



Structural insight into the ZFAND1–p97 interaction involved in stress granule clearance

Received for publication, December 22, 2023, and in revised form, March 18, 2024. Published, Papers in Press, March 25, 2024.
<https://doi.org/10.1016/j.jbc.2024.107230>

Chih-Hsuan Lai^{1,‡}, Kuang-Ting Ko^{1,‡}, Pei-Ju Fan², Tsun-Ai Yu², Chi-Fon Chang², Piotr Draczkowski¹ , and Shang-Te Danny Hsu^{1,3,4,*} 

From the ¹Institute of Biological Chemistry, and ²High-Field Nuclear Magnetic Resonance Center, Academia Sinica, Taipei, Taiwan; ³Institute of Biochemical Sciences, National Taiwan University, Taipei, Taiwan; ⁴International Institute for Sustainability With Knotted Chiral Meta Matter (SKCM2), Hiroshima University, Higashihiroshima, Japan

Reviewed by members of the JBC Editorial Board. Edited by Wolfgang Peti

Arsenite-induced stress granule (SG) formation can be cleared by the ubiquitin-proteasome system aided by the ATP-dependent unfoldase p97. ZFAND1 participates in this pathway by recruiting p97 to trigger SG clearance. ZFAND1 contains two An1-type zinc finger domains (ZF1 and ZF2), followed by a ubiquitin-like domain (UBL); but their structures are not experimentally determined. To shed light on the structural basis of the ZFAND1–p97 interaction, we determined the atomic structures of the individual domains of ZFAND1 by solution-state NMR spectroscopy and X-ray crystallography. We further characterized the interaction between ZFAND1 and p97 by methyl NMR spectroscopy and cryo-EM. ¹⁵N spin relaxation dynamics analysis indicated independent domain motions for ZF1, ZF2, and UBL. The crystal structure and NMR structure of UBL showed a conserved β -grasp fold homologous to ubiquitin and other UBLs. Nevertheless, the UBL of ZFAND1 contains an additional N-terminal helix that adopts different conformations in the crystalline and solution states. ZFAND1 uses the C-terminal UBL to bind to p97, evidenced by the pronounced line-broadening of the UBL domain during the p97 titration monitored by methyl NMR spectroscopy. ZFAND1 binding induces pronounced conformational heterogeneity in the N-terminal domain of p97, leading to a partial loss of the cryo-EM density of the N-terminal domain of p97. In conclusion, this work paved the way for a better understanding of the interplay between p97 and ZFAND1 in the context of SG clearance.

Stress granules (SGs) are translationally inactive messenger ribonucleoprotein complexes formed in the cytoplasm of eukaryotic cells under various environmental stresses (1–4). SGs are dynamic and reversible structures. SG assembly helps protect cells from acute environmental stresses; SGs are disassembled upon the disappearance of stress factors (2–4). The disassembly of SGs is crucial for preventing the formation of pathological aggregates (4, 5). Such a process is facilitated by the p97 activation (4, 6–8). p97, also known as a valosin-

containing protein (VCP), acts as a segregase in the ubiquitin-proteasome system to trigger proteolytic degradation for the clearance of irreparably damaged proteins (9). Mutations in p97 are implicated in the accumulation of SGs in neurodegenerative diseases, such as amyotrophic lateral sclerosis (1, 10). Over-expression of disease-related p97 mutants in mammalian cells results in the accumulation of SGs (1). Recruitment of p97 to SGs is observed under various stress conditions, such as exposure to arsenite and heat shock (6, 11). The subcellular localization of p97 depends on how p97 interacts with diverse cofactors (12). In the case of arsenic stress, the environmental carcinogen induces protein misfolding, protein aggregation, and SG formation in eukaryotic cells (13). p97 is recruited to arsenite-induced SGs and endoplasmic reticulum by three cofactors, An1-type zinc finger protein 1 (ZFAND1), ZFAND2A (also known as AIRAP), and ZFAND2B (also known as AIRAPL) (6, 14, 15). ZFAND1, in particular, is a key contributing factor for the clearance of arsenite-induced SGs (6). In addition, a nonsense mutation in ZFAND1, R130*, is implicated in ovarian carcinoma and breast cancer, but the underlying mechanism remains unclear (6, 16, 17).

The functional implication of ZFAND1 in arsenite-induced SG clearance has been identified based on the earlier studies of its yeast homolog, Cuz1, essential for cell survival in an arsenite-exposure environment (18, 19). Similar to ZFAND1, Cuz1 has an AN1-like zinc finger (ZF) domain at the N terminus and a ubiquitin-like (UBL) domain at the C terminus (Fig. 1A). However, Cuz1 lacks the second ZF present in ZFAND1. Instead, the N-terminal ZF is connected to the C-terminal UBL through a long linker (residues 59–169) (20). Cuz1 is found to interact with the 26S proteasome and Cdc48 (a homologous protein of p97) (19). ZFAND2A and ZFAND2B are human paralogs of ZFAND1 (Fig. 1A). They share the first two ZFs at the N termini and a UBZ-like ZF between the two ZFs. Additionally, ZFAND2B has two ubiquitin-interacting motifs at the C terminus (6, 9, 14). The deletion of one of the ZFs of ZFAND2A results in the loss of binding to the 19S regulatory particle of the proteasome (15). The C-terminal UBL of Cuz1 is used to bind to Cdc48 according to a pull-down assay (19). In contrast, ZFAND1 UBL shows a

[‡] These authors contributed equally to this work.

* For correspondence: Shang-Te Danny Hsu, sthsu@gate.sinica.edu.tw.

Structural basis of ZFAND1-p97 interaction

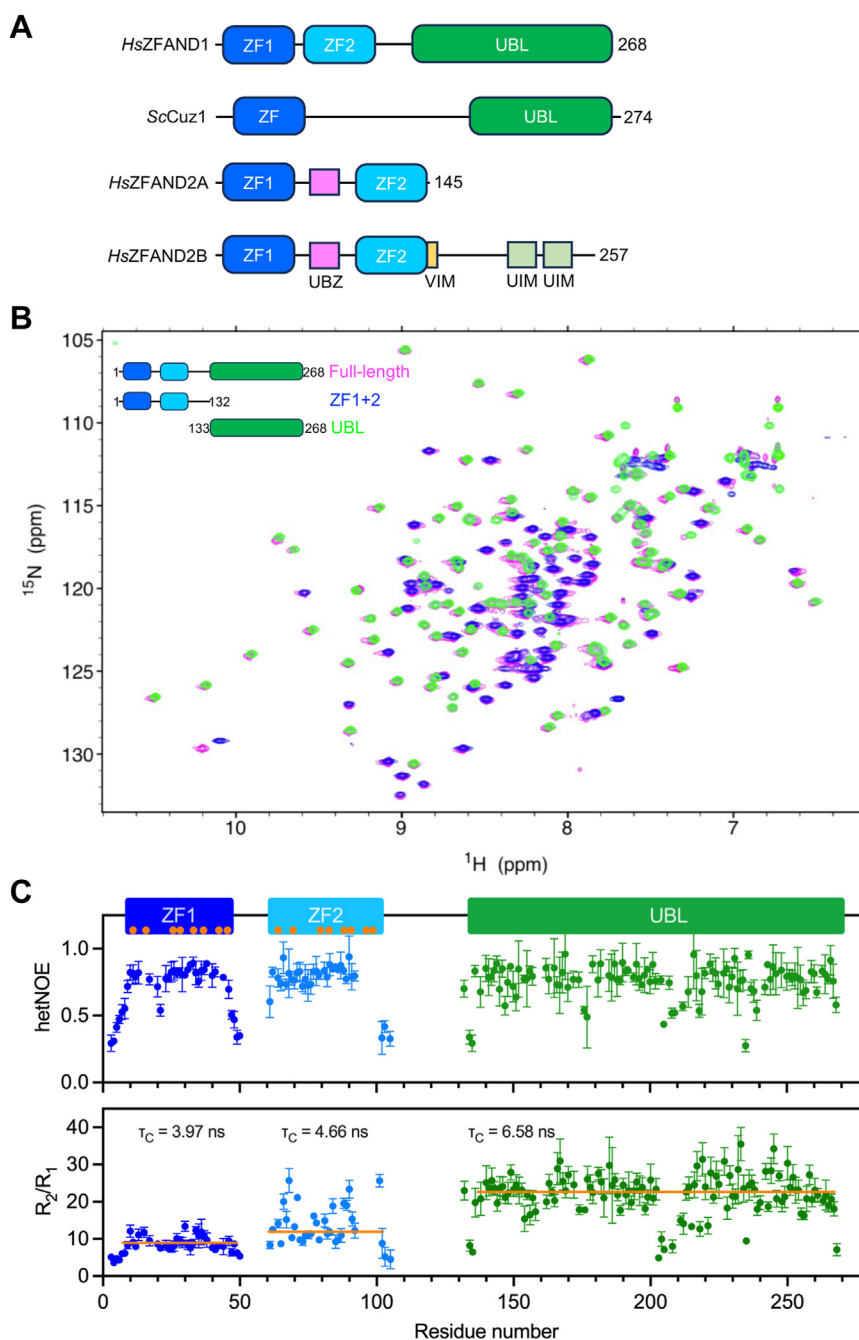


Figure 1. Folding and dynamics of individual domains of ZFAND1. A, domain structures for *Homo sapiens* ZFAND1 (HsZFAND1), *Saccharomyces cerevisiae* Cuz1 (ScCuz1), HsZFAND2A, and HsZFAND2B. The An1-type zinc finger domains (ZF1, ZF2, or ZF) are shown as *marine* and *cyan*. The ubiquitin-like domain (UBL) is shown in *green*. The UBZ-like zinc finger (UBZ), VCP-interacting motif (VIM), and ubiquitin-interacting motif (UIM) are shown in *violet*, *light yellow*, and *light green*, respectively. The lengths of the individual proteins are indicated at the C-termini. B, overlay of the ¹⁵N-¹H HSQC spectra of full-length ZFAND1 (*magenta*), ZF1+2 (residues 1–132; *blue*), and UBL (residues 133–268; *green*) showed minimal spectral perturbations upon domain truncations, indicating that the individual domains fold autonomously without significant interdomain interactions. C, ¹⁵N spin relaxation dynamics analysis of full-length ZFAND1. The {¹H}-¹⁵N heteronuclear NOE (hetNOE) and the ratio of ¹⁵N R₂/R₁ as a function of residue number showed well-folded structures of the individual domains and independent domain dynamics, respectively. The hetNOE and R₂/R₁ ratios of ZF1, ZF2, and UBL are colored *blue*, *cyan*, and *green*, respectively, according to the domain schematics above. The mean values of R₂/R₁ ratios in each domain are shown as *horizontal orange lines* to derive the τ_C values as indicated above. The *orange dots* indicate the positions of the C3H1 residues involved in zinc ion coordination. All the NMR spectra were recorded at a proton Larmor frequency of 850 MHz at 300 K. UBL, ubiquitin-like domain; VCP, valosin-containing protein; ZF, zinc finger.

