Acta Cryst. (1997). D**53**, 426–433

Procedure to Calculate the Molecular Envelope from a Partial Model†

RENÉ COULOMBE AND MIROSLAW CYGLER

Biotechnology Research Institute, NRC, 6100 Royalmount Avenue, Montréal, Québec H4P 2R2, Canada.
E-mail: mirek.cygler@bri.nrc.ca

(*Received 21 October 1996; accepted 21 January 1997*)

Abstract

Density-modification algorithms have gained, in recent years, a widespread use in the early stages of protein structure determination, especially in combination with the single isomorphous replacement, the multiple isomorphous replacement and the multiple-wavelength anomalous dispersion methods, where density modification usually leads to a significant improvement in the quality and interpretability of the initial electron-density map. The current computer programs which are used to perform this task combine several approaches, an important component of which is the solvent-flattening procedure. The latter procedure depends crucially on the correct determination of the molecular envelope. The solvent-flattening procedure has also been applied to the electron-density maps calculated from partial models obtained from the molecular replacement method. In such case the envelope calculated in the standard way does not always encompass entirely the missing part. It has been found that the standard application of the density-modification method (as implemented by programs *SQUASH* and *DM*) to a map calculated from a molecular replacement model containing ~60% of the molecule, led to little improvement in the map interpretability. Here, it is shown that a significant improvement of the map can be achieved when a better envelope is used in the procedure. Various methods of calculating the molecular envelope have been evaluated, the effect of the shape of the envelope on the modified electron-density map has been investigated and an improved procedure to calculate the envelope from a partial molecular replacement model is proposed.

1. Introduction

The electron density *a priori* has to conform to certain conditions, like non-negativity and boundedness. These conditions are usually not strictly fulfilled by the electron-density distributions calculated from the experimental data due to errors in the amplitudes and phases. The idea of modification of the experimental electron density to make it conform to the theoretical expectations was exploited for the purpose of improving the phases more than 30 years ago (*e.g.*, Hoppe & Gassmann, 1968) and was initially applied to small-molecule structures with the hope of leading to a direct approach of solving crystal structures. The first suggestions for the usefulness of density modification to improve the protein electron-density maps were associated with the concept of non-crystallographic symmetry averaging (Rossmann & Blow, 1963; Main & Rossmann, 1966) and were mathematically analyzed by Bricogne (1974). He later developed a suite of programs for direct-space non-crystallographic symmetry averaging and formalized the concept of the molecular envelope (Bricogne, 1976). The application of density modification to the resolution of the phase ambiguity inherent to single isomorphous replacement or single anomalous substitution phases for proteins was developed by Wang (1985) following the premise that the electron density of the solvent, outside of the protein, should be flat. An implemented algorithm for an automated molecular envelope determination contributed to a widespread application of his procedure. Since a protein crystal contains usually at least 45%, and often more, of its volume filled by solvent, a large part of the map is affected by density modification. This technique, called solvent flattening, has since been successfully used for improving electron-density maps obtained from various sources of phases. The method strongly depends on the proper assignment of the volume that is occupied by the molecule (molecular envelope) and that occupied by the solvent. The usefulness of solvent flattening, especially when combined with other approaches such as histogram matching (Lunin, 1993) and Sayre's equation (Cowtan & Main, 1993), has been well proven in recent years when applied to electron-density maps calculated with initial phases obtained from the multiple or single isomorphous replacement (MIR, SIR) or from the multiple-wavelength anomalous dispersion (MAD) techniques and this method is very commonly applied in crystallographic practice. Density modification can also be a powerful tool when the initial phases are obtained from a molecular replacement (MR) model. The rapidly increasing number of known protein structures means that there is an ever increasing chance for a protein to have at least one domain in common with a related structure in the Protein Data Bank. The large database of known structures makes for the ever increasing application of the molecular replacement method. However, the phases obtained from this method are bi-

† NRCC publication 39960.

ased toward the molecular replacement model. When the molecular replacement model (known part) encompasses only a part of the whole structure or when the model is only modestly similar to the structure of the protein of interest, the resulting electron-density map is not easily interpretable beyond the initial molecular replacement model. Density-modification methods could significantly improve this map and simplify the tracing of the missing parts of the model. Since all present density-modification programs include the solvent-flattening step as a major part of the algorithm, the challenging problem is the proper choice of the molecular envelope starting from a map that is strongly biased toward a partial model. Here, we analyze an example of the application of density-modification method to the structure of procathepsin L determined from a partial model, recount the difficulties related to the proper definition of the molecular envelope and evaluate the impact of the envelope on the improvement of the electron density by the procedure.

