









## Fragment screening using bilayer interferometry reveals ligands targeting the SHP-motif binding site of the AAA+ ATPase p97

Sebastian Bothe <sup>1,2</sup>, Petra Hänzelmann<sup>2</sup>, Stephan Böhler <sup>1,2</sup>, Josef Kehrein<sup>1</sup>, Markus Zehe <sup>1</sup>, Christoph Wiedemann <sup>3</sup>, Ute A. Hellmich <sup>3,4,5</sup>, Ruth Brenk <sup>6</sup>, Hermann Schindelin <sup>2</sup>✉ & Christoph Sotriffer <sup>1</sup>✉

Biosensor techniques have become increasingly important for fragment-based drug discovery during the last years. The AAA+ ATPase p97 is an essential protein with key roles in protein homeostasis and a possible target for cancer chemotherapy. Currently available p97 inhibitors address its ATPase activity and globally impair p97-mediated processes. In contrast, inhibition of cofactor binding to the N-domain by a protein-protein-interaction inhibitor would enable the selective targeting of specific p97 functions. Here, we describe a bilayer interferometry-based fragment screen targeting the N-domain of p97 and demonstrate that a region known as SHP-motif binding site can be targeted with small molecules. Guided by molecular dynamics simulations, the binding sites of selected screening hits were postulated and experimentally validated using protein- and ligand-based NMR techniques, as well as X-ray crystallography, ultimately resulting in the first structure of a small molecule in complex with the N-domain of p97. The identified fragments provide insights into how this region could be targeted and present first chemical starting points for the development of a protein-protein interaction inhibitor preventing the binding of selected cofactors to p97.

<sup>1</sup>Institute of Pharmacy and Food Chemistry, University of Würzburg, Am Hubland, 97074 Würzburg, Germany. <sup>2</sup>Rudolf Virchow Center for Integrative and Translational Bioimaging, University of Würzburg, Josef-Schneider-Str. 2, Haus D15, 97070 Würzburg, Germany. <sup>3</sup>Faculty of Chemistry and Earth Sciences, Institute of Organic Chemistry and Macromolecular Chemistry, Biostructural Interactions, Friedrich Schiller University Jena, Humboldtstr. 10, 07743 Jena, Germany. <sup>4</sup>Cluster of Excellence "Balance of the Microverse", Friedrich Schiller University Jena, Jena, Germany. <sup>5</sup>Center for Biomolecular Magnetic Resonance (BMRZ), Goethe University Frankfurt, Max von Laue Str. 9, 60438 Frankfurt, Germany. <sup>6</sup>Department of Biomedicine, University of Bergen, Jonas Lies Vei 91, 5020 Bergen, Norway. ✉email: [hermann.schindelin@virchow.uni-wuerzburg.de](mailto:hermann.schindelin@virchow.uni-wuerzburg.de); [sotriffer@uni-wuerzburg.de](mailto:sotriffer@uni-wuerzburg.de)

The AAA+ ATPase p97 is an essential protein involved in numerous cellular processes and plays multiple key roles during protein homeostasis<sup>1,2</sup>. p97 consists of six identical monomers that assemble into a functional C<sub>6</sub>-symmetrical hexamer. Each monomer can be subdivided into an N-terminal domain, the Walker A and Walker B motif-containing ATPase domains D1 and D2, and an extended unstructured C-terminal tail (Fig. 1a). The ATPase functions convert the energy of ATP hydrolysis into mechanical energy required to extract ubiquitylated proteins from membranes, chromatin or macromolecular complexes<sup>1</sup>. The functional diversity of p97 is mediated through the interaction with a large number of distinct protein cofactors, which mainly interact with the N-terminal domain of p97 and also the C-terminal tail<sup>3,4</sup>. Cofactors binding to the N-domain interact with two different regions. UBX (ubiquitin regulatory X) or UBXL (UBX-like) domain-containing cofactors as well as cofactors harboring a VIM (VCP-interacting motif) or VBM (VCP-binding motif) bind into a cleft between the Nn- and Nc-subdomains of the N-domain. In contrast, cofactors with a SHP-motif interact with an extended region located in the Nc-subdomain<sup>5</sup>.

Point mutations in p97 lead to inclusion body myopathy associated with Paget's disease of the bone and frontotemporal dementia<sup>6</sup> and familial amyotrophic lateral sclerosis<sup>7</sup>, now collectively referred to as multisystem proteinopathy. Furthermore, due to its significant role in regulating a variety of physiological responses, especially as part of the ubiquitin-proteasome-system, p97 has emerged as a potential therapeutic target, in particular to treat cancer<sup>8</sup>. Inhibition of p97 leads to cell death in HCT116, A549<sup>9</sup> and B-ALL<sup>10</sup> tumor cells.

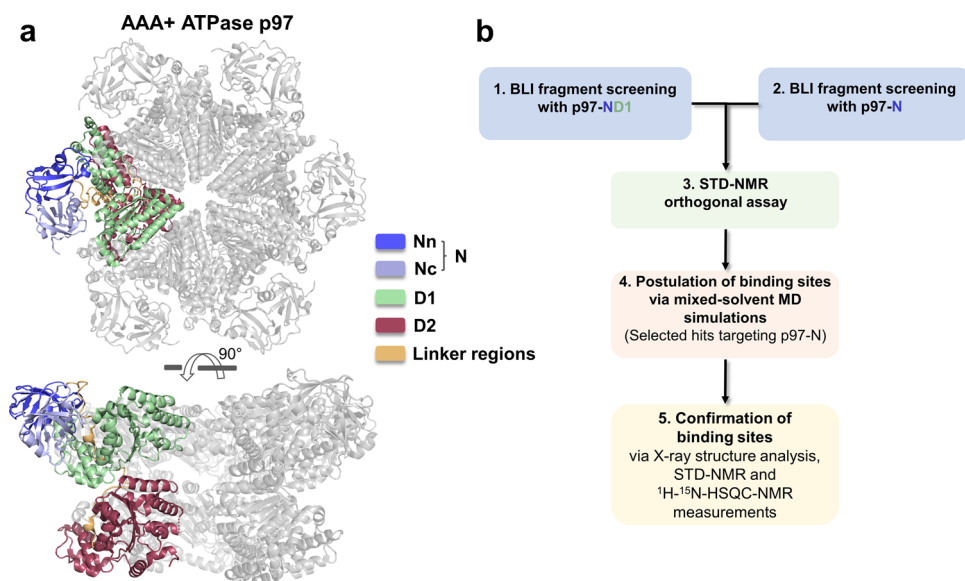
Based on their mechanism of action, known p97 inhibitors can be divided into three categories<sup>7,11,12</sup>: (i) Competitive D1 and/or D2 ATPase inhibitors (e.g., DBEQ<sup>13</sup>, ML240/241<sup>14</sup>, NMS-859<sup>15</sup> or CB-5083<sup>16</sup>); (ii) allosteric inhibitors of the ATPase function by a non-competitive mechanism (e.g. NMS 873<sup>17</sup>, UPCDC30245<sup>18</sup> or MSC1094308<sup>19</sup>); and (iii) inhibitors with an unknown mechanism of action (e.g. clotrimazol<sup>20</sup> or eeyarestatin I<sup>21</sup>). Furthermore, inhibitors can be divided into non-covalent ligands

and covalent binders, the latter mainly modifying residue C522 in the nucleotide-binding site of the D2 domain<sup>11,15</sup>. So far, only the natural product xanthohumol, occurring in hop plants, has been reported to bind to the N-domain of p97<sup>22</sup>.

As demonstrated by the failure of the CB-5083 inhibitor in clinical phase I trials<sup>12</sup>, inhibitors that competitively target the ATPase functions of p97 may exhibit limited selectivity due to the inhibition of other nucleotide-dependent enzymes. Furthermore, inhibition of the ATPase function leads to an indiscriminate loss of cellular p97 functions. In contrast, inhibition of cofactor binding to the N-domain by a protein-protein-interaction (PPI) inhibitor would enable the selective targeting of specific p97 functions. Beyond the long-term goal of using them as therapeutics, these inhibitors would also be important to dissect the molecular and cellular functions of p97 and its cofactors, thus helping to unravel how certain cofactors control p97 activity.

To identify chemical starting points for the development of PPI inhibitors targeting the N-domain of p97 (p97-N), we used a fragment-based screening approach (Fig. 1b). In contrast to screenings with drug-like compounds, the use of fragments (i.e., smaller compounds with 10–20 heavy atoms) can sample the regions of chemical space more efficiently<sup>23</sup>. A fragment-based approach is also best suited to address more challenging targets, such as interfaces of protein-protein complexes, and has already been used in the successful development of PPI inhibitors against IL-2<sup>24</sup>, TNF- $\alpha$  or BCL-X<sub>L</sub><sup>25</sup>. In the single hitherto reported fragment screening with p97<sup>26</sup>, six fragments binding to the N-domain were identified by surface plasmon resonance spectroscopy (SPR) and saturation transfer difference (STD)-NMR methods.

The screen reported here was conducted by biolayer-interferometry (BLI), followed by STD-NMR experiments. To map the binding sites of the identified fragments mixed-solvent molecular dynamics (MD) simulations were used as well as X-ray crystallography and <sup>1</sup>H-<sup>15</sup>N-HSQC (heteronuclear single quantum coherence) NMR measurements to confirm the binding sites predicted by the simulations. Importantly, the first structure of a small molecule in complex with the N-domain of p97 could be



**Fig. 1** Structure of p97 and workflow of the fragment screen. **a** Hexameric assembly of p97 (PDB entry: 5C1B) with one monomer shown in color. Each monomer can be divided into the two ATPase domains, D1 (green) and D2 (red), and an N-terminal domain consisting of two subdomains, Nn (blue) and Nc (light blue), where most of the known cofactors bind. The domains are connected by two linkers (orange). **b** Workflow of the fragment screen. To validate the screen, it was first performed using a p97-ND1 construct. The established experimental conditions were transferred to the isolated N-domain, followed up by an orthogonal STD-NMR assay and the identification of the binding region of selected hits.

obtained by X-ray crystallography, presenting a starting point for further optimization and structure-based drug design. Moreover, this study demonstrates that the BLI technique, which has been reported so far only sporadically in the literature<sup>27</sup>, is well suited for detecting fragments and is a viable option besides the commonly used SPR method.

## Results

**Initial biolayer interferometry screenings.** As starting point for the initial screen, a BLI binding assay was developed with a p97-ND1 construct (amino acids 2-481) to allow validation with adenosine diphosphate (ADP) as a positive control. The p97-ND1 construct was enzymatically biotinylated via an Avi-Tag and *Escherichia coli* biotin ligase<sup>28</sup> and loaded to high response (>5 nm) on Super Streptavidin Biosensors (SSA), yielding a homogeneous sensor surface. The affinity ( $K_D$ ) of ADP to p97-ND1 was measured to be  $200 \pm 14$  nM using a Langmuir model based on steady state measurements of a dilution series. This value is in good agreement with reported affinities for ND1 constructs of similar lengths; e.g., a value of 200 nM was reported for a ND1 fragment spanning amino acids 1-458 and a value of 98 nM for a ND1 fragment spanning amino acids 1-480<sup>26,29</sup>. The kinetic analysis of the sensorgrams indicated a clear 1:1 binding model for ADP. The global exponential regression resulted in an affinity of  $239 \pm 9$  nM, in agreement with previous results. A comparison of the kinetic rate constants determined by SPR measurements with a comparable ND1 construct<sup>29</sup> revealed  $k_{on}$  and  $k_{off}$  rates in a similar range for ADP (see Supplementary Fig. 1). Another surprising result were the high responses upon ADP-binding with shifts over 0.7 nm. These high values indicate that the BLI signal is not only dependent on the molar weight as in SPR but may also reflect conformational changes induced in p97 upon ADP binding.