transient interaction with p97. Note, however, that the physical interaction between ZFAND1 and p97 is only observed when an E578Q mutation is introduced to p97 to abrogate its ATPase activity or by chemical cross-linking to covalently link ZFAND1 and p97 (6). Despite the transient nature of

ZFAND1-p97 interaction, truncation of UBL by a nonsense mutation at R130 (R130*) reduces ZFAND1-p97 interaction to a background-level (6).

ZFAND1 is highly conserved among vertebrates (Fig. S1A), but its UBL shares limited sequence similarity with other UBLs

and ubiquitin regulator X (UBX) (Fig. S1B). Emerging structural information of p97–Cdc48 interaction with their adaptor proteins has revealed a diverse interaction mode (*cf.*, PDB IDs 6CHS, 6OA9, 7LMY, 7R7S, 8B5R, 8DAT, 8FCL) (21–27). However, the lack of structural information on ZFAND1 has hindered a detailed understanding of how ZFAND1 may help recruit ubiquitinated substrates to p97 and the 26S proteasome for arsenite-induced SG clearance. To address this issue, we employed solution-state NMR spectroscopy and X-ray crystallography to characterize the structures and dynamics of the individual domains of ZFAND1 and full-length ZFAND1. We also used methyl NMR spectroscopy to probe the interaction between full-length ZFAND1 and p97, which has an apparent molecular weight of 0.6 MDa as a homohexamer. Additionally, we used cryo-EM single particle analysis (SPA) to investigate conformational changes in p97 upon binding to ZFAND1. The integrated structural analyses revealed the transient and highly dynamic nature of ZFAND1–p97 interaction mediated by the UBL of ZFAND1, which triggers significant conformational changes and increased dynamics of the N-terminal domain (NTD) of p97.

Results

Folding dynamics of ZFAND1

HsZFAND1 comprises of three domains, namely ZF1, ZF2, and UBL (Fig. 1). The fingerprint ^{15}N - ^1H HSQC spectra of different combinations of domains exhibited well-dispersed cross peaks, indicating well-folded structures. Importantly, most of the cross peaks can be superimposed with each other, suggesting that the individual domains fold autonomously without appreciable interdomain interactions (Fig. 1B). Indeed, complete backbone resonance assignments showed that the affected cross peaks corresponded to those that are located at the domain boundaries, and the peak shifts were caused by the domain truncations. We next carried out ^{15}N spin relaxation dynamics experiments—longitudinal relaxation (R_1), transverse relaxation (R_2), and $\{^1\text{H}\}$ - ^{15}N heteronuclear NOE (hetNOE) experiments—of full-length ZFAND1 to evaluate the backbone dynamics of the individual domains. While the hetNOE values in the folded regions were similarly high (>0.7) for ZF1, ZF2, and UBL, indicating limited fast backbone motions on the ps–ns timescale, the R_2/R_1 ratios were distinctly different for the three domains (Fig. 1C). The N-terminal ZF1 exhibited the lowest overall R_2/R_1 ratio, while the C-terminal UBL exhibited the highest overall R_2/R_1 ratio. The rotational correlation times (τ_C) of ZF1, ZF2, and UBL were 3.97, 4.66, and 6.58 ns, respectively (Fig. 1C). The results indicated that the three domains tumble independently in solution, with UBL showing the slowest molecular tumbling rate as expected from the largest domain size (136 residues in length, whereas the folded regions of both ZF1 and ZF2 are about 40 residues in length). Compared to the uniform R_2/R_1 ratio of ZF1, ZF2 exhibited elevated values in some parts of the structure, which may be attributed to conformation changes that lead to enhanced transverse relaxation. Indeed, the backbone amide assignments for residues 93 to 100 were

missing due to the lack of corresponding ^{15}N - ^1H cross peaks. These residues are involved in zinc ion coordination, and they may exhibit unfavorable dynamics that lead to severe line-broadening beyond detection (see below). In the case of UBL, residues 203 to 205 and 208 exhibited significantly lower R_2/R_1 ratios than the rest of the UBL domain. These residues are located in a loop region that may exhibit additional motions leading to reduced R_2 relaxation. Collectively, the ^{15}N spin relaxation analysis and the spectral comparison of the individual domains indicated that the three domains of ZFAND1 fold and tumble independently like beads in a string.

Solution structures of ZF1, ZF2, and UBL

We prepared [^{15}N / ^{13}C] ZF1, ZF2, and UBL to determine their solution structures by NMR spectroscopy (Fig. 2 and Experimental procedures). All three domains exhibited well-dispersed cross peaks in their respective ^{15}N - ^1H HSQC spectra (Fig. 2, A–C). The isolation of ZF1 and ZF2 had minimal structural perturbation compared to the full-length ZFAND1 spectrum evidenced by the minimal chemical shift perturbations of the ^{15}N - ^1H HSQC spectra of the isolated ZF1 and ZF2 compared to that of full-length ZFAND1 and ZF1+ZF2 (Fig. S2). For ZF1 and ZF2, in particular, the chemical shifts of the C β atoms of the cysteine residues fell in the range of around 30 ppm, consistent with the expectation for zinc-ligated cysteines (Fig. S3) (28). Both ZF1 and ZF2 contain six cysteines and two histidines to form two sets of C3H1 ZFs. The solution structures of ZF1 and ZF2 were highly converged with an average positional RMSD of the backbone (N, C α , C', and O) of the secondary structures to the mean of 0.46 Å and 0.64 Å, respectively. ZF1 contained two β -strands and one α -helix (Fig. 2D), while ZF2 contained two β -strands and two α -helices (Fig. 2E). For ZF1, the first C3H1 ZF consists of Cys10 and Cys15 from the N-terminal loop, Cys33 from the loop between β -strand 2 (β_2) and α -helix 1 (α_1), and His36 from α_1 ; the second C3H1 ZF consists of Cys25 from β_1 and Cys28 from β_2 and His42 and Cys44 from the C-terminal loop. For ZF2, the first C3H1 ZF consists of Cys64 and Cys69 from the N-terminal loop, Cys87 from the loop between β_4 and α_2 , and His90 from α_2 ; the second C3H1 ZF consists of Cys79 from β_3 and Cys82 from the loop between β_3 and β_4 , His96 from α_3 , and Cys98 from the C-terminal loop, although the backbone amide assignments for residues 94, 96 to 98 were missing due to the lack of corresponding ^{15}N - ^1H cross peaks.

We next determined the solution structure of UBL by NMR. The average backbone RMSD to mean was 0.69 Å with a highly convergent structural ensemble even for the N-terminal helix (α_4) that had limited contacts with the remaining globular domain (Fig. 2F). UBL adopts a β -grasp fold that is commonly found in ubiquitin-like proteins. The loop connecting α_5 and β_7 (residues 201–212) was less converged than the other secondary structure, which may be attributed to increased loop dynamics consistent with the reduced ^{15}N R_2/R_1 ratio (Fig. 1C).

