SUPERPOSITION OF MOLECULAR STRUCTURES USING QUATERNIONS

GERALD R. KNELLER

IBM France, 94-96, rue Réaumur, F-75002 Paris, France and Département de Biologie Cellulaire et Moléculaire SBPM CEA, CEN Saclay, F-91191 Gif-sur-Yvette, France

An algorithm is developed that finds the optimal orientation of a rigid molecular structure, represented by N reference sites, with respect to the same number of sites in an observed structure. The optimal orientation is found by minimizing the weighted sum of squared deviations of the rotated reference site positions from the observed site positions. The rotation is parametrized by a quaternion whose components, written as a column vector, are shown to be an eigenvector of a characteristic matrix which is defined in terms of the coordinate sets to be superimposed.

The presented algorithm is particularly useful with respect to the calculation of orientational correlations of molecular structures.

KEY WORDS: Rigid-body, quaternions, superposition of protein structures, orientational correlations of molecules.

1 INTRODUCTION

The simulation or analysis of molecular systems often requires the rotational fit of a given molecular reference structure, defined by a set of atomic coordinates, to a corresponding observed structure. Good examples are the investigation of rigid-body motions in proteins by the analysis of molecular dynamics (MD) simulations, the comparison of protein structures [1,2], and the removal of unwanted protein rotations in MD simulations or their analysis. The algorithm described in the following to solve such problems is an efficient general purpose algorithm that has a particular advantage for the study of orientational correlations: Given the observed site positions $\{\vec{r}_\alpha\}$, $\alpha = 1...N$, and the corresponding positions $\{\vec{r}_\alpha^{(0)}\}$ of the reference structure, both coordinate sets related to a suitable chosen rotation center, we define for each site an 'error'

$$\vec{\varepsilon}_\alpha(q) \doteq \mathbf{D}(q)\vec{r}_\alpha^{(0)} - \vec{r}_\alpha. \tag{1}$$

The rotation to be optimized is described by an orthogonal matrix $\mathbf{D}(q)$ which is parametrized in terms of four quaternion parameters (q_0, q_1, q_2, q_3), written in compact notation as $q^\mathsf{T} \equiv (q_0, \vec{q}^\mathsf{T})$ [3]:

$$\mathbf{D}(q) = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(-q_0q_3 + q_1q_2) & 2(q_0q_2 + q_1q_3) \\ 2(q_0q_3 + q_1q_2) & q_0^2 + q_2^2 - q_1^2 - q_3^2 & 2(-q_0q_1 + q_2q_3) \\ 2(-q_0q_2 + q_1q_3) & 2(q_0q_1 + q_2q_3) & q_0^2 + q_3^2 - q_1^2 - q_2^2 \end{pmatrix}. \tag{2}$$

For the notation of vectors we keep the following convention throughout this paper:

Arrows label three-dimensional column vectors, underscores label four-dimensional column vectors, and an upper 'T' stands for the transposition of a matrix or a vector.

The quaternion parameters have to satisfy the normalization constraint

$$q_0^2 + q_1^2 + q_2^2 + q_3^2 = \underline{q} \cdot \underline{q} = 1. \tag{3}$$

To find the optimal superposition of the coordinate sets $\{\vec{r}_\alpha\}$ and $\{\vec{r}_\alpha^{(0)}\}$ in a least-squares sense the following function has to be minimized with the constraint (3):

$$m(\underline{q}) = \sum_\alpha w_\alpha |\vec{\varepsilon}_\alpha(\underline{q})|^2. \tag{4}$$

The w_α are positive weights for each site.

