Structural and Functional Analysis of Disease-Linked p97 ATPase Mutant Complexes

Purbasha Nandi 1,†, Shan Li 2,†, Rod Carlo A. Columbres 2, Feng Wang 2, Dewight R. Williams 3, Yu-Ping Poh 2, Tsui-Fen Chou 2,* and Po-Lin Chiu 1,*,†

1 Biodesign Center for Applied Structural Discovery, School of Molecular Sciences, Arizona State University, Tempe, AZ 85287, USA; pnandi@asu.edu
2 Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125, USA; lflshan@caltech.edu (S.L.); rodcarlo.columbres@gmail.com (R.C.A.C.); fengwang@caltech.edu (F.W.); yu-ping.poh@asu.edu (Y.-P.P.)
3 Eyring Materials Center, Arizona State University, Tempe, AZ 85287, USA; dewight.williams@asu.edu
* Correspondence: tfchou@caltech.edu (T.-F.C.); plchiu@asu.edu (P.-L.C.)
† These authors contributed equally to this work.

Abstract: IBMPFD/ALS is a genetic disorder caused by a single amino acid mutation on the p97 ATPase, promoting ATPase activity and cofactor dysregulation. The disease mechanism underlying p97 ATPase malfunction remains unclear. To understand how the mutation alters the ATPase regulation, we assembled a full-length p97 R155H with its p47 cofactor and first visualized their structures using single-particle cryo-EM. More than one-third of the population was the dodecameric form. Nucleotide presence dissociates the dodecamer into two hexamers for its highly elevated function. The N-domains of the p97 R155H mutant all show up configurations in ADP- or ATPγS-bound states. Our functional and structural analyses showed that the p47 binding is likely to impact the p97 R155H ATPase activities via changing the conformations of arginine fingers. These functional and structural analyses underline the ATPase dysregulation with the miscommunication between the functional modules of the p97 R155H.

Keywords: p97 ATPase; p97 R155H mutation; p47 cofactor; arginine finger; IBMPFD; single-particle cryo-EM

1. Introduction

Human p97/VCP (valosin-containing protein) belongs to the type II AAA+ (ATPase Associated with various cellular Activities) protein family, which has two AAA+ ATPase domains [1,2]. The orthologs are known as Cdc48 in Saccharomyces cerevisiae and transitional endoplasmic reticulum (TER) ATPase in archaea and eukaryotes. p97 acts as a molecular hub, interacting with various cofactors to perform a wide variety of cellular functions, including autophagy, cell-cycle regulation, ubiquitin-dependent proteostasis, and reassembly of Golgi and nuclear membranes [3–7]. p97 is abundant in the cytosol, comprising 1% of the cytosolic proteins [8,9], and carries out ATP hydrolysis to gain energy to fuel the conformational change underlying its activities [10–13]. Abnormal functions or uncontrolled regulations of p97 can cause serious diseases. Single amino acid mutations in human p97 have been linked to diseases of aging and neurodegeneration, including IBM/PFD (inclusion body myopathy with Paget’s disease of bone and frontotemporal dementia) and familial amyotrophic lateral sclerosis (ALS) [14–17].

p97 comprises an N-terminal domain (NTD), two ATPase domains, D1 and D2, containing conserved Walker A and B motifs, and an unstructured C-terminal tail [2]. The NTD and C-terminal tail are the key domains regulating ATPase function [10,13,18,19]. Six p97 monomers form a functional complex, featuring a double-ring structure based on the association of D1 and D2. The D1 domain is responsible for p97 oligomerization,
regardless of their nucleotide-binding states. Upon ATP binding to the D1 domain [20–22],
the D2 domain is the primary mediator of ATP hydrolysis [21,23–27]. The two stacked
rings create a central channel in the p97 hexamer that binds substrates, facilitating the
segregation of substrates destined for proteolysis from their usual location in multi-subunit
complexes, organelar membranes, and chromatin [28,29]. Previous cryogenic electron
microscopic (cryo-EM) structures of the AAA+ ATPases revealed that the substrate process-
ing by the Cdc48 ATPase results in a broken ring and an asymmetric structure of the six
p97 monomers. These structures have been used to suggest a hand-over-hand model for
the substrate’s induced motion [30–32]. The substrate movement along the central channel
is dependent on the D1 nucleotide state and is mainly driven by the ATP hydrolysis on the
D2 domain [33,34]. This movement would call for cooperation between domains within the
AAA+ ATPase complex and cofactor bindings [10,35,36]. However, it is not certain whether
the hand-over-hand model applies to other p97 functions, such as membrane fusion.

The p97 NTD mediates the binding of various cofactors and ubiquitylated substrates
and modulates ATPase activity [13,19]. Although the p97 NTD does not directly relate to
nucleotide binding, it has two distinct configurations that are affected by the D1 nucleotide-
binding states: the ‘up’ (ATP-bound D1) configuration, rising above the plane of the D1
ring, and the ‘down’ (ADP-bound D1) configuration, coplanar with the D1 ring [26,27,37].
However, previous EM structural analysis revealed variant NTD positions, rather than the
only two distinct states [19,38]. One hypothesis suggests that the NTD configuration affects
linker conformation, initiating inter-subunit or intra-domain communication, thus influ-
encing ATPase activity [39–41]. This is supported by the position of IBMPFD-associated
disease mutants that result in elevated ATPase activities, mostly from the D2 ATPase [19,42].

The IBMPFD mutations cluster on the NTD or the N-D1 linker of the p97 [26,43]. Because
the mutation sites are far away from the D2 nucleotide-binding site, it has been suggested
that the NTD mutation changes the interactions between p97 inter-domains, leading to
functional alterations [34]. However, the connection between the pathological effects of the
IBMPFD mutations and elevated p97 ATPase activity remains unclear.

In addition to mutations, the binding of p97 cofactors on the NTDs also impacts
ATPase activities. p47 is one of the p97 cofactors involved in p97-mediated membrane
fusion [7]. It contains an N-terminal UBA (ubiquitin-associated) domain, followed by a SEP
(shp1, eyc, and p47) and UBX (ubiquitin-regulatory X) domain, which interacts with the
p97 NTD (Figure 1A) [24,40,44]. The interaction reduces the D1 ATPase activities of both
wild-type and p97 mutants [45]. Also, p47 enhances the D2 ATPase activity of wild-type
p97 (p97\textsuperscript{WT}), whereas it reduces the activities of the p97 disease mutants [45]. Thus, p47
has different effects on the two ATPase domains of both p97\textsuperscript{WT} and its mutants, resulting
in that the mutant p97-p47 complex blocks the ATP hydrolysis in one single sigmoidal
phase, rather than a biphasic response of the wild-type p97-p47 complex [45].

The binding mode of the p47 UBX domain to the p97 NTD is conserved among
the UBX or UBX-like (UBX-L) containing cofactors [46]. Previous structures of partial
or complete complexes reveal this interaction as seen in p97\textsuperscript{N-D1-p47UBX}, p97\textsuperscript{N-FAF1UBX},
p97\textsuperscript{N-OTU1UBX}, and p97-ASPL [24,47–51]. The high-resolution crystal structure of the
p97\textsuperscript{N-D1-p47UBX} truncated complex showed that a conserved β sheet with an S3/S4 loop
in the UBX domain binds to a hydrophobic pocket formed by the two p97 N-subdomains
(comprised of an N\textsubscript{a} (four-stranded β barrel) and an N\textsubscript{b} (ψ barrel) domain) [24]. Recently,
the bipartite and tripartite binding modes have been proposed for the interactions between
multiple p47 and wild type p97 NTDs [51]. However, how this binding in the combination
of disease NTD mutations modulates the D1 and D2 ATPase activities remain uncertain.
Most of the current data for p97 mutants were based on the structures of truncated do-
 mains [52]. Because the structural information of the full-length diseased complexes is still
lacking, more efforts on studying their functions and structures are needed to understand
the mechanistic details.
Because the structural information of the full-length diseased complexes is still lacking, more efforts on studying their functions and structures are needed to understand the mechanistic details.

Figure 1. Characterization of the p97 R155H-p47 and p97 WT-p47 complex formation. (A) Domain structures of the p97 ATPase and its cofactor p47. The disease-linked mutation of R155H is labeled on the NTD of p97. (B) Size-exclusion chromatographic (SEC) profiles of the p97 WT-p47 (left) and p97 R155H-p47 (right) assemblies. Blue and purple curves are for the p97-p47 assembly and p47 protein alone. Peak fractions are labeled. (C) SDS-PAGE analysis of the SEC eluted peaks of the p97-p47 assemblies. (D) Determination of dissociation constants, $K_d$, for the p97 WT-p47 and p97 R155H-p47 complexes. Unlabeled p97 was titrated against 10 nM of RED-NHS-labeled p47 in two-fold steps from 2.80 µM to 1.37 nM. The temperature-related intensity change (TRIC) signals were recorded and plotted against p47 concentrations.

