The General Definition of the p97/Valosin-containing Protein (VCP)-interacting Motif (VIM) Delineates a New Family of p97 Cofactors

Received for publication, June 21, 2011, and in revised form, August 30, 2011. Published, JBC Papers in Press, September 6, 2011, DOI 10.1074/jbc.M111.274472

Christopher Stapf‡, Edward Cartwright§, Mark Bycroft‡, Kay Hofmann‡, and Alexander Buchberger‡

From the ‡Department of Biochemistry, Biocenter, University of Würzburg, Am Hubland, 97074 Würzburg, Germany, the §Medical Research Council Centre for Protein Engineering, Hills Road, Cambridge, CB2 0QH, United Kingdom, and †Miltenyi Biotec GmbH, Friedrich-Ebert-Strasse 68, 51429 Bergisch-Gladbach, Germany

Cellular functions of the essential, ubiquitin-selective AAA ATPase p97/valosin-containing protein (VCP) are controlled by regulatory cofactors determining substrate specificity and fate. Most cofactors bind p97 through a ubiquitin regulatory X (UBX) or UBX-like domain or linear sequence motifs, including the hitherto ill defined p97/VCP-interacting motif (VIM). Here, we present the new, minimal consensus sequence RXAAXAXAAXXH as a general definition of the VIM that unites a novel family of known and putative p97 cofactors, among them UBXD1 and ZNF744/ANKZF1. We demonstrate that this minimal VIM consensus sequence is necessary and sufficient for p97 binding. Using NMR chemical shift mapping, we identified several residues of the p97 N-terminal domain (N domain) that are critical for VIM binding. Importantly, we show that cellular stress resistance conferred by the yeast VIM-containing cofactor Vms1 depends on the physical interaction between its VIM and the critical N domain residues of the yeast p97 homolog, Cdc48. Thus, the VIM-N domain interaction characterized in this study is required for the physiological function of Vms1 and most likely other members of the newly defined VIM family of cofactors.

The chaperone-related ATPase p97 (also known as valosin-containing protein (VCP) and Cdc48) participates in a variety of cellular processes involving protein ubiquitylation, including protein quality control, cell cycle progression, autophagy, homotypic membrane fusion, transcriptional activation, DNA repair, and apoptosis (1–5). The common function of p97 underlying these processes is believed to be its ability to extract ubiquitylated substrate proteins from stable associations with other proteins, membranes, or chromatin.

p97 is a hexameric, ring-shaped member of the AAA (ATPases associated with various cellular activities) ATPase superfamily (6). The p97 protomer consists of an N-terminal N domain, the two ATPase domains D1 and D2, and an unstructured C-terminal tail (7). The diverse cellular functions of p97 are tightly regulated by a large number of cofactors. These control important aspects of substrate turnover, including p97 subcellular localization, substrate specificity, and ubiquitin chain plasticity (2, 8). Consequently, altered cofactor interactions with p97 have been linked to the disease inclusion body myopathy, Paget disease of the bone, and frontotemporal dementia (IBMPFD) (9, 10), a fatal proteinopathy caused by mutations in the VCPl gene encoding p97 (11). Most cofactors interact with p97 through a limited set of binding modules. The largest group of cofactors binds to the p97 N domain by virtue of a ubiquitin regulatory X (UBX) domain or structurally related ubiquitin- or UBX-like domains (8). Cofactors containing PUB or PUL domains bind to the C-terminal tail of p97 (2, 12). In addition to these structural domains, several short, linear binding motifs mediating cofactor binding to the N domain have been identified, including binding site 1 (BS1; also known as the SHP box) (13–15), the VCP-binding motif (VBM) (16), and the p97/VCP-interacting motif (VIM) (17).

The VIM was originally identified in the mammalian ERAD ubiquitin ligase gp78 (also known as AMFR and RNF45), where it serves to recruit p97 to the endoplasmic reticulum membrane in order to assist in the retrotranslocation of gp78 substrates like CD3δ and the Z variant of α-1-antitrypsin (17, 18). Subsequently, the small VCP-inhibiting protein (SVIP), a membrane-anchored negative regulator of ERAD, was shown to interact with p97 through a VIM as well (19). This led to the definition of a VIM consensus sequence of 30 residues in length based solely on the p97 interaction sites found in gp78 and SVIP (12, 17).

