Acta Cryst. (1995). D**51**, 888–895

On the Solution of the Molecular-Replacement Problem at Very Low Resolution: Application to Large Complexes

BY ALEXANDR URZHUMTSEV* AND ALBERTO PODJARNY

UPR de Biologie Structurale, IGBMC, BP 163-67404 *Illkirch CEDEX, CU de Strasbourg, France*

(*Received* 28 *December* 1994; *accepted* 19 *May* 1995)

Abstract

The applicability of the molecular-replacement (MR) method, implemented through the *AMoRe* package [Navaza (1994). *Acta Cryst.* A**50**, 157–163], is studied at very low resolution ($d > 20$ Å) and for very large molecular complexes. Due to the nature of the low-resolution data, specific problems appear. In particular, rotation-function peaks are very broad and translation functions based on Patterson overlap show large spurious peaks. To solve these problems, the translation function is replaced by a search using amplitude correlation and a systematic three-dimensional angular search is performed around each rotation-function peak. Furthermore, these functions are applied in different resolution ranges during the same search. The corresponding algorithms are applied to two cases: the $tRNA^{Asp}$–synthetase complex (neutron diffraction data) and a ribosome model crystal (calculated data). This new implementation is shown to solve the problem for a variety of search models, ranging from a detailed atomic model to a rough envelope.

1. Introduction

The molecular-replacement (MR) method (Rossmann & Blow, 1962; Crowther & Blow, 1967) has become a widespread and powerful technique for solving the phase problem when a structure closely related to the one under study (or a fragment of it) is available. This method is generally applied to diffraction data of 6 Å or higher resolution. However, there are cases where the method can only be applied at much lower resolution, *i.e.* (1) if only low-resolution data are available; (2) if no high-resolution homologous structure is available, but a low-resolution image can be obtained; this is the case when an electron-microscopy (EM) image is known, or when the known homologous structure is not similar enough and can provide only the molecular envelope.

The correct placement of such an image in the crystallographic unit cell is useful in several ways.

(1) For the case of a new structure, it provides a starting phase set for phase-extension procedures, as well as information useful for their application (*e.g.* the molecular envelope, placement of non-crystallographic symmetry-related units).

(2) For the case when the model of an homologous structure is known, but it is significantly different and high-resolution MR fails, the placement of a low-resolution envelope might provide a solution which can be extended to higher resolution.

Low-resolution data are strongly influenced by solvent. To check the reproducibility of low-resolution amplitudes with model data, preliminary tests have been conducted on structures solved at high resolution (Urzhumtsev & Podjarny, 1995). They showed that at resolutions lower than 20 Å there is good agreement between the observed diffraction amplitudes and the ones calculated from a model. The solvent contribution is large, but since it is essentially collinear with the contribution from the model, it amounts to a scale factor. At higher resolution (10–15 Å) the solvent region cannot be considered flat, and, therefore, the diffraction amplitudes have a significant non-collinear solvent contribution and cannot be easily reproduced with a molecular model only.

The cubic form of the $tRNA^{Asp}$–synthetase complex provides an ideal test case for low-resolution molecular replacement, since:

(1) a high-resolution structure of the same complex is available (Ruff *et al.*, 1991), and, therefore, a search can be carried out with either model or envelope at different resolutions;

(2) because of the high solvent content (Lorber *et al.*, 1983) the envelope is well defined inside the unit cell;

(3) neutron data are available with excellent very low resolution completeness (Moras *et al.*, 1983).

The work described below shows that it is possible to solve the MR problem at very low resolution. It also describes problems specific to low-resolution work, a way to solve them and its application to the cases of the $tRNA^{Asp}$–synthetase complex and of a ribosome model crystal.

2. Molecular-replacement strategy

In order to test and develop the methods for low-resolution MR, the program package *AMoRe* (Navaza, 1994) was used. While consisting of the same steps

* Permanent position: Institute of Mathematical Problems of Biology, Russian Academy of Sciences, Pushchino, Moscow Region, 142 292, Russia.

of rotation function, translation function and rigid-body refinement, *AMoRe* has proved more efficient than other MR packages [*e.g. MERLOT* by Fitzgerald (1988) or *X-PLOR*, Brünger (1990)] in solving the MR problem. Following the usual needs, its parameters are tuned for high- and medium-resolution ranges. In particular, the translation function is based on the Patterson overlap (OverF^2, see equation in Table 1) for fast calculation. Other measures of model fitness, like amplitude correlation and R factor, are calculated only for the maxima of OverF^2. Since the nature of this signal varies when lowering the resolution, the nature of the function used to detect it should also vary. The purpose of this paper is to develop some molecular-replacement methods that will work at very low resolution. For the purpose of clarity, the original package is called '*AMoRe*' and the resulting programs are called 'modified *AMoRe*' in the following text.

3. Low-resolution MR. Tests with an atomic model

3.1. *Initial check of AMoRe*

The neutron data from the cubic form of the $tRNA^{Asp}$–synthetase complex (Moras *et al.*, 1983) were used for these tests. This structure had been solved by *AMoRe* using a high-resolution model from a different crystal form (Ruff *et al.*, 1991) and X-ray diffraction data at 8 Å resolution (Ruff *et al.*, 1988). The solution showed very clearly one single dimer in the asymmetric unit, placed in a general position (Urzhumtsev, Podjarny & Navaza, 1994).