If needed, the choice of the rotation centers can easily be included into the above minimization problem since the corresponding fit is decoupled from the rotational fit: With $\sum_\alpha w_\alpha = 1$ and the coordinate sets $\{\vec{x}_\alpha\}$ and $\{\vec{x}_\alpha^{(0)}\}$, $\alpha = 1...N$, representing the observed structure and the reference structure, the optimal rotation centers are found to be

$$\left.\begin{aligned} \vec{x}_c &= \sum_\alpha w_\alpha \vec{x}_\alpha, \\ \vec{x}_c^{(0)} &= \sum_\alpha w_\alpha \vec{x}_\alpha^{(0)} \end{aligned}\right\}. \tag{5}$$

The coordinate sets for the following rotational fit are then given by

$$\left.\begin{aligned} \vec{r}_\alpha &= \vec{x}_\alpha - \vec{x}_c, \\ \vec{r}_\alpha^{(0)} &= \vec{x}_\alpha^{(0)} - \vec{x}_c^{(0)} \end{aligned}\right\}. \tag{6}$$

It is convenient to use quaternion parameters for the rotational minimization problem not only because their usage leads to an algorithm that is numerically easy to handle and efficient, as we will see, but also because they allow to express easily *relative* orientations between molecules or parts of them. The second property is particularly useful if one wants to study orientational correlations [4,5]. A comprehensive treatise on quaternions can be found in [3] and [6].

The quaternion parameters are related in the following way to the familiar Euler angles:

$$\left.\begin{aligned} q_0(\alpha,\beta,\gamma) &= \cos\left(\frac{\beta}{2}\right)\cos\left(\frac{\gamma + \alpha}{2}\right) \\[2mm] q_1(\alpha,\beta,\gamma) &= \sin\left(\frac{\beta}{2}\right)\sin\left(\frac{\gamma - \alpha}{2}\right) \\[2mm] q_2(\alpha,\beta,\gamma) &= \sin\left(\frac{\beta}{2}\right)\cos\left(\frac{\gamma - \alpha}{2}\right) \\[2mm] q_3(\alpha,\beta,\gamma) &= \cos\left(\frac{\beta}{2}\right)\sin\left(\frac{\gamma + \alpha}{2}\right) \end{aligned}\right\}. \tag{7}$$

The convention for the Euler-angles used here is the 'y-convention' where the rotation is defined by successive rotations about the z,y',z''-axis with angles α,β,γ. The quaternion parameters can be as well expressed in terms of the rotation axis $\vec{n}$ and the

rotation angle Φ of a rotation:

$$\left.\begin{array}{rcl} q_0(\vec{n},\Phi) &=& \cos\ (\Phi/2), \\ q_1(\vec{n},\Phi) &=& \sin\ (\Phi/2)n_x, \\ q_2(\vec{n},\Phi) &=& \sin\ (\Phi/2)n_y, \\ q_3(\vec{n},\Phi) &=& \sin\ (\Phi/2)n_z \end{array}\right\}. \tag{8}$$

2 FITTING THE QUATERNION PARAMETERS

2.1 Preliminary Definitions

The algorithm to find the quaternion parameters that minimize $m(q)$ defined in (4), subject to the constraint $q \cdot q = 1$, is most conveniently derived in the framework of quaternion algebra. An arbitrary quaternion $\mathbf{A}$ is determined by four real parameters $(a_0, a_1, a_2, a_3) \equiv \underline{a}^T \equiv (a_0, \vec{a}^T)$ and is written as:

$$\mathbf{A} = a_0 \mathbf{1} + a_1 \mathbf{I} + a_2 \mathbf{J} + a_3 \mathbf{K}. \tag{9}$$

$$\mathbf{I}^2 = \mathbf{J}^2 = \mathbf{K}^2 = -1. \tag{10}$$

$$\mathbf{IJ} = \mathbf{K}, \text{ cycl.} \tag{11}$$

A possible matrix representation of $\mathbf{A}$ is [6]:

$$\mathbf{A} = \begin{pmatrix} a_0 & -a_1 & -a_2 & -a_3 \\ a_1 & a_0 & -a_3 & a_2 \\ a_2 & a_3 & a_0 & -a_1 \\ a_3 & -a_2 & a_1 & a_0 \end{pmatrix}. \tag{12}$$

From (12) we derive the following multiplication rule for the product $C = AB$ of two quaternions:

$$\left.\begin{array}{rcl} c_0 &=& a_0 b_0 - \vec{a} \cdot \vec{b}, \\ \vec{c} &=& a_0 \vec{b} + b_0 \vec{a} + \vec{a} \wedge \vec{b} \end{array}\right\}. \tag{13}$$

