Mechanistic insight into substrate processing and allosteric inhibition of human p97

Man Pan[1#*], Yuanyuan Yu[1#], Huasong Ai[2], Qingyun Zheng[2], Yuan Xie[1], Lei Liu[2*], Minglei Zhao[1*]

[1] Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL 60637, USA.

[2] Tsinghua-Peking Center for Life Sciences, Department of Chemistry, Tsinghua University, Beijing 100084, China.

[#]These authors contributed equally to the work.

[*]Corresponding authors:

Minglei Zhao: mlzhao@uchicago.edu

Lei Liu: lliu@mail.tsinghua.edu.cn

Man Pan: panm@uchicago.edu

ABSTRACT

p97, also known as valosin-containing protein (VCP), processes ubiquitinated substrates and plays a central role in cellular protein homeostasis. Mutations in human p97 are associated with multisystem proteinopathy (MSP), a dominantly inherited degenerative disorder that can affect muscle, bone and the central nervous system. It is also a drug target for cancer therapy with various inhibitors developed over the past decade. Despite significant structural insights into the fungal homologue of p97, Cdc48, little is known about how human p97 processes its substrates and how the activity is allosterically affected by inhibitors. Here, we report a series of cryo-electron microscopy (cryo-EM) structures of substrate-engaged human p97 complex with resolutions ranging from 2.9 to 3.8 Å that captured “power stroke”-like motions of both the D1 and D2 ATPase rings of p97. The structures elucidated how the unfolded substrate is engaged in the pore at atomic level. Critical conformational changes of the inter-subunit signaling (ISS) motifs were revealed, providing molecular insights into substrate translocation. Furthermore, we also determined cryo-EM structures of human p97 in complex with NMS-873, the most potent p97 inhibitor, at a resolution of 2.4 Å. The structures showed that NMS-873 binds at a cryptic groove in the D2 domain and interacts with the ISS motif, preventing its conformational change, thus blocking substrate translocation allosterically. Finally, using NMS-873 at a substoichiometric concentration, we captured a series of intermediate states, suggesting how the cofactor Npl4 coordinates with the D1 ring of p97 to initiate the translocation.

Keywords

p97, NMS-873, single-particle cryo-EM, translocation, allosteric inhibition

INTRODUCTION

p97 is a 500 kDa protein that belongs to the family of ATPases Associated with diverse cellular Activities (AAA)[1-3]. It undergoes repetitive ATP hydrolysis and conformational changes to extract ubiquitinated substrates from various cellular components, such as membranes, ribosomes, and chromatin[4]. p97 plays a central role in many cellular processes, including endoplasmic reticulum-associated degradation (ERAD), mitochondrial-associated degradation, chromatin-associated degradation, autophagy, and endosomal trafficking[5]. Mutations in human p97 have been found in multisystem proteinopathy (MSP), a dominantly inherited degenerative disorder that can affect muscle, bone and the central nervous system, that can manifest clinically as inclusion body myopathy with early-onset Paget disease and frontotemporal dementia (IBMPFD) and familial amyotrophic lateral sclerosis (ALS)[6,7]. Moreover, recent studies have also targeted p97 for the treatment of cancer and viral infections[8,9].

p97 has two tandem AAA domains named D1 and D2 and an additional N domain at the N-terminus (**Fig. 1a**). This protein functions as a homomeric hexamer, with twelve copies of ATPase domains organized into two rings[10]. Crystal structures of full-length p97 were determined more than a decade ago[11,12]. High-resolution single-particle cryo-electron microscopy (cryo-EM) structures have also been reported[13,14]. However, all the structures of p97 determined thus far adopt a sixfold symmetric conformation, providing limited information regarding the mechanism since almost all AAA proteins, including close homologues of p97, break symmetry when engaged with substrates[15-17]. Recent cryo-EM studies of Cdc48, the yeast homolog of p97, in complex with either in vivo[18] or in vitro[19] substrates provided crucial insight into the working states. However, with medium resolution (3.7 Å to 4.2 Å) and limited conformational states, critical mechanistic questions, such as how the AAA domains in two rings coordinate to achieve translocation, remain to be answered. It is also not clear whether human p97 shares the same mechanism as its yeast homologue. In addition, owing to its biomedical importance, many small-molecule inhibitors of p97 have been developed in the past

decade[8,20,21]. Little is known regarding the allosteric inhibition mechanism due to the lack of a p97 structure in a working state, which is of considerable biological and clinical interest.

In this study, starting from a fully assembled p97 complex with the cofactor Npl4/Ufd1 and a substrate, we triggered ATP hydrolysis and determined a series of substrate-engaged p97 structures with resolutions ranging from 2.9 to 3.8 Å using single-particle cryo-EM. The structures captured “power stroke”-like motions of both the D1 and D2 rings of p97 and elucidated how the two rings coordinate to achieve translocation through the pore loops, intersubunit signaling (ISS) motifs, and N- and C-terminal tails. We further determined the cryo-EM structures of p97 in the presence of NMS-873, an allosteric noncompetitive inhibitor[21], at a resolution of 2.4 Å and uncovered its inhibition mechanism. It is worth noting that NMS-873 has been shown to be active in tumor cell lines and IBMPFD patient fibroblasts[21,22]. More interestingly, low nanomolar concentrations of NMS-873 was recently reported to inhibit the replication of several viruses, including severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and influenza virus, through its inhibition on p97[23-25]. Finally, using NMS-873 at a substoichiometric concentration, we captured a series of intermediate states, suggesting how p97 coordinates with its cofactor Npl4 to initiate substrate translocation. Taken together, our results established a comprehensive understanding of substrate processing and allosteric inhibition of human p97.

RESULTS

Capture the working states of human p97

First, we assembled complexes of p97 from recombinantly overexpressed components without introducing additional ATP. Two mutations, A232E and E578Q, were introduced into the human p97 construct to stabilize the complex as previously described [26]. The A232E mutation is found in MSP patients and is known as an activating mutant with its N domains locked in the "up conformation"[27]. E578Q is a mutation in the Walker B motif of the D2 domain that slows substrate unfolding[19,28]. Combined, the two mutations decreased but maintained the unfolding activity of p97 (**Extended Data Fig. 1a**) and minimized the conformational heterogeneity resulting from the N domains. All the structural work in this study used the p97 (A232E/E578Q) mutant and Npl4/Ufd1 as the cofactors. For substrates, either polyubiquitinated Ub-Eos (Ub_n-Eos) or K48-linked hexa-ubiquitin (Ub_6) was used. The assembled complex, named p97-Npl4/Ufd1-Ub_n-Eos or p97-Npl4/Ufd1-Ub_6, eluted as single peaks in the size-exclusion chromatograms (**Extended Data Fig. 1b and c**). To capture the working states of p97, ATP hydrolysis was triggered by adding 5 mM ATP to the purified complexes followed by incubation at room temperature for 5 minutes. To relieve the orientation preference, either CHAPSO or fluorinated octyl maltoside (FOM) was added immediately before sample vitrification (**Fig. 1b**). Three datasets, namely, p97-Npl4/Ufd1-Ub_n-Eos-CHAPSO, p97-Npl4/Ufd1-Ub_n-Eos-FOM, and p97-Npl4/Ufd1-Ub_6-FOM, were collected, and a total of 9 different cryo-EM maps with resolutions ranging from 2.9 to 3.8 Å were reconstructed using single-particle analysis (**Extended Data Figs. 2-4, Extended Data Table 1**). Atomic models were built based on existing p97 structures (**Extended Data Table 2**). For well-resolved regions, individual residues were manually built into the density. For flexible regions, rigid body docking was used to fit individual secondary structure elements or subdomains (see Methods for details).

Open and closed states of the D2 ring

In the p97-Npl4/Ufd1-Ub_n-Eos-CHAPSO dataset, two conformational states of the D2 ring were resolved (**Fig. 1c and d**). In the open state (3.6 Å, **Fig. 1c**, corresponding to class 2 in **Extended Data Fig. 2**), all six D2 domains interacted with the substrate and formed an open, right-handed spiral, whereas in the closed state (3.1 Å, **Fig. 1d**, corresponding to class 1 in **Extended Data Fig. 2**), the bottom D2 domain (termed chain F) detached from the substrate and shifted up to interact with the top D2 domain (termed chain E), closing the gap of the spiral in the open state. Only five D2 domains (chains A-E) interacted with the substrate in the closed state. The two states suggested a "power stroke"-like motion of the D2 domain (**Supplemental Movie 1**). Compared to the same complex structure in the presence of ATPγS that was previously published[26] (PDB ID: 7JY5), there was a relative 20 degree rotation of the D2 ring to the D1 ring in both states (**Fig. 1e and Extended Data Fig. 5a**). A superimposition of the two states showed that the D2 domain immediately above chain F (chain A) also shifted up from the open state to the closed state (**Extended Data Fig. 6a** and **Supplemental Movie 1**). The particles in the open state were significantly less than those in the closed state (13% vs. 70%), suggesting that the open state is relatively transient. A third class was resolved from the 3D classification (3.8 Å, class 3 in **Extended Data Fig. 2**), with better resolved density for the cofactor Npl4 but otherwise the same conformation as the closed state.

Conformational changes in the D1 ring

In the p97-Npl4/Ufd1-Ub_n-Eos-FOM dataset, three conformational states of the D1 ring were resolved (**Extended Data Fig. 3**). All the D2 rings were in the same conformation as the closed state described above. Using the same naming convention, we discovered from the superimposed structures that chain A shifted up slightly from class 1 (3.6 Å) to class 2 (3.4 Å) and much further from class 2 to class 3 (3.0 Å) (**Extended Data Fig. 7a**). No open spiral conformation similar to the open state of the D2 ring was observed for the D1 ring, although a right-handed spiral was still formed by chains B-F. The three states of the D1 ring also

suggested a “power stroke”-like motion despite the differences from the D2 ring (**Supplemental Movie 2**). It is worth noting that no conformational changes in the D1 ring were observed for the p97-Npl4/Ufd1-Ub_n-Eos-CHAPSO dataset. The conformation of the D1 ring in the open and closed states was very similar to class 3 of the p97-Npl4/Ufd1-Ub_n-Eos-FOM dataset (see Discussion).

Sequential but asynchronous hydrolysis in D1 and D2 rings

The resolutions of the maps allowed us to build atomic models and identify the nucleotides in most nucleotide binding sites (**Extended Data Figs. 6d and 7d**). For the less well resolved chains A and F, the identity of the nucleotides could not be determined with certainty, but the density of a nucleotide was clearly visible (**Extended Data Figs. 6d and 7d**). All nucleotide binding sites were occupied, suggesting that the nucleotide exchange might be very fast. When superimposing all the AAA domains within the same ring, two major conformations were observed: one with a larger angle between the α/β and the α subdomain, potentially corresponding to an ATP binding conformation, and the other with a smaller angle between the α/β and the α subdomain, potentially corresponding to an ADP binding conformation (**Extended Data Figs. 6c and 7c**). Such binary conformations have been observed for N-ethylmaleimide sensitive factor (NSF), a close homologue of human p97[17].

When focusing on the well-resolved sites, we discovered a new compact ATP binding mode in both the D1 and D2 rings involving a previously defined ISS motif[29]. Specifically, the ISS motif changed from a helical conformation to a triangular loop that inserted into the neighboring nucleotide binding site (**Fig. 1e and f**). L335 in the D1 domain and M611 in the D2 domain are the key residues interacting with the hydrophobic residues in the Walker B motifs of the neighboring subunits (**Fig. 1f**). As a result, ATP is coordinated by three more residues than the ATPγS-bound structure (PDB ID: 7JY5), including one additional arginine finger residue (R362 in D1 and R638 in D2), an aspartic acid from the ISS motif (D333 in D1 and D609 in D2),

and a basic residue following the Walker A motif (R256 in D1 and K529 in D2) (**Fig. 1f**). This binding mode was observed only when ATP was engaged and only for two to four consecutive binding sites in both rings (**Fig. 1e**, **Extended Data Figs. 6b and 7b**). We hypothesize that this tight binding mode results from rotational compression in ATPase rings essential for substrate translocation. For ATP hydrolysis to occur, the ISS motif needs to be retracted. For example, the ISS motif in the D2 domain of chain A retracted in the closed state (not in the open state) since the neighboring D2 domain in chain B is the next subunit to undergo ATP hydrolysis (**Fig. 1e and Extended Data Fig. 6b**).

Another unexpected observation is that the hydrolysis in the D1 and D2 rings is not in sync, which means that at a given time, the D1 and D2 domains of the same chain may engage different nucleotides. If viewed from the top, hydrolysis propagates counterclockwise in both ATPase rings, with the D1 ring one or two subunits ahead of the D2 ring (**Fig. 1e**). This feature is unique for human p97 and has not been observed for other type II AAA proteins, including yeast Cdc48.

Interactions between the pore loops and the substrate

In all six maps discussed above, we observed the density corresponding to a translocating peptide going through the central pore of p97 (**Fig. 2a**). A continuous extended peptide could be modeled into the density. The peptide went through the pore from the N-terminus (bottom in **Fig. 2a**) to the C-terminus (top in **Fig. 2a**), as the other direction fit much worse into the density. This directionality of translocating peptide is consistent with that of Cdc48[19,30]. The density corresponding to the side chains was averaged, and polyalanine was modeled.

In general, the density of the translocating peptide in the D2 ring was better than that in the D1 ring. Three residues in pore loop-I (PL-I) of the D2 domain, namely, M550, W551, and F552, interacted with the peptide like a “pitching grip” (**Fig. 2a**). Each PL-I formed two hydrogen

bonds with the nitrogen and carbonyl oxygen of one residue in the translocating peptide through the carbonyl oxygen of M550 and nitrogen of F552. Every other residue in the translocating peptide was hydrogen bonded by a PL-I so 12 residues were engaged with six PL-I in the D2 ring, which formed a spiral in the open state, with chain F at the bottom and chain E at the top (**Fig. 2a**). In the closed state, the PL-I of chain F detached from the translocating peptide. The remaining five PL-I formed the spiral in a very similar way except for the side chain of M550 in chain A (now the bottom chain), which was pointing downwards and would not allow an additional PL-I to engage (**Fig. 2a inset**). In both the open and closed states, the density of five pore loop-II (PL-II) in the D2 ring was resolved (chains A to E) but at lower resolution (**Fig. 2a inset**). PL-II in the D2 ring did not form hydrogen bonds with the translocating peptide and likely played a less important role in translocation. We mutated the residues of PL-I and PL-II and performed the substrate unfolding assay[1,26,31]. Indeed, the PL-I mutation (W551A/F552A) abolished the unfolding activity, whereas the PL-II mutation (I590A/D592A) showed no effect (**Fig. 2d**). A closer examination revealed that R599 was π-stacked with F552 of PL-I (**Fig. 2b**). This interaction might be critical to stabilize the “pitching grip” conformation of PL-I since the R599A mutant did not show any unfolding activity, similar to the PL-I mutation (**Fig. 2d**). In contrast, the pore loops in the D2 domain of the nontranslocating structure (PDB ID: 7JY5) were mostly unstructured (**Fig. 2c**).

PL-I in the D1 ring interreacted with the translocating peptide in the same fashion as that in the D2 ring, except that only four PL-I (chains B-E) were engaged and the three residues were K277, L278, and A279 (**Fig. 2a**). The “pitching grip” and hydrogen bonding pattern still existed, but the interaction was not as strong as that of the D2 ring since the average length of the hydrogen bonds was 3.9 Å in the D1 ring and 3.4 Å in the D2 ring. Two PL-II of the D1 ring (chains C and D) were involved in the interaction with the peptide by forming a hydrogen bond through H317 (**Fig. 2a**). Interestingly, H317 of chains B and C also stacked with W551 (PL-I of

D2) of chains D and E, respectively (**Fig. 2a**, red double arrow). This π-π stacking might be involved in the coordination of two ATPase rings.