Here, we present a minimal, general VIM consensus sequence based on unbiased bioinformatic analyses, which is necessary and sufficient for p97 binding. The redefined VIM consensus guided the identification of a number of additional VIM-containing proteins, including previously known as well as novel bona fide p97 cofactors. Importantly, we mapped the VIM binding site on the p97 N domain and demonstrate that impairing the VIM-p97 interaction by mutation of either binding partner causes similar defects in vivo.
**General Definition of the p97/VCP-interacting Motif (VIM)**

**EXPERIMENTAL PROCEDURES**

**Bioinformatics**—Multiple sequence alignments were calculated with the L-INS-I algorithm of the MAFFT software (20). Construction of generalized sequence profiles and profile-to-sequence comparison was performed by the pftools software (21). Generation of hidden Markov models (HMMs) from sequence alignments and HMM-to-HMM comparison were performed by the HHPRED software package (22). Disordered regions in proteins were predicted by GlobPlot (23).

**Plasmids**—Plasmids for the bacterial expression of N-terminal hexahistidine fusion proteins of human p97, p97ΔN, and p97ΔC (24); for the bacterial expression of N-terminal GST fusion proteins of human UBXD1, UBXD115-NTA, and UBXD1(1-150) and human p47 (25); for the mammalian expression of N-terminally FLAG epitope-tagged UBXD1 and UBXD115-NTA (24); for the expression of p97 as a Gal4 activation domain fusion in yeast (24); and for the expression of Saccharomyces cerevisiae CDC48 in yeast (26) were described previously. Molecular cloning of the coding regions of human UBXD1 into pET28-His6-SUMO1 (27), of human ZNF744 (ANKZF1) into pGBDU (28) and pCMV-Tag2B (Stratagene), of yeast VMS1 (YDR049W) with and without C-terminal 3HA epitope tag into YCplac22 (29), and of the N domain (residues 1-213) of p97 into pHLT (30) was performed according to standard procedures. Details are available from the authors upon request. Site-directed mutagenesis of ZNF744 (R655A, A581/A59L, and R62A/L63A), p97 (D35A/N36A/S37A/V38A, S37A/V38A, D55A/T56A, and D150A), Vms1 (R617A, A623L/A624L, and R627A), and Cdc48 (D45A/N46A/S47A/V48A, S47A/V48A, D65A/T66A, and D160A) was performed using the QuikChange II XL kit (Stratagene) according to the manufacturer’s instructions and verified by sequencing of the entire coding regions.

**Yeast Strains and Media**—Reporter strain PJ69-4a (28) was used for yeast two-hybrid interaction assays. All other yeast experiments were performed in the DF5a strain background (31). The Δvms1::hphNT1 mutant was generated by disruption of the coding sequence using standard procedures (32). CDC48 shuffle strains were generated exactly as described (26). Yeast cells were grown in standard YPD and in synthetic complete media lacking the appropriate nutrients.

**Protein Expression and Purification**—Bacterial expression and affinity purification of His6-p97 (24, 25) and of GST fusions of UBXD1 (24) and p47 (25) were performed exactly as described. His6-SUMO1-UBXD1 was expressed in Escherichia coli BL21(DE3) pRIL (Novagen) and purified by Ni2+NTA affinity chromatography using standard protocols. The His6 SUMO1 moiety was removed by incubation with recombinant, purified His6-SenP2(364-489) protease (27) during overnight dialysis at 4 °C against 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM DTT. After dialysis, the reaction mixture was reapplied to a Ni2+NTA affinity column, and untagged UBXD1 was recovered as flow-through. Isotope-labeled p97 N domain was prepared by bacterial expression in K-MOPS minimal medium containing 15NH4Cl. Cells were lysed by sonication, and the hexahistidine fusion protein was purified by Ni2+-NTA affinity chromatography using standard protocols. The lipoyl domain fusion tag was removed by TEV protease digestion followed by a second Ni2+NTA affinity chromatography. The final purification step was size exclusion chromatography using a HiLoad 26/60 Superdex 75 column (GE Healthcare). The purified, isotope-labeled p97 N domain was concentrated to 50–75 μM in 25 mM HEPES-NaOH, pH 7.5, 125 mM NaCl, 5 mM DTT, 0.01% NaN3, 5% 3H2O.

**Binding Assays**—In vitro pull-down assays using immobilized biotinylated peptide (Biotin-GGSDREKRALAAERRLAAQ-COOH; PANATecs GmbH, Tübingen, Germany) or GST fusion proteins were performed as described (26), using 10 μl of beads, 0.76 nmol of GST or GST fusion proteins, and 0.2 nmol of His6-p97. For the in vitro co-immunoprecipitation of p97 and UBXD1, 0.2 nmol of p97 were incubated overnight at 4 °C in immunoprecipitation (IP) buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 2 mM MgCl2, 0.1% Nonidet P-40, 10% glycerol) with an affinity-purified, polyclonal anti-Cdc48 antibody (26). p97-antibody complexes were immobilized by incubation with 10 μl of Protein A-Sepharose (GE Healthcare), followed by two wash steps with IP buffer and one wash step with 1× TBST (25 mM Tris-HCl pH 7.5, 137 mM NaCl, 2.6 mM KCl, 0.1% Triton X-100). UBXD1 (0.2 nmol) was incubated with immobilized p97 for 1 h at 4 °C in 1× TBST, followed by four wash steps in the same buffer and Western blot analysis of the p97-bound fraction using a mouse monoclonal antibody against UBXD1 (33). Yeast two-hybrid assays (24) and immunoprecipitation experiments using yeast (26) or HEK293T (24) cell lysates were performed exactly as described.