The symbol '$\wedge$' denotes the vector-product. An immediate consequence from (12) and (13) is the following formula for the scalar product $\underline{a} \cdot \underline{b}$:

$$\underline{a} \cdot \underline{b} = a_0 b_0 + \vec{a} \cdot \vec{b} = \frac{1}{4} \text{tr}\{\mathbf{A}^T \mathbf{B}\}. \tag{14}$$

With tr$\{...\}$ the trace is denoted. The most important feature of *normalized* quaternions is that they represent rotations: Introducing the quaternions $\mathbf{X} \Leftrightarrow \underline{x} \equiv (0, \vec{x})^T$ and $\mathbf{Q} \Leftrightarrow \underline{q} \equiv (q_0 \vec{q})^T$, with $q \cdot q = 1$, we have for the components of

$$\mathbf{X}' = \mathbf{QXQ}^T \tag{15}$$

the following relations:

$$x'_0 = 0, \left.\vphantom{\begin{matrix}a\\b\end{matrix}}\right\}$$
$$\vec{x'} = \mathbf{D}(q)\vec{x},$$

$$(16)$$

where $\mathbf{D}(q)$ is given by (2).

2.2 The Algorithm

Using eqs. (14), (15), and (16) of the last subsection we can write the sum of squared deviations $m(q)$ as

$$m(q) = \frac{1}{4} \sum_{\alpha} w_{\alpha} tr\{\mathbf{E}_{\alpha}^{T}(q)\mathbf{E}_{\alpha}(q)\}, \qquad (17)$$

where $\mathbf{E}_{\alpha}(q)$ is the quaternion representation of $\vec{\varepsilon}_{\alpha}(q)$ defined in (1):

$$\mathbf{E}_{\alpha}(q) = \mathbf{Q}\mathbf{X}_{\alpha}^{(0)}\mathbf{Q}^{T} - \mathbf{X}_{\alpha}. \qquad (18)$$

Instead of minimizing $m(q)$, with respect to q it is, more convenient to minimize the function $\tilde{m}(q)$, defined by

$$\tilde{m}(q) = \frac{1}{4} \sum_{\alpha} w_{\alpha} tr\{\tilde{\mathbf{E}}_{\alpha}^{T}(q)\tilde{\mathbf{E}}_{\alpha}(q)\}, \qquad (19)$$

$$\tilde{\mathbf{E}}_{\alpha}(q) = \mathbf{Q}\mathbf{X}_{\alpha}^{(0)} - \mathbf{X}_{\alpha}\mathbf{Q}. \qquad (20)$$

From equation (20) and the invariance of the trace under cyclic permutation of the arguments it follows immediately that

$$\tilde{m}(q) = m(q) \qquad \text{if } q \cdot q = 1. \qquad (21)$$

The advantage of using $\tilde{m}(q)$ instead of $m(q)$ is due to the fact that the quaternions $\tilde{\mathbf{E}}_{\alpha}(q)$ are *linear* in q, i.e. its component vectors $\tilde{\underline{\varepsilon}}_{\alpha}$ can be related to the component vector q of the quaternion $\mathbf{Q}$ by a matrix $\mathbf{K}_{\alpha}$ whose explicit form will be derived later:

$$\tilde{\underline{\varepsilon}}_{\alpha} = \mathbf{K}_{\alpha}q. \qquad (22)$$

Consequently $\tilde{m}(q)$ can be written as a quadratic form in q, regardless if q is normalized or not:

$$\tilde{m}(q) = q \cdot \mathbf{M}q, \qquad (23)$$

$$\mathbf{M} = \sum_{\alpha} w_{\alpha} \mathbf{K}_{\alpha}^{T} \mathbf{K}_{\alpha}. \qquad (24)$$

