Induced polyubiquitination of proteins mediated by overexpression of a peptide: a novel tool for targeted protein degradation (TPD) research.

Girish Nallur (girish@vistarabio.com)
Vistara Bioscience

Research Article

Keywords: Targeted protein degradation (TPD), E3 ligase, PINTAC, peptide, targeting, polyubiquitination, MAPK1, MEK1/2, amyotropic lateral sclerosis, fronto-temporal dementia, myelodysplastic syndrome, multiple myeloma, lenalidomide, cereblon, CRL4, TRIM28, substrate, FUS, TD-43, LUC7L2, PARK7

Posted Date: April 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1547607/v1

Abstract

A novel strategy for targeting proteins with peptides which give rise to stable ubiquitination has been identified with proteomics. Several bioactive peptides gave rise to characteristic patterns of polyubiquitinated proteins, containing distinct sets of proteins in a cell culture model. Differential mass spectrometry of the ubiquitin-conjugated proteins with one such peptide namely, VB-009, identified proteins predominantly having roles in RNA metabolism, including some that are frequently mutated in diseases, such as, Amyotropic Lateral Sclerosis, Fronto-Temporal Dementia, and myelodysplastic syndrome. The applications of this serendipitous finding for targeting specific targets in TPD are discussed.

Introduction

Current targeted protein degradation (TPD) efforts such as PROTACS (Bekes et al, 2022) and Molecular Glues (Wertz and Wang, 2019) have received widespread attention, registering notable progress to date, but have frequently been faced with challenges in the clinic. The efforts thus far have identified areas of strength for PROTACs and glues, and also highlighted a number of development needs for ushering in the new era of TPD. Besides co-opting new E3 ligases, a deeper understanding of ubiquitin biology can help, such as, possible outcomes of ubiquitin-tagged substrates (Oh et al, 2018); target specificities and selectivity of biological E3 ligases towards their substrates, and post-translational modifications. Mechanistically, accessory proteins of the ubiquitin system, changes in intracellular localization of the proteins, or their expression in specific tissues (or at particular stages of development or disease) can have an impact as well on the outcomes of a TPD solution. Understanding the clinical translation inflexion points for the set of proteins suggested to be amenable for targeting with PROTAC technology (Schneider et al. 2021) can help find the best TPD solutions.

E3 ligases recognize their substrates via amino acid signatures in their primary sequence, termed motifs, which may be contiguous or not along the sequence of the substrate protein (Duan and Pagano, 2021). Finding these motifs on a much larger scale than currently possible can help us understand the functional complexity and/or the degradation network of E3 ligases, thus allowing mapping the substrates into distinct pathways, characterizing mutations which may affect E3 ligase-target interactions, or detecting how different E3 ligases collaborate in the cells. Besides, motif information can help greatly in expanding the TPD toolbox by providing a choice of targets paired with their corresponding natural biological E3 ligase, and afford the choice of degrading a cohort of proteins with a single E3 ligase.

The motifs can be potentially targeted with peptides for TPD research and developed further as a toolkit. Peptides offer many advantages for targeting proteins in TPD. They can help bring into the TPD realm any of the nearly 620 E3 ligases in humans. They can help extend TPD to any disease area by overcoming limitations on the poor “ligandability” of most E3 ligases. Other important applications include: assay development for identifying small molecules, engineering cells harboring the motifs, and

degrading heterologous proteins by tagging motifs to proteins of clinical interest, if they can function as a ‘degron’. So, the key discovery aspect is to expand the repertoire of targets engaged by E3 ligases.

Proteomic investigations identified several bioactive peptides with interesting properties towards the ubiquitome. Herein, this report describes using some of the peptides for targeting proteins in a cell culture model. Serendipitously, the peptides consistently gave rise to polyubiquitinated proteins which accumulated in peptide treated cells. This key feature may have important implications in the TPD arena. That the polyubiquitinated proteins produced in response to one bioactive peptide, VB-009, were a distinct set of putative substrates, and other UPS proteins, further emphasized their utility, and merited a deeper investigation. Many additional peptides showed a similar pattern, indicating that they interact with the ubiquitome in a similar fashion, the mechanism of which need to be elucidated with further investigation.

These findings show the utility of peptides for altering the function of proteins in cells via polyubiquitination, and may present an alternative strategy in TPD by disrupting the function of proteins in cells, rather than degradation. Although much work lies ahead with characterizing the peptides, understanding the underlying mechanisms of polyubiquitination, and developing them as useful tools for drug discovery research and therapeutics, they do hold the exciting prospect of a paradigm shift in our thinking about TPD.

Results

Identification and preliminary characterization of peptides which induce polyubiquitination of proteins when expressed in HEK293T cells: A series of peptides were designed which displayed the property of inducing polyubiquitination when overexpressed in HEK293T cells. As shown in Fig. 1a, peptides to target 3 distinct functional classes of proteins were used: 1) 4 members of a family of enzymes (VB-001 through VB-004); 2) MAPK1 (VB012) or TRIM28, an E3 ligase; or 3) 5 additional ligandable E3 ligases (VB-009-11 and VB-013-015).

Each peptide induced polyubiquitination of multiple proteins, producing a smear of proteins larger than 100 kd when overexpressed in HEK293T cells and detected by western blotting with an anti-ubiquitin antibody (Fig. 1b). As shown in Fig. 1c, the smear of ubiquitin-conjugated proteins was pronouncedly larger and brighter in peptide-treated cells over control (pCMV vector), indicating the presence of multiple ubiquitin conjugated proteins. Moreover, each peptide produced a characteristic pattern of bands within the smear suggesting that there may be some specificity as to which proteins resided in the polyubiquitinated fractions. The proteins targeted by VB-002 and VB-003 produced significantly larger smears, which likely stems from their co-chaperone function, and may involve extensive interactions with other proteins and/or E3 ligase substrates as well. These initial observations suggested that the peptides shown in Fig 1a consistently give rise to polyubiquitinated proteins in cells.

Polyubiquitinated proteins have been described in the literature, for instance, in neurodegenerative diseases (Johnston et al., 1998; Dehvari et al., 2012; Maynard et al., 2009), with mutant protein expression (Chadchankar et al., 2009), during oxidative stress or apoptosis (Reeg et al., 2015; Canu et al., 2000), and also during viral infections (Kobayashi et al., 2020). So, while a serendipitous observation, polyubiquitination was, in itself, not surprising.

That the polyubiquitinating property can be localized to such small regions as 25 – 30 amino acids and the number of peptides which consistently show this effect suggest an underlying mechanism, which need to be elucidated. While polyubiquitin-conjugated proteins do not normally accumulate in healthy cells, being rapidly transported to the proteasome for degradation, the peptide-induced polyubiquitinated proteins are stable. So, they likely also interfere with the process of delivering polyubiquitinated proteins to the proteasome for degradation, which might explain their accumulation in cells. For this reason, the peptides were termed, PINTACs, [Proteasome Inhibiting Targeting Construct].

