Protein Science (1992), *1*, 1279–1287. Cambridge University Press. Printed in the USA.

On the multiple simultaneous superposition of molecular structures by rigid body transformations

R. DIAMOND

MRC Laboratory for Molecular Biology, Hills Road, Cambridge CB2 2QH, England

(RECEIVED January 31, 1992; REVISED MANUSCRIPT RECEIVED May 12, 1992)

Abstract

A method of optimally superimposing n coordinate sets on each other by rigid body transformations, which minimizes the sum of all $n(n-1)/2$ pairwise residuals, is presented. In the solution phase the work load is approximately linear on n, is independent of the size of the structures, is independent of their initial orientations, and terminates in one cycle if $n = 2$ or if the coordinate sets are exactly superposable, and otherwise takes a number of cycles dependent only on genuine shape differences. Enantiomorphism, if present, is detected, in which case the option exists to reverse or not to reverse the chirality of relevant coordinate sets. The method also offers a rational approach to the problem of multiple minima and has successfully identified four distinct minima in such a case. Source code, which is arranged to enable the study of the disposition of domains in multidomain structures, is available from the author.

Keywords: algorithms; comparisons; quarternions; rotations; superpositions

With the development of NMR and molecular dynamics techniques it is becoming increasingly common to wish to compare a number of actual or putative structures by superimposing their coordinates using rigid-body transformations (translation and strain-free rotation) and to measure the root mean square (rms) coordinate differences that result.

The superposition of one structure on one other is a problem to which many solutions have been offered, such as those of McLachlan (1972, 1979, 1982), Kabsch (1976, 1978), Diamond (1976, 1988), Lesk (1986), and Kearsley (1989), but the optimal superposition of ensembles of structures has only recently received similar attention. The situation is complicated by the fact that if structure A is superimposed on structure C, and structure B is superimposed on structure C, then, in general, structure A is not optimally superimposed on structure B. In these circumstances the superposition of A on B is only optimal if two of the three structures are identical in shape.

If the sum of the squares of the coordinate differences between A and B is designated E_{AB}, and if it is wished to compare n structures by superposition, there are $n(n-1)/2$ interactions such as E_{AB}, and the method described in this paper minimizes the sum of all of these in a process that (apart from initialization) is of order n, rather than n^2, and places all n structures on an exactly equivalent footing, except that the first structure has the most influence on the orientation of the entire ensemble after all relationships within the ensemble are determined.

No "average" structure of any kind is involved at any stage. Average structures are best avoided because two identical structures in different orientations have an average structure that is not exactly superposable on either.

The method proposed here allows each structure to be optimally superimposed on all other structures in the ensemble, even if they are no longer in their original orientations, using a one-step method which does not require reference to the coordinates except in the initialization phase. Iteration is only required to cycle over the various structures, bringing each in turn to its best relationship with all $(n-1)$ others in the ensemble simultaneously.

The alternative approach of reorienting all n structures simultaneously has been adopted by Kearsley (1990). His method entails an optimization in $4n$ dimensions, for which the Hessian matrix, **H**, may be large. The objective function is not a quadratic function of the unknowns, so that iteration is a feature of his method also, as many as 50 cycles being required in some cases. Like the method described below, it uses unit vectors in four dimensions (quarternions) to define the rotations, but their unit character is only maintained by the application

Reprint requests to: R. Diamond, MRC Laboratory for Molecular Biology, Hills Road, Cambridge CB2 2QH, England.

of a penalty function. In contrast, the method described here uses no matrices larger than 4 × 4 (although their total storage requirements are the same as for **H**), the variables involved are intrinsically normalized, and, in current experience, convergence has always been achieved within nine cycles even in the most adverse circumstances.

More recently, Shapiro et al. (1992) have developed a method comparable to the present one in that it adjusts one orientation at a time in an iterative fashion, but working with 3 × 3 orthogonal matrices as independent variables. They also address the question of the uniqueness of the solution obtained and show that the minimum residual reached by their method is so close to the theoretical limit that it is almost certain that no minimum exists that is lower than the minimum found. This question of uniqueness is also addressed in the Multiple solutions and the global minimum section in this paper, wherein a means of characterizing the problem and of locating alternative solutions is offered. As is typical of techniques using 3 × 3 orthogonal matrices, matrices with negative determinants may arise in some circumstances, and special attention is required to ensure that unwanted reversals of chirality do not occur. By contrast, the method described below is immune to these effects because the structure of the matrices involved ensures that negative determinants can never arise.

Theory

The treatment is based on a four-dimensional column vector

$$\boldsymbol{\rho} = \begin{pmatrix} \lambda \\ \mu \\ \nu \\ \sigma \end{pmatrix} = \begin{pmatrix} l\sin(\theta/2) \\ m\sin(\theta/2) \\ n\sin(\theta/2) \\ \cos(\theta/2) \end{pmatrix}, \tag{1}$$

in which l, m, and n are the direction cosines of an axis of rotation, and θ is an angle through which a rotation is being made. The sign convention for θ is such that, when viewed along the rotation axis from the origin toward the point lmn, an object is rotated clockwise for positive θ for a fixed right-handed axial system. It is evident that a $\boldsymbol{\rho}$ vector completely specifies a rotation, and it provides a very much more powerful means of dealing with combinations of rotations and with the interactions, E, than does the more conventional representation of rotations in terms of the 3 × 3 orthogonal matrices which effect coordinate transformations.