For hit discovery, a fragment library located at the Facility for Biophysics, Structural Biology and Screening at the University of Bergen (BiSS) containing 679 fragments was screened. Due to the lack of a positive control for p97-N, the fragment library was first screened with p97-ND1 to validate the experimental conditions. The screen resulted in an overall Z-value<sup>30</sup> of 0.67, demonstrating its high quality (values between 0.5 and 1.0 represent excellent assays). A fraction of 82 fragments was identified above the chosen threshold (>1.0 sigma, Supplementary Fig. 2). Fragments showing abnormally high or negative signals or untypical shapes of the BLI curves were eliminated after visual inspection, leaving 48 compounds (7.0 %) for subsequent hit validation. Out of these, 29 fragments exhibited a dose-dependent response, demonstrating the functionality of the assay. Therefore, the same conditions were transferred to a screen with p97-N and the aforementioned library. 66 fragments were above the threshold (>1.0 sigma), and 42 compounds (6.5%) were selected for dose-response experiments after visual inspection of the sensorgrams (Supplementary Fig. 2).

**Biolayer interferometry dose-response analysis of N-domain hits.** The selected 42 fragments of the screen with the N-domain were tested in a dose-dependent manner using concentrations from 1000.0  $\mu$ M to 31.3  $\mu$ M in 1:1 dilution steps. 36 compounds exhibited dose-dependent response signals. While some sensorgrams displayed a behavior characteristic of unspecific binding, 22 curves displayed binding curves indicative of fragment binding (Supplementary Fig. 3). The sensorgrams of these fragments were fitted with a 1:1 Langmuir model to estimate their affinities and analyzed with respect to the following parameters: (i) The presence of a stationary phase; (ii) curve shapes according to either a homogeneous (1:1) or a heterogeneous (2:1) binding model, and

(iii) justification of a 1:1 binding model via the obtained empirical constants  $k_{obs}$  of the association phase. Based on these analyses, a  $Score_{BLI}$  was calculated (for details see Supplementary Methods in the Supporting Information). The calculated  $Score_{BLI}$  should be close to or above 1.0 for ligands that can be described with a 1:1 model, whereas ligands showing unspecific binding yield values  $\ll$  1.0. Based on the  $Score_{BLI}$ , we divided the fragments into a low-scoring (<0.70) and a high-scoring ( $\geq$ 0.70) group.

**Confirmation of selected fragments by STD-NMR.** To confirm the hits of the BLI screening, fragments with p97-N  $K_D$  values below 700  $\mu$ M were investigated by STD-NMR. A total of 19 fragments satisfied this affinity cutoff. The STD-NMR method provides insights into the transient binding of a small molecule to a protein in solution. Therefore, the method is suitable to eliminate false positive hits arising from the immobilization of the target protein in the BLI-based screen. However, one has to keep in mind that the strength of the STD effect does not directly correlate with the affinity of a molecule and is influenced by other factors as well<sup>31</sup>. Accordingly, the STD-NMR measurements were not used for rank-ordering or affinity-based prioritization.

The NMR measurements were carried out with the biologically more relevant p97-ND1 construct in the presence of 500  $\mu$ M ADP to block the nucleotide binding site. The STD effect was calculated by dividing the intensity of the difference spectrum by the intensity of the off-spectrum ( $I_{diff}/I_{off} \times 100$ ) for each signal<sup>32</sup>. For further analysis only resonances in the aromatic region of the <sup>1</sup>H NMR spectrum of the individual ligand were considered. A mean STD effect of each fragment interacting with p97-ND1 was calculated using the sum of the individual STD effects divided by the number of signals considered. Furthermore, the maximum STD effect of each compound was determined. Three out of the 19 fragments exhibited no significant effect ( $S/N < 3.0$ ) and were therefore not confirmed by the STD-NMR assay. The remaining fragments showed maximum STD effects in a range from 10.5 % (TROLL9) up to 72.5 % (TROLL6). Similar results were obtained in the previously reported fragment screen with p97<sup>26</sup>. An overview of the STD effects is shown in Supplementary Fig. 4.

**Selection of the identified fragments for further analysis.** For further selection of the fragments the  $Score_{BLI}$  was used as criterion. All fragments showing a  $Score_{BLI}$  value greater than 0.7 were further investigated, resulting in 13 molecules (Supplementary Fig. 4).

To rule out the possibility that the fragments bind to regions of the N-domain that are normally blocked by the D1-domain of p97, dose-dependent measurements were carried out with the ND1 construct, again using the BLI method. For 12 fragments, a dose-dependent binding was observed and a 1:1 Langmuir model could be applied to obtain the respective  $K_D$  value. The determined  $K_D$  (p97-ND1) values mostly agreed well with those previously determined for the respective fragments with the isolated N-domain. Only TROLL1 showed no response with the p97-ND1 construct, and VIK40 bound the ND1 construct with only low affinity ( $K_D > 700$   $\mu$ M). TROLL1 and VIK40 were therefore not considered further.

The sensorgrams of the fragments showed surprisingly slow on and off rates as mentioned above for the binding of ADP. Therefore, based on the kinetic data the affinity was determined and plotted against the affinity deduced from the steady state analysis (Supplementary Fig. 5). 7 fragments showed comparable affinities indicating that the observed slow on and off rates result from the BLI assay. 3 fragments (TROLL13, TROLL14 and TROLL18) showed a higher discrepancy for the calculated affinities with up to 3.5 times lower affinity deduced from the

kinetic data. TROLL10 showed the highest discrepancy with a 9.5 times lower affinity and a poor fit to the underlying model and was therefore rejected. The remaining 10 fragments were subjected to additional experimental characterization. An overview of these 10 fragments is given in Table 1. Figure 2 shows the sensorgrams and Langmuir regressions for four fragments of particular interest for the subsequent analyses (as described below).

**Chemical space analysis.** To characterize the chemical properties of the fragments identified in this study, also in relation to those reported by Chimenti et al.<sup>26</sup>, the positions of the fragments within the chemical space of the BiSS fragment library were determined. As a standard tool to analyze the chemical diversity of compound libraries<sup>33–35</sup>, a principal component analysis (PCA) was applied. The PCA was based on 10 molecular descriptors (as specified in Materials and Methods).

Figure 3 shows a plot of the first two principal components, which together describe 47.6% of the variance in the BiSS library. The positions of the previously reported fragments<sup>26</sup> and the compounds identified in this work are mostly located in the same area in the lower right quartile of the plot: Eight of the 10 fragments from the present study and five of the six compounds previously reported can be found here. Only TROLL2 and its close analogue TROLL7, as well as ID2 are located in different regions of chemical space. The PCA analysis also indicates that there are additional enrichment options for further enhancements of screens against p97-N since the fragment library under investigation does not fully coincide with the chemical space where most of the p97-N ligands can be found.

**Ligand-based pharmacophore models.** The fragments identified here as well as those reported earlier<sup>26</sup> were further analyzed for common or similar functional groups in terms of putative pharmacophore models. High structural agreement was found between TROLL12, TROLL14, VIK20 and ID4 regarding the position of an aromatic feature (i.e., a phenyl ring), an aromatic heterocyclic scaffold and the position of a hydrogen-bond donor function represented by a hydroxyl- or amino-group (Fig. 4, superpositions a to c). TROLL18 and ID2 share a 3-amino-pyrazole substructure in combination with a para-substituted phenyl ring (Fig. 4d). The overlay of TROLL2 or its analogue TROLL7 with TROLL8 revealed two conserved hydrogen-bond donor functions, in combination with a hydrophobic, substituted aromatic ring system (Fig. 4e). Furthermore, TROLL6 and ID3 display similar structural properties (Fig. 4f). It can be hypothesized that the identified common features may be involved in interactions with residues in their p97 binding site.

**Ligand efficiencies of the identified fragments.** One of the main questions in the early phase of drug discovery is the selection of compounds best suited for optimization towards lead structures. Besides an increase in affinity and specificity, especially for weakly binding molecules such as fragments, the optimization towards favorable physicochemical and ADME (Absorption, Distribution, Metabolism, and Excretion) profiles is of high importance. For this purpose, different ligand efficiency metrics were developed to guide the process of molecular optimization<sup>36</sup>.

In this work, four different ligand efficiencies were calculated for the 10 selected molecules shown in Table 1 based on their  $K_D$  value for p97-N, the heavy atom count and the clogP value: the ligand efficiency (LE)<sup>37</sup>, the size-independent ligand efficiency (SILE)<sup>38</sup>, the lipophilicity corrected efficiency (LELP)<sup>39</sup>, and the modified lipophilic ligand efficiency by Mortenson and Murray (LLE<sub>AT</sub>)<sup>40</sup>. The four ligand efficiency metrics include indices for

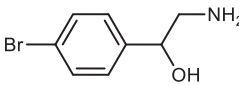
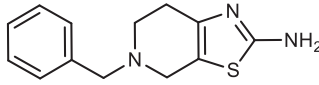
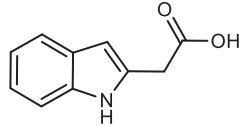
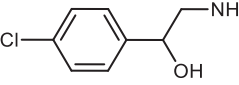
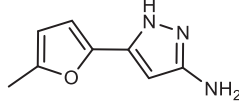
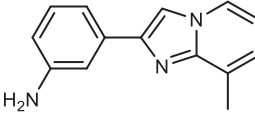
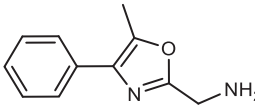
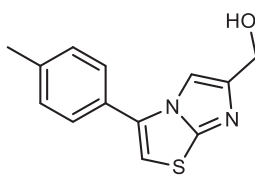
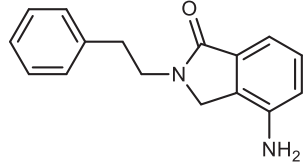
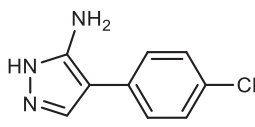
controlling the size (LE, SILE) and lipophilicity (LELP, LLE<sub>AT</sub>) of the fragments. For LE and SILE higher values indicate more promising candidates. A lower limit of 0.3 is commonly suggested for LE. LELP is calculated as  $(c)\log P/LE$  and outlines the cost of the LE paid in lipophilicity; with a value of 0.3 for LE, LELP should lie between  $-10$  and  $10$ . The LLE<sub>AT</sub> represents a more suitable metric for fragments, taking size and lipophilicity of the fragments into account. As with LE, its suggested lower cut-off is 0.3, and fragments with higher values are more promising<sup>40</sup>. The calculated indices were used to compare our fragment screen with screening campaigns reported in the literature. Table 2 summarizes average values for selected physicochemical descriptors and ligand efficiency metrics of published hits from fragment screens<sup>36</sup> and compares them with the values for the fragments binding to the N-domain reported by Chimenti et al.<sup>26</sup> and the fragments identified in this study.

The average values for the 10 hits of our p97-N fragment screen are very close to literature averages of successful fragment screens. In comparison to the fragments reported by Chimenti et al., the compounds identified in this study show superior properties. Especially the LELP and LLE<sub>AT</sub> parameters (which are dependent both on clogP and affinity) indicate that binding of the fragments identified by Chimenti et al. may be dominated by hydrophobic interactions, which is not ideal for further optimization. The hits identified here appear more promising because of a better balance between affinity and lipophilicity. In particular, the hit compounds TROLL2, TROLL7, TROLL8 and TROLL12 show an LLE<sub>AT</sub> value above the cut-off of 0.3, which indicates that the binding is not dominated by hydrophobic interactions.