Here we performed functional and structural analyses on the disease-linked complex, p97 R155H-p47, to understand how the cofactor binding and NTD mutation leads to functional alterations. We chose p97 R155H as our target, since the NTD mutation of arginine 155 to histidine (R155H) is the most commonly found in the IBMPFD phenotype [53]. We have previously identified that p97 disease mutants are defective in the cofactor-regulated ATPase hydrolysis cycle [45]. Also, the increased p47 expression can correct the autophagy defect caused by the p97 disease mutant [45]. To understand how p47 impacts the function of the p97 R155H disease mutant, we characterized the interactions between the p97 R155H and p47 using single-particle cryo-EM image analysis. The structural results revealed two major forms in the molecular population: the p97 R155H dodecamer and the p97 R155H hexamer-p47.
complex. The nucleotide-binding pockets are empty in the p97\textsuperscript{R155H} dodecamer. The presence of the nucleotides, such as ADP or ATP\textsubscript{γ}S, destabilizes the dodecameric formation. The NTDs showed all up configurations when bound to ATP\textsubscript{γ}S or in the absence of nucleotides. However, in the presence of ADP, the NTDs of the mutant allow variable intermediate positions, which may affect the accessibility of the p47 cofactor. Our structures also showed that the arginine fingers exhibit different conformations than those of the wild type when p97\textsuperscript{R155H} binds to p47. We propose a possible model that impacts functional alteration of the p97\textsuperscript{R155H} upon p47 binding, providing an insight into the pathological function of the IBMPFD-associated p97 mutants.

2. Results

2.1. Two Molecular Populations Are in the p97\textsuperscript{R155H}-p47 Assembly

To analyze the activity of mutated p97, we first assembled an in vitro complex of full-length p97\textsuperscript{R155H} with p47 in the absence of nucleotides (molar ratio of p97\textsuperscript{R155H} hexamer:p47 monomer 1:10). A peak representing the p97\textsuperscript{R155H}-p47 complex, corresponding to a molecular weight of about 670 kDa, was detected in the size-exclusion chromatographic (SEC) profile (Figure 1B). The peak profile showed a leading shoulder, implying that multiple molecular species were present in the peak fraction. We further characterized the peak fractions using SDS-PAGE, which confirmed the presence of both p97\textsuperscript{R155H} and p47 proteins in the main peak fraction and only the p97\textsuperscript{R155H} proteins in the leading shoulder (Figure 1C). This observation corroborates the previous report on the IBMPFD mutants, which showed that a full-length p97\textsuperscript{R155H} dodecamer was present in the leading fraction [19]. In contrast, the p97\textsuperscript{WT}-p47 assembly did not show an asymmetric leading peak profile (Figure 1B). Thus, the p97\textsuperscript{R155H} mutant is found in additional forms, forming complexes beyond the hexameric arrangement of the wild-type protein. Although the higher-order form has been reported for the wild type [19], inhibitor-bound [54,55], or engineered p97 [56], we did not find the same pattern in the SEC of the wild type, possibly because the high-order complex was in a less proportion in the wild type that could not be detected. In the SDS-PAGE analysis, these higher-order complexes do not associate with p47, even in the presence of excess p47 (Figure 1B). This suggests that the higher-order arrangement of p97\textsuperscript{R155H} may hinder the interaction with the p47 cofactor.

If the high order p97\textsuperscript{R155H} does not interact with p47 and forms in a larger proportion than the wild type, the p47 binding affinity to the p97\textsuperscript{R155H} on average is lower than that of the wild type p97\textsuperscript{WT}. To test this idea, we further employed fluorescent labeling and measured the temperature-related intensity change (TRIC) signal to determine the affinities of the p47 binding on p97\textsuperscript{WT} and p97\textsuperscript{R155H} and their dissociation constants, \( K_d \). In the absence of nucleotides, the resulting \( K_d \) averages of p97\textsuperscript{WT}-p47 and p97\textsuperscript{R155H}-p47 are 80 nM and 132 nM, respectively, suggesting a slightly higher binding affinity for the wild type (Figure 1D).

2.2. Up NTDs of the p97\textsuperscript{R155H} Mutant and p47 Are Structurally Disordered

To visualize the interaction between the p97\textsuperscript{R155H} mutant and the p47 cofactor, we plunge-froze the sample of the p97\textsuperscript{R155H}-p47 assembly for single-particle cryo-EM analysis. Notably, two different particle populations were discernible in the two-dimensional (2D) class averages (Figure S1A). One showed a typical side view of the p97 hexameric form, and the other showed two stacked p97 hexameric double rings (Figure S1A). Ab initio map generation produced two different three-dimensional (3D) densities presenting one with general p97 hexameric features and the other containing a dimer of hexamers (Figure S1B). We reconstructed the 3D density maps of the p97\textsuperscript{R155H} dodecamer complex at 3.34 Å resolution and the p97\textsuperscript{R155H}-p47 complex at 4.30 Å (Figure 2 and Figure S1B,C). The ratio of the single-particle images of the p97\textsuperscript{R155H}-p47 complex to the p97\textsuperscript{R155H} dodecamer is 1.56:1.
Local resolution estimation showed that the density of the p97\textsuperscript{R155H} dodecamer or the p97\textsuperscript{R155H}-p47 complex is well resolved in D1 and D2 domains but not the NTDs (Figure S2D). For p97\textsuperscript{R155H} dodecamer, we did not find densities that could be assigned to p47. Based on specific map features of the p97\textsuperscript{R155H}-p47 complex, we were able to identify the NTD and p47\textsuperscript{UBX} domains at a lower contour level (2.7σ) by docking available atomic models (PDB code: 1S3S) into the density map (Figure 2B) [24]. Four NTD densities could be assigned and shown in up configurations, and they are rendered into slightly different orientations in the azimuthal direction along the C6 pseudo-symmetrical axis (Figure 2B). Only one p47\textsuperscript{UBX} domain was found to fit into the density attached to the up NTD (Figure 2B). We did not identify any high-order form of the p47 bound to p97, and it is consistent with the previous finding that the p97 NTDs interfere with p47.
oligomerization [51,57]. In addition, the flattened densities of the NTD of p97 and p47 are possible due to their dynamic and partially disordered nature, leading to the weak scattering signals and limiting resolution of these components (Figure 2B and Figure S1D). We also performed image classification in 3D, but the result did not yield high-resolution densities for modeling the SEP or UBA domain of the p47 (Figure S1D). This observation is also corroborated by a neural network data analysis [58], showing high variations in the NTD and p47 densities (Figure 2C). The densities of D1 and D2 rings were invariable across the reconstructions. Unlike the p97WT-p47 complex [51], the mutant NTD may be too disordered or mobile to present a structural feature with the p47 factor to be resolved by cryo-EM.

In this complex assembling, we did not supply any nucleotides in the purification steps, and nucleotide was not detectable in the D1 and D2 nucleotide-binding sites of the p97R155H dodecamer or the p97R155H-p47 cryo-EM density maps. This observation differs from the previous study where ADP is always observed in the wild type p97 D1 ATPase, even when the nucleotide was not supplied in the sample buffers [23,26,59–61]. It is the first time that the full-length p97R155H disease mutant structure is revealed by single-particle cryo-EM. One possibility could be that the nucleotide binding for the p97R155H mutant is weak in the extremely low concentrations of nucleotides, so the densities of the nucleotides were insufficient to be detected in the averages of the cryo-EM images. Since the previous analyses were conducted using p97WT or truncated p97 mutants, it is likely that the R155H mutation or the domain-domain interactions in the full-length p97R155H D1 ATPase is two-fold faster than that of the p97WT D1 ATPase and the nucleotide binding in the D2 domain depends on the D1 nucleotide state, it indicates a high likelihood of the ADP binding instability in D1 ATPase of the full-length p97R155H [62]. Apart from this, the previous SPR experiments showed that the R155H mutation reduces the ADP-binding affinity to p97 [61,63]. Thus, the p97R155H mutant is likely to be structurally stable when the nucleotide-binding pockets are empty.

2.3. P97R155H Dodecamer Is Stabilized by the Two Oppositely Stacked D2 Rings

Our SEC profile and SDS-PAGE analysis showed that the dimeric p97R155H hexamers do not bind p47 (Figure 1B,C). To elucidate how the two hexamers organize into a higher-order complex, we built an atomic model along with the cryo-EM density map of the p97R155H dodecamer (Figure S2). The two p97R155H hexamers are oppositely packed against their D2 rings with a D6 symmetrical arrangement (Figures 2A and 3A). We did not find any density assigned for the p47 cofactor, as corroborated with our biochemical characterization and the $K_d$ determination of the p97-p47 binding (Figure 1B–D and Figure 2A). Thus, this suggests that the p47 cofactor does not access the p97R155H dodecamer.

The C-terminal tail of the p97 (residues Q764 to G806) that interacts with the VCP cofactors was previously reported as flexible and has not previously been structurally resolved for a full-length p97R155H mutant [19,64,65]. The crystal structure of the C-terminus up to P774 of an engineered p97 variant has also been determined, but at a weak signal level [56]. This is because in our p97R155H dodecamer, the C-terminal tail is sandwiched by the two hexameric D2 rings, limiting its mobility and allowing visualization of high-resolution details (Figure 3A). Our cryo-EM density map of the p97R155H dodecamer was able to show an apparent density of the partial C-terminal tail (residues Q764 to F773) (Figure 3A). The C-terminal tail points away from the pore center, interacting with another C-terminal tail of the opposite neighbor to stabilize the packing of the two p97R155H hexamers (Figure 3B). Packing occurs via largely polar-polar interactions between the D2 $\alpha$ helices and a hydrogen-bonding network of the C-terminal residues (D749, E756, Q760, Q764, and R766) (Figure 3B). $\pi$-stacking forces between the aromatic side chains of the C-terminal tails and the D2 $\alpha$ helices (F674, F682, F768, F771 and F773) also contribute to the packing of the two p97R155H hexamers.
Figure 3. Cryo-EM structure of the p97R155H dodecamer. Orange, orange-red, and grey indicate the D1, D2, and linker domains of the p97 R155H. (A) Cryo-EM density of the partial C-terminal tail. The purple surface shows the cryo-EM density (4.6σ). The side chains are shown in stick representation. (B) Interaction network of the binding interface of the two p97 R155H hexamers. The two hexamers interact via hydrogen bonds and π-π aromatic packing. (C) Superposition of the p97 R155H with p97 structures in different nucleotide states. Dark green, light green, and purple are p97 ATPγS, p97 ADP, and p97 CB-5083, respectively. RMSDs are 1.127 Å, 0.919 Å, and 0.747 Å for p97 ATPγS, p97 ADP, and p97 CB-5083, respectively. Blue rectangles highlight conformational changes in the D1 helix-turn-helix motif. The direction of the D2 movement upon ATPγS binding is shown in dots and an arrow.