The general form of $\mathbf{M}$ given in (24) follows from eqs. (22), (19), and (14). Taking into account the side-constraint $q \cdot q = 1$ by the method of Lagrange multipliers we have to minimize the following extended function:

$$\tilde{m}'(q,\lambda) = \tilde{m}(q) - \lambda(q \cdot q - 1). \qquad (25)$$

The necessary condition for a minimum of $\tilde{m}'(q,\lambda)$ with respect to q and λ leads

immediately to the following eigenvector equation for q:

$$\mathbf{M}\underline{q} = \lambda\underline{q}, \tag{26}$$

$$\underline{q}\cdot\underline{q} = 1. \tag{27}$$

According to (24) $\mathbf{M}$ is a symmetric and positive semidefinite matrix and therefore all its eigenvalues are real and greater than or equal to zero. Since $m(\underline{q})$ is equal to $\tilde{m}(q)$ if $\underline{q}\cdot\underline{q} = 1$, we have for all normalized solutions $\underline{q}_i$, $i = 0...3$, of (26):

$$m(\underline{q}_i) = \lambda_i. \tag{28}$$

Therefore the normalized eigenvector corresponding to the smallest eigenvalue of $\mathbf{M}$ is the solution we are looking for.

It remains to find the explicit form of the matrix $\mathbf{M}$. To do this we start from eq. (24) which expresses $\mathbf{M}$ in terms of the matrices $\mathbf{K}_\alpha$ which are themselves defined by the relation (22). The explicit form of the matrices $\mathbf{K}_\alpha$ can be easily found from the definition of $\bar{\mathbf{E}}_\alpha\,(q)$ given in (20) and the multiplication rules for quaternions listed in (13). Introducing the vectors $\vec{d}_\alpha$ and $\vec{s}_\alpha$, defined by

$$\vec{d}_\alpha \doteq \vec{r}_\alpha - \vec{r}_\alpha^{(0)}, \tag{29}$$

$$\vec{s}_\alpha \doteq \vec{r}_\alpha + \vec{r}_\alpha^{(0)}, \tag{30}$$

and the antisymmetric matrix $\mathbf{W}_\alpha$,

$$\mathbf{W}_\alpha \doteq \begin{pmatrix} 0 & -s_{\alpha z} & s_{\alpha y} \\ s_{\alpha z} & 0 & -s_{\alpha x} \\ -s_{\alpha y} & s_{\alpha x} & 0 \end{pmatrix}, \tag{31}$$

$\mathbf{K}_\alpha$ can be written as

$$\mathbf{K}_\alpha = \begin{pmatrix} 0 & \vec{d}_\alpha^T \\ -\vec{d}_\alpha & -\mathbf{W}_\alpha \end{pmatrix}. \tag{32}$$

The explicit form of $\mathbf{M}$ can now be obtained from straight-forward insertion of the above expression for $\mathbf{K}_\alpha$ into eq. (24). Defining the quantities

$$\vec{u}_\alpha \doteq \vec{r}_\alpha \wedge \vec{r}_\alpha^{(0)}, \tag{33}$$

$$\mathbf{P}_\alpha \doteq \vec{r}_\alpha \otimes \vec{r}_\alpha^{(0)} + \vec{r}_\alpha^{(0)} \otimes \vec{r}_\alpha, \tag{34}$$

where the symbol '$\otimes$' stands for the dyadic product, we arrive at the following expression for $\mathbf{M}$:

$$\mathbf{M} = \sum_\alpha w_\alpha \mathbf{M}_\alpha, \tag{35}$$

$$M_\alpha = \begin{pmatrix} |\vec{d}_\alpha|^2 & 2\vec{u}_\alpha^T \\ 2\vec{u}_\alpha & |\vec{s}_\alpha|^2 \mathbf{1} - 2\mathbf{P}_\alpha \end{pmatrix}. \tag{36}$$

Here $\mathbf{1}$ denotes the three by three unit matrix.