So, these initial observations indicated a role for PINTACs in ubiquitin modification of substrates, and may be a strategy for harnessing E3 ligases by designing appropriate PINTACs, and developing assays for use in TPD drug discovery research. To realize those opportunities, it was important to determine the fidelity of ubiquitination induced by the peptides by associating a PINTAC with the distinct set of proteins polyubiquitinated, regardless of whether the ubiquitinated proteins are the natural substrates (or not) of the E3 ligase targeted by the PINTAC.

Targeted ubiquitination is highly selective towards an E3 ligase substrate: The specificity of a PINTAC to differentially target an E3 ligase and/or its substrate was next investigated. Two PINTACs were designed, one directed to MAPK1, (VB-012, 42 kd native size) and the other to the Sumo/E3 ligase, TRIM28 (VB-016), which likely targets MAPK1 for degradation (VB-015) in HEK293T cells (unpublished). As shown in Fig. 2a, the two PINTACs produced an identical pattern of ubiquitinated MAPK1, with discrete bands larger than 100 kd. Moreover, there was a near complete disappearance of the native MAPK1 protein indicating extensive ubiquitination. In the same samples, however, TRIM28 was not significantly altered nor ubiquitinated, and many other proteins related to the MAPK1 function were not conjugated with ubiquitin (Fig. 2c). Only PIK3CD showed a band of slightly lower mobility with VB-012, which is likely not a ubiquitinated form but may be an alternatively spliced isoform. In sharp contrast to VB-012, the VB015 sample showed a large smear of high molecular weight proteins when detected with an anti-ubiquitin antibody, such as in Fig. 1c, indicating the presence of multiple proteins in this fraction.

Taken together, these observations indicated that each of the PINTACs possesses the ability to induce polyubiquitination. It appears that a PINTAC directed to the substrate of an E3 ligase, in this case MAPK1, predominantly modifies the substrate, but a PINTAC directed to an E3 ligase modifies multiple proteins via ubiquitination.

Differential targeting of MEK1/2 by PINTACs directed to each of 4 E3 ligases: Next, an attempt was made to apply the serendipitous observation of polyubiquitination induced by PINTACs to purposefully target E3 ligases. Accordingly, PINTACs were designed to target three different E3 ligases (VB-009, -010, -011, -013 and -014). Polyubiquitinated proteins from the transfected HEK293T cells were observed as a smear of high molecular weight proteins as described above.

The samples were further characterized by western blotting with a panel of antibodies to detect specific proteins, and, by challenging the transfected cells with a panel of compounds to observe any alterations to the composition of polyubiquitinated proteins. As expected, most proteins remained unchanged with any of the 5 PINTAC treatments. However, an antibody to MAP2K1 (MEK1) and MAP2K2 (MEK2) readily detected MEK1 and/or MEK2 in the polyubiquitinated fractions with VB-009, -010, and -011 PINTAC treatments, as shown in Fig. 4a. MAPK1/Erk2 or B-Raf were not altered in these samples, nor many other proteins amongst the panel of antibodies used (not shown). To confirm the ubiquitination, total proteins from each of the samples, plus a control PINTAC which did not polyubiquitinate MEK1 or MEK2 (VB-014, pCMV) were fractionated on a column having a molecular weight cutoff of 100 kd, so as to exclude unconjugated MEK1 (43 kd) and MEK2 (44 kd). The partially purified 100 kd supernatants from the column fractionation were immunoprecipitated with the anti-ubiquitin antibody, and probed by western blotting with the MEK1/2 antibody. As shown in Fig. 4b, the antibody detected MEK1/2 bands of significantly higher molecular weights than the native proteins. MAPK1 was not detectable in the same supernatants, indicating that the column fractionation completely removed MAPK1 or unconjugated MEK1/2. Importantly, the process also showed more direct proof of ubiquitin conjugation of MEK1/2.

Lenalidomide alters the pattern of proteins polyubiquitinated with PINTACs: Lenalidomide, which is actively being applied in the clinic as a molecular glue, is particularly effective in multiple myeloma. The compound, and other thalidomide analogs, target cereblon (CRBN), a substrate receptor for a CRL4-type E3 ligase complex that was originally identified as a gene associated with mild intellectual disability (reviewed in: Ito et al 2021). Upon binding to lenalidomide CRBN mediates its pharmacological activities by engaging over a dozen neosubstrates and targeting them for degradation.

Several PINTAC-transfected HEK293T cells were treated with lenalidomide. As shown in Fig. 3, lenalidomide markedly decreased polyubiquitination by the MAPK1 PINTAC, VB-012, and also VB-015, but

enhanced the effect with VB-009 and VB-010, both of which are designed to target the same E3 ligase (not cereblon). The pronounced effects of lenalidomide on VB-009 and VB-010 is highly intriguing, and needs to be fully understood. Perhaps, competition for ubiquitin reagents or enzymes, alteration of protein networks necessary for CRL4 function, or even direct crosstalk between the E3 ligase targeted by the PINTACs and CRL4 E3 ligase may be the reasons for the observed effects.

Since PINTACs polyubiquitinate many substrates, it is highly advantageous if they can be applied for targeting CRL E3 ligase subunits, cullins, or other essential proteins necessary for their function. If so, it may become possible to apply the combined effects of the PINTACs and lenalidomide in the clinic for improving selectivity towards neotargets, finding new substrates, or overcoming drug resistance in multiple myeloma, such as, aberrant Wnt signaling (van Andel et al, 2019). Enhancement of VB-009 and VB-010 polyubiquitination by lenalidomide may suggest crosstalk between the E3 ligase targeted by them and CRL4, but the mechanism needs to be fully understood. Characterizing the polyubiquitinated fractions from these samples with or without lenalidomide treatment may identified the proteins responsible, and the pathways involved. Also, these studies need to be extended with additional thalidomide analogs to understand their specificities.

From the foregoing, it is clear that PINTACs can be used to deliberately interfere with the functioning of the UPS, bear some selectivity towards the proteins which are polyubiquitinated by each, and can be further modulated with compound treatments.

Mass spectrometric detection of proteins in the polyubiquitinated fraction: Identifying the proteins which reside in the polyubiquitinated fraction can help understand the mechanism and substrate specificity of the PINTAC, identify E3 ligase substrates, help design assays for drug discovery research, and screen for drug candidates with therapeutic applications. While some ubiquitinated proteins are expected in the 100 kd supernatants even from normal cells, the stimulation of ubiquitination provided by PINTAC treatments substantially increases their relative levels, and enables their detection over the cellular background. Toward this goal, the 100 kd supernatants from VB-011 PINTAC and two additional control PINTACs (directed to other E3 ligases) were subjected to differential mass spectrometry.

