C-End binding by TPR co-chaperones links proteolysis and protein homeostasis

by
Matthew Anthony Ravalin

DISSERTATION
Submitted in partial satisfaction of the requirements for degree of
DOCTOR OF PHILOSOPHY
in
Chemistry and Chemical Biology
in the
GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

Approved:

Jason Gestwicki
Chair

Charles Craik

Michelle Arkin

Committee Members
Dedicated Christine Richardson and Lana

Thank you for helping me keep it together(ish)
Acknowledgements

I would like to thank my advisors Jason Gestwicki and Charles Craik for their guidance, flexibility, and trust. I would also like to thank my thesis committee member and collaborator Michelle Arkin for support and guidance during my time at UCSF.

I am grateful to all current and former members of the Gestwicki and Craik labs, particularly those who have served as mentors and collaborators along the way.

I could not have completed this work without the love and support of my family including my parents Michael and Dayna Ravalin, my sister Maddisen Ravalin, my partner Christine Richardson and our loving and enthusiastic street dog Lana.
Acknowledgement of Previously Published Material

Chapters 1 and 2 are reprints of the material as it appears in:


C-end binding by TPR co-chaperones links proteolysis and protein homeostasis

By Matthew Anthony Ravalin

Abstract: Proteolysis is unique among post translational modifications because it is irreversible. The specific recognition of protein termini is an essential mechanism by which proteolytic enzymes mediate signaling in biology. The identification of terminal recognition elements with complimentary specificity to proteases, often as part of a E3 ubiquitin ligase, has facilitated the elucidation of protease signaling pathways that function in critical cellular processes. The caspase family of protease are well known for their role in directing the fate of the cell by initiating and executing inflammatory and death-related signaling cascades. The biochemical feature that unites this family of enzymes is the ability to hydrolyze a peptide bond following and aspartic acid. We have determined that the TPR family of co-chaperones, including the E3 ubiquitin ligase CHIP, can also bind to a subset of new C-termini generated by caspase activity. These co-chaperones were thought to function predominantly, if not exclusively, by interacting with a conserved Glu-Glu-Val-Asp at the C-terminus of cytosolic Hsp70s and 90s. This work has uncovered new functions for both caspases and TPR co-chaperones and has identified a dedicated molecular framework for caspases and protein homeostasis networks to exchange information. These systems play a critical role in balancing cell fate in development and disease making it likely that critical signaling nodes occur at this interface.
# Table of Contents

Chapter 1: ........................................................................................................................................... 1

End-Binding E3 ubiquitin ligases enable protease signaling

References........................................................................................................................................... 18

Chapter 2: ........................................................................................................................................... 37

Specificity for latent C-termini links the E3 ubiquitin ligase CHIP to caspases

References........................................................................................................................................... 67

Chapter 3: ........................................................................................................................................... 100

Probing selectivity among TPR co-chaperones using peptide photoprobes

References........................................................................................................................................... 110
List of Figures

Figure 1.1: End binding E3 ubiquitin ligases regulate protease product stability to enable signaling.................................................................30

Figure 1.2: Specific determinants of N-end binding link protease and E3 function in signaling pathways..................................................................32

Figure 1.3: Figure 3. N-end binding is critical for the regulation of apoptosis........34

Figure 1.4: C-end binding adaptors to Cullin E3 ubiquitin ligases link degradation and non-prime protease specificity........................................................................35

Figure 1.5: C-end binding by the E3 ubiquitin ligase CHIP links caspases and protein homeostasis....................................................................36

Figure 2.1: CHIP function is rooted in the recognition of C termini..................74

Figure 2.2: Structural basis for CHIP TPR domain specificity..........................76

Figure 2.3: Proteome-wide prediction of CHIP TPR interactions with C-termini....78

Figure 2.4: Tau^{D421} specifically recruits CHIP...........................................79

Figure 2.5: CHIP specifically binds and inhibits mature caspase-6....................80

Figure 2.6: CHIP interactions at latent C-termini suggest a role in the progression of AD.................................................................................82

Figure S2.1: DSF PSSCL Screen results..........................................................88

Figure S2.2: Competition FP PSSCL Screen......................................................89

Figure S2.3: Competition binding of acetylated 5mer peptides..........................90

Figure S2.4: Structural studies of peptide-bound CHIP TPR..............................92

Figure S2.5: CHIPScore of the C-terminome, Degradome, and proteome........93
Figure S2.6: Caspase cleavage at tau^{D421} directs binding of and ubiquitination by CHIP.................................................................................................................................95

Figure S2.7: Caspase6^{D179} is bound and inhibited by CHIP.................................................................97

Figure S2.8: CHIP mRNA is not downregulated in AD.............................................................................98

Figure S2.9: Un-cropped blots from figure 4 and 5..................................................................................99

Figure 3.1: TPR photoprobes bind and label TPR co-chaperones in vitro..........................112

Figure 3.2: TPR photoprobes specifically label CHIP in cell lysate.................................................113

Figure S3.1: TPR photoprobe labeling time course.................................................................114

Figure S3.2: Biotin-TPR photoprobe competition binding to CHIP.................................115
List of Tables

Table S2.1: X-ray crystallography statistics.................................................................83
Table S2.2: Rosetta ddg_monomer values for CHIP TPR ligand.............................84
Table S2.3: CHIPScore hits from the explicit C-terminome.....................................85
Table S2.4: CHIPScore hits from the Degrabase.......................................................86
Table S2.5: Predicted CHIP interactions from the proteome....................................87
Chapter 1

End-Binding E3 ubiquitin ligases enable protease signaling

Matthew Ravalin, Koli Basu, Jason E. Gestwicki and Charles S. Craik
Abstract. The modular recognition of post-translational modifications directs the assembly of macromolecular complexes in cells. This mechanism is indispensable for signaling events critical to all life. Proteolysis is a unique PTM because it is irreversible and generates two protein products in the form of a nascent N and C terminus. This distinction from other PTMs underscores unique challenges and opportunities in the study of proteolysis as a signaling modality. We highlight the critical biological function for end-binding E3 ubiquitin ligases that specifically engage in protein-protein interactions with the termini that arise from proteolysis. These interactions potentiate proteolytic signaling by tuning the duration and magnitude of signaling events, a role fulfilled by the catalysis of PTM removal in other systems. We describe recently characterized examples wherein biological functions depend on these interactions and formalize a conceptual framework within which to parse these pathways to elucidate the biology they direct.

1. The unique attributes of proteolysis as a PTM. The biophysical underpinnings of protein-protein interactions (PPIs) are diverse.¹ ² Among PPIs, those that occur between a globular protein domain and a peptide have garnered significant interest because of the role they play in signaling and their amenability to chemical modulation.³ ⁴ This class of PPI is often regulated by post-translational modifications (PTMs). When this is the case, the formation of the PPI is a composite readout of the activities of the PTM placing (E) and removing enzymes (E*), as well the specificity of the PTM-binding modular recognition element (MRE).⁵ This framework has been widely employed in the study of chromatin modification. In these signaling networks the interplay between
readers (MREs), writers (E), and erasers (E*) directs chromatin state and the transcriptional landscape of the cell.\textsuperscript{6,7} The development of chemical probes for the components of these networks has facilitated a better understanding of the biology they direct.\textsuperscript{8}

Proteolysis, the hydrolysis of an amide bond in a polypeptide chain, is a unique PTM because it is irreversible, and thus represents a committed biological step. Given this irreversibility, proteolytic signals exist in flux rather than in a dynamic equilibrium.\textsuperscript{9,10} Each proteolytic event generates two products, a nascent N and C terminus with opposing charges (\textbf{Figure 1.1A}). Proteolytic activity ranges from complete degradation to limited cleavage. It is capable of transmitting signals by removing whole domains, facilitating the maturation of structural features, destroying entire proteins, or by removing a single amino acid.\textsuperscript{11} The way to recover the function of a cleaved protein is through re-synthesis. Conversely the way to remove the function of cleavage products is through degradation. This feature positions the cellular effectors of protease product stability at a critical regulatory node for proteolytic signaling. The ubiquitin proteasome system (UPS) is a major pathway for protein degradation in eukaryotes.\textsuperscript{12} Generally, the propensity for a protein to be degraded by the UPS is controlled by its ubiquitination and delivery to the proteasome. The specificity of ubiquitination and degradation in the cell is controlled by the activity of the 500-1000 E3 ubiquitin ligases expressed in mammals.\textsuperscript{13} A subset of these E3 ubiquitin ligases recognize protein termini (i.e. are end-binding) and are poised to regulate the stability of the products of proteolysis (\textbf{Figure 1.1B and Figure 1.1C}).\textsuperscript{14,15}
2. Shared specificity links proteolysis and E3 function. The unique features of proteolysis as a PTM have enabled the evolution of signaling cascades that exploit it as a means of transmitting information.\textsuperscript{16} Indeed, biological processes rooted in proteolysis are ubiquitous in living systems. Additionally, links between terminal recognition and protein degradation have been characterized in bacteria, plants, yeast, and mammals.\textsuperscript{17} In eukaryotes E3 ubiquitin ligases that contain a terminal recognition element (TRE) couple the terminal chemical composition of a polypeptide substrate to its degradation. An implicit connection between a protease and E3 ligase arises from the physical interaction of each with a shared substrate. The strength of this connection depends on the overlap of the specificity of each (Figure 1.1D). Explicitly, a link between a protease and a C-end binding E3 depends on the complementarity between the C-end TRE and amino acids N-terminal to the cleaved bond (the non-prime sites). Accordingly, the link between a protease and an N-end binding E3 depends on the complementarity between the N-end TRE and the amino acids C-terminal to the cleaved bond (the prime sites) (Figure 1.1E). Given that product degradation is a regulator of protease signaling flux, the development of a comprehensive understanding of protease-E3 crosstalk is likely to identify regulatory nodes in protease signaling. A key challenge in understanding the role of proteases in biological processes is the differentiation of functional substrates from bystanders (i.e. cleavages that yield no change in function). Focusing on cleavage events coupled to subsequent end-binding E3 activity may facilitate the prioritization of functional cleavages. With this framework in hand we can hope to leverage an understanding of the specificity of proteases and end-binding E3 ubiquitin ligases determined \textit{in vitro} to illuminate biological signaling pathways \textit{in vivo}. 
3. **N-degrons tune protease substrate function.** A degron is a protein feature that potentiates the degradation of a given substrate. A N-degron is such a feature that is present at the extreme N-terminus of a polypeptide. In total, the N-degron pathways link *in vivo* protein stability to the physiochemical properties of the N-terminus of a polypeptide.\(^{18}\) N-degron pathways are present in prokaryotes and eukaryotes. In eukaryotes, these pathways were classically divided into two arms. In the Ac/N-degron pathways, protein degradation is triggered by the recognition of an acetylated N-terminal sequence by dedicated E3 ubiquitin ligases.\(^{19}\) The bulk of eukaryotic proteins are irreversibly acetylated co-translationally.\(^{20,21}\) Acetylation can be preceded by the removal of the N-terminal Met generated by the start codon via the activity of methionine aminopeptidases (MetAPs).\(^{22,23}\) The convolution of mRNA encoded N-terminal sequence, acetylation, and demethionylation impart newly translated proteins with the ability to adopt an array of physiochemical features at the N terminus. The propensity for an encoded N-terminal sequence to be acetylated or demethionylated is dictated by the specificity of the MetAPs and N-terminal acetylases. It general, the removal of methionine from the N-terminus is favored in instances when a sterically small (Ala, Val, Ser, Thr, Cys, Gly, and Pro) residue is present at the position directly adjacent to the N-terminal Met. N-terminal acetylation is favored for all of these residues with the exception of Gly and Pro.\(^{24}\)

The other classical arm of the N-degron pathways are the Arg/N-degron pathways. In this arm, N-termini bearing positively charged (Arg, Lys, His, e.g. Type I degrons) or large hydrophobic (Leu, Phe, Trp, Tyr, Ile, e.g. Type II degrons) are
ubiquitinated by a subset of the UBR family of E3 ubiquitin ligases (UBR 1, 2, 4, and 5 in mammals). These E3 ligases depend on the specific recognition of the positively charged N-terminal amine to bind to their substrates. Type 1 degrons are recognized by the UBR domain that contains a binding site capable of coordinating the positive charges of both the N-terminus and the side chain of the N-terminal Arg, Lys, or His (Figure 1.2A). The structural aspects of the recognition of Type II degrons is less understood, however a discrete recognition domain in UBR1 and UBR2 is responsible for binding these sequences. This domain is homologous to a TRE found in bacteria, ClpS, which binds to N-terminal hydrophobic and aromatic amino acids, directing degradation by ClpXP (Figure 1.2B). Collectively, N-termini that are directly recognized and ubiquitinated by the UBR E3s without being modified are referred to as “primary” N-degrons.

Some N-termini require chemical modification prior to being shunted through the Arg/N-degron pathways. For example, N-terminal Asp and Glu residues must first be modified with an N-terminal Arg by the Ate family of Arg transferases. Given the necessity for a single modification to sequences bearing terminal Asp and Glu residues they are referred to as “secondary” degrons. N-terminal Cys residues require oxidation by nitic oxide prior to argenylation by Ate. N-terminal Asn and Gln are converted to Asp and Glu respectively via side-chain deamidation by Ntaq1 before argenylation can occur. As they require two modifications prior to recognition by the Arg/N-degron E3 ligases, sequences bearing terminal Cys, Asn, or Gln are referred to as “tertiary” degrons. The generation of secondary and tertiary degrons by proteolysis decouples the link between protease and TRE specificity so that the sequences generated by
proteolysis must be considered in the context of the intermediary activity of the additional enzymes necessary to produce a UBR binding sequence. A poignant example of this logic is evident in the processing of the Ubiquitin specific peptidase 1 (USP1). USP1 auto-proteolyzes to yield a nascent N terminus bearing a conserved Gln residue. The degradation of this USP1 fragment requires the activity of both Ntaq1 and Ate1. Therefore, although USP1 does not generate a N-terminal Arg directly, its prime-side specificity still links it to the UBR E3 ligases.

Although N-degron pathways were first described over thirty years ago, facets of the biological functions they direct are still being uncovered. Given the propensity for N-terminal Gly and Pro residues to be generated by MetAP activity but resist acetylation and degradation by both the Ac/N-degron Arg/N-degron pathways, the fate of these N-termini has remained a mystery. Recently it was discovered that GID4, a component of the GID E3 ubiquitin ligase complex in *S. cerevisiae* specifically bound proteins bearing a free N-terminal proline and shunted them to the UPS. This pathway proved critical for the rapid degradation of gluconeogenic enzymes when *S. cerevisiae* were transitioned from acetate or ethanol-based growth media to glucose-based media. Further work provided structural and biophysical evidence that demonstrated that human GID4 specifically recognized N-terminal proline and required a free amine at the N-terminus (Figure 1.2C).

An N-degron pathway for unacetylated Gly residues has also recently been described. Briefly, a functional genomics screen identified Cul2<sup>ZYG11b</sup> and Clu2<sup>ZER1</sup>, as the E3 ubiquitin ligase complexes responsible for the degradation of N-terminal Gly bearing polypeptides. An analysis of the specificity of degradation determined that the
presence of a proximal Phe, Gly, His, Leu, Met, or Tyr residue favored degradation. These Gly N-degrons were depleted in the native human proteome but enriched in the N-termini produced by caspase cleavages. Furthermore, N-myristoylation of N-terminal Gly was shown to compete with E3 ligase activity so that proteins or protein fragments bearing N-terminal Gly which do not readily undergo myristoylation and subsequent membrane localization are likely to be destroyed through this N-degron pathway. When considered in context with previous data indicating that N-degrons are conserved in toxic caspase-generated fragments, this data points to intricate crosstalk between caspase activity and N-degron pathways.\textsuperscript{36}

Considering the importance of limited proteolysis to cell health, it is not surprising that pathogens employ limited proteolysis to hijack degradation pathways. It was recently demonstrated that the mechanism of cell killing for Anthrax Lethal Factor (LF), a secreted virulence factor from \textit{Bacillus anthracis}, depends on the N-end rule E3 ubiquitin ligases UBR2 and UBR4. It was known that cell killing by LF depended on the activation of the NLRP1B inflammasome.\textsuperscript{37,38} LF is a metalloprotease that is trafficked to the cytosol where it cleaves the N-terminus of NLRPB1 between Lys44 and Leu45, creating a type-II Arg/N-degron. NLRP1B exists as an autoinhibited heterodimer that arises from autoproteolysis in the FIND domain of the protein. The unmasking of Leu45 at the N-terminus by LF destabilizes the autoinhibitory fragment, leading to the oligomerization of the NLRP1B CARD (CAspase-Recruitment Domain) domain containing fragment, the activation of caspase-1, and pyroptosis. In essence, an effector reprograms the stability of a host protein via proteolysis to elicit a cellular response that favors pathogenesis. Given that many pathogens have evolved proteolytic effectors, it
seems probable that reshaping host protein stability in this manner is common at the host-pathogen interface.

