Journal of Computer-Aided Molecular Design, 10 (1996) 558–566
ESCOM

J-CAMD 363

Reconstruction of the 3D coordinates of α-carbon atoms of proteins from a pair of stereographic figures

Yoriko Iwata, Atsushi Kasuya and Shuichi Miyamoto*

Exploratory Chemistry Research Laboratories, Sankyo Co., Ltd., 1-2-58, Hiromachi, Shinagawa-ku, Tokyo 140, Japan

Received 1 December 1995
Accepted 6 July 1996

Keywords: Molecular modeling; Stereographic figure; C^{α} trace; Root-mean-square deviation

Summary

In an attempt to promptly use the experimentally determined structures of proteins in modeling studies, we have developed the program ReconstCα to generate the 3D coordinates of α-carbon atoms from a pair of stereographic figures. Calculations of the 3D coordinates were performed by estimating the stereo parameters systematically. Geometrical features of C^{α} traces were used to evaluate the integrity of the calculated structure. The program was applied to four kinds of protein structures to examine the performance. It was found that the root-mean-square deviation of atomic positions between constructed and reference crystal structures ranged from 0.36 to 0.78 Å. The range represents a reasonable accuracy and its automatic feature suggests that our approach would be expedient for providing initial structures for protein modeling studies.

Introduction

In line with the current pace of X-ray crystallographic and NMR spectroscopic technologies, the number of research articles pertaining to the three-dimensional (3D) structural studies of proteins has been growing in an exponential-like fashion in the past few years. One of the frustrations in reading these articles is that in most cases the coordinates are not available for modeling studies when they are published. Much of the crystallographic information for proteins is compiled in the Brookhaven Protein DataBank (PDB) [1] and this databank has proved valuable to a wide variety of investigations on protein structure and function. However, more often than not entry to this database is either restricted or time-consuming. Occasionally, the coordinates are unavailable for commercial reasons or because the investigators wish to exploit the data for further structural studies.

In the fierce competition amongst pharmaceutical companies in pursuing drug development, it is of great importance to obtain timely 3D structures of proteins that would serve as models for drug targets. In our work on human immunodeficiency virus 1 (HIV-1) protease inhibitors, it was indeed desirable to build a 3D model of the protease based on the first X-ray crystallographic analyses appearing in the literature [2,3]. Such literature gave much impetus to create an automatic and accurate modeling algorithm which could generate the 3D coordinates of α-carbon atoms from stereo diagrams of proteins. Furthermore, this reconstruction of the C^{α} atoms was considered imperative for building the backbone of proteins.

A computer program for extracting 3D coordinates from C^{α} stereo figures was produced by Michael Rossmann many years ago. The program STEREO with assembler codes is distributed as an unsupported program from the PDB [1]. The detailed algorithm and application results are still obscure at this stage. We, therefore, have independently developed a program, ReconstCα, to reconstruct accurate 3D structures from published stereographic α-carbon traces of protein structures. The first step in our approach is to measure the 2D atomic coordinates in the figures corresponding to the left and right eye views of the molecular system. The program then automatically calculates the 3D coordinates of the C^{α} atoms by systematically estimating the stereo parameters.

*To whom correspondence should be addressed.

