

Macromolecular Dynamics by Hybrid Electron Microscopy Normal Mode Analysis

Qiyu Jin¹, Carlos Oscar S. Sorzano², Jose Miguel de la Rosa-Trevín², Florence Tama³, and Slavica Jonić^{1,*}

¹. IMPMC, Sorbonne Universités - CNRS UMR 7590, UPMC, MNHN, IRD, 75005 Paris, France.

². Biocomputing Unit, Centro Nacional de Biotecnología — CSIC, 28049 Madrid, Spain.

³. RIKEN, Advanced Institute for Computational Sciences, Kobe, Hyogo, 650-0047,

Japan

* Corresponding author (Slavica.Jonic@impmc.upmc.fr)

Structural changes are critical for biological functions of proteins and describing conformational changes in large macromolecular complexes is a major challenge. To allow full studies of gradual conformational changes, we have recently developed Hybrid Electron Microscopy Normal Mode Analysis (HEMNMA) [1]. HEMNMA uses normal modes to elastically align electron microscopy (EM) images with a reference structure, in order to determine the conformations present in images and evaluate their pertinence. HEMNMA allows analyzing gradual changes more extensively than other EM methods, as HEMNMA gives full dynamics while other EM methods give only a few conformations. The computed conformational distribution allows modeling of transition pathways. Here, we show the performance of HEMNMA using simulated data (**Fig. 1**). The results of the method with experimental EM images will be shown at the symposium. Also, we will show a user-friendly graphical interface to the method that we have recently built within the Xmipp software.

Figure 1 shows the potential of the method to analyze any gradual-type conformational change. The synthetic conformational changes were computed using a crystal structure of Tomato Bushy Stunt Virus (TBSV) [2]. A set of 4000 images (size: 128^2 pixels; pixel size: $(3.2 \text{ \AA})^2$) was generated for random orientations and positions of the reference structure displaced along a linear combination of mode 28 (mainly related to the capsid expansion), mode 80 (mainly related to the movement of subunits away and towards the 5-fold symmetry axis) and mode 107 (mainly related to the rotation of subunits around the 5-fold symmetry axis). The contribution of mode 28 (deformation amplitude along the mode 28) was arbitrarily chosen to be distributed uniformly in the range $[0, 2000]$ and the contributions of the modes 80 and 107 were chosen to be linearly related to the contribution of the mode 28 (**Fig. 1a**). Two out-of-plane rotations were distributed uniformly in the ranges $[0^\circ, 360^\circ]$ and $[0^\circ, 180^\circ]$, the in-plane rotation was 0° , and the in-plane translations were distributed uniformly in the range $[-5, 5]$ pixels. Noise (SNR=0.05) and CTF (defocus of $2 \mu\text{m}$, 200 kV microscope, spherical aberration of 0.5 mm) were applied onto the computed projections using the model with noise before and after the CTF [3]. Each synthetic image was aligned elastically with the reference structure to estimate five rigid-body parameters and three deformation amplitudes (for three used modes). We compared the results of principal component analysis (PCA) of the computed deformation amplitudes with the results of PCA of the ground-truth deformation amplitudes, using the first two principal axes (**Fig. 1b-c**). The mean RMSD of 0.54 \AA was obtained along the linear-regression lines through the computed and ground-truth conformations (**Fig. 1b-c**). We also compared the structures computed by 3D reconstruction (**Fig. 1d-g**) from 7 images groups of similar sizes (**Fig. 1b-c**). The cross-correlation of 0.98-0.99 was obtained between the structures reconstructed for the ground-truth and computed alignments. The conformational change for both the ground-truth parameters (**Fig. 1d** for the sequence of classes 1-4; **Fig. 1e** for the sequence of classes 1-2'-3'-4') and the computed parameters (**Fig. 1f** for the sequence of classes 1-4;

Fig. 1g for the sequence of classes 1-2'-3'-4') can be mainly described as an increase in the virus' radius and a clockwise (classes 1-4) or counter-clockwise (classes 1-2'-3'-4') rotation of subunits around the 5-fold symmetry axis. Similar conformational changes were observed in the movies of the reference-structure displacement along the linear-regression lines. However, these movies also showed movements of subunits away and towards the 5-fold symmetry axis, which are less visible on the reconstructed volumes. This example shows that HEMNMA can identify the main conformational changes from highly heterogeneous data. Also, it shows that 3D reconstruction can hide some details of the conformational change due to mixing of different but close conformations into the same volume. This example also suggests that the method could have a high impact in the EM field in the future. The authors acknowledge funding [4].