Structural basis of ZFAND1-p97 interaction

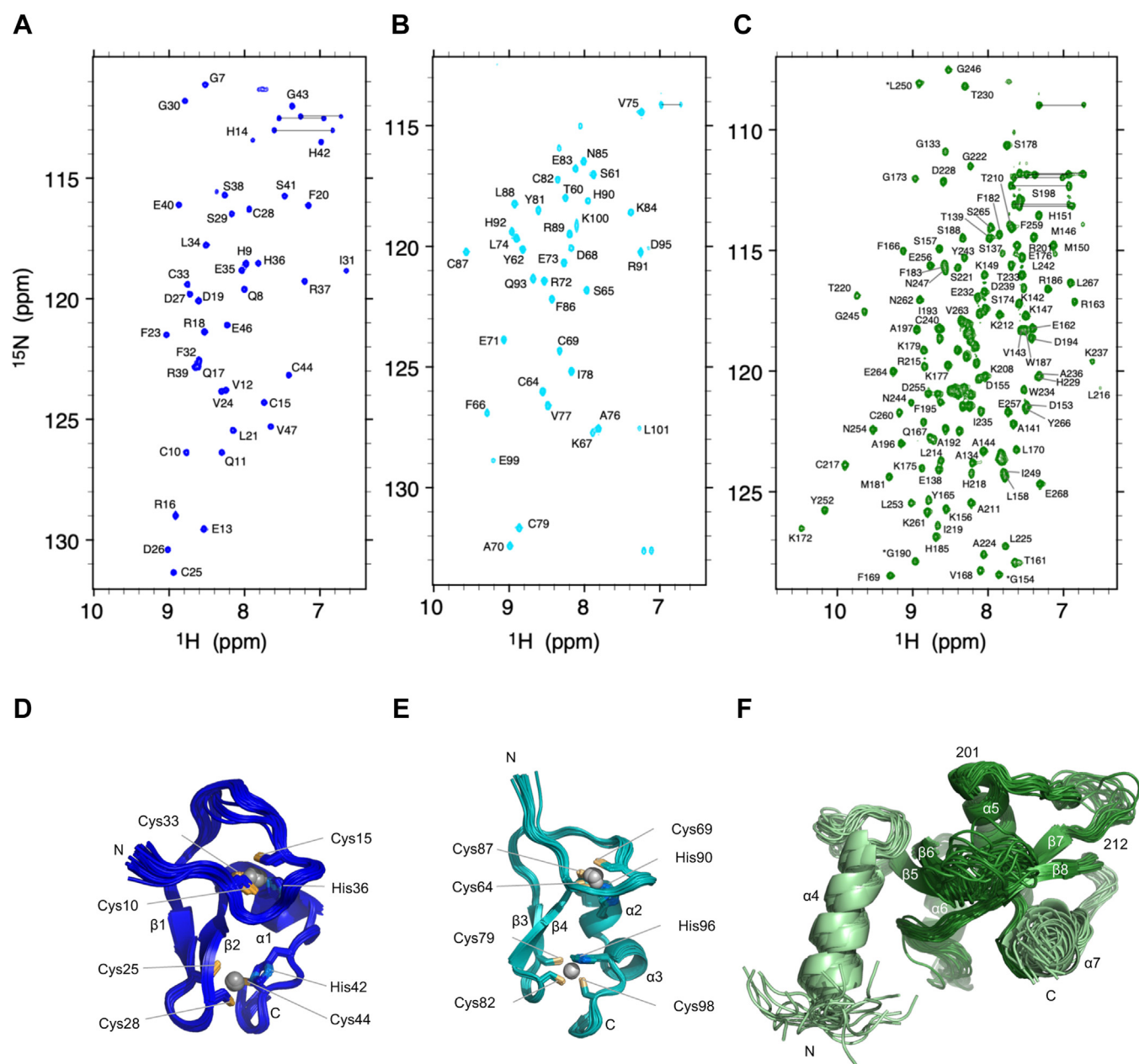


Figure 2. Backbone NMR spectral assignments and solution structures of individual domains of ZFAND1. The ^{15}N - ^1H HSQC spectra of isolated ZF1 (A), ZF2 (B), and UBL (C) are labeled with their corresponding residue-specific assignments. All NMR spectra were recorded at a proton Larmor frequency of 850 MHz at 300 K. Cartoon representations of the ensemble of 20 structures of ZF1 (D), ZF2 (E), and UBL (F) determined by CYANA. Both ZF1 and ZF2 contain two C3H1 ZFs. The coordinated zinc ions are shown in gray spheres, and the coordinating side chains are shown in sticks with the sulfur and nitrogen atoms colored in gold and blue, respectively. The identities of the individual C3H1 motifs are labeled. The N-terminal helix ($\alpha 4$) of UBL is colored light green with the remaining UBL fold colored olive. The individual secondary structures of ZF1, ZF2, and UBL are labeled and numbered consecutively corresponding to the full-length construct.

Crystal structure of UBL

We next crystallized UBL to collect the X-ray diffraction data to the maximum resolution of 1.88 Å (Table S2). The crystal structure of UBL was determined by molecular replacement (MR) using a search model predicated by AlphaFold 2 (AF2) (29, 30). Different search models were generated by serial truncations of the N-terminal helix ($\alpha 4$) with different truncation lengths. Eventually, the MR solution was successfully generated by an N-terminal truncation at Met146, and the $\alpha 4$ was rebuilt in further refinements. A comparison of the crystal

structure, NMR structure, and AF2 prediction showed a general agreement in the globular β -grasp fold (Fig. 3, A–C). Nevertheless, significant deviations could be observed in the relative orientation of the loop region and the N- and C-termini, including $\alpha 4$, $\alpha 4$ - $\beta 5$ loop (residues 152–159), $\beta 5$ - $\beta 6$ loop (residues 170–177), $\alpha 5$ - $\beta 7$ loop (residues 201–212), and $\alpha 7$ (Figs. 3D and S4). The primary structural difference between the crystal structure and NMR structure lies in the N-terminal $\alpha 4$ and the $\alpha 5$ - $\beta 7$ loop (Fig. S4). Indeed, the $\alpha 4$ - $\beta 5$ loop and the $\alpha 5$ - $\beta 7$ loop exhibited significantly reduced ^{15}N

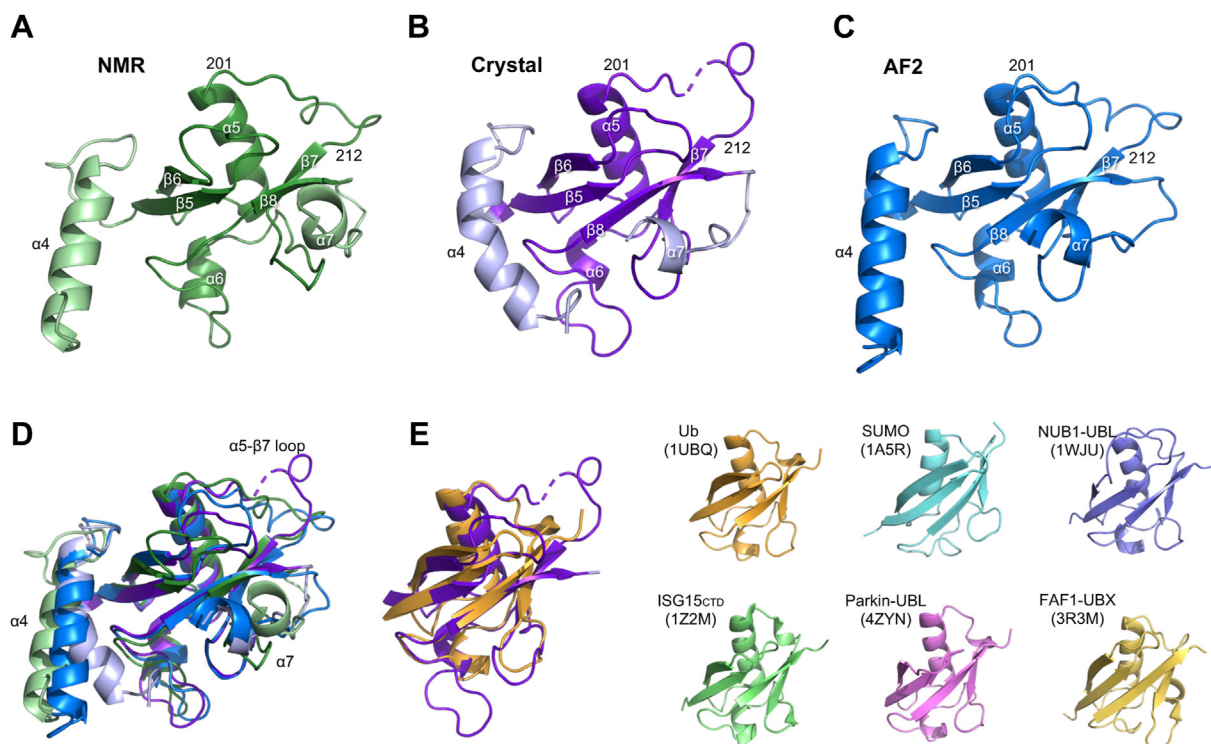


Figure 3. Comparison of the NMR structure, crystal structure, and AF2-predicted structure of UBL with other structural homologs. The representative NMR structure (A), crystal structure (B), and AF2-predicted structure (C) of UBL are shown in *olive, purple, and marine*, respectively. The N-terminal helix $\alpha 4$ of the NMR solution structure and that of the crystal structure are colored *light green* and *pale purple*, respectively, to highlight the difference in the relative orientations. Also highlighted is the difference in the relative orientation of the $\alpha 5$ - $\beta 7$ loop (residues 201–212), which is partially resolved in the crystal structure (indicated by a *dashed line*). For the remaining core structure, which adopts a conserved β -grasp fold commonly found in other ubiquitin-like proteins, the pair-wise RMSD values are 1.54, 1.52, and 0.52 Å for NMR-crystal, NMR-AF2, and crystal-AF2, respectively. D, structure alignments of NMR, crystal, and AF2-predicted structures showed the structure differences on the $\alpha 4$, $\alpha 5$ - $\beta 7$ loop, and $\alpha 7$. E, structural alignment of the core structure of ZFAND1_UBL and ubiquitin (Ub) revealed a conserved β -grasp fold. The other representative ubiquitin-like proteins are shown on the *right part*. Cartoon representations of the crystal structures of ubiquitin (Ub), small ubiquitin-like modifier (SUMO), neural-precursor-cell-expressed developmentally down-regulated (NEED4)-like ubiquitin E3 ligases, interferon-stimulated gene 15 (ISG15), UBL of Parkin E3 ligase, and the UBX domain of FAS-associated factor 1 (FAF1). The core structure of ZFAND1_UBL is Glu162-Leu253, excluding the $\alpha 4$, $\alpha 7$, and their connecting loops. The structures are aligned with respect to the UBL of ZFAND1 in (A–C) to have the same orientations. The corresponding PDB entries of the individual structures are shown in parentheses. AF2, AlphaFold 2; UBL, ubiquitin-like domain.

R_2/R_1 ratios, indicative of local dynamics and conformational exchanges. Despite the clear difference in the relative orientations of $\alpha 4$, it did not show a clear difference in the hetNOE values nor in the R_2/R_1 ratio, suggesting that this helix is relatively ordered in solution. The difference in the conformation of $\alpha 4$ of UBL among the NMR structure, crystal structure, and AF2 prediction may be attributed to the crystal packing of UBL, of which $\alpha 4$ was tightly packed between the core structures intramolecularly and intermolecularly. A number of polar side chains within the core structure, namely Gln167, Asn244, and Asn247, were packed against the hydrophobic residue of $\alpha 4$, namely Ala141, Ala144, Leu145, and Leu148 (Fig. S5). While such an unfavorable interaction may be compensated by the crystal packing, it is likely to be responsible for the separation between $\alpha 4$ and the core structure observed in the NMR structure (Fig. 3A).