2. Materials and methods

Procathepsin L crystallized in space group $P2_12_12_1$ with cell dimensions $a = 40.1$, $b = 88.1$, $c = 94.9$ Å and one molecule in the asymmetric unit. A native data set to 2.2 Å resolution was collected at the station BL6A2, Photon Factory synchrotron facility (Tsukuba, Japan), using a Weissenberg camera and the wavelength $\lambda = 1.00$ Å. The frames were processed with the program *WEIS* and yielded a total of 46 120 observations that merged to 13 872 unique reflections with $R_{\rm merge} = 0.069$. This data set is 87% complete to 2.5 Å resolution and 78% complete to 2.2 Å resolution (Coulombe, Li *et al.*, 1996).

Mature cathepsin L shares ~40% amino-acid identity with three other members of the papain superfamily with known three-dimensional structures: papain (PDB code 9pap), actinidin (1aec) and caricain (1ppo). Using each of them as a molecular replacement model yielded a clear solution for both the rotation and translation function (*AMoRe*, Navaza, 1994). The highest correlation coefficient and the lowest R factor were for the actinidin model. Further improvement of the agreement indices was obtained by removing three loops that differed significantly between the three known structures (Coulombe, Grochulski *et al.*, 1996). The $2F_o - F_c$ and $3F_o - F_c$ electron-density maps calculated with the phases derived from this model were reasonably good within the model but unclear outside of it.

The actinidin side chains were changed to those of cathepsin L guided by the electron-density maps and the resulting model was rigid-body minimized with *X-PLOR* (Brünger, 1992). This partial model included residues 1–57, 62–98, 109–172 and 181–219 of cathepsin L (195 residues in total) of which ten were represented as alanines. The R factor for this model was 0.40 for reflections in the 8–2.2 Å resolution range. We refer to this coordinate set as the MR model. Further minimization of this model with *X-PLOR* resulted in a decrease of the R factor to 0.33 but with no significant decrease of $R_{\rm free}$ (0.41) suggesting an overfitting. For this reason we have used the rigid-body-refined MR model as a starting point in all calculations described below, unless stated otherwise (see flowchart in Fig. 1). The fully refined model of procathepsin L used for comparative analysis contained residues 5p–96p of the proregion and 1–174, 180–220 of mature cathepsin L. Residues 1p–4p and 175–179 were disordered in the crystal and were not included in the model. The R factor for the final model, which included 71 solvent molecules, is 0.182 ($R_{\rm free} = 0.241$).

The electron-density maps were calculated with the programs *SFALL* and *FFT* from the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994). Molecular envelopes were determined with the programs *ENVELOPE* (Leslie, 1987) or *SOLOMON* (Abrahams & Leslie, 1996). For the former an averaging radius of 8 Å was used in all calculations. For *SOLOMON* the radius parameter was set to 2.5 Å for the $F_{\rm obs}$ map and to 4.5 Å for the $(3F_{\rm obs} - 2F_c)$ map. Map manipulation was performed with the program *MAPMAN* from the *O* package (Kleywegt & Jones, 1996). Correlation coefficients between the maps were calculated with programs specially written for this purpose. Apart from comparing the entire map, a comparison of only the region corresponding to the partial model or only the missing part of the model was carried out as well, as this was a more sensitive indicator of the map quality. These comparisons were performed for densities within a specific envelope, either for the partial model or for the missing part. The electron densities at atomic positions were calculated by interpolating the electron-density map. These values were used to obtain an average electron density for the backbone atoms for the known and unknown parts of the model. Since all the electron-density maps were calculated with the same structure factor amplitudes, $F_{\rm obs}$, the average densities were expressed relative to those derived from the final electron-density map (based on the fully refined model, without solvent). The program *SIGMAA* (Read, 1986) was used to obtain estimates of weights for phases derived from the model for phase combination. Density modification was initially performed with the program *DM* (Cowtan, 1994), which combines solvent flattening and histogram matching, and with the program *SQUASH* (Zhang, 1993). In both cases a comparably small improvement in the interpretability of the resulting maps was found. Subsequent calculations described below which tested various molecular envelopes were performed only with the *DM* program. The program allows for easy combination of a mask calculated from one map to be applied for the density modification of another map. The flowchart of the *DM* procedure is

shown in Fig. 1. Resulting maps and masks were visually inspected with the program *O*.

