JOURNAL OF COMPUTATIONAL BIOLOGY
Volume 25, Number 1, 2018

Pp. 72–88
DOI: 10.1089/cmb.2017.0166

Symmetrical Parameterization of Rigid Body Transformations for Biomolecular Structures

JIN SEOB KIM and GREGORY S. CHIRIKJIAN

ABSTRACT

Assessing preferred relative rigid body position and orientation is important in the description of biomolecular structures (such as proteins) and their interactions. In this article, we extend and apply the "symmetrical parameterization," which we recently introduced in the kinematics community, to address problems in structural biology. We also review parameterization methods that are widely used in structural biology to describe relative rigid body motions (in particular, orientations) as a basis for comparison. The new symmetrical parameterization is useful in describing the relative biomolecular rigid body motions, where the parameters are symmetrical in the sense that the subunits of a complex biomolecular structure are described in the same way for the corresponding motion and its inverse. The properties of this new parameterization, singularity analysis, and inverse kinematics are also investigated in more detail. Finally, parameterization is applied to real biomolecular structures and a potential application to structure modeling of symmetric macromolecules to show the efficacy of the symmetrical parameterization in the field of computational structural biology.

Keywords: biomolecular structure, rigid body transformation, rotation, symmetrical parameterization.

1. INTRODUCTION

STRUCTURES OF LARGE BIOMOLECULAR COMPLEXES consist of several subunits of molecules, of which the movements are important in performing the functions in biological cells and tissues. Also relative rigid body motions are important in describing protein–protein interactions such as ligand–receptor binding, which in turn is essential in drug therapy for therapeutic purposes [see Gohlke and Klebe (2002); Kitchen et al. (2004) and references therein]. Earlier work on protein–protein associations treated each molecule as a spherical particle to determine the relative positions (Levinthal et al., 1975; Wodak and Janin, 1978). Hence, assessment of preferred relative rigid body position and orientation between subunits in large biomolecular complexes becomes crucial. It is often very useful to treat each subunit of the complex as a rigid body to describe the relative rigid body motions (i.e., positions and orientations). For that purpose, techniques from kinematics have been widely used to determine the motions between subunits treated as rigid bodies (Schuyler and Chirikjian, 2003; Kim et al., 2005), as well as describing the atomistic molecular conformations (Zhang and Kavraki, 2002). For example, axis-angle parameters and Euler angles have been popular in describing the relative orientations between subunits.

Department of Mechanical Engineering, Johns Hopkins University, Baltimore, Maryland.

In those parameterization methods, one needs to fix one molecule in the space (base frame) and describe the relative position and orientation of another molecule (moving frame) by using one of the methods. However, when one wants to switch the roles of base and moving frames (i.e., changing the view), then the conventional methods require a little bit of manipulation, which might be cumbersome in actual computations. Recently, a new method, called the symmetrical parameterization, has been developed where the corresponding parameters are free from changing the view between reference frames (Chirikjian, 2014; Kim and Chirikjian, 2015). The name ''symmetrical parameterization'' comes from the fact that the parameters are symmetrical in a given rigid body motion and its inverse. Hence, this new parameterization method has a high potential in computational structural biology, as well as in the robotic community.

This article builds on and is expanded from the prior work that was presented in the associated conference (Kim and Chirikjian, 2016). The organization of this article is as follows. First, we review the parameterization methods that have been popular in computational structural biology in Section 2. In particular, we emphasize the possibility of the symmetric description of a rigid body position and orientation. Then, in Section 3, we give an extended and updated analysis on the symmetrical parameterization. In particular, the updated definition and properties of the symmetrical parameterization are presented. In Section 4, the application of the symmetrical parameterization to relevant biological macromolecules is presented, where we describe how to extract corresponding symmetrical parameters, given the symmetric subunits, and discuss potential application of symmetrical parameterization to the calculation of energy functions, which is important in the process of structural prediction of symmetrical large biocomplexes. Finally, Section 5 concludes the article.

2. CLASSICAL PARAMETERIZATION ON THE ROTATIONS

Description of rotations and translations (i.e., orientations and positions) in terms of kinematics has become a standard in robotic and computational structural biology communities. A reader can refer to the in-depth classical treatment of rotations and rigid body motions (Angeles, 1988; Bottema and Roth, reprinted 1990; McCarthy, 1990; Selig, 2004, 2005; Chirikjian and Kyatkin, 2016). Here we review the classical methods to parameterize three-dimensional (3D) rotations, $SO(3)$, that have been popular in molecular modeling (i.e., axis-angle parameters and Euler angles). These two parameterizations have been widely used in that the physical interpretations of the parameters are relatively straightforward to understand. Here we emphasize three-parameter descriptions rather than those that require four parameters and a constraint such as unit quaternions and Euler parameters. As is well known, any rotation in 3D space can be described by a 3×3 real matrix R satisfying the conditions

$$RR^T = \mathbb{I}_3 \text{ and } \det R = +1$$

where $\mathbb{I}_3$ denotes the 3×3 identity matrix, the superscript T denotes the transpose, and ''det'' denotes the determinant of a matrix. We also express it as $R\in SO(3)$. Due to the first of the above two constraints, even though there are nine entries in R, there are only three degrees of freedom, while the second of the above constraints does not contribute to reducing this dimension further.

2.1. Euler angles

The Euler-angle parameterization starts from the consideration of counterclockwise rotations about the $x=x_1$, $y=x_2$, and $z=x_3$ axes of a given coordinate system by the angle θ, which are given as

$$R_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \tag{1}$$

$$R_2(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}, \tag{2}$$

and

$$R_3(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3}$$

With these three fundamental matrices, Euler angles are defined by three such noncommutative rotations consecutively applied. For example, the *ZXZ*, *ZYZ*, and *ZYX* Euler angle parameterizations are defined as

$$R_{ZXZ}(\alpha, \beta, \gamma) = R_3(\alpha)R_1(\beta)R_3(\gamma); \tag{4}$$

$$R_{ZYZ}(\alpha, \beta, \gamma) = R_3(\alpha)R_2(\beta)R_3(\gamma); \tag{5}$$

$$R_{ZYX}(\alpha, \beta, \gamma) = R_3(\alpha)R_2(\beta)R_1(\gamma). \tag{6}$$

Note that each of three fundamental rotations in Equations (1), (2), and (3) is symmetrical in the sense that $R_i^{-1}(\theta) = R_i^T(\theta) = R_i(-\theta)$ $(i = 1, 2, 3)$. To be more specific, the corresponding inversion operation is seen as $\theta \rightarrow -\theta$, which is a simple non-transcendental transformation, because a sign change is just an algebraic operation. Furthermore, if the range of the angle θ is taken as $-\pi \le \theta \le \pi$, this simple transformation holds as its current form, whereas if the range is taken as $0 \le \theta \le 2\pi$, then the inversion operation must be modified as $\theta \rightarrow 2\pi - \theta$. In both cases, the end points of the ranges ($-\pi$ and π, or 0 and 2π, respectively) are identical with each other because they correspond to the same rotation.

However, the *ZYX* Euler angles are not symmetrical since

$$(R_{ZYX}(\alpha, \beta, \gamma))^T = R_1(-\gamma)R_2(-\beta)R_3(-\alpha).$$

If one wants to express this in the form of R_{ZYX} $(\alpha', \beta', \gamma')$, then one finds that the resulting form, which requires a daunting task, would involve arc-trigonometric functions that are transcendental.

The most popular choices of Euler angles are the *ZXZ* and *ZYZ* angles. In both cases, the ranges of angles for these choices are $0 \le \alpha \le 2\pi$, $0 \le \beta \le \pi$, and $0 \le \gamma \le 2\pi$. One can find the relationships between the *ZXZ* and *ZYZ* Euler angles as follows. First when the *ZYZ* Euler angles are used, one can see that

$$\begin{aligned} &R_3(\alpha)R_2(\beta)R_3(\gamma) \\ &= R_3(\alpha)(R_3(\pi/2)R_1(\beta)R_3(-\pi/2))R_3(\gamma) \\ &= R_3(\alpha+\pi/2)R_1(\beta)R_3(-\pi/2+\gamma) \end{aligned}$$

hence, it follows that

$$R_{ZYZ}(\alpha, \beta, \gamma) = R_{ZXZ}(\alpha+\pi/2, \beta, \gamma-\pi/2).$$

Now let us investigate the possibility of the symmetric property in the inversion operation. One might think that the *ZXZ* Euler angles are symmetrical such as $(\alpha, \beta, \gamma) \rightarrow (-\gamma, -\beta, -\alpha)$ just because

$$(R_{ZXZ}(\alpha, \beta, \gamma))^T = R_3(-\gamma)R_1(-\beta)R_3(-\alpha).$$

However, it follows that the resulting angles are no longer in the correct range as mentioned earlier. Partial corrections in terms of matching the range can be done by changing $-\gamma \rightarrow 2\pi - \gamma$ and $-\alpha \rightarrow 2\pi - \alpha$. Yet, how to put $-\beta$ back into the correct range is not clear. Furthermore, as will be discussed later, this inversion formula turns out to be not consistent with the inverse kinematics of the corresponding Euler angles, which shows the complexity in terms of the inversion of this parameterization. Therefore, Euler angles do not have the symmetric property in terms of the corresponding inversion operation.