References:

- [1] Q Jin *et al*, *Structure* (2014), Available online 6 February 2014, ISSN 0969-2126, <http://dx.doi.org/10.1016/j.str.2014.01.004>.
 [2] SC Harrison *et al*, *Nature* **276** (1978), p. 368.
 [3] JA Velazquez-Muriel *et al*, *Ultramicroscopy* **96** (2003), p. 17.
 [4] The work was partially funded by the CNRS and the CSIC [PICS 2011 to S.J. and C.O.S.S.]; the French National Research Agency ANR [ANR-11-BSV8-010-04 to S.J.]; and the European Social Fund and the Ministerio de Educación y Ciencia ["Ramón y Cajal" fellowship to C.O.S.S.];

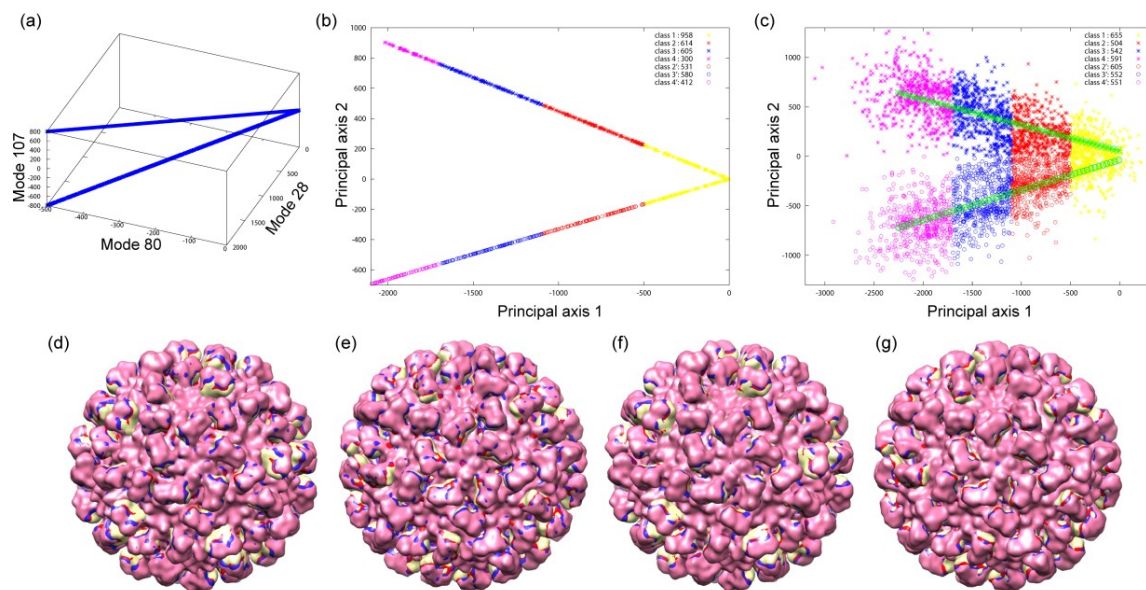


Figure 1. Illustration of the HEMNMA performance using simulated data. (a) Ground-truth synthetic deformation amplitudes along three modes of the TBSV. (b-c) PCA of the ground-truth deformation amplitudes (b) and PCA of the computed deformation amplitudes (c), with two linear-regression lines (green asterisks and circles for positive and negative coordinates on the principal axis 2, respectively (c)). (d-g) Overlap of the reconstructed structures, for the ground-truth alignment parameters [(d) for the classes 1-4; (e) for the classes 1, 2'-4'] and for the computed parameters [(f) for the classes 1-4; (g) for the classes 1, 2'-4']. The meaning of "class N : M" is that the class N contains M images (b-c). Same color code is used for the volumes (d-g) as for the image classes (b-c). The TBSV view is aligned with the 5-fold symmetry axis (d-g).