Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Probing Cooperativity of N-Terminal Domain Orientations in the p97 Molecular Machine: Synergy Between NMR Spectroscopy and Cryo-EM

Rui Huang[+],* *Zev A. Ripstein*[+],* *John L. Rubinstein,** *and Lewis E. Kay**

anie_202009767_sm_miscellaneous_information.pdf

Author Contributions

R.H. Conceptualization: Lead; Data curation: Lead; Formal analysis: Lead; Investigation: Lead; Methodology: Lead; Validation: Lead; Writing—Original Draft: Lead; Writing—Review & Editing: Lead
Z.R. Conceptualization: Lead; Data curation: Lead; Formal analysis: Lead; Investigation: Lead; Methodology: Lead; Validation: Lead; Visualization: Lead; Writing—Original Draft: Lead; Writing—Review & Editing: Lead
J.R. Funding acquisition: Lead; Project administration: Lead; Resources: Lead; Supervision: Lead; Writing—Review & Editing: Lead
L.K. Conceptualization: Lead; Funding acquisition: Lead; Investigation: Lead; Project administration: Lead; Resources: Lead; Supervision: Lead; Validation: Lead; Writing—Original Draft: Lead; Writing—Review & Editing: Lead.

Supporting Information

69451 Weinheim, Germany

DOI: 10.1002/anie.202009767

Table of Contents

Materials and Methods

Cloning, protein expression and purification

The R95G mutant of full-length p97 from *Mus musculus* was cloned, expressed and purified as previously described.[1]

Sample vitrification

Purified R95G p97 was treated with apyrase (New England Biolabs; 5 units per mM of protomers) for *ca.* 16 hr at 25 °C, followed by size exclusion chromatography on a HiLoad 16/60 Superdex 200 gel filtration column to remove apyrase. Fractions corresponding to the p97 peak were concentrated to ~15 mg/mL in a buffer comprising 50 mM Hepes (pH 7.5), 200 mM NaCl, 1 mM $MgCl_2$, and 5 mM TCEP. Following concentration, the complex was incubated with 10 mM ADP and immediately before freezing 0.025 % (wt/vol) IGEPAL CA-630 (Sigma-Aldrich) was added to increase the number of particles adopting side-views on the grid. A total of 2.5 μL of sample was applied to nanofabricated holey sputtered gold grids[2] with a hole size of ~1 μm and blotted using a modified FEI Vitribot for 4.5 s before plunge freezing in a liquid ethane/propane mixture (ratio of ~3:2)[3] held at liquid nitrogen temperature.

Electron Microscopy

Movies consisting of 30 frames over a 60 second exposure were collected with defocuses ranging from 1.0 to 3.9 μm using a FEI Titan Krios G3 operating at 300 kV and equipped with a FEI Falcon 3EC DDD camera. Movies were collected in counting mode at a nominal magnification of 75000 ×, corresponding to a calibrated pixel size of 1.06 Å and with an exposure rate of 0.8 electrons/pixel/s, giving a total exposure of ~43 electrons/$Å^2$. 2796 movies were collected semi-automatically using EPU software.

EM image analysis

Whole frame alignment was performed with an implementation of *alignframes_lmbfgs*[4] in the cryoSPARC v2 package[5] and the resulting averages of frames were used for contrast transfer function determination with CTFFIND4[6] and automated particle selection with RELION.[7] Particle coordinates were used to extract particles in 256 pixel boxes from the aligned movies, while performing individual particle alignment and exposure weighting with a reimplementation of *alignparts_lmbfgs*[4] in cryoSPARC v2. The R95G p97-ADP dataset of 757,274 particle images was initially cleaned with 2D classification and heterogeneous refinement, leading to a dataset of 411,221 particle images. These images were used to refine a C6 symmetry map to 3.6 Å resolution, with interpretable density for the D1 and D2 domains, but with poor density for the NTDs (Figure S1A). As the NTD portion of the map was at lower resolution, for the subsequent analysis particle images were Fourier-cropped to a box size 160×160, corresponding to a pixel size of 1.70 Å. In order to further resolve the NTD region, classification of the up/down state for each protomer was necessary, with the dataset first split into four independent groups to provide an error estimate for the classification procedure, described below.

The approach that we have taken makes use of an algorithm described in detail elsewhere.[8] The configuration of each protomer was determined using the "3D variability analysis job"[8] in cryoSPARC[5] to assign each protomer to either an 'up' or a 'down' conformation. A mask was used to limit the analysis to the NTDs of the protomers. The mask was created by superimposing the NTDs from previously solved up and down states (PDB IDs 5FTK[9] (down) and 5FTN[9] (up)), with the *molmap* function in UCSF Chimera[10] so as to create a single volume with density for both. This volume was then converted into a mask using cryoSPARC's volume tools (Figure S1B). Symmetry expansion[11] was used so that each

protomer in the pseudo C6 hexamer contributed to the analysis. The resulting reconstructed volumes for a protomer show either a clear "up" conformation for the NTD, or a clear "down" configuration (Figure S1D). The symmetry expansion approach allows classification of each NTD in a particle as belonging to either the up or down state, and thus each particle image was classified as being derived from one of the 14 possible non-degenerate states of the p97 complex (Figure 2A) using an in-house written python program. The images sorted in this manner were subsequently used to reconstruct 3D densities for the different states using the prior Euler angles describing the relative position of an image in the context of the 3D map, obtained from the previous refinement in cryoSPARC (Figure S1E). Notably, the resulting reconstructions faithfully represent the expected p97 distributions of Figure 2A, with all 14 states resolved by this analysis.

