Supplementary Information for

An intrinsically disordered motif regulates the interaction between the p47 adaptor and the p97 AAA+ ATPase

This PDF file includes:

Supplementary text
Figs. S1 to S14
Tables S1 to S4
References for SI reference citations

Materials and Methods

Cloning, protein expression and purification The expression plasmid of full-length p47 from *Mus musculus* was purchased from GenScript, with an N-terminal His_6-tag and a tobacco etch virus (TEV) cleavage site between the tag and the protein sequence. All point mutations and fragments of p47 were generated with *Phusion* DNA polymerase using the QuikChange site-directed mutagenesis method (Agilent, Santa Clara, CA, USA). p47 fragments, used in a variety of different experiments described in the main text, were prepared using the *NEB* Gibson Assembly Master Mix via the Gibson assembly approach (1). Plasmids of full-length p97 and $p97_{ND1L}$ *from Mus musculus* were obtained as previously described (2).

All proteins were expressed in Codon+ *Escherichia coli* BL21(DE3) cells. To produce unlabeled proteins, transformed cells were grown in LB media. For uniformly ^{15}N,^{13}C-labeled samples, proteins were expressed in M9 minimal media with $^{15}NH_4Cl$ and ^{13}C-glucose as the sole nitrogen and carbon sources. Perdeuterated samples were prepared using D_2O-based M9 media with either [^{2}H,^{12}C]-glucose or [^{2}H,^{13}C]-glucose as the carbon source. Selective labeling of Ile-δ1, Leu-δ1/δ2, Val-γ1/γ2, and Met-ε1 methyl groups, in an otherwise highly deuterated environment, was achieved as previously described (3). For the production of unlabeled and uniformly ^{15}N,^{13}C-labeled proteins expression was induced with 0.25 mM IPTG at OD_{600} = 0.7. For selective methyl labeled proteins, precursors were added into the culture at OD_{600} = 0.7 followed by

www.pnas.org/cgi/doi/10.1073/pnas.2013920117

induction with 0.25 mM IPTG 45 min after addition of precursors. Expression was continued for *ca.* 16 hrs at 25 ^{o}C before harvest.

Purification of p47 and p47 constructs was achieved through a procedure in which harvested cells were initially resuspended in Ni-A buffer [50 mM TrisHCl (pH 7.4), 500 mM NaCl, 30 mM imidazole], followed by lysis via sonication. After centrifugation at 14,000 × g for 20 min, the supernatant was filtered with a 1.0 μm filter and applied to a HisTrap affinity column (GE Healthcare). The protein was then washed with Ni-A buffer and eluted with Ni-B buffer [25 mM TrisHCl (pH 7.8), 300 mM imidazole]. Subsequently, the His_6-tag was cleaved by the addition of TEV protease (*ca.* 1:25 molar ratio of TEV to p47), followed by dialysis against a buffer containing 50 mM TrisHCl (pH 7.5), 100 mM NaCl, 1 mM EDTA, and 2 mM DTT, overnight at 4 ^{o}C. The cleaved His_6-tag and TEV protease, containing an uncleavable His_6-tag, were then removed by applying the protein solution to a Ni affinity column and collecting the flow-through. The protein was further purified on a Superdex 200 Increase 10/300 or a Superdex 75 10/300 gel filtration column (GE Healthcare) in a buffer containing 50 mM HEPES (pH 7.5), 100 mM NaCl, and 1 mM EDTA. p97 constructs were purified as previously described (2). To generate apo-$p97_{ND1L}$, purified $p97_{ND1L}$ was treated with apyrase (New England Biolabs, 10 units per mmol of p97 protomer) overnight at room temperature, followed by purification on a Superdex 200 Increase 10/300 gel filtration column (GE Healthcare) to remove apyrase. The apo state of $p97_{ND1L}$ is indicated by an absorbance ratio at 260 nm and 280 nm (A_{260}/A_{280}) of $\leq$ 0.7. After purification, protein samples were subsequently exchanged into appropriate buffers for measurements using Amicon centrifugal concentrators (Millipore) or frozen at -80 ^{o}C for storage. A list of buffers used for various experiments is provided in Table S4.

NMR experiments NMR experiments were performed on Bruker Avance III HD 18.8 T, Bruker Avance Neo 23.5 T, and Varian 14.1 T spectrometers, all equipped with cryogenically cooled, pulse-field gradient, triple-resonance probes. All NMR spectra were processed using the *NMRPipe* suite of programs (4) and visualized using NMRFAM-SPARKY (5). Peak fitting was performed using the *nlinLS* routine in *NMRPipe*, and errors in peak heights obtained from uncertainties in the fits.

The sequential backbone resonance assignment of full-length p47 was performed using standard TROSY-based (6) triple resonance experiments (7), including HNCA, HNCO, HNCACO, HNCACB,

HNCOCACB and HNN recorded on a [U-^{2}H,^{15}N,^{13}C]-p47 sample. Backbone assignments of p47$_{1\text{-}174}$ and p47$_{171\text{-}370}$ were obtained previously (8), and verified using data sets recorded on [U-^{15}N,^{13}C]-samples with standard (non-TROSY) triple resonance experiments (HBCBCACONH, HNCACB, HNCO HNCACO, and HNN for p47$_{1\text{-}174}$; HNCACB for p47$_{171\text{-}370}$, p47$_{171\text{-}266}$, and p47$_{291\text{-}371}$). Backbone resonance assignments were mapped to different pHs when needed through a series of HNCO experiments acquired at buffer pH values of 6.0, 6.5, and 7.0. Assignments of methyl sidechains of Ile, Leu, Val and Met residues were obtained using (H)C(CO)NH-TOCSY, H(CCO)NH-TOCSY experiments (9) recorded on [U-^{15}N,^{13}C]-p47 fragments (p47$_{1\text{-}174}$, p47$_{171\text{-}370}$, p47$_{171\text{-}266}$ and p47$_{291\text{-}371}$) and transferred to full-length p47. Methyl assignments were further confirmed by recording 3D ^{13}C-edited NOSEY experiments (200 ms mixing time) on [U-^{2}H, ILVM-$^{13}CH_3$]-labeled samples of full-length p47 and p47$_{101\text{-}266}$ as well as ^{15}N-edited TROSY-based 3D NOE experiments that record either a pair of ^{1}H (or ^{15}N) chemical shifts (^{1}H(t$_1$)-NOE-^{15}N(t$_2$)-^{1}H(t$_3$), and ^{15}N(t$_1$)-NOE-^{15}N(t$_2$)-^{1}H(t$_3$)), using a [U-^{2}H,^{15}N]-full-length p47 sample. Triple resonance datasets on full-length p47 were acquired via non-uniform sampling (NUS) with Poisson gap schemes (10), using sampling densities between 15-25%; other experiments were obtained using uniform sampling in each of the indirect dimensions. For the NUS acquired datasets, either SMILE reconstruction (11) or iterative soft thresholding (12) approaches were used for data processing, both of which are integrated into the *NMRPipe* package. All assignment experiments (25 °C, an interscan delay of 1.5 s) were obtained at 18.8 T, except for HNCACB datasets for p47$_{171\text{-}266}$ and p47$_{291\text{-}370}$ which were acquired at 14.1 T.

Analysis of NMR data

(i) Chemical shift perturbations

Combined chemical shift perturbations were calculated from ^{15}N-^{1}H (or ^{13}C-^{1}H) correlated spectra according to $CSP = \sqrt{\Delta v_H^2 + \Delta v_X^2}$, in which Δv_H and Δv_X are chemical shift differences (Hz) in the ^{1}H and ^{15}N (or ^{13}C) dimensions, respectively. ^{13}C-^{1}H HMQC experiments exploiting a methyl-TROSY effect (13) were recorded on [U-^{2}H, ILVM-$^{13}CH_3$]-p47 constructs in the presence or absence of interacting partners (*e.g.* p97$_{ND1L}$, p97$_{NTD}$) to achieve high spectra quality, while for smaller proteins or protein fragments ^{15}N-^{1}H HSQC (either TROSY (6) or non-TROSY) and ^{13}C-^{1}H CT-HSQC spectra (14, 15), recorded on either perdeuterated or protonated samples, were analyzed.

(ii) *Trans* interaction between $p47_{171-370}$ and $p47_{1-174}$

In order to measure a K_d value for the SIM-SEP interaction (*trans*; in intact p47 the interaction is *cis* and therefore expected to be stronger) a titration was performed in which [U-^{2}H]-$p47_{1-174}$ was added to a solution of ILVM-$p47_{171-370}$ (maintained at 70.4 μM) in a series of 11 steps from 0 to 341 μM (*SI Appendix*, Fig. S4*A*). Assuming a one-site binding model, K_d was extracted from simultaneous nonlinear least squares fits of 9 binding curves via the following equation:

$$CSP = CSP_{max}\frac{P_T+L_T+K_d-\sqrt{(P_T+L_T+K_d)^2-4P_TL_T}}{2P_T} \qquad [S1]$$

where *CSP* is the chemical shift perturbation for a chosen peak calculated as described above, CSP_{max} is its chemical shift difference between free and fully bound $p47_{171-370}$, P_T and L_T are the total concentrations of $p47_{171-370}$ and $p47_{1-174,}$ respectively. The error in K_d was estimated from a covariance matrix analysis (16).

(iii) Titration of $p97_{ND1L}^{apo}$ with p47

Before describing the model used to fit the p97/p47 titration data let us consider first a simpler binding reaction in which ligands bind non-cooperatively to a receptor comprising *k* sites. We write the binding equilibrium for the case where *j* ligands bind as

$$P + jL \rightleftarrows PL_j \qquad K_{mac,j} = {}_kC_jK^j \qquad [S2]$$

where *P* and *L* denote the receptor and ligand, respectively, *K* is the microscopic association constant for a single binding event that is assumed to be the same for each of the *j* equivalent binding steps, $K_{mac,j}$ is the effective macroscopic association constant, and ${}_kC_j = \frac{k!}{j!(k-j)!}$. The concentration of PL_j follows directly from Eq [S2] as $[PL_j] = K_{mac,j}[P][L]^j$. The total protein and ligand concentrations are thus given by

$$\begin{aligned}[P_T] &= \sum_{j=0}^{k}[PL_j] = \sum_{j=0}^{k}K_{mac,j}[P][L]^j \\ [L_T] &= [L] + \sum_{j=0}^{k}j[PL_j] = [L] + \sum_{j=0}^{k}jK_{mac,j}[P][L]^j\end{aligned} \qquad [S3]$$

from which it follows that

$$\frac{[P_T]}{[P]} = \sum_{j=0}^{k} K_{mac,j}[L]^j$$
$$\frac{[L_T]-[L]}{[P]} = \sum_{j=0}^{k} jK_{mac,j}[L]^j \qquad \text{[S4]}$$

Eq. [S4] can be rearranged to yield

$$\sum_{j=0}^{k} K_{mac,j}[L]^{j+1} + \sum_{j=0}^{k} ([P_T]j - [L_T])K_{mac,j}[L]^j = 0 \qquad \text{[S5]}$$

which can be solved to give [*L*] and, subsequently [*P*] (Eq. [S4]) and thus the values of [PL_j] that are required in fits of the titration profile.

It is also possible to present the above binding model in a simple manner schematically by grouping the configurations that are possible for each receptor with *j* bound ligands, PL_j in the above notation, into separate states that are non-degenerate, as is done in *SI Appendix*, Fig. S9*A* for the case of a hexameric receptor (*i.e.*, p97), where each of the 6 circles illustrated in each hexamer denotes a ligand (*i.e.* p47) binding site that is either occupied (*i.e.*, UBX+SHP_C, orange) or unoccupied (blue). There are 64 (2^6) different ways that a ligand can bind but many of these binding configurations are degenerate. For example, there are 6 equivalent ways of binding a single ligand to an empty hexamer (state A), so the degeneracy for state B is 6 (*i.e.*, there are 6 equivalent configurations for state B, as depicted in *SI Appendix*, Fig. S9*A*). This can be seen simply by moving the orange circle successive steps until it returns to its initial position – this process takes 6 moves and hence state B has a degeneracy of 6. In contrast, the 15 different ways that 2 ligands can bind to state A ($_6C_j = \frac{6!}{j!(6-j)!}$ with *j*=2), can be captured by 3 different states (C-E), with degeneracies of 6, 6, and 3 for states C, D, and E, respectively. The 14 non-degenerate states A-N are indicated in *SI Appendix*, Fig. S9*A*, with the number of degenerate configurations in each state highlighted.

The above model is satisfactory for describing the binding of p47 with $p97^{ADP}$, where a single p47 ligand binds to one p97 protomer (via UBX+SHP_C motifs). However, it is inadequate when the receptor is $p97^{apo}$ as in this case a single p47 protomer can associate with the p97 receptor via either bipartite (binding of p47 UBX+SHP_C to a single NTD of p97) or tripartite (binding of UBX+SHP_C to a common NTD *and* binding of

SHP_N to an adjacent NTD) interactions. The most complete model would be one in which the binding of any combination of UBX, SHP_C, and SHP_N is taken into account (*i.e.*, just 1, any 2 or all 3), however our titrations of full length p47 with isolated NTD domains establish that UBX and SHP_C bind to p47 as a pair (*SI Appendix*, Fig. S5). Further, binding of $p47_{1\text{-}170}$ (containing the SHP_N motif) with $p97_{ND1L}$ could not be observed in ITC studies, suggesting that the isolated interaction involving SHP_N is very weak, and a comparative analysis of the binding of full length p47 with $p97^{apo}$ and $p47_{171\text{-}370}$ with $p97^{apo}$ based on ITC data (Fig. 5*B*) is also consistent with weak binding of SHP_N alone. We have therefore considered only bipartite and tripartite interactions in the formulation of the binding model for $p97^{apo}$.

Although there is no advantage to separating non-degenerate states as is done in *SI Appendix*, Fig. S9*A* when each binding event involves only a single site on the receptor (*i.e.*, all binding events are bipartite, for example), the advantage becomes clear for the analysis of the p47-$p97^{apo}$ titration where each interaction is either bipartite or tripartite. Thus, the assumption that each binding event is equivalent, an inherent feature of the above model, is no longer valid. The utility of describing binding in terms of the non-degenerate states of *SI Appendix*, Fig. S9*A* when both bipartite and tripartite interactions are allowed can be made most clear by considering an example. Let us focus on the binding of 3 p47 molecules to $p97^{apo}$. There are 4 unique states corresponding to this situation, denoted by F-I in *SI Appendix*, Fig. S9*A*. States F-H are each six-fold degenerate, as there are six equivalent representations for each, while state I is two-fold degenerate, for a total of 20 different ways of binding 3 ligands ($_6C_3$). However, when binding of SHP_N is also included it becomes clear that additional configurations are possible that vary with each state. In order to count the number of configurations in each state we assume that, as described above, each binding event must include UBX+SHP_C and that if binding of SHP_N occurs it must be to an adjacent unoccupied protomer, as simultaneous binding of SHP_N and SHP_C to a single NTD domain cannot occur (both SHP motifs have the same p97 binding site). *SI Appendix*, Fig. S9*B*&*C* focuses on states F and G, both with three bound p47 ligands, illustrating how the SHP_N binding permutations are taken into account. We denote SHP_N domains by arrows; an arrow pointing to a blue (unoccupied) site indicates SHP_N binding to that site, while an arrow that is pointing away from the receptor indicates no binding of the SHP_N. Thus, in the case of state F, where there is a 6-fold degeneracy associated with UBX+SHP_C binding in the manner indicated by the orange circles, there are additional configurations depending on whether 0 (1

configuration), 1 (2 configurations) or 2 (1 configuration) SHP_N motifs bind. These additional configurations are highlighted in *SI Appendix*, Fig. S9*B* and summarized in Table S2. Note that it is not possible for 3 SHP_N motifs to bind simultaneously in this case. In total there are 6×1 + 6×2 + 6×1 = 24 configurations for state F that includes all cases where SHP_N is or is not bound. The situation is more complex for state G. As for state F, there are 6 initial configurations for binding involving only UBX+SHP_C (orange circles), yet 6×1, 6×4, 6×4 and 6×1 configurations when 0, 1, 2, or 3 SHP_N motifs bind, for a total of 60 configurations. By expanding each of the states indicated in *SI Appendix*, Fig. S9*A* to take into account SHP_N binding as is illustrated in *SI Appendix*, Fig. S9*B*&*C*, expressions for $[PL_j]$ can be derived, as above for the simpler model. In this case