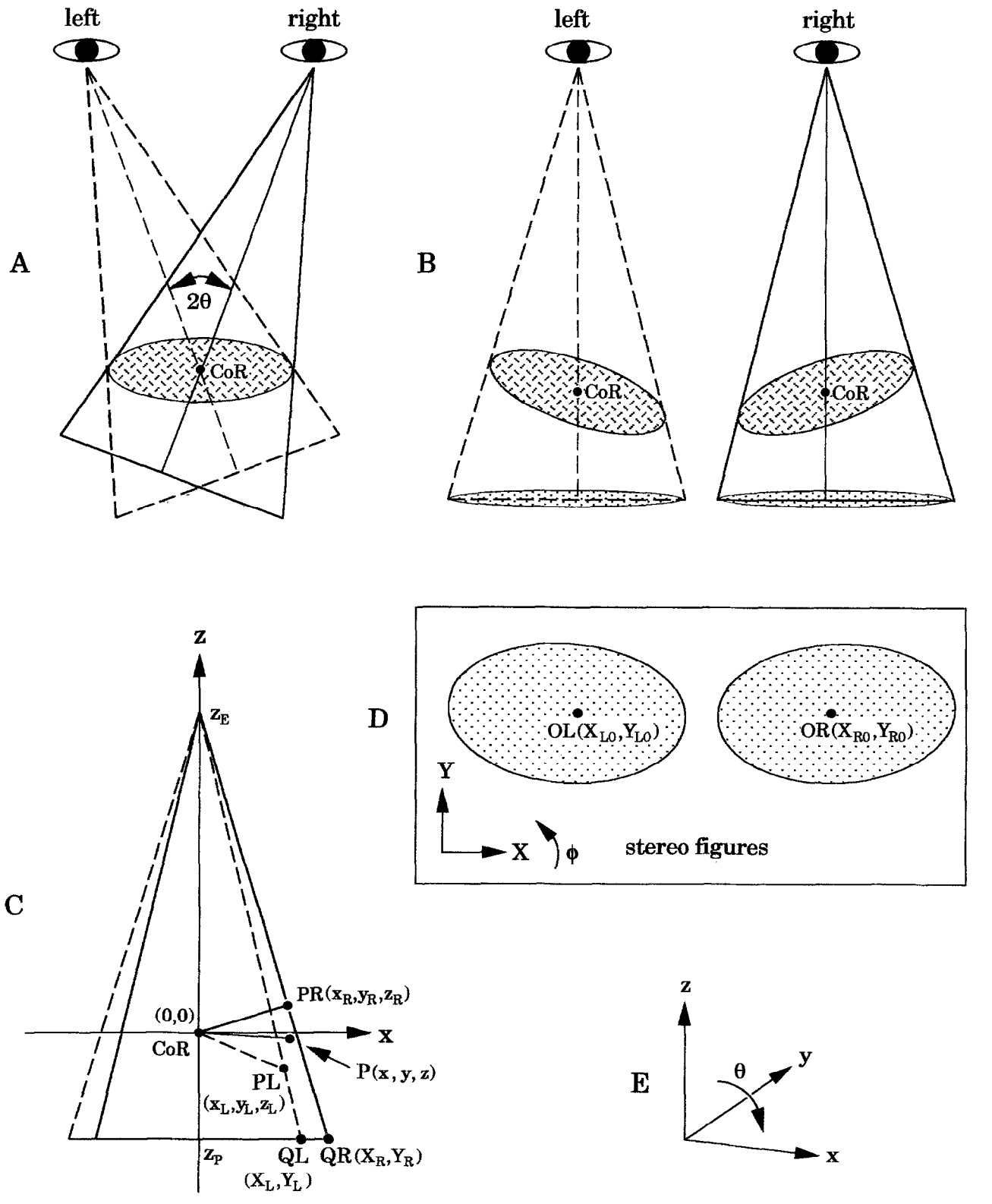

Fig. 1. Schematic drawing of the geometric basis of the stereographic figure generation. (A) The left and right eyes are looking into a molecular model. The stereo angle is twice that made by two viewing directions. (B) The left and right eyes are looking at the stereo figures to get a 3D molecular image. (C) The relationship between a certain point P in the model, its rotated points PL and PR and their projection points QL and QR is shown. When the model is rotated for the left and right eye views by θ, P(x,y,z) is translated to PL(x_L,y_L,z_L) and PR(x_R,y_R,z_R), respectively. (D) The 2D stereo figures corresponding to (B). OL(X_{L0},Y_{L0}) and OR(X_{R0}, Y_{R0}=Y_{L0}) are the projected points of CoR (origin). Since it was difficult to define the exact orientation of the 2D coordinate system, another parameter designated as the minute rotation angle ϕ was employed in adjusting the orientation. (E) A positive value of θ corresponds to the model rotation in the designated direction.

In this process, geometrical features of protein structures are used to evaluate the integrity of the calculated model.

Very recently, Oldfield and Hubbard [4] have presented a method for extracting the 3D coordinates from α-carbon stereo diagrams. The first stage in their procedure is to digitize the stereo figures with an image scanner and to display the image on a stereo-equipped workstation. The scale and stereo angle are then adjusted manually so that a template piece of canonical β-sheet or α-helix appears to fit well with an appropriate section of the image. The main stage involves building C^{α} atoms individually, one at a time, on the workstation using the diagram as a guide. If the new C^{α} position cannot be constructed satisfactorily, it is necessary to start again from the adjustment of the scale and stereo angle.

In this article, we present the details of an algorithm to generate 3D coordinates based on the published stereo α-carbon traces of protein structures. The results obtained by applying the program to four protein structures are also shown and these are compared with those provided by Oldfield and Hubbard [4].

Methods

As shown in Fig. 1, stereo figures of molecular models are produced by placing a left eye view and a right eye view (Fig. 1A) on the paper plane separately (Fig. 1B). In practice, these projections are generated by rotating the model itself as illustrated in Fig. 1C. In this process, the center of rotation, CoR, the rotation angle (stereo angle), θ, the distance from the view point to the rotation center, z_E, and the distance from the rotation center to the projection plane, $-z_P$, become variable parameters. As revealed in the Appendix, 3D coordinates of model atoms can be calculated if the corresponding 2D coordinates in the figures and the above stereo parameters are known. Since those parameters are not shown in the literature, we employed the method to calculate the 3D coordinates by a systematic derivation of those values. The problem we needed to address was how to evaluate the reliability of the estimated parameters. We noticed that the distance between two contiguous α-carbons is constant in protein structures. This implies that the α-carbon distances should be constant if the estimated stereo parameters are reasonable. An assumption was made that the lower the distance deviation, the greater the accuracy of the parameters and therefore of the atomic positions. In other words, the standard deviation (SD) of inter-C^{α} distances was used to assess the integrity of the model and, indirectly, the estimated parameters.

Presumably, the errors in the calculated coordinates are much larger in the Z direction perpendicular to the paper plane compared to those in the X and Y directions. In the calculation, therefore, simple local corrections in the Z coordinates were carried out as shown in Fig. 2. After computing the average α-carbon distances, the Z coordinates of each pair of the adjacent α-carbons were corrected according to the distance deviations from the average.