$\boldsymbol{\rho}$ is always a unit vector in the sense that

$$\boldsymbol{\rho}^T\boldsymbol{\rho} = \lambda^2 + \mu^2 + \nu^2 + \sigma^2 = 1, \tag{2}$$

and it also has the property

$$\boldsymbol{\rho} \equiv -\boldsymbol{\rho}, \tag{3}$$

where the equivalence sign is used to express the fact that $\boldsymbol{\rho}$ and $-\boldsymbol{\rho}$ give rise to the same orientation. They may be thought of as generating rotations through angles θ and $\theta - 2\pi$, respectively, i.e., to represent the short and the long way round to a new orientation. $\bar{\boldsymbol{\rho}}$, the inverse of $\boldsymbol{\rho}$, is given by negating θ,

$$\bar{\boldsymbol{\rho}} = \begin{pmatrix} -\lambda \\ -\mu \\ -\nu \\ \sigma \end{pmatrix}, \tag{4}$$

and an inversion operation $\bar{\mathbf{I}}$ may be defined by

$$\bar{\mathbf{I}} = \begin{pmatrix} -\mathbf{I} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix}, \tag{5}$$

in which $\mathbf{I}$ is the 3 × 3 identity (and later, also, the 4 × 4 identity) in terms of which

$$\bar{\boldsymbol{\rho}} = \bar{\mathbf{I}}\boldsymbol{\rho}, \tag{6}$$

and by Equation 3 it follows that

$$\bar{\mathbf{I}} \equiv -\bar{\mathbf{I}}. \tag{7}$$

The vector $\boldsymbol{\rho}$ may also be expanded into a 4 × 4 matrix, written $[\boldsymbol{\rho}]$, defined by

$$[\boldsymbol{\rho}] = \begin{pmatrix} \sigma & -\nu & \mu & \lambda \\ \nu & \sigma & -\lambda & \mu \\ -\mu & \lambda & \sigma & \nu \\ -\lambda & -\mu & -\nu & \sigma \end{pmatrix}. \tag{8}$$

Such matrices have a number of important properties, among them that

$$[\boldsymbol{\rho}]^T[\boldsymbol{\rho}] = \mathbf{I}, \tag{9}$$

i.e., $[\boldsymbol{\rho}]$ is orthogonal, and

$$[\boldsymbol{\rho}]^T = [\bar{\boldsymbol{\rho}}] = [\bar{\mathbf{I}}\boldsymbol{\rho}], \tag{10}$$

so that the transpose of $[\boldsymbol{\rho}]$ relates to the inverse rotation, as is the case with the more familiar coordinate transformations. $[\boldsymbol{\rho}]$ always has a positive determinant.

In an earlier paper (Diamond, 1988) it was shown that if rotation 1 is followed by rotation 2, then the $\boldsymbol{\rho}$ vector corresponding to the product rotation is given by

$$\boldsymbol{\lambda} = \boldsymbol{\lambda}_2\sigma_1 + \boldsymbol{\lambda}_1\sigma_2 + \boldsymbol{\lambda}_2 \times \boldsymbol{\lambda}_1$$
$$\sigma = \sigma_1\sigma_2 - \boldsymbol{\lambda}_2 \cdot \boldsymbol{\lambda}_1 \tag{11}$$

in which $\boldsymbol{\lambda}$ is a three-dimensional vector consisting of the first three components of $\boldsymbol{\rho}$. This is equivalent to

$$\boldsymbol{\rho} = [\boldsymbol{\rho}_2]\boldsymbol{\rho}_1, \tag{12}$$

and since

$$\boldsymbol{\rho}_1 = [\boldsymbol{\rho}_1]\boldsymbol{\rho}_0, \tag{13}$$

in which $\boldsymbol{\rho}_0$ is the identity rotation given by

$$\boldsymbol{\rho}_0 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \tag{14}$$

it follows that

$$\boldsymbol{\rho} = [\boldsymbol{\rho}_2][\boldsymbol{\rho}_1]\boldsymbol{\rho}_0, \tag{15}$$

which, for n rotations compounded, generalizes to

$$\boldsymbol{\rho} = [\boldsymbol{\rho}_n][\boldsymbol{\rho}_{n-1}] \ldots [\boldsymbol{\rho}_1]\boldsymbol{\rho}_0 \tag{16}$$

(which is Equation 46 of Diamond [1988], wherein the symbol **T** was used for $[\boldsymbol{\rho}]$). Equation 15 says that the product rotation vector, $\boldsymbol{\rho}$, is the fourth column of the matrix product $[\boldsymbol{\rho}_2][\boldsymbol{\rho}_1]$. Evaluation of this product shows it to be structured in the same way as Equation 8, so that we may generalize Equation 16 to

$$[\boldsymbol{\rho}] = [\boldsymbol{\rho}_n][\boldsymbol{\rho}_{n-1}] \ldots [\boldsymbol{\rho}_1]. \tag{17}$$

Diamond (1988) showed that the weighted sum of squares of coordinate differences between a vector set **X** and a rotated vector set **Rx**, is given by

$$E = E_0 - 2\boldsymbol{\rho}^T\mathbf{P}\boldsymbol{\rho}, \tag{18}$$

in which E_0 is the value associated with **X** and the unrotated **x**, i.e., for $\mathbf{R} = \mathbf{I}$, **P** is a real symmetric 4×4 matrix bilinear on **X** and on **x**, and $\boldsymbol{\rho}$ is the rotation vector specifying the rotation effected by the orthogonal 3×3 matrix **R**. The importance of this equation is that it provides an analytical link between E and any rotation one may wish to specify, and gives access to the value of E before any coordinates are transformed. In particular, it is clear that E is minimized if $\boldsymbol{\rho}$ is the top eigenvector of **P**. Equation 18 holds whatever choice of origin is made, but the lowest E values arise if each structure is referred to its centroid as origin.