If the number of sites N in the minimization problem is equal to one the matrix $\mathbf{M} \equiv \mathbf{M}_1$ has two twofold degenerate eigenvalues:

$$\lambda_{1,2} = |\vec{r}|^2 + |\vec{r}^{(0)}|^2 \mp 2|\vec{r}|\,|\vec{r}^{(0)}|. \tag{37}$$

The eivenvalue λ_1 is equal to the minimum of $m(q)$, which is achieved when the rotated reference vector $\vec{r}^{(0)}$ is *parallel* to $\vec{r}$, and λ_2 is equal to the maximum $m(q)$, when the rotated reference vector $\vec{r}^{(0)}$ and $\vec{r}$ are *antiparallel*. The degeneracy of the eigenvalues is due to the fact that the rotations which lead to the minimum or maximum in $m(q)$ are evidently not unique: If we choose a coordinate system in which $\vec{r}^{(0)}$ is parallel to the z-axis and express the active rotation of $\vec{r}^{(0)}$ by the Euler-angles α, β, and γ, the rotation is entirely determined by α and β, which are the azimuthal and polar angle of the vector $\vec{r}$. The third angle – γ – is redundant since the axis for the third rotation is (anti)parallel to $\vec{r}$. Correspondingly any linear combination

$$q = c_1 q_1^{(1,2)} + c_2 q_2^{(1,2)}, \quad \text{with } c_1^2 + c_2^2 = 1, \tag{38}$$

where $q_1^{(1,2)}$ and $q_2^{(1,2)}$ are orthonormal basis-vectors which span the two-dimensional eigenspaces belonging to λ_1 and λ_2 respectively, leads to the same $m(q)$. In this case the solution of our minimization problem is given by an arbitrary normalized linear combination of the eigenvectors $q_1^{(1)}$ and $q_2^{(1)}$ which are the basis of the eigenspace belonging to λ_1.

The above case can be immediately generalized for the fit of linear structures with a number of sites, $N > 1$: For linear structures the optimal rotation for each single site in the reference structure, and therefore for the whole reference structure, is the same. Consequently the two twofold degenerate eigenvalues read:

$$\lambda_{1,2} = \sum_{\alpha} w_{\alpha} \{|\vec{r}_{\alpha}|^2 + |\vec{r}_{\alpha}^{(0)}|^2 \mp 2|\vec{r}_{\alpha}| \, |\vec{r}_{\alpha}^{(0)}|\}. \tag{39}$$

3 CONCLUSION AND NUMERICAL CONSIDERATIONS

As was shown above the optimal rotational superposition of two molecular structures – regardless of the arrangement of the atoms – can be found by the calculation of the normalized eigenvector belonging to the smallest eigenvalue of the real and symmetric 4×4-matrix $\mathbf{M} = \sum_{\alpha} w_{\alpha} \mathbf{M}_{\alpha}$, defined in eqs. (35,36). The components of this eigenvector represent the quaternion parameters describing the optimal orientation of the reference structure with respect to the target structure. For the numerical solution of the eigenvalue problem the single components of the matrices $\mathbf{M}_{\alpha}$ are given below. Since the $\mathbf{M}_{\alpha}$ are symmetric matrices, it suffices to list – and to calculate – only the upper triangle $\mathbf{M}_{\alpha,i,j}$, $i \leqslant j$. With $\vec{r}_{\alpha} \equiv (x_{\alpha},y_{\alpha},z_{\alpha})^T$ and $\vec{r}_{\alpha}^{(0)} \equiv (x_{0\alpha},y_{0\alpha},z_{0\alpha})^T$ we have:

$$M_{\alpha,11} = x_{\alpha}^2 + y_{\alpha}^2 + z_{\alpha}^2 + x_{0\alpha}^2 + y_{0\alpha}^2 + z_{0\alpha}^2 - 2x_{\alpha}x_{0\alpha} - 2y_{\alpha}y_{0\alpha} - 2z_{\alpha}z_{0\alpha}$$