From our analysis of C^{α} geometry in high-resolution protein structures, it was found that the average distance between two neighboring α-carbons is 3.8 Å for the trans-configuration and the virtual C^{α}-C^{α}-C^{α} bond angles (VBA) range from 70° to 157°. Violations of these angle values were also used as a guide to assess the modeled structure. In summary, the model that gave the minimum score was selected as the optimum one, and the score employed as a criterion for the model structure can be defined as

$$\text{score} = \text{SD} \times \text{Coeff} + f_p(\text{VBA})$$

where SD is the standard deviation of inter-C^{α} distances in Å after scaling the coordinates such that the average distance coincides with 3.8 Å. f_p(VBA) is a penalty function for the VBA violations and it is simply the sum of violations of each VBA. Values of 1.01, 2.11 and 3.21 are assigned to f_p(VBA) when $70° > \text{VBA} \geq 60°$ or $157° < \text{VBA} \leq 167°$, $60° > \text{VBA} \geq 50°$ or $167° < \text{VBA} \leq 167°$, and $\text{VBA} < 50°$ or $\text{VBA} > 177°$, respectively. Coeff is a coefficient to balance the two kinds of errors and it has a tentative value of 20.

A pair of stereo figures displayed in the literature was magnified roughly 1.7 times using a typical photocopying machine. After placing a piece of graph paper underneath the figures, the 2D coordinates of both figures were measured through marking of the atom positions, by pushing a pin through the top sheet into the graph paper below. Alternatively, a pair of stereo diagrams could be magnified on to a piece of graph paper directly, which would

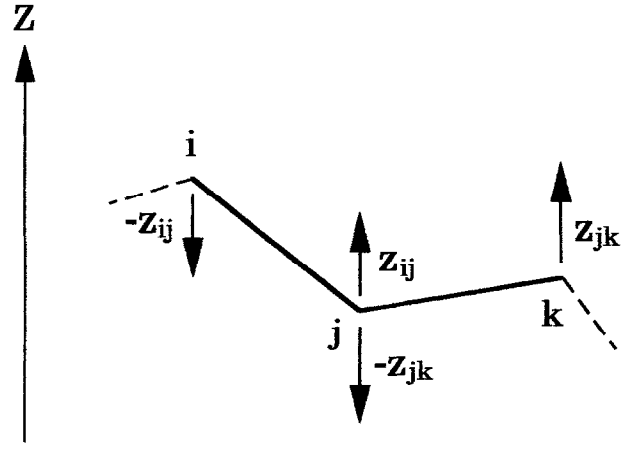

Fig. 2. Schematic drawing for local corrections of the Z coordinates. For each pair of two adjacent α-carbons, two correction vectors (z_{ij},$-z_{ij}$,...) with equal magnitudes and opposite orientations along the Z-axis are calculated to make the distance equal to the average. The net vector for each atom ($z_{ij} - z_{jk}$ for atom j, for example) is added to its position.

TABLE 1
CHARACTERISTICS OF PROTEIN STRUCTURES USED IN THIS STUDY[a]

Code name[b]	Protein name	Number of residues	Resolution[c] (Å)	SD of C^{α} distances[d] (Å)	Reference
1SN3	Scorpion neurotoxin	65	1.8	0.07	5
1CTX	α-Cobratoxin	71	2.8	0.09	6
3HVP	HIV-1 protease	99	2.8	0.06	2
2APP	Penicillopepsin	323[e]	1.8	0.03	7

[a] Data are taken from the PDB.
[b] Code name in the PDB.
[c] Resolution in X-ray crystallographic analysis.
[d] Standard deviation of the distances between two contiguous C^{α} atoms.
[e] Two hundred out of 323 residues were used in the calculation.

probably make the measurement of the coordinates easier. When defining the coordinate system, the original stereo parameter CoR was set to the origin (Fig. 1C) as described in the Appendix. In place of CoR, therefore, its projected points, $OL(X_{L0},Y_{L0})$ and $OR(X_{R0}, Y_{R0} = Y_{L0})$, become stereo parameters to be estimated (Fig. 1D). These two points were strongly anticipated to be located near the respective figure centers, and so their initial positions were taken from the geometrical center of each diagram. As it was difficult to define the exact orientation of the 2D coordinate system, another parameter designated as the minute rotation angle ϕ was employed in adjusting the orientation. A detailed derivation of the 3D coordinate calculations based on the measured XY coordinates and stereo parameters is given in the Appendix.

An estimation of the stereo parameters was performed as follows. The rotation angle ϕ was changed from $-2°$ to $2°$ with an increment of $1°$. The stereo angle θ was varied five times around $3°$ with a user-specified increment (with a default value of $0.25°$). In order to reduce the computation time, the other stereo parameters were searched in a stepwise fashion. Initially, such parameters were scanned five times with larger increments and those giving the minimum score were identified. In subsequent steps, they were reexamined five times using a smaller grid confined to the best values obtained in the previous cycle. This procedure is repeated to optimize the stereo parameters. As for the initial approximations of these parameters, the program supplies default values based on the figure size.

Most backbone peptides in proteins are of a trans-configuration; however, those between a proline and the preceding residue bear occasionally cis-configurations. If the positions of these *cis*-peptides are described in the literature, they can be specified as such in the program input. This avoids the introduction of extra errors when calculating the SD of α-carbon distances.

Results and Discussion

On the basis of the methods described above, the FORTRAN program ReconstCα has been developed. In order to examine the performance of the program, we applied it to rebuild the α-carbon coordinates of four proteins, namely, scorpion neurotoxin [5], α-cobratoxin [6], HIV-1 protease [2] and penicillopepsin [7]. These proteins consist of 65, 71, 99 and 323 residues and their code names in the PDB are 1SN3, 1CTX, 3HVP and 2APP, respectively. The structures of 1SN3 and 2APP are solved with a high resolution of 1.8 Å, while the crystallographic resolution of the remaining ones is 2.8 Å. Every molecule contains some β-strands. One or more α-helix structures are included in the proteins except for 1CTX. The secondary structure contents are not very high, except for 2APP. These structural features are summarized in Table 1. The table also provides a list of the SDs of α-carbon distances, since this is used as a guide in evaluating the model.

