Hybrid Electron Microscopy Normal Mode Analysis graphical interface and protocol

Carlos Oscar S. Sorzano [a], José Miguel de la Rosa-Trevín [a], Florence Tama [b], Slavica Jonić [c,*]

[a] Biocomputing Unit, Centro Nacional de Biotecnología – CSIC, 28049 Madrid, Spain
[b] RIKEN, Advanced Institute for Computational Sciences, Kobe, Hyogo 650-0047, Japan
[c] IMPMC, Sorbonne Universités – CNRS UMR 7590, UPMC Univ Paris 6, MNHN, IRD UMR 206, 75005 Paris, France

ABSTRACT

This article presents an integral graphical interface to the Hybrid Electron Microscopy Normal Mode Analysis (HEMNMA) approach that was developed for capturing continuous motions of large macromolecular complexes from single-particle EM images. HEMNMA was shown to be a good approach to analyze multiple conformations of a macromolecular complex but it could not be widely used in the EM field due to a lack of an integral interface. In particular, its use required switching among different software sources as well as selecting modes for image analysis was difficult without the graphical interface. The graphical interface was thus developed to simplify the practical use of HEMNMA. It is implemented in the open-source software package Xmipp 3.1 (http://xmipp.cnb.csic.es) and only a small part of it relies on MATLAB that is accessible through the main interface. Such integration provides the user with an easy way to perform the analysis of macromolecular dynamics and forms a direct connection to the single-particle reconstruction process. A step-by-step HEMNMA protocol with the graphical interface is given in full details in Supplementary material. The graphical interface will be useful to experimentalists who are interested in studies of continuous conformational changes of macromolecular complexes beyond the modeling of continuous heterogeneity in single particle reconstruction.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

To perform biological functions, macromolecular assemblies change their shapes driven by different factors such as hydrogen bonding, ionic interactions, Van der Waals forces, hydrophobic packing and interactions with different partners. The alternative structures are referred to as different conformations, and the transitions between them conformational changes. Structural dynamics elucidation by determining the alternative conformations can be essential for understanding functions of large macromolecular complexes. Although single-particle electron microscopy (EM) has shown to be promising in capturing the alternative conformations (Elad et al., 2008; Fu et al., 2007, 2011; Grob et al., 2006; Penczek et al., 2006; Scheres et al., 2007; Simonetti et al., 2008), describing gradual (continuous) conformational changes is still a major challenge. This is the case when the complex cannot be stabilized in one or a few discrete conformational states or the

microscope is used to study full dynamics of the complex that is allowed to freely change its conformation.

In our previous work, we have shown that Normal Mode Analysis (NMA) of a reference structure can be used to analyze EM images of large macromolecular complexes and obtain an overall view of the conformational distribution for elucidating continuous conformational changes (Jin et al., 2014). In the approach referred to as Hybrid Electron Microscopy Normal Mode Analysis (HEMNMA), normal modes are used within elastic alignment of EM images with a reference structure (atomic-resolution structure or EM structure), without image classification or 3D reconstruction, as each image is allowed to contain a unique particle conformation. Structures can be reconstructed from images after discretizing the output, full conformational distribution. The full conformational distribution allows modeling of transition pathways and analyzing the conformational changes more extensively than with the classic, discrete methods (*e.g.*, Scheres et al., 2007), as the classic methods give only a few conformations. The discrete methods can be combined with HEMNMA to extend the analyses, as was shown in Jin et al. (2014).

Although HEMNMA is a powerful methodology, its usage has until now been limited to researchers who are at ease with

* Corresponding author at: IMPMC-UMR 7590, Université Pierre & Marie Curie, Case courrier 115, 4 Place Jussieu, 75005 Paris, France. Fax: +33 1 44 27 37 85.
   E-mail address: Slavica.Jonic@impmc.upmc.fr (S. Jonić).

installation of different software on Linux operating system and learning how to use it without a user-friendly graphical interface or guidance. Indeed, HEMNMA required the usage of a collection of methods that were not coming from the same software package and that were run on the console without a graphical interface or, in some cases, graphical interface was available on web servers. For instance, the elastic image alignment method is a part of the open-source software package Xmipp since Xmipp version 3.0.1 (http://xmipp.cnb.csic.es; for general references to the package, please see de la Rosa-Trevin et al. (2013), Scheres et al. (2008), Sorzano et al. (2004b)). It used to be run without graphical interface, through a Linux command-line program obtained after the Xmipp compilation. The values of a number of parameters had to be set by the user according to the documentation provided on the Xmipp web site. Also, normal modes of atomic structures used to be computed either using *ElNemo* web server (Suhre and Sanejouand, 2004) or by locally compiling and running several programs involved in the *ElNemo* methodology. Normal modes of EM structures used to be computed either using *3DEM Loupe* web server (Nogales-Cadenas et al., 2013) or by locally compiling and running several programs involved in the *3DEM Loupe* methodology. The advantage of using the web servers was the available user-friendly graphical interface and user guidance. However, the web servers could provide help for only one step of the HEMNMA methodology. Also, the results obtained on the web server had to be integrated in the rest of the workflow of the HEMNMA methodology manually by the user.