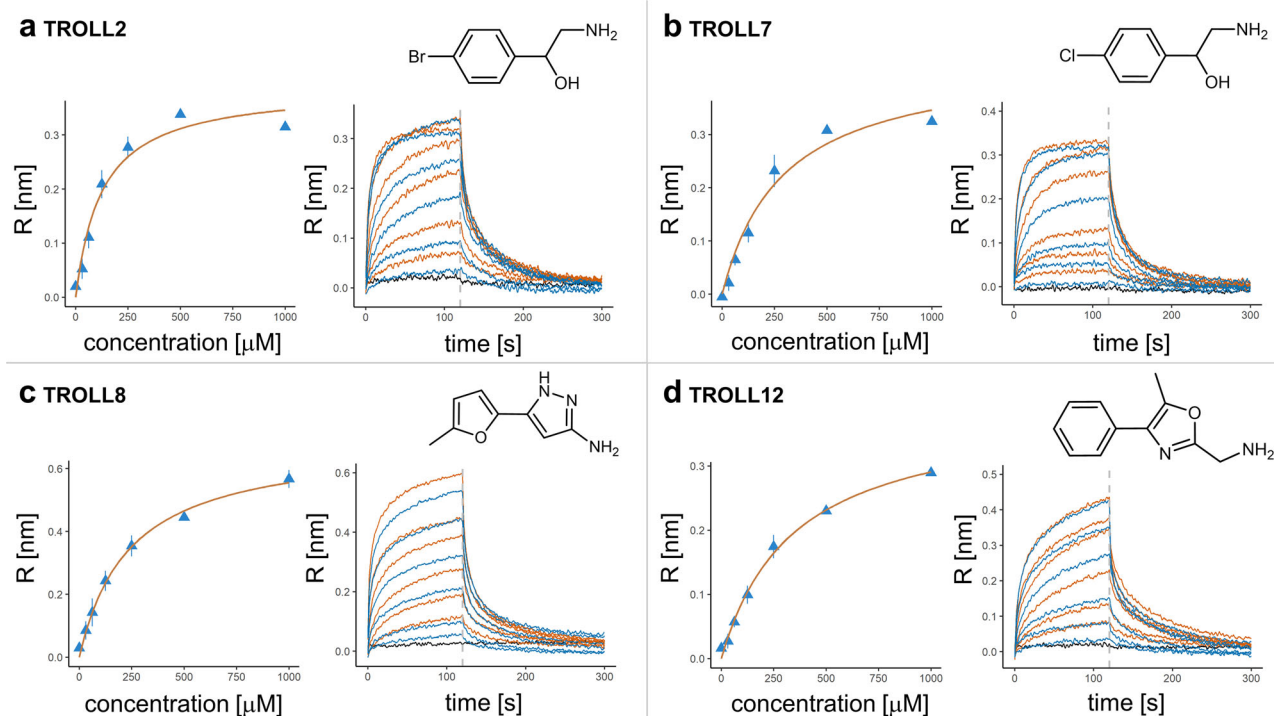
**Identification of p97 N-domain binding sites via mixed-solvent MD simulations.** Mixed-solvent MD simulations are a common tool to analyze possible interaction hot spots for small molecules in protein binding sites<sup>41–45</sup>. In contrast to blind docking approaches to the entire protein surface, where the protein is mostly rigid and no explicit solvent is used, mixed-solvent simulations provide a more realistic model of the protein in solution. Typically, fragments employed in this technology are even smaller in size than the fragments used in biophysical screens or the hits identified here. Nevertheless, Martinez-Rosell et al.<sup>46</sup> showed that the mixed-solvent simulation approach can handle such larger molecules as well. Consequently, mixed-solvent simulations were carried out to identify putative binding sites. The approach used here is based on the “Site Identification by Ligand Competitive Saturation” (SILCS) methodology<sup>41</sup>, where a high fragment concentration is subjected to an all-atom explicit-solvent MD simulation. The aggregation of fragments is prevented by dummy atoms that are placed into the molecules to exert a repulsive force between the fragments.

We simulated each fragment at a concentration of 1 M and calculated occupancies of the fragments on the surface of p97-N to identify possible binding regions. For every fragment, the two sites with the highest occupancies were selected, and the binding site was defined by all amino acids within 6 Å of the fragment's dummy atom. Most of the residues identified by this approach are in the Nc-subdomain of p97-N, especially in the area of the SHP binding region (Fig. 5), while only a few residues are located in the Nn-subdomain. This suggests the Nc-subdomain and particularly its SHP-binding region as the most druggable part of the N-domain. Based on the observations in the mixed-solvent simulations, the SHP binding region was subdivided into three regions (SHP I, SHP II and SHP III). SHP I is the region addressed by the conserved leucine residue of the SHP binding motif. SHP II is targeted by the aromatic amino acids F228 of

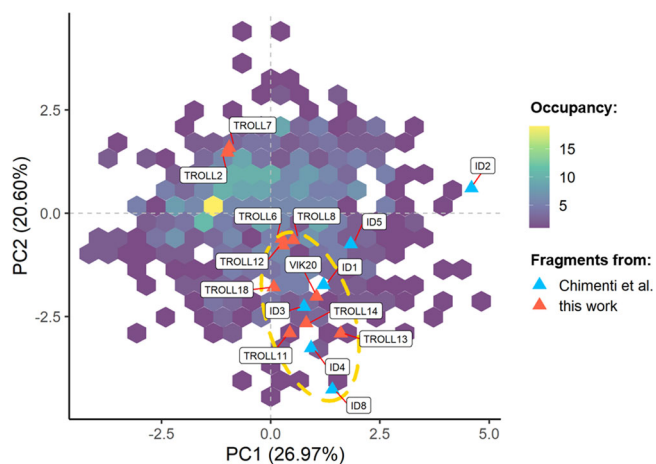
**Table 1 Overview of the identified fragments binding to the N-domain of p97.**

ID	Structure	$K_D$ p97-N [ $\mu$ M]	$K_D$ p97-ND1 [ $\mu$ M]	MW [Da]	HA	clogP	STD [%]	Score <sub>BLLI</sub>	LE	LLE <sub>AT</sub>
TROLL2		56 ± 46	321 ± 67	216	11	1.39	37.7	0.83	0.52	0.47
VIK20		147 ± 32	192 ± 92	245	17	1.80	18.3	0.94	0.30	0.27
TROLL6		160 ± 22	243 ± 16	175	13	2.04	72.5	0.75	0.39	0.30
TROLL7		222 ± 83	718 ± 57	172	11	1.18	37.3	0.91	0.45	0.42
TROLL8		245 ± 43	598 ± 214	200	12	1.09	41.0	0.91	0.41	0.40
TROLL11		323 ± 48	290 ± 179	223	17	2.08	43.7	0.90	0.28	0.22
TROLL12		354 ± 17	192 ± 92	118	14	1.09	47.9	0.84	0.33	0.34
TROLL14		409 ± 61	290 ± 33	244	17	2.29	67.4	0.93	0.27	0.16
TROLL13		409 ± 80	784 ± 118	252	19	2.42	38.7	0.92	0.24	0.18
TROLL18		625 ± 89	445 ± 223	194	13	2.78	57.4	0.89	0.33	0.27

The compounds are sorted by their affinity for p97-N. The affinity obtained with the p97-ND1 construct is shown for comparison. Errors are derived from the fitting model. Molecular weight (MW), heavy atom count (HA) and calculated logP (clogP) are provided as descriptors. STD refers to the maximum STD signal measured with p97-ND1. The calculated Score<sub>BLLI</sub> (based on the sensorgrams from measurements with p97-N) and two ligand efficiency metrics (LE and LLE<sub>AT</sub>, see Discussion section for definition) are based on the affinities for p97-N. IDs are internal screening identification numbers.



**Fig. 2 Sensorgrams and Langmuir models of selected fragments.** Results of the dose-response analysis of selected N-domain hits. For each fragment (a–d) the obtained shifts are plotted against the fragment concentration, with the fit to the Langmuir model shown in orange. Every concentration was measured twice and both values were used for fitting (with bars indicating the range of the two data points). Additionally, the sensorgrams of the fragments are shown after alignment and double referencing (first measurement: orange; second: blue). The negative control containing assay buffer only is shown in black. The dashed lines indicate the start of the dissociation phase. Sensorgrams which did not reach a plateau phase, because of unspecific binding or biphasic curve courses, received a penalty term in the calculated  $\text{Score}_{\text{BLI}}$  (see method description in Supplementary Methods).



**Fig. 3 PCA analysis of the fragment library employed in this study.** The locations of the fragments identified in this screen (red triangles) and the fragments reported by Chimenti et al.<sup>26</sup> (blue triangles) are highlighted. The hexagons (colored purple to yellow) indicate how many compounds were available in the library for a particular region of the chemical space. The dashed orange ellipse highlights the region where most of the fragments of both screens are located, indicating that these fragments share similar physicochemical parameters.

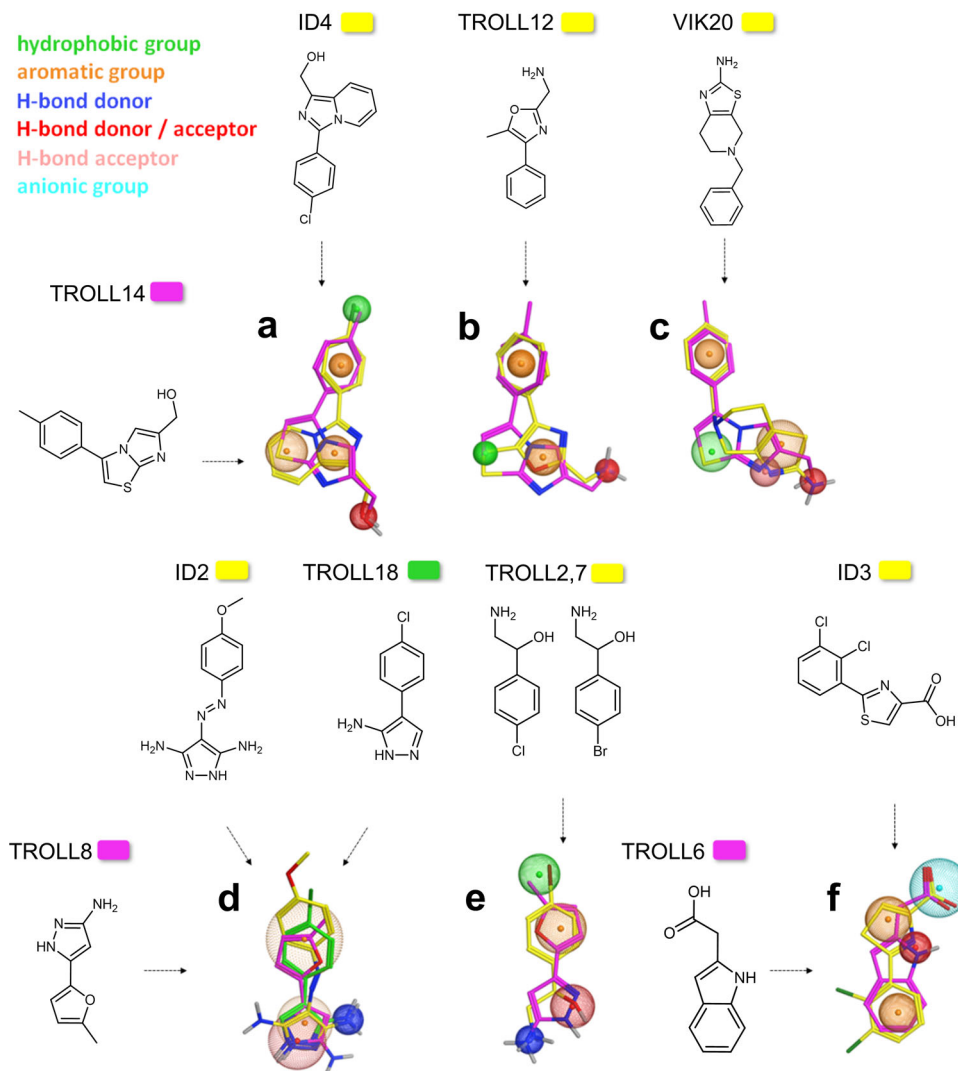
UFD1 and W242 of Derlin-1<sup>47,48</sup>. SHP III is an adjacent hydrophobic cavity not directly addressed by SHP motif containing cofactors, which harbors a solvent-exposed cysteine residue (C184).