The C^{α} trace stereo diagrams of 1SN3, 1CTX, 3HVP and 2APP were taken from Figs. 8a, 2, 6 and 3b contained in Refs. 5, 6, 2 and 7, respectively. After enlargement (160–180%) of a pair of the published stereo figures by a conventional photocopying machine, the 2D coordinates of the α-carbons were measured in submillimeter resolution with a graph paper placed underneath the figures. The XY orientation was initially adjusted when placing the graph paper under the figures because the adjustable rotation angle ϕ was incorporated into the program afterwards. We therefore did not simply magnify the diagram directly on to a graph paper this time, because this direct enlargement usually needed some adjustment of the XY orientation.

A couple of the measured 2D coordinate sets for the left and right figures were supplied to the program ReconstCα to calculate the 3D coordinates of the α-carbon atoms. For practical purposes, it was found that three cycles of parameter estimations were sufficient. A typical run on 3HVP took 12 min on a Silicon Graphics Indy/R4000 workstation. The CPU cost linearly increases with the size of the protein.

The accuracy of the calculated atomic coordinates can be derived by comparison with the respective X-ray structure. Since the coordinate system of stereo diagrams is different from that of the corresponding PDB file, the deviations of atom positions were calculated using coordinate superposition. Thus, the derived root-mean-square deviation (rmsd) of atoms was considered to be equivalent to the errors incurred by the constructed model. The distance SD was also considered to reflect the accuracy,

although in this case it was secondary. These data are compiled in column A of Table 2. Figure 3 illustrates the results obtained using the superpositions of the calculated and the reference crystal structures. The viewing directions of half the figures are the same as those of the published figures. The other diagrams are obtained by rotating the models by 90° along the Y-axis.

1SN3 The literature has assigned a *cis*-proline to 1SN3 as the 59th residue [5], which was thus specified in the input. As can be seen from the data in Table 2, the rmsd between the α-carbons in the constructed model and those in the crystal structure is 0.37 Å and the SD of α-carbon distances is 0.11 Å, both being the smallest among the four test cases. No violation of the VBA is observed. Figures 3A and B show how well the calculated model reproduces the X-ray structure, inferring that the algorithm and the scoring scheme have worked properly. The small errors appearing among the test cases presumably arise from the small number of protein residues used in composing the stereo figures.

1CTX The second-best accuracy was obtained in the reconstruction of 1CTX. Although no records of *cis*-peptides for α-cobratoxin were found in Ref. 5 and the PDB file, there seemed to be three *cis*-prolines in the protein based on our structural analysis. Table 2 includes both results with and without taking the *cis*-peptides into account. Contrary to our expectation, similar rmsd values of 0.58 and 0.60 Å with one VBA violation were attained for the two cases. This can be rationalized on the basis of the rmsd of ~0.59 Å being not so small compared to the difference (ca. 0.9 Å) of inter-C^{α} distances between cis- and trans-configurations.

3HVP The HIV-1 protease exists as a stable dimer with C_2 symmetry and the coordinates of a monomer were compiled in the PDB. The calculation was then performed for the monomer while the dimer structure was displayed in the literature [2]. The C^{α} distance SD is rather small (0.14 Å) with no VBA violation, but the atomic rmsd is 0.62 Å, which is larger than that of 1CTX. As can be seen in Fig. 3E, errors in the XY direction are small. On the other hand, the Z coordinate errors are large, indicating that the modeled protein molecule contracts slightly in the Z direction compared to the native one (Fig. 3F). The deformation is relatively small at the active site which is located at the dimer interface. In the real modeling study, several atomic coordinates of the second monomer were measured and reconstructed to produce the whole dimer structure.

2APP In some cases only a portion of the protein structure, such as the ligand binding site, is necessary to be subjected to modeling. As for penicillopepsin, the C^{α} positions of 200 residues (three polypeptide chains), which correspond to two-thirds of the whole protein structure, were generated. In the calculation run, three C^{α} chains were handled simultaneously. The quality of the reconstructed coordinates is less satisfactory than that obtained from other operations. Larger errors of the calculated positions are not due to errors in the experimental ones because crystallographic resolution is high, with a small rmsd of 0.03 Å for the inter-C^{α} distances. 2APP has more residues than the other proteins tested and so it is reasonable to suggest that the resolution of the stereo figures in the literature is relatively low and thus may be responsible for the lower accuracy of the modeled structure.

When compared to the reference structure, XY coordinates are calculated larger or smaller in the region with large positive or negative XY values, respectively, as shown in Fig. 3G. The Z coordinates of the rebuilt model are smaller or larger than the original ones where these values are highly positive or negative (Fig. 3H). In other words, the modeled protein molecule is pressed in the Z direction and swells in the XY direction by a small amount. This may give rise to the larger rmsd of 0.90 Å with respect to the C^{α} positions.