4. The IAP N-end-binding E3s regulate caspase activation and activity. As proteolysis is irreversible, the regulation of protease activity is often redundant and multi-faceted. Many proteases are produced in a zymogen form to restrict the context of their activity. These proteases must mature into an active enzyme prior to efficient cleavage of their biological substrates. This maturation usually requires proteolysis such that the active enzyme contains N and/or C termini not present in the zymogen. The caspase family of cystine proteases are well understood in this respect. The inflammatory caspases (-1, -4, and -5 in humans) and initiator caspases (-2, -8, -9, and -10), must dimerize through scaffolding interactions with the CARD or death effector domains (DED) present at their N-terminus prior to maturation. This scaffolding is mediated by signaling events that cue the assembly of multicomponent complexes (i.e. inflammasome, apoptosome, DISC complex, etc.) which facilitate dimerization through these accessory domains. Upon dimerization, autoproteolysis leads to the release of the mature enzyme bearing new N and C termini on both the large and small subunits. The executioner caspases (-3, -6, and -7) exist as constitutive homodimers in their zymogen form. Maturation is canonically induced by the activity of initiator or inflammatory caspases and subsequent maturation via the proteolytic excision of the pro-domain and the inter-subunit linker. The maturation of caspases -3, -7, and -9 yields a new N-terminus on the small subunit that engages in a regulatory interaction with the X-linked inhibitor of apoptosis
(XIAP). This E3 ubiquitin ligase consists of three Baculovirus IAP Repeat (BIR) domains and a RING domain which mediates the formation of a homodimer. BIR1 is devoid of caspase binding activity and seems to support dimerization. BIR2 mediates the inhibition of the active forms of caspases -3 and -7. This activity depends on the capacity of the peptide linker region between BIR1 and BIR2 to occupy the active site of the caspase in a reverse orientation relative to substrate. This orients Asp148 so that it occupies the S1 subsite. This binding mode is supported by the ability of the BIR2 domain to bind to the new N-terminus of the small subunit generated during maturation (Figure 1.3A). Coordination of the N-terminus is mediated by Asp214 and Glu219 in the BIR2 Domain. Structural and mutagenesis studies have confirmed that both the end-binding activity and substrate mimicry are important for BIR2 inhibition of caspase-3 and caspase-7. The BIR3 domain mediates the inhibition of caspase-9 in a similar fashion, with the distinction that binding of the N-terminus of the small subunit is sufficient to inhibit activity (Figure 1.3B). Although the role of the RING domain and ubiquitin ligase activity of XIAP in regulating caspase degradation remains unclear, it is clear that the end-binding activity of XIAP is critical for directing cell fate.

XIAP deters the activity of caspases that may lead to undesirable programmed cell death. Accordingly, this activity is neutralized during apoptosis. This is achieved by the generation of nascent N-termini that compete with caspases for the N-end binding sites in XIAP. IAP Binding Motifs (IBMs) are present at the nascent N-termini of proteins released from the mitochondria during apoptosis. These proteins include Smac, HTRA2, and PGAM5 among others. The most extensively characterized of these proteins is Smac (Second mitochondrial activator of caspases). In a healthy
mitochondrion, Smac is tethered to the inter-membrane by an N-terminal transmembrane anchor. Upon outer-membrane permeabilization, Smac is cleaved by the integral-membrane metalloprotease PARL.\textsuperscript{48} PARL cleavage imparts SMAC with a nascent N-terminus (NH\textsubscript{2}-AVPI) capable of binding BIR2 and BIR3 in XIAP (Figure 1.3C and Figure 1.3D).\textsuperscript{49,50} Smac dimerizes upon release and diffuses into the cytosol where it sequesters XIAP and potentiates the propagation of caspase signaling cascades. In total it is the interplay between IBM motif generation and the binding partners they engage with that tunes the probability of a cell surviving or undergoing programmed cell death. The Smac-IAP axis has proven amenable to the generation of peptidomimetic chemical probes with activity in cells and \textit{in vivo}.\textsuperscript{51} Medicinal chemistry efforts have yielded multiple scaffolds with degrees of selectivity for different BIR domains among the IAPs with the goal of potentiating cell death for the treatment of cancer.\textsuperscript{52} The coupling of PARL and XIAP function by Smac serves as an effective exemplar for the molecular logic that underpins substrate-linked functional interactions between proteases and end-binding E3 ligases. The competition between caspases and cleaved Smac for XIAP effectively links the two signaling modules and creates a higher-order signaling network (Figure 1.3E).

5. C-degrons: new opportunities in protease signaling. Until recently, knowledge of end-binding E3 ubiquitin ligases was largely restricted to E3s with the capacity to recognize an N-terminus. In 2018 Koren \textit{et al} and Lin \textit{et al} used global protein-stability (GPS) profiling and CRISPER-based functional genomics screens to systematically identify C-degrons and match said degrons to the components of the UPS that
mediated their degradation.\textsuperscript{53–55} This global cellular profiling effort led to the identification of eleven C-end binding E3 ubiquitin ligases with unique specificities that mediate degradation. The bulk of the E3 ligases identified were of Cullin-RING E3 ligases (CRLs e.g. Cul2 or Cul4) bound to adaptors that mediated C-end binding. All identified adaptors contained repeats of protein domains that form solenoids (Kelch, Ankyrin, Tetratricopeptide, WD40, and Armadillo-like repeats). Structures of the Kelch domain adaptor KLHDC2 bound to C-degron sequences were solved by Rusnac \textit{et al.}\textsuperscript{56} These studies identified the molecular determinants of C-end binding in this adaptor to a Gly-Gly sequence at the C-terminus (\textbf{Figure 1.4A and Figure 1.4B}). One such Gly-Gly C-end degron in USP1 arises from the same autoproteolytic cleavage event that also generates a tertiary Type-1 N degron (via an N-terminal Asn residue). This coincidence of N and C degron pathways on a single cleavage site highlights the amount of information that can be irreversibly unmasked by a single proteolytic event (\textbf{Figure 1.4C}).

The regulatory logic of C-degron pathways is likely quite different from N-degron pathways. Free C-termini are far more abundant than free N-termini in the cell. While C-terminal modifications including methylation, amination and adenylation occur, they are generally transient or limited to specific biochemical or cellular context.\textsuperscript{57} Additionally, the specificity of soluble proteases is often dominated by the non-prime substrate specificities. This makes the identification of co-evolved protease and C-end binding E3 ligase specificities more likely, even absent knowledge of physiological substrates. Further work is required to understand the molecular and cellular determinants of C-
degron pathway function and to what degree they are shaped by proteolysis and other PTMs.

6. Crosstalk between the C-end binding E3 ligase CHIP and caspases. Few E3 ubiquitin ligases containing C-end TREs had been identified prior to the discovery of the C-degron pathways mentioned above. The carboxy terminus of Hsp70 interacting protein (CHIP) was among few with annotated specificity and function. CHIP is a Ubox domain-containing E3 ubiquitin ligase that forms a homodimer.\textsuperscript{58,59} In addition to the Ubox domain CHIP also contains a tetratricopeptide repeat (TPR) domain at its N terminus. The TPR domain binds to a conserved motif at the C-terminus of cytosolic Hsp70s (IEEVD-CO\textsubscript{2}H) and Hsp90s (MEEVD-CO\textsubscript{2}H) in eukaryotes. (Figure 1.5A) This PPI positions CHIP to facilitate transfer of ubiquitin from E2 to a chaperone-bound client. (Figure 1.5B) Specificity for the terminal Asp residue is achieved through a conserved carboxylate clamp consisting of two Lys residues that form a bidentate interaction with the negatively charged carboxylates.\textsuperscript{60,61} In addition to the coordination of the C-terminal Asp at P1, interactions with the Ile or Met at P5 and the Val at P2 are also critical for chaperone binding.\textsuperscript{62} CHIP has been attributed with a wide array of cellular functions. The presumption in a majority of these cases has been that CHIP is functioning in cooperation with cytosolic chaperones. While the degradation of multiple chaperone clients has been attributed to CHIP function, a wholistic model of CHIP regulation has remained elusive. For example, it has been shown that mono-ubiquitination of CHIP is essential for it to be “active” and degrade aggregation-prone mutants of Huntington.\textsuperscript{63–65} The absence of monoubiquitinated CHIP in neurons is
thought to mediate neuron-specific selective vulnerability to Huntington toxicity relative to astrocytes which contain a pool of monoubiquitinated CHIP. Additionally, it has been shown that in the context of cellular stress CHIP can repartition to the membrane or mitochondria.\textsuperscript{66,67} In short it seems that the regulation of CHIP activity is far more complex than the model wherein CHIP binds to the C-terminus of chaperones and potentiates the degradation of their clients by ubiquitination.

We recently reported that the CHIP TPR domain was capable of interacting with a far broader range of C-termini than previously thought.\textsuperscript{68} Briefly we used a positional scanning synthetic combinatorial library (PSSCL) of peptides to assign the subsite specificity of the CHIP TPR domain at the four positions preceding the C-terminal Asp. This approach identified that in addition to the specificities mentioned above, the CHIP TPR could accommodate Leu at P5, preferred aromatic amino acids at P4, preferred bulky residues at P3 and preferred Pro at P2. When the optimal amino acids for each position were combined (CHIPOpt, Ac-LWWPD), the resulting peptide exhibited superior binding properties to the best known sequence (Hsp70, Ac-IEEVD). (Figure 1.5C and Figure 1.5D) We also determined that we could use the specificity profile to accurately predict proteome derived interactors of CHIP. Given that caspases have a strong preference for P1 Asp residues, resulting in proteolytic product with C-terminal Asp, we developed an algorithm to predict CHIP binding sequences that would arise from caspase activity from among the nearly 500,000 Asp residues in the reference human proteome. These efforts led to the identification of over 2700 putative CHIP binding sequences likely to arise from caspase activity. We subsequently validated cleavage-site sequences in the microtubule associated protein tau (tau cleaved at Asp
421, tau^{D421}) and active caspase-6 (caspase-6 cleaved at Asp179, caspase-6^{D179}), both of which accumulate in Alzheimer’s disease. In vitro characterization of the interactions demonstrated that CHIP could rapidly ubiquitinate tau^{D421} in the absence of a chaperone. Additionally, binding to mature active caspase-6^{D179} mediated ubiquitin-independent inhibition of caspase-6 activity akin to XIAP inhibition of caspases-3, -7 and -9. In total these data suggest a complex interplay between protein homeostasis networks and caspases mediated by the interactions of CHIP with chaperones as well as caspases, and caspase substrates. The inhibition of caspase-6^{D179} activity establishes a network architecture wherein the accumulation of CHIP binding sequences may potentiate cellular capase-6 via a feed-forward mechanism. (Figure 1.5E)

It is interesting to consider the co-evolution of caspases and CHIP in the context of multicellularity. The TPR family of co-chaperones and the corresponding C-terminal interaction motif in Hsp70s and 90s are absent in prokaryotes but present in yeast. However, there is no yeast ortholog of CHIP. Yeast are also devoid of caspases or similar proteases with specificity for a P1 Asp that mediate programmed cell death. It is enticing to consider that CHIP and caspases evolved together in higher organisms to facilitate the transfer of information between caspases and protein homeostasis networks. Mutual antagonism between caspases and protein homeostasis has been established. It seems likely that the recruitment of proteostasis network components to caspase substrates (i.e. the association of CHIP function to caspase-6^{D179} and tau^{D421}) and the concomitant disruption of critical PPIs in protein homeostasis networks (i.e. the disruption of CHIP:Hsp70/90 complexes by caspase substrates) will contribute to this
antagonism. Further work is necessary to ascertain the scope of CHIP activity as a caspase inhibitor, identify additional caspase substrates that bind to CHIP, and determine if this mechanism is common among the TPR co-chaperones. The development of chemical probes based on the CHIPOpt sequence will likely aid this endeavor and allow the probing of CHIP function in a manner inaccessible to genetic perturbations.

7. Implications of linking degradation to end binding in the proteome

End-binding E3 ubiquitin ligase signaling is intimately intertwined with protease signaling wherein the activity of the E3 primes the emerging function of nascent termini. Additionally, some of these E3s (XIAP and CHIP) are capable of directly inhibiting the activity of proteases. While at this point these relationships seems to be enriched for the caspase family of proteases it is difficult to assess if this enrichment is representative or simply a biproduct of the robust phenotypes and available reagents for these proteases. Further work is essential to building a comprehensive understanding of how endo- and exo- proteolysis in the cell are linked to protein stability through end-binding E3s. These efforts will likely integrate modern proteomic methods for detecting cellular protein termini, activity-based protein profiling (ABPP) methods, as well as cellular and biochemical tools for assessing protease and end-binding E3 ligase specificity.\textsuperscript{70–75}

Notwithstanding any bias in our existing knowledge of these systems it is has become clear that protein termini, explicit or nascent, are a hotspot for regulation. The continued development of high-quality probes for individual components of these regulatory networks will be critical for elaborating useful models of these aspects of cell signaling.
However, these probes need not be limited to small molecules such as Smac mimetics or caspase-inhibitors. Drawing inspiration from anthrax LF one could imagine harnessing an exogenous protease to selectively alter the stability of proteins by changing the composition of their termini. In total it would seem that being mindful of the ends or proteins of interest is likely an important component of understanding biological function.
References


(8) Tough, D. F.; Tak, P. P.; Tarakhovsky, A.; Prinjha, R. K. Epigenetic Drug


31 (7), 1630–1643. https://doi.org/10.1038/emboj.2012.42.


https://doi.org/10.1002/pmic.201000397.


https://doi.org/10.1021/bi1005464.


(42) Lin, S. C.; Huang, Y.; Lo, Y. C.; Lu, M.; Wu, H. Crystal Structure of the BIR1


(49) Wu, G.; Chai, J.; Suber, T. L.; Wu, J.-W.; Du, C.; Wang, X.; Shi, Y. Structural

(50) Liu, Z.; Sun, C.; Olejniczak, E. T.; Meadows, R. P.; Betz, S. F.; Oost, T.; Herrmann, J.; Wu, J. C.; Fesik, S. W.; Division, P. D.; et al. Structural Basis for Binding of Smac / DIABLO to the XIAP BIR3 Domain. 2000, 408 (December).


Assimon, V. A.; Southworth, D. R.; Gestwicki, J. E. Specific Binding of
Tetratricopeptide Repeat Proteins to Heat Shock Protein 70 (Hsp70) and Heat Shock Protein 90 (Hsp90) Is Regulated by Affinity and Phosphorylation.


https://doi.org/10.1021/acs.biochem.5b00801.


https://doi.org/10.1073/pnas.1710549114.


Varshavsky, A. Life Sciences Breakthrough Prize Lecture

https://www.youtube.com/watch?v=JfXMRFxS3kI.
Figure 1.1. End binding E3 ubiquitin ligases regulate protease product stability to enable signaling. A, The balance of activity between PTM placing (E) and PTM removing (E*) enzymes controls the partitioning of function between modified and unmodified species for most PTMs. The irreversibility of proteolysis requires a different regulatory mechanism. B, End-binding E3 ubiquitin ligases can degrade the products of proteolysis and control proteolytic signal intensity and duration. C, E3 ubiquitin ligases provide specificity to the ubiquitin proteasome system by binding to degrons (yellow circles) and scaffolding substrate ubiquitination and subsequent delivery to the proteasome. D, When a protease (P) cleaves a substrate (S) that is subsequently recognized by and end-binding E3 ubiquitin ligase (E3), a functional interaction is created between protease and E3. These elements form a minimal signaling network. E, The strength of the functional link between a protease and an end-binding E3 is related to the overlap in the specificity between the protease active site and the terminal...
recognition element (TRE) in the E3. This links S1-S4 subsites to C-terminal TREs and S1'-S4' to N-terminal TREs.
Figure 1.2. Specific determinants of N-end binding link protease and E3 function in signaling pathways. A, The X-ray crystal structure (PDB 3NY3, left panel) of the UBR N-end binding TRE from UBR2 (surface, grey) bound to a tetrapeptide (NH₂-RIFS-CO₂H, sticks, yellow). Specificity for the N-terminal Arg is mediated by Asp150 and Asp153 (middle panel). UBR binds to the N-degron at the nascent N-terminus of a self-cleavage product of USP1 after the N-terminal Gln is deamidated and argenylated. B, The X-ray crystal structure (PDB 3O2H, left panel) of the E. coli ClpS (surface, grey) bound to a tetrapeptide (NH₂-PGLW-CO₂H, sticks, yellow). Specificity for the N-terminal Pro is mediated by coordination of the proline amine by Asn34, Asp35, and His66 (middle panel). A homologous domain in human UBR 1 and 2
is thought to direct the binding of Type II N-degrons. Generation of a Type II N-degron by Anthrax LF activity mediates NLRP1b inflammasome activation and cell death. C, The X-ray crystal structure (PDB 6CDC, left panel) of the GID4 N-end binding TRE from GID E3 ubiquitin ligase complex (surface, grey) bound to a tetrapeptide (NH$_2$-PGLW-CO$_2$H, sticks, yellow). Specificity for the N-terminal Pro is mediated by coordination of the proline amine by Glu237 (not shown) and backbone contacts with Gln132, Ser253, Ser278, Gln282 (middle panel). GID4 binds to the N-degron at the nascent N-terminus of a methionine aminopeptidase cleavage product of gluconeogenic enzymes including Fbp1 to mediate their rapid decay.
Figure 1.3. N-end binding is critical for the regulation of apoptosis. A, The X-ray crystal structure (PDB 1I3O) of the BIR2 domain of XIAP (grey, surface and cartoon) bound to caspase-3 dimer (yellow, surface and cartoon). BIR2 inhibition of caspase-3 activity depends on its ability to occupy the active site (blue square), and bind the nascent N-terminal IBM motif generated by maturation of the enzyme (red square). B, The X-ray crystal structure (PDB 1NW9) of the BIR3 domain of XIAP (grey, cartoon) bound to caspase-9 monomer. BIR3 inhibition of caspase-9 depends on its ability to bind the nascent N-terminal IBM motif generated by maturation of the enzyme (red square) but not the active site (blue square). C, The X-ray crystal structure (PDB 1G37) of Smac dimer (Second mitochondrial activator of caspases, yellow, cartoon and surface) bound to BIR3 domain of XIAP (grey, cartoon and surface). Cleavage of Smac by PARL and release form the mitochondria allows the IBM motif generated by cleavage to sequester XIAP and relieve caspase inhibition. D, The N-terminal Ala in the IBM motif of Smac (NH₂-AVPI) is specifically coordinated by Asp314 and Gln319. E, Complementary specificities of caspase-9, PARL, and XIAP establish a network architecture that encodes the signaling function of this system.
Figure 1.4. C-end binding adaptors to Cullin E3 ubiquitin ligases link degradation and non-prime protease specificity. A, The X-ray crystal structure (PDB 6DO5) the Cul2 adaptor KLHDC2 (grey, surface) bound to a C-terminal peptide derived from a self-cleavage event in USP1 (NH2-IGLLGG-CO2H, yellow sticks, red square). B, The C-terminal Gly is coordinated by Arg236, Arg241, and Ser269. C, Self-cleavage by USP1 generates coincident N and C degrons with the potential to regulate USP1 function.
Figure 1.5. C-end binding by the E3 ubiquitin ligase CHIP links caspases and protein homeostasis. A, Canonical CHIP (grey) function stems from an ability to bind to ubiquitin charged E2s (black) and cytosolic chaperones (i.e Hsp70s, red) to mediate ubiquitination of chaperone bound client (white). B, This function is rooted in the ability of CHIP (PDB 2C12, grey, cartoon and surface) to bind a conserved C-terminal motif (IEEVD-H in Hsp70s and MEEVD-H in Hsp90s, yellow, spheres) through its TPR domain (red square) and bring it into the proximity of E2 bound to the Ubox domain (PDB 20XQ, black, cartoon). C, Global profiling of the TPR domain specificity for proteinogenic amino acids expanded probable binders beyond canonical chaperone sequences. D, The C-terminal Asp of an optimized peptide bound to the CHIP TPR (PDB 6NS) is coordinated by Lys 30, Asn34, Asn65 and Lys95. E, Shared specificity with caspases mediates direct binding and inhibition of active caspase-6 (caspase-6D179) and binding and ubiquitination of caspase-cleaved tau (tauD421).
Chapter 2