2.4. Nucleotide Binding Destabilizes the p97R155H Dodecameric Formation

We did not find nucleotide binding in the p97R155H dodecamer structure, which occupies about 39.1% of the population (Figure 1B,C and Figure S1). In contrast, p97 WT is always associated with ADP in the D1 ATPase and has a trace amount of dodecameric form [19]. This could be possible because the nucleotide binding affects p97 oligomerization. We would like to test whether the nucleotide in the solution affects the dodecameric formation. We then mutated p97R155H on the Walker A motif of the D1 (K251R) or D2 (K524R) ATPase to disrupt the nucleotide binding. The Walker A mutants do not bind any nucleotides nor perform the function. These double mutants were then imaged using negative-stain electron microscopy (EM) for single-particle image analysis. 2D class averages of these mutants showed that the ratios of the dodecamer to hexamer are 6.94:1 and 6.69:1 for p97R155H-K251R and p97R155H-K524R, respectively (Figure S3). Compared to the above cryo-EM analysis of the p97R155H-p47 assembling, the ratios of the p97R155H
dodecamer in D1 and D2 Walker A mutants increase 10.8 and 10.4 times, respectively. This suggests that the p97R155H dodecamer is stable when it does not bind nucleotides. Moreover, the double mutants lack ATPase activity, suggesting that the dodecameric p97R155H is likely to represent an inactive state.

2.5. P97R155H Dodecamer Is Likely to Be an Inactive Form

To compare the monomeric structure of the p97R155H dodecamer with others, we superimposed our structure with p97WT in ADP or ATPγS bound state and a CB-5083-bound p97 (PDB codes: 5FTK (ADP), 5FTN (ATPγS), and 6MCK (CB-5083)) (Figure 3C) [37,54]. The superposition of p97R155H and p97ADPγS revealed a downward rigid-body movement of the D2 domain, which may be induced by the torsional change of the D1-D2 linker (Figure 3C). Another significant structural change occurs in the C-terminal tail conformation. The C-terminal tail of the p97ADPγS points to the pore center, which is in the opposite direction, against that of the p97R155H dodecamer (Figure 3C). However, the HTH motif in the p97R155H seems to not interact with the NTD, allowing NTD to be less spatially restricted in the up configuration. Although the NTD densities were resolved in low resolutions partly due to its mobility, the NTD densities in low χ contours (2.4σ) could still be identified in the up configuration (Figure S1D).

CB-5083 is a highly selective D2 ATP-competitive inhibitor on p97 ATPase, and the packing of the dimeric hexamers was also found in the crystal structure of the NTD-truncated p97 with CB-5083 [54]. Based on these superpositions, the structure of the p97R155H dodecamer more closely resembles p97CB-5083 (RMSD 0.747 Å) than p97ADP (RMSD 0.919 Å) or p97ATPγS (RMSD 1.127 Å) (Figure 3C). The crystal structure is consistent with our dodecamer structure, suggesting that the p97R155H dodecamer could be functionally equivalent to the inhibited p97 with CB-5083 (Figure 3C) [54].

2.6. Nucleotide Binding Influences the p47 Binding onto p97R155H

The nucleotide-bound state of the p97WT D1 domain affects the D2 ATPase activity and influences the binding of the p47 cofactor [2,62]. In addition, p47 has a differential influence on the D1 and D2 ATPase functions of wild-type and p97 mutants [45]. To uncover how nucleotide binding modulates p47 binding to the p97R155H mutant, we assembled the complexes in the presence of ADP or ATPγS and biochemically analyzed the resulting complexes. ATPγS was used to mimic the ATP-bound site since it cannot be hydrolyzed when bound to the nucleotide-binding site. The SEC profiles of the two assemblies in the presence of ADP (p97R155H|ADP-p47) and ATPγS (p97R155H|ATPγS-p47) were similar, and the peaks suggested a uniform distribution for both complexes (Figure S4A). SDS-PAGE gel analyses verified the peak fractions containing p97R155H and p47 (Figure S4B).

We then collected the peak fractions and froze the samples for cryo-EM structural analysis (Figures S5 and S6). Although no leading fraction was observed in the SEC profile of the p97R155H|ADP-p47, a small fraction of the p97R155H dodecamers were identified in the 2D image class averages (Figure S5). After 3D classification, the number of particle images representing the p97R155H dodecamer was only 2219 (1.37%) among an overall 162,269 particle images. Like the 3D reconstruction of the p97R155H without nucleotides, the dodecamer does not show any sign of bound nucleotide or p47 cofactor. However, the p97R155H|ATPγS-p47 image dataset did not have the p97R155H dodecameric population, suggesting that p97R155H dodecamer does not form in the presence of ATPγS (Figure S6). Thus, these suggest that the nucleotide presence, particularly ATPγS, favors the dissociation of the p97R155H dodecamer into hexamers. Furthermore, because the ATPase activity of
the p97R155H is higher than p97WT [3,19,42,45,61,66] but the p97R155H dodecamer does not interact with nucleotides, it is very likely that the increased activity of the p97R155H is contributed from the functional hexamers, rather than dodecamers.

We calculated the 3D reconstructions of the p97R155H|ADP-p47 and p97R155H|ATPγS-p47 complexes at 4.50 Å and 4.23 Å resolution, respectively (Figures S5 and S6). Cryo-EM densities of the nucleotide can be identified in both the D1 and D2 nucleotide-binding pockets of the p97R155H|ATPγS-p47 density (Figure S7). Unlike the p97R155H-p47 and p97R155H|ATPγS-p47 densities, the p97R155H|ADP-p47 density map showed a heterogeneous distribution of NTD orientations (Figure S8A). An NTD density with bound p47UBX domain was identified (1.0σ), located at the highest height among all the other NTDs (Figure S8B). The tilting angles and heights of the NTDs relative to the ring plane varied sequentially (Figure S8B). This observation of the high structural variabilities is consistent with the deep coordinate neural network analysis (Figure S9).

For the p97R155H|ATPγS-p47 complex, four NTDs and two p47UBX structures can be assigned to the cryo-EM density, and the unassigned two NTD densities were fragmented (Figure 4A). The four up NTD are consistent with the crystal structure of the truncated N-D1 domain of the p97R155H|ATPγS [26]. Like the p97R155H|ADP-p47 complex, the densities of NTDs and p47 were highly variable, but those for D1 and D2 domains were almost invariant, as shown in the local resolution estimation (Figure S6D).

2.7. P47 Binding Impacts p97R155H Function via an Allosteric Effect on ATPases

Previous functional analyses demonstrated that cofactor binding or NTD mutation alters p97 ATPase function, mainly the D2 ATPase [42,45]. However, the NTD is far away from the nucleotide-binding pocket and how do NTDs allosterically regulate p97 ATPase functions? Because D1 and D2 ATPase activities are correlated, one possibility is that the NTD regulates ATPase functions via inter-domain or interprotomer communications between these functional modules [42,67]. Because the cryo-EM density of p97R155H|ATPγS-p47 allows us to visualize the side chains in its D1 and D2 domains, we compared it with the model with p97WT|ATPγS structure (PDB code: 5FTN) [37]. We identified two p47UBX bound onto the p97R155H NTDs, implying majority of the p97R155H particles bind two p47 molecules. The superposition of the structures of the p47-bound p97R155H|ATPγS and p97WT|ATPγS alone showed that the D2 domain slightly moves and most of the critical residues in the nucleotide-binding pockets do not exhibit major conformational changes (RMSD: 1.015 Å) (Figure 4B). However, the D2 R635, the D2 arginine finger, of the p97R155H|ATPγS changes its conformation upon p47 binding, although the conformations of the D1 arginine fingers seem unchanged (Figure 4B). It is also known that the arginine fingers extending from the neighbor protomer can affect the ATP hydrolysis efficiency [3,68,69].

