



HAL
open science

Iterative Elastic 3D-to-2D Alignment Method Using Normal Modes for Studying Structural Dynamics of Large Macromolecular Complexes

Qiyu Jin, Carlos Oscar S. Sorzano, José Miguel de La Rosa-Trevín, José Román Bilbao-Castro, Rafael Nunez-Ramirez, Oscar Llorca, Florence Tama, Slavica Jonic

► **To cite this version:**

Qiyu Jin, Carlos Oscar S. Sorzano, José Miguel de La Rosa-Trevín, José Román Bilbao-Castro, Rafael Nunez-Ramirez, et al.. Iterative Elastic 3D-to-2D Alignment Method Using Normal Modes for Studying Structural Dynamics of Large Macromolecular Complexes. *Structure*, 2014, 22 (3), pp.496-506. 10.1016/j.str.2014.01.004 . hal-01430531

HAL Id: hal-01430531

<https://hal.science/hal-01430531>

Submitted on 13 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Iterative elastic 3D-to-2D alignment method using normal modes for studying structural dynamics of large macromolecular complexes

Qiyu Jin ¹, Carlos Oscar S. Sorzano ², Jose Miguel de la Rosa-Trevín ², José Román
Bilbao-Castro ³, Rafael Núñez-Ramírez ⁴, Oscar Llorca ⁴, Florence Tama ⁵ and
Slavica Jonić ^{1,*}

¹ IMPMC, UMR 7590, CNRS-Université Pierre & Marie Curie-IRD, 75252 Paris Cedex 05,
France.

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

³ Dept. Informática, Universidad de Almería, 04120 Almería, Spain.

⁴ Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas
(CSIC), 28040 Madrid, Spain.

⁵ RIKEN, Advanced Institute for Computational Sciences, Kobe, Hyogo, 650-0047, Japan.

* To whom correspondence should be addressed. Tel : +33 1 44 27 72 05 ; Fax : +33 1 44
27 37 85 ; Email : Slavica.Jonic@impmc.upmc.fr

Present address: Qiyu Jin, Institute of Image Processing and Pattern Recognition, Shanghai
Jiao Tong University, 200240 Shanghai, China; Rafael Núñez-Ramírez, Instituto de
Estructura de la Materia, IEM-CSIC, 28006 Madrid, Spain.

Running title: Iterative elastic 3D-to-2D alignment

Highlights:

- HEMNMA fully analyzes gradual conformational changes of complexes from EM images
- Computed conformational distribution allows modeling of transition pathways
- HEMNMA gives a full dynamics while other EM methods give only a few conformations
- Gradual changes are analyzed more extensively by HEMNMA than by other EM methods

Summary: This paper presents a method to study large-scale conformational changes by combining electron microscopy (EM) single-particle image analysis and normal mode analysis (NMA). It is referred to as HEMNMA, which stands for Hybrid Electron Microscopy Normal Mode Analysis. NMA of a reference structure (atomic-resolution structure or EM volume) is used to predict possible motions that are then confronted with EM images within an automatic iterative elastic 3D-to-2D alignment procedure to identify actual motions in the imaged samples. HEMNMA can be used to extensively analyze the conformational changes and may be used in combination with classic discrete procedures. The identified conformations allow modeling of deformation pathways compatible with the experimental data. HEMNMA was tested with synthetic and experimental datasets of *E. coli* 70S ribosome, DNA polymerase Pol α - B subunit complex of the eukaryotic primosome, and Tomato Bushy Stunt Virus.

INTRODUCTION

Normal mode analysis (NMA) of structures relies on modeling complex dynamics by a linear combination of harmonic oscillations around an equilibrium structural conformation. NMA was extensively used in electron microscopy (EM), for predicting functional motions from non-atomic-resolution EM structures (Chacon et al., 2003; Ming et al., 2002; Tama et al.,

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

2002) and for flexible fitting (Suhre et al., 2006; Tama et al., 2004a; Tama et al., 2004b) of atomic-resolution structures (e.g., from X-ray crystallography) into EM structures (EM density volumes). Normal modes of EM density volumes were shown to provide a good approximation of atomic-resolution normal modes in the low-frequency range that generally contains the modes reflecting experimentally observed large-scale conformational changes (Tama et al., 2002). They are thus useful for predicting conformational dynamics of complexes whose structure at atomic resolution is unavailable but an intermediate-resolution structure can be obtained by EM (Chacon et al., 2003; Ming et al., 2002; Tama et al., 2002).

Single particle analysis (SPA) EM techniques have been traditionally used in determining structures of large macromolecular complexes (diameter: 10-30 nm) from samples prepared biochemically to ideally contain all complexes in the same steady conformation. The high homogeneity of conformations is required for a high resolution of the structure reconstructed from images. As the homogeneous population is sometimes difficult to obtain even with the most improved biochemical procedures, image processing methods have been developed to separate images into a few classes and determine the ones corresponding to the steady conformations that are being studied (Elad et al., 2008; Grob et al., 2006; Penczek et al., 2006; Scheres et al., 2007; Simonetti et al., 2008). The goal of these standard approaches is thus to identify the most homogeneous classes of images in terms of conformations and compute the average structure from the classes. More accurate classification methods combined with the biochemical sample preparation to reduce the number of possible conformations should result in a higher homogeneity of classes and a higher resolution of computed structures. As the expected number of conformations is small in such cases, the standard methods are generally used with a small number of classes defined by the user initially. The standard methods thus allow studying only a few discrete conformations of the complex and new developments are needed for studying a large range of conformations with many intermediate conformations. This problem arises when studying a full dynamics of the complex that can freely change the conformation or when the biochemical stabilization of the

1 complex in a few steady conformations is impossible. In such cases, the standard methods
2 provide the dynamics information biased by the mentioned initial discretization into a small
3 number of classes (the reconstructed structures are not necessarily discrete samples of the
4 conformational change but they may result from mixing different conformations into the
5 selected number of classes).
6
7
8
9

10
11
12 In this paper, we propose a method that was specifically designed for studying a large range
13 of conformations with many intermediate conformations. It allows analyzing highly
14 heterogeneous populations of complexes with slow and continuous conformational changes
15 (gradual conformational changes where a discrete number of states cannot fully describe the
16 continuum of conformations). Given a reference structure (an atomic-resolution structure or a
17 reduced representation of an EM structure), a set of normal modes of the structure, and a set
18 of images, projection-matching-based elastic alignment of each single-particle image with the
19 reference structure provides an overall view of the conformational distribution. A potential of
20 combining NMA with EM image analysis was demonstrated by Brink et al (2004) who
21 compared several different conformations observed in images with the conformational
22 changes predicted by NMA, through a visual evaluation of their consistency. Here, we
23 describe an automatic method to determine conformations from images using normal modes.
24 The proposed method is based on a procedure for iterative elastic projection matching. This
25 procedure allows processing a high number of images to determine all actual confirmations
26 and evaluate their pertinence, which is impossible to achieve with other EM methods or X-ray
27 crystallography. The proposed method has the following four main features: 1) the
28 conformational modeling is done by displacing the reference structure along combinations of
29 normal modes and the amplitudes of the displacement are computed through the elastic
30 image alignment; 2) each single-particle image may represent a unique conformation; 3)
31 classification and 3D reconstruction are not mandatory; and 4) classic discrete EM methods
32 may be used to compute the reference structure when it is unavailable at atomic resolution
33 and/or to identify a subset of images to analyze. Thanks to these features, this new method
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 allows modeling of deformation pathways compatible with the experimental data and
2 analyzing the conformational changes more extensively than using the classic discrete EM
3 methods. The proposed methodology is referred to as HEMNMA, which stands for Hybrid
4 Electron Microscopy Normal Mode Analysis. It was tested here using several synthetic and
5 experimental datasets (synthetic and EM images of *E. coli* 70S ribosome (70S), EM images
6 of DNA polymerase Pol α and B subunit complex (Pol α – B) of the eukaryotic primosome,
7 and EM images of Tomato Bushy Stunt Virus (TBSV)).
8
9
10
11
12
13
14
15
16

17 RESULTS

22 HEMNMA design

23
24
25
26 The workflow (**Fig. 1A**) comprises the following four steps: 1) NMA of a reference structure at
27 atomic or pseudo-atomic resolution (see **Experimental procedures** for the volume-to-
28 pseudoatoms conversion and NMA methods used here); 2) selection of normal modes; 3)
29 iterative elastic 3D-to-2D alignment of single-particle images with the reference structure;
30 and 4) visualization of computed conformations. The preprocessing block in **Figure 1A** may
31 include any preprocessing of single-particle images (*e.g.*, correction of the effects of the
32 electron-microscope contrast transfer function (CTF)).
33
34
35
36
37
38
39
40
41
42
43

44 ***Selection of normal modes:*** The computation of normal modes takes negligible time with
45 respect to the analysis of a long series of images using normal modes (*e.g.*, 10000 single-
46 particle images). To reduce image analysis time while focusing on the most biologically
47 relevant conformational changes, normal modes can be selected based on some *a priori*
48 knowledge about the conformational change (*e.g.*, two known motions of 70S, ratchet-like
49 and L1-stalk motions, were used in **Experiments 1-2**). If no *a priori* knowledge is available,
50 lowest-frequency modes with the collectivity degree above a threshold can be selected as
51 highly collective low-frequency modes have been shown to be relevant to functional
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 conformational changes (Delarue and Dumas, 2004; Suhre et al., 2006; Tama and Brooks,
2 2006; Tama et al., 2004a; Tama and Sanejouand, 2001; Wang et al., 2004). For instance,
3
4 the overlap between experimentally observed conformational changes (difference between
5
6 two atomic conformations of the same complex) and the modes computed for one of the
7
8 conformations of the complex has showed that 1-3 low-frequency modes with the highest
9
10 overlap are highly collective and usually capture 60 to 70% of the conformational change
11
12 (Tama and Sanejouand, 2001; Tama et al., 2002). The most relevant modes are generally
13
14 among the 20 lowest-frequency modes for low-symmetry structures (Tama and Sanejouand,
15
16 2001; Tama et al., 2002). They move to higher frequencies (due to degeneracy) for
17
18 symmetric structures but remain among the 100 lowest-frequency modes, even for highly
19
20 symmetric structures such as icosahedral viruses (Tama and Brooks, 2005). The number of
21
22 lowest-frequency modes to compute can thus generally be fixed to 100. The collectivity
23
24 degree is computed to count the number of (pseudo-)atoms that are significantly affected by
25
26 the mode (Bruschweiler, 1995). Given the number of (pseudo-)atoms N , the collectivity
27
28 degree is normalized between $1/N$ and 1. The collectivity degree approaches 1 for maximally
29
30 collective movements, whereas for localized motions (few atoms move), it approaches 0. The
31
32 most collective modes capturing functional conformational changes can generally be
33
34 selected by setting the collectivity threshold to 0.5 (**Experiment 3**). Higher values of the
35
36 threshold should be chosen to select a reduced number of modes for highly symmetric
37
38 structures (**Experiment 4**). For instance, in the case of “swelling” icosahedral-symmetry
39
40 viruses, many modes have collectivities above 0.5 (**Experiment 4**). However, a few of them
41
42 are usually enough to describe functional conformational changes (e.g., the mode describing
43
44 best the experimentally observed conformational changes is related to a radially symmetric
45
46 expansion of the capsid (Tama and Brooks, 2005)). The most relevant modes can be
47
48 selected using the collectivity threshold of 0.75 (**Experiment 4**).

55
56
57 ***Iterative elastic 3D-to-2D alignment method:*** The amplitudes of the reference-structure
58
59 displacement (deformation) along normal modes and the orientation and the position of the
60
61

1 deformed structure are refined until it is found a projection of the deformed-reference density
2 volume that is the most similar to the analyzed single-particle image. The method is based
3 on a local-search Powell's UOBYQA optimization (Powell, 2002) that uses quadratic
4 approximation for building a local model of the objective function. It consists of repeating the
5 following three steps iteratively until the change in elastic parameters (deformation
6 amplitudes) and rigid-body parameters (orientation and position of the deformed structure)
7 between two successive iterations becomes insignificant or a specified number of iterations
8 is reached (**Fig. 1B**):

- 17 • Step 1: Find the minimum of the current local model of the objective function and
18 move the current point (current vector of deformation amplitudes) to this minimum.
- 22 • Step 2: Evaluate the objective function at the new vector of deformation amplitudes
23 (new point). The structure is first deformed with the new deformation amplitudes and
24 the modified atomic or pseudo-atomic coordinates are converted into a volume. The
25 computation of volumes from pseudo-atomic coordinates is based on the volume
26 representation by a sum of Gaussian functions (Nogales-Cadenas et al., 2013). The
27 computation of volumes from atomic coordinates uses electronic-form atomic factors
28 (Peng et al., 1996). The volume is subsequently aligned (rigid-body) with the single-
29 particle image using a rough alignment based on a discrete library of reference
30 projections of the volume (Sorzano et al., 2004a) and a continuous refinement using
31 a gradient-based approach (Jonic et al., 2005). The alignment quality determined by
32 a measure of similarity between the single-particle image and the corresponding best-
33 matching volume projection is considered as the new evaluation of the objective
34 function for the optimization.
- 51 • Step 3: Reconstruct the local model of the objective function around the new point
52 with the new evaluation. Go back to Step 1.

58 The source code of the method is available in the open-source software Xmipp (Sorzano et
59 al., 2004b), starting from the version *Xmipp 3.0.1*. The current implementation takes on

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

average around 2, 4, and 6 minutes to analyze an image of size 128x128 pixels using 1, 2, and 3 normal modes, respectively, on a core of a Dual Intel Xeon X5472 processor (3.00GHz). Note that the execution time was measured with 1-3 modes but the program allows using any number of modes. An MPI-parallelized version of the method allows an efficient analysis of a large number of images (xmipp_mpi_nma_alignment whose tutorial is at http://xmipp.cnb.csic.es/twiki/bin/view/Xmipp/Nma_alignment_v3), which was tested on standard-size clusters and super-computer centers.