$$[PL_j] = [PL_j]_0 + [PL_j]_1 + [PL_j]_2 + [PL_j]_3 \qquad \text{[S6]}$$

where $[PL_j]_q$ is the concentration of p97, *P*, with *j* bound p47 ligands (UBX+SHP_C), *L*, and *q* bound SHP_N motifs. Note that $q<=j$ and, in many cases, such as for state F, described above, $q<j$. Thus,

$$[PL_j] = \sum_{s\in states} (a_s b_{s,0} K_B^j + a_s b_{s,1} K_B^{j-1} K_T + a_s b_{s,2} K_B^{j-2} K_T^2 + a_s b_{s,3} K_B^{j-3} K_T^3)[P][L]^j \qquad \text{[S7]}$$

where the summation is over each of the states for which *j* ligands are bound (for example, 3 or 4 states when *j* = 3 or 4, *SI Appendix*, Fig. S9*A*), a_s is the number of configurations associated with each of the states (*i.e.*, the degeneracy, *SI Appendix*, Fig. S9*A*, a_s = 1, 6, 6, 6, 3,… for states A, B, C, D, E, …), neglecting SHP_N binding, $b_{s,q}$ is a factor listed in Table S2, that takes into account the binding of *q* SHP_N motifs (for example, for state F, $b_{s,0}$ = 1, $b_{s,1}$ = 2, $b_{s,2}$ = 1, $b_{s,3}$ = 0), and K_B and K_T are microscopic association constants for bipartite and tripartite binding, respectively. Therefore, when *j* = 3,

$$\begin{aligned}
[PL_3] &= [\sum_{s\in states} (a_s b_{s,0} K_B^3 + a_s b_{s,1} K_B^2 K_T + a_s b_{s,2} K_B^1 K_T^2 + a_s b_{s,3} K_T^3)][P][L]^3 \\
&= (6\times 1 K_B^3 + 6\times 2 K_B^2 K_T + 6\times 1 K_B^1 K_T^2)[P][L]^3 \quad \text{state F} \\
&+ (6\times 1 K_B^3 + 6\times 4 K_B^2 K_T + 6\times 4 K_B^1 K_T^2 + 6\times 1 K_T^3)[P][L]^3 \quad \text{state G} \\
&+ (6\times 1 K_B^3 + 6\times 4 K_B^2 K_T + 6\times 4 K_B^1 K_T^2 + 6\times 1 K_T^3)[P][L]^3 \quad \text{state H} \\
&+ (2\times 1 K_B^3 + 2\times 2 K_B^2 K_T + 2\times 3 K_B^1 K_T^2 + 2\times 2 K_T^3)[P][L]^3 \quad \text{state I} \\
&= (20 K_B^3 + 64 K_B^2 K_T + 60 K_B^1 K_T^2 + 16 K_T^3)[P][L]^3
\end{aligned} \qquad \text{[S8]}$$

To summarize:

$$[PL_0] = [P] = \alpha_0[P]$$
$$[PL_1] = (6K_B + 12K_T)[P][L]^1 = \alpha_1[P][L]^1$$
$$[PL_2] = (15K_B^2 + 42K_BK_T + 33K_T^2)[P][L]^2 = \alpha_2[P][L]^2$$
$$[PL_3] = (20K_B^3 + 64K_B^2K_T + 60K_B^1K_T^2 + 16K_T^3)[P][L]^3 = \alpha_3[P][L]^3 \quad \text{[S9]}$$
$$[PL_4] = (15K_B^4 + 42K_B^3K_T + 33K_B^2K_T^2)[P][L]^4 = \alpha_4[P][L]^4$$
$$[PL_5] = (6K_B^5 + 12K_B^4K_T)[P][L]^5 = \alpha_5[P][L]^5$$
$$[PL_6] = K_B^6[P][L]^6 = \alpha_6[P][L]^6$$

Note that in the limit that $K_T \rightarrow 0$, that is, when there is no binding of the SHP_N motif such that each ligand binds in only a single way (UBX+SHP_C; 1 ligand binds to 1 NTD), this model reverts to the simple binding model considered initially (Eq [S2]) and $[PL_j] = {}_6C_jK^j[P][L]^j$.

Finally, we can write

$$[P_T] = \sum_{j=0}^{6}[PL_j] = \sum_{j=0}^{6}\alpha_j[P][L]^j$$
$$[L_T] = [L] + \sum_{j=0}^{6} j[PL_j] = [L] + \sum_{j=0}^{6} j\alpha_j[P][L]^j \quad \text{[S10]}$$

from which it follows that

$$\sum_{j=0}^{6}\alpha_j[L]^{j+1} + \sum_{j=0}^{6}([P_T]j - [L_T])\alpha_j[L]^j = 0 \quad \text{[S11]}$$

The solution to Eq. [S11] yields [*L*] from which [*P*] is obtained from Eq [S10], and [PL_j] from Eq [S9].

In the NMR titration data of Fig. 6 the intensities of cross-peaks corresponding to the free state are fit to an equation of the form $I_{free} = I_o[L]$, where [*L*] is calculated as described above, to obtain the fitting parameters K_B, K_T and I_o. Reduced χ^2 values were calculated according to $\chi^2 = \frac{1}{pn-m}\sum_{j=1}^{p}\sum_{i=1}^{n}(\frac{I_{i,j}^{\exp} - I_{i,j}^{fit}}{\sigma_{i,j}})^2$, where *n* is the number of titration points, *m* is the number of fitting parameters, *p* is the number of peaks fit globally, and $\sigma_{i,j}$ is the error in the measured peak intensities in NMR spectra, estimated by recording multiple repeats of spectra (5×).

(iv) Measurement of inter-molecular NOEs

Purified [U-^{2}H,Ile(δ1),Met-$^{13}CH_3$]-p47 and [U-^{2}H,^{15}N,Leu,Val-$^{13}CH_3$,$^{12}CD_3$]- p47 samples were mixed in a 1:1 molar ratio in a buffer containing 6 M GdnHCl, 50 mM HEPES (pH 7.5), 100 mM NaCl, and 1 mM EDTA, to a final protein concentration of 50 μM in a total volume of 8 mL. The unfolded mixture was then refolded via a 20-fold fast dilution step into a refolding buffer containing 50 mM HEPES (pH 7.5), 100 mM NaCl and 1 mM EDTA, and concentrated prior to purification via a Superdex 200 Increase 10/300 gel filtration column (GE Healthcare). Fractions were collected and exchanged to a buffer containing 25 mM HEPES (pD 7.4), 25 mM NaCl, 1 mM EDTA, 99.9% D_2O for NMR measurements. A 3D ^{13}C-edited NOSEY experiment (200 ms mixing time) was recorded on this GdnHCl-mixed p47 sample at 18.8 T, 25 °C.

GdnHCl-mixed p47 and p47$_{1\text{-}290}$ pull down assays Purified His$_6$-tagged full-length p47 and His$_6$-less (no tag) p47$_{1\text{-}290}$ were mixed in a 1:1 molar ratio in a buffer containing 6 M GdnHCl, 50 mM HEPES (pH 7.5), 100 mM NaCl, and 1 mM EDTA, to a final protein concentration of 60 μM. The unfolded mixture was then refolded via an 18-fold fast dilution step into a refolding buffer containing 50 mM HEPES (pH 7.5), 100 mM NaCl and 1 mM EDTA, and concentrated prior to separation via a Ni affinity column. After loading the refolded mixture, the column was washed with 3 column volumes of Ni-A buffer and the His$_6$-tagged full-length p47 eluted with Ni-B buffer. The flow-through and wash fractions were combined and concentrated to the same volume as the eluate. A control sample was made by mixing the same amount of His$_6$-tagged p47 and His$_6$-less (no tag) p47$_{1\text{-}290}$ under native conditions (no GdnHCl in the buffer) immediately followed by Ni purification. The results are readout by SDS-PAGE.

Dynamic light scattering Translational diffusion rates of p47 and p47$_{GB1}$ were measured by dynamic light scattering on a DynaPro DLS Plate Reader III (Wyatt Technology). Purified proteins were exchanged into a buffer containing 25 mM HEPES (pH 7.4), 50 mM NaCl and 1 mM EDTA and diluted to designated concentrations. Samples were spun down at 20,000 × g at 4 °C for 30 min to remove large aggregates or impurities prior to loading on a 384-well plate. The plate was then subjected to a 5-min 3,000 × g centrifugation to remove air bubbles from the wells. Measurements were carried out 20 times per well before averaging, with 4.8-s acquisitions at 25 °C. Resulting autocorrelation functions were fitted with the equation $g^{(2)}(\tau) = B + \beta e^{-2\Gamma\tau}\left(1 + \frac{\mu_2}{2}\tau^2\right)$(17), where B is a term accounting for the baseline, β is the correlation function amplitude at

$\tau = 0$, $\Gamma = Dq^2$, and μ_2 is the 2nd moment of the distribution of Γ values about their mean. In the equation for Γ, D is the translational diffusion coefficient and q is the scattering wave vector given by $q = \frac{4\pi n}{\lambda_0} sin\left(\frac{\theta}{2}\right)$, where n is the solvent refractive index (n=1.3347 was used), λ_0 is the wavelength of the light used by the instrument, and θ is the scattering angle. D, β and B and μ_2 were output as fitted parameters.

SAXS Small angle X-ray scattering (SAXS) profiles were collected on an Anton Paar SAXSpace instrument, configured for line collimation and equipped with a 1D Mythen2 CMOS detector and a temperature-controlled TCStage. p47, $p97_{NTD}$, and $p97_{ND1L}$ samples for SAXS measurements were dialyzed overnight at 4 °C into a buffer containing 25 mM HEPES (pH 7.4), 150 mM NaCl, 1 mM EDTA, and 1 mM TCEP. Dialyzed samples were harvested, diluted to concentrations used in experiments with the same buffer, and centrifuged for 10 minutes at 20,000 × g, 4 °C to pellet any impurities. In the case of p47-$p97_{ND1L}$ complexes, p47 was mixed at a 3:6 molar ratio (protomer:protomer) with $p97_{ND1L}$, such that the final concentration of the complex was either 5 mg/mL or 10 mg/mL. Sample polydispersity was assessed by DLS as described above, and was found to be < 5 %. 60 µL sample volumes were then loaded into a 1 mM quartz capillary, transferred to the TCStage, and irradiated for a total acquisition time of 1 hour. The sample temperature was maintained at 25 °C. Data were de-smeared using the Lake method (18), and the reference SAXS profile of HEPES buffer was subtracted from profiles of the protein samples. R_g values were obtained by the Guinier approximation (19). For p47 and $p97_{NTD}$ samples, SAXS profiles were further analyzed with the V_c method described by Rambo and Tainer (20) in order to obtain assessments of protein molecular weight. Standard errors of the calculated molecular weights were obtained from Monte Carlo simulations of 1000 synthetic intensity profiles generated on the basis of experimental uncertainties of scattering intensity reported by SAXSpace software (16).

Analytical ultracentrifugation Sedimentation velocity experiments were carried out using a ProteomeLab analytical ultracentrifuge (Beckman Coulter, Indianapolis). Samples were diluted from a concentrated p47 stock to final concentrations between 0.2 to 20.4 mg/mL in a buffer containing 25 mM HEPES (pH 7.4), 50 mM NaCl and 1 mM EDTA. Depending on the concentration, samples were loaded into cell assemblies with sapphire windows and double sector 12 mm Epon, or 4 mm narrow or standard width 3D printed Microfine Green centerpieces (21) at volumes to generate 12 mm long solution columns. The cell assemblies were placed in an 8-

hole An-50 Ti rotor and temperature equilibrated at rest in the evacuated rotor chamber at 19.6 °C for at least 2 h. Acceleration to 49,000 rpm was immediately followed by data acquisition using the Rayleigh interference optical detection system. Data were corrected for scan time errors (22) using the software *REDATE* and analyzed using the *c*(*s*) model in *SEDFIT* (23). *REDATE* was kindly provided by Dr. Chad Brautigam. For direct comparison with DLS data of p47 as described in the main text, sedimentation coefficients were corrected to 25°C, assuming the buffer viscosity to be approximated by that of water.

Isothermal titration calorimetry ITC measurements were performed using a Malvern Auto-ITC200 calorimeter. Protein samples comprising $p97_{ND1L}$, p47, $p47_{171-370}$, $p47_{F150A}$, $p47_{1-290}$, and $p47_{F253A}$ were dialyzed overnight into a buffer containing 25 mM HEPES (pH 7.4), 150 mM NaCl, 1 mM EDTA, and 2 mM TCEP. Dialyzed samples were harvested and centrifuged for 10 minutes at 20,000 × g, 4 °C to pellet any impurities. For titrations of p47, $p47_{171-370}$ and $p47_{F150A}$ with $p97_{ND1L}$, samples were diluted to concentrations of 15 μM for $p97_{ND1L}$ (monomer concentration) and 150 μM for p47 (and p47 variants) with the same buffer, while in the case of titrations involving $p47_{1-290}$ and $p47_{F253A}$, 30 μM samples of $p97_{ND1L}$ (monomer concentration) and 600 μM samples of p47 variants were prepared. $p97_{ND1L}$ and p47 solutions were then loaded into the ITC200 cell and syringe, respectively, using the Auto-ITC200 auto-sampling and transfer apparatus. In cases where the ADP state of $p97_{ND1L}$ was desired, samples were supplemented with ADP to 2 mM. Titration experiments consisted of a 1 × 0.2 μL injection, followed by 19 × 2 μL injections with inter-injection delays of 180 s, using a reference power of 12 μcal/s. ITC profiles were corrected for baseline distortions and integrated to obtain heats of injection and further corrected for contributions from heats of dilution. Titration data were fit (i) to an N-equivalent sites binding model, parameterized by an equilibrium association constant ($K_a = K_d^{-1}$), an enthalpy of binding (ΔH), and a binding stoichiometry (N) or (ii) to a more complex scheme in which ligand can bind via either a bipartite or a tripartite mechanism, as described in detail above. Notably, model (i) can be readily derived from model (ii) by setting $K_T = 0$ in Eqs [S7-S9].

Below we outline the fitting procedure, initially for the case of an N-equivalent sites model. Consider an ITC cell of volume V_o with an initial protein concentration of $[P]_{T,0}$ and without added ligand. Each injection comprises the addition of a volume Δ of a ligand solution of concentration M_S to the cell; during this addition a volume Δ of solution in the cell is removed. It can be shown that after the i^{th} addition (24, 25)

$$[P]_{T,i} = [P]_{T,0}(1-\frac{\Delta}{V_o})^i = \sum_{j=0}^{6}[PL_j]$$
$$[L]_{T,i} = M_S\{1-(1-\frac{\Delta}{V_o})^i\} = [L] + \sum_{j=1}^{6} j[PL_j] \qquad \text{[S12]}$$

where $[PL_j]$ is the concentration of p97 bound with j ligands, as described above. Eq [S12] provides the initial conditions from which the concentration of bound ligands after the ith titration point can be calculated, using the model of choice via $[L]_{Bd,i} = \sum_{j=1}^{6} j[PL_j]$, where $[PL_j]$ values are obtained from the binding model that is considered. Immediately after addition of the ith aliquot but before the binding reaction can occur the concentration of bound ligand is $[L]_{Bd,i}^{before} = [L]_{Bd,i-1}(1-\frac{\Delta}{V_o})$, where the factor $(1-\frac{\Delta}{V_o})$ takes into account the dilution from the addition of the ith aliquot and concomitant removal of a volume Δ of the solution. The change in concentration of bound ligand from the addition of the ith aliquot is thus given by

$$\delta = [L]_{Bd,i} - [L]_{Bd,i}^{before}, \qquad \text{[S13]}$$

which is related to the heat released from the ith addition, q_i, via, $q_i = \delta V_o \Delta H^o$, where ΔH^o is the molar enthalpy of the binding reaction. The heat per mole of added ligand in the ith titration step is thus given by $q_i/(M_S\Delta)$.