3. Results and discussion

Procathepsin L is the proform of the cysteine protease cathepsin L, which belongs to the papain superfamily. It is synthesized in the Golgi apparatus and directed to lysosomes. In the acidic environment of the lysosomal compartment the proregion is cleaved, liberating the active protease, whose main role is in degradation of proteins in the lysosome. The mature protein has 220 residues and the prosegment is 96 residues long. The orthorhombic crystals of procathepsin L contain 48% solvent and diffraction data to 2.2 Å resolution were collected. These data were used in our initial attempts to solve the structure. There are good models for the structure of cathepsin L and we thought that, despite the initial model containing only ~62% of the procathepsin L, density modification should improve the initial electron-density map to allow the tracing of the proregion. This was not so for the orthorhombic crystal form. The electron-density map was not interpretable beyond the molecular replacement model. Standard application of density modification method using programs like *SQUASH* (Zhang, 1993), *PHASES* (Furey & Swaminathan, 1990) and *DM* (Cowtan, 1994) did not improve the interpretability of the map. The structure was eventually solved when a different crystal form with a higher solvent content became available (Coulombe, Grochulski *et al.*, 1996). Fig. 2 shows the structure of procathepsin L with the parts not included in the MR model shown in light shades. A similar case of a MR solution with a large missing fragment was reported recently by Rudenko, Bonten, d'Azzo & Hol (1996), where the solvent flattening with various, manually modified envelopes, was aided by the twofold non-crystallographic symmetry to help solve the structure. We have undertaken a detailed analysis to understand the reasons for the failure of density-modification methods to improve sufficiently the initial electron-density map. As a result of this analysis we propose several ways by which a significantly improved molecular envelope can be obtained.

Although it was possible that solvent flattening made little improvement due to a relatively low solvent content, approximately 48%, we thought that a more severe problem was an inaccurate envelope determination. In

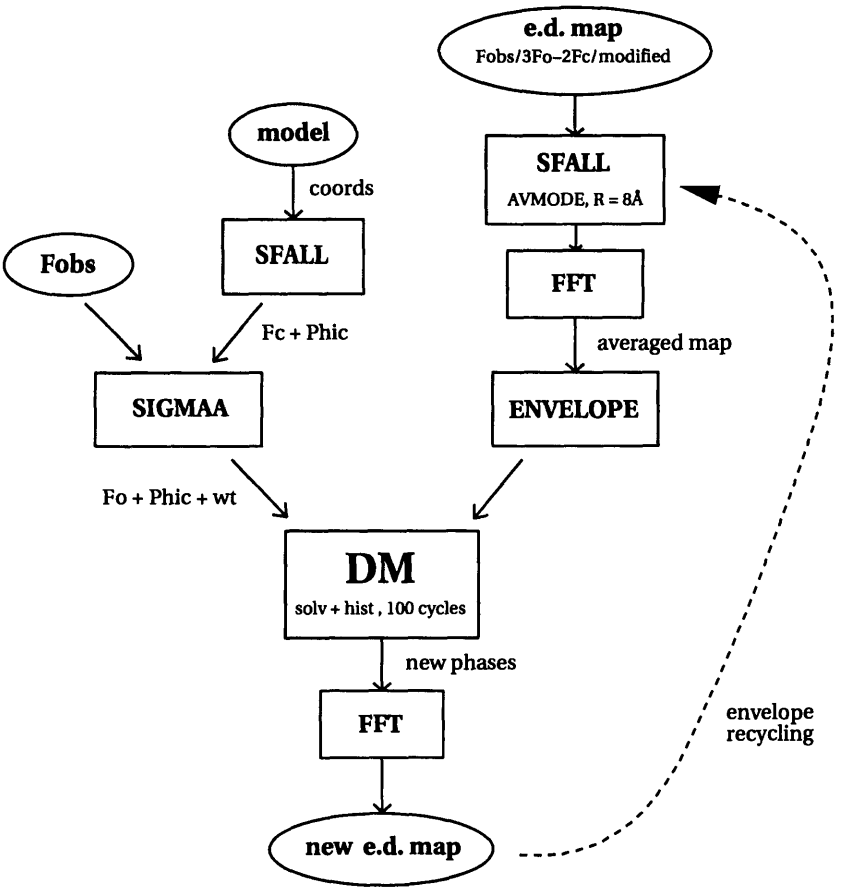

Fig. 1. Flowchart of the density-modification procedure as executed with the *DM* program.

the trials described below we have used a slightly higher solvent content, 52%, which produced the best results. The differences were not large but the improvement correlated with a decrease of the R_{free} index. Thus, the R_{free} calculated by the *DM* procedure could be used as an indicator to select the optimal solvent content for further calculations.

To start, we compared the map calculated with phases derived from the MR model to the final map which was calculated from the refined coordinates, excluding solvent atoms. The correlation coefficient for the whole map was 0.72; within the known part the correlation was much higher, 0.84, but within the missing part the two maps showed only a correlation of 0.56. The MR derived map was not interpretable within the missing part volume.

3.1. *Reference and standard envelopes*

In the next step of our analysis we calculated the best molecular envelope (mask) for the procathepsin L from the final electron-density map calculated with the phases derived from the refined coordinates (excluding solvent molecules). As in the calculations of other envelopes we have assumed a solvent content of 52% and an averaging radius of 8 Å (Fig. 3*a*). This was our reference envelope to which other envelopes were compared in subsequent calculations. Since all the envelopes had roughly the same volume, we have judged the similarity of two envelopes by the percentage of the volume that was common to both of them (see Table 1).

