Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Meaningful refinement of polyalanine models using rigid-body simulated annealing: application to the structure determination of the A31P Rop mutant

Nicholas M. Glykos[a] and Michael Kokkinidis[a,b]*

[a]Foundation for Research and Technology-Hellas, Institute of Molecular Biology and Biotechnology, PO BOX 1527, 71110 Heraklion, Crete, Greece, and [b]Department of Biology, University of Crete, PO Box 2208, 71409 Heraklion, Crete, Greece

Correspondence e-mail: kokkinid@crystal1.imbb.forth.gr

Received 19 February 1999
Accepted 6 April 1999

PDB Reference: A31P Rop mutant, 1b6q.

Conventional refinement methods, when applied to even correctly positioned polyalanine models of a target structure, result in a systematic distortion of the molecular geometry and to a concomitant increase in the mean phase difference from the correct phase set. Here, it is shown that iterative rigid-body simulated-annealing refinement of polyalanine models employing successively fewer residues per rigid body (down to one alanine residue per body) at a very high initial temperature (of the order of $T_0 = 10000$ K) and with the geometric energy terms switched on, not only preserves the geometry of the model but can also converge to an essentially correct polyalanine trace of the target structure, even when the starting model deviates systematically and significantly from the sought structure. As an example of the application of the method, details are presented of the structure determination of the Ala31Pro mutant of the Rop protein, where an initial roughly positioned polyalanine model (giving an average phase difference of 78.2° from the final phase set) was successfully refined against a 1.8 Å resolution native data set, leading to an essentially correct model of the main chain with an average displacement of its atomic positions from the final model of 0.275 Å. The phases calculated from this refined polyalanine model had an average difference of 43.8° from the final phase set (corresponding to a mean figure of merit of 0.72) and gave a readily interpretable electron-density map.

1. Introduction

The repressor of primer (Rop) protein is a homodimeric RNA-binding protein involved in the regulation of the copy number of the ColE1 plasmid (Polisky, 1988). Rop is the paradigm of a canonical 4-α-helical bundle (Banner *et al.*, 1987; Cohen & Parry, 1990; Kamtekar & Hecht, 1995; Eberle *et al.*, 1991) and has been the target of an extensive structure–function investigation programme combining site-directed and random mutagenesis with thermodynamic and structural characterization of the resulting mutants (Castagnoli *et al.*, 1989, 1994; Vlassi *et al.*, 1994; Predki *et al.*, 1995, 1996; Steif *et al.*, 1995; Munson *et al.*, 1996, 1997; Nagi & Regan, 1997; Peters *et al.*, 1997; Lassalle *et al.*, 1998).

We have recently completed the determination of the crystal structure of the Ala31Pro mutant of Rop and have found that this single amino-acid substitution in the turn region results in a complete reorganization of the whole protein, which is converted from the canonical left-handed all-antiparallel form to a right-handed mixed parallel and anti-parallel 4-α-helical bundle (Glykos *et al.*, 1999).

Here, we discuss the experiences gained from this structure determination, which confronted us with the problem of having only one derivative, isomorphous to approximately 4 Å, with no anomalous signal recorded for it, and only 35% solvent content for the crystal form under investigation. The application of a novel procedure involving a one-residue-per-rigid-body simulated annealing of polyalanine models will be discussed in some detail.

