniques, using the reflection-omit method for phase combination, are shown in Fig. 15.1.6.1.

Solvent flattening alone has slightly improved the phases at low resolution but has not lead to significant phase extension. The solvent-flattening function in Fig. 15.1.3.1 only has nonzero amplitudes close to the origin. It relates structure factors only in a very thin resolution shell. Therefore, solvent flattening is weak on phase extension.

Histogram matching alone improves the low-resolution phases and gives significant phase extension to higher resolutions. The histogram-matching function in Fig. 15.1.3.1 showed much stronger high-resolution amplitudes. Therefore, it could relate structure factors in a larger resolution shell. Moreover, there is always an ideal histogram specified at a given target resolution for phase extension. These two reasons combined make histogram matching a more powerful technique in phase extension than solvent flattening.

The combination of histogram matching and solvent flattening is slightly more powerful than histogram matching alone; since histogram matching sharpens the protein density, it implies an element of solvent flattening. Solvent flattening and averaging give a significant improvement at low resolution, but little phase extension. Averaging is powerful for phase refinement, but is weak for phase extension if no special precautions are taken. If there are flexible loop regions on the protein surface, these regions should be excluded from the molecular mask for averaging. The phasing power of averaging weakens at high resolution when the differences between NCS-related molecules become significant. Solvent flattening, histogram matching and averaging combined give a dramatic improvement at all resolutions. The addition of Sayre's equation gives a slight further improvement at high resolution.

Sayre's equation is very effective for phase refinement and extension at atomic or near atomic resolution. It becomes ineffective at low resolution or when the initial map is poor. Under these circumstances, it is better to apply other density-modification methods first to refine the phases and extend them to a higher resolution before Sayre's equation is applied. Sayre's

equation also decreases in power as the solvent content increases, since it is only applicable to the protein regions of the map.

The fact that the best results were obtained when all the constraints were combined indicates that each constraint contains some degree of independent phasing information. Moreover, it also suggests that the strengths of these constraints are complementary. Each constraint, when applied in isolation, may introduce systematic errors that are difficult to overcome when a different constraint is subsequently applied. This problem is greatly reduced when the constraints are applied simultaneously and the combined process iterates much further towards the desired density map.

Density-modification methods have become sufficiently powerful that it is possible to solve structures from comparatively poor initial maps. This has reduced the amount of effort required to find more heavy-atom derivatives and to collect additional diffraction data sets. Density modification may simplify the process of map interpretation, even when good phase information is available. Density modification can also be used to obtain phases *ab initio* when high-order noncrystallographic symmetry is present.

Acknowledgements

KYJZ acknowledges the US National Institutes of Health for support (grant GM55663). KDC acknowledges the UK BBSRC for support (grant 87/B03785). Some of the material used in this article is reprinted from Cowtan & Zhang (1999) with permission from Elsevier Science.