To further establish the robustness of the classification approach we generated a simulated dataset starting from the previously solved all NTD up and down structures (PDB IDs 5FTN and 5FTK). These two structures were converted to density models and subsequently projected into 2D images with CTF variation and added noise using the "simulate data" job in cryoSPARC. CTF values were allowed to vary between 1-2 μm defocus, and the signal-to-noise ratio was set to 0.5. 10,000 particle images were created from each density model (all NTD up and all NTD down), and then mixed together for refinement with C6 symmetry. The same 3D variability analysis described above in the context of the R95G p97-ADP experimental dataset was carried out for this simulated dataset. The results of the classification are detailed in Figure S2. Notably for this simulated dataset the largest principle component vector was nearly identical to that observed for the experimental R95G p97-ADP dataset. More than 99% of the images in the simulated dataset were classified as either in the all up (49.3%) or all down (49.9%)

conformations, in agreement with expectations based on the input data (where half of the particles were fully up and half fully down).

Discussion

Effect of cooling on the up/down NTD equilibrium

A concern with any cryo-EM analysis of population distributions is how these are affected by the freezing procedure in which the specimen temperature is cooled from 4 °C or room temperature to, in our case, a bath of liquid ethane/propane maintained at near liquid nitrogen temperature (77 K). In addition how the air-water interface influences subtle thermodynamics is also an open question. Cooling rates were estimated[12] to be on the order of 10^5 K/s, based on the average rate at which a thermocouple was cooled from 273 K to 100 K. However, in that experiment the thermocouple was a 25 μm wire that is expected to cool much more slowly than a thin aqueous biological film (typically ~500 Å in thickness), so the estimate of 10^5 K/s (or 2 ms to cool over a range of 200 K) must be considered a lower bound on the cooling rate.[12] Nevertheless, as the life-times of the NTD up and down states for R95G p97-ADP are less than approximately 100 μs[13] it is not *a priori* clear whether the thermodynamics will be affected by the freezing process. In the present case there is some evidence to suggest that there is minimal perturbation. This has been evaluated by comparing populations of the up state, p_u, obtained from NMR (at 323 K), based on the position of cross-peaks that report on the up/down equilibrium,[14] with p_u calculated from the number of NTDs classified as up by the 2D cryo-EM image classification algorithm. The values of p_u obtained, 47% (NMR) *vs* 51.3±0.6% (cryo-EM), are in good agreement. The estimated error in the cryo-EM derived p_u value is the standard deviation of the measurement after dividing the dataset into 4. Notably, in a separate comparative NMR and cryo-EM study of an archaeal homologue of p97 from *Thermoplasma acidophilum*, called VAT,

similar populations were also obtained.[15] In this application, the populations of two different conformations of VAT-ATPγS were compared, corresponding to split-ring and stacked-ring structures. Fractional populations of 28±4% (split) and 72±4% (stacked) were obtained via NMR, in reasonable agreement with 42±3% and 58±3% measured by cryo-EM, despite the large temperature difference (the NMR measurements were performed at 333K).

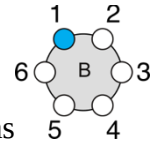
Fitting fractional populations of states to assess up/down NTD cooperativity

In what follows we present the model that we used to fit the fractional populations of the 14 R95G p97-ADP states, A-N (Figure 2A), as determined by our cryo-EM analysis (Figure 2D). In this model, we assume that the probability of a given NTD in an up or down conformation depends only on the status of the NTDs from the two immediately adjacent protomers (*i.e.*, whether they are up or down, Figure 2E; see below). We define the conditional probability of an NTD adopting the down conformation to be p_0, p_1 or p_2 for the case where 0, 1 or 2 of the nearest NTDs are in the down position, respectively, and we denote the NTD of protomer i (N_i) in the down(up) conformation by $N_i = 1 (N_i = 0)$. Thus, the conditional probabilities p_0, p_1 and p_2 are given by the following expressions:

$$\begin{aligned} p_0 &= P(N_i = 1 \mid N_{i-1} = 0, N_{i+1} = 0) \\ p_1 &= P(N_i = 1 \mid N_{i-1} = 1, N_{i+1} = 0) = P(N_i = 1 \mid N_{i-1} = 0, N_{i+1} = 1) \\ p_2 &= P(N_i = 1 \mid N_{i-1} = 1, N_{i+1} = 1) \end{aligned} \qquad \text{[S1]}$$

Here we use the canonical notation for conditional probability whereby P(A|B) refers to the probability of event A occurring given condition B. Thus, p_0 is the probability that NTD$_i$ is in the down configuration given that NTDs of the i-1 and i+1 protomers are both up, for example. Based on these definitions, the fractional populations of the 14 different NTD states (Figure 2A) can be written in terms of p_0, p_1 and p_2, as illustrated in the following three examples. Before we do this, it is important to realize that each state has a different level of degeneracy. For example,

state A has a degeneracy of 1 (*i.e.*, only a single configuration), while state B has a degeneracy of 6, as one can rotate the ring 60º clockwise/counter-clockwise (see Figure 2A) in five consecutive steps before it returns to the original position, generating five other rotationally symmetric configurations that are equivalent to the one shown. Table S1 lists the degeneracies of each state.