A total of 3,084 proteins were identified across the three samples, represented by 37,593 peptides from VB-009 sample, and 32,065 and 5,599 peptides, respectively, for two control samples. The mass spectrometry output from VB-009 sample was queried for proteins with a native molecular weight under 100 kd, represented by at least 20 peptides per protein, and, with the number of peptides being in 2-fold excess (or higher) when compared with at least one of the control samples (evidence of stimulation).

Such proteins were considered specifically polyubiquitinated in response to VB-009 expression and selected for further analysis. This identified 40 proteins (Table I) which were substantially enriched in the VB-009 sample. Consistent with the western blot data in Fig. 4, MEK1 and MEK2 were identified in the mass spectrometry data from VB-009 and VB-009, but not in the control sample which did not detect MEK1/2 by western blotting.

A number of proteins involved in the ubiquitination process were also enriched with VB-009. These include the E2 enzymes, UBE2D2, UBE2K, and UBE2V2, all of which are known interactors of the intended E3 ligase target of VB-009. Among the ubiquitin proteins likely conjugated to the proteins in this study, UBA1, RPS27A, SAE1, and the ubiquitin fold containing protein, GABARAPL2 were highly enriched, and UBL4A was represented. The deubiquitinases, OTUD6B and UCHL1 were enriched in VB-009 sample compared to the control samples.

As shown in Table I, 18 proteins possess either an RNA binding property or are components of the splicing or translation machinery. 5 additional proteins associate with DNA. Curiously, 15 proteins are known to associate with Stress granules (SGs), which are cytosolic membraneless organelles involved in RNA metabolism, post-transcriptional regulation, and translational control [Reviewed in: Youn et al, 2019]. Believed to form through phase separation enabled by a combination of interactions among different molecular entities, SGs exhibit a very large number of inter-molecular interactions, including, RNA-RNA interactions (Van Treeck and Parker, 2018), protein-protein interactions, and RNA-protein interactions.

SGs have been implicated in neurodegenerative diseases: Cellular ubiquitination processes are involved in the maintenance of SGs, and may be dysregulated in Amyotropic Lateral Sclerosis (ALS) and Frontotemporal dementia (FTD) (Maxwell et al, 2021; Farawell et al, 2020). Aberrant SG dynamics and a growing number of RNA binding proteins are being investigated as candidates in both diseases (Olney, N.T et al, 2017). Besides FUS, low levels of TDP43, EWSR1, and SMN1 were identified in the VB-009 samples, as well. Tar-binding protein (TDP43), FUS, EWS RNA Binding Protein 1 (EWSR1), TAF15, hnRNPA1, hnRNPA2B1, ATXN2, and TIA1 are the prime candidates which cause or influence disease (Beradan-Heravi et al, 2019).

RNA binding proteins are frequently mutated in ALS and FTD: FUS and Tar binding protein (TDP-43) rank 1st and 10th among the disease candidates forming cytoplasmic inclusions in the degenerating motor neurons of ALS patients and mutations in TDP-43 and FUS causes familial ALS. So, the finding of both these proteins (including RPL3 [Tar-RNA binding protein]) among the ubiquitin-conjugated candidates

with VB-009 suggests a potential link from these proteins to the ubiquitin system, and likely an E3 ligase which may affect the degradation of these proteins.

FUS and TDP43 mutation spectrum in ALS and FTD:

To date, more than 50 different FUS mutations have been described in patients with ALS (Deng et al., 2014), of which many disrupt the nuclear localization signal and result in mislocalization of FUS to the cytoplasm (Dormann et al., 2010; Lagier-Tourenne et al., 2010). The expression in neuronal-like cells of either mutant TDP-43 (M337V) or FUS (R495X) mutant led to UPS dysfunction, suggesting a dysregulation of the UPS system as an additional feature of ALS pathology (Farrawell et al, 2020). TDP-43 is depleted from the nucleus and found as hyperphosphorylated, aggregated cytoplasmic inclusions in ~97% of ALS and ~50% of FTD patients (Giordana et al., 2010). Most of the ALS associated mutations appear in the exon 6 representing the C-terminal glycine-rich region of TDP-43. N-terminal mutations are rare, but the missense mutations A90V and D169G are causative in ALS as well as FTD.

Splicing factor genes are mutated in myeloid malignancies: LUC7L2, SRSF2, and U2AF1 are among the proteins mutated at frequencies ranging between 40% and 85% in different subtypes of myelodysplastic syndrome (MDS) (Visconte et al, 2019). Mutations in U2AF1 at codon S34 and Q157 are found in about 11% of patients with MDS. Likewise, the expression of the L166P mutated form of PARK7 leads to enhanced degradation through the ubiquitin-proteasome system.

Identification of peptide signatures or 'motifs' in FUS shared with other proteins in the VB-009 polyubiquitinated fraction: It is highly intriguing that several of the proteins ubiquitinated with VB-009 expression are frequently mutated in neurological diseases and myeloid malignancies, and may be candidates for targeted degradation in the clinic. Since a peptide directed against a single E3 ligase was used to stimulate polyubiquitination of the 40 proteins identified in this study, there exists the possibility that they may interact with the UPS, or a component of it, in a common manner. Therefore, the protein sequences were investigated further to explore if some common theme emerges, such as, structural or functional motifs that may be shared by some proteins, which may enable targeting them individually or as a group.

E3 ligases recognize their targets through specific motifs referred to as degrons, which may either be a stretch of linear amino acids (physical degrons), or comprised of discontinuous sequences brought in close proximity by the folding of the protein (structural degrons). Degrons have been identified in some E3 ligases, such as, SCF^{FBXL17}, APC/C, SCF^{bTrCP}, and SPOP (reviewed in: Jevtic et al, 2021). Similarly, the

substrate proteins contain conserved sequences, or motifs, which are recognized by the cognate E3 ligase for target engagement prior to ubiquitination. For substrate engagement, degrons may require posttranslational modifications, such as phosphorylation (Winston et al., 1999), acetylation (Shemorry et al., 2013), hydroxylation (Ivan et al., 2001; Jaakkola et al., 2001), ADP ribosylation (Zhang et al., 2011), or arginylation (Yoo et al., 2018), or be inactivated by oxidation (Manford et al., 2020).

A variety of approaches have been effectively employed for identifying motifs in E3 ligases, including, protein interactions (House, 2003; Venables, 2004; and Buchwald, 2013), structural studies (Santelli et al, 2005), and miRNA knockdown (Murphy Schafer et al, 2020) with the E3 ligase Siah1. The polyubiquitinated proteins reported here represent a signature of the action by one or more E3 ligases within the complex cellular environment. Accessory factors, protein modifications, proximity, protein network alterations or other factors may determine the range or substrate specificities and many of these interactions can be motif-based as well. To search for such sequences in the output from VB-009, protein sequences of the candidates in Table I were aligned using COBALT and refined to shorter stretches of ~100 amino acids which exhibited maximal homology. Only contiguous sequence homologies were considered, bipartite sequences were not searched. Additionally, the signatures of proteins within any group may be structurally similar in native proteins or after modifications (structural degrons), which were not searched.