Visualization of computed conformations: The computed conformations can be visualized by displacing the reference atomic or pseudo-atomic structure along normal modes using the amplitudes of the displacement computed through the elastic image alignment (direct visualization). Also, the animated volumes can be obtained by converting displaced atoms or pseudo-atoms into volumes as in Step 2 of the elastic 3D-to-2D alignment method. Optionally, 3D reconstruction from images can be obtained by separating images into classes with similar conformations (similar values of the computed displacement amplitudes). However, for highly heterogeneous samples, the underlying dynamics can only be partially elucidated by analyzing these reconstructed structures due to a discretization of the conformational distribution (discrete number of classes) and averaging during the reconstruction. A direct visualization shall be a preferred choice in such cases, as we show with Pol α – B (**Experiment 3**). For more details on the visualization methods, see the corresponding section of **Supplemental Experimental Procedure S1**.

Experiment 1: Synthetic continuous-type conformational change

In this experiment, we show robustness of HEMNMA to noise and to the CTF effects remaining after the phase inversion that is usually used for CTF correction. A variety of conformations were sampled from a synthetic continuous-type conformational change determined by deforming an atomic-resolution *E. coli* 70S structure (**Fig. S1A**) referred to as

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

3I1OP (without ligands, in the conformation in which the 30S subunit is non-rotated with respect to the 50S subunit (Zhang et al., 2009)). Synthetic images (size: 128^2 pixels, pixel size: $(3 \text{ \AA})^3$) were computed as 300 random projections of the atomic structure displaced with random displacement amplitudes along two of its normal modes while preserving a constant ratio between the two displacement amplitudes (for details on the random image generation with a linear relationship between the contributions of the two modes, see **Supplemental Experimental Procedure S2**). The used modes were mode 11 as it describes the ratchet-like motion (rotation of 30S with respect to 50S visible in **Movie S1**) and mode 12 as it describes the L1-stalk motion. The modes responsible for these two motions have already been characterized computationally (Tama et al., 2003; Wang et al., 2004) and they were chosen here as the combined motion was already observed experimentally (Fu et al., 2011; Valle et al., 2003). The overall characteristics of the modes used in this paper are similar to these reported previously although the mode numbers are not the same because of different structures and methods used for calculations.

The amount of noise before and after CTF was adjusted for each tested defocus value (0.5 μm , 1 μm , and 2 μm) so that the images were obtained with the signal-to-noise ratio (SNR) of 0.3 or 0.1. The different image sets (one without noise and six for different defocus and SNR values (**Fig. S1D**)) were analyzed using only the mode capturing ratchet-like motion to see whether the proposed method can recover both synthesized motions, ratchet-like and L1-stalk motions. The elastic alignment of images was done with respect to the given atomic structure and the pseudo-atomic structure obtained from a synthetic volume of the atomic structure (**Fig. S1C**). **Movie S2** shows the animated pseudo-atomic mode 13 (ratchet-like mode) used in image analysis. Details of the pseudo-atomic structure and normal modes computations can be found in **Supplemental Experimental Procedure S2**.

The image alignment with a pseudo-atomic structure is more challenging than the alignment with an atomic structure as pseudo-atomic and atomic normal modes have different

1 resolutions and the same movement often appears at a different frequency in the two sets of
2 modes (Tama et al., 2002). Here, a more difficult alignment in the pseudo-atomic case was
3 expected also because different structures were used to generate images (atomic structure)
4 and to analyze them (pseudo-atomic structure) while the same structure was used to
5 generate and analyze images in the atomic case. However, such a difficulty is expected in
6 practice as the resolution of experimental images is usually higher than the resolution of the
7 reference structure when using EM structures. The exploration of only one movement
8 instead of two that are present in images is also challenging but close to realistic as, in
9 practice, images are expected to be analyzed using only a few normal modes (sometimes,
10 only one as in **Experiment 2**). Interestingly enough, although the used mode was selected
11 based on previously observed 70S motions instead of using the modes collectivity criterion,
12 this mode is actually the most collective one for each of the used 70S structures in
13 **Experiments 1-2**.

14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31 The results in the atomic and pseudo-atomic cases show that the accuracy of estimation of
32 rigid-body and elastic parameters decreases as SNR decreases and the defocus increases.
33 Although the maximum value (over 7 image sets) of the relative deformation-amplitude error
34 doubled in the pseudo-atomic case compared to the atomic case (8.9% for pseudo-atoms
35 with respect to 4.2% for atoms (**Table S1A**)), the accuracy is globally good, with the absolute
36 angular and shift errors that are similar in the two cases (below 0.6° and 0.15 pixel for
37 pseudo-atoms; below 0.5° and 0.15 pixel for atoms (**Table S1B**)). In the most difficult case
38 (SNR=0.1, defocus=2 μm), the RMSD between the structures with computed and ground-
39 truth deformation amplitudes along the mode describing the ratchet-like motion is about 0.46
40 Å for pseudo-atoms and 0.22 Å for atoms, which is much below the commonly achievable
41 resolution in EM. **Table S1C** summarizes how the key parameters of the volume-to-
42 pseudoatoms conversion and NMA methods and the resulting alignment error depend on the
43 volume approximation error (for more details on the influence of these parameters, see
44 **Supplemental Experimental Procedure S2**).

Experiment 2: Mixed EM images of *E. coli* 70S ribosome with and without EF-G

In this experiment, HEMNMA was used to analyze cryo-EM images of a mixture of *E. coli* 70S ribosome with and without Elongation Factor G (EF-G) bound. These images were used in previous EM studies where ratchet-like and L1-stalk motions were observed on the reconstructions corresponding to the ribosome with and without EF-G (Elad et al., 2008; Scheres et al., 2007). Here, the images were analyzed with the ratchet-like mode and separated into the classes with and without EF-G using the computed deformation amplitudes along the mode. Indeed, our tests with synthetic data (generated using EM structures of three different 70S conformations) showed that this mode alone can be used to classify images of the ribosome at distinct states of the conformational change combining ratchet-like and L1-stalk motions, and to reconstruct the structure at distinct states of the conformational change (for details of these tests, see **Supplemental experimental procedure S3** and **Figure S2**).

The dataset contained 10000 cryo-EM images (size 130x130 pixels, pixel size $(2.82 \text{ \AA})^2$) and a result of a supervised image classification into two subsets, with or without EF-G (5000 images per subset). Ribosomes without EF-G contain three tRNAs in the classical A/A, P/P and E/E positions and are in the normal (unratcheted) conformation. Ribosomes with EF-G contain a single tRNA in the hybrid P/E position, and are ratcheted. The unratcheted-conformation EM structure EMD-5262 (Fu et al., 2011) was used to compute the pseudo-atomic structure and its normal modes. Details of the computations are given in **Supplemental Experimental Procedure S3** as the same structure and normal modes were used there with synthetic images capturing ratchet-like and L1-stalk motions from authentic EM structures. Image analysis was done using the pseudo-atomic structure and its mode 9 that describes ratchet-like motion (**Supplemental Experimental Procedure S3** and **Movie S3**).

1
2 The histogram of the computed amplitudes along the pseudo-atomic mode 9 (**Fig. 2A**
3 (green)) may be interpreted as a mixture of at least two Gaussian-type distributions as two
4 peaks can be noticed. **Figure 2A** also shows the histogram portions corresponding to the
5 image separation by supervised classification (with EF-G: red; without EF-G: blue). The bins
6 containing one third of images with the smallest amplitudes (mean amplitude: -401) are
7 indicated by red asterisks while the bins containing one third of images with the largest
8 amplitudes (mean amplitude: 237) are indicated by blue asterisks (**Fig. 2A**). The structures
9 reconstructed from images in the red and blue bins clearly show two distinct conformations,
10 with and without EF-G, respectively (**Fig. 2B-C**). The structure from the mid amplitude range
11 (mean amplitude: -78) seemed to result from a mixture of different conformations with
12 partially bound EF-G and was not further analyzed in this paper.
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27

28 The reconstructed structures are in excellent agreement with those obtained with other
29 methods and the same dataset (Elad et al., 2008; Scheres et al., 2007) although the overall
30 quality of the obtained structures cannot be fully compared because of a different number of
31 analyzed images or a different number of images used in reconstruction (*e.g.*, around 90000
32 particle images were available in the first paper using this data (Scheres et al., 2007) while
33 only 10000 images were available here). A particularly interesting result is that the mode
34 related to the movement of the L1 stalk with respect to the 30S subunit was not explicitly
35 explored through the image analysis but this movement was recovered (**Fig. 2E**) by
36 classifying images based on the ratchet-like motion visible in **Fig. 2C-D**.
37
38
39
40
41
42
43
44
45
46
47
48
49
50

51 These results show robustness of the proposed method to the use of a reference structure
52 that comes from a different dataset or a different composition of the studied complex (the
53 number of tRNAs in the reference EMD-5262 structure is different from the number of tRNAs
54 in the analyzed images as well as the EF-G factor is present in some of the images but not in
55 the EMD-5262 structure).
56
57
58
59
60
61
62
63
64
65

Experiment 3: EM images of Pol α – B

In the absence of a structure at atomic resolution, Pol α – B was previously studied by analyzing negative-stain single-particle images of the complex with 3D maximum likelihood (ML3D) method (Scheres et al., 2007), which yielded to a structure composed of two lobes connected by a flexible linker where the larger and smaller lobes roughly correspond to Pol α and B subunit, respectively (Klinge et al., 2009). More precisely, ML3D revealed three different classes of images that contained 42%, 30%, and 28% of the total number of images and yielded to the reconstructions at 22.9 Å, 24.5 Å, and 25.5 Å resolution, respectively, each corresponding to a different degree of flexion of the linker (Klinge et al., 2009). The same images (12000 single particles) were analyzed here using HEMNMA. The structure obtained by ML3D from the largest class of images was converted into a pseudo-atomic structure that was used as the reference structure to analyze images. The volume and the images were of size 64^3 voxels (voxel size: $(3.8 \text{ \AA})^3$) and 64^2 pixels (pixel size: $(3.8 \text{ \AA})^2$), respectively. A view of the reference conformation is shown in **Figure 3A-C** (a structure with 968 pseudo-atoms was obtained for the EM volume approximation error of 5% and the standard deviation of Gaussian functions $\sigma = 3.8$).

From 100 computed normal modes, we selected six lowest-frequency non-rigid-body modes as their collectivity degree was above 0.5 (**Fig. 3D**). These 6 modes were used for the image analysis with the proposed method. The computed deformation amplitudes along the modes were then analyzed by principal component analysis (PCA) to reduce the complexity from 6 to 3 dimensions. To visualize how the conformation changes along a mode or a principal axis (PA), a volumetric form of the pseudo-atomic structure (obtained by pseudoatoms-to-volume conversion) was displaced in positive and negative directions of each of six normal modes and six principal axes (**Figure S3**). As the principal axes are linear combinations of normal modes, a contribution of different normal modes to the principal axes can be observed in

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Figure S3G-L. For example, the most important PA (principal axis 1: PA1; **Fig. S3G**) is mainly contributed by mode 7 (**Fig. S3A**) (a similar movement of the upper lobe is visible in **Fig. S3G** and **Fig. S3A**), although some other modes contribute as well (e.g., the linker is moving in **Fig. S3G** but not in **Fig. S3A**). Similarly, the second most important PA (principal axis 2: PA2; **Fig. S3H**) is mostly contributed by mode 8 (**Fig. S3B**). The remaining principal axes are more difficult to interpret visually and their quantitative analysis showed that they were more “mixed” than PA1-2 (*i.e.*, a normal mode has a similar contribution to a PA as the other modes).

The computed deformation amplitudes are shown in **Figure 3E** as projections onto PA1-3 (each image is represented by a point), together with linear regression lines through three identified clusters of points (**Fig. 3E,I**). Displacements of the volumetric form of the pseudo-atomic structure along each of the three lines (**Fig. 3F-H**) show a movement similar to that along PA2 (**Fig. S3H**). The main difference between the displacements along the three lines is related to the degree of an additional movement similar to that along PA1 (**Fig. S3G**). **Movies S4-5** show the displacement of the pseudo-atomic structure and its volumetric form along the red (central) line as the highest number of images was distributed along this line.

