Journal of Computer-Aided Molecular Design, **12:** 605–614, 1998.
KLUWER/ESCOM

The reconstruction of atomic co-ordinates from a protein stereo ribbon diagram when additional information for sufficient sidechain positions is available

Paulo Sérgio Lopes de Oliveira[a,b] & Richard Charles Garratt[a,*]
[a]*Departamento de Fìsica e Informática, Instituto de Fìsica de São Carlos, São Carlos, Brazil;* [b]*Departamento de Fìsico-Quìmica, Instituto de Quìmica de São Carlos, Universidade de São Paulo, Caixa Postal 369, São Carlos - SP, CEP 13560-970, Brazil*

Received 20 November 1997; Accepted 6 July 1998

Key words: jacalin, lectin, 3D structure

Summary

We describe the application of a method for the reconstruction of three-dimensional atomic co-ordinates from a stereo ribbon diagram of a protein when additional information for some of the sidechain positions is available. The method has applications in cases where the 3D co-ordinates have not been made available by any means other than the original publication and are of interest as models for molecular replacement, homology modelling etc. The approach is, on the one hand, more general than other methods which are based on stereo figures which present specific atomic positions, but on the other hand relies on input from a specialist. Its exact implementation will depend on the figure of interest. We have applied the method to the case of the α-D-galactose-binding lectin jacalin with a resultant RMS deviation, compared to the crystal structure, of 1.5 Å for the 133 C_α positions of the α-chain and 2.6 Å for the less regular β-chain. The success of the method depends on the secondary structure of the protein under consideration and the orientation of the stereo diagram itself but can be expected to reproduce the mainchain co-ordinates more accurately than the sidechains. Some ways in which the method may be generalised to other cases are discussed.

Introduction

A recent estimate indicates that by the end of this century the crystallographic and NMR communities will be responsible for the determination of a new protein structure every 30 min [1]. For the crystallographer it is to be expected that a growing proportion of these new structures will be determined by molecular replacement. The importance of such methods will be due to improvements in computational molecular replacement methods [2] together with the very growth of the Protein DataBank (PDB, [3]) itself which increases the probability of there being an homologous protein of known 3D structure available. Future developments in so-called threading techniques [4] may extend the use of molecular replacement to yet more difficult cases in the future.

One of the pre-requisites for the normal application of molecular replacement in the solution of a new protein structure is access to the 3D co-ordinates of an homologous molecule. Often such co-ordinates are not made generally available despite the fact that frequently the structure has already been reported in the literature. Furthermore molecular replacement is far from being the only potential application for such structures and their inaccessibility can be frustrating for researchers interested in their use for molecular modelling applications or for the design of future experiments.

Several years ago Michael Rossmann proposed a solution to this problem by suggesting that 3D co-ordinates could be readily recovered from a 2D stereo

*To whom correspondence should be addressed.

image of the molecule in question. Given that a stereo pair is nothing more than two different views of the same three-dimensional object related by a small rotation, all of the information necessary for the regeneration of the orthogonal co-ordinates (*xyz*) of any point on the object is present within the figure if the relative rotation angle between the two views (θ) and the scale factor (S_f) are known.

Several programs have subsequently been developed founded on the original idea but to our knowledge all have been based on the reconstruction of the molecular structure from a stereo image which displays specific atomic positions (normally C_α) [5–7]. One of the retrograde consequences of such a capacity has been a reluctance on the part of authors to publish stereo images of C_α traces in cases where there is no intention to make the co-ordinates publicly available in the short term. As an alternative either single views of C_α-traces or stereo views of ribbon diagrams are frequently published.

In this paper we describe briefly an attempt to recreate the three-dimensional co-ordinates of the α-D-galactose-binding lectin jacalin from a stereo image of a ribbon diagram of the structure [8]. Such a diagram provides no explicit atomic positions (Figure 1, upper panel). We compare the derived co-ordinates with those of the genuine crystal structure which were subsequently released by the PDB.

Methods

In the reconstruction of 3D co-ordinates from any 2D stereo figure it is necessary to determine the co-ordinates (x_L,y_L) and (x_R,y_R) of equivalent points on the left and right-hand images respectively. Here the X and Y directions are taken to have their normal meanings, i.e. horizontal and vertical in the plane of the figure respectively. These 2D co-ordinates can be most readily obtained by digitally scanning the stereo figure as it appears in the original paper. If we assume that the rotation which relates the two images has been made parallel to the Y axis, then the final co-ordinates (x_F,y_F,z_F) of the point in 3D space are given by:

$$x_F = \frac{1/2(x_L + x_R)}{S_f}$$

$$y_F = \frac{y_L}{S_f} = \frac{y_R}{S_f}$$

$$z_F = \frac{(D - (x_R - x_L))}{2\tan(\theta/2) \cdot S_f}$$

where S_f is a scale factor which converts the co-ordinates from the units of measurement (pixels, mm, inches…) into Å, θ is the angle which relates the two images and D is the distance between the mean x co-ordinate of the left- and right-hand images. These equations are thus simplifications of those derived for the more general case where no assumption need be made about the y co-ordinate being constant [7].