To further investigate the behavior of p97 with substrates of different sizes, we used a much smaller substrate, Ub_6 (**Extended Data Fig. 1c**), and collected a third dataset, p97-Npl4/Ufd1-Ub_6-FOM (**Extended Data Fig. 4**). Again, three conformations of the D1 ring that were very similar to those of the p97-Npl4/Ufd1-Ub_n-Eos-FOM dataset were resolved. Intriguingly, the density corresponding to the translocating peptide was resolved only in the D2 ring (**Fig. 2e**). The D1 ring was essentially void of the translocating peptide despite adopting the same conformation. One possible explanation is that the D1 ring is engaged only when a "complicated" substrate such as Ub_n-Eos is encountered.

We noticed that the D2 ring of the closed state is similar to one of the Cdc48 structures[30]. Both have five PL-I engaging with the translocating peptide. When the D2 rings from two structures were superimposed, it was obvious that the two rings of p97 are out of register, that is, the lowest D2 and D1 domains are not in the same chain (**Fig. 2f**), whereas in Cdc48, the two rings essentially formed a single "split washer" (**Fig. 2g**). This out-of-register feature is consistent with the asynchronous hydrolysis occurring in the D1 and D2 rings.

N- and C-terminal tails and the linker between D1 and D2 rings

In addition to the ATPase rings, some of the N- and C-terminal tails (**Fig. 1a**) of p97 were also resolved in the cryo-EM maps. Both the resolved N-terminal (L12 to R22) and C-terminal (G767-S775) tails interacted with the neighboring subunit in a counterclockwise manner if viewed from the top (N-terminal tail, **Fig. 3a**) or a clockwise manner if viewed from the bottom (C-terminal tail, **Fig. 3c**). The N-terminal tail interacted with an acidic patch in the α subunit of the D1 domain (**Fig. 3b**). The C-terminal tail interacted with a groove in the α subunit of the D2 domain. Such interactions have not been described before and may play important roles in the coordination between different chains. Truncations of either the N- or C-terminal tails greatly

affected the unfolding activity of p97 (**Fig. 3f**). The linker between the D1 and D2 rings was resolved for chains B-E. Compared to the structure of nontranslocating p97 (PDB ID: 7JY5), a 19 Å shift of L464 was observed (**Fig. 3e**). In the translocating structure, L464 interacted with A569 in the D2 domain, which contributed to the relative twist of the D2 ring (**Fig. 1e**). The L464A mutation decreased the unfolding activity of p97 (**Fig. 3f**). Notably, the N- and C-terminal tails and the linker regions are less conserved between human p97 and yeast Cdc48, which might contribute to the different conformations of the ATPase rings.

NMS-873 inhibits the translocation by locking the ISS motif

Inhibition of p97 has emerged as a promising strategy in anticancer and antiviral studies[8]. NMS-873, the first discovered noncompetitive inhibitor of p97 (**Fig. 4a**), was recently characterized as a candidate for the treatment of cancers, neurodegenerative diseases, and viral infections[22-24]. However, the mechanism of inhibition remained unclear. We introduced NMS-873 at saturation (80 μM) after triggering ATP hydrolysis and substrate translocation (**Fig. 4b**). Two datasets were collected using Ub_n-Eos and Ub_6 as the substrate. A single map was obtained from each dataset with overall resolutions of 2.4 Å (Ub_6, **Extended Data Fig. 8**) and 2.8 Å (Ub_n-Eos, **Extended Data Fig. 9**). The two maps are essentially identical; therefore, only the map with the higher resolution is shown (**Fig. 4c**). NMS-873 binds at a cryptic groove next to the ISS motif of the D2 domain (**Fig. 4d**) and is surrounded by many hydrophobic residues (**Fig. 4e**). Through the interaction with the alkyl chain of K615, NMS-873 locked the ISS motif and prevented the essential conformational change discovered in the translocating structures (**Fig. 4f**). As a result, no translocating peptides were found in the central pore, and the structure was 6-fold symmetric, although a substrate was present, and ATP hydrolysis was triggered before adding NMS-873.

Compared to the ATPγS-bound nontranslocating structure (PDB ID: 7JY5), the relative positions of the ATPase rings in the NMS-873-bound structure are similar (**Extended Data Fig.**

5b). The bottom opening of the D2 ring was expanded, which is similar to that of the ADP-bound structures[13], although ATP was engaged (**Extended Data Fig. 5c**). The pore loops in the D2 domain of the NMS-873-bound structure were resolved but retracted from the pore compared to the translocating structure (**Extended Data Fig. 5e**). The smaller angle between the α/β and the α subdomains suggested that it is more similar to an ADP-bound state (**Extended Data Fig. 5d**). Indeed, superimposing the NMS-873-bound structure with another inhibitor-bound structure (UPCDC30245, PDB ID: 5FTJ [13]) showed surprising similarities, including the conformations of pore loops and a shared binding pocket of the two inhibitors in the D2 domain (**Extended Data Figs. 5f and 5g**). The difference is that the UPCDC30245-bound structure was determined in the presence of ADP without a cofactor and a substrate, and the compound was not fully resolved in the cryo-EM map[13]. Importantly, the interaction between K615 and the biphenyl group of NMS-873, which essentially locks the ISS motif, does not occur in the UPCDC30245-bound structure (**Fig. 4f and Extended Data Fig. 5g**). Previous studies found that UPCDC30245, a phenyl indole amide-based allosteric inhibitor, is an uncompetitive inhibitor and binds only in the presence of ADP, whereas triazole-based NMS-873 is a noncompetitive inhibitor and can bind in the absence of nucleotides[21]. This locking mechanism of NMS-873 through K615 might be the key for noncompetitive allosteric inhibition.

Conformational states of Npl4 revealed by NMS-873 at a substoichiometric concentration

We further discovered that NMS-873 did not fully inhibit the unfolding activity of p97 at a substoichiometric concentration (6 equivalents (eq) is the stoichiometric concentration, **Fig. 4g**). On the reasoning that intermediate structures could be captured using NMS-873 as a modulator, a dataset was collected following the same protocol using Ub_n-Eos as the substrate and in the presence of 10 μM NMS-873 (~0.2 eq, **Fig. 5a**). As expected, multiple intermediate structures were resolved after 3D classification (**Extended Data Fig. 10**). Two out of six maps (~11% particles) are in the nontranslocating state, with flat and symmetric ATPase rings and no density

for a translocating peptide (**Fig. 5c**). One of the nontranslocating maps (Class 1, ~8% particles) showed the density of NMS-873 (**Fig. 5c**, top) in all six binding pockets, suggesting that the binding of NMS-873 is highly cooperative. The cofactor Npl4 was resolved in four out of six maps (**Extended Data Fig. 10**). Both nontranslocating maps showed density for Npl4, with either one or two of its zinc finger motifs (ZF1 and ZF2, **Fig. 5b**) binding on top of the D1 ring (**Fig. 5c**). Such seesaw conformations of Npl4 have been described in a previous study[26]. The four translocating structures all had spiral ATPase rings and a translocating peptide similar to the conformation of the closed state (**Extended Data Fig. 10**). Two maps (class 5 and class 6) had very poor densities of Npl4 and differed in the D1 ring in a similar way to class 1 and class 3 of the p97-Npl4/Ufd1-Ub_n-Eos-FOM dataset. The other two maps showed density for Npl4, but with its zinc finger motifs interacting with different chains (**Fig. 5d**), suggesting a rotation of Npl4 on top of the D1 ring. Although the resolution of Npl4 is not sufficient for atomic model building, a trace density connected to the translocating peptide can be seen to be engaged in a groove of Npl4 (**Fig. 5d**, top), which is reminiscent of a substrate processing complex structure of Cdc48[19].

DISCUSSION

Here, we report the cryo-EM structures of human p97 in several working states, with a translocating peptide going through the central pore. Together, the structures revealed “power stroke”-like conformational changes in both the D1 and D2 rings. Surprisingly, the two AAA domains undergoing power stroke motion in each ring were not from the same chain, which seemed to be a unique feature of human p97 and was not observed for Cdc48 and other Type-II AAA proteins[15,16,19,30,32,33]. The unusual structural feature corresponds to the asynchronous hydrolysis taking place in the two ATPase rings (**Fig. 1e**). Regarding translocation, the D2 ring seemed to play a more important role by tightly interacting with the peptide through backbone hydrogen bonds. The D1 ring was engaged when encountering a much larger substrate. Previous biochemical studies have shown that the D2 ring of p97 had higher ATPase activity and provided the driving force for translocation, which is consistent with our structures[34]. The structures of Cdc48 also showed fewer interactions between the translocating peptide and the pore loops of the D1 ring[19], further supporting this hypothesis. Although the structures strongly support a sequential hydrolysis model[35], it is not clear whether the power strokes in the D1 and D2 rings take place concurrently. Multiple structures from a single dataset are needed to address this question. In this study, the motions of the D1 and D2 rings were resolved from two datasets, one using CHAPSO and the other using FOM as the detergent to fix the orientation preference. CHAPSO required a much higher concentration of the complex than FOM (8x) to keep the samples in the hole but seemed to preserve the complex better and yielded more particles with Npl4. Since the open state is a minor conformation (7% of the total particles, **Extended Data Fig. 2**), the higher concentration of the CHAPSO dataset may have contributed to its discovery.

The cryo-EM structures of human p97 in complex with NMS-873 revealed the mechanism of allosteric inhibition by locking the ISS motif in the D2 domain, which together with the ISS motif in the D1 domain played critical roles in forming the compressed ATPase rings

essential for substrate translocation (**Fig. 1f**). Targeting the ISS motif of p97, therefore, may serve as a potential strategy for the development of new inhibitors.

Multiple structures in the presence of NMS-873 at a substoichiometric concentration suggested that Npl4 may play a critical role in the initiation of translocation by unfolding the ubiquitin chain and threading it into the central pore through a seesaw motion. The pore loops in the D1 and D2 rings catch the thread-in peptide, and the two rings are compressed to spiral conformations. Following sequential ATP hydrolysis and power stroke motions circulating in both rings, the peptide in the pore is translocated downward. Npl4 might be unnecessary once the translocation starts.

Mutations in p97 have been linked to MSP[6,36]. Structures of p97 in working states provided crucial insights into the consequences of the mutations (**Extended Data Table 3, Supplemental Movie 3**). The resolutions of the cryo-EM maps in this study are among the highest of type II AAA proteins[15,17,19,30,33,37], allowing a comprehensive understanding of substrate processing and allosteric inhibition of human p97 and shedding light on the mechanism of other type II AAA proteins.

METHODS

Protein expression and purification

Wildtype human p97, human p97 mutants, Npl4/Ufd1, His_6-Ub^{G76V}-Ub^{G76V}-mEos3.2 (Ub-Eos), human UBA1, Ube25k and gp78RING-Ube2g2 were purified as previously described[26,31,38]. Briefly, all proteins were expressed in *E. coli* BL21(DE3) cells and purified through Ni-NTA resin at 4 °C. All purified proteins were further buffer-exchanged to the Storage Buffer containing 50 mM HEPES, pH 8.0, 150 mM NaCl, 1 mM $MgCl_2$, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP) before snap freezing.

Preparation of polyubiquitinated Ub-Eos and polyubiquitin chain

For cryo-EM studies, His-tagged Ub-Eos was directly used for polyubiquitination. For the substrate unfolding assay, Ub-Eos was first irradiated under LED UV light (385-395 nm, uvBeast V2) for 1 hour at 4 ºC to induce the photoconversion of mEos3.2, followed by polyubiquitination. The polyubiquitination reaction was performed as previously described[31]. Briefly, 20 μM Ub-Eos was mixed with 1 μM UBA1, 20 μM gp78RING-Ube2g2, and 500 μM ubiquitin in the Ubiquitination Buffer containing 20 mM Hepes, pH 7.4, 150mM KCl, 10 mM ATP, and 10 mM $MgCl_2$. After incubation at 37 ºC overnight, the mixture was incubated with Ni-NTA resin to remove free ubiquitin chains. The elution was further separated using a Superdex 200 (GE Healthcare) size-exclusion column equilibrated in the Storage Buffer. Polyubiquitinated Ub-Eos (Ub_n-Eos) with longer ubiquitin chains (estimated over 10 ubiquitin subunits) was collected and flash frozen for the substrate unfolding assay.

K48-linked polyubiquitin chains were assembled by mixing 1 μM UBA1, 10 μM Ube25k, and 1 mM ubiquitin in the Ubiquitination Buffer at 37 ºC for 4 hours. The reaction mixture was then diluted 10 folds with 50 mM NaOAc, pH 4.5, and separated by a Mono S cation exchange column. The corresponding peak for hexa-ubiquitin chain (Ub_6) was collected and further

purified with a Superdex 75 (GE Healthcare) size-exclusion column equilibrated in the Storage Buffer.

Assembly of substrate-engaged p97 complex

p97-Npl4/Ufd1 complex was assembled as previously described[26]. A p97 mutant bearing A232E and E578Q mutations was used in order to decrease the heterogeneity of N domains and reduce the unfolding activity of the complex. Two-fold molar excess of Ub_6 or Ub_n-Eos was added to the p97-Npl4/Ufd1 complex, followed by gel filtration using a Superose 6 column (GE Healthcare) equilibrated in the Storage Buffer (**Extended Data Fig. 1b and 1c**). The assembled substrate-engaged p97 complex was flash frozen in liquid nitrogen for structural studies. No nucleotide was supplemented during the assembly.

Specimen preparation for single-particle cryo-EM

All samples for cryo-EM were prepared as previously described[26], with the following modifications. Either fluorinated octyl maltoside (FOM) or CHAPSO was used to relieve the preferred orientations instead of IGEPAL CA-630. The complex was concentrated to ~2.5 mg/mL followed by incubation with 5 mM ATP at room temperature for 5 minutes before adding FOM (final concentration 0.01% v/v, Fig. 1b). Glow discharged Quantifoil Cu 1.2/1.3 grids were used for FOM supplemented samples. Alternatively, the samples were concentrated to ~20mg/mL followed by incubation with 5 mM ATP at room temperature for 5 minutes before adding CHAPSO (final concentration 4 mM, Fig. 1b). Non-glow-discharged Quantifoil Cu 1.2/1.3 grids were used for CHAPSO supplemented samples. In both cases, 3.5 μL sample was applied to the grid and was blotted for 1 second using standard Vitrobot filter paper (Ted Pella, 47000-100) before plunge freezing in liquid ethane. For substrate-engaged p97-Npl4/Ufd1 complex in the presence of NMS-873, besides the aforementioned procedures, NMS-873 was supplemented right after the incubation with ATP at either 10 μM or 80 μM final concentration.

The samples were incubated for another 30 min at room temperature before adding the detergents and plunge freezing (Fig. 4b and 5a).

Data collection for single-particle cryo-EM

Data collection was performed either in Advanced Electron Microscopy Facility at the University of Chicago or National Cryo-Electron Microscopy Facility at the National Cancer Institute. All datasets were acquired as movie stacks with a Titan Krios electron microscope operating at 300 kV, equipped with either a Gatan K2 Summit or K3 direct detector camera. A single stack typically consists of 40 frames with a total exposure around 50 electrons/$Å^2$. The defocus range was set at -1.0 to -2.5 µm. See **Extended Data Table 1** for the details.