These results clearly show a high conformational heterogeneity of the complex, which suggests that any structure obtained by 3D reconstruction from these images would be of low resolution and partially inaccurate as the actual flexibility would be limited by averaging of different conformations. Here, we present some reconstructed structures to show these effects. A structure was computed from each of the three clusters of images (**Fig. 3I,K**). The overlapped structures (**Fig. 3K**) show a movement similar to that along PA1 (**Fig. S3G**). An overlap between the structures reconstructed along any of the three regression lines (**Fig. 3E**) shows a movement similar to that along PA2 (**Fig. S3H**). For example, **Figure 3L** shows an overlap of the structures reconstructed from three subsets (**Fig. 3J**) of the yellow (central) cluster in **Fig. 3I**. These reconstructed structures (**Fig. 3L**) are similar to those shown for the

1
2 displaced pseudo-atomic structure (**Fig. 3G**) along the corresponding central regression line
3
4 in **Fig. 3E** but their resolution is low due to a variability of the incorporated conformations in
5
6 the reconstruction.
7

8
9 As the reconstructed structures show movements similar to those along two most important
10
11 principal axes and these axes are mostly contributed by modes 7-8, we can conclude that
12
13 these two modes are the most dominant modes for this dataset. Interestingly enough, it has
14
15 been shown that 1-3 modes can be sufficient to globally describe experimentally observed
16
17 conformational changes (Tama et al., 2004b; Tama and Sanejouand, 2001). Our method
18
19 thus can help to find the most dominant modes, which may be particularly useful when they
20
21 are not among a few lowest-frequency modes (e.g., **Experiment 4**).
22
23
24
25

26
27 Regarding the reconstructions in **Figure 3K**, a similar movement was observed previously
28
29 (Klinge et al., 2009) but with a smaller amplitude. To check if the smaller amplitude was due
30
31 to an incorporation of different conformations into the same reconstructed structure, we
32
33 designed a test involving three more reconstructions. We split the central cluster in **Fig. 3I**
34
35 into three subsets of images according to their coordinate on PA1 (each subset contained
36
37 around 30% of images). Two structures were computed from two artificially mixed sets (each
38
39 containing a side cluster and its neighboring central-cluster subset). The third structure was
40
41 computed from the remaining central-cluster subset. Interestingly enough, these last three
42
43 structures overlap quite well with the conformations obtained previously (Klinge et al., 2009)
44
45 (similar movements with similar amplitudes can be observed in **Figure 3M**). The smaller
46
47 amplitude of the movement observed previously can thus be explained by a heterogeneity of
48
49 conformations incorporated in the reconstructed average structures (the flexibility could not
50
51 be fully explored with ML3D using a small number of classes and the full exploration would
52
53 require much larger number of classes and images (Klinge et al., 2009)). HEMNMA
54
55 described more extensively the conformational heterogeneity that was detected by ML3D
56
57 using the same dataset. For example, it offered the possibility to additionally visualize the
58
59
60
61
62
63
64
65

1 changes of the linker length coupled with the changes of the relative distance between the
2 lobes (movements reflecting mainly displacements along PA2 (**Fig. 3G,L** and **Fig. S3H**)).
3
4
5

6 **Experiment 4: EM images of TBSV**

7
8 Swelling of many icosahedral plant viruses has been observed upon changes of pH and
9 EDTA chelation of divalent cations on the interface between subunits (Ca²⁺ or Mg²⁺ that
10 appear to play a critical role in virion stability) but the swelling mechanism is still not well
11 understood (Witz and Brown, 2001). In a previous EM study of the TBSV swelling, a
12 structure of the native (compact) TBSV form was obtained at 13 Å resolution and a
13 continuous particle size variation was found on micrographs of the virus after its dialysis in
14 Tris–EDTA buffer (Aramayo et al., 2005). Also, a structure of the most swollen form was
15 obtained at 19 Å resolution by an iterative procedure whose each iteration comprised image
16 alignment, sorting of the largest particles (based on a threshold cross-correlation with the last
17 reference structure), and 3D reconstruction to compute a new reference structure for the next
18 iteration (Aramayo et al., 2005). HEMNMA was used here to analyze a set of 7000 TBSV
19 images.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

37 The EM structure and image sizes were reduced to 128³ voxels and 128² pixels, respectively
38 (pixel size: (3.2 Å)²). The pseudo-atomic structure for NMA was obtained from the structure
39 of the compact form (6096 pseudo-atoms for the volume approximation error of 5% and for
40 the standard deviation of Gaussian functions $\sigma = 3.2$) (**Fig. 4A**). From 100 computed normal
41 modes, image analysis was done using four lowest-frequency non-rigid-body modes whose
42 collectivity degree was above 0.75 (modes 13, 24, 40, and 77 marked by asterisks in **Figure**
43 **4B**). No information was used about potential symmetry of the conformations during the
44 image analysis and the used modes allowed for asymmetric geometrical transformations. In
45 fact, only one of the used modes deforms the structure while perfectly preserving its
46 symmetry (mode 24 that is related to the radially symmetric expansion). The other three
47 modes result in asymmetric changes. The deformation amplitudes computed through image
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 analysis for each of the four modes were then mapped onto the 3D PCA space (**Fig. 4C**).
2 The data were fitted with a linear regression line in the PCA space and the reference
3 structure was displaced along this line (blue asterisks in **Figure 4C, Movie S6**). Also, the
4 data were split into four classes according to the data coordinate on PA1 (4 equal segments
5 of the axis) and 3D reconstruction was done from the images in each class (**Fig. 4D**). Red,
6 magenta, green, and yellow colors in **Figure 4C-D** correspond to the four classes with 21%,
7 25%, 24%, and 30% of all images, respectively. Although the image analysis was done with
8 no *a priori* information about the conformational symmetry, the reconstructed structures were
9 icosahedrally symmetrized to reduce noise as it was done in the previous study (Aramayo et
10 al., 2005). The difference between the diameter of the structure from the most swollen
11 particles (yellow in **Figure 4D**) and that of the compact-form structure (grey in **Figure 4D**) is
12 about 4 nm, which is consistent with the previously reported differences (Aramayo et al.,
13 2005; Kruse et al., 1982). The difference in the diameter of the structure from the most and
14 least swollen particles (yellow and red in **Figure 4D**, respectively) is about 3.2 nm. This
15 indicates that the diameter of the least swollen particles in the analyzed images is larger than
16 that of the compact-form particle and that the particles were all more or less swollen in the
17 analyzed samples.

18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40 The conformational change can be mainly described as a combination of an increase in the
41 radius of the particle (increase in the negative direction of PA1), a rotation of the subunits
42 around the 5-fold symmetry axis, and an increase in the size of the intersubunit space at the
43 subunits quasi-trimer (central part of the capsid in **Fig. 4D** and **Movie S6**). These results are
44 consistent with the models of the TBSV conformational change proposed in the literature
45 (Aramayo et al., 2005; Kruse et al., 1982; Robinson and Harrison, 1982). Most importantly,
46 the visualization of the full conformational distribution revealed many intermediate states of
47 the change, which describes more extensively the conformational heterogeneity that was
48 detected previously (Aramayo et al., 2005).

DISCUSSION

1
2
3
4 The paper describes the method HEMNMA that allows analyzing dynamics of large
5 macromolecular complexes by experimental EM image analysis of possible conformational-
6 change directions predicted by normal mode analysis. HEMNMA is particularly useful in the
7 cases in which the reference structure at atomic resolution is unavailable but a structure can
8 be obtained by EM.
9

Full dynamics can be studied

10
11 HEMNMA is ideally suited to study full dynamics as it does not use classification during the
12 image analysis. The classification of the computed deformation amplitudes can be done at
13 the last step to compute structures with similar conformations. We showed that the
14 conformational transition conveyed by the reconstructed structures is consistent with the one
15 conveyed by the reference structure displaced along the trajectories fitted through the
16 computed deformation amplitudes. However, the full conformational distribution conveys
17 additional information that is not contained in the reconstructions obtained from the classes.
18 Therefore, when flexibility is described by a large range of conformations, a clear advantage
19 of our method is in looking for a full range of conformations contrary to the standard methods
20 that only look for discrete and limited number of conformational states.
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

HEMNMA robustness

44
45 In the ribosome experiments (**Experiments 1-2**), we analyzed images with the proposed
46 method using only the mode related to ratchet-like motion to see whether we could recover
47 both ratchet-like and L1-stalk motions. The motions were successfully recovered and
48 robustness of the method was showed to different realistic factors (e.g., different signal-to-
49 noise ratios, different contrast transfer functions, different resolutions and compositions of the
50 reference structure with respect to those of the imaged complexes).
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Complexes of different composition can be studied

1
2 The experiment with EM images of 70S ribosome (**Experiment 2**) was principally used to
3
4 show that the method can work with different compositions of complexes in the same sample
5
6 (with EF-G, without EF-G, and with partially bound EF-G) that, at the same time, produce
7
8 different conformations of the complexes. We refer to this kind of conformational variability to
9
10 as “discrete-type conformational changes” because the sample is prepared so that a small
11
12 number of conformational classes are expected and the problem can be solved by classical
13
14 (“discrete”) methods. This experiment shows that, in the case of such “discrete-type”
15
16 problems, our method can reproduce the results obtained using classical methods and the
17
18 same data (Elad et al., 2008; Scheres et al., 2007). The functional significance of these
19
20 results was reported in the first paper in which this dataset was analyzed (Scheres et al.,
21
22 2007).
23
24
25
26
27

Functional significance of observed motions

28
29
30 The experiments with Pol α – B and TBSV highlight the main characteristics of the
31
32 methodology, showing that it can be used to explore truly continuous-type conformational
33
34 changes where the complex cannot be stabilized in one or a few particular conformations.
35
36 They also showed how the method works when no *a priori* knowledge about the
37
38 conformational change is used for the selection of normal modes. The method can find
39
40 contributions of the used modes to the principal axes as well as their contributions to the
41
42 principal trajectories that determine the most dominant motions. Moreover, the methodology
43
44 is particularly useful in identifying the most dominant modes when they are not among a few
45
46 lowest-frequency modes (*e.g.*, TBSV case). These experiments are also interesting as they
47
48 show that the method allows studying conformational pathways that may actually exist but
49
50 may be difficult to solve by X-ray crystallography or classical EM methods that rely on a pre-
51
52 defined initial number of classes.
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Pol α – B is arranged as an elongated structure organized in two lobes connected by a flexible linker. This quaternary organization has been observed for other eukaryotic replicative polymerases such as DNA polymerase ϵ (Asturias et al., 2006) and DNA Polymerase δ (Jain et al., 2009), which are responsible for the replication of the leading and lagging strands respectively. In all these replicative complexes the flexibility between the different lobes has been identified as a functional characteristic that allows the accommodation of different conformations of DNA during its replication and facilitates the interactions of these complexes with other components of the replication machinery. An EM analysis of the structure of the complete eukaryotic primosome, consisting of Pol α , the B subunit and the two components of the primase (PriL and PriS) revealed a very flexible connection between the two lobes and that this complex was capable of adopting a large range of different conformations (Nunez-Ramirez et al., 2011). This flexibility was postulated as a mechanism for the transfer of the DNA substrate from the primase catalytic site placed in one of the lobes to the DNA polymerase catalytic site located in the other lobe. Surprisingly, ML3D image classification of Pol α – B, a sub-complex, comprising only two of the components of the primosome, revealed a limited flexibility between the catalytic core of Pol α (the big lobe) and the B-CTD platform (the small lobe) (Klinge et al., 2009). The distinct behavior of the full primosome complex and the Pol α – B sub-complex was difficult to rationalize. Now, the analysis of the same dataset with HEMNMA has revealed that the high degree of flexibility showed in the full complex is also present in the Pol α – B sub-complex. Intermediate conformations of the Pol α – B sub-complex were undetected by ML3D probably due to the difficulty to assign a large range of conformations to a discrete number of solutions. In addition, the classical refinement-based methods use an initial template and typically discard images that poorly correlate with the average 3D model obtained during the template refinement. Thus, if a dominant conformation was used to build the initial reference, we suspect that the particles corresponding to distant conformations would be removed during the refinement by classical methods. HEMNMA has provided a more precise

1 description of the continuous range of conformations of Pol α – B complexes, allowing a
2 better understanding of their structure-function relationship.
3
4
5

6 The visualization of the full conformational distribution of the TBSV helped to reveal many
7 intermediate conformations which was previously hypothesized but was impossible to
8 visualize (Aramayo et al., 2005). Indeed, many intermediate conformations might explain a
9 difficult crystallization of such swelling viruses (e.g., crystallography produced a swollen
10 TBSV structure at resolution of only 8 Å (Robinson and Harrison, 1982)). The Ca²⁺ ions are
11 known to be present between the subunits forming the TBSV capsid in order to strengthen
12 the links between them. Many different intermediate conformations may be explained by an
13 incomplete removal of these ions by EDTA (sample preparation is described in (Aramayo et
14 al., 2005)) that could lead to a partial opening of the viruses, partially hold by remaining Ca²⁺
15 ions. Similar conformations may exist as mechanical intermediate states permitting the
16 release of the viral RNA within the host cell. The genome of the TBSV consists of a single
17 RNA filament shifting towards the capsid in an environment depleted of Ca²⁺ ions and such
18 intermediate conformations could be enough for the release of the single RNA filament *in*
19 *vivo*.
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

40 Thus, HEMNMA seems highly suitable for macromolecular complexes displaying a
41 continuous range of conformations. This may be the case of a large number of complexes in
42 biology. In these conditions, HEMNMA seems significantly more powerful than the methods
43 with a pre-defined initial number of classes, which would require larger datasets and larger
44 computing times to evaluate the presence of a large number of possible conformers.
45
46 However, the classification-based methods can be combined with HEMNMA to better
47 analyze overall heterogeneity (e.g., by providing one or several reference structures as in the
48 Pol α – B case).
49
50
51
52
53
54
55
56
57
58
59

60 **EXPERIMENTAL PROCEDURES**

61
62
63
64
65

Pseudo-atomic structures from EM volumes

To obtain a reduced EM density representation, our recently published method for volume-to-pseudoatoms conversion (Nogales-Cadenas et al., 2013) was used as it allows controlling the volume approximation error and, thus, controlling quality of the projections for the elastic projection-matching-based image alignment. The used conversion method represents a volume by a collection of Gaussian functions whose number, positions, and amplitudes are adjusted for a given volume approximation error and a given Gaussian standard deviation. The Gaussian functions are referred to as pseudo-atoms although their positions do not generally coincide with the atomic positions, and the structure with pseudo-atoms is referred to as pseudo-atomic structure (Nogales-Cadenas et al., 2013).

Normal mode analysis

To compute normal modes, a simplified elastic network representation of the potential energy function (Tirion, 1996) was used, as implemented on EINemo web server (Suhre and Sanejouand, 2004) for atomic structures and on 3DEM Loupe web server (Nogales-Cadenas et al., 2013) for EM structures. Before computing normal modes, 3DEM Loupe converts the input EM volume into a pseudo-atomic structure that can be directly used as the reference structure for the method proposed here. The coordinates of the normal mode vectors are computed in Angstroms while the deformation amplitudes have no units. Six lowest-frequency modes (modes 1-6) are not used as they are related to rigid-body movements.