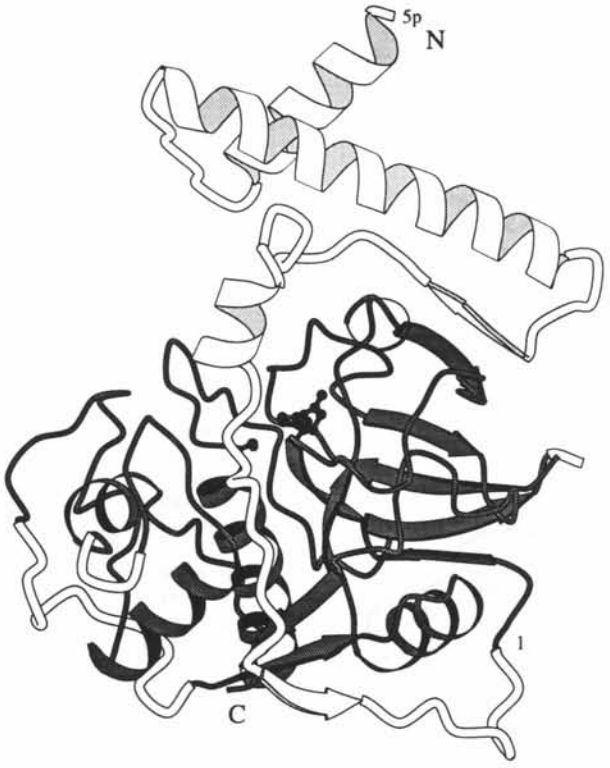

Fig. 2. The cartoon representation of the procathepsin L structure. The parts that corresponding to the MR model are shown by dark shading, those not included in the model are lightly shaded.

Next, the molecular envelope was calculated in a standard way, from the F_{obs} map phased with the MR model. We will call it the standard envelope. When the standard envelope was compared to the reference envelope their common volume was 77.6% of the reference envelope volume (Table 1). This relatively low agreement immediately indicated that the problems we were experiencing were largely related to an inadequate envelope. On visual inspection it became obvious that the standard envelope covers only a small part of the prosegment (Fig. 3*b*). The modified map resulting from the *DM* procedure with the standard envelope was also compared with the final electron-density map. The overall correlation coefficient between the two maps was 0.75; while within the known part the correlation between the maps was 0.85 it dropped to only 0.62 in the prosegment area (Table 1).

3.2. *MR model with reference envelope*

To substantiate our conviction that the correct envelope should lead to an interpretable map we have applied the *DM* procedure with the same starting phases from the MR model but using the reference envelope. The resulting map was significantly better then the previously calculated map and had a correlation coefficient of 0.80 to the final F_{obs} map. The improvement was mostly in the prosegment region where the correlation between the maps increased to 0.74, while within the known part the correlation increased only slightly to 0.87 (Table 1). The relative average electron density for the backbone atoms of the known part was 0.95 while it was 0.42 for the unknown part. A visual inspection of this map indicated that the prosegment could be traced in this map with minor difficulties.

3.3. *Partial cathepsin L model with standard envelope*

To find out to what extent the inaccuracies of the partial model influence the modified map we applied the *DM* procedure to the map calculated with the phases derived from the same residues as the MR model but taken from the final refined procathepsin L structure (partial pcatL), and used the standard envelope. The correlation coefficient of this map was 0.86 for the whole map and 0.75 for the prosegment part. The relative average electron density for the backbone atoms of the known and unknown parts was 0.93 and 0.29, respectively. The improvement from the best partial phases was then of the same order as that from the best envelope. However, the MR model, which originated from a close relative, already represented a rebuilt model and its further improvement was not straightforward.

Table 1. *Comparison of molecular envelopes (masks) and electron-density maps obtained from various starting phases and masks*

DM calculations				Correlation coefficient*				Relative average electron density†	
Phases from model	Envelope	R_{free} (*DM*)	Envelope overlap‡ (%)	Entire map	Known part	Unknown part	Solvent	Known part	Unknown part
pcatL§	—	—	100	1.000	1.000	1.000	1.000	1.00	1.00
MR model	None¶	—	—	0.718	0.838	0.558	0.350	1.02	0.22
MR model	Standard**	0.399	77.6	0.746	0.853	0.620	0.390	0.93	0.25
	cycle 2	0.397	78.3	0.748	0.854	0.625	0.393	0.93	0.26
MR model	Reference††	0.353	100	0.804	0.872	0.741	0.450	0.95	0.42
pcatL part‡‡	Standard**	0.407	77.6	0.856	0.950	0.752	0.575	0.93	0.29
MR model	From $3F_o - 2F_c$§§	0.378	84.7	0.771	0.861	0.674	0.412	0.94	0.34
MR model	From $2mF_o - DF_c$	0.386	86.3	0.769	0.859	0.675	0.414	0.93	0.31
MR model	Combined *A*¶¶	0.381	80.0	0.766	0.859	0.663	0.408	0.95	0.31
MR model	Combined *B****	0.382	86.9	0.768	0.858	0.664	0.397	0.94	0.32
MR model	Combined *C*†††	0.373	87.5	0.776	0.862	0.688	0.419	0.94	0.35
	cycle 2	0.366	91.1	0.789	0.867	0.709	0.438	0.95	0.38
	cycle 3	0.362	92.0	0.792	0.868	0.716	0.440	0.95	0.39
MR model	Solomon (F_o)‡‡‡	0.394	79.6	0.754	0.854	0.637	0.440	0.93	0.28
MR model	Solomon ($3F_o - 2F_c$)‡‡‡	0.371	87.0	0.772	0.860	0.678	0.415	0.94	0.32
	cycle 2	0.399	87.4	0.781	0.860	0.702	0.423	0.92	0.39
pcatL part‡‡	Reference††	0.331	100	0.900	0.961	0.850	0.642	1.04	0.49
pcatL part‡‡	Combined *C*†††	0.363	87.5	0.879	0.955	0.798	0.611	1.03	0.41