During the initial stages of the structure solution (Podjarny *et al.*, 1987), neutron data had been collected in the resolution range ∞–20 Å (Moras *et al.*, 1983). These data are very complete from ∞ to 26 Å; moreover, the number of reflections (31) at 50 Å resolution is large enough to have a significant observations/parameters ratio in a six-dimensional search.

An *AMoRe* search with the complex dimer using the centered Patterson overlap translation function was conducted against the neutron diffraction data from 20 Å to ∞ (although the high-resolution end is not complete). No peak in this search was close to the correct position and orientation. The highest one [Corr(F) = 81%, R factor = 47%] corresponded to a spurious position where the dimer is placed on the origin of the cubic space group. The reason for this failure is neither the quality of the neutron data nor the appropriateness of the search model, since the correct solution gives values of Corr(F) = 92% and of R factor = 28%. The problem is that the use of a translation function based on the centered Patterson overlap missed the global maxima of the correlation function between F_{obs} and $F_{calc}(h;\mathbf{R},\mathbf{t})$.

It should be noted that *AMoRe* did give the correct solution when using the full-symmetry phased-translation function (Colman, Fehlhammer & Bartels, 1974; Bentley & Houdusse, 1992), but the peak for the correct position (which has the largest full correlation factor) appeared beyond the 50th position of this phased-translation function. Therefore, this procedure could also miss the right solution if only the top peaks of the translation function were investigated, as is usually the case.

3.2. *Tests of standard AMoRe with calculated diffraction data*

To identify the reasons of the problems described in §3.1 independently of experimental errors, the neutron diffraction data were replaced by values calculated from the atomic model without solvent modelization.

Several resolution ranges were tested. Table 1(*a*) (lines 1–5) shows the result of these tests, the high-resolution end varying between 15 and 30 Å and the low-resolution end varying between 30 Å and ∞. *AMoRe* works for all cases with two exceptions.

(*a*) When the higher resolution limit is 30 Å or lower (Table 1*a*, lines 3, 5, see note 2).

(*b*) When the 'inner core' of reflections ($d > 50$ Å) is included (Table 1*a*, lines 4, 5, see note 3).

The failure in (*a*) is linked to large errors in the rotation function. Fig. 1 shows the (β,γ) sections of the rotation function for different conditions (including data to 10 Å for comparison). For the resolution range 10–15 Å (Fig. 1*a*) a clear signal is seen for the right orientation. For the resolution range 30–40 Å, the peak is displaced (Fig. 1*b*). Including lower terms of the spherical harmonics series (Fig. 1*c*) does not improve the situation. These large errors (about 20°) in the orientation lead to the failure of the whole procedure. Note that the correct orientation is in the neighbourhood of a large peak, and could possibly be identified by exploring the peak neighbourhood.

For case (*b*), inclusion of the inner core of reflections broadens the rotation-function peaks (Fig. 1*d*) and can change their order (Table 1*a*, lines 4–5) but does not displace them. The overall failure concerns here the translation function, which fails even for the low error in the model orientation (5°).

The solution of these problems could not be obtained by simple adjustments of the parameters (*i.e.* resolution ranges) in the existing algorithm. In particular, problem (*b*) precluded the use of the 'central reciprocal-space zone' terms to diminish the sensitivity of the translation searches to the errors in the rotation parameters.

3.3. *Development of modified AMoRe*

To solve the problem highlighted in (*b*), the Patterson overlap calculation was replaced by searches (Urzhumtsev & Podjarny, 1994) which use the correlation coefficient (of either F or F^2, see equation in Table 1) directly as the target value [for a previous discussion, see Harada, Lifchitz, Berthou & Jollès (1981)

Table 1. *Application of AMoRe to model data and to experimental data from AspRS*

N_c, N_w, sequential numbers for the correct and first incorrect peak of the rotation function. H_c, Δ_c; H_w, Δ_w, corresponding height and orientation errors. OverF^2, CorrF^2, Corr$F$, different translation functions (see formulae): Patterson overlap, intensities correlation, amplitude correlation. Corr$F$,ex, translation search with amplitude correlation starting at expanded rotation peaks. Cf_c, Cf_w, highest amplitude correlations found with in the translation search with the correct and first incorrect rotation peaks, respectively. Rf_c, Rf_w, R-factor values after rigid-body refinement of the positions found in the translation search with the correct and first incorrect rotation peaks, respectively.

Patterson overlap OverF^2,

$$\sum_h [F^2_{obs}(h) - \langle F^2_{obs}\rangle][F^2_{calc}(h) - \langle F^2_{calc}\rangle].$$

Intensity correlation CorrF^2,

$$\left\{\sum_h [F^2_{obs}(h) - \langle F^2_{obs}\rangle][F^2_{calc}(h) - \langle F^2_{calc}\rangle]\right\} \Big/ \left\{\sum_h [F^2_{obs}(h) - \langle F^2_{obs}\rangle]^2\right\}^{1/2} \left\{\sum_h [F^2_{calc}(h) - \langle F^2_{calc}\rangle]^2\right\}^{1/2}.$$

Amplitude correlation Corr$F$,

$$\left\{\sum_h [F_{obs}(h) - \langle F_{obs}\rangle][F_{calc}(h) - \langle F_{calc}\rangle]\right\} \Big/ \left\{\sum_h [F_{obs}(h) - \langle F_{obs}\rangle]^2\right\}^{1/2} \left\{\sum_h [F_{calc}(h) - \langle F_{calc}\rangle]^2\right\}^{1/2}.$$

$F_{obs}(h)$ and $F_{calc}(h) = F_{calc}(h; \mathbf{R}, \mathbf{t})$ are observed and calculated amplitudes; $\mathbf{R}$, $\mathbf{t}$ represent the three rotation and the three translation parameters of the rigid-body model, respectively.