The situation is only slightly more complicated for the bipartite/tripartite model. In this case it is necessary to separate $[L]_{Bd,i}$ into contributions from ligands bound in a bipartite (B; $[L]_{Bd,i,B}$) or tripartite (T; $[L]_{Bd,i,T}$) manner, as the ΔH^o values associated with these separate binding interactions can be different. For example, for the case of j=3, and following Eq [S8], we obtain

$$[PL_3] = (20K_B^3 + 64K_B^2K_T + 60K_B^1K_T^2 + 16K_T^3)[P][L]^3$$
$$[L]_{Bd,i}^{j=3} = 3[PL_3]$$
$$[L]_{Bd,i,B}^{j=3} = 3(20K_B^3 + \frac{2}{3}64K_B^2K_T + \frac{1}{3}60K_B^1K_T^2)[P][L]^3 \qquad \text{[S14]}$$
$$[L]_{Bd,i,T}^{j=3} = 3(\frac{1}{3}64K_B^2K_T + \frac{2}{3}60K_B^1K_T^2 + 16K_T^3)[P][L]^3$$

from which values of δ_B and δ_T can be calculated and $q_i = \delta_B V_o \Delta H_B^o + \delta_T V_o \Delta H_T^o$.

Best-fit parameters for K_a, ΔH, and N (for N-equivalent sites), or K_B, K_T, ΔH_B, ΔH_T (for bipartite/tripartite model, including an additional parameter, ζ, accounting for potentially small concentration errors in $p97_{ND1L}$ such that $[P]_{T,0} = \zeta[p97]_{ND1L}$) were obtained by minimization of the sum of squared residuals $\chi^2 = \sum_{i=1}^{k} (q_i^{calc} - q_i^{exp})^2$, where k is the number of aliquots and q_i^{calc} and q_i^{exp} are the calculated and experimental heats, respectively.

Sample preparation for electron cryo-microscopy The ADP bound p47-p97 complex was prepared by mixing apo p97 with a molar excess of p47, followed by incubation for 5 minutes before purification via a Superdex 200 size exclusion column in a buffer containing 50 mM HEPES (pH 7.5) and 150 mM NaCl. The peaks corresponding to the p47-p97 complex were then pooled and concentrated to ~8 mg/mL before addition of 5 mM ADP and 5 mM $MgCl_2$. The ATP bound p47-p97 complex was prepared by mixing apo p97 with an equimolar amount of p47 at a final concentration of ~8 mg/mL and adding 5 mM ATP and 5 mM $MgCl_2$ before freezing. For both samples 0.025% v/v IGEPAL-CA630 was added to increase the proportion of particles adopting side views on the grid. 2.5 μL samples were applied to nanofabricated holey gold grids (26–28) with a hole size of ~1 μm. Grids were blotted on both sides using a modified FEI Vitrobot mark III for 4.5 seconds at 4 °C and ~100% relative humidity before freezing in a liquid ethane/propane mixture (29).

Electron microscopy data collection The ATP bound p47-p97 complex was imaged with a Thermo Fisher Scientific Titan Krios G3 microscope operating at 300 kV and equipped with a prototype FEI Falcon 4 DDD camera used in electron counting mode at 250 frames/sec. Counting mode movies consisting of 30 exposure fractions, were obtained over 9.6 s with defocuses ranging from 0.9 to 2.5 μm. Movies were at a nominal magnification of 75000× corresponding to a calibrated pixel size of 1.03 Å and with an exposure rate of ~5 electrons/pixel/s, giving a total exposure of ~45 electrons/Å^2. 4702 movies were collected using the microscope's *EPU* software. During collection, data quality was monitored using *cryoSPARC Live* (30). Both ADP bound and apo p47-p97 maps were calculated from data obtained using a FEI Tecnai F20 electron microscope operating at 200 kV and equipped with a Gatan K2 Summit direct detector device camera used in electron counting mode at 400 frames/sec. Movies consisting of 30 exposure fractions over 15 s were obtained

with defocuses ranging from 1.5 to 3.0 μm. Movies were at a nominal magnification of 25000× corresponding to a calibrated pixel size of 1.45 Å and with an exposure rate of 5 electrons/pixel/s, and a total exposure of 35 electrons/Å^2. For the ADP and apo datasets 330 and 278 movies were collected, respectively, using Digital Micrograph software.

EM image analysis Except for use of a custom Python program as described below, all image analysis was done within the *cryoSPARC* v2 package (30). For the ATP bound dataset, patch based alignment and exposure weighting was done with a 10×10 grid, and the resulting averages of frames were used for patch based contrast transfer function (CTF) determination and particle picking. Templates for particle selection were generated by 2D classification of manually selected particles. Particle images were extracted, and local motion correction was performed in 350×350-pixel boxes with an implementation of the alignparts_lmbfgs algorithm (31). 527,145 particle images were initially picked, and after 2D classification to remove images of ice or damaged particles, 420,679 particles were carried forward for further analysis. An *Ab initio* map was calculated using a subset of ~60,000 particle images and used as a template for refinement. Refinement with all 420,679 particles was run with C6 symmetry and led to a reconstruction at 3.6 Å. Refinement of the same particles with C1 symmetry led to a reconstruction at 3.9 Å (as shown in Fig. 7*B*). We then focused on classifying different configurations of the complex based upon whether NTDs of p97 are bound with p47 UBX domains. As the UBX and SEP portions of p47 were at much lower resolution, for the subsequent analysis particle images were Fourier-cropped to a box size of 160×160. Symmetry expansion (32) was performed on the dataset to classify the asymmetric units. The dataset was then divided into three independent groups and 3D variability analysis based on a principle component analysis algorithm (33) was performed for each group within a masked area including the NTD and UBX densities. The resulting clusters were differentiated by the presence or absence of density for the p47 UBX domain, with ~57% classified as having a UBX bound. Symmetry related asymmetric units in a complex were then compared using a custom-written Python program, and complexes were identified as belonging to one of the 14 possible UBX-p97 conformations (*SI Appendix*, Fig. S12*A*&*B*). The symmetry expansion was then "collapsed" by choosing a single particle image for each complex, and the sorted particles were reconstructed using the prior Euler angles from the previous refinement, with local refinement of angles in *cryoSPARC*, recapitulating the expected conformations (*SI Appendix*, Fig. S12*C*).

For ADP bound p47-p97 the same preprocessing steps were applied as described above and 49,246 particle images were extracted in 256×256 pixel boxes for further analysis. After 2D classification 38,495 particles images were used in refinement with C1 symmetry which led to a reconstruction at 4.7 Å resolution. Similar symmetry expansion analysis described for the ATP bound dataset was performed. In the ADP bound dataset 56% of the NTD positions had UBX bound, however not enough particles images were available to accurately reconstruct the conformations as was done for the ATP dataset.

For apo p47-p97 the same preprocessing steps were applied as described for the other datasets. 59,635 particle images were extracted in 256×256 pixel boxes for further analysis. After 2D classification 29,180 particles were used in refinement with C1 symmetry leading to a reconstruction at 6.14 Å resolution.

Molecular dynamics simulations

(i) Modelling monomeric p47

There are no available structures of full-length p47. Therefore, an atomic model of monomeric p47 was generated, using XPLOR-NIH version 2.52 (34), in which the flexible linkers are fully extended in order to minimize the probability of linker entanglement upon docking to p97. Individual p47 UBA (PDB ID: 1V92 (35)) and SEP (PDB ID: 1VAZ (35)) domains, as well as the UBX domain from the p97 ND1-UBX X-ray structure (PDB ID: 1S3S (36)) were manually positioned in Cartesian space such that inter-domain distances were approximately 100 Å and 25 Å between the UBA and SEP domains and between the SEP and UBX domains, respectively. Positions of non-hydrogen atoms from homologous residues in the individual domains were used as initial coordinates for the corresponding atoms in the full-length mouse p47 polypeptide, as well as coordinates for harmonic positional restraints with a force constant of 1 kcal mol^{-1} $Å^{-2}$; these were applied to enforce the correct structures of the domains in full length p47 during the course of the MD trajectories that followed. Atoms corresponding to residues within flexible linker regions were positioned based on their atomic indices such that they did not overlap with the folded domains. The system was then minimized using successive rounds of molecular dynamics (MD) within the IVM coordinate space (37), using uniform atomic masses of 100 amu and frictional coefficients of 10 ps^{-1}. In each round of MD, the Verlet algorithm was used for integration at timesteps

of 0.5 fs for a total simulation time of 5 ps, and the temperature was maintained at 300 K. For the first MD simulation, only bond length, bond angle, and harmonic energy terms were considered. The improper dihedral term and the repulsive van der Waals term (with a force constant of 4 kcal mol^{-1} $Å^{-2}$ and scale factor of 0.9) were applied in succession over the second and third rounds of MD. A final round of minimization was then performed using the Powell method (38) implemented in XPLOR-NIH.

This full-length p47 model was then used as a starting point for extension along one axis of Cartesian space. In this instance, the UBA, SEP and UBX domains were treated as rigid bodies within the IVM. Harmonic positional restraints were applied to the UBA and UBX domains, such that the UBA and UBX domains would experience potential forces toward the origin or toward 1000 Å from the origin, respectively, with corresponding force constants of 1 kcal mol^{-1} $Å^{-2}$ and 10^{-4} kcal mol^{-1} $Å^{-2}$. The difference in force constants between the UBA and UBX domains is necessary to prevent any significant deviation of the UBA domain from the origin.

(ii) Docking of the 3:6 p47-p97^{apo} complex

An initial atomic model of a 3:6 p47-p97^{apo} complex was then built with XPLOR-NIH (34, 39), using a stepwise docking and refinement protocol detailed below. A 3:6 stoichiometry is only one of several that are possible for a given concentration of ligand and receptor. For example, assuming physiological concentrations for p47 and p97 (see text) and that all p97 is available for interaction with p47 exclusively, the fraction of p97 bound with 0, 1, 2, or 3 ligands can be calculated to be 25%, 49%, 23% and 2.7% in the apo state and 69%, 26%, 4.1% and 0.3% in the ADP state, using bipartite and tripartite affinities measured via ITC. At higher p47/p97 concentrations, such as those used in some of our NMR studies and achieved at the endpoints of ITC titrations, p97 is fully bound, as described in the text. Indeed, cryo-EM models show that p97 in the ATP state can be fully bound with p47 (6:6). Initial coordinates for the p97 NTD-up state (referred to in what follows as p97^{up}) were obtained from the 3.3 Å cryo-EM structure of a p97$^{ATP\gamma S}$ complex (PDB ID: 5FTN (40)), and regularized using built-in XPLOR-NIH routines with default parameters. However, the NTDs of the cryo-EM structure show poor correspondence with NTDs in either of the published X-ray structures of $p47_{UBX}$-$p97_{ND1}$ (PDB ID: 1S3S, (36)) or $UFD1_{SHP}$-p97 (PDB ID: 5C1B, (41)), with C^{α} RMSDs of 2.2 Å and 4.1 Å for the $p47_{UBX}$-$p97_{ND1}$ and $UFD1_{SHP}$-p97 models, respectively. To ensure that there is no atomic overlap or non-ideal covalent geometry upon p47 docking (to be discussed subsequently), the NTDs of the regularized cryo-EM model were first refined against

those in the $p47_{UBX}$-$p97_{ND1}$ X-ray structure, minimizing the overall NTD RMSD without altering the positions of these domains with respect to the D1 ring. This was accomplished through the XPLOR-NIH posDiffPot term which enforces a square-well potential based on the RMSD of equivalent atom selections between the simulated structure and the target structure within a simulated annealing molecular dynamics (MD) protocol (42). In this instance, only NTD atoms corresponding to residues 22 through 188 were used in the posDiffPot calculation with a force constant of 1 kcal mol^{-1}; full translation and rotation of this NTD region was allowed. The remaining atoms were fixed in position to reduce computational complexity, as well as to ensure that the NTDs do not deviate from the up conformation. Bond lengths, angles, and improper dihedrals were enforced with their appropriate potential energy terms, and the RepelPot potential term was used to prevent atomic overlap (42). An initial MD annealing was conducted for 5 ps at a bath temperature of 3000 K, a step size of 1 fs, and a RepelPot force constant of 0.002 kcal mol^{-1}. The bath temperature was reduced to 100 K in steps of 30 K, with 0.2 ps of molecular dynamics at each temperature step. Concurrently, the force constant of the RepelPot term was scaled up to 1 kcal mol^{-1} in steps of 0.01 kcal mol^{-1}. A final round of Powell minimization was then performed; the final C^{α} RMSD between NTDs in the $p97^{up}$ model and NTDs in the $p47_{UBX}$-$p97_{ND1}$ structure was ~0.5 Å.

In order to construct an initial model of the p47-$p97^{up}$ complex, intermolecular distance restraints for the p47-$p97_{NTD}$ interactions (which includes both UBX and SHP components) were first generated based on the published X-ray structures of $p47_{UBX}$-$p97_{ND1}$ and $UFD1_{SHP}$-p97 mentioned above. In either case, only intermolecular atomic distances of 6 Å or less for homologous residues were used as input and enforced with the XPLOR NOE potential term, with a restraint tolerance of ± 1 Å. (Note that the p47 used in our experiments and in all calculations is from *Mus musculus*, with a SHP_C domain distinct from $UFD1_{SHP}$ used to model the SHP-NTD interactions. In contrast, the NTD and UBX domains are identical between mouse and rat proteins used in the present work and in the X-ray studies, respectively). The coordinates of the aforementioned extended p47 monomer structure were imported into the $p97^{up}$ model, duplicated twice and positioned such that three p47 molecules were separated by 120° about the axis of the D1/D2 central pore. The intermolecular distance restraints between p47 and $p97_{NTD}$ were applied to each of the p47 molecules, so as to encourage binding of the UBX and SHP_C to a single NTD, and binding of SHP_N to an adjacent NTD. Notably, we have no experimental evidence to support a defined handedness of SHP_N binding with respect to the UBX+SHP_C binding site. In our

simulations we constrained SHP_N to interact with the NTD located counterclockwise from the NTD bound to UBX+SHP_C of the same p47 ligand. Intramolecular restraints for the SIM-SEP interaction, quantified in the present study on the basis of CSPs and NOEs, were also included in the initial model building phase. Ambiguous distance restraints for SIM residues F123, K124, G125, G130 and SEP residues L184, K185, Q202, A207, I213, R215, G216, E217, V218, H227, D234, M235, E236, and D237 based on CSPs were calculated and enforced as described by Clore and Schwieters (43). At this stage the NOEs were limited (more NOEs were subsequently included) to an interaction between the Cδ1/δ2 of L122 and Cδ1/δ2 of L210 using ambiguous restraints of 4 Å, with upper and lower tolerances of 1 Å and 2.8 Å, respectively; stereospecific assignments were initially not available for p47 methyl groups. Restraints were enforced using the XPLOR NOE soft square-well potential term with default parameters. The three p47 monomers were then each simultaneously docked across two neighboring p97 protomers in a 120 ps MD simulation with a bath temperature of 300 K and a time step of 2 fs. As before, bond length, angle, and improper dihedrals were enforced, and the RepelPot potential term was used with a force constant of 1 kcal mol^{-1} to prevent atomic overlap. p47 UBA, SEP and UBX domains were treated as rigid bodies, and the p97 D1 and D2 domains as well as the backbone atoms of the NTDs were fixed in position to reduce computational complexity. The system was then further minimized with the Powell method. A final round of MD and minimization was then performed, wherein the XPLOR COLLapse potential was applied to p47 and to atoms from p97 residues within the NTD and D1 domains and appended to the list of potential energy terms so as to enforce agreement with the measured experimental radius of gyration (56.5 Å) obtained from SAXS measurements. In this case, the MD consisted of 40 ps with 2 fs steps at a bath temperature of 300 K.

The initial structure, obtained as described above, was then aligned with the cryo-EM density map of the p47-$p97^{ATP}$ complex, after which the coordinates for the D2 domain (residues 461-768) and their corresponding voxels in the electron density map were removed to reduce the computational expense for further simulations. Additionally, the UBA domain (for which we have no direct restraints to position it with respect to the remainder of the complex) was moved away from the D1 ring by rotating backbone bonds in the disordered UBA-SEP linker region. The cryo-EM density map was then converted into a Gaussian Mixture Model (GMM) through an iterative divide-and-conquer process (44). The map was fit to an initial GMM composed of 220 Gaussians,

which was then refined into the final GMM composed of 48383 Gaussians. The final GMM's cross correlation with the cryo-EM map was 0.993.