The above results indicate that the program has a tendency to generate a 'shrunken' model that is distorted

TABLE 2
ACCURACY OF THE COORDINATES CALCULATED BY THE PROGRAM ReconstCα

Code name[a]	A[b]		B[c]		C[d]	
	SD of C^{α} distance (Å)	Rmsd of C^{α} position[e] (Å)	SD of C^{α} distance (Å)	Rmsd of C^{α} position (Å)	SD of C^{α} distance (Å)	Rmsd of C^{α} position (Å)
1SN3	0.11	0.37	0.11	0.37	0.01	0.36
1CTX[f]	0.18	0.58	0.18	0.57	0.01	0.56
	0.16	0.60	0.15	0.57	0.02	0.54
3HVP	0.14	0.62	0.14	0.62	0.01	0.61
2APP	0.16	0.90	0.16	0.80	0.01	0.78

[a] Code name in the PDB.
[b] Local adjustments of the Z coordinates have been carried out.
[c] Global expansion of the model in the Z direction has been taken into account.
[d] Final refinement of C^{α} geometry has been performed.
[e] Rmsd of C^{α} positions calculated versus reference PDB structures.
[f] The upper and lower rows correspond to the results obtained with and without taking the *cis*-peptides into account, respectively.

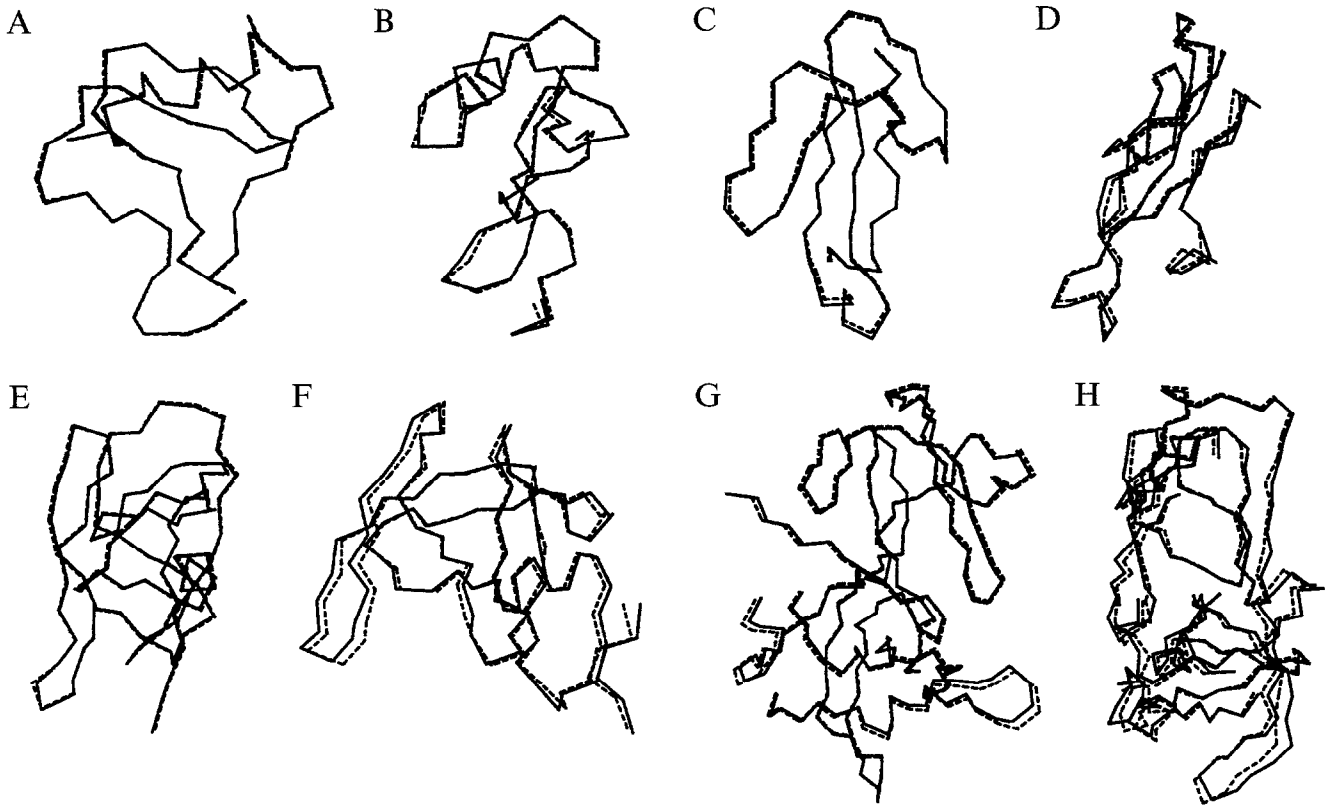

Fig. 3. The superpositions of the calculated (dotted line) and original (solid line) crystal structures. (A, B) 1SN3; (C, D) 1CTX; (E, F) 3HVP; (G, H) 2APP. (A, C, E, G) Viewed from the same direction as that of the published stereo figures. (B, D, F, H) Viewed along the X-axis from the negative to the positive direction (see Fig. 1). In the case of 1CTX, results without taking the *cis*-peptides into account are shown.

in the Z direction, although any verification of this deformation is presently unavailable. We then examined if the accuracy was improved when the global enlargement of the model in the Z direction was applied. The calculated Z coordinates were magnified by up to 5% in increments of 1%. After recalculating the score, the model with the best score was selected. The revised results are given as column B of Table 2. The atomic rmsd between the constructed and original structures ranges from 0.37 to 0.80 Å. Some improvement for the rmsd of the C^{α} positions was observed with respect to 1CTX and 2APP (by 2–3% enlargement). Since the score did not decrease by the expansion for 1SN3 and 3HVP, the modeled structure remained intact. The introduction of the model enlargement in the Z direction was thus successful and was incorporated into the ReconstCα program.