(i) Consider state B, which is denoted as 

(Figure 2A) for the initial example. In this representation each of the 6 NTDs is denoted by a circle, that, in turn, is coloured either blue or white to indicate a down or an up NTD conformation, respectively. The fractional population of one of the 6 degenerate configurations of state B (*i.e.*, probability of this configuration), p_B, corresponding to the NTD of protomer 1 in the down position and all other NTDs up, can be related to the conditional probability values, p_i, as follows, where we focus on the NTD from protomer 1:

$$p_B = P(N_1 = 1 \mid N_{2\text{-}6} = 0) \times P(N_{2\text{-}6} = 0). \qquad [S2]$$

Noting that

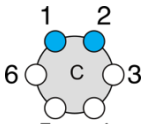
$$P(N_{2\text{-}6} = 0) = \{P(N_1 = 0, N_{2\text{-}6} = 0) + P(N_1 = 1, N_{2\text{-}6} = 0)\} \qquad [S3]$$

it follows that

$$\begin{aligned} p_B &= P(N_1 = 1 \mid N_{2\text{-}6} = 0) \times \{P(N_1 = 0, N_{2\text{-}6} = 0) + P(N_1 = 1, N_{2\text{-}6} = 0)\} \\ &= p_0 \times (p_A + p_B) \end{aligned} \qquad [S4]$$

where $P(N_1 = 1 \mid N_{2\text{-}6} = 0) = P(N_1 = 1 \mid N_2 = 0, N_6 = 0)$ since we have assumed that only the nearest neighbours affect the up/down orientation of NTD1. From Eq. [S4], we obtain

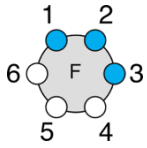
$$p_B = \frac{p_0}{1 - p_0} p_A = q_0 p_A \text{ where } q_0 = \frac{p_0}{1 - p_0}.$$

(ii) As a second example, the fractional population of one of the 6 degenerate configurations of state C 

is calculated, by again focusing on the NTD of protomer 1(N_1):

$$\begin{aligned} p_C &= P(N_1=1 \mid N_2=1, N_{3\text{-}6}=0) \times P(N_2=1, N_{3\text{-}6}=0) \\ &= P(N_1=1 \mid N_2=1, N_{3\text{-}6}=0) \times \{P(N_1=0, N_2=1, N_{3\text{-}6}=0) + P(N_1=1, N_2=1, N_{3\text{-}6}=0)\} \\ &= P(N_1=1 \mid N_2=1, N_6=0) \times \{P(N_1=0, N_2=1, N_{3\text{-}6}=0) + P(N_1=1, N_2=1, N_{3\text{-}6}=0)\} \\ &= p_1 \times (p_B + p_C) \end{aligned} \quad [S5]$$

from which it follows that $p_C = \frac{p_1}{1 - p_1} p_B = q_1 p_B = q_0 q_1 p_A$, where $q_1 = \frac{p_1}{1 - p_1}$.

(iii) As a final example, we calculate the probability of a configuration of state F 

(degeneracy of 6) by first focusing on the NTD of protomer 1 (N_1):

$$\begin{aligned} p_F &= P(N_1=1 \mid N_{2\text{-}3}=1, N_{4\text{-}6}=0) \times P(N_{2\text{-}3}=1, N_{4\text{-}6}=0) \\ &= P(N_1=1 \mid N_{2\text{-}3}=1, N_{4\text{-}6}=0) \times \{P(N_1=0, N_{2\text{-}3}=1, N_{4\text{-}6}=0) + P(N_1=1, N_{2\text{-}3}=1, N_{4\text{-}6}=0)\} \\ &= P(N_1=1 \mid N_2=1, N_6=0) \times \{P(N_1=0, N_{2\text{-}3}=1, N_{4\text{-}6}=0) + P(N_1=1, N_{2\text{-}3}=1, N_{4\text{-}6}=0)\} \\ &= p_1 \times (p_C + p_F) \end{aligned} \quad [S6]$$

so that

$$p_F = \frac{p_1}{1 - p_1} p_C = q_1 p_C = q_0 q_1^2 p_A \ . \quad [S7]$$

Alternatively, p_F could also be calculated by focusing on the NTD of protomer 2 (N_2):

$$\begin{aligned} p_F &= P(N_2=1 \mid N_{1,3}=1, N_{4\text{-}6}=0) \times P(N_{1,3}=1, N_{4\text{-}6}=0) \\ &= P(N_2=1 \mid N_{1,3}=1, N_{4\text{-}6}=0) \times \{P(N_2=0, N_{1,3}=1, N_{4\text{-}6}=0) + P(N_2=1, N_{1,3}=1, N_{4\text{-}6}=0)\} \\ &= P(N_2=1 \mid N_1=1, N_3=1) \times \{P(N_2=0, N_{1,3}=1, N_{4\text{-}6}=0) + P(N_2=1, N_{1,3}=1, N_{4\text{-}6}=0)\} \\ &= p_2 \times (p_D + p_F) \end{aligned} \quad [S8]$$

and, therefore, $p_F = \frac{p_2}{1- p_2} p_D = q_2 p_D$, where $q_2 = \frac{p_2}{1- p_2}$. Following the procedure illustrated in examples (i)-(iii) it can be shown that $p_D = q_0^2 p_A$, so that