(*a*) Model data

		Rotation function						Translation function			Rigid-body refinement			
Line	Resolution (Å)	N_c	H_c	Δ_c (°)	N_w	H_w	Δ_w (°)	Type	Resolution (Å)	Cf_c (%)	Cf_w (%)	Resolution (Å)	Rf_c (%)	Rf_w (%)	Notes
						Standard *AMoRe*									
1	15–30	1	19	6	2	15	25	OverF^2	15–30	68	46	15–30	12*	47	(1)
2	20–30	1	21	8	2	19	60	OverF^2	20–20	48	37	20–30	10*	43	(1)
3	30–40	3	21	18	1	32	38	OverF^2	30–40	48	49	30–40	37	34	(2)
4	15–∞	2	38	5	1	42	25	OverF^2	15–∞	77	71	15–∞	77	78	(3)
5	30–∞	1	54	27	2	39	44	OverF^2	30–∞	71	70	30–∞	43	53	(2),(3)
						Searches with different translation functions									
6	30–40	3	21	18	1	32	38	Corr$F$	30–40	61	59	30–40	37	34	(4)
7	30–40	3	21	18	1	32	38	Corr$F$	30–∞	93	88	30–∞	10*	37	(5),(6)
8	30–40	3	21	18	1	32	38	CorrF^2	30–∞	93	96	30–∞	8*	45	(6),(7)
9	30–40	3	21	18	1	32	38	Corr$F$	30–∞	93	88	30–40	14*	34	(7)
10	30–40	3	21	18	1	32	38	Corr$F$	50–∞	92	87	30–40	14*	34	(8)
						Correlation searches using 'inner-core' reflections									
11	50–∞	2	45	20	1	72	34	Corr$F$	50–∞	96	87	50–∞	2*	40	—
12	30–∞	1	54	27	2	39	44	Corr$F$	50–∞	90	87	30–∞	68	42	—
13	15–∞	1	42	5	2	38	25	Corr$F$	50–∞	97	88	15–∞	10*	50	—
14	50–∞	1	68	27	2	45	46	Corr$F$	50–∞	84	82	50–∞	40	36	(9)
15	50–∞	1	68	27	2	45	46	Corr$F$,ex	50–∞	88	87	50–∞	1*	29	(9)

(*b*) Experimental data

		Rotation function						Translation	Rigid-body refinement		
Line	Resolution (Å)	N_c	H_c	Δ_c (°)	N_w	H_w	Δ_w (°)	Type	Resolution (Å)	Rf_c (%)	Rf_w (%)	Notes
						Searches with atomic models						
1	20–∞	1	48	27	2	33	44	Corr$F$	20–∞	18*	49	—
2	50–∞	2	45	22	1	72	35	Corr$F$	50–∞	13*	30	—
3	50–∞	1	68	28	2	45	48	Corr$F$	50–∞	40	40	(9)
4	50–∞	1	68	28	2	45	48	Corr$F$,ex	50–∞	13*	40	(9)
						Searches with envelope models						
5	50–∞	1	72	17	2	50	48	Corr$F$,ex	50–∞	22*	40	(10)
6	50–∞	1	72	18	2	56	45	Corr$F$,ex	50–∞	33*	40	(11)
7	20–∞	1	40	25	2	26	44	Corr$F$,ex	20–∞	20*	45	—

Notes: (1) the rotation function works (errors of 8° or less, first peak in the rotation function) if the higher resolution limit is 15–20 Å and the inner core of reflections is excluded. (2) The translation function fails when very low resolution terms are included; this is true even for cases where the rotation function has worked properly. (3) When the higher resolution limit is 30 Å or lower, the errors raise to 18° or more. (4) The searches using correlation functions without the inner core of reflections repeat the one using the Patterson overlap, where the right peak was not found if the rotation error was large. (5) The search using correlation functions with the lowest resolution 'inner core' of reflections finds the right solution. (6) Both the correlation in F and the correlation in F^2 work; the correlation in F gives a somewhat better contrast. (7) The refinement step corrects the large rotation errors: exclusion of the lowest resolution reflections in this step can increase the contrast in correlation value for the right solution against the next peak (99:84% at the 30 Å–∞ and 96:70% at 30–40 Å resolution) and may be useful in some difficult cases. (8) The signal/noise ratio in the search with the 'inner core' of reflections is low; it comes mostly from this 'inner core' and does not significantly improve with the inclusion of higher resolution data. (9) for the case of changed Patterson radius the rotation error is higher (27–28°) and causes the failure of the translation function; the search using the expanded rotation peaks solves the problem. (10–11) For the flat and modulated envelope searches at 50 Å resolution the corresponding final errors are (9°, 3 Å) and (12°, 2 Å), respectively; for the second case ten more refinement cycles decreased the error to (6°, 2 Å) with practically the same values of correlation and R factor; the search with modulated envelope (11) gives less contrast in R factor than one with the flat envelope (10) but it less sensitive to the choice of the envelope. The corresponding contrast in correlation coefficient against the next peak is (84:70%) for the modulated envelope and (92:80%) for the flat one.