As mentioned in the Introduction section, the reconstruction of the α-carbon coordinates is considered to lead to the generation of the backbone structure of proteins. To this end, it is desirable to reduce the deformation of C^{α} geometry in the model. Final refinement of the coordinates was, therefore, achieved as illustrated in Fig. 4. The displacement vectors to set the C^{α} distance equal to 3.8 Å and to relieve any VBA violations were calculated. From the analysis of C^{α} geometry in the PDB files, almost all the distances between nonadjacent α-carbons are found to be longer than 3.7 Å. The correction vectors to release these violations, if any, were also computed. The sum of these displacement vectors was added to the atom positions and this procedure was repeated a few times until convergence. The SD of C^{α} distances and the rmsd from the PDB coordinates after the refinement are listed in column C of Table 2. The distance SDs were revised to be within the range of 0.01–0.02 Å, which is lower than those of experimental values. Furthermore, the rmsd values were slightly improved by the refinement, ranging from 0.36 to 0.78 Å.

Very recently, Oldfield and Hubbard [4] have devised a different approach to the extraction of 3D coordinates from α-carbon stereo diagrams. They modeled myoglobin (153 residues) [8] and interleukin 1β (151 residues) [9] using the program EXTRACT. The rmsd values of their models were 0.72 and 0.64 Å, respectively. As shown in Table 2, our results (values of 0.36–0.78 Å) are comparable to, and sometimes even better than, those obtained by Oldfield and Hubbard. We used the actual published figures as input, although it appears that they have employed the stereo diagrams produced from their native coordinates. In other words, stereo figures of myoglobin and interleukin 1β were not available and our method could not be applied to them.