Specificity for latent C-termini links the E3 ubiquitin ligase CHIP to caspases

Abstract. Protein-protein interactions between E3 ubiquitin ligases and protein termini shape the proteome. These interactions are sensitive to proteolysis, which alters the ensemble of cellular N- and C-termini. We have identified a mechanism wherein caspase activity reveals latent C-termini that are recognized by the E3 ubiquitin ligase CHIP. We expanded established CHIP specificity and identified hundreds of putative CHIP interactions arising from caspase activity. Subsequent validation confirmed that CHIP binds the latent C-termini at tau<sup>D421</sup> and caspase-6<sup>D179</sup> in vitro and in cells. CHIP binding to tau<sup>D421</sup>, but not tau<sup>FL</sup>, promotes its ubiquitination, while binding to caspase-6<sup>D179</sup> mediates ubiquitin-independent inhibition. Given that caspase activity generates tau<sup>D421</sup> in Alzheimer’s Disease (AD), these results suggest a concise model for CHIP regulation of tau homeostasis. Indeed, we find that a loss of CHIP expression in AD coincides with the accumulation of tau<sup>D421</sup> and caspase-6<sup>D179</sup>. These results illustrate an unanticipated link between caspases and protein homeostasis.

1. Introduction. An important subset of protein-protein interactions (PPIs) involve the recognition of a free amino (N) or carboxy (C) terminus by a binding partner. This type of PPI is sensitive to proteolysis, which irreversibly modifies the chemical composition of a protein’s terminus. Indeed, proteolysis can initiate binding events at protein termini and subsequently activate signaling cascades, including protease-activated receptor (PAR) signaling, apoptosis, and the unfolded protein response. Similar recognition events are important in maintenance of protein homeostasis. E3 ubiquitin ligases provide specificity to the ubiquitin proteasome system (UPS) by linking substrate-recognition and ubiquitination. A subset of E3 ubiquitin
ligase recognition domains are selective for protein termini; for example, the E3 ubiquitin ligases of the N-end rule pathways (e.g. the UBR E3 ligases) recognize N-termini bearing specific amino acids (N-degrons), prioritizing them for destruction via the UPS. In the N-end rule, proteolysis by methionine aminopeptidases plays a critical gatekeeping role by removing stabilizing N-terminal methionine residues to reveal UBR binding motifs. Analogous recognition of C-termini by adaptor proteins to Cullin-RING E3 ubiquitin ligases has recently been shown to couple C-terminal recognition to turnover, perhaps through a protease-dependent pathway. Additionally, E3 ubiquitin ligases that recognize protein termini are central to the regulation of cell survival. Specifically, XIAP is a RING E3 ubiquitin ligase that inhibits the cystine proteases, caspase-3, -7 and -9. This inhibition deters the activation of apoptosis in a healthy cell. However, upon activation of apoptotic pathways, second mitochondrial activator of caspases (SMAC) is cleaved by the mitochondrial protease, PARL, allowing the free N-terminus of SMAC to sequester the inhibitory domain in XIAP. Through these mechanisms, proteases enable recognition of latent termini to regulate critical cellular processes.

Here, we report a protease-coupled recognition event wherein caspase activity reveals latent C-termini that are bound by carboxy-terminus of Hsp70-interacting protein (CHIP). CHIP is a homodimeric E3 ubiquitin ligase that contains a Ubox domain and a tetratricopeptide repeat (TPR) domain. The TPR domain was thought to be selective for the (I/M)EEVD motif present at the C-terminus of cytosolic members of the heat shock protein 70 (Hsp70) and 90 (Hsp90) families. These interactions favor the ubiquitination and turnover of both the chaperones and chaperone-bound substrates.
Moreover, similar TPR domains are found in other Hsp70/Hsp90 co-chaperones such that competition for chaperone is thought to direct client fate. Here, we find that CHIP’s TPR domain has a broader specificity than previously appreciated, so that hundreds of latent C-termini may compete with Hsp70/Hsp90 in the context of caspase activation. We show that caspase cleavage of microtubule-associated protein tau (MAPT/tau) generates a high-affinity CHIP ligand. Strikingly, CHIP mediates *in vitro* poly-ubiquitination of cleaved tau (tau\(^{D421}\)), but not full-length tau (tau\(^{FL}\)), without the need for chaperone. In addition, we find that cleavage of caspase-6 during its maturation generates a strong CHIP recognition motif, which mediates potent ubiquitin-independent inhibition of caspase-6\(^{D179}\). As both tau and caspase-6 cleavage are linked to AD, these findings suggested a possible feedback mechanism that links tau proteostasis and cell survival. In support of this model, we find that CHIP levels are decreased during progression of AD in patients, concomitant with an increase in cleaved tau\(^{D421}\) and active caspase-6\(^{D179}\). These results point to a much broader role for CHIP than its classical collaboration with Hsp70s/Hsp90s.

2. **CHIP TPR domain specificity extends beyond chaperones.** Co-crystal structures have shown that Hsp70’s IEEVD motif binds to a groove in CHIP’s TPR domain, such that the two carboxylates of the C-terminal Asp form a bidentate interaction with CHIP’s carboxylate clamp formed by residues Lys30 and Lys95 (**Figure 2.1A**)\(^{23}\). This interaction is common among TPR co-chaperones (**Figure 2.1B**). In addition, previous mutagenesis studies have identified the P5 (Ile in Hsp70s or Met in Hsp90s), P2 (Val), and P1 (Asp) as the critical positions for affinity and specificity\(^{24}\). However, the global
binding specificity of CHIP's TPR domain has not been explored. To achieve this, we designed a positional scanning synthetic combinatorial library (PSSCL) of acetylated five amino acid peptides. In this library, we fixed the C-terminal amino acid as an Asp and synthesized pools of peptides in which a second position was fixed, while each of the others contained a mixture of the twenty proteinogenic amino acids (see Methods). The resulting library consisted of eighty pools of eight thousand peptide sequences, each with two fixed positions (Figure 2.1C). To measure binding, we used two approaches. In the first, we determined the change in thermal stability ($T_m$) of CHIP in the presence of peptide pools by differential scanning fluorimetry (DSF) (Figure S2.1). We then confirmed these findings in fluorescence polarization (FP) experiments by measuring the capacity of each peptide pool to displace FP tracers derived from either the Hsc70 (HSPA8) or Hsp90α (HSP90AA) sequences (Figure S2.2). These experiments reproduced the known determinants of chaperone binding to CHIP's TPR domain and the results from the three biochemical platforms correlated well with each other (Figure 2.1D). In addition to the established specificities mentioned above, we identified unexpected preferences for Leu at P5, aromatic amino acids at P4, bulky residues at P3 and Pro at P2. Interestingly, we found that the best-known CHIP TPR binder (Hsp70's IEEVD) was sub-optimal at all four positions queried. Using this information, we synthesized an optimized FP tracer (CHIPOpt; FITC-Ahx-LWWPD) and found that it bound CHIP with an apparent p$K_a$ of 7.8 ± 0.1 (Figure 2.1E). This affinity is significantly higher than that measured for tracers based on Hsp70s (FITC-Ahx-IEEVD; p$K_a$ of 6.8 ± 0.1) or Hsp90s (FITC-Ahx-MEEVD; p$K_a$ of 6.1 ± 0.1). An alanine scan of the CHIPOpt sequence in competitive FP experiments established that the positional
hierarchy of binding contributions was P2 > P1 > P4 > P5 > P3 based on perturbations to the pKᵢ of the peptide (Figure 2.1F and Figure S2.3). This data suggests that the determinants of CHIP binding extend well beyond the established I/MEEVD motifs.

3. CHIPOpt engages distinct interactions with CHIP. To better understand the molecular determinants of CHIP specificity, we solved X-ray crystal structures of the human CHIP TPR domain bound to acetylated 5mer peptides corresponding to the Hsp70 sequence (Ac-IEEVD, PDB 6EFK) and CHIPOpt (Ac-LWWPD, PDB 6NSV) at 1.5Å and 1.3Å respectively (Figure 2.2A, Figure S2.4A, B, C, Table S2.1). It is worth noting that the CHIPOpt structure included density at the crystallographic interface of the two peptide-bound TPR domains in the asymmetric unit, which is not present in the Hsp70 peptide-bound structure (Figure S2.4C). This density appeared as a linear chain that wrapped around the epsilon nitrogen of Lys72 in each TPR domain. Given this typical host-guest orientation and the presence of polyethylene glycol (PEG) in the crystallization conditions we modeled the densities as two PEG-6 molecules.

As mentioned above, CHIP belongs to a family of proteins that use TPR domains to bind Hsp70 and Hsp90 C-termini. However, CHIP’s TPR domain is unique among these proteins because it contains a discrete hydrophobic pocket formed by an extra turn in helix 6 and an elongated linker between helices 6 and 7, which orient residues Phe98, Phe131 and Ile135 around P5 in the bound peptides (Figure 2.2B, Fig S2.4D). This arrangement favors a kinked conformation of bound peptides, which contrasts with the linear arrangement of similar peptides in the TPR domain of HOP (Figure S2.4E)²⁷. Consistent with this idea, we found that both Hsp70-derived peptide and CHIPOpt
peptide had the expected kinked binding mode. A closer examination of the relative orientation of these peptides helped explain why CHIPOpt was a superior ligand. To further that effort, we also performed a computational analysis of the bound peptide in Rosetta. Briefly, using the Hsp70-bound structure as a template, we computationally estimated perturbations to the free energy of folding ($\Delta G_{\text{fold}}$) for the protein-peptide complex upon saturation mutagenesis of the peptide. By this analysis, the amino acids that provided the greatest increases in affinity, such as P4 aromatics and P2 proline, agreed between the experimental results and computational predictions (Table S2.2). Moreover, direct comparison of the P2 position in the Hsp70 and CHIPOpt crystal structures revealed a +19° perturbation to the $\phi$ dihedral angle, suggesting that the conformational rigidity of the proline may enhance favorable contacts elsewhere in the ligand and reduce the entropic cost of binding (Figure 2.2C, P2-V & P2-P). We also observed that the P4 Trp of CHIPOpt packs on a hydrophobic shelf formed by F98 and F99, which is not engaged by the corresponding P4 Glu in the Hsp70-bound structure. Given that F99 bridges the P4 and P5 binding sites, it is likely that the P4 aromatic also further reinforces the kinked peptide binding mode (Figure 2.2C, P4-E & P4-W). Another striking difference between the Hsp70 and CHIPOpt structures is an apparent shift in hydrogen bonding at the peptide N-terminus. In our Hsp70 structure and all published CHIP-peptide co-crystal structures, D134 in the TPR domain forms a hydrogen bond with the amide nitrogen of the P5 amino acid (Figure 2.2D, grey). However, in the CHIPOpt structure, this hydrogen bond does not form. Instead, the amide oxygen of the acetyl group forms a hydrogen bond with K95 (Figure 2.2D, yellow). This shift seems to support an exaggerated kink in the CHIPOpt conformation. In total, the structural and
computational data suggests that the enhanced affinity of CHIPOpt likely arises from optimized interactions at P2, P4 and the N-terminus. Together, these results indicate that previously unknown, high-affinity interactions with CHIP exist within the human proteome.

4. **Proteome-wide prediction of CHIP interactions.** We next used the expanded specificity of the CHIP TPR domain to generate an algorithm for predicting peptide sequences that would bind to CHIP (termed CHIPScore). Briefly, PSSCL data from the DSF dataset were normalized to the peptide pool with the highest signal (P2 Pro) (Figure S2.5A). Then, each value for an amino acid at a given position was summed so that a CHIPScore is equal to the sum of the normalized values for P5, P4, P3 and P2 (Figure 2.3A). To validate the scoring function, we compared the CHIPScore to the measured pKᵢ values of the CHIPOpt peptide and its variants as well as Hsp70 and Hsp90 sequences. In this analysis, CHIPScore correlated well with the experimental pKᵢ values (R² = 0.86) (Figure 2.3B). Then, we used CHIPScore to search for binders in the human proteome. First, of the 20,246 protein sequences in the reference proteome, we found that 973 have a C-terminal Asp residue. Among these, only cytosolic Hsp70s and Hsp90s terminate in the canonical (I/M)EEVD motif. However, using CHIPScore, we found an additional eight predicted interactions whose: (i) score exceeded Hsp90 (MEEVD CHIPScore = 1.79) and (ii) were expressed in the cytosol or nucleus (Figure 2.3C, Figure S2.5B, Table S2.3). Two proteins on this list (NADE and TXLNG) had previously been identified in proteomic studies as potential CHIP binders (Taipale et al.), supporting the predictive power of the scoring function.
This finding suggested that only a small number of possible interactions could be identified in the intact C-terminome. However, taking inspiration from the central role of proteolysis in revealing N-termini, we expanded our search space to include latent C-termini that are only unmasked by the activity of proteases. Here, “latent” is used to describe a new C-terminus that can be produced by proteolytic activity. Of the human proteases with well-annotated specificity, the caspase family of cysteine proteases was of particular interest because they exhibit a strong preference for Asp at the P1 position, such that their activity yields a new C-terminus ending in this residue\(^{29,30}\). Moreover, extensive experimental and bioinformatic efforts have been previously undertaken to identify and predict caspase substrates. Specifically, 1651 unique P1 Asp caspase cleavage sites have been experimentally validated by N-terminomics\(^{31}\). When we searched this dataset with CHIPScore, we identified 84 C-termini that are likely to bind better than Hsp90 to CHIP (Figure S2.5C, Table S2.4). Encouraged by these results, we expanded the search to include a proteome-wide dataset of predicted caspase cleavage sites, which were generated from a support vector machine (SVM) scoring function\(^{32}\). To set a threshold SVM score, we compared the distribution of 493,321 proteome-derived SVM scores to the 1,651 experimentally validated caspase cleavage sites and set a cutoff that would include 75% of the validated sites (SVM=-0.464) (Figure S2.5D, E). This process yielded 2,757 candidates with latent C-termini that are predicted to be high affinity ligands for CHIP (Figure 2.3D, Figure S2.5F, Table S2.5). We anticipate that this list includes many new PPIs between CHIP and caspase substrates, perhaps providing a rich source of regulatory interactions.
To validate a subset of this data, we curated the lists of predicted candidates from the explicit and latent C-terminome datasets. This process directed our attention to tau (cleaved at Asp421; tau$^{D421}$) and an auto-proteolytic activation site of the inter-domain linker in the caspase-6 zymogen (cleaved at Asp179; caspase-6$^{D179}$) \(^{33}\). These sites were of interest because tau$^{D421}$ and caspase-6$^{D179}$ have both been shown to accumulate in the brains of AD patients \(^{34–36}\). However, the mechanism underlying the accumulation of these specific fragments has been unclear. Moreover, CHIP had previously been linked to the clearance of tau$^{D421}$ in cellular and animal models \(^{37–39}\). Thus, we considered a model in which CHIP might serve as a key regulator of both tau homeostasis and caspase-6 activity through recognition of latent C-termini in each protein. Before proceeding, we confirmed that peptides based on tau$^{D421}$ and caspase-6$^{D179}$ indeed bound (pIC$_{50}$ values $7.1 \pm 0.1$ and $5.9 \pm 0.1$ respectively) to CHIP’s TPR domain using competitive FP experiments (Figure 2.3E and Figure S2.5G).