In the case of a C_α trace for example, S_f and θ can be readily determined by the imposition of stereochemical restraints, the most commonly used being the inter-C_α separation of 3.8 Å. This also means that there is no absolute requirement for the values of y_L and y_R of a given point to be identical. For a ribbon diagram, however, several additional difficulties arise. Firstly it is necessary to ensure that in reading the original co-ordinates, (x_L, y_L) and (x_R, y_R), equivalent points are being selected from the two images. Since there are no explicit atomic positions represented on the figure this is most conveniently done by tracing a grid of horizontal guidelines (lines of constant y) across the stereo figure. The co-ordinates for equivalent points can then be read at the intersection of the guidelines with the ribbon. Implicit in this method is the assumption that the figure is perfectly orientated such that the y co-ordinate is a constant for equivalent points on the two images and for this purpose it is necessary that the figure be accurately scanned. It is convenient to use the N- and C-termini of the ribbon as reference points for orientating the figure when scanning.

The second major difficulty is in the determination of S_f and θ. Once again this arises as a consequence of not having specific atomic positions on the figure which can be used for defining stereochemical restraints. There are several possible solutions to this problem which depend on the additional information available. These include, for example, estimating their values from the pitch of an α-helix or the minimal separation between two or more adjacent β-strands. Alternatively, if the number of residues in each element of secondary structure is known (and such information is often provided in the article in question) then the unknown parameters can be obtained by measuring the length of one or more such elements in the reconstituted structure and relating it to the expected length in Å, employing standard values for the rise per residue along the axis of the secondary structural element. Since there are only two parameters to be determined this may readily be achieved by a grid search for the minimum in the difference between the determined distance and its target value on the two-dimensional