A stereo-equipped workstation is necessary to use the

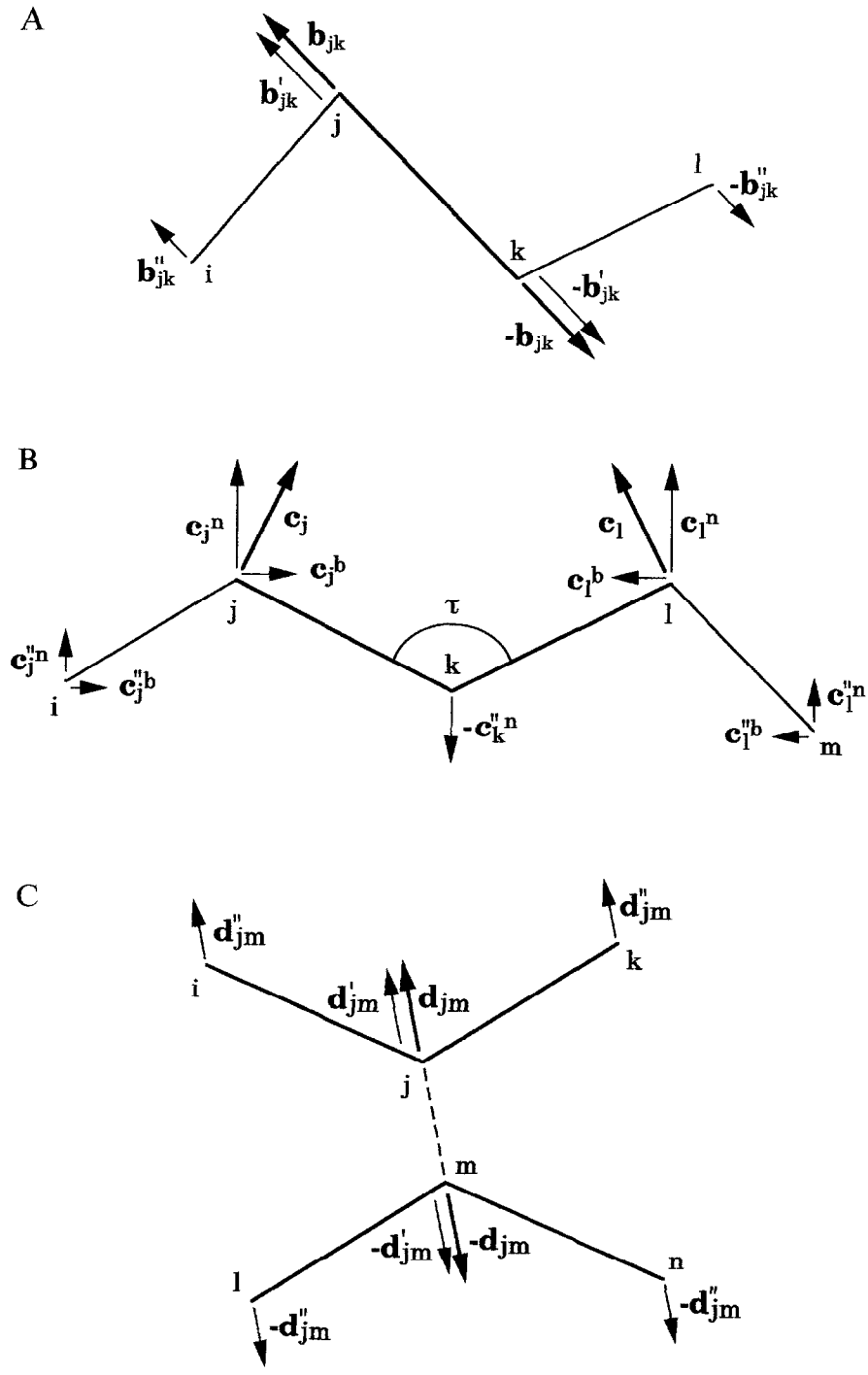

Fig. 4. Schematic drawing for the final refinement of the coordinates. (A) Two displacement vectors ($\mathbf{b}_{jk}$,$-\mathbf{b}_{jk}$) required to make the C^{α} distance equal to 3.8 Å are initially calculated for each pair of two contiguous α-carbons. They are of the same magnitude and of opposite orientation along the virtual C^{α}-C^{α} bond. The reduced vectors with 90% magnitude ($\mathbf{b}'_{jk}$,$-\mathbf{b}'_{jk}$) are assigned to atoms j and k and those with 45% magnitude ($\mathbf{b}''_{jk}$,$-\mathbf{b}''_{jk}$) are assigned to their neighboring atoms i and l. (B) Two correction vectors ($\mathbf{c}_j$ and $\mathbf{c}_l$) to relieve the VBA violation are initially computed for the angle τ formed by atoms j, k and l. These vectors with equal magnitude are perpendicular to the fictitious C^{α}-C^{α} bonds between atoms j and k, and atoms k and l. They are then divided into two vectors ($\mathbf{c}_j^n$ and $\mathbf{c}_j^b$, for example) which are nearly perpendicular to each other. One ($\mathbf{c}_j^b = -\mathbf{c}_l^b$) is along atoms j to l and the other ($\mathbf{c}_j^n = \mathbf{c}_l^n$) is along atom k to the middle point of atoms j and l. The reduced vectors with 90% magnitude of $\mathbf{c}_j^b$ and $\mathbf{c}_l^b$ are assigned to atoms j and l and those with 45% magnitude ($\mathbf{c}_j''^b$ or $\mathbf{c}_l''^b$) are assigned to their neighboring atoms i and m. The contracted vector with a 30% magnitude of $\mathbf{c}_j^n$ is assigned to atoms j and l and those with 15% magnitude ($\mathbf{c}_j'''^n = \mathbf{c}_l'''^n$) are assigned to their adjacent atoms i and m. The reduced vector with 60% magnitude and an opposite orientation of $\mathbf{c}_j^n$, i.e., $-\mathbf{c}_k'''^n$, is assigned to atom k. (C) Two displacement vectors ($\mathbf{d}_{jm}$,$-\mathbf{d}_{jm}$) to release the short contacts of C^{α} atoms, if any, are computed. They are of the same magnitude and of opposite orientation along atoms j to m. The reduced vectors with 90% magnitude ($\mathbf{d}'_{jm}$,$-\mathbf{d}'_{jm}$) are assigned to atoms j and m and those with 45% magnitude ($\mathbf{d}''_{jm}$,$-\mathbf{d}''_{jm}$) are assigned to their neighboring atoms i and k, and l and n. The sum of all displacement vectors acting on an atom is added to its position. This calculation and application cycle of refinement is repeated a few times until convergence.

EXTRACT program. It is not mandatory in our case however. In the approach by Oldfield and Hubbard, the scale and stereo angle are adjusted manually so that a template piece of canonical β-sheet or α-helix structure appears to fit well with an appropriate section that is considered to be a β-sheet or an α-helix. Although their method might be suitable for proteins such as myoglobin with a high α-helix content, it would seem difficult for it to be applied to proteins that have no distinct secondary structure. Actually, they mentioned that defining the correct scale and stereo angle is one of the most arduous aspects of their procedure. In our approach, on the other hand, there is no restriction of the secondary structure content and the adjustment of stereo parameters is performed automatically based on the SD of inter-C^{α} distances so that an unbiased model is built. When applying the method of Oldfield and Hubbard, positioning of the C^{α} atoms is carried out manually, one at a time, on the workstation using the diagram as a guide. If the new C^{α} position cannot be constructed satisfactorily, it is necessary to start again from the adjustment of the scale and stereo angle. In contrast, the calculation of the coordinates is done quickly at the same time as the estimation of stereo parameters in the case of the ReconstCα program.

It is hoped that the method described here can be used to provide a good starting point for the main chain construction. An efficient approach to backbone building is currently under investigation. The ReconstCα program will be available from JCPE [10], to which the program is to be submitted.

Conclusions

In this paper we have described an algorithm for generating the 3D coordinates of α-carbons from a pair of C^{α} trace figures. To carry out the calculations, the ReconstCα program was developed and applied to four proteins, namely scorpion neurotoxin, α-cobratoxin, HIV-1 protease and penicillopepsin. A comparison of the atom positions in the constructed and crystal reference structures gave rmsd values between 0.36 and 0.78 Å. This reasonable accuracy indicates the effectiveness of the program. Our approach would be useful to provide an initial structure for protein modeling studies.

References

1 a. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer Jr., E.F., Brice, M.D., Rodgers, J.R., Kennard, O., Shimanouchi, T. and Tasumi, M., J. Mol. Biol., 112 (1977) 535.
b. Abola, E.E., Bernstein, F.C., Bryant, S.H., Koetzle, T.F. and Weng, J., In Allen, F.H., Bergerhoff, G. and Sievers, R. (Eds.) Crystallographic Database – Information Content, Software Systems, Scientific Applications, Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester, 1987, pp. 107–132.

2 Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B.K., Baldwin, E., Weber, I.T., Selk, L.M., Clawson, L., Schneider, J. and Kent, S.B.H., Science, 245 (1989) 616.

3 Navia, M.A., Fitzgerald, P.M.D., McKeever, B.M., Leu, C.-T., Heimbach, J.C., Herber, W.K., Sigal, I.S., Darke, P.L. and Springer, J.P., Nature, 337 (1989) 615.