* Correlation coefficient between the specified map and final electron-density map calculated from the refined pcatL model. † Relative average electron density interpolated at positions corresponding to the location of the atoms in the final pcatL model in relative units. The electron-density maps were calculated with unscaled F_{obs} and the average density is expressed as a fraction of that for the final map, scaled independently for the known and unknown part. ‡ Percentage of the volume in common with the reference envelope. § Final refined model of procathepsin L. ¶ Initial map, no density modification. ** Standard envelope calculated from the F_{obs} map with phases derived from the MR model. †† Envelope calculated from the final F_{obs} map based on refined pcatL coordinates. ‡‡ pcatL part - counterpart of MR model in the final pcatL model. §§ Envelope calculated from the $(3F_o - 2F_c)$ map with phases derived from the MR model. ¶¶ Combined *A* envelope calculated from $(F_o - F_c)$ map as described in the text. *** Combined *B* envelope calculated from F_{obs} map as described in the text. ††† Combined *C* envelope calculated from F_{obs} map and $(3F_o - 2F_c)$ map as described in the text. ‡‡‡ Envelope calculated by the program *SOLOMON* either from the F_{obs} map or from the $(3F_o - 2F_c)$ map phased by the MR model.

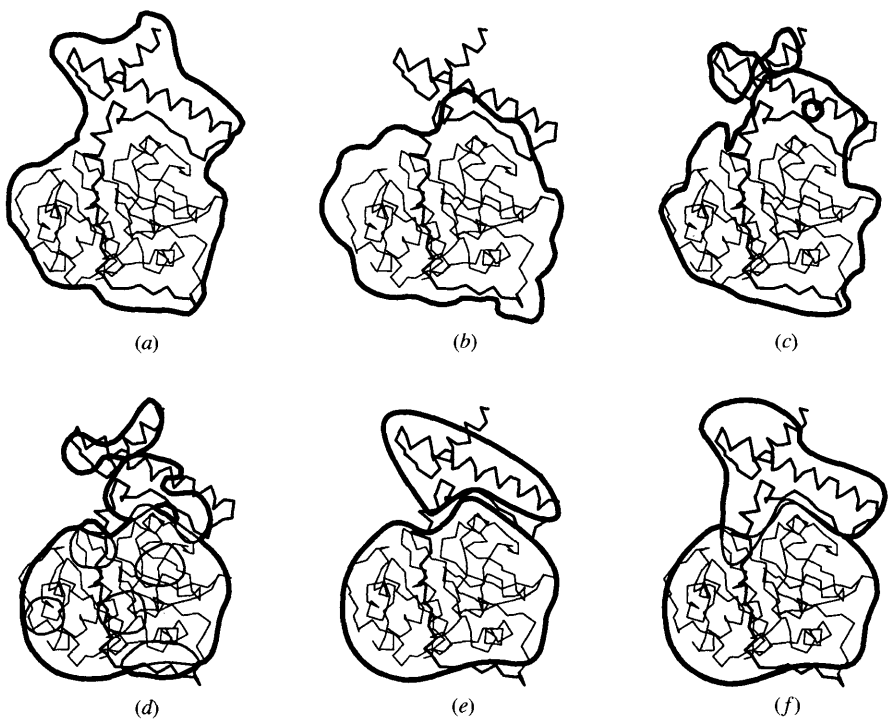

Fig. 3. Qualitative representation of the molecular envelopes calculated from various maps using identical parameters. The integration radius was 8 Å in each case. (*a*) Reference envelope, based on the final electron-density map of the refined structure without solvent molecules. (*b*) Standard envelope, based on the F_{obs} map phased by the MR model, (*c*) Envelope based on the $3F_{obs} - 2F_{calc}$ map phased by the MR model. (*d*) 'Combined *A*' envelope. (*e*) 'Combined *B*' envelope. (*f*) 'Combined *C*' envelope.