ACKNOWLEDGMENTS

This work was supported by the French National Center for Scientific Research and the Spanish National Research Council [CSIC2009FR0015 and PICS 2011 to S.J. and C.O.S.S.]; the Spanish Ministry of Economy and Competitiveness [ACI2009-10220, ACI2010-1088, BIO2010-16566 to J.M. Carazo, SAF2011-22988 to O.L.]; the European Social Fund and the Ministerio de Educación y Ciencia [“Ramón y Cajal” fellowship to

1 C.O.S.S. and J.R.B.C.]; the European Commission [DECI-6 to C.O.S.S. and S.J.]; and the
2 French National Research Agency [ANR-11-BSV8-010-04 to S.J.]. We thank GENCI-
3 CINES/IDRIS (France) for HPC resources [2011-x2010072174, 2012-x2010072174 to S.J.],
4 J. Frank and H. Gao (Columbia University, New York, USA) for depositing the 70S/EF-G
5 images on the EMDDataBank web site, L. Pellegrini and S. Klinge (Cambridge University, UK)
6 for providing the Pol α – B samples, E. Larquet and R. Aramayo (IMPMC, Paris, France) for
7 providing the TBSV data, and I. Callebaut (IMPMC, Paris, France) and J.M. Carazo (CNB,
8 Madrid, Spain) for stimulating discussions.

9
10
11
12
13
14
15
16
17 *Conflict of interest statement.* None declared.

18 19 20 21 22 **FIGURE LEGENDS**

23
24 **Figure 1:** Flowchart of the method for studying macromolecular dynamics (A) using a 3D-to-
25 2D normal-mode-based alignment of atomic or pseudo-atomic structures with single-particle
26 images (B).

27
28
29
30
31
32
33 **Figure 2:** Experiment with EM images of a mixture of *E. coli* 70S ribosome with and without
34 EF-G factor, and with the reference pseudo-atomic structure and normal mode shown in **Fig.**
35 **S2D** and **Movie S3**, respectively. (A) Histogram of computed deformation amplitudes (green)
36 along the normal mode with the contribution of images classified by supervised classification
37 as without EF-G (non-rotated conformation) and with EF-G (rotated conformation) indicated
38 in blue and red, respectively. The bins corresponding to the images used for computing the
39 two 3D reconstructions shown in (B)-(E) are indicated by two different colors of asterisks (red
40 and blue). (B) Structure reconstructed from the images corresponding to the bins indicated
41 by red asterisks, with EF-G and P/E-position tRNA indicated with blue and red mesh,
42 respectively. (C) Overlapped structure shown in (B) (blue mesh) with the structure
43 reconstructed from the images corresponding to the bins with blue asterisks (grey solid), with
44 A/A-, P/P- and E/E-position tRNAs indicated with mesh in green, yellow, and red,
45 respectively (one large part of the EF-G region is indicated by an arrow on the blue mesh,
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

which helps to show that the region is occupied by no density from the grey solid structure).

(D) Overlapped structures from (C) showing different relative orientations of the 30S and 50S subunits in the two reconstructions *i.e.* ratchet-like movement (grey: with EF-G; tan: without EF-G; magenta asterisk: density belonging to EF-G). (E) Second view of the overlapped structures from (D) showing (arrows) the 30S rotation and the L1 movement (grey solid: with EF-G; tan mesh: without EF-G). Although the L1 movement is rotational, it is shown as a shift to designate an inward direction of the movement. See also **Figure S2**.

Figure 3: Experiment with EM images of Pol α – B. (A) EM volume obtained by ML3D method (from the largest class of images). (B) Pseudo-atomic structure (spheres) from the volume in (A) overlapped with the volume (mesh). (C) Approximation of the volume shown in (A) by the pseudo-atomic structure shown in (B). (D) Collectivity degree of normal modes computed for the pseudo-atomic structure (the first six modes related to rigid-body movements are not shown; for better visibility, only 40 modes are shown; the asterisks denote the modes with the collectivity degree above 0.5; see also **Figure S3A-F**). (E) Computed deformation amplitudes projected onto the three most important principal axes (the points denote single-particle images and the most important principal axis is 1; see also **Figure S3G-L**), and linear regression lines (yellow, red, and green) through three clusters (see also (I)). (F-H) Volumetric form of the pseudo-atomic structure displaced along the lines shown in (E) ((F), (G), and (H) for the yellow, red, and green lines, respectively), with the overlapped structures from two endpoints of the line (blue mesh and red mesh) and the midpoint (yellow solid) (see also **Movies S4-S5**). (I) Clusters of images denoted by cylinders. (J) Three equal-size subsets of the yellow (central) cluster in (I). (K) Structures reconstructed from the clusters in (I) (the color of the structure is that of the cluster). (L) Structures reconstructed from the subsets in (J) (same coloring as in (K)). (M) Structures from mixed subsets (blue mesh, see the text for more details) overlapped with the structures obtained with ML3D method (orange solid). The red asterisk denotes an absence of the complete

1 mass in the linker region obtained by ML3D, which is an evidence of the conformational
2 heterogeneity of the image set used to compute this structure. Arrows show movements.
3
4
5

6 **Figure 4:** Experiment with EM images of TBSV. (A) Pseudo-atomic structure (spheres)
7 computed using a compact-form TBSV structure from a previous EM study (the EM structure
8 (Aramayo et al., 2005) is shown in transparent grey). (B) Collectivity degree of the pseudo-
9 atomic normal modes (the first six modes related to rigid-body movements are not shown;
10 the horizontal line corresponds to the collectivity of 0.75; the asterisks denote the modes with
11 the collectivity degree above 0.75). (C) Computed deformation amplitudes projected onto the
12 three most important principal axes (the points denote single-particle images and the most
13 important principal axis is 1), linear regression line (blue), and four groups of images (red,
14 magenta, green, and yellow) determined by splitting the full range of the data coordinates on
15 the first principal axis into four equal-length intervals. (D) Overlapped structures
16 reconstructed from the four groups of images shown in (C) (the volumes are colored with the
17 color of the corresponding group in (C)). See also **Movie S6** for the animated pseudo-atomic
18 structure along the regression line (C).
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36

37 REFERENCES

38 Aramayo, R., Merigoux, C., Larquet, E., Bron, P., Perez, J., Dumas, C., Vachette, P., and Boisset, N.
39 (2005). Divalent ion-dependent swelling of tomato bushy stunt virus: a multi-approach study.
40 *Biochim Biophys Acta* 1724, 345-354.
41
42 Asturias, F. J., Cheung, I. K., Sabouri, N., Chilkova, O., Wepplo, D., and Johansson, E. (2006). Structure
43 of *Saccharomyces cerevisiae* DNA polymerase epsilon by cryo-electron microscopy. *Nat Struct Mol*
44 *Biol* 13, 35-43.
45
46 Brink, J., Ludtke, S. J., Kong, Y., Wakil, S. J., Ma, J., and Chiu, W. (2004). Experimental verification of
47 conformational variation of human fatty acid synthase as predicted by normal mode analysis.
48 *Structure* 12, 185-191.
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1
2 Bruschweiler, R. (1995). Collective protein dynamics and nuclear spin relaxation. *J Chem Phys* *102*,
3 3396–3403.
4
5 Chacon, P., Tama, F., and Wriggers, W. (2003). Mega-Dalton biomolecular motion captured from
6
7 electron microscopy reconstructions. *J Mol Biol* *326*, 485-492.
8
9 Delarue, M., and Dumas, P. (2004). On the use of low-frequency normal modes to enforce collective
10
11 movements in refining macromolecular structural models. *Proc Natl Acad Sci U S A* *101*, 6957-6962.
12
13 Elad, N., Clare, D. K., Saibil, H. R., and Orlova, E. V. (2008). Detection and separation of heterogeneity
14
15 in molecular complexes by statistical analysis of their two-dimensional projections. *J Struct Biol* *162*,
16
17 108-120.
18
19
20
21 Fu, J., Munro, J. B., Blanchard, S. C., and Frank, J. (2011). Cryoelectron microscopy structures of the
22
23 ribosome complex in intermediate states during tRNA translocation. *Proc Natl Acad Sci U S A* *108*,
24
25 4817-4821.
26
27
28 Grob, P., Cruse, M. J., Inouye, C., Peris, M., Penczek, P. A., Tjian, R., and Nogales, E. (2006). Cryo-
29
30 Electron Microscopy Studies of Human TFIID: Conformational Breathing in the Integration of Gene
31
32 Regulatory Cues. *Structure* *14*, 511-520.
33
34
35 Jain, R., Hammel, M., Johnson, R. E., Prakash, L., Prakash, S., and Aggarwal, A. K. (2009). Structural
36
37 insights into yeast DNA polymerase delta by small angle X-ray scattering. *J Mol Biol* *394*, 377-382.
38
39
40 Jonic, S., Sorzano, C. O. S., Thevenaz, P., El-Bez, C., De Carlo, S., and Unser, M. (2005). Spline-based
41
42 image-to-volume registration for three-dimensional electron microscopy. *Ultramicroscopy* *103*, 303-
43
44 317.
45
46
47 Klinge, S., Nunez-Ramirez, R., Llorca, O., and Pellegrini, L. (2009). 3D architecture of DNA Pol alpha
48
49 reveals the functional core of multi-subunit replicative polymerases. *Embo J* *28*, 1978-1987.
50
51
52 Kruse, J., Kruse, K. M., Witz, J., Chauvin, C., Jacrot, B., and Tardieu, A. (1982). Divalent ion-dependent
53
54 reversible swelling of tomato bushy stunt virus and organization of the expanded virion. *J Mol Biol*
55
56 *162*, 393-414.
57
58
59
60
61
62
63
64
65

1 Ming, D., Kong, Y., Lambert, M. A., Huang, Z., and Ma, J. (2002). How to describe protein motion
2 without amino acid sequence and atomic coordinates. *Proceedings of the National Academy of*
3
4 *Sciences* **99**, 8620-8625.
5
6
7 Nogales-Cadenas, R., Jonic, S., Tama, F., Arteni, A. A., Tabas-Madrid, D., Vázquez, M., Pascual-
8
9 Montano, A., and Sorzano, C. O. S. (2013). 3DEM Loupe: analysis of macromolecular dynamics using
10
11 structures from electron microscopy. *Nucleic Acids Research* **41**, W363-W367.
12
13
14 Nunez-Ramirez, R., Klinge, S., Sauguet, L., Melero, R., Recuero-Checa, M. A., Kilkenny, M., Perera, R.
15
16 L., Garcia-Alvarez, B., Hall, R. J., Nogales, E., *et al.* (2011). Flexible tethering of primase and DNA Pol
17
18 alpha in the eukaryotic primosome. *Nucleic Acids Res* **39**, 8187-8199.
19
20
21 Penczek, P. A., Frank, J., and Spahn, C. M. (2006). A method of focused classification, based on the
22
23 bootstrap 3D variance analysis, and its application to EF-G-dependent translocation. *J Struct Biol* **154**,
24
25 184-194.
26
27
28 Peng, L. M., Ren, G., Dudarev, S. L., and Whelan, M. J. (1996). Robust Parameterization of Elastic and
29
30 Absorptive Electron Atomic Scattering Factors. *Acta Crystallographica Section A* **52**, 257-276.
31
32
33 Powell, M. J. D. (2002). UOBYQA: unconstrained optimization by quadratic approximation.
34
35 *Mathematical Programming* **92**, 555-582.
36
37
38 Robinson, I. K., and Harrison, S. C. (1982). Structure of the expanded state of tomato bushy stunt
39
40 virus. *Nature* **297**, 563-568.
41
42
43 Scheres, S. H., Gao, H., Valle, M., Herman, G. T., Eggermont, P. P., Frank, J., and Carazo, J. M. (2007).
44
45 Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization.
46
47 *Nat Methods* **4**, 27-29.
48
49
50 Simonetti, A., Marzi, S., Myasnikov, A. G., Fabbretti, A., Yusupov, M., Gualerzi, C. O., and Klaholz, B. P.
51
52 (2008). Structure of the 30S translation initiation complex. *Nature* **455**, 416-420.
53
54
55 Sorzano, C. O. S., Jonic, S., El-Bez, C., Carazo, J. M., De Carlo, S., Thevenaz, P., and Unser, M. (2004a).
56
57 A multiresolution approach to orientation assignment in 3D electron microscopy of single particles. *J*
58
59 *Struct Biol* **146**, 381-392.
60
61
62
63
64
65

1 Sorzano, C. O. S., Marabini, R., Velazquez-Muriel, J., Bilbao-Castro, J. R., Scheres, S. H., Carazo, J. M.,
2 and Pascual-Montano, A. (2004b). XMIPP: a new generation of an open-source image processing
3 package for electron microscopy. *J Struct Biol* *148*, 194-204.
4
5
6
7 Suhre, K., Navaza, J., and Sanejouand, Y. H. (2006). NORMA: a tool for flexible fitting of high-
8 resolution protein structures into low-resolution electron-microscopy-derived density maps. *Acta*
9 *Crystallogr D Biol Crystallogr* *62*, 1098-1100.
10
11
12
13
14 Suhre, K., and Sanejouand, Y. H. (2004). ElNemo: a normal mode web server for protein movement
15 analysis and the generation of templates for molecular replacement. *Nucleic Acids Res* *32*, W610-
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 Tirion, M. M. (1996). Large Amplitude Elastic Motions in Proteins from a Single-Parameter, Atomic
2 Analysis. *Phys Rev Lett* 77, 1905-1908.

3
4 Valle, M., Zavialov, A., Sengupta, J., Rawat, U., Ehrenberg, M., and Frank, J. (2003). Locking and
5
6 unlocking of ribosomal motions. *Cell* 114, 123-134.

7
8
9 Wang, Y., Rader, A. J., Bahar, I., and Jernigan, R. L. (2004). Global ribosome motions revealed with
10
11 elastic network model. *J Struct Biol* 147, 302-314.