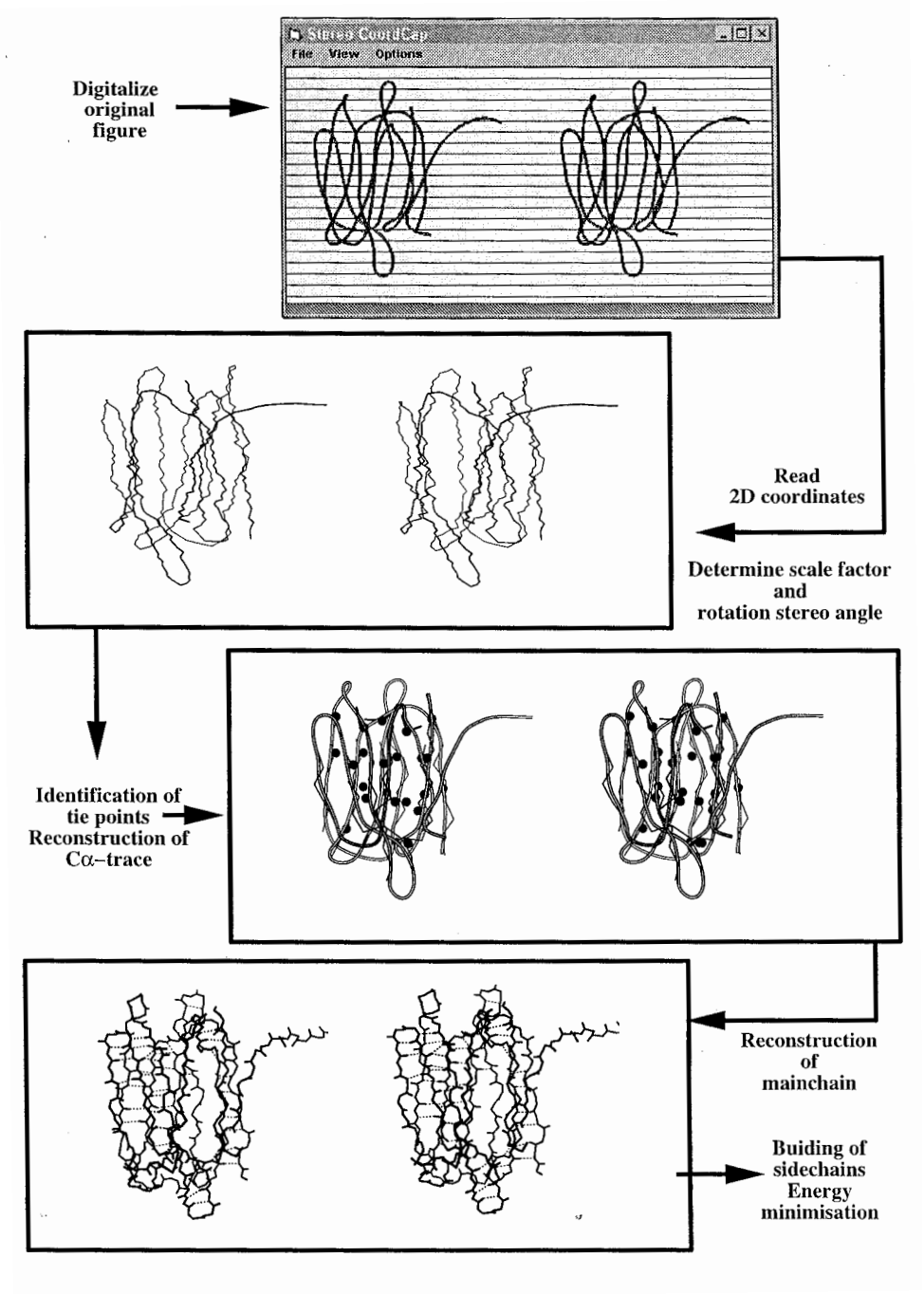

Figure 1. A general scheme for the reconstruction of the jacalin atomic co-ordinates from the original stereo image. Initially the figure is oriented with reference to the N- and C-termini and scanned at high resolution (first panel). Horizontal guidelines (in practice at higher density to those shown) are used to read 2D co-ordinates from the left and right-hand images. Recalculation of the 3D co-ordinates (x_F,y_F,z_F) is possible once the scale factor (S_f) and rotation angle (θ) have been established (second panel). The following stage involves the identification of tie points (shown as solid spheres) along the structure which serve as reference residue markers for the subsequent C_α chain tracing (third panel) from which the mainchain (fourth panel) and finally the sidechains may be built. The molecule is shown in the same orientation and in stereo in all parts of the figure with the exception of the second panel which has been rotated by 45° around the Y-axis in order to emphasise the oscillatory nature of the recalculated z co-ordinate (which is the least well determined) and to demonstrate that the figure is not simply a copy of the original.

surface defined by S_f and θ. However, in the case of the lectin jacalin the determination of S_f and θ was simplified by taking advantage of a second superimposed ribbon structure for which 3D co-ordinates were available.

The figure used for the greater part of the reconstruction of the jacalin co-ordinates appears as Figure 6a in the paper by Sankaranarayanan et al. (1996) [8]. It is a stereo ribbon figure of the jacalin monomer superposed on that of the second domain of *Bacillus thuringiensis* δ-endotoxin (residues 291–500) which presents a similar β-prism fold despite no detectable sequence similarity. The co-ordinates for the δ-endotoxin were available from the Brookhaven DataBank (PDB code 1dlc) [9].

Reconstruction of the jacalin structure (Figure 1)

(1) Initially the image was scanned at high resolution and co-ordinates read for 271 equivalent points along the δ-endotoxin ribbon from the two images using a grid of guidelines as described above and employing the program STERCAP written specifically for the purpose.

(2) Since the final x_F and y_F co-ordinates depend only on the scale factor S_f and not the rotation angle, the scale factor could be fixed by viewing the image of the reconstituted ribbon (calculated for an arbitrary value of θ) in the XY plane and comparing it to a manually superposed C_α-trace derived from the available 3D co-ordinates.

(3) Once the scale factor had been fixed it was possible to determine the rotation angle by viewing the reconstituted structure perpendicular to the Z axis and once again manually refining it by comparison with the known 3D co-ordinates for the δ-endotoxin.

(4) The jacalin 'monomer' consists of an α-chain of 133 residues and a β-chain of 20 residues of which only 15 were observed in the electron density map [8] (Figure 2). Pixel co-ordinates were read from the jacalin ribbon as described in stage 1 but this time at a higher resolution leading to a final co-ordinate set of 609 pairs of points. The values of S_f and θ determined in stages 2 and 3 were applied to these data resulting in a 3D co-ordinate set which described the overall trace of the jacalin mainchain (Figure 1).

At this stage the reconstituted ribbon was examined on the graphics workstation and the z co-ordinate of several of the points were manually adjusted by simultaneously viewing the original stereo figure. These points occur in regions of the structure where the ribbon runs approximately parallel to the guidelines. In such circumstances it is difficult to determine with precision the exactly equivalent point of intersection on the two images and this leads to large errors in the calculated z co-ordinate.

(5) The conversion of the reconstituted ribbon into genuine atomic co-ordinates started with the identification of several residues in the amino acid sequence which would serve as tie points. At this stage advantage was taken of a second stereo figure present in the original reference (Figure 4 of Sankaranarayanan et al. [8]) which shows 24 side chains forming the hydrophobic core of the β-prism. Since this figure provides explicit atomic positions, the program STEREO [5] was used for recalculating their co-ordinates.