Within the identified binding sites, the occupancies of the heavy atoms of each fragment were further inspected to obtain

information about the orientation of the molecules and to distinguish between specific and unspecific binding. In case of the four fragments TROLL2, TROLL7, TROLL8 and TROLL12, the resulting occupancies indicated an orientational preference, thus allowing to predict a binding mode based on the observed densities (Fig. 5e). For TROLL2 (which was simulated as both the *R*- and the *S*-enantiomer) only the *S*-enantiomer was predicted to bind within a pocket formed by residues R113, H115, E167 and H183 (SHP II, Fig. 5e, i). The closely related analogue TROLL7 presumably exhibits the same interactions (Fig. 5e, ii). TROLL12 was also found in this region, but oriented differently (Fig. 5e, iv), whereas TROLL8 bound to a sub-pocket near C184 in the SHP III region (Fig. 5e, iii). Additional occupancies for the methylfuran ring of TROLL8 were also identified in the SHP II region. The binding poses of TROLL2 and TROLL7 were identified in all replicas of the MD simulations, whereas for TROLL8 and TROLL12 a binding event to the aforementioned regions could be observed in three of the four replicas.

**Crystal structure of TROLL2 in complex with the N-domain of p97.** To confirm the predicted binding sites, X-ray crystallographic studies were performed. As the fragment with the highest binding affinity in the BLI screen, a high-resolution structure of TROLL2 in complex with p97-ND1 could be obtained. The identification of the Br-containing TROLL2 fragment and its binding site was assisted by the anomalous signal of the halogen atom. The structure was solved by molecular replacement using PDB structure 5DYG as the search model. In analogy to Tang & Xia, the same hexagonal space group (P622) with one ND1-monomer in the asymmetric unit was obtained<sup>49</sup>.

In four data sets a significant anomalous signal corresponding to one binding site as determined by the SHELXCD pipeline<sup>50</sup>



**Fig. 4 Pharmacophore models derived from selected fragments binding to p97-N.** Fragments of this study and from Chimenti et al.<sup>26</sup> were superimposed with respect to similar chemical properties or recognition elements. Spheres represent similar functional groups present in both fragments. The chemical properties of the spheres are explained in the colour code (top left). The resulting pharmacophore models **a-f** are described in the main text.

**Table 2 Comparison of descriptor values and ligand efficiency metrics for hits of fragment screens reported in the literature<sup>36</sup>, hits of the p97-N screen reported by Chimenti et al.<sup>26</sup> and the hits of this work as listed in Table 1.**

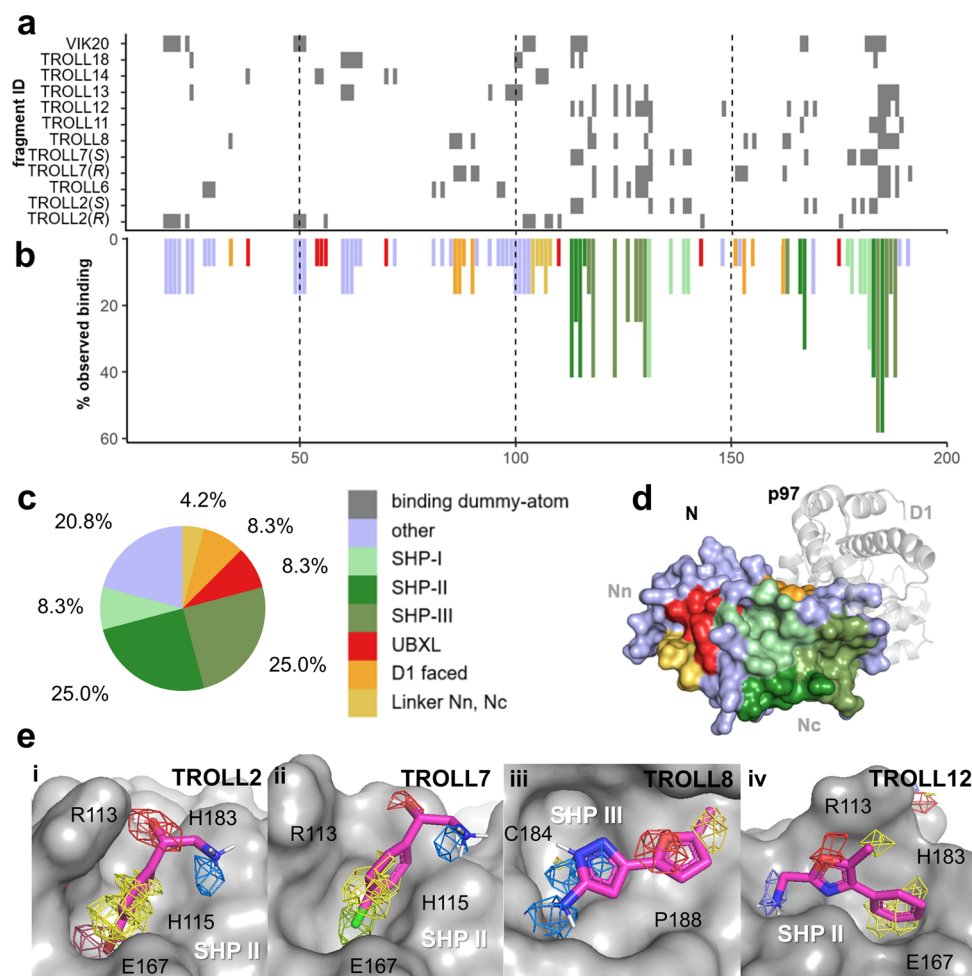
**Mean values**

Descriptor	Lit. <sup>36</sup> ( <i>n</i> = 145)	p97-N Lit. <sup>26</sup> ( <i>n</i> = 6)	p97-N ( <i>n</i> = 10)
MW	221	243	202
HA	15	17	15
pK <sub>D</sub> (pIC <sub>50</sub> )	4.1	3.0	3.5
(c)logP	1.8	2.4	1.8
LE	0.37	0.26	0.35
SILE	1.80	1.36	1.64
LELP	4.60	9.68	5.43
LLE <sub>AT</sub>	0.32	0.14	0.30

MW molecular weight (in g/mol), HA number of heavy (i.e., non-hydrogen) atoms; the four ligand efficiency indices are described in the text.

could be identified. To confirm that the anomalous signal is not an artefact of the crystallization conditions, data from apo-crystals (*n* = 43) obtained under identical conditions, but without any fragment in the crystallized solution, were analyzed for anomalous signals. No anomalous signal could be identified in the same region, and any observed anomalous signal was significantly weaker (by 6σ) compared to the TROLL2 data. These data demonstrate that the observed anomalous signal is due to the presence of the TROLL2 fragment. The data set with the strongest anomalous signal was refined to a crystallographic R-factor of 19.20% and a free R-factor of 23.53% at 1.73 Å resolution with Phenix<sup>51</sup> (see Supplementary Table 1 for data collection and refinement statistics). A superposition of PDB structure 5DYG with our structure in complex with TROLL2 showed a root mean square deviation (RMSD) of only 0.2 Å for the C<sub>α</sub>-atoms.

The anomalous signal of the Br-atom was unambiguously assigned to the binding pocket identified in the mixed-solvent simulations (Fig. 6a). A polder omit map contoured at 3σ showed a clear density for the Br-atom and most parts of the neighboring carbon atoms of the phenyl ring. Furthermore, weak densities are visible for the oxygen and the nitrogen atom of the ligand (Fig. 6b). The hydrophobic substructure of TROLL2 interacts



**Fig. 5 Results of mixed-solvent MD simulations.** The 10 hit fragments of Table 1 were simulated, including both enantiomers of TROLL2 and TROLL7. **a** Binding-site residue map. Grey rectangles specify amino acids along the primary sequence (horizontal axis) which are in the vicinity ( $<6 \text{ \AA}$  from the dummy atom) of the two sites with highest occupancies of the corresponding fragment. **b** Relative propensity of observed binding interactions. The bars indicate in how many of the 12 individual simulations the residue was identified as binding-site residue according to panel **a**, using the color code of panel **c**. **c** Pie chart with the distribution of the two sites with highest occupancy for each fragment across the N-domain surface ( $N = 24$ ). **d** Location of the corresponding regions on the surface, color-coded according to the pie chart. **e** Illustration of fragments for which a binding mode could be postulated from the mixed-solvent simulations: **i** TROLL2 and **ii** TROLL7 in the SHP II site; **iii** TROLL8 in a cavity near the SHP binding region (SHP III); and **iv** TROLL12 in SHP II, in a different orientation compared to TROLL2 and TROLL7. A representative pose in stick representation as well as the occupancies of selected heavy atoms in mesh representation are shown. The poses were selected as described in the Materials and Methods section.

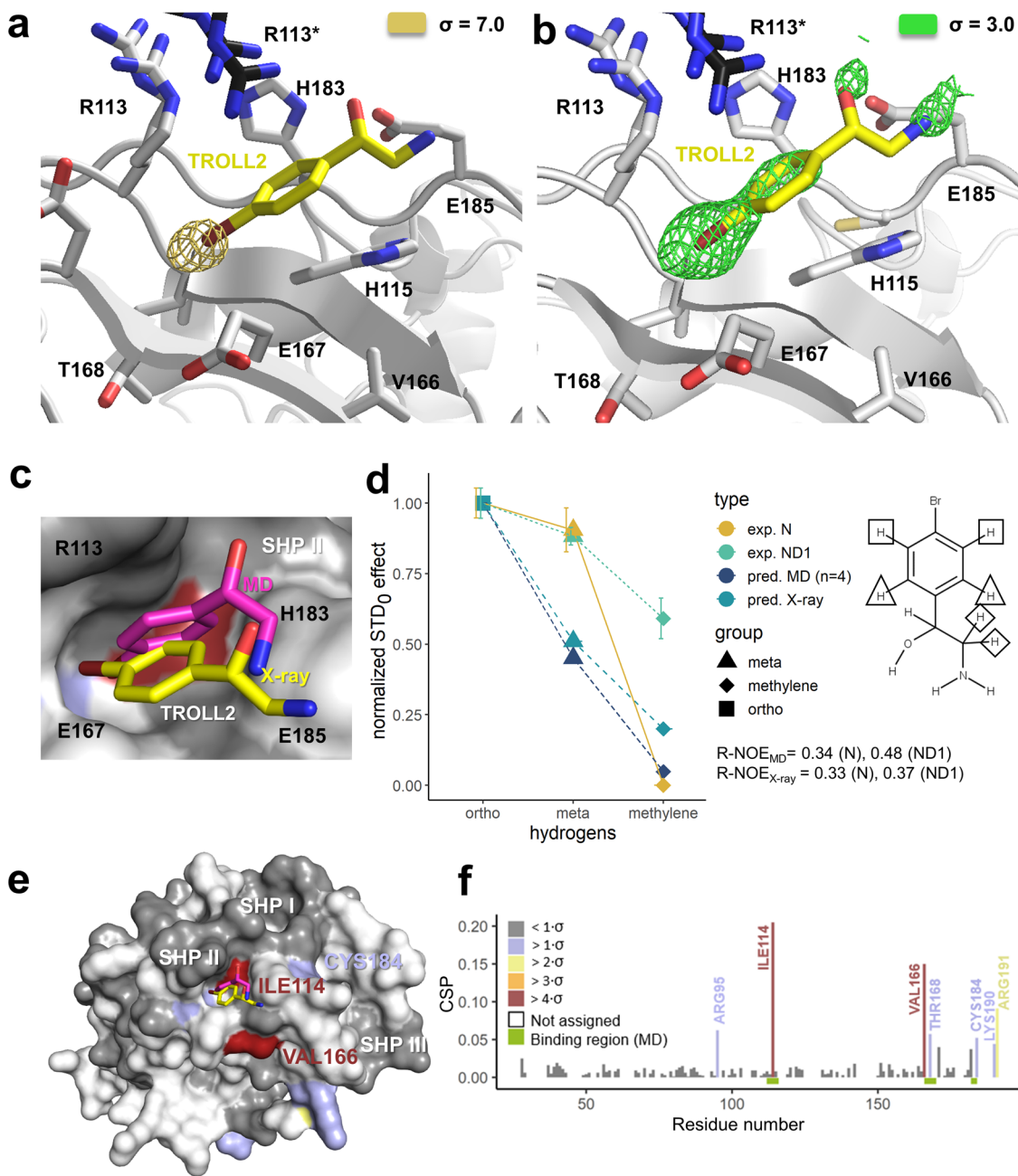
with a region formed by the aliphatic parts of residues R113, H115, E167 and E185. The phenyl ring of TROLL2 adopts a favorable orientation for a T-shaped interaction with H183. The charged head groups of R113, E167 and the adjacent D169 are all oriented away from the pocket, whereas the pocket itself is outlined by the alkyl chains of the corresponding side chains. To analyze whether there is an energetic binding preference for bromine or aromatic carbon atoms in this environment, a grid-based interaction potential analysis for bromine and aromatic carbon probes was conducted using MOE (Molecular Operating Environment, 2019.01, Chemical Computing Group) on the basis of the X-ray crystal structure (Supplementary Fig. 6). For both probes, interaction hot spots were predicted within this SHP II site, which is also in line with the aromatic residues of the SHP binding motif interacting with this area (Supplementary Fig. 6b).

