Curvilinear dynamics of protein complexes

K. O. E. Henriksson

Department of Physics
University of Helsinki
P.O.Box 43
FIN-00014 University of Helsinki
krister.henriksson@helsinki.fi

J. Pesonen

Department of Chemistry
University of Helsinki
P.O.Box 55
FIN-00014 University of Helsinki

P. Chacon

Centro de Investigaciones Biologicas
Ramiro de Maeztu 9
28040 Madrid
Spain

The suitability of a recently developed extension to the polyspherical parametrization of protein complexes has been investigated by molecular dynamics simulations. The extension makes it possible to study the motion of proteins consisting of several chains, which are internally covalently bonded. Tests indicate that time steps of up to at least 10 fs for simulations lasting up to 100 ps are possible, at least for small proteins.

Keywords: protein; dynamics; torsion; conformation

1. Introduction

In our previous studies [1,2] we have laid the theoretical foundations for simulation of single-chain proteins using internal coordinates. In this study we present an extension to that polyspherical approach (PA) so that molecules consisting of several chains can also be described. The extension allows for e.g. free rotation of the entire molecule or only some of the chains. In the extension of the polyspherical approach the relative distances and orientations of the chains have their own parameters. Additional parameters are provided for the total center of mass, and the orientation of the arbitrarily chosen first chain, relative to a fixed external frame (the laboratory). Each chain has its own body-fixed local frame, and internal coordinates such as bond lengths, bond angles, and torsion angles.

In the present study we implement the polyspherical approach to investigate the dynamics of a test protein containing three chains. As before, [2] the bond lengths and bond angles are considered fixed, allowing only torsion angles as well as the chain-related degree of freedom to change with time. The present approach is somewhat similar to the separation into subsystems described in the review by Gatti and Iung [3], but the formalism is Lagrangian, and especially adopted for very large complexes. The present study is not aimed at performing simulations using realistic potentials on dense proteins, but merely verifying the suitability of the parametrization for multi-chain protein dynamics. Therefore the test is limited to using a simple Morse potential. The protein motion is integrated with the standard velocity Verlet, using several different time steps. The conservation of the total energy and (angular) momentum is also investigated.

The general coordinate system presented here can be applied to multiple interesting problems. For example, we are actively working on normal mode analysis [4] and inverse kinematic problems, but other applications could also be possible (e.g. refinement in crystallography [5]).

2. Theory

2.1. *Curvilinear coordinates*

In the polyspherical approach [1,6] the molecule is located in an external fixed (laboratory) frame with the orthonormal basis $\{\mathbf{u}_k\}$, $k = 1, 2, 3$, which can be taken as the usual Cartesian basis. The molecule is divided into N_C chains, each with its own body-fixed frame $\{\mathbf{u}_k^{(I)}\}$. Each chain is internally parametrized with the bond lengths $r_i^{(I)}$, bond angles $\widetilde{\theta}_j^{(I)}$, and torsion angles $\phi_k^{(I)}$.

The frame $\{\mathbf{u}_1^{(I)}, \mathbf{u}_2^{(I)}, \mathbf{u}_3^{(I)}\}$, fixed to Ith chain ($I = 2, 3, \ldots, N_C$) is defined by

$$\mathbf{u}_3^{(I)} = \hat{\mathbf{r}}_1^{(I)} \tag{1}$$

$$\mathbf{u}_1^{(I)} = \frac{\hat{\mathbf{r}}_2^{(I)} + \cos\tilde{\theta}_{12}^{(I)}\hat{\mathbf{r}}_1^{(I)}}{\sin\tilde{\theta}_{12}^{(I)}} \tag{2}$$

$$\mathbf{u}_2^{(I)} = \mathbf{u}_3^{(I)} \times \mathbf{u}_1^{(I)} \tag{3}$$

using the definitions from Ref. [1]. This frame can be obtained from the frame $\{\mathbf{u}_1^{(1)}, \mathbf{u}_2^{(1)}, \mathbf{u}_3^{(1)}\}$ attached to the chain 1 by rotating it about the instantaneous rotation axis

$$\hat{\boldsymbol{\beta}}_{1I} = \sin\theta_{1I}\sin\varphi_{1I}\mathbf{u}_1^{(1)} - \sin\theta_{1I}\cos\varphi_{1I}\mathbf{u}_2^{(1)} + \cos\theta_{1I}\mathbf{u}_3^{(1)} \tag{4}$$

as

$$\mathbf{u}_j^{(I)} = \left(2\alpha_{1I}^2 - 1\right)\mathbf{u}_j^{(1)} + 2\alpha_{1I}\boldsymbol{\beta}_{1I} \times \mathbf{u}_j^{(1)} + 2\mathbf{u}_j^{(1)} \cdot \boldsymbol{\beta}_{1I}\boldsymbol{\beta}_{1I} \tag{5}$$

where

$$\alpha_{1I} = \cos\frac{\chi_{1I}}{2} \tag{6}$$

$$\boldsymbol{\beta}_{1I} = \hat{\boldsymbol{\beta}}_{1I}\sin\frac{\chi_{1I}}{2} \tag{7}$$

are the Euler scalar and vector, respectively. Note that the frame $\{\mathbf{u}_1^{(I)}, \mathbf{u}_2^{(I)}, \mathbf{u}_3^{(I)}\}$ is oriented relative to the frame $\{\mathbf{u}_1^{(1)}, \mathbf{u}_2^{(1)}, \mathbf{u}_3^{(1)}\}$ and not the laboratory frame $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$. This has the advantage that $\{\mathbf{u}_1^{(I)}, \mathbf{u}_2^{(I)}, \mathbf{u}_3^{(I)}\}$ depends on $\{\mathbf{u}_1^{(1)}, \mathbf{u}_2^{(1)}, \mathbf{u}_3^{(1)}\}$, but *not* vice versa.

Given the two frames $\{\mathbf{u}_1^{(I)}, \mathbf{u}_2^{(I)}, \mathbf{u}_3^{(I)}\}$ and $\{\mathbf{u}_1^{(1)}, \mathbf{u}_2^{(1)}, \mathbf{u}_3^{(1)}\}$ as an input, the Euler parameters can be obtained as

$$\alpha_{1I} = \frac{\gamma_{1I}}{\sqrt{\gamma_{1I}^2 + \boldsymbol{\kappa}_{1I}\cdot\boldsymbol{\kappa}_{1I}}} \tag{8}$$

$$\boldsymbol{\beta}_{1I} = -\frac{\boldsymbol{\kappa}_{1I}}{\sqrt{\gamma_{1I}^2 + \boldsymbol{\kappa}_{1I}\cdot\boldsymbol{\kappa}_{1I}}} \tag{9}$$

where

$$\gamma_{1I} = 1 + \sum_{k=1}^{3} \mathbf{u}_k^{(I)} \cdot \mathbf{u}_k^{(1)} \tag{10}$$

$$\boldsymbol{\kappa}_{1I} = \sum_{k=1}^{3} \mathbf{u}_k^{(I)} \times \mathbf{u}_k^{(1)} \tag{11}$$

When the axis of rotation $\hat{\boldsymbol{\beta}}_{1I}$ is known, the angles φ_{1I} and θ_{1I} are easily obtained as follows. Because $0 \le \theta_{1I} \le \pi$, we can obtain θ_{1I} as

$$\theta_{1I} = \arccos\left(\hat{\boldsymbol{\beta}}_{1I}\cdot\mathbf{u}_3^{(1)}\right) \tag{12}$$

If the value of θ_{1I} differs from zero, the cosines and sines of the angles φ_{1I} can be solved as

$$\sin\varphi_{1I} = \frac{\hat{\boldsymbol{\beta}}_{1I}\cdot\mathbf{u}_1^{(1)}}{\sin\theta_{1I}} \tag{13}$$

$$\cos\varphi_{1I} = -\frac{\hat{\boldsymbol{\beta}}_{1I}\cdot\mathbf{u}_2^{(1)}}{\sin\theta_{1I}} \tag{14}$$

Because $0 \le \varphi_{1I} < 2\pi$, the value of $\sin\varphi_{1I}$ is needed in order to recover the appropriate value of φ_{1I} from the given value of $\arccos(\cos\varphi_{1I})$. It is obtained as

$$\sin\varphi_{1I} \ge 0 \Rightarrow \varphi_{1I} = \arccos(\cos\varphi_{1I}) \in [0,\pi] \tag{15}$$

$$\sin\varphi_{1I} < 0 \Rightarrow \varphi_{1I} = 2\pi - \arccos(\cos\varphi_{1I}) \in (\pi, 2\pi) \tag{16}$$

where $[\alpha, \beta]$ stands for a closed, and (α, β) stands for an open interval from α to β. If $\theta_{1I} = 0$, then the value of φ_{1I} is inconsequential (any value will do). The angle of rotation $0 \leq \chi_{1I} < 2\pi$ can be easily recovered from the given Euler scalar as

$$\frac{\chi_{1I}}{2} = \arccos \alpha_{1I} \tag{17}$$

The angles θ_{1I}, χ_{1I}, and φ_{1I} are determined uniquely for almost every relative orientation of the frames $\{\mathbf{u}_1^{(I)}, \mathbf{u}_2^{(I)}, \mathbf{u}_3^{(I)}\}$ and $\{\mathbf{u}_1^{(1)}, \mathbf{u}_2^{(1)}, \mathbf{u}_3^{(1)}\}$, excluding the case that the axes $\mathbf{u}_3^{(I)}$ and $\mathbf{u}_3^{(1)}$ coincide, i.e., $\theta_{1I} = 0$. In that case φ_{1I} can be chosen freely.