In the context of the present work Equation 18 will be generalized to

$$E_{AB} = E_{0AB} - 2\boldsymbol{\rho}_A^T\mathbf{P}_{AB}\boldsymbol{\rho}_A \tag{19}$$

for the pair of structures A and B, $\boldsymbol{\rho}_A$ is the rotation to be applied to structure A and $\mathbf{P}_{AB}$ is as defined in Diamond (1988), $\mathbf{X}_A$ corresponding to **x** and $\mathbf{X}_B$ to **X**. As pointed out in Diamond (1988), interchanging **x** and **X** negates the last row and last column of **P**, i.e.,

$$\mathbf{P}_{BA} = \bar{\mathbf{I}}\mathbf{P}_{AB}\bar{\mathbf{I}}, \tag{20}$$

in terms of which E_{AB} may alternatively be written

$$E_{AB} = E_{BA} = E_{0BA} - 2\boldsymbol{\rho}_B^T\mathbf{P}_{BA}\boldsymbol{\rho}_B \tag{21}$$

so that $\mathbf{P}_{BA}$ may be used to control the rotation of B onto A in the same way that $\mathbf{P}_{AB}$ may control the rotation of A onto B.

Suppose that structure A has already been rotated by $\boldsymbol{\rho}_A$ and it is required to find a rotation $\boldsymbol{\rho}_B$ that will optimally rotate the unrotated B structure onto the rotated A structure, then the rotation $\boldsymbol{\rho}_B$ of structure B, followed by the rotation $\bar{\boldsymbol{\rho}}_A$ of both structures, would rotate the unrotated B structure onto the unrotated A structure. Therefore in these circumstances $[\bar{\boldsymbol{\rho}}_A]\boldsymbol{\rho}_B$ must be the top eigenvector of $\mathbf{P}_{BA}$, i.e., we require to maximize

$$\boldsymbol{\rho}_B^T[\bar{\boldsymbol{\rho}}_A]^T\mathbf{P}_{BA}[\bar{\boldsymbol{\rho}}_A]\boldsymbol{\rho}_B = \boldsymbol{\rho}_B^T[\boldsymbol{\rho}_A]\mathbf{P}_{BA}[\boldsymbol{\rho}_A]^T\boldsymbol{\rho}_B, \tag{22}$$

and $\boldsymbol{\rho}_B$ must therefore be the top eigenvector of

$$[\boldsymbol{\rho}_A]\mathbf{P}_{BA}[\boldsymbol{\rho}_A]^T. \tag{23}$$

Thus it is possible to direct the orientation of structure B to fit structure A, even if structure A is not in its original orientation, still without referring to the coordinates.

In order to minimize E_{tot} $(=\sum_A \sum_{B \neq A} E_{AB})$ it is worth noting that, in the case of three structures, if A is first rotated to minimize $E_{AB} + E_{AC}$, then the rotated structure A is in some sense intermediate between B and C. If B is then rotated to minimize $E_{BA} + E_{BC}$, structure B will then be intermediate between the rotated A and the unrotated C. If A, B, and C initially have widely differing orientations, then this process condenses the orientations to closely similar ones more slowly than if each structure is superimposed on a chosen member of the ensemble on the first cycle, with each being optimally superimposed on all others simultaneously in second and subsequent cycles.

An optimization algorithm for an ensemble, having this characteristic, is illustrated below for the case of four structures; its generalization to larger numbers being obvious. We begin by calculating all $n(n-1)/2$ matrices

$\mathbf{P}_{AB}$ and their complements $\mathbf{P}_{BA}$, this being the only stage at which the coordinates need to be consulted. These matrices are themselves arranged in an array:

$$\begin{matrix} . & \mathbf{P}_{12} & \mathbf{P}_{13} & \mathbf{P}_{14} \\ \mathbf{P}_{21} & . & \mathbf{P}_{23} & \mathbf{P}_{24} \\ \mathbf{P}_{31} & \mathbf{P}_{32} & . & \mathbf{P}_{34} \\ \mathbf{P}_{41} & \mathbf{P}_{42} & \mathbf{P}_{43} & . \end{matrix} \tag{24}$$

We begin by finding the four eigenvalues of $\mathbf{P}_{21}$ using the methods of Householder (Wilkinson, 1960) and Ortega (1960) coded by D.W. Matula. These are designated p_1 to p_4 in descending order, and we evaluate

$$E'_{\min} - E_{\min} = p_1 - p_2 - p_3 + p_4, \tag{25}$$

in which $E_{\min}$ is the minimum residual attainable in rotating structure 2 onto structure 1 and $E'_{\min}$ is the minimum attainable by rotation of structure 2 onto structure 1 inverted through the origin (Diamond, 1990, Equation 10). Hence, if the right-hand side of Equation 25 is negative, an enantiomorphous relationship has been detected between structures 1 and 2. Several options are then open, such as

1. Disregard the fact and obtain the best superposition of structures of opposite hand.
2. Reverse the hand of structure 2 before proceeding. This may be done by replacing E_{012} by

$$\begin{aligned} E'_{012} &= E_{012} - \operatorname{tr} \mathbf{P}_{12} \\ E'_{02J} &= E_{02J} - \operatorname{tr} \mathbf{P}_{2J} \end{aligned} \tag{26}$$

   (Diamond, 1990) and then negating every **P** matrix with a 2 in either subscript.
3. Remove structure 2 from the ensemble.

Assuming option 1 or 2 is taken, $\mathbf{P}_{21}$ is then solved for $\boldsymbol{\rho}_2$, the top eigenvector of $\mathbf{P}_{21}$, using the method of Wilkinson (1958). All other structures must now recognize that structure 2 is a moved target, so that all matrices in the second column of the array of **P** matrices must now be transformed according to expression 23.