The binding region contains a contact from an arginine (R113\*) of a symmetry-related N-domain which is not part of the p97-ND1 hexamer. This arginine engages in a cation- $\pi$  interaction with the phenyl ring of TROLL2. A detailed comparison of the results of the mixed-solvent MD simulations shows a high

level of agreement in the orientation of the resulting occupancy densities with the experimentally obtained structure (Fig. 6c). Additionally, the simulations revealed that R113 from the same monomer can also interact with the phenyl ring of the fragment via a cation- $\pi$  interaction; in solution, it should thus be able to replace the R113\* interaction resulting from crystal packing.

**Binding epitope mapping with STD-NMR.** To obtain additional information about the putative binding mode of the selected fragments, full STD-NMR build-up curves were measured for the fragments TROLL2, TROLL7, TROLL8 and TROLL12 and the initial rates of STD build-up ( $\text{STD}_0$ ) were calculated. Besides measurements with the ND1 construct, additional STD measurements with the isolated N-domain were carried out, which again confirmed the binding of all four fragments to the N-domain of p97 detected in the BLI measurements. As expected, the saturation effects were higher for the ND1 construct compared to the isolated N-domain due to the stronger nuclear Overhauser effect (NOE) of the much larger protein. Because of





**Fig. 6** Crystal structure of TROLL2 in complex with the SHP II binding site of p97-N and binding sites identified via  $^1\text{H}$ - $^{15}\text{N}$ -HSQC NMR measurements. **a** Anomalous map contoured at a sigma level of 7 highlighting the single signal in the asymmetric unit, which harbors one ND1 protomer. Only the binding region superimposed with the final refined model is shown, with TROLL2 carbon atoms highlighted in yellow (bromine atom in dark red). **b** Polder omit map of TROLL2 calculated with Phenix at a sigma level of 3. **c** Close-up view of the SHP II region with the MD-pose in magenta and the crystal structure in yellow. **d** Comparison of experimental (ND1, N) and predicted  $STD_0$  values for TROLL2 and assignment of protons visible in the STD-NMR spectra. Errors are derived from the fitting model. **e** Chemical shift perturbations (CSP) mapped onto the surface representation of p97-N for TROLL2. **f** CSP observed for TROLL2 plotted against the primary sequence.

water suppression, not all hydrogens of the selected fragments were experimentally accessible. For the remaining hydrogen atoms, different  $STD_0$  values were determined, indicating a defined orientation of the fragments upon binding to p97 (Supplementary Fig. 7). For TROLL2, TROLL7 and TROLL12 a clear dominance in the saturation of the protons at the hydrophobic aromatic parts of the molecules was visible.

To compare these observations from the STD experiments with the results of the mixed-solvent MD simulations, a representative binding pose of each fragment was analyzed with the

CORCEMA-ST program to calculate theoretical  $STD_0$  values (see Supplementary Methods in the Supporting Information). For TROLL2 the theoretical  $STD_0$  values were additionally predicted based on the crystal structure. The best agreement between calculated and experimental  $STD_0$  values was found for TROLL2 measured with the isolated N-domain. The protons in the ortho position received the highest amount of saturation, followed by the protons in the meta position and the methylene group, which is in line with the predicted pose of the MD simulation (NOE R-factor of 0.34). The calculation based on the experimental

crystal structure resulted in a nearly identical NOE R-factor of 0.33 (Fig. 6d) and showed high agreement with the predictions based on the MD simulation pose (NOE R-factor of 0.14).

It should be kept in mind that these weakly interacting fragments could have additional binding sites on p97 which are not covered by the conducted MD simulations. Moreover, even at a given binding site the fragments may show considerable flexibility or even occur in multiple binding modes which are not adequately represented by a single pose selected for the calculations. This is also obvious from the crystal structure, where the density shows binding but is not sharply defined for the entire fragment. Accordingly, the binding model used to predict STD effects is certainly incomplete, and a full agreement with the experimentally determined  $STD_0$  values can hardly be expected. Nonetheless, the CORCEMA predictions support the binding site of TROLL2 at the N-domain of p97. The close analog TROLL7, on the other hand, showed a higher saturation effect for the meta position, in contrast to the predicted effects based on the MD simulation, resulting in high NOE R-factors (Supplementary Table 2). An overlaid additional binding mode of TROLL7 not detected by the MD simulations represents a possible explanation. Similarly, the disagreement with the MD model for TROLL8 may indicate an additional or alternative binding site(s) (which is also suggested by the  $^1H$ - $^{15}N$ -HSQC measurements described in the next section). Finally, for TROLL12 the predicted binding pose shows a fair agreement in the trend of the STD effects, which is more pronounced with the measurements for the isolated N-domain. Taken together, the quantitative STD-NMR experiments confirm the binding to the N-domain of p97 and the CORCEMA-ST calculations support the predicted binding mode for TROLL2 and TROLL12, but not for TROLL7 and TROLL8. As mentioned, additional binding sites or binding poses (not atypical for fragments), which may not have been captured by the MD simulations, render quantitative predictions with CORCEMA-ST difficult.

#### $^1H$ - $^{15}N$ -HSQC NMR measurements with the N-domain of p97.

To clarify the results of the quantitative STD-NMR measurements and since no crystal structures of p97-ND1 in complex with TROLL7, TROLL8 and TROLL12 could be obtained,  $^1H$ - $^{15}N$ -HSQC NMR measurements with the isolated N-domain of p97 were performed to further complement the mixed-solvent MD simulations. The  $^1H$ - $^{15}N$ -HSQC NMR experiments of all fragments showed chemical shift changes within the regions postulated by the mixed-solvent MD simulations (Fig. 6e, f and Supplementary Fig. 8). Specifically, the largest chemical shift changes could be observed for TROLL7 and TROLL2 for residues I114 and V166, which are both located in the SHP II region. Small shifts for both fragments were also detected for T168 within the SHP II site and C184 in the neighboring SHP III region. Additional chemical shifts were assigned to R95, K190 and R191. In full-length p97 these three residues are close to the D1-domain and are either not accessible or exhibit a different conformation. Due to this fact the observed chemical shifts of these amino acids were not further considered. The remaining chemical shift changes were all in or in close proximity of the SHP II region, in agreement with the binding region identified by the mixed-solvent MD simulations and the crystal structure with TROLL2. TROLL12 showed only for I114 a slight chemical shift, originating within the predicted region of the mixed-solvent MD simulations, but the data here were not as robust as they were for TROLL2 or TROLL7.

For TROLL8 a significant chemical shift was observed for C184 in the SHP III region predicted by the simulations and for E141. E141 is located between the SHP I site and the UBXL binding

region. Both regions were not identified as possible binding sites by the mixed-solvent MD simulations; however, because of the lack of a suitable pocket in the UBXL region, a binding to the SHP I region seems to be more likely. Additional minor chemical shifts were located again within the SHP II region. Here, the simulations showed a low occupancy for the compound's methylfuran ring.

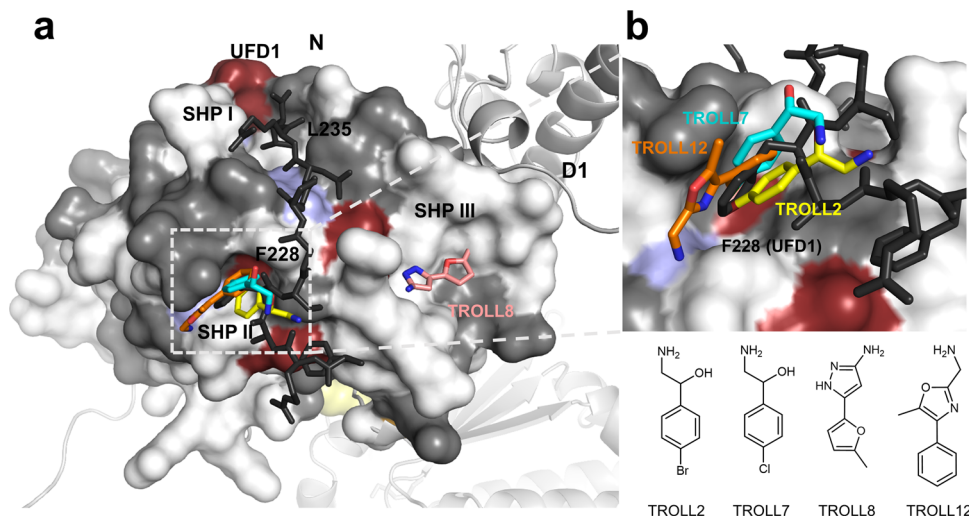
The observed chemical shift changes were not particularly strong and caution is certainly warranted in the interpretation. It should also be noted that parts of the SHP II site and most of the SHP III site could not be assigned and so no experimental data were available for these areas. Titration experiments for affinity determination were technically not possible, due to the poor solubility of the fragments. However, the observed chemical shifts were remarkably focused to single amino acids, and their location is consistent with the presumed fragment binding in the SHP region. Interestingly, no strong chemical shift changes were observed within the N terminal subdomain, in agreement with the general outcome of the mixed-solvent MD simulations. At the very least, the results of the  $^1H$ - $^{15}N$ -HSQC NMR measurements confirm the addressability of the SHP binding site with small molecules. Furthermore, by indicating additional interaction sites for TROLL8, the experiments provide a possible explanation for the discrepancy of the CORCEMA-ST predictions in case of this fragment.

## Discussion

**BLI is a well-suited method for fragment screening.** So far the BLI method has been reported only sporadically in the field of fragment-based drug design<sup>27</sup>. Here, we established BLI as a robust and well-performing fragment-screening platform. Key advantages of BLI as a fluid-free system are an easier setup and handling in comparison to other biosensor platforms. The present study demonstrates that this method is a suitable biophysical approach for detecting the binding of low molecular-weight compounds, thus providing promising candidates for a fragment-based drug design approach, also in the context of PPI inhibitors.

The N-domain of p97 as a main cofactor interaction site and potential target for PPI inhibitors was the focus of this work. Due to the absence of a specific binder as a positive control for the isolated N-domain, the quality of the screen was first assessed with the p97-ND1 construct. The screen with p97-ND1 revealed a Z-factor of 0.67 and a signal distribution that reflects a high overall quality. The optimized screening conditions for the ND1 domain were then transferred to the isolated N-domain. An overlap of 18 fragments between the two screens was detected, where most of the N-domain hits were also identified with the ND1 construct (see Supplementary Table 3). Stringent screening allowed to identify 10 high-confidence hits, which were subsequently confirmed by STD-NMR using the ND1 construct as well as the isolated N-domain and additional BLI measurements using the ND1 construct. The satisfactory performance of the screen was further demonstrated when comparing our results with the SPR-based screen of the ND1 domain conducted by Chimenti et al.<sup>26</sup> using principal component and pharmacophore analyses. This comparison showed that the BLI-based screen was capable of identifying similar fragments in terms of their position in the chemical space, but also revealed new binders exhibiting different properties.