2.2. *Axis-angle and exponential coordinates*

The axis-angle parameterization results from the Rodrigues formula, which is written as

$$R(\theta, \mathbf{n}) = \mathbb{I}_3 + \sin\theta N + (1 - \cos\theta)N^2 \tag{7}$$

where $\mathbf{n} = [n_1, n_2, n_3]^T$ denotes the unit vector of the axis of rotation, and θ denotes the angle of rotation. The matrix N denotes the 3×3 skew-symmetric matrix corresponding to the vector **n,** defined as

$$N = \begin{pmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{pmatrix},$$

and the corresponding operation to construct N from $\mathbf{n}$ is denoted as $N = \hat{\mathbf{n}}$. Here the parameterization of $\mathbf{n}$ requires simply two coordinates because it lies in the unit sphere $\mathbb{S}^2$. Hence, in total, there are three parameters for the axis-angle parameterization. Sometimes, the notation rot($\mathbf{n}$, θ) is used to represent the axis-angle parameterization. This latter notation will be used in later sections.

Regarding the ranges of the parameters, there are several choices. One can choose $\mathbf{n}$ from the upper hemisphere and set $\theta \in [0, 2\pi]$. Note that 0 and 2π correspond to the same angle. Another choice is such that one chooses $\mathbf{n}$ from the whole unit sphere, $\mathbb{S}^2$, and set $\theta \in [0, \pi]$. In the latter case, (θ, $\mathbf{n}$) and ($-\theta$, $-\mathbf{n}$) correspond to the same rotation. Hereafter we use the latter choice, that is, $(\theta, n) \in [0, \pi] \times \mathbb{S}^2$. Note that it is well known that the axis-angle parameterization is equivalent to the matrix exponential

$$R(\theta, \mathbf{n}) = \exp(\theta N).$$

This parameterization is almost symmetric in the sense that given $(\theta, \ \mathbf{n}) \in [0, \pi] \times \mathbb{S}^2$, it follows that

$$(R(\theta, \mathbf{n}))^T = R(\theta, -\mathbf{n}).$$

However, an issue arises when we consider the range of the parameters for $\mathbf{n}$ Specifically, let $(\Phi, \Theta) \in [0, 2\pi] \times [0, \pi]$ denote two angles to parameterize **n.** Then, one can readily see that

$$\mathbf{n}(\Phi, \Theta) = \begin{pmatrix} \cos\Phi \sin\Theta \\ \sin\Phi \sin\Theta \\ \cos\Theta \end{pmatrix}.$$

Again the similar convention is applied such that $\Phi = 0$ and $\Phi = 2\pi$ always correspond to the same point, and all values of Φ map to the same point when $\Theta = 0$ or π. Then, one can find that

$$-\mathbf{n}(\Phi, \Theta) = \mathbf{n}(\Phi \pm \pi, \pi - \Theta)$$

which means that $(\theta, \Phi, \Theta) \rightarrow (\theta, \Phi \pm \pi, \pi - \Theta)$. This obviously raises a problem that the correct value in $\pm$signs must be checked to make sure that the parameters are in the correct range.

The above problem can be resolved when the Lie algebra $so(3)$ (a set of all 3×3 real skew-symmetric matrices) is used. To be more specific, let $\mathbf{x} = [x_1, x_2, x_3]^T$ be an arbitrary vector in $\mathbb{R}^3$. Then, the 3×3 skew-symmetric matrix associated with $\mathbf{x}$ (i.e., $X = \hat{\mathbf{x}}$) can be expressed as

$$X = \begin{pmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{pmatrix} \doteq \sum_{k=1}^{3} x_i E_i.$$

Here E_i ($i = 1, 2, 3$) denotes a skew-symmetric matrix corresponding to the i^{th} natural basis vector $\mathbf{e}_i$ for $\mathbb{R}^3$ (i.e., $E_i = \hat{\mathbf{e}}_i$). Note that the vectorization operator is defined as (Chirikjian, 2011; Chirikjian and Kyatkin, 2016)

$$X^{\vee} \doteq \mathrm{x}.$$

Then, as is well known in kinematics, $\exp X$ becomes a rotation matrix. These three parameters (x_1, x_2, x_3) form the exponential coordinates. In fact, $\exp X$ can be interpreted as the rotation around axis $\mathbf{x}/\parallel \mathbf{x} \parallel$ by the angle $\parallel \mathbf{x} \parallel$. In other words, $\mathbf{x}/\parallel \mathbf{x} \parallel$ and $\parallel \mathbf{x} \parallel$, respectively, correspond to $\mathbf{n}$ and θ. Then, Equation (7) can be written as

$$\exp X = \mathbb{I}_3 + \frac{\sin \parallel \mathbf{x} \parallel}{\parallel \mathbf{x} \parallel} X + \frac{1 - \cos \parallel \mathbf{x} \parallel}{\parallel \mathbf{x} \parallel^2} X^2 \tag{8}$$

where $X = \theta N$. This parameterization is very convenient for many applications, as it is well behaved over the open ball of radius defined by $\parallel \mathbf{x} \parallel < \pi$. Furthermore, it behaves well under the inversion as

$$(\exp X)^T = \exp(-X),$$

that is, the inversion operation is simply $\mathbf{x} \rightarrow -\mathbf{x}$, and all the values stay in the correct range. However, when $\parallel \mathbf{x} \parallel = \pi$ (or $\theta = \pi$), a singularity occurs where the inverse operation of exponentiation [the log operation defined as $X = \log(\exp X)$] becomes multivalued. This can be problematic when, for example, two molecules are oriented and positioned along their respective z axes, since in this case it can be the case that $\parallel \mathbf{x} \parallel \approx \pi$.

2.3. Singularity analysis

The determinant of the Jacobian matrix corresponding to the parameters of interest is important in assessing the quality of parameterization. It is well known [see Chirikjian and Kyatkin (2016)] that given rotation with any three parameters $\mathbf{p}=[p_1, p_2, p_3]^T$ of the form $R(\mathbf{p})=R(p_1, p_2, p_3)$, the ''right'' Jacobian matrix is computed as

$$J_r=\left[\left(R^T\frac{\partial R}{\partial p_1}\right)^\vee, \left(R^T\frac{\partial R}{\partial p_2}\right)^\vee, \left(R^T\frac{\partial R}{\partial p_3}\right)^\vee\right]$$

and the ''left'' Jacobian is

$$J_l=\left[\left(\frac{\partial R}{\partial p_1}R^T\right)^\vee, \left(\frac{\partial R}{\partial p_2}R^T\right)^\vee, \left(\frac{\partial R}{\partial p_3}R^T\right)^\vee\right].$$

These are related to each other as

$$J_l=RJ_r. \tag{9}$$

For the Euler angle parameterization, when we use the *ZXZ* Euler angles, then we have

$$J_r=\begin{pmatrix}\sin\beta\sin\gamma & \cos\gamma & 0\\ \sin\beta\cos\gamma & -\sin\gamma & 0\\ \cos\beta & 0 & 1\end{pmatrix} \tag{10}$$

and from Equation (9), it follows that

$$\det J_r=\det J_l=\sin\beta \tag{11}$$

which indicates that the singularity occurs when $\beta=0$ or π.