**NMR Spectroscopy**—1H-15N HSQC spectra of the isotope-labeled p97 N domain were recorded on a Bruker DRX-600 spectrometer at 298 K (25 °C). Sample and acquisition conditions were identical to those in a previous report (34), allowing for the use of the published cross-peak assignments. Peptide binding to the p97 N domain was detected by monitoring chemical shift changes upon titration of peptide to final molar ratios (N domain/peptide) of 1:0.26, 1:0.8, 1:1.8, 1:3.6, 1:7, and 1:12. Chemical shift changes were quantified as root mean squared weighted chemical shift perturbation, ((δ1H)2 + (δ15N/5))1/2.

**RESULTS**

**A General Consensus Sequence for the VIM**—In order to assess the role of short sequence motifs in the specific recognition of p97, we performed a bioinformatics screen of candidate motifs. To that end, multiple-sequence alignments of all known or suspected p97 interactors and their orthologs were constructed. These alignments were visually scanned for short patches of local residue conservation, embedded within unstructured or poorly conserved regions. Next, all candidate motifs were used for the construction of short HMMs, which were then used for a comparison against each other and against a collection of precalculated HMMs derived from genome-wide sets of multiple-ortholog alignments.4 Several motifs were found to be present in multiple p97-binding proteins, where they often overlap with known BS1, VBM, and VIM motifs. Of

4 K. Hofmann, unpublished observations.
**General Definition of the p97/VCP-interacting Motif (VIM)**

**FIGURE 1. A minimal VIM consensus sequence.** A multiple-sequence alignment shows the VIMs of selected homologs of the eight VIM-containing proteins indicated at the left. (See supplemental Fig. S1 for a comprehensive alignment of VIMs.) The most highly conserved residues are boxed in red and define the newly identified, minimal consensus sequence RX_xAAX_xR shown at the bottom. X denotes positions without significant conservation or with preference for unusual amino acid groupings of unclear significance (e.g. E/R/L/I at the first position of X). All residues must be compatible with a helical structure of the motif. Residues conserved within but not between the eight subgroups are boxed in black or gray, depending on the degree of conservation. Note that the first arginine residue of the consensus sequence is not boxed or gray.

| A5FR2 | HUMAN | VTRpvGAAAGAALER | Q289CO | XENF | TTGRRKGGAGAALER |
| AMVR2 | HUMAN | VTRpvGAAAGAALER | Q297CO | XENT | TTGRRKGGAGAALER |
| AMVR2 | HUMAN | VTRpvGAAAGAALER | Q297CO | XENT | TTGRRKGGAGAALER |
| Q5575X | CRYN | DNDX GGGG |
| Q5575X | CRYN | DNDX GGGG |
| SVIP | HUMAN | LEpALKRAEAALER |
| SVIP | HUMAN | LEpALKRAEAALER |
| SVIP | HUMAN | LEpALKRAEAALER |
| Q5575X | CRYN | DNDX GGGG |
| Q5575X | CRYN | DNDX GGGG |
| SELS | HUMAN | VAGRAAAAX |
| SELS | HUMAN | VAGRAAAAX |
| SELS | HUMAN | VAGRAAAAX |
| Q5575X | CRYN | DNDX GGGG |
| Q5575X | CRYN | DNDX GGGG |
| ANKZF1 (ZNF744) | HUMAN | ALXpALX |
| ANKZF1 (ZNF744) | HUMAN | ALXpALX |
| ANKZF1 (ZNF744) | HUMAN | ALXpALX |
| Q5575X | CRYN | DNDX GGGG |
| Q5575X | CRYN | DNDX GGGG |
| YGL240c | HUMAN | RAKRAEAALER |
| YGL240c | HUMAN | RAKRAEAALER |
| YGL240c | HUMAN | RAKRAEAALER |
| Q5575X | CRYN | DNDX GGGG |
| Q5575X | CRYN | DNDX GGGG |
| Wss1 | YEAST | KSARPA |
| Wss1 | YEAST | KSARPA |
| Wss1 | YEAST | KSARPA |
| Q5575X | CRYN | DNDX GGGG |
| Q5575X | CRYN | DNDX GGGG |
| UBXD1 | HUMAN | APXARX |
| UBXD1 | HUMAN | APXARX |
| UBXD1 | HUMAN | APXARX |
| Q5575X | CRYN | DNDX GGGG |
| Q5575X | CRYN | DNDX GGGG |
| ANKZF1 (ZNF744) | HUMAN | VAGRAAAAX |
| ANKZF1 (ZNF744) | HUMAN | VAGRAAAAX |
| ANKZF1 (ZNF744) | HUMAN | VAGRAAAAX |
| Q5575X | CRYN | DNDX GGGG |
| Q5575X | CRYN | DNDX GGGG |
| ANKZF1 (ZNF744) | HUMAN | VAGRAAAAX |
| ANKZF1 (ZNF744) | HUMAN | VAGRAAAAX |
| ANKZF1 (ZNF744) | HUMAN | VAGRAAAAX |
| Q5575X | CRYN | DNDX GGGG |
| Q5575X | CRYN | DNDX GGGG |