Nonetheless, even when promising fragment candidates have been identified for a target protein, the selection of candidates for a more detailed experimental characterization and further optimization is a challenging task<sup>52</sup>. The  $Score_{BLI}$  applied in this work proved helpful for eliminating hits with unwanted properties, such as unspecific binding. With respect to ligand efficiencies



**Fig. 7** Overlap between postulated and experimentally verified fragment binding sites on p97 with the interaction sites of UFD1. **a** Binding sites of TROLL7, TROLL8 and TROLL12 identified in the mixed-solvent MD simulations, as well as the crystal structure of TROLL2 superimposed with the bound UFD1-SHP-peptide (PDB entry 5B6C) mapped onto a surface representation of the N-domain. Additionally, the highest chemical shifts of the  $^1\text{H}$ - $^{15}\text{N}$ -HSQC-NMR measurements are shown, colored in terms of standard deviation as in Fig. 6. **b** Zoom into the SHP II binding site, which is occupied by F228 in the structure of the p97-N-UFD1 peptide complex and is addressed by TROLL2, TROLL7 and TROLL12. Based on the  $^1\text{H}$ - $^{15}\text{N}$ -HSQC-NMR measurements, and, in contrast to the findings from the mixed-solvent MD simulations, TROLL8 seems to bind to a sub-pocket (SHP I) which is occupied by L235 in the structure of the p97-N-UFD1 peptide complex.

the fragments identified here display superior properties compared to those reported earlier<sup>26</sup>, in particular, our hits are not dominated by hydrophobic interactions and exhibit more favorable ligand efficiencies for subsequent optimization. The use of mixed-solvent MD simulations for the further selection of fragments binding to specific regions of the target protein proved to be a highly beneficial step as the postulated binding region could essentially be confirmed by X-ray and NMR measurements. Taken together, our BLI-based fragment screening strategy in combination with subsequent *in silico* tools turned out to be a successful approach.

#### Identification of starting points for PPI inhibitors targeting the p97 SHP-motif binding site.

A selective inhibitor of cofactor binding to the N-domain of p97 would represent a new and valuable avenue in the field of p97 inhibitors. The work described here represents a significant advance towards this goal. The structural relevance of the identified fragments, TROLL2, TROLL7, TROLL8 and TROLL12, as starting points for the development of a PPI inhibitor is shown in Fig. 7. A comparison of the binding sites in the crystal structures of the SHP-motif containing cofactors Derlin-1 and UFD1 with the identified target sites of the fragments reveals a high degree of overlap, in particular in the SHP-II region. The identified fragments address two important hotspots of the PPI of SHP-motif containing cofactors: (i) The postulated binding poses of TROLL7 and TROLL12 derived from the mixed-solvent MD simulations as well as the binding pose found in the crystal structure of TROLL2 mimic the binding of the aromatic amino acids F228 of UFD1 and W242 of Derlin-1 in the SHP-II region, respectively. The importance of this hotspot interaction for the binding of the SHP-motif containing cofactor UFD1 is demonstrated in the reduced affinity of the UFD1 F228A mutant towards p97<sup>47,53</sup>. (ii) Furthermore, the  $^1\text{H}$ - $^{15}\text{N}$ -HSQC NMR measurements indicated that TROLL8 might bind adjacent to the SHP-III region into a sub-pocket referred to as SHP-I region, which is addressed by L235 (UFD1) or L248 (Derlin-1). Like the aforementioned aromatic residues these aliphatic side chains are hotspot residues and are important

for the binding to p97 as shown by their respective replacements with alanine<sup>48,53</sup>.

Taken together, our results suggest that the design of a PPI inhibitor addressing this site should be possible, since the SHP-motif binding region on p97 has the ability to bind to small molecules. Besides the potential to more comprehensively understand the regulation of p97 by its cofactors, a PPI inhibitor targeting the interaction between p97 and SHP-motif containing cofactors would yield a promising strategy to modulate the degradation of ubiquitylated proteins, with the ultimate promise of a new strategy to treat cancer.

Our successful fragment screen against the N-domain of p97 used the previously underemployed BLI approach and establishes it as a viable alternative to SPR. Importantly, the characterization of the obtained hits resulted in the first crystal structure of a small molecule in complex with the SHP-motif binding region of p97. Mixed-solvent MD simulations and NMR measurements also indicated this region as the most likely target site of the top-ranked fragments. The compounds identified in this study display favorable physicochemical properties for further drug design and provide promising starting points for the development of a PPI inhibitor addressing the SHP-motif binding site in the N-domain of p97.

#### Materials and methods

**Cloning, protein expression and purification.** Constructs of p97-N (aa 21-199, N-terminal 3 C protease-cleavable His<sub>6</sub>-tag) and p97-ND1 (aa 2-481, N-terminal TEV protease-cleavable His<sub>6</sub>-tag) harbouring a C-terminal Avi-tag were generated by inverse mutagenesis as described<sup>28</sup>. The correct nucleotide sequence of all constructs was verified by DNA sequencing (Microsynth Seqlab, Göttingen, Germany). Expression and purification were carried out as described in Hänzelmann et al.<sup>54</sup>. Biotinylation of Avi-tagged p97-N and p97-ND1 was conducted as described<sup>28</sup>, followed by size exclusion chromatography.

For crystallization p97-ND1(L198W) (residues 1-460) encoding a C-terminal His<sub>6</sub>-tag with a two amino acid linker (RS) according to the construct used by the Xia lab<sup>49</sup> was generated by the restriction free cloning method<sup>55</sup> and site-directed mutagenesis (QuickChange Site-Directed Mutagenesis Kit, Agilent). Expression and purification were carried out as described in Tang and Xia<sup>49</sup>.

Bacterial expression of  $^{15}\text{N}$ -labelled protein was performed according to a published protocol<sup>56</sup>. p97-N (aa 1-208, N-terminal TEV protease-cleavable His<sub>6</sub>-tag<sup>57</sup>) was expressed in *E. coli* BL21 (DE3) RIL cells. Cells were grown in LB medium until the OD<sub>600</sub> reached a value between 0.6 and 0.7. After centrifugation, the medium was

removed and the cells were resuspended in M9 minimal medium supplemented with  $^{15}\text{NH}_4\text{Cl}$  as the sole nitrogen source, where the volume of the M9 medium corresponded to one quarter of the amount of LB medium previously used. After incubating the cells for 1 h at room temperature, p97 expression was induced with 1 mM IPTG overnight at 16 °C. The protein was purified in a buffer containing 50 mM Tris pH 8, 150 mM KCl, 5% glycerol (v/v), 5 mM  $\text{MgCl}_2$  and 5 mM  $\beta$ -mercaptoethanol by Ni-IDA affinity chromatography and size exclusion chromatography (Superdex 200, Cytiva). Prior to size exclusion chromatography, the protein was dialyzed overnight at 4 °C in the presence of TEV protease (1:50) followed by Ni-IDA affinity chromatography to remove uncleaved His-tagged protein.

### BLI fragment screening

**Fragment library.** For fragment screening, the library compiled by the Facility for Biophysics, Structural Biology and Screening at the University of Bergen (BiSS) was used. This fragment library is a 679 compound subset of the OTAVA chemicals solubility library which was filtered to exclude similar compounds and compounds with unwanted functionalities<sup>58</sup>. The fragments have a mean molecular weight of 202 Da and an average clogP (calculated partition coefficient) of 1.52. The fragments contain on average two ring systems, one hydrogen-bond donor, two hydrogen-bond acceptors and two rotatable bonds. Hits from the initial screenings were ordered from OTAVA chemicals for further experimental characterizations.

**BLI sensor preparation.** Protein constructs for BLI measurements were enzymatically biotinylated using BirA<sup>28</sup>. Biotinylated p97 was loaded on Super Streptavidin Biosensors (SSA) obtained from Sartorius as follows. SSA-sensors were equilibrated in PBS-buffer, loaded with a protein solution containing 50  $\mu\text{g}/\text{mL}$  (p97-ND1) or 100  $\mu\text{g}/\text{mL}$  (p97-N) protein, blocked with biocytin and washed in PBS buffer. Reference SSA-sensors were set up by blocking them with a 10  $\mu\text{g}/\text{mL}$  biocytin solution for 5 min<sup>27</sup>. The SSA-sensors were loaded to a shift of up to 7 nm with p97-ND1 and 10 nm with p97-N, respectively.

**General screening setup.** Screening plates were prepared by adding 10  $\mu\text{L}$  of a 10 mM DMSO stock solution of library compound to 190  $\mu\text{L}$  of PBS buffer (containing 0.05% TWEEN 20, 1 mM DTT and 5 mM  $\text{MgCl}_2$  [only for p97-ND1]) in 96 well plates (Greiner). The final DMSO concentration was 5% (v/v). All BLI measurements were performed on an Octet RED96e (Sartorius) device.

All fragments were measured twice at a concentration of 500  $\mu\text{M}$  by screening every plate in the forward and backward direction. The assay settings were as follows: Baseline measurement 15 s; association time 60 s and dissociation time 300 s. The resulting data were processed using the double reference method of the Octet Analysis software.

**Initial screening using the ND1-domain.** Signals of positive (500 nM ADP) and negative controls (buffer) were collected before and after each screening plate. The signals of each sensor were corrected for partial inactivation of the protein over time and for different protein loading states of the sensors using the signals of the 500 nM ADP positive control (see Supplementary Fig. 9 for details on the correction approach).

**Initial screening using the N-domain.** One fragment (ID5) reported to bind to the N-domain<sup>26</sup> could be purchased from Enamine and was tested as a positive control. The compound showed an uninterpretable behavior in the BLI assay using the ND1- as well as the N-domain and no valid dose response could be obtained. Because of the lack of another valid positive control (small molecule) for the isolated N-domain of p97, the assay conditions of the ND1 screen were transferred to the N-domain screen. Every sensor with immobilized p97-N was used for measuring only two screening plates and then discarded. Sensors showing artificial signals during the screening were replaced. The signals were normalized against the protein loading signal of each sensor. Signals with artificially high or large negative values based on visual inspection were eliminated.

**Concentration dependencies.** Selected fragments of the initial screenings were confirmed by dose-response-titrations using six concentrations ranging from 1000 to 31.3  $\mu\text{M}$  in a 1:1 dilution series. Every concentration was measured twice. The assay settings were as follows: baseline measurement 15 s; association time 120 s and dissociation time 180 s. The double referenced data were further analyzed in Origin Pro to estimate the affinity by fitting the signals of the steady states to a 1:1 Langmuir model. For the fitting all measured data points ( $n = 2$ ) were considered.

**Principal component analysis (PCA) and ligand-based pharmacophore models.** The PCA for fragments of the BiSS library was carried out in R (R Core Team, 2020, Vienna, Austria: R Foundation for Statistical Computing) using the prcomp function. It was based on the following descriptors: HA, number of heavy atoms; HDon, number of hydrogen-bond donors; HAcc, number of hydrogen-bond acceptors; TPSA, topological polar surface area; Nrot, number of rotatable bonds; NChir, number of chiral atoms; NRings, number of ring systems; clogP, calculated logarithmic partition coefficient; Narom, number of aromatic atoms; Fsp3, fraction of sp3 hybridized atoms. All descriptors were calculated with MOE (Molecular

Operating Environment, 2019.01, Chemical Computing Group ULC, 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A2R7, 2021). A flexible molecular alignment was carried out in MOE with the selected fragments of this study and the fragments reported by Chimenti et al.<sup>26</sup>. The aligned molecules were analyzed using the 3D pharmacophore builder in MOE.