3. SYMMETRICAL PARAMETERIZATION ON THE ROTATIONS AND RIGID BODY MOTIONS

As we have seen in the previous sections, the classical parameterization methods do not possess the symmetric property in forward and inverse operations. In this section, a symmetrical parameterization is presented. The main purpose of a symmetrical parameterization is to parameterize rotations and rigid body motions in a ''symmetrical'' way, that is, given a rotation or a rigid body motion, the corresponding inverse looks the same way. Mathematically, given a 3×3 rotation matrix $R(\mathbf{p})$ or a 4×4 homogeneous matrix $H(\mathbf{p}')$, a symmetrical parameterization allows us to write $[\mathrm{R}(\mathbf{p}_1)]^T=\mathrm{R}(\mathbf{p}_2)$ and $[H(\mathbf{p}'_1)]^{-1}=H(\mathbf{p}'_2)$, where the relationships between $\mathbf{p}_1$ and $\mathbf{p}_2$ and between $\mathbf{p}'_1$ and $\mathbf{p}'_2$ involve simple symmetry operations on their arguments. Here we present the extended analysis on this symmetrical parameterization. See Figure 1 for graphical illustration. For detailed explanation, including the application of this parameterization to the planar rigid body motions, see Kim and Chirikjian (2015). The main advantage of this parameterization is that, due to its symmetrical property, it enables us to describe the interaction between subunits more efficiently (i.e., without need to fix one subunit as a base frame and the other as a moving frame).

3.1. Rotations in 3D

To define the symmetrical parameterization in $SO(3)$, we need to define a special form of the rotation matrix $\mathrm{R}(\mathbf{a}, \mathbf{b})$. This matrix defines the rotation of which the direction of rotation is $\mathrm{a}\times\mathrm{b}$, and transfers $\mathbf{a}$ to $\mathbf{b}$ such that $\mathrm{R}(\mathbf{a}, \mathbf{b})\ \mathbf{a}=\mathbf{b}$. The explicit form of this matrix is given as (Chirikjian and Kyatkin, 2016)

$$R(\mathbf{a}, \mathbf{b}) \doteq \exp\left(\theta_{ab}\cdot\frac{\widehat{\mathbf{a}\times\mathbf{b}}}{\|\mathbf{a}\times\mathbf{b}\|}\right) \tag{12}$$

$$=\mathbb{I}_3+\widehat{\mathbf{a}\times\mathbf{b}}+\frac{(1-\mathbf{a}\cdot\mathbf{b})}{\|\mathbf{a}\times\mathbf{b}\|^2}\left(\widehat{\mathbf{a}\times\mathbf{b}}\right)^2. \tag{13}$$

Note that θ_{ab} should be defined as the unique angle in the range $[0, \pi]$ such that

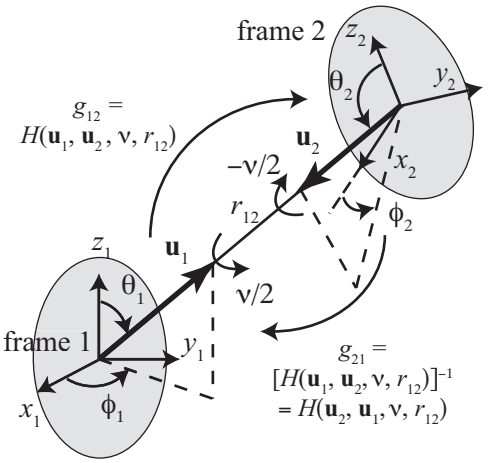

FIG. 1. Symmetrical parameterization for a three-dimensional full rigid body motion.

$$\sin\theta_{ab} = \| \mathbf{a}\times\mathbf{b} \| \text{ and } \cos\theta_{ab} = \mathbf{a}\cdot\mathbf{b}.$$

Some properties of the transference matrix $R(\mathbf{a}, \mathbf{b})$ can be found as

$$(R(\mathbf{a}, \mathbf{b}))^{-1} = R(\mathbf{b}, \mathbf{a}) = R(-\mathbf{b}, -\mathbf{a}) \tag{14}$$

and

$$R(\mathbf{a}, -\mathbf{b}) = R(-\mathbf{a}, \mathbf{b}) \tag{15}$$

which can be shown from Equation (13) and the elementary properties of the cross product. Another important property is

$$A\ R(\mathbf{a}, \mathbf{b})\ A^T = R(A\mathbf{a}, A\mathbf{b}) \tag{16}$$

where $A \in SO(3)$ is an arbitrary rotation, and $\mathrm{a}, \mathrm{b} \in \mathbb{S}^2$ are arbitrary unit vectors. This can be shown by applying Equation (13) and the classical kinematic equality

$$R(\mathbf{a}\times\mathbf{b}) = (R\mathbf{a})\times(R\mathbf{b}).$$

Also, the following expression holds

$$\mathrm{rot}(\mathbf{b}, \beta)\ R(\mathbf{a}, \mathbf{b})\ \mathrm{rot}(\mathbf{a}, \alpha)\mathbf{a} = \mathbf{b} \tag{17}$$

for any angles α and β, because rot $(\mathbf{a}, \alpha)$ $\mathbf{a} = \mathbf{a}$ for any rotation angle α.

With the information so far, our symmetrical parameterization for rotations is defined as

$$R_{12}(\mathbf{u}_1, \mathbf{u}_2, \nu) \doteq \mathrm{rot}(\mathbf{u}_1, \nu/2) R(\mathbf{u}_2, -\mathbf{u}_1)\mathrm{rot}(\mathbf{u}_2, -\nu/2). \tag{18}$$

As depicted in Figure 1, let $\mathbf{u}_1$ be the unit vector pointing from the origin of frame 1 to the origin of frame 2 (as viewed in frame 1). Also, let $\mathbf{u}_2$ be the unit vector pointing from the origin of frame 2 to the origin of frame 1 (as viewed in frame 2). Note the order and signs of $\mathbf{u}_1$ and $\mathbf{u}_2$ inside the transference rotation matrix in Equation (18), which will become clear in the later section on rigid body motions. R_{12} can be decomposed as

$$R_{12} = R_1 R_2 = R_1 (R_2^{-1})^{-1}$$

where

$$R_1 \doteq \mathrm{rot}(\mathbf{u}_1, \nu/2)[R(\mathbf{u}_2, -\mathbf{u}_1)]^{1/2},$$

and

$$R_2 \doteq [R(\mathbf{u}_2, -\mathbf{u}_1)]^{1/2}\mathrm{rot}(\mathbf{u}_2, -\nu/2).$$

Physically, R_1 represents the rotation half way starting at frame 1 and moving toward frame 2. And R_2^{-1} represents a rotation half way from frame 2 back to frame 1. Taking the inverse gives the expression of R_2

as shown above. Note that $\mathbf{u}_1$ and $\mathbf{u}_2$ should be treated as 3×1 arrays of numbers, not as ''physical'' vectors, with which we can perform cross products.

Now let us consider the inverse as

$$\begin{aligned} R_{12}^{-1} &= [\mathrm{rot}(\mathbf{u}_2, -\nu/2)]^{-1}[R(\mathbf{u}_2, -\mathbf{u}_1)]^{-1}[\mathrm{rot}(\mathbf{u}_1, \nu/2)]^{-1} \\ &= \mathrm{rot}(\mathbf{u}_2, \nu/2)R(\mathbf{u}_1, -\mathbf{u}_2)\mathrm{rot}(\mathbf{u}_1, -\nu/2) \\ &= R_{21}. \end{aligned} \tag{19}$$

In other words, inversion operation is simply written as

$$(\mathbf{u}_1, \mathbf{u}_2, \nu) \rightarrow (\mathbf{u}_2, \mathbf{u}_1, \nu),$$

or

$$(\theta_1, \phi_1, \theta_2, \phi_2, \nu) \rightarrow (\theta_2, \phi_2, \theta_1, \phi_1, \nu)$$

in its component form, which shows that parameterization is indeed symmetrical.