12
13
14 Witz, J., and Brown, F. (2001). Structural dynamics, an intrinsic property of viral capsids. *Arch Virol*
15
16 146, 2263-2274.

17
18
19 Zhang, W., Dunkle, J. A., and Cate, J. H. (2009). Structures of the ribosome in intermediate states of
20
21 ratcheting. *Science* 325, 1014-1017.

Figure 1

[Click here to download high resolution image](#)

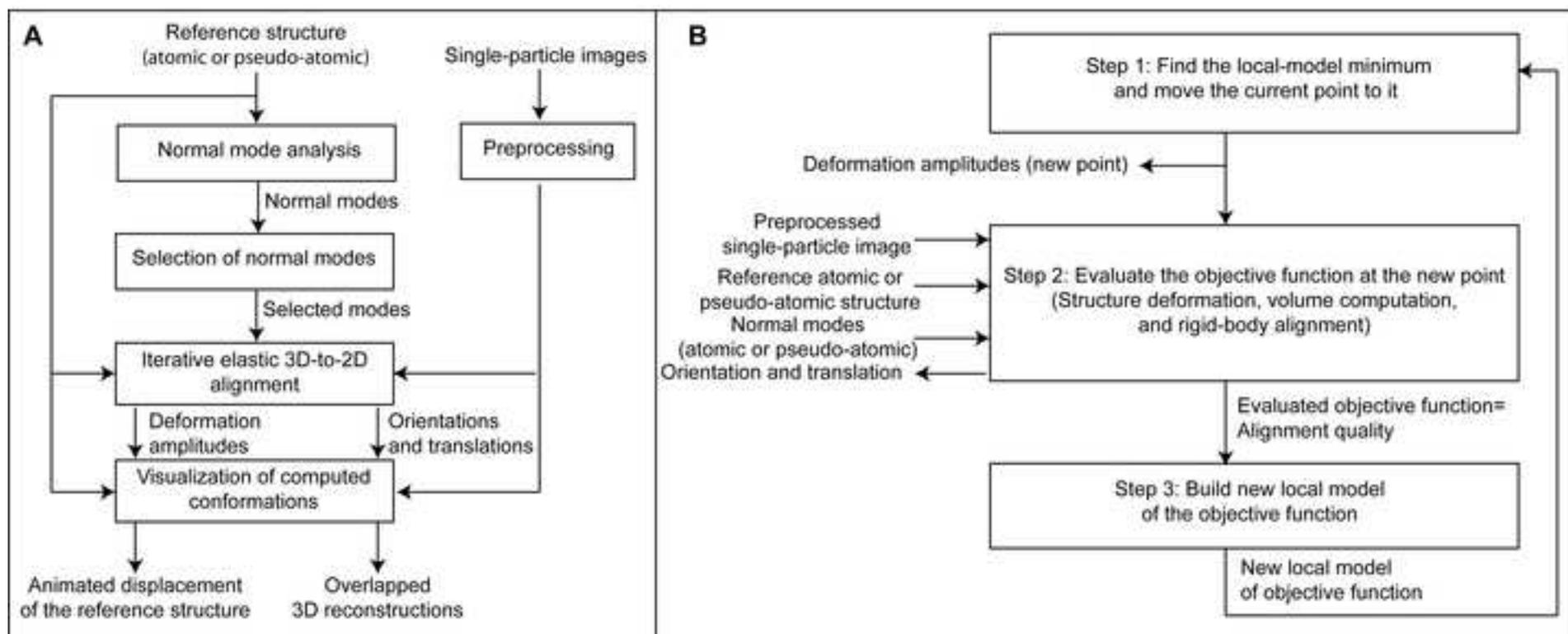


Figure 2
[Click here to download high resolution image](#)

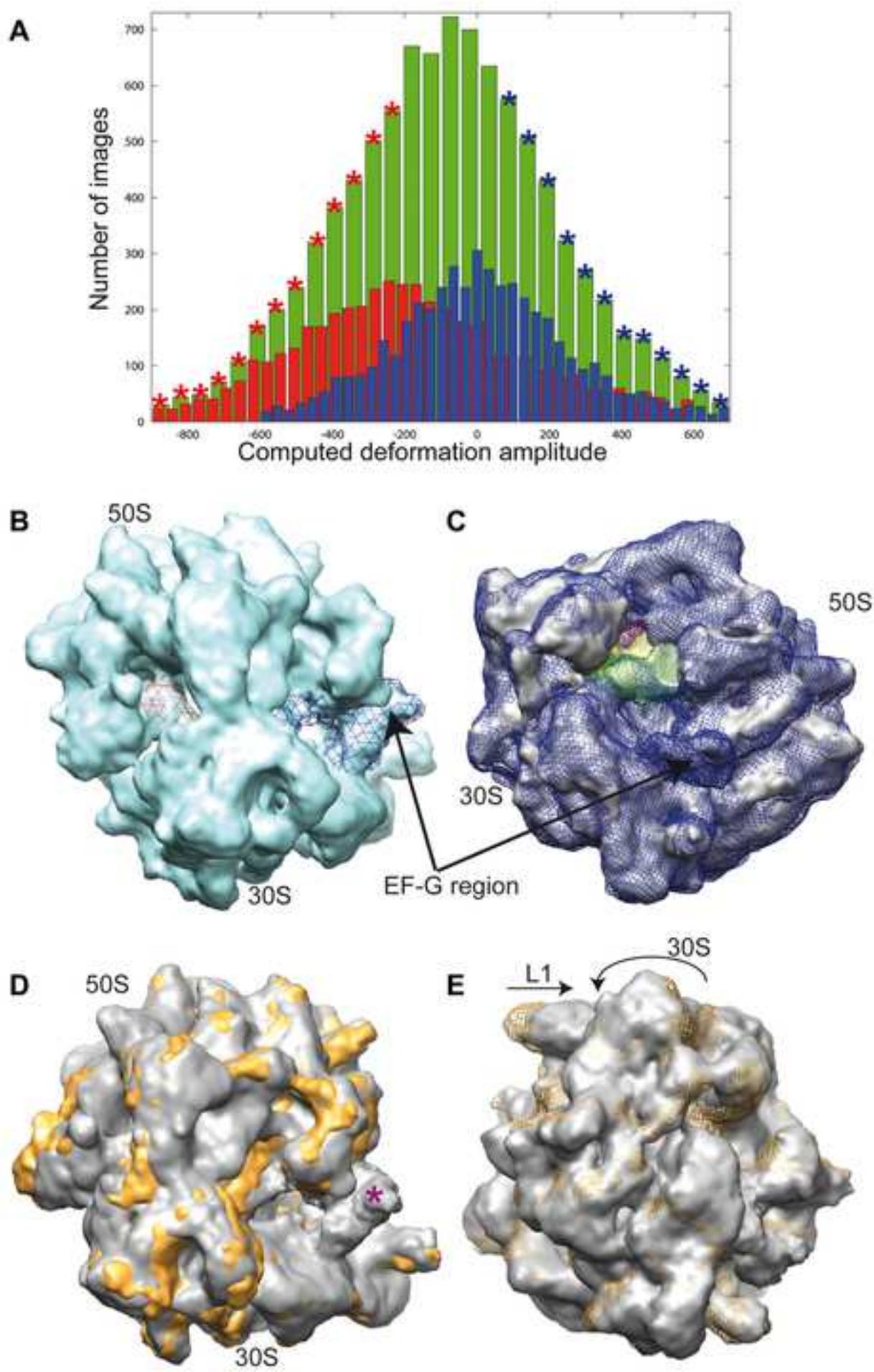


Figure 3
[Click here to download high resolution image](#)

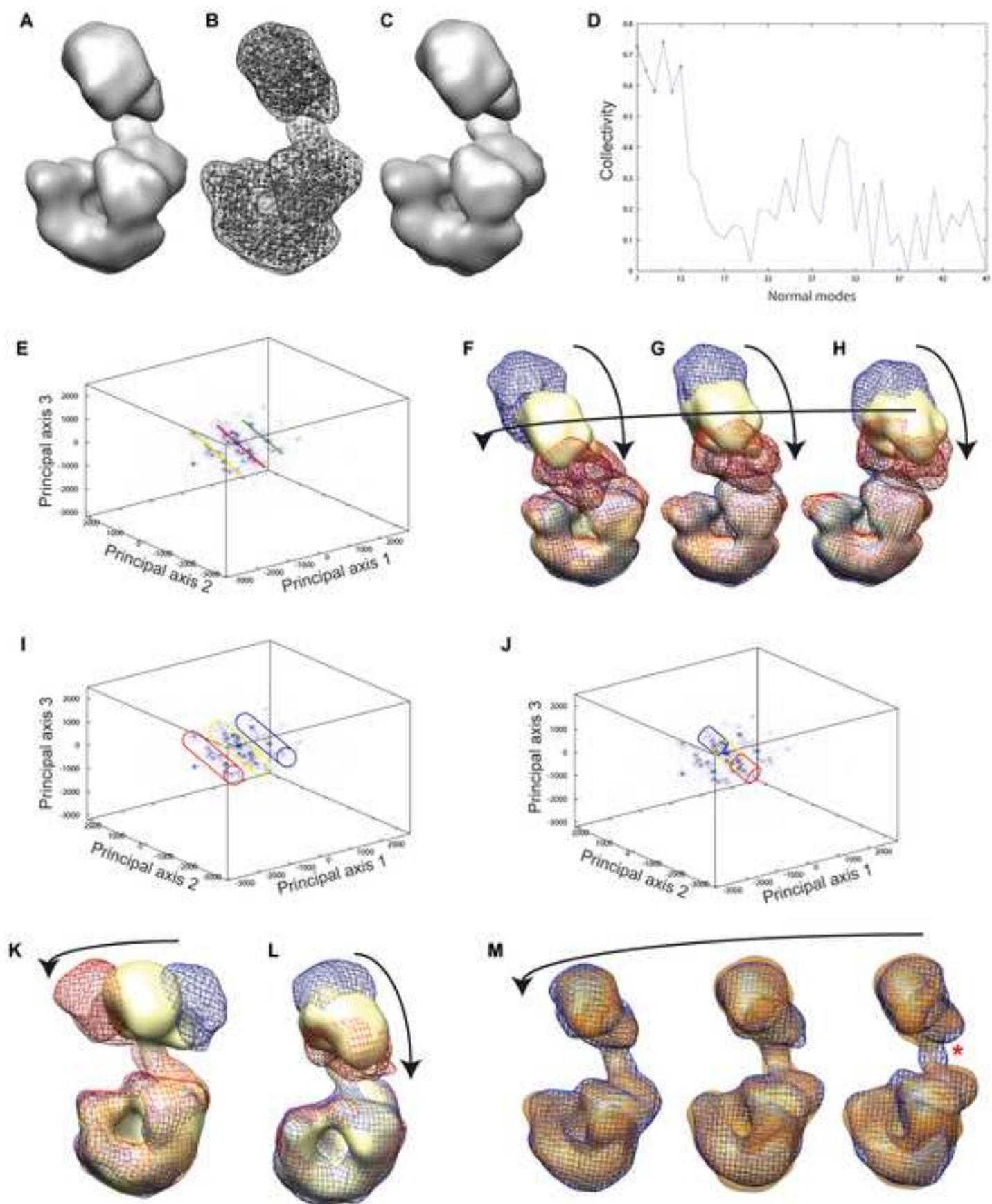
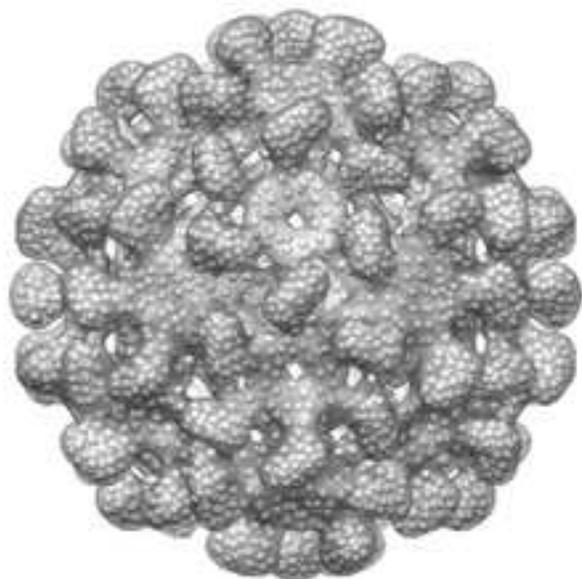
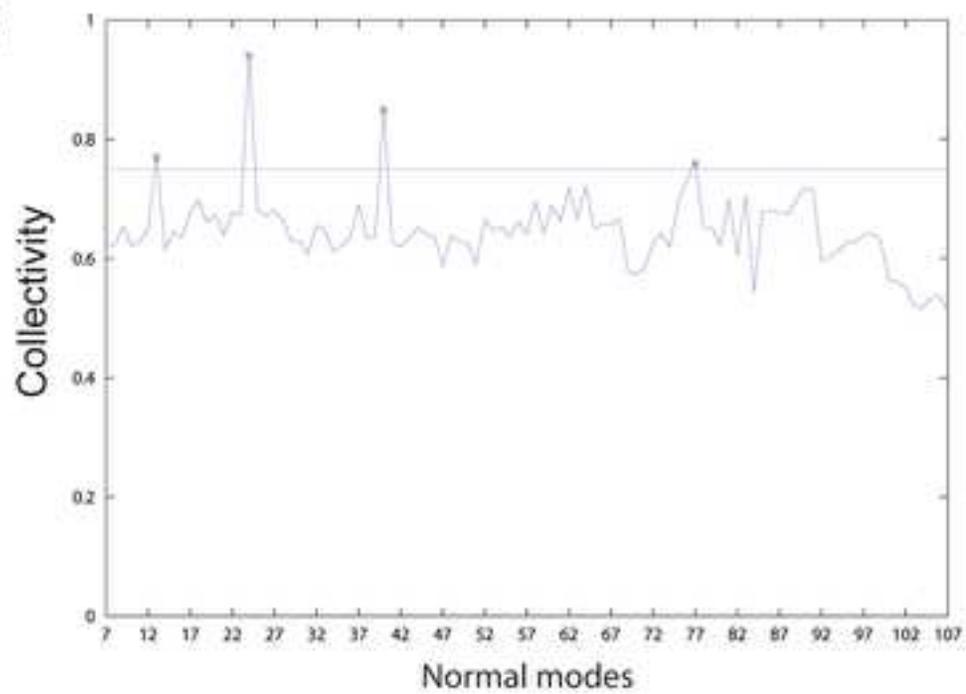