(iii) Preliminary MD simulations

Simulations were subsequently carried out in GROMACS 2018.4 (45, 46) patched with the open-source PLUMED library version 2.5 (47, 48) and the EEF1-modified CHARMM36 force field (49, 50). The structure was energy minimized using the steepest descent and conjugate gradient methods with target maximum forces of 1000 kJ mol^{-1} and 100 kJ mol^{-1}, respectively. The minimized structure was equilibrated in the constant number, volume, and temperature (NVT) ensemble for 50 ps at 298 K using a Langevin dynamic integrator with a friction coefficient of 1 ps^{-1}. Position restraints were applied to heavy atoms. Solvent was modeled with the EEF1 solvent-accessible surface area based model (50, 51) applied to the protein heavy atoms using the EEFSOLV keyword in PLUMED. A timestep of 1 fs was used with LINCS constraints on all bonds (52). van der Waals interactions were switched smoothly between 0.7 and 0.9 nm. Coulomb interactions were tabulated with a distance dependent dielectric constant and a cutoff at 0.9 nm, as in Löhr et al (53).

From the equilibrated structure, a preliminary 1 ns simulation with a timestep of 2 fs was carried out to generate initial conformations. Neighbor lists for van der Waals and Coulomb interactions were updated every 10 steps, with cutoffs of 0.9 nm. Distance restraints between α-carbons of UBA and SEP domains were used to maintain their structures to within a tolerance of 0.1 nm using a force constant of 500 kJ mol^{-1} nm^{-2}, while UBX, SHP_C, SHP_N, and p97 were fixed in position using the freeze group option in GROMACS. In order to more closely match the measured R_g value, UBA domains were pulled toward the bottom of the D1 ring by enforcing an intermolecular distance constraint of 0.6 nm between residues K20 and T375. As discussed subsequently, none of the simulations we have done allowed the UBA domains to fully explore their allowed conformational space; R_g values can also be satisfied when these domains are positioned above D1, proximal to the SEP domains.

(iv) Production MD simulations

Conformations were extracted from the preliminary simulations as starting points for production simulations, using metainference (54), a Bayesian framework that models the structural ensemble while correcting for force field errors with experimental data. Thus, restraints were required to be satisfied as averages

over the complete ensemble, rather than by any single structure, so as to account for the dynamic nature of the linker regions of p47. During these computations, p97 was fixed in position using the freeze group option in GROMACS, while the UBA, SEP, and UBX domains were restrained to their solution structures using distance restraints between alpha carbons (as done in the preliminary simulation). Inter-molecular UBX/SHP-NTD distances were restrained as described for the generation of the initial p47-p97 model. A simulation time step of 2 fs was employed, and configurations were saved every 1 ps. Seven independent runs, each containing 48 replicates, were carried out for a total time of 1008 ns.

Experimental restraints derived from cryo-EM, chemical shifts, NOEs, CSPs, SHP_C-NTD, SHP_N-NTD and UBX-NTD distances, and the SAXS-based R_g value (*SI Appendix* Table S3) were applied with metainference (54), such that the ensemble average quantity of each restraint is in agreement with the experimental value (44). Additionally, restraints on all atoms of p97, and on all of the structured p47 domains were imposed as during the preliminary simulations described above. Cryo-EM restraints were implemented by calculating a GMM using the heavy atoms of residues 180-370 of p47 and all heavy atoms of $p97_{ND1}$; this is subsequently compared with the cryo-EM based GMM every 2 steps with a neighbor list cutoff of 0.01. The neighbor list itself was updated every 50 steps. Chemical shifts of the α- and β-carbons in p47 were back-calculated at each time step using CamShift (55) and compared with experimental values. Methyl-methyl NOEs constraining the SIM/SEP interaction included connections between (i) L122 (on SIM) and residues L184, L186, L210, I213, L233 and M235 (on SEP) and (ii) L118 (on SIM) and L186, I213 and L233 (on SEP). NOEs were enforced using harmonic potentials restraining contacts between methyl carbons in the SIM and SEP domains with a distance cutoff of 1.0 nm. This was accomplished through the use of the PLUMED COORDINATION keyword, which computes the contact value s_{ij} between selected atoms i and j, which in the discrete case returns either a 1 if the atomic distance is within the specified cutoff, or 0 otherwise. The potential energy term is then proportional to

$$\sum_i \frac{(\langle s_{ij} \rangle - s_{ij}^{res})^2}{\sigma^2} \qquad \text{[S15]}$$

where i and j are methyl residues (with i-j NOEs) belonging to the SIM sequence and SEP domain, respectively, the summation is over all SIM methyl groups with NOEs to methyl groups in the SEP domain, $<s_{ij}>$ is the

contact value ($s_{ij,}$) for atoms *i* and *j* averaged over the metainference ensemble, s_{ij}^{res} is the corresponding restraint parameter (set to 1 for all methyl-methyl NOEs), and σ encompasses all of the metainference uncertainty parameters (54). Therefore, the potential energy is minimized when the distance between atoms *i* and *j* are within the cutoff of 1.0 nm. In order to ensure that the potential term is continuous, and therefore differentiable, a switching function is used to calculate s_{ij}:

$$s_{ij} = \frac{1-\left(\frac{r_{ij}}{r_0}\right)^6}{1-\left(\frac{r_{ij}}{r_0}\right)^{12}} \qquad \text{[S16]}$$

where r_{ij} is the interatomic distance between atoms *i* and *j* and r_0 is the cutoff (1.0 nm). A number of SIM/SEP amide resonances showed significant CSPs and/or line broadening during the course of the titration of $p47_{1\text{-}174}$ (which includes the SIM sequence) and $p47_{171\text{-}300}$ (which includes the SEP domain); these include resonances derived from S114, E117, L118, V119, D120, L122, F123, K124, G125, G130 of the SIM sequence and resonances from residues L184, K185, L186, Q202, A207, L210, I213, R215, G216, E217, V218, L222, L225, H227, L233, D234, M235, Q236, D237 of the SEP domain. These interactions were modeled using ambiguous distances between the appropriate residues according to

$$r_{\text{ambi}} = \left(\sum_{i>j}\sum_{j} r_{ij}^{-6}\right)^{-1/6} \qquad \text{[S17]}$$

where *i* and *j* are nitrogen/oxygen atoms within SIM and SEP, respectively, including all three p47 polypeptide chains. Thus, instead of a particular interatomic distance, we consider SIM-SEP interactions over all the residues which show perturbations in NMR spectra. In this manner potential intermolecular interactions are included. All ambiguous distances were restrained to 0.5 nm. Finally, an R_g value was calculated using the alpha carbons of the entire complex and restrained to the experimental value of 5.65 nm.

Convergence of the ensemble was assessed by performing a cluster analysis of the last 336 ns of the simulations. This section of the trajectory was split into first and second halves, and clustering analysis performed on each half with the GROMOS clustering algorithm (56) with a cutoff of 1.0 nm, using the backbone

root-mean-square-deviation (RMSD) of residues 114-370 of p47 as a measure of conformational similarity. The system was determined to have converged after ~ 670 ns of aggregate simulation time.

As described above, we have included an R_g potential in an attempt to enforce the effective size of the p47-p97 complex and hence place bounds on the position of the UBA domain from p47, for which experimental data is not available. The simulations that we have performed establish, however, that the R_g value can be satisfied equally well when UBA domains are placed to the sides of the D1 ring or immediately above the D1 ring. Moreover, the R_g measured for the p47-p97$_{ND1L}$ complex reflects that of the ensemble of p47-p97$_{ND1L}$ stoichiometries and binding configurations for the concentrations used, and is therefore not a particularly useful metric by which to assess the size of only one state that was simulated here (corresponding to 3:6 stoichiometry and SHP$_N$ motifs fully bound). Although we illustrate the UBA domain in our structural models, we note that our representations may be artificially compact.

In order to generate an atomic model for the p47-p97$_{ND1}$ ADP/down state, the procedure used was similar to that for the apo complex described above, with a few notable exceptions. Initial coordinates for the p97 down conformation were obtained from an X-ray structure of p97 in complex with ADP (PDB ID: 3CF3, (57)). Given the observation that only two SEP domains are visible in the cryo-EM map of the p47-p97^{ADP} complex, only two molecules of p47 were used in the initial XPLOR-NIH docking procedure and subsequent simulations. Moreover, as our model predicts that SHP$_N$ is not bound in the ADP/down conformation, SHP$_N$/NTD contacts were not enforced. To ensure that the p47 SEP domains are in close proximity to the p97 NTDs, as observed in the cryo-EM density maps, separate XPLOR COLLapse potential terms were applied to each of the SEP/NTD pairs, with a target radius of 22 Å and a force constant of 1 kcal mol^{-1} Å^{-2}.

Preliminary and production GROMACS MD simulations were then conducted as described for the apo/NTD up case, with the following differences. First, the SHP$_N$-NTD restraints and R_g potential term were not enforced. Second, a new GMM cryo-EM restraint was generated on the basis of the p47-p97^{ADP} cryo-EM density map. Seven independent runs, each containing 48 replicates, were simulated for a total simulation time of 1008 ns, and convergence was reached by 672 ns.

Cartoon representations of the structures were prepared using VMD (58) and POV-ray 3.6 (59), and cryo-EM density maps were prepared using *UCSF Chimera* (60).

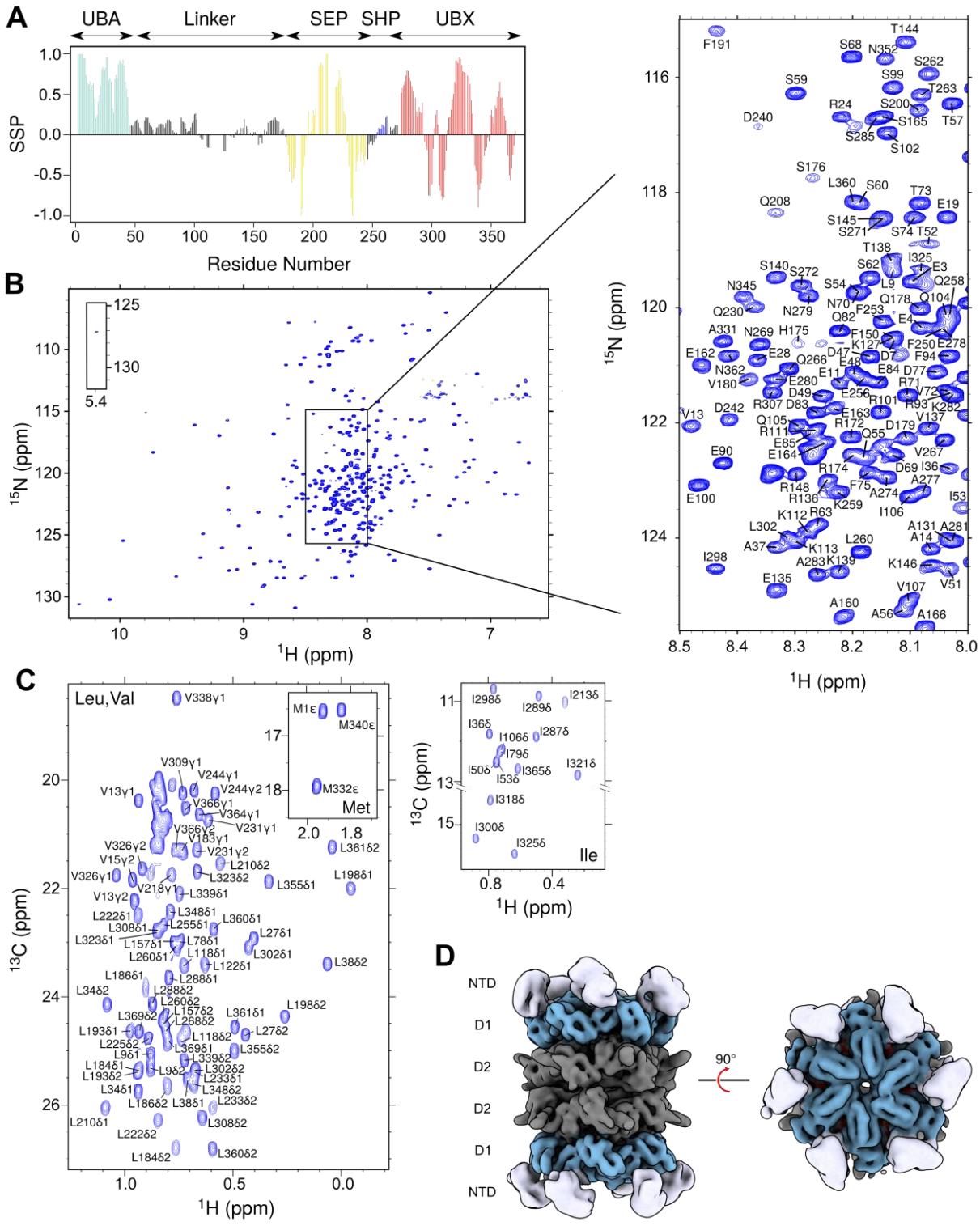

Fig. S1. (A) Secondary structure propensity (SSP (61)) of p47 calculated from backbone chemical shifts. Folded domains are colored according to Fig. 1*B*. Positive and negative SSP scores indicate propensities to form α-helices and β-strands, receptively. (B) ^{15}N-^{1}H TROSY-HSQC spectrum of [U-^{2}H,^{15}N]-p47 recorded at 18.8 T, 25 °C. (C) ^{13}C-^{1}H HMQC spectrum of [U-^{2}H, ILVM-^{13}CH$_3$]-p47 recorded at 18.8 T, 25 °C. (D) Electron cryo-microscopy map of full-length p97 in the apo-form (*i.e.*, no bound nucleotide) with the NTDs adopting the up

conformation, shown in side (left) and top (right) views. NTD, D1 and D2 are coloured white, blue and gray, respectively. Note that full-length $p97^{apo}$ forms a tail-to-tail dodecamer under the conditions of the cryo-EM experiment.