3.2. Rigid body motions in 3D

Suppose there are two rigid bodies, and reference frame 1 and 2 are attached to each body (Fig. 1). The position and orientation of the second body (frame 2) viewed from the first body (frame 1) are expressed as $g_{12}=(R_{12}, r_{12}\mathbf{u}_1)\in SE(3)$, of which the associated homogeneous transformation matrix is expressed as

$$H(g_{12}) = H(R_{12}, r_{12}\mathbf{u}_1) = \begin{pmatrix} R_{12} & r_{12}\mathbf{u}_1 \\ \mathbf{0}^T & 1 \end{pmatrix}.$$

Here the translation vector is described by the spherical coordinates (i.e., $\mathbf{r}=r\mathbf{u}$; $r\in\mathbb{R}_{\geq0}$ and $\mathbf{u}(\theta, \phi)\in\mathbb{S}^2$). In this case, symmetrical parameterization that we seek takes the form as

$$\begin{aligned} &H(\mathbf{u}_1, \mathbf{u}_2, \nu, r_{12}) = H(\theta_1, \phi_1, \theta_2, \phi_2, \nu, r_{12}) \\ &= \begin{pmatrix} R(\mathbf{u}(\theta_1, \phi_1), \mathbf{u}(\theta_2, \phi_2), \nu) & r_{12}\mathbf{u}(\theta_1, \phi_1) \\ \mathbf{0}^T & 1 \end{pmatrix} \end{aligned} \tag{20}$$

where we use the notations $\mathbf{u}_i=\mathbf{u}\ (\theta_i, \phi_i)$ $(i=1, 2)$. In Equation (20), we need only one additional radial parameter r_{12}, compared to the parameters for $SO(3)$, which denotes the distance between the origins of the two frames under consideration. This is due to the fact that the directional information on the translation is naturally enciphered in $\mathbf{u}_1$ and $\mathbf{u}_2$.

On the other hand, the position and orientation of frame 1 (the first body) from frame 2 (the second body) become $g_{21}=(R_{21}, r_{12}\mathbf{u}_2)$ by definition. Then, by using the fact that $g_{12}\circ g_{21}=g_{21}\circ g_{12}=e$ with e being the identity element in $SE(3)$, it should be true that

$$g_{21} = g_{12}^{-1} = (R_{12}^T, -r_{12}R_{12}^T\mathbf{u}_1).$$

Note that the vectors $\mathbf{u}_i$ $(i=1, 2)$ lie on the same line that connects the origins of frame 1 and 2, while their directions are opposite. Also, they are defined in their respective reference frames. From Equations (14) and (19), it can be shown that

$$\begin{aligned} -r_{12}R_{12}^T\mathbf{u}_1 &= -r_{12}\,\mathrm{rot}(\mathbf{u}_2, \nu/2)R(-\mathbf{u}_1, \mathbf{u}_2)\mathrm{rot}(\mathbf{u}_1, -\nu/2)\mathbf{u}_1 \\ &= -r_{12}\,\mathrm{rot}(\mathbf{u}_2, \nu/2)R(\mathbf{u}_1, -\mathbf{u}_2)\mathrm{rot}(\mathbf{u}_1, -\nu/2)\mathbf{u}_1 \\ &= -r_{12}\,\mathrm{rot}(u_2, \nu/2)R(\mathbf{u}_1, -\mathbf{u}_2)\mathbf{u}_1 \\ &= -r_{12}\,\mathrm{rot}(\mathbf{u}_2, \nu/2)(-\mathbf{u}_2) \\ &= r_{12}\,\mathbf{u}_2 \end{aligned} \tag{21}$$

from which one can see why the transference matrix in 3D rotation is defined in the form of $R(\mathbf{u}_2, -\mathbf{u}_1)$. Note that the unit vector $\mathbf{u}_1$ is no longer an arbitrary fixed vector, unlike as in the case of rotations in 3D space, because it is chosen as the direction of the translation for frame 1.

Finally, one can see that the inversion formula for these parameters is written as

$$(\mathbf{u}_1, \mathbf{u}_2, \nu, r_{12}) \rightarrow (\mathbf{u}_2, \mathbf{u}_1, \nu, r_{12}), \tag{22}$$

or in the component form,

$$(\theta_1, \phi_1, \theta_2, \phi_2, \nu, r_{12}) \rightarrow (\theta_2, \phi_2, \theta_1, \phi_1, \nu, r_{12}),$$

which shows that parameterization is indeed symmetrical.

3.3. Singularity analysis

In this section, we investigate the singularities of parameterization in Equation (18) when $\mathbf{u}_1$ is a fixed unit vector, and $\mathbf{u}_2 = \mathbf{u}$ (θ, ϕ) is a general unit vector. In particular, without loss of generality, we can set $\mathbf{u}_1 = \mathbf{e}_3 = [0, 0, 1]^T$ and $\mathbf{u}_2 = \mathbf{u} = [u_1, u_2, u_3]^T$, because we can use Equation (16) to expand the result to arbitrary $\mathbf{u}_1$. First, we compute the following terms:

$$R(\mathbf{u}, -\mathbf{e}_3) = \begin{pmatrix} 1 - \frac{u_1^2}{1-u_3} & -\frac{u_1 u_2}{1-u_3} & u_1 \\ -\frac{u_1 u_2}{1-u_3} & 1 - \frac{u_2^2}{1-u_3} & u_2 \\ -u_1 & -u_2 & -u_3 \end{pmatrix}, \tag{23}$$

$$\mathrm{rot}(\mathbf{e}_3, \nu/2) = \begin{pmatrix} \cos(\nu/2) & -\sin(\nu/2) & 0 \\ \sin(\nu/2) & \cos(\nu/2) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{24}$$

and

$$\mathrm{rot}\ (\mathbf{u}, \nu/2) = [\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3] \tag{25}$$

where

$$\mathbf{c}_1 = \begin{pmatrix} u_1^2 \mathrm{v}(\nu/2) + \mathrm{c}(\nu/2) \\ u_1 u_2 \mathrm{v}(\nu/2) + u_3 \mathrm{s}(\nu/2) \\ u_1 u_3 \mathrm{v}(\nu/2) - u_2 \mathrm{s}(\nu/2) \end{pmatrix},$$

$$\mathbf{c}_2 = \begin{pmatrix} u_2 u_1 \mathrm{v}(\nu/2) - u_3 \mathrm{s}(\nu/2) \\ u_2^2 \mathrm{v}(\nu/2) + \mathrm{c}(\nu/2) \\ u_2 u_3 \mathrm{v}(\nu/2) + u_1 \mathrm{s}(\nu/2) \end{pmatrix},$$

and

$$c_3 = \begin{pmatrix} u_3 u_1 \mathrm{v}(\nu/2) + u_2 \mathrm{s}(\nu/2) \\ u_3 u_2 \mathrm{v}(\nu/2) - u_1 \mathrm{s}(\nu/2) \\ u_3^2 \mathrm{v}(\nu/2) + \mathrm{c}(\nu/2) \end{pmatrix},$$

with $\mathrm{s}\theta = \sin\theta$, $\mathrm{c}\theta = \cos\theta$, and $\mathrm{v}\theta = 1 - \cos\theta$.

Then, when we consider $R(\mathbf{e}_3, \mathbf{u}(\theta, \phi), \nu)$ as in Equation (18), the corresponding Jacobian can be computed as

$$J(\theta, \phi, \nu) = \left[\left(R^T \frac{\partial R}{\partial \theta}\right)^\vee, \left(R^T \frac{\partial R}{\partial \phi}\right)^\vee, \left(R^T \frac{\partial R}{\partial \nu}\right)^\vee \right]$$
$$= \begin{pmatrix} \sin\phi & -\cos\phi \sin\theta & -\cos\phi \sin\theta \\ -\cos\phi & -\sin\phi \sin\theta & -\sin\phi \sin\theta \\ 0 & -1 - \cos\theta & -\cos\theta \end{pmatrix}.$$

Finally, it follows that

$$|\det J\ (\theta, \phi, \nu)| = \sin\theta.$$

Note that $|\det J\ |$ does not depend on φ and ν. This shows that as long as $\mathbf{u} \neq \pm\mathbf{e}_3$, the corresponding symmetrical parameters do not experience any singularity.