3.4. *Partial cathepsin L model and reference envelope*

For comparison, we have also performed these calculations with the partial pcatL model using the reference envelope. This is the best that can be achieved starting from this partial model and provides a reference point for other comparisons. The resulting map agreed very well with the final map: the correlation coefficient was 0.90 for the whole map and 0.85 for the prosegment area (Table 1). The relative average electron density was 1.04 for the known part and 0.49 for the unknown. The latter value was significantly higher than when the standard envelope was used.

The calculations above show clearly that in the case of initial phases derived from a partial model, the crucial factor determining the success or failure of the density modifiation method is the shape of the envelope. It is also evident that automated mask calculation from the F_{obs} map by the standard algorithm was not able to provide good coverage of the missing prosegment and hence little improvement in the interpretability of the modified map was noted. While there was not much more that could be done to improve the MR model at this stage, there was much that could be done to improve the envelope.

3.5. *Envelope based on a $(3F_{obs}-2F_{calc})$ map or a $(2mF_{obs}-DF_{calc})$ map*

The electron-density map calculated from F_{obs} coefficients and phases based on a partial model is strongly biased toward the partial model. The electron density for parts of the molecule not included in the model are significantly weaker than the known parts. A map calculated with Sim weights (Sim, 1960) or with the modified coefficients such as $(2F_{obs}-F_{calc})$, $(2mF_{obs}-DF_{calc})$ or even better $(3F_{obs}-2F_{calc})$, should be less biased toward the model and have more pronounced electron density corresponding to the unknown part. Such a map should provide a better starting point for the calculation of the molecular envelope. Table 1 shows that, indeed, the mask obtained from a $(3F_{obs}-2F_{calc})$ map has 84.7% of its volume in common with the reference mask. The density-modification algorithm with this envelope resulted in a map that had a correlation coefficient of 0.77 to the final map and 0.67 in the prosegment area. Visual inspection of this molecular envelope showed that it covers some parts of the prosegment but parts with weak density are not covered (Fig. 3*c*). Starting with the $(2mF_{obs}-DF_{calc})$ map led to almost the same results (Table 1).

3.6. *Decoupling of envelope calculation for the known and unknown parts*

The calculation of an envelope for the missing prosegment in the F_{obs} map is hampered by its weak electron density that resides next to a much stronger electron density of the partial model. The map calculated with $(3F_{obs}-2F_{calc})$ coefficients has an elevated electron density for the unknown part and thus leads to an envelope that covers more of the missing part of the molecule (Fig. 3*c*). However, we thought that a better result would be obtained if the calculation of the envelope for the unknown part could be carried out separately from the calculation of the envelope for the known part. For this purpose a map should be constructed with a negligible contribution from the known part of the model. This task can be performed either in reciprocal space by calculating the $(F_{obs}-F_{calc})$ difference map, or in real space, by eliminating the density in the volume taken by the partial model. We have tried both aproaches and concluded that working in real space provides more flexibility in map manipulation.

3.6.1. *Reciprocal space.* The envelope was calculated in two steps. First, the envelope for the known part was calculated in a standard way based on an F_{obs} map phased with the MR model. Since the known part is approximately $p = 0.62$ of the whole molecule and the fraction of the unit-cell volume taken by the molecule is $V_{mol} = 0.52$, the known part takes up approximately $V_{known} = pV_{mol} = 0.62 \times 0.52\ (0.35)$ of the total volume of the asymmetric unit. The missing prosegment takes approximately $V_{mol} - V_{known} = V_{missing}\ (0.15)$ of the volume. To calculate the envelope for the known part we have used, then, a solvent content of $(1 - V_{known}) = 0.65$. Second, we used the $(F_{obs}-F_{calc})$ difference map to calculate the envelope for the missing part. The solvent volume assumed for this calculation was $(1 - V_{missing}) = 0.85$. The combined envelope for the whole molecule (combined *A*) was obtained by the addition (union) of the two envelopes described above (Fig. 3*d*). Since the envelope calculated from the difference map covered some volume of the known part, the total volume of the 'combined *A*' envelope was somewhat smaller than other envelopes. Density-modification calculations with phases from the MR model and the 'combined *A*' envelope produced a map that had a correlation coefficient of 0.77 to the final F_{obs} map and 0.66 in the prosegment area and the relative average electron density of 0.95 and 0.31 for the known and unknown parts, respectively (Table 1).