$$p_F = q_0^2 q_2 p_A. \qquad \text{[S9]}$$

Equating the two expressions for p_F (Eqs. [S7]&[S9]), obtained by focussing on protomers N_1 and N_2, respectively, it follows that $q_1^2 = q_0 q_2$. Thus, p_0, p_1 and p_2 (or q_0, q_1 and q_2) are restrained and there are only two free parameters that are needed to calculate all the populations, p_j, via

$$\begin{aligned}
p_B &= q_0 p_A \\
p_C &= q_0 q_1 p_A \\
p_D &= q_0^2 p_A \\
p_E &= q_0^2 p_A \\
p_F &= q_0 q_1^2 p_A \qquad \text{[S10]} \\
p_G &= q_0^2 q_1 p_A \\
p_H &= q_0^2 q_1 p_A \\
p_I &= q_0^3 p_A \\
p_J &= q_0 q_1^3 p_A \\
p_K &= q_0^2 q_1^2 p_A \\
p_L &= q_0^2 q_1^2 p_A \\
p_M &= q_0 q_1^4 p_A \\
p_N &= q_1^6 p_A
\end{aligned}$$

with $q_i = \frac{p_i}{1- p_i} \quad i \hat{\mathrm{I}} \{0,1,2\}.$

In examples (i)-(iii) above we have focused on protomer N_1 in the derivations since then all of the $p_j \; j \hat{\mathrm{I}} \{A,...,N\}$ can be readily expressed in terms of p_A. However, consistent results are

obtained irrespective of how the derivations are performed. For example, if instead we focus on N_5 in the derivation of p_F

$$\begin{aligned} p_F &= P(N_5 = 0 \mid N_{1\text{-}3} = 1, N_{4,6} = 0) \times P(N_{1\text{-}3} = 1, N_{4,6} = 0) \\ &= P(N_5 = 0 \mid N_{1\text{-}3} = 1, N_{4,6} = 0) \times \{P(N_5 = 0, N_{1\text{-}3} = 1, N_{4,6} = 0) + P(N_5 = 1, N_{1\text{-}3} = 1, N_{4,6} = 0)\} \\ &= P(N_5 = 0 \mid N_4 = 0, N_6 = 0) \times \{P(N_5 = 0, N_{1\text{-}3} = 1, N_{4,6} = 0) + P(N_5 = 1, N_{1\text{-}3} = 1, N_{4,6} = 0)\} \\ &= (1 - p_0) \times (p_F + p_K) \end{aligned} \quad \text{[S11]}$$

which leads to $p_K = q_0 p_F$. Since $p_F = q_0 q_1^2 p_A$ (Eq. [S10]) it follows that $p_K = q_0 p_F = q_0^2 q_1^2 p_A$, as expected (Eq. [S10]). Finally, the fractional populations of each of the configurations of state $j \in \{A,...,N\}$, p_j, can be converted to the fractional populations of the 14 states, P_j, by including a factor that accounts for the degeneracy of each configuration due to rotational symmetry, $P_j = \kappa p_j$, where κ is given in Table S1. The fractional populations of the states, in turn, sum to unity,

$$\sum_{j=A}^{N} P_j = 1 \quad \text{[S12]}$$

so that the value of p_A can be expressed in terms of q_0 and q_1, and hence each P_j is uniquely defined by these q_i values. The fractional populations of the 14 states obtained from the analysis of the cryo-EM 2D images are fit to expressions for P_j with fitting parameters p_i such that i is any two elements of the set {0,1,2}, via

$\chi^2 = \frac{1}{n-m} \sum_{j \in \{A,B,C,..N\}} \left(\frac{P_j^{obs} - P_j^{fit}}{\sigma_j} \right)^2$, where n is the number of data points (n=14), m is the number of fitting parameters (m=2), P_j^{obs} and P_j^{fit} are the observed and fitted fractional populations of state j, respectively, and σ_j is the error in the observed fractional population, estimated from analyses of four independent subsets of the cryo-EM data.

As described above, the model that we have used assumes that the probability of the up/down conformation of a given protomer is only dependant on the up/down status of the

nearest neighbours. More complex models are, of course, also possible. For example, one could assume that the both the nearest and next nearest neighbours influence the up/down equilibrium (*i.e.*, protomers 2&6 and 3&5 affect protomer 1, where the protomer numbering increases by 1 for each successive protomer around the ring). However, this would introduce a significant number of additional probabilities (p_i values; from 3 in the present model to 10). The model that we have chosen fits the data to within error, and there is thus no reason *a priori* to consider more complex scenarios. Our model of choice is also supported by experimental data[1]. NMR spectra recorded on samples of hetero-hexamers composed of methyl-labeled R95G and unlabeled WT protomers at different molar ratios show as many as three peaks for each methyl probe reporting on the up/down NTD equilibrium, with peak intensities varying depending on relative ratios of protomer-types. One of the three peaks is at the position expected for a fully R95G p97 sample, a second peak is located at a position expected when an R95G protomer is surrounded by WT subunits and a third peak is in-between. The in-between peak increases initially in intensity as the fraction of R95G subunits grows, and then disappears when the population of R95G protomers approaches 100%. The simplest interpretation of these results is that this 'in-between' peak derives from a probe attached to an R95G protomer with one each of R95G and WT protomers as immediate neighbours. This chemical shift data provides strong evidence that the nearest neighbours have the largest effect on the NTD up/down equilibrium of a given protomer.