$\mathbf{P}_{31}$ and then $\mathbf{P}_{41}$ are similarly processed, after which the array has become

$$\begin{matrix} . & [\boldsymbol{\rho}_2]\mathbf{P}_{12}[\boldsymbol{\rho}_2]^T & [\boldsymbol{\rho}_3]\mathbf{P}_{13}[\boldsymbol{\rho}_3]^T & [\boldsymbol{\rho}_4]\mathbf{P}_{14}[\boldsymbol{\rho}_4]^T \\ \mathbf{P}_{21} & . & [\boldsymbol{\rho}_3]\mathbf{P}_{23}[\boldsymbol{\rho}_3]^T & [\boldsymbol{\rho}_4]\mathbf{P}_{24}[\boldsymbol{\rho}_4]^T \\ \mathbf{P}_{31} & [\boldsymbol{\rho}_2]\mathbf{P}_{32}[\boldsymbol{\rho}_2]^T & . & [\boldsymbol{\rho}_4]\mathbf{P}_{34}[\boldsymbol{\rho}_4]^T \\ \mathbf{P}_{41} & [\boldsymbol{\rho}_2]\mathbf{P}_{42}[\boldsymbol{\rho}_2]^T & [\boldsymbol{\rho}_3]\mathbf{P}_{43}[\boldsymbol{\rho}_3]^T & . \end{matrix} \tag{27}$$

At this stage every structure has been individually rotated onto structure 1. If the structures are identical except for their initial orientations the problem is now solved, as may be revealed by Equation 19 evaluating to zero for each interaction with structure 1. If shape differences exist among the structures, the problem is not yet solved and further work is needed, as follows.

Each of the transformed matrices on the top row of 27 gives the residual as between the unrotated structure 1 and each of the other rotated structures as a function of any rotation now to be applied to structure 1. Therefore the rotation to be applied to structure 1 to minimize the sum of its residuals with all the other rotated structures is the top eigenvector of $\boldsymbol{\Pi}_1$ where

$$\boldsymbol{\Pi}_1 = \sum_{J \neq 1} [\boldsymbol{\rho}_J]\mathbf{P}_{1J}[\boldsymbol{\rho}_J]^T. \tag{28}$$

This generates a $\boldsymbol{\rho}_1$ by which the first column of 27 may be transformed.

A new vector $\boldsymbol{\rho}_2$ is then formed as the top eigenvector of

$$\boldsymbol{\Pi}_2 = \sum_{J \neq 2} [\boldsymbol{\rho}_J]\mathbf{P}_{2J}[\boldsymbol{\rho}_J]^T, \tag{29}$$

after which the second column of the array of **P** matrices is replaced. The new $\boldsymbol{\rho}_2$ vector found at this stage is still a rotation from the original orientation of structure 2, to superimpose 2 on all the other structures in their current orientations. Hence the new $\boldsymbol{\rho}_2$ replaces the old $\boldsymbol{\rho}_2$ and is not to be concatenated with it. This, incidentally, minimizes the accumulation of round-off error.

The process of using row sums to determine $\boldsymbol{\rho}$ vectors followed by changes to corresponding columns may be pursued to convergence, which is normally very rapid.

The scheme outlined above began by individually optimizing each structure against structure 1. It follows that when the top row is processed for the first time there is no rotation of structure 1 required, i.e.,

$$\boldsymbol{\rho}_1 = \boldsymbol{\rho}_0. \tag{30}$$

That this is computationally the case is a strong check on both the algebra and the coding. $\boldsymbol{\rho}_1$ may not be the identity on subsequent cycles, however, and would not be the identity if the step of rotating each structure individually onto structure 1 to begin with were omitted.

On completion, the quantities

$$E_A = \left(\sum_{B \neq A} E_{0AB} \right) - 2\boldsymbol{\rho}_A^T \boldsymbol{\Pi}_A \boldsymbol{\rho}_A \tag{31}$$

give the total squared errors as between structure A and all other members of the ensemble and serve to detect outliers of different shape.

The corresponding rotation matrices, **R**, by which the coordinates must be transformed for each structure are then given by

$$\begin{pmatrix} \mathbf{R}_A & \mathbf{0} \\ \mathbf{0}^T & 1 \end{pmatrix} = [\boldsymbol{\rho}_A]\bar{\mathbf{I}}[\boldsymbol{\rho}_A]\bar{\mathbf{I}}. \tag{32}$$

Note that the determinant of $\mathbf{R}_A$ is the square of that of $[\boldsymbol{\rho}_A]\bar{\mathbf{I}}$ and is therefore always positive.

Multiple solutions and the global minimum

Shapiro et al. (1992) have drawn attention to the nonlinear character of the multiple superposition problem and the implication this carries that there may be several minima of E_{tot} in the space of the rotational variables and they have developed a criterion to establish that the minimum found in their process is indeed the global minimum.

In the present context, it is clear that each E_{AB} is a quadratic function of $\boldsymbol{\rho}_A$ in which the coefficients are themselves quadratic functions of the elements of $\boldsymbol{\rho}_B$, and that the converse is also true. This means that the minimum of E with respect to any one rotation, under constancy of the others, is always unique or a continuum, but leaves open the question of whether alternative combinations of $\boldsymbol{\rho}$ vectors may also be stable under the optimizing scheme given above, because the $\mathbf{\Pi}$ matrices that have to be solved are quadratic functions of the current orientation vectors.

In order to search for alternative minima, and to develop some insight into the character of the problem, the approach adopted here is to change the combinations of rotations found in ways that alter the orientations grossly, while at the same time altering E as little as possible, and then to initialize the iterative phase of the optimization to such a situation.