The jacalin fold is based on a β-prism structure in which each of the three sides of the prism are formed by a β-sheet of four strands with the topology of a $(4,0)_C$ Greek key [10]. The only residues specifically indicated on this second stereo figure were three conserved isoleucine residues present at the N-terminus of the first strand of each Greek key (Ile12, Ile65, Ile113). These three residues could be used to roughly orientate the 24 sidechains within the core of the reconstituted ribbon structure. With the help of the amino acid sequence it was possible to unambiguously identify the remaining 21 residues. These served as further tie points along the ribbon which aided in maintaining the reconstituted structure in register with the amino acid sequence during the following stages.

(6) The mainchain secondary structure of the jacalin molecule was initially built in fragments. From the stereo ribbon diagram regions of the structure which were clearly similar in both jacalin and δ-endotoxin were identified. These included three β-bulges characteristic of the fold as well as β-strands 4-6, 9 and 10 of the jacalin α-chain (Figures 1 and 2). These were initially built by rigid body rotation and translation of the appropriate fragments taken from δ-endotoxin and modelled onto the reconstructed jacalin ribbon. The first β-bulge permitted the initial positioning of strands 1 and 2 which the bulge unites and a further stereo figure aided in positioning strand 7. The remainder of the β-strands were built by following the hydrogen bonding pattern of the previously established sheet and employing standard strand geometry whilst taking into account peculiarities of the ribbon (such as a fourth bulge in strands 5 and 6) and in the sequence (such as the prolines in strand 8). Once again the fragments being used were positioned with reference to the ribbon.

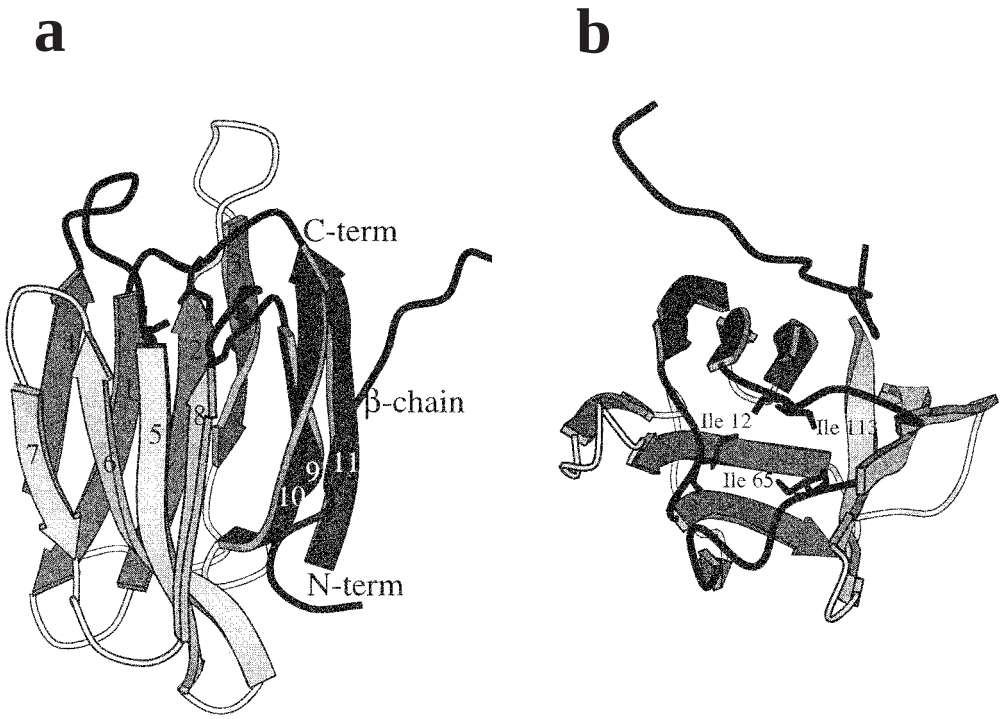

Figure 2. Ribbon figure showing two orthogonal views of the jacalin structure. (a) The β-prism, composed of three four-stranded β-sheets of Greek key topology, showing the conventional strand numbering [8]. The three sheets are distinguished by different tones of grey. (b) The β-arches which unite the three Greek keys are shown in bold. Each possesses a conserved isoleucine residue (indicated) followed by a classic type β-bulge.

(7) The most taxing and speculative part of the model reconstruction was the building of the loops. There are two reasons for this. Firstly due to their less regular structure and secondly due to the orientation of the figure which meant that the ribbon for the majority of the loops lay approximately parallel to the guidelines. Peptide flips may be particularly common in these regions. The three β-arches which cross the top of the prism from one Greek key to the next did not present serious ambiguities for the majority of their component residues due to the well defined isoleucines at positions 12, 65 and 113. The connections between strands 2 and 3 and between 5 and 6 were modelled as type I′ β-turns based on their sequences [11]. The connections between strands 1 and 2 and between strands 6 and 7 were modelled as type I turns based on the same argument. The loop between strands 9 and 10 was modelled in such a way that Gly121 donates a hydrogen bond to Trp123 and with residues 119, 125 and 126 forming a classic type β-bulge. The most difficult loops were those connecting strand 3 to 4 and 7 to 8. In these cases C_α guide positions separated by 3.8 Å and positioned by following the ribbon were used as the basis for loop searches using the algorithm of Jones and Thirup [12]. The final conformation was chosen according to stereochemical criteria and its compatibility with the amino acid sequence, particularly in terms of hydrophobicities.