*The correct solution (with positional error of about 1 Å and rotation error of about 1°, if not stated otherwise).

and Fujinaga & Read (1987)]. The correlation coefficient is deduced from the Patterson overlap essentially by dividing by,

$$\sigma(F_{\rm calc}) = \{\sum_h [F_{\rm calc}(h) - \langle F_{\rm calc}\rangle]^2\}^{1/2}$$
$$= \{\sum_h [F_{\rm calc}(h)]^2 - N\langle F_{\rm calc}\rangle^2\}^{1/2}. \quad (1)$$

As shown by Harada *et al.* (1981) the term,

$$\sum(\mathbf{R},\mathbf{t}) = \sum_h [F_{\rm calc}(h;\mathbf{R},\mathbf{t})]^2$$
$$= K\int [\rho_{\rm calc}(r;\mathbf{R},\mathbf{t})]^2 {\rm d}^3 r, \quad (2)$$

can increase considerably as a function of **t** when molecules overlap with their symmetry-related images, due to the squaring of the density. These variations are reflected in the translation function; the lack of normalization of the Patterson overlap leads to very strong spurious peaks which correspond to those of $\sigma(F_{\rm calc})$ and not to the true correlation signal (see Fig. 2). Note also that Fig. 2(*b*) shows that the highest spurious peak ($x = 0.25$, $y = 0$, $z = 0$) is larger for Corr(F^2) than for Corr(F). The signal contrast is thus better for Corr(F).

This analysis led to the following strategy.

(1) Replace the Patterson overlap by the correlation of amplitudes in the translation-function searches.

(2) Use lowest resolution data for translation searches.

Table 1(*a*) (lines 6–10) shows that these searches with different translation functions work, starting with the model orientations obtained from the rotation search using 30–40 Å data. Note that the resolution ranges for rotation and translation searches can be different in the course of the same run, in order to obtain the best signal in each case. All cases confirm the previous observation that the translation searches using the lowest resolution ranges and based on correlation functions show the signal clearly; the refinement step can then correct large rotation errors, if necessary. The correlation in F worked slightly better than the correlation in F^2. Table 1(*a*) (lines 11–15) extends these observations to rotation searches using data including the 'inner core' reflections, with several high-resolution limits.

Further tests showed that the rotation-error limit for the success of ulterior translation searches usually lies around 20°. However, the rotation error can be as large as 30°, particularly when the high-resolution end is only 40 or 50 Å. To assure that the error in the rotation parameters is small enough to fall within the 20° limit, a scanning around the rotation peaks in the three angular directions with a step of ±20° was introduced. This resulting set of orientations is called the 'expanded rotation peak'. Table 1(*a*) (lines 14–15) shows the application of these searches to a model case at 50 Å resolution, intentionally changing the Patterson radius from the optimal value (from 60 to 80 Å). This increased the rotation error from 20 to 27°, but even this larger value can be corrected for by using the 'expanded rotation peak' as input to the translation function, and by refining with the 'inner core' reflections.

Modified *AMoRe* is thus obtained from *AMoRe* by using the 'expanded rotation peak' and translation searches based in the amplitude correlation.

3.4. *Tests with experimental neutron diffraction data*

In order to test modified *AMoRe* using real diffraction data, the same procedure was applied for the experi-

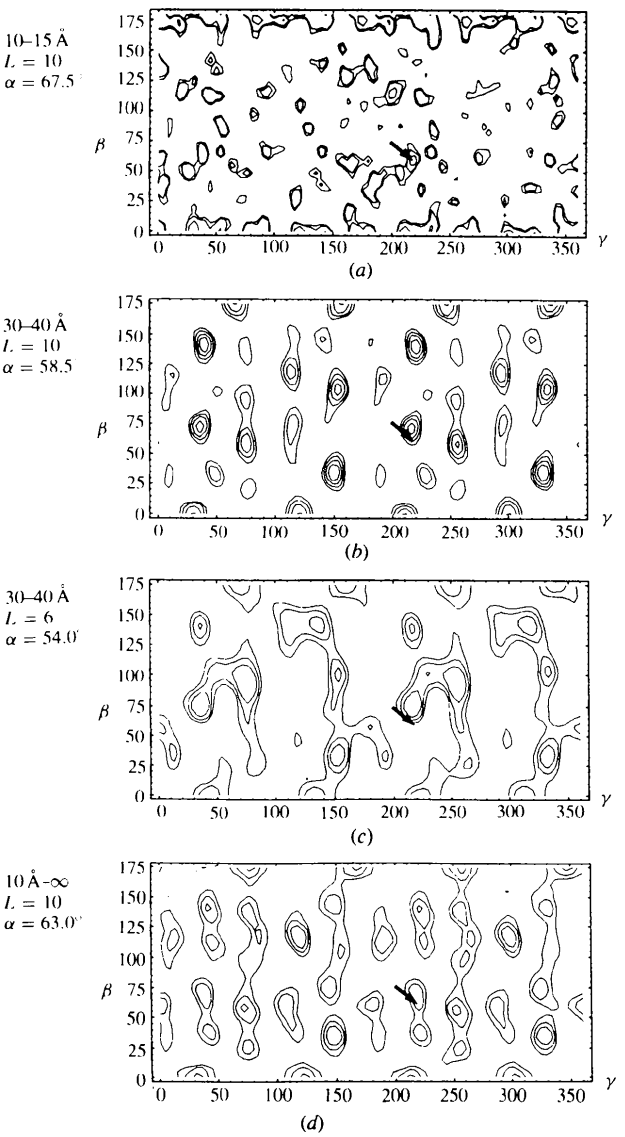


Fig. 1. Rotation-function analysis for the calculated data for the $tRNA^{Asp}$–synthetase complex. The rotation function is calculated for several resolution ranges and minimal L values for the included spherical harmonics. α-Sections of the functions are shown crossing the centre of the rotation-function peak closest to the correct orientation (corresponding β and γ values are indicated by the arrows). Note the quasi-periodicity in γ due to the dimeric nature of the model.

mental neutron data at very low resolution. Table 1(*b*) (lines 1–2) shows that it converges to the correct solution in all cases. Note that, as before, the limitation of data at the high-resolution end causes a large error in the rotation function, but the translation search gives the correct solution and the rigid-body refinement corrects the error.

