8 Solvent Flattening and Molecular Envelope Determination

8.1 The Method of Solvent Flattening

Protein crystals typically contain large volumes of solvent. Because of the dynamic structure of the solvent, its electron density has a low constant value when imaged by X-ray diffraction. Once the protein regions of the crystal are identified, the noise peaks in the solvent region can be removed by setting this electron density to a constant low value. Removal of solvent noise peak allows calculation of improved phases and thus the image of the protein can be enhanced. Schervitz et al. (1981) used the solvent flattening procedure to improve isomorphous replacement phases to dramatically improve the electron density map of tRNA\textsubscript{Met}. A novel automated method for defining the molecular envelope around the protein molecules has been proposed in real space by Wang (1985) and modified for reciprocal space by Leslie (1987b). A typical solvent flattening protocol is shown in flowchart form in Figure 17, for both Wang’s real space and Leslie’s reciprocal space algorithms. In Wang’s real space procedure, the electron density is replaced at each grid point by the weighted average electron density of all surrounding grid points within a sphere of radius $R$, using the weighting function

$$w(i) = \begin{cases} 
1 - \frac{r(i)}{R} & \text{for } \rho(i) > 0 \\
0 & \text{for } \rho(i) < 0 
\end{cases} \quad (13)$$

where $\rho(i)$ is the electron density at the grid point $i$ at a distance $r(i)$ from the center of the sphere. The best value of $R$ depends on the resolution, and was
found to be 9 Å and 12 Å for data of 3 Å and 6 Å resolution, respectively (Wang, 1985). In the resulting new map (which should have large connected volumes of relatively high protein electron density and low solvent electron density) the molecular boundary is revealed by tracing a threshold “solvent level” (step 3a in Figure 17).

In order to remove the noise in the map (filtering of errors in direct space), a constant density is added to the protein region, negative density is set to zero and a new average density value is substituted for the solvent region (outside the envelope; step 4a in Figure 17). New structure factor amplitudes and improved phases are then calculated by fast Fourier transform technique (step 5a in Figure 17). A new electron density map is calculated with observed structure factor amplitudes and together with the improved angles obtained from either the solvent flattening procedure alone or by phase combination of the improved solvent flattening phases with the original experimental phases from isomorphous replacement. From the modified Sim probability curve the best phases and figures of merit can be derived and used for calculation of the best electron density map; also this permits one to merge the phase information obtained from density modification with that originally obtained from SIR by multiplying the two phase probability curves. The centroid phase and the corresponding figure of merit can be obtained from the joint phase probability curve. The phase probability distribution $P(\alpha)$ is obtained by Bricogne’s (1976) adoption of Sim’s (1959) weighting scheme

$$P(\alpha) = N \exp \left\{ \frac{2 |F_{\text{obs}}| |F_{\text{calc}}|}{<|F_{\text{obs}}^2| - |F_{\text{calc}}^2|>} \cos(\alpha P - \alpha_{\text{calc}}) \right\} \quad (14)$$
Figure 17: Solvent flattening procedure according to Wang (1985; steps 1a through 7a) and Leslie (1987b; steps 1b through 9b) adopted from Drenth (1994).

* by multiplying $|F_{mod}|$ with the Fourier Transform of $w(i) = 1 - \frac{r(i)}{R}$. 
where $N$ is a normalising constant and $\alpha_P$ and $\alpha_{\text{calc}}$ are the current protein and the phase angle calculated for the map with the flattened solvent density respectively, and $\langle |F_{\text{obs}}|^2 - |F_{\text{calc}}|^2 \rangle$ is the mean difference between the observed and calculated intensities (filtering of errors in reciprocal space; step 6a in Figure 17).
8.2 Detailed Description of the Solvent Flattening Protocol