The over-all orientation of the chain complex in space can be described by the orientation of the body-frame $\{\mathbf{u}_1^{(1)}, \mathbf{u}_2^{(1)}, \mathbf{u}_3^{(1)}\}$ of the first chain with respect to the laboratory frame $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$. The body frame is parametrized by the instantaneous rotation angle χ, and the two angles θ and φ that parametrize the orientation of the instantaneous rotation axis. The relevant formulas are presented in Ref. [1], using $\mathbf{u}_i^{(1)}$ instead of $\mathbf{u}_i'$.

The above definitions allow a description of the orientation of all the N_C body frames and the internal state of each chain I. However, the mutual distances between the body frames are still not accounted for. The missing degrees of freedom can be derived from the center of mass of the molecule,

$$\mathbf{X} = \sum_{I=1}^{N_C} \sum_{j=1}^{N_I} \frac{m_j^{(I)} \mathbf{x}_j^{(I)}}{M} = \sum_{I=1}^{N_C} \frac{M_I \mathbf{X}_I}{M}, \tag{18}$$

where $M = \sum_{I=1}^{N_C} \sum_{j=1}^{N_I} m_j^{(I)} = \sum_{I=1}^{N_C} M_I$. We simply chose the projection scalars

$$X^{(k)} = \mathbf{X} \cdot \mathbf{u}_k. \tag{19}$$

and

$$X_{1I}^{(k)} = (\mathbf{X}_I - \mathbf{X}_1) \cdot \mathbf{u}_k^{(1)}, \quad I \geq 2. \tag{20}$$

as the required degrees of freedom.

The coordinates defined above take full care of the positions and orientations of the chains and their atoms. The parameters $\{r_i^{(J)}\}$, $\{\widetilde{\theta}_i^{(J)}\}$, $\{\phi_{jk}^{(Ji)}\}$, $\{X_{1J}^{(k)}\}$, and $\{\theta_{1J}, \varphi_{1J}, \chi_{1J}\}$ can be said to be *shape coordinates*. In contrast, the two sets of parameters $\{X^{(k)}\}$ and $\{\theta, \varphi, \chi\}$ describe the location of the total center of mass, and the overall orientation of the molecule, respectively.

2.2. *Atomic positions*

By using the result derived in Ref. [1], it is obvious that the position vectors $\{\mathbf{x}_1^{(I)}, \mathbf{x}_2^{(I)}, \ldots, \mathbf{x}_{N_I}^{(I)}\}$ can be written as functions of the bond vectors $\{\mathbf{r}_1^{(I)}, \mathbf{r}_2^{(I)}, \ldots, \mathbf{r}_{N-1}^{(I)}\}$ and the center of mass vector $\mathbf{X}_I$ as

$$\mathbf{x}_i^{(I)} = \mathbf{X}_I + \mathbf{s}_i^{(I)} - \sum_{j=2}^{N_I} \frac{m_j^{(I)}}{M_I} \mathbf{s}_j^{(I)} \tag{21}$$

Now, $\mathbf{s}_j^{(I)} = \sum_k \mathbf{r}_k^{(I)}$ is the sum of the bond vectors, which form the chain connecting $\mathbf{x}_1^{(I)}$ and $\mathbf{x}_j^{(I)}$ (i.e., the tail of the first bond vector in the chain is on first atom, and the head of the last bond vector points to the jth atom), and $i = 1, 2, \ldots, N_I$. However, we still need to evaluate $\mathbf{X}_I$ in terms of the relative centers of mass $\{\mathbf{X}_{1I}\}$ and the total center of mass $\mathbf{X}$. First, the definition for $\mathbf{X}_{1I}$ can be written

$$\mathbf{X}_I = \mathbf{X}_1 + \mathbf{X}_{1I} \tag{22}$$

Second, we note that $\sum_{J=1}^{N_C} M_J \mathbf{X}_{1J}/M = \sum_{J=1}^{N_C} M_J \left(\mathbf{X}_J - \mathbf{X}_1\right)/M = \mathbf{X} - \mathbf{X}_1$, which gives

$$\mathbf{X}_1 = \mathbf{X} - \sum_{J=2}^{N_C} \frac{M_J \mathbf{X}_{1J}}{M}, \tag{23}$$

(note that $\mathbf{X}_{11} = 0$). Consequently,

$$\mathbf{X}_I = \mathbf{X} + \mathbf{X}_{1I} - \sum_{J=2}^{N_C} \frac{M_J \mathbf{X}_{1J}}{M}, \tag{24}$$

and only fundamental curvilinear coordinates are now present in Eq. (21).

2.3. *Equations of motion*

As before, ^{2}, the Lagrangian equations of motion are

$$\ddot{q}_p = -\sum_{ijk}^{A} g^{(pi)} \Gamma_{ijk} \dot{q}_j \dot{q}_k + \sum_{i}^{A} g^{(pi)} F_{(q),i}. \tag{25}$$

Here $g^{(pi)}$ is the inverse metric tensor, Γ_{ijk} is the Christoffel symbol, and $F_{(q),i}$ is the force acting on coordinate q_i. The metric tensor and the Christoffel symbol are defined as

$$g_{ij} = \sum_{I=1}^{N_c} \sum_{k}^{N_I} m_k^{(I)} \frac{\partial \mathbf{x}_k^{(I)}}{\partial q_i} \cdot \frac{\partial \mathbf{x}_k^{(I)}}{\partial q_j}, \tag{26}$$

$$\Gamma_{ijk} = \sum_{I=1}^{N_c} \sum_{k'}^{N_I} m_{k'}^{(I)} \frac{\partial \mathbf{x}_{k'}^{(I)}}{\partial q_i} \cdot \left(\frac{\partial^2 \mathbf{x}_{k'}^{(I)}}{\partial q_j \partial q_k} \right), \tag{27}$$

Here I denotes the chain, N_c is the total number of chains in the complex, N_I is the total number of atoms in chain I, $m_k^{(I)}$ is the mass of atom k in chain I, and $\mathbf{x}_k^{(I)}$ is the position of that atom. All the necessary derivatives are presented in Appendices A and B. The coordinates $\{q_i\}$ include the shape coordinates $\{r_k^{(I)}\}$, $\{\widetilde{\theta}_k^{(I)}\}$, $\{\phi_{jk}^{(Ii)}\}$, $\{X_{1I}^{(k)}\}$, $\{\theta_{1I}, \varphi_{1I}, \chi_{1I}\}$, as well as the overall positional and orientation coordinates $\{X^{(k)}\}$ and (θ, φ, χ), respectively.

The force acting on coordinate q_i is

$$F_{(q),i} = -\frac{\partial V}{\partial q_i} = -\sum_k \frac{\partial \mathbf{x}_k}{\partial q_i} \cdot \boldsymbol{\nabla}_{\mathbf{x}_k} V. \tag{28}$$

The evaluation of $\boldsymbol{\nabla}_{\mathbf{x}_k} V$ is conceptually straightforward, but can be laborious in practice.

In the present study the focus is on demonstrating the validity of the extension of the polyspherical parametrization to polymer complexes, hence a simple potential will be used. The choice is the Morse potential

$$\begin{aligned} V_{\mathrm{M}}^{(\mathrm{tot})} &= \sum_{ij}^{N} f_c(r_{ij}) V_{\mathrm{M}}(r_{ij}) \\ &= \sum_{ij}^{N} f_c(r_{ij}) D_e \bigg(\exp\big[-2a(r_{ij} - r_{ij}^{(0)})\big] - 2\exp\big[-a(r_{ij} - r_{ij}^{(0)})\big] \bigg), \end{aligned} \tag{29}$$

where $r_{ij} = |\mathbf{r}_{ij}| = |\mathbf{x}_j - \mathbf{x}_i|$ and $f_c(r) = 1$ when $r \le R - D$, $f_c(r) = 0$ when $r \ge R + D$, and

$$f_c(r) = \frac{1}{2} - \frac{1}{2}\sin\Big(\frac{\pi}{2D}(r - R)\Big) \tag{30}$$

otherwise. The potential cutoff function f_c guarantees that the potential energy contribution for pair i, j go towards zero when r_{ij} approaches $R + D$ from below.

The equilibrium distance between the atoms i and j is denoted $r_{ij}^{(0)}$. The Morse potential is more "natural" than the harmonic potential we have used earlier [2], since the potential energy contribution of far-way atoms automatically approach zero, and approach a large value when the atoms come close to one another.

The integration of Eq. (25) is performed using velocity Verlet [7]. The work flow for solving the equations of motion in torsion space is spelled out in Ref. [2].