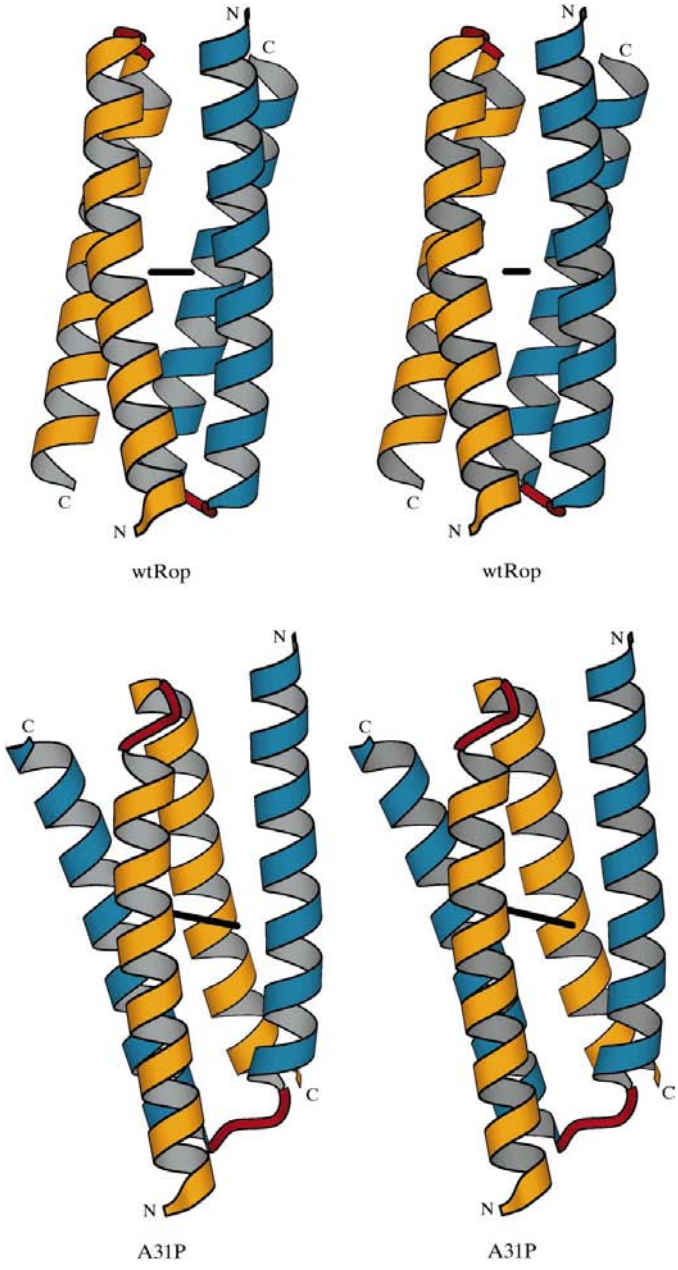

Figure 1
Comparative stereodiagrams of the overall structures of wild-type Rop and the A31P mutant. The colour coding is yellow and blue for the helices of each monomer, red for the connective strands. The position and orientation of the intramolecular (crystallographic) dyad axis is noted in both structure schematics. The orientations of the wild-type and mutant structures is such that the N-terminal helices of the blue-coloured monomers are in exactly the same position and orientation in both diagrams. The structure schematics were prepared with the program *BOBSCRIPT* (Esnouf, 1997).

2. The dead-end stories

The expression, purification, crystallization and preliminary crystallographic characterization of the A31P Rop mutant was reported quite some time ago (Castagnoli *et al.*, 1989; Kokkinidis *et al.*, 1993), but its structure determination turned out to be somewhat more difficult than originally anticipated. Numerous attempts to attack the problem with molecular-replacement methods all failed, leading with increased confidence to the conclusion that this mutant was significantly different from the wild-type protein. The refined mutant structure (compared in Fig. 1 with the wild-type Rop structure) more than justified these expectations: with an r.m.s. deviation between the equivalent atoms of the least-squares superimposed wild-type and A31P structures of σ_Δ = 5.97 Å and a maximum separation of $\max_\Delta$ = 13.99 Å (*LSQKAB* program; Collaborative Computational Project, Number 4, 1994), it is hardly surprising that even sophisticated molecular-replacement methods (such as Patterson correlation refinement in *X-PLOR*; Brünger, 1990, 1992*a*) could not converge to the correct solution when using the Rop monomer as a search model. The degree of structural difference observed between individual helices from the wild-type and mutant structures (with σ_Δ = 1.15 Å, $\max_\Delta$ = 5.28 Å for the first helix, σ_Δ = 2.16 Å, $\max_\Delta$ = 9.85 Å for the second helix), explains our failure to identify a correct solution even when these helices were used individually as search models in molecular-replacement calculations. Attempts using isolated polyalanine helix models could in principle have fared better, but the signal from a search model accounting for only ∼25% of the scattering material in the asymmetric unit (containing one monomer, *i.e.* two connected helices) and in such a tightly packed crystal form (35% solvent content) was too weak to be detected. When an exhaustive systematic search with *AMoRe* (Navaza, 1987; *CCP*4 suite of programs) using two polyalanine helices as search models[1] resulted in a more-or-less uniform distribution of the linear correlation coefficients from the translation functions, it meant an end to our efforts using molecular-replacement methods.

The discovery that a data set collected from a platinum-soaked crystal (Papanikolaou *et al.*, 1999) was isomorphous with a high-resolution native data set collected on a CAD4 diffractometer (Papanikolaou *et al.*, 1999; data set collected as described by Banner *et al.*, 1987; see Table 1 for statistical information), set the stage for the final structure determination. This single-site derivative was, as shown in Table 2, highly isomorphous at low angles, but the lack of isomorphism at higher angles together with the complete absence of anomalous differences (only 16 reflections to 3.8 Å had a Bijvoet pair measurement) and the very low solvent content compli-

[1] This search was conducted as follows: the first polyalanine helix was fixed in orientation and position by combining the best 99 orientations from its rotation function with the top 20 peaks from each of the corresponding translation functions, giving a total of 1980 starting models for the first helix. For each of these models, we calculated the translation functions corresponding to each of the 99 best orientations for the second helix, giving a grand total of 1980 × 99 = 196020 translation functions or a list of 3920400 correlation coefficients.

Table 1
Statistical information for the native (CAD4) data set.

Space group $C222_1$ with unit-cell parameters $a = 30.40$, $b = 42.10$, $c = 81.40$ Å, one monomer (residues 1–63) per asymmetric unit and 35% solvent content. The complete (biologically active) homodimer is formed through the application of a crystallographic twofold axis running parallel to *b*.

d_{max} (Å)	d_{min} (Å)	Number of reflections	Completeness (%)	Mean $F/\sigma(F)$
∞	3.88	556	99.8	51.79
3.88	3.08	527	100.0	26.56
3.08	2.69	515	100.0	11.19
2.69	2.44	512	100.0	9.40
2.44	2.27	505	100.0	7.18
2.27	2.13	492	100.0	5.95
2.13	2.03	496	100.0	4.35
2.03	1.94	508	100.0	3.26
1.94	1.86	496	100.0	2.80
1.86	1.80	499	99.6	2.48

Table 2
Statistical information from the refinement of the platinum derivative.