UBLs are prevalent in nature, and many UBL structures have been reported in the literature. We used DALI (31) to align experimentally determined protein structures that share a similar fold with the UBL of ZFAND1. Ubiquitin (Ub), SUMO-1 (SUMO), NEDD8 ultimate buster 1 (NUB1), ISG15, E3 ligase UBL domain of Parkin (Parkin-UBL), and FAF1 UBX

domain (FAF1-UBX) were found to be highly similar in their overall fold with the UBL of ZFAND1 despite their very low sequence identities (<18%) (Fig. 3E and Table 1). Note, however, that when the structures of these structurally homologous proteins were used as inputs for MR to solve the phase problem of the X-ray diffraction, we were unable to find a converging solution. Successfully MR solution was only obtained when the AF2 prediction of ZFAND1 structure was made available in 2022.

Probing ZFAND1-p97 interaction by methyl NMR spectroscopy and cryo-EM

We generated [U - 2H , ^{13}C / 1H_m -Ala- β /Ile- $\delta 1$ /Leu- $\delta 1$ /2/Val- $\gamma 1$ /2] full-length ZFAND1 to monitor its interaction with p97 using ^{13}C - 1H SOFAST-HMQC, which is relaxation-optimized for methyl NMR spectroscopy. To minimize the dipolar interaction-induced line-broadening, we also perdeuterated p97 for the NMR titration experiments. The methyl group NMR assignments of full-length ZFAND1 were transferred from the assignments derived from the individual domain constructs (Figs. 2 and 4A). Upon the addition of increasing

Structural basis of ZFAND1-p97 interaction

Table 1
Structural homology analysis of UBLs by DALI

PDB ID ^a	Protein	Residue	Sequence identity (%)	Z-score	RMSD (Å)
1UBQ	Ubiquitin	1–73	14	7.8	2.3
1A5R	SUMO-1	19–94	14	3.6	2.9
1WJU	NEDD8 ultimate buster 1 UBL domain	14–92	18	7.5	1.8
1Z2M	ISG15 C-terminal domain	80–154	14	2.1	8.3
4ZYN	E3 ubiquitin protein ligase parkin UBL domain	1–73	14	7.7	2.3
3R3M	Fas-associated factor 1 UBX domain	571–649	16	5.7	3.0

^a Alignments using the crystal structure of 7Y39. All structure alignments are based on chain A.

amounts of perdeuterated p97, the methyl resonances of ZFAND1 exhibited differential decreases in the peak intensities without appreciable chemical shift changes. By plotting the resonance intensity ratio with respect to the free ZFAND1 as a function of sequence number, it was clear that UBL exhibited more pronounced line-broadening than ZF1 and ZF2. In the presence of a six-fold molar excess of p97, the average signal reduction of UBL was $56 \pm 5\%$ compared to $26 \pm 6\%$ and $21 \pm 5\%$ for ZF1 and ZF2, respectively (Fig. 4B). There was no significant difference within error in the degree of signal reduction for different methyl groups within UBL, suggesting that the overall line broadening was due to the slower overall tumbling rate of UBL when it transiently

interacts with p97, while ZF1 and ZF2 are less likely to make direct contact with p97. The localized NMR line-broadening within UBL compared to ZF1 and ZF2 indicated that UBL was the primary interaction site with p97. This is consistent with the previous biochemical observation that ZFAND1 binding to p97 is strongly reduced upon truncation of UBL, ZFAND1- Δ UBL, according to immunoprecipitation analysis (6). Likewise, the disease-associated nonsense mutation, R130*, also results in reduced p97 recruitment (6). Although the number of titration points was limited, one could infer the lower bound of the dissociation constant of ZFAND1 binding to p97 to be higher than $60 \mu\text{M}$ as the signal reduction of UBL has not yet reached 50% (Fig. 4C).

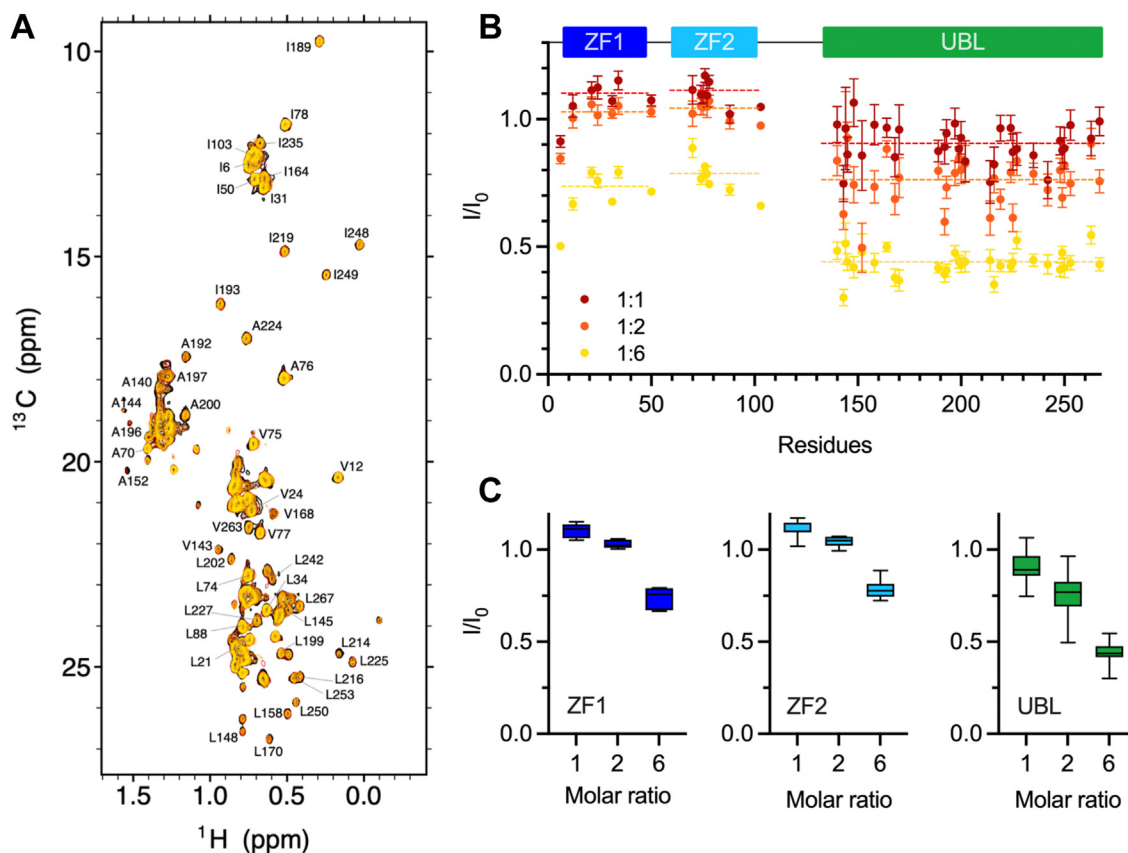


Figure 4. Methyl NMR analysis of p97 interaction with ZFAND1. A, overlay of methyl NMR spectra of $[\text{U-}^2\text{H}, ^{13}\text{C-}^1\text{H}_m\text{-Ile-}\delta 1, \text{Leu-}\delta 1/2, \text{Val-}\gamma 1/2]$ ZFAND1 in the absence and in the presence of 1, 2, and 6 M equivalents of perdeuterated monomeric p97. The spectra are colored according to the labels in the upper right corner. B, intensity ratios of individual methyl groups with respect to the reference spectrum in the absence of p97. The averaged values within the individual domains are shown in horizontal lines to indicate the differential responses upon the addition of p97. C, the box plots of the intensity ratio of individual methyl groups as a function of the molar ratio of monomeric p97 to ZFAND1.

We next attempted to use cryo-EM to determine the structure of p97 in complex with ZFAND1. Our initial attempt failed to identify ZFAND1 in complex with p97 as the reconstructed cryo-EM maps were essentially identical to previously reported EM maps of apo p97 (32). Given the transient nature of ZFAND1-p97 interaction observed by methyl NMR spectroscopy (Fig. 4), we chemically cross-linked ZFAND1 and p97 with an excess amount of ZFAND1 in the presence of ATP γ S (Fig. S6). After extensive three-dimensional image classification, we observed distinct conformation of p97 in the active and inactive states defined by the relative orientation of the NTD: the active state corresponds to the NTD-up conformation, while the inactive state corresponds to the NTD-down conformation (33) (Fig. S7). In both states, however, a significant loss of the NTD EM map was observed (Fig. 5B). In the NTD-up state, three of the six NTDs were resolved; in the NTD-down state, only one NTD was resolved. The loss of the NTD EM map density may be attributed to increased conformational heterogeneity of the NTD induced by ZFAND1 binding such that the particle images of the conformational ensemble cannot be classified into defined conformations through the 3D variability analysis during the image process workflow (Fig. S7). By imposing a C6 symmetry during the image refinement, the NTD completely disappeared despite the increased overall resolution (Fig. S7). For the well-resolved NTDs, it was not obvious as to where ZFAND1 may be located despite it being covalently linked to p97. The results indicated that despite the transient nature of ZFAND1 binding to p97, ZFAND1 binding significantly increases the conformational heterogeneity and potential dynamics of the NTD of p97.