2.4. *Importance of orientation coordinates*

The potential will depend only on the shape coordinates, and is invariant with respect to the changes in $\{\chi, \theta, \varphi\}$. However, the kinetic energy part contains terms where these coordinates are mixed with other coordinates. Hence, $\{\chi, \theta, \varphi\}$ *must* be included to the equations of motion. Especially, they are needed to ensure that the total angular momentum is conserved. To understand this, it suffices to write the body-frame components $l_i' = \mathbf{u}_i' \cdot \mathbf{l}$ of the angular momentum

$$\mathbf{l} = \sum_{I}^{N_c} \sum_{\alpha}^{N_I} m_\alpha^{(I)} \mathbf{y}_\alpha^{(I)} \times \dot{\mathbf{y}}_\alpha^{(I)}, \tag{31}$$

where $\mathbf{y}_\alpha^{(I)} = \mathbf{x}_\alpha^{(I)} - \mathbf{X}$, and $\mathbf{X}$ is the center of mass position of the entire complex, explicitly as [8]

$$l'_i = \frac{\partial L}{\partial \omega'_i} = \sum_{j=1}^{3} g'_{ij}\omega'_j + \sum_{j=1}^{A} \widetilde{g}'_{ij}\dot{s}_j, \tag{32}$$

where g'_{ij} denotes the element (ω'_i, ω'_j) and $\tilde{g}'_{ij}$ denotes the element (ω'_i, s_j) of the covariant metric tensor g'. It is related to the metric tensor g used in the present work by the transformation $g = \Omega g' \Omega^T$, where the superscript T stands for the transpose. The matrix Ω relates the rotational velocities $\omega'_i = \boldsymbol{\omega} \cdot u_i^{(1)}$ and the time derivatives of the rotational angles $\{\chi, \theta, \varphi\}$ (denoted now for compactness by $\{\Theta_j\}$) as $\omega'_i = \sum_{j=1}^{3} \Omega^T_{ij} \dot{\Theta}_j$. Although Ω is not needed explicitly here, it can be evaluated easily for the present rotational parametrization. It is given for the fixed Z-X-Z Euler angle parametrization, or the "Goldstein convention", in Eq. (34), p. 278 in Ref. [8]. Consider the case $\mathbf{l} = 0$. Because the values of $\{\dot{s}_j\}$ (and the element (ω'_i, s_j) of the metric tensor g) generally differ from zero, the total angular momentum does *not* vanish by setting the coordinates$\{\chi, \theta, \varphi\}$ to constant value (which would imply $\omega'_j = 0$ for $j = 1, 2, 3$). In fact, the *only* way to make l'_i $(i = 1, 2, 3)$ vanish identically is to let the coordinates $\{\chi, \theta, \varphi\}$ change in the manner that the value of the first term,

$$j'_i = \sum_{j=1}^{3} g'_{ij}\omega'_j, \tag{33}$$

(often referred as the rotational contribution to angular momentum) compensates the value of the second term,

$$k'_i = \sum_{j=1}^{A} \widetilde{g}'_{ij}\dot{s}_j, \tag{34}$$

which is often referred as the vibrational contribution to angular momentum. This, in turn, means that the coordinates $\{\chi, \theta, \varphi\}$ may change, and consequently, ω'_j may differ from zero, when the Lagrange's equation of motion are solved (note that $\mathbf{l}$ is an invariant for conservative systems). Note also that the force needed to constrain $\{\chi, \theta, \varphi\}$ to constant value can *not* be derived from a potential, which is solely a function of the shape coordinates. Hence, the required constraint force is not conservative, and therefore does not preserve angular momentum.

3. Results and discussion

3.1. *Static test of parametrization*

The polyspherical parametrization for multiple chain molecules has been tested by applying it to the photosynthetic enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (the version with PDB code 2VDH). This protein contains 15 chains, making a total of 37804 atoms. The internal degrees of freedom were calculated and

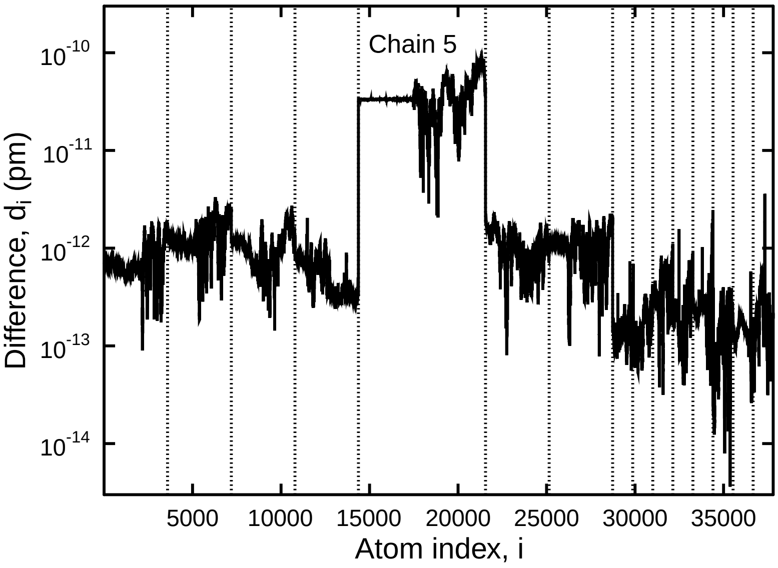

Figure 1. A plot of the difference d_i between read-in and reconstructed Cartesian atomic positions for the chain with the PDB code 2VDH. The vertical lines indicate that a new chain is started.

used to reconstruct the Cartesian atomic positions. A plot of the differences d_i ($i = 1, 2, \ldots, N$), given by the expression

$$d_i = |\mathbf{x}_i - \mathbf{y}_i| \tag{35}$$

where $\mathbf{y}_i$ denotes the reconstructed position of atom i, is shown in Fig. 1.

As can be seen in Fig. (1), the difference between original and reconstructed positions is always less than $\sim 10^{-10}$ pm in this particular case. The atomic positions themselves are given in Ångström, which equals 0.1 pm. Also seen from the figure, there are no cumulative round-off errors between the chains. This can be understood from Eq. (21), where it is clear that all positions inside a chain I are given as a sum over bonds in that chain only. Therefore round-off errors accumulate inside chains only, and not from one chain to the next chain, as can also be observed from the figure. From this relationship it is also clear that the longer the chain is the greater is the round-off error for positions in that chain, as compared to shorter chains. This fact is illustrated in Fig. (1) by the relatively larger error in atomic positions for chain 5, which contains roughly twice the number of residues as the preceding chains.

3.2. *Dynamic test of parametrization*

For performing molecular dynamics the code TOD used in our previous studies [1,2] was extended to handle dynamics of polymers with multiple chains. The simulated multi-chain protein was extracted from the single-chain protein oxymyoglobin (PDB code 1A6M). Residues 1-5, 10-15, and 20-24 were used. The nitrogen atoms in residues 6, 16, and 25 were kept and labelled as oxygen atoms in order to properly terminate the chains.

A maximum of three backbone angle types per residue, ϕ, ψ, χ_1, were considered free to move. These angles involve the bond sets (i) C-N-CA-C, (ii) N-CA-C-N, as well as (iii) N-CA-CB with CA connecting to N, respectively. Here CA is the alpha carbon atom in a residue and CB is the first carbon atom in the side-chain, if such a one is present. This gave a total of 13, 17, and 13 torsion angles in chain 1, 2, and 3, respectively. The initial values of the torsion angles were set as $\phi_i = \phi_i^{(\mathrm{eq})} + u\Delta\phi$, where $\phi_i^{(\mathrm{eq})}$ is the equilibrium value, u is a uniformly distributed random number in the interval $[-1, 1]$, and $\Delta\phi$ is an amplitude. The amplitude $\Delta\phi = 3$ (degrees) was used. In addition, all the coordinates describing the orientation of the chains and their distance from each other — namely θ, φ, χ, $\{\theta_{1J}, \varphi_{1J}, \chi_{1J}\}, J = 2, 3$, $\{X^{(k)}\}, k = 1, 2, 3$, $\{X_{1J}^{(k)}\}, J = 2, 3, k = 1, 2, 3$ — were used. These contributed a total of 18 degrees of freedom. The initial values of these coordinates were not modified. All initial velocities $\dot{q}_i$ were set to zero. No resetting or scaling was performed on any of the velocities during the simulation. It should be noted that no hydrogen atoms and no solvent was used in the simulations, *i.e.* the protein complex was considered to be in vacuum.

In the potential the parameter values $D_e = 10^{-4}$ eV and $a = 1.0$ Å were used. The D_e value sets the overall energy scale, which is not very important in the present study. A small value for this parameter makes the energy drift clearer. The equilibrium distances $r_{ij}^{(0)}$ in the potential was calculated from the conformation in the given PDB file. Cutoff function parameters of $R = 9.5$ Å and $D = 0.5$ Å were used. The system was simulated for 100 ps. It should be noted that the time is shorter than in our previous study for single-chain polymers [2], which already demonstrated the validity of the single-chain dynamics. Actually, in the present case the number of degrees of freedom is 61, which is five times as much as in the previous study, where the number of degrees of freedom was 12. This makes the round-off errors in the total energy and the coordinate values pile up faster. For this reason the simulation time was reduced.

The total energies for different values of the time step are shown in Fig. 2. Explicit measures for the energy drift are shown in Table 1. Denoting the initial total energy with E_{ini} and final total energy with E_{fin}, the energy drift is here defined as

$$\delta_E = (E_{\mathrm{fin}} - E_{\mathrm{ini}})/E_{\mathrm{ini}}, \tag{36}$$

As can be seen in the figure and the table, the energy drift wanders more and more away from zero when the time step is increased. This is a general problem for numerical integrators. Using a temperature control method, which e.g. scales the velocities to achieve a stable temperature, would alleviate the drift problems.