Image processing

Movie stacks were subjected to beam-induced motion correction using MotionCor2[39]. CTF parameters for each micrograph were determined using CTFFIND4[40]. Particle picking, two- and three-dimensional classifications, three-dimensional refinement, and local resolution estimation were performed in RELION-3[41]. Particle picking was performed by manually choosing ~2,000 particles and generating templates through reference free 2D classification, followed by automatic template based picking. False-positive particles or particles classified in poorly defined classes were discarded after 2D classification. The initial 3D classification was performed on a binned dataset with the previously reported p97 structures as the reference model[26]. The detailed data processing flows are shown in Extended Data Figs. 2, 3, 4, 8, 9, and 10. To make sure that the 3D classification did not miss any major conformations, additional runs were performed with different number of classes (ranges from 3-8) and different regularization parameters (ranges from 2-6). Since we did not observe a dependence on the regularization parameter, the results from the default value (T=4) were shown in the figures.

Data processing statistics are summarized in **Extended Data Table 1**. Reported resolutions are based on Fourier shell correlation (FSC) using the FSC=0.143 criterion.

Model building, refinement, and validation

Model building was based on the existing cryo-EM structures of human p97[26] (PDB ID: 7JY5). The models of individual domains (N, D1, and D2) were first docked into the cryo-EM maps as rigid bodies using UCSF Chimera[42] followed by further adjustment using COOT[43]. For well-resolved regions, sharpened maps were used and individual residues were manually adjusted and fit into the density. For flexible regions, unsharpened maps were used and individual secondary structure elements or subdomains were fit into the density as rigid bodies. A short β strand was automatically built into the density corresponding to the translocating peptide using COOT, followed by manual extension on either ends and real-space refinement until the all the density was covered. The final models were subjected to global refinement and minimization in real space using the real-space refinement module in Phenix[44]. Model validation was performed using the comprehensive validation tool in Phenix. The statistics of model refinement is shown in **Extended Data Table 2**.

Substrate unfolding assay

Substrate unfolding assay was performed as previously described[26,31]. Briefly, 20 nM photoconverted polyubiquitinated Ub-Eos was mixed with 400 nM p97 or p97 mutants and 500 nM Npl4/Ufd1 in the Assay Buffer containing 50 mM Tris pH 7.4, 5 mM KCl, 20 mM $MgCl_2$, 1 mM EDTA, 0.5 mM TCEP, and 0.01% Triton. Proteins were pre-incubated in a 96-well plate (Costar 3694) for 10 minutes at 37 °C before initiating the reaction by supplementing the ATP regeneration mixture (5 mM ATP, 30 mM creatine phosphate, and 50 µg/mL creatine phosphokinase). Fluorescence signal was monitored using a TECAN safire2 plate reader at 540 nm excitation and 580 nm emission wavelengths with 30 seconds intervals for 60 min. Each

reaction condition was repeated three times. Background fluorescence was measured by mixing the same amount of substrate with 6 M guanidine-HCl and was subtracted from the average of the experimental groups. Normalized fluorescence was plotted using GraphPad Prism 8.4.2.

DATA AVAILABILITY

Cryo-EM maps have been deposited in the Electron Microscopy Data Bank (EMDB) under the accession codes EMDB-23449 (p97-Npl4/Ufd1-Ub_n-Eos-CHAPSO, Class 1), EMDB-23450 (p97-Npl4/Ufd1-Ub_n-Eos-CHAPSO, Class 2), EMDB-23451 (p97-Npl4/Ufd1-Ub_n-Eos-CHAPSO, Class 3), EMDB-23446 (p97-Npl4/Ufd1-Ub_n-Eos-FOM, Class 1), EMDB-23447 (p97-Npl4/Ufd1-Ub_n-Eos-FOM, Class 2), EMDB-23448 (p97-Npl4/Ufd1-Ub_n-Eos-FOM, Class 3), EMDB-23443 (p97-Npl4/Ufd1-Ub_6 -FOM, Class 1), EMDB-23444 (p97-Npl4/Ufd1-Ub_6-FOM, Class 2), EMDB-23445 (p97-Npl4/Ufd1-Ub_6-FOM, Class 3), EMDB-23442 (p97-Npl4/Ufd1-Ub_6-NMS-873-FOM), EMDB-23452 (p97-Npl4/Ufd1-Ub_n-Eos-NMS-873-CHAPSO), EMDB-23453 (p97-Npl4/Ufd1-Ub_n-Eos-NMS-873(substoichiometric)-CHAPSO, Class 1), EMDB-23454 (p97-Npl4/Ufd1-Ub_n-Eos-NMS-873(substoichiometric)-CHAPSO, Class 2), EMDB-23455 (p97-Npl4/Ufd1-Ub_n-Eos-NMS-873(substoichiometric)-CHAPSO, Class 3), EMDB-23456 (p97-Npl4/Ufd1-Ub_n-Eos-NMS-873(substoichiometric)-CHAPSO, Class 4), EMDB-23457 (p97-Npl4/Ufd1-Ub_n-Eos-NMS-873(substoichiometric)-CHAPSO, Class 5), EMDB-23458 (p97-Npl4/Ufd1-Ub_n-Eos-NMS-873(substoichiometric)-CHAPSO, Class 6). The atomic models have been deposited in the Protein Data Bank (PDB) under the accession codes 7LN5 (p97-Npl4/Ufd1-Ub_n-Eos-CHAPSO, Class 1), 7LN6 (p97-Npl4/Ufd1-Ub_n-Eos-CHAPSO, Class 2), 7LN2 (p97-Npl4/Ufd1-Ub_n-Eos-FOM, Class 1), 7LN3 (p97-Npl4/Ufd1-Ub_n-Eos-FOM, Class 2), 7LN4 (p97-Npl4/Ufd1-Ub_n-Eos-FOM, Class 3), 7LMZ (p97-Npl4/Ufd1-Ub_6 -FOM, Class 1), 7LN0 (p97-Npl4/Ufd1-Ub_6-FOM, Class 2), 7LN1 (p97-Npl4/Ufd1-Ub_6-FOM, Class 3), 7LMY (p97-Npl4/Ufd1-Ub_6-NMS-873-FOM).

ACKNOWLEGEMENT

Funding for this work was, in part, provided by the Catalyst Award from the Chicago Biomedical Consortium**.** This work was supported by Chicago Biomedical Consortium Catalyst Award C-086 to M.Z. We thank the National Key R&D Program of China (No.

2017YFA0505200), NSFC (91753205) for financial support. We thank staff at the National Cryo-Electron Microscopy Facility at the Frederick National Laboratory and the Advanced Electron Microscopy Facility at the University of Chicago for the help in cryo-EM data collection. This research was, in part, supported by the National Cancer Institute's National Cryo-EM Facility at the Frederick National Laboratory for Cancer Research under contract HSSN261200800001E.

AUTHOR CONTRIBUTIONS

M.P. and M.Z. designed all the experiments and interpreted the results. M.P., Y.Y., H.A., Q.Z., and Y.X. cloned, expressed, and purified all the complexes and carried out related biochemical characterizations. M.P., Y.Y., and M.Z. performed cryo-EM data collection and processing. M.Z., M.P., and L.L. wrote the paper. M.Z., L.L., and M.P. supervised the project.

COMPETING INTERESTS

The authors declare no competing interests.

FIGURE LEGENDS

FIGURE 1

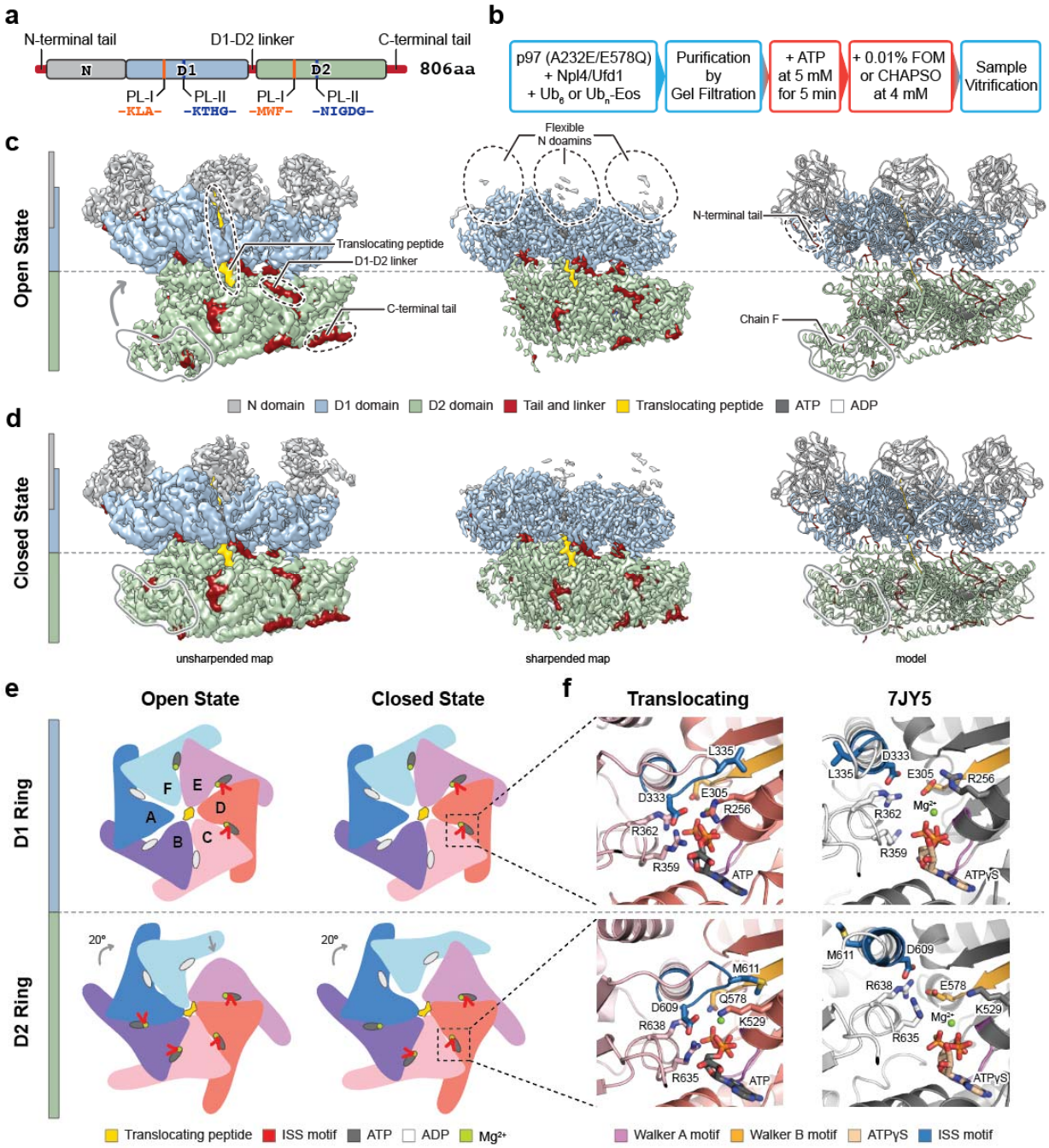


Figure 1: Structures of human p97 in working states.

a, Domain architecture of human p97. The color code of individual domains is used whenever possible. Pore loop residues in D1 and D2 domains are listed. **b,** Sample preparation procedures. Steps performed at 4 °C are in blue boxes. Steps performed at room temperature are in red boxes. **c and d,** Cryo-EM maps and models of human p97 engaged with a translocating peptide. The maps and models are aligned. Contour level: unsharpened map, 0.015; sharpened map, 0.035. **c,** Open state. **d,** Closed state. **e,** Illustrations showing the arrangement of the D1 and D2 rings in the open and closed states, respectively. **f,** Magnified views of compressed ATP binding sites in the D1 and D2 rings. The corresponding ATPγS binding sites from the nontranslocating structure (PDB ID: 7JY5) are shown for comparison.

FIGURE 2

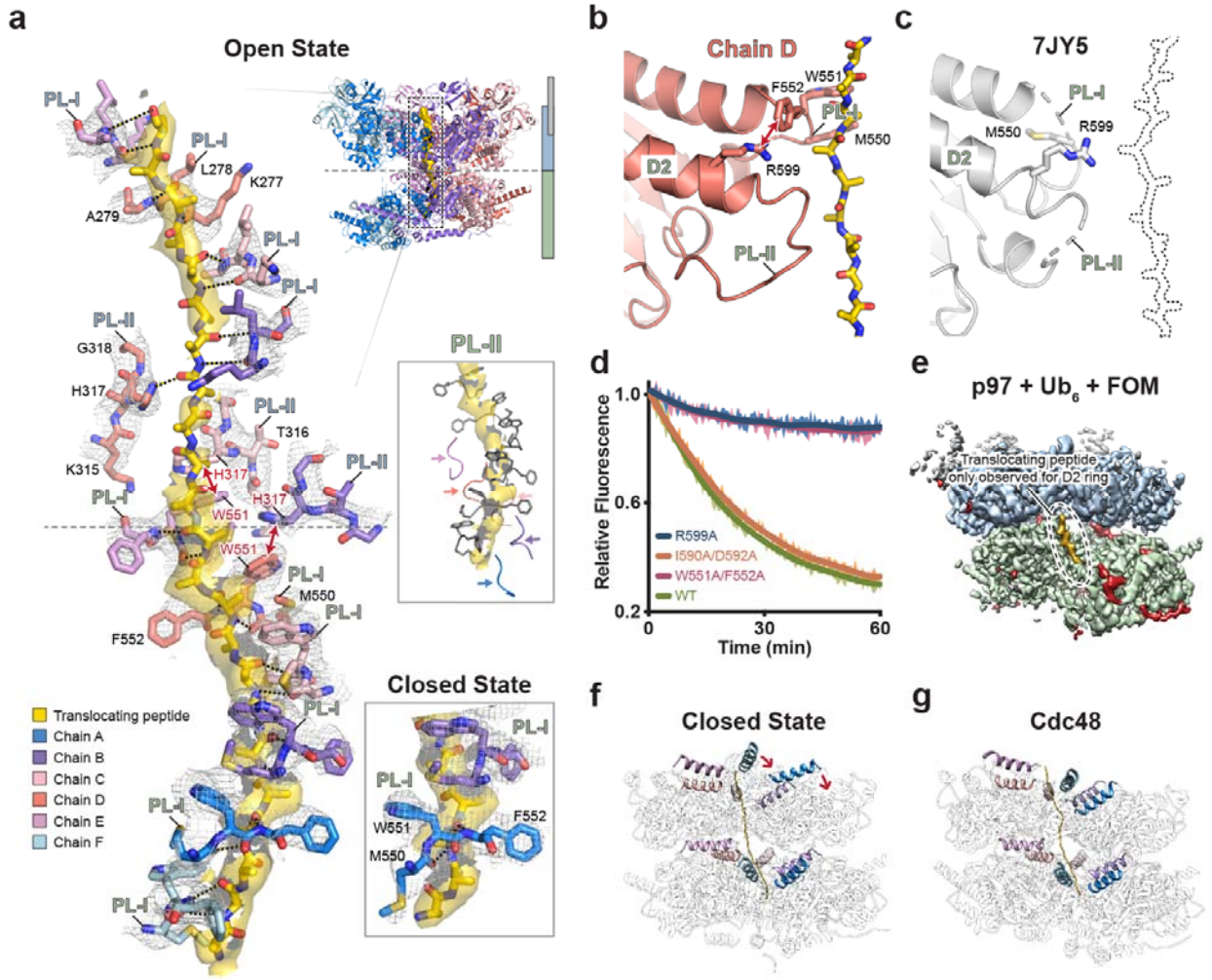


Figure 2: Interactions between the pore loops and the substrate.

a, Interactions between the translocating peptide and the interacting pore loops in the D1 and D2 rings. The density of the pore loops and the translocating peptide are shown as a gray mesh and yellow surface, respectively. Contour level: 6.0 root-mean-square deviation (rmsd). The closed state and PL-II of the D2 ring are shown in the insets. Hydrogen bonds between the pore loops and the peptide are labeled with black dotted lines. π-π interactions between H317 and W551 are labeled with red double arrows. **b and c,** A comparison of the D2 domain in chain D of the open state (panel **b**) and that of the nontranslocating structure (panel **c**, PDB ID: 7JY5). The D2 domains are aligned. **d,** Substrate unfolding assay for the pore loop mutants of p97. The

error bands represent the standard deviation from triplicate experiments. **e,** Unsharpened cryo-EM map (class 1) of the p97-Npl4/Ufd1-Ub_6-FOM dataset (**Extended Data Fig. 4**). The map is shown in a similar orientation to **Fig. 1c and 1d**. Contour level: 0.015. **f and g**, A comparison of cryo-EM models of translocating structures of p97 (closed state, panel **f**) and Cdc48 (PDB ID: 6OPC, panel **g**). The two models are aligned based on the D2 ring. The □2 helices immediately before PL-I in D1 and D2 domains are colored. The color code is the same as that in panel **a**.