Output from the program *MLPHARE* (Otwinowski, 1991; Collaborative Computational Project, Number 4, 1994).

d_{min} (Å)	Phasing power (centric)	Cullis *R* factor (centric)
11.88	5.60	0.17
8.77	4.39	0.26
6.95	4.27	0.25
5.76	2.64	0.37
4.91	1.66	0.52
4.29	1.03	0.70
3.80	0.83	0.74
Total	1.99	0.45

cated the phase-determination process. A 5 Å solvent-flattened map (shown in Fig. 2, solvent flattening was performed with the program *SOLOMON* from the *CCP*4 suite of programs; Wang, 1985; Abrahams & Leslie, 1996), clearly defined both the boundary of the bundle and the position of the individual helices. In retrospect, this is an excellent low-resolution map: not only are the positions of all four helices clearly and correctly identified, but additionally the angle and proximity of the helices (which – as it turned out later – belong to the same monomer) could have guided us to the (correct) conclusion that what we have is not a localized change in the turn region, but a change of topology. Unfortunately, at this stage in the analysis we had not been expecting the unexpected, and this, together with the presence of erroneous density connecting helices on the same site of the crystallographic twofold axis (shown by an arrow in Fig. 2) and a break in the density connecting the correct monomer helices, postponed the structure solution for quite some time. Numerous attempts with *SOLOMON* and *DM* (both programs from the *CCP*4 suite) to refine the *SIR* phases at 3.8 Å resolution [using, in the case of *DM* (Cowtan, 1994), both solvent flattening and histogram mapping (Zhang & Main, 1990)] all gave electron-density maps with a clear definition of the position of the helices, but very little side-chain information. Attempts to fit (using the program *O*; Jones *et al.*, 1991) and refine (with *X-PLOR*) polyalanine or complete (with side chains) helices from the wild-type structure were all unsuccessful.

The map (shown in Fig. 3) which actually made us realise that the bundle topology may indeed be different (and that the helices lying on the same site of the twofold axis could be parallel after all) was the result of a rather questionable procedure: the phases and figures of merit obtained from a low-resolution (5 Å) run of *SOLOMON* were treated as observed and were merged with the experimentally determined (*SIR*) phases in the 5–4 Å resolution range. This procedure gives excessive weight to the low-resolution phases by treating as experimentally observed the notoriously overestimated figures of merit from solvent flattening. Our hope was that the (assumed) reasonably accurate low-resolution phases would drive the development of the rest of the phase set in the correct direction. The map obtained by transforming this 4 Å mixed-phase data set was used as a starting point for several rounds of solvent flattening (with *SOLOMON*) and phase combination (with *SIGMAA* from the *CCP*4 suite of programs; Read, 1986, 1997) at constant resolution. Although the resulting map (Fig. 3) was again uninterpretable in terms of the protein sequence,[2] its appearance did suggest that two parallel helices (on the same site of the crystallographic twofold axis) could account for the observed density better than the expected antiparallel arrangement. The problem, again, was that the density corresponding to the correct inter-helix connection (shown by the red arrow in Fig. 3), was weaker than a false connection (shown by the black arrow in Fig. 3) between helices lying on the same site of the twofold axis. This ambiguity forced us to treat the parallel and antiparallel arrangements as being equally probable until the very end of the determination, and to repeat all calculations that followed with two models, one with parallel and the other with antiparallel helices.[3]

3. Hit, not miss

To summarize, the problem at this stage of the structure determination was that although we knew the position of both helices and the orientation of one of them, the lack of convincing and consistent side-chain density made it impossible to decide which helix was which, what their connectivity was and what their orientation about the helical axes was. Several attempts to fit (using the program *O*) individual helices (with side chains) from the wild-type Rop structure into the 4 Å map discussed above, followed by conjugate-gradient rigid-body refinement and simulated-

[2] The lack of meaningful side-chain information is not too surprising given that this is a low-resolution (4 Å) map and that the phases used for calculating it had a mean difference from the final phase set of 56°.

[3] A reader may object that even with the origin and enantiomorph fixed (from the derivative), there are still four (instead of two) possible arrangements for the helices: two of these models consist of parallel helices (schematically, ↑↑ and ↓↓) and the other two have the antiparallel arrangement (↑↓ and ↓↑). We could reduce the number of possible models to two only because the density corresponding to one of the helices in the 4 Å map had a convincing and consistent 'Christmas-tree' appearance which allowed us to fix its polarity for all subsequent calculations.

annealing refinement (both in *X-PLOR*), all resulted in unrealistically distorted models giving unacceptably high $R_{\rm free}$ values (Brünger, 1992*b*, 1997). Considerable effort was also expended on attempts to fit individual polyalanine helices (again obtained from the wild-type Rop structure) in the 4 Å map, followed by conjugate-gradient rigid-body refinement (in *X-PLOR*), phase combination with the experimental phases (using *SIGMAA* from the *CCP*4 suite) and calculation of a $(2mF_o - DF_c)\exp(i\varphi_{\sigma_A})$ map (Read, 1997). None of the maps examined showed any improvement in the side-chain density.

The procedure which finally opened the way to the structure determination was essentially a hit-or-miss approach, whose basic steps are outlined below. Two polyalanine helices (each 26 residues long, extracted from the wild-type Rop structure) were placed manually in the 4 Å map (using *O*) in both the parallel and the antiparallel orientations and the resulting models were refined (with *X-PLOR*) using rigid-body simulated annealing at a very high temperature ($T_0 = 10000$ K) and with the geometric energy terms switched on, as follows: in the first step, both helices were treated as one rigid body and their orientation and position was refined, followed by optimization of the orientation and position of each helix separately, then of collections of three alanine residues per body and, finally, of two residues per body. This procedure was successively applied against 5, 4, 3 and 2 Å resolution data from the native data set. Whereas all solutions with antiparallel helices converged to R and $R_{\rm free}$ values of approximately 0.50, the best parallel-helix solution converged to $R = 0.407$, $R_{\rm free} = 0.414$ for all data in the 8–2 Å resolution range. This incomplete polyalanine model was further refined with one-residue-per-rigid-body conjugate-gradient minimization in *X-PLOR* (again with the geometric energy terms switched on) and converged to $R = 0.403$ and $R_{\rm free} = 0.413$ against all data from 8 to 1.8 Å. As shown in Fig. 4, the 1.8 Å σ_A-weighted map of the form $(2mF_o - DF_{\rm Ala52})\exp(i\varphi_{\rm Ala52})$ was readily interpretable in terms of the protein sequence and the inter-helix connectivity.