Synergy between NMR and cryo-EM approaches

We have presented an example where cryo-EM has been used to evaluate the extent of NTD up/down cooperativity in the context of a homo-hexameric R95G severe disease mutant of p97. A similar analysis using NMR approaches would be difficult as only bulk magnetization is recorded (*i.e.*, from all six protomers). It might be possible, however, to consider a titration

whereby different fractions of NMR visible full-length R95G protomers and NMR invisible protomers lacking an NTD are mixed, with quantification of how the up/down NTD equilibria of the full length protomers are affected, following an approach we have described previously for studies of gating cooperativity in the proteasome[16]. However, the cryo-EM based study is likely simpler in this case. On the other hand, as discussed above, it is not clear *a priori* whether the up/down NTD thermodynamics are affected by the freezing process that is necessary for cryo-EM sample preparation or by interactions between the particles and the air-water interface. In this regard, the NMR measurement of the populations of up/down NTDs is important for validation of the 'accuracy' of the cryo-EM data, emphasizing the synergy between the two technologies. A further example of synergy is provided by our previous study, which illustrates the advantages of NMR, where we considered mixtures of WT and R95G protomers, as is likely present in MSP1 patients, showing that the presence of the WT subunits has a marked effect on the NTD orientations of the R95G subunits when they co-exist in the same hexameric ring. We further extended this type of analysis to consider how the NTD equilibria in R95G protomers are affected when the other subunits in the p97 particle have more or less skewed NTD up/down populations but are not WT[1]. These experiments were made possible by differential labeling of the subunit-types (*i.e.*, only 1 subunit class was NMR visible), so that the R95G protomers could be separated from the others spectroscopically. It would be very difficult to distinguish between such protomers in a cryo-EM analysis because the only difference is a single amino-acid substitution out of ~800 residues in each p97 protomer. Moreover, even subtle differences in NTD equilibria can be readily observed by NMR through chemical shift changes that are easily measured. Once again, a strong synergy exists between the two methods, with cryo-EM

favouring studies of homo-hexameric particles and NMR preferred in studies of the structural dynamics of the biologically relevant hetero-hexameric p97 mutants involved in disease.