Consider again the case of one structure being rotated onto one other, and let the rotating structure have three orientations, an initial orientation, and two others designated ξ and ζ, with associated $\boldsymbol{\rho}$ vectors $\boldsymbol{\rho}_\xi$ and $\boldsymbol{\rho}_\zeta$. Let the corresponding **P** matrix be diagonalized by an orthogonal matrix **A** such that

$$\mathbf{\Lambda} = \mathbf{A}^T\mathbf{P}\mathbf{A} = \mathrm{diag}(p_1, p_2, p_3, p_4), \tag{33}$$

and let the columns of **A** be designated $\mathbf{a}_1 \ldots \mathbf{a}_4$. Then resolve $\boldsymbol{\rho}_\xi$ and $\boldsymbol{\rho}_\zeta$ onto these eigenvectors according to

$$\boldsymbol{\rho}_\xi = \mathbf{A}\boldsymbol{\xi}, \qquad \boldsymbol{\rho}_\zeta = \mathbf{A}\boldsymbol{\zeta}, \tag{34}$$

then necessarily

$$\boldsymbol{\xi}^T\boldsymbol{\xi} = \boldsymbol{\zeta}^T\boldsymbol{\zeta} = 1. \tag{35}$$

By Equation 17 the rotation from ξ to ζ is given by

$$\begin{aligned}[\boldsymbol{\rho}_{\xi\to\zeta}] &= [\mathbf{A}\boldsymbol{\zeta}][\mathbf{A}\boldsymbol{\xi}]^T \\ &= \sum_{i=1}^{4}\sum_{j=1}^{4}[\mathbf{a}_i][\mathbf{a}_j]^T\zeta_i\xi_j.\end{aligned} \tag{36}$$

Now suppose that ξ is the best superposition, so that $\boldsymbol{\rho}_\xi$ is the top eigenvector of **P**, i.e., $\boldsymbol{\xi}^T = (1,0,0,0)$, and let $\boldsymbol{\zeta}^T = (\alpha,\beta,0,0)$ then

$$[\boldsymbol{\rho}_{\xi\to\zeta}] = \alpha[\mathbf{a}_1][\mathbf{a}_1]^T + \beta[\mathbf{a}_2][\mathbf{a}_1]^T, \tag{37}$$

in which $[\mathbf{a}_1][\mathbf{a}_1]^T$ is the identity, and $[\mathbf{a}_2][\mathbf{a}_1]^T$ is a $[\cdot]$ matrix having zeros on the diagonal, because $\mathbf{a}_2^T\mathbf{a}_1 = 0$, and therefore corresponds to a rotation of 180°. Comparison with Equation 8 shows that

$$\begin{aligned}\alpha &= \cos\frac{\theta_{\xi\to\zeta}}{2} \\ \beta &= \sin\frac{\theta_{\xi\to\zeta}}{2}.\end{aligned} \tag{38}$$

By Equation 18

$$\begin{aligned}E_\zeta - E_\xi &= 2(\boldsymbol{\xi}^T\mathbf{\Lambda}\boldsymbol{\xi} - \boldsymbol{\zeta}^T\mathbf{\Lambda}\boldsymbol{\zeta}) \\ &= 2(p_1 - (\alpha^2 p_1 + \beta^2 p_2)) \\ &= 2\beta^2(p_1 - p_2).\end{aligned} \tag{39}$$

Thus, the rotation $\xi \to \zeta$ corresponds to a rotation of $2\arctan(\beta,\alpha)$ away from the orientation of best fit, about an axis that incurs least penalty, because $p_1 - p_J$ is least when J = 2.

It is clear, therefore, that widely differing solutions may be expected to arise if two or more of the matrices $\mathbf{P}_{AB}$ have small differences between their first and second eigenvalues. This possibility is illustrated in Figure 1, in which Figure 1a shows three cubes, identical except for the labeling of their vertices, which specify which corners are to be brought together by superposition. Cube B has its floor coinciding with that of A, but its ceiling is rotated 180° about the z axis relative to that of A. This means that E_{AB} is independent of any rotation of cube A or cube B about the z axis, and that the eigenvalues of $\mathbf{P}_{AB}$ are equal in pairs. Similarly, E_{AC} is independent of rotations about y, and E_{BC} is independent of a certain compound rotation. Thus, all three **P** matrices and their complements are degenerate, but the $\mathbf{\Pi}$ matrices that arise are well conditioned, two distinct solutions being obtainable, shown in Figure 1b,c. The two solutions found involve rotating B and C about z and y respectively, two combinations of such rotations being found that satisfy the requirements of the BC interaction also.

An algorithm has been developed, based on these observations, which is an extension to the algorithm described in the Theory section, which is executed first in any case. During the processing of the first column, the

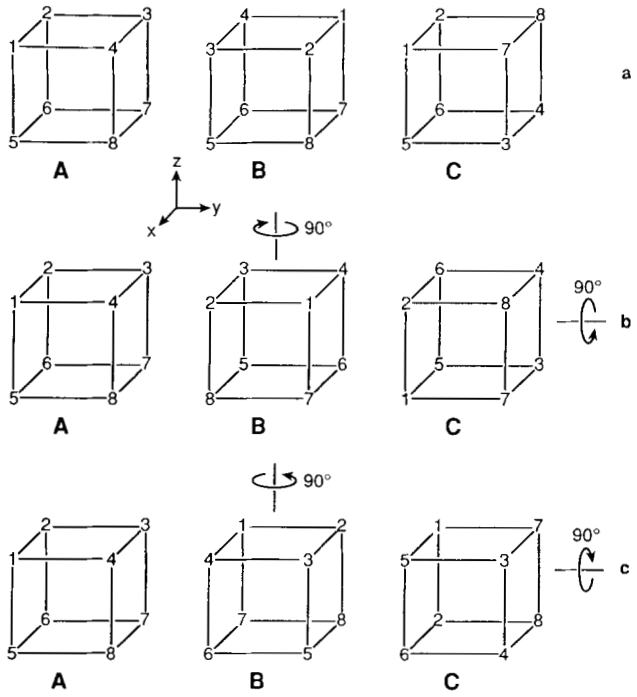


Fig. 1. **a** (top row): Three structures in their initial orientations for which E_{tot} is $28l^2$, where l is the length of the cube edge. **b**, **c** (second, third rows): Two distinct optimal superpositions for each of which E_{tot} is $24l^2$. This confirms that optimal superpositions are not necessarily unique. Both solutions are found automatically by the algorithm proposed, and a four-solution case has also been solved. The numbers indicate which corners are to be superimposed on which.