**Mixed-solvent MD simulations.** For the mixed-solvent MD simulations the structure of p97-N in complex with the SHP-motif of UFD1 (PDB:5B6C)<sup>47</sup> was used. All water and buffer molecules as well as the UFD1 peptide were removed. The structure was prepared and protonated (at pH 7) in MOE using default values. For mixed-solvent simulations the Site Identification by Ligand Competitive Saturation (SILCS) method<sup>41</sup> was used. To prevent aggregation of lipophilic fragments, dummy atoms were inserted at the center of mass of each ligand. These atoms served as virtual sites for repulsive interactions (only between fragments) using the Lennard-Jones parameters adopted from the SILCS method ( $\epsilon = -0.01$  kcal/mol;  $R_{\text{min}} = 24.0$  Å), but without applying a switching function. 1.14\*CM1A-based charges for the ligands were retrieved from the LigParGen web server<sup>59-61</sup>. A cubic simulation box with a 1 M fragment concentration and 5 nm edge length was solvated with TIP3P water. The solvation box was energy-minimized for 50,000 steps using the steepest descent minimization, equilibrated under NVT and, subsequently, NPT conditions for 100 ps each, and then simulated for 1 ns (NPT) using GROMACS [2019.01]<sup>62</sup> with the OPLS-AA force field<sup>63</sup> in order to obtain a mixed-solvent box with a homogenous ligand concentration. Coupling of temperature and pressure was performed using the velocity rescaling thermostat and the Parrinello-Rahman barostat<sup>64</sup>. The protonated protein structure was then solvated with this mixed-solvent box, resulting in a cubic starting system with at least a 1 nm distance from the simulation box border to the protein. The system was again minimized and equilibrated as described above and then subjected to a 40 ns production run with 2 fs time steps, writing out snapshots every 10 ps. Positional restraints on protein heavy atoms were applied during the equilibration phase. Four replicas were performed for each ligand, resulting in a total simulation time of 160 ns per fragment. For TROLL2 and TROLL7 simulations of both stereoisomers were carried out. The trajectories were aligned on the  $C_\alpha$  atoms and occupancy densities were calculated for the dummy atoms and selected heavy atoms of the fragments with a resolution of 1.0 Å using the volmap plugin in VMD<sup>65</sup>. The densities of individual replicas were added using the function `volutil` resulting in a sum density map for each fragment. In cases where the resulting occupancies indicated an orientational preference, thus allowing to predict a binding mode based on the observed densities, a representative binding pose of the corresponding fragment was selected as follows: For each replica, an average structure was calculated from the trajectory, starting at the time point where the binding event occurred. The snapshot showing the smallest RMSD with respect to this average structure was taken as binding pose of this replica. The binding poses from each replica of a given fragment were compared with the calculated occupancies, and the pose showing the highest agreement with the occupancies was chosen as representative binding pose of that fragment.

**X-ray crystallography.** The p97-ND1 L198W variant was incubated with ADP (molar ratio 1:13) and crystals were grown at 20 °C by vapor diffusion in hanging drops containing equal volumes (1:1) of protein solution (15 mg/ml) and reservoir solution consisting of 5% PEG 600 (v/v), 4.0 M sodium formate (pH 6.0), 5% glycerol (v/v) and TROLL2 in a final molar ratio of 1:10.

p97-ND1 crystals were cryo-protected by soaking in mother liquor containing increasing amounts (10–30% [v/v]) of glycerol and TROLL2 in a molar ratio of 1:10. The crystals were flash-cooled in liquid nitrogen, and data collection was performed at 100 K at the ESRF in Grenoble (beamline ID30B, wavelength 0.9184 Å). Data were processed using XDS<sup>66</sup> and STARANISO<sup>67</sup>. The anomalous signal of the bromine atom was detected using SHELX<sup>50</sup> and AnDe<sup>68</sup>. Further calculations were performed using the CCP4 Suite<sup>69</sup> and Phenix<sup>51</sup>. Phases were obtained by molecular replacement using Phaser with the ND1-domain of p97 (PDB: 5DYG<sup>49</sup>) as search model. TROLL2 was modelled based on the difference map. The structure was ultimately refined in Phenix<sup>51</sup>. TLS groups were obtained from the TLSMD-Webserver<sup>70</sup>. Restraints for TROLL2 were calculated using `eLBOW` in Phenix.

**STD-NMR spectroscopy.** All Saturation-Transfer-Difference (STD)<sup>32</sup> measurements were performed on a Bruker Avance III 400 MHz spectrometer (Karlsruhe, Germany) operating at 400.13 MHz with a PABBI 1H/D-BB Z-GRD inverse probe at 300 K adjusted with a BSVT (Bruker) temperature control unit. Temperature calibration was carried out with 99.8%  $\text{CD}_3\text{OD}$  (282–330 K). The `stdiffesgp.3` pulse program from the Bruker pulseprogram library was used for data acquisition. The suppression of the HDO signal (4.703 ppm) was achieved by the excitation sculpting method. The saturation frequency for the on- and off-resonances was set to  $-400$  Hz and  $-16,000$  Hz, respectively. The saturation time was 3 s with a relaxation delay D1 of 4 s.

The protein stock solutions were transferred to a deuterated buffer containing 35 mM potassium phosphate and 25 mM NaCl in  $\text{D}_2\text{O}$  at a pD of 7.5 using 0.5 mL Centricon devices (Merck Millipore) with a MWCO of 30 kDa by washing and centrifuging several times at  $10,000 \times g$  and 4 °C. Buffer-exchanged protein was

adjusted to a concentration of 30  $\mu\text{M}$  with NMR buffer. The respective fragment concentration was 428  $\mu\text{M}$  with a final DMSO- $d_6$  fraction of 5% (v/v). All measurements were performed in the presence of 500  $\mu\text{M}$  ADP (in case of the ND1 domain) and 5 mM  $\text{MgCl}_2$ . All data were analyzed using TopSpin 4.0.8 (Bruker Corporation, Billerica, MA, USA).

For binding epitope mapping of selected fragments, the initial growth rates approach was used as described<sup>71</sup>. For determination of the initial rate of STD build up (STD<sub>0</sub>) the STD intensities for seven saturation times (0.5, 0.75, 1.0, 1.5, 2.0, 3.0, 5.0 s) were measured for the ND1 as well as for the isolated N-domain and fitted by an exponential equation using R. The recycle delay (d1) was set to 5 s. For each fragment also a negative control without protein was measured with a saturation time of 5.0 s. Theoretical STD<sub>0</sub> values on the basis of the obtained X-ray crystal structure as well as for representative mixed-solvent MD poses were calculated using the Complete Relaxation and Conformational Exchange Matrix Analysis of Saturation Transfer (CORCEMA-ST) software version 3.8<sup>72</sup> and ShiftX version 1.1<sup>73</sup>. Predicted and experimental STD-data were evaluated using the NOE R-factor (see Supplementary Methods in the Supporting Information).

**HSQC NMR measurements.** <sup>15</sup>N-labelled p97-N was concentrated to 50–60  $\mu\text{M}$  in 25 mM HEPES (pH 7.5), 125 mM NaCl, 5 mM DTT and 0.01%  $\text{NaN}_3$ . Before the measurements were conducted, 10%  $\text{D}_2\text{O}$  and 3-(trimethylsilyl)propane-1-sulfonate (DSS) at a final concentration of 0.1 mM were added to the samples. DSS was used for direct <sup>1</sup>H chemical shift referencing as 0.00 ppm. <sup>15</sup>N chemical shifts were indirectly referenced by the gyromagnetic ratio<sup>74</sup>.

<sup>1</sup>H-<sup>15</sup>N-HSQC spectra were recorded on a Bruker Avance III (600 MHz) spectrometer at 298 K. DMSO-dissolved ligands were added to the proteins, resulting in molar ratios of 1:80 or 1:160 (protein:ligand). As a control, equivalent volumes of pure DMSO were added. Assignments were obtained from previously recorded spectra of the p97 N-domain<sup>75,76</sup>. Spectra were plotted and analysed using CcpNMR Analysis<sup>77,78</sup> version 3.0.4. Chemical Shift Perturbations (CSP) were calculated according to the equation:

$$\text{CSP} = \sqrt{\Delta\delta_{\text{H}}^2 + (0.13 \cdot \Delta\delta_{\text{N}})^2}$$

**Reporting summary.** Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

Supplementary Information (Figures, Tables, Methods, chemical identity of the four selected fragments, and References) is provided as a single separate file in PDF format. Atomic coordinates and structure factors of the crystal structure of TROLL2 in complex with p97-ND1 have been deposited with the Protein Data Bank with PDB accession code 7PUX (<https://www.rcsb.org/structure/7PUX>). Supplementary Data have been deposited at the Zenodo repository, where they are accessible under the following DOI link: <https://doi.org/10.5281/zenodo.7300018>. These include: i) the BLI sensorgrams of the 10 binding fragments identified in the screening (shown in Table 1), both for the N-domain and the ND1 construct of p97, along with the measurements of ADP as positive control; ii) for TROLL2, TROLL7, TROLL8 and TROLL12 the STD-NMR build-up data, the HSQC-NMR spectra and the representative poses from the mixed solvent MD simulations used for CORCEMA predictions; and iii) for TROLL2 the anomalous map of bromine for the crystal structure of the TROLL2 - p97-ND1 complex (PDB code 7PUX). All additional data are available from the corresponding authors on reasonable request.

Received: 4 May 2022; Accepted: 17 November 2022;

Published online: 07 December 2022

## References

- Stach, L. & Freemont, P. S. The AAA+ ATPase p97, a cellular multitool. *Biochem. J.* **474**, 2953–2976 (2017).
- Xia, D., Tang, W. K. & Ye, Y. Structure and function of the AAA+ ATPase p97/Cdc48p. *Gene* **583**, 64–77 (2016).
- Hänzelmann, P. & Schindelin, H. Structural Basis of ATP Hydrolysis and Intersubunit Signaling in the AAA+ ATPase p97. *Structure* **24**, 127–139 (2016).
- Buchberger, A., Schindelin, H. & Hänzelmann, P. Control of p97 function by cofactor binding. *FEBS Lett.* **589**, 2578–2589 (2015).
- Hänzelmann, P. & Schindelin, H. The Interplay of Cofactor Interactions and Post-translational Modifications in the Regulation of the AAA+ ATPase p97. *Front. Mol. Biosci.* **4**, 21 (2017).
- Tang, W. K. & Xia, D. Mutations in the human AAA+ chaperone p97 and related diseases. *Front. Mol. Biosci.* **3**, 1–12 (2016).
- Huryn, D. M., Kornfilt, D. J. P. & Wipf, P. P97: An Emerging Target for Cancer, Neurodegenerative Diseases, and Viral Infections. *J. Med. Chem.* **63**, 1892–1907 (2020).
- Bastola, P., Oien, D. B., Cooley, M. & Chien, J. Emerging Cancer Therapeutic Targets in Protein Homeostasis. *AAFS J.* **20**, 94 (2018).
- Anderson, D. J. et al. Targeting the AAA ATPase p97 as an approach to treat cancer through disruption of protein homeostasis. *Cancer Cell.* **28**, 653–665 (2015).
- Gugliotta, G. et al. Valosin-Containing Protein/p97 as a Novel Therapeutic Target in Acute Lymphoblastic Leukemia. *Neoplasia USA* **19**, 750–761 (2017).
- Chapman, E., Maksim, N., De La Cruz, F. & La Clair, J. J. Inhibitors of the AAA+ chaperone p97. *Molecules* **20**, 3027–3049 (2015).
- Sui, X., Pan, M. & Li, Y.-M. Insights into Design of p97-targeting Small Molecules from Structural Studies on p97 Functional Mechanism. *Curr. Med. Chem.* **26**, 1–18 (2019).
- Chou, T. F. et al. Reversible inhibitor of p97, DBEq, impairs both ubiquitin-dependent and autophagic protein clearance pathways. *Proc. Natl. Acad. Sci.* **108**, 4834–4839 (2011).
- Chou, T. F., Li, K., Frankowski, K. J., Schoenen, F. J. & Deshaies, R. J. Structure-Activity Relationship Study Reveals ML240 and ML241 as Potent and Selective Inhibitors of p97 ATPase. *ChemMedChem* **8**, 297–312 (2013).
- Magnaghi, P. et al. Covalent and allosteric inhibitors of the ATPase VCP/p97 induce cancer cell death. *Nat. Chem. Biol.* **9**, 548–559 (2013).
- Tang, W. K., Odzorig, T., Jin, W. & Xia, D. Structural basis of p97 inhibition by the site-selective anticancer compound CB-5083. *Mol. Pharmacol.* **95**, 286–293 (2019).
- Polucci, P. et al. Alkylsulfanyl-1,2,4-triazoles, a new class of allosteric valosine containing protein inhibitors. Synthesis and structure-activity relationships. *J. Med. Chem.* **56**, 437–450 (2013).
- Alvarez, C. et al. Structure-Activity Study of Bioisosteric Trifluoromethyl and Pentafluorosulfanyl Indole Inhibitors of the AAA ATPase p97. *ACS Med. Chem. Lett.* **6**, 1225–1230 (2015).
- Pöhler, R. et al. A Non-Competitive Inhibitor of VCP/p97 and VPS4 Reveals Conserved Allosteric Circuits in Type I and II AAA ATPases. *Angew. Chem. Int. Ed.* **57**, 1576–1580 (2018).
- Segura-Cabrera, A. et al. A structure- and chemical genomics-based approach for repositioning of drugs against VCP/p97 ATPase. *Sci. Rep.* **7**, 44912 (2017).
- Wang, Q. et al. The ERAD inhibitor eeyarestatin I is a bifunctional compound with a membrane-binding domain and a p97/VCP inhibitory group. *PLoS One* **5**, e15479 (2010).
- Sasazawa, Y. et al. Xanthohumol impairs autophagosome maturation through direct inhibition of valosin-containing protein. *ACS Chem. Biol.* **7**, 892–900 (2012).
- Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W. & Jhoti, H. Twenty years on: The impact of fragments on drug discovery. *Nat. Rev. Drug Discov.* **15**, 605–619 (2016).
- Braisted, A. C. et al. Discovery of a potent small molecule IL-2 inhibitor through fragment assembly. *J. Am. Chem. Soc.* **125**, 3714–3715 (2003).
- Guo, W., Wisniewski, J. A. & Ji, H. Hot spot-based design of small-molecule inhibitors for protein-protein interactions. *Bioorg. Med. Chem. Lett.* **24**, 2546–2554 (2014).
- Chimenti, M. S. et al. A fragment-based ligand screen against part of a large protein machine: The ND1 domains of the AAA+ ATPase p97/VCP. *J. Biomol. Screen.* **20**, 788–800 (2015).
- Wartchow, C. A. et al. Biosensor-based small molecule fragment screening with bilayer interferometry. *J. Comput. Aided Mol. Des.* **25**, 669–676 (2011).
- Fairhead, M. & Howarth, M. Site-Specific Biotinylation of Purified Proteins Using BirA. In: Gautier, A. & Hinner, M. (eds) Site-Specific Protein Labeling. *Methods Mol. Biol.* **1266**, 171–184 (2015).
- Chou, T.-F. et al. Specific Inhibition of p97/VCP ATPase and Kinetic Analysis Demonstrate Interaction between D1 and D2 ATPase domains. *J. Mol. Biol.* **426**, 2886–2899 (2014).
- Zhang, J. H., Chung, T. D. Y. & Oldenburg, K. R. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen.* **4**, 67–73 (1999).
- Aretz, J. & Rademacher, C. Ranking Hits From Saturation Transfer Difference Nuclear Magnetic Resonance-Based Fragment Screening. *Front. Chem.* **7**, 215 (2019).
- Viegas, A., Manso, J., Nobrega, F. L. & Cabrita, E. J. Saturation-transfer difference (STD) NMR: A simple and fast method for ligand screening and characterization of protein binding. *J. Chem. Educ.* **88**, 990–994 (2011).
- Jinuraj, K. R. et al. Feature optimization in high dimensional chemical space: Statistical and data mining solutions. *BMC Res. Notes.* **11**, 1–7 (2018).
- Reymond, J. L. The Chemical Space Project. *Acc. Chem. Res.* **48**, 722–730 (2015).
- Akella, L. B. & DeCaprio, D. Cheminformatics approaches to analyze diversity in compound screening libraries. *Curr. Opin. Chem. Biol.* **14**, 325–330 (2010).