4 Oldfield, T.J. and Hubbard, R.E., J. Mol. Graph., 13 (1995) 18.

5 Almassy, R.J., Fontecilla-Camps, J.C., Suddath, F.L. and Bugg, C.E., J. Mol. Biol., 170 (1983) 497.

6 Walkinshaw, M.D., Saenger, W. and Maelicke, A., Proc. Natl. Acad. Sci. USA, 77 (1980) 2400.

7 James, M.N.G. and Sielecki, A.R., J. Mol. Biol., 163 (1983) 299.

8 Smerdon, S., Oldfield, T.J., Dodson, E.J., Dodson, G.G., Hubbard, R.E. and Wilkinson, A.J., Acta Crystallogr., B45 (1990) 327.

9 Finzel, B.C., Clancy, L.L., Holland, D.R., Muchmore, S.W., Watenpaugh, K.D. and Einspahr, H.M., J. Mol. Biol., 209 (1989) 779.

10 Japan Chemistry Program Exchange (JCPE), 1-7-12 Nishinenishi Tsuchiura-shi 300, Japan. If any difficulty is experienced in reaching JCPE, interested parties can contact us directly by e-mail: miya@sankyo.shina.co.jp.

Appendix

In this section we derive equations to calculate the 3D coordinates (x,y,z) from the 2D coordinates (X,Y) of stereo figures and stereo parameters (Fig. 1). Let us define the coordinate system so that the Z direction is perpendicular to the paper plane of the stereo diagrams and the X direction is along the left eye view to the right eye view as shown in Fig. 1. Suppose we look into a molecular model from the positive Z direction and CoR coincides with the origin. The paper plane, therefore, corresponds to the projection plane ($z = z_P$).

Let us denote a certain point in the model as P(x,y,z). When the model is rotated for the left and right eye views by θ, P is translated to $PL(x_L,y_L,z_L)$ and $PR(x_R,y_R,z_R)$, respectively, as shown in Fig. 1C. Figure 1E shows the correspondence of the sign of θ with the direction of the model rotation. The coordinates of PL and PR are given by

$$x_L = z \sin\theta + x \cos\theta \quad \text{(A1)}$$

$$z_L = z \cos\theta - x \sin\theta \quad \text{(A2)}$$

$$y_L = y \quad \text{(A3)}$$

$$x_R = -z \sin\theta + x \cos\theta \quad \text{(A4)}$$

$$z_R = z \cos\theta + x \sin\theta \quad \text{(A5)}$$

$$y_R = y \quad \text{(A6)}$$

The relationships between the coordinates of PL and PR and those of their projection points QL (X_L,Y_L) and QR (X_R,Y_R) are written as

$$x_L / X_L = (z_E - z_L) / (z_E - z_P) \quad \text{(A7)}$$
$$y_L / Y_L = (z_E - z_L) / (z_E - z_P) \quad \text{(A8)}$$
$$x_R / X_R = (z_E - z_R) / (z_E - z_P) \quad \text{(A9)}$$
$$y_R / Y_R = (z_E - z_R) / (z_E - z_P) \quad \text{(A10)}$$

where z_E is the z coordinate of the view point and z_P is that of the projection plane. Substituting Eqs. A1 and A2 into Eq. A7, we find

$$(z \sin\theta + x \cos\theta)z_{EP} = (z_E - z \cos\theta + x \sin\theta)X_L \quad \text{(A11)}$$

where $z_{EP} = z_E - z_P$.

In the same way, from Eqs. A4, A5 and A9 we obtain

$$(-z \sin\theta + x \cos\theta)z_{EP} = (z_E - z \cos\theta - x \sin\theta)X_R \quad \text{(A12)}$$

The rearrangements of Eqs. A11 and A12 lead to

$$(z_{EP} \cos\theta - X_L \sin\theta)x + (X_L \cos\theta + z_{EP} \sin\theta) = z_E X_L \quad \text{(A13)}$$

$$(z_{EP} \cos\theta + X_R \sin\theta)x + (X_R \cos\theta - z_{EP} \sin\theta) = z_E X_R \quad \text{(A14)}$$

The solution of the above simultaneous linear equations is given by

$$x = z_E z_{EP}(X_L + X_R)\sin\theta / \{2(z_{EP}^2 + X_L X_R)\sin\theta\cos\theta + z_{EP}(X_L - X_R)(\cos^2\theta - \sin^2\theta)\} \quad \text{(A15)}$$

$$z = \{2z_E X_L X_R \sin\theta + z_E z_{EP}(X_L - X_R)\cos\theta\} / \{2(z_{EP}^2 + X_L X_R)\sin\theta\cos\theta + z_{EP}(X_L - X_R)(\cos^2\theta - \sin^2\theta)\} \quad \text{(A16)}$$

On the other hand, substituting Eqs. A2 and A8 into Eq. A3, we find

$$y = (z_E - z \cos\theta + x \sin\theta)Y_L / z_{EP} \quad \text{(A17)}$$

In the same way, from Eqs. A5, A10 and A6 we obtain

$$y = (z_E - z \cos\theta - x \sin\theta)Y_R / z_{EP} \quad \text{(A18)}$$

Substituting x and z calculated from Eqs. A15 and A16 into Eqs. A17 and A18, two values of y are obtained. Although these should be the same theoretically, they are slightly different in practice. We have, therefore, taken the average of the two as the final y value. Thus, 3D coordinates of P(x,y,z) are calculated from the coordinates of $PL(x_L,y_L,z_L)$ and $PR(x_R,y_R,z_R)$ and the estimated stereo parameters (θ,z_E,z_P).