Table 1(*b*) (line 3) shows the effect of changing the Patterson radius from 60 to 80 Å. In this case, the rotation function finds a position with larger error (28°), enough to cause failure of the translation function. Therefore, the expansion of rotation peaks (Table 1*b*, line 4) is essential to find the right solution.

3.5. *Summary of searches with an atomic model*

The low-resolution searches described above show that different resolution ranges should be used for different types of searches. For rotation searches, the highest available resolution is preferable, and very low resolution terms can be excluded. Translation searches should be carried out using correlation searches against very low resolution data. The neighbourhood of rotation peaks should be explored. Rigid-body refinements should be carried out against all available data. Using lowest resolution data may be important for correcting large errors, even when it may also lower the signal/noise contrast.

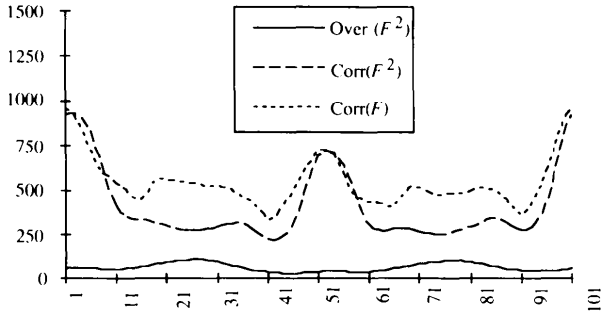
4. Tests with envelopes.

The object available for conducting low-resolution searches might be an envelope (and not a detailed atomic model), for example coming from electron microscopy or from very approximate models. Therefore, it is important to check whether modified *AMoRe* can be used to correctly place a molecular envelope inside the unit cell.

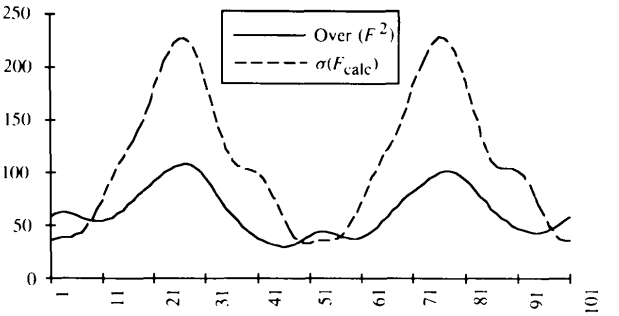
As before, the studies with envelopes were conducted using the neutron data from the cubic form of the $tRNA^{Asp}$–synthetase complex.

4.1. *Exact flat envelope, 50 Å neutron data*

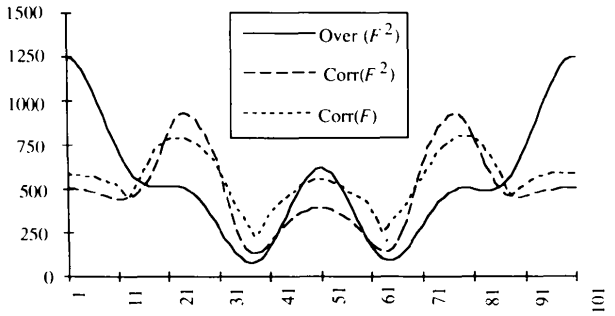
To obtain an exact envelope, each atom of the model was surrounded by a sphere of 2.5 Å radius on a three-dimensional grid (2 Å step). For the purpose of using modified *AMoRe*, the resulting grid points were placed



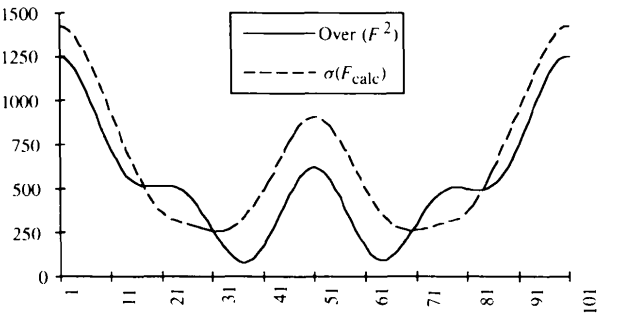
$Y = y_0$, $Z = z_0$, $0 \le X \le 1$. $X = 0.0$ is close to the correct position x_0.
(*a*)



$Y = y_0$, $Z = z_0$, $0 \le X \le 1$. $X = 0.0$ is close to the correct position x_0.
(*c*)



$Y = 0$, $Z = 0$, $0 \le X \le 1$.
(*b*)



$Y = 0$, $Z = 0$, $0 \le X \le 1$.
(*d*)