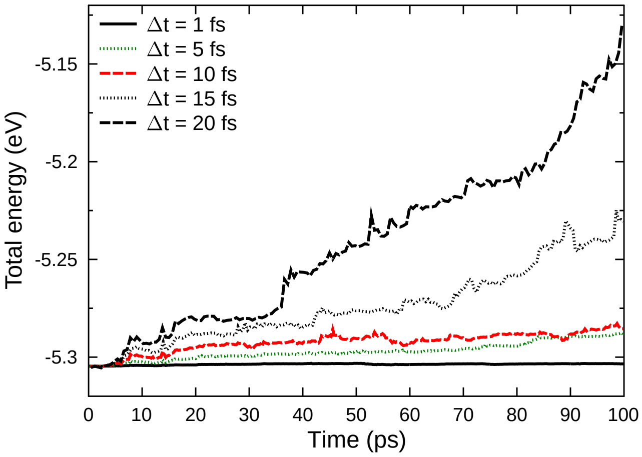

Figure 2. Total energy as a function of time and time step size. The system consisted of three chains with five, six, and five residues per chain, respectively. 43 torsion angles (ϕ, ψ, χ), and 18 oriental and shape coordinates, were used as degrees of freedom.

Table 1. Energy drift δ_E (in per cent) for different time steps.

Time step Δt (fs)	Energy drift δ_E in per cent
1	-0.02
5	-0.31
10	-0.36
15	-1.40
20	-3.28

The time evolution of some torsion angles during the last 10 ps of the simulations is shown in Fig. 3. The evolution of the same angle (e.g.ϕ_4) is different for different sizes of the time step. This is an effect of using a relatively small cutoff value, 10 Å, and using a cutoff function which goes from 1 to 0 over a distance of 1 Å. If no cutoff function is used, then the time evolution of an angle for different sizes of the time step is mainly affected by inherent round-off effects of the Verlet integrator.

The orientation angles and center of mass projection scalars are shown in Fig. 4-5 for the entire duration of the simulation, with the time step equal to 10 fs. Fig. 4-5 illustrate that the individual chains are indeed moving relative to each other. It should be noted that the orientation of the entire complex changes over time, as demonstrated in Fig. 4(a) and Fig. 6. But the center of mass of the complex does not move significantly, see Fig. 5(a), since there is no external force acting on the complex.

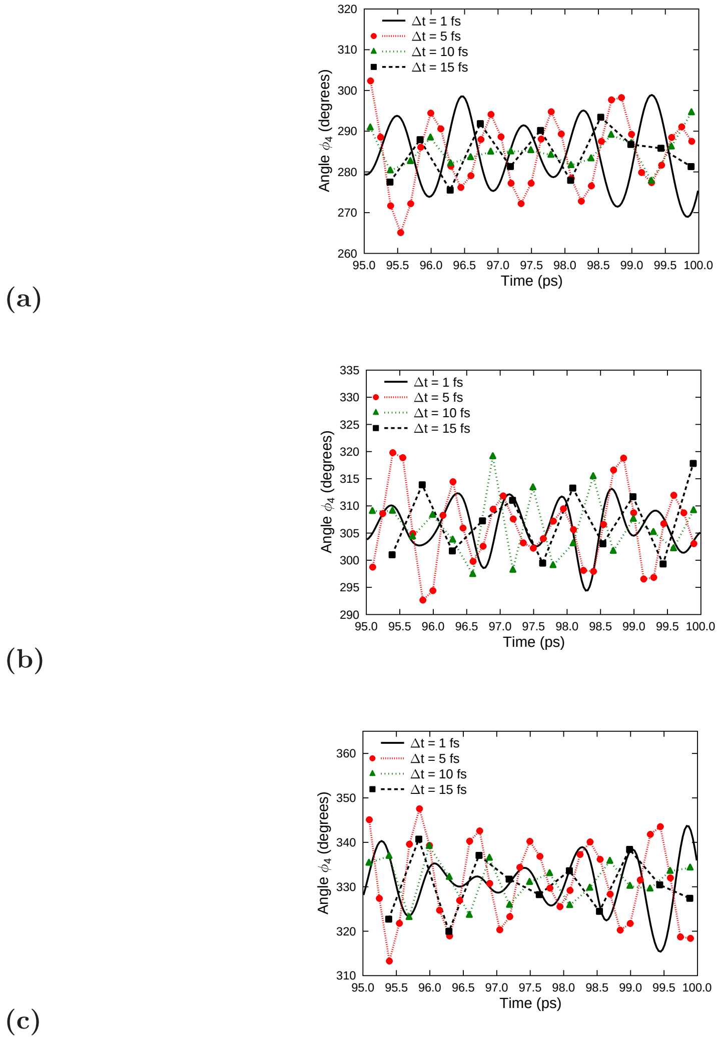

Figure 3. Torsion angles (a) ϕ_4, (b) ϕ_{14}, and (c) ϕ_{24} as a function of time and time step size. Other details are as in Fig. 2.

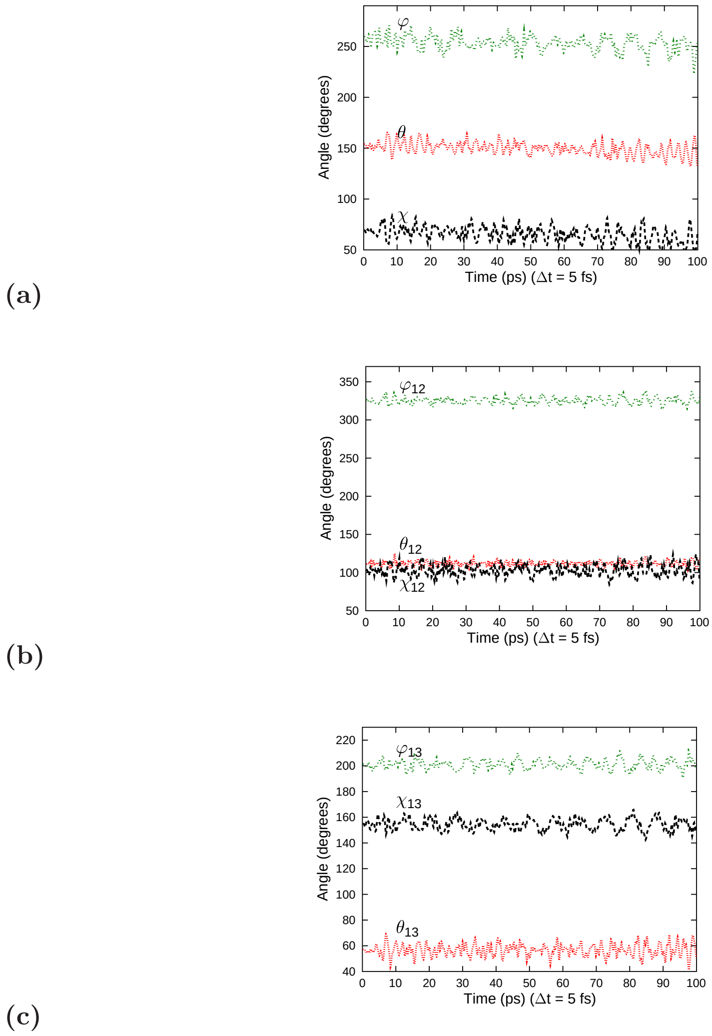

Figure 4. Orientation angles (a) θ, φ, χ, (b) $\theta_{12}, \varphi_{12}, \chi_{12}$, and (c) $\theta_{13}, \varphi_{13}, \chi_{13}$ as a function of time, for time step $\Delta t = 5$ fs.

The total linear momentum $\mathbf{p}$ and angular momentum $\mathbf{L}_{\mathrm{CM}}$ relative to the center of mass are defined as

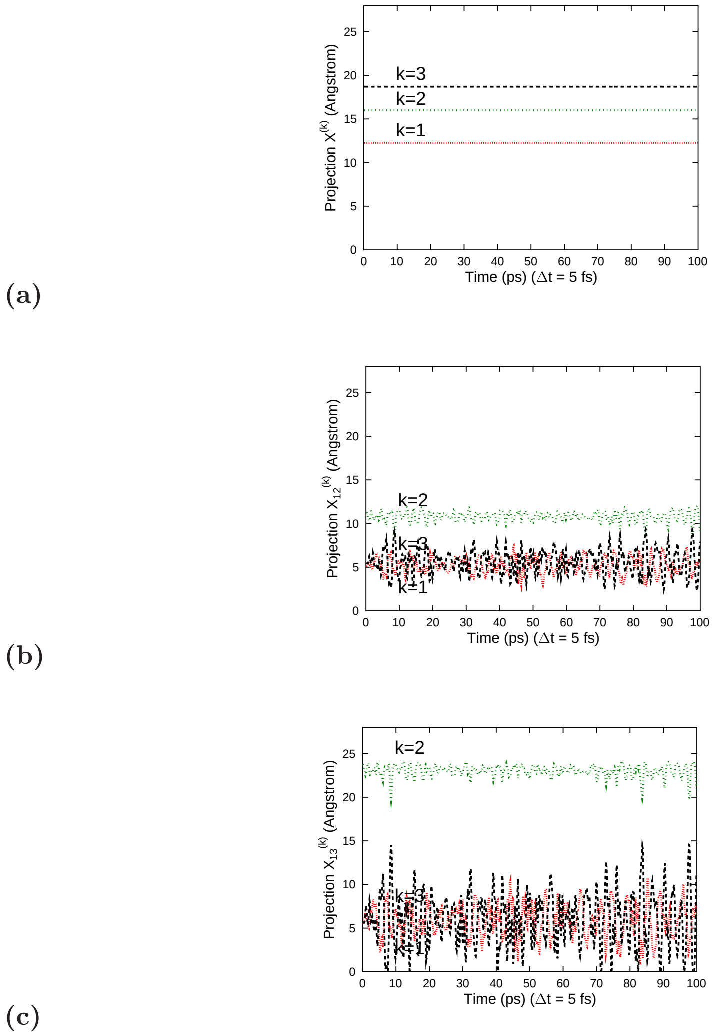

Figure 5. Center of mass projection scalars (a) $X^{(k)}$, (b) $X_{12}^{(k)}$, and (c) $X_{13}^{(k)}$ as a function of time, for time step $\Delta t = 5$ fs.

