Journal of Structural Biology xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Local analysis of strains and rotations for macromolecular electron microscopy maps

C.O.S. Sorzano [a,b,*], A. Martín-Ramos [b], F. Prieto [b], R. Melero [a], J. Martín-Benito [a], S. Jonic [c], J. Navas-Calvente [a], J. Vargas [a], J. Otón [a], V. Abrishami [a], J.M. de la Rosa-Trevín [a], J. Gómez-Blanco [a], J.L. Vilas [a], R. Marabini [d], J.M. Carazo [a]

[a] Centro Nac. Biotecnología (CSIC), c/Darwin, 3, 28049 Cantoblanco, Madrid, Spain
[b] Univ. San Pablo – CEU, Campus Urb. Montepríncipe, 28668 Boadilla del Monte, Madrid, Spain
[c] IMPMC, Sorbonne Universités – CNRS UMR 7590, UPMC University Paris 6, MNHN, IRDUMR206, 75005 Paris, France
[d] Univ. Autónoma de Madrid, 28049 Cantoblanco, Madrid, Spain

ARTICLE INFO

Article history:
Received 22 February 2016
Received in revised form 1 April 2016
Accepted 6 April 2016
Available online xxxx

Keywords:
Single particle analysis
Local deformation analysis
Local rotations
Local strains
Elastic registration

ABSTRACT

Macromolecular complexes perform their physiological functions by local rearrangements of their constituents and biochemically interacting with their reaction partners. These rearrangements may involve local rotations and the induction of local strains causing different mechanical efforts and stretches at the different areas of the protein. The analysis of these local deformations may reveal important insight into the way proteins perform their tasks. In this paper we introduce a method to perform this kind of local analysis using Electron Microscopy volumes in a fully objective and automatic manner. For doing so, we exploit the continuous nature of the result of an elastic image registration using B-splines as its basis functions. We show that the results obtained by the new automatic method are consistent with previous observations on these macromolecules.

1. Introduction

Biological macromolecules are intrinsically flexible entities that may adopt a variety of conformations as they perform defined biological functions. At high resolution, these conformational changes can be studied through a detailed analysis of the atomic locations in two different conformational states of the macromolecule. This analysis provides precise information about which areas are stretching, which areas are compressing, and which areas are rotating. However, performing the analysis at the atomic level is not always possible simply because many times structural results produce data that are better treated as 3D maps, as it is the case of many macromolecules solved by Electron Microscopy. Electron Microscopy has shown to be a very powerful tool to analyze different macromolecular discrete (Penczek et al., 2006; Penczek et al., 2006; Leschziner and Nogales, 2007; Scheres et al., 2007; Spahn and Penczek, 2009; Shatsky et al., 2010; Scheres, 2012; Klaholz, 2015) or continuous states (Dashti et al., 2014; Jin et al., 2014). However, given two medium-low resolution conformational states of the macromolecule (in our experimental section we show examples with resolutions ranging from 3.5 Å to 20 Å), an approach to automatically analyse the differences between the maps is not available. The most common action is to perform a morphing (e.g., with Chimera) between the two states so that the researcher gets an idea of the regions of the proteins primarily changing. However, from the video itself it is impossible to know if a certain region is compressing, stretching or simply rotating with respect to its surroundings.

In this article we introduce the first computational tool, to the best of our knowledge, in which the two different kinds of movements (local rotations and local strains) can be separated. The method works at the level of the EM reconstructed maps with the advantage that it can handle a wide variety of resolutions, and the disadvantage that the deformations explored are not stereochemically constrained. The method is based on the differential analysis of the deformation field between two conformational states as calculated with an elastic registration algorithm. Elastic registration is a well-established field in Image Processing in which the correspondence between equivalent coordinates in two different maps is estimated by computational means (Holden, 2008). The term elastic is used as opposed to rigid registration in which only global rotations and translations are allowed. In elastic

---

* Corresponding author at: Centro Nac. Biotecnología (CSIC), c/Darwin, 3, 28049 Cantoblanco, Madrid, Spain.