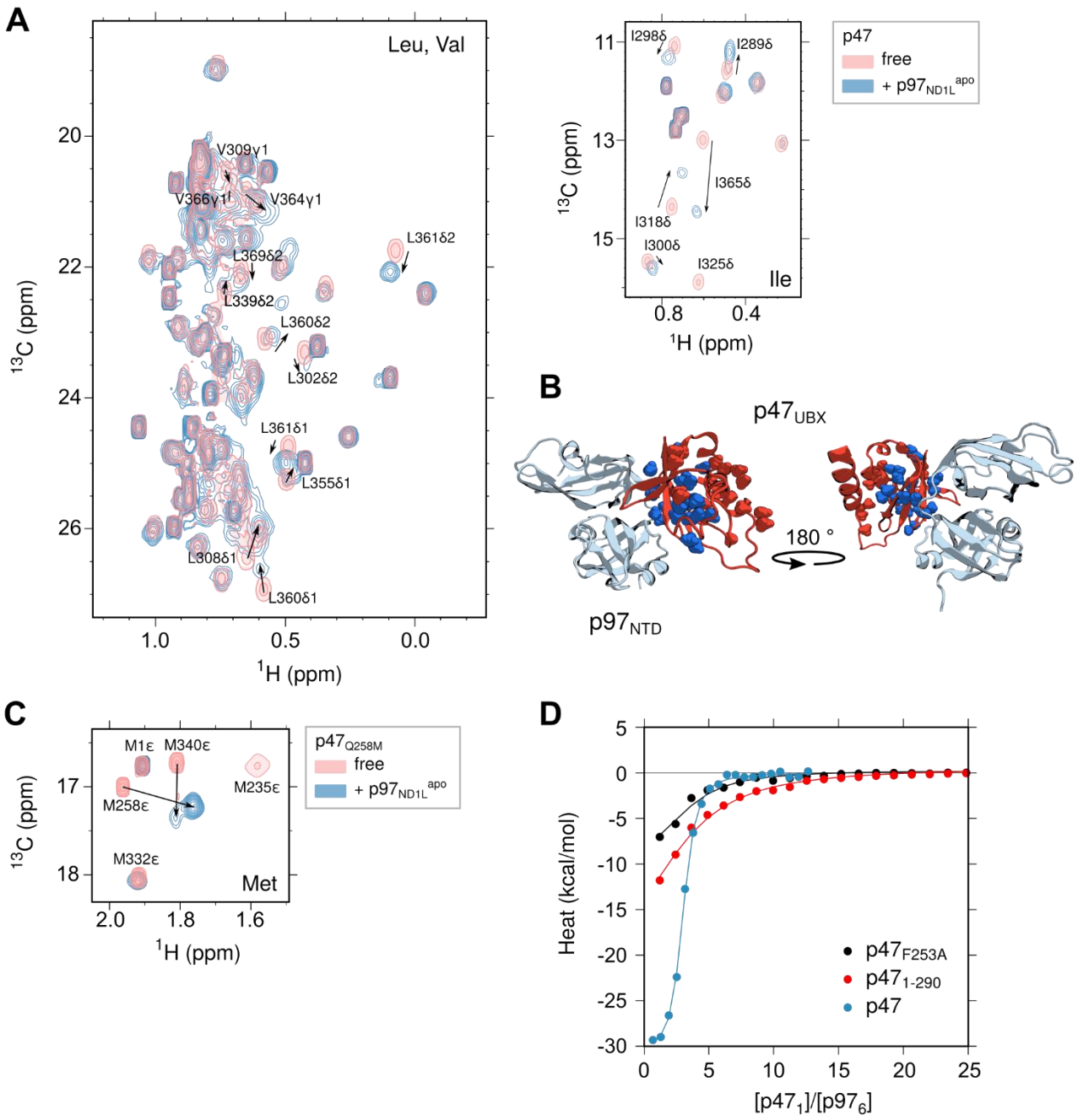

Fig. S2. (A) Superposition of selected regions of ^{13}C-^{1}H HMQC spectra of [U-^{2}H, ILVM-$^{13}CH_3$]-labeled full-length p47 in the absence (pink) and presence (light blue) of 5.3-fold excess [U-^{2}H]-$p97_{ND1L}{}^{apo}$. (B) Structure of $p47_{UBX}$-$p97_{NTD}$ (PDB ID 1S3S (36)), with ILVM-methyl groups of $p47_{UBX}$ shown as spheres; methyl groups with CSPs > 1σ above the mean are highlighted in blue and the remaining are in red. (C) Superposition of the Met methyl region of ^{13}C-^{1}H HMQC spectra of ILVM-$p47_{Q258M}$ in the absence (pink) and presence (blue) of 2-fold excess [U-^{2}H]-$p97_{ND1L}{}^{apo}$, highlighting a large chemical shift perturbation for the M258 methyl resonance; residue 258 is located on the SHP_C motif of p47 and the chemical shift change observed provides strong

evidence of SHP_C binding to p97, as expected. NMR data recorded at 18.8 T, 40 °C. (D) Integrated heats from ITC experiments in which solutions of WT p47 (blue), $p47_{F253A}$ (black; mutated SHP_C motif), or $p47_{1\text{-}290}$ (red; absence of UBX domain) were titrated into a solution of $p97_{ND1L}{}^{apo}$. Solid lines represent best fits to a two-state equivalent sites model as described in *SI Appendix*, Materials and Methods, with stoichiometry parameter *N* (fixed to 3 for $p47_{F253A}$ and $p47_{1\text{-}290}$); the absence of baselines at low $[p47_1]/[p97_6]$ ratios makes it difficult to fit accurate *N* values. Fitted K_d values of 0.23 ± 0.16, 6.4 ± 2.4, and 13.9 ± 3.3 μM, ΔH values of -30.8 ± 1.9, -10.8 ± 0.9, and -24.4 ± 1.1 kcal/mol were obtained for p47, $p47_{F253A}$ and $p47_{1\text{-}290}$, respectively. The significant differences in the three titration isotherms establishes the importance of SHP_C and UBX binding to $p97_{ND1L}{}^{apo}$, as has been observed in studies of full length p97.

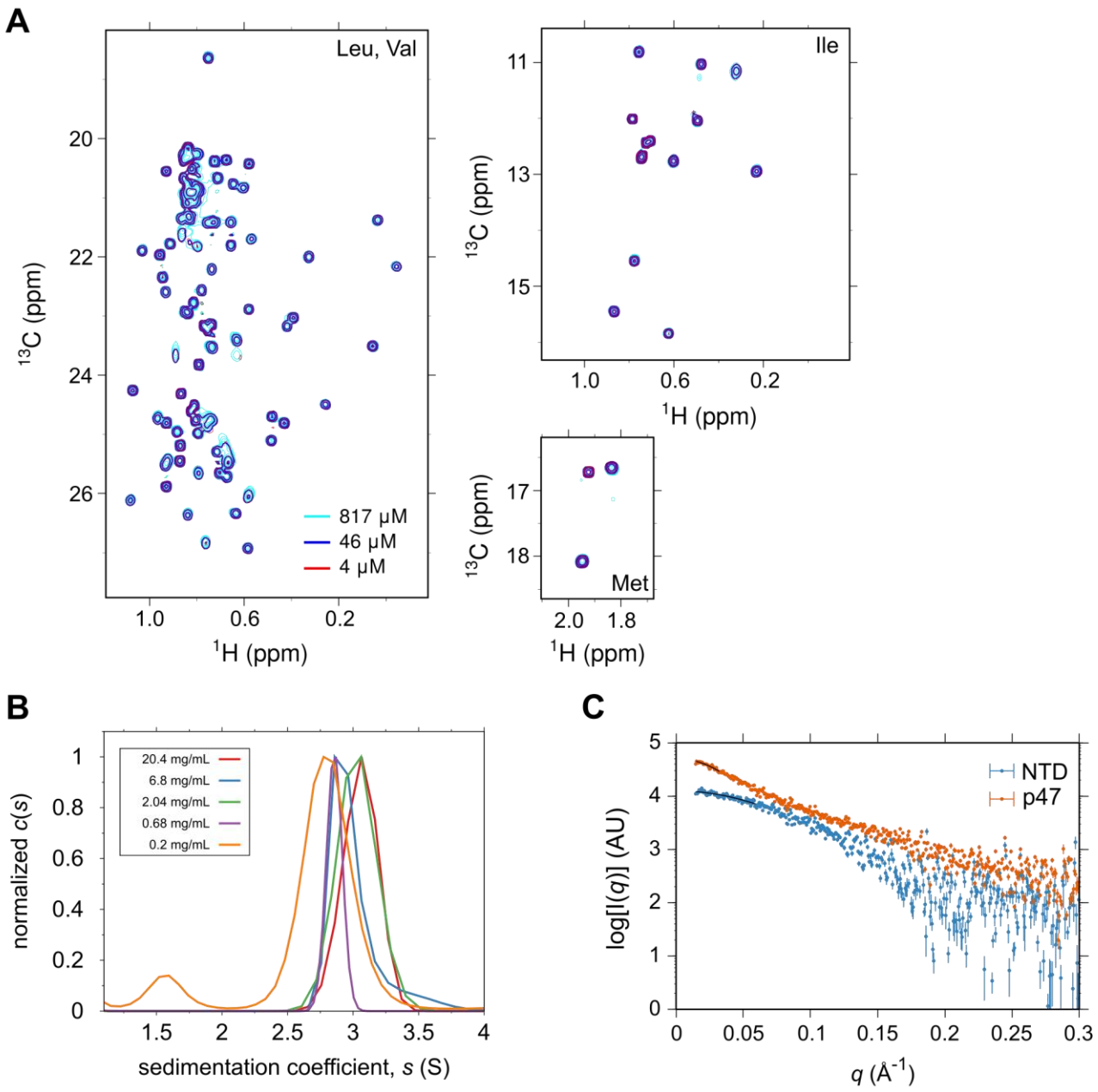

Fig. S3. (A) Superposition of ^{13}C-^{1}H HMQC spectra of ILVM-p47 at 4 µM (red), 46 µM (blue) and 817 µM (cyan). Data recorded at 18.8 T, 25 °C. (B) Sedimentation coefficient distributions c(*s*) calculated from sedimentation velocity AUC experiments of p47 at sample concentrations ranging from 0.2 mg/mL (~5 µM) to 20 mg/mL (~500 µM). *s*-values are corrected for temperature, and *c*(*s*) values are normalized to the maximum height of the largest profile (corresponding to the major species at ~2.9 S). (C) SAXS profiles for 13 mg/mL p47 (orange) and 3 mg/mL p97$_{NTD}$ (blue). Solid black lines are best fits of exponential decays based on the Guinier approximation.

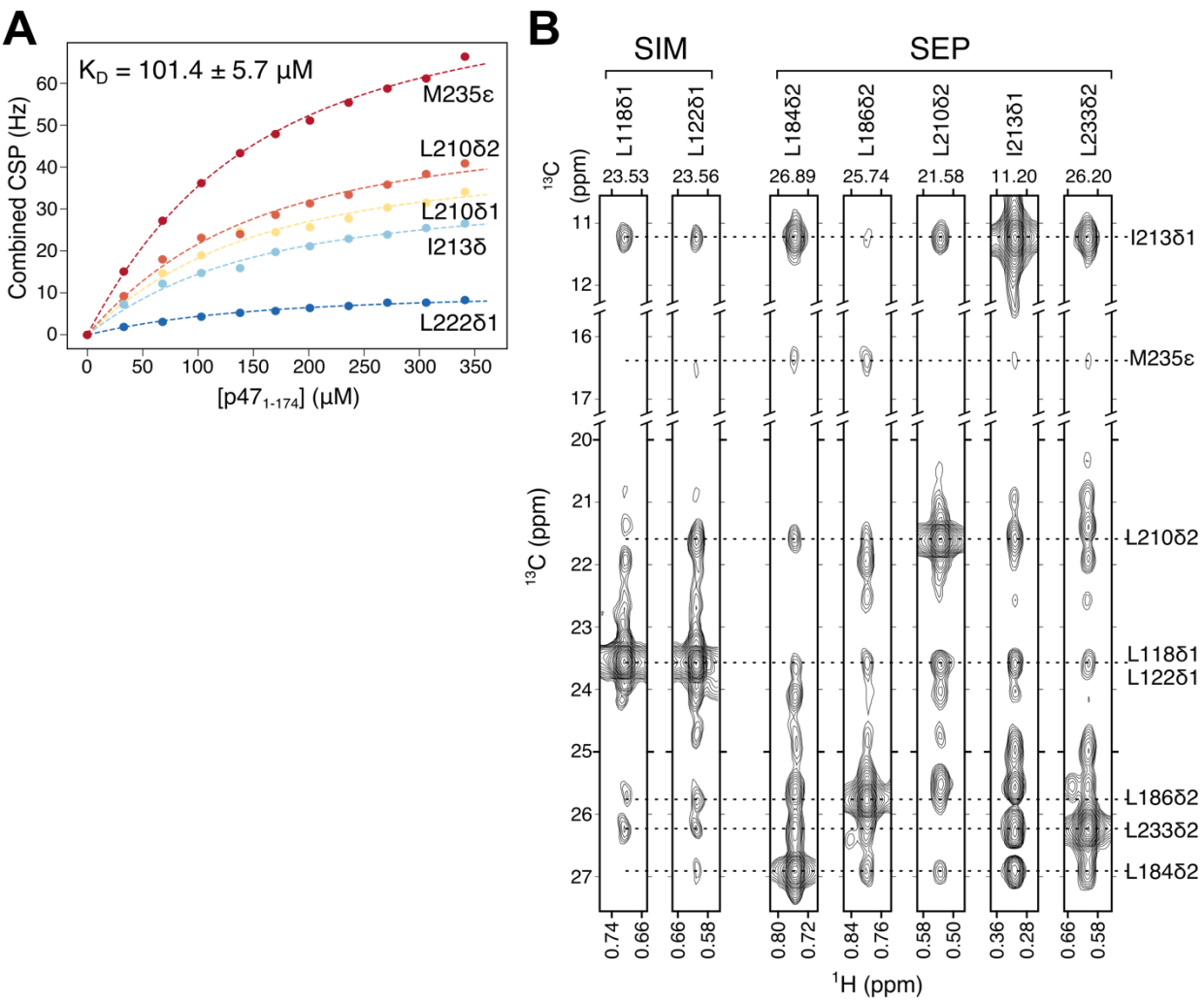

Fig. S4. (A) Binding curves derived from CSPs of selected methyl groups in [U-^{2}H,ILVM-^{13}CH$_3$]-p47$_{171-370}$ upon addition of [U-^{2}H]-p47$_{1-174}$ (circles), with the best fits shown in dashed lines. Combined CSPs were calculated as described in *SI Appendix*, Materials and Methods. All binding data were recorded at 14.1 T, 25 °C. (B) NOEs observed between methyl groups of residues L118 and L122 located on the SIM sequence and methyl groups of residues located on the SEP domain, including L184, L186, L210, I213 and L233. The NOESY experiment was recorded on an ILVM-p47$_{101-266}$ sample (200 ms mixing time) at 18.8 T, 25 °C.

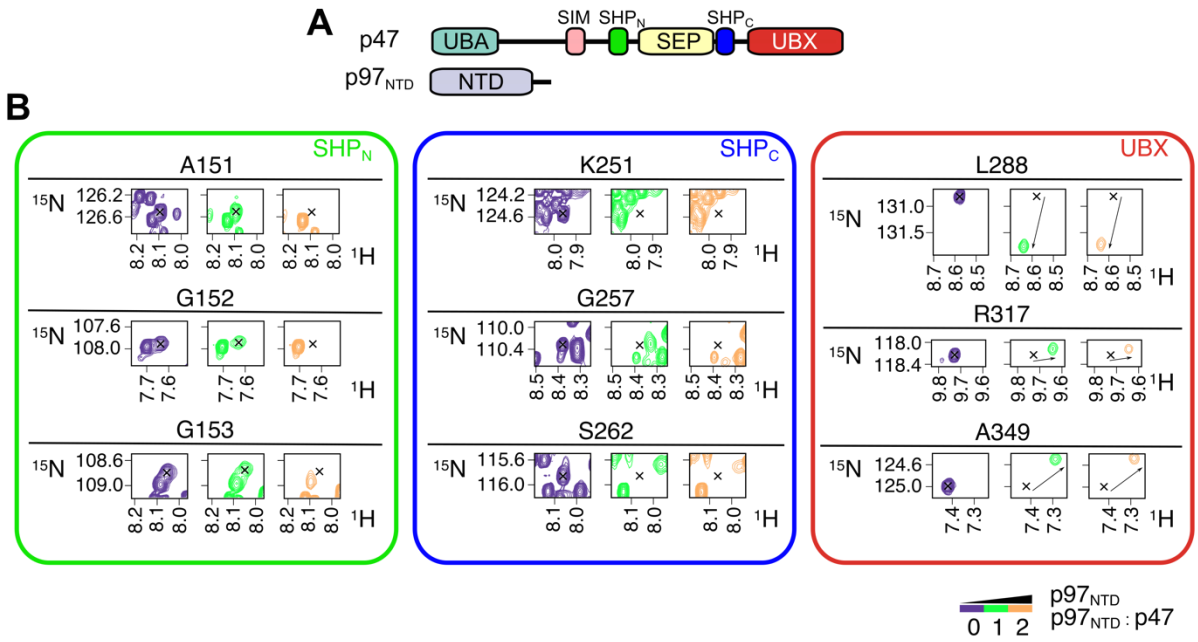

Fig. S5. Full-length p47 interacts with $p97_{NTD}$ via both SHP motifs and a UBX domain. (A) Domain architecture of p47 and $p97_{NTD}$ constructs that are used to record the data in this figure. (B) Selected regions of ^{15}N-^{1}H HSQC spectra of [U-^{2}H,^{15}N,^{13}C]-p47 in the absence (purple) and presence of equimolar (green) and two-fold excess (orange) [U-^{2}H]-$p97_{NTD}$, highlighting residues located on SHP_N (green box, left), SHP_C (blue box, middle) and the UBX domain (red box, right) of p47 that are probes of binding to $p97_{NTD}$. Data recorded at 18.8 T, 25 °C.