**Note.** The residues contributing most to the bioinformatically defined similarity score were not always those highlighted in the previously published consensus sequences (data not shown). The most prevalent of the bioinformatically defined motifs is characterized by two flanking arginine and two central alanine residues in an invariant spacing (RX_xAAX_xR; Fig. 1 and supplemental Fig. S1). Besides other proteins, this motif is found in the p97 cofactors gp78 and SVIP (Fig. 1), where it overlaps with the previously defined VIM (12, 17). We therefore propose the minimal consensus sequence RX_xAAX_xR as the new, general definition of the VIM.

Importantly, the revised RX_xAAX_xR consensus allowed the identification of several additional proteins as true VIM-containing cofactors (Fig. 1). These include the ERAD regulator VIMP (also known as SELS) (35); the zinc finger-containing protein of unknown function ZNF744 (also known as ANKZF1), whose yeast homolog Vms1 (Ydr049w) was recently found to be involved in mitochondrial stress response and ERAD (36, 37); the fungal SUMO-dependent isopeptidase Wss1 implicated in proteasomal degradation of SUMOylated substrates (38, 39); the uncharacterized fungal Ygl108c protein; the known p97 cofactor UBXD1 (also known as UBXN6) (24, 33, 40); and the AN1-type zinc finger protein ZFAND2b (also known as AIRAPL), a proteasomal regulator (41) that had not been linked to p97 before. Importantly, the latter two proteins lack the otherwise highly conserved first arginine residue of the RX_xAAX_xR consensus, suggesting that it may be dispensable for p97 binding at least in certain contexts. Taken together, our bioinformatic analysis provides a general definition of the previously ill defined VIM and unites a number of hitherto unlinked proteins in a new family of known and putative p97 cofactors.

**Figure 2. The New VIM Consensus Sequence Is Necessary and Sufficient for p97 Binding**. In order to provide experimental evidence for the direct binding of the redefined VIM to p97, we performed in vitro pull-down assays using purified, recombinant p97 and a biotinylated, minimal VIM peptide. We chose a peptide derived from ZNF744 because its VIM is highly conserved from yeast to humans (Fig. 1). The immobilized VIM peptide bound full-length p97 and the truncated variant p97ΔC lacking the six C-terminal residues but not the p97ΔN variant lacking the entire N domain (Fig. 2A), showing that the new VIM is indeed sufficient for p97 binding. To analyze the importance of the most highly conserved residues within the VIM for p97 binding, we monitored binding of full-length ZNF744 to p97 in a yeast two-hybrid assay (Fig. 2B). Whereas wild-type ZNF744 interacted robustly with p97, mutating the VIM at the flanking residues Arg617 and Arg627-Leu628 to alanine or at the two central residues Ala622-Ala623 to leucine completely abrogated the two-hybrid interaction with p97. Similarly, the co-immunoprecipitation of endogenous p97 with ZNF744 ectopically expressed in human HEK293T cells was reduced to background levels by the same mutations (Fig. 2C). Together, these results demonstrate that the ZNF744 residues defining the minimal VIM consensus are required for p97 binding.