eigenvalue differences $p_1 - p_2$ are recorded for each of the matrices $\mathbf{P}_{l1}$ and sorted into ascending order. After the process of the Theory section has terminated, it is then reentered and the processing of the first column, which initializes the iterative phase of the work, is then done using the second eigenvector of selected structures, so that the iteration begins with certain structures rotated 180° about their least effective axis relative to structure 1. The software permits the user to declare how many structures, t, may be turned in this way, and lower and upper limits on the number of these that may be turned simultaneously. The number of such trials is then $\sum_{r=l}^{u} t!/((r!\,(t-r)!)$, the t structures having smallest $(p_1 - p_2)$ being used. This approach finds both the solutions shown in Figure 1, and finds four distinct solutions if a fourth cube is included that is similarly twisted about x.

Whether or not this process is the best that may be devised, it is clear that this organization of the problem in terms of **P** and **Π** matrices provides a convenient framework for experimentation with alternatives, for example, by using eigenvalues of the final **Π** matrices to select candidates for reversal, or even exploring the consequences of initializing the iterative phase by maximally disturbing the system using third or fourth eigenvectors.

The example given here, of course, is highly contrived, although it serves to illustrate the considerations involved. The likelihood that an ensemble of protein structures may possess more than one stable minimal superposition seems remote, although structures containing helices possessing differing numbers of turns for the same number of residues would seem to be candidates. Even in those circumstances, however, it is difficult to see how any third structure might stabilize two or more minima, because twists about axes normal to the helix axis are not a realistic possibility.

Root mean square deviations

There are several possible measures that may be used to measure the closeness of fit obtained in multiple superpositions. One that is commonly used is to perform $n(n-1)/2$ independent pairwise superpositions and to evaluate

$$R_0 = \left(\sum_{A} \sum_{B \neq A} E^*_{AB}/mn(n-1) \right)^{1/2} \tag{40}$$

for m atoms and n structures, each E^*_{AB} being determined in the absence of all $(n-2)$ other structures (i.e., by using **P** matrices rather than **Π**). This is necessarily the smallest such measure because E^*_{AB} is, by definition, the best possible fit of A on B, and can therefore only be degraded if the fit of A on B is compromised by the requirements of the other structures, but it does not correspond to a superposition, because many different orientations are associated with each structure. R_0 may be obtained in the present context as

$$R_0 = \left(\sum_{A} \sum_{B \neq A} (E_{0AB} - 2p_{1AB})/mn(n-1) \right)^{1/2}, \tag{41}$$

p_{1AB} being the top eigenvalue of $\mathbf{P}_{AB}$.

The residual adopted here is

$$R_1 = \left(\sum_{A} \left(\left(\sum_{B \neq A} E_{0AB} \right) - 2\pi_{1A} \right) \Big/ mn(n-1) \right)^{1/2}, \tag{42}$$

in which π_{1A} is the top eigenvalue of $\mathbf{\Pi}_A$. R_1 differs from R_0 because of the difference between the sum of the top eigenvalues of matrices, and the top eigenvalue of the matrix sum. A large value of $R_1 - R_0$ may be an indication that the minimum residual, R_1, found is not the global minimum, and would suggest that the methods of the Multiple solutions and the global minimum section might be invoked. However, multiple minima can exist when $R_1 = R_0$, as shown above.

If an average structure is calculated from the superposed structures then an rms difference from the mean structure is given by

$$R_2 = R_1((n-1)/2n)^{1/2}. \quad (43)$$

No comparable residual derived from R_0 can be given because in that context the ith atom of the jth structure does not have a defined position. The jth structure is present in many orientations, an arbitrary number of which are unrotated and the remainder of which are matched to unrotated structures whose orientations are themselves arbitrary.

Multidomain cases

The method described above has been implemented in software that permits an arbitrary number of structures, each possessing the same arbitrary number of domains, to be superimposed using each domain in turn to control the superposition. The domains must possess equal numbers of atoms from structure to structure (with an established correspondence) but with variable numbers from domain to domain. This generates a set of $\boldsymbol{\rho}$ vectors, $\boldsymbol{\rho}_{SD}$, referring to the rotation applied to structure S to rotate it from its initial orientation to its superposed one when all structures are superposed by reference to domain D.

Any structure, S_0, may be nominated as the structure which, on output, is to retain its initial orientation, the rotation vectors delivered being

$$\boldsymbol{\rho}'_{SD} = [\boldsymbol{\rho}_{S_0D}]^T \boldsymbol{\rho}_{SD}. \quad (44)$$

If S_0 is declared to be zero then $[\boldsymbol{\rho}_{0D}]^T$ is set equal to the rotation which rotates the superposed ensemble (all structures, one domain) from the orientation given by the superposing algorithm to an orientation in which the principal axes of inertia of the superposed ensemble are aligned on the axes of the coordinate system. Equation 44 then gives the individual orientations for each of the structures on this principal axial system. This axial system is found by including the matrices $\mathbf{P}_{AA}$ associated with the diagonal of expression 24 when that array is established, because the upper left 3×3 partition of each $\mathbf{P}_{AA}$ is minus two times the inertia tensor for the structural element A, in its original orientation. When the superposition is completed, each inertia tensor is transformed in accordance with the rotation found for the relevant structure, the resulting tensors are added together to give the inertia tensor for the superposed ensemble, and $\bar{\boldsymbol{\rho}}_{0D}$ is then the rotation which diagonalizes this tensor sum. This facility provides for the automatic orientation of results when initial orientations are random, so that, for example, if the superposition is based on a selected helix or a heme group as controlling domain, then setting $S_0 = 0$ causes the helix to align its longest dimension along x or the heme group to lie in the xy plane.