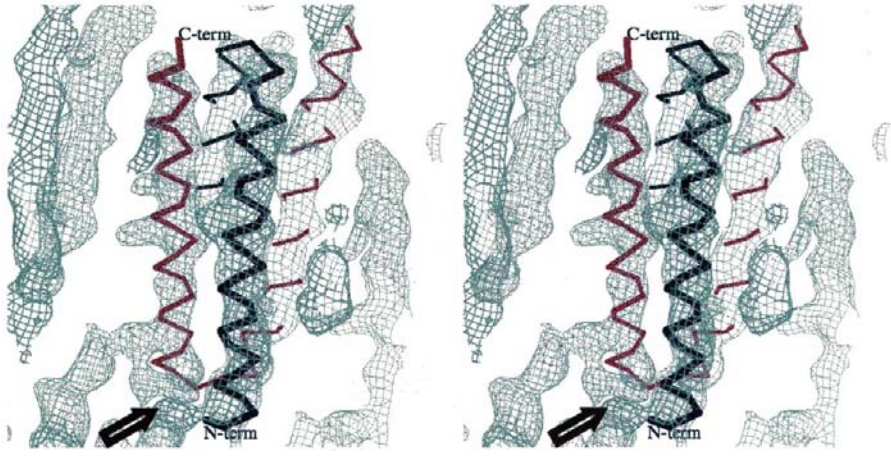

Figure 2
Stereodiagram of the 5 Å solvent-flattened electron-density distribution with the C^{α} trace of the final model superimposed (see text for details). Individual monomers are colour coded and the N- and C-termini are noted. The arrow points to a false connection between helices which belong to different monomers. The map is contoured at 1.0σ above the mean, and all density features in the volume shown are drawn. All electron-density plots were prepared with the programs *O* and *Oplot* (Jones *et al.*, 1991).

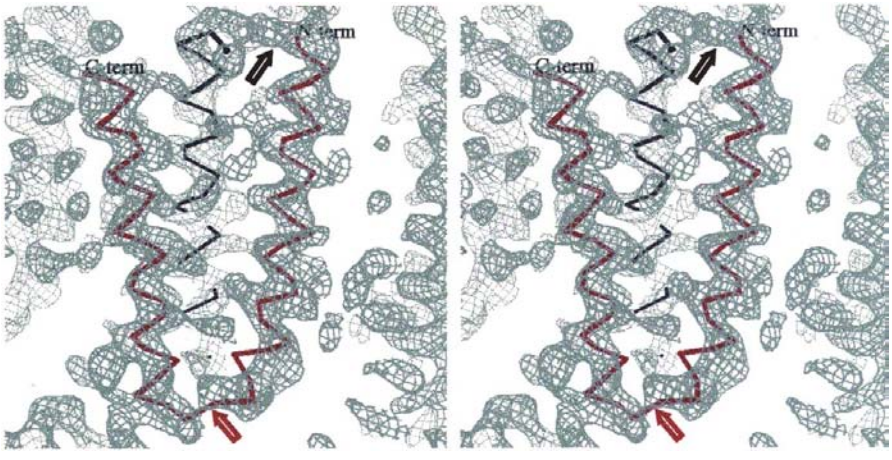

Figure 3
Stereodiagram of the 4 Å solvent-flattened electron-density distribution calculated as described in the text. The C^{α} trace of the final model is superimposed and the individual monomers are colour coded. The red arrow points to the correct inter-helix connection, the black arrow points to a false connection. The map is contoured at 1.0σ above the mean and all density features in the volume shown are drawn.

The availability of a medium- to high-resolution data set made it possible to improve this map even further – and before actually constructing a complete atomic model – by using the *ARP* program (Lamzin & Wilson, 1993, 1997) as implemented in the *wARP* procedure (Perrakis *et al.*, 1997). Fig. 5 shows a relatively large volume from a *wARP* map calculated assuming 600 atoms per asymmetric unit: if the original map was not difficult to interpret, the *wARP* map was definitely difficult to misinterpret. Side chains were built using *O*, with reference to both the original $(2mF_o - DF_{\rm Ala52})\exp(i\varphi_{\rm Ala52})$ map and a collection of *wARP* maps calculated with different refinement protocols and assuming different number of atoms per asymmetric unit. The refinement was completed with rounds of model building in *O* and conjugate-gradient refinement in *X-PLOR*, which included bulk-solvent correction, overall anisotropic scaling between the F_os and the F_cs and a resolution-dependent two-line