$$\mathbf{p} = \sum_{I}^{N_c} \sum_{\alpha}^{N_I} m_\alpha^{(I)} \mathbf{v}_\alpha^{(I)}, \tag{37}$$

$$\begin{aligned}\mathbf{L}_{\mathrm{CM}} &= \sum_{I}^{N_c} \sum_{\alpha}^{N_I} m_\alpha^{(I)} (\mathbf{X} - \mathbf{x}_\alpha^{(I)}) \times \mathbf{v}_\alpha^{(I)} \\ &= \mathbf{X} \times \mathbf{p} - \sum_{I}^{N_c} \sum_{\alpha}^{N_I} m_\alpha^{(I)} \mathbf{x}_\alpha^{(I)} \times \mathbf{v}_\alpha^{(I)} \\ &= \mathbf{X} \times \mathbf{p} - \mathbf{L}, \end{aligned} \tag{38}$$

where $m_\alpha^{(I)}, \mathbf{v}_\alpha^{(I)}, \mathbf{x}_\alpha^{(I)}$ is the mass, velocity, and position of atom α in chain I, given in the external (laboratory) frame, and $\mathbf{L}$ is the angular momentum relative to the origin of the laboratory system. Only $|\mathbf{L}_{\text{CM}}|$ as a function of time is shown in Fig. 6, since the evolution for $|\mathbf{p}|$ is much more stable during the entire simulation.

The force acting on the molecule is conservative, which means that besides the energy, the total linear and angular momenta are to be conserved. The increase with time, as seen in Fig. 6, is due to round-off errors. The situation is similar to what happens with the total energy in Fig. 2.

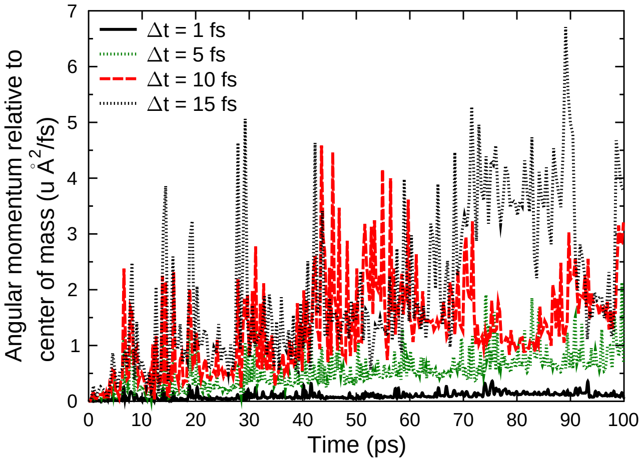

Figure 6. As in Fig. 2, but for the total angular momentum relative to the center of mass of the complex.

4. Conclusion

We have presented a detailed test of the implementation of the extended polyspherical approach for multi-chain protein complexes. In this approach the molecules and their motion is parametrized with torsion angles as well as oriental and other shape-related coordinates. A static test shows that the parametrization reproduce original Cartesian coordinates of even proteins with 15 chains. The suitability of the parametrization for molecular dynamics simulations has been successfully tested using the velocity Verlet integrator and several different time steps. The total number of degrees of freedom was 61, and the total simulation time was 100 ps. The largest safest time step was found to be 5 fs, if an energy drift of less than 1% and a very low drift in total angular momentum is desired. A time step of 10 fs also provides a small energy drift, but the angular momentum is more unstable in this case. The tests presented here show that the polyspherical approach can be used for molecular dynamics simulations, especially when the potential is written explicitly in terms of natural degrees of freedom such as bond lengths, bond angles, and torsion angles.

Acknowledgement. K. O. E. H. is funded from the Research Grant *Motions in Macromolecular Function: New Approaches to Visualize and Simulate Protein Flexibility*, awarded 2008 by the *Human Frontier of Science Program* (*HFSP*).

Appendix A. First order derivatives

Now, it follows by directly differentiating $\mathbf{u}_i^{(1)} = R_1^\dagger \mathbf{u}_i R_1$, where $R_1 = \exp\left(\chi\imath\hat{\boldsymbol{\beta}}/2\right) = \alpha + \imath\boldsymbol{\beta} = \cos(\chi/2) + \imath\sin(\chi/2)\,\hat{\boldsymbol{\beta}}$, that

$$\begin{aligned}\frac{\partial \mathbf{u}_i^{(1)}}{\partial\chi} &= \left(\frac{\partial R_1^\dagger}{\partial\chi}\right)\mathbf{u}_i R_1 + R_1^\dagger \mathbf{u}_i \frac{\partial R_1}{\partial\chi} = \frac{\imath}{2}(-\hat{\boldsymbol{\beta}} R_1^\dagger \mathbf{u}_i R_1 + R_1^\dagger \mathbf{u}_i R_1 \hat{\boldsymbol{\beta}}) \\ &= \frac{\imath}{2}(-\hat{\boldsymbol{\beta}}\mathbf{u}_i^{(1)} + \mathbf{u}_i^{(1)}\hat{\boldsymbol{\beta}}) = \hat{\boldsymbol{\beta}} \times \mathbf{u}_i^{(1)} \end{aligned} \tag{A.1}$$

By using this result and

$$\hat{\boldsymbol{\beta}}_{1I} = \sin\theta_{1I}\sin\varphi_{1I}\mathbf{u}_1^{(1)} - \sin\theta_{1I}\cos\varphi_{1I}\mathbf{u}_2^{(1)} + \cos\theta_{1I}\mathbf{u}_3^{(1)} \tag{A.2}$$

we also obtain

$$\frac{\partial \boldsymbol{\beta}_{1I}}{\partial\chi} = \hat{\boldsymbol{\beta}} \times \boldsymbol{\beta}_{1I} \tag{A.3}$$

Because $\alpha_{1I} = \cos(\chi_{1I}/2)$ and the expansion coefficients $\{\mathbf{u}_j^{(1)} \cdot \boldsymbol{\beta}_{1I}\}$ do not depend on χ, and that because it follows from the Jacobi identity $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = 0$ that $2\alpha_{1I}\left[\left(\hat{\boldsymbol{\beta}} \times \boldsymbol{\beta}_{1I}\right) \times \mathbf{u}_j^{(1)} + \boldsymbol{\beta}_{1I} \times \left(\hat{\boldsymbol{\beta}} \times \mathbf{u}_j^{(1)}\right)\right] = 2\alpha_{1I}\hat{\boldsymbol{\beta}} \times \left(\boldsymbol{\beta}_{1I} \times \mathbf{u}_j^{(1)}\right)$, we see that the derivative of

$$\mathbf{u}_j^{(I)} = \left(2\alpha_{1I}^2 - 1\right)\mathbf{u}_j^{(1)} + 2\alpha_{1I}\boldsymbol{\beta}_{1I} \times \mathbf{u}_j^{(1)} + 2\mathbf{u}_j^{(1)} \cdot \boldsymbol{\beta}_{1I}\boldsymbol{\beta}_{1I} \tag{A.4}$$

with respect to χ is

$$\frac{\partial \mathbf{u}_j^{(I)}}{\partial\chi} = \hat{\boldsymbol{\beta}} \times \mathbf{u}_i^{(I)} \tag{A.5}$$

By taking the advantage of the coordinate representation $\mathbf{r}_i^{(I)} = \sum_{j=1}^{3}\left(\mathbf{r}_i^{(I)} \cdot \mathbf{u}_j^{(I)}\right)\mathbf{u}_j^{(I)}$ (the expansion coefficients $\left\{\mathbf{r}_i^{(I)} \cdot \mathbf{u}_j^{(I)}\right\}$ depend only on the torsion angles of the Ith protein for all values of the index i), it also follows that

$$\frac{\partial \mathbf{r}_i^{(I)}}{\partial\chi} = \hat{\boldsymbol{\beta}} \times \mathbf{r}_i^{(I)} \tag{A.6}$$

Analogously, we also have

$$\frac{\partial \mathbf{r}_i^{(I)}}{\partial \chi_{1J}} = \delta_{IJ} \hat{\boldsymbol{\beta}}_{1I} \times \mathbf{r}_i^{(I)} \tag{A.7}$$

where Kronecker's delta δ_{IJ} is one, if $I = J$, and zero otherwise.