Leslie's reciprocal space implementation (CCP4, 1994) of Wang's solvent flattening approach was applied. The program TRUNCMAP (step 2b in Figure 17) sets all negative density to zero. It also removed large positive peaks in the map (arising, for example, from ripples about the heavy-atom positions, build-up of errors on crystallographic symmetry axes or the presence of metal ions bound to the protein) which could possibly distort the molecular boundary. The program UNIQUE then listed all possible independent reflections to the desired resolution limit for the specified space group, as it is essential for the reciprocal space algorithm to calculate all possible structure factors (even low resolution terms which were not included in the original SIRAS map) and it is then necessary to provide the back transform program SFC with a list of these reflections. The electron density map was transformed using the fast Fourier transform technique (program SFC) to obtain calculated structure factors (step 3b in Figure 17). The program HKLWEIGHT applies weighting to Fourier coefficients to produce an averaged map [using CCP4 program FFT (CCP4, 1994; Ten Eyck, 1977); step 4b in Figure 17] for the determination of the protein–solvent boundary using Wang's algorithm [program ENVELOPE (Wang, 1985; Leslie, 1987b); step 5b in Figure 17]. The resultant molecular envelope was used to modify the SIRAS map by setting the solvent density to zero and truncating negative density in the protein region (program FLATMAP; step 6b in Figure 17). The solvent–flattened density was back–transformed (program BACKTRANS; step 7b in Figure 17) to give improved phases. The
calculated (solvent–flattened) and observed structure factors were scaled with
the program SCALENEW which also calculates Sim weights. Bricogne’s program
COMBINE combined the calculated phases with the SIRAS phases (step 8b in
Figure 17). The phase probability profiles were conveniently held in the form
of Hendrickson–Lattman coefficients (Hendrickson & Lattman, 1970), which
allows phase combination by simple addition of the coefficients. An ISIRAS
(iterative isomorphous replacement method with anomalous scattering) map
was calculated (program FFT; step 9b in Figure 17) and the noise filtering pro-
cedure could be repeated iteratively [programs FLATMAP (Wang, 1985; Leslie,
1987b), BACKTRANS, SCALENEW and COMBINE]. The results converged after four
cycles of iteration, after which a new molecular envelope was determined. The
ISIRAS map was returned to the program TRUNCMAP, and the entire procedure
described was followed repeatedly until no further improvement occurred in
the quality of the electron density.
8.3 Solvent Flattening of Neu5Ac lyase and Results

The solvent flattening procedure was performed on the SIRAS map calculated in space group \(P3_121\). The location of the tetramer within the unit cell was confirmed by inspection of the resultant solvent–flattened 6 Å resolution ISIRAS (iterative single isomorphous replacement with anomalous scattering) map (shown in Figure 1.1). Only low resolution data were used, as the phase angles of the high resolution data were still very poor. In order to determine a molecular envelope, a SIRAS map was subjected to Leslie's automatic envelope determination algorithm using a radius \(R\) of 15 Å and a solvent content of 0.55, which is slightly less than that estimated based on one tetramer in the asymmetric unit (0.56) in order not to cut off the outer loops of the protein molecule. The envelope was updated every four cycles, as the improved electron density map allows tracing of a better envelope. The map improved when the envelope was changed more substantially than after further iteration. The initial envelope was then edited manually on the basis of the SIRAS and the ISIRAS map and finally symmetrised to be consistent with the local 222 point group and placed from the skewed frame into the asymmetric unit in the trigonal space group. The average figure of merit in the final cycle was 0.98 with an \(R\)-factor \(\left(\frac{\sum|F_1-F_2|}{\sum 0.5|F_1+F_2|}\right)\) of 0.31 for 4096 reflections. The envelope did not seem to overlap with any symmetrically related envelopes. Figure 1.1 shows a 10 Å thick slab of the new ISIRAS map after three rounds of the noise-filtering solvent–flattening procedure with four cycles of iteration each.
Figure 18: Projection of a 10 Å thick slab of the map calculated from phases obtained using data between 6 Å and 20 Å resolution contoured at 1σ with 1σ increments. Compared with the same sections of the SIRAS map shown in Figure 16, this map now clearly shows enhancement of the density in protein regions.