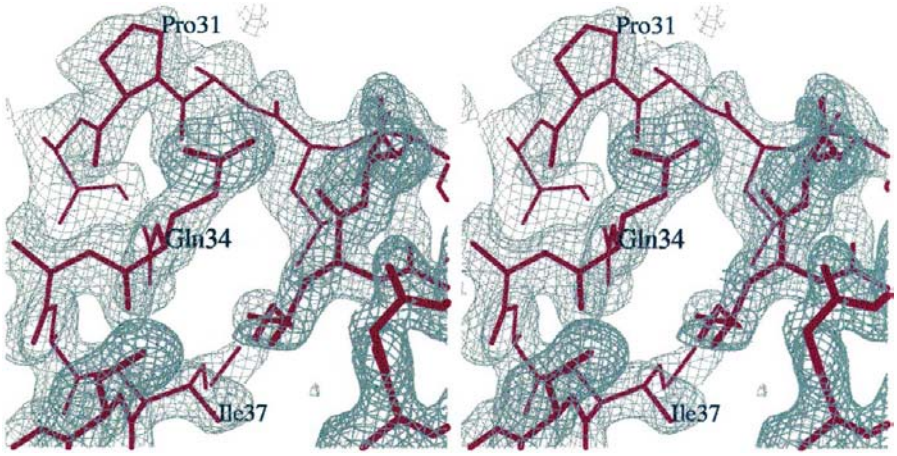

Figure 4
1.8 Å electron-density map of the form $(2mF_o - DF_{\rm Ala52})\exp(i\varphi_{\rm Ala52})$ with phases from the incomplete (but refined) polyalanine model. The map is contoured around the mutated proline and the turn region (which was not part of the polyalanine model). The final model is shown superimposed and residues Pro31, Gln34 and Ile37 have been labelled. The map is contoured at 1.0σ above the mean and for clarity contours have only been drawn around the protein model.

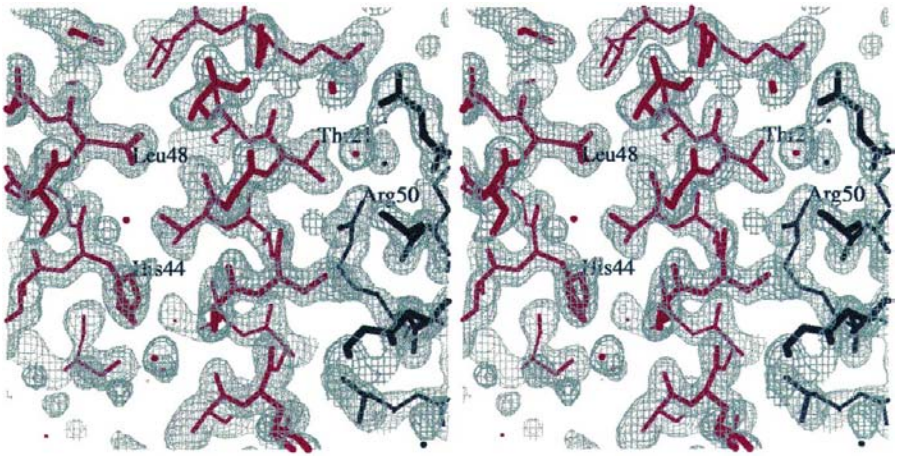

Figure 5
1.8 Å *wARP* map (see text for details) contoured around a crystal packing contact between the first and second helix of the monomer (coloured black and red, respectively). Residues Leu48, His44, Thr21 and Arg50 are labelled. The map is contoured at 1.0σ above the mean and all density features in the volume shown are drawn.

weighting scheme as suggested by Smith (1996). Fig. 6 shows part (in a view similar to that used in Fig. 5) of the final sharpened σ_A-weighted difference map of the form $(2mF_o - DF_c)^{0.6}E^{0.4}_{2mF_o-DF_c}\exp(i\varphi_c)$, where $E_{2mF_o-DF_c}$ are the normalized structure-factor amplitudes corresponding to $2mF_o - DF_c$ (Butterworth *et al.*, 1996; Gamblin *et al.*, 1996). A comparison between Figs. 5 and 6 (taking into account that the map in Fig. 6 is sharpened), provides convincing evidence for our statement that the 1.8 Å *wARP* maps were difficult to misinterpret.

The final model comprises 56 residues (Met1–Phe56) and 55 water molecules, with the terminal atoms of the solvent-exposed side chains of Leu9, Arg16, Asn27 and Arg55 excluded owing to lack of convincing density. This model has an R factor of 0.188 and an $R_{\rm free}$ of 0.240 for all data between infinity and 1.8 Å [R = 0.149, $R_{\rm free}$ = 0.191 for all data with $F/\sigma(F) > 3.0$]. The model scores better than average on all *PROCHECK* tests (Laskowksi *et al.*, 1993; *CCP*4 program suite), giving an overall G factor of +0.56, with 100% of the residues in the core Ramachandran regions, an average standard deviation for main-chain bond lengths and angles of 0.007 Å and 0.909°, respectively, and B-factor r.m.s. deviations for main-chain and side-chain bonds of 1.55 and 2.77 Å^2, respectively.

4. Much ado about nothing?

In Fig. 7 we compare the starting (red) and the refined (green) polyalanine models. At first sight this comparison reveals nothing exciting: it would even appear that most differences between the two models could be accounted for by a rigid-body translation and rotation of the whole helices. In the following paragraphs, we will try to provide evidence supporting our thesis that the calculations described in the previous section constitute a meaningful polyalanine refinement procedure and amount to more than just moving the helices (as rigid bodies) to the correct position in the unit cell.

Our first and most convincing – at least to us – argument is that the method converged to two absolutely normal (with respect to their geometry) polyalanine helices only because this was the underlying structure: when exactly the same calculations were performed on the model with the antiparallel helical arrangement, the resulting structure was so unrealistically distorted as to leave no space for misinterpretations.[4] This finding, together with the observation that all false solutions gave consistently high values for both R and $R_{\rm free}$ (>0.49) whereas all correct solutions were in the region of 0.40, suggests that this is a dependable method. The closeness and reliability of both the R and $R_{\rm free}$ values is, we believe, a direct

[4] It is worth noting here that the starting (before any refinement) parallel and antiparallel models had one helix in common (identical in the two structures), whereas their second helices had opposite direction but with their C$^\alpha$ atoms as close in space as possible. This resulted in an r.m.s. deviation in the atomic positions of the C$^\alpha$ atoms from the two arrangements of only 0.91 Å, with an average displacement of 0.65 Å.