We would then like to test whether the D2 ATPase activity change is affected by NTD mutation or p47 binding via the arginine fingers. We introduced two mutations of the arginine fingers on R359A and R635A and measured the mutant ATPase activities (Figure 4C). The p97R359A and p97R635A mutants have reduced activities compared to the wild type (59% and 5%, respectively), and the mutation of the D2 R635 nearly abolished the p97 ATPase activity, consistent with the previous findings [70]. We next prepared the two double mutants, p97R155H-R359A and p97R155H-R635A, and the results showed that both double mutants reduced p97R155H activities to 26% and 2.3%, respectively. Because the p97R155H has a higher ATPase activity than p97WT (Figure 4C) [3,19,42,45,61,66], the results showed that the increased p97R155H ATPase activity due to its R155H mutation was diminished by the D1 and D2 arginine mutations. Both of the D1 and D2 arginine fingers appear to affect the communications within p97 subunits for the R155H functional enhancement. Thus, the gained ATPase activity from the R155H mutation of the NTD is very likely to associate with functional arginine fingers.
Figure 4. Cryo-EM structural analysis of the p97<sup>R155H</sup>|ATP<sub>γ</sub>S-p47 complex and functional measurements of the p97 arginine finger mutants. (A) Cryo-EM density of the p97<sup>R155H</sup>|ATP<sub>γ</sub>S-p47 complex. Color coding is the same as those in Figure 2. Enclosed is the p97<sup>R155H</sup>|ATP<sub>γ</sub>S-p47 monomer used for structural comparison with p97<sup>WT</sup>|ATP<sub>γ</sub>S. (B) Superposition of the structures of p97<sup>R155H</sup>|ATP<sub>γ</sub>S-p47 and p97<sup>WT</sup>|ATP<sub>γ</sub>S (PDB code: 5FTN) [37]. Bound ATP<sub>γ</sub>S and arginine fingers are shown in stick representation. Upper panels are the superpositions of the two structures and white are the p97<sup>WT</sup>|ATP<sub>γ</sub>S structure. Lower panels are the map-model fitting of the two arginine fingers in D1 and D2 nucleotide-binding sites. White surfaces are cryo-EM densities. (C) Full-length (FL) p97 ATPase activities of wild type and R155H mutants. Activity measurements were normalized relative to the p97<sup>WT</sup> activity and measured in the presence of 200 µM ATP (<i>n</i> = 4). (D) Normalized ATPase activities of p97 ATPase: WT (blue circles), R155H (red circles), R359A (green triangles) and R155H-R359A (magenta triangles). The R359A mutation is located in the D1 arginine finger. p97 ATPase activity was normalized relative to its basal activity in the absence of p47. The addition of p47 increased from 0 to 800 nM. ATPase activity measurements were performed in the presence of 200 µM ATP. (E) Experiments as in (B), but the mutated residue was located at the R635A, the D2 arginine finger. p97 ATPase activity in R635A and R155H-R635A mutants is shown in green squares and magenta squares, respectively. All error bars indicate ±SD (<i>n</i> = 4).

Next, we would like to know whether p47 impacts p97<sup>R155H</sup> function via arginine fingers in the same manner as R155H mutation. As seen previously, upon binding to p47, p97<sup>WT</sup> exhibits a biphasic response, but p97<sup>R155H</sup> does not exhibit the rebound Phase 2 (Figure 4D) [45]. However, the D1 arginine finger mutants, p97<sup>R359A</sup>, lacked...
Phase 1 inhibition but showed significant activation in the later phase (1.9-fold), similar to the D1 Walker B mutant (E305Q), which binds to nucleotide but does not perform the catalysis (Figure 4D) [45]. Similar to p97R359A, the D2 arginine finger mutants, p97R635A, showed a significant Phase 2 activation (3-fold), but different from the D2 Walker B mutant (E578Q) (Figure 4E) [45]. These data indicate that upon p47 binding, the Phase 1 inhibition of the p97WT associates with both D1 and D2 arginine fingers, but the Phase 2 activation is mainly contributed by the D2 ATPase activity that associates with the D2 arginine finger (Figure 4D,E). Thus, p47 likely impacts both wild type and p97R155H mutant via arginine fingers, particular in Phase 2.

The responses of the double mutants showed both impacts from R155H mutation and p47 binding. The p47 inhibitory regulation on the p97R155H was abolished for the two double mutants, the p97R155H-R359A and p97R155H-R635A mutants (Figure 4C,D). Especially for the p97R155H-R635A mutant for the D2 arginine finger, the curve is flat, implying that the p47 binding has no impact on the ATP hydrolysis. Thus, this may indicate a functional connection between p47 binding and the arginine fingers, R359 and R635, of the p97R155H mutant. The structural superpositions showed that the change of the ATPase function may result from the slight conformational changes of the arginine fingers (Figure 4A).

Thus, these findings suggest that the arginine fingers are critical for the p47-induced communications between the NTD and ATPase domain.

In addition to the functional modulation of p47 via arginine fingers, we would like to test whether the p47 binding is affected when the arginine finger is mutated. We fractionated a mixture of p97R155H-R359A or p97R155H-R635A (1.67 µM in hexameric form) and p47 (80 µM in monomeric form) proteins in a gel-filtration column, using p97R155H as a positive control (Figure S10). We found that the double mutant p97R155H-R359A and p97R155H-R635A retained the ability to bind p47, eluting in the same fraction as the p97R155H-p47 complex (Figure S10). We also determined the binding affinities, $K_d$, of the p47 to p97R155H, p97R155H-R359A, and p97R155H-R635A as 132, 179, and 228 nM, respectively (Figure S11). Thus, these results showed that the gained ATPase activities of the p97R155H-R635A from p97R155H impacted by p47 may be partly due to weak p47 binding affinity to the p97R155H-R635A double mutant (Figures 1 and 4E and Figure S11B). In addition, the impaired arginine fingers reduce the p47 binding affinity to p97R155H, implying that the domain-domain communications within p97 ATPase affect how p97 responds to cofactor binding.

3. Discussion

Single amino acid mutations in p97 have long been linked to diseases, including IBMPFD and ALS [71], and these disease mutants alter the p97 ATPase activity and cofactor binding on the NTDs [45]. Because the NTD is far away from the D2 ATPase, where ATP hydrolysis mostly occurs, it is unclear how the NTD allosterically affects the p97 ATPase function. The molecular mechanism underlying the disease is still unresolved today. Our goal aims to understand (1) how NTD conformational change, such as up and down configurations, connects to ATPase function, (2) the structural change of p97 ATPase caused by R155H mutation, and (3) how the mutant responds to p47 binding in the context of a complete full-length complex. Here we conducted biochemical and structural analyses to study potential pathological changes in the p97R155H mutant, the most prevalent disease-linked mutation [53]. To our knowledge, our report reveals the first full-length p97R155H mutant structures in dodecameric and hexameric forms, as well as the p97R155H-p47 complex in different nucleotide states. Previous studies have placed less emphasis on the p97R155H dodecamer. Our report found that the p97R155H dodecamer seems unstable in the presence of nucleotides and insensitive to bind p47. Although we do not know if the p97R155H dodecamer plays a physiological role in vivo or leads to IBMPFD disease, our cryo-EM image analysis showed that this high order form can be stabilized in an aqueous solution and occupies about 40%. We also found a possible connection for the p47 and NTD regulation on ATPase function, likely arginine fingers, and the miscommunication between these domains alters the normal ATPase activity. This may interfere with the
ability of the p97 to carry out the p47-related function such as Golgi-membrane reassembly or autophagy [30]. These new findings thus reveal how cellular functions may be so profoundly impacted in the p97\textsuperscript{R155H} mutant.

The formation of a p97\textsuperscript{WT} dodecamer has been previously reported [47,56,72]. However, our SEC analysis did not show the dodecameric fraction in the wild-type assembly, and the proportion of the p97\textsuperscript{WT} dodecamer may be too low to be detected in our SEC analysis (Figure 1B). Although it is not known whether the formation of the p97\textsuperscript{R155H} dodecamer occurs in vivo or whether it plays any role in a cell, the dodecamer at least can be stably formed in the solution and likely in the cell as well. In addition, the structural comparison of the CB-5083-bound structure [54] and its empty nucleotide-binding pockets highly suggested the dodecameric form as an inactive state. The inactivation of the p97 dodecamer could be caused by its high-order organization, which may hinder the mobility of the monomers, interfere with the hexameric ring to break, prevent hand-over-hand movements and possibly, in turn, down-regulate normal p97 ATPase function during substrate processing [31,32]. Only when the nucleotides are present, the p97\textsuperscript{R155H} dodecamer will dissociate into two functional hexamers with an elevated ATPase activity. However, we have no information about why the R155H mutation favors the dodecameric formation. One possible explanation would be that the dodecameric formation induced by R155H mutation may keep p97 inactive while under the stress of the limited ATP concentrations in the cell [54]. The definition of its functional role in vivo requires further investigations using cell biological tools.

Our data showed that the p47 binding or NTD mutation affects the communications between the functional modules of the full-length p97\textsuperscript{R155H} ATPase in either an intra-domain or an inter-domain manner or both. In addition, the affinity of p47 to p97\textsuperscript{R155H} is affected by the D1 or D2 ATPase activities. We also showed that either D1 or D2 arginine finger mutation can affect the p47 binding affinity. These views might not be observed using the truncated or incomplete p97 complexes to show the effects caused by domain communications. The partial p97 ATPase may also lead to inconsistent observations on the stoichiometry of the p97-p47 bindings. For example, the crystal structure of the p97\textsuperscript{N-D1}-p47\textsuperscript{UBX} showed that two p47 UBX domains bind to two adjacent p97 monomers, while a third is poorly resolved located away from those two bound NTDs [24]. However, this view is not consistent with our observation from single-particle cryo-EM. Therefore, a complete full-length complex will be required to provide a view of the interactions between overall functional modules that are close to their native states.