FIGURE 3

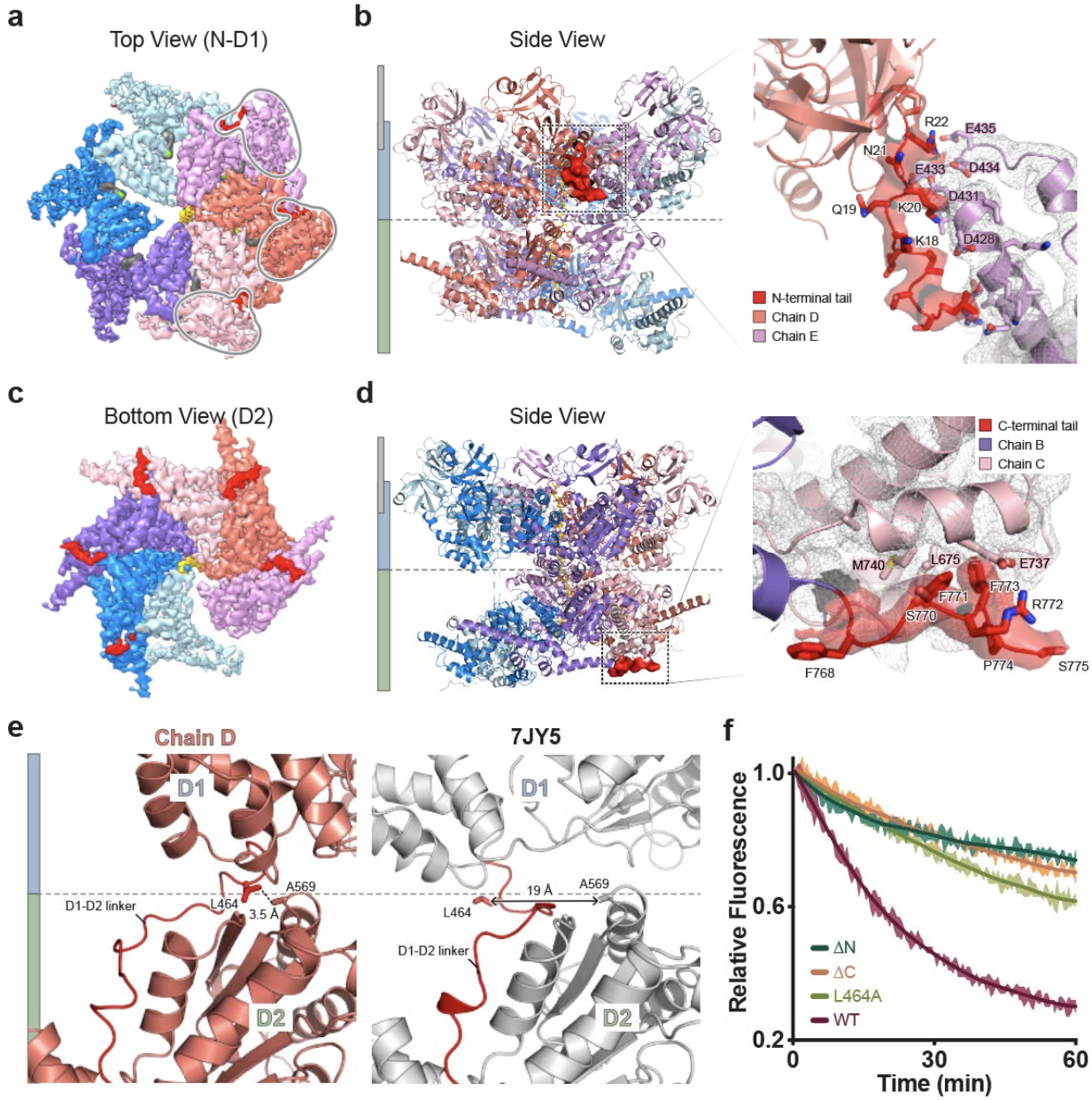


Figure 3: N- and C-terminal tails and the linker between the D1 and D2 rings.

a, Top view of the unsharpened cryo-EM map focused on the N and D1 domains (open state). N-terminal tails are colored in red. Contour level: 0.015. **b,** A side view of the cryo-EM model of the open state with the density of one N-terminal tail highlighted and magnified. Contour level:

6.0 rmsd. **c,** A bottom view of the unsharpened cryo-EM map focused on the D2 ring (open state). C-terminal tails are shown in red. Contour level: 0.015. **d,** A side view of the cryo-EM model of the open state with the density of one C-terminal tail highlighted and magnified. Contour level: 6.0 rmsd. **e,** A comparison of the D1-D2 linker in chain D of the open state and that of the nontranslocating structure (PDB ID: 7JY5). The D2 domains are aligned. **f,** Substrate unfolding assay for tail and linker mutants of p97. The error bands represent the standard deviation from triplicate experiments.

FIGURE 4

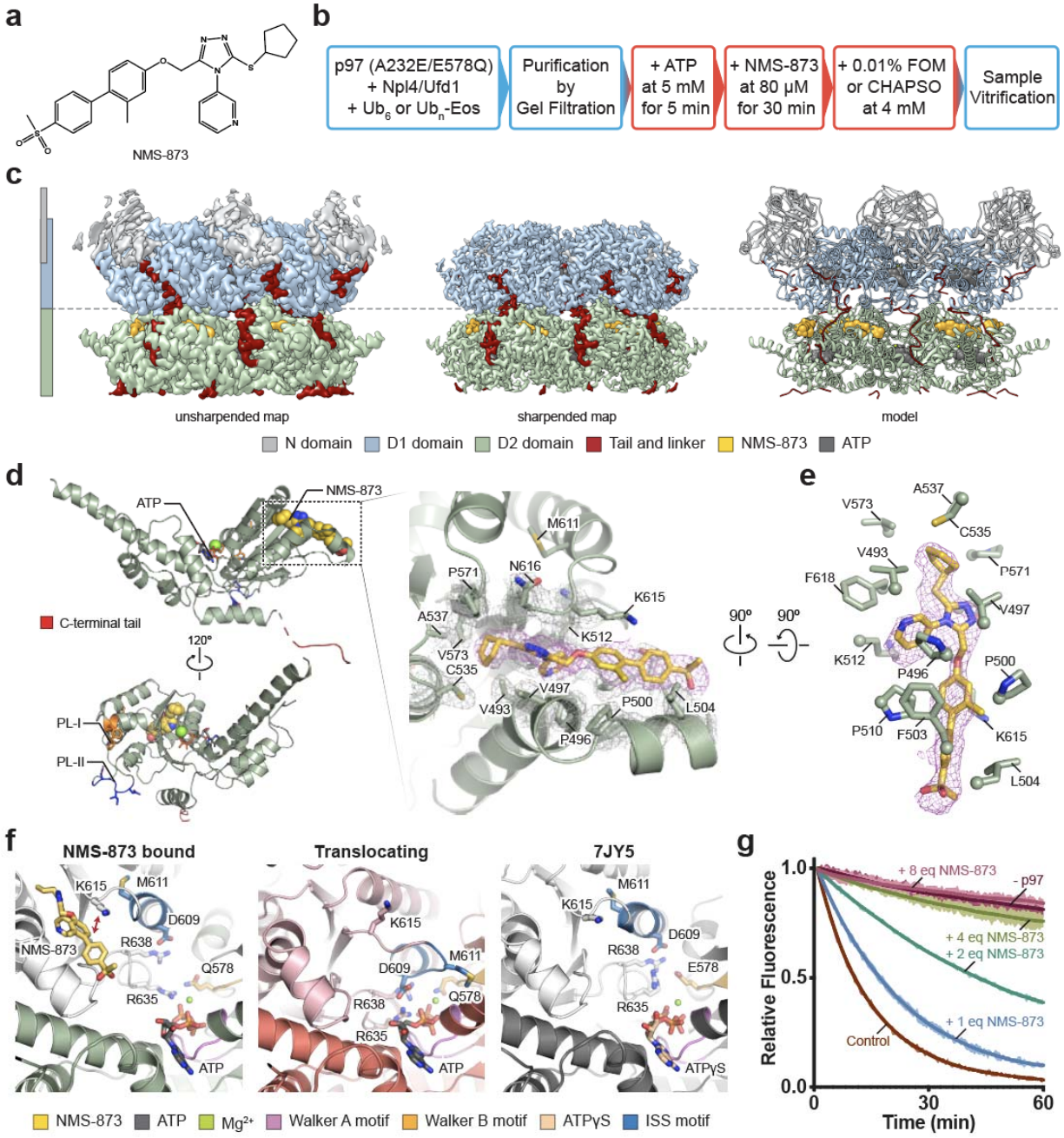


Figure 4: NMS-873 inhibits translocation by locking the ISS motif.

a, Chemical structure of NMS-873. **b,** Sample preparation procedures. Steps performed at 4 °C are in blue boxes. Steps performed at room temperature are in red boxes. **c,** Cryo-EM maps and models of human p97 in complex with NMS-873. The maps and models are aligned.

Contour level: unsharpened map, 0.018; sharpened map, 0.052. **d,** D2 domain of the complex structure in panel **c**, with the NMS-873 binding site zoomed in. The density of NMS-873 and the interacting residues are shown as magenta and gray mesh, respectively. Contour level: 6.0 rmsd. **e,** A different view of the NMS-873 binding site, with side chains of interacting residues shown. Contour level: 6.0 rmsd. **f,** A comparison of nucleotide binding sites in NMS-873 bound, translocating (chain D in open state), and nontranslocating (PDB ID: 7JY5) structures. The bottom D2 domains (green, salmon, and gray) are aligned. **g,** Substrate unfolding assay for wild-type p97 in the presence of different concentrations of NMS-873. One p97 hexamer can bind 6 equivalents (eq) of NMS-873.

FIGURE 5

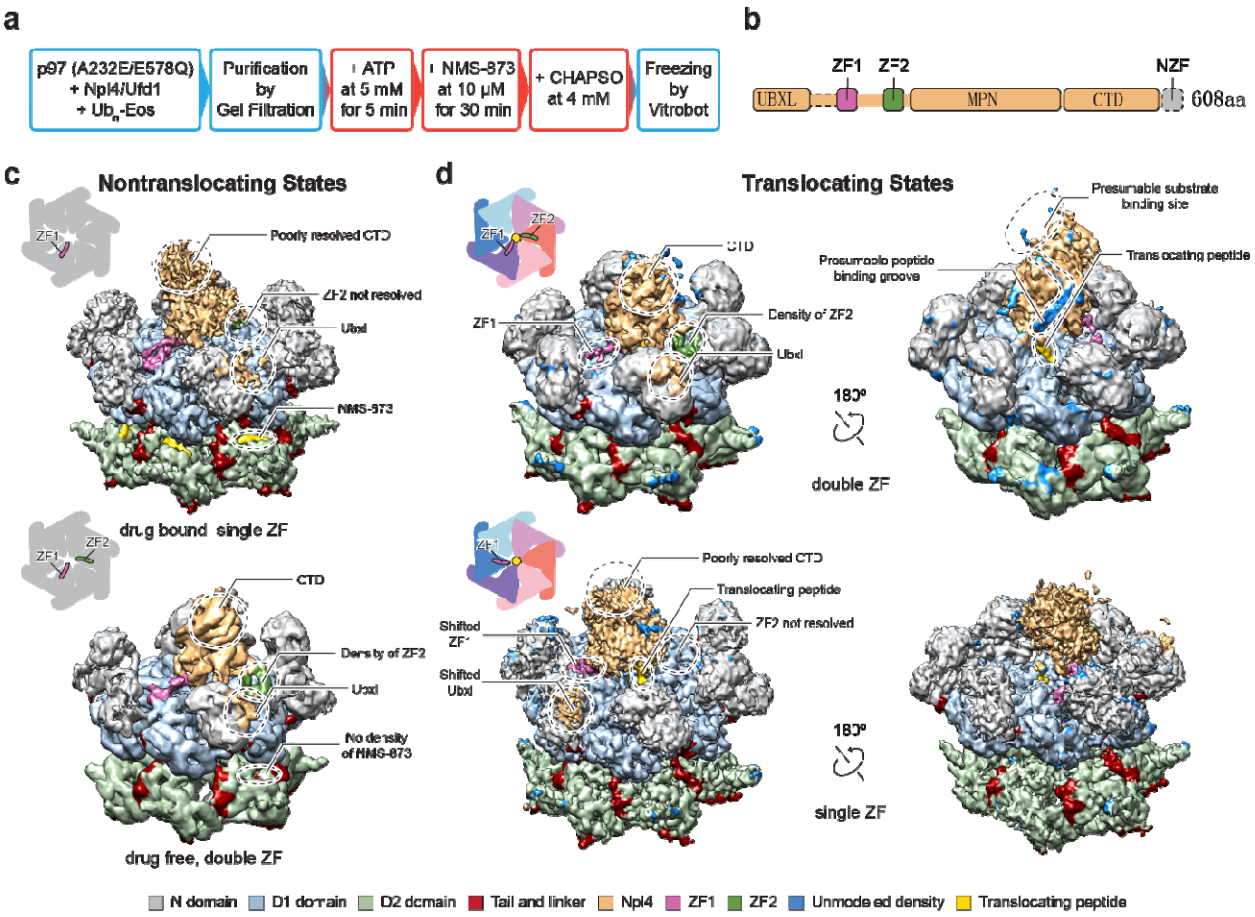
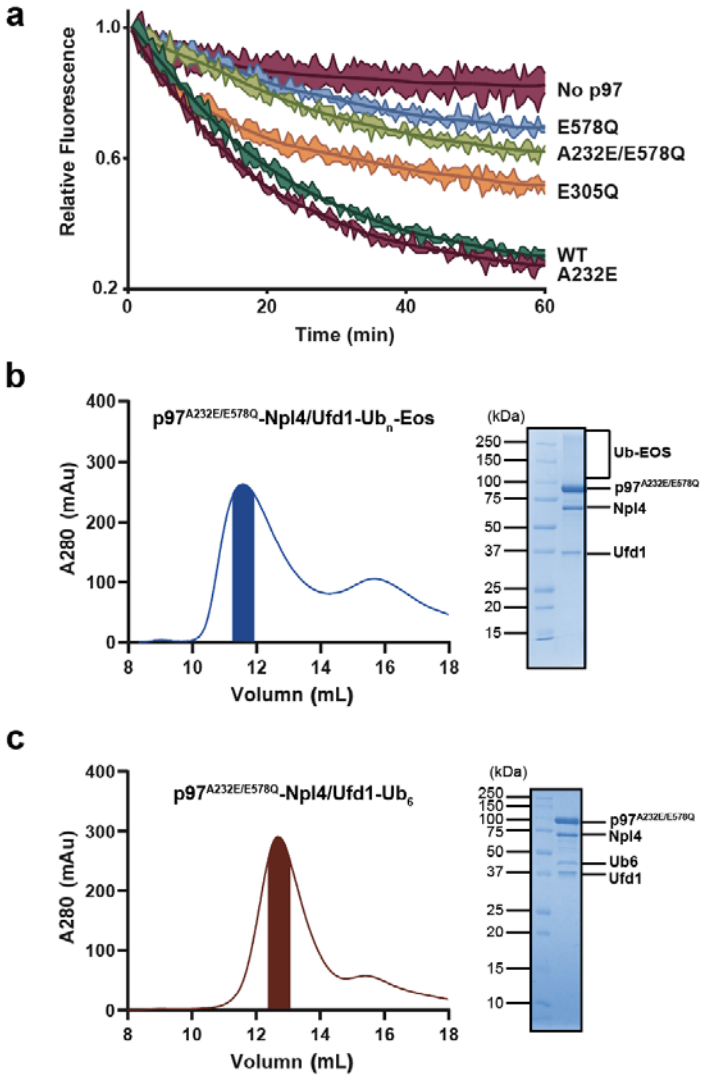


Figure 5: Conformational states of Npl4 in the presence of NMS873 at a substoichiometric concentration.

a, Sample preparation procedures. Steps performed at 4 °C are in blue boxes. Steps performed at room temperature are in red boxes. **b,** Domain architecture of human Npl4. The color code of the individual motifs are the same as that in panels **c** and **d**. Dotted parts are not resolved in the cryo-EM maps. **c,** Two cryo-EM maps in nontranslocating states resolved from the dataset (**Extended Data Fig. 10**), one bound with NMS-873 (top). **d,** Two cryo-EM maps in translocating states resolved from the dataset. Different positions of zinc finger motifs of Npl4 are illustrated in the insets.