Figure 4
[Click here to download high resolution image](#)

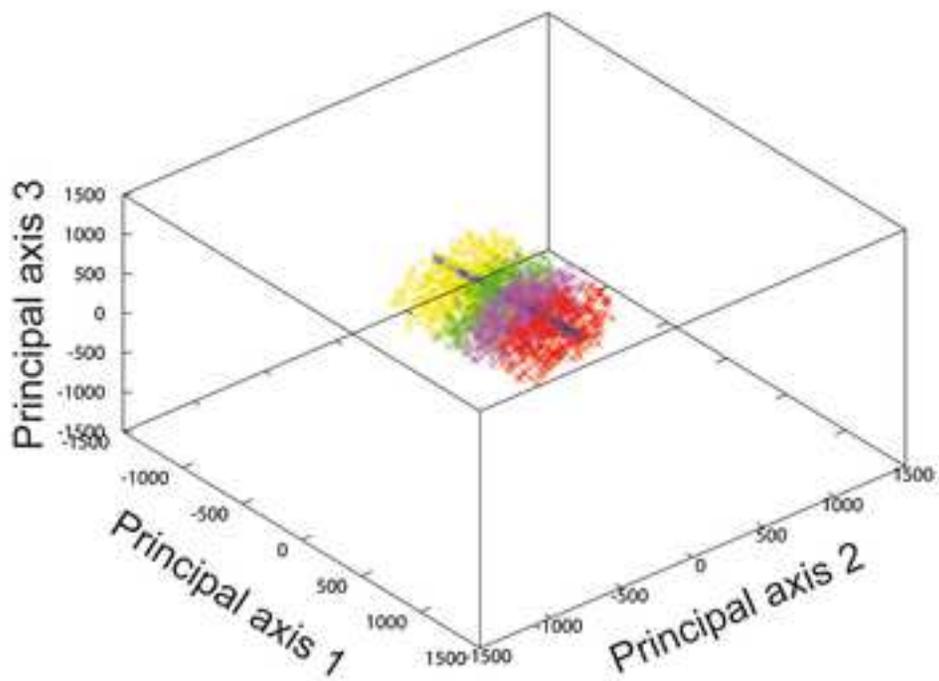
A



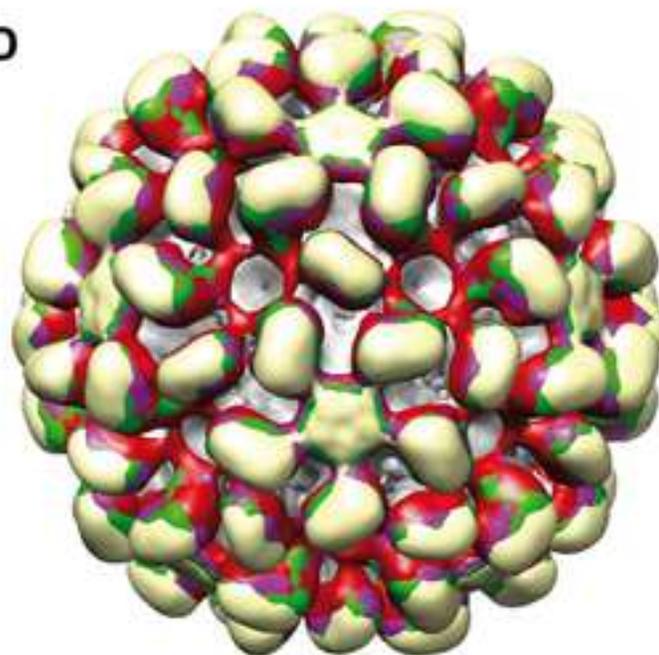
B



C



D



Inventory of Supplemental Information:

Figure S1, related to **Experiment 1** and **Supplemental experimental procedure S2**

Figure S2, related to **Figure 2 (Experiment 2)** and **Supplemental experimental procedure S3**

Figure S3, related to **Figure 3 (Experiment 3)**

Table S1, related to **Experiment 1** and **Supplemental experimental procedure S2**

Movie S1, related to **Experiment 1** and **Supplemental experimental procedure S2**

Movie S2, related to **Experiment 1** and **Supplemental experimental procedure S2**

Movie S3, related to **Figure 2 (Experiment 2)** and **Figure S2 (Supplemental experimental procedure S3)**

Movie S4, related to **Figure 3 (Experiment 3)**

Movie S5, related to **Figure 3 (Experiment 3)**

Movie S6, related to **Figure 4 (Experiment 4)**

Supplemental experimental procedure S1, related to **Methodology**

Supplemental experimental procedure S2, related to **Experiment 1**

Supplemental experimental procedure S3, related to **Experiment 2**

*Please note that the legends of the supplemental figures, tables, and movies are in the file **Jonic_SupplementalTextFiguresAndTables.pdf**, together with the supplemental experimental procedures, figures, tables, and references.*

Supplemental Data

SUPPLEMENTARY FIGURE LEGENDS

Figure S1: Structure used to compute images with synthetic continuous conformational change of the *E. coli* 70S ribosome, the images, and structure used to analyze them. (A) Atomic-resolution structure of the 70S ribosome in the conformation in which the subunit 30S is not rotated with respect to the subunit 50S (unratcheted conformation). (B) Synthetic volume at the resolution of 15 Å from the structure shown in (A). (C) Pseudo-atomic structure (spheres) from the volume in (B) overlapped with the volume (mesh). (D) Synthetic images without noise and CTF effects (row 1), images with SNR=0.3 and the defocus of 0.5 μm, 1 μm, and 2 μm (rows 2-4, respectively), and images with SNR=0.1 and the defocus of 0.5 μm, 1 μm, and 2 μm (rows 5-7, respectively). See the main text (**Experiment 1**) and **Supplemental experimental procedure S2** for more details.

Figure S2: Experiment with synthetic discrete conformational change given by EM structures of three different *E. coli* 70S conformations: (A-C) EM structures of unratcheted (EMD-5262 (A)), intermediate (EMD-5267 (B)), and ratcheted (EMD-5265 (C)) conformations used for computing synthetic images. (D) Pseudo-atomic structure (spheres) overlapped with the structure shown in (A) (mesh). (E) Histogram of the estimated deformation amplitudes (green) for the most difficult of the seven tested cases (SNR=0.1, defocus=2μm), where blue, magenta, and red indicate the contribution of images of the structures shown in (A), (B), and (C), respectively, and the vertical tan and cyan lines indicate the separation of images with respect to the computed deformation amplitudes into three classes of equal size (300 images per class). (F) Overlap between the structure reconstructed from the class with the smallest deformation amplitudes (red mesh for the reconstruction from the images at the left side of the tan class-separation line in (E)), the structure reconstructed from the class with the intermediate values of the deformation amplitudes (grey solid for the reconstruction from the images at the right side of the tan class-separation line and at the left side of the cyan class-separation line in (E)), and the structure reconstructed from the class with the largest deformation amplitudes (blue mesh for the reconstruction from the images at the right side of the cyan class-separation line in (E)). (G) Overlap between the structure reconstructed from the class with the smallest deformation amplitudes (red solid in (G) is red mesh in (F)) and the structure reconstructed from the class with the largest deformation amplitudes (blue solid in (G) is blue mesh in (F)). Rotation of the 30S subunit and the movement of the L1 stalk are denoted with arrows (although the L1 movement is rotational, it is shown as a shift to designate an inward direction of the movement). See the main text (**Experiment 2, Figure 2**) and **Supplemental experimental procedure S3** for more details.

Figure S3: Displacement of a volumetric form of the Pol α – B pseudo-atomic structure along normal modes and along principal component axes. (A-F) Displacement along normal modes, from mode 7 (A) to mode 12 (F). (G-L) Displacement along principal component axes, from the most important principal axis (G) to the least important principal axis (L). The displacement amplitude was chosen large to show clearly the movement directions. Displacement amplitude +/-1000 (blue mesh and red mesh) and zero displacement amplitude (reference conformation, yellow). See the main text for more details (**Experiment 3**) as well as **Figure 3**.

SUPPLEMENTARY TABLE LEGENDS

Table S1: Results of the experiment in which the ground-truth synthetic continuous-type conformational change of *E. coli* 70S ribosome was determined by two atomic normal modes (modes 11 and 12) but it was estimated by analyzing images with only one atomic (mode 11 of the atomic reference structure (**Movie S1**)) or pseudo-atomic normal mode (mode 13 of

1 the pseudo-atomic synthetic reference structure (**Movie S2**). (A,B) Error of the estimated
2 elastic (A) and rigid-body (B) parameters with respect to the ground-truth values for no noise,
3 SNR=0.3, SNR=0.1, and the defocus of 0.5 μm , 1 μm , and 2 μm . (C) Dependence of
4 parameters of the volume-to-pseudoatoms conversion and NMA methods and that of the
5 image alignment error on the volume approximation error and the size of pseudo-atoms (the
6 standard deviation of the Gaussian functions), for the images without noise and CTF effects.
7 See the main text (**Experiment 1**) and **Supplemental experimental procedure S2** for more
8 details.

9 **SUPPLEMENTARY MOVIE LEGENDS**

10
11 **Movie S1:** Displacement along the atomic mode 11 related to the ratchet-like motion of the
12 70S ribosome (the atomic structure is 3I1OP (**Fig. S1A**)). The mode was used to analyze
13 images in **Experiment 1**. See also **Supplemental experimental procedure S2**. **Movies S1-**
14 **S3** were produced using large amplitudes of the displacement along the mode to better
15 visualize the movements.
16

17
18 **Movie S2:** Displacement along the pseudo-atomic mode 13 related to the ratchet-like motion
19 of the 70S ribosome, from a synthetic density volume of the atomic (3I1OP) structure (the
20 pseudo-atomic structure (**Fig. S1C**) was obtained by converting the 15 Å-resolution 3I1OP
21 density volume (**Fig. S1B**)). The mode was used to analyze images in **Experiment 1**. See
22 also **Supplemental experimental procedure S2**.
23

24
25 **Movie S3:** Displacement along the pseudo-atomic mode 9 related to the ratchet-like motion
26 of the 70S ribosome, from the EM structure EMD_5262 (the EM and pseudo-atomic
27 structures are shown in **Fig. S2A,D**). The mode was used to analyze images in **Experiment**
28 **2**. See also **Supplemental experimental procedure S3**.
29

30
31 **Movie S4:** Animated displacements of the reference pseudo-atomic Pol α – B structure along
32 the red line in **Figure 3E**. The amplitude of the displacement is a result of the image analysis
33 in **Experiment 3**. The same view of the structure is showed during the displacement along
34 the line.
35

36
37 **Movie S5:** Animated displacements of a volumetric form of the reference pseudo-atomic Pol
38 α – B structure along the red line in **Figure 3E**. The amplitude of the displacement is a result
39 of the image analysis in **Experiment 3**. The volume is rotated during the displacement along
40 the line so that it can be viewed from different directions.
41

42
43 **Movie S6:** Animated displacements of the reference pseudo-atomic TBSV structure along
44 the blue line in **Figure 4C**. The amplitude of the displacement is a result of the image
45 analysis in **Experiment 4**. The same view of the structure is showed during the displacement
46 along the line.
47

48 **Supplemental experimental procedures**

49 **Supplemental experimental procedure S1 (related to Methodology)**

50 **Visualization of computed conformations**

51 *Three-dimensional reconstruction*

52 The 3D-reconstructed structures shown throughout the paper were obtained using the
53 computed deformation amplitudes (for the classification), the computed orientations and
54 translations of particles (for the reconstruction), and a standard 3D reconstruction method
55 (here, ART implemented in Xmipp (Sorzano et al., 2004b)), without any further refinement of
56
57
58
59
60
61
62
63
64
65

1 the alignment parameters, noise reduction, or resolution improvement (e.g., by adding more
2 images).