(8) For the sake of completeness an attempt was made to build sidechains onto the entire model in accordance with the amino acid sequence. The sidechains which appeared in stereo figures in the original article were oriented accordingly but the vast majority had to be model built with standard rotamers using the graphics program O [13]. The final model was subject to steepest descent energy minimisation until there was no further improvement in the stereochemical and residue packing parameters.

Results and discussion

The values determined for θ and S_f were 5°and 17 pixels/Å respectively. The resulting reconstructed model

presented acceptable stereochemistry as assessed by the program PROCHECK [14]. No residue lies outside of the most favoured and additional allowed regions of the Ramachandran plot and the model presents an overall G-factor equivalent to that of a crystal structure of approximately 1.7 Å resolution. The model is also apparently respectable in terms of residue environment as evaluated by the program VERIFY_3D [15] which gave a total score of 55, only slightly inferior to the expected value of 67 for proteins of the same number of residues and well within the range of acceptability. This suggests that the overall chemical environment of the component residues as a whole have been well reproduced. The structure fared much worse in terms of atomic contacts [16] as assessed by the program WHATIF [17] where the overall Z-score was −4.18, corresponding to the equivalent of a poorly refined experimental structure or low-homology model. This is clearly due to the obvious difficulty in modelling the sidechains and is a consequence of the sensitivity of the method to details of the structure in terms of badly placed atomic co-ordinates. The corresponding Z-score for the mainchain contacts of the rebuilt structure was much more acceptable at −0.53.

Subsequent to our reconstruction of the jacalin co-ordinates the crystal structure of Sankaranarayanan et al. was released by the Brookhaven DataBank (PDB code 1jac). As anticipated both the VERIFY_3D residue environment analysis and WHATIF quality Z-scores for the crystal structure are superior to those of the reconstituted co-ordinates, being 68 and −0.31 respectively (compared to 55 and −4.0). Once again it is apparent that the principal difficulties of the model reconstruction lie in the accurate building of the sidechains and not in their overall positioning in the context of the rest of the structure. This is clear from the fact that the atomic contact WHATIF quality for the reconstituted model differs greatly from that of the crystal structure whilst the residue environment score (VERIFY_3D) differs only marginally.

Comparison of the rebuilt model with the crystal structure by least squares superposition of the C_α co-ordinates of the α-chain results in an RMS deviation of 1.5 Å. When the superposition is repeated to include the atoms of the extended mainchain (C_α, C_β, C, O and N) this increases marginally to 1.6 Å and when all atoms are included rises to 2.3 Å. Due to several inconsistencies between the deposited co-ordinates and the sequence as it appears in the original paper [8] seven residues had to be reduced to glycines prior to the all-atom superposition.

Once again these results clearly demonstrate that the major difficulty in the reconstruction of the α-chain, which is dominated by the Greek key β-structure, was in the modelling of the side chains. This conclusion is supported by the relatively low mean errors in the mainchain torsion angles, $\Delta\phi_\alpha$ and $\Delta\iota_\alpha$, of 31° and 33° respectively. As expected, if the comparison is limited to just the secondary structure a marginal improvement is observed in the RMS deviations; 1.4 Å, 1.5 Å and 2.2 Å for the C_α's, extended mainchain and all atoms respectively. Since the secondary structures were largely orientated perpendicular to the guidelines whilst the loops were not, a comparison of the above RMS deviations gives some feeling for the dependence of the coordinate accuracy on the ribbon orientation.

The above values are comparable to those obtained by Chung and Subbiah [18] for homology modelling purposes using a sidechain packing algorithm based on previously established mainchain coordinates. They demonstrate that for an RMS deviation of 1.6–1.8 Å in the backbone atoms, the expected RMS error for the core sidechains is of the order of 2.2 Å. These are similar to the values of 1.6 for the backbone and 2.3 Å for the sidechains quoted above for our reconstructed α-chain, where the latter refers to the RMS deviation for all sidechains and not only those for the core, which are normally the easiest to correctly predict. This suggests that any improvement that may have been possible by adopting a more sophisticated strategy for sidechain modelling is probably marginal.

The jacalin β-chain is 20 residues long and is separated from the α-chain as a result of an internal cleavage of a single polypeptide precursor. The crystal structure includes co-ordinates for 15 of the 20 residues and when compared to the reconstructed model they present an RMS deviation of 2.6 Å on C_α's, 3.1 Å for the extended mainchain and 4.9 Å for all atom co-ordinates. This considerably poorer performance compared with the α-chain is a consequence of its more irregular structure and is clear from Figure 3 which shows a superposition of the reconstituted model and the crystal structure.