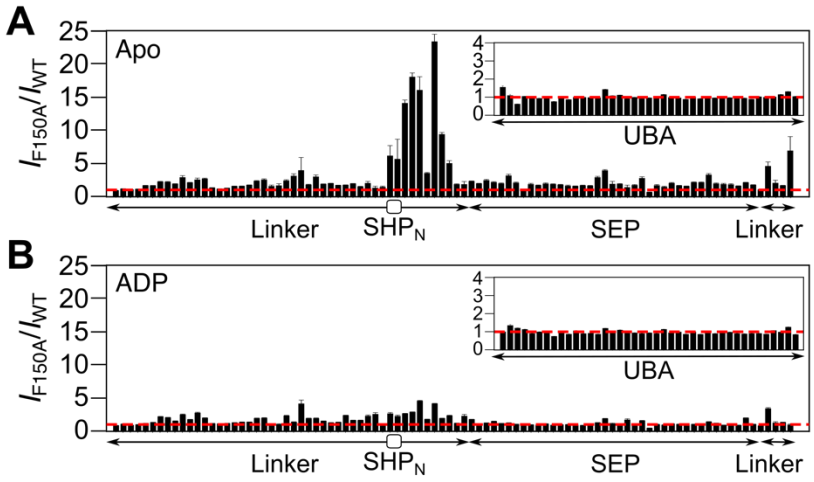

Fig. S6. Cross-peak intensity ratios measured in spectra recorded of [U-^{2}H,^{15}N]-p47$_{F150A}$ or WT [U-^{2}H,^{15}N]-p47, complexed with [U-^{2}H]-p97$_{ND1L}$ (3:6 ratio) in the apo (A) or ADP (B) states. The red dashed lines indicate intensity ratios of 1.0 and the box above 'SHP_N' denotes residues in this motif. Residues located in SHP_C and UBX broaden beyond detection in ^{15}N-based spectra (but not in methyl-TROSY data sets) and are thus not shown here. These results, taken together with those in Fig. 5*C*, provide strong support that the WT SHP_N motif is bound to p97 in the apo (NTD up) but not ADP (NTD down) states, with mutation of the key F150 residue eliminating binding, and hence increasing flexibility. Data recorded at 18.8T, 40 °C.

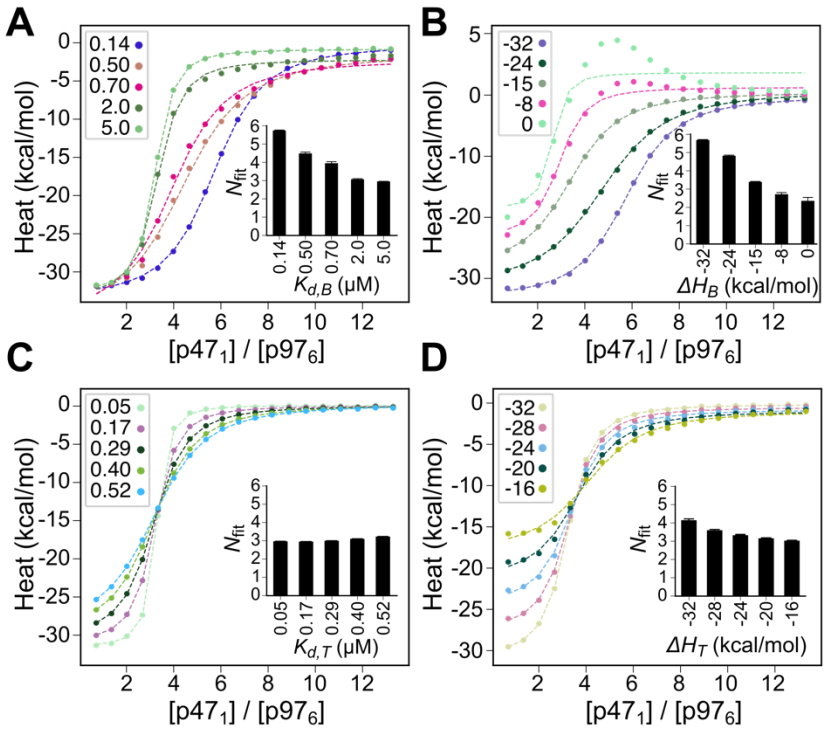

Fig. S7. Simulation of ITC titration profiles assuming a bipartite-tripartite binding model. The resulting isotherms were fit assuming an N-equivalent sites model (dashed lines) (62) with the obtained N_{fit} values reported in insets. Details of the binding model are provided in *SI* Appendix, Material and methods. (A) Integrated ITC heats were simulated with the dissociation constant of tripartite binding (involving UBX, SHP_C and SHP_N), $K_{d,T}$, set to 0.14 μM, the binding enthalpies of tripartite and bipartite binding (UBX+SHP_C), ΔH_T and ΔH_B, both set to -32 kcal/mol, and the dissociation constant of bipartite binding, $K_{d,B}$, varying from 0.14 to 5.0 μM. For the case where $K_{d,T} = K_{d,B}$ the model reduces to one where all sites are equivalent and the N_{fit} value reflects the stoichiometry. In contrast, as $K_{d,B}$ increases, tripartite binding is favored such that for $K_{d,B} >> K_{d,T}$ only three sites become occupied and N_{fit} = 3. (B) Integrated ITC heats were simulated with $K_{d,T}$ and $K_{d,B}$, both set to 0.14 μM, ΔH_T set to -32 kcal/mol, and ΔH_B varying from -32 to 0 kcal/mol. As $|\Delta H_B|$ decreases the effective amount of heat released upon binding decreases and each successive isotherm is shifted to the left. When ΔH_B = 0 kcal/mol positive heats can be obtained as release of SHP_N to accommodate a bipartite interaction involves a positive enthalpy change. Note that the N_{fit} values vary from 6 to 3, despite the fact that $K_{d,T} = K_{d,B}$ and six molecules of ligand can bind to p97 in all cases. Thus, for complex binding mechanisms N_{fit} values do not necessarily report the stoichiometry of the interaction. This phenomenon is also observed in simulations in panels (C) and (D), where N_{fit} values are all close to 3, as observed in the analysis of experimental data, Fig. 5*B*. For (C) and (D), $K_{d,B}$ and ΔH_B, are set to 0.52 μM and -16 kcal/mol, respectively. Integrated ITC heats were simulated either with ΔH_T set to -32 kcal/mol and $K_{d,T}$ varying from 0.05 to 0.52 μM (C) or with $K_{d,T}$ set to 0.14 μM and ΔH_T varying from -32 to -16 kcal/mol (D). Data were simulated assuming a volume of the ITC cell and an injection volume of 0.201 mL and 2 μL, respectively, and concentrations of $p97_{ND1L}^{apo}$ hexamer in the cell and p47 monomer in the syringe of 2.5 μM and 160 μM, respectively.

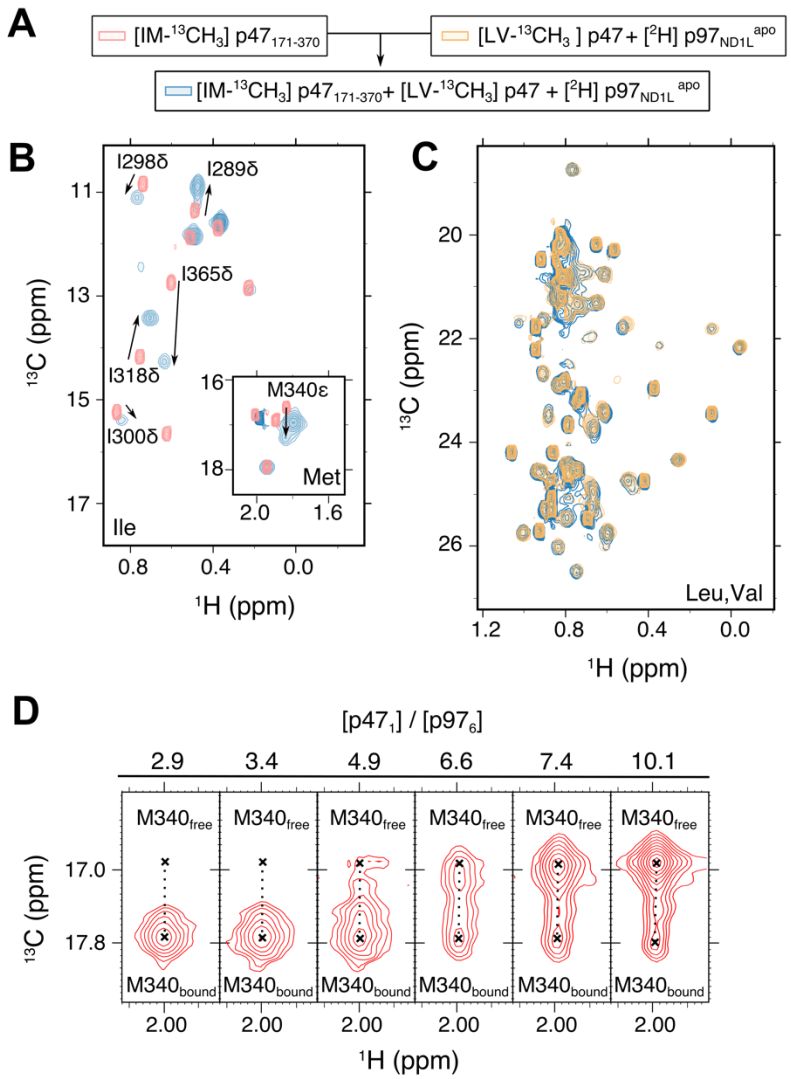

Fig. S8. (A) Schematic of a mixing experiment in which [U-^{2}H,IM-^{13}CH$_3$]-p47$_{171\text{-}370}$ is added to a 3:6 p47-p97$_{ND1L}^{apo}$ complex formed between [U-^{2}H,LV-^{13}CH$_3$]-p47 and [U-^{2}H]-p97$_{ND1L}^{apo}$. An equal amount of p47$_{171\text{-}370}$ is added as full-length p47, to form a 6:6 complex. Superposition of Ile, Met (B) and Leu, Val (C) methyl regions of ^{13}C-^{1}H HMQC spectra, with free [U-^{2}H,IM-^{13}CH$_3$]-p47$_{171\text{-}370}$ in pink, the 3:6 [U-^{2}H,LV-^{13}CH$_3$]-p47:[U-^{2}H]-p97$_{ND1L}^{apo}$ complex in yellow, and the 3:3:6 [U-^{2}H,IM-^{13}CH$_3$]-p47$_{171\text{-}370}$:[U-^{2}H,LV-^{13}CH$_3$]-p47:[U-^{2}H] p97$_{ND1L}^{apo}$ complex in blue. Data recorded at 18.8 T, 40 °C. (D) A selective region of ^{13}C-^{1}H HMQC spectra of [U-^{2}H,IM-^{13}CH$_3$]-p47 as a function of [p47$_1$]/[U-^{2}H-p97$_{ND1L}^{apo}$], highlighting UBX domain residue M340 of p47, with crosshairs indicating the peak positions in the free and bound states. Data recorded at 23.5 T, 40 °C.

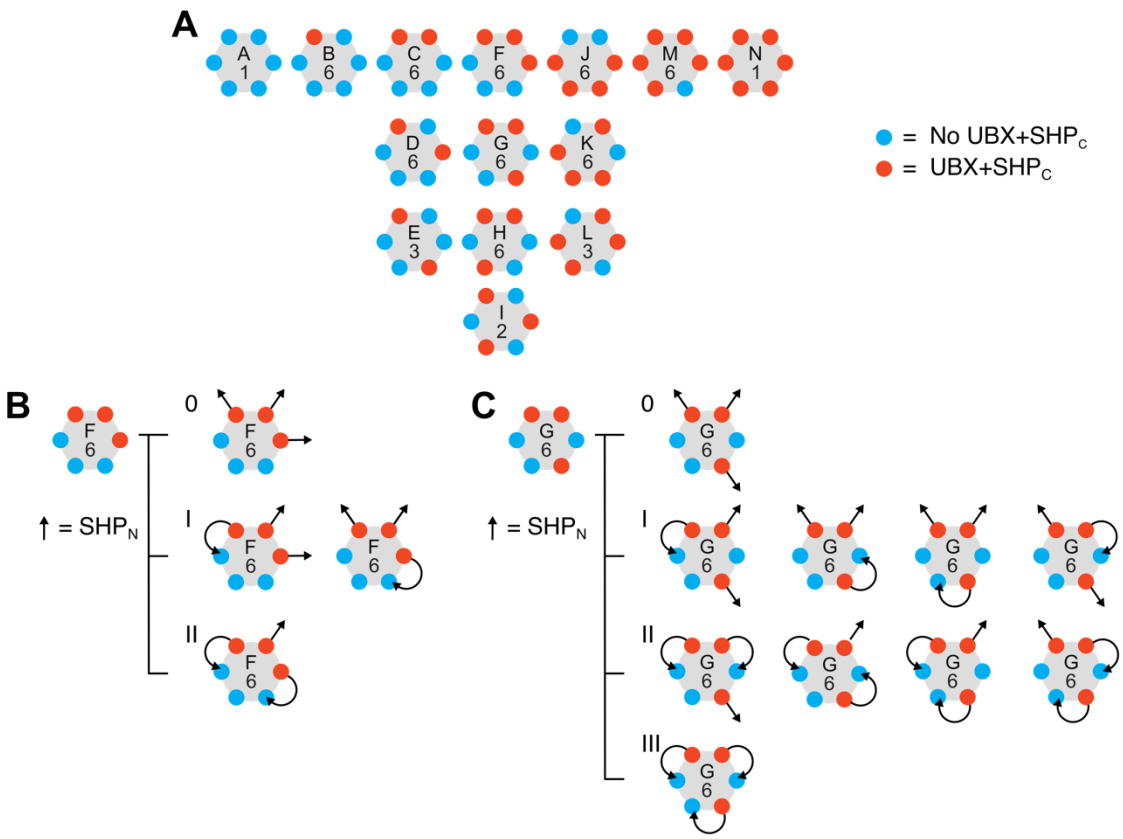

Fig. S9. Schematic of the bipartite/tripartite binding model describing the interaction of full-length p47 with $p97^{apo}$. (A) Non-degenerate states (A-N) of hexameric p97 with NTDs either bound with UBX+SHP$_C$ of p47 (orange circle; bipartite binding) or unbound (blue circle). The degeneracy of each state (*i.e.* different number of ways of binding each p47 ligand leading to a unique configuration) is denoted by the number in the circle (*e.g.* 1 for state A and 6 for state B). (B, C) Including the possibility of tripartite binding, focusing on states F (panel B) and G (panel C), where SHP$_N$, denoted by an arrow, can either be unbound (arrow pointing away from the ring), or bound to an unoccupied NTD (arrow pointing to blue circle) from a protomer that is adjacent to the site bound with UBX+SHP$_C$ of the p47 ligand from which the SHP$_N$ is derived (orange circle). For state F (panel B), there can be 0 (0), 1 (I) or 2 (II) SHP$_N$s in the bound state, with two distinct configurations for the singly bound SHP$_N$, while for state G (panel C) up to 3 (III) SHP$_N$s can engage with NTD, with 4 distinct bound configurations for the cases where 1 or 2 SHP$_N$ motifs are bound.

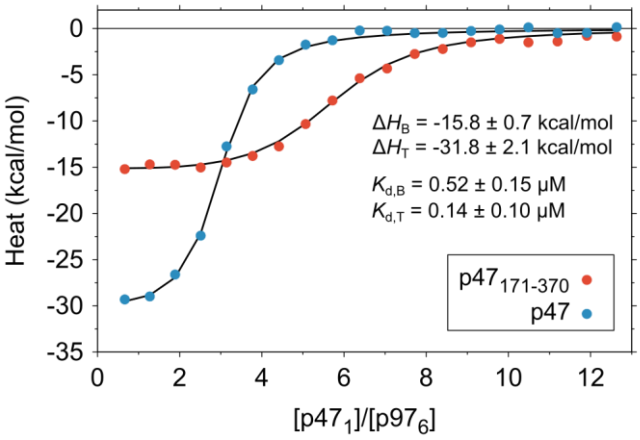

Fig. S10. Simultaneous fit of the p47 bipartite/tripartite binding model to ITC data for the titration of full length p47 (blue) and $p47_{171-370}$ (red) into $p97_{ND1L}^{apo}$. Experimental data points are reproduced from Fig. 5*B*. Bipartite and tripartite thermodynamic binding parameters are indicated with the subscripts B and T, respectively.