Discussion

p97 is an integral part of cellular proteostasis, but it requires a variety of cofactors to ensure timely and precise subcellular

localization with accurate substrate specificities (34). In the context of arsenite-induced SG clearance, ZFAND1 has been identified as a key cofactor for p97 recruitment. Nevertheless, the molecular basis of p97 recruitment by ZFAND1 remains elusive, not least because of the lack of detailed structural information on the individual functional domains of ZFAND1. In this study, we have shown by solution-state NMR spectroscopy and X-ray crystallography that the C-terminal UBL of ZFAND1 adopts a β -grasp fold similar to ubiquitin and other UBLs, despite the low sequence homology (<20%; Table 1 and Fig. S1B). It is remarkable how AF2 accurately predicts the overall structure of the individual domains of ZFAND1 (Fig. S8). In the case of ZF1 and ZF2, the AF2-predicted structures align well with the NMR structures even though the zinc ions are not present in the AF2 prediction, which are central to the coordination of the side chains of the C3H1 motifs. The overall folds of ZF1 and ZF2 contain limited secondary structures. As such, the loop structures of the AF2 prediction deviate from the NMR structures more significantly than the secondary structures. In fact, among all the ZF domain sequences of the AN1 ZF family curated by Pfam (PF01428, Pfam 36.0) (35), the C3H1 motifs and the flanking residues are the only strictly conserved residues. Nevertheless, the overall folds of all reported experimental structures of AN1 ZFs are highly similar, illustrating how structural conservation can be achieved with minimal sequential conservation (Fig. S9). In the case of the UBL, the overall structure predicted by AF2 agrees well with the NMR structure, and the core structure is in good agreement with the crystal structure such that the AF2 prediction was the one that could yield a reliable MR solution for the crystal structure determination while the other crystal structures of UBL failed to yield satisfactory MR solutions (Fig. 3). Also remarkable is the ability of AF2 to predict the orientation of α 4 that closely resembles that of the NMR structure, which is a relaxed conformation in solution

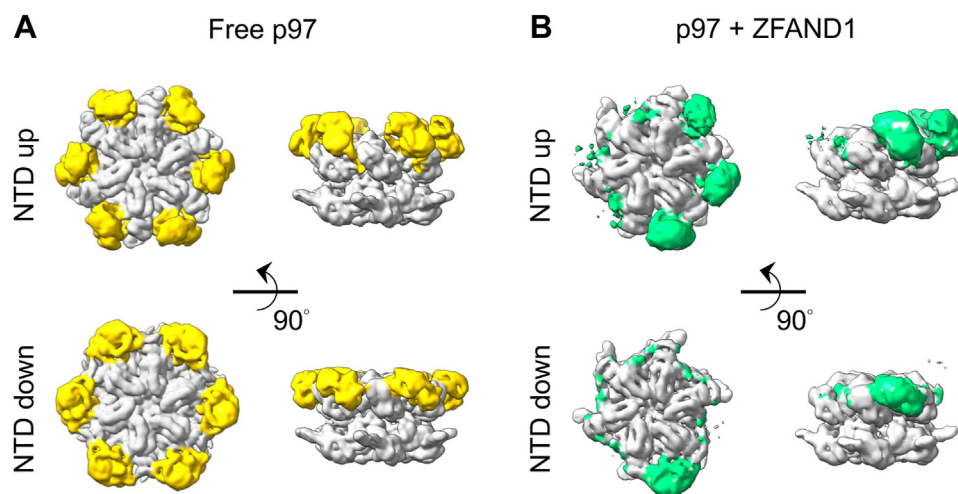


Figure 5. Cryo-EM analysis of conformational changes in p97 in response to ZFAND1 binding. The cryo-EM maps of free p97 (A) and p97 cross-linked with ZFAND1 (B), which indicated the conformational difference primarily located on the NTDs. The NTDs of the hexameric p97 in the absence and presence of ZFAND1 are colored *yellow* and *green*, respectively. In the case of free p97, all six NTDs are well-defined both in the NTD-up and NTD-down states. In contrast, only three of the six NTDs are resolved in the NTD-up state when cross-linked with ZFAND1, whereas only one of the six NTDs is resolved in the NTD-down state, indicating enhanced and asynchronous domain motions in response to ZFAND1 binding. The EM maps of free p97 are taken from EMD-3298 (3-NTD up) and EMD-3299 (1-NTD down).

Structural basis of ZFAND1-p97 interaction

that may be more physiologically relevant, while the bent conformation observed in the crystal structure would be an artefact due to crystal packing (Figs. 3 and S5).

We demonstrated by methyl NMR spectroscopy that UBL is the primary contact point for p97 binding, while ZF1 and ZF2 are less involved in p97 interactions (Fig. 4). Our finding is consistent with the biochemical evidence that truncation of the UBL domain abolishes the p97 recruitment function of ZFAND1 (6). According to COSMIC (36), the nonsense mutation of R130* in ZFAND1 is associated with ovarian carcinoma and breast cancer. The mutation also leads to the loss of the UBL domain of ZFAND1, thereby abrogating its ability to recruit p97 to facilitate the clearance of arsenite-induced SG (6). It is therefore plausible that ZFAND1 may be implicated in other proteostasis functions pertinent to oncogenesis that resemble the physiological stress induced by arsenite treatments. The NTD of p97 is the canonical docking site for its auxiliary factors, such as ubiquitin regulator X domains, VCP-interacting motif (VIM), VCP-binding motif, SHP motif/box, and peptide N-glycosidase/ubiquitin-associated (PUB) domain (9). The UBL of ZFAND1 shares very limited sequence identity with UBX and PUB domains, and it does not contain VCP-binding motif and SHP motif/box. Nevertheless, ZFAND1_UBL contains an RX₅AAX₂R-like sequence—Lys190-Arg201—that is considered a potential VIM (Fig. S10A). The VIM-containing proteins can be classified into two subgroups according to their conserved sequences, and the two subgroups exhibit different binding affinities towards the NTD of p97 (9). The subgroup harboring an RX₅AAX₂R sequence would bind to the NTD of p97 with a higher binding affinity. Examples of such proteins are gp78 and ANKZF1. Their VIM peptides bind to the NTD of p97 with a K_D value of 21 nM and 16 nM, respectively (37). In contrast, another subgroup harboring an AAX₂R sequence binds to the NTD of p97 with a lower binding affinity. Examples include ZFAND2B and UBXD1 that bind to full-length p97 with a K_D value of 1 μM and 3.5 μM, respectively (14, 37). The full-length protein also includes another binding domain that probably contributes to the binding affinity, such as a PUB domain in UBXD1. The interaction between UBXD1-VIM and the NTD of p97 is much weaker with a K_D value of 22 μM according to methyl NMR spectroscopy (38).

Even though ZFAND1 contains an RX₅AAX₂R-like sequence, there are two factors that might interfere with its function, the replacement of a conserved alanine and the structural barrier of a dynamic loop. Note that the sequence of Ala195-Leu202 in ZFAND1 is similar to the VIM of UBXD1, with the exception that the conserved alanine (Ala633 in UBXD1) is replaced by a serine (Fig. S10A). A study of the conserved alanine mutations in the ANKZF1-VIM reveals that the Ala-to-Leu mutation can abolish the p97 interaction (39), suggesting that the Ala-to-Ser replacement in ZFAND1_UBL may contribute to the decrease of binding affinity because the narrow cleft of p97NTD may not tolerate a larger side chain (Fig. S11). For ZFAND1, the VIM-like motif is located at the α5 of the UBL domain before the flexible α5-β7 loop. The α5 is flanked by β-strands β6 and β7 of the well-folded β-sheets

(Fig. 3) with limited backbone motions according to the ¹⁵N spin relaxation dynamics analysis (Fig. 1C). In contrast, the VIMs in gp78, ANKZF1, ZFAND2B, and UBXD1 are on a free α-helix (Fig. S10B). However, the structural alignment between gp78-VIM and ZFAND1_UBL-VIM suggested that using α5 to interact with p97NTD is possible, but the structural clash from the α5-β7 loop (residues 201–212) can interfere with the interaction (Fig. S11B). A cofactor-induced conformational change of residues 201 to 212 may be necessary to improve the ZFAND1_UBL-VIM and p97-NTD interaction.

Despite the similarity to other known VIMs, the UBL of ZFAND1 may have a noncanonical p97-binding motif. Judging from the methyl peak intensity loss as a function of p97 concentration, ZFAND1 binding to p97 was far from being saturated in the presence of a six-fold molar excess of p97 (Fig. 4C), suggesting that the apparent K_D of ZFAND1-p97 binding would be higher than 60 μM. The caveat of our experimental design was that the methyl NMR titration experiment was carried out in the absence of SG-like ubiquitinated substrates, which may be the linchpin to recruiting p97 to bind to ZFAND1. Likewise, the primary function of ZFAND1 is to recruit p97 to disentangle SG for 26S proteasome-mediated proteolysis. In this context, the presence of the 26S proteasome may also be required to enhance the ZFAND1 binding to p97 through multivalency. Despite the limitations of our experimental design to investigate the binding between ZFAND1 and p97, our methyl NMR titration clearly indicated that the UBL is chiefly responsible for p97 binding. The reciprocal structural evidence of the implication of the NTD of p97 in ZFAND1 binding was generated by cryo-EM SPA, which revealed increased conformational heterogeneity of the NTD of p97 upon cross-linking with ZFAND1.