A previous study of two proteins of similar size to that of jacalin but employing the program EXTRACT [6] which utilises explicit atomic positions for reconstruction, produced an average RMS deviation of 0.68 Å for C_α coordinates. The result obtained here of 1.5 Å does not compare unfavourably with that obtained by EXTRACT allowing for the limited information present within the ribbon diagram. Perhaps the

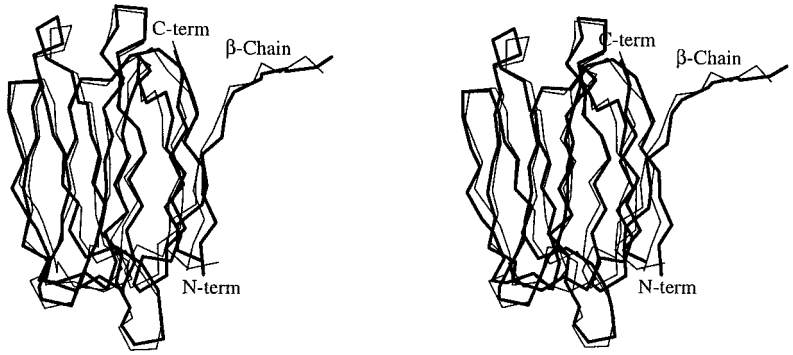

Figure 3. Superposed C_α-trace of the crystal structure of jacalin and the reconstructed model (bold), calculated on the basis of the C_α co-ordinates of the α-chain.

most impressive aspect of Figure 3 is the fact that at no point along the 133 residues of the α-chain does the predicted structure become out of register with the crystal structure, itself a critical factor in determining the respectable RMS deviation obtained. This is not so for the smaller β-chain leading to the larger deviations cited above.

Figure 4 shows a hydrogen bonding diagram for both the rebuilt and crystal structures produced with the program HERA [19]. The H-bonding patterns are in general very similar with a large percentage of the secondary structure (H-bonding partners) identical in the two figures. Of note are the three aforementioned β-bulges at the beginning of each Greek key, whose position and nature were correctly predicted and which are shown in bold in the figure. The two type I′ β-turns (between strands 2 and 3 and between 5 and 6) were correctly predicted as was the type I turn between strands 6 and 7. However, there remain some notable differences between the two figures. In the first Greek key the β-chain is missing from the reconstituted model due to its irregular structure. Several hydrogen bonds do exist between the β-chain and strand 10 but these are largely incorrect. β-strand 7 is reduced in size due to difficulties in correctly modelling a wide type β-bulge (residues 94 and 95). Strand 8 is apparently missing altogether despite the presence of several correctly modelled hydrogen bonds indicated on the figure. Its apparent absence is due to a rather irregular geometry in the reconstructed model which resulted in disorientation of several H-bonding partners.

One of the more fortunate aspects of the ribbon figure employed in the present work was the superposed structure of the second domain of δ-endotoxin for which atomic co-ordinates were available. This information was principally of value in the determination of the parameters S_f and θ. It is instructive to ask if it would have been possible to determine the unknown parameters directly from the jacalin ribbon without the use of the δ-endotoxin co-ordinates. This is expected to be the far more common situation and a solution to the problem would generalise our approach considerably.

In order to address this question we have redetermined S_f and θ using an alternative procedure which depends on fixing the minimal separation between pairs of adjacent β-strands. Since most β-sheets present a left-handed twist (when viewed perpendicular to the strands) the inter-strand separation varies along their length and is a minimum at the centre where they cross over. A mean value for this minimal separation was determined from 11 very high-resolution β-sheet containing protein structures [20] to be 4.74 Å. This is in very good agreement with both experimental and calculated values for the flat β-sheet of silk fibroin [21]. It was taken to be the target separation between the recalculated co-ordinates of points at the centre of adjacent β-strands. These points were selected by initially viewing the structure calculated using the typical value of 6° for θ and an arbitrary S_f. At this stage the exact value of S_f is irrelevant as the model is merely used for determining the cross-over points on the β-strands.