5. Caspase cleavage of tau at D$^{421}$ recruits CHIP. In healthy neurons, tau binds to microtubules through its microtubule binding repeats (MBRs) \(^{40}\). However, in AD, tau is known to dissociate from microtubules to form neurofibrillary tangles (NFTs) \(^{41}\). In the NFTs, tau is often post-translationally modified, including by proteolytic cleavage \(^{42}\). Specifically, tau contains a number of validated caspase cleavage sites that accumulate during disease progression (Figure 2.4A) \(^{43}\). Our CHIPScore predictions suggested that the D421 site, in particular, would create a high affinity CHIP-binding site. To explore this idea, we first determined whether CHIP selectively bound the tau$^{D421}$ epitope over others. Using competitive FP experiments, we screened peptides corresponding to the
reported caspase cleavage sites, establishing that only the peptide from tau\textsuperscript{D421} interacts with CHIP (Figure S2.6A). To validate binding in the context of a full-length protein, we expressed and purified the 0N4R splice isoform of tau (tau\textsuperscript{FL}) that is expressed in adult neurons, as well as the corresponding tau\textsuperscript{D421} truncation. We found that only tau\textsuperscript{D421}, and not tau\textsuperscript{FL}, could bind to CHIP (pIC\textsubscript{50} of 6.2 ± 0.03); (Figure 2.4B). This specificity was confirmed using synthetic FP tracers corresponding to the ten amino acids at the C-terminus of tau\textsuperscript{FL} and tau\textsuperscript{D421} (Figure 2.6B). Independent analyses by DSF corroborated the selectivity for cleaved tau (Fig S2.6C,D) in the context of full-length protein. Multiple caspases have been implicated in tau processing during AD. To determine which ones might produce the tau\textsuperscript{D421} epitope, we created a corresponding 10 amino acid fluorogenic substrate (see Methods) and monitored its cleavage in the presence of a commercially-available panel of active caspase enzymes. This experiment showed that caspase-3 had the most specific activity for the site in the conditions tested, although caspases-1, -6 and -7 also had activity (Figure S2.6E). Indeed, the addition of active caspase-3 (at t = 0) was able to convert tau\textsuperscript{FL} into a CHIP-binding species in a kinetic FP assay (Figure S2.6F). Next, we wondered whether the direct interaction of CHIP with tau\textsuperscript{D421} might promote ubiquitination \textit{in vitro}. These studies showed that tau\textsuperscript{D421}, but not tau\textsuperscript{FL}, was rapidly (~3 min) poly-ubiquitinated by CHIP in the presence of an E2/E1/Ub/ATP mixture. Use of CHIP\textsuperscript{K30A}, a mutation in the carboxylate clamp, significantly reduced the ubiquitination of tau\textsuperscript{D421} (Figure 2.4C), confirming that the interaction required binding to the TPR domain. This robust ubiquitination was also observed for tau\textsuperscript{FL} processed by caspase-3 \textit{in situ} to generate
tau\textsuperscript{D421} (Figure S2.6G). Thus, when tau\textsuperscript{D421} is produced by caspase activity, it directly binds to the CHIP TPR domain and the protein can be poly-ubiquitinated.

To validate the CHIP-tau\textsuperscript{D421} interaction in a cellular context, we generated stable cell lines expressing a doxycycline inducible GFP-0N4R construct terminating in either the tau\textsuperscript{FL} sequence, the tau\textsuperscript{D421} truncation, or tau\textsuperscript{D421A}. Live-cell fluorescence microscopy confirmed that all three proteins were properly associated with cytoskeletal structures consistent with microtubules (Figure 2.4D). Also, microtubule fractionation showed that each of them was associated with the insoluble pool (Figure 2.4E). In the absence of tau expression, endogenous CHIP was present in the soluble fraction in all three cell lines (lanes 1,3,5). This localization was unperturbed by the expression of either tau\textsuperscript{FL} or tau\textsuperscript{D421A} (lanes 2,6). However, induced expression of GFP-tau\textsuperscript{D421} (lane 4) dramatically partitions CHIP to the insoluble, microtubule fraction. These results suggest that CHIP tightly and specifically associates with tau\textsuperscript{D421} in a cellular context.

6. CHIP specifically inhibits mature caspase-6\textsuperscript{D179}. Active caspase-3, -7 and -9 are inhibited by the E3 ubiquitin ligase XIAP; however, an equivalent endogenous inhibitor for caspase-6 is not known\textsuperscript{44}. Caspase-6 undergoes a maturation process similar to caspase-3 and caspase-7 wherein the cleavage of the pro-domain and inter-subunit linker yield the mature, fully activated enzyme. The fully mature enzyme bears the predicted CHIP binding site, caspase-6\textsuperscript{D179}, at the C-terminus of the large subunit (Figure 2.5A). To understand whether CHIP might act as an endogenous caspase-6 inhibitor, we expressed and purified active enzyme as two polypeptide chains with the large subunit terminating at Asp179 and the small subunit terminating with a 6His tag.
Competition binding experiments confirmed that active caspase-6D179 binds to CHIP (pIC₅₀ = 6.6 ± 0.08) (Figure 2.5B). The specificity of this interaction to C-terminal recognition was confirmed using biolayer interferometry (BLI) and comparing immobilized CHIP to CHIPK30A (Figure S2.7A). We predicted that CHIP would poly-ubiquitinate caspase-6D179 in vitro, as was observed for tauD421. However, similar experiments showed that active caspase-6D179 was not robustly ubiquitinated under these conditions (Figure S2.7B). However, we found that CHIP acted as an inhibitor of caspase-6D179 substrate hydrolysis. Specifically, we used proteolysis of a fluorogenic peptide substrate (Ac-VEID-AMC) to monitor caspase-6D179 activity in the presence of CHIP. A kinetic analysis demonstrated potent inhibition of substrate turnover and a mode of action consistent with competitive inhibition (an increase in Kᵥ) (Figure 2.5C). No significant inhibition was detected with the addition of a CHIPK30A (Figure S2.7C). Thus, CHIP inhibits caspase-6 function, but through a mechanism distinct from XIAP regulation of other caspases.

To determine whether CHIP binds mature caspase-6 in a cellular context, we first generated active caspase-6 in Jurkat cells by treating them with staurosporine (1μM) for two hours. Control and staurosporine treated cell lysates were supplemented with recombinant biotinylated CHIP or biotinylated CHIPK30A. Bound proteins were enriched on streptavidin coated magnetic beads, washed, and eluted. We used antibodies specific for caspase-6D179 (e.g. active) or caspase-6FL (e.g. inactive) to perform Western blot analyses on cell lysates and eluates from pulldowns. We found that caspase-6FL was depleted upon treatment with staurosporine and that neither CHIP or CHIPK30A were capable of pulling it down (Figure 2.5D, left panel). Conversely, active caspase-
6\textsuperscript{D179} was specifically enriched by pulled-down with CHIP but not CHIP\textsuperscript{K30A} (Figure 2.5D, right panel). These experiments demonstrate that the CHIP-caspase-6 interaction is dependent on the maturation of the enzyme and integrity of the CHIP TPR carboxylate clamp.

**7. CHIP loss coincides with neoepitope accumulation in AD.** These results suggest a molecular mechanism to explain why tau\textsuperscript{D421} and active caspase-6\textsuperscript{D179} accumulate in neurons during AD progression. Specifically, we hypothesized that CHIP levels or function might be diminished during disease progression, contributing to increases in active caspase-6\textsuperscript{D179} and failure to clear tau\textsuperscript{D421} (Figure 2.6A). To test this idea, we developed a multiplexed fluorescent immunostaining method to interrogate the levels of tau\textsuperscript{D421}, active caspase-6\textsuperscript{D179}, and CHIP in the hippocampus of fixed human tissue from donors with AD pathology in early, middle and late stages of the disease (Braak I, III, and VI, respectively). Briefly, Braak staging characterizes the spread of neurofibrillary tangles across cortical regions and is widely used to stage AD pathology\textsuperscript{45}. Consistent with the literature, we found that hippocampal Braak VI tissue exhibited the most significant accumulation of tau\textsuperscript{D421} and active caspase-6\textsuperscript{D179}. In addition, we observed a striking loss in CHIP expression in Braak III and VI (Figure 2.6B). Analysis of RNAseq data from the Allen Brain Atlas revealed that CHIP mRNA does not significantly change as a function of Braak stage, ApoE status, or age (Figure S.8A, B and C)\textsuperscript{46}, suggesting that the observed loss of CHIP at Braak III and VI occurs post transcriptionally. Together, these findings support a model in which loss of CHIP expression allows
accumulation of tau\textsuperscript{D421} and potentiates the activity of caspase-6, contributing to the accumulation of both C-termini in AD.

8. **Discussion.** Proteostasis mechanisms that act on protein termini, such as the N-End Rule pathways, are fundamentally important in sculpting the proteome. The work described here identifies a C-terminal recognition process that recruits CHIP to sites of caspase cleavage. By determining the specificity of CHIP’s TPR domain, we identified \(~2,700\) latent C-termini that are predicted to be regulated by this protein. Further work is needed to probe the scope and specific roles of these interactions; however, as an initial proof-of-principle, we were drawn to the observation that both tau\textsuperscript{D421} and caspase-6\textsuperscript{D179} were putative CHIP ligands. These proteins had been linked to AD and to each other, yet it was not clear how or if they might be co-regulated and what this might mean for the disease. Indeed, we found that CHIP binds both tau\textsuperscript{D421} and caspase-6\textsuperscript{D179}, directing the ubiquitination of tau\textsuperscript{D421} and inhibiting caspase activity in a ubiquitin-independent way. Both interactions were of high affinity and could be readily observed in cell-based models. Linking these observations together, histopathology studies in AD brains suggested that loss of CHIP could be a contributor to the accumulation of these latent C-termini in advanced disease. Thus, these findings may generate new ideas for the treatment of AD. Given the clinical failures in this therapeutic area, new targets and mechanisms are of utmost importance.

It is worth noting that the identification of CHIP as an endogenous inhibitor of caspase-6 fits well into existing paradigms of caspase signaling. Caspase activity in a cell is limited at multiple levels\textsuperscript{47}. However, among the broader caspase family, the
regulation of caspase-6 is less understood. Like the other executioner caspases, caspase-6 is capable of auto-activation and exists as a constitutive homodimer \(^{48}\). However caspase-6 is insensitive to XIAP and it has a unique attributes that separate it from the other caspases \(^{49,50}\). Thus, our finding that CHIP inhibits caspase-6 suggests a novel mechanism of caspase regulation and fills a gap in the regulatory framework of these important signaling proteases.

Finally, CHIP was previously thought to act primarily (or even exclusively) as a co-chaperone for cytosolic Hsp70/Hsp90. Our work shows that, in some cases CHIP circumvents the requirement for chaperone and binds directly to protein targets. We found that few C-termini within explicit open reading frames (ORFs) are likely to compete with the Hsp70/Hsp90 interaction, but that caspase cleavage created a large number of possible binding events at latent C-termini. We used minimal affinity cut-off values based on the CHIP-Hsp90 interaction, such that the predicted contacts might be expected to compete with the canonical chaperone-dependent roles. However, given the abundance of chaperones in cells, more work will be needed to understand which of these putative contacts displace chaperones from CHIP. An additional complexity is the roles of the other TPR co-chaperones, such as FKBP51 and PPP5C, which also converge on the shared (I/M)EEVD motifs. It seems likely that caspase activity could broadly remodel chaperone and non-chaperone PPIs, creating new opportunities to rapidly and post-translationally regulate critical signaling nodes in cells.

**Acknowledgments:** This work is dedicated to Ron Raines on the occasion of his 60\(^{th}\) birthday. This work was supported by grants from the Tau Consortium and NIH
Additional support included an NSF GRFP fellowship (to K.A.O), an ARCS Foundation fellowship (to M.R.), institutional grants UL1 TR001872 and K01AG053433 (to P.T.), and a Program for Breakthrough Biomedical Science funded by the Sandler Foundation (to C.S.C.). The authors would like to thank the laboratories of K. Matthew Scaglione (Medical College of Wisconsin) and James A. Wells (University of California San Francisco) for technical support, as well as Sue-Ann Mok for input on the manuscript.


**Competing interests:** None declared

### 10. Methods

**Protein Purification**

Unless otherwise stated all proteins were produced in *E. coli* BL21(DE3) and stored at -80°C.

**CHIP and CHIP K30A** (human, His-tagged) were expressed from a pET151 construct with a N-terminal TEV-cleavable 6His Tag. *E. coli* were grown in terrific broth (TB) at 37°C, induced with 500 μM IPTG in log phase, cooled to 18°C and grown overnight. Cells were harvested, resuspended in binding buffer (50 mM Tris pH 8.0, 10 mM imidazole,
and 500 mM NaCl) supplemented with protease inhibitors, sonicated, clarified, and supernatant was applied to Ni-NTA His-Bind Resin (Novagen). Resin was washed with binding buffer and His wash buffer (50 mM Tris pH 8.0, 30 mM imidazole, and 300 mM NaCl). Protein was eluted from the resin with His elution buffer (50 mM Tris pH 8.0, 300 mM imidazole, and 300 mM NaCl). N-terminal His tag was removed by overnight dialysis with TEV protease at 4° C. Digested material was applied to His-Bind resin to remove cleaved His tag, undigested material, and TEV protease. Protein was further purified by size exclusion chromatography (Superdex 200) in 50 mM HEPES pH 7.4, 10 mM NaCl.

**CHIP TPR Domain** (human, AA 22-154, His-tagged) was expressed from a pMCSG7 construct with an N-terminal TEV-cleavable 6His tag. *E. coli* were grown in TB at 37 °C, induced with 1mM IPTG in log phase, cooled to 24° C and grown overnight. Ni-NTA purification and tag removal were conducted as for FL CHIP. Protein was further purified on a Mono S cation exchange column and stored in 10 mM Tris pH 8.0, 150 mM NaCl, and 2 mM DTT.

**0N4R tau**<sup>FL</sup> and **0N4R tau**<sup>D421</sup> (human) were expressed from a pET28 construct. *E. coli* were grown in TB at 37° C. NaCl (500 mM) and betaine (10 mM) were added to media prior to log phase induction with 200 µM IPTG for 3.5 h at 30 °C. Cells were harvested, lysed by microfluidizer, and lysate was boiled for 20 min. Clarified lysate was dialyzed into 20 mM MES, pH 6.8, 50 mM NaCl, 1 mM EGTA, 1 mM MgCl, 2 mM DTT, 0.1 mM PMSF and purified by cation exchange.
**Caspase-6**\textsuperscript{D179} (human, His-tag) Large subunit (AA 31–179) and small subunit (AA 194–293, 6XHis) in pET24b (+) and pET23b (+) vectors were co-transformed into Rosetta BL2 (DE3) cells. *E. coli* were grown in 2XYT medium at 37°C, and induced in log phase with 0.2 mM IPTG for 20 h at 16 °C. Cells were harvested, resuspended in 100 mM Tris, pH 8.0, 100 mM NaCl, and lysed by microfluidizer. Supernatant was loaded to Ni-NTA His-Bind Resin for 1 h at 4°C. Resin was washed 2X with 100 mM Tris, pH 8.0, 100 mM NaCl, and 1X with 100 mM Tris, pH 8.0, 100 mM NaCl 40 mM imidazole. Protein was eluted with 50 mL 100 mM Tris, pH 8.0, 100 mM NaCl 200 mM imidazole. Further purification was achieved by anion exchange over a linear gradient of 0 – 1 M NaCl in 20 mM Tris pH 8.0.

**UbcH5c** (human) was expressed from a pET2 construct. *E. coli* were grown in TB at 37°C and induced in log phase with 500 μM IPTG overnight at 16 °C. Cells were harvested, washed with PBS, and resuspended in 30 mM MES pH 6.0, 1 mM DTT, 1 μg/mL Leupeptin, 1 μg/mL Aprotinin and 1 mM PSMF, sonicated and clarified. Purification was achieved by cation exchange chromatography (SP Sepharose) followed by size exclusion chromatography (Superdex 200) in 50mM Tris, 50 mM KCl, pH 8.0.

**Ube1** (human, His-tagged) was expressed from a pET21d construct. *E. coli* were grown in TB at 37°C and induced in log phase with 500 μM IPTG overnight at 16 °C. Cells were harvested, washed with PBS, and resuspended in 50 mM Tris, 150 mM NaCl, 1 mM EDTA-NaOH, pH 8.0, 1 mM DTT, 0.1% (w/v) Triton X-100, 1 mM PSMF, 1 protease inhibitor tablet and sonicated. The soluble fraction was loaded onto Ni-NTA His-Bind
Resin for 2 h at 4 °C, washed with 50 mM sodium phosphate pH 8.0 150 mM NaCl and eluted in 50 mM sodium phosphate pH 8.0, 150 mM NaCl, 100 mM imidazole. Further purification was achieved by anion exchange chromatography in 10 mM Tris pH 8.0, 1 mM EDTA, 1 mM DTT and size exclusion chromatography (Superdex 200) in 20 mM Tris pH 8.0, 100 mM NaCl, 1 mM EDTA, 1 mM DTT.