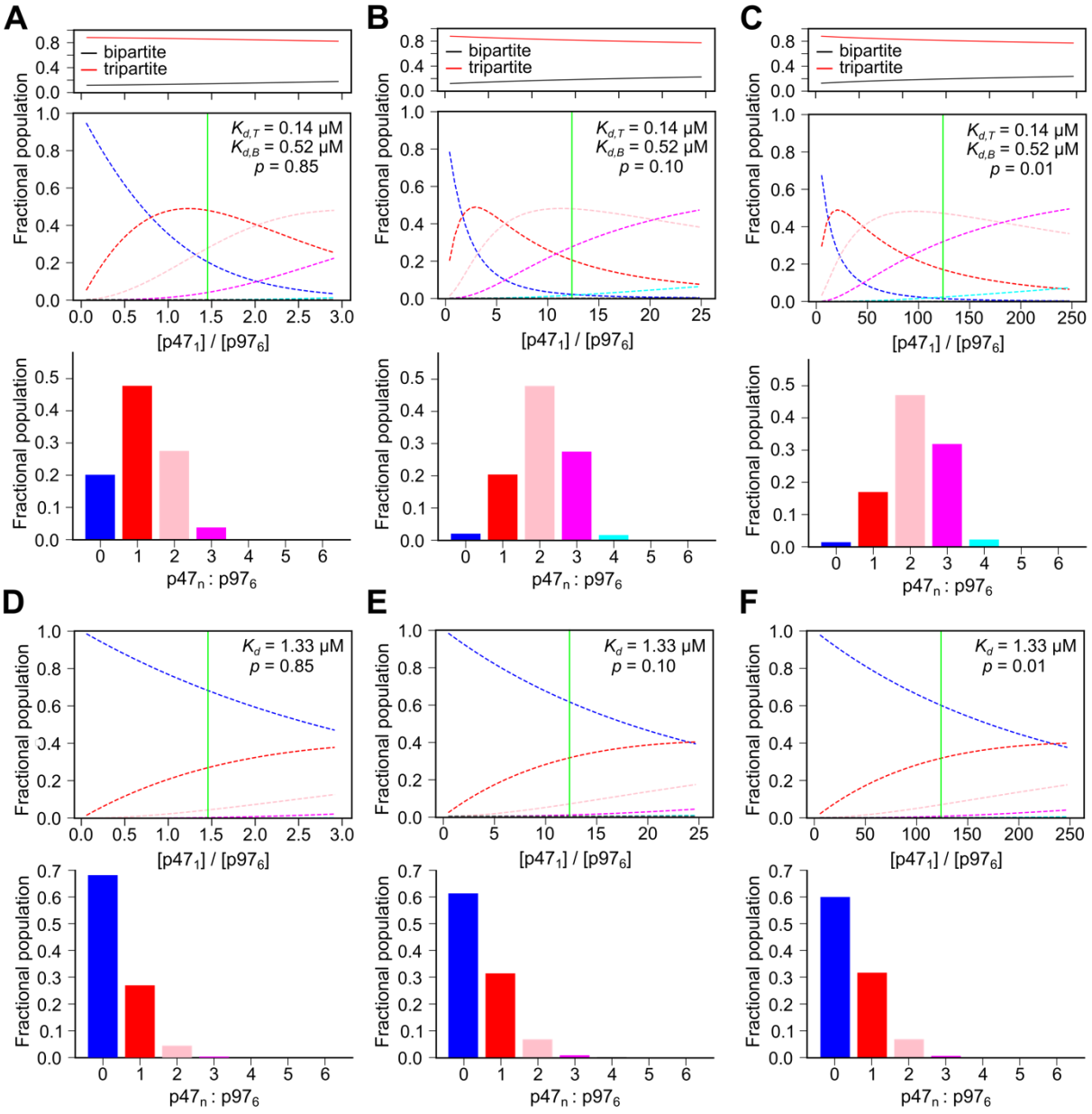

Fig. S11. Simulation of fractional populations of different p47-p97 bound states under near-physiological conditions. The total concentration of p97 hexamer is set to its concentration in mammalian cells 95.5 nM (63), and the fraction of p97 available to bind p47, *p*, is varied between 0.85 (A&D), 0.10 (B&E) and 0.01 (C&F) in the simulations. Simulations are carried out assuming p97 is either in the apo state (A-C), where a bipartite-tripartite binding model is employed (details of the model provided in *SI appendix*, Materials and Methods) with dissociation constants for tripartite and bipartite binding, $K_{d,T}$ and $K_{d,B}$, of 0.14 and 0.52 μM, respectively, or in the ADP state (C-F) where only bipartite binding occurs and a simple six-equivalent-sites model was used with a binding constant, K_d, of 1.33 μM, as measured via ITC. Fractional populations of the bound ligand interacting in bipartite- (black solid line) and tripartite (red solid line) modes as a function of the total concentration of p47 (varied between 0 and 0.236 μM) are shown in A-C, top. Fractional populations of p97 hexamers with 0 (blue dash line), 1 (red dash line), 2 (pink dash line), 3 (magenta dash line), 4 (cyan dash line), 5 (grey dash line) and 6

(black dash line) bound p47 molecules (either bipartite or tripartite) are shown for both $p97^{apo}$ (A-C, middle) and $p97^{ADP}$ (D-F, top) with the total concentration of p47 varied between 0 and 0.236 μM. Bar plots (A-C, bottom for $p97^{apo}$ and D-F, bottom for $p97^{ADP}$) show fractional populations of p97 hexamer bound with *n* p47 molecules when the concentration of p47 is set to be 0.118 μM (indicated by the green solid line in A-C, middle, and D-F, top) to reflect the total concentration of p47 in vivo (63).

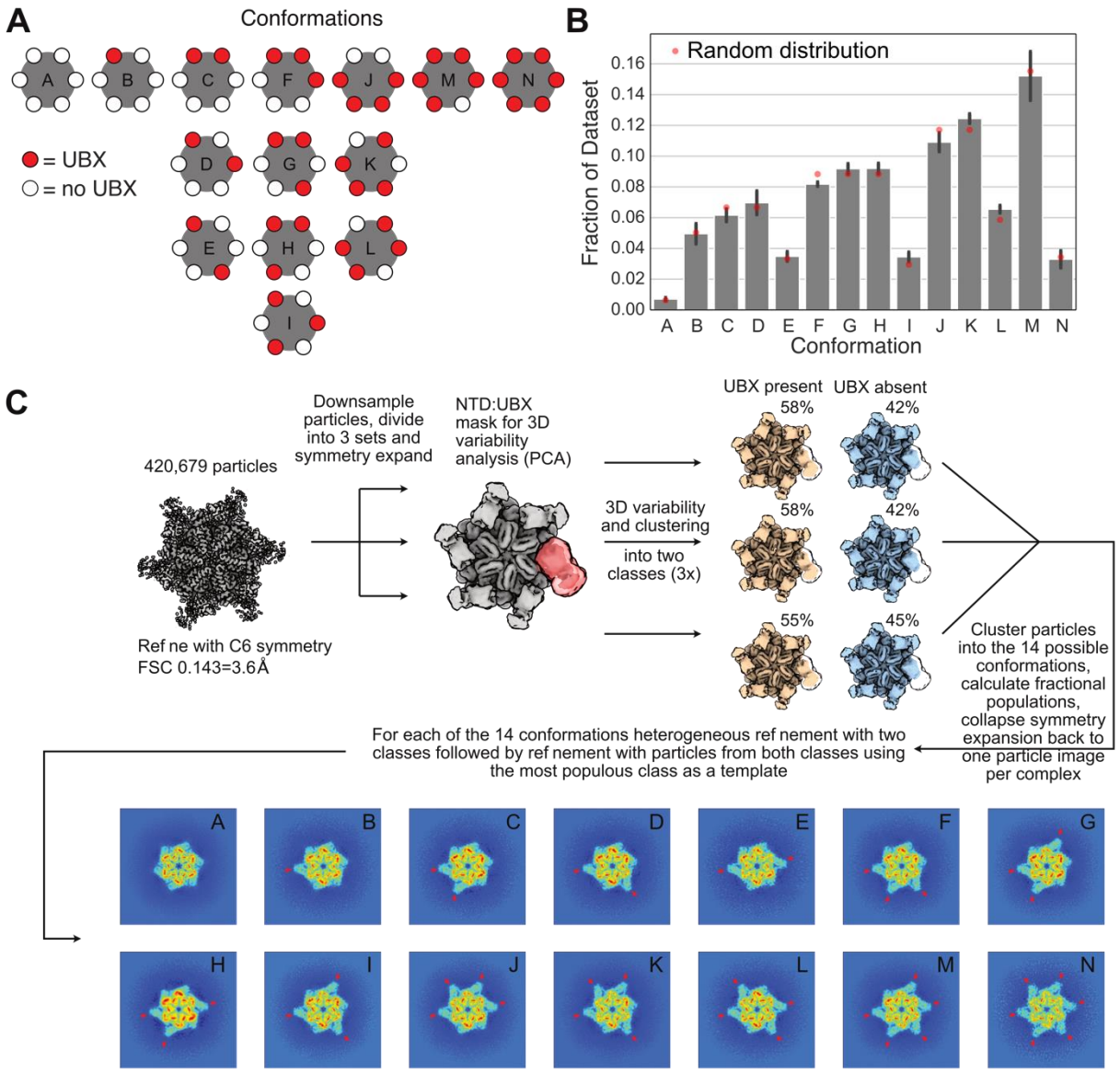

Fig. S12. Classification and reconstruction of p47-p97^{ATP} conformations by cryoEM. (A) The 14 possible states of p47 UBX domains bound to p97 NTDs. The 64 possible conformations can be reduced to 14 states by taking into account rotational symmetry related conformations, as described in *SI appendix*, Materials and Methods. (B) Fractional populations of the 14 p97-p47 states show a random distribution (*i.e.*, probability of k bound states = $_cC_k a^{6-k} b^k$, where $a(b)$ is the probability of an unbound(bound) NTD and $_6C_k = \frac{6!}{k!(6-k)!}$; values of a and b are given by the fraction populations of bound and free NTDs over the complete distribution of particles). This distribution likely reflects the random loss of bound p47 ligands during cryoEM specimen preparation stages, possibly due to blotting or through interactions with the air-water interface. The expected binding stoichiometry in the absence of any loss of bound ligands during preparation is 6:6, based on p47 and p97 concentrations used and measured affinities. Asymmetric units of p97-NTDs were classified based on the density of UBX binding,

with 2D particle images subsequently organized into the 14 possible conformations for the hexameric complex, prior to reconstruction of 3D maps. Three separate classifications (UBX on/off) were performed on independent thirds of the dataset. Red dots indicate an expected stochastic binomial distribution taking into account the population of total bound UBX (in this case ~58%), and grey bars are averaged values of the measured populations with error bars in black (standard deviation of the three classifications). (C) Image analysis workflow for classification of UBX bound conformations (see *SI Appendix*, Materials and Methods). Boxes at the bottom are slices of the 3D reconstructions of the 14 conformations with UBX positions denoted by red arrows.

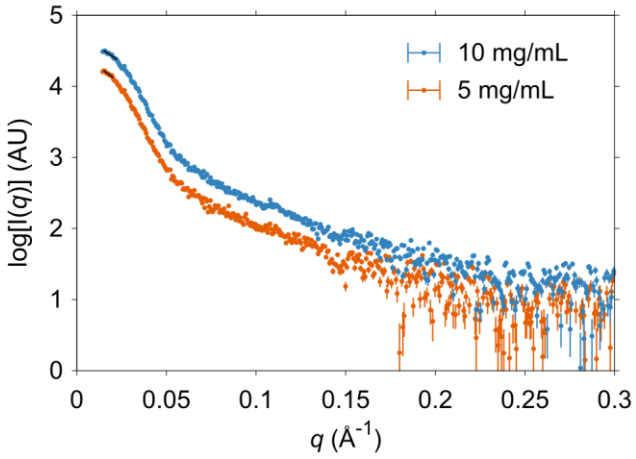

Fig. S13. SAXS intensity profiles measured for 10 mg/mL (blue) and 5 mg/mL (orange) p47-p97$_{ND1L}^{apo}$ complexes at 3:6 p47:p97$_{ND1L}$ (monomer:monomer) molar ratio. Solid black lines are best fits of exponential decays based on the Guinier approximation.

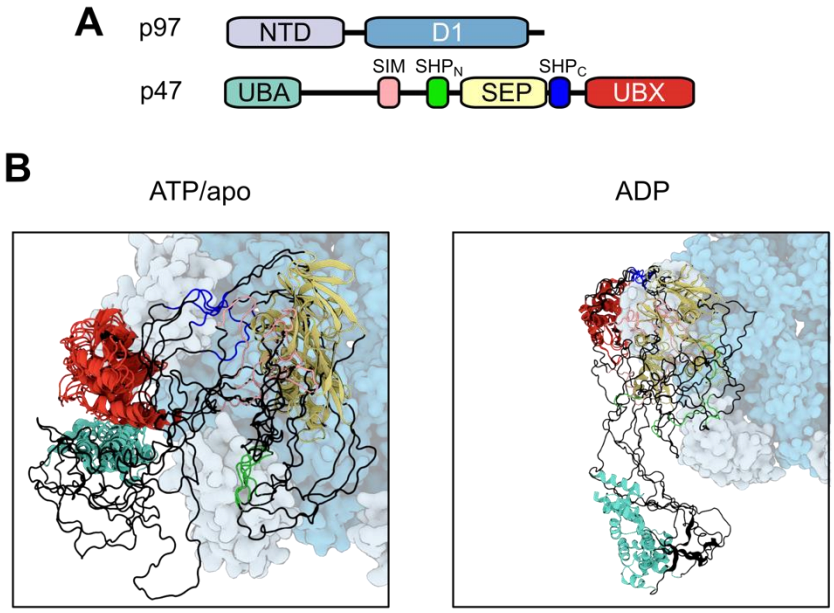

Fig. S14. Structural ensembles from p47-p97$_{ND1}$ metainference molecular dynamics (MD) simulations highlight the dynamic nature of the p47-p97 complex. (A) Domain architectures for p97 and p47 fragments used in the MD simulations. (B) Representative structures from the five most-populated clusters obtained for one p47 protomer in the ATP/apo (left) and ADP (right) p97$_{ND1}$ ensembles are superimposed above the ND1 ring. Positions of the UBA and SEP domains are more variable in the ADP state than in the ATP/apo state, as these regions are, in part, no longer tethered to p97 via SHP$_N$-NTD intermolecular contacts.

Table S1A Thermodynamic parameters obtained from fits of an N-equivalent sites model to ITC data

Sample	K_d (μM)	ΔH (kcal/mol)	N
p47-p97$_{ND1L}^{apo}$	0.23 ± 0.16	-30.8 ± 1.9	2.7 ± 0.3
p47-p97$_{ND1L}^{ADP}$	1.33 ± 0.28	-16.6 ± 1.0	5.5 ± 0.4
p47$_{171-370}$-p97$_{ND1L}^{apo}$	0.55 ± 0.14	-15.9 ± 0.4	5.5 ± 0.2
p47$_{171-370}$-p97$_{ND1L}^{ADP}$	0.97 ± 0.27	-15.5 ± 1.4	6.1 ± 0.3
p47$_{F150A}$-p97$_{ND1L}^{apo}$	1.01 ± 0.29	-16.0 ± 1.3	4.7 ± 0.3
p47$_{F150A}$-p97$_{ND1L}^{ADP}$	0.88 ± 0.17	-15.4 ± 0.6	5.3 ± 0.2
p47$_{F253A}$-p97$_{ND1L}^{apo}$	6.4 ± 2.4	-10.8 ± 0.9	3*
p47$_{1-290}$-p97$_{ND1L}^{apo}$	13.9 ± 3.3	-24.4 ± 1.1	3*

* Parameter fixed during fitting routine.