**Additional Remarks**

**A Cryptic VIM Is the Second p97 Binding Site of UBXD1**—UBXD1 is an atypical UBX protein inasmuch as its UBX domain lacks key residues critical for p97 binding (24, 33). Instead, UBXD1 interacts with p97 through two independent sites, a central PUB domain binding to the C terminus of p97 and the truncated variant p97ΔC lacking the six C-terminal residues but not the p97ΔN variant lacking the entire N domain (Fig. 2A), showing that the new VIM is indeed sufficient for p97 binding. To analyze the importance of the most highly conserved residues within the VIM for p97 binding, we monitored binding of full-length ZNF744 to p97 in a yeast two-hybrid assay (Fig. 2B). Whereas wild-type ZNF744 interacted robustly with p97, mutating the VIM at the flanking residues Arg617 and Arg627-Leu628 to alanine or at the two central residues Ala622-Ala623 to leucine completely abrogated the two-hybrid interaction with p97. Similarly, the co-immunoprecipitation of endogenous p97 with ZNF744 ectopically expressed in human HEK293T cells was reduced to background levels by the same mutations (Fig. 2C). Together, these results demonstrate that the ZNF744 residues defining the minimal VIM consensus are required for p97 binding.
residues other than the conserved consensus residues are not required for VIM function. p97 binding assays using GST fusions of full-length UBXD1 yielded highly consistent results (Fig. 3B). Whereas mutations within the VIM alone had no effect on p97 binding of UBXD1 due to the presence of its intact PUB domain, the additional mutation of five critical residues within the PUB domain (UBXD1^{5×PUB}) (24, 25) resulted in the loss of p97 binding of the AA and RL but not of the QM variant of UBXD1^{5×PUB}. Moreover, the addition of an excess of the ZNF744-derived VIM peptide abolished p97 binding to UBXD1(1–150) and full-length UBXD1^{5×PUB}, both containing their wild-type VIM (data not shown), further supporting the importance of the VIM of UBXD1 for p97 binding.

In order to confirm the observed binding behavior also for untagged UBXD1, we immunoprecipitated recombinant p97 in vitro and analyzed the binding of recombinant, full-length UBXD1 to the immobilized p97 (Fig. 3C). Consistent with the GSH pull-down assays, wild-type UBXD1, UBXD1^{5×PUB}, and the UBXD1-RL variant mutated in the VIM showed comparable binding to full-length p97, whereas UBXD1^{5×PUB}-RL did not bind above background (Fig. 3C, top), indicating that the PUB domain and the VIM of UBXD1 are each sufficient for p97 binding. This conclusion was further supported by the finding that wild-type UBXD1 and UBXD1^{5×PUB} bound to a C-terminally truncated p97 variant unable to interact with the PUB domain (p97^{C}) in a VIM-dependent manner (Fig. 3C, center left) and that wild-type UBXD1 and UBXD1-RL bound to an N-terminally truncated p97 variant unable to interact with the VIM (p97^{ΔN}) in a PUB-domain-dependent manner (Fig. 3C, center right).

Finally, the in vivo relevance of the VIM in UBXD1 for p97 binding was investigated in immunoprecipitation experiments using HEK293T cell lysates (Fig. 3D). Again, mutation of the critical consensus VIM residues, but not of QM, resulted in loss of p97 binding for the UBXD1^{5×PUB} variant but not for the otherwise wild-type UBXD1. Taken together, these data demonstrate that a minimal VIM is indeed the elusive second p97 binding site of UBXD1 and that it is the only other binding site besides the PUB domain.

Mapping of the VIM Binding Site on p97—After establishing that the redefined, minimal VIM consensus sequence is necessary and sufficient for p97 binding, we sought to identify regions on the p97 N domain critical for VIM binding. To that end, we recorded $^1$H-$^1$N HSQC NMR spectra of uniformly $^{15}$N-labeled N domain (residues 1–213), an approach that had previously proven useful for the mapping of N domain interactions (34). The obtained $^1$H-$^1$N HSQC spectrum (Fig. 4A) was superimposable to the previously published spectrum of the N domain (34). Titration of the unlabeled ZNF744-derived VIM peptide to the isolated N domain resulted in saturable, distinct shifts of several cross-peaks (see insets in Fig. 4A), whereas the majority of peaks were not at all or not significantly shifted (Fig. 4, A and B). These data are indicative of the specific binding of the VIM peptide to the isolated N domain and, in fact, show for the first time that the N domain alone is sufficient for VIM binding, in contrast to previous reports (17, 42). When residues exhibiting strongly shifted cross-peaks were mapped onto the known three-dimensional structure of the N domain, most were found to cluster in two neighboring regions (Fig. 4C).

Next, we wanted to analyze the relevance of these regions for VIM binding. In order to identify candidate residues for site-directed mutagenesis of full-length p97, we reasoned that VIM-specific contact sites (i) should contain negatively charged and/or polar side chains given the critical importance of the flanking arginine residues for the VIM and (ii) should not extensively overlap with known binding sites of other N domain interactors to facilitate the interpretation of subsequent in vivo experiments. We therefore chose the following sets of residues for replacement with alanine. (i) Asp^{35}, Asn^{36}, Ser^{37}, Val^{38}, Ser^{77} and Val^{38} show the strongest chemical shift perturbation within a stretch of shifted residues (Fig. 4B) and face the cleft separating Nn and Nc. Although the neighboring residues Asp^{35} and Asn^{36} could not be unambiguously assigned in the NMR spectrum, they were included because they are negatively charged and polar, respectively, and face the central cleft as well. The corresponding p97 variants were termed p97^{35SV} and p97^{35SV}, respectively. (ii) We chose Gly^{54}, Asp^{55}, Thr^{56}, Gly^{57}, and Thr^{56} belong to the most strongly shifted cross-peaks (Fig. 4A).
and are also oriented toward the central cleft. Because substitution of glycine residues can be difficult to interpret, residue Asp was chosen based on its negative charge and proximity to Thr. The corresponding p97 variant is p97DT, (iii) Asp, a negatively charged residue at the bottom of the subdomain cleft, was mutated to generate p97DT. We disregarded the interconnecting loop Asp–Lys and some neighboring residues, mainly because previous structural studies had identified this region as the major interaction site for other N domain binding modules, including the UBX domain, the UBX-like domain of Npl4, and the BS1 motif of Ufd1.