Ubiquitin (human) was expressed from a pET15 construct. *E. coli* were grown in TB at 37° C and induced in log phase with 500 μM IPTG overnight at 16 °C. Cells were harvested, washed with PBS, resuspended in Lysis buffer (50 mM Tris, 1mM EDTA pH 7.6 + 0.05% Tween-20) and sonicated. The soluble fraction was dialyzed twice into 3.5 L 50 mM sodium acetate pH 4.5. Dialyzed sample was pelleted by centrifugation at 18,000 RPM for 20 min. Supernatant was loaded onto a cation exchange column (SP Sepharose) and eluted in 50 mM sodium phosphate pH 4.5 and NaCl. Fractions containing ubiquitin were purified further by SEC (Superdex 75) in 50 mM Tris pH 8.0, 50 mM KCl.

Protein modification with maleimide-biotin or maleimide-FAM

Tau and CHIP were dialyzed into 25 mM HEPES pH 7.4 50 mM KCl, 1 mM TCEP and labeled with 1.1 eq of maleimide-biotin or maleimide-FAM for 1 h at room temperature. Caspase-6 was labeled with 2 eq maleimide-FAM overnight at 4° C in purification buffer supplemented with 1 mM TCEP. Excess reagent was removed by iterative dilution and concentration.

Peptide Synthesis
Peptides were synthesized by Fmoc solid phase peptide synthesis on a Syro II peptide synthesizer (Biotage) at ambient temperature and atmosphere on a 12.5 μmol scale using pre-loaded Wang resin. Coupling reactions were conducted with 4.9 eq of HCTU (O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluoro- phosphate), 5 eq of Fmoc-AA-OH and 20 eq of N-methylmorpholine (NMM) in 500 uL of N,N dimethyl formamide (DMF). Reactions were run for 8 min while shaking. Each position was double coupled. Fmoc deprotection was conducted with 500 uL 40% 4-methypiperadine in DMF for 3 minutes, followed by 500 μL 20% 4-methypiperadine in DMF for 10 minutes, and six washes with 500 μL of DMF for three minutes. Acetylation was achieved by reaction with 20 eq acetic anhydride and 20 eq NMM in 500 μL DMF for 1 h while shaking. Peptides were cleaved with 500 μL of cleavage solution (95% Trifluoroacetic acid 2.5% Water 2.5% Triisopropylsilane) while shaking for 1 h. Crudes were precipitated in 10 mL cold 1:1 diethyl ether : hexanes. Peptide crudes were solubilized in a 1:1:1 mixture DMSO: water: acetonitrile and purified by HPLC on a Agilent Pursuit 5 C18 column (5 mm bead size, 150 mm x 21.2 mm) using an Agilent PrepStar 218 series preparative HPLC. The mobile phase consisted of A: Water 0.1% Trifluoroacetic acid and B: Acetonitrile 0.1% Trifluoroacetic acetic acid. Solvent was removed under reduced atmosphere and 10 mM DMSO stocks were made based on the gross peptide mass. Purity was confirmed by LC/MS. Stocks were stored at -20 °C. Fluorescein isothiocyanate (FITC) was added to FP tracers by reacting 1.7 eq of FITC and 20 eq of NMM in 500 uL of DMF for while shaking for 1 h. FITC capping reactions were run twice. The quenched fluorogenic peptide corresponding to the tau$^{0421}$ site (NH$_2$-K(MCA)IDMVD/SPQLAK(DNP)-COOH) was synthesized as above where fluorophore (7-Methoxycoumarin-4-acetic acid (MCA)) and...
the quencher (2,4-dinitrophenyl (DNP)) are linked to the epsilon nitrogen of a lysine. Peptide was synthesized using Wang resin pre-loaded with Fmoc-Lys(DNP)-OH (Millipore). Fmoc-Lys(MCA)-OH coupling reactions were conducted with 1.7 eq monomer and HCTU.

**The PSSCL library** was synthesized using standard coupling conditions outlined above. “Mixed” positions in the library were coupled with an equimolar mixture of all twenty proteinogenic Fmoc protected amino acid monomers. Pools were cleaved and used without further purification. Molecular weights were estimated by using 110 Da (the average MW of a natural amino acid) for mixed positions.

**DSF**

**PSSCL** screens were conducted in a 96-well qRT-PCR plate on a Stratagene Mx3005P-rtPCR. Each well contained 10 μL 10 μM CHIP in 25 mM HEPES pH 7.4 50 mM KCl 1 mM TCEP, 10 μL of 200 μM PSSCL pool in the same buffer + 0.2% CHAPS + 2% DMSO, and 5 μL 25X Sypro Orange dye. Fluorescence intensity readings were taken over 70 cycles where reactions were heated to desired temp and then cooled to 25° C before reading. Temperature was increased at a rate of 1° C per cycle. Fluorescence Intensity data was normalized and truncated at 55° C to fit to a Boltzmann Sigmoid in Graphpad Prism 7.0.

\[ Y = \text{Bottom} + \frac{(\text{Top} - \text{Bottom})}{1 + \exp\left(\frac{T_m - T}{\text{Slope}}\right)} \]  

 Binding to CHIP was assessed by the change in T_m.

\[ \Delta T_m = \Delta T_{\text{POOL}} - \Delta T_{\text{DMSO}} \]
Tau DSF was conducted as above with 10 μL of 10 μM (1eq) tau in lieu of the peptide library.

**FP Binding Assay**

FP assays were run in 18 μL in a Corning black 384 well round bottom low volume plate and read on a BioTek H4 multimode plate reader at 21°C.

**PSSCL** screens were performed by measuring displacement of CHIP tracers corresponding to the Hsc70 or Hsp90α sequences (FAM-Ahx-SSGPTIEEVD and FAM-Ahx-DDTSRMEEVD). 2X Protein-tracer complexes (20 nM Tracer +2X EC85 CHIP) were prepared in 25 mM HEPES pH 7.4 50 mM KCl 1 mM TCEP. 2X (200 μM) PSSCL pool was prepared in the same buffer + .2% CHAPS + 2% DMSO. Each solution was added to the well and incubated for 15 min. ΔPolarization (mP) was calculated relative to DMSO control. Final concentration of CHIP was 1.58 μM for Hsc70 and 3.48 μM for Hsp90α.

**Saturation binding** assays with 1 nM FITC-CHIPOpt (FITC-Ahx-LWWPD), FITC-Hsp70 (FITC-Ahx-IEEVD), and FITC-Hsp90 (FITC-Ahx-MEEVD) were run in a final buffer composition of 25 mM HEPES pH 7.4, 50 mM KCl, .01% Triton X-100 and 1% DMSO. Raw polarization (mP) data were normalized and plotted relative to log_{10}[CHIP]M. Data was fit to the model for log(agonist) vs response (variable slope) in Graphpad Prism 7.0. Kd was extrapolated from EC_{50}. Tau tracer saturation binding was identical except that 0.1% CHAPS was substituted for .01% Triton X-100 and tracer concentration was 20 nM.

\[
Y = \text{Bottom} + \frac{(\text{Top} - \text{Bottom})}{(1+10^{((\text{LogEC}_{50}-X) \times \text{HillSlope})})}
\] (3)
**Competition binding assays** A 2X stock of CHIP+Tracer was made in HEPES pH 7.4, 50 mM KCl so that the final concentration of protein would be EC_X and tracer would be 1 nM. 2X peptide stock was prepared in 25 mM HEPES pH 7.4 50 mM KCl .02% Triton X-100 and 2% DMSO. Raw polarization (mP) data were plotted relative to log_{10}[Peptide]M. Data was fit to the model for log(antagonist) vs response (variable slope) in Graphpad Prism 7.0.

\[
Y = \frac{100}{1 + 10^{([\text{LogIC}_{50} - X] \times \text{HillSlope})}}
\] (4)

K_i was calculated from EC_{50}.

\[
K_i = \frac{[I]_{50}/([L]_{50}/K_d+[P]_0/K_d+1)}
\] (5)

**Kinetic FP** A 2X stock of CHIP+Tracer (1000 nM CHIP 20 μM FAM-Ahx-SSGPTIEEVD) was prepared in 25 mM HEPES pH 7.4 50mM KCl 1 mM TCEP, 0.1% CHAPS. A 2X stock of tau^{FL} or tau^{D421} (20 μM) was prepared in the same buffer. 9 μL of each 2X stock were added to the wells in triplicate. An initial FP reading was recorded. 2 μL of active caspase-3 (10X or 50nM) in the same assay buffer and added to each well. FP readings were taken every 46 s for 20 min. A well with no tau was included to account for effects of caspase on CHIP binding to tracer.

**Competition binding with FL Tau and Caspase-6** were run identically to peptide competition binding experiments with the Hsp70 tracer (FITC-Ahx-IEEVD) but in 50 mM
HEPES pH 7.5, 25mM MgSO₄, 0.5 mM EGTA, 5 mM reduced glutathione, and 0.01% Triton X100. Tracer was used at a final concentration of 20 nM.

**X-Ray Crystallography**

The protein solution was prepared by mixing a 1:1 molar ratio of human CHIP-TPR, at 7 mg/ml, and the 5mer 70 or CHIPOpt peptide in protein buffer (10 mM Tris-HCl pH 8.0, 150 mM NaCl and 2 mM DTT), and incubated on ice for 30 min. Crystals of the complex were grown at room temperature by hanging-drop by mixing 100 nL of the protein solution with 100 μL of the crystallization condition (Hsp70: 0.4 M CaCl₂, 0.1 M HEPES (pH 7.5), 25% PEG 4K) CHIPOpt: 0.05M CaCl₂, 0.1 HEPES (pH 7.0), 28% PEG 4K, 0.01M CoCl₂) by TTPLabtech Mosquito Nanoliter Dropsetter. Crystals appear within 24 h and were harvested ~ 1 week after setup by flash-freezing in liquid nitrogen using a cryogenic solution of 50% MPD in the crystallization condition. Data were collected at Lawrence Berkeley National Laboratory Advanced Light Source beamline 8.3.1. Diffraction images were processed using Xia2 with the Dials pipeline. Automatic molecular replacement was performed using the online Balbes tool. The resulting structure models were refined over multiple rounds of restrained refinement and isotropic B-factor minimization with Phenix. For 6EFK Ramachandran favored=98.5%, allowed=1.5%, outliers=0%. For 6NSV Ramachandran favored=98.46%, allowed=1.54%, outliers=0%.

**Rosetta ΔΔG calculations**

The energetics of particular mutations were assessed computationally using ddg_monomer from the Rosetta suite using the solved structure of CHIP TPR bound to Ac-IEEVD peptide as starting model. Structures of starting model and mutants underwent
three rounds of energy minimization. Fifty iterations of the ΔΔG optimization process were performed and the model with the lowest overall energy was used for analysis.

**Protease Activity Assays**

**Caspase screening** was conducted with a panel of commercially available active caspase -1, -2, -3, -6, -7, -8, -9 and -10 (Enzo Life Sciences, ALK-850-243-KI01). 1 unit of enzyme was added to 10 μM of fluorogenic substrate corresponding to the tau_{D421} cleavage site in 25 mM HEPES pH 7.4 50 mM KCl 0.1% CHAPS 1 mM TCEP. Fluorescence intensity was monitored as a function of time. Enzymes were compared based on the initial velocity of the reaction.

**Caspase 6 kinetic assays** The activity of a final enzyme concentration of 10 nM was monitored using a fluorogenic substrate (Ac-VEID-AMC). To each well was added a 4X stock of caspase-6_{D179} and a 4X stock of CHIP in 50 mM HEPES pH 7.5, 25 mM MgSO₄, 0.5 mM EGTA, 5 mM reduced glutathione, and 0.01% Triton X100. This mixture was incubated at ambient temperature for 10 min. Substrate was added immediately prior to reading as a 2X stock in 50 mM HEPES pH 7.5, 25 mM MgSO₄, 0.5 mM EGTA, 5 mM reduced glutathione, and 0.01% Triton X100 +2% DMSO. Activity was monitored at 30 s intervals for 10min. V₀ was calculated from data collected in the interval between 30 s and 360 s in arbitrary units (RFU/s).

**Ubiquitination Assays**
Four 4X stocks were prepared: (A) Ube1 + UbcH5c (400 nM Ube1 and 4 μM UbcH5c), (B) Ubiquitin (1 mM Ub), (C) ATP + MgCl$_2$ (10 mM ATP and 10 mM MgCl$_2$), and (D) CHIP + substrate (4 μM CHIP and 4 μM FAM-MAPT or FAM-caspase-6) in 50 mM Tris, 50 mM KCl, pH 8.0. Ubiquitination reactions were set up by adding 27.5 μL of each 4X stock, in order from A to D, for a final volume of 110 μL (100 nM Ube1, 1 μM UbcH5c, 250 μM ubiquitin, 2.5 mM ATP, 2.5 mM MgCl$_2$, 1 μM CHIP, 1 μM FAM-substrate). Reactions incubated at room temperature, quenching 10 μL aliquots in 5 μL 3X SDS-PAGE running buffer after the appropriate incubation period. Gel was imaged at Alexa488 channel on Biorad Imager.

**Biolayer Interferometry**

A qualitative analysis of the selective interaction between CHIP and caspase-6$^{D179}$ was conducted using biolayer interferometry (BLI) on an Octet Red 384well system (ForteBio). Biotinylated CHIP or CHIP K30A were loaded onto a BLI sensor coated with streptavidin. Association was observed by immersing loaded tips into solutions of caspase-6$^{D179}$ at varying concentrations in 50 mM HEPES pH 7.5, 25mM MgSO$_4$, 0.5 mM EGTA, 5 mM reduced glutathione, and 1% BSA for 240 s. Dissociation was observed by transferring the tip of a well with the same buffer and no caspase-6 for 450 s. Relative binding was assessed qualitatively.

**SDS PAGE Gels**

All SDS PAGE was run on a precast 15% gel (Biorad) for 33 min at 200V
**Western Blotting**

Transfers were conducted on nitrocellulose membranes using a Trans-Blot Turbo Transfer system (Bio-Rad). Membranes were blocked with Odyssey TBS blocking buffer (Licor) for 1h at room temperature. Primary antibodies were diluted 1:2000 in blocking buffer and incubated overnight at 4C. Membranes were washed 5x3min with TBST and probed with secondary for 1h at room temperature. Secondaries were washed 5x3min with TBST and imaged on a Licor Fc imaging system.

**Imaging**

Cells were seeded in a clear bottom, TC treated 96 well plate at 10K cells/well in 100μL complete growth medium and allowed to adhere overnight. Cells were induced with 10 μg/mL Doxycycline for 72h. Nuclei were counter stained with 1 μg/mL Hoechst 33342 in complete growth media for 30 min before imaging. Live-cell images were taken at 40X magnification on a GE InCell 6000.

**Microtubule Pulldowns**

Flp-In T-Rx 293 cell lines expressing GFP-0N4R tau constructs were generated, propagated, and stored according to the manufacturer’s instructions (https://www.thermofisher.com/order/catalog/product/R78007). 150K cells were seeded in a 6-well plate in 3 mL complete growth medium (DMEM, 10% FBS, 1% Penicillin/Streptomycin), allowed to adhere for 12h, and induced with 10 μg/mL Doxycycline for 72h. Media was removed and cells were washed with DPBS. Cell membranes were lysed with 90 μL of a microtubule-stabilizing lysis buffer (20 mM MOPS
pH 6.8, 50 mM NaCl, .5% NP-40, 2 mM EGTA, 1 mM MgCl₂, 2 mM TCEP, 2 μg/mL paclitaxel, 2 mM GTP, 250 unit/mL benzonase) and rocked gently for 3 min. Soluble lysate was removed gently with a cut pipette tip. The well was rinsed gently with 45 μL of lysis buffer and the soluble fraction and centrifuged at 15K RPM at room temp. 50 μL of was diluted in 3X SDS-PAGE loading buffer. 200 μL of 1X SDS-PAGE loading buffer was added to the well to solubilize the microtubule fraction. Samples were boiled and 15 μL loaded for Western Blot analysis. Membranes were probed with mouse αGFP (Santa Cruz), mouse αTubulin (Santa Cruz), and Rabbit αCHIP (Abcam). Blots were developed using Licor secondary antibodies (Goat-αRabbit 680RD and Goat-αmouse 800CW).

**Biotin-CHIP Pulldowns**

Jurkat cells were cultured according to ATCC (RPMI-1640, 10% FBS, 1% Penicillin/Streptomycin). For pull-down experiments cells were seeded at 500K/mL in complete growth media and treated with either 1 μM Staurosporine (from a 1000X DMSO stock) or DMSO alone for 2h. Cells were harvested by centrifugation and washed with DPBS. Washed cell pellets were flash frozen. Cells pellets were thawed and lysed at room temp for 5 min while shaking in 10 μL of M-PER (Thermo-Fischer Scientific) per million cells. 50 μL (5 million cells) of lysate was diluted with 150 μL of TBS pH 8.0. Biotinylated CHIP (WT or K30) was added to the diluted lysate so that the final concentration of CHIP was 1 μM and the mixture was allowed to incubate for 1h. After incubation Biotinylated CHIP was enriched by incubation with 100 μL of streptavidin coated magnetic bead slurry (MagneSphere, Promega) for 1h at room temperature.
Beads were washed 3x with TBS and eluted in 50 μL of 1X loading buffer. 15 μL of eluate was for Western Blot analysis with Licor secondary antibody (Goat αRabbit 800CW).