Fig. 2. Comparison of the translation functions for one-dimensional searches. The correctly oriented model is placed in the right position (x_0 = 0.0003, y_0 = 0.3588, z_0 = 0.1636) (*a*) and at the origin (x = 0.0, y = 0.0, z = 0.0) (*b*). In both cases it is then displaced along the x axis, and the values of Over(F^2), Corr(F^2) and Corr(F) are calculated. the separate analysis (*c* and *d*) of the contribution of the numerator, Over(F^2), and the denominator, $\sigma(F_{calc})$, to Corr(F^2) [see equation (1) and Table 1] shows that the peaks of $\sigma(F_{calc})$ correspond to the molecule moving over special symmetry positions. Note that these peaks are reflected in Over(F^2), creating a strong spurious signal (see text).

in a box six times larger than the model in each linear direction, where it occupied 0.0012 of the volume, and assigned a density value of 1; the rest of the points are set to 0. This density was then used to calculate structure factors. The R factor at 50 Å between amplitudes calculated from the atomic model and from this flat envelope was 4%. This shows that at this very low resolution range, the transform of an atomic model is virtually the same as that of the corresponding exact envelope. The results were, as expected, virtually the same as those obtained using an atomic model.

4.2. *Density-based envelope, 50 Å neutron data*

To obtain an envelope closer to that of a real case, the model structure factors were used to calculate a density distribution at 50 Å, which was placed in the same box as before. The envelope was defined by all the points above a cut-off level. In this synthesis, the envelope can be expanded to 0.0035 of the volume by lowering the cut-off level before it includes isolated noise peaks.

Two possibilities were considered: a flat envelope or a modulated one, where the density values inside the molecular envelope are kept. In both cases, two different cut-off levels were considered, giving the exact molecular volume and three times the molecular volume, respectively. For the flat envelope the exact molecular volume fits the data well ($R = 15\%$) while the expanded volume is not so good ($R = 39\%$). The situation is reversed for the modulated envelope, where the larger molecular volume is better ($R = 12\%$ *versus* 31% for the smaller volume).

Searches were performed using both the flat and the modulated envelopes, at their optimal volumes. Table 1(*b*), lines 5–6, shows that the solution was clearly found in both cases. Again, the orientation errors are large but they are within the limits of the rotation-peak expansion.

4.3. *Modulated envelope; 20 Å neutron data*

These tests were repeated at higher resolution. The model structure factors were used to calculate a density distribution at 20 Å. In this synthesis, the envelope can be expanded to 0.0020 of the volume by lowering the cut-off level before it includes noise peaks. Like in the 50 Å case, for the flat envelope the exact molecular volume fits the data better, while for the modulated envelope the larger molecular volume is better. Searches were conducted using the modulated envelope only. Table 1(*b*) (line 7) shows that the correct solution is found. It should be noted that the signal contrast here is higher than in the 50 Å resolution searches.

4.4. *Protocol for low-resolution searches with envelopes*

From this work the following rules can be derived.

(1) The same procedure of MR which was used for atomic models at very low resolution can be applied also for envelopes.

(2) Both modulated and flat envelopes work; however, the result of searching modulated envelopes is less dependent on the cut-off level.

(3) To get a modulated envelope, a synthesis should be calculated from the available amplitudes and phases, then cut at the 'noise level' (maybe substantially larger than the estimated model volume).

5. Tests on ribosome model crystals

A case where finding the position and orientation of an envelope is very important is that of the ribosome particle. Previous to applications with measured experimental data, it was necessary to test the ideas described above. These tests had two objectives: to confirm the results of the experiences with the $tRNA^{Asp}$–synthetase complex in a different case, and to find the best parameters to be applied to the ribosomal experimental data. To do so, a model density distribution was constructed for the tetragonal crystal form of the 50S particle (T50S, space group $P4_12_12$, $a = b = 498$, $c = 198$ Å) for *Thermus thermophilus* (Yonath, 1992). It was based on a low-resolution EM image of *Bacillus stearothermophilus* (B50S), provided by Yonath and coworkers (Berkovitch-Yellin, Wittmann & Yonath, 1990). This image was obtained using diffraction information from tilt-series of two-dimensional crystalline arrays. The resulting image was packed in the tetragonal crystal lattice observed for the T50S particles. This artificial packing simulated proper crystalline contacts. Structure factors were calculated from this model crystal at 20 Å resolution.

5.1. *Low-resolution MR runs on simulated ribosome data*

Modified *AMoRe* found the right solution with very high contrast at 30 Å resolution. The height (in σ) of the translation-function peaks is 19 and 5 for the correct position and first spurious peak, respectively; the corresponding positional errors are 1 and 51 Å, and the R factors after rigid-body refinement are 7.9 and 53.5%, respectively. This result is not surprising, since the data has no errors and the model is perfect.

5.2. *Systematic study of signal/noise ratio and tolerance to errors*

In order to optimize the parameters for the different steps of modified *AMoRe*, a systematic analysis was carried out for each of them. The most crucial variable is the resolution range (D_{high}–D_{low} cut-offs); therefore, the searches were repeated systematically for a wide variety of them.

5.2.1. *Rotation functions.* Rotation functions were calculated varying both ends of the resolution range. The highest resolution (D_{high}) was varied between 20 and 100 Å, and the lowest resolution (D_{low}) was varied between 30 Å and infinity. In each case the ratio of the

height of the 'correct' peak to that of the first incorrect peak (contrast) was calculated (Fig. 3*a*) and the error of the 'correct' peak (D_φ) was noted (Fig. 3*b*).