The recombinant p97 mutant proteins were analyzed in a peptide pull-down assay for their ability to bind to the ZNF744-derived VIM peptide (Fig. 5A). Intriguingly, all four mutant p97 proteins exhibited strongly reduced VIM binding, indicating that the NMR chemical shift mapping indeed had identified regions of the N domain critical for the interaction with the VIM. The interaction with the VIM peptide was virtually abolished for p97SV and p97DNSV, whereas p97DT and particularly p97D150A showed some residual binding. We also analyzed the p97 mutant proteins in pull-down experiments using immobilized GST-UBXD1(1–150) with its intact VIM (Fig. 5B, top). Although p97DT and p97DNSV showed no reduction and moderate reduction in UBXD1(1–150) binding, respectively, the interaction was strongly reduced for p97SV and almost completely lost for p97DNSV. In this assay, p97DNSV appeared to be significantly more defective than p97SV in VIM binding, suggesting that residues Asp and/or Asn indeed contribute to the interaction. Importantly, all four p97 mutant proteins were proficient in binding to the major N domain-interacting cofactor p47 (Fig. 5B, bottom). Taken together, these results show that the p97 mutant proteins that were engineered on the basis of the NMR chemical shift mapping are specifically impaired in VIM binding.

The VIM of Yeast Vms1 Is Important for Cellular Stress Resistance—In order to test the functional relevance of VIM-mediated cofactor binding in vivo, we chose the yeast S. cerevisiae as a model. We first confirmed that the yeast homolog of ZNF744, Vms1 (Ydr049w), interacts with the yeast p97 homolog, Cdc48, in a VIM-dependent manner (Fig. 6A) (37). Coimmunoprecipitation of Cdc48 with Vms1 was strongly reduced to background levels upon mutation of the central alanine residues 623/624 or of the C-terminal arginine residue 627, whereas mutation of the N-terminal arginine residue 617 reduced but did not abolish Cdc48 binding. Vms1 was recently reported to be involved in mitochondrial stress resistance and ERAD (36, 37). Consequently, yeast mutants lacking Vms1
exhibit a number of stress phenotypes, among them hypersensitivity against the translation inhibitor cycloheximide (Fig. 6B) (45). Importantly, the differential reduction in Cdc48 binding observed for the VIM mutants of \textit{VMS1} correlated well with their sensitivity toward cycloheximide. Whereas the \textit{vms1R627A} and \textit{vms1AA} mutants were almost as sensitive as the \textit{vms1} knock-out strain, \textit{vms1R617A} exhibited hardly any growth defect in the presence of cycloheximide. These results indicate that the physiological function of Vms1 in cellular stress resistance critically depends on Cdc48 binding via the VIM, in agreement with a recent report (37).

A Yeast cdc48 Mutant Defective in VIM Binding Is Stress-sensitive—We next addressed the importance of the VIM binding region mapped on p97 for cellular stress resistance. To that end, we used shuffle strains deleted for the chromosomal Cdc48 locus and expressing untagged wild-type or mutant Cdc48 from a centromeric plasmid at endogenous levels (26). Strains expressing the \textit{cdc48SV}, \textit{cdc48DNSV}, and \textit{cdc48D160A} mutant alleles (corresponding to the respective p97 variants described above) were viable and did not exhibit any general growth defect. Immunoprecipitation experiments showed that Cdc48SV and Cdc48DNSV failed to stably interact with Vms1 \textit{in vivo}, whereas Cdc48D160A showed residual binding to Vms1 (Fig. 6C), in excellent agreement with the \textit{in vitro} analysis of the corresponding p97 variants. Importantly, the Cdc48SV, Cdc48DNSV, and Cdc48D160A mutant proteins were unaffected in their ability to interact with the Cdc48 cofactors Ubx2 and Ufd1 \textit{in vivo} (Fig. 6D), whereas Cdc48DT showed reduced binding to both cofactors (data not shown) and was therefore excluded from further analyses. Together, these results show that the DNSV turn facing the subdomain cleft of the Cdc48 N domain is required for the binding of the yeast VIM proteins Vms1 and Ygl108c (data not shown) but not of cofactors interacting through the UBX domain or the BS1 motif. Finally, we tested the cycloheximide sensitivity of the cdc48 shuffle strains (Fig. 6E). Intriguingly, the cdc48DNSV