**Brain Immunofluorescence**

The de-identified post mortem tissues were sourced from the Brain Bank of the Brazilian Brain Aging Study Group and the Neurodegenerative Disease Brain Bank at UCSF as previously described ([https://memory.ucsf.edu/neurodegenerative-disease-brain-bank](https://memory.ucsf.edu/neurodegenerative-disease-brain-bank))

AD was categorized according to the Braak staging system and CERAD neuropathology criteria. All cases represented sporadic AD and included Braak 1 (early AD), Braak 3 (mid-AD), and Braak 6 (late AD). Blocks of the entorhinal cortex-hippocampus were embedded in paraffin and cut in serial sections. To visualize the interaction of different markers and their overlapping positivity, the sections were stained with multiplex immunofluorescence (IF). Sections were autoclaved in citrate buffer retrieval solution at 121 °C. Primary antibodies (aCasp6 from Aviva Scientific, CHIP from Sigma Aldrich and TauC3 from Life Technologies; all 1:200) were incubated overnight, followed by species-specific secondary antibodies for 1 h. Neuronal cell bodies were labeled with Neurotrace 435/455 for 30 min (1:50; Life Tech), a fluorescent Nissl stain. Sections were visualized with a 20x objective (Plan Apo N.A. 0.75, Nikon, Japan) using a Nikon 6D high throughput wide field epifluorescence microscope (Nikon, Tokyo, Japan) at the UCSF Nikon Imaging Center.
References


Figure 2.1. CHIP function is rooted in the recognition of C-termini. A, A composite structural model depicting how the CHIP homodimer (grey surface/cartoon) positions the E2 ubiquitin conjugating enzyme UBCH5 (black cartoon) so that the chaperone tail is 58 Å away from the activated ubiquitin thioester (yellow spheres) (PDB: 2C2L and 2OXQ). The critical motif for chaperone binding to CHIP spans the last five amino acids (P5,4,3,2,1, yellow) of Hsp70s and Hsp90s (PDB: 3Q49). B, Lines represent experimentally validated interactions between cytosolic Hsp70 (red) or Hsp90 (blue) chaperones with TPR co-chaperones (black circles). CHIP (yellow) is the only E3 ubiquitin ligase amongst these TPR co-chaperones. C, The terminal sequences of Hsp70s (red) and 90s (blue) bind to CHIP’s TPR domain. Positions P5, P2, and P1 (yellow) are critical for binding. A PSSCL library was designed to sample amino acid diversity at P5, P4, P3 and P2. D, Each peptide pool was analyzed by DSF and two competitive fluorescence polarization (FP) assays, yielding Δ(Tm) and Δ(mP) values. Pearson’s r was used to compare the data sets to one another. These experiments indicated that affinity is not optimized in the natural peptides (IEEVD or MEEVD). E, A
saturation FP experiment with tracers corresponding to the optimized sequence “CHIPOpt” (FITC-Ahx-LWWPD, black), Hsp70 (FITC-Ahx-IEEVD, red), and Hsp90 (FITC-Ahx-MEEVD, blue). Normalized polarization is plotted for individual replicate samples (n=3) in a representative experiment. Reported pKd is a mean of pKd values from independent experiments (n=3) with error reported as SEM. F, Competition binding FP experiments with an alanine scanning library of CHIPOpt. pKi values calculated from independent experiments (n=3) are plotted individually (open circles). Bars represent mean pKi. Error is plotted as SEM.
Hsp70 Ac-IEEVD

Hsp70 Ac-LWWPD

CHIPOpt Ac-IEEVD

CHIPOpt Ac-LWWPD

P2-V

P2-P

P4-E

P4-W

D134

K72

Q102

V61

N65

F98

K95

F131

I135

L68

N34

Y49

K30

Hsp70

CHIPOpt

2.9 Å

3.2 Å
**Figure. 2.2. Structural basis for CHIP TPR domain specificity.** A, X-ray crystal structures of Hsp70 peptide (PDB ID 6EFK) or CHIPOpt peptide (PDB ID 6NSV) bound to CHIP’s TPR domain were solved at 1.5 and 1.3 Å respectively. B, The overall binding mode is conserved between Hsp70 and CHIPOpt peptides. C, At the P2 position the proline in CHIPOpt fixes the $\phi$ at -67.5° vs -86.4 for the P2 valine in Hsp70. The P4 tryptophan in CHIPOpt occupies a hydrophobic shelf formed by F98, F99, and L68. D, A kinked backbone orientation supports hydrogen bonding between the acetyl oxygen and K95 in the carboxylate clamp by CHIPOpt (yellow) rather than the acetyl nitrogen and D134 in Hsp70 (grey).
Figure 2.3. Proteome-wide prediction of CHIP TPR interactions with C-termini. A, A model of CHIP TPR specificity where the contribution of P5-P2 is denoted by the size of the circle (where Area ∝ ΔpK_i for a given Ala variant relative to CHIPOpt). CHIPScore is defined as the normalized sum of thermal shifts from the PSSCL DSF dataset for a given peptide sequence. B, Correlation (R^2) between CHIPScore and experimental pK_i (n=3, see Figure 1f) for the CHIPOpt alanine scanning library, as well as Hsp70 (red) and Hsp90 (blue) peptides. The degree of correlation was determined by linear regression. The fit and the 95% confidence interval are represented by the dotted and solid lines, respectively. C, The application of CHIPScore to the reference proteome identified proteins (black) with C-termini that are predicted to bind CHIP. Proteins that are predicted to be outside the cytosol or nucleus (dashed) would not be expected to interact with CHIP. Proteins previously identified as putative CHIP interaction partners are shown by arrows (Taipale et al.). D, Proteome wide prediction of latent CHIP binding C-termini likely to be unmasked by caspase activity, based on CHIPScore (x-axis) and caspase scoring (y-axis). Caspase scoring was based on a support vector machine (SVM) prediction function for caspase cleavage (Barkan et al. https://modbase.compbio.ucsf.edu/peptide/). Hits (yellow) fall above a CHIPScore threshold of 1.79 (the value for Hsp90’s MEEVD) and SVM threshold of (~0.464) (see methods and Supplementary Fig 5). E, The location of predicted CHIP interaction sites on caspase-6^D179 and tau^D421. pIC_{50} values calculated from independent experiments (n=3) are plotted individually (open circles). Bars represent mean pIC_{50}. Error is plotted as SEM.
Figure 2.4. Tau^{D421} specifically recruits CHIP A, Schematic of how tau dissociates from microtubules and forms high-order aggregates during the progression of AD. Multiple caspase cleavages sites in tau, including D421, have been reported to accumulate in AD. B, In a competition binding FP experiment, tau^{D421} but not tau^{FL} displaces tracer bound to CHIP. Polarization (mP) is plotted for individual replicate samples (n=3) in a representative experiment. Reported \( pIC_{50} \) is mean of three independent experiments with error reported as SEM. C, In vitro ubiquitination experiment with CHIP(E3), UBCH5c(E2), Ubiquitin, and ATP demonstrates selective ubiquitination of tau^{D421} but not tau^{FL} by CHIP. Use of CHIP^{K30A} diminishes ubiquitination. Results are representative of three independent experiments. D, Tau^{FL}, tau^{D421}, or tau^{D421A} were stably expressed with an N-terminal GFP tag under the control of a Tet repressor ((TetO2)2) in TREx HEK293 cells. Representative live cell image shows GFP-tau (green, false color) associated with cytoskeletal structures. Scale bar is 30 μm. Nuclei are labeled with Hoechst (blue, false color). E, A microtubule fractionation assay for all three GFP-tau cell lines +/- 10 μg doxycycline for 72h after seeding. Probing with anti-GFP reveals that GFP-tau is in the microtubule-bound fraction (bottom panel). A low MW GFP positive band (*) was detected in the soluble fraction (top panel). Tubulin is distributed between the soluble and microtubule-bound fractions. CHIP is predominantly in the soluble fraction but is re-partitioned to the microtubule-bound fraction when tau^{D421} is expressed. Results are representative of three independent experiments. Uncropped blots in Supplementary Fig 9.
Figure 2.5 CHIP specifically binds and inhibits mature caspase-6. A, Caspase-6 is a constitutive homodimer that undergoes iterative proteolytic cleavages at three distinct sites (yellow circles) during maturation to yield a CHIP binding site (D179) at the C-terminus of the large subunit. B, A competition binding FP experiment, caspase-6\(^{D179}\) displaces tracer bound to CHIP. Polarization (mP) is plotted for individual replicate samples (n=3) in a representative experiment. pIC\(_{50}\) is reported as a mean of three independent experiments with error reported as SEM. C, In a fluorogenic substrate cleavage assay, caspase-6\(^{D179}\) is competitively inhibited by increasing concentrations of CHIP. The reciprocal of initial velocity (1/V\(_0\)) is plotted relative to 1/[S] for individual replicate samples (n=3) in a given experiment. Results are representative of three
independent experiments. D, Western blots show that caspase-6 zymogen (top panel, C-6^FL) does not pull down with recombinant biotinylated CHIP^WT or CHIP^K30A from Jurkat cell lysate. Caspase-6^D179 is specifically generated in cells treated with staurosporin (1 \( \mu \)M) for 2h and is enriched in eluates from CHIP^WT but not CHIP^K30A (bottom panel, C-6^D179). Results are representative of three independent experiments. Uncropped blots in Supplementary Figure 2.9.
Fig. 2.6 CHIP interactions at latent C-termini suggest a role in the progression of AD. A, An integrated model of expanded CHIP functions including chaperone- and caspase-coupled ubiquitination of tau (MAPT$^{D421}$), as well as inhibition of caspase-6. B, Immunofluorescence signals and high-magnification insets of caspase-6$^{D179}$ (green), tau$^{D421}$ (red) and CHIP (yellow) positive neurons (blue Neurotrace/Nissl) in the hippocampal area CA1 of human brains from early (Braak I), middle (Braak III) and late (Braak VI) stages of AD. Scale bars: 50 μm; insets: 10 μm. Qualitative analysis for this study was conducted based on trends within the three samples presented in this figure.
<table>
<thead>
<tr>
<th></th>
<th>Hsp70(6EFK)</th>
<th>CHIPOpt(6NSV)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P 2 1 2 1 21</td>
<td>P 2 2 1 21</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a, b, c$ (Å)</td>
<td>46.71 70.34 77.07</td>
<td>46.22 71.92 77.79</td>
</tr>
<tr>
<td>$\alpha, \beta, \gamma$ (°)</td>
<td>90 90 90</td>
<td>90 90 90</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>39.95-1.50</td>
<td>39.73-1.31</td>
</tr>
<tr>
<td></td>
<td>(1.55-1.50)*</td>
<td>(1.35-1.31)*</td>
</tr>
<tr>
<td>$R_{sym}$ or $R_{merge}$</td>
<td>0.1539 (1.749)</td>
<td>0.03455 (0.6336)</td>
</tr>
<tr>
<td>$I / \sigma I$</td>
<td>9.42 (1.55)</td>
<td>11.09 (1.24)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>96.64 (86.34)</td>
<td>94.66 (89.46)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>11.5 (7.1)</td>
<td>2.0 (2.0)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>1.50</td>
<td>1.31</td>
</tr>
<tr>
<td>No. reflections</td>
<td>39996</td>
<td>60397</td>
</tr>
<tr>
<td>$R_{work}$ / $R_{free}$</td>
<td>0.2107 / 0.2442</td>
<td>0.2145 / 0.2427</td>
</tr>
<tr>
<td>No. atoms</td>
<td>2534</td>
<td>2447</td>
</tr>
<tr>
<td>Protein</td>
<td>2209</td>
<td>2172</td>
</tr>
<tr>
<td>Ligand / ion</td>
<td>7</td>
<td>41</td>
</tr>
<tr>
<td>Water</td>
<td>318</td>
<td>234</td>
</tr>
<tr>
<td>$B$-factors</td>
<td>19.33</td>
<td>25.09</td>
</tr>
<tr>
<td>Protein</td>
<td>17.73</td>
<td>24.22</td>
</tr>
<tr>
<td>Ligand / ion</td>
<td>14.78</td>
<td>32.86</td>
</tr>
<tr>
<td>Water</td>
<td>30.54</td>
<td>31.73</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.008</td>
<td>0.01</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.15</td>
<td>1.05</td>
</tr>
</tbody>
</table>

*Data was collected from 1 xtal for both structures*
Table S2.2
CHIP-TPR with 5mer ddg_monomer values for the parent sequence (red) and predictions at P4 and P2 (yellow).

<table>
<thead>
<tr>
<th>AA</th>
<th>P5</th>
<th>P4</th>
<th>P3</th>
<th>P2</th>
<th>P1</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>-3.575</td>
<td>0.362</td>
<td>-0.674</td>
<td>1.048</td>
<td>-1.33</td>
</tr>
<tr>
<td>C</td>
<td>-2.28</td>
<td>-0.08</td>
<td>-0.98</td>
<td>1.616</td>
<td>-0.89</td>
</tr>
<tr>
<td>D</td>
<td>-0.298</td>
<td>2.674</td>
<td>1.33</td>
<td>7.444</td>
<td>0.019</td>
</tr>
<tr>
<td>E</td>
<td>-0.555</td>
<td>-0.021</td>
<td>-0.05</td>
<td>6.386</td>
<td>-0.559</td>
</tr>
<tr>
<td>F</td>
<td>-6.41</td>
<td>-1.779</td>
<td>-1.836</td>
<td>7.365</td>
<td>0.534</td>
</tr>
<tr>
<td>G</td>
<td>-4.823</td>
<td>1.268</td>
<td>0.521</td>
<td>2.745</td>
<td>-1.259</td>
</tr>
<tr>
<td>H</td>
<td>-4.647</td>
<td>-0.481</td>
<td>-0.865</td>
<td>3.863</td>
<td>-0.162</td>
</tr>
<tr>
<td>I</td>
<td>0.13</td>
<td>-0.546</td>
<td>-0.488</td>
<td>3.271</td>
<td>-1.891</td>
</tr>
<tr>
<td>K</td>
<td>3.832</td>
<td>0.349</td>
<td>0.096</td>
<td>13.283</td>
<td>-0.943</td>
</tr>
<tr>
<td>L</td>
<td>-3.204</td>
<td>-1.997</td>
<td>-0.893</td>
<td>5.378</td>
<td>-0.931</td>
</tr>
<tr>
<td>M</td>
<td>-3.78</td>
<td>-0.871</td>
<td>0.129</td>
<td>1.629</td>
<td>-2.13</td>
</tr>
<tr>
<td>N</td>
<td>-3.216</td>
<td>0.07</td>
<td>-0.457</td>
<td>2.618</td>
<td>-0.499</td>
</tr>
<tr>
<td>P</td>
<td>-2.216</td>
<td>2.406</td>
<td>10.33</td>
<td>-1.726</td>
<td>5.08</td>
</tr>
<tr>
<td>Q</td>
<td>-1.279</td>
<td>-0.273</td>
<td>0.18</td>
<td>2.506</td>
<td>-1.013</td>
</tr>
<tr>
<td>R</td>
<td>9.239</td>
<td>1.929</td>
<td>-1.398</td>
<td>17.138</td>
<td>-0.056</td>
</tr>
<tr>
<td>S</td>
<td>-3.866</td>
<td>1.237</td>
<td>-1.117</td>
<td>2.174</td>
<td>-0.955</td>
</tr>
<tr>
<td>T</td>
<td>-2.486</td>
<td>-0.422</td>
<td>-0.919</td>
<td>0.586</td>
<td>-1.567</td>
</tr>
<tr>
<td>V</td>
<td>-2.932</td>
<td>-0.675</td>
<td>-0.667</td>
<td>0</td>
<td>-0.632</td>
</tr>
<tr>
<td>W</td>
<td>-0.372</td>
<td>-2.24</td>
<td>-1.72</td>
<td>5.149</td>
<td>-1.183</td>
</tr>
<tr>
<td>Y</td>
<td>-1.393</td>
<td>-1.151</td>
<td>-2.376</td>
<td>8.253</td>
<td>-0.079</td>
</tr>
</tbody>
</table>
**Table S2.3**  
CHIPScore hits from the explicit C-terminome.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Sequence</th>
<th>CHIPScore</th>
<th>Entry name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q86WH2</td>
<td>VWKPD</td>
<td>2.44</td>
<td>RASF3_HUMAN</td>
</tr>
<tr>
<td>Q8TW40</td>
<td>LYQVD</td>
<td>2.36</td>
<td>CC103_HUMAN</td>
</tr>
<tr>
<td>P02747</td>
<td>LLFPD</td>
<td>2.30</td>
<td>C1QC_HUMAN</td>
</tr>
<tr>
<td>Q7Z5L3</td>
<td>LLYPD</td>
<td>2.19</td>
<td>C1QL2_HUMAN</td>
</tr>
<tr>
<td>Q5VXU3</td>
<td>IFRPD</td>
<td>2.19</td>
<td>CHIC1_HUMAN</td>
</tr>
<tr>
<td>Q9UKJ5</td>
<td>IFRPD</td>
<td>2.19</td>
<td>CHIC2_HUMAN</td>
</tr>
<tr>
<td>Q8N3L3</td>
<td>LEGVD</td>
<td>2.12</td>
<td>TXLINB_HUMAN</td>
</tr>
<tr>
<td>Q9H5X1</td>
<td>VLEPD</td>
<td>2.09</td>
<td>FA96A_HUMAN</td>
</tr>
<tr>
<td>Q6IA69</td>
<td>LDGVD</td>
<td>2.07</td>
<td>NADE_HUMAN</td>
</tr>
<tr>
<td>P17066</td>
<td>IEEVD</td>
<td>2.05</td>
<td>HSP76_HUMAN</td>
</tr>
<tr>
<td>P0DMV9</td>
<td>IEEVD</td>
<td>2.05</td>
<td>HSP71B_HUMAN</td>
</tr>
<tr>
<td>P54652</td>
<td>IEEVD</td>
<td>2.05</td>
<td>HSP72_HUMAN</td>
</tr>
<tr>
<td>P34931</td>
<td>IEEVD</td>
<td>2.05</td>
<td>HSP71L_HUMAN</td>
</tr>
<tr>
<td>P0DMV8</td>
<td>IEEVD</td>
<td>2.05</td>
<td>HSP71A_HUMAN</td>
</tr>
<tr>
<td>P11142</td>
<td>IEEVD</td>
<td>2.05</td>
<td>HSP7C_HUMAN</td>
</tr>
<tr>
<td>Q8NFB2</td>
<td>IEMPD</td>
<td>2.03</td>
<td>T185A_HUMAN</td>
</tr>
<tr>
<td>Q05D32</td>
<td>LLFPD</td>
<td>2.03</td>
<td>CTS1L2_HUMAN</td>
</tr>
<tr>
<td>P07510</td>
<td>LPSPD</td>
<td>2.00</td>
<td>ACHG_HUMAN</td>
</tr>
<tr>
<td>Q9H7F4</td>
<td>IDMPD</td>
<td>1.99</td>
<td>T185B_HUMAN</td>
</tr>
<tr>
<td>Q6Z223</td>
<td>IYIPD</td>
<td>1.96</td>
<td>C1QL4_HUMAN</td>
</tr>
<tr>
<td>P0C671</td>
<td>VESPD</td>
<td>1.93</td>
<td>CF222_HUMAN</td>
</tr>
<tr>
<td>P21854</td>
<td>FRFPD</td>
<td>1.89</td>
<td>CD72_HUMAN</td>
</tr>
<tr>
<td>Q9NUQ3</td>
<td>IESVD</td>
<td>1.86</td>
<td>TXLNG_HUMAN</td>
</tr>
<tr>
<td>Q6PL45</td>
<td>YLPD</td>
<td>1.85</td>
<td>BRID5_HUMAN</td>
</tr>
<tr>
<td>Q58FF7</td>
<td>MEEVD</td>
<td>1.79</td>
<td>H90B3_HUMAN</td>
</tr>
<tr>
<td>F07900</td>
<td>MEEVD</td>
<td>1.79</td>
<td>HS90A_HUMAN</td>
</tr>
<tr>
<td>F08238</td>
<td>MEEVD</td>
<td>1.79</td>
<td>HS90B_HUMAN</td>
</tr>
</tbody>
</table>
Table S2.4
CHIPS Score and SVM Scores for hits identified in the Degrabase. Entries with no SVM (denoted #N/A) Score were not present in the SVM database.