3.4. Inverse kinematics

Inverse kinematics is especially useful when one wants to extract the symmetrical parameters, given rigid body motions of biomolecules. Given a homogeneous transformation matrix

$$H = \begin{pmatrix} A & \mathbf{a} \\ \mathbf{0}^T & 1 \end{pmatrix}$$

($A \in SO(3)$ and $\mathbf{a} \in \mathbb{R}^3$), one can find the corresponding symmetrical parameters as follows. First, r_{12}, θ_1, and ϕ_1 can be obtained from considering $\mathbf{a} = [a_1\ a_2\ a_3]^T$ as

$$\begin{aligned} r_{12} &= \| \mathbf{a} \| = \sqrt{a_1^2 + a_2^2 + a_3^2} \\ \phi_1 &= \arctan2(a_1, a_2) \\ \theta_1 &= \arctan2\left(a_3, \sqrt{a_1^2 + a_2^2}\right). \end{aligned}$$

Then, by considering the fact that $\mathbf{u}_2 = -A^T\mathbf{u}_1 \doteq [u_1^{(2)}\ u_2^{(2)}\ u_3^{(2)}]^T$, θ_2 and ϕ_2 can be obtained as

$$\begin{aligned} \theta_2 &= \arctan2\left(u_3^{(2)}, \sqrt{(u_1^{(2)})^2 + (u_2^{(2)})^2}\right) \\ \phi_2 &= \arctan2\left(u_1^{(2)}, u_2^{(2)}\right). \end{aligned}$$

Next, the rotational part is expressed as

$$\begin{aligned} &(A - R(\mathbf{u}_2, -\mathbf{u}_1)) + \sin\frac{\nu}{2}(A\hat{\mathbf{u}}_2 - \hat{\mathbf{u}}_1 R(\mathbf{u}_2, -\mathbf{u}_1)) \\ &+ \left(1 - \cos\frac{\nu}{2}\right)\left(A\hat{\mathbf{u}}_2^2 - \hat{\mathbf{u}}_1^2 R(\mathbf{u}_2, -\mathbf{u}_1)\right) = 0_{3\times 3}. \end{aligned}$$

Let $T_1 = \text{trace}(A - R(\mathbf{u}_2, -\mathbf{u}_1))$, $T_2 = \text{trace}\left(A\hat{\mathbf{u}}_2 - \hat{\mathbf{u}}_1 R(\mathbf{u}_2, -\mathbf{u}_1)\right)$, and $T_3 = \text{trace}\left(A\hat{\mathbf{u}}_2^2 - \hat{\mathbf{u}}_1^2 R(\mathbf{u}_2, -\mathbf{u}_1)\right)$. Then, we obtain the following equation

$$T_1 + \sin\frac{\nu}{2}T_2 + \left(1 - \cos\frac{\nu}{2}\right)T_3 = 0.$$

Applying the half-tangent rule as

$$\begin{aligned} \sin\varphi &= \frac{2\tan\frac{\varphi}{2}}{1 + \tan^2\frac{\varphi}{2}} \\ \cos\varphi &= \frac{1 - \tan^2\frac{\varphi}{2}}{1 + \tan^2\frac{\varphi}{2}} \end{aligned}$$

enables us to obtain the following equation

$$(T_1 + 2T_3)\ \tan^2\frac{\nu}{4} + 2T_2\tan\frac{\nu}{4} + T_1 = 0,$$

from which the angular parameter is calculated as

$$\nu = 4\ \arctan\left(\frac{-T_2 \pm \sqrt{T_2^2 - T_1^2 - 2T_1T_3}}{T_1 + 2T_3}\right)$$

among which the value that falls into $(0, 2\pi)$ is chosen so that $\nu/2 \in [0, \pi)$ (as long as $T_2^2 - T_1^2 - 2T_1T_3 \geq 0$).

4. APPLICATION TO BIOMOLECULAR STRUCTURES

In this section, we discuss the application of symmetrical parameterization to biomolecular structures. In the first subsection, we consider three molecules, a chaperon (GroEL/GroES complex), the Prohead domain of the HK97 viral capsid, and the structure of an actin filament. In these examples, the corresponding symmetrical parameters are obtained by considering a pair of subunits in the symmetric complex structures, through which we compare the symmetrical parameterization with other methods. In the subsequent subsection, we discuss a potential application of symmetrical parameterization to the calculation of energy scoring functions, which is one of the important steps in the process of structural prediction of symmetrical large biocomplexes (André et al., 2007; DiMaio et al., 2011).

4.1. Extracting the symmetrical parameters

First, let us consider a chaperon (GroEL/GroES complex), which is a large biomolecular complex that guides proper protein folding. X-ray crystallography has revealed the structure of chaperonin complex: it

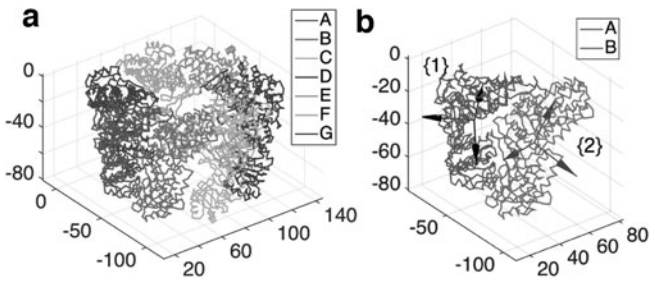

FIG. 2. **(a)** Seven subunits (chain A to G) that form a ring structure in GroEL; **(b)** two subunits (chain A and B) in **(a)** that are considered for parameterization.

consists of several subunits, including GroEL, GroES, and seven bound adenosine diphosphate (ADP) molecules [Protein Data Bank (PDB) entry: 1AON] (Xu et al., 1997). In particular, we consider a ring structure made by seven GroEL subunits (*cis*-GroEL). See Figure 2a for the structure of *cis*-GroEL.

Let us consider the rigid body motion from chain A to chain B. As shown in Figure 2b, frame 1 and 2 are attached to the center of mass of chain A and B, respectively. Frame 1 and 2 are obtained by the method of standardizing the coordinates (Kim et al., 2002). The relative rigid body motion from frame 1 to frame 2 is described by the following homogeneous transformation matrix

$$H = \begin{pmatrix} 0.9448 & 0.3267 & -0.0236 & 12.9677 \\ -0.2374 & 0.6333 & -0.7366 & -29.4770 \\ -0.2257 & 0.7015 & 0.6760 & 23.5122 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

For the comparison, we calculate the *ZXZ* Euler angles and axis-angle parameters for the rotation between frame 1 and 2. First, the *ZXZ* Euler angles are calculated by using the following formula. Given $R = (R_{ij}) \in SO(3)$ $(i, j = 1, 2, 3)$, angles are computed as

$$\beta = \arccos(R_{33}); \quad \alpha = \arctan2(-R_{23}, R_{13});$$
$$\gamma = \arctan2(R_{32}, R_{31})$$

which gives

$$\alpha = 358.16^\circ; \quad \beta = 47.47^\circ; \quad \gamma = 342.17^\circ.$$

For the inverse orientation, we obtain

$$\alpha' = 197.83^\circ; \quad \beta' = 47.47^\circ; \quad \gamma' = 181.84^\circ.$$

Hence, the *ZXZ* Euler angles are not symmetrical. Also, in comparison with a possible inversion formula as discussed earlier, this shows the complication of finding parameters between forward and inverse orientations. If we consider the axis-angle parameters, then we obtain

$$\theta = 51.17^\circ; \quad \Phi = 8.00^\circ; \quad \Theta = 111.23^\circ$$

for the orientation from frame 1 and frame 2, and

$$\theta' = 51.17^\circ; \quad \Phi' = 188.00^\circ; \quad \Theta' = 68.77^\circ$$

for the inverse orientation. Note that $\Theta + \Theta' = \pi$ and $\Phi' - \Phi = \pi$ in this case. However, in general, one needs to further check for Φ and Φ' as discussed earlier.

Now let us consider the symmetrical parameterization shown in Equations (18) and (20). By applying the inverse kinematics in Section 3.4, we obtain

$$\theta_1 = 53.87^\circ; \quad \phi_1 = 293.75^\circ;$$
$$\theta_2 = 159.30^\circ; \quad \phi_2 = 188.42^\circ;$$
$$r_{12} = 39.87\ [\text{Å}]; \quad \nu = 359.50^\circ.$$

Since these parameters are symmetrical, this set is complete in the sense that it also determines the inverse motion.