36. Ferenczy, G. G. & Keseru, G. M. How are fragments optimized? A retrospective analysis of 145 fragment optimizations. *J. Med. Chem.* **56**, 2478–2486 (2013).
37. Kuntz, I. D., Chen, K., Sharp, K. A. & Kollman, P. A. The maximal affinity of ligands. *Proc. Natl Acad. Sci.* **96**, 9997–10002 (1999).
38. Nissink, J. W. M. Simple size-independent measure of ligand efficiency. *J. Chem. Inf. Model.* **49**, 1617–1622 (2009).
39. Keseru, G. M. & Makara, G. M. The influence of lead discovery strategies on the properties of drug candidates. *Nat. Rev. Drug Discov.* **8**, 203–212 (2009).
40. Mortenson, P. N. & Murray, C. W. Assessing the lipophilicity of fragments and early hits. *J. Comput. Aided Mol. Des.* **25**, 663–667 (2011).
41. Guvench, O. & MacKerell, A. D. Computational fragment-based binding site identification by ligand competitive saturation. *PLoS Comput. Biol.* **5**, e1000435 (2009).
42. Alvarez-Garcia, D. & Barril, X. Molecular simulations with solvent competition quantify water displaceability and provide accurate interaction maps of protein binding sites. *J. Med. Chem.* **57**, 8530–8539 (2014).
43. Ung, P. M. U., Ghanakota, P., Graham, S. E., Lexa, K. W. & Carlson, H. A. Identifying binding hot spots on protein surfaces by mixed-solvent molecular dynamics: HIV-1 protease as a test case. *Biopolymers* **105**, 21–34 (2016).
44. Ghanakota, P. & Carlson, H. A. Driving Structure-Based Drug Discovery through Cosolvent Molecular Dynamics. *J. Med. Chem.* **59**, 10383–10399 (2016).
45. Defelipe, L. A. et al. Solvents to fragments to drugs: MD applications in drug design. *Molecules* **23**, 1–14 (2018).
46. Martinez-Rosell, G., Harvey, M. J. & De Fabritiis, G. Molecular-Simulation-Driven Fragment Screening for the Discovery of New CXCL12 Inhibitors. *J. Chem. Inf. Model.* **58**, 683–691 (2018).
47. Thi, L. et al. Structural Details of Ufd1 Binding to p97 and Their Functional Implications in ER-Associated Degradation. *PLoS ONE* **11**, e0163394 (2016).
48. Lim, J. J. et al. Structural insights into the interaction of human p97 N-terminal domain and SHP motif in Derlin-1 rhomboid pseudoprotease. *FEBS Lett.* **590**, 4402–4413 (2016).
49. Tang, W. K. & Xia, D. Role of the D1-D2 Linker of Human VCP/p97 in the Asymmetry and ATPase Activity of the D1-domain. *Sci. Rep.* **6**, 20037 (2016).
50. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. Sect. A Found. Crystallogr.* **64**, 112–122 (2008).
51. Adams, P. D. et al. PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **66**, 213–221 (2010).
52. Schultes, S. et al. Ligand efficiency as a guide in fragment hit selection and optimization. *Drug Discov. Today Technol.* **7**, 157–162 (2010).
53. Hänzelmann, P. & Schindelin, H. Characterization of an Additional Binding Surface on the p97 N-Terminal Domain Involved in Bipartite Cofactor Interactions. *Structure* **24**, 140–147 (2016).
54. Hänzelmann, P., Buchberger, A. & Schindelin, H. Hierarchical binding of cofactors to the AAA ATPase p97. *Structure* **19**, 833–843 (2011).
55. Van Den Ent, F. & Löwe, J. RF cloning: A restriction-free method for inserting target genes into plasmids. *J. Biochem. Biophys. Methods* **67**, 67–74 (2006).
56. Marley, J., Lu, M. & Bracken, C. A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR* **20**, 71–75 (2001).
57. Hänzelmann, P. & Schindelin, H. 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 (2011).
58. Brenk, R. et al. Lessons learnt from assembling screening libraries for drug discovery for neglected diseases. *ChemMedChem* **3**, 435–444 (2008).
59. Jorgensen, W. L. & Tirado-Rives, J. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proc. Natl Acad. Sci.* **102**, 6665–6670 (2005).
60. Dodda, L. S., Vilseck, J. Z., Tirado-Rives, J. & Jorgensen, W. L. 1.14\*CM1A-LBCC: Localized Bond-Charge Corrected CM1A Charges for Condensed-Phase Simulations. *J. Phys. Chem. B.* **121**, 3864–3870 (2017).
61. Dodda, L. S., Cabeza De Vaca, I., Tirado-Rives, J. & Jorgensen, W. L. LigParGen web server: an automatic OPLS-AA parameter generator for organic ligands. *Nucl. Acids Res.* **45**, 331–336 (2017).
62. Abraham, M. J. et al. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
63. Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **118**, 11225–11236 (1996).
64. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
65. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
66. Kabsch, W. XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **66**, 125–132 (2010).
67. Tickle, I. J. et al. STARANISO (<http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi>). (Cambridge, United Kingdom: Global Phasing Ltd.)
68. Thorn, A. & Sheldrick, G. M. ANODE: Anomalous and heavy-atom density calculation. *J. Appl. Crystallogr.* **44**, 1285–1287 (2011).
69. Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **67**, 235–242 (2011).
70. Painter, J. & Merritt, E. A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **62**, 439–450 (2006).
71. Walpole, S., Monaco, S., Nepravishta, R. & Angulo, J. STD-NMR as a Technique for Ligand Screening and Structural Studies. *Methods Enzymol.* **615**, 423–451 (2019).
72. Jayalakshmi, V. & Krishna, N. R. Complete relaxation and conformational exchange matrix (CORCEMA) analysis of intermolecular saturation transfer effects in reversibly forming ligand-receptor complexes. *J. Magn. Reson.* **155**, 106–118 (2002).
73. Neal, S., Nip, A. M., Zhang, H. & Wishart, D. S. Rapid and accurate calculation of protein 1H, 13C and 15N chemical shifts. *J. Biomol. NMR* **26**, 215–240 (2003).
74. Wishart, D. S. et al. 1H, 13C and 15N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* **6**, 135–140 (1995).
75. Isaacson, R. L. et al. Detailed structural insights into the p97-Npl4-Ufd1 interface. *J. Biol. Chem.* **282**, 21361–21369 (2007).
76. Stapf, C., Cartwright, E., Bycroft, M., Hofmann, K. & Buchberger, A. 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 (2011).
77. Skinner, S. P. et al. CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J. Biomol. NMR* **66**, 111–124 (2016).
78. Mureddu, L. & Vuister, G. W. Simple high-resolution NMR spectroscopy as a tool in molecular biology. *FEBS J.* **286**, 2035–2042 (2019).

## Acknowledgements

This work was supported by the Deutsche Forschungsgemeinschaft through GRK2243 (H.S. and C.S.) and EXC 2051 - Project ID 390713860 (U.A.H.) and by the Rudolf Virchow Center for Integrative and Translational Bioimaging (P.H. and H.S.). The fragment screen was performed at the facility for Biophysics, Structural Biology and Screening at the University of Bergen (BiSS), which received funding from the Research Council of Norway (RCN) through the NOR-OPENSCREEN consortium (grant number 245922). We thank Khanh Dim Dao and Emil Hausvik (University of Bergen) for assistance with the screening, Curd Schollmayer (Institute of Pharmacy and Food Chemistry) for assistance with NMR measurements, Monika Kuhn (Rudolf Virchow Center) for assistance in protein expression and purification, Alexander Buchberger (University of Würzburg) for helpful discussions and the beamline scientists at the ESRF. We thank the Centre of Biomolecular Magnetic Resonance (BMRZ) at the Goethe University Frankfurt funded by the state of Hesse for support. This publication was supported by the Open Access Publication Fund of the University of Würzburg.

## Author contributions

C.S., H.S. and P.H. initiated and supervised the project. Se.B., P.H. and St.B. performed cloning, protein expression and purification. R.B. initiated fragment screening. Se.B. performed BLI fragment screening and data analysis, CSP analysis and crystallization. Se.B. and J.K. performed MD simulations and analysis. Se.B. and M.Z. performed STD-NMR experiments and analysis. Se.B. and H.S. performed data collection and X-ray data analysis. St.B., U.A.H. and C.W. performed HSQC-NMR measurements. Se.B. drafted the manuscript and all authors contributed additions and corrections. All authors have read and approved the final manuscript.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42004-022-00782-5>.

**Correspondence** and requests for materials should be addressed to Hermann Schindelin or Christoph Sotriffer.

**Peer review information** *Communications Chemistry* thanks the anonymous reviewers for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022