<table>
<thead>
<tr>
<th>Entry Nam</th>
<th>5mer</th>
<th>P1Site</th>
<th>SVMScore</th>
<th>CHIPS Score</th>
<th>Entry Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q15477</td>
<td>LEEP</td>
<td>800</td>
<td>-0.46</td>
<td>2.37</td>
<td>SKIV2_HUMAN</td>
</tr>
<tr>
<td>Q9H9J4</td>
<td>LEEP</td>
<td>764</td>
<td>1.27</td>
<td>2.37</td>
<td>UBP42_HUMAN</td>
</tr>
<tr>
<td>Q8WYQ5</td>
<td>LEEP</td>
<td>396</td>
<td>2.64</td>
<td>2.32</td>
<td>DCR8_HUMAN</td>
</tr>
<tr>
<td>Q06013</td>
<td>LYAP</td>
<td>256</td>
<td>0.26</td>
<td>2.32</td>
<td>HSF1_HUMAN</td>
</tr>
<tr>
<td>A7XYQ1</td>
<td>LTPD</td>
<td>298</td>
<td>-0.96</td>
<td>2.25</td>
<td>SOBP_HUMAN</td>
</tr>
<tr>
<td>Q9UE87</td>
<td>LDDP</td>
<td>219</td>
<td>0.95</td>
<td>2.24</td>
<td>DAXX_HUMAN</td>
</tr>
<tr>
<td>Q95347</td>
<td>LDVD</td>
<td>1116</td>
<td>2.66</td>
<td>2.19</td>
<td>SM2C_HUMAN</td>
</tr>
<tr>
<td>Q92619</td>
<td>LVDP</td>
<td>662</td>
<td>1.27</td>
<td>2.18</td>
<td>HMHA1_HUMAN</td>
</tr>
<tr>
<td>P00558</td>
<td>LGRP</td>
<td>68</td>
<td>0.65</td>
<td>2.18</td>
<td>PGK1_HUMAN</td>
</tr>
<tr>
<td>Q87DS5</td>
<td>LDVP</td>
<td>179</td>
<td>1.94</td>
<td>2.17</td>
<td>PKHO2_HUMAN</td>
</tr>
<tr>
<td>Q15042</td>
<td>LEDV</td>
<td>712</td>
<td>1.20</td>
<td>2.15</td>
<td>SR140_HUMAN</td>
</tr>
<tr>
<td>Q7L2B7</td>
<td>ILPD</td>
<td>77</td>
<td>-1.25</td>
<td>2.15</td>
<td>EIF3M_HUMAN</td>
</tr>
<tr>
<td>Q43194</td>
<td>LDQV</td>
<td>86</td>
<td>1.76</td>
<td>2.14</td>
<td>PJA2_HUMAN</td>
</tr>
<tr>
<td>Q9Y5V3</td>
<td>ISEP</td>
<td>222</td>
<td>3.38</td>
<td>2.13</td>
<td>MAGD1_HUMAN</td>
</tr>
<tr>
<td>Q96SB3</td>
<td>LVEV</td>
<td>551</td>
<td>2.05</td>
<td>2.13</td>
<td>NEB2_HUMAN</td>
</tr>
<tr>
<td>Q13601</td>
<td>LTVP</td>
<td>38</td>
<td>1.00</td>
<td>2.13</td>
<td>KR1_HUMAN</td>
</tr>
<tr>
<td>Q2TA86</td>
<td>LLTV</td>
<td>290</td>
<td>-0.15</td>
<td>2.12</td>
<td>QRI1_HUMAN</td>
</tr>
<tr>
<td>Q14980</td>
<td>VEEP</td>
<td>1829</td>
<td>1.06</td>
<td>2.12</td>
<td>NUMA1_HUMAN</td>
</tr>
<tr>
<td>Q14818</td>
<td>VFSV</td>
<td>13</td>
<td>-0.35</td>
<td>2.11</td>
<td>PSA7_HUMAN</td>
</tr>
<tr>
<td>Q12888</td>
<td>LSDV</td>
<td>211</td>
<td>0.66</td>
<td>2.11</td>
<td>TP53B_HUMAN</td>
</tr>
<tr>
<td>Q9Y2K2</td>
<td>IVEP</td>
<td>383</td>
<td>1.23</td>
<td>2.08</td>
<td>SIK3_HUMAN</td>
</tr>
<tr>
<td>Q96NH2</td>
<td>LVSP</td>
<td>109</td>
<td>0.57</td>
<td>2.07</td>
<td>SAH3_HUMAN</td>
</tr>
<tr>
<td>Q9UL51</td>
<td>LFEI</td>
<td>54</td>
<td>1.06</td>
<td>2.07</td>
<td>BAG5_HUMAN</td>
</tr>
<tr>
<td>Q9UKA4</td>
<td>LTVV</td>
<td>1081</td>
<td>0.95</td>
<td>2.06</td>
<td>ARA11_HUMAN</td>
</tr>
<tr>
<td>Q09666</td>
<td>IDTP</td>
<td>3717</td>
<td>-0.34</td>
<td>2.05</td>
<td>AHNK_HUMAN</td>
</tr>
<tr>
<td>Q60496</td>
<td>LDVY</td>
<td>1008</td>
<td>1.99</td>
<td>2.04</td>
<td>PLX1_HUMAN</td>
</tr>
<tr>
<td>Q9WR50</td>
<td>LVEV</td>
<td>355</td>
<td>-0.07</td>
<td>2.01</td>
<td>EZBG_HUMAN</td>
</tr>
<tr>
<td>Q9P2N5</td>
<td>TBPV</td>
<td>487</td>
<td>1.59</td>
<td>2.01</td>
<td>RBM27_HUMAN</td>
</tr>
<tr>
<td>Q66K74</td>
<td>LDRV</td>
<td>286</td>
<td>1.81</td>
<td>2.01</td>
<td>MAP1S_HUMAN</td>
</tr>
<tr>
<td>P31350</td>
<td>LSLV</td>
<td>29</td>
<td>-0.88</td>
<td>2.00</td>
<td>RIR2_HUMAN</td>
</tr>
<tr>
<td>Q9P2M7</td>
<td>LSSY</td>
<td>173</td>
<td>-0.27</td>
<td>2.00</td>
<td>CING_HUMAN</td>
</tr>
<tr>
<td>Q723T8</td>
<td>LSSV</td>
<td>107</td>
<td>1.54</td>
<td>2.00</td>
<td>ZFY16_HUMAN</td>
</tr>
<tr>
<td>P27797</td>
<td>KDPD</td>
<td>212</td>
<td>0.28</td>
<td>1.99</td>
<td>CALR_HUMAN</td>
</tr>
<tr>
<td>Q15181</td>
<td>VDDV</td>
<td>165</td>
<td>0.91</td>
<td>1.99</td>
<td>IPYR_HUMAN</td>
</tr>
<tr>
<td>P24666</td>
<td>NWVR</td>
<td>43</td>
<td>-0.42</td>
<td>1.99</td>
<td>PPCA_HUMAN</td>
</tr>
<tr>
<td>P09651</td>
<td>VEOR</td>
<td>69</td>
<td>0.35</td>
<td>1.99</td>
<td>ROA1_HUMAN</td>
</tr>
<tr>
<td>Q96686</td>
<td>ISMV</td>
<td>1583</td>
<td>-1.09</td>
<td>1.99</td>
<td>AHNK_HUMAN</td>
</tr>
<tr>
<td>Q09666</td>
<td>ISMP</td>
<td>2711</td>
<td>-2.12</td>
<td>1.99</td>
<td>AHNK_HUMAN</td>
</tr>
<tr>
<td>Q96666</td>
<td>LDVP</td>
<td>865</td>
<td>-0.70</td>
<td>1.98</td>
<td>AHNK_HUMAN</td>
</tr>
<tr>
<td>Q96666</td>
<td>LDVP</td>
<td>2882</td>
<td>-0.38</td>
<td>1.98</td>
<td>AHNK_HUMAN</td>
</tr>
<tr>
<td>Q9Y695</td>
<td>GWPD</td>
<td>1012</td>
<td>0.22</td>
<td>1.98</td>
<td>NCOA3_HUMAN</td>
</tr>
<tr>
<td>Q9H673</td>
<td>LDVP</td>
<td>451</td>
<td>1.00</td>
<td>1.98</td>
<td>RPAP3_HUMAN</td>
</tr>
</tbody>
</table>
Table S2.5
CHIPScore and SVM Scores for hits identified in the Proteome.
Supplementary Table 5 is available free of charge at the following link
https://www.nature.com/articles/s41589-019-0322-6#Sec29
Figure S2.1. DSF PSSCL Screen of CHIP TPR Specificity

Normalized DSF data (y=Normalized Fluorescence intensity, x=° C) in the range of 30 °C to 50 °C. ΔTm is values were calculated from the difference of calculated Tm for CHIP with 100 µM peptide pool (red) vs a DMSO control (black). Optimal residues at each position are highlighted in yellow boxes.
Figure S2.2 Competition Fluorescence Polarization PSSCL Screen. Competition FP assays for CHIP binding to tracers corresponding to a, Hsc70 (HspA8) and b, Hsp90α (Hsp90AA). The observed change in polarization (ΔmP) is plotted relative to the dynamic range of the assay (dotted line). Each peptide pool was incubated with tracer bound to CHIP at 100 µM and the tracer concentration was set to EC₈₅. Delta mP values are plotted for individual replicate samples (open circles, n=4). Means are represented as bars with error plotted as SD.
**Figure S2.3. Competition binding of acetylated 5mer peptides.** Representative competition binding curves for acetylated 5mer peptides CHIPopt (Ac-LWWPD), CHIPopt Ala Scan (P1-P5), Hsp70 peptide (Ac-IEEVD) and Hsp90 peptide (Ac-MEEVD). CHIPopt, CHIPopt P5A, and CHIPopt P3A competitions were conducted with CHIP bound to a FITC-CHIPopt tracer at EC_{85}. Polarization (mP) values for individual replicate samples (n=4) are plotted. CHIPopt P4A, CHIPopt P2A, CHIPopt P1A, Hsp90 and Hsp70 competitions were conducted with CHIP bound to a FITC-Hsp70 tracer at EC_{50}. 
Polarization (mP) values for individual replicate samples (n=3) are plotted. Results are representative of three independent experiments.
Figure S2.4. Structural studies of peptide-bound CHIP-TPR. A, Crystal structure of the human CHIP TPR domain bound to Hsp70 (6EFK) and B, CHIPOpt (6NSV). Peptide is presented in sticks and CHIP-TPR shown as surface. C, A density unique to the crystallographic interface in the CHIPOpt structure (6NSV) was modeled as two PEG6 molecules based on a host-guest interaction with K72 in both of the TPR protomers in the asymmetric unit. D, An overlay of ten structures of TPR co-chaperone TPR domains (AIP (4AIF), CHIP (3Q49), FKBP4 (1KT0), PPP5 (1A17), RPAP3_TPR1 (3CGV), RPAP3_TPR2 (3CGW), SGTA (2VYI), STIP1_TPR1 (1ELW) STIP1_TPR2 (1ELR), and UNC45A (2DBA), highlights the unique loop in CHIP (red, ribbon) that forms the P5 binding site for the Ile in the Hsp70 peptide (yellow, sticks). E, A comparison of the peptide binding modes of an Hsp70-derived peptide (PTIEEVDCO2-) bound to HOP TPR1 (grey, 1ELW) and CHIP (yellow) highlights the kink in the CHIP bound peptide relative to the extended peptide conformation when peptide is bound by HOP TPR1.
Figure S2.5. CHIPScore of the C-terminome, degradome, and proteome. A, "CHIPScore" was based on the sum of normalized thermal shift values calculated from the PSSCL DSF data set, where CHIPScore = P5+P4+P3+P2. For a given amino acid sequence B, CHIPScores for the native C-terminome identified 27 C-termini with a CHIPScore that exceeded that of HSP90. C, CHIPScores for C-termini generated by experimentally validated caspase cleavage sites from the Degrabase (https://wellslab.ucsf.edu/degrbase/) identified 84 C-termini with scores that exceeded Hsp90. D, To identify latent C-termini with aspartates likely to be unmasked by caspase activity, we used a dataset generated by a support vector machine (SVM) scoring function designated to weigh the likelihood of cleavage at a given aspartate in the proteome (https://modbase.compbio.ucsf.edu/peptide/). Comparison of SVM scores for the proteome (black) and experimentally validated cleavage sites from the Degrabase (grey) showed a significant enrichment of higher SVM scores in the true positives. A SVM
threshold capturing 75% (horizontal line) of validated cleavages (SVM=-.464, vertical line) was set to identify probable cleavage sites. E, Scoring of the Degradome dataset with CHIPScore and SVM scoring confirmed that a majority of CHIPScore hits fell above the SVM threshold (yellow). F, CHIPScore and SVM scoring of 493,321 aspartic acids in the proteome yielded 2,757 hits including tau\textsuperscript{D421} and caspase-6\textsuperscript{D179} (inset, yellow). G, Competition binding curves for acetylated 5mer peptides Tau\textsuperscript{D421} (Ac-IDMVD) and caspase-6\textsuperscript{D179}. Polarization (mP) values for individual replicate samples (n=3) are plotted. Competitions were conducted with CHIP bound to a FITC-Hsp70 tracer at EC\textsubscript{50}. Results are representative of three independent experiments.
Figure S2.6. Caspase cleavage at tau\textsuperscript{D421} directs binding of and ubiquitination by CHIP. A, A screen of known caspase-generated latent C-termini (grey bars) from tau in a competition FP binding assay at 100 µM showed that only tau\textsuperscript{D421} is capable of binding the CHIP TPR domain. Hsp70 (red) and Hsp90 (blue) are included as positive controls. Delta mP values are plotted for individual replicate samples in a given experiment (open circles, n=4). Means are represented as bars with error plotted as SD. B, Saturation FP experiment with tracers corresponding to the explicit C-terminus (grey) and the latent C-terminus tau\textsuperscript{D421} (black) confirmed selectivity. Polarization is plotted for individual replicate samples (n=3) in a given experiment. Results are representative of three independent experiments. C, DSF confirmed that CHIP (black) was not stabilized by the addition of FL 0N4R tau (light grey) but was significantly stabilized (>5 °C) by tau\textsuperscript{D421} (dark grey) Normalized fluorescence intensity is plotted for individual replicate samples (n=3). D, Calculated Tm for Fig S6c based on a fit of the data to a Boltzmann sigmoidal. Error is plotted as the SEM. E, A screen of a commercial panel of recombinant active caspases determined that caspase-3 was the most efficient at cleaving an internally quenched fluorogenic substrate corresponding to the tau\textsuperscript{D421} site. Initial velocity is plotted for individual replicate samples (n=3) in the experiment F.
Addition of 5 nM caspase-3 (at t=0, dotted line) to a competition binding assay containing tau$^{FL}$ (black) induces the time dependent generation of tau$^{D421}$ and the displacement of bound FP tracer *in situ*. This displacement did not occur when tau was excluded (dark grey) and the plateau of the reaction occurred at a similar mP value as equimolar tau$^{D421}$. Polarization (mP) is plotted for individual replicate samples in the experiment (n=3). Results are representative of three independent experiments G, Caspase-3 cleavage of tau$^{FL}$ induces tau ubiquitination (left panel). When CHIP is excluded from the reaction cleaved tau$^{D421}$ accumulates but is not ubiquitinated (right panel). Results are representative of three independent experiments.
**Figure S2.7. Caspase6<sup>D179</sup> is bound and inhibited by CHIP.**

A, BLI confirmed that biotinylated WT CHIP coated sensors bound to caspase-6<sup>D179</sup> (left panel, greyscale), but that no significant interaction was observed for sensors coated with biotinylated CHIP<sup>K30A</sup> at 1000 nM caspase 6 (right panel). Results are representative of three independent experiments.