Let us now evaluate $\partial \mathbf{r}_i^{(I)} / \partial \varphi$. It follows by directly differentiating

$$\hat{\boldsymbol{\beta}} = \sin\theta \sin\varphi \mathbf{u}_1 - \sin\theta \cos\varphi \mathbf{u}_2 + \cos\theta \mathbf{u}_3 \tag{A.8}$$

that

$$\frac{\partial \hat{\boldsymbol{\beta}}}{\partial \varphi} = \mathbf{u}_3 \times \hat{\boldsymbol{\beta}} \tag{A.9}$$

As a consequence,

$$\frac{\partial R_1}{\partial \varphi} = \frac{\partial}{\partial \varphi} \left[\cos\left(\frac{\chi}{2}\right) + \imath \sin\left(\frac{\chi}{2}\right) \hat{\boldsymbol{\beta}} \right] = \imath \sin\left(\frac{\chi}{2}\right) \mathbf{u}_3 \times \hat{\boldsymbol{\beta}} \tag{A.10}$$

$$\frac{\partial R_1^\dagger}{\partial \varphi} = \frac{\partial}{\partial \varphi} \left[\cos\left(\frac{\chi}{2}\right) - \imath \sin\left(\frac{\chi}{2}\right) \hat{\boldsymbol{\beta}} \right] = -\imath \sin\left(\frac{\chi}{2}\right) \mathbf{u}_3 \times \hat{\boldsymbol{\beta}} \tag{A.11}$$

and hence

$$\begin{aligned}
\frac{\partial \mathbf{u}_i^{(1)}}{\partial \varphi} &= \frac{\partial \left(R_1^\dagger \mathbf{u}_i R_1 \right)}{\partial \varphi} = \frac{\partial R_1^\dagger}{\partial \varphi} \mathbf{u}_i R_1 + R_1^\dagger \mathbf{u}_i \frac{\partial R_1}{\partial \varphi} \\
&= \imath \sin\left(\frac{\chi}{2}\right) \left(-\mathbf{u}_3 \times \hat{\boldsymbol{\beta}} \mathbf{u}_i R_1 + R_1^\dagger \mathbf{u}_i \mathbf{u}_3 \times \hat{\boldsymbol{\beta}} \right) \\
&= \imath \sin\left(\frac{\chi}{2}\right) \left(-R_1 \mathbf{a}^{(\varphi)} \mathbf{u}_i^{(1)} + \mathbf{u}_i^{(1)} \mathbf{a}^{(\varphi)} R_1^\dagger \right)
\end{aligned} \tag{A.12}$$

where

$$\mathbf{a}^{(\varphi)} = \mathbf{u}_3^{(1)} \times \hat{\boldsymbol{\beta}} \tag{A.13}$$

The last equality follows by taking the advantage of $\mathbf{u}_3 \times \hat{\boldsymbol{\beta}} \mathbf{u}_i R_1 = R_1 R_1^\dagger \mathbf{u}_3 \times \hat{\boldsymbol{\beta}} R_1 R_1^\dagger \mathbf{u}_i R_1 = R_1 \left(R_1^\dagger \mathbf{u}_3 R_1 \right) \times \hat{\boldsymbol{\beta}} R_1^\dagger \mathbf{u}_i R_1 = R_1 \mathbf{u}_3^{(1)} \times \hat{\boldsymbol{\beta}} \mathbf{u}_i^{(1)}$ (and its reverse). By explicitly writing $R_1 = \alpha + \imath \boldsymbol{\beta}$, we have $\partial \mathbf{u}_i^{(1)} / \partial \varphi = \imath \sin(\chi/2) [-(\alpha + \imath \boldsymbol{\beta}) \mathbf{a}^{(\varphi)} \cdot \mathbf{u}_i^{(1)} + \mathbf{u}_i^{(1)} \cdot \mathbf{a}^{(\varphi)} (\alpha - \imath \boldsymbol{\beta}) - (\alpha + \imath \boldsymbol{\beta}) \mathbf{a}^{(\varphi)} \wedge \mathbf{u}_i^{(1)} + \mathbf{u}_i^{(1)} \wedge \mathbf{a}^{(\varphi)} (\alpha - \imath \boldsymbol{\beta})]$. Hence,

$$\begin{aligned}
\frac{\partial \mathbf{u}_i^{(1)}}{\partial \varphi} &= \imath \sin\left(\frac{\chi}{2}\right) [-2\imath \mathbf{a}^{(\varphi)} \cdot \mathbf{u}_i^{(1)} \boldsymbol{\beta} + 2\alpha \mathbf{u}_i^{(1)} \wedge \mathbf{a}^{(\varphi)} + 2\imath \mathbf{a}^{(\varphi)} \wedge \mathbf{u}_i^{(1)} \cdot \boldsymbol{\beta}] \\
&= 2 \sin\left(\frac{\chi}{2}\right) [\mathbf{u}_3^{(1)} \times \hat{\boldsymbol{\beta}} \cdot \mathbf{u}_i^{(1)} \boldsymbol{\beta} + \alpha \left(\mathbf{u}_3^{(1)} \times \hat{\boldsymbol{\beta}} \right) \times \mathbf{u}_i^{(1)} \\
&\quad - \left(\mathbf{u}_3^{(1)} \times \hat{\boldsymbol{\beta}} \right) \wedge \mathbf{u}_i^{(1)} \cdot \boldsymbol{\beta}]
\end{aligned} \tag{A.14}$$

or

$$\frac{\partial \mathbf{u}_i^{(1)}}{\partial \varphi} = 2 \sin\left(\frac{\chi}{2}\right) \left(\mathbf{a}^{(\varphi)} \cdot \mathbf{u}_i^{(1)} \boldsymbol{\beta} + \alpha \mathbf{a}^{(\varphi)} \times \mathbf{u}_i^{(1)} - \mathbf{a}^{(\varphi)} \mathbf{u}_i^{(1)} \cdot \boldsymbol{\beta} \right) \tag{A.15}$$

which further assumes the form

$$\frac{\partial \mathbf{u}_i^{(1)}}{\partial \varphi} = 2\sin\left(\frac{\chi}{2}\right)\left(\alpha \mathbf{a}^{(\varphi)} + \mathbf{a}^{(\varphi)} \times \boldsymbol{\beta}\right) \times \mathbf{u}_i^{(1)} \tag{A.16}$$

Because

$$\mathbf{u}_i^{(I)} = \sum_j f_{ij}^{(I)} \mathbf{u}_j^{(1)} \tag{A.17}$$

where the expansion coefficients $\left\{f_{ij}^{(I)}\right\}$ depend *only* on the angles χ_{1I}, θ_{1I} and φ_{1I}, we furthermore have

$$\begin{aligned}\frac{\partial \mathbf{u}_i^{(I)}}{\partial \varphi} &= \sum_j f_{ij}^{(I)} \frac{\partial \mathbf{u}_j^{(1)}}{\partial \varphi} = 2\sin\left(\frac{\chi}{2}\right)\sum_j f_{ij}^{(I)}\left[\alpha \mathbf{a}^{(\varphi)} + \left(\mathbf{a}^{(\varphi)} \times \boldsymbol{\beta}\right)\right] \times \mathbf{u}_j^{(1)} \\ &= 2\sin\left(\frac{\chi}{2}\right)\left[\alpha \mathbf{a}^{(\varphi)} + \left(\mathbf{a}^{(\varphi)} \times \boldsymbol{\beta}\right)\right] \times \mathbf{u}_i^{(I)}\end{aligned} \tag{A.18}$$

This in turn implies (because $\left\{\mathbf{r}_i^{(I)} \cdot \mathbf{u}_j^{(I)}\right\}$ depend only on the torsion angles) that

$$\frac{\partial \mathbf{r}_i^{(I)}}{\partial \varphi} = 2\sin\left(\frac{\chi}{2}\right)\left(\alpha \mathbf{a}^{(\varphi)} + \mathbf{a}^{(\varphi)} \times \boldsymbol{\beta}\right) \times \mathbf{r}_i^{(I)} \tag{A.19}$$

The derivative $\partial \mathbf{r}_i^{(I)}/\partial \varphi_{1J}$ can be evaluated in an analogous manner. The result is

$$\frac{\partial \mathbf{r}_i^{(I)}}{\partial \varphi_{1J}} = \delta_{IJ} 2\sin\left(\frac{\chi_{1I}}{2}\right)\left(\alpha_{1I} \mathbf{a}_{1I}^{(\varphi)} + \mathbf{a}_{1I}^{(\varphi)} \times \boldsymbol{\beta}_{1I}\right) \times \mathbf{r}_i^{(I)} \tag{A.20}$$

where

$$\mathbf{a}_{1I}^{(\varphi)} = \mathbf{u}_3^{(I)} \times \hat{\boldsymbol{\beta}}_{1I} \tag{A.21}$$

Let us now evaluate $\partial \mathbf{r}_i^{(I)}/\partial \theta$. We define the (non-orthogonal and planar) basis $\mathbf{e}_1 = \mathbf{u}_1 \times \hat{\boldsymbol{\beta}} = -\cos\theta \mathbf{u}_2 - \sin\theta\cos\varphi \mathbf{u}_3$, $\mathbf{e}_2 = \mathbf{u}_2 \times \hat{\boldsymbol{\beta}} = \cos\theta \mathbf{u}_1 - \sin\theta\sin\varphi \mathbf{u}_3$. The derivative

$$\frac{\partial \hat{\boldsymbol{\beta}}}{\partial \theta} = (\cos\theta\sin\varphi \mathbf{u}_1 - \cos\theta\cos\varphi \mathbf{u}_2 - \sin\theta \mathbf{u}_3) \tag{A.22}$$

can be written in this basis as

$$\frac{\partial \hat{\boldsymbol{\beta}}}{\partial \theta} = \cos\varphi \mathbf{e}_1 + \sin\varphi \mathbf{e}_2 = (\cos\varphi \mathbf{u}_1 + \sin\varphi \mathbf{u}_2) \times \hat{\boldsymbol{\beta}} \tag{A.23}$$