Next, we consider the structure of the HK97 viral capsid (Wikoff et al., 2000; Conway et al., 2001). We take the Prohead II domain (PDB entry: 1IF0), which consists of a hexamer (chain A to F) and a part of a

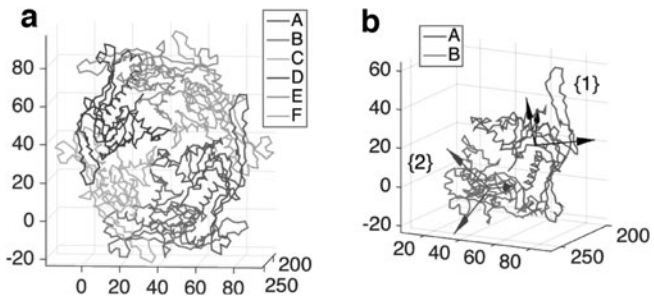

FIG. 3. **(a)** Six subunits (chain A to F) that form a hexamer assembly in the HK viral capsid Prohead II domain; **(b)** two subunits (chain A and B) in **(a)** that are considered for parameterization.

pentamer (chain G) (Conway et al., 2001). Here we consider a hexamer assembly as shown in Figure 3a. As in the previous example, we select chain A and B to describe the relative motion from A to B (Fig. 3b). The relative rigid body motion is described by the homogeneous transformation matrix as

$$H = \begin{pmatrix} 0.5267 & 0.7736 & -0.3523 & -18.2295 \\ 0.4851 & -0.6139 & -0.6228 & -25.2412 \\ -0.6980 & 0.1571 & -0.6986 & -8.0927 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The *ZXZ* Euler angles for the rotation are obtained as

$$\alpha = 330.50^\circ; \quad \beta = 134.32^\circ; \quad \gamma = 282.69^\circ.$$

For the inverse rotation, we obtain

$$\alpha' = 257.31^\circ; \quad \beta' = 134.32^\circ; \quad \gamma' = 209.50^\circ.$$

Also, the axis-angle parameters for the rotation are obtained as

$$\theta = 153.24^\circ; \quad \Phi = 23.91^\circ; \quad \Theta = 108.68^\circ,$$

and for the inverse rotation, we obtain

$$\theta' = 153.24^\circ; \quad \Phi' = 203.91^\circ; \quad \Theta' = 71.32^\circ,$$

which all together show the complexity of the parameters between forward and inverse rotations.

The symmetrical parameters are obtained as

$$\begin{array}{ll} \theta_1 = 104.57^\circ; & \phi_1 = 234.16^\circ; \\ \theta_2 = 149.77^\circ; & \phi_2 = 359.57^\circ; \\ r_{12} = 32.17\ [\mathring{A}]; & \nu = 216.98^\circ \end{array}$$

and these parameters are all that need to determine the inverse motion.

Finally, let us consider the structure of an actin filament (PDB entry: 3G37, see Fig. 4a), which has been revealed through the cryo-EM method (Murakami et al., 2010). Actin filaments are important in many cellular functions such as cell motility and maintenance of cell cortex. This is an example of helical symmetries where, unlike the previous two examples, a rigid body motion (i.e., rotation and translation) along a single axis is enough to represent the structure. Here we consider two subunits as in Figure 4b. The relative rigid body motion is described by the homogeneous transformation matrix as

$$H = \begin{pmatrix} 0.7980 & -0.4984 & -0.3388 & 26.6600 \\ -0.5962 & -0.7353 & -0.3225 & 24.7723 \\ -0.0884 & 0.4593 & -0.8839 & 17.5743 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The *ZXZ* Euler angles for the rotation are obtained as

$$\alpha = 313.58^\circ; \quad \beta = 152.11^\circ; \quad \gamma = 349.11^\circ.$$

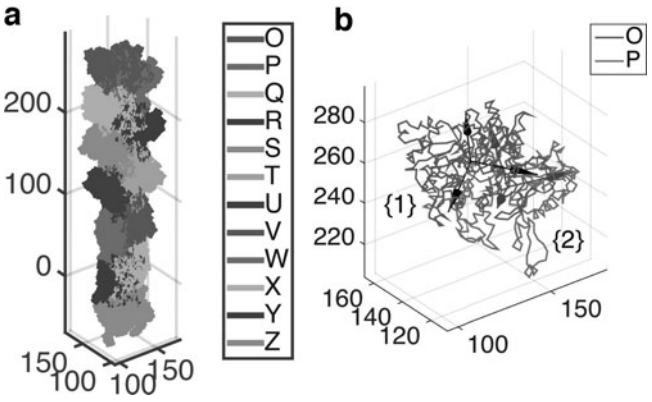

FIG. 4. **(a)** Twelve subunits (chain O to Z) that form a part of F-actin; **(b)** two subunits (chain O and P) in **(a)** that are considered for parameterization.

For the inverse rotation, we obtain

$$\alpha' = 190.89^\circ; \quad \beta' = 152.11^\circ; \quad \gamma' = 226.42^\circ.$$

Also the axis-angle parameters for the rotation are obtained as

$$\theta = 155.58^\circ; \quad \Phi = 342.24^\circ; \quad \Theta = 96.79^\circ,$$

and for the inverse rotation, we obtain

$$\theta' = 155.58^\circ; \quad \Phi' = 162.24^\circ; \quad \Theta' = 83.21^\circ,$$

which again all together show the complexity of calculating the parameters between forward and inverse rotations. However, the symmetrical parameters are obtained as

$$\begin{aligned} &\theta_1 = 64.22^\circ; \quad \phi_1 = 42.90^\circ; \\ &\theta_2 = 36.34^\circ; \quad \phi_2 = 101.94^\circ; \\ &r_{12} = 40.41\ [\mathring{A}]; \quad \nu = 121.55^\circ \end{aligned}$$

which again shows that this set of parameters is what we need to describe forward and inverse rigid body transformations between two subunits.

This as a whole emphasizes the usefulness of symmetrical parameterization in describing the structures of relative positions and orientations between subunits in complex biomolecules regardless of the symmetry type.

4.2. Calculating energy functions between symmetric subunits

In this section, we discuss a potential advantage of symmetrical parameterization on structure determination of symmetric large biomolecules.

In modeling symmetric large macromolecules, due to the structural symmetry, one does not need to consider the whole subunits, rather one subunit and the relative rigid body transformation between subunits are often enough. In practice, one subunit, called the ''master subunit,'' is often enough for the structure prediction (DiMaio et al., 2011). This master subunit is copied and pasted according to the symmetry to generate the entire symmetric complex structure. In doing so, apparently, attaching a reference frame at each copy of the master subunit has a big advantage. This is particularly true when considering interactions between subunits, which is important in predicting the structure of a large protein assembly. Instead of considering all the individual atom coordinates, one can use the parameters for the rigid body transformation between subunits in computing energy scoring functions between subunits. Another advantage of using a reference frame to each subunit is that it can facilitate the computation of the gradient of the energy function during structure prediction process than when to use the coordinates of a single subunit and the information on symmetry (André et al., 2007).

In calculating the energy function of a symmetric large complex, as well as interactions between atoms/residues inside each subunit, one also has to consider interactions between subunits. Due to the symmetry, however, one does not need to consider all the pairs. Rather pairs generated between the master subunit and the adjacent subunits (called the ''slave subunits''), which are copied from the master subunit, are usually enough (DiMaio et al., 2011). There are many kinds of energy functions included in the protein structure

prediction. For example, Rosetta, a software that is widely used in structure prediction, includes more than 15 energy scoring functions (Rohl et al., 2004). Many functions involve interactions between residues, secondary structural motifs, and so on. Hence, to explain how symmetrical parameterization contributes to the modeling of symmetric large macromolecules, let us consider a function of the distance between residues between subunits, which can represent the energy scoring functions such as Lennard-Jones potential energy, electric interactions, steric repulsion, and hydrogen bonding.

Let $f(r_{i,j}^{(m,1)}, r_{i,k}^{(m,2)})$ $(i, j, k \in \{1, 2, \cdots, N\}$, N being number of atoms/residues in one subunit) denote such a function, where

$$r_{i,j}^{(m,1)} = \| \mathbf{r}_i^{(m)} - \mathbf{r}_j^{(1)} \| \tag{26}$$

denotes the distance between the i-th residue in the master subunit ($\mathbf{r}_i^{(m)}$) and the j-th residue in the first copy of the master subunit ($\mathbf{r}_j^{(1)}$), and

$$r_{i,k}^{(m,2)} = \| \mathbf{r}_i^{(m)} - \mathbf{r}_k^{(2)} \| \tag{27}$$

denotes the distance between the i-th residue in the master subunit ($\mathbf{r}_i^{(m)}$) and the k-th residue in the second copy of the master subunit ($\mathbf{r}_k^{(2)}$). See Figure 5 for a graphical illustration of the current situation. For example, the Lennard-Jones potential between $\mathbf{r}_i^{(m)}$ and $\mathbf{r}_j^{(1)}$ can be of the form as

$$V_{LJ} = \varepsilon \left(\left(\frac{d_m}{r_{i,j}^{(m,1)}} \right)^{12} - 2 \left(\frac{d_m}{r_{i,j}^{(m,1)}} \right)^{6} \right) \tag{28}$$

where d_m denotes the distance at the minimum potential, and ε denotes the depth of the potential well. Note that the first and second copies are exactly the same as the master subunit, as explained earlier. Let $g_{m,1}$ denote the rigid body transformation between the master subunit and the first copy. We parameterize it as

$$g_{m,1} = H(\theta_1, \phi_1, \theta_2, \phi_2, r_{12}, \nu) \tag{29}$$

where H denotes the homogeneous transformation matrix of $g_{m,1}$ with given symmetrical parameters. Then, the j-th atom or residue in the copy 1 subunit from the reference frame attached to the master subunit can be simply computed as

$$\begin{pmatrix} \mathbf{r}_j^{(1)} \\ 1 \end{pmatrix} = H(\theta_1, \phi_1, \theta_2, \phi_2, r_{12}, \nu) \begin{pmatrix} \mathbf{r}_j^{(m)} \\ 1 \end{pmatrix} \tag{30}$$

or

$$\mathbf{r}_j^{(1)} = g_{m,1} \cdot \mathbf{r}_j^{(m)} \tag{31}$$

where $\cdot$ denotes the action of $SE(3)$ on $\mathbb{R}^3$ (i.e., given $g = (R, \mathbf{r}) \in SE(3)$, $g \cdot \mathbf{b} = R\mathbf{b} + \mathrm{r}$). In other words, the coordinates of the j-th residue in the first copy from the reference frame attached to the master subunit are determined solely by the symmetrical parameters.