FIGURE 4. NMR chemical shift mapping of the VIM binding site on p97. A, \textsuperscript{1}H-\textsuperscript{15}N HSQC NMR spectra of the p97 N domain. Assigned cross-peaks (34) are labeled with residue numbers. The black spectrum was recorded in the absence of peptide. Spectra recorded in the presence of increasing concentrations of a ZNF744-derived VIM peptide are overlaid in spectral colors as indicated. The insets show peptide-induced peak shifts for selected residues. B, VIM peptide-induced shifted cross-peaks. Root mean square weighted chemical shift perturbations at saturating peptide concentration were plotted against residue number. C, mapping to the N domain structure. Chemical shift perturbation values from B were mapped onto the three-dimensional structure of the N domain (highest values in red, lowest values above threshold in blue). Shown is a surface representation of one protomer from the hexameric N-D1 crystal structure (Protein Data Bank entry 1E32 (55)). The arrow marks the central cleft separating the Nn and Nc subdomains of the N domain. The enlargement shows the N domain in ribbon representation, with residues subject to chemical shift perturbation as sticks. Selected residues discussed here are labeled.
DISCUSSION

This study reports the general definition of the linear VIM p97 binding motif based on a new, minimal consensus sequence and establishes VIM-containing proteins as a diverse family of p97 cofactors consisting of at least eight members. We identified residues of the VIM and of the p97 N domain critical for the VIM-p97 interaction, and we generated for the first time p97/Cdc48 binding variants specifically defective in VIM binding. Importantly, using these variants, we were able to demonstrate that the VIM-Cdc48 interaction is required for the in vivo function of a VIM-containing cofactor involved in cellular stress resistance.

The VIM was originally based on sequence alignments of the N domain of p97 specifically abolish VIM binding. This study reports the general definition of the linear VIM p97 binding motif based on a new, minimal consensus sequence and establishes VIM-containing proteins as a diverse family of p97 cofactors consisting of at least eight members. We identified residues of the VIM and of the p97 N domain critical for the VIM-p97 interaction, and we generated for the first time p97/Cdc48 binding variants specifically defective in VIM binding. Importantly, using these variants, we were able to demonstrate that the VIM-Cdc48 interaction is required for the in vivo function of a VIM-containing cofactor involved in cellular stress resistance. The VIM was originally based on sequence alignments of only two different proteins, gp78 and SVIP (12, 17), and thus warranted a thorough bioinformatic analysis. The two major improvements of our de novo approach are the use of an extended database of p97 interactors and of HMM-HMM searches that are unbiased by previous motif knowledge. The power of the approach is evident from the identification of a number of previously overlooked VIMs based on the new RX₅₋₆X₅₋₆ consensus sequence (Fig. 1). Whereas VIMP (35) and UBXD1 (36) has been reported to possess binding sites for the N domain of p97, the identity of these sites as VIMs was revealed for the first time in this study. Similarly, ZNF744 and its yeast homolog Vms1 (Ydr049w) were predicted to be p97/Cdc48 interactors based on proteomics approaches without recognizing their VIM (46, 47). Our identification of a minimal VIM in these proteins was corroborated in a recent study on the function of Vms1 in mitochondrial stress response (37), which was published while this work was in progress. The SUMO-specific isopeptidase Wss1 has previously been shown to possess a BS1 p97/Cdc48 binding motif (13).

The additional presence of a VIM suggests that this interesting but poorly characterized cofactor interacts with Cdc48 through a bipartite N domain binding mechanism reminiscent of the cofactors p47 and Ufd1-Npl4 (14). In addition to the characterization of known p97 interactors as VIM-containing cofactors, our analysis also led to the identification of VIMs in the proteasomal regulator ZFAND2b (41) and in the uncharacterized fungal protein Ygl108c. This unexpected finding strongly suggests that both proteins are p97/Cdc48 cofactors as well and opens avenues toward their future characterization. Finally, the results of our bioinformatic approach clarified the identity of the VIM in gp78. Its classification as VIM had been put into question in a recent study, where it was speculated, based on the occurrence of several flanking arginine residues, to rather resemble a “bipartite VBM” (48). In contrast to this interpretation, the minimal consensus sequence RX₅₋₆X₅₋₆ firmly establishes the p97 binding site in gp78 as a prototypical VIM (Fig. 1) that is clearly distinct from the VBM consensus sequence as published (12, 16) or as redefined by a bioinformatic approach analogous to the one described above.4