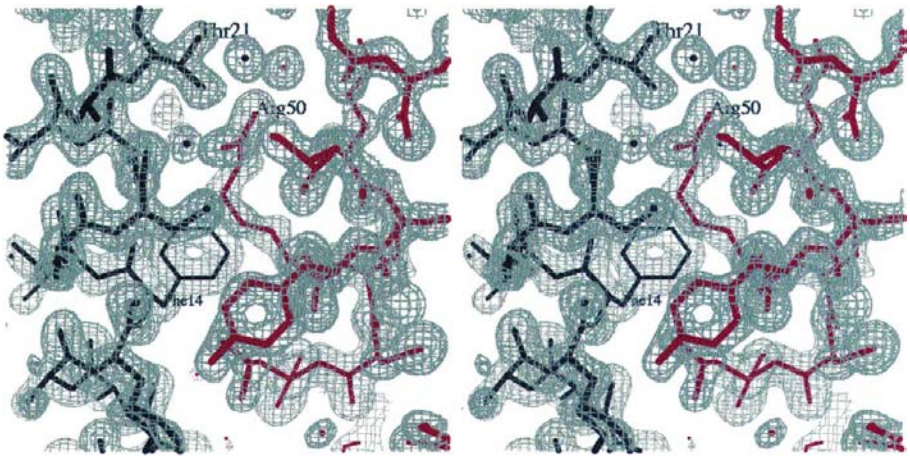

Figure 6
Final 1.8 Å electron-density map calculated with coefficients $(2mF_o - DF_c)^{0.6} E^{0.4}_{2mF_o-DF_c} \exp(i\varphi_c)$, where $E_{2mF_o-DF_c}$ are the normalized structure-factor amplitudes corresponding to $2mF_o - DF_c$. The map is contoured at 1.5σ above the mean and all density features in the volume shown are drawn. Residues Thr21, Phe14 and Arg50 have been labelled.

consequence of the reduced number of the degrees of freedom of the model combined with the overdeterminancy of the problem.[5]

Our second argument is essentially a refutation of the statement which opened this section: the differences between the two models cannot be accounted for by a mere rigid-body operation applied to the whole helices. To start with, the arrows in Fig. 7 point to similarities and differences between the two models which are inconsistent with a simple rigid-body movement of the whole helices. To put this observation into figures, we compared (using least-squares superposition, program *LSQKAB*, *CCP*4 program suite) the atomic arrangement of individual polyalanine helices before and after the annealing procedure: for the first polyalanine helix we obtained $\sigma_\Delta = 0.67$ Å, $\max_\Delta = 2.28$ Å and an average displacement $\bar{\Delta} = 0.57$ Å, and for the second helix we obtained $\sigma_\Delta = 0.69$ Å, $\max_\Delta = 2.20$ Å and $\bar{\Delta} = 0.59$ Å. The important thing, of course, is that these changes in the atomic arrangement are in the correct direction (towards the target structure): if we use these least-squares superimposed polyalanine helices (obtained from the wild-type Rop) to calculate phases, we increase the mean phase difference from the final phase set from 43.8 to 56.9° (corresponding to a reduction of the mean figure of merit[6] from 0.72 to 0.54).

The last piece of evidence supporting our view that one-residue-per-rigid-body refinement of polyalanine models is meaningful comes from a comparison with the results obtained from conventional positional refinement at high resolution: when the final rigid-body-refined polyalanine model (obtained as discussed in §3) was further refined at 1.8 Å resolution using conjugate-gradient least-squares minimization (in *X-PLOR*), the resulting model was very similar to the original with a $\sigma_\Delta = 0.19$ Å, $\max_\Delta = 0.65$ Å and $\bar{\Delta} = 0.15$ Å (which is less than the estimated coordinate error). This is not to imply that positional refinement improved the polyalanine model; it improved the R factor (which was reduced from 0.403 to 0.393), but it also increased $R_{\rm free}$ (from 0.413 to 0.433), the mean phase difference from the final model phases (from 43.8 to 46.0°) and the mean phase difference from the correct polyalanine trace phases (from 28.7 to 31.2°).

5. Discussion

In the preceding sections, we have shown that iterative rigid-body simulated-annealing refinement of a roughly positioned polyalanine model, employing successively fewer residues per rigid body and at successively higher resolution, can converge to an essentially correct polyalanine trace of a target structure, producing phase information sufficient for a complete structure determination. Although the method certainly did work in the case examined, there are a number of questions related to its applicability to other problems.

The first is whether the method could be equally efficient with a larger not all-α structure. Although we have no experimental evidence for this, we do think that as long as the starting polyalanine model has the correct secondary-structure elements at approximately the right place, the method could be equally efficient. On the other hand, we would be surprised if, for example, starting from an α-helical polyalanine model the method could (correctly) converge to a β structure.

The second important question is whether the method could have worked in the absence of 1.8 Å data. Model calculations with error-free data showed that this refinement strategy can converge to a correct solution even at 3 Å resolution (data not shown). The problem, of course, is that as we move towards lower resolution, the electron-density maps phased from an even perfectly correct polyalanine model contain less information about the side chains (which is true even for maps calculated with perfect phases). These same model calculations showed that at 3 Å it might just be possible to gain some side-chain information from a polyalanine-phased electron-density map.