Several trials were performed in which either 3, 5 or 7 pairs of points (one from each strand) were used for each set of adjacent strands. Frequently edge strands are more distorted than central strands and we therefore limited the choice of points to the pair of central strands in each Greek key. θ was systematically

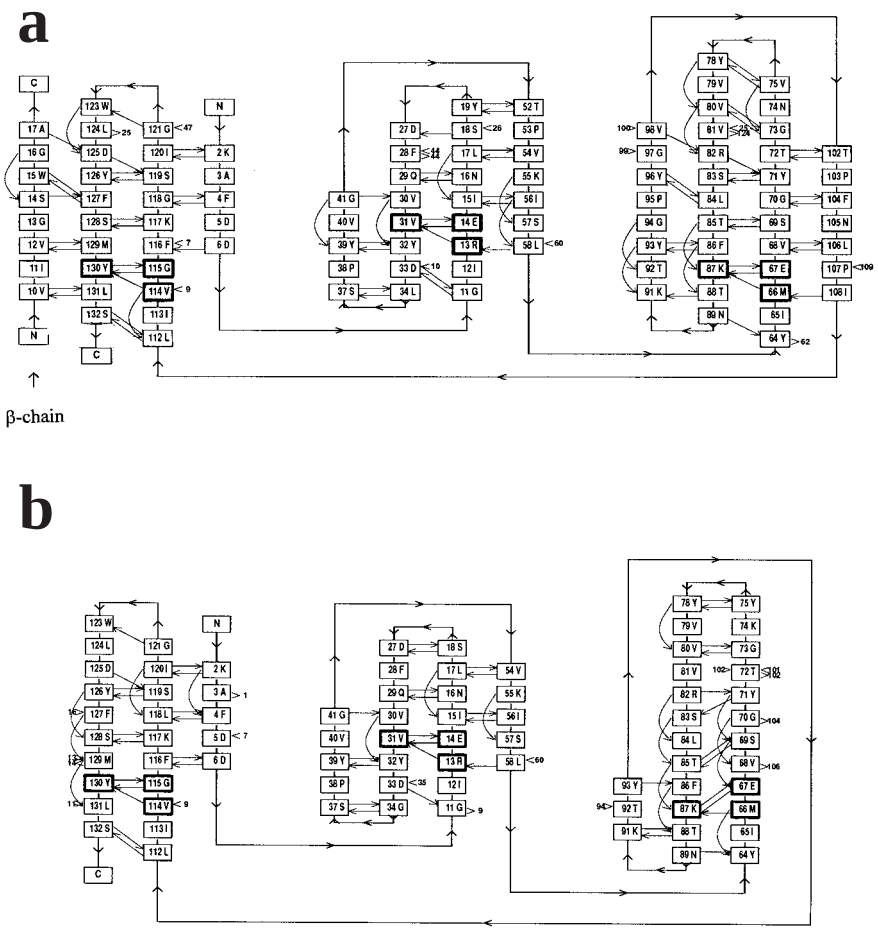

Figure 4. Mainchain H-bonding diagram for the jacalin crystal structure (a) and rebuilt model (b) produced with the program HERA. The three β-bulges commented on in the text and used as the initial tie points for retracing the protein backbone are indicated by the bold boxes. The overall pattern of hydrogen bonds is very similar in both cases.

varied from 3–15° in steps of 0.25° and S_f from 1 to 30 pixels/Å in steps of 0.5 pixels/Å and the RMS difference between the interstrand separation and the target value of 4.74 Å calculated at each point. When 3 interstrand distances were used for each Greek key (a total of 9) a shallow minimum was observed at $\theta = 5.0°$ and $S_f = 17$ pixels/Å, identical values to those determined in the original reconstruction. If the number of distance measurements is extended from the cross-over point to include 5 or 7 per Greek key (15 or 21 in total) two adjacent minima are observed slightly shifted to $\theta = 4.5°$, $S_f = 17.0$ pixels/Å and $\theta = 4.25°$, $S_f = 17.5$ pixels/Å respectively. Recalculation of the 3D co-ordinates based on these altered parameters results in an RMS deviations of 0.84 and 1.1 Å when compared to the original values of 5° and 17 pixels/Å. These results suggest that it would have been feasible to produce a sufficiently accurate 3D reconstruction of the ribbon even in the absence of the additional information provided by the superposed structure of the second domain of δ-endotoxin.

In the above procedure it was necessary to use the minimal inter-strand separation as the criterion for the determination of θ and S_f as a consequence of the fact that jacalin does not possess a single α-helix. For α-helical proteins the situation may be rather more simple as the pitch of the helix may be used for the same purpose. This is advantageous as it does not re-

quire that any special point on the helix (analogous to the point of closest approach of two β-strands) be previously identified.

Conclusions

Our initial objective in attempting to reconstruct the jacalin co-ordinates was to employ them as a search model for the solution of a related lectin structure (lectin KM+) [22] by molecular replacement. Thus far we have been unsuccessful in so doing but it is still unclear if this is a consequence of the co-ordinate inaccuracy or not since it has not been possible to solve the km+ structure using the crystallographic co-ordinates either. Be that as it may, the reconstruction procedure has been to a large extent a success as judged by the RMS deviations obtained which serve to demonstrate that at least in some cases sufficient information for such purposes is present within a stereo ribbon figure together with a sufficient number of side-chain positions to serve as tie points.