3.6.2. *Real space.* In real space the mask for the known part was calculated exactly as described above, from the F_{obs} map phased with the MR model and assuming a solvent content of $(1 - V_{known}) = 0.65$. To build the envelope for the unknown part we calculated the $(3F_o - 2F_c)$ map, which was then modified by replacing the electron density at every grid point within the envelope for the known part by a value of zero. We used a similar approach previously in the treatment of a translation problem (Cygler & Desrochers, 1989). This map is conceptually equivalent to a difference map. The envelope for the unknown part was calculated from this map assuming a solvent content of $(1 - V_{missing}) = 0.85$. The combined envelope ('combined *B*') was an addition of the two (Fig. 3*e*) and had 86.9% volume in common

with the reference envelope. The modified map obtained with the 'combined *B*' envelope had a correlation coefficient of 0.77 to the final F_{obs} map and 0.66 in the prosegment area while the relative average electron density for the backbone was 0.94 and 0.32, respectively (Table 1). Inspection of the envelopes calculated as described above showed that there was a gap between the envelope of the partial model and that for the unknown part, resulting in the 'combined *B*' envelope having holes at the interface between the prosegment and the cathepsin L (Fig. 3*e*). This is an artifact resulting from close contacts between a significant portion of the unknown domain and the partial model. After blanking the density inside the partial model envelope, the step of electron-density averaging over a large radius artificially reduces the density values at the grid points near the partial model. The calculation of the envelope for the unknown part then automatically excludes many of these grid points.

To correct this problem we have redesigned the procedure that removes the density for the partial model from the electron-density map. Instead of truncating the map within the whole volume of the partial model we have modified the density only within the 'core' of this model. This 'core' was defined by a 'tight' envelope of the partial model and was calculated from the F_{obs} map using, somewhat arbitrarily, $r = 3/5$ (60%) of the volume occupied by the partial model. That led to the value of $V_{core} = rV_{known} = 0.20$ for the core of the partial model and $(1 - V_{core}) = 0.80$ for the solvent volume. The $(3F_{obs} - 2F_{calc})$ map was then blanked to zero within this 'tight' envelope and used to calculate the envelope for the unknown part with the solvent content of $(1 - V_{missing}) = 0.85$, as in the previously described procedure. This envelope was then added to the original envelope for the partial model to form the 'combined *C*' envelope (Fig. 3*f*). The 'combined *C*' and the reference envelopes have 87.5% volume in common. The *DM* procedure with the 'combined *C*' envelope led to a map with a correlation coeficient of 0.78 to the final F_{obs} map and 0.69 within the prosegment area. The relative average electron density for the backbone atoms was 0.94 and 0.35 for the known and unknown parts, respectively.

Visual inspection of the 'combined *C*' envelope showed that the holes at the interface between partial model and prosegment, that were present in the 'combined *B*' envelope, had been eliminated. The value of the partial model volume fraction which is suitable for the calculation of the 'core' volume depends on the radius of averaging used in subsequent envelope calculation. If a smaller averaging radius is chosen, a larger volume fraction will suffice.

3.7. *Envelope calculations by the SOLOMON program*

A new approach for calculating the molecular envelope was recently proposed and applied successfully to the solution of the F1-ATPase (Abrahams & Leslie, 1996). The program *SOLOMON* that performs the calculations became available as part of the latest version of *CCP*4 suite of crystallographic programs (J. P. Abrahams, *CCP*4 suite, 1996 version). For comparison with other approaches described here, we have calculated the envelope using the *SOLOMON* program and applied the same *DM* procedure as in other cases. The input to the *SOLOMON* program was either the F_{obs} or the $(3F_{obs} - 2F_{calc})$ map calculated with the phases derived from the MR model. The F_{obs} map led to an envelope that had 79.6% volume in common with the reference envelope. Although this envelope was somewhat better than the standard envelope, it suffered similar problems, namely poor coverage of the proregion. As a consequence, the modified map resulting from the *DM* procedure had a correlation of 0.75 to the final map and the correlation within the proregion was only 0.64, only slightly better than in the case of the standard envelope. The results for the $(3F_{obs} - 2F_{calc})$ were significantly better. The volume in common with the reference envelope increased to 87.0%, slightly worse than the 'combined C' envelope. The modified map based on this envelope had correlation coefficients to the reference map of 0.77 for the whole map, 0.86 for the known part and 0.68 for the unknown part. The relative average electron density was 0.93 and 0.32 for the known and unknown parts, respectively.

3.8. *Iterating DM procedure with recalculation of the envelope*

Since there is an improvement in the electron-density map after one cycle of the *DM* procedure, the new map and/or phases could be used to calculate an improved molecular envelope. We have tested the iterative application of the *DM* procedure for three cases: (*a*) standard, (*b*) 'combined *C*' and (*c*) *SOLOMON* procedure envelope calculation. The maps used for deriving the new envelopes were calculated with the coefficients as described above but with the phases obtained at the end of the previous *DM* cycle. This new mask and the phases from the MR model were then input to the next cycle of *DM*.

Iteration of *DM* cycles with the standard envelope led to no further improvement of the electron-density map due to little change in the envelope. On the other hand, a significant improvement was achieved with the 'combined *C*' envelope. The correlation coefficient between the final map and the *DM*-modified map within the volume of the unknown part increased from 0.69 to 0.71 in the second cycle and to 0.72 in the third cycle. The latter value is nearly equal to the correlation for the map that was produced using the reference envelope. The relative average electron density for the backbone atoms in this map was 0.95 for the known part and 0.39 for the unknown part, a significant increase of the latter as compared to the starting F_{obs} map phased by

the MR model. As expected, this iterative procedure improved also the maps when the mask was obtained by the *SOLOMON* procedure (Table 1).