Experimental procedures

Plasmid, expression, and purification of ZFAND1 and p97

The DNA sequence corresponding to ZFAND1 (UniProt ID: Q8TCF1) was codon-optimized, synthesized, and subcloned into a modified pGEX-4t3 expression vector. This modified vector contained an N-terminal glutathione S-transferase (GST) fusion tag followed by a tobacco etch virus protease cleavage site. The constructs corresponding to the individual domains—ZF1 (residues 7–47), ZF2 (residues 60–101), and UBL (residues 133–268)—and the 2 ZFs, namely ZF1+2 (residues 1–132) were amplified by PCR and subcloned into the same pGEX-4t3 expression vector. The DNA sequence of full-length p97 was cloned into a pRSFduet-1 expression vector with an N-terminal fusion hexahistidine (His₆) tag as described previously (40). All plasmids were transformed into *Escherichia coli* BL21 (DE3) for the overexpression of recombinant proteins. For full-length ZFAND1 and its truncated protein expression, the transformed cells were grown in lysogeny broth at 37 °C until the absorbance at 600 nm (A₆₀₀) reached 0.6 to 0.8 followed by the addition of IPTG to a final concentration of 0.1 mM. The expression of full-length ZFAND1, ZF1+2, ZF1, and ZF2 required the addition of zinc acetate (or zinc sulfate) at the final

concentration of 50 μM in the cell culture. The overexpression of the recombinant proteins was continued at 20 °C for 18 to 20 h. The cells were harvested by centrifugation. For stable isotope labeling, the *E. coli* cultures were grown in M9 media at 37 °C until A_{600} reached 0.4 to 0.5 as described previously (41). The ^{13}C -labeled D-glucose (Cambridge Isotope Laboratory) and ^{15}N -labeled ammonium chloride (Cambridge Isotope Laboratory) were used as the sole carbon and nitrogen sources for uniform labeling. For selective ^{13}C - and ^1H -methyl labeling at the Ile- δ 1, Leu-C δ 1/2, Val-C γ 1/2, and Ala-C β positions under a perdeuterated background, the M9 media was prepared in 100% D_2O with uniformly $^2\text{H}/^{12}\text{C}$ -labeled D-glucose (Cambridge Isotope Laboratory) as the sole carbon source and the appropriate precursors for methyl labeling (*In vivo* ILVA methyl labeling kit; CDLM-8806-KIT; Cambridge Isotope Laboratory) added 1 h prior to the IPTG induction. For perdeuterated p97, the *E. coli* culture was grown in perdeuterated M9 medium with uniformly ^2H -labeled D-glucose (Cambridge Isotope Laboratory) as the sole carbon source at 37 °C until A_{600} reached 0.6 to 0.7. The p97 overexpression was triggered by the addition of 0.1 mM IPTG and continued at 25 °C for 16 h.

For recombinant protein purification, the cell pellets were resuspended in lysis buffer (p97: 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 5 mM MgCl_2 , and 1 mM DTT; Full-length ZFAND1/ZF1+2/UBL: 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 0.5 mM TCEP, 10% (v/v) glycerol, and protease inhibitor; ZF1/ZF2: 50 mM Tris (pH 8.0), 200 mM NaCl, 1 mM β -ME, 5% (v/v) glycerol, and 0.01% (v/v) TritonX-100) and lysed by a Q700 Sonicator (Qsonica). The cell debris was removed by centrifugation at 20,000 rpm for 30 min, and the supernatants were collected for purification. His-tagged p97 was purified by Ni-NTA affinity chromatography (HisTrap HP column, Cytiva) in the Ni-NTA elution buffer (50 mM Tris (pH 8.0), 500 mM NaCl, 1 mM DTT, 300 mM imidazole) and followed by size-exclusion chromatography (SEC) using a Superdex 200 Increase 10/300 column (Cytiva) in the SEC buffer (50 mM Hepes (pH 7.8), 150 mM NaCl, 5 mM MgCl_2 , 0.5 mM TCEP, and 10% (v/v) glycerol). The full-length ZFAND1 and individual variants were purified by GST affinity resin (Glutathione Sepharose 4B (Cytiva)) in the GST elution buffer (Full-length ZFAND1/ZF1+2/UBL: 50 mM Tris (pH 8.0), 150 mM NaCl, 0.5 mM TCEP, 10% (v/v) glycerol, and 20 mM glutathione; ZF1/ZF2: 50 mM Tris-HCl (pH 8.0), 200 mM NaCl, 10 mM β -ME, 5% (v/v) glycerol, and 20 mM glutathione). The GST-tag was removed by tobacco etch virus protease digestion and followed by SEC using a Hiload 16/600 Superdex 75 column (Cytiva). For ZFAND1, ZF1+2 elution, and UBL, the SEC buffer was 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 0.5 mM TCEP, and 10% (v/v) glycerol. For ZF1 and ZF2 elution, the SEC buffer contained 50 mM Tris (pH 8.0) and 200 mM NaCl.

NMR spectroscopy and structural determination

The NMR spectra were collected on a Bruker Avance III 850 spectrometer equipped with a TCI cryo-probe and a Bruker Avance III 600 spectrometer equipped with a TXI cryo-probe.

[U- ^{15}N] full-length ZFAND1, [U- ^{15}N] ZF1+2, [U- ^{15}N] UBL, and [U- ^2H , $^{13}\text{C}/^1\text{H}_m$ -Ala- β /Ile- δ 1/Leu- δ 1/2/Val- γ 1/2] full-length ZFAND1 were prepared in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.5 mM TCEP. For structure determination, [U- $^{13}\text{C}/^{15}\text{N}$] ZF1 and [U- $^{13}\text{C}/^{15}\text{N}$] ZF2 were prepared in 50 mM sodium phosphate buffer (pH 6.5), and [U- $^{13}\text{C}/^{15}\text{N}$] UBL was prepared in 50 mM Tris-HCl (pH 7.1), 150 mM NaCl, and 0.5 mM TCEP. The triple resonance data and NOESY data were collected at 300 K. The backbone resonance assignments were achieved by the combination of ^{15}N - ^1H HSQC, HNCACB, CBCA(CO)NH, and HNCO. The side-chain resonance assignments were completed by using ^{13}C - ^1H constant-time HSQC ^{15}N -, aliphatic ^{13}C -, and aromatic ^{13}C -edited NOESY-HSQC, HCCH-TOCY, CCH-COSY, HBHA(CO)NH. (HB)CB(CGCDCE)HE (42), and (HB)CB(CGCD)HD (42) in addition to the backbone triple resonance experiments. All NMR spectra were processed by NMRPipe (43) and analyzed by NMRFAM-SPARKY (44) and CARI (45), and the NMR structure determinations were carried out using CYANA (46) (Table S1). The zinc ions in the ZF1 and ZF2 structures were created by the ION of CYANA.lib. The distance of SG-Zn was fixed in 2.4 Å by using ambiguous restraints. The distance restraint was based on a study of zinc protein structures in NMR and crystal by Patel *et al.* (47). The NMR structures of ZF1, ZF2, and UBL have been deposited in the PDB databank under the accession numbers of 7YAB, 7Y7L, and 8XE8, respectively. The NMR assignments of ZF1, ZF2, and UBL have been deposited in the Biological Magnetic Resonance Bank under the accession numbers of 51499, 36494, and 36622, respectively. The ZF1, ZF2, and UBL structures were rendered by PyMol 2.0 (Schrodinger Scientific) (48).

^{15}N backbone relaxation measurements were recorded using full-length [U- ^{15}N] ZFAND1. The sample concentration was 250 μM and buffered in 50 mM Tris-HCl (pH 7.1), 100 mM NaCl, and 0.5 mM TCEP. The experiments were carried out at 300 K and at a proton Larmor frequency of 850 MHz. The experimental setup followed previously described procedures (49). The inversion recovery delays for longitudinal (R_1) relaxation were set to 100, 200*, 500, 750*, 1000, 1250, and 1500 ms (asterisks indicate that the experiments with corresponding delays were carried in duplicate for improved estimates of error). Carr-Purcell-Meiboom-Gill spin echo pulses with a field strength of 1.56 kHz and a spacing of 450 μs were applied during the transverse (R_2) relaxation delays that were set to 16.96*, 33.92, 50.88, 84.80*, 101.76, and 117.72 ms (asterisks indicate that the experiments with corresponding delays were carried in duplicate for improved estimates of error). Additional Carr-Purcell-Meiboom-Gill pulses were applied during the relaxation delays to compensate for sample heating caused by repetitive inversion pulse trains. For steady-state hetNOE measurements, proton saturation was achieved by means of a 5-s pulse train of 120 ^1H pulses, separated by an interval of 5 ms. The heteronuclear NOE experiments were recorded in triplicate for better error estimation. The relaxation rates were derived by extracting peak

Structural basis of ZFAND1-p97 interaction

heights followed by fitting the peak heights as a function of the relaxation delay using the built-in rate analysis module in SPARKY. The hetNOE values and R_2/R_1 ratio were plotted by Prism 9 (GraphPad Software; <https://www.graphpad.com>) (50).

The overall correlation times of individual domains were estimated from the R_2/R_1 ratio according to the following equation (51):

$$\tau_C = \frac{1}{2\omega_N} \sqrt{\frac{6R_2}{R_1} - 7}$$

where ω_N is the Larmor frequency of ^{15}N , which is 86.176 MHz. As conformational exchanges in loop regions may lead to enhanced R_2 relaxation (52, 53), the R_2/R_1 ratios corresponding to the basal levels of individual domains were manually curated and averaged to obtain an average R_2/R_1 ratio of each domain to estimate the overall correlation time.

Methyl NMR spectroscopy for probing p97 interaction with ZFAND1

The p97 titration experiments were conducted on a Bruker Avance III 850 spectrometer equipped with a TCI cryo-probe. The [$U\text{-}^2\text{H}$, $^{13}\text{C}/^1\text{H}_m\text{-Ala-}\beta/\text{Ile-}\delta 1/\text{Leu-}\delta 1/2/\text{Val-}\gamma 1/2$] full-length ZFAND1 and the [$U\text{-}^2\text{H}$] p97 were prepared in the buffer system containing 50 mM Tris-HCl (pH 7.1), 100 mM NaCl, and 0.5 mM TCEP. Due to the high molecular weight of p97 and its limited expression yield in the perdeuterated form, the concentration of [$U\text{-}^2\text{H}$, $^{13}\text{C}/^1\text{H}_m\text{-Ala-}\beta/\text{Ile-}\delta 1/\text{Leu-}\delta 1/2/\text{Val-}\gamma 1/2$] full-length ZFAND1 was set to 10 μM ; the [$U\text{-}^2\text{H}$] p97 was concentrated to 100 to 270 μM for the NMR titration experiment. The ZFAND1-p97 interaction was monitored by edited ^{13}C -SOFAS-T-HMQC at 300 K in the molar ratio ranging from 1:1 to 1:6. The titration spectra were processed by NMRPipe (43), and intensity changes were analyzed by NMRFAM-SPARKY (44). The I/I_0 values were visualized by Prism 9 (GraphPad Software) (50).