E-mail address: coss@cnb.csic.es (C.O.S. Sorzano).

http://dx.doi.org/10.1016/j.jsb.2016.04.001
1047-8477/

registration each part of the images is allowed to move independently, although some regularization may be imposed to constrain the deformation field to be smooth. As we show below, the differential analysis of the local properties of the deformation field allows distinguishing between local strains (compression and stretching) and local rotations. We show in three different experimental cases that our method provides results consistent with the conformational changes described in their corresponding reference publications.

2. Materials and methods

In this section the methodology for the local analysis of structural strains and rotations is described. This theory is well established in Continuum Mechanics and it is here presented in connection with the issue of volume registration as it is employed to identify the deformation field between two volumes. The input to the method is a pair of EM maps describing two different, although similar, states of a given macromolecule. Let us call these maps as $V_0(\mathbf{r})$ and $V_F(\mathbf{r})$, standing for the initial and final state ($\mathbf{r} = (x, y, z)^T \in \mathbb{R}^3$ is the spatial location at which the maps are evaluated). In the following, we assume that any rigid motion between V_0 and V_F has already been removed (in our examples we use Scipion's protocol "align volume", http://scipion.cnb.csic.es, de la Rosa-Trevín et al., in press). At this stage, we search for a continuous deformation field, $\mathbf{g}(\mathbf{r}) = (g_x(\mathbf{r}), g_y(\mathbf{r}), g_z(\mathbf{r}))^T \in \mathbb{R}^3$, that transforms coordinates from the final to the initial state such that

$$V_F(\mathbf{r}) \approx V_0(\mathbf{g}(\mathbf{r}))$$

If the initial and the final state were identical, $\mathbf{g}(\mathbf{r})$ would be the identity transformation ($\mathbf{g}(\mathbf{r}) = \mathbf{r}$). We define the displacement vector field as

$$\mathbf{u}(\mathbf{r}) = \mathbf{g}(\mathbf{r}) - \mathbf{r}$$

and at each point we can calculate the displacement gradient as

$$U(\mathbf{r}) = \nabla\mathbf{u}(\mathbf{r}) = \begin{pmatrix} \frac{\partial g_x}{\partial x}(\mathbf{r}) - 1 & \frac{\partial g_x}{\partial y}(\mathbf{r}) & \frac{\partial g_x}{\partial z}(\mathbf{r}) \\ \frac{\partial g_y}{\partial x}(\mathbf{r}) & \frac{\partial g_y}{\partial y}(\mathbf{r}) - 1 & \frac{\partial g_y}{\partial z}(\mathbf{r}) \\ \frac{\partial g_z}{\partial x}(\mathbf{r}) & \frac{\partial g_z}{\partial y}(\mathbf{r}) & \frac{\partial g_z}{\partial z}(\mathbf{r}) - 1 \end{pmatrix}$$

We may decompose this tensor in its symmetric, $D(\mathbf{r})$, and antisymmetric, $H(\mathbf{r})$, parts (see Sadd (2005) [Chap. 2.1] for more details):

$$\begin{aligned} D(\mathbf{r}) &= \tfrac{1}{2}(U(\mathbf{r}) + U^T(\mathbf{r})) \\ H(\mathbf{r}) &= \tfrac{1}{2}(U(\mathbf{r}) - U^T(\mathbf{r})) \end{aligned}$$

such that

$$U(\mathbf{r}) = D(\mathbf{r}) + H(\mathbf{r})$$

D is a tensor that describes local deformations while H is a tensor that describes local rotations. Note that the D, H decomposition is not the Helmholtz decomposition of a vector field ($\mathbf{u}(\mathbf{r}) = \nabla\phi(\mathbf{r}) + \nabla \times \boldsymbol{\psi}(\mathbf{r})$), although the D, H decomposition could be calculated from the Helmholtz potentials. The eigenvalues of D describe how the final structure, V_F, is locally deformed (suffering local strains that promote local stretching or compression) to match the initial state, V_0. The sign of the eigenvalue denotes whether there is a compression or a stretching, while its magnitude indicates the strength of the deformation. Similarly, the eigenvalues of H describe local rotations. Note that global rotations have already been taken away and global shifts result in null derivatives (so that they do not appear in the tensor U). We should note that due to the antisymmetric nature of H, its eigenvalues are always of the form $0, i\alpha, -i\alpha$, that is, one of the eigenvalues is always 0, and the other two are purely imaginary, complex conjugates of each other (the larger the local rotation, the larger α).