3 *Visualization of pseudo-atomic models*

4 Figures and movies of pseudo-atomic models were obtained by approximating Gaussian
5 functions with spheres. Overlapping spheres were used in the movies (obtained with VMD
6 (Humphrey et al., 1996)) to enhance the structure visually (to give impression of density
7 without converting the pseudo-atomic structure into a density volume) and to get closer to the
8 reality as the Gaussian functions overlap (they have unlimited support).
9

10 **Supplemental experimental procedure S2 (related to Experiment 1)**

11 *Atomic structure and atomic modes*

12
13 The ground-truth synthetic conformational changes were obtained by deforming a reference
14 atomic-resolution structure (**Fig. S1A**) that was a coarse-grained model (10204 atoms
15 comprising only C α and phosphate (P) atoms) of a 0.35 nm-resolution X-ray crystal structure
16 of the *E. coli* 70S ribosome without ligands, in the conformation in which the small subunit of
17 the ribosome (30S) is non-rotated with respect to the large subunit (50S) (a composite of the
18 PDB entries 3I1O and 3I1P that will be referred to as 3I1OP (Zhang et al., 2009)). Twenty
19 atomic lowest-frequency normal modes were computed for the atomic reference structure
20 using Elnemo (Suhre and Sanejouand, 2004) web server.
21
22

23 *Synthetic images*

24 The synthetic images were computed as random projections of the atomic reference
25 structure displaced with a combination of random displacement amplitudes along two atomic
26 normal modes describing ratchet-like motion (mode 11) and L1-stalk motion (mode 12). The
27 images were computed for the deformation amplitude along mode 11 that was random and
28 had a uniform distribution (equal probability of different values) in the range [-500,0]. The
29 amplitude along mode 12 was chosen to be 2.5 times smaller than the amplitude along mode
30 11, which means that mode 11 contributes 2.5 times more than mode 12 to the final motion.
31 In this synthetic data experiment, the ratio of 2.5 was chosen to generate the total
32 deformation (weighted linear combination of two modes with their deformation amplitudes as
33 the weights) describing the structural rearrangements similar to those already observed for
34 the 70S ribosome (30S rotates with respect to 50S by $\sim 8^\circ$ and the distal part of the L1 stalk
35 moves by $\sim 20 \text{ \AA}$ (Fu et al., 2011; Valle et al., 2003)).
36
37
38
39

40 Density volumes were computed from atomic coordinates using electronic-form atomic
41 factors (Peng et al., 1996) and they were then projected onto the image plane using random
42 values of Euler angles determining the volume orientation and using random values of the
43 particle coordinates in the image (two out-of-plane rotations were distributed uniformly in the
44 ranges $[0^\circ, 360^\circ]$ and $[0^\circ, 180^\circ]$, the in-plane rotation was 0° , and each of the two in-plane
45 translations was distributed uniformly in the range $[-5, 5]$ pixels). The CTF and noise were
46 simulated for a 200 kV microscope with a spherical aberration of 0.5 mm using the method of
47 Velazquez-Muriel et al. (2003). The computed data thus comprised a set of 300 images
48 without noise and CTF effects (image size: 128^2 pixels; pixel size: $(3 \text{ \AA})^2$) and six versions of
49 these images affected by different levels of noise (SNR=0.3 or SNR=0.1) and different CTF
50 effects (defocus values of 0.5 μm , 1 μm , or 2 μm) (**Fig. S1D**). The chosen defocus and SNR
51 values are commonly encountered in EM practice.
52
53

54 *Pseudo-atomic structure and pseudo-atomic modes*

55 The reference atomic-resolution structure (**Fig. S1A**) was converted into volumetric data
56 (size: 128^3 voxels, voxel size: $(3 \text{ \AA})^3$) that was low-pass filtered at 15 \AA (**Fig. S1B**). This
57 volume was converted into a pseudo-atomic structure with 6083 pseudo-atoms for the
58 standard deviation of Gaussian functions $\sigma = 3$ and the volume approximation error of 2 %
59 (**Fig. S1C**) and twenty lowest-frequency normal modes were computed for the pseudo-
60
61
62
63
64
65

1 atomic structure with the pseudo-atom interaction cutoff distance of 22 Å, using the methods
2 of our 3DEM Loupe web server (Nogales-Cadenas et al., 2013) (the values of the volume-to-
3 pseudoatoms conversion and NMA parameters were chosen as described in (Nogales-
4 Cadenas et al., 2013)).

5 *Image analysis*

6 The best overlap between the atomic mode responsible for the ratchet-like motion (mode 11)
7 and the set of pseudo-atomic modes was found for the pseudo-atomic mode 13 (the movies
8 animating displacements along the atomic mode 11 and the pseudo-atomic mode 13 are
9 given in **Movies S1-S2**). The elastic 3D-to-2D alignment of the synthetic images with the
10 atomic reference structure was thus done using the atomic mode 11 and the one with the
11 pseudo-atomic reference structure was done using the pseudo-atomic mode 13.
12
13

14 *Influence of the quality of pseudo-atomic representation on image alignment error*

15 **Table S1C** summarizes how the key parameters of the used volume-to-pseudoatoms
16 conversion and NMA methods (see also **Experimental procedures**) and the resulting
17 alignment error depend on the volume approximation error and the size of pseudo-atoms
18 (the standard deviation of the Gaussian function). The results in **Table S1C** were obtained
19 using images without noise and CTF effects and can be compared with the results in **Table**
20 **S1A,B** for no-noise/no-CTF case (6083 pseudo-atoms, $\sigma = 3$, volume approximation error of
21 2 %, pseudo-atom interaction cutoff distance of 22 Å). We can notice that the number of
22 pseudo-atoms decreases as the volume approximation error or the Gaussian standard
23 deviation increases (e.g., from 6083 pseudo-atoms for the target approximation error of 2 %
24 to 450 pseudo-atoms for the target approximation error of 7 %, when using $\sigma = 3$ in both
25 cases). Also, the pseudo-atom interaction cutoff distance increases as the number of
26 pseudo-atoms decreases (from 22 Å for 6083 pseudo-atoms to 40 Å for 450 pseudo-atoms)
27 and the alignment error increases as the volume approximation error increases (e.g., the
28 relative deformation-amplitude error of 4.3% (**Table S1A**) increases to 27.3% (**Table S1C**)
29 when the volume approximation error increases from 2 % to 7 % for the same Gaussian
30 standard deviation of 3). Similar results may be expected with noise- and CTF-affected
31 images.
32
33
34
35

36 **Supplemental experimental procedure S3 (related to Experiment 2)**

37 *EM structures*

38 We analyzed synthetic images of EM structures representing the following three
39 conformations of the *E. coli* 70S ribosome with two tRNAs, published by Fu et al. (2011): 1)
40 unratcheted conformation EMD-5262 with resolution of around 13 Å (**Fig. S2A**) (the subunits
41 30S and 50S are not rotated with respect to each other and the A-site and P-site tRNAs are
42 in the classical A/A and P/P positions, respectively, 2) ratcheted conformation EMD-5265
43 with resolution of around 14 Å (**Fig. S2C**) (the two subunits are fully rotated with respect to
44 each other and the A-site and P-site tRNAs are in the A/A and P/E positions, respectively),
45 and 3) structure with an intermediate relative rotation of the two subunits and an intermediate
46 position of the P-site tRNA called position P-flip (EMD-5267, resolution of around 18 Å, **Fig.**
47 **S2B**).
48
49
50

51 *Synthetic images*

52 The synthetic images were computed as random projections of the EM structures. The given
53 EM volumes (EMD-5262, EMD-5265, and EMD-5267) were padded from the size 125^3
54 voxels (voxel size $(3 \text{ \AA})^3$) to the size of 128^3 voxels (the image size of 128^2 pixels and the
55 volume size of 128^3 voxels were used throughout this paper regarding the experiments with
56 the ribosome). We computed 300 images (128^2 -pixels images with the pixel size of $(3 \text{ \AA})^2$) for
57 each of the three EM structures, using the parameters of its orientation and position (rigid-
58 body parameters) whose values were generated randomly, as in **Experiment 1** (two out-of-
59 plane rotations were distributed uniformly in the ranges $[0^\circ, 360^\circ]$ and $[0^\circ, 180^\circ]$, the in-plane
60
61
62
63
64
65

1 rotation was 0°, and each of the two in-plane translations was distributed uniformly in the
2 range [-5,5] pixels). The noise (SNR of 0.3 or 0.1) and the CTF (defocus of 0.5 μm, 1 μm, or
3 2 μm) were applied onto the ideal projections, as in **Experiment 1**. The total number of
4 images was 900 in each of 7 analyzed datasets (one set without and six sets with noise and
5 CTF effects).

6 *Pseudo-atomic structure and pseudo-atomic modes*

7 The volume with the unratcheted conformation (EMD-5262) was converted into a pseudo-
8 atomic structure with 9411 pseudo-atoms for $\sigma = 3$ and the volume approximation error of
9 3.4% (**Fig. S2D**) and twenty lowest-frequency normal modes were computed for the pseudo-
10 atomic structure with the pseudo-atom interaction cutoff distance of 22 Å using the methods
11 of 3DEM Loupe web server (Nogales-Cadenas et al., 2013). The volume approximation error
12 was larger when approximating this volume than the synthetic one (**Supplemental**
13 **experimental procedure S2**), for the same voxel size and the same value of σ , because of a
14 higher level of noise of the EM volume with respect to the synthetic one.
15
16

17 *Image analysis*

18 The images were analyzed using the reference pseudo-atomic structure and its mode
19 reflecting the ratchet-like motion of the 30S and 50S subunits. The pseudo-atomic mode 9
20 (**Movie S3**) was thus used here as it overlapped the best with the atomic mode responsible
21 for the ratchet-like motion (**Experiment 1, Movie S1**). Note that the atomic structure 3I1OP
22 was aligned with the EMD-5262 structure before any further processing. As in **Experiment 1**,
23 the results showed that the accuracy of the estimated rigid-body parameters decreases as
24 the SNR decreases and the defocus increases and, in the most difficult tested case
25 (SNR=0.1, defocus of 2 μm), the angular and shift errors were 1.3° and 0.2 pixel,
26 respectively. Accuracy of the estimated elastic parameters could not be checked here as we
27 had no knowledge about the ground-truth real conformational change but the analysis of the
28 estimated deformation amplitudes showed distinct mean values for the three conformations
29 (the average deformation amplitude over different SNR and defocus values is around -1 for
30 the unratcheted conformation, -135 for the intermediate conformation, and -280 for the
31 ratcheted conformation). The images were thus separated into three equal-size classes of
32 estimated deformation-amplitude values (300 highest, 300 lowest, and 300 intermediate
33 values).
34
35
36

37
38 Even for this small number of images per class (300), 3D reconstruction from the classes
39 produced structures that were similar to the given EM structures (e.g., **Figure S2E-G** shows
40 the results in the most difficult tested case). **Figure S2F** clearly shows a rotation of the 30S
41 subunit with respect to the 50S subunit (ratchet-like movement). **Figure S2F** also shows a
42 movement of the L1 stalk with respect to the 30S subunit, which is an important result
43 because the L1-stalk movement was not explicitly used to analyze images. Distinct degrees
44 of rotation of the 30S subunit with respect to the 50S subunit and distinct positions of the L1
45 stalk are even better visualized in **Figure S2G** that shows an overlap between the structure
46 reconstructed from the class with the smallest deformation amplitudes (images at the left
47 side of the tan class-separation line in **Figure S2E**) and the structure reconstructed from the
48 class with the largest deformation amplitudes (images at the right side of the cyan class-
49 separation line in **Figure S2E**). Thus, this experiment with synthetic data showed that the
50 ratchet-like mode alone can be used to classify images of the ribosome at distinct states of
51 the conformational change combining ratchet-like and L1-stalk motions, and to reconstruct
52 the structure at distinct states of the conformational change
53
54
55

56 **Supplemental References**

57
58 Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. *J Mol Graph* 14,
59 33-38, 27-38.
60
61
62
63
64
65

Velazquez-Muriel, J. A., Sorzano, C. O., Fernandez, J. J., and Carazo, J. M. (2003). A method for estimating the CTF in electron microscopy based on ARMA models and parameter adjustment. *Ultramicroscopy* 96, 17-35.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