Previous structural studies have revealed the interactions between p97\textsuperscript{WT} and p47 proteins [24,51,57,73]. NMR (nuclear magnetic resonance) spectroscopic analysis of the p97\textsuperscript{N-D1-L}-p47 (residues 1-480 of p97) indicated a possible tripartite p47 binding mode [51]. This interaction involves the UBX and the C-terminal SHP of p47 binding to one p97 NTD, while a third p47 N-terminal SHP domain binds the adjacent NTD of the p97. The up NTD configurations allow organizing p47 in this arrangement, rather than down NTDs. On the other hand, the p97\textsuperscript{R155H} mutant, especially when bound to ATP, should interact with p47 cofactors followed by the proposed tripartite binding mode, since the NTDs are all in up configuration [51,62]. However, our p97\textsuperscript{R155H}-p47 complex structures indicate a disordered or flexible structure of NTD or p47 SHP domain, implying that p97 NTD mutant may either interact with p47 more transiently or require a concerted motion on NTDs to adapt the proposed binding mode [40,62]. In this sense, the oscillating up NTDs of the p97\textsuperscript{R155H}\textsuperscript{ADP} could easily access the p47 cofactors than the p97\textsuperscript{WT}\textsuperscript{ADP}, which has all down NTDs [37]. This corroborates with the previous finding that the binding of p47 on the p97\textsuperscript{WT} ATPase in the presence of ADP is much weaker than those in the presence of ATP or the absence of nucleotides [62]. However, the p47 binding for the p97\textsuperscript{R155H} mutant is independent of the nucleotide states [62].

Recent cryo-EM structures of the AAA+ ATPase with p47 cofactor or its homolog showed an asymmetric arrangement of the hexameric structure [31,35]. These structures have shown the substrate binding in the central channel of the ATPase double
modulate the ATPase function. Our analyses suggest the importance of using a full-length, rather than truncated, p97 for functional or structural characterization to gain a complete view of the domain-domain communications within the p97 hexamer. To build on our finding regarding dysregulation of interprotomer communication in p97 R155H hexamer, we expect the incorporation of additional physiological substrates for p97-p47 assembly for the next critical step to understand the p97 pathological mechanism.

Figure 5. Proposed model for p47 binding to the disease-linked p97 R155H mutant. The NTD, D1, and D2 domains of the p97 R155H are colored green, orange, and orange-red. In the absence of nucleotides, p97 R155H can form a dodecamer and become inactive. p47 does not access the p97 R155H dodecamer in the absence of nucleotides. The p97 R155H hexamer does not stably bind to p47 in the absence of nucleotides. Once p97 R155H binds nucleotides, it stably interacts with p47. The p47 binding blocks the p97 R155H ATPase activity. Although the NTDs are all above the D1 ring plane in the presence of ADP, they adopt slightly different tilting angles and heights. However, in the presence of ATP, the NTDs are all in the same up configuration. The p47 interacts with the up-NTD, but not the down-NTD. Thicker lines indicate favor direction for reaction.
4. Materials and Methods

4.1. Overexpression of the Wild Type p97, p97R155H Disease Mutant, and p47 Proteins

Plasmid constructs used for generating p97 and its mutants and p47 proteins are listed in Table S1. Overexpression and purification of the p97 and its mutants and p47 followed previous methods [61].

4.2. ATPase Activity Measurements

Detection of ATPase activity using Biomol Green reagent (Enzo Life Sciences, Farmingdale, NY, USA) was performed as previously described with slight modifications [45]. To compare the ATPase activity of wild type and mutant p97 proteins, each purified protein was diluted to a final monomer concentration of 25 nM in 50 µL ATPase assay buffer (50 mM Tris (pH 7.4), 20 mM MgCl$_2$, 1 mM EDTA, 0.5 mM TCEP, 0.01% Triton X-100 and 80 nM BSA). After adding 200 µM ATP, the reaction was carried out at room temperature for optimal reaction times. 50 µL Biomol Green reagent (Enzo Life Sciences, Farmingdale, NY, USA) was added to stop the reaction and the absorbance at 635 nm was measured using a BioTek Synergy Neo 2 plate reader (BioTek, Winooski, VT, USA). The eleven-dose titrations with p47 cofactor were performed by adding the varying amount of p47 protein in ATPase buffer with p97. The results were calculated from five replicates using GraphPad Prism 7.0.

4.3. P47 Binding Affinity Measurements

The binding affinity (K$_d$ values) between p47 and p97$_{WT}$ or p97 mutants was determined by measuring the temperature-related intensity change (TRIC) signals using a Dianthus NT.23 instrument (Nano-Temper Technologies, München, Germany). p47 was labeled with a RED-NHS dye using a Monolith Protein Labeling Kit RED-NHS second-generation (Nano-Temper Technologies, CAT# LO-L011). The full-length wild type or mutant p97 was titrated against 10 nM of p47-NHS in two-fold steps from 2.8 µM to 1.37 nM in 20 µL working buffer (20 mM HEPES (pH 7.4), 150 mM KCl, 1 mM MgCl$_2$, 5% (w/v) glycerol, and 0.0025% (v/v) Tween 20). Assays were performed in a Dianthus 384-well plate and reproduced in three independent experiments. Data from the three independent measurements were fitted using non-linear regression analysis in Prism 7.0.

4.4. Assembling p97-p47 Complexes

To assemble p47 with p97$_{WT}$ or p97 mutants, 1.67 µM p97 hexamer was mixed with 80 µM p47 for 10 min and fractionated in 20 mM HEPES (pH 7.4), 150 mM KCl, 1 mM MgCl$_2$, and 5% glycerol. Three different assemblies were prepared in the absence of nucleotides and the presence of 100 µM ADP or ATP$\gamma$S in the SEC elution buffer. The mixture of the p97-p47 complex was loaded onto a Superdex 200 10/300 GL (GE Healthcare, Chicago, IL, USA) for size-exclusion chromatography (SEC). When the samples were used for cryo-EM analysis, 0.25% (v/v) glutaraldehyde (Electron Microscopy Sciences, Hatfield, PA, USA) was applied for on-column cross-linking by following the previous method [78]. Peak fractions were characterized using SDS-PAGE, Western blotting, and negative-stain electron microscopy (EM).

4.5. Negative-Stain Electron Microscopy for Single-Particle Analysis

The negatively stained specimens were prepared using the previous method [79]. 0.01 mg/mL of the protein sample was used to applied onto a continuous carbon-coated copper grid. The specimen was imaged using a Philips CM12 transmission electron microscope (TEM) (80 keV) with a side-mounted CCD camera (Model 791, Gatan, Pleasanton, CA, USA) or a FEI Tecnai TF20 TEM with a CCD camera.

For imaging the Walker A mutants, the specimens were imaged under a FEI Tecnai TF20 TEM, recording at a pixel size of 1.04 Å/pixel at the specimen level. 153 and 144 electron images were collected for p97$^{R155H-K251R}$ and p97$^{R155H-K524R}$, respectively, and imported into Relion (version 3.1-beta-commit-ca101f) [80] for image processing. 23,342
and 23,020 particles of p97$^{R155H-K251R}$ and p97$^{R155H-K524R}$ were manually selected from the electron images, respectively, and the two-dimensional (2D) class averages with an assigned $k$ of 50 were calculated. For the p97$^{R155H-K251R}$ dataset, the numbers of side views for dodecamer and hexamer were 2173 and 313, respectively. For the p97$^{R155H-K524R}$ dataset, the numbers are 2007 and 300, respectively.

4.6. Cryo-EM Data Collection

A holey-carbon C-flat grid (2/1 4C; Protochips, Morrisville, NC, USA) was glow-discharged for 15 s using a Pelco easiGlow glow-discharge system (Ted Pella, Redding, CA, USA). A 6 µL protein sample was applied to the pretreated grid, and the excess solution was blotted using a filter paper (retention 20 µm) (Product #:47000-100, Ted Pella, Redding, CA, USA). The grid specimen was quickly plunge frozen into liquid ethane using a Thermo Fisher/FEI Vitrobot Mark IV automated freeze plunger (Thermo Fisher/FEI, Hillsborough, OR, USA) for 6 s in a chamber with a humidity of 100%. Particle homogeneity and ice backgrounds of the grid specimen were screened using a FEI Tecnai TF20 TEM. Grids with thin ice and a homogenous particle dispersion were used for subsequent cryo-EM data collection.

All the cryo-EM data in this report were collected at the Eyring Materials Center (EMC) at Arizona State University (ASU) (Tempe, AZ, USA) using a Thermo Fisher/FEI Titan Krios TEM (Thermo Fisher/FEI, Hillsborough, OR, USA) at an accelerating voltage of 300 keV. Cryo-EM movies were recorded using a Gatan K2 Summit direct electron detector (DED) camera (Gatan, Pleasanton, CA, USA). Defocus range was set to −0.8 to −2.5 µm. Nominal magnification was 48,077X, resulting in a physical pixel size of 1.04 Å/pixel at the specimen level. The movie data was recorded at a counted rate of 2 e⁻/sub-pixel/sec and a sub-frame rate of 200 msec in super-resolution mode. Total exposure was 6 s, accumulating to a dosage of 44.4 e⁻/Å². The beam-image shift was applied to accelerate data acquisition [81]. Data collection was automated using the customized SerialEM macros (version 3.9) [82].

4.7. Image Processing

For the p97$^{R155H}$-p47 assembly, 4223 collected movies were unpacked and gain-normalized using Relion (version 3.1-beta-commit-ca101f) [80]. Image frames of a movie were translationally registered and averaged with a dose-weighting scheme. The final frame average was Fourier-cropped at the spatial frequency of 1.04 Å⁻¹. The defocus and astigmatism of the images were estimated using the CTFFIND4 program (version 4.1.13) [83]. A few particles were manually selected from the images, and their average was then served as a searching template for the subsequent automated particle picking in Relion [80]. 368,575 selected particle images were automatically selected, and the data curation was completed using iterative 2D unsupervised classification. The classes with discernible features were selected for ab initio volume generation using cryoSPARC software (version 2.15) [84]. The two three-dimensional (3D) densities of a p97$^{R155H}$ dodecamer and a p97$^{R155H}$-p47 assembly were generated.