As shown in Table I, 33 proteins from the polyubiquitinated fraction with VB-009 could be assigned to one of 5 distinct consensus sequences. Group 1, consisting of 7 proteins, are rich in Arg and Gly residues, with some Ser residues as well. Only FUS contains the Tri-RGG motif within the homology region (RGG(X0–4)RGG(X0–4)RGG), while SRSF1 contains the di-RG motif (RG(X0–4)RG) (Thandapani et al). Group 1 proteins contain RRM domains in their structure [FUS, RBM26, EIF4B, EIF4H, EIF5A, SRSF1, and SRSF6], however, the regions of maximal sequence homology among these proteins did not consistently map to the RRM domains. Further work is needed to ascertain if the RGG motifs constitute degrons or are involved in interacting with an E3 ligase.

Evidence from literature suggests that the Group 1 signature may possess biological function. For instance, several of the amino acids in FUS homology region are located within the nuclear localization signal (NLS) and are frequently mutated in Amyotropic Lateral Sclerosis (ALS) [Deng and Jankovich]. The GRG triplet (residues 486-488), DRG (502 to 505), G (507), S (513), and RP (524-525) amino acids share homologies with other proteins in Group 1 [Chong et al]. Some clinically significant FUS mutations are truncated at G466X, R495X,and G456vfsx, or may alter the secondary structure of the motif (G472X, or R521G, R521L, R521C). These peptide signatures are being tested if they can serve as an E3 ligase interaction site, more specifically the E3 ligase targeted by VB-009.

The Group 2 signature sequence was not as striking as that of Group 1, but were generally rich in charged or modifiable amino acids (K, R, Q, N, or T) with interspersed serine, glycine or alanine. Importantly, the amino acid region of PARK7 (DJ-1) exhibiting homology with other members of this group is known from literature to contain mutations which cause autosomal recessive forms of Parkinson's disease (PD). A107, E113, and P158, are frequently mutated in PD, and I105, L116, L122 and T154 are located within the (Hering et al, 2004) homologous regions of Group 2. Earlier studies of the E64D variant in fibroblasts from a patient bearing the homozygous mutation showed that levels of the protein are decreased. Likewise, the E163K mutation reduces the stability of the protein *in vitro*, and the P158del variant is unstable when expressed in cells. Besides HMGB1, this group also includes HMGB2 and HMGB3 which were identified in the sample, but not included in the homology search on account of the high level of sequence conservation with HMGB1. This group also consists of FABP5, wherein the G114R and N124S polymorphisms have been implicated in schizophrenia and autism (Shimamoto et al, 2014).

The Group 3 signature sequence exhibited a high degree of sequence conservation, and may involve a peroxiredoxin fold. The remaining homology groups also consisted of at least one protein mutated (or possessing a causative polymorphism) in hematological or neurological diseases. This includes the S34 and Q157 mutations in U2AF1 found in about 11% of patients with myelodysplastic syndrome. The significance of these mutations in relation to any biological activity of the homologies identified in this study needs to be fully understood, particularly in the context of the ubiquitin proteasome system function.

Finally, it remains to be determined why several putative motifs were observed in this study, although the PINTAC was directed to a single E3 ligase. The typical protein interaction motif is around 6-12 amino acids in length or shorter. So, the PINTACs may contain a few tandem or overlapping motifs in their sequence (currently 30 amino acids), each with the ability of recruiting distinct sets of proteins. Additional proteins, such as ubiquitin accessory factors or chaperones may partially determine interaction specificity, and may have recruited some proteins. G3BP2, DNAJC8, TBCA, SUGT1, STIP1, or U2AF1 are likely candidates and additional proteins not identified in this study may also be involved. Many other RNA binding proteins and chaperones were identified in this study but are not presented here since they did not meet the selection threshold of at least 20 peptides represented in the mass spectrometry data. Optimization of PINTAC sequence, deeper sequencing of the polyubiquitinated fraction with mass spectrometry, computational analysis of the candidate motifs, and comparisons of the ubiquitinated fraction across more samples (directed towards different E3 ligases) may aid their identification, as well as the search for motifs.

Conclusion

Our study highlights the fact that polyubiquitination of proteins can be triggered by overexpressing a PINTAC, which are peptidic in nature. The consistency with which several peptides gave rise to polyubiquitinated proteins and the distinct sets of proteins which each targets, suggest the existence of a common theme or mechanism, which needs to be elucidated. Meanwhile, this report explores the design and deliberate use of the peptides to target proteins with a specific E3 ligase for each.

While the first clinical success of TPD tools, such as, PROTACs and glues are anxiously awaited, this study highlights an alternative strategy of functionally inactivating (or modifying the function) of proteins, rather than total degradation. Since there was excellent congruence between polyubiquitinated proteins and the intended specificity of each peptide, they have the potential for development as tools for discovery and/or disruption of proteins of interest. They can also be targeted to sub-cellular compartments for disrupting proteins active in those compartments, while retaining the essential functions of the same protein intact in other parts of the cells. This aspect is an area of need in TPD, but are not achievable by PROTACs or glues.

The peptides may also serve as reporters of the specific roles played by E3 ligases. For instance, proteins polyubiquitinated by VB-009 were found to have roles predominantly in RNA metabolism and some that are components of stress granules, which may suggest a functional link between the E3 ligase and these cellular processes. Several of the same proteins are clinically important in ALS, FTD, and hematological malignancies. Curiously, stress granules are also known to be differentially regulated in these diseases. Hence, the peptides may serve to elucidate the fine structure of protein networks and regulatory feedback mechanisms in these diseases.

The peptides can also be used to identify small molecules for combining with PROTACs or gules. The polyubiquitinated proteins represent biomarker sets specifically targeted to the ubiquitome. These can help identify small molecules, in *in silico* design and structural studies via co-crystallization, or identify new glue candidates. With improvements, the peptides may serve as a tool for detecting aberrations in ubiquitin mediated degradation of select proteins in clinical samples, or the impact of mutations, particularly in E3 ligases. Doing so can help improve the design of TPD solutions and evaluate outcomes. Further optimization of the peptide sequences can help target smaller subsets, or individual proteins, with peptides.

Methods

Reagents:

DMEM (Cat. No. 11965-092), Penicillin-Streptomycin (Cat. No. 15140-122), Trypsin-EDTA (Cat. No. 25300-054) and Fetal bovine serum (Cat. No. A31604-02) were purchased from Thermo Fisher Scientific. Lipofectamine2000 (Cat. No. 2352136) and Novex WedgeWell (Cat. No. XP00165) precast polyacrylamide gels were purchased from Thermo Fisher Scientific. Rabbit polyclonal antibodies against

human Ubiquitin (Cat. No. 3933; Acc. No. P62987) and MEK1/2 (Cat. No. 9122) were purchased from Cell Signaling Technology. Mouse monoclonal antibodies for MAPK1 (sc-514302), GAPDH (Cat. No. sc-47724), and B-RAF (sc-55522) were purchased from Santa Cruz Biotechnology. Custom oligonucleotides were purchased from Thermofisher and pCMV6-AC-Myc-DDK, mammalian expression vector (Cat. No. 100007) was purchased from Origene Technologies. The restriction enzymes AsiSI (Cat. No. R0630S) and EcoRV (Cat. No. R0195S) were purchased from New England Biolabs.