$$M_{\alpha,12} = 2(y_{\alpha}z_{0\alpha} - z_{\alpha}y_{0\alpha})$$

$$M_{\alpha,13} = 2(-x_{\alpha}z_{0\alpha} + z_{\alpha}x_{0\alpha})$$

$$M_{\alpha,14} = 2(x_{\alpha}y_{0\alpha} - y_{\alpha}x_{0\alpha})$$

$$M_{\alpha,22} = x_{\alpha}^2 + y_{\alpha}^2 + z_{\alpha}^2 + x_{0\alpha}^2 + y_{0\alpha}^2 + z_{0\alpha}^2 - 2x_{\alpha}x_{0\alpha} + 2y_{\alpha}y_{0\alpha} + 2z_{\alpha}z_{0\alpha}$$

$$M_{\alpha,23} = -2(x_{\alpha}y_{0\alpha} + y_{\alpha}x_{0\alpha})$$

$$M_{\alpha,24} = -2(x_{\alpha}z_{0\alpha} + z_{\alpha}x_{0\alpha})$$

$$M_{\alpha,33} = x_{\alpha}^2 + y_{\alpha}^2 + z_{\alpha}^2 + x_{0\alpha}^2 + y_{0\alpha}^2 + z_{0\alpha}^2 + 2x_{\alpha}x_{0\alpha} - 2y_{\alpha}y_{0\alpha} + 2z_{\alpha}z_{0\alpha}$$

$$M_{\alpha,34} = -2(y_{\alpha}z_{0\alpha} + z_{\alpha}y_{0\alpha})$$

$$M_{\alpha,44} = x_{\alpha}^2 + y_{\alpha}^2 + z_{\alpha}^2 + x_{0\alpha}^2 + y_{0\alpha}^2 + z_{0\alpha}^2 + 2x_{\alpha}x_{0\alpha} + 2y_{\alpha}y_{0\alpha} - 2z_{\alpha}z_{0\alpha} \tag{40}$$

The complete matrix M is obtained by a loop over all sites $\alpha = 1...N$, accumulating the matrices M_α, multiplied by the weights w_α. Because of its structure this loop vectorizes well on vector machines if the number of sites is sufficiently large.

To solve the eigenvector problem for M one could in principle use a routine that calculates by inserve iteration [7] only its smallest eigenvalue and the corresponding eigenvector. However, for a 4×4-matrix no gain in speed can be expected [7] since it is only an advantage to use inverse iteration if less than approximately 25% of the eigenvalues and eigenvectors are needed. Tests with the routes DSPEV (full solution) and DSPSV (inserve iteration) from the IBM ESSL-library [8] confirm this.

The superposition of two peptid conformations, each described by 310 atom positions, takes 2.1 ms on an IBM 3090/600 G computer in vector mode, and 15.0 ms in scalar mode. Approximately 0.4 ms of the above CPU-times are used to diagonalize the matrix M.

References

[1] A.D. McLachlan, 'A mathetical procedure for superimposing atomic coordinates of proteins', *Acta Cryst.*, **A28**, 656-7 (1972).

[2] W. Kabsch, 'A solution for the best rotation to relate two sets of vectors', *Acta Cryst.*, **A32**, 922-3 (1976).

[3] S.L. Altmann, 'Rotations, Quaternions, and Double Groups', *Clarendon Press*, Oxford (1986).

[4] R.M. Lynden-Bell and A.J. Stone, 'Reorientational correlation functions, quaternions and Wigner rotation matrices', *Molecular Simulation*, **3**, 271-81 (1989).

[5] G.R. Kneller and A. Geiger, 'A method to calculate the g-coefficients of the molecular pair correlation function from molecular dynamics simulations'. *Molecular Simulation*, **3**, 283-300 (1989).

[6] B. Artmann, 'The Concept of Number: From Quaternions to Monads and Topological Fields', edited and translated from German by H.B. Griffiths, *Ellis Horwood Limited*, Chichester, England (1988).

[7] W.H. Press, B.P. Flannery, S.A. Teukolski, W.T. Vetterling, 'Numerical Recipes, The Art of Scientific Programming', *Cambridge University Press*, Cambridge (1986).

[8] IBM Engineering and Scientific Subroutine Library, Guide and Reference, Rel. 4 (SC23-0184-4).