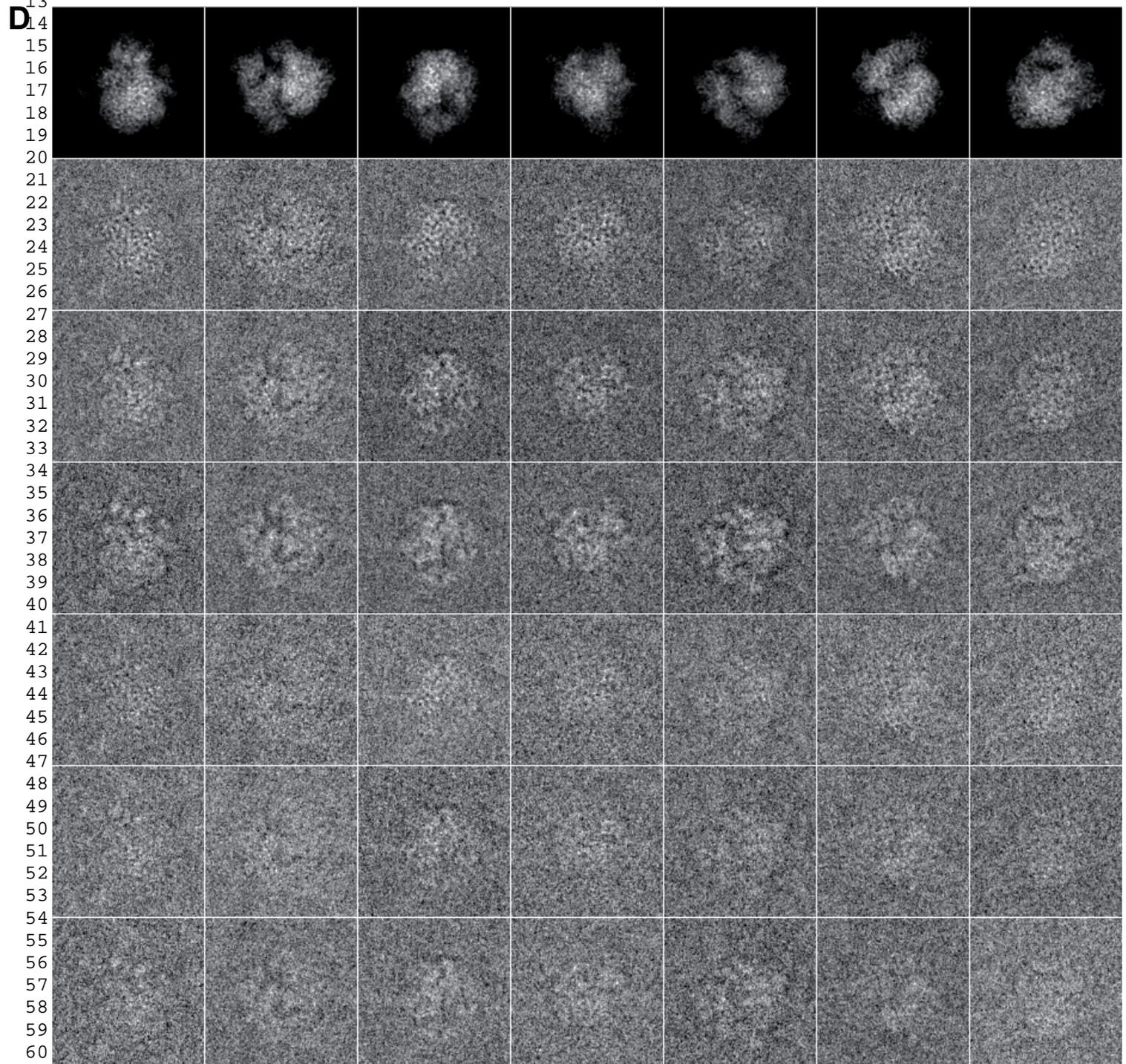
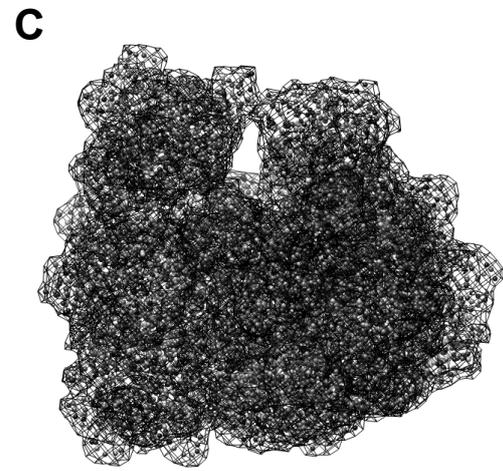
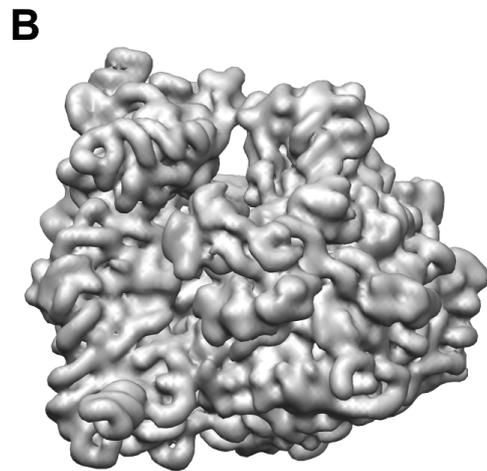
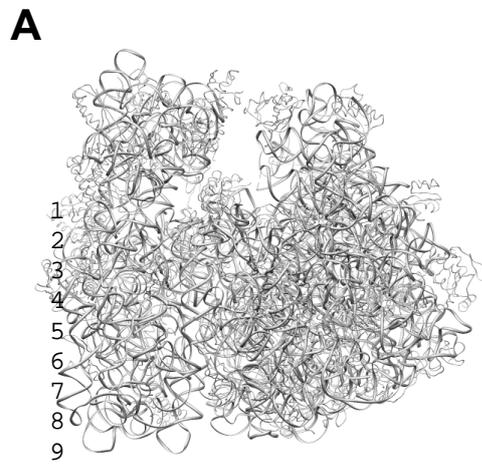


Figure S1

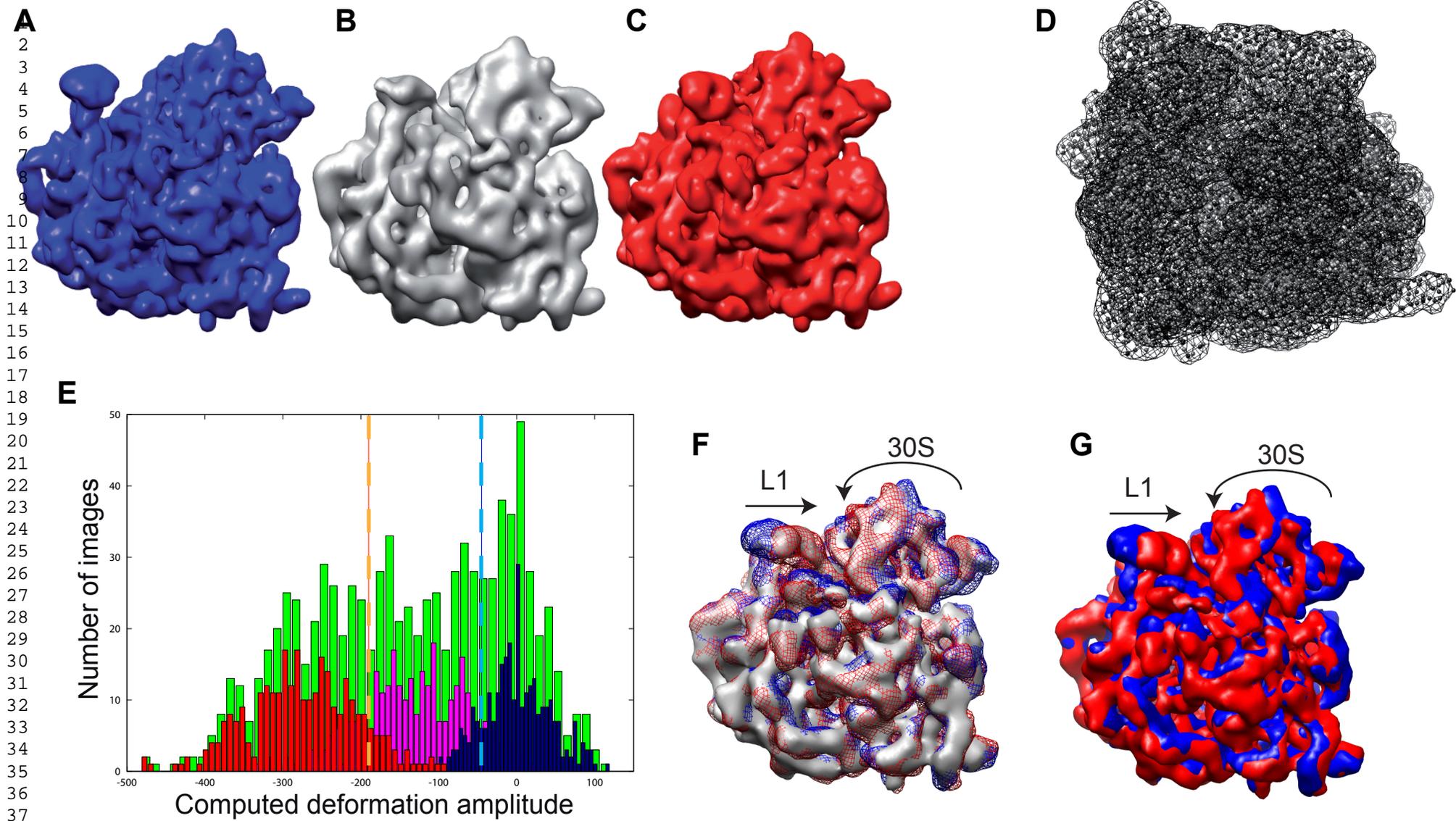


Figure S2

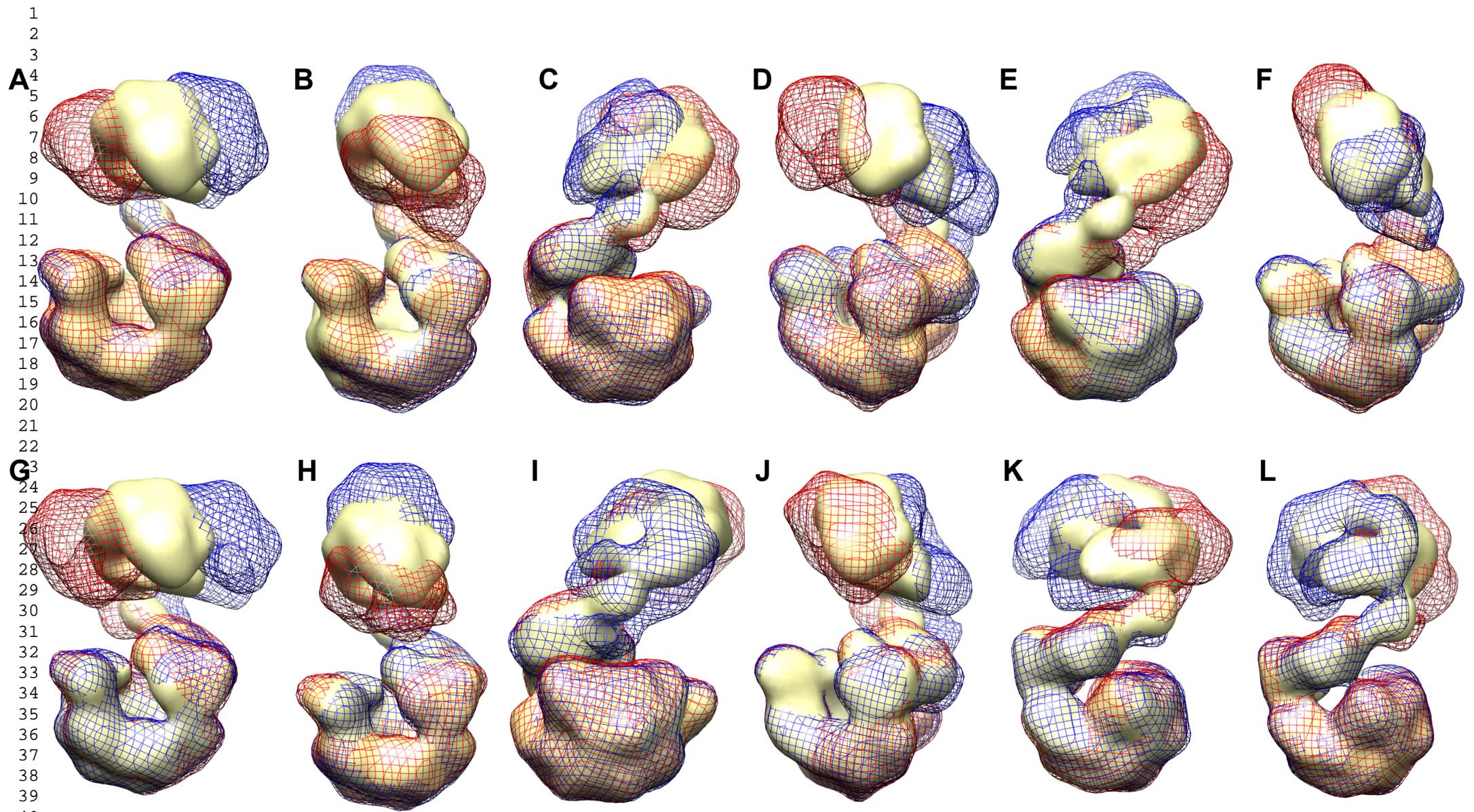


Figure S3

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49

A		Relative amplitude error along mode for no noise [%]		Relative amplitude error along mode for SNR=0.3 [%]		Relative amplitude error along mode for SNR=0.1 [%]								
		Atoms	Pseudo-atoms	Atoms	Pseudo-atoms	Atoms	Pseudo-atoms							
		Defocus 0.5 μm	NA	NA	3.4	6.7	4.1	8.8						
		Defocus 1 μm	NA	NA	3.4	7.3	3.9	8.2						
		Defocus 2 μm	NA	NA	3.7	8.3	4.2	8.9						
	No CTF	1.6	4.3	NA	NA	NA	NA							
B		Global absolute angular and shift error for no noise				Global absolute angular and shift error for SNR=0.3				Global absolute angular and shift error for SNR=0.1				
		Orientation [deg]		Position [pix]		Orientation [deg]		Position [pix]		Orientation [deg]		Position [pix]		
		Atoms	Pseudo-atoms	Atoms	Pseudo-atoms	Atoms	Pseudo-atoms	Atoms	Pseudo-atoms	Atoms	Pseudo-atoms	Atoms	Pseudo-atoms	
		Defocus 0.5 μm	NA	NA	NA	NA	0.30	0.37	0.07	0.08	0.45	0.54	0.11	0.12
		Defocus 1 μm	NA	NA	NA	NA	0.32	0.43	0.07	0.09	0.37	0.51	0.08	0.11
		Defocus 2 μm	NA	NA	NA	NA	0.38	0.48	0.09	0.10	0.50	0.59	0.12	0.12
		No CTF	0.15	0.15	0.03	0.04	NA	NA	NA	NA	NA	NA	NA	NA
	C		Volume-to-pseudoatoms conversion			Normal mode analysis		Image analysis (no noise case)						
			Volume approximation error	Gaussian-function standard deviation	Number of pseudo-atoms	Pseudo-atoms interaction cutoff distance	30S-rotation mode	Relative deformation-amplitude error along mode [%]	Global absolute angular error [deg]	Global absolute shift error [pix]				
			0.05	2	5155	22	12	7.0	0.27	0.07				
			0.05	3	1510	22	13	6.9	0.38	0.09				
			0.07	2	1246	35	10	24.9	8.69	0.28				
		0.07	3	450	40	17	27.3	13.65	0.46					

Table S1

Movie S1

[Click here to download Supplemental Movies and Spreadsheets: Jonic_MovieS1.mov](#)

Movie S2

[Click here to download Supplemental Movies and Spreadsheets: Jonic_MovieS2.mov](#)

Movie S3

[Click here to download Supplemental Movies and Spreadsheets: Jonic_MovieS3.mov](#)

Movie S4

[Click here to download Supplemental Movies and Spreadsheets: Jonic_MovieS4.mov](#)

Movie S5

[Click here to download Supplemental Movies and Spreadsheets: Jonic_MovieS5.mov](#)

Movie S6

[Click here to download Supplemental Movies and Spreadsheets: Jonic_MovieS6.mov](#)