EXTENDED DATA FIGURES

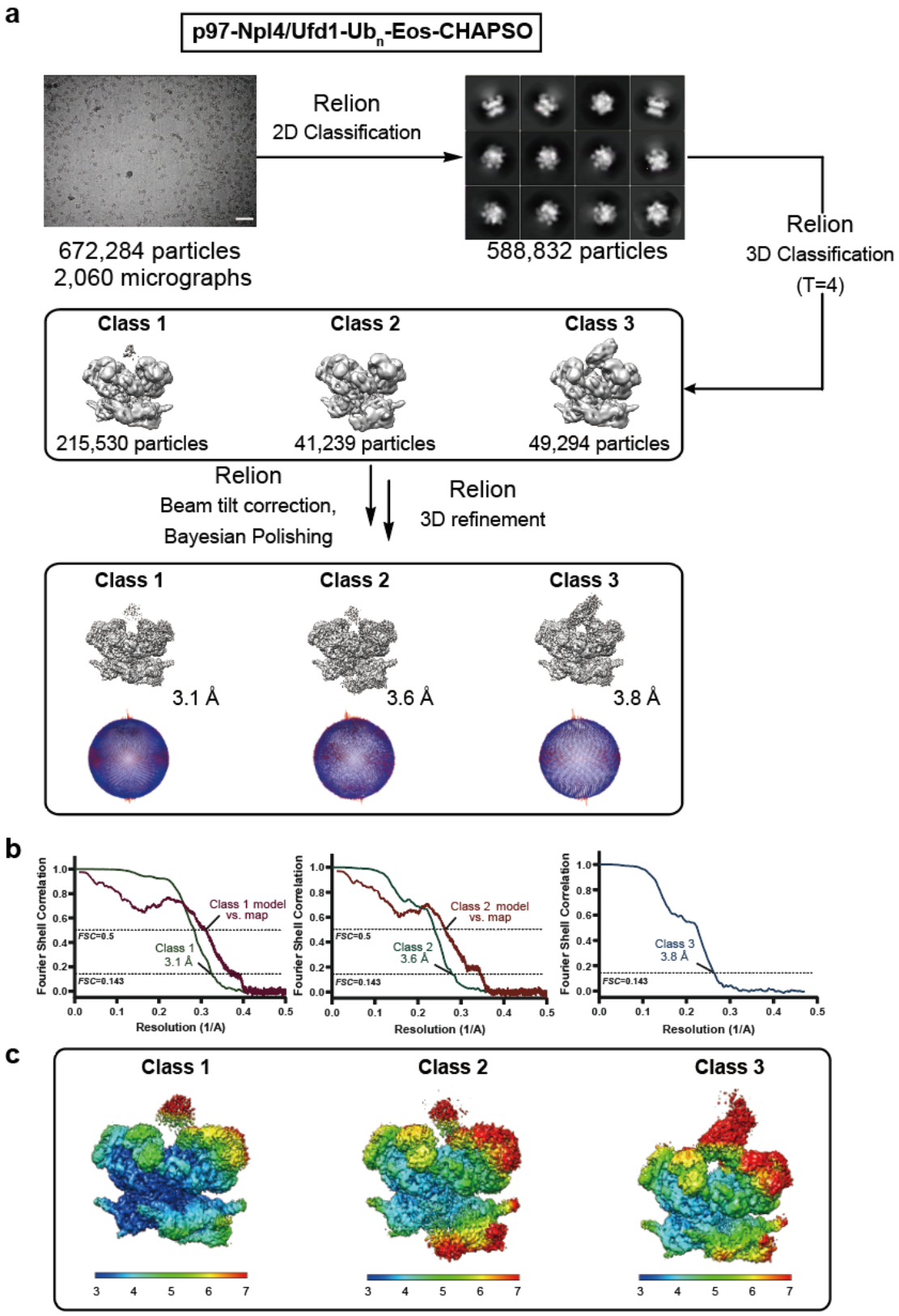
EXTENDED DATA FIGURE 1



Extended Data Figure 1: Substrate unfolding activity of various p97 mutants and the purification of cofactor- and substrate-engaged p97 complexes.

a, Substrate unfolding assay for wild-type (WT) p97 and various mutations. E305Q and E578Q are the two Walker B mutations in the D1 and D2 domains, respectively. A232E is a disease mutation found in multisystem proteinopathy (MSP)[27,31]. **b,** Gel filtration chromatogram and SDS-PAGE gel of the p97$^{A232E/E578Q}$-Npl4/Ufd1-Ub$_n$-Eos complex. **c,** Gel filtration chromatogram and SDS-PAGE gel of the p97$^{A232E/E578Q}$-Npl4/Ufd1-Ub$_n$-Eos complex.

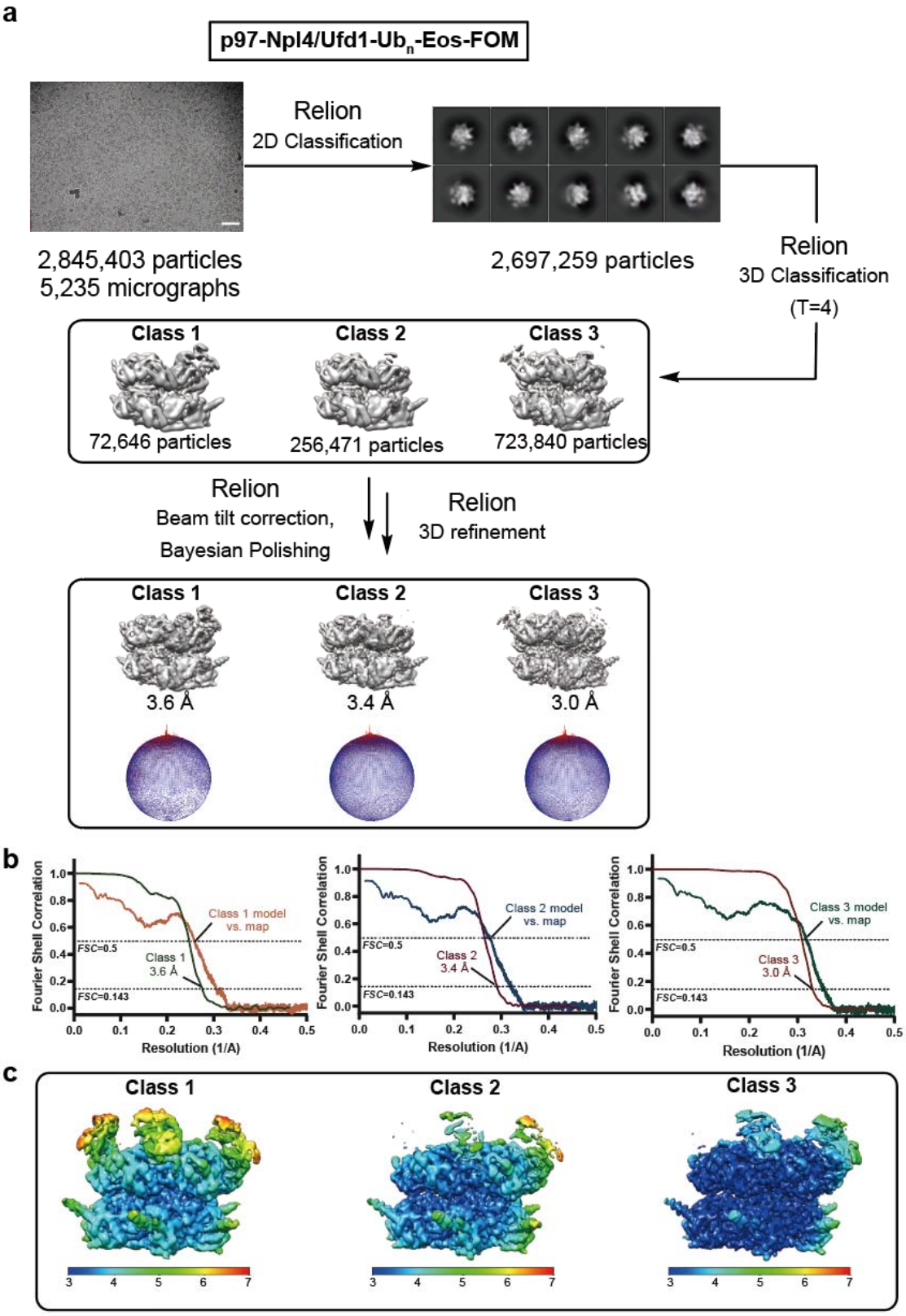
EXTENDED DATA FIGURE 2



Extended Data Figure 2: Single-particle cryo-EM analyses for the p97-Npl4/Ufd1-Ub_n-Eos-CHAPSO dataset.

a, The workflow of data processing. The dataset was subjected to particle selection, 2D classification, and multiple rounds of 3D classification. A representative micrograph (scale bar corresponds to 50 nm) and representative 2D class averages are shown. Three classes were resolved from the dataset, including class 1 (closed state), class 2 (open state), and class 3 (similar to the closed state but with the Npl4 density). The distributions of the Euler angles for each reconstruction are shown below the maps. **b,** Fourier shell correlation (FSC) curves of the masked maps after Relion postprocessing. The resolutions were determined by the FSC=0.143 criterion. The model vs. map FSC curves are also shown for class 1 and class 2 (red). **c,** Local resolutions of the maps calculated using Relion.

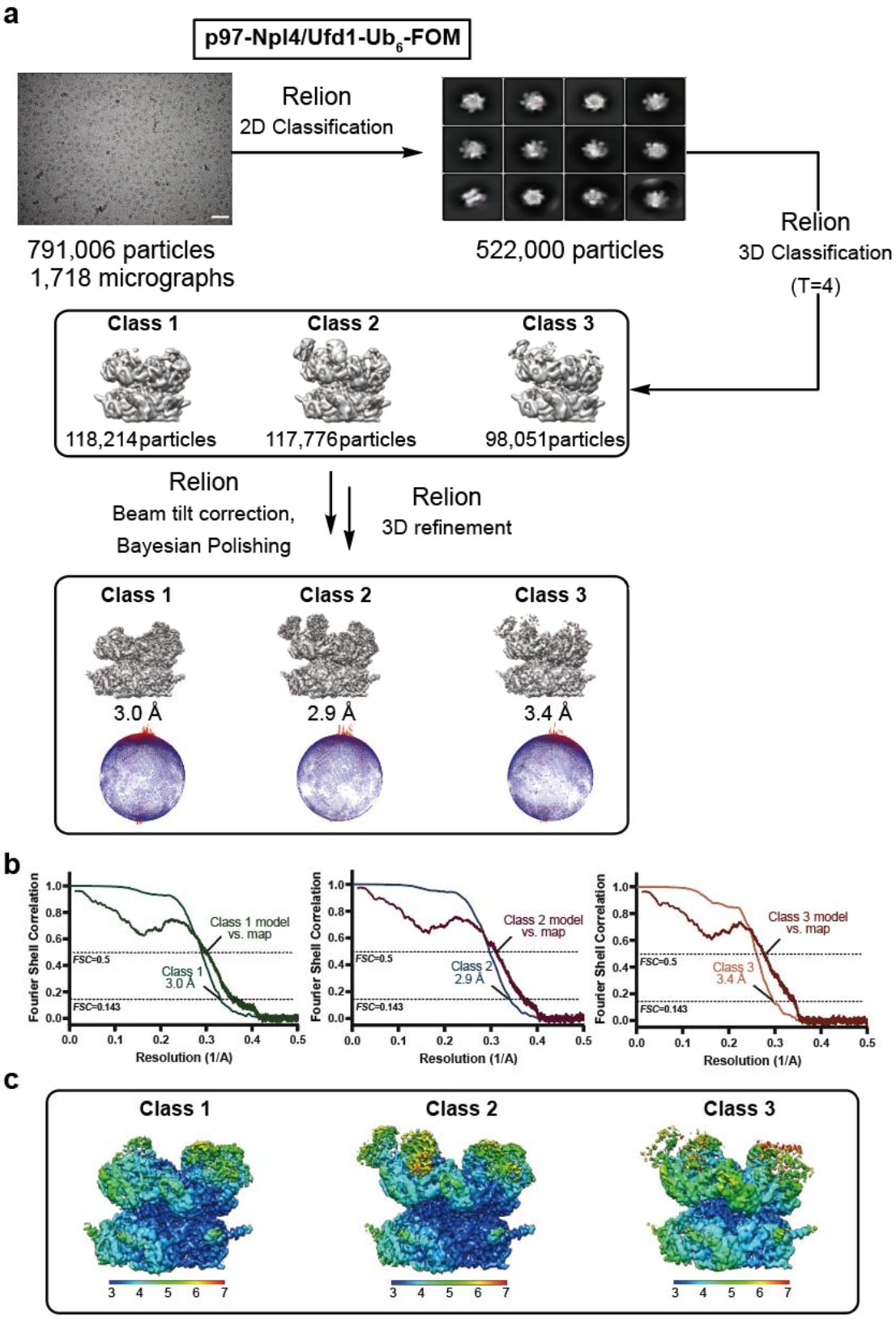
EXTENDED DATA FIGURE 3



Extended Data Figure 3: Single-particle cryo-EM analyses for the p97-Npl4/Ufd1-Ub_n-Eos-FOM dataset.

a, The workflow of data processing. The dataset was subjected to particle selection, 2D classification, and multiple rounds of 3D classification. A representative micrograph (scale bar corresponds to 50 nm) and representative 2D class averages are shown. Three classes were resolved from the dataset. The distributions of the Euler angles for each reconstruction are shown below the maps. **b,** FSC curves of the masked maps after Relion postprocessing. The resolutions were determined by the FSC=0.143 criterion. The model vs. map FSC curves for each class are also shown. **c,** Local resolutions of the maps calculated using Relion.