As noted above, the best superpositions are achieved if each structural element to be superposed is first referred to its centroid as origin, so that the position vector $\bar{\mathbf{x}}_{SD}$ of the centroid of domain D of structure S in its original coordinate system must be found for each S and D as a preliminary. Then, in the coordinate system of the output, the translation from the centroid of domain D_1 of structure S_1 to the centroid of domain D_2 of structure S_2 when superposed by domain D_0 is

$$\mathbf{R}'_{S_2D_0}(\bar{\mathbf{x}}_{S_2D_2} - \bar{\mathbf{x}}_{S_2D_0}) - \mathbf{R}'_{S_1D_0}(\bar{\mathbf{x}}_{S_1D_1} - \bar{\mathbf{x}}_{S_1D_0}), \quad (45)$$

in which $\mathbf{R}'$ is the 3×3 rotation matrix given by Equation 32 using the relevant primed rotation vector of Equation 44. By setting $D_1 = D_0$ the distribution of locations of the centroid of domain D_2 (over all structures) relative to the domain D_0 may be observed, and domain D_2 may be a single atom if it is required to track the displacement of individual atoms. The case $D_1 = D_2 \neq D_0$ allows relative displacements of domains to be studied, and the completely general case is also available.

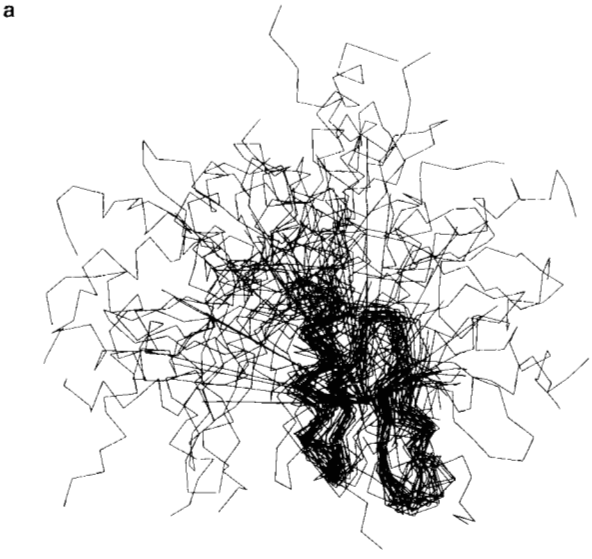
Also in the coordinate system of the output, the rotation

$$\boldsymbol{\rho}' = [\boldsymbol{\rho}'_{S_2D_0}][\bar{\boldsymbol{\rho}}'_{S_2D_1}][\boldsymbol{\rho}'_{S_1D_1}]\,\bar{\boldsymbol{\rho}}'_{S_1D_0} \quad (46)$$

rotates domain D_1 of structure S_1 to match domain D_1 of structure S_2 when all structures are superposed by reference to domain D_0, so that relative changes of orientation may also be observed. The evaluation of Expressions 45 and 46 is controlled by a dialogue in which values of D_0, etc., are provided.

Application

The method has been applied to two protein structural problems of interest. The first of these concerns a fragment from a yeast transcriptional activator protein, SWI5, that has been extensively studied in this laboratory. The 70-residue, 1,173-atom fragment contains two so-called zinc-finger motifs, the first of which spans approximately residues 1–37, and the second of which spans approximately residues 42–66. Each of these motifs represents a relatively rigid structural unit based around a tetrahedrally coordinated zinc ion, but the linker between motifs (residues 38–41) and the C-terminal tail (residues 67–70) are apparently flexible in solution. These properties are reflected in the results of a set of 29 simulated annealing calculations based on NMR-derived distance and dihedral angle constraints (Neuhaus et al., 1992), making these structures a useful test bed for the superposition methods developed in this paper. If superposition of the structures is restricted to the region of either one of the zinc-finger motifs, where the structures are similar to one another, convergence requires only three cycles, the last of which merely establishes that there is no further progress to be made. These superpositions, based on Cα atoms alone, are illustrated in Figure 2, which includes



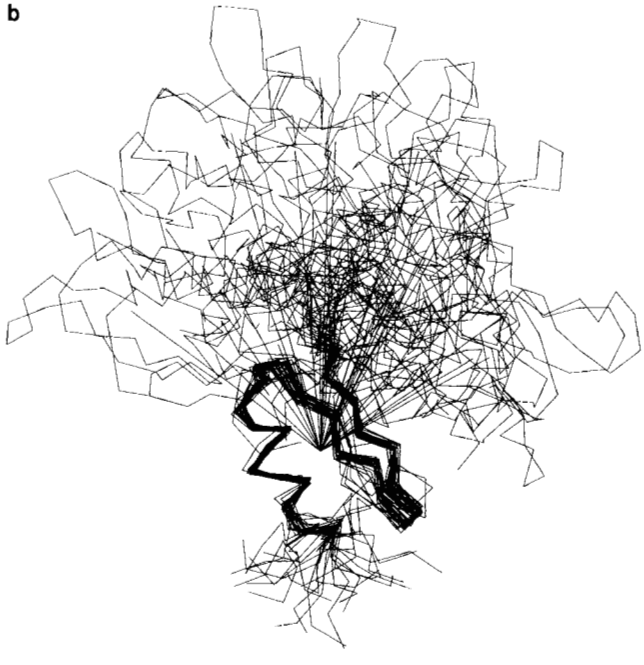


Fig. 2. Superposition (by Cα atoms only) of 29 SWI5 structures derived from NMR, drawn as Cα chain traces. In **a** the superposition is by residues 1–37, and in **b** by residues 42–66, the residuals R_0, R_1, and R_2 being 3.996, 4.024, and 2.796 Å for a and 1.333, 1.333, and 0.926 Å for b.