At each location of the final state, we summarize the information of these two tensors by a local strain and local rotation as

$$\begin{aligned} \text{Local Strain }(\mathbf{r}) &= |\det D(\mathbf{r})| \\ \text{Local Rotation }(\mathbf{r}) &= |\alpha(\mathbf{r})| \end{aligned}$$

If there is no local strain, all eigenvalues of D should be 0. The further these eigenvalues are from 0, the larger the local strain. If an eigenvalue is positive, it means that, locally, the overall effect is a stretching of the molecule. If an eigenvalue is negative, then, locally, the dominant effect is a compression. Note that the determinant is combining the strains in three perpendicular directions, so that there could be compression in one direction and stretching in the other two. The determinant is the multiplication of the three eigenvalues and, consequently, reduces a richer information to a single number that can be easily visualized and interpreted. Similarly, the local rotation should be the local torsion angle in radians.

This local analysis requires knowing the deformation field $\mathbf{g}(\mathbf{r})$. This is a well-known problem in image processing called elastic or non-rigid image registration, and many algorithms exist to solve it (Zitova and Flusser, 2003). Spline descriptions of the deformation field have shown to be very effective in describing rich deformation fields as well as vector field regularizers (Sorzano et al., 2005; Arganda-Carreras et al., 2006; Arganda-Carreras et al., 2008) (see Appendix). Additionally, they allow the analytical calculation at any spatial point of the derivatives needed for the displacement gradient. In this work we consider deformation fields of the form

$$\mathbf{g}(\mathbf{r}) = \sum_{jkl} \mathbf{c}_{jkl}\beta_3\left(\frac{\mathbf{r}}{h} - \begin{pmatrix} j \\ k \\ l \end{pmatrix}\right) \tag{1}$$

where j, k and l are dummy indexes to run over a cubic grid of tensor product splines (see the Appendix for more details), $\mathbf{c}_{jkl} \in \mathbb{R}^3$ is a vector of spline coefficients, h is a scale factor, and $\beta_3(\mathbf{r})$ is the three-dimensional, tensor product spline defined as

$$\beta_3(\mathbf{r}) = \beta_3(x)\beta_3(y)\beta_3(z)$$

where $\beta_3(x)$ is the cubic B-spline.

The goal of the image registration problem is to find the coefficients $\mathbf{c}_{jkl}$ such that the initial and final maps are as similar as possible:

$$\underset{\mathbf{c}_{jkl}}{\operatorname{argmin}} \|V_F(\mathbf{r}) - V_0(\mathbf{g}(\mathbf{r}))\|^2$$

For this minimization we follow the method described by Myronenko and Song (2010), in which a pyramidal coarse-to-fine approach to the image registration is followed. h is progressively decreased as the pyramid progresses (in our case, we employ three levels of the pyramid, which results in $h = 5$, that is, in the finest detail level of the pyramid there is a spline every 5 voxels in each direction, X, Y, Z).

The main two alternatives to describing the deformation field with B-splines are thin-plate splines (Rohr et al., 2004) and pixel-by-pixel defined deformation fields (Guimond et al., 2001). The first approach requires knowledge of corresponding coordinates (that is, pairs of points in the two structures that correspond to the same molecular location), which prevents automating the algorithm since it requires user input. The second approach has proved to be a powerful registration technique, especially for the definition of diffeomorphisms. However, it is not well suited for our purpose since we need to calculate derivatives of the

Please cite this article in press as: Sorzano, C.O.S., et al. Local analysis of strains and rotations for macromolecular electron microscopy maps. J. Struct. Biol. (2016), http://dx.doi.org/10.1016/j.jsb.2016.04.001