Our NMR chemical shift analysis mapped the VIM binding site to the central cleft separating the Nn and Nc subdomains of the p97 N domain and to adjacent exposed areas of Nn and Nc, in particular the linker connecting them (Fig. 4C). Of these, the linker region has been shown to be important for interactions with other N domain binding modules, including the UBX domains of p47 (43) and FAF1 (44), the ubiquitin-/UBX-like domain of Npl4, and BS1 of Ufd1 (34). Moreover, the region around residue Gly25 (next to Asp55-Thr56) was also reported to be involved in binding of Npl4 (34) and the UBX domain of p47 (43), consistent with the general defects in cofactor binding observed for the corresponding Cdc48¹⁷⁵ variant in vivo (data not shown). In contrast, the DNSV turn is located deeper in the subdomain cleft, whereas residue Asp¹⁵⁰ (Asp¹⁶⁰ in Cdc48) is distant from the aforementioned residues, strongly suggesting that VIM proteins interact with the N domain through a binding mode distinctly different from other cofactors. Consistent with this interpretation, mutations in the DNSV turn and residue Asp¹⁶⁰ of Cdc48 resulted in specific defects in the binding of Vms1 but not Ubx2 and Ufd1 (Fig. 6, C and D). Of note, the redefined VIM has a strong α-helical secondary structure prediction, and all four residues defining the RX₅₋₆X₅₋₆ consensus...
sus cluster to one side of the helix in a helical wheel projection (data not shown). Considering the importance of residues lining the central cleft for VIM binding, it is tempting to hypothesize that the VIM forms an α-helix that fits snugly into the subdomain cleft and that the invariant alanine residues of the VIM consensus are an adaptation to spatial constrictions within the narrow cleft. This prediction is indeed verified by the crystal structure of the gp78-derived VIM in complex with the p97 N domain (see the accompanying paper by Hänzelmann and Schindelin (56)). In this structure, the helical VIM binds to the central cleft for VIM binding, it is tempting to hypothesize that the VIM forms an α-helix that fits snugly into the subdomain cleft and that the invariant alanine residues of the VIM consensus are an adaptation to spatial constrictions within the narrow cleft. This prediction is indeed verified by the crystal structure of the gp78-derived VIM in complex with the p97 N domain (see the accompanying paper by Hänzelmann and Schindelin (56)).

The RX$_5$AAX$_4$R consensus sequence is highly conserved throughout the VIM family and defines the minimal requirements for p97 binding. Interestingly, however, the VIMs of UBBD1 and ZFAND2b lacking the N-terminal arginine residue are still proficient in p97 binding (Fig. 3) (data not shown), suggesting that the N-terminal arginine residue of the consensus sequence contributes less than the C-terminal arginine residue to p97 binding. This possibility is further supported by the finding that the R617A mutant of Vms1 was not completely impaired in p97 binding, in contrast to the R627A mutant (Fig. 6A). It is therefore likely that the loss of the first arginine residue in the VIMs of UBBD1, ZFAND2b, and the Vms1$^{R617A}$ mutant protein can be partially compensated by residues at or N-terminal of the position of the first arginine.

As discussed above, there is significant overlap of the VIM binding region on the p97 N domain with binding sites for UBX and ubiquitin-/UBX-like domains and the linear BS1 motif (see also the accompanying paper by Hänzelmann and Schindelin (56)). Consistent with such overlapping sites, binding of the (then unknown) VIMs of SVIP and UBBD1 has been shown to be mutually exclusive with p47 binding (24, 42). This competitive binding can be rationalized on the basis of the presence of six N domain binding sites per p47 trimer (three UBBD domains and three BS1 motifs), which are likely to occupy all available N domains of the p97 hexamer (49). Interestingly, however, p97 binding of the VIM proteins gp78, Vms1, and UBBD1 appears to be also mutually exclusive with Ufd1 but not Npl4 (17, 37, 50, 51). These findings are unexpected, given that one Ufd1-Npl4 heterodimer interacts with the p97 hexamer, presumably through binding of two neighboring N domains (52). This arrangement leaves space, and in fact has been shown to be a prerequisite, for the additional binding of UBBD domain proteins to free N domains within the same hexamer (40, 44, 53). It therefore remains unclear why gp78 and Vms1 cannot similarly bind via their VIMs to unoccupied N domains of the p97-Ufd1-Npl4 complex, in particular when considering that recombinant VIM proteins are proficient in p97 binding in the absence of other cofactors (24, 35, 42, 54). One possible explanation would be a competition of gp78 and Vms1 with Ufd1 for Npl4 binding (rather than for p97 binding), although there so far exists no evidence for a direct interaction of VIM proteins with Npl4. Thus, although this and the accompanying report (56) provide a framework for rationalizing the interaction of VIM cofactors with p97, elucidation of the molecular basis for the mutual influence of different cofactors on p97 binding will clearly require further structural and mechanistic studies.