[5] In our case and with 52 alanine residues, we had $52 \times 6 = 312$ parameters (three positional and three rotational for each residue) which were refined against 5103 reflections (at 1.8 Å), giving a ratio of 16 observations per parameter; this is without taking the geometric energy restraints into account.
[6] Formally, it is not correct to use the 'figure of merit' in this context: the target of our refinement is not the whole crystal structure, but only the main-chain protein atoms. It would probably be acceptable to use the figure of merit (with its $m = \langle \cos \Delta\varphi \rangle$ meaning) when comparing the annealed polyalanine model with the same subset of atoms from the final structure (in which case the figure of merit turns out to be 0.87, with a mean phase difference of 28.7°), but not otherwise. We have been using this (incorrect) expression only because most crystallographers find it easier to estimate the difference (in terms of electron-density map quality) between $\langle m \rangle = 0.72$ and $\langle m \rangle = 0.54$ than between $\langle \Delta\varphi \rangle = 43.8°$ and $\langle \Delta\varphi \rangle = 56.9°$.

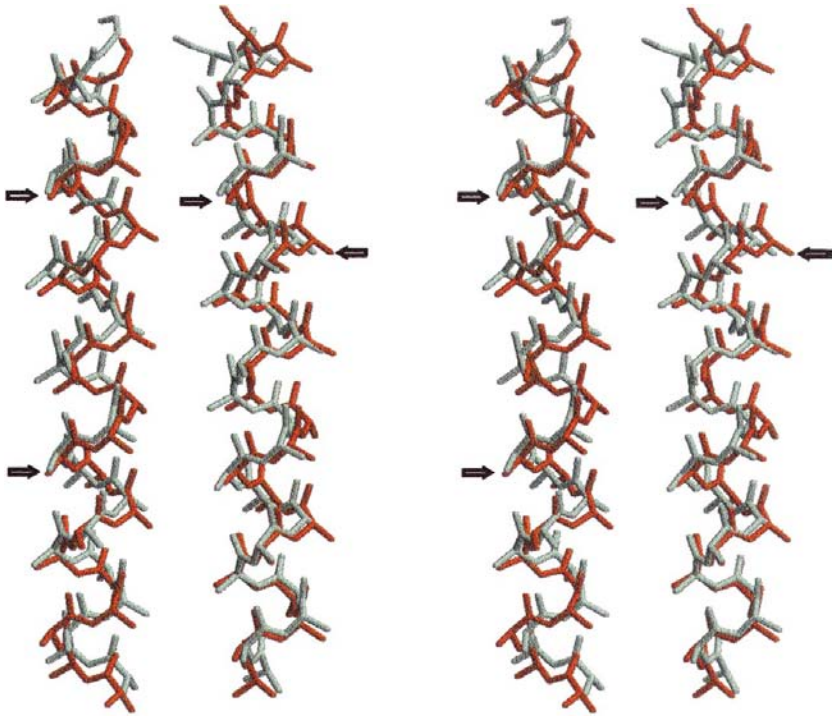

Figure 7
Comparative stereodiagram of the starting (red) and final (green) polyalanine model before and after rigid-body simulated annealing. The arrows point to similarities and differences between the two models which are inconsistent with a simple rigid-body movement of the whole helices.

One final question is whether this refinement strategy is the best available or even the only one possible. We think that the answer to both questions is clearly no. The most obvious improvement is a change of the target function for the minimization: neither conventional simulated-annealing nor conjugate-gradient least-squares methods can deal in a satisfactory manner with missing structure, and this is exactly the name of the game with the refinement of polyalanine models. Switching from atomic to rigid-body refinement improved, as we have shown, the accuracy of the refinement, but this is, in a way, more avoiding the refinement than actually performing it. A maximum-likelihood target function would definitely be more suited to the problem under investigation and might even make the rigid-body approximation (during the last stages of the refinement) unneccessary (Pannu & Read, 1996; Bricogne & Irwin, 1996; Murshudov *et al.*, 1996; Read, 1997; Adams *et al.*, 1997; Brünger *et al.*, 1998). A second possible improvement is changing the refinement method: rigid-body simulated annealing with one alanine residue per body and with the geometric energy terms switched on is equivalent to an inefficient φ–ψ torsion-angle refinement. Recent calculations[7] seem to confirm the expectation that torsion-angle dynamics is the most natural and efficient solution to the final stages of polyalanine refinement (Rice & Brünger, 1994; Adams *et al.*, 1996).

[7] Four torsion-angle dynamics minimizations were performed using a slow-cooling protocol in *X-PLOR*. The best of these minimizations (based on the R_{free} value) gave R = 0.401 and R_{free} = 0.424 against all data in the 8–1.8 Å resolution range and a mean phase difference from the final phase set of 44.5°, which is very similar to the results obtained from the rigid-body simulated-annealing method (R = 0.403, R_{free} = 0.413, $\Delta\varphi$ = 43.8°).

References

Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* D**52**, 30–42.

Adams, P. D., Braig, K., Rice, L. M. & Brünger, A. T. (1996). In *Proceedings of the CCP4 Study Weekend. Macromolecular Refinement*, edited by E. Dodson, M. Moore, A. Ralph & S. Bailey. Warrington: Daresbury Laboratory.

Adams, P. D., Pannu, N. S., Read, R. J. & Brünger, A. T. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 5018–5023.

Banner, D. W., Kokkinidis, M. & Tsernoglou, D. (1987). *J. Mol. Biol.* **196**, 657–675.

Bricogne, G. & Irwin, J. (1996). In *Proceedings of the CCP4 Study Weekend. Macromolecular Refinement*, edited by E. Dodson, M. Moore, A. Ralph & S. Bailey. Warrington: Daresbury Laboratory.