This implies (by the derivation analogous to that for $\partial \mathbf{r}_i^{(I)}/\partial \varphi$) that $\partial \mathbf{u}_i^{(I)}/\partial \theta = 2\sin(\chi/2)\left[\alpha \mathbf{a}^{(\theta)} + \left(\mathbf{a}^{(\theta)} \times \boldsymbol{\beta}\right)\right] \times \mathbf{u}_i^{(I)}$, and consequently,

$$\frac{\partial \mathbf{r}_i^{(I)}}{\partial \theta} = 2\sin\left(\frac{\chi}{2}\right)\left(\alpha \mathbf{a}^{(\theta)} + \mathbf{a}^{(\theta)} \times \boldsymbol{\beta}\right) \times \mathbf{r}_i^{(I)} \tag{A.24}$$

with

$$\mathbf{a}^{(\theta)} = \left(\cos\varphi \mathbf{u}_1^{(1)} + \sin\varphi \mathbf{u}_2^{(1)}\right) \times \hat{\boldsymbol{\beta}} \tag{A.25}$$

The derivative $\partial \mathbf{r}_i^{(I)}/\partial\theta_{1J}$ can be evaluated in an analogous manner. The result is

$$\frac{\partial \mathbf{r}_i^{(I)}}{\partial\theta_{1J}} = \delta_{IJ} 2\sin\left(\frac{\chi_{1I}}{2}\right)\left(\alpha_{1I}\mathbf{a}_{1I}^{(\theta)} + \mathbf{a}_{1I}^{(\theta)}\times\boldsymbol{\beta}_{1I}\right)\times\mathbf{r}_i^{(I)} \tag{A.26}$$

where

$$\mathbf{a}_{1I}^{(\theta)} = \left(\cos\varphi_{1I}\mathbf{u}_1^{(I)} + \sin\varphi_{1I}\mathbf{u}_2^{(I)}\right)\times\hat{\boldsymbol{\beta}}_{1I} \tag{A.27}$$

Let now $\left\{\mathbf{r}_l^{(I)}\right\}$ form the bond tree for a protein I. Then

$$\frac{\partial \mathbf{r}_l^{(J)}}{\partial\phi_{jk}^{(Ii)}} = \delta_{IJ} d\left(\mathbf{r}_k^{(J)},\mathbf{r}_l^{(J)}\right)\hat{\mathbf{r}}_i^{(J)}\times\mathbf{r}_l^{(J)} \tag{A.28}$$

where $d\left(\mathbf{r}_k^{(J)},\mathbf{r}_l^{(J)}\right)$ is one, if the coordinate $\phi_{jk}^{(Ji)}$ parametrizes a bond $\mathbf{r}_k^{(J)}$ in a chain, which connects $\mathbf{r}_i^{(J)}$ to $\mathbf{r}_l^{(J)}$ (in that order) in the bond tree. The fact that $\partial\mathbf{r}_l^{(J)}/\partial\phi_{jk}^{(Ii)}$ vanishes if $I \neq J$ can be deduced from the basis representation $\mathbf{r}_l^{(J)} = \sum_{j=1}^3\left(\mathbf{r}_l^{(J)}\cdot\mathbf{u}_j^{(J)}\right)\mathbf{u}_j^{(J)}$ (the expansion coefficients $\left\{\mathbf{r}_l^{(J)}\cdot\mathbf{u}_j^{(J)}\right\}$ depend only on the torsion angles of the Jth protein for all values of the index l).

Because the parametrization of an arbitrary bond vector $\mathbf{r}_i^{(J)}$ does not depend on $X_{1J}^{(j)}$, we have identically

$$\frac{\partial \mathbf{r}_i^{(I)}}{\partial X_{1J}^{(j)}} = 0 \tag{A.29}$$

for all values of the indexes.

The derivatives of the relative center of masses w.r.t χ, θ and φ are given by

$$\frac{\partial \mathbf{X}_{1I}}{\partial\vartheta} = \sum_j X_{1I}^{(j)}\frac{\partial\mathbf{u}_j^{(1)}}{\partial\vartheta} \tag{A.30}$$

(with $\vartheta = \chi,\theta,\varphi$), where

$$\frac{\partial\mathbf{u}_j^{(1)}}{\partial\vartheta} = \mathbf{b}^{(\vartheta)}\times\mathbf{u}_j^{(1)} \tag{A.31}$$

with

$$\mathbf{b}^{(\chi)} = \hat{\boldsymbol{\beta}} \tag{A.32}$$

$$\mathbf{b}^{(\varphi)} = 2\sin\left(\frac{\chi}{2}\right)\left(\alpha\mathbf{a}^{(\varphi)} + \mathbf{a}^{(\varphi)}\times\boldsymbol{\beta}\right) \tag{A.33}$$

$$\mathbf{b}^{(\theta)} = 2\sin\left(\frac{\chi}{2}\right)\left(\alpha\mathbf{a}^{(\theta)} + \mathbf{a}^{(\theta)}\times\boldsymbol{\beta}\right) \tag{A.34}$$

Because $\mathbf{X}_{1I}$ does not depend parametrically on any other shape coordinate besides $X_{1I}^{(1)}$, $X_{1I}^{(2)}$ and $X_{1I}^{(3)}$, it follows that

$$\frac{\partial \mathbf{X}_{1I}}{\partial X_{1J}^{(i)}} = \delta_{IJ}\mathbf{u}_i^{(1)} \tag{A.35}$$

for $i = 1, 2, 3$, and

$$\frac{\partial \mathbf{X}_{1I}}{\partial \vartheta_{1J}} = \frac{\partial \mathbf{X}_{1I}}{\partial \phi_{jk}^{(Ji)}} = 0 \tag{A.36}$$

for all values of the indexes.

Appendix B. Second order derivatives

The second order derivatives are obtained from the first order derivatives as follows.

Now,

$$\frac{\partial^2 \mathbf{r}_l^{(J)}}{\partial q_n \partial \phi_{jk}^{(Ii)}} = \delta_{IJ}\, d\left(\mathbf{r}_k^{(J)}, \mathbf{r}_l^{(J)}\right)\left(\frac{\partial \mathbf{r}_i^{(J)}/\partial q_n}{r_i^{(J)}} \times \mathbf{r}_l^{(J)} + \hat{\mathbf{r}}_i^{(J)} \times \frac{\partial \mathbf{r}_l^{(J)}}{\partial q_n}\right) \tag{B.1}$$

where q_n can be any coordinate in our set of coordinates.

Let now $\vartheta, \vartheta' = \chi, \theta, \varphi$. Then

$$\frac{\partial^2 \mathbf{r}_i^{(I)}}{\partial \vartheta' \partial \vartheta} = \frac{\partial \mathbf{b}^{(\vartheta)}}{\partial \vartheta'} \times \mathbf{r}_i^{(I)} + \mathbf{b}^{(\vartheta)} \times \frac{\partial \mathbf{r}_i^{(I)}}{\partial \vartheta'} \tag{B.2}$$

The derivatives $\partial \mathbf{b}^{(\chi)}/\partial \varphi = \partial \hat{\boldsymbol{\beta}}/\partial \varphi = \mathbf{u}_3 \times \hat{\boldsymbol{\beta}}$ and $\partial \mathbf{b}^{(\chi)}/\partial \theta = \partial \hat{\boldsymbol{\beta}}/\partial \theta = (\cos\varphi \mathbf{u}_1 + \sin\varphi \mathbf{u}_2) \times \hat{\boldsymbol{\beta}}$ were derived in the previous section, $\partial \mathbf{b}^{(\chi)}/\partial \chi = \partial \hat{\boldsymbol{\beta}}/\partial \chi = 0$, and

$$\begin{aligned}\frac{\partial \mathbf{b}^{(\vartheta)}}{\partial \vartheta'} &= 2\sin\left(\frac{\chi}{2}\right)\left[\alpha \frac{\partial \mathbf{a}^{(\vartheta)}}{\partial \vartheta'} - \frac{\delta_{\vartheta'\chi}}{2}\sin\left(\frac{\chi}{2}\right)\mathbf{a}^{(\vartheta)}\right.\\ &\quad \left. + \frac{\partial \mathbf{a}^{(\vartheta)}}{\partial \vartheta'} \times \boldsymbol{\beta} + \mathbf{a}^{(\vartheta)} \times \frac{\partial \boldsymbol{\beta}}{\partial \vartheta'}\right]\\ &\quad + \delta_{\vartheta'\chi}\cos\left(\frac{\chi}{2}\right)\left(\alpha \mathbf{a}^{(\vartheta)} + \mathbf{a}^{(\vartheta)} \times \boldsymbol{\beta}\right)\end{aligned} \tag{B.3}$$

for $\vartheta = \theta, \varphi$, where

$$\frac{\partial \mathbf{a}^{(\varphi)}}{\partial \vartheta'} = \frac{\partial \mathbf{u}_3^{(1)}}{\partial \vartheta'} \times \hat{\boldsymbol{\beta}} + \mathbf{u}_3^{(1)} \times \frac{\partial \hat{\boldsymbol{\beta}}}{\partial \vartheta'} \tag{B.4}$$

$$\begin{aligned}\frac{\partial \mathbf{a}^{(\theta)}}{\partial \vartheta'} &= \left[\delta_{\varphi\vartheta'}\left(-\sin\varphi \mathbf{u}_1^{(1)} + \cos\varphi \mathbf{u}_2^{(1)}\right) + \cos\varphi \frac{\partial \mathbf{u}_1^{(1)}}{\partial \vartheta'} + \sin\varphi \frac{\partial \mathbf{u}_2^{(1)}}{\partial \vartheta'}\right] \times \hat{\boldsymbol{\beta}}\\ &\quad + \left(\cos\varphi \mathbf{u}_1^{(1)} + \sin\varphi \mathbf{u}_2^{(1)}\right) \times \frac{\partial \hat{\boldsymbol{\beta}}}{\partial \vartheta'}\end{aligned} \tag{B.5}$$

The Kronecker's delta $\delta_{\varphi\vartheta'}$ is one, if ϑ' equals φ, and zero otherwise.