Protein X-ray crystallography

Recombinant UBL was concentrated to 11 mg/ml in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.5 mM TCEP and mixed with reservoir solutions at a volume ratio of 1:1. The crystals were grown in sitting plates, and the successful reservoir solution contained 0.1 M MES (pH 6.5) and 1.2 M sodium citrate. The X-ray diffraction dataset was collected at the National Synchrotron Radiation Research Center (NSRRRC) using the beamline BL13B1. The diffraction data were indexed by HKL2000 (54). The crystal structure was determined by MR in Phenix (55) using the AF2 predicated model (AF-Q8TCF1-F1) (29) as the template. N-terminal truncation of the AlphaFold model (keeping only residues 149–268) was necessary to yield successful MR results. The structure was manually optimized by Coot (56), and subsequent refinement was performed by Phenix (55) and ccp4i (57). The

final crystal structure of UBL has been deposited in the PDB databank under the accession number 7Y39.

Cryo-EM SPA

Chemical cross-linking of p97 and ZFAND1 was necessary to maintain complex formation during grid preparation. Recombinant p97 and ZFAND1 were mixed at a molar ratio of 1:6 in 20 mM Hepes (pH 7.5), 150 mM NaCl, 5 mM MgCl_2 , 0.5 mM TCEP, and 1 mM ATP γS . Glutaraldehyde was added to the mixture to a final concentration of 0.0625% (v/v) and followed by incubation at 37 °C for 5 min before the addition of 1 M Tris-HCl (pH 8.0) to a final concentration of 30 mM to quench the cross-linking reaction. The cross-linked samples were analyzed using Coomassie blue staining and Western blotting. For Western blotting, the reacted samples were resolved by NuPAGE 3 to 8% Tris-Acetate Protein Gel (Invitrogen) and transferred onto the PVDF blotting membrane (GE HealthCare). The cross-linked p97-ZFAND1 was probed by a primary rabbit anti-ZFAND1 antibody (Sigma, Cat#HPA023383) diluted 1:500 in TBST (20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.05% (v/v) Tween20) with 5% milk and probed by a secondary HRP-conjugated anti-rabbit antibody (Abcam) diluted 1:5000 in TBST with 5% milk. The chemiluminescence was generated by incubating with an ECL reagent (Bio-Rad) and detected by the iBright imaging system (Thermo Fisher Scientific) with an exposure time of 150 ms. The cross-linked sample was subject to SEC purification using a Superpose 6 increase 10/300 column (Cytiva) in 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 0.5 mM TCEP to remove higher oligomers and unreacted ZFAND1 monomer. The fractions corresponding to cross-linked p97-ZFAND1 were collected and concentrated to 1 mg/ml.

The concentrated p97-ZFAND1 complex was supplemented with 1 mM ATP γS and incubated on ice for 30 min prior to grid preparation. Four microliters of the stock solution was applied on the glow-discharged UltrAuFoil R1.2/1.3 Au grids, blotted for 3 s at 4 °C with 100% humidity, and plunge-frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific). Cryo-EM micrographs were collected using a 300 kV Titan Krios microscope (Thermo Fisher Scientific) with a K2 direct electron detector (Gatan). Data collection was operated in a counting mode with a defocus range between -1.5 and -2 μm at a magnification of 150000 \times , corresponding to a pixel size of 0.52 Å. The dose rate was 9.2 $\text{e}^-/\text{px}/\text{s}$, and the total exposure time was 4.5 s, resulting in a total electron exposure of 58.7 $\text{e}^-/\text{Å}^2$, fractionated into 60 frames.

The micrographs were motion-corrected by MotionCor2 v.1.3.2 (58) followed by CTF estimation by Gctf v.1.0.6 and CTFFIND v.4.1 for particle picking all within Relion 3.0 (59). The particle images (957k particles) were exported to CryoSPARC 3.1 (60) for 2D classification (293k particles were selected), ab initio model building, and 3D classification with heterogeneous refinement to select a subset of particle images (153k particles) for subsequent homogeneous refinement without imposing symmetry. The selected particles were

further processed in two different ways: first, a C6 symmetry was imposed for homogeneous refinement leading a map of 3.3 Å, which was deposited to the Electron Microscopy Data Bank under the accession ID of EMD-38772. Second, the same particles were analyzed by 3D variability function with CryoSPARC to classify them into five distinct classes of which one corresponded to three NTD-up conformation and the other corresponded to one NTD-down conformation. These two classes were further refined by non-uniform refinement without symmetry. The final cryo-EM maps had an overall resolution of 3.8 Å for 3 NTD-up and 3.9 Å for 1 NTD-down and were deposited to the Electron Microscopy Data Bank under the accession IDs of EMD-38770 and EMD-38771, respectively.

Data availability

Additional data could be shared by contacting Shang-Te Danny Hsu (thsu@gate.sinica.edu.tw).

Supporting information—This article contains supporting information (28–30, 35, 48, 61–66).

Acknowledgments—We thank Academia Sinica High-Field NMR Center (AS-CFII-111-214) and Academia Sinica Cryo-EM Center (AS-CFII-111-210) for their technical support in data collection and analyses; both core facilities are funded by the Academia Sinica Core Facility and Innovative Instrument Project. Taiwan Protein Project (AS-KPQ-109-TPP2) is also acknowledged for supporting the cryo-EM facility. We thank the protein core facility of the Institute of Biological Chemistry, Academia Sinica (IBC-AS), for supporting protein production, Dr Kai-Fa Huang of the protein crystallography core of IBC-AS for supporting X-ray diffraction data collection and processing, and Dr Kuen-Phon Wu of IBC-AS for the kind gift of the expression vector for p97.

Author contributions—C.-H. L., K.-T. K., P.-J. F., T.-A. Y., and S.-T. D. H. data curation; C.-H. L., K.-T. K., P.-J. F., T.-A. Y., C.-F. C., P. D., and S.-T. D. H. formal analysis; C.-H. L., K.-T. K., P. D., and S.-T. D. H. investigation; C.-H. L., K.-T. K., and S.-T. D. H. validation; C.-H. L. and S.-T. D. H. writing—original draft; C.-F. C. and S.-T. D. H. supervision; P. D. and S.-T. D. H. methodology; S.-T. D. H. conceptualization; S.-T. D. H. funding acquisition; S.-T. D. H. project administration; S.-T. D. H. resources; S.-T. D. H. visualization; S.-T. D. H. writing—review and editing.

Funding and additional information—This work was supported by Academia Sinica intramural fund, an Academia Sinica Career Development Award, Academia Sinica to S.-T. D. H. (AS-CDA-109-L08), and the National Science and Technology Council (NSTC), Taiwan (110-2113-M-001-050-MY3 and 110-2311-B-001-013-MY3) to S.-T. D. H.

Conflict of interest—The authors declare that they have no conflicts of interests with the contents of this article.

Abbreviations—The abbreviations used are: AF2, AlphaFold 2; GST, glutathione-S-transferase; MR, molecular replacement; NTD, N-terminal domain; SEC, size-exclusion chromatography; SG, stress

granule; SPA, single particle analysis; UBL, ubiquitin-like domain; UBX, ubiquitin regulator X; VCP, valosin-containing protein; VIM, VCP-interacting motif; ZF, zinc finger.

References

- Buchan, J. R., Kolaitis, R. M., Taylor, J. P., and Parker, R. (2013) Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. *Cell* **153**, 1461–1474
- Ivanov, P., Kedersha, N., and Anderson, P. (2019) Stress granules and processing bodies in translational control. *Cold Spring Harb. Perspect. Biol.* **11**, a032813
- Hofmann, S., Kedersha, N., Anderson, P., and Ivanov, P. (2021) Molecular mechanisms of stress granule assembly and disassembly. *Biochim. Biophys. Acta Mol. Cell Res.* **1868**, 118876
- Krause, L. J., Herrera, M. G., and Winkhofer, K. F. (2022) The role of ubiquitin in regulating stress granule dynamics. *Front. Physiol.* **13**, 910759
- Wolozin, B., and Ivanov, P. (2019) Stress granules and neurodegeneration. *Nat. Rev. Neurosci.* **20**, 649–666
- Turakhiya, A., Meyer, S. R., Marincola, G., Bohm, S., Vanselow, J. T., Schlosser, A., et al. (2018) ZFAND1 recruits p97 and the 26S proteasome to promote the clearance of arsenite-induced stress granules. *Mol. Cell* **70**, 906–919.e7
- Tolay, N., and Buchberger, A. (2021) Comparative profiling of stress granule clearance reveals differential contributions of the ubiquitin system. *Life Sci. Alliance* **4**, e202000927
- Maxwell, B. A., Gwon, Y., Mishra, A., Peng, J., Nakamura, H., Zhang, K., et al. (2021) Ubiquitination is essential for recovery of cellular activities after heat shock. *Science* **372**, eabc3593
- Buchberger, A., Schindelin, H., and Hanzelmann, P. (2015) Control of p97 function by cofactor binding. *FEBS Lett.* **589**, 2578–2589
- Johnson, J. O., Mandrioli, J., Benatar, M., Abramzon, Y., Van Deerlin, V. M., Trojanowski, J. Q., et al. (2010) Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* **68**, 857–864
- Gwon, Y., Maxwell, B. A., Kolaitis, R. M., Zhang, P., Kim, H. J., and Taylor, J. P. (2021) Ubiquitination of G3BP1 mediates stress granule disassembly in a context-specific manner. *Science* **372**, eabf6548
- Ye, Y., Tang, W. K., Zhang, T., and Xia, D. (2017) A mighty “protein extractor” of the cell: structure and function of the p97/CDC48 ATPase. *Front. Mol. Biosci.* **4**, 39
- Jacobson, T., Navarrete, C., Sharma, S. K., Sideri, T. C., Ibstedt, S., Priya, S., et al. (2012) Arsenite interferes with protein folding and triggers formation of protein aggregates in yeast. *J. Cell Sci.* **125**, 5073–5083
- Glinka, T., Alter, J., Braunstein, I., Tzach, L., Wei Sheng, C., Geifman, S., et al. (2014) Signal-peptide-mediated translocation is regulated by a p97-AIRAPL complex. *Biochem. J.* **457**, 253–261
- Stanhill, A., Haynes, C. M., Zhang, Y., Min, G., Steele, M. C., Kalinina, J., et al. (2006) An arsenite-inducible 19S regulatory particle-associated protein adapts proteasomes to proteotoxicity. *Mol. Cell* **23**, 875–885
- Pharoah, P. D., Tsai, Y. Y., Ramus, S. J., Phelan, C. M., Goode, E. L., Lawrenson, K., et al. (2013) GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. *Nat. Genet.* **45**, 362–370
- Dumont, M., Weber-Lassalle, N., Joly-Beauparlant, C., Ernst, C., Droit, A., Feng, B. J., et al. (2022) Uncovering the contribution of moderate-penetrance susceptibility genes to breast cancer by whole-exome sequencing and targeted enrichment sequencing of candidate genes in women of European ancestry. *Cancers (Basel)* **14**, 3363
- Hanna, J., Waterman, D., Isasa, M., Elsasser, S., Shi, Y., Gygi, S., et al. (2014) Cuz1/Ynl155w, a zinc-dependent ubiquitin-binding protein, protects cells from metalloinduced proteotoxicity. *J. Biol. Chem.* **289**, 1876–1885
- Sa-Moura, B., Funakoshi, M., Tomko, R. J., Jr., Dohmen, R. J., Wu, Z., Peng, J., et al. (2013) A conserved protein with AN1 zinc finger and ubiquitin-like domains modulates Cdc48 (p97) function in the ubiquitin-proteasome pathway. *J. Biol. Chem.* **288**, 33682–33696