Clone construction: Peptide sequences to be used for clone construction were reverse translated using Expasy Translate, with reference to the human genetic code. Oligonucleotide pairs comprising the peptide sequences were synthesized bearing the appropriate overhangs for SgfI and EcoRV, annealed, and inserted into the pCMV6-AC-Myc-DDK vector digested with AsiSI and EcoRV. Bacterial transformation and clone selection were conducted according to standard procedures. The sequences of the peptides were confirmed by sequencing the recombinant clones.

Cells, Transfection, whole cell protein preparation, and western analysis:

HEK293T cells were purchased from ATCC (Cet. No. CRL-3216) and expanded according to manufacturer's recommendations. For transfection: Cells were seeded in 100 mm tissue culture dishes at a density of 2.2 x 10^6 in 2 ml of growth medium (DMEM with 10% fetal bovine serum), and grown to a confluency of 90 per cent. Plasmid DNAs were introduced into the cells via transfection using Lipofectamine2000 according to the manufacturer's protocol. Transfected cells grown for 72 hours at 37 °C under 5% CO2 without antibiotic selection.

For whole cell protein preparation, the transfection plates were washed once with cold PBS. The adherent cells were resuspended in 400 ul of 1x Cell Lysis buffer (9803, Cell Signaling Technology) and incubated at room temperature for 15 min with occasional mixing. Lysates were cleared by centrifugation and the supernatants stored in aliquots at –20 °C.

Whole cell proteins were separated on SDS-polyacrylamide gels, transferred to nitrocellulose membranes and probed with the indicated antibodies according to western blotting protocols. Secondary antibodies tagged with infrared probes were purchased from Licor. Images were acquired using Odyssey infrared scanner according to manufacturer's recommended protocol.

Lenalidomide treatment of cells expressing PINTACs:

Transfected cells incubated without antibiotic selection for 48 hr were treated with Lenalidomide (Cat. No. A4211, ApexBio Tech) at a final concentration of 650nM, and incubated for a further 4 hr. Cell extracts were prepared as above.

Sample preparation and differential mass spectrometry analysis:

100 ul of whole cell proteins from transfected cells were mixed with 300 ul of PBS, clarified by centrifugation, and heated to 56 ºC for 15 min in a heat block. The sample was immediately applied to an Amicon Ultra-0.5 Centrifugal Filter Unit (Cat. No. UFC510024) and centrifuged at 9k rpm for 10 min. The supernatants were washed thrice with PBS at 56 ºC, and collected for mass spectrometry analysis. The 100k supernatants were confirmed by western blotting with the anti-ubiquitin antibody.

The 100k supernatants from VB-009 sample, 4 additional samples with control PINTACs, and 1 transfected with pCMV6 vector alone were analyzed by mass spectrometry. Mass spectrometry analysis was performed by the Biological Mass Spectrometry Facility of Robert Wood Johnson Medical School and Rutgers, The State University of New Jersey. Mascot software was used to identify proteins from LC-MS/MS data against the Swiss-Prot human database using a 1% false positive discovery rate (FDR).

Mass spectrometry data analysis:

The output from mass spectrometry (in Excel format) was sorted with respect to VB-009 (from high to low values) for the number of peptides identified against each protein. Proteins represented by at least 20 peptides, which is greater than 2 -fold as compared with the same protein in the pCMV6 control sample plus at least 2 other control PINTAC-treated cells were considered derived from proteins preferentially ubiquitinated in presence of VB-009.

Sequence analysis and alignments:

Amino acid sequences of the candidate proteins ubiquitinated by VB-009 were aligned using COBALT alignment software, iterated by truncation to the regions exhibiting maximal sequence similarity, and realigned. The grouped sequences were downloaded into a .png file.

References

1. Bekes, M, *et al* 2022. PROTAC targeted protein degraders: the past is prologue. Nat. Rev. 21:181–200.

2. Wertz, IE and Wang, X. 2019. From Discovery to Bedside: Targeting the Ubiquitin System. Cell Chem. Biol. 26:1–22.
3. Schneider, M, *et al* (2021). The PROTACtable genome. Nat. Rev. Drug Disc. 20:789–797.
4. Bard JAM, *et al* 2018. Structure and Function of the 26S Proteasome. Annu Rev Biochem. 20;87:697–724.
5. Oh E, *et al* 2018. Principles of Ubiquitin-Dependent Signaling. Annu Rev Cell Dev Biol. 6;34:137–162.
6. Kannt A, and Đikić I 2021. Expanding the arsenal of E3 ubiquitin ligases for proximity-induced protein degradation. Cell Chem Biol. 15;28(7):1014–1031.
7. Duan, S, and Pagano, M 2021. Ubiquitin ligases in cancer: Functions and clinical potentials. Cell Chemical Biology, 28: 918–933.
8. Johnston, JA *et al*., 1998. Aggresomes: A Cellular Response to Misfolded Proteins. J. Cell Biol. 143: 1813–1898.
9. Dehvari, N *et al*., 2012. Amyloid precursor protein accumulates in aggresomes in response to proteasome inhibitor. Neurochem. Intl. 60:533–542.
10. Maynard, Christa J *et al*., 2009. Accumulation of ubiquitin conjugates in a polyglutamine disease model occurs without global ubiquitin/proteasome system impairment. Proc. Natl. Acad. Sci. USA. 106: 13986-91.
11. Chadchankar J, *et al*., 2019. Inactive USP14 and inactive UCHL5 cause accumulation of distinct ubiquitinated proteins in mammalian cells. PLoS One. 2019:e0225145.
12. Reeg S, and Grune T, 2015. Protein Oxidation in Aging: Does It Play a Role in Aging Progression? Antioxid. Redox Signal. 23:239–255.
13. Canu, N *et al*., 2000. Proteasome involvement and accumulation of ubiquitinated proteins in cerebellar granule neurons undergoing apoptosis." J. Neurosci. 20: 589–99.
14. Kobayashi S, *et al*., 2020. West Nile virus capsid protein inhibits autophagy by AMP-activated protein kinase degradation in neurological disease development. PLoS Pathog. 16:e1008238.
15. Ito, T *et al*., 2021. Exploiting ubiquitin ligase cereblon as a target for small-molecule compounds in medicine and chemical biology. Cell Chem. Biol. 28:987–999.
16. van Andel, H, *et al*., 2019. Aberrant Wnt signaling in multiple myeloma: molecular mechanisms and targeting options. Leukemia. 33:1063–1075.
17. Youn, J-Y *et al*., 2019. Properties of Stress Granule and P-Body Proteomes. Mol. Cell 76:286–294.
18. Van Treeck B and Parker R, 2018. Emerging Roles for Intermolecular RNA-RNA Interactions in RNP Assemblies. Cell. 9:791–802.
19. Maxwell, BA, *et al*., 2021. Ubiquitination is essential for recovery of cellular activities after heat shock. Science 372: DOI: 10.1126/science.abc3593.
20. Farrawell NE, *et al*., 2020. Ubiquitin Homeostasis Is Disrupted in TDP-43 and FUS Cell Models of ALS. iScience. 20:101700.
21. Olney NT *et al*., 2017. Frontotemporal Dementia. Neurol Clin. 35:339–374.