For particle images of the p97$^{R155H}$ dodecamer, one additional round of 3D classification was performed to remove poorly aligned particle images. The selected 3D class average was then refined against 64,252 experimental particle images by enforcing a D6 symmetry. The density was further improved using CTF refinement (fit of the defocus and astigmatism per particle and estimation of beam tilt) [85] and Bayesian polishing [86] procedures to a final resolution of 3.34 Å. The final map was sharpened using a $b$-factor of −78.6 Å².

For the p97$^{R155H}$-p47 complex particle images, the consensus 3D volume of the p97$^{R155H}$-p47 complex was calculated at 4.30 Å resolution with a sharpened $b$-factor of −125.7 Å². The C6 symmetry was applied to further improve the densities of the D1 and D2 domains. The generated map reached a final resolution of 3.98 Å and was sharpened using a $b$-factor of −139.3 Å². The final resolutions of all the density maps were determined using.
the golden FSC criteria at 0.143 cutoff [87], and the local resolution estimations were performed using the implementation in Relion software. The processing schematic is shown in Figure S1. On the other hand, further 3D classification procedures with \( k = 6 \) or larger or multiple hierarchical layers were attempted (Figure S1B). However, no major structural differences between the generated class averages were discernible, and no improvements were shown in the density quality in local regions of NTDs and p47 (Figure S1B).

For the \( \text{p97}^{\text{R155H}}_{|\text{ADP-p47}} \) and \( \text{p97}^{\text{R155H}}_{|\text{ATP_\gammaS-p47}} \) datasets, image processing was generally conducted using cryoSPARC software (version 2.15) [84]. 3512 and 2796 movies of the \( \text{p97}^{\text{R155H}}_{|\text{ADP-p47}} \) and \( \text{p97}^{\text{R155H}}_{|\text{ATP_\gammaS-p47}} \) were unpacked, gain-normalized and imported into cryoSPARC, respectively. Frame registration and averaging were performed using patch motion correction. Defocus and astigmatism parameters were estimated using patch CTF estimation. Particle locations were automatically selected using the Topaz program (version 0.2.3) [88].

3219 and 160,050 particle images were used to calculate the final volumes of the \( \text{p97}^{\text{R155H}}_{|\text{ADP-p47}} \) dodecamer and \( \text{p97}^{\text{R155H}}_{|\text{ATP_\gammaS-p47}} \), respectively. The two volumes were refined against their experimental particle images and reached resolutions of 6.10 Å (\( \text{p97}^{\text{R155H}}_{|\text{ADP-p47}} \) dodecamer) and 4.50 Å (\( \text{p97}^{\text{R155H}}_{|\text{ATP_\gammaS-p47}} \)). Like the \( \text{p97}^{\text{R155H}}_{|\text{ADP-p47}} \) dataset, further 3D classification did not yield discernible features in the local regions for NTDs and p47. Processing schematic and local resolution estimation are shown in Figure S5. For the \( \text{p97}^{\text{R155H}}_{|\text{ATP_\gammaS-p47}} \) dataset, we reconstruct the volume following the same procedure as above. Final 63,353 particle images were used to reconstruct a 3D volume at 4.23 Å resolution, sharpened using a \( b \)-factor of \(-77.6 \text{ Å}^2\). Processing schematic and local resolution estimation are shown in Figure S6.

Structural heterogeneity of the three \( \text{p97}^{\text{R155H}}_{|\text{ADP-p47}} \) datasets was further analyzed using a deep coordinate neural network by the cryoDRGN program (version 0.3) [58]. Before latent encoding, the particle images were Fourier-cropped to a box size corresponding to pixel sizes of 3.90 and 3.71 Å/pixel for \( \text{p97}^{\text{R155H}}_{|\text{ADP-p47}} \) and \( \text{p97}^{\text{R155H}}_{|\text{ADP-p47}} \), respectively. For each group, an eight-dimensional latent variable model was trained for 25 epochs. The encoder and decoder architectures were three layers with 1024 nodes. The latent encodings were visualized using a principal component factor plot or UMAP (uniform manifold approximation and projection) representation [89]. After training, \( k \)-means clustering with \( k = 20 \) was conducted on the latent encodings, and reconstructions were calculated at the cluster centers using the decoder network.

4.8. Modeling

The previous p97 and p47 coordinates (PDB code: 5FTK and 1S3S) were used as the templates for atomic modeling or molecular docking [24,37]. The templates were first docked into individual cryo-EM densities using the ‘Fit in the Map’ function in UCSF Chimera (version 1.14) [90]. The fitted model was manually rebuilt using Coot (version 0.9.1) [91] and then refined against the cryo-EM densities using the ‘phenix.real_space_refine’ program in Phenix software package (version 1.18.2-3874) [92]. Hydrogen atoms were added using the ‘phenix.reduce’ program for the model refinement and removed after the refinement. Secondary structure restraints were applied during the model refinement. The refinement and validation statistics were listed in Table S3. The figures for the cryo-EM density maps and atomic models were prepared using UCSF Chimera and ChimeraX (version 0.91) [93].

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/ijms22158079/s1. Figure S1: Cryo-EM analysis of the full-length \( \text{p97}^{\text{R155H}}_{|\text{p47}} \) assembly in the absence of nucleotides. Figure S2: Modelling of atomic coordinates in the cryo-EM densities of the \( \text{p97}^{\text{R155H}}_{|\text{dodecamer}} \) and \( \text{p97}^{\text{R155H}}_{|\text{p47}} \) complexes in the absence of nucleotides. Figure S3: Single-particle EM image analysis on the negatively stained Walker A mutants in the context of \( \text{p97}^{\text{R155H}}_{|\text{p47}} \).
Figure S4: Biochemical characterization of the p97R155H|ADP-p47 and p97R155H|ATPγS-p47. Figure S5: Cryo-EM structural analysis of the full-length p97R155H|ADP-p47 assembly. Figure S6: Cryo-EM structural analysis of the full-length p97R155H|ATPγS-p47 assembly. Figure S7: Modelling of atomic coordinates into the cryo-EM densities of p97R155H|ATPγS-p47 complexes. Figure S8: Cryo-EM structure of the p97R155H|ADP-p47 complex. Figure S9: Deep coordinate neural network analysis of the p97R155H|ADP-p47 single-particle cryo-EM images. Figure S10: Western blotting analyses of the SEC-eluted fractions of the p97-p47 complexes. Figure S11: Determination of the binding constant, $K_d$, for the interaction between p97 arginine finger mutants and p47 proteins.

Table S1: List of plasmids used for overexpression of p97 and p47 proteins. Table S2: Statistics of single-particle cryo-EM structure determination on p97-p47 assembly.

Author Contributions: Conceptualization, T.-F.C. and P.-L.C.; investigation, formal analysis, and data curation, P.N., S.L., R.C.A.C., F.W., Y.-P.P., D.R.W. and P.-L.C.; supervision and validation, T.-F.C. and P.-L.C.; writing, P.N., T.-F.C. and P.-L.C. All authors have read and agreed to the published version of the manuscript.

Funding: We acknowledge using the Titan Krios TEM in the Eyring Materials Center (EMC) at Arizona State University (ASU) and funding for the instrumentation from grant number NSF MRI 1531991. We thank the NVIDIA GPU Grant Program to P.-L.C for GPU device support. The project was partially supported by a grant from the National Institute of Neurological Disorders and Stroke (NINDS) (R01NS102279) to T.-F.C. and the ASU startup fund to P.-L.C.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Cryo-EM density maps have been deposited in the Electron Microscopy Data Bank (EMDB) under accession numbers EMD-24302 (p97R155H|ATPγS-p47), EMD-24305 (p97R155H|dodecamer), EMD-23191 (p97R155H|ADP-p47), and EMD-23192 (p97R155H|dodecamer II). Model coordinates were deposited in the Protein Data Bank (PDB) under accession numbers 7R7S (p97R155H|ATPγS-p47), 7R7U (p97R155H|p47), 7L5W (p97R155H|dodecamer), and 7R7T (p97R155H|ADP-p47). All data are available from the corresponding authors upon request.

Acknowledgments: We thank Ellen Zhong (Massachusetts Institute of Technology) for the brief discussion on the data analysis using cryoDRGN neural networks. We thank Xiaoyi Zhang, Gui Lin, and Daniel Wong for preliminary ATPase activity determination for p97 mutants. We thank Andrey Malyutin for assisting in sample screening at the Beckman Caltech Cryo-EM Center.

Conflicts of Interest: The authors declare no competing interests.