4. Conclusions

All present density-modification programs include the solvent-flattening step as a major part of the algorithm. We have shown here that to take the full advantage of the solvent-flattening procedure in the case of a molecular replacement solution when a significant portion of the model is missing, the molecular envelope should be calculated in a different way than for the MIR case. While in the MIR maps the errors are distributed more or less evenly throughout the whole asymmetric unit, in the MR maps there is a strong bias toward the partial model. The analysis of the reasons of the failure of the standard application of the density-modification method to improve the initial electron-density map obtained from a partial molecular replacement model led us to propose an improved procedure to determine the molecular envelope. The most straightforward and quite effective is to use $(2F_{obs} - F_{calc})$, $(2mF_{obs} - DF_{calc})$ or $(3F_{obs} - 2F_{calc})$ maps for envelope determination in place the F_{obs}. However, in the MR case it is advantageous to decouple the process of envelope determination for the known part (MR model) from the calculation of the envelope for the unknown part. While this can be achieved conceptually either in reciprocal or real space, we believe that manipulation of the map in real space provides more flexibility and leads to somewhat better results. Many different schemes can be thought of to acheive this goal. We have investigated a few possibilities and found that among those the best results were obtained by calculating the envelope for the partial model from the F_{obs} map while that for the unknown part from the $(3F_{obs} - 2F_{calc})$ map modified to zero within the core of the partial model. A significant improvement in the electron-density map derived from such an envelope suggested an iterative approach. Indeed, iteration of the *DM* procedure with recalculation of the envelope after each cycle of density modification provided a further improvement of the map making the interpretation of electron density relatively easy. In the end, by applying this procedure to determine the envelope, we were able to improve significantly the electron-density map as compared with the standard iterative application of density modification: the correlation of the resulting map to the final map in the unknown portion increased from 0.620 to 0.716 and the relative average density for the backbone atoms increased from 0.25 to 0.39. Finally, we have observed that the algorithm used by the *SOLOMON* program to calculate envelopes was superior to the traditional approach and when applied to $(3F_{obs} - 2F_{calc})$ map gave results almost as good as the 'combined C' envelope.

The authors are grateful to Mr Marc Desrochers for writing some of the programs used for the analysis of masks and maps.

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.
Bricogne, G. (1974). *Acta Cryst.* A**30**, 395–405.
Bricogne, G. (1976). *Acta Cryst.* A**32**, 832–847.
Brünger, A. T. (1992). *X-PLOR* Version 3.1, *A system for X-ray crystallography and NMR*, Yale University Press, New Haven, USA.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Coulombe, R., Grochulski, P., Sivaraman, J., Ménard, R., Mort, J. S. & Cygler, M. (1996). *EMBO J.* **15**, 5492–5503.
Coulombe, R., Li, Y., Takebe, S., Ménard, R., Mason, P., Mort, J. S. & Cygler, M. (1996). *Proteins Struct. Funct. Genet.* **25**, 398–400.
Cowtan, K. D. (1994). *Jnt CCP4 ESF-EACBM Newslett. Protein Crystallogr.* **31**, 34–38.
Cowtan, K. D. & Main, P. (1993). *Acta Cryst.* D**49**, 148–157.
Cygler, M. & Desrochers, M. (1989). *Acta Cryst.* A**45**, 563–72.
Furey, W. & Swaminathan, S. (1990). Am. Crystallogr. Assoc. Meet. Abstracts, Series 2, Vol. 18, p. 17.
Hoppe, W. & Gassmann, J. (1968). *Acta Cryst.* B**24**, 97–107.
Kleywegt, G. J. & Jones, T. A. (1996). *Acta Cryst.* D**52**, 826–828.
Leslie, A. (1987). *Acta Cryst.* A**43**, 134–136.
Lunin, V. Yu. (1993). *Acta Cryst.* D**49**, 90–99.
Main, P. & Rossmann, M. G. (1966). *Acta Cryst.* **21**, 67–72.
Navaza, J. (1994). *Acta Cryst*, A**50**, 157–163.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Rossmann, M. G. & Blow, D. M. (1963). *Acta Cryst.* **16**, 39–45.
Rudenko, G., Bonten, E., d'Azzo, A. & Hol, W. G. J. (1996). *Acta Cryst.* D**52**, 923–936.
Sim, G. A. (1960). *Acta Cryst.* **13**, 511–512.
Wang, B.-C. (1985). *Methods Enzymol.* **115**, 90–112.
Zhang, K. Y. J. (1993). *Acta Cryst.* D**49**, 213–222.