Next, we want to compute the second term, $r_{i,k}^{(m,2)}$. To do that, we need the inverse rigid body transformation of $g_{m,1}$, because

$$\mathbf{r}_k^{(2)} = g_{m,2} \cdot \mathbf{r}_k^{(m)} = g_{m,1}^{-1} \cdot \mathbf{r}_k^{(m)}. \tag{32}$$

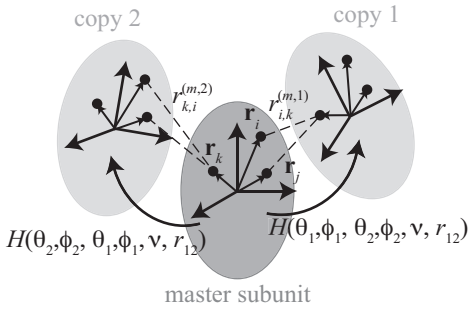

FIG. 5. A schematic that illustrates the concept of how to consider interactions between the master subunit and two adjacent slave subunits (copy 1 and 2).

If one utilizes other parameterizations (such as Euler angles for the orientation and the spherical coordinates for the translation), then as we have seen in the previous section, we have to calculate the corresponding parameters for the inverse transformation. However, when we use symmetrical parameterization, we can simply perform the following:

$$\begin{pmatrix} \mathbf{r}_k^{(2)} \\ 1 \end{pmatrix} = H(\theta_2, \phi_2, \theta_1, \phi_1, r_{12}, \nu) \begin{pmatrix} \mathbf{r}_k^{(m)} \\ 1 \end{pmatrix} \tag{33}$$

which means that we do not need to compute the corresponding parameters for the inverse rigid body transformation. Rather, we can still use the same set of symmetrical parameters for that purpose.

To illustrate this, Figure 6 shows two conformations of three subunits in GroEL (red and magenta, respectively, denote the master subunit and slave subunits), given two different sets of symmetrical parameters. The symmetrical parameters are as follows:

$$r_{12} = 49.87\mathring{A}; \quad \theta_1 = 53.87^\circ; \quad \phi_1 = 293.75^\circ;$$
$$\theta_2 = 159.30^\circ; \quad \phi_2 = 188.42^\circ; \quad \nu = 330.85^\circ$$

for Figure 6a, and

$$r_{12} = 39.87\mathring{A}; \quad \theta_1 = 42.41^\circ; \quad \phi_1 = 288.02^\circ;$$
$$\theta_2 = 170.76^\circ; \quad \phi_2 = 182.69^\circ; \quad \nu = 330.85^\circ$$

for Figure 6b.

For comparison, let us apply other parameterizations. For the rotation, we use the *ZXZ* Euler angles (α, β, and γ) and axis-angle parameters (θ, Φ, and Θ). For the translation, we use spherical coordinates (r, θ, and φ). First, let us consider the configuration in Figure 6a. The rigid body transformation from the master unit to copy 1 (i.e., $g_{m,1}$) has

$$\alpha = 4.42^\circ; \quad \beta = 38.38^\circ; \quad \gamma = 310.70^\circ \text{ (}ZXZ\text{ Euler angles)}$$
$$\theta = 58.40^\circ; \quad \Phi = 26.86^\circ; \quad \Theta = 137.64^\circ \text{ (axis}-\text{angle parameters)}$$
$$r = 49.87\mathring{A}; \quad \theta = 53.87^\circ; \quad \phi = 293.75^\circ \text{ (spherical coordinates).}$$

On the contrary, the inverse rigid body transformation (i.e., $g_{m,2}$) has

$$\alpha' = 229.30^\circ; \quad \beta' = 38.38^\circ; \quad \gamma' = 175.58^\circ \text{ (}ZXZ\text{ Euler angles)}$$
$$\theta' = 58.40^\circ; \quad \Phi' = 206.86^\circ; \quad \Theta' = 42.36^\circ \text{ (axis}-\text{angle parameters)}$$
$$r' = 49.87\mathring{A}; \quad \theta' = 159.30^\circ; \quad \phi' = 188.42^\circ \text{ (spherical coordinates).}$$

Next, let us consider Figure 6b. $g_{m,1}$ has the following parameters:

$$\alpha = 7.46^\circ; \quad \beta = 35.98^\circ; \quad \gamma = 317.60^\circ \text{ (}ZXZ\text{ Euler angles)}$$
$$\theta = 49.75^\circ; \quad \Phi = 24.93^\circ; \quad \Theta = 132.75^\circ \text{ (axis}-\text{angle parameters)}$$
$$r = 39.87\mathring{A}; \quad \theta = 42.41^\circ; \quad \phi = 288.02^\circ \text{ (spherical coordinates),}$$

while the corresponding inverse $g_{m,2}$ has

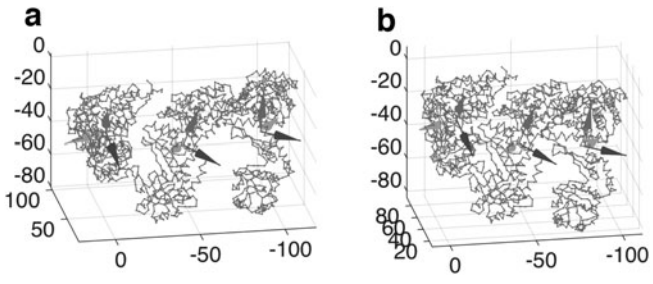

FIG. 6. Two conformations of three subunits in GroEL structure. The red-colored subunit denotes the master subunit, whereas the magenta-colored ones (copy 1 and 2) denote the slave subunits, which are copied and pasted from the master subunit by a rigid body transformation. **(a, b)** correspond to two different sets of symmetrical parameters as explained in the main text, which illustrates the efficiency of symmetrical parameterization on the construction of symmetrical subunits.

$$\alpha' = 222.40^\circ; \quad \beta' = 35.98^\circ; \quad \gamma' = 172.54^\circ \text{ (}ZXZ\text{ Euler angles)}$$
$$\theta' = 49.75^\circ; \quad \Phi' = 204.93^\circ; \quad \Theta' = 47.25^\circ \text{ (axis} - \text{angle parameters)}$$
$$r' = 39.87\text{Å}; \quad \theta' = 170.76^\circ; \quad \phi' = 182.69^\circ \text{ (spherical coordinates).}$$

This clearly illustrates that, given a set of symmetrical parameters, we can describe the conformation of three adjacent subunits very efficiently.