22. Baradaran-Heravi Y, *et al*., 2020. Stress granule mediated protein aggregation and underlying gene defects in the FTD-ALS spectrum. Neurobiol Dis. 134:104639.
23. Deng H, *et al*., 2014. The role of FUS gene variants in neurodegenerative diseases. Nat. Rev. Neurol. 10:337–48.
24. Dormann D, *et al*., 2010. ALS-associated fused in sarcoma (FUS) mutations disrupt Transportin-mediated nuclear import. EMBO J. 29:2841–57.
25. Lagier-Tourenne C, *et al*., 2010. TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration. Hum Mol Genet. 2010 19:R46-64.
26. Giordana MT, *et al*., 2010. TDP-43 redistribution is an early event in sporadic amyotrophic lateral sclerosis. Brain Pathol. 20:351–360.
27. Visconte V, *et al*., 2019. Mutations in Splicing Factor Genes in Myeloid Malignancies: Significance and Impact on Clinical Features. Cancers (Basel). 22:1844.
28. Jevtic, P, *et al*., 2021. An E3 ligase guide to the galaxy of small-molecule-induced protein degradation. Cell Chem. Biol. 28:https://doi.org/10.1016/j.chembiol.2021.04.002.
29. Winston, J T *et al*., 1999. The SCFbeta-TRCP-ubiquitin ligase complex associates specifically with phosphorylated destruction motifs in IkappaBalpha and beta-catenin and stimulates IkappaBalpha ubiquitination in vitro. Genes & development 13: 270–83.
30. Shemorry A, *et al*., 2013. Control of protein quality and stoichiometries by N-terminal acetylation and the N-end rule pathway. Mol Cell. 50:540–551.
31. Ivan M, *et al*., 2001. HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. Science. 292:464–8.
32. Jaakkola P, *et al*., 2001. Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. Science. 292:468–72.
33. Zhang Y, *et al*., 2011. RNF146 is a poly(ADP-ribose)-directed E3 ligase that regulates axin degradation and Wnt signalling. Nat Cell Biol. 13:623–9.
34. Yoo, YD, *et al* 2018. N-terminal arginylation generates a bimodal degron that modulates autophagic proteolysis. Proc. Natl. Acad. Sci. USA 115: E2716-E2724.
35. Manford AG, *et al*., 2020. A Cellular Mechanism to Detect and Alleviate Reductive Stress. Cell. 183:46–61.
36. House, CM, *et al*., 2003. A binding motif for Siah ubiquitin ligase. Proc. Natl. Acad. Sci. 100:3101–3106.
37. Venables, JP, *et al*., 2004. SIAH1 targets the alternative splicing factor T-STAR for degradation by the proteasome. Human Molecular Genetics, 13: 1525–1534.
38. Buchwald, M., *et al*., 2013. SIAH ubiquitin ligases target the nonreceptor tyrosine kinase ACK1 for ubiquitinylation and proteasomal degradation. Oncogene 32, 4913–4920.
39. Santelli, E et al., (2005). Structural Analysis of Siah1-Siah-interacting Protein Interactions and Insights into the Assembly of an E3 Ligase Multiprotein Complex. J. Biol. Chem., 280:34278–34287.

40. Murphy Schafer, AR, *et al*, 2020. The E3 Ubiquitin Ligase SIAH1 Targets MyD88 for Proteasomal Degradation During Dengue Virus Infection. Front. Microbiol., https://doi.org/10.3389/fmicb.2020.00024.
41. Thandapani, P, *et* al., 2013. Defining the RGG/RG Motif. Mol. Cell, 50:613–623.
42. Hering, R, *et al*., 2004. Novel homozygous p.E64D mutation in DJ1 in early onset Parkinson disease (PARK7). Hum. Mutat. 24:321–329.
43. Shimamoto, C, *et* al., 2014. Functional characterization of FABP3, 5 and 7 gene variants identified in schizophrenia and autism spectrum disorder and mouse behavioral studies. Human Molecular Genetics, 23: 6495–6511.

Declarations

Conflict of interest statements: The author declares no conflict of interest.

Tables

Table 1 is available in the Supplementary Files section.

Figures

a. PINTACs employed in this study and their intended targets.

PINTAC	Length [amino acids]	Intended Target of PINTAC
VB-001	27	Enzyme 1
VB-002	27	Enzyme 2
VB-003	27	Enzyme 3
VB004	27	Enzyme 4
VB-009	26	E3 ligase 1
VB-010	26	E3 ligase 1
VB-011	27	E3 ligase 2
VB-012	27	Protein kinase
VB-013	27	E3 ligase 3
VB-014	26	E3 ligase 4
VB-015	29	E3 ligase 5

b. Process for overexpression of PINTACs in HEK293T cells and analysis of polyubiquitinated proteins.

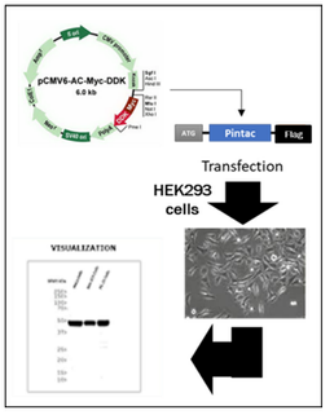

Fig. 1c.

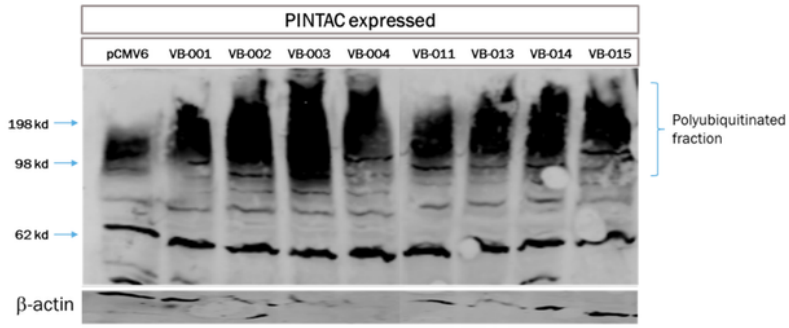

Figure 1

PINTACs employed in this study and their intended targets. The peptides used in this study, methods employed for introducing them into cells and analysis by western blotting are shown. 1a). 11 peptides were designed for targeting the proteins indicated in column 3 are shown. A pair of complementary oligonucleotides corresponding to the peptide sequences were synthesized with the appropriate

overhangs, and inserted in the SgfI and EcoRV sites of pCMV6-AC-Myc-DDK mammalian expression vector. The clones generated from them were confirmed by sequencing the recombinant plasmids.