Supplementary Figures

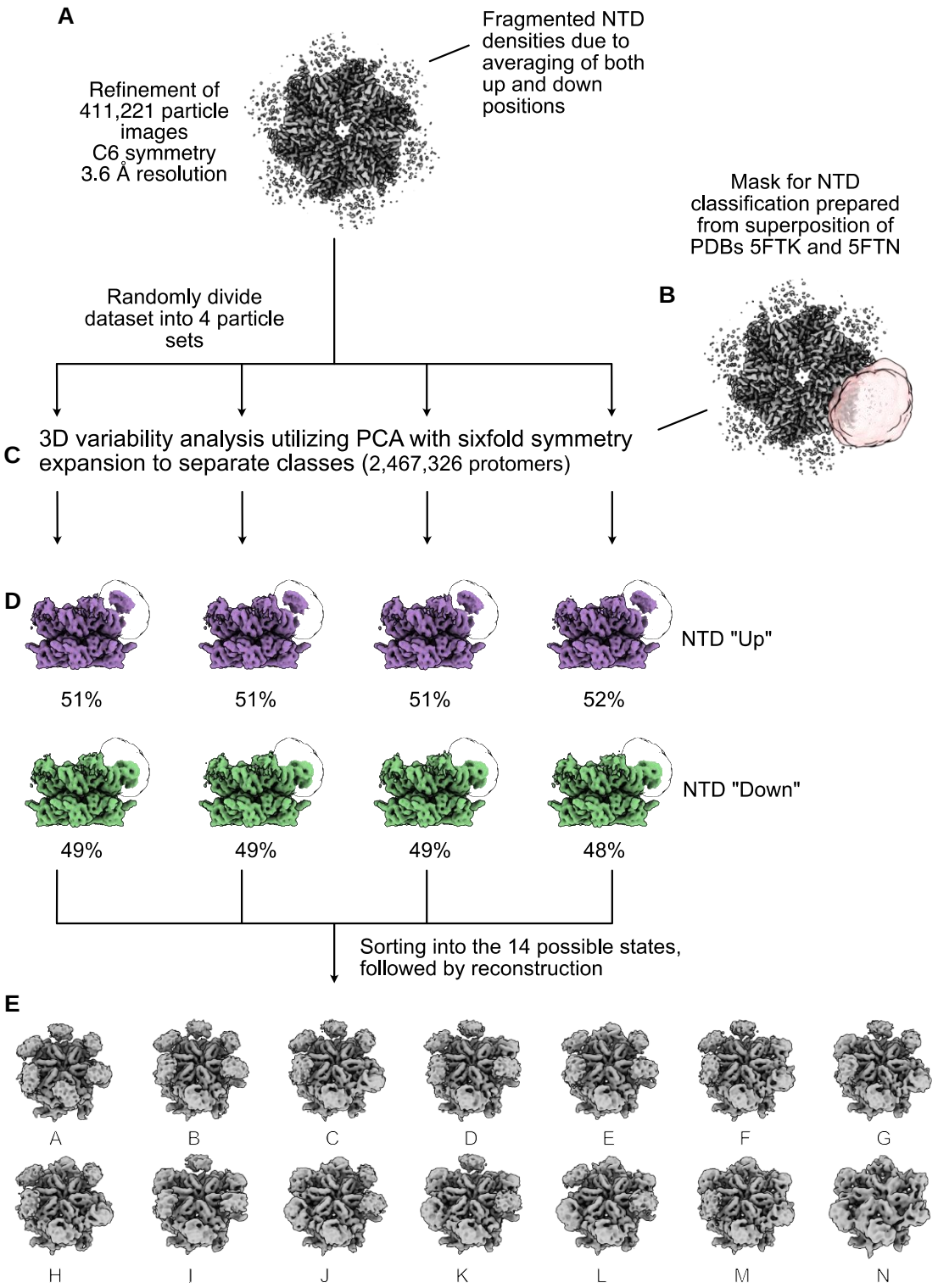


Figure S1. Workflow for the classification of cryoEM particle images. (A) After junk particles were removed by successive rounds of 2D classification, a total of 411,221 particle images were refined to a resolution of 3.6 Å using C6 symmetry. In this reconstruction the D1 and D2 domains are well resolved, however the NTDs show fragmented density, characteristic of a low-resolution feature due to mixing of multiple conformations. (B-C) In order to separate the up and down NTDs states of each protomer a 3D variability analysis/PCA was performed, as described in the SI text, on four randomly divided subsets of the data, focusing on the volume regions containing only the NTDs and with symmetry expansion. This clustering was achieved by masking the region of interest. Particles were then clustered along the PCA vectors into two classes. (D) Consensus volumes for these sorted particle sets are shown for each replicate classification, with every classification faithfully recreating either an up or down NTD state. As a result of this analysis, the configurations of all 6 NTDs are known for each particle image, as well as their relative positions, and thus each particle can be sorted into one of the 14 possible p97 states (Figure 2A). These sorted images are then reconstructed using the known poses of the particles, and the resulting volumes are shown in (E).

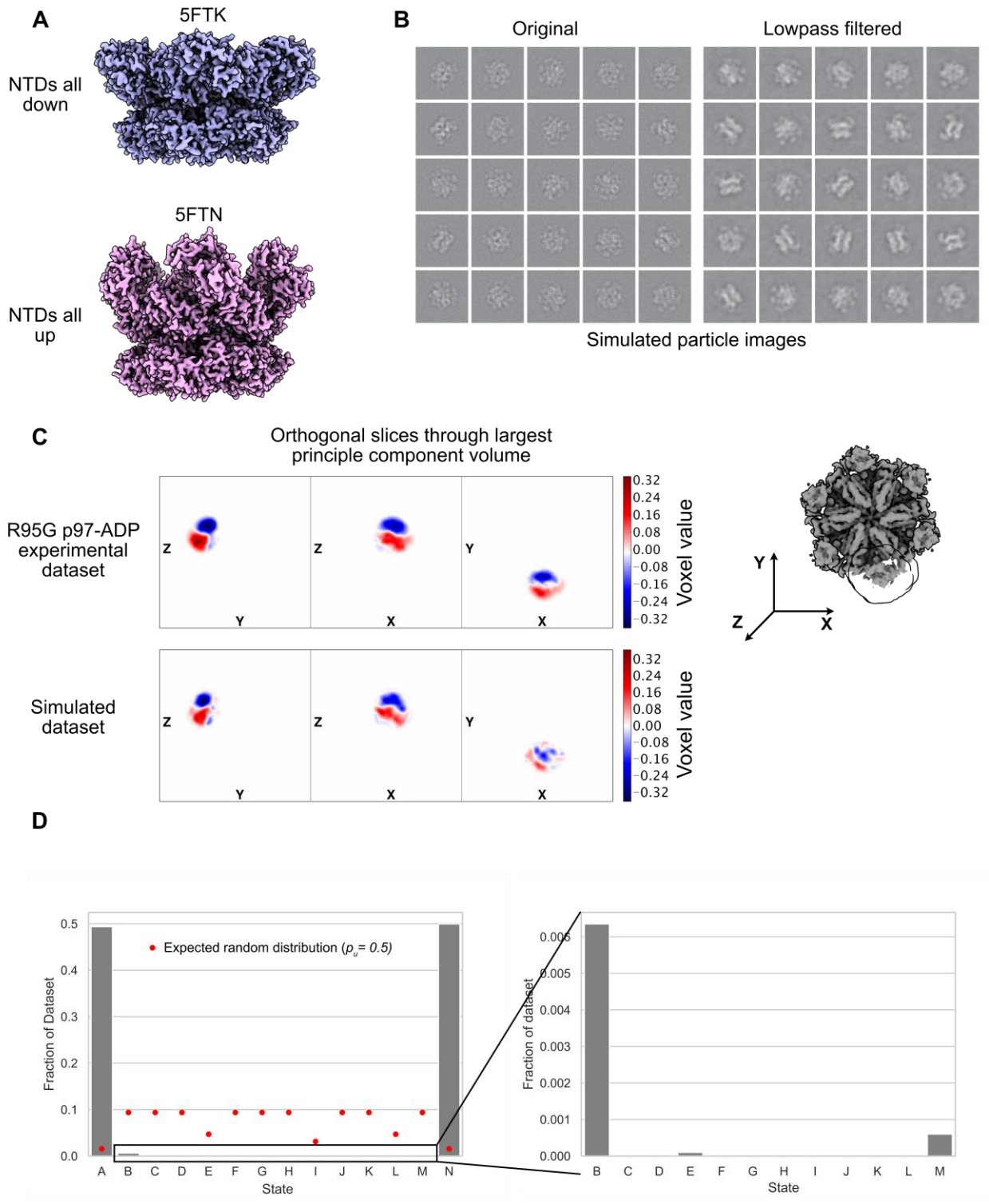


Figure S2. Analysis of simulated dataset verifies robustness of classification algorithm. (A) 3D density maps of p97 generated from PDBs 5FTK[9] and 5FTN[9], corresponding to NTDs all