Finally, using the symmetrical parameterization, the energy function becomes the function of the symmetrical parameters, together with a set of atom/residue coordinates of the master subunit $\{\mathbf{r}_i^{(m)}\}$, written as

$$f = f(\theta_1, \phi_1, \theta_2, \phi_2, r_{12}, \nu). \tag{34}$$

For a detailed discussion, let us consider the Lennard-Jones potential energy again. Suppose there are N residues in the master subunit. Total potential energy from the interactions between the master subunit and slave subunits can be written, assuming that constant parameters (ε and d_m) are the same regardless of pairs of residues, as

$$f = \sum_{i,j=1}^{N} \varepsilon \left[\left(\frac{d_m}{r_{i,j}^{(m,1)}} \right)^{12} - 2 \left(\frac{d_m}{r_{i,j}^{(m,1)}} \right)^{6} \right] + \sum_{i,k=1}^{N} \varepsilon \left[\left(\frac{d_m}{r_{i,k}^{(m,2)}} \right)^{12} - 2 \left(\frac{d_m}{r_{i,k}^{(m,2)}} \right)^{6} \right]$$

where $r_{i,j}^{(m,1)}$ and $r_{i,k}^{(m,2)}$ are defined in Equations (26) and (27), respectively. The first summation term is from considering the interaction between the master subunit and its first copy subunit, while the second term is from the interaction between the master subunit and its second copy. As we have seen, $r_{i,j}^{(m,1)}$ includes the forward rigid body transformation $g_{m,1}$. When choosing the ZXZ Euler angles (α, β, γ) for the rotation and spherical coordinates (r, θ, φ) for the translation, then the first term becomes

$$\sum_{i,j=1}^{N} \varepsilon \left[\left(\frac{d_m}{r_{i,j}^{(m,1)}} \right)^{12} - 2 \left(\frac{d_m}{r_{i,j}^{(m,1)}} \right)^{6} \right] \doteq f^{(m,1)}(\alpha, \beta, \gamma; r, \theta, \phi).$$

On the contrary, to compute $r_{i,k}^{(m,2)}$, we need $g_{m,2}$, which is the inverse rigid body transformation of $g_{m,1}$. At this step, if one chooses one of the classical parameterization methods (e.g., the ZXZ Euler angles (α', β', γ') for the rotation and spherical coordinates (r', θ', ϕ') for the translation), then one can write

$$\sum_{i,k=1}^{N} \varepsilon \left[\left(\frac{d_m}{r_{i,k}^{(m,2)}} \right)^{12} - 2 \left(\frac{d_m}{r_{i,k}^{(m,2)}} \right)^{6} \right] \doteq f^{(m,2)}(\alpha', \beta', \gamma'; r', \theta', \phi'),$$

and hence one can write

$$f = f^{(m,1)}(\alpha, \beta, \gamma; r, \theta, \phi) + f^{(m,2)}(\alpha', \beta', \gamma'; r', \theta', \phi').$$

As we have seen in the previous examples, the relationships between the parameters of the forward and inverse rigid body transformations are complicated in any of the classical parameterization methods. However, if one uses the symmetrical parameterization method, then one can write

$$\begin{aligned} f &= f^{(m,1)}(\theta_1, \phi_1, \theta_2, \phi_2, r_{12}, \nu) + f^{(m,2)}(\theta_2, \phi_2, \theta_1, \phi_1, r_{12}, \nu) \\ &\doteq f(\theta_1, \phi_1, \theta_2, \phi_2, r_{12}, \nu). \end{aligned}$$

This is of a great advantage when it comes to computing energy scoring functions. All together, this demonstrates that using symmetrical parameterization has a high potential in the structure prediction of a large symmetric macromolecule.

5. CONCLUSIONS

In this article, we presented an extension of our previous analysis of symmetrical parameterization for rotations and rigid body motions in 3D space, and applied this parameterization to model spatial relationships between rigid subunits in biomolecular structures and complexes. To compare with current

alternatives, we presented a review of the most common parameterizations of 3D rotations (axis-angle parameters and Euler angles) used in the structural biology field. Our new parameterization has a symmetric property such that the inversion of a given rigid body rotation and translation is described by the same set of parameters used in describing the original motion. Singularity analysis for parameterization on 3D rotations was presented, and an inverse kinematic procedure for a given homogeneous transformation matrix followed accordingly. Finally, we applied symmetrical parameterization to a couple of biological macromolecules that consist of several subunits, and discussed its potential application to structure prediction of symmetric macromolecules, which emphasizes the efficacy of presented symmetrical parameterization on the structural description of large complex biological macromolecules.

ACKNOWLEDGMENTS

This work was supported by NSF grants (award numbers: CCF-1640970 and IIS-1619050) and the National Institute of General Medical Sciences of the NIH (award number: R01GM113240).

AUTHOR DISCLOSURE STATEMENT

No competing financial interests exist.

REFERENCES

André, I., Bradley, P., Wang, C., et al. 2007. Prediction of the structure of symmetrical protein assemblies. *Proc. Nat. Acad. Sci. U. S. A.* 104, 17656–17661.

Angeles, J. 1988. *Rational Kinematics*. Springer-Verlag, New York, New York.

Bottema, O., and Roth, B. reprinted 1990. *Theoretical Kinematics*. Dover Publications, Inc., New York, New York.

Chirikjian, G. 2011. *Stochastic Models, Information Theory, and Lie Groups: Volume 2—Analytic Methods and Modern Applications*. Birkhäuser.

Chirikjian, G. 2014. Rigid-body parameters for molecular docking applications. In *Proceedings of the ASME 2014 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference*. Buffalo, New York.

Chirikjian, G., and Kyatkin, A. 2016. *Harmonic Analysis for Engineers and Applied Scientists*. Dover, New York.

Conway, J.F., Wikoff, W.R., Cheng, N., et al. 2001. Virus maturation involving large subunit rotations and local refolding. *Science* 292, 744–748.

DiMaio, F., Leaver-Fay, A., Bradley, P., et al. 2011. Modeling symmetric macromolecular structures in Rosetta3. *PLoS One* 6, e20450.

Gohlke, H., and Klebe, G. 2002. Approaches to the description and prediction of the binding affinity of small-molecule ligands and macromolecular receptors. *Angew. Chem. Int. Ed.* 41, 2644–2676.

Kim, J.S., and Chirikjian, G.S. 2015. Principles of transference in theoretical kinematics. In *Proceedings of the ASME 2015 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Boston, Massachusetts.

Kim, J.S., and Chirikjian, G.S. 2016. Symmetrical rigid body parameterization for biomolecular structures. In *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. pp. 126–131. IEEE, Shenzhen, China.

Kim, M., Jernigan, R., and Chirikjian, G. 2005. Rigid-cluster models of conformational transitions in macromolecular machines and assemblies. *Biophys. J.* 89, 43–55.

Kim, M.K., Chirikjian, G.S., and Jernigan, R.L. 2002. Elastic models of conformational transitions in macromolecules. *J. Mol. Graph. Model.* 21, 151–160.

Kitchen, D.B., Decornez, H., Furr, J.R., et al. 2004. Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nat. Rev. Drug Discov.* 3, 935–949.

Levinthal, C., Wodak, S.J., Kahn, P., et al. 1975. Homoglobin interaction in sickle cell fibers I: Theoretical approaches to the molecular contacts. *Proc. Nat. Acad. Sci. U. S. A.* 72, 1330–1334.

McCarthy, J.M. 1990. *Introduction to Theoretical Kinematics*. MIT Press, Cambridge, Massachusetts.

Murakami, K., Yasunaga, T., Noguchi, T., et al. 2010. Structural basis for actin assembly, activation of atp hydrolysis, and delayed phosphate release. *Cell* 143, 275–287.

Rohl, C.A., Strauss, C.E.M., Misura, K., et al. 2004. Protein structure prediction using rosetta. *Methods Enzymol.* 383, 66–93.

Schuyler, A.D., and Chirikjian, G.S. 2003. Normal mode analysis of proteins: A comparison of rigid cluster modes with C_α coarse graining. *J. Mol. Graph. Model.* 22, 183–193.

Selig, J. 2004. Lie groups and Lie algebra in robotics, 101–125. In Byrnes, J., ed., *Computational Noncommutative Algebra and Applications*, volume 136. Springer; Dordrecht.

Selig, J.M. 2005. *Geometric Fundamentals of Robotics*. Springer, New York.

Wikoff, W.R., Liljas, L., Duda, R.L., et al. 2000. Topologically linked protein rings in the bacteriophage HK97 capsid. *Science* 289, 2129–2133.

Wodak, S.J., and Janin, J. 1978. Computer analysis of protein-protein interaction. *J. Mol. Biol.* 124, 323–342.

Xu, Z., Horwich, A.L., and Sigler, P.B. 1997. The crystal structure of the asymmetric GroEL-GroES-(ADP)7 chaperonin complex. *Nature* 388, 741–750.

Zhang, M., and Kavraki, L.E. 2002. A new method for fast and accurate derivation of molecular conformations. *J. Chem. Inf. Comput. Sci.* 42, 64–70.

Address correspondence to:
Prof. Gregory S. Chirikjian
Department of Mechanical Engineering
Johns Hopkins University
Baltimore, MD 21218

E-mail: gregc@jhu.edu