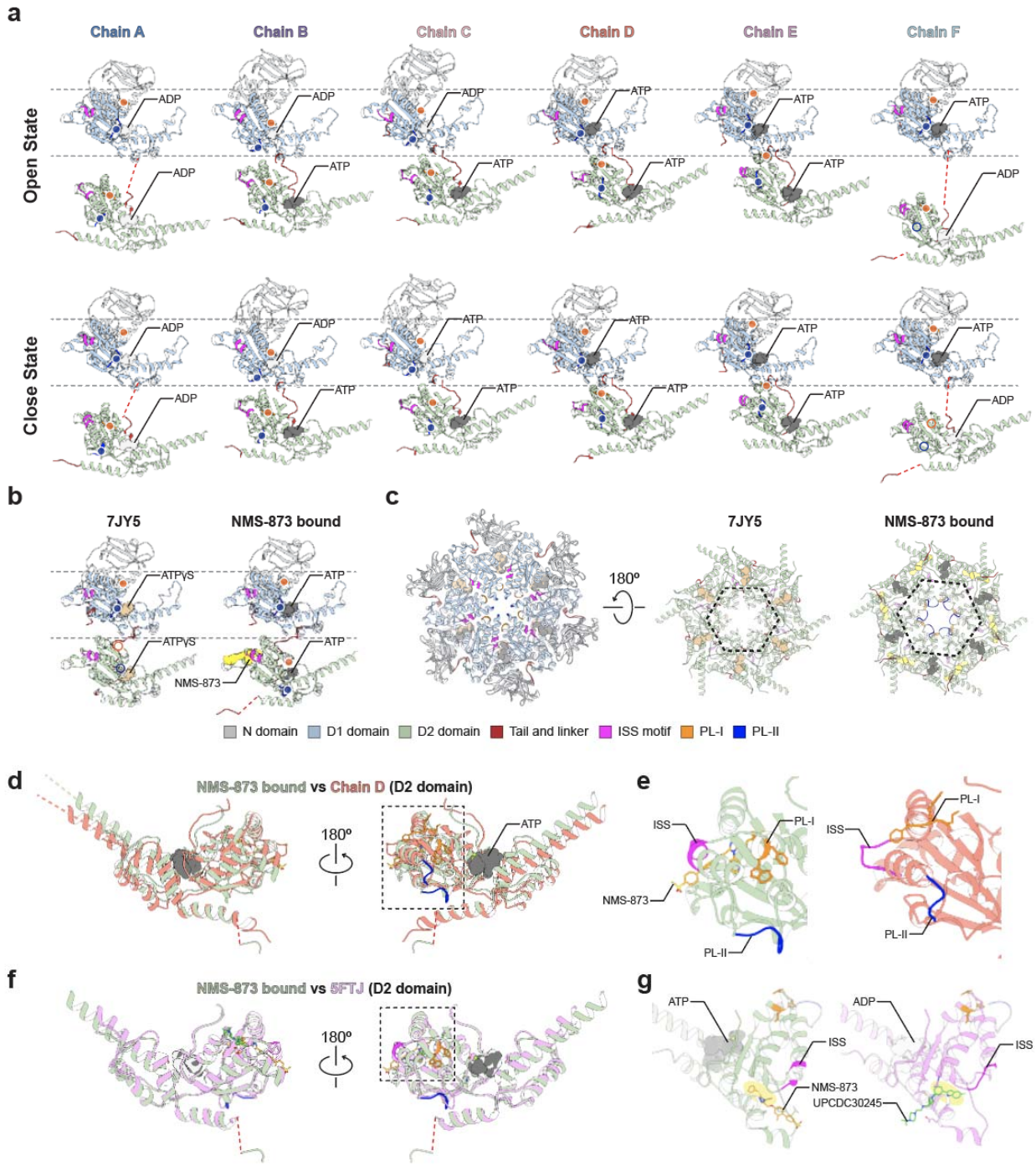
EXTENDED DATA FIGURE 4



Extended Data Figure 4: Single-particle cryo-EM analyses for the p97-Npl4/Ufd1-Ub_6-FOM dataset.

a, The workflow of data processing. The dataset was subjected to particle selection, 2D classification, and multiple rounds of 3D classification. A representative micrograph (scale bar corresponds to 50 nm) and representative 2D class averages are shown. Three classes were resolved from the dataset. The distributions of the Euler angles for each reconstruction are shown below the maps. **b,** FSC curves of the masked maps after Relion postprocessing. The resolutions were determined by the FSC=0.143 criterion. The model vs. map FSC curves for each class are also shown. **c,** Local resolutions of the maps calculated using Relion.

EXTENDED DATA FIGURE 5

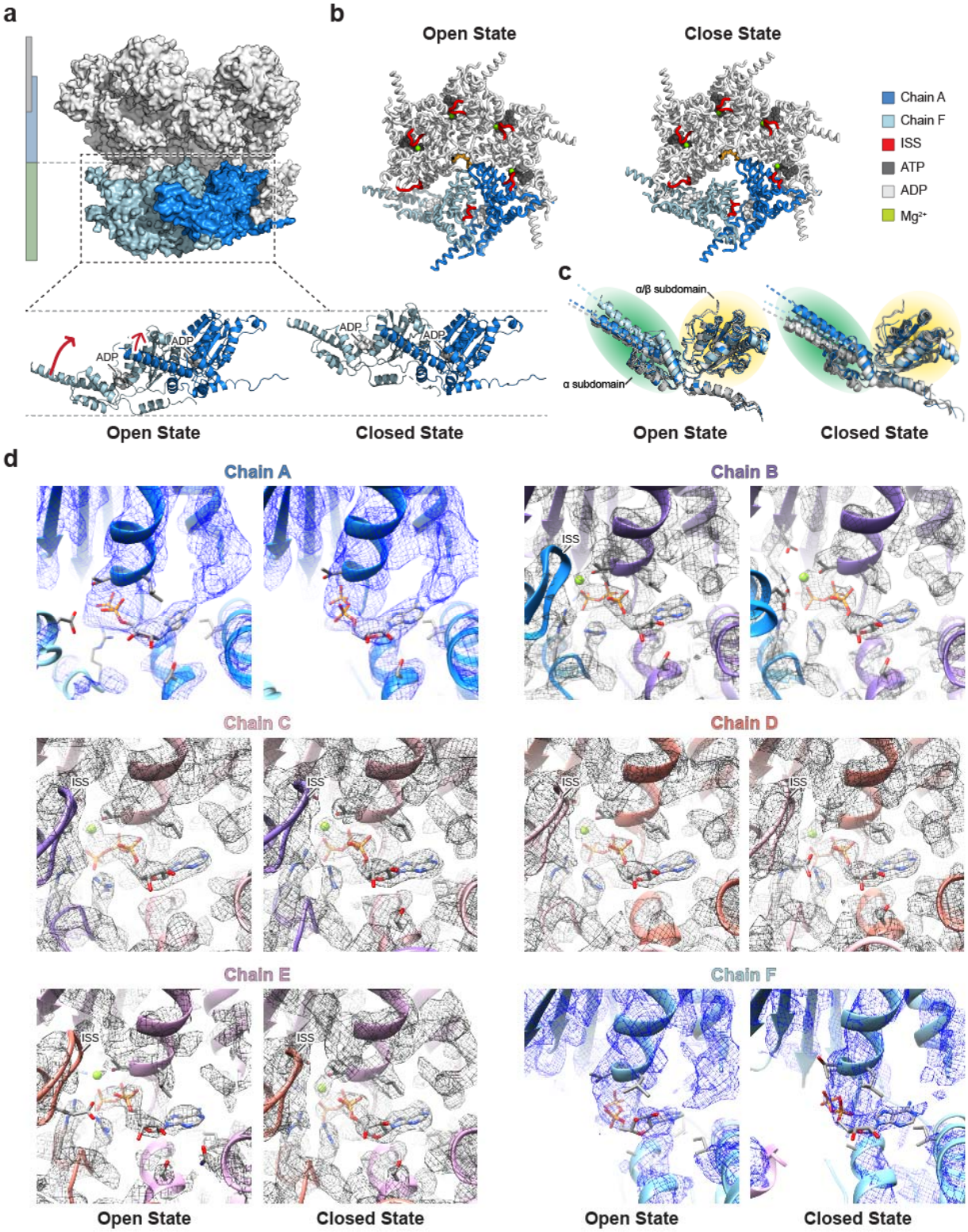


Extended Data Figure 5: Comparison of various p97 structures.

a, Unrolling of the open (top row) and closed (bottom row) states of p97. The lateral and vertical movement of D2 domains can be visualized. The positions of pore loops are marked with solid

dots. Unresolved pore loops are marked by open circles. **b,** Unrolling of the ATPγS-bound nontranslocating structure (PDB ID: 7JY5) and NMS-873-bound structure. Only one chain of each structure is shown since both structures are sixfold symmetric. **c,** A superimposition of 7JY5- and NMS-873-bound structures. Left: a top view of the N and D1 domains; Right: a bottom view of the D2 ring. The opening of the D2 ring is marked by a dotted hexagon. Panels **a, b,** and **c** share the same color code. **d,** A comparison of the D2 domain in the NMS-873-bound structure and that in chain D of the open state. The superimposition is based on the α/β subdomain. **e,** Magnified views of the dotted box in panel **d**, showing the conformations of the PL-I, PL-II, and ISS motifs. **f,** A comparison of the D2 domain in NMS-873- and UPCDC30245 (PDB ID: 5FTJ)-bound structures. The superimposition is based on the α/β subdomain. **g,** Magnified views of the dotted box in panel **f**, showing the binding sites of the compounds and the conformations of the ISS motif.

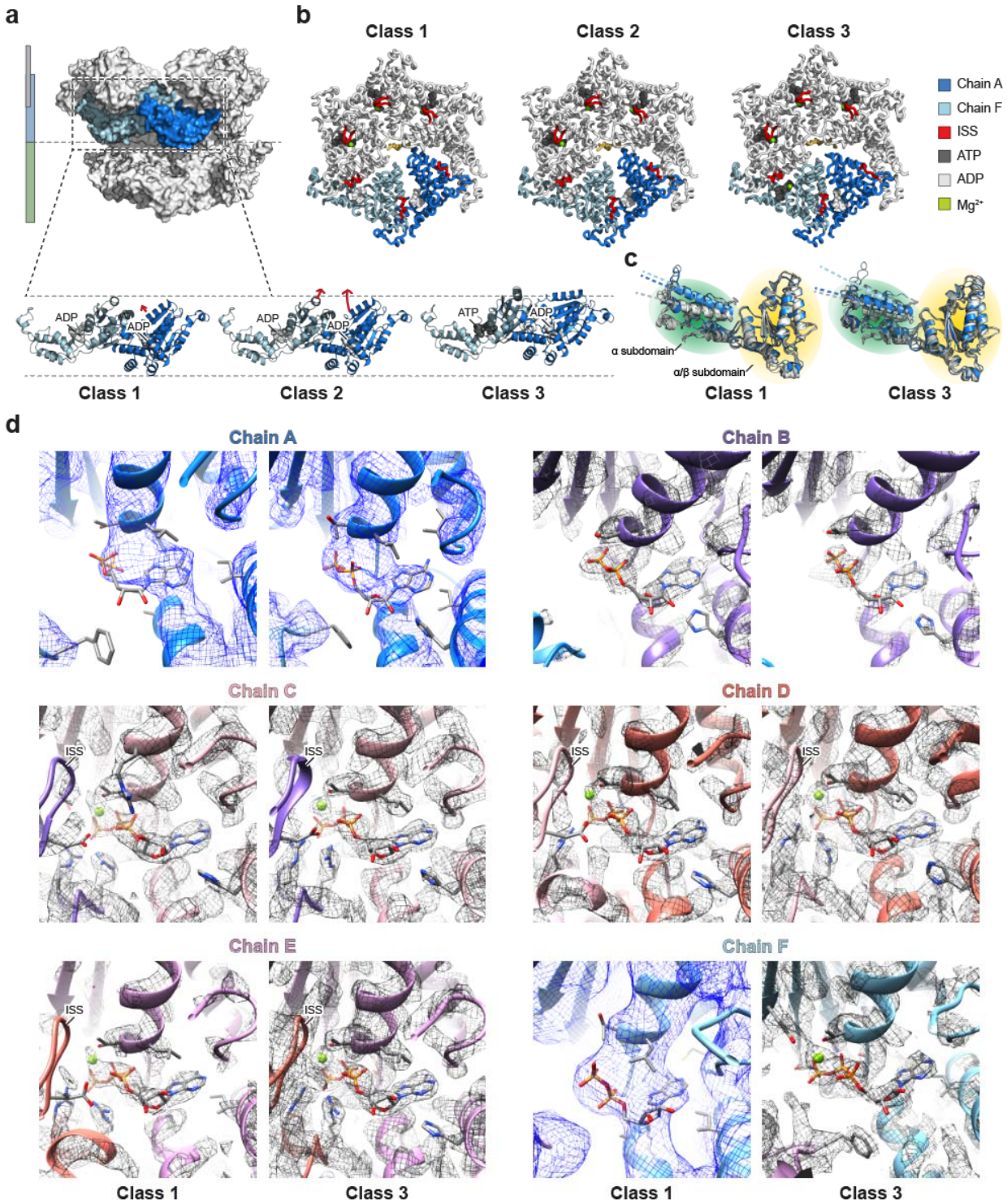
EXTENDED DATA FIGURE 6



Extended Data Figure 6: The power stroke motion and the nucleotide binding sites of the D2 domains.

a, A superimposition of the open and closed state structures, with two D2 domains (chain A and chain F) magnified. The relative positions of the D2 domains in chain A and chain F can be visualized. **b,** Top views of the D2 ring in the open and closed states, highlighting chain A, chain F, and the ISS motifs. **c,** Superimpositions of all D2 domains in the open and closed states based on the α/β subdomain. The directions of α7 helices are highlighted by dotted lines for comparison (dark blue, chain A; light blue, chain F; and gray, other chains). **d,** Individual nucleotide binding sites of the D2 domains in the open and closed states. Blue mesh: unsharpened map at contour level; Gray mesh: sharpened map at contour level.;