Structural basis of ZFAND1-p97 interaction

- Sun, Z. J., Bhanu, M. K., Allan, M. G., Arthanari, H., Wagner, G., and Hanna, J. (2016) Solution structure of the Cuz1 AN1 zinc finger domain: an exposed LDFLP motif defines a subfamily of AN1 proteins. *PLoS One* **11**, e0163660
- Bodnar, N. O., Kim, K. H., Ji, Z., Wales, T. E., Svetlov, V., Nudler, E., *et al.* (2018) Structure of the Cdc48 ATPase with its ubiquitin-binding cofactor Ufd1-Npl4. *Nat. Struct. Mol. Biol.* **25**, 616–622
- Twomey, E. C., Ji, Z., Wales, T. E., Bodnar, N. O., Ficarro, S. B., Marto, J. A., *et al.* (2019) Substrate processing by the Cdc48 ATPase complex is initiated by ubiquitin unfolding. *Science* **365**, eaax1033
- Pan, M., Yu, Y., Ai, H., Zheng, Q., Xie, Y., Liu, L., *et al.* (2021) Mechanistic insight into substrate processing and allosteric inhibition of human p97. *Nat. Struct. Mol. Biol.* **28**, 614–625
- Nandi, P., Li, S., Columbres, R. C. A., Wang, F., Williams, D. R., Poh, Y. P., *et al.* (2021) Structural and functional analysis of disease-linked p97 ATPase mutant complexes. *Int. J. Mol. Sci.* **22**, 8079
- van den Boom, J., Marini, G., Meyer, H., and Saibil, H. R. (2023) Structural basis of ubiquitin-independent PP1 complex disassembly by p97. *EMBO J.* **42**, e113110
- Lee, H. G., Lemmon, A. A., and Lima, C. D. (2023) SUMO enhances unfolding of SUMO-polyubiquitin-modified substrates by the Ufd1/Npl4/Cdc48 complex. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2213703120
- Braxton, J. R., Altobelli, C. R., Tucker, M. R., Tse, E., Thwin, A. C., Arkin, M. R., *et al.* (2023) The p97/VCP adapter UBXD1 drives AAA+ remodeling and ring opening through multi-domain tethered interactions. *Nat. Struct. Mol. Biol.* **30**, 2009–2019
- Kornhaber, G. J., Snyder, D., Moseley, H. N., and Montelione, G. T. (2006) Identification of zinc-ligated cysteine residues based on ¹³Cα and ¹³Cβ chemical shift data. *J. Biomol. NMR* **34**, 259–269
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589
- Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., *et al.* (2022) AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444
- Holm, L., Laiho, A., Toronen, P., and Salgado, M. (2023) DALI shines a light on remote homologs: one hundred discoveries. *Protein Sci.* **32**, e4519
- Banerjee, S., Bartesaghi, A., Merk, A., Rao, P., Bulfer, S. L., Yan, Y., *et al.* (2016) 2.3 Å resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. *Science* **351**, 871–875
- van den Boom, J., and Meyer, H. (2018) VCP/p97-Mediated unfolding as a principle in protein homeostasis and signaling. *Mol. Cell* **69**, 182–194
- Conicella, A. E., Huang, R., Ripstein, Z. A., Nguyen, A., Wang, E., Lohr, T., *et al.* (2020) An intrinsically disordered motif regulates the interaction between the p47 adaptor and the p97 AAA+ ATPase. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 26226–26236
- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., *et al.* (2021) Pfam: the protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419
- Sondka, Z., Dhir, N. B., Carvalho-Silva, D., Jupe, S., Madhumita, McLaren, K., *et al.* (2024) COSMIC: a curated database of somatic variants and clinical data for cancer. *Nucleic Acids Res.* **52**, D1210–D1217
- Hanzelmann, P., and Schindelin, H. (2011) The structural and functional basis of the p97/valosin-containing protein (VCP)-interacting motif (VIM): mutually exclusive binding of cofactors to the N-terminal domain of p97. *J. Biol. Chem.* **286**, 38679–38690
- Schuetz, A. K., and Kay, L. E. (2016) A dynamic molecular basis for malfunction in disease mutants of p97/VCP. *Elife* **5**, e20143
- Stapf, C., Cartwright, E., Bycroft, M., Hofmann, K., and Buchberger, A. (2011) The general definition of the p97/valosin-containing protein (VCP)-interacting motif (VIM) delineates a new family of p97 cofactors. *J. Biol. Chem.* **286**, 38670–38678
- Hung, T. I., Hsieh, Y. J., Lu, W. L., Wu, K. P., and Chang, C. A. (2023) What strengthens protein-protein interactions: analysis and applications of residue correlation networks. *J. Mol. Biol.* **435**, 168337
- Lou, S. C., Wetzel, S., Zhang, H., Crone, E. W., Lee, Y. T., Jackson, S. E., *et al.* (2016) The knotted protein UCH-L1 exhibits partially unfolded forms under native conditions that share common structural features with its kinetic folding intermediates. *J. Mol. Biol.* **428**, 2507–2520
- Yamazaki, T., Forman-Kay, J. D., and Kay, L. E. (1993) Two-dimensional NMR experiments for correlating ¹³Cβ and ¹Hδ/ε chemical shifts of aromatic residues in ¹³C-labeled proteins via scalar coupling. *J. Am. Chem. Soc.* **115**, 11054–11055
- Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293
- Lee, W., Tonelli, M., and Markley, J. L. (2015) NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327
- Keller, R. L. J. (2004) *The Computer Aided Resonance Assignment Tutorial*. Cantina Verlag, Goldau, Switzerland
- Guntert, P., and Buchner, L. (2015) Combined automated NOE assignment and structure calculation with CYANA. *J. Biomol. NMR* **62**, 453–471
- Patel, K., Kumar, A., and Durani, S. (2007) Analysis of the structural consensus of the zinc coordination centers of metalloprotein structures. *Biochim. Biophys. Acta* **1774**, 1247–1253
- The PyMOL Molecular Graphics System, Version 2.4.1*. (2020). Schrödinger, LLC, New York, NY
- Hsu, S.-T. D., Cabrita, L. D., Fucini, P., Dobson, C. M., and Christodoulou, J. (2009) Structure, dynamics and folding of an immunoglobulin domain of the gelation factor (ABP-120) from *Dictyostelium discoideum*. *J. Mol. Biol.* **388**, 865–879
- GraphPad Prism Version 9.5.1 for macOS*. (2023). GraphPad Software, San Diego, CA
- Fushman, D., Weisemann, R., Thuring, H., and Ruterjans, H. (1994) Backbone dynamics of ribonuclease T1 and its complex with 2'GMP studied by two-dimensional heteronuclear NMR spectroscopy. *J. Biomol. NMR* **4**, 61–78
- Hsu, S.-T. D., Lee, Y. C., Mikula, K. M., Backlund, S. M., Tascon, I., Goldman, A., *et al.* (2021) Tying up the loose ends: a mathematically knotted protein. *Front. Chem.* **9**, 663241
- Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T., *et al.* (2002) Long-range interactions within a nonnative protein. *Science* **295**, 1719–1722
- Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326
- Liebschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., *et al.* (2019) Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D Struct. Biol.* **75**, 861–877
- Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501
- Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242
- Zheng, S. Q., Palovcak, E., Armache, J. P., Verba, K. A., Cheng, Y., and Agard, D. A. (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332
- Scheres, S. H. W. (2020) Amyloid structure determination in RELION-3.1. *Acta Crystallogr. D Struct. Biol.* **76**, 94–101
- Punjani, A., Rubinstein, J. L., Fleet, D. J., and Brubaker, M. A. (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296
- Magrane, M., and UniProt Consortium (2011) UniProt Knowledgebase: a hub of integrated protein data. *Database (Oxford)* **2011**, bar009
- Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., *et al.* (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* **38**, W695–W699

63. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539
64. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324
65. Theobald, D. L., and Steindel, P. A. (2012) Optimal simultaneous superpositioning of multiple structures with missing data. *Bioinformatics* **28**, 1972–1979
66. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190