This analysis confirms the previous observations that the best contrast is obtained using the highest possible resolution and without the inner core of reflections, and that the error of the rotation peaks increases when the inner core of reflections is included.

5.2.2. *Translation functions.* Translation functions were also calculated for a wide range of resolutions. In this case, an additional dimension was added: the error in the orientation of the input model, which was varied from 0 to 40°. In particular, Figs. 3(*c*) and 3(*d*) show the behaviour of the signal/noise ratio for the cases of 0° rotation error and 30° rotation error. If the error is less than 20°, the best resolution range is the highest one, and the introduction of the inner core of reflections lowers the contrast (Fig. 3*c*). On the other side, if the error is larger than 20°, as it is expected in a real case at this resolution, the situation is reversed. The signal/noise ratio decreases drastically for the higher resolution data ranges (to the limit of complete lack of signal) while it stays constant for the lowest resolution data ranges (Fig. 3*d*).

5.2.3. *Refinement.* Since the role of this step is mainly to correct rotation and translation errors, the maximum allowable translation error (DX_{max}) was analysed as a function of resolution range and of orientation error from 0 to 30°). In general, the radius of convergence was very high (up to 30° and 30 Å). However, for high rotation error the inclusion of the reflections in the 20–30 Å resolution range causes overall failure.

6. Concluding remarks

This study highlights the importance of the use of correlation functions as the search criterion. It also shows that, in the very low resolution case, the optimum resolution ranges are not the same for the rotation function, the translation function and the refinement steps. The rotation function works best when using higher resolution terms; however even in this case it can be very inaccurate for experimental data. The correlation searches and subsequent refinement based on the lowest resolution data only are capable of correcting this error. In some cases, the rotation error is too large and the neighbourhood of the rotation peaks needs to be scanned to fall within the radius of convergence

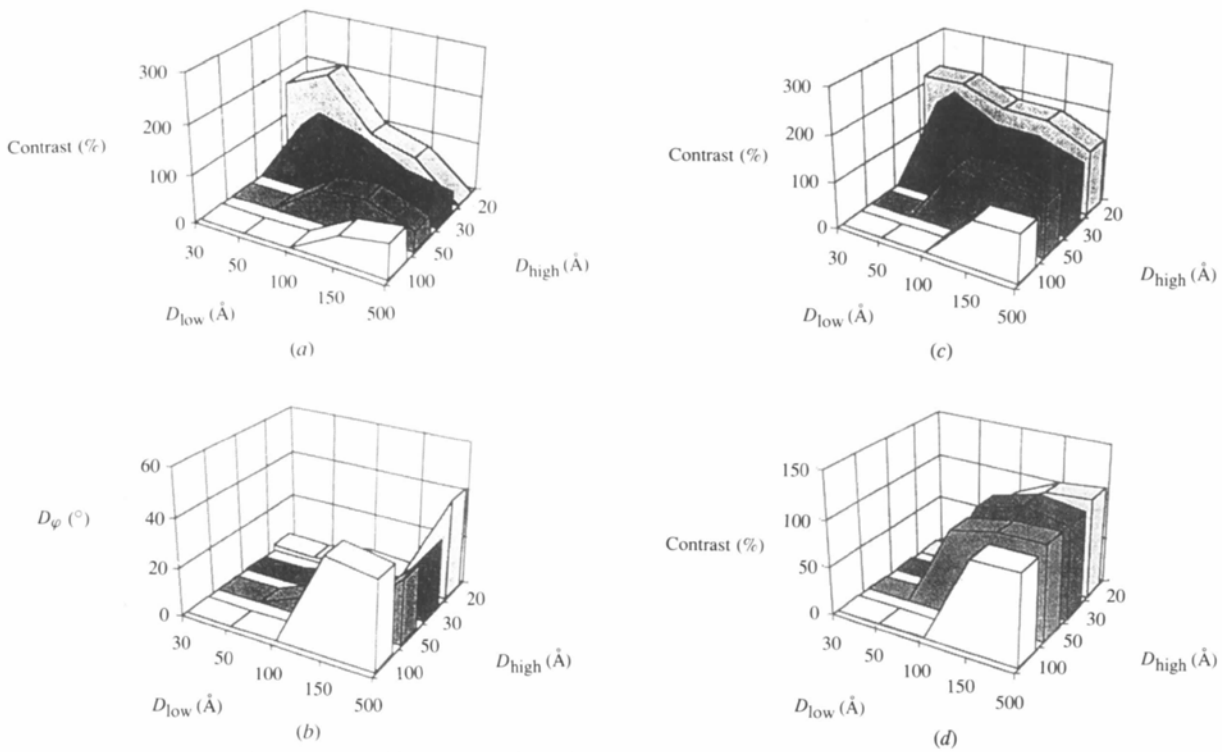


Fig. 3. Analysis of signal and error for the different molecular-replacement steps for the model ribosome crystals as a function of the resolution range ($D_{high} \leq d \leq D_{low}$) of the data. (*a*) The contrast (height of the correct peak/height of first incorrect data) of the rotation function as a function of the resolution range of the data. A contrast value >100% means that the first peak of the rotation is correct; a value <100% means that the correct peak appears but not as the first one. (*b*) The error (D_φ) measured as the angular distance between the rotation-function peak and the correct rotation values as a function of the resolution range of the data. (*c*) Contrast of the translation function for 0° rotation error. (*d*) Contrast of the translation function for 30° rotation error.

of the translation and refinement procedures. The low-resolution case therefore needs specific changes. With these changes, MR packages can be used successfully to solve the molecular-replacement problem at very low resolution. In particular, the F^2-correlation searches can be calculated with Fourier transforms (Navaza & Vernoslova, 1995) and have been already implemented in the last version of the *AMoRe* package.