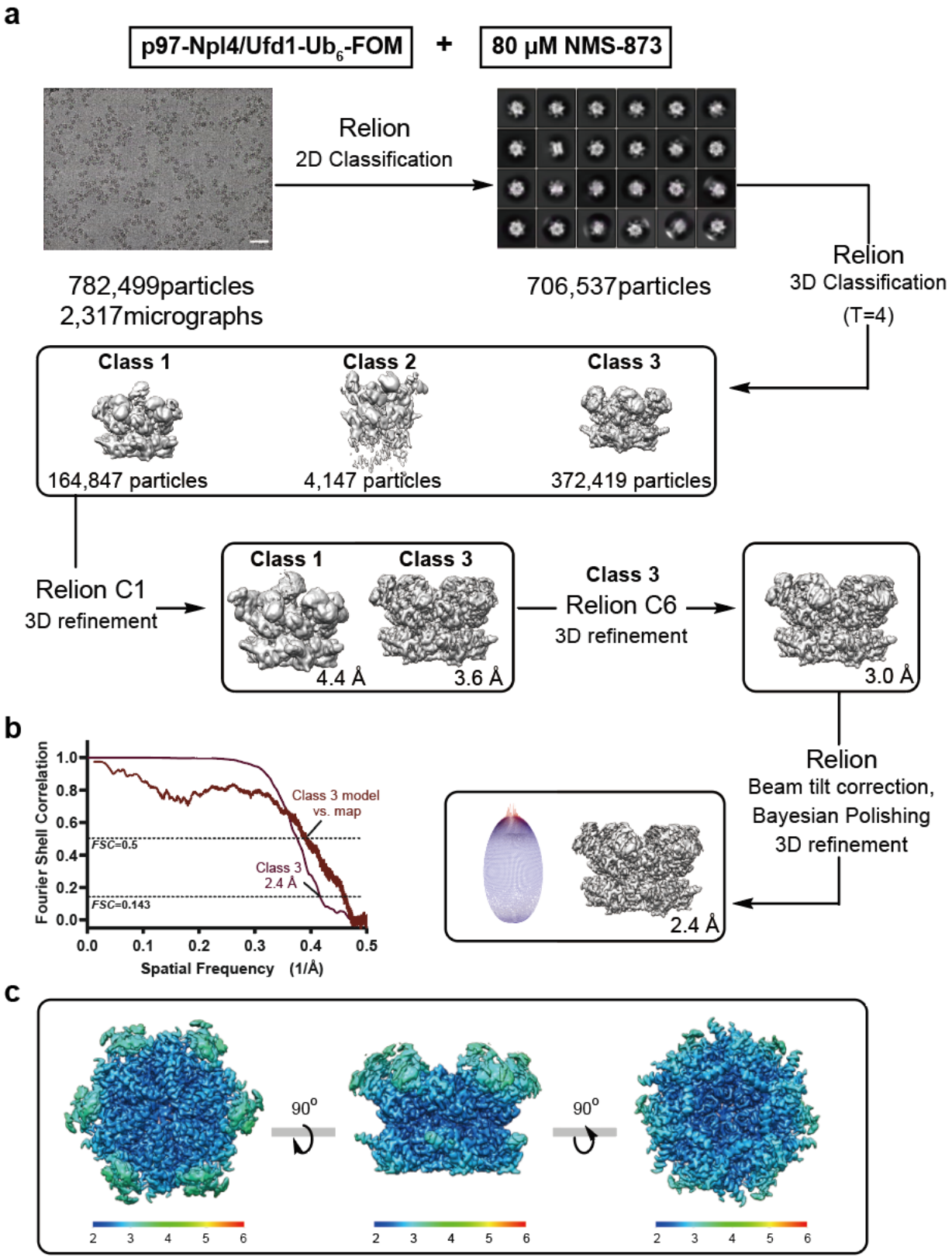
EXTENDED DATA FIGURE 7



Extended Data Figure 7: The power stroke motion and the nucleotide binding sites of the D1 domains.

a, A superimposition of three structures (classes 1, 2, and 3) from the p97-Npl4/Ufd1-Ub_n-Eos-FOM dataset, with two D1 domains (chain A and chain F) magnified. The relative positions of the D1 domains in chain A and chain F can be visualized. **b,** Top views of the D1 ring in the three classes, highlighting chain A, chain F, and the ISS motifs. **c,** Superimpositions of all D1 domains in class 1 and class 3 based on the α/β subdomain. The directions of α7 helices are highlighted by dotted lines for comparison (dark blue, chain A; light blue, chain F; and gray, other chains). **d,** Individual nucleotide binding sites of the D1 domains in class 1 and class 3. Blue mesh: Unsharpened map at contour level; Gray mesh: sharpened map at contour level.;

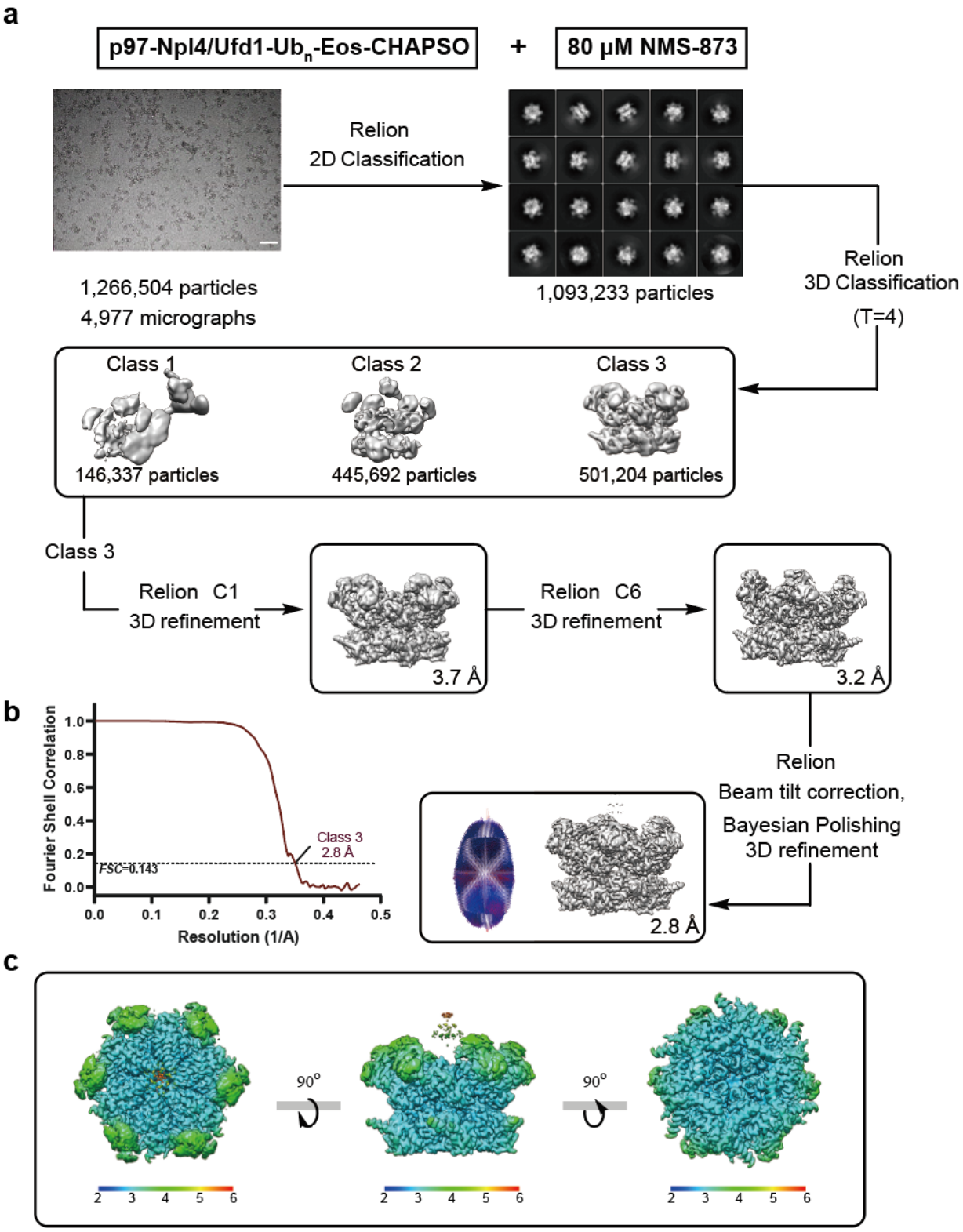
EXTENDED DATA FIGURE 8



Extended Data Figure 8: Single-particle cryo-EM analyses for the p97-Npl4/Ufd1-Ub_6 complex in the presence of 80 μM NMS-873.

a, The workflow of data processing. The dataset was subjected to particle selection, 2D classification, and multiple rounds of 3D classification. A representative micrograph (scale bar corresponds to 50 nm) and representative 2D class averages are shown. A single class was resolved from the dataset. The distribution of the Euler angles is shown next to the map. **b,** The FSC curve of the masked map after Relion postprocessing. The resolution was determined by the FSC=0.143 criterion. The model vs. map FSC curve is also shown. **c,** Local resolution of the map calculated using Relion.

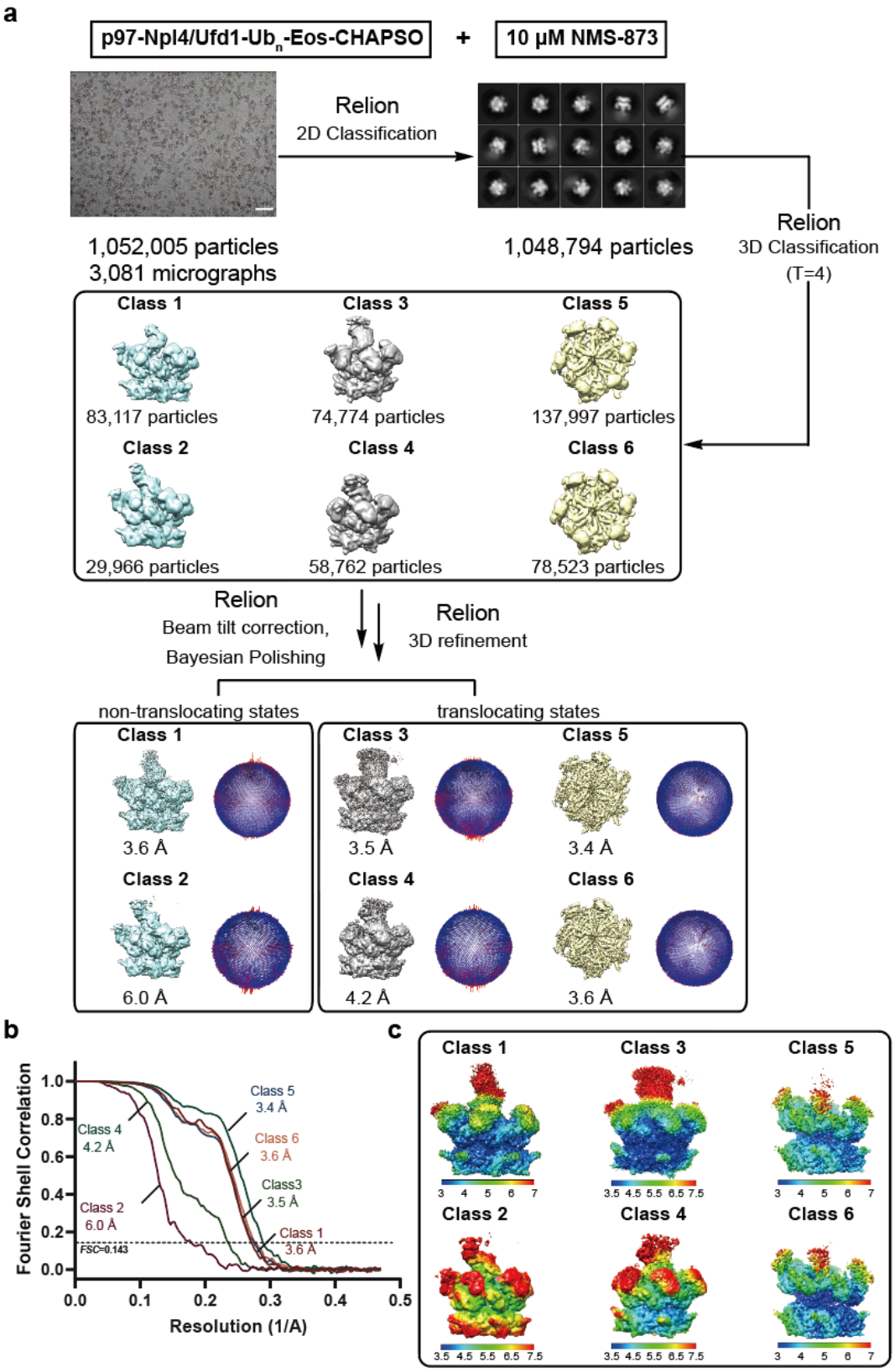
EXTENDED DATA FIGURE 9



Extended Data Figure 9: Single-particle cryo-EM analyses for the p97-Npl4/Ufd1-Ub_n-Eos complex in the presence of 80 μM NMS-873.

a, The workflow of data processing. The dataset was subjected to particle selection, 2D classification, and multiple rounds of 3D classification. A representative micrograph (scale bar corresponds to 50 nm) and representative 2D class averages are shown. A single class was resolved from the dataset. The distribution of the Euler angles is shown next to the map. **b,** The FSC curve of the masked map after Relion postprocessing. The resolution was determined by the FSC=0.143 criterion. **c,** Local resolution of the map calculated using Relion.

EXTENDED DATA FIGURE 10:



Extended Data Figure 10: Single-particle cryo-EM analyses for the p97-Npl4/Ufd1-Ub_n-Eos complex in the presence of 10 μM NMS-873.

a, The workflow of data processing. The dataset was subjected to particle selection, 2D classification, and multiple rounds of 3D classification. A representative micrograph (scale bar corresponds to 50 nm) and representative 2D class averages are shown. Six classes were resolved from the dataset, including two in nontranslocating states (Class 1 and Class 2) and four in translocating states (Class 3 through Class 6). The distributions of the Euler angles for each reconstruction are shown next to the maps. **b,** FSC curves of the masked maps after Relion postprocessing. The resolutions were determined by the FSC=0.143 criterion. **c,** Local resolutions of the maps calculated using Relion.

EXTENDED DATA TABLE 1: Statistics of cryo-EM data collection and processing

	p97-Npl4/Ufd1-Ub$_n$-Eos			**p97-Npl4/Ufd1-Ub$_n$-Eos**		
Detergent	CHAPSO (4 mM)			FOM (0.001%)		
Microscope	Krios (UChicago)			Krios (NCI)		
Magnification	81,000			81,000		
Voltage (kV)	300			300		
Spherical aberration (mm)	2.7			2.7		
Detector	K3			K3		
Camera mode	Super resolution counting			Super resolution counting		
Exposure rate (e^-/pixel/s)	15			15		
Total exposure (e^-/Å^2)	50			50		
Defocus range (μm)	-1.0 to -2.5			-1.0 to -2.5		
Pixel size (Å)	1.063			1.0794		
Mode of data collection	Image shift			Image shift		
Energy filter	20 eV slit			20 eV slit		
Software for data collection	EPU			Latitude S		
Number of micrographs	2,060			5,235		
Symmetry imposed	C1			C1		
Box size (pixel)	320			320		
Initial particle images (no.)	672,284			2,845,403		
Particle images for 3D (no.)	588,832			2,697,259		
	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
Final particle images (no.)	215,530	41,239	49,294	72,646	256,471	723,840
Map resolution, masked (Å)	3.09	3.58	3.77	3.63	3.45	3.02
B-factor estimated (Å^2)	79.7	79.7	80.3	97.6	109	96.8
EMD accession code	23449	23450	23451	23446	23447	23448

	p97-Npl4/Ufd1-Ub$_6$		
Detergent	FOM (0.001%)		
Microscope	Krios (UChicago)		
Magnification	81,000		
Voltage (kV)	300		
Spherical aberration (mm)	2.7		
Detector	K3		
Camera mode	Super resolution counting		
Exposure rate (e^-/pixel/s)	15		
Total exposure (e^-/Å^2)	50		
Defocus range (μm)	-1.0 to -2.5		
Pixel size (Å)	1.063		
Mode of data collection	Image shift		
Energy filter	20 eV slit		
Software for data collection	EPU		
Number of micrographs	1,718		
Symmetry imposed	C1		
Box size (pixel)	320		
Initial particle images (no.)	791,006		
Particle images for 3D (no.)	522,000		
	Class 1	Class 2	Class 3
Final particle images (no.)	118,214	117,776	98,051
Map resolution, masked (Å)	3.06	2.98	3.40
B-factor estimated (Å^2)	69.1	64.7	77.3
EMD accession code	23443	23444	23445

	p97-Npl4/Ufd1-Ub_6-NMS-873 (80 µM)	**p97-Npl4/Ufd1-Ub_n-Eos-NMS-873 (80 µM)**
Detergent	FOM (0.001%)	CHAPSO (4 mM)
Microscope	Krios (UChicago)	Krios (NCI)
Magnification	81,000	81,000
Voltage (kV)	300	300
Spherical aberration (mm)	2.7	2.7
Detector	K3	K3
Camera mode	Super resolution counting	Super resolution counting
Exposure rate (e^-/pixel/s)	15	15
Total exposure (e^-/Å^2)	50	50
Defocus range (µm)	-1.0 to -2.5	-1.0 to -2.5
Pixel size (Å)	1.063	1.0794
Mode of data collection	Image shift	Image shift
Energy filter	20 eV slit	20 eV slit
Software for data collection	EPU	Latitude S
Number of micrographs	2,317	4,977
Symmetry imposed	C6	C6
Box size (pixel)	320	320
Initial particle images (no.)	782,499	1,266,504
Particle images for 3D (no.)	706,537	1,093,233
Final particle images (no.)	372,419	501,204
Map resolution, masked (Å)	2.41	2.87
B-factor estimated (Å^2)	58.8	96.2
EMD accession code	23442	23452

	p97-Npl4/Ufd1-Ub$_n$-Eos-NMS-873 (10 μM)					
Detergent	CHAPSO (4 mM)					
Microscope	Krios (UChicago)					
Magnification	81,000					
Voltage (kV)	300					
Spherical aberration (mm)	2.7					
Detector	K3					
Camera mode	Super resolution counting					
Exposure rate (e^-/pixel/s)	15					
Total exposure (e^-/Å^2)	50					
Defocus range (μm)	-1.0 to -2.5					
Pixel size (Å)	1.063					
Mode of data collection	Image shift					
Energy filter	20 eV slit					
Software for data collection	Latitude S					
Number of micrographs	3081					
Symmetry imposed	C1					
Box size (pixel)	320					
Initial particle images (no.)	1,052,005					
Particle images for 3D (no.)	1,048,794					
	Non-Translocating		Translocating			
	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
Final particle images (no.)	83,317	29,966	74,774	58,762	137,997	78,523
Map resolution, masked (Å)	3.69	6.07	3.58	4.25	3.47	3.65
B-factor estimated (Å^2)	105	198	93.0	129	91.7	92.0
EMD accession code	23453	23454	23455	23456	23457	23458