1b). Recombinant plasmids were introduced into HEK293T cells by lipofection, and incubated without antibiotic selection for 48 hr. Whole cell proteins from the transfected cells were analyzed by western blotting.

1c). Total proteins from PINTAC transfected cells as indicated were fractionated on a 12% polyacrylamide gel with SDS, and probed by western blotting with an rabbit antibody to human ubiquitin. The image was acquired using the Odyssey infra-red imager and analyzed with ImageStudio software.

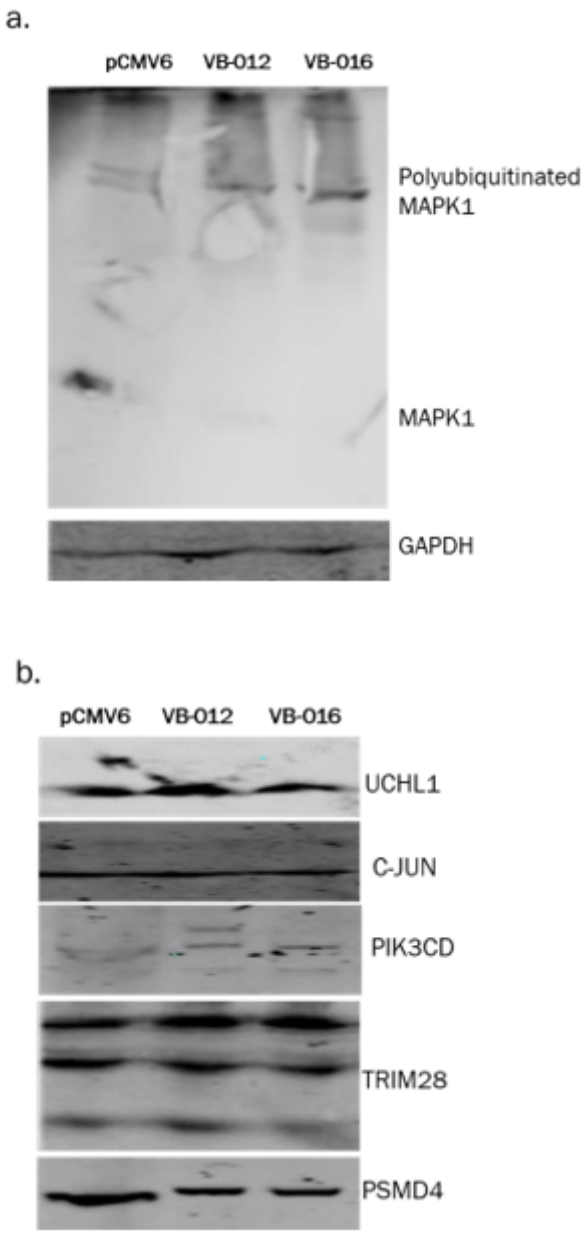

Figure 2

Target selectivity of polyubiquitination with PINTAC expression. Western blotting of total proteins from cells treated with PINTACs for MAPK1 (VB-012) or TRIM28 (VB-016). pCMV represents the control vector without PINTAC sequence.

2a). The blots were probed with a mouse monoclonal antibody to human MAPK1. Native MAPK1 and the higher molecular weight, polyubiquitinated form of MAPK1 are indicated.

2b). The same samples were probed with additional antibodies. The polyubiquitinated regions of these images (not shown) did not reveal any bands or smear.

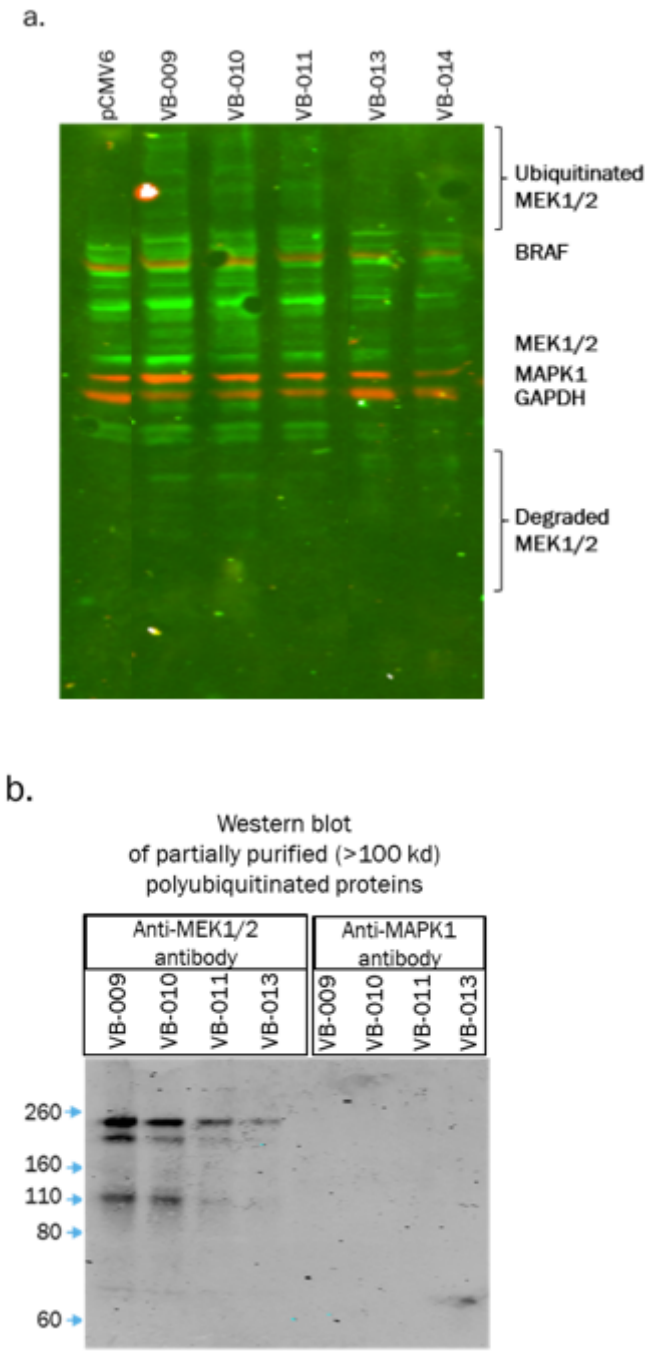

Figure 3

Polyubiquitination of MEK1 and MEK2 by PINTACs. Western blotting of total proteins treated with the PINTACs shown.

3a). Multiplexed probing of western blots with a rabbit polyclonal antibody recognizing MEK1 and MEK2 (shown in green), as well as mouse monoclonal antibodies to BRAF, MAPK1, and GAPDH (in red).