Table S1B Thermodynamic parameters obtained from fits of the bipartite/tripartite binding model to ITC data

Sample	$K_{d,B}$ (μM)	$K_{d,T}$ (μM)	ΔH_B (kcal/mol)	ΔH_T (kcal/mol)	ζ
p47-p97$_{ND1L}^{apo}$ & p47$_{171-370}$-p97$_{ND1L}^{apo}$	0.52 ± 0.15	0.14 ± 0.10	-15.8 ± 0.7	-31.8 ± 2.1	0.91 ± 0.06

Table S2. Number of configurations in each state, $b_{s,j}$ (Eq [S7]) resulting from binding of q={0-3} SHP_N motifs.

# of bound SHP_N / State	0	1	2	3
A	1	0	0	0
B	1	2	0	0
C	1	2	1	0
D	1	4	3	0
E	1	2	3	0
F	1	2	1	0
G	1	4	4	1
H	1	4	4	1
I	1	2	3	2
J	1	2	1	0
K	1	4	3	0
L	1	2	3	0
M	1	2	0	0
N	1	0	0	0

Table S3. Experimental restraints used for molecular dynamics simulations of p47-p97$_{ND1}$

Experimental Observation	Utility	Reference
Backbone NH CSPs & Line Broadening (SIM): S114, E117, V119, D120, F123, K124, G125, G130	Limit intramolecular distance between p47 SIM and SEP domain	This work
Methyl CSPs (SIM): L118, L122	Limit intramolecular distance between p47 SIM and SEP domain	This work
Backbone NH CSPs & Line Broadening (SEP): L184, K185, Q202, A207, I213, R215, G216, E217, V218, H227, D234, M235, Q236, D237	Limit intramolecular distance between p47 SIM and SEP domain	This work
Methyl CSPs (SEP): L184, L186, L210, I213, L222, L225, M235	Limit intramolecular distance between p47 SIM and SEP domain	This work
NOEs from L122 Cδ1/δ2 to L184 Cδ1/δ2, L186 Cδ1/δ2, L210 Cδ1/δ2, I213 Cδ1, L233 Cδ1/δ2, and M235 Cε	Limit intramolecular distance between p47 SIM and SEP domain	This work
NOEs from L118 Cδ1/δ2 to L186 Cδ1/δ2, I213 Cδ1, and L233 Cδ1/δ2	Limit intramolecular distance between p47 SIM and SEP domain	This work
cryo-EM density maps for p47-p97^{ATP} or p47-p97^{ADP}	Limit conformational space of the p47 UBX and SEP domains with respect to p97.	This work
R_g derived from SAXS measurements of p47-p97$_{ND1L}{}^{apo}$ complex	Limit conformational space of the UBA domain and UBA-SEP linker (p97$_{ND1}{}^{apo/up}$ state only).	This work
Intermolecular p47 UBX-p97 NTD atomic distances within 6 Å, obtained from the X-ray structure of p47$_{UBX}$-p97$_{ND1}$ complex	Define intermolecular p47$_{UBX}$-p97 NTD contacts.	PDB ID: 1S3S (36)
Intermolecular UFD1$_{SHP}$-p97 NTD	Define intermolecular p47$_{SHP}$-p97	PDB ID: 5C1B (41)

atomic distances within 6 Å, obtained from the X-ray structure of $UFD1_{SHP}$-p97 complex	NTD contacts (conserved residues only).	

Table S4. Sample conditions used for the various experiments in this study.

Experiments/Samples	Buffer components	Location
NMR		
Backbone assignment of full-length p47, NMR titration of [U-^{2}H,^{15}N,^{13}C]-p47 with [U-^{2}H]-p97$_{NTD}$	25 mM Bis-Tris, 25 mM NaCl, 1 mM EDTA, 3 mM TCEP, 10 % D_2O, pH 6.6	Fig. S1*A*&*B*, Fig. S5*B*
Backbone assignment of p47$_{1-174}$	25 mM citric acid - sodium phosphate, 1 mM EDTA, 10 % D_2O, pH 5.5, 6.0, 6.5, and 7.0	
Methyl assignments of p47 fragments, including p47$_{1-174}$, p47$_{171-370}$, p47$_{171-266}$, p47$_{291-370}$	20 mM sodium acetate, 1 mM EDTA, 10 % D_2O, pH 5.53	
^{13}C-^{1}H HMQC of ILVM-p47 and ILVM-p47$_{171-370}$ in the presence and absence of [U-^{2}H] p97^{ND1L}	25 mM HEPES, 25 mM NaCl, 1 mM EDTA, 3 mM TCEP, 99.9% D_2O, pD 7.4	Fig. 3*H*, Fig. 6, Fig. S1*C*
NMR experiments studying interactions between p47$_{1-174}$ and p47$_{171-370}$	25 mM HEPES, 25 mM NaCl, 1 mM EDTA, 10 % D_2O, pH 7.4	Fig. 3*B*-*E*
^{13}C-edited NOESY experiment on p47$_{101-266}$	25 mM HEPES, 25 mM NaCl, 1 mM EDTA, 99.9% D_2O, pD 7.4	Fig. 3*G*
[U-^{15}N,^{13}C]-p47$_{1-174}$ + p97$_{NTD}$ binding	25 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM TCEP, 10 % D_2O, pH 7.4	Fig. 4*B*
[U-^{2}H,^{15}N,^{13}C]-p97$_{NTD}$ + [U-^{2}H]-p47$_{1-174}$ binding	25 mM MES, 50 mM NaCl, 1 mM EDTA, 1 mM TCEP, 10 % D_2O, pH 6.0	Fig. 4*D*
^{15}N-^{1}H HSQC of [U-^{2}H,^{15}N]-p47 (or p47$_{F150A}$) + [U-^{2}H]- p97^{ND1L}	25 mM HEPES, 25 mM NaCl, 1 mM EDTA, 3 mM TCEP, 10 % D_2O, pH 7.4	Fig. 5*C*
Pull-down assay	50 mM HEPES, 100 mM NaCl, 1 mM EDTA, pH 7.5	Fig. 2*B*
Dynamic light scattering	25 mM HEPES, 50 mM NaCl, 1 mM EDTA, pH 7.4	Fig. 2*C*
Analytical ultracentrifugation	25 mM HEPES, 50 mM NaCl and 1 mM EDTA, pH 7.4	Fig. *3B*
ITC	25 mM HEPES, 150 mM NaCl, 1 mM EDTA, 2 mM TCEP, pH 7.4	Fig. 5*B*, Fig. S2*D*, Fig. S10
SAXS	25 mM HEPES, 150 mM NaCl, 1 mM EDTA, 2 mM TCEP, pH 7.4	Fig. S3*C*, Fig. S13
Cryo-EM	50 mM HEPES, 150 mM NaCl, 5 mM $MgCl_2$, 5 mM ADP (or ATP), pH 7.5	Fig. 7*B*

References

1. D. G. Gibson, *et al.*, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).

2. A. K. Schuetz, L. E. Kay, A dynamic molecular basis for malfunction in disease mutants of p97/VCP. *Elife* **5**, e20143 (2016).

3. R. Huang, F. Pérez, L. E. Kay, Probing the cooperativity of Thermoplasma acidophilum proteasome core particle gating by NMR spectroscopy. *Proc. Natl. Acad. Sci.* **114**, E9846–E9854 (2017).

4. F. Delaglio, *et al.*, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).

5. W. Lee, M. Tonelli, J. L. Markley, NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).

6. K. Pervushin, R. Riek, G. Wider, K. Wuthrich, Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci.* **94**, 12366–12371 (1997).

7. M. Sattler, J. Schleucher, C. Griesinger, Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**, 93–158 (1999).

8. X. Yuan, *et al.*, Letter to the Editor: Complete backbone resonance assignments of p47: The 41kDa adaptor protein of the AAA ATPase p97. *J. Biomol. NMR* **28**, 309–310 (2004).

9. S. Grzesiek, J. Anglister, A. Bax, Correlation of backbone amide and aliphatic side-chain resonances in 13C/15N-enriched proteins by isotropic mixing of 13C magnetization. *J. Magn. Reson. Ser. B* **101**, 114–119 (1993).

10. S. G. Hyberts, K. Takeuchi, G. Wagner, Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data. *J Am Chem Soc* **132**, 2145–2147 (2010).

11. J. Ying, F. Delaglio, D. A. Torchia, A. Bax, Sparse multidimensional iterative lineshape-enhanced (SMILE) reconstruction of both non-uniformly sampled and conventional NMR data. *J. Biomol. NMR* **68**, 101–118 (2017).

12. S. G. Hyberts, A. G. Milbradt, A. B. Wagner, H. Arthanari, G. Wagner, Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling. *J. Biomol. NMR* **52**, 315–327 (2012).

13. V. Tugarinov, P. M. Hwang, J. E. Ollerenshaw, L. E. Kay, Cross-correlated relaxation enhanced 1H-13C NMR spectroscopy of methyl groups in very high molecular weight proteins and protein complexes. *J. Am. Chem. Soc.* **125**, 10420–10428 (2003).

14. J. Santoro, G. C. King, A constant-time 2D overbodenhausen experiment for inverse correlation of isotopically enriched species. *J. Magn. Reson.* **97**, 202–207 (1992).

15. G. W. Vuister, A. Bax, Resolution enhancement and spectral editing of uniformly 13C-enriched proteins by homonuclear broadband 13C decoupling. *J. Magn. Reson.* **98**, 428–435 (1992).

16. W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *Numerical recipes in C* (Cambridge University Press, 1998).

17. B. J. Frisken, Revisiting the method of cumulants for the analysis of dynamic light-scattering data. *Appl. Opt.* **40**, 4087–4091 (2001).

18. J. A. Lake, An iterative method of slit-correcting small angle X-ray data. *Acta Crystallogr.* **23**, 191–194 (1967).

19. A. Guinier, G. Fournet, *Small-angle scattering of X-rays* (Wiley, 1955).

20. R. P. Rambo, J. A. Tainer, Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496**, 477–481 (2013).

21. S. C. To, *et al.*, Enhanced sample handling for snalytical ultracentrifugation with 3D-printed centerpieces. *Anal. Chem.* **91**, 5866–5873 (2019).

22. H. Zhao, *et al.*, Recorded scan times can limit the accuracy of sedimentation coefficients in analytical ultracentrifugation. *Anal. Biochem.* **437**, 104–108 (2013).

23. P. Schuck, Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophys. J.* **78**, 1606–1619 (2000).

24. E. Freire, A. Schön, A. Velazquez-Campoy, Isothermal titration calorimetry: general formalism using binding polynomials. *Methods Enzymol.* **455**, 127–155 (2009).

25. A. Sekhar, J. Nagesh, R. Rosenzweig, L. E. Kay, Conformational heterogeneity in the Hsp70 chaperone-substrate ensemble identified from analysis of NMR-detected titration data. *Protein Sci.* **26**, 2207–2220 (2017).

26. C. R. Marr, S. Benlekbir, J. L. Rubinstein, Fabrication of carbon films with ~500nm holes for cryo-EM with a direct detector device. *J. Struct. Biol.* **185**, 42–47 (2014).

27. C. J. Russo, L. A. Passmore, Ultrastable gold subtrates for electron cryomicroscopy. *Science* **346**, 1377–1380 (2014).

28. J. R. Meyerson, *et al.*, Self-assembled monolayers improve protein distribution on holey carbon cryo-EM supports. *Sci. Rep.* **4**, 7084 (2014).

29. W. F. Tivol, A. Briegel, G. J. Jensen, An improved cryogen for plunge freezing. *Microsc. Microanal.* **14**, 375–379 (2008).

30. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

31. J. L. Rubinstein, M. A. Brubaker, Alignment of cryo-EM movies of individual particles by optimization of image translations. *J. Struct. Biol.* **192**, 188–195 (2015).

32. M. Zhou, *et al.*, Atomic structure of the apoptosome: mechanism of cytochrome c- and dATP-mediated activation of Apaf-1. *Genes Dev.* **29**, 2349–2361 (2015).

33. A. Punjani, D. J. Fleet, 3D Variability Analysis: Directly resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM images. *bioRxiv*, 2020.04.08.032466 (2020).

34. C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson.* **160**, 65–73 (2003).

35. X. Yuan, *et al.*, Structure, dynamics and interactions of p47, a major adaptor of the AAA ATPase, p97. *EMBO J.* **23**, 1463–1473 (2004).

36. I. Dreveny, *et al.*, Structural basis of the interaction between the AAA ATPase p97/VCP and its adaptor protein p47. *EMBO J.* **23**, 1030–1039 (2004).

37. C. D. Schwieters, G. M. Clore, Internal coordinates for molecular dynamics and minimization in structure determination and refinement. *J. Magn. Reson.* **152**, 288–302 (2001).

38. M. J. D. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives. *Comput. J.* **7**, 155–162 (1964).

39. C. D. Schwieters, J. J. Kuszewski, G. Marius Clore, Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **48**, 47–62 (2006).

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

41. P. Hänzelmann, H. Schindelin, Characterization of an additional binding surface on the p97 N-terminal domain involved in bipartite cofactor interactions. *Structure* **24**, 140–147 (2016).

42. C. D. Schwieters, G. A. Bermejo, G. M. Clore, Xplor-NIH for molecular structure determination from NMR and other data sources. *Protein Sci.* **27**, 26–40 (2018).

43. G. M. Clore, C. D. Schwieters, Docking of protein−protein complexes on the basis of highly ambiguous intermolecular distance restraints derived from 1HN/15N chemical shift mapping and backbone 15N−1H residual dipolar couplings using conjoined rigid body/torsion angle dynamics. *J Am Chem Soc* **125**, 2902–2912 (2003).

44. M. Bonomi, R. Pellarin, M. Vendruscolo, Simultaneous determination of protein structure and dynamics using cryo-electron microscopy. *Biophys. J.* **114**, 1604–1613 (2018).

45. H. J. C. Berendsen, D. van der Spoel, R. van Drunen, GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* **91**, 43–56 (1995).

46. S. Pronk, *et al.*, GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).

47. M. Bonomi, *et al.*, Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **16**, 670–673 (2019).

48. G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, G. Bussi, PLUMED 2: New feathers for an old bird. *Comput. Phys. Commun.* **185**, 604–613 (2014).

49. R. B. Best, *et al.*, Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ(1) and χ(2) dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).

50. S. Bottaro, K. Lindorff-Larsen, R. B. Best, Variational optimization of an all-atom implicit solvent force field to match explicit solvent simulation data. *J. Chem. Theory Comput.* **9**, 5641–5652 (2013).

51. T. Lazaridis, M. Karplus, Effective energy function for proteins in solution. *Proteins Struct. Funct. Bioinforma.* **35**, 133–152 (1999).

52. B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).

53. T. Löhr, A. Jussupow, C. Camilloni, Metadynamic metainference: Convergence towards force field independent structural ensembles of a disordered peptide. *J. Chem. Phys.* **146**, 165102 (2017).

54. M. Bonomi, C. Camilloni, A. Cavalli, M. Vendruscolo, Metainference: A Bayesian inference method for heterogeneous systems. *Sci. Adv.* **2**, e1501177 (2016).

55. P. Robustelli, K. Kohlhoff, A. Cavalli, M. Vendruscolo, Using NMR Chemical Shifts as Structural Restraints in Molecular Dynamics Simulations of Proteins. *Structure* **18**, 923–933 (2010).

56. X. Daura, *et al.*, Peptide folding: when simulation meets experiment. *Angew. Chemie Int. Ed.* **38**, 236–240 (1999).

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

58. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

59. Persistence of Vision Pty Ltd, Persistence of Vision (TM) Raytracer (2004).

60. E. F. Pettersen, *et al.*, UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–12 (2004).

61. J. A. Marsh, V. K. Singh, Z. Jia, J. D. Forman-Kay, Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: Implications for fibrillation. *Protein Sci.* **15**, 2795–2804 (2006).

62. S. Williston, J. F. Brandts, Rapid Measurement of Binding Constants and Heats of Binding Using a New Titration Calorimeter. **137**, 131–137 (1989).

63. M. Beck, *et al.*, The quantitative proteome of a human cell line. *Mol. Syst. Biol.* **7**, 549 (2011).