The experience gained in the atomic model searches has been applied to searches with molecular envelopes. These searches are very important since they open a whole new field, that of using experimentally determined envelopes as the search model in the MR procedure. This problem has been solved for the case of the cubic form of the $tRNA^{Asp}$–synthetase complex. The behaviour of the envelope searches is similar to that of searches using atomic models. Thus, this work has succeeded in finding a molecular-replacement method that can use envelopes as the search object at very low resolution.

This experience is currently being applied to the case of the ribosome experimental X-ray diffraction data. The first part of this work, concerning tests with a ribosome model crystal, shows clearly that the rules developed from the experience with the $tRNA^{Asp}$–synthetase complex are more generally applicable. Following these tests, the application of modified *AMoRe* to experimental ribosome data has been started.

We thank Dr J. Navaza for his extensive collaboration with the use of the *AMoRe* package, and useful discussions and criticism for all the aspects of the present work. The experimental data of the $tRNA^{Asp}$–synthetase complex used in this work were collected by Drs J. Cavarelli, J.-C. Thierry and D. Moras (X-ray data) and by Drs M. Roth, A. Lewitt-Bentley and D. Moras (neutron data). We thank them for making these data available to us. The EM image used for generating the ribosome model crystal was provided by Dr A. Yonath and coworkers; we thank them for making these data available to us, and for continuous discussions and encouragement. We also thank Dr B. Rees for useful discussions and making available all his previous work on the different aspects of the crystallography of the cubic form of the $tRNA^{Asp}$–synthetase complex, and Dr P. Dumas and Dr M. Roth for their continuing interest and encouragement on the current work. This work is supported by the CNRS through the UPR 9004, by the EMBL through a fellowship to AGU, by the Institut National de la Santé et de la Recherche Médicale and the Centre Hospitalier Universitaire Régional.

References

BENTLEY, G A. & HOUDUSSE, A. (1992). *Acta Cryst.* A**48**, 312–322.

BERKOVITCH-YELLIN, Z., WITTMANN, H. G. & YONATH, A. (1990). *Acta Cryst.* B**46**, 637–643.

BRÜNGER, A. T. (1990). *X-PLOR Manual, Version* 2.1, Yale Univ., CT, USA.

COLMAN, P. M., FEHLHAMMER, H. & BARTELS, H. (1974). *Crystallographic Computing Techniques*, edited by F. R. AHMED, K. HUML & B. SEDLACEK, pp. 248–258. Copenhagen: Munksgaard.

CROWTHER, R. A. & BLOW, D. M. (1967). *Acta Cryst.* **23**, 544–548.

FITZGERALD, P. M. D. (1988). *J. Appl. Cryst.* **21**, 273–278.

FUJINAGA, M. & READ, R. J. (1987). *J. Appl. Cryst.* **20**, 517–521.

HARADA, Y., LIFCHITZ, A., BERTHOU, J. & JOLLÉS, P. (1981). *Acta Cryst.* A**37**, 398–406.

LORBER, B., GIEGÉ, R., EBEL, J.-P., BERTHET, C., THIERRY, J.-C. & MORAS, D. (1983). *J. Biol. Chem.* **258**, 8429–8435.

MORAS, D., LORBER, B., ROMBY, P., EBEL, J.-P., GIEGÉ, R., LEWITT-BENTLEY, A. & ROTH, M. (1983). *J. Biomol. Struct. Dynam.* **1**, 209–223.

NAVAZA, J. (1994). *Acta Cryst.* A**50**, 157–163.

NAVAZA, J. & VERNOSLOVA, E. (1995). *Acta Cryst.* A**51**, 445–449.

PODJARNY, A. D., REES, B., THIERRY, J.-C., CAVARELLI, J., JESIOR, J. C., ROTH, M., LEWITT-BENTLEY, A., KAHN, R., LORBER, B., EBEL, J.-P., GIEGÉ, R. & MORAS, D. (1987). *J. Biomol. Struct. Dynam.* **5**, 187–198.

ROSSMANN, M. G. & BLOW, D. M. (1962). *Acta Cryst.* **15**, 24–31.

RUFF, M., CAVARELLI, J., MIKOL, V., LORBER, B., MITSCHLER, A., GIEGÉ, R., THIERRY, J.-C. & MORAS, D. (1988). *J. Mol. Biol.* **201**, 235–236.

RUFF, M., KRISHNASWAMY, S., BOEGLIN, M., POTERSZMAN, A., MITSCHLER, A., PODJARNY, A., REES, B., THIERRY, J.-C. & MORAS, D. (1991). *Science*, **252**, 1682–1689.

URZHUMTSEV, A. G. & PODJARNY, A. D. (1994). *J. Appl. Cryst.* **27**, 122–124.

URZHUMTSEV, A. G. & PODJARNY, A. D. (1995). *Jt CCP4 ESF-EACBM Newslett. Protein Crystallogr.* **32**, 12–16.

URZHUMTSEV, A. G., PODJARNY, A. D. & NAVAZA, J. (1994). *Jt CCP4 ESF-EACBM Newslett. Protein Crystallogr.* **30**, 29–36.

YONATH, A. (1992). *Ann. Rev. Biophys. Biomol. Struct.* **21**, 77–93.