3b). Whole cell extracts were applied to an Amicon cartridge with a size exclusion of 100 kd. The supernatants were used for immunoprecipitation with a rabbit polyclonal antibody to ubiquitin and analyzed by western blotting with the MEK1/2 antibody.

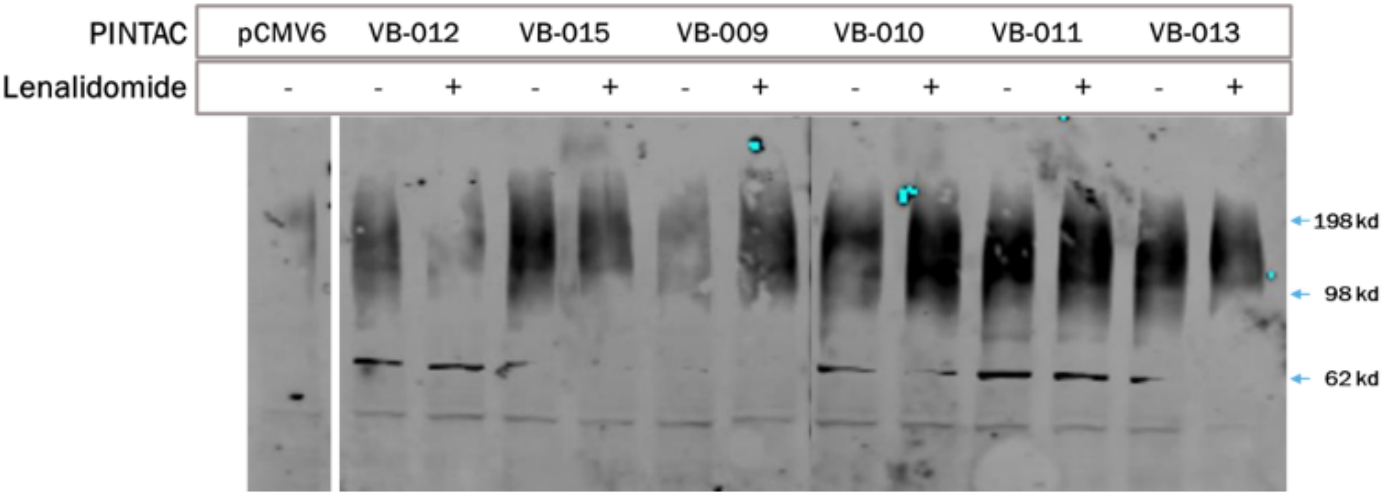

Figure 4

PINTAC mediated polyubiquitination is sensitive to lenalidomide treatment. HEK293T cells transfected with the appropriate PINTACs for 48 hr were incubated with lenalidomide. Whole cell extracts from the cells were analyzed by western blotting with a rabbit polyclonal antibody to human ubiquitin. + indicates lenalidomide treatment; - indicates DMSO controls.

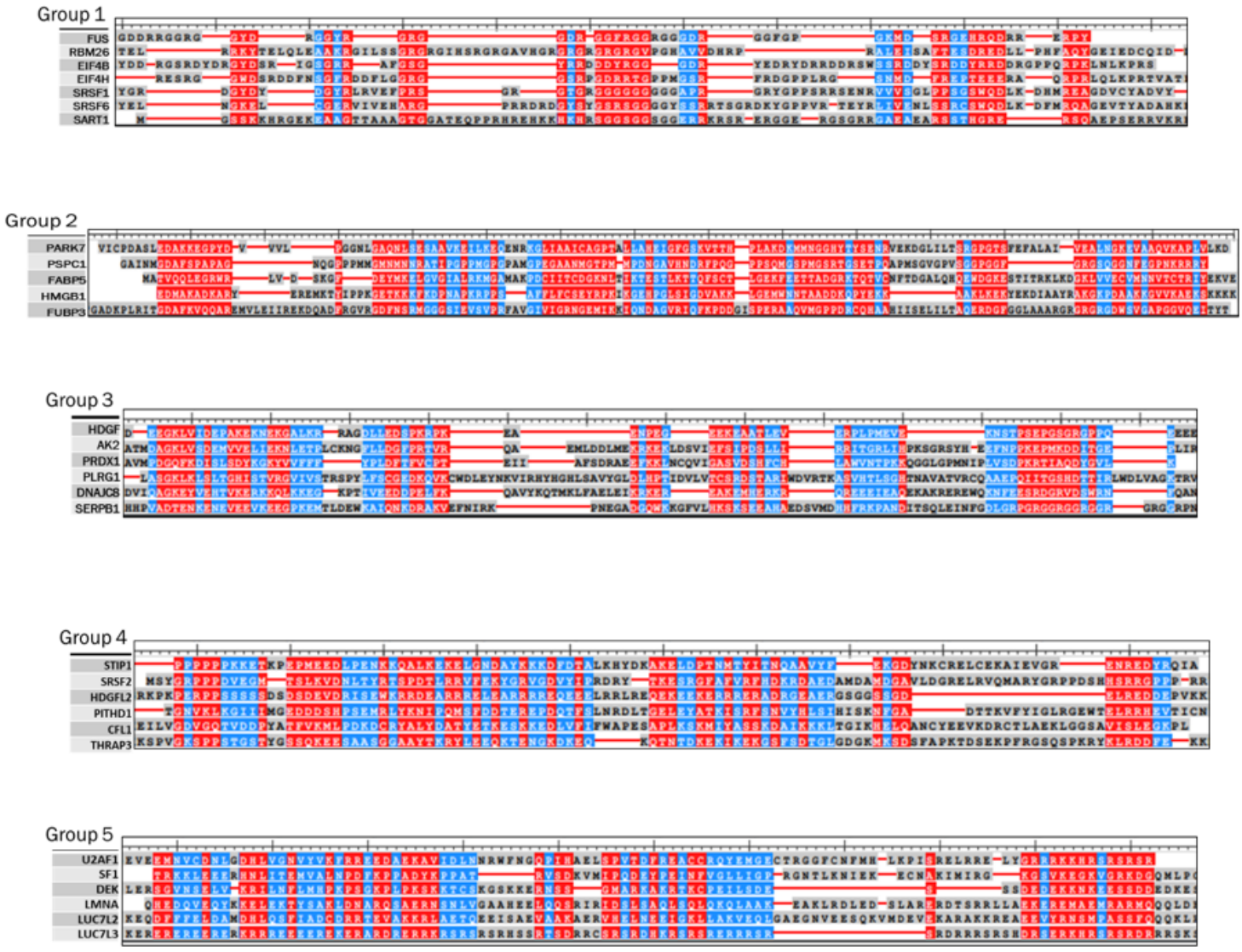

Figure 5

Candidate motifs identified in proteins polyubiquitinated with VB-009. Each alignment shows the proteins and their amino acid segments as indicated in Table I.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table1.docx