Brünger, A. T. (1990). *Acta Cryst.* A**46**, 46–57.

Brünger, A. T. (1992*a*). *X-PLOR Version* 3.1. *A System for X-ray Crystallography and NMR.* Connecticut: Yale University Press.

Brünger, A. T. (1992*b*). *Nature (London)*, **355**, 472–474.

Brünger, A. T. (1997). *Methods Enzymol.* **277**, 366–396.

Brünger, A. T., Adams, P. D. & Rice, L. M. (1998). *Curr. Opin. Struct. Biol.* **8**, 606–611.

Butterworth, S., Lamzin, V. S. & Wilson, K. S. (1996). In *Proceedings of the CCP4 Study Weekend. Macromolecular Refinement*, edited by E. Dodson, M. Moore, A. Ralph & S. Bailey. Warrington: Daresbury Laboratory.

Castagnoli, L., Scarpa, M., Kokkinidis, M., Banner, D. W., Tsernoglou, D. & Cesareni, G. (1989). *EMBO J.* **8**, 621–629.

Castagnoli, L., Vetriani, C. & Cesareni, G. (1994). *J. Mol. Biol.* **237**, 378–387.

Cohen, C. & Parry, D. A. D. (1990). *Proteins*, **7**, 1–15.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cowtan, K. (1994). *Jnt CCP4 ESF/EACBM Newslett. Protein Crystallogr.* **31**, 34–38.

Eberle, W., Pastore, A., Sander, C. & Rosch, P. (1991). *J. Biomol. NMR*, **1**, 71–82.

Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–134.

Gamblin, S. J., Rodgers, D. W. & Stehle, T. (1996). In *Proceedings of the CCP4 Study Weekend. Macromolecular Refinement*, edited by E. Dodson, M. Moore, A. Ralph & S. Bailey. Warrington: Daresbury Laboratory.

Glykos, N. M., Cesareni, G. & Kokkinidis, M. (1999). In the press.

Jones, T. A., Zou, J. Y., Cowan, S. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kamtekar, S. & Hecht, M. H. (1995). *FASEB J.* **9**, 1013–1022.

Kokkinidis, M., Vlassi, M., Papanikolaou, Y., Kotsifaki, D., Kingswell, A., Tsernoglou, D. & Hinz, H.-J. (1993). *Proteins*, **16**, 214–216.

Lamzin, V. S. & Wilson, K. S. (1993). *Acta Cryst.* D**49**, 129–147.

Lamzin, V. S. & Wilson, K. S. (1997). *Methods Enzymol.* **277**, 269–305.

Laskowksi, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Lassalle, M. W., Hinz, H.-J., Wenzel, H., Vlassi, M., Kokkinidis, M. & Cesareni, G. (1998). *J. Mol. Biol.* **279**, 987–1000.

Munson, M., Anderson, K. S. & Regan, L. (1997). *Fold. Des.* **2**, 77–87.

Munson, M., Balasubramanian, S., Fleming, K. G., Nagi, A. D., O'Brien, R., Sturtevant, J. M. & Regan, L. (1996). *Protein Sci.* **5**, 1584–1593.
Murshudov, G. N., Dodson, E. J. & Vagin, A. A. (1996). In *Proceedings of the CCP4 Study Weekend. Macromolecular Refinement*, edited by E. Dodson, M. Moore, A. Ralph & S. Bailey. Warrington: Daresbury Laboratory.
Nagi, A. D. & Regan, L. (1997). *Fold. Des.* **2**, 67–75.
Navaza, J. (1987). *Acta Cryst.* A**43**, 645–653.
Otwinowski, Z. (1991). In *Proceedings of the CCP4 Study Weekend. Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie. Warrington: Daresbury Laboratory.
Pannu, N. S. & Read, R. J. (1996). *Acta Cryst.* A**47**, 110–119.
Papanikolaou, Y., Kotsifaki, D., Fadouloglou, V., Glykos, N. M., Cesareni, G. & Kokkinidis, M. (1999). Submitted.
Perrakis, A., Sixma, T. K., Wilson, K. S. & Lamzin, V. S. (1997). *Acta Cryst.* D**52**, 659–668.
Peters, K., Hinz, H.-J. & Cesareni, G. (1997). *Biol. Chem.* **378**, 1141–1152.
Polisky, B. (1988). *Cell*, **55**, 929–932.
Predki, P. F., Agrawal, V., Brünger, A. T. & Regan, L. (1996). *Nature Struct. Biol.* **3**, 54–58.
Predki, P. F., Nayak, L. M., Gottlieb, M. B. C. & Regan, L. (1995). *Cell*, **80**, 41–50.
Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.
Read, R. J. (1997). *Methods Enzymol.* **277**, 110–128.
Rice, L. M. & Brünger, A. T. (1994). *Proteins*, **19**, 277–290.
Smith, G. D. (1996). In *Proceedings of the CCP4 Study Weekend. Macromolecular Refinement*, edited by E. Dodson, M. Moore, A. Ralph & S. Bailey. Warrington: Daresbury Laboratory.
Steif, C., Hinz, H.-J. & Cesareni, G. (1995). *Proteins*, **23**, 83–96.
Vlassi, M., Steif, C., Weber, P., Tsernoglou, D., Wilson, K. S., Hinz, H.-J. & Kokkinidis, M. (1994). *Nature Struct. Biol.* **1**, 706–716.
Wang, B.-C. (1985). *Methods Enzymol.* **115**, 90–112.
Zhang, K. Y. J. & Main, P. (1990). *Acta Cryst.* A**46**, 377–381.