down (ADP-loaded) or up (ATP bound), respectively. (B) Sample 2D projected images with microscope contrast transfer functions applied and with added noise, as described in SI text. Equal numbers of images from each of the two density maps were generated and combined to produce a simulated dataset that is used in the analysis. (C) Orthogonal slices through the largest principle component volume element calculated from analysis of the R95G p97-ADP experimental data and from the simulated dataset, using a PCA analysis approach described in SI text. The coordinate frame used (with the z-axis along the 6-fold pseudo-symmetry axis of the structure) is shown to the right. (D) (left) Relative populations of the 14 states (A-N, Figure 2A), calculated from analysis of the simulated dataset, are shown as gray bars. Assuming equal up/down NTD populations, the predicted populations of the 14 states for a random distribution of up and down NTD configurations are shown in red dots. (right) Expanded region including states B-M.

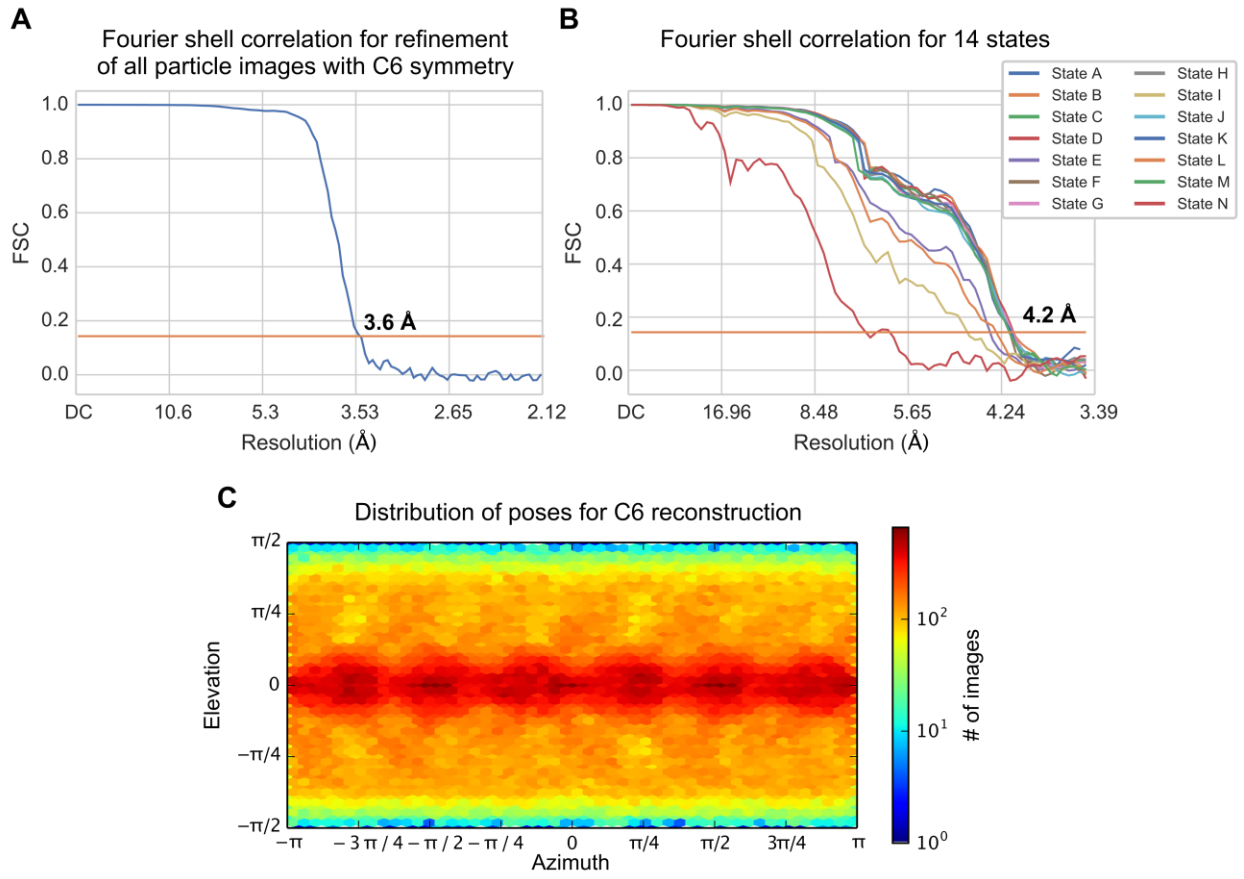


Figure S3. Fourier shell correlation curves for 3D reconstructions analyzed and angular distribution of particle image poses. (A) Fourier shell correlation curve for the 3D reconstruction with all 411,221 particle images and C6 symmetry. (B) Fourier shell correlation curves for all 14 sorted states. (C) Distribution of calculated particle image poses for the reconstruction with all 411,221 particle images and C6 symmetry.