References

1. Iyer, L.M.; Leipe, D.D.; Koonin, E.V.; Aravind, L. Evolutionary history and higher order classification of AAA+ ATPases. J. Struct. Biol. 2004, 146, 11–31. [CrossRef]
2. Xia, D.; Tang, W.K.; Ye, Y. Structure and function of the AAA+ ATPase p97/Cdc48p. Gene 2016, 583, 64–77. [CrossRef]
3. Wendler, P.; Ciniaewsky, S.; Kock, M.; Kube, S. Structure and function of the AAA+ nucleotide binding pocket. Biochim. Biophys. Acta 2012, 1823, 2–14. [CrossRef] [PubMed]
4. Meyer, H.H.; Shorter, J.G.; Seemann, J.; Pappin, D.; Warren, G. A complex of mammalian ufd1 and npl4 links the AAA-ATPase, p97, to ubiquitin and nuclear transport pathways. EMBO J. 2000, 19, 2181–2192. [CrossRef] [PubMed]
5. Hetzer, M.; Meyer, H.H.; Walther, T.C.; Bilbao-Cortes, D.; Warren, G.; Mattaj, I.W. Distinct AAA-ATPase p97 complexes function in discrete steps of nuclear assembly. Nat. Cell Biol. 2001, 3, 1086–1091. [CrossRef] [PubMed]
6. Snider, J.; Thibault, G.; Houry, W.A. The AAA+ superfamily of functionally diverse proteins. Genome Biol. 2008, 9, 216. [CrossRef]
7. Kondo, H.; Rabouille, C.; Newman, R.; Levine, T.P.; Pappin, D.; Freemont, P.; Warren, G. p47 is a cofactor for p97-mediated membrane fusion. Nature 1997, 388, 75–78. [CrossRef]
8. Koller, K.J.; Brownstein, M.J. Use of a cDNA clone to identify a supposed precursor protein containing valosin. Nature 1987, 325, 542–545. [CrossRef]
9. Peters, J.M.; Walsh, M.J.; Franke, W.W. An abundant and ubiquitous homo-oligomeric ring-shaped ATPase particle related to the putative vesicle fusion proteins Sec18p and NSF. EMBO J. 1990, 9, 1757–1767. [CrossRef]
10. Buchberger, A.; Schindelin, H.; Hänelzmann, P. Control of p97 function by cofactor binding. FEBS Lett. 2015, 589, 2578–2589. [CrossRef]
11. Meyer, H.; Bug, M.; Bremer, S. Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system. Nat. Cell Biol. 2012, 14, 117–123. [CrossRef] [PubMed]
12. Stolz, A.; Hilt, W.; Buchberger, A.; Wolf, D.H. Cdc48: A power machine in protein degradation. Trends Biochem. Sci. 2011, 36, 515–523. [CrossRef]

13. Yeung, H.O.; Kloppsteck, P.; Niwa, H.; Isacson, R.L.; Matthews, S.; Zhang, X.; Freemont, P.S. Insights into adaptor binding to the AAA protein p97. Biochem. Soc. Trans. 2008, 36, 62–67. [CrossRef] [PubMed]

14. Abramzon, Y.; Johnson, J.O.; Scholz, S.W.; Taylor, J.P.; Brunetti, M.; Calvo, A.; Mandrioli, J.; Benatar, M.; Mora, G.; Restagno, G.; et al. Valosin-containing protein (VCP) mutations in sporadic amyotrophic lateral sclerosis. Neurobiol. Aging 2012, 33, 2231.e1–2231.e6. [CrossRef]

15. Chapman, E.; Fry, A.N.; Kang, M. The complexities of p97 function in health and disease. Mol. Biosyst. 2011, 7, 700–710. [CrossRef] [PubMed]

16. Watts, G.D.J.; Wymer, J.; Kovach, M.J.; Mehta, S.G.; Mumm, S.; Darvish, D.; Pestrhonk, A.; Whyte, M.P.; Kimonis, V.E. Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. Nat. Genet. 2004, 36, 377–381. [CrossRef]

17. Johnson, J.O.; Mandrioli, J.; Benatar, M.; Abramzon, Y.; Van Deerlin, V.M.; Trojanowski, J.Q.; Gibbs, J.R.; Brunetti, M.; Gronka, S.; Wuu, J.; et al. Exome sequencing reveals VCP mutations as a cause of familial ALS. Neuron 2010, 68, 857–864. [CrossRef]

18. Ogura, T.; Wilkinson, A.J. AAA+ superfamily ATPases: Common structure–Diverse function. Genes Cells 2001, 6, 575–597. [CrossRef]

19. Niwa, H.; Ewens, C.A.; Tsang, C.; Yeung, H.O.; Zhang, X.; Freemont, P.S. The role of the N-domain in the ATPase activity of the mammalian AAA ATPase p97/VCP. J. Biol. Chem. 2002, 277, 8561–8570. [CrossRef]

20. Nishikori, S.; Esaki, M.; Yamanaka, K.; Sugimoto, S.; Ogura, T. Positive cooperativity of the p97 AAA ATPase is critical for essential functions. J. Biol. Chem. 2011, 286, 15815–15820. [CrossRef]

21. Ye, Y.; Meyer, H.H.; Rapoport, T.A. Function of the p97-Ufd1-Npl4 complex in retrotranslocation from the ER to the cytosol: Dual recognition of nonubiquitinated polypeptide segments and polyubiquitin chains. J. Cell Biol. 2003, 162, 71–84. [CrossRef]

22. Davies, J.M.; Tsuruta, H.; May, A.P.; Weis, W.I. Conformational changes of p97 during nucleotide hydrolysis determined by small-angle X-Ray scattering. Structure 2005, 13, 183–195. [CrossRef] [PubMed]

23. Torrecilla, I.; Oehler, J.; Ramadan, K. The role of ubiquitin-dependent segregase p97 (VCP or Cdc48) in chromatin dynamics after DNA double strand breaks. Philos. Trans. R. Soc. Lond. B Biol. Sci. 2017, 372, 20160282. [CrossRef]

24. van den Boom, J.; Meyer, H.; VCP/p97-Mediated Unfolding as a Principle in Protein Homeostasis and Signaling. Mol. Cell 2018, 69, 182–194. [CrossRef]

25. Stolz, A.; Freemont, P.S. The AAA+ ATPase p97, a cellular multilool. Biochem. J. 2017, 474, 2953–2976. [CrossRef] [PubMed]

26. Cooney, I.; Han, H.; Stewart, M.G.; Carson, R.H.; Hansen, D.T.; Iwasa, J.H.; Price, J.C.; Hill, C.P.; Shen, P.S. Structure of the Cdc48 ATPase complex is initiated by ubiquitin unfolding. EMBO J. 2004, 23, 1030–1039. [CrossRef]

27. Cdc48 ATPase complex is initiated by ubiquitin unfolding. EMBO J. 2004, 23, 1030–1039. [CrossRef]

28. Huang, R.; Ripstein, Z.A.; Rubinstein, J.L.; Kay, L.E. Cooperative subunit dynamics modulate p97 function. Proc. Natl. Acad. Sci. USA 2019, 116, 158–167. [CrossRef]

29. Gates, S.N.; Martin, A. Stairway to translocation: AAA+ motor structures reveal the mechanisms of ATP-dependent substrate translocation. Protein Sci. 2020, 29, 407–419. [CrossRef] [PubMed]

30. Hänzelmann, P.; Schindelin, H. The Interplay of Cofactor Interactions and Post-translational Modifications in the Regulation of the AAA+ ATPase p97. Front. Mol. Biosci. 2017, 4, 21. [CrossRef]

31. Banerjee, S.; Bartesaghi, A.; Merk, A.; Rao, P.; Bulfer, S.L.; Yan, Y.; Green, N.; Mroczkowski, B.; Neitz, R.J.; Wipf, P.; et al. 2.3 Å resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. Science 2016, 351, 871–875. [CrossRef] [PubMed]

32. 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 2019, 27, 1820–1829.e4. [CrossRef]

33. 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 2019, 27, 1820–1829.e4. [CrossRef]
40. Beuron, F.; Dreveny, I.; Yuan, X.; Pye, V.E.; McKeown, C.; Briggs, L.C.; Cliff, M.J.; Kaneko, Y.; Wallis, R.; Isaacson, R.L.; et al. Conformational changes in the AAA ATPase p97-p47 adapter complex. EMBO J. 2006, 25, 1967–1976. [CrossRef]

41. Huang, C.; Li, G.; Lennarz, W.J. Dynamic flexibility of the ATPase p97 is important for its interprotomer motion transmission. Proc. Natl. Acad. Sci. USA 2012, 109, 9792–9797. [CrossRef] [PubMed]

42. Halawani, D.; LeBlanc, A.C.; Rouiller, I.; Michnick, S.W.; Servant, M.J.; Latterich, M. Hereditary inclusion body myopathy-linked p97/VCP mutations in the NH2 domain and the D1 ring modulate p97/VCP ATPase activity and D2 ring conformation. Mol. Cell. Biol. 2009, 29, 4484–4494. [CrossRef] [PubMed]

43. Saracino, D.; Clot, F.; Camuzat, A.; Anquetil, V.; Hannequin, D.; Guyant-Maréchal, L.; Didic, M.; Guillot-Noël, L.; Rinaldi, D.; Latouche, M., et al. Novel VCP mutations expand the mutational spectrum of frontotemporal dementia. Neurobiol. Aging 2018, 72, e11–e187. [CrossRef]

44. Buchberger, A.; Howard, M.J.; Proctor, M.; Bycroft, M. The UBX domain: A widespread ubiquitin-like module. J. Mol. Biol. 2001, 307, 17–24. [CrossRef] [PubMed]

45. Zhang, X.; Gu, L.; Zhang, X.; Bulfer, S.L.; Sanghez, V.; Wong, D.E.; Lee, Y.; Lehmann, L.; Lee, J.S.; Shih, P.-Y.; et al. Altered cofactor regulation with disease-associated p97/VCP mutations. Proc. Natl. Acad. Sci. USA 2015, 112, E1705–E1714. [CrossRef]

46. Schuberth, C.; Buchberger, A. UBX domain proteins: Major regulators of the AAA ATPase Cdc48/p97. Cell. Mol. Life Sci. 2008, 65, 2360–2371. [CrossRef]