Fig. 3. The first 37 residues of two of the structures included in Figure 2a, whose enantiomorphous relationship was detected by Equation 25. For description see text.

also vectors from the (common) centroid of the superimposed domain to the individual centroids of the other domain, which serves to show the great variability of this vector, but also that it is confined to rather less than one hemisphere of possible directions. Equation 25 showed that two of the structures in Figure 2a were enantiomorphous to the first structure in the list, but chirality reversal was not permitted in making the superposition, so that these two structures had rms differences from the remaining 28 structures of 7.418 and 7.026 Å, respectively. One such pair is illustrated in Figure 3, which shows, in the lower part of the picture, two α helices, both of them right handed and consisting of L amino acids. The near end of the right helix then connects through the strand on the left of the figure to its N-terminus, which is in the distance at the top right of the picture. The near end of the left helix similarly connects through the strand on the right to its N-terminus in the distance at the top left. Clearly, a complete reflection from left to right of one of these structures would have permitted a closer fit, at the expense of including a left-handed helix and D amino acids. Difference vectors between helices of opposite hand need not exceed one helix diameter, which is much less than the differences found here. The essential difference between this pair of structures is that the gross topology of the β-sheet region of the structure (upper part of Figure 3) is inverted relative to the other. Both structures consist entirely of L amino acids, as a consequence of the chirality constraints active during the simulated annealing calculations used to generate them, and consequently the α-helix in both structures is right handed. The local inversion of the β-sheet in this case arises because the NMR constraint list from which the structures were calculated is, for technical reasons, deficient in entries relevant to the relative disposition of the sheet and the helix. Such problems of local mirror images in sparsely constrained regions of a structure determined by NMR are widespread (Pastore et al., 1991).

No such cases were found in the second domain, Figure 2b, where the corresponding NMR constraint list contained more entries restricting the relative disposition of the helix and the sheet.

Superpositions based on the C-terminal tail alone, where the conformations are essentially random, have never required more than six cycles to converge. Superposition over all 70 residues is a most demanding test with this protein, because the random *relative* orientations of

the two individually rigid zinc-finger motifs result in very large residuals and a shallow minimum. Consequently, R_0 is 10.185 Å and R_1 is 10.736 Å in these cases, and convergence is correspondingly slow, requiring up to nine cycles, although the result is stable and completely independent of which structure serves as structure 1. Some 45 additional optimizations were also done in each of which 2 structures out of the 10 for which $(p_1 - p_2)$ is least were initialized back to front, all of which led to the same minimum, suggesting that the minimum found is unique. Many enantiomorphous relationships were detected (but not altered) when all 70 residues were superposed together.

The method has also been applied to the same five structures that were studied by Shapiro et al. (1992). The methods of the Multiple solutions and the global minimum section were used with $l = 1$, $u = t = 4$, so that all possible combinations of turned structures were used in the initialization. Every case produced the same result with $R_0 = 2.0466$ Å, $R_1 = 2.0470$ Å, $R_2 = 1.2947$ Å, which supports the conclusion of Shapiro et al. (1992) that no alternative solution exists.

Acknowledgments

I am indebted to Dr. D. Neuhaus for his interest and support and for the coordinates of the SW15 structures, to Dr. A.M. Lesk for a preprint of his paper with Shapiro et al. (1992), and to Dr. A. Pastore who kindly supplied the coordinates used in their study.

References

Diamond, R. (1976). On the comparison of conformations using linear and quadratic transformations. *Acta Crystallogr. A32*, 1-10.

Diamond, R. (1988). A note on the rotational superposition problem. *Acta Crystallogr. A44*, 211-216.

Diamond, R. (1990). Chirality in rotational superposition. *Acta Crystallogr. A46*, 423.

Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. A32*, 922-923.

Kabsch, W. (1978). A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. A34*, 827-828.

Kearsley, S.K. (1989). On the orthogonal transformation used for structural comparisons. *Acta Crystallogr. A45*, 208-210.

Kearsley, S.K. (1990). An algorithm for the simultaneous superposition of a structural series. *J. Comput. Chem. 11*(10), 1187-1192.

Lesk, A.M. (1986). A toolkit for computational molecular biology. II. On the optimal superposition of two sets of coordinates. *Acta Crystallogr. A42*, 110-113.

McLachlan, A.D. (1972). A mathematical procedure for superimposing atomic coordinates of proteins. *Acta Crystallogr. A28*, 656-657.

McLachlan, A.D. (1979). [An appendix to] Gene duplications in the structural evolution of chymotrypsin. *J. Mol. Biol. 128*, 49-79.

McLachlan, A.D. (1982). Rapid comparison of protein structures. *Acta Crystallogr. A38*, 871-873.

Neuhaus, D., Nakaseko, Y., Schwabe, J.W.R., & Klug, A. (1992). Solution structures of two zinc finger domains from SW15 obtained using two-dimensional ^{1}H NMR spectroscopy; a zinc finger structure with a third strand of β-sheet. *J. Mol. Biol.*, in press.

Ortega, J.M. (1960). On Sturm sequences for tridiagonal matrices. *J. Assoc. Comput. Machinery 7*, 260-263.

Pastore, A., Atkinson, A., Saudek, V., & Williams, R.J.P. (1991). Topological mirror images in protein structure computation: An underestimated problem. *Proteins Struc. Funct. Genet. 10*, 22-32.

Shapiro, A., Botha, J.D., Pastore, A., & Lesk, A.M. (1992). A method for multiple superposition of structures. *Acta Crystallogr. A48*, 11-14.

Wilkinson, J.H. (1958). The calculation of the eigenvectors of codiagonal matrices. *Comput. J. 1*, 90-96.

Wilkinson, J.H. (1960). Householder's method for the solution of the algebraic eigenproblem. *Comput. J. 3*, 23-27.