Because all the derivatives $\partial \mathbf{b}^{(\vartheta)}/\partial \vartheta_{1I}$ with $\vartheta = \chi, \theta, \varphi$ and $\vartheta_{1I} = \theta_{1I}, \varphi_{1I}, \chi_{1I}$ are identically zero, we have

$$\frac{\partial^2 \mathbf{r}_i^{(I)}}{\partial \vartheta_{1J} \partial \vartheta} = \delta_{IJ} \mathbf{b}^{(\vartheta)} \times \frac{\partial \mathbf{r}_i^{(I)}}{\partial \vartheta_{1I}} \tag{B.6}$$

Also,

$$\frac{\partial^2 \mathbf{r}_i^{(I)}}{\partial \vartheta_{1K} \partial \vartheta_{1J}} = \delta_{JK}\delta_{IJ} \left(\frac{\partial \mathbf{b}_{1I}^{(\vartheta)}}{\partial \vartheta_{1K}} \times \mathbf{r}_i^{(I)} + \mathbf{b}_{1I}^{(\vartheta)} \times \frac{\partial \mathbf{r}_i^{(I)}}{\partial \vartheta_{1K}} \right) \tag{B.7}$$

where

$$\mathbf{b}_{1I}^{(\chi)} = \hat{\boldsymbol{\beta}}_{1I} \tag{B.8}$$

and, for, $\vartheta = \theta, \varphi$,

$$\mathbf{b}_{1I}^{(\vartheta)} = 2\sin\left(\frac{\chi_{1I}}{2}\right) \left[\alpha_{1I} \mathbf{a}_{1I}^{(\vartheta)} + \left(\mathbf{a}_{1I}^{(\vartheta)} \times \boldsymbol{\beta}_{1I} \right) \right] \tag{B.9}$$

$$\begin{aligned} \frac{\partial \mathbf{b}_{1I}^{(\vartheta)}}{\partial \vartheta_{1K}} &= \delta_{IK} 2\sin\left(\frac{\chi_{1I}}{2}\right) \left[\alpha_{1I} \frac{\partial \mathbf{a}_{1I}^{(\vartheta)}}{\partial \vartheta_{1K}} - \frac{\delta_{\vartheta_{1K}\chi_{1I}}}{2} \sin\left(\frac{\chi_{1I}}{2}\right) \mathbf{a}_{1I}^{(\vartheta)} \right. \\ &\quad \left. + \frac{\partial \mathbf{a}_{1I}^{(\vartheta)}}{\partial \vartheta_{1K}} \times \boldsymbol{\beta}_{1I} + \mathbf{a}_{1I}^{(\vartheta)} \times \frac{\partial \boldsymbol{\beta}_{1I}}{\partial \vartheta_{1K}} \right] \\ &\quad + \delta_{\vartheta_{1K}\chi_{1I}} \cos\left(\frac{\chi_{1I}}{2}\right) \left(\alpha_{1I} \mathbf{a}_{1I}^{(\vartheta)} + \mathbf{a}_{1I}^{(\vartheta)} \times \boldsymbol{\beta}_{1I} \right) \end{aligned} \tag{B.10}$$

where

$$\frac{\partial \mathbf{a}_{1I}^{(\varphi)}}{\partial \vartheta_{1K}} = \delta_{IK} \left[\frac{\partial \mathbf{u}_3^{(I)}}{\partial \vartheta_{1K}} \times \hat{\boldsymbol{\beta}}_{1I} + \mathbf{u}_3^{(I)} \times \frac{\partial \hat{\boldsymbol{\beta}}_{1I}}{\partial \vartheta_{1K}} \right] \tag{B.11}$$

$$\begin{aligned} \frac{\partial \mathbf{a}_{1I}^{(\theta)}}{\partial \vartheta_{1K}} &= \delta_{IK} \left[\delta_{\varphi_{1I}\vartheta_{1K}} \left(-\sin\varphi_{1I} \mathbf{u}_1^{(I)} + \cos\varphi_{1I} \mathbf{u}_2^{(I)} \right) + \cos\varphi_{1I} \frac{\partial \mathbf{u}_1^{(I)}}{\partial \vartheta_{1K}} + \sin\varphi_{1I} \frac{\partial \mathbf{u}_2^{(I)}}{\partial \vartheta_{1K}} \right] \times \hat{\boldsymbol{\beta}}_{1I} \\ &\quad + \delta_{IK} \left(\cos\varphi_{1I} \mathbf{u}_1^{(I)} + \sin\varphi_{1I} \mathbf{u}_2^{(I)} \right) \times \frac{\partial \hat{\boldsymbol{\beta}}_{1I}}{\partial \vartheta_{1K}} \end{aligned} \tag{B.12}$$

and

$$\frac{\partial \hat{\boldsymbol{\beta}}_{1I}}{\partial \chi_{1K}} = 0 \tag{B.13}$$

$$\frac{\partial \hat{\boldsymbol{\beta}}_{1I}}{\partial \varphi_{1K}} = \delta_{IK} \mathbf{u}_3^{(1)} \times \hat{\boldsymbol{\beta}}_{1I} \tag{B.14}$$

$$\frac{\partial \hat{\boldsymbol{\beta}}_{1I}}{\partial \theta_{1K}} = \delta_{IK} \left(\cos\varphi_{1I} \mathbf{u}_1^{(1)} + \sin\varphi_{1I} \mathbf{u}_2^{(1)} \right) \times \hat{\boldsymbol{\beta}}_{1I} \tag{B.15}$$

Finally, we have

$$\frac{\partial^2 \mathbf{X}_{1I}}{\partial q_j \partial q_i} = 0 \tag{B.16}$$

if q_i and q_j are in $\{\phi_{jk}^{(Il)}, X_{1K}^{(k)}, \chi_{1I}, \theta_{1I}, \varphi_{1I}\}$,

$$\frac{\partial^2 \mathbf{X}_{1I}}{\partial \vartheta \partial X_{1J}^{(j)}} = \delta_{IJ} \frac{\partial \mathbf{u}_j^{(1)}}{\partial \vartheta} \tag{B.17}$$

and

$$\frac{\partial^2 \mathbf{X}_{1I}}{\partial\vartheta'\partial\vartheta} = \sum_j X_{1I}^{(j)} \frac{\partial^2 \mathbf{u}_j^{(1)}}{\partial\vartheta'\partial\vartheta} \tag{B.18}$$

The derivatives $\partial^2 \mathbf{u}_j^{(1)}/\partial\vartheta'\partial\vartheta$ are given by

$$\frac{\partial^2 \mathbf{u}_i^{(1)}}{\partial\vartheta'\partial\vartheta} = \frac{\partial \mathbf{b}^{(\vartheta)}}{\partial\vartheta'} \times \mathbf{u}_i^{(1)} + \mathbf{b}^{(\vartheta)} \times \frac{\partial \mathbf{u}_i^{(1)}}{\partial\vartheta'} \tag{B.19}$$

1. J. Pesonen, K.O.E. Henriksson, J. Comput. Chem. **31**(9), 1873 (2010)
2. K.O.E. Henriksson, J. Pesonen, J. Comput. Chem. **31**(9), 1882 (2010)
3. F. Gatti, C. Iung, Physics Reports **484**(1-2), 1 (2009). DOI DOI:10.1016/j.physrep.2009.05.003. URL `http://www.sciencedirect.com/science/article/B6TVP-4WXBM5N-1/2/9d39f720f4758309422b394a656077dd`
4. Q.Cui, I.Bahar (eds.), *Normal mode analysis — theory and practice to biological and chemical systems* (Chapman & Hall, CRC, 2006)
5. L.M.Rice, A.T.Brünger, Proteins: Structure, function, and genetics **19**, 277 (1994)
6. J. Pesonen, K.O.E. Henriksson, P. Chacon, J. Comput. Chem. (2010)
7. T. Schlick, *Molecular modeling and simulation — An interdisciplinary guide* (Springer, 2006)
8. J. Pesonen, L. Halonen, Advances in chemical physics **125**, 269 (2003)