47. Esbenshade, C.A.; Drokow, S.; Klopsteck, P.; McKeown, C.; Ewens, J.; Robinson, C.; Zhang, X.; Freemont, P.S. The p97-FAF1 protein complex reveals a common mode of adaptor binding. J. Biol. Chem. 2014, 289, 12077–12084. [CrossRef]

48. Arumughan, A.; Roske, Y.; Barth, C.; Forero, L.L.; Bravo-Rodriguez, K.; Redel, A.; Kostova, S.; McShane, E.; Opitz, R.; Faebler, K.; et al. Quantitative interaction mapping reveals an extended UBX domain in ASPL that disrupts functional p97 hexamers. Nat. Commun. 2016, 7, 13047. [CrossRef]

49. Kim, K.H.; Kang, W.; Suh, S.W.; Yang, J.K. Crystal structure of FAF1 UBX domain in complex with p97/VCP N domain reveals a conformational change in the conserved FcisP touch-turn motif of UBX domain. Proteins 2011, 79, 2583–2587. [CrossRef]

50. Kim, S.J.; Cho, J.; Song, E.J.; Kim, S.J.; Kim, H.M.; Lee, K.E.; Suh, S.W.; Kim, E.E. Structural basis for ovarian tumor domain-containing protein 1 (OTU1) binding to p97/valosin-containing protein (VCP). J. Biol. Chem. 2014, 289, 12264–12274. [CrossRef] [PubMed]

51. Conicella, A.E.; Huang, R.; Ripstein, Z.A.; Nguyen, A.; Wang, E.; Löhr, T.; Schuck, P.; Vendruscolo, M.; Rubinstein, J.L.; Kay, L.E. A dynamic molecular basis for malfunction in disease mutants of p97/VCP. eLife 2016, 5, e20143. [CrossRef] [PubMed]

52. Schuetz, A.K.; Kay, L.E. A dynamic molecular basis for malfunction in disease mutants of p97/VCP. eLife 2016, 5, e20143. [CrossRef] [PubMed]

53. Tang, W.K.; Odzorgi, T.; Jin, W.; Xia, D. Structural Basis of p97 Inhibition by the Site-Selective Anticancer Compound CB-5083. ACS Chem. Biol. 2016, 11, 2112–2116. [CrossRef] [PubMed]

54. Walker, J.E.; Saraste, M.; Runswick, M.J.; Gay, N.J. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. 1982, 1, 945–951. [CrossRef] [PubMed]

55. Li, G.; Zhao, G.; Schindelin, H.; Lennarz, W.J. Tyrosine phosphorylation of ATPase p97 regulates its activity during ERAD. Biochem. Biophys. Res. Commun. 2008, 375, 247–251. [CrossRef] [PubMed]
65.  
Ewens, C.A.; Kloppsteck, P.; Förster, A.; Zhang, X.; Freemont, P.S. Structural and functional implications of phosphorylation and acetylation in the regulation of the AAA+ protein p97. Biochem. Cell Biol. 2010, 88, 41-48. [CrossRef] [PubMed]

66.  
Manno, A.; Noguchi, M.; Fukushima, J.; Motohashi, Y.; Kakizuka, A. Enhanced ATPase activities as a primary defect of mutant valosin-containing proteins that cause inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia. Genes Cells 2010, 15, 911-922. [CrossRef]

67.  
Trusch, F.; Matena, A.; Vuk, M.; Koerver, J.; Knaedelsrud, H.; Freemont, P.S.; Meyer, H.; Bayer, P. The N-terminal Region of the Ubiquitin Regulatory X (UBX) Domain-containing Protein 1 (UBXDI) Modulates Interdomain Communication within the Valosin-containing Protein p97. J. Biol. Chem. 2015, 290, 29414-29427. [CrossRef]

68.  
Guo, P.; Driver, D.; Zhao, Z.; Zheng, Z.; Chan, C.; Cheng, X. Controlling the Revolving and Rotating Motion Direction of Asymmetric Hexameric Nanomotor by Arginine Finger and Channel Chirality. ACS Nano 2019, 13, 6207–6223. [CrossRef]

69.  
Ogura, T.; Whiteheart, S.W.; Wilkinson, A.J. Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and AAA+ ATPases. J. Struct. Biol. 2004, 146, 106–112. [CrossRef]

70.  
Wang, Q.; Song, C.; Irizarry, L.; Dai, R.; Zhang, X.; Li, C.-C.H. Multifunctional roles of the conserved Arg residues in the second region of homology of p97/valosin-containing protein. J. Biol. Chem. 2005, 280, 40515–40523. [CrossRef]

71.  
Blythe, E.E.; Olson, K.C.; Chau, V.; Deshaies, R.J. Ubiquitin- and ATP-dependent unfoldase activity of P97/VCP

72.  
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 2017, 6. [CrossRef]

73.  
Uchiyama, K.; Jokitalo, E.; Kano, F.; Murata, M.; Zhang, X.; Canas, B.; Newman, R.; Rabouille, C.; Pappin, D.; Freemont, P.S.; et al. The N-terminal Region of VCIP135, a novel essential factor for p97/p47-mediated membrane fusion, is required for Golgi and ER assembly in vivo. J. Cell Biol. 2002, 159, 855–866. [CrossRef]

74.  
Yuan, X.; Shaw, A.; Zhang, X.; Kondo, H.; Lally, J.; Freemont, P.S.; Matthews, S. Solution structure and interaction surface of the C-terminal domain from p47: A major p97-colator involved in SNARE disassembly. J. Mol. Biol. 2001, 311, 255–263. [CrossRef]

75.  
Stach, L.; Morgan, R.M.; Makhlouf, L.; Douangamath, A.; von Delft, F.; Zhang, X.; Freemont, P.S. Crystal structure of the catalytic D2 domain of the AAA+ ATPase p97 reveals a putative helical split-washer-type mechanism for substrate unfolding. FEBS Lett. 2020, 594, 933–943. [CrossRef] [PubMed]

76.  
Sato, Y.; Tsuchiya, H.; Yamagata, A.; Okatsu, K.; Tanaka, K.; Saeki, Y.; Fukai, S. Structural insights into ubiquitin recognition and Ufd1 interaction of Npl4. Nat. Commun. 2019, 10, 5708. [CrossRef]

77.  
Uchiyama, K.; Jokitalo, E.; Kano, F.; Murata, M.; Zhang, X.; Canas, B.; Newman, R.; Rabouille, C.; Pappin, D.; Freemont, P.; et al. VCIP135, a novel essential factor for p97/p47-mediated membrane fusion, is required for Golgi and ER assembly in vivo. J. Cell Biol. 2002, 159, 855–866. [CrossRef]

78.  
Shukla, A.K.; Westfield, G.H.; Xiao, K.; Reis, R.I.; Huang, L.-Y.; Tripathi-Shukla, P.; Qian, J.; Li, S.; Blanc, A.; Oleksie, A.N.; et al. Visualization of arrestin recruitment by a G-protein-coupled receptor. Nature 2014, 512, 218–222. [CrossRef]

79.  
Oh, M.; Li, Y.; Cheng, Y.; Walz, T. Negative Staining and Image Classification—Powerful Tools in Modern Electron Microscopy. Biol. Proced. Online 2004, 6, 23–34. [CrossRef] [PubMed]

80.  
Schers, S.H.W. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. J. Struct. Biol. 2012, 180, 519–530. [CrossRef] [PubMed]

81.  
Cheng, A.; Eng, E.T.; Alink, L.; Rice, W.J.; Jordan, K.D.; Kim, L.Y.; Potter, C.S.; Carragher, B. High resolution single particle cryo-electron microscopy using beam-image shift. J. Struct. Biol. 2018, 204, 270–275. [CrossRef] [PubMed]

82.  
Mastronarde, D.N. Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. 2005, 152, 36–51. [CrossRef] [PubMed]

83.  
Rohou, A.; Grigorieff, N. CTFIND4: Fast and accurate defocus estimation from electron micrographs. J. Struct. Biol. 2015, 192, 216–221. [CrossRef]

84.  
Punjani, A.; Rubinstein, J.L.; Fleet, D.J.; Brubaker, M.A. cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods 2017, 14, 290–296. [CrossRef]

85.  
Zivanov, J.; Nakane, T.; Schers, S.H.W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. IUCrJ 2020, 7, 253–267. [CrossRef]

86.  
Zivanov, J.; Nakane, T.; Schers, S.H.W. A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. IUCrJ 2019, 6, 5–17. [CrossRef]

87.  
Schers, S.H.W.; Chen, S. Prevention of overfitting in cryo-EM structure determination. Nat. Methods 2012, 9, 853–854. [CrossRef] [PubMed]

88.  
Bepler, T.; Morin, A.; Rapp, M.; Brasch, J.; Shapiro, L.; Noble, A.J.; Berger, B. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. Nat. Methods 2019, 16, 1153–1160. [CrossRef] [PubMed]

89.  
McInnes, L.; Healy, J.; Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv 2018, arXiv:1802.03426.

90.  
Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—a visualization system for exploratory research and analysis. J. Comput. Chem. 2004, 25, 1605–1612. [CrossRef]

91.  
Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 2010, 66, 486–501. [CrossRef] [PubMed]
92. Adams, P.D.; Afonine, P.V.; Bunkóczi, G.; Chen, V.B.; Davis, I.W.; Echols, N.; Headd, J.J.; Hung, L.-W.; Kapral, G.J.; Grosse-Kunstleve, R.W.; et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 2010, 66, 213–221. [CrossRef] [PubMed]

93. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Meng, E.C.; Couch, G.S.; Croll, T.I.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* 2021, 30, 70–82. [CrossRef] [PubMed]