The HEMNMA methodology involves other steps such as statistical analysis of the results of the elastic alignment, clustering, 3D reconstruction, and identification and animation of characteristic trajectories of the conformational changes. These steps used to be done in GNU Octave (Eaton et al., 2009), with scripts written by the user. Thus, the entire workflow used to be managed by the user fully and required his/her full attention.

In this article, we describe the graphical interface of HEMNMA that was recently developed to facilitate the practical use of the methodology. The interface simplifies the use of HEMNMA and helps understanding the workflows behind different HEMNMA steps. Also, it allows easier setting of values of different parameters of a run, and easier start and comparison of different runs. The main part of the graphical interface was implemented in Xmipp version 3.1. The source files for NMA of atomic and EM structures (Nogales-Cadenas et al., 2013; Suhre and Sanejouand, 2004) were incorporated into Xmipp (version 3.1) so that the NMA programs can now be accessed and run as any other Xmipp program. Only a part of the interface was implemented in MATLAB (MATLAB and Statistics Toolbox Release 2011b, The MathWorks, Inc., Natick, Massachusetts, United States), and it allows an interactive analysis of results of an advanced statistical analysis step. The graphical interface uses Chimera (Goddard et al., 2007; Pettersen et al., 2004; Yang et al., 2012) and VMD (Hsin et al., 2008; Humphrey et al., 1996) to visualize the input and output PDB-format files and EM density volumes, and output animations.

This integral graphical interface provides the experimentalists in single-particle EM with an easy way to explore continuous heterogeneity of samples and it forms a direct connection to the single-particle reconstruction process. As an integral tool for capturing continuous motions of macromolecular complexes from single-particle EM images, it is an important addition to the single-particle EM field.

In the article, the main workflow of the interface is described together with the logic implemented in three main modules, which is necessary to know for a correct interpretation and usage of the output of each module. A detailed step-by-step protocol for using HEMNMA with the graphical interface is given in Supplementary material. The usage is shown with a simple and intuitive synthetic data set of adenalyte kinase that was specifically designed so that the readers can easily catch the whole workflow and see how to interpret the results of many available graphics outputs.

2. Results and discussion

The developed graphical interface to HEMNMA comprises the following three modules: (1) *Normal Mode Analysis*; (2) *Flexible Alignment*; and (3) *Advanced Results Analysis*. The main workflow of this integral interface is shown in Fig. 1. The workflow requires an intermediate-resolution EM structure or atomic-resolution structure (reference structure) and a set of single particle images. The ML3D, Multireference Projection Matching or Relion methods (Scheres, 2012; Scheres et al., 2008, 2007) can be used to obtain a reference EM structure if the structure at atomic resolution is unavailable. ML3D and Multireference Projection Matching are a part of Xmipp and Relion can be accessed through it.

After computing and analyzing normal modes, in the *Normal Mode Analysis* module, the user can either stop the processing or continue it by selecting the normal modes to be used for image analysis. Finally, the user can perform the image analysis and visualize the image analysis results in the *Flexible Alignment* module. The image analysis results can further be analyzed using the *Advanced Results Analysis* module. One can run in parallel multiple copies of the same workflow using different data sets or parameters. The image analysis part of the *Flexible Alignment* module is MPI-parallelized, as it is the most time-consuming part of the workflow. The other parts of the workflow perform a single-core processing. We start this section by describing each of the three modules of the workflow and we finish it by showing an example of using the graphical interface and some of the most important graphics output.

2.1. Normal Mode Analysis module

The *Normal Mode Analysis* module allows computing normal modes of an input atomic or EM structure as well as an interactive visualization of the computed modes by animating displacements of the structure along the modes. Finally, it allows selecting the modes to be used in the *Flexible Alignment* module.

2.1.1. EM volume-to-pseudoatoms conversion

If an EM structure is used to compute normal modes, the EM volume is converted into a pseudoatomic structure with the method that also runs on *3DEM Loupe* web server (Nogales-Cadenas et al., 2013). The input volume is converted into a collection of Gaussian functions (pseudoatoms) with a given standard deviation (pseudoatom size) so that the pseudoatomic structure converted back to a volume approximates the input volume with a given error. In the context of using normal modes for image alignment, it has been shown that smaller radius and approximation errors (higher numbers of pseudoatoms) result in higher resolutions of computed projections, and thus, more precise image alignments (Jin et al., 2014). In the case of heavy background noise, a masked EM structure can be used for the volume-to-pseudoatoms conversion. The pseudoatomic structure and the computed modes are written in the format that is readable by the *Flexible Alignment* module.

2.1.2. Computation of normal modes, collectivity, and score

We use the elastic network model to compute normal modes, which does not require energy minimization (the reference conformation is considered to be at the energy minimum). Normal Mode Analysis requires the diagonalization of a $3N \times 3N$ matrix of second derivatives of the potential energy (Hessian matrix), where N is the