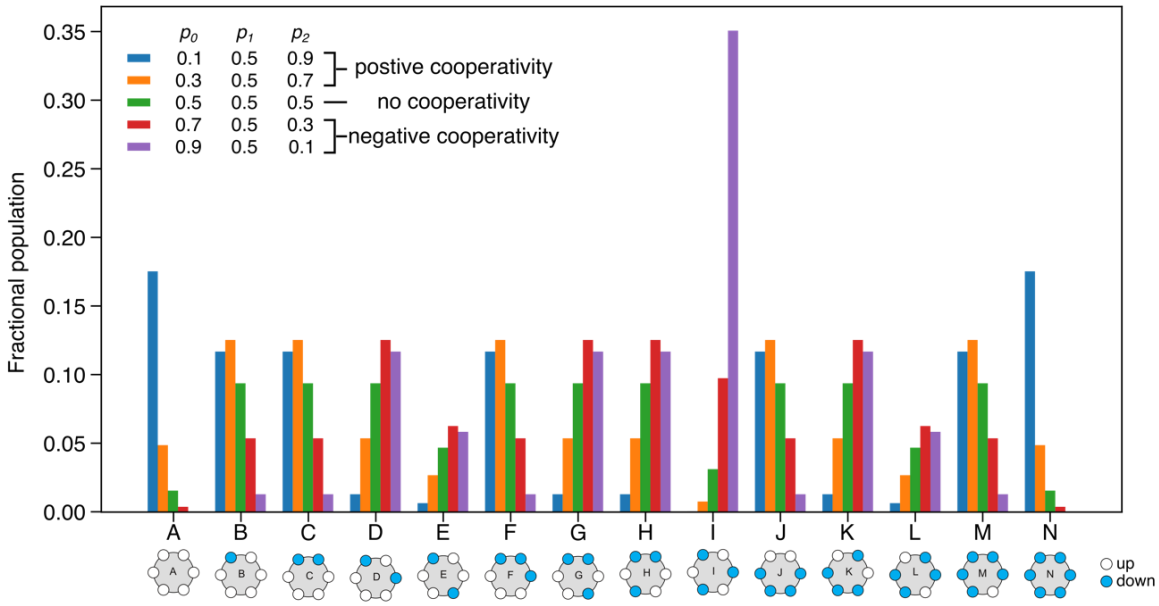


Figure S4. Simulation of the fractional populations of the 14 R95G p97-ADP states for different p_0, p_1, p_2 values. Distinct distributions are noted for cases of positive ($p_0<p_1<p_2$, blue and orange bars), negative ($p_0>p_1>p_2$, red and purple bars), and no cooperativity ($p_0=p_1=p_2$, green bars). Thus, the analysis of population distributions provides a robust means for assessing the role of nearest neighbour cooperativity in the up/down equilibrium of a given NTD.

Supplementary Tables

Table S1. Number of degeneracies (κ) for each of the 14 p97-ADP states (Figure 2A).

State	Degeneracy	State	Degeneracy
A	1	H	6
B	6	I	2
C	6	J	6
D	6	K	6
E	3	L	3
F	6	M	6
G	6	N	1

References

[1] R. Huang, Z. A. Ripstein, J. L. Rubinstein, L. E. Kay, *Proc Natl Acad Sci U S A* **2019**, *116*, 158–167.

[2] C. R. Marr, S. Benlekbir, J. L. Rubinstein, *J. Struct. Biol.* **2014**, *185*, 42–47.

[3] W. F. Tivol, A. Briegel, G. J. Jensen, *Microsc. Microanal.* **2008**, *14*, 375–379.

[4] J. L. Rubinstein, M. a. Brubaker, *J. Struct. Biol.* **2015**, *192*, 188–195.

[5] A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, *Nat. Methods* **2017**, *14*, 290–296.

[6] A. Rohou, N. Grigorieff, *J. Struct. Biol.* **2015**, *192*, 216–221.

[7] S. H. W. Scheres, *J. Struct. Biol.* **2012**, *180*, 519–530.

[8] A. Punjani, D. J. Fleet, *bioRxiv* **2020**, 2020.04.08.032466.

[9] S. Banerjee, A. Bartesaghi, A. Merk, P. Rao, S. L. Bulfer, Y. Yan, N. Green, B. Mroczkowski, R. J. Neitz, P. Wipf, et al., *Science* **2016**, *351*, 871–875.

[10] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J. Comput. Chem.* **2004**, *25*, 1605–1612.

[11] M. Zhou, Y. Li, Q. Hu, X.-C. Bai, W. Huang, C. Yan, S. H. W. Scheres, Y. Shi, *Genes Dev.* **2015**, *29*, 2349–2361.

[12] R. B. G. Ravelli, F. J. T. Nijpels, R. J. M. Henderikx, G. Weissenberger, S. Thewessem, A. Gijsbers, B. W. A. M. M. Beulen, C. López-Iglesias, P. J. Peters, *Nat. Commun.* **2020**, *11*, 2563.

[13] A. K. Schütz, E. Rennella, L. E. Kay, *Proc. Natl. Acad. Sci.* **2017**, *114*, E6822-6829.

[14] A. K. Schuetz, L. E. Kay, *Elife* **2016**, *5*, e20143.

[15] R. Huang, Z. A. Ripstein, R. Augustyniak, M. Lazniewski, K. Ginalski, L. E. Kay, J. L. Rubinstein, *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, E4190–E4199.

[16] R. Huang, F. Pérez, L. E. Kay, *Proc. Natl. Acad. Sci.* **2017**, *114*, E9846–E9854.