EXTENDED DATA TABLE 2: Statistics of cryo-EM model refinement and geometry

Dataset: p97-Npl4/Ufd1-Ub$_n$-Eos-CHAPSO		
	Class1 (Closed State)	**Class2 (Open State)**
Composition (#)		
Chains	8	8
Atoms	35,051	35,072
Residues	Protein: 4,447	Protein: 4,456
Water	0	0
Ligands	MG: 7	MG: 7
	ADP: 5	ADP: 5
	ATP: 7	ATP: 7
Bonds (RMSD)		
Length (Å)	0.003	0.004
Angles (°)	0.617	0.673
MolProbity score	1.52	1.68
Clash score	9.95	10.93
Ramachandran plot (%)		
Outliers	0.00	0.02
Allowed	1.54	2.63
Favored	98.46	97.35
Rotamer outliers (%)	0.03	0.03
Cβ outliers (%)	0.00	0.00
Peptide plane (%)		
Cis proline/general	10.1/0.0	10.2/0.0
Twisted proline/general	0.0/0.0	0.0/0.0
CaBLAM outliers (%)	1.83	1.58
ADP (B-factors)		
Iso/Aniso (#)	35,051/0	35072/0
min/max/mean		
Protein	0.48/75.21/37.58	0.36/49.31/19.03
Ligand	11.67/65.09/33.22	2.26/41.30/13.85
PDB accession code	7LN5	7LN6

Dataset: p97-Npl4/Ufd1-Ub$_n$-Eos-FOM			
	Class1	**Class2**	**Class3**
Composition (#)			
Chains	8	8	8
Atoms	34,941	34,920	35,103
Residues	Protein: 4,433	Protein: 4,431	Protein: 4,455
Water	0	0	0
Ligands	MG: 7	MG: 7	MG: 8
	ADP: 5	ADP: 5	ADP: 4
	ATP: 7	ATP: 7	ATP: 8
Bonds (RMSD)			
Length (Å)	0.004	0.005	0.003
Angles (°)	0.651	0.674	0.604
MolProbity score	1.57	1.55	1.43
Clash score	11.31	10.70	7.52
Ramachandran plot (%)			
Outliers	0.07	0.07	0.00
Allowed	1.61	1.75	2.07
Favored	98.32	98.18	97.93
Rotamer outliers (%)	0.00	0.00	0.03
Cβ outliers (%)	0.00	0.00	0.00
Peptide plane (%)			
Cis proline/general	8.5/0.0	8.5/0.0	10.1/0.0
Twisted proline/general	0.0/0.0	0.0/0.0	0.0/0.0
CaBLAM outliers (%)	1.58	1.65	1.48
ADP (B-factors)			
Iso/Aniso (#)	34,941/0	34,920/0	35,103/0
min/max/mean			
Protein	0.79/73.39/37.41	0.45/71.85/36.71	0.39/86.71/50.32
Ligand	7.76/90.54/36.11	7.16/72.98/34.80	15.72/92.31/44.90
PDB accession code	7LN2	7LN3	7LN4

Dataset: p97-Npl4/Ufd1-Ub$_6$-FOM			
	Class1	**Class2**	**Class3**
Composition (#)			
Chains	8	8	8
Atoms	34,994	34,983	34,796
Residues	Protein: 4,434	Protein: 4,434	Protein:
Water	0	0	0
Ligands	MG: 8	MG: 7	MG: 7
	ADP: 4	ADP: 4	ADP: 5
	ATP: 8	ATP: 8	ATP: 7
Bonds (RMSD)			
Length (Å)	0.004	0.003	0.006
Angles (°)	0.654	0.646	0.759
MolProbity score	1.60	1.59	1.79
Clash score	12.30	11.95	14.43
Ramachandran plot (%)			
Outliers	0.09	0.05	0.05
Allowed	1.64	1.75	2.63
Favored	98.27	98.20	97.32
Rotamer outliers (%)	0.00	0.03	0.00
Cβ outliers (%)	0.00	0.00	0.00
Peptide plane (%)			
Cis proline/general	8.9/0.0	8.9/0.0	9.0/0.0
Twisted proline/general	0.0/0.0	0.0/0.0	0.0/0.0
CaBLAM outliers (%)	1.60	1.67	
ADP (B-factors)			
Iso/Aniso (#)	34,994/0	34,983/0	34,796/0
min/max/mean			
Protein	3.48/79.40/39.29	2.27/77.11/40.21	2.28/70.59/33.60
Ligand	12.21/68.94/37.64	8.22/74.08/36.83	11.82/52.39/30.60
PDB accession code	7LMZ	7LN0	7LN1

Dataset: p97-Npl4/Ufd1-Ub_6-NMS-873 (80 µM)	
Composition (#)	
Chains	6
Atoms	35,646
Residues	Protein: 4,470
Water	0
Ligands	MG: 12
	NM3: 6 (NMS-873)
	ATP: 6
Bonds (RMSD)	
Length (Å)	0.003 (0)
Angles (°)	0.845 (3)
MolProbity score	1.79
Clash score	8.11
Ramachandran plot (%)	
Outliers	0.00
Allowed	2.71
Favored	97.29
Rotamer outliers (%)	1.89
Cβ outliers (%)	0.00
Peptide plane (%)	
Cis proline/general	10.0/0.0
Twisted proline/general	0.0/0.0
CaBLAM outliers (%)	0.95
ADP (B-factors)	
Iso/Aniso (#)	35,646/0
min/max/mean	
Protein	1.58/57.10/22.25
Ligand	2.92/28.45/15.85
PDB accession code	7LMY

EXTENDED DATA TABLE 3: A summary of human p97 disease mutations and potentially affected structures

Disease Mutations	Potentially Affected Structures
R93C/R93H, I114V, I151V	N domain structure
I126F, T127A, P137L, M158V, R159G/R159C/R159H, E185K	N-cofactor interaction
I27V, R95C/R95G/R95H, G97E, R155S/R155H/R155L, G156C/G156S, G157R, M158V, R159G/R159C/R159H, R191G/R191Q, L198W, G202W, A232E, T262A, K386E, N387H/N387S/N387T, A439S/A439P/A439G	N-D1 interaction
I206F	D1 nucleotide binding site
A232E	D1-D1 interaction
D395G	D1 domain structure
N401S	D1-D2 interaction
R487H, R662C	D2-D2 interaction
D592N	Pore loop II of D2 domain
N750S	C terminal tail

REFERENCES

1 Bodnar, N. & Rapoport, T. Toward an understanding of the Cdc48/p97 ATPase. *F1000Res* **6**, 1318 (2017).

2 Erzberger, J. P. & Berger, J. M. Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annu. Rev. Biophys. Biomol. Struct.* ***35****, 93-114* (2006).

3 Rabouille, C., Levine, T. P., Peters, J.-M. & Warren, G. An NSF-like ATPase, p97, and NSF mediate cisternal regrowth from mitotic golgi fragments. *Cell* **82**, 905-914 (1995).

4 van den Boom, J. & Meyer, H. VCP/p97-Mediated Unfolding as a Principle in Protein Homeostasis and Signaling. *Mol Cell.* **69**, 182-194 (2018).

5 Ye, Y. H., Tang, W. K., Zhang, T. & Xia, D. A Mighty "Protein Extractor" of the Cell: Structure and Function of the p97/CDC48 ATPase. *Front Mol Biosci.* **4**, 39 (2017).

6 Tang, W. K. & Xia, D. Mutations in the Human AAA+ Chaperone p97 and Related Diseases. *Front Mol Biosci.* **3**, 79 (2016).

7 Darwich, N. F. *et al.* Autosomal dominant VCP hypomorph mutation impairs disaggregation of PHF-tau. *Science* **370**, eaay8826 (2020).

8 Huryn, D. M., Kornfilt, D. J. P. & Wipf, P. p97: An Emerging Target for Cancer, Neurodegenerative Diseases, and Viral Infections. *J Med Chem.* **63**, 1892-1907 (2020).

9 Skrott, Z. *et al.* Alcohol-abuse drug disulfiram targets cancer via p97 segregase adaptor NPL4. *Nature* **552**, 194-199 (2017).

10 Xia, D., Tang, W. K. & Ye, Y. Structure and function of the AAA+ ATPase p97/Cdc48p. *Gene* **583**, 64-77 (2016).

11 Huyton, T. *et al.* The crystal structure of murine p97/VCP at 3.6 angstrom. *J Struct Biol* **144**, 337-348 (2003).

12 Davies, J. M., Brunger, A. T. & Weis, W. I. Improved structures of full-length p97, an AAA ATPase: Implications for mechanisms of nucleotide-dependent conformational change. *Structure* **16**, 715-726 (2008).

13 Banerjee, S. *et al.* 2.3 A resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. *Science* **351**, 871-875 (2016).

14 Zhang, X. *et al.* Structure of the AAA ATPase p97. *Mol Cell* **6**, 1473-1484 (2000).

15 Ripstein, Z. A., Huang, R., Augustyniak, R., Kay, L. E. & Rubinstein, J. L. Structure of a AAA+ unfoldase in the process of unfolding substrate. *Elife* **6**, e25754 (2017).

16 Puchades, C. *et al.* Structure of the mitochondrial inner membrane AAA+ protease YME1 gives insight into substrate processing. *Science* **358**, eaao0464 (2017).

17 Zhao, M. L. *et al.* Mechanistic insights into the recycling machine of the SNARE complex. *Nature* **518**, 61-67 (2015).

18 Cooney, I. *et al.* Structure of the Cdc48 segregase in the act of unfolding an authentic substrate. *Science* **365**, 502-505 (2019).

19 Twomey, E. C. *et al.* Substrate processing by the Cdc48 ATPase complex is initiated by ubiquitin unfolding. *Science* **365**, eaax1033 (2019).

20 Alverez, C. *et al.* Allosteric Indole Amide Inhibitors of p97: Identification of a Novel Probe of the Ubiquitin Pathway. *Acs Med Chem Lett.* **7**, 182-187 (2016).

21 Magnaghi, P. *et al.* Covalent and allosteric inhibitors of the ATPase VCP/p97 induce cancer cell death. *Nat Chem Biol.* **9**, 548-556 (2013).

22 Zhang, T., Midshra, P., Hay, B. A., Chan, D. & Guo, M. Valosin-containing protein (VCP/p97) inhibitors relieve Mitofusin-dependent mitochondrial defects due to VCP disease mutants. *Elife* **6**, e17834 (2017).

23 Zhang, J. T. *et al.* Identification of NMS-873, an allosteric and specific p97 inhibitor, as a broad antiviral against both influenza A and B viruses. *Eur J Pharm Sci* **133**, 86-94 (2019).

24 Bojkova, D. *et al.* Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* **583**, 469-472 (2020).
25 Hao, Q. *et al.* A non-canonical role of the p97 complex in RIG-I antiviral signaling. *Embo J.* **34**, 2903-2920 (2015).
26 Pan, M. *et al.* Seesaw conformations of Npl4 in the human p97 complex and the inhibitory mechanism of a disulfiram derivative. *Nat Commun.* **12**, 121, doi:10.1038/s41467-020-20359-x (2021).
27 Blythe, E. E., Gates, S. N., Deshaies, R. J. & Martin, A. Multisystem Proteinopathy Mutations in VCP/p97 Increase NPLOC4.UFD1L Binding and Substrate Processing. *Structure* **27**, 1820-1829 (2019).
28 Bodnar, N. O. & Rapoport, T. A. Molecular Mechanism of Substrate Processing by the Cdc48 ATPase Complex. *Cell* **169**, 722-735 (2017).
29 Hanzelmann, P. & Schindelin, H. Structural Basis of ATP Hydrolysis and Intersubunit Signaling in the AAA+ ATPase p97. *Structure* **24**, 127-139 (2016).
30 Cooney, I. *et al.* Structure of the Cdc48 segregase in the act of unfolding an authentic substrate. *Science* **365**, 502-505 (2019).
31 Blythe, E. E., Olson, K. C., Chau, V. & Deshaies, R. J. Ubiquitin- and ATP-dependent unfoldase activity of P97/VCP center dot NPLOC4 center dot UFD1L is enhanced by a mutation that causes multisystem proteinopathy. *Proc Natl Acad Sci USA.* **114**, E4380-E4388 (2017).
32 Monroe, N., Han, H., Shen, P. S., Sundquist, W. I. & Hill, C. P. Structural basis of protein translocation by the Vps4-Vta1 AAA ATPase. *Elife* **6**, e24487 (2017).
33 Yokom, A. L. *et al.* Spiral architecture of the Hsp104 disaggregase reveals the basis for polypeptide translocation. *Nat Struct Mol Biol.* **23**, 830-837 (2016).
34 Song, C., Wang, Q. & Li, C.-C. H. ATPase Activity of p97-Valosin-containing Protein (VCP). *J Biol Chem.* **278**, 3648-3655 (2003).
35 Martin, A., Baker, T. A. & Sauer, R. T. Rebuilt AAA+ motors reveal operating principles for ATP-fuelled machines. *Nature* **437**, 1115-1120 (2005).
36 Meyer, H. & Weihl, C. C. The VCP/p97 system at a glance: connecting cellular function to disease pathogenesis. *J Cell Sci.* **127**, 3877-3883 (2014).
37 Tan, D., Blok, N. B., Rapoport, T. A. & Walz, T. Structures of the double-ring AAA ATPase Pex1/Pex6 involved in peroxisome biogenesis. *FEBS J.* **283**, 986-992 (2015).
38 Pickart, C. M. & Raasi, S. Controlled synthesis of polyubiquitin chains. *Methods Enzymol* **399**, 21-36 (2005).
39 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017).
40 Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* **142**, 334-347 (2003).
41 Zivanov, J. *et al.* New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).
42 Pettersen, E. F. *et al.* UCSF chimera - A visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004).
43 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **60**, 2126-2132 (2004).
44 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D* **66**, 213-221 (2010).