It could be quite justifiably argued that the procedure we have described is not entirely general and thus may not be applicable in the exact same form in other cases. In the reconstruction of the jacalin co-ordinates we used all of the information available to us and present in the original reference. It is highly unlikely that no information other than the ribbon figure itself would be available in other comparable cases. One of the principal difficulties to be overcome will always be the determination of the scale factor and rotation angle. Here we have demonstrated that two different approaches yielded very similar results. The first approach is specific to the case at hand as it required the use of a second stereo figure of a structure for which co-ordinates were available. The second method is more general in nature as it can be applied directly to the ribbon of interest by imposing limitations on the possible values of θ and S_f via stereochemical restraints, namely standard parameters which describe elements of secondary structure.

If the figure has no specific labels on the N- and C-termini it is possible that the chain could be traced backwards. However, it is unlikely that in any real situation insufficient information would be present in the text/figure to resolve this potential ambiguity. A further factor worthy of mention is the exact form of the ribbon representation itself since several alternatives are commonly used in the literature. In principle the method can be applied to all such representations (be they tapeworms, arrows, coils etc.) but clearly the closer the approximation of the ribbon to the genuine course of the mainchain, the better. One parameter which is probably critical is the ribbon width since this will clearly affect the precision in reading the original 2D coordinates which are taken from the centre of the ribbon.

One final factor which will probably determine the accuracy of the method in other cases is the orientation of the figure. In the case of jacalin, a structure dominated by antiparallel β-sheet, the ease in reading the 2D co-ordinates was greatly facilitated by having the majority of the β-strands aligned approximately parallel to the Y axis. This means that the guidelines cut the ribbon in these regions at approximately right angles. This can not be guaranteed to be the case for other examples and, as described above, it is the regions for which the ribbon lies parallel to the X axis that are subject to the greatest error in the determination of the z co-ordinate. It is therefore true that the accuracy of the results will probably vary from case to case but at least for one example we have demonstrated the viability of the approach. The usefulness and limitations of the method will undoubtedly become more apparent as it is applied to other examples including proteins of different structural classes, different sizes, less favourable orientations and different types of ribbon representation. We hope that our experience reported here will serve to stimulate others.

Acknowledgements

We would like to thank the following Brazilian funding agencies; FAPESP, CNPq, PADCT, FINEP for providing the financial support for the computational facilities used in this work. The program STERCAP for reading the 2D coordinates from the ribbon figure is available on request from the authors.

References

1. Wodak, S.J., Nat. Struct. Biol., 3 (1996) 575.
2. Navaza, J., Acta Crystallogr., A50 (1994) 157.
3. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F.Jr., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M., J. Mol. Biol., 112 (1977) 535.
4. Lemer, C. M.R., Rooman, M.J. and Wodak, S.J., Proteins, 23 (1995) 337.
5. CCP4 Suite: SERC Daresbury Laboratory, Warrington, UK, Acta Crystallogr., D50 (1994) 760.
6. Oldfield, T.J. and Hubbard, R.E., J. Mol. Graph., 13 (1995) 18.

7. Hooft, R.W.W. and Vriend, G., J. Mol. Graph., 14 (1996) 168.
8. Sankaranarayanan, R., Sekar, K., Banerjee, R., Sharma, V., Surolia, A. and Vijayan, M., Nat. Struct. Biol., 3 (1996) 596.
9. Li, J., Carroll, J. and Ellar, D.J., Nature, 353 (1991) 815.
10. Hutchinson, E.G. and Thornton, J.M., Protein Eng., 6 (1993) 233.
11. Sibanda, B.L., Blundell, T.L. and Thornton, J.M., J. Mol. Biol., 206 (1989) 759.
12. Jones, T.A. and Thirup, S., EMBO J., 5 (1986) 819.
13. Jones, T.A., Zou, J.-Y., Cowan, S.W. and Kjeldgaard, M., Acta Crystallogr., A47 (1991) 110.
14. Laskowski, R.A., MacArthur, M.W. and Thornton, J.M., J. Appl. Crystallogr., 26 (1993) 283.
15. Lüthy, R., Bowie, J.U. and Eisenberg, D., Nature, 356 (1992) 83.
16. Vriend, G. and Sander, C., J. Appl. Crystallogr., 26 (1993) 47.
17. Vriend, G., J. Mol. Graphics, 8 (1990) 52.
18. Chung, S.Y. and Subbiah, S., Protein Sci., 4 (1995) 2300.
19. Hutchinson, E.G. and Thornton, J.M., Proteins, 8 (1990) 203.
20. Fabiola, G.F., Krishnaswamy, S., Nagarajan, V. and Pattabhi, V., Acta Crystallogr., D53 (1997) 316.
21. Fossey, S.A., Némethy, G., Gibson, K.D. and Scheraga, H.A., Biopolymers, 31 (1991) 1529.
22. Oliveira, P.S.L., Garratt, R.C., Mascarenhas, Y.P., Beltramini, L.M., Roque-Barreira, M.C., Harvey, I. and Oliva, G., Proteins, 27 (1997) 157.