B, *In vitro* ubiquitination of caspase<sup>D179</sup> showed that neither the large (L) or small (S) subunits were ubiquitinated in these conditions. (* denotes a contaminant band). Results are representative of three independent experiments.

C, CHIP inhibits the turnover of 12.5 μM Ac-VEID-AMC fluorogenic substrate by active 10 nM caspase-6. This inhibition is not observed with CHIP<sup>K30A</sup>. Initial velocity $V_0$ is plotted for individual replicate samples in the experiment (n=3). Data is representative of three independent experiments.
Figure S2.8. CHIP mRNA is not downregulated in AD. A, A comparative analysis of CHIP mRNA levels (z-score) in RNAseq datasets from patients in the Allen Brain Atlas Aging, Dementia, and TBI dataset (http://aging.brain-map.org/) at Braak 0, I, II, III, IV, V, VI. There was no significant downregulation of CHIP mRNA in the four brain regions (Hippocampus, Parietal Neocortex, Temporal Neocortex, or White Matter). B, Comparison of CHIP mRNA z-score in patients +/- the ApoE4 allele also revealed no significant change in CHIP mRNA levels between the groups. C, Additionally, no significant variation of CHIP mRNA levels occurred with age. Significance was assessed by ordinary One-way ANOVA for Braak Stage and Age. ApoE4 dataset was analyzed using unpaired t-test.
Figure S2.9. Uncropped blots from main text figure 4e and 5d. *Blots from figure 5d include a replicate experiment not shown in the main text.
Chapter 3

Probing selectivity among TPR co-chaperones using peptide photoprobes

Matthew Ravalin, Cory Nadel, Jason E. Gestwicki and Charles S. Craik
Abstract. The human proteome contains 273 proteins with an annotated TPR domain. Of these seventeen have been identified to be TPR co-chaperones that 1) co-immunoprecipitated with cytosolic chaperones contain the conserved (M/I)EEVD motif at the extreme C-terminus and 2) contain a properly oriented carboxylate clamp to specifically engage the C-terminal aspartate present in this motif (Figure 3.2A). The discovery that the TPR co-chaperone CHIP engages in chaperone competitive interactions with proteins containing a C-terminal Asp establishes the possibility that all TPR Co-chaperones can engage in such interactions with explicit and latent C-termini. This observation increases the theoretical diversity of the PPI network anchored in TPR co-chaperone interactions by two orders of magnitude. The assessment of the selectivity of a given TPR binder will be critical to understanding the features of this PPI network. To this end we have developed a TPR footprinting platform that uses pernicious TPR binding sequences modified with an alkyl diazirine to covalently label available TPR binding sites in a cell lysate. Competition labeling experiments demonstrated that this platform could rank order candidate CHIP binders in a cell lysate and justify a transition to chemo-proteomic experiments to address TPR co-chaperone selectivity.

1. Introduction. Photo-crosslinkers have found utility mapping protein-ligand and protein-protein interactions. Briefly, a protein or putative ligand is functionalized with a chemical moiety (usually a biphenyl acetone (BPA), azo-benzene, alkyl or aryl diazirine) through synthetic modification or metabolic incorporation.¹ Each of these moieties is chemically inert in the absence of photo-stimulated uncaging of a reactive moiety. In the
case of the alkyl diazirine, exposure to light in the 330-370nm range generates an alkyl-diazo intermediate that readily decays via the generation of molecular nitrogen and a reactive carbene. Canonically the carbene is capable of reacting with a wide array of functionalities present in biomolecules, although recent work indicates that the diazo intermediate may also covalently label carboxylic acids to form alky esters.² Regardless of mechanism, diazirine photoprobes have proven useful for mapping PPIs, identifying ensembles of metabolite binding proteins, target identification of natural products, and mapping ligandable sites in the proteome³–⁶.

Modern methods in mass spectrometry including stable isotope labeling by amino acids (SILAC) and tandem mass tagging (TMT) have made the quantitative comparative analysis of chemically modified and enriched proteomes tractable and multiplexable. More than a decade of work in this area has settled on an experimental paradigm wherein a high confidence interaction for a given bait is not determined by it’s ability to label a candidate prey but rater it ability to be competed specifically.⁷ This paradigm emphasizes the identification of high confidence interactions. This approach often results in the use of a promiscuous reactive moiety with the capacity to be visualized or enriched being used to footprint the specificity of competitors that lack the ability to be enriched (i.e do not contain an alkyne, azide, or other handle that can be used for affinity purification.). Borrowing form this paradigm we sought to develop promiscuous TPR photoprobes based on the sequences of Hsp70s and Hsp90s known to be critical for this interaction (IEEVD and MEEVD respectively). In this manner we could label all TPR co-chaperones and query candidate interactors against this background in a competitive paradigm.
2. TPR targeted photoprobes bind and label recombinant CHIP and HOP. Previous work from our lab and others indicates that while most TPR co-chaperones can bind to the Hsp70 derived IEEVD sequence as well as the Hsp90 derived MEEVD sequence, intrinsic preferences for one over the other are common. To ensure that we could maximize the probability of labeling as many TPR co-chaperones as possible we synthesized photoprobes containing a specificity element based on both sequences. In addition to the specificity element probes contained a lysine that was functionalized at the epsilon nitrogen with FITC (detection element) and an N-terminal 4,4'-azipentanoate (X-link) (Figure 3.1B). To assess the ability of the probes to bind to TPR co-chaperones we took advantage of the presence of the fluorophore and monitored the fluorescence polarization (FP) of the probe in a saturating binding experiment absent crosslinking. These experiments confirmed that the probes bound to both the E3 ubiquitin ligase CHIP and the scaffolding protein HOP (Heat-shock protein organizing protein) with relative affinities in line with previously reported values.

Given the apparent affinities of the photoprobes for CHIP and HOP we chose to incubate the proteins and photoprobes at a 1:1 ratio at 5uM. After pre-incubation probes were activated with a 40W LED array with an emission maximum at 365nM (see methods). Optimization experiments determined that labeling reached a maximum after 10 min of exposure based on visualization by in-gel fluorescence after SDS-PAGE (Figure S3.1). To confirm the specificity of labeling we co-incubated an unfunctionalized 10 amino acid peptide corresponding to the C-terminus of either Hsp72 or Hsp90α at various concentrations to compete away the photoprobe and inhibit labeling. These experiments confirmed that both photoprobes (hv-(FAM)-70 and hv-
(FAM)-90) were sensitive to the presence of a competing ligand when labeling both CHIP and HOP (**Figure 3.1C and Figure 3.1D**). Interestingly for HOP, competitors seemed, qualitatively, to be better at competing matched photoprobes (i.e. hv-(FAM)-70 and Hsp72 or hv-(FAM)-90 and Hsp90α). This is in line with previous work that suggests that TPRs within HOP exhibit specificity between Hsp90s and Hsp70s. These experiments provided confidence that these probes would have utility in profiling a wider range of competitors in the more complex environment of a cell lysate.

**3. TPR photoprobes specifically label CHIP in cell lysate.** To validate the of the photoprobes to specifically label TPR co-chaperones in the context of a cell lysate we generated a HEK 293T cell line that constitutively overexpressed CHIP with a C-terminal FLAG tag (**Figure 3.2A**). Cell lysates were prepared and labeled with hv-(FAM)-70 at 10 µM, hv-(FAM)-90 at 10 µM, or a 1:1 mixture of the two at 5 µM each and analyzed by on-membrane fluorescence in the absence and presence of photo-crosslinking. To assess which bands corresponded to specific labeling events we also included competitors corresponding to the 10 C-terminal amino acids in Hsp72 and Hsp90α (**Figure 3.2B, top panel**). This analysis identified two prominent bands that were labeled by the photoprobe and diminished in intensity in the condition containing competitor. The top band, with an approximate molecular weight of 75 kDa, was preferentially labeled by and competed by Hsp72 derived sequences. The other, with an approximate molecular weight of 37 kDa was competed by both Hsp72 and Hsp90α derived sequences. Based on the molecular weight of the 37 kDa band it seemed likely that this was CHIP. This hypothesis was supported by the fact that a western blot with and anti-FLAG antibody overlaid with the 37 kDa band. To further confirm the specificity...
of this labeling event photo labeled lysate was immunoprecipitated with anti-FLAG magnetic beads and subjected to a similar analysis (Figure 3.2C). These experiments confirmed that while CHIP-FLAG was enriched in all three conditions, labeling was only achieved upon exposure to light. Moreover, signal was competed away with excess peptide corresponding to the C-terminus of Hsp72. In total these results give us confidence that our probes are specifically labeling TPR co-chaperones in a cell lysate.

4. TPR photoprobe competitions can rank order CHIP binders. A possible application for this technology is to compare the ability of a synthetic or endogenous ligand to compete for binding to a given TPR sequence. To gauge the sensitivity of the platform in this context we took advantage of the range of affinities of CHIP binding sequences at our disposal. Previous work (see Chapter 2) identified sequences other than Hsp70 and Hsp90 derived sequences that are able to compete for the CHIP interaction. One of these sequences corresponds to a neo-C-terminus generated by caspase cleavage of the microtubule associated protein tau (tau) cleaved at Asp 421 (tau^{D421}). The other, CHIPOpt (Ac-LWWPD) was determined to be an optimized acetylated five amino acid sequence for CHIP TPR binding. CHIPOpt competes for the chaperone binding site with a $K_i$ roughly 50-fold lower than the Hsp70 derived sequence. To validate that the photoprobe platform could capture these differences we labeled CHIP-FLAG lysate in the presence of increasing concentrations of Hsp90$\alpha$, Hsp72, tau^{D421}, and CHIPOpt peptides ranging from .01 to 10 $\mu$M. These experiments showed minimal competition of the CHIP band with Hsp90$\alpha$ and moderate competition with Hsp72 and tau^{D421} at the 10 $\mu$M concentration. However, CHIPOpt was capable of completely preventing labeling of the CHIP band at 1$\mu$M, 10 fold less than the
concentration of photoprobe (Figure 3.2D). These experiments demonstrate that the photoprobès are useful for rank ordering TPR binders and provide further support that CHIPOpt is a superior ligand for the CHIP TPR domain.

5. Discussion and future directions. Based on the successful experiments discussed above we have high confidence that these photoprobès will prove useful for profiling the specificity of TPR co-chaperone binding sequences. In order to enable the affinity enrichment of labeled proteins we synthesized biotin functionalize (hv-(Bio)-70 and hv-(Bio)-90) photoprobès analogous to the FAM functionalized probes. These probes were able to non-covalently compete with the FAM functionalized probe hv-(FAM)-70 for CHIP binding with pKi of 5.8 and 5.1 for hv-(Bio)-70 and hv-(Bio)-90, respectively (Figure S3.2). Optimization of photolabeling parameters for these probes will be required to maximize the number of TPR co-chaperones engaged and enriched in a given cell lysate. If effective in this sense, these probès would be able to assess the probable TPR interactome of a given sequence in a single experiment. It may be possible to use recombinant or genetically encoded competitors in situations where peptides were undesirable. One could imagine using this technology in conjunction with the PSSCL platform described in Chapter 2 to iteratively identify new binders for a given TPR co-chaperone and subsequently characterize the TPR co-chaperone footprint of that sequence. In total we anticipate these platforms will synergize to enable a quantitative understanding of TPR co-chaperone PPI networks.

6. Methods

Protein purification

CHIP was expressed and purified as described in chapter 2 (page 53)
**HOP** was purified as describe in Assimon et al 2015

**Peptide synthesis**

Peptide synthesis was conducted as described in chapter 2 (page 57)

Functionalized photoprobes were synthesized by the incorporation of an orthogonally protected lysine (Fmoc-Lys(Mtt)-OH, EMD Millipore) in the 6th position relative to the C-terminus. The diazirine was appended to the N-terminus using succinimidyl 4,4’ azipentanoate (SDA diazirine, Thermo Fisher Scientific). Diazirine capping was conducted with 1.7eq SDA diazirine and 20eq NMM in 500 µL DMF overnight at room temperature while shaking. After capping the Lys-MTT was deprotected with 500 µL 3% Trifluoroacetic acid in dichloromethane (DCM) 3x30min. Resin was washed 3x with 1mL DCM and allowed to dry on a vacuumed manifold. Resin was re-swollen in DMF for 30min prior to modification of the deprotected lysine with either FITC as described in chapter 2 (page 57) or NHS-Biotin (EZ Link NHS-Biotin, Thermo Fisher Scientific). Reaction conditions for NHS biotin were identical to those for SDA diazirine. Peptides were cleaved, purified and characterized as described in chapter 2.

**Saturation binding and competition binding FP**

FP experiments were conducted as noted in chapter 2 (page 59) with the exception that experiments were conducted with 10 nM tracer and read on a SpectraMax M5 plate reader

**Crosslinking**

All crosslinking reactions were conducted in a PCR plate using a Tfscl coin 40W UV LED Nail Dryer Lamp. Proteins were crosslinked at 10 µM. Cell lysates were crosslinked at 1 mg/mL final protein concentration.

**Lysate preparation**
All cell lysates were prepared using M-PER (Thermo Fisher Scientific) supplemented with cOmplete protease inhibitor tablet (Roche) as per the manufacturers instructions. Protein concentration was measured by BCA. Lysates were adjusted to 2 mg/mL and snap frozen in liquid nitrogen.

**SDS-PAGE**

See chapter 2 (page 64)

**In-gel fluorescence**

See chapter 2 (page 63)

**On-membrane fluorescence and western blotting**

50 µg of lysate was labeled for standard on membrane fluorescence experiments. After labeling proteins were precipitated with 9 volumes of cold acetonitrile and centrifuged at 21,000 x g for 10 minutes. The solution was removed and the precipitate allowed to dry. The pellet was resuspended in 15 µL 1X SDS loading buffer, vortexed briefly and boiled prior to SDS PAGE. For general transfer conditions and blotting conditions see chapter 2 (page 64). CHIP Flag was detected using a rabbit anti DYKDDDDK antibody (Cell Signaling D6W5B) at 1:2000 in Licor blocking buffer. Licor anti rabbit 800 (1:10000) in Licor blocking buffer was used as the secondary. FAM and FLAG signal were detected sequentially using the 600 nm and 800 nm channels on the Licor Fc imaging system.

**Immunoprecipitation**

100 ug of HEK293T CHIP-FLAG Lysate was labeled for IP. After labeling, reaction was diluted to 300uL with M-PER and incubated for 1h with 25uL anti FLAG magnetic beads (Sigma) that had been washed 3X with 1mL M-PER After binding, Supernatants were removed and frozen. Beads were washed 5X w/ 1mL M-PER and Eluted with 40uL of
1X SDS PAGE loading buffer. 50 ug of Input (unlabeled lysate) and supernatants were precipitated as described above. On membrane fluorescence and western blotting were conducted as described above.
References


(7) Backus, K. M.; Correia, B. E.; Lum, K. M.; Forli, S.; Horning, B. D.; González-

Figure 3.1. TPR photoprobes bind and label TPR co-chaperones in vitro. A, Cytosolic Hsp70s and Hsp90s engage in protein-protein interactions with a large network of TPR co-chaperones via a conserved interaction at a conserved motif (I/MEEVD) at the extreme C-terminus of the chaperone. B, Photoprobes containing a detection module, a crosslinking module (X-link), and a specificity module based on the C-terminus of Hsp70s (red) and Hsp90s) were synthesized to enable covalent labeling of TPR co-chaperones. C, Fluorescence polarization (FP) saturation binding experiments were used to determine the dissociation constant (Kd) of each photoprobe with CHIP and HOP. CHIP contains a one TPR domain and one Ubox domain in each protomer. The TPR coordinates the C-terminal Asp via a carboxy clamp formed by K30 and K95 (yellow circles). Both hv-(FAM)-70 (left panel) and hv-(FAM)-90 (right panel) are capable of labeling CHIP. Both probes can be competed with noncovalent ligand in the form of a 10 amino acid peptide from either Hsp72(SGGPTIEEVD, top panel) or Hsp90α (DDTSMEEVD, bottom panel). D, HOP contains a Three TPR domains and one. The TPRs coordinate the C-terminal Asp via a carboxy clamps formed by K8 and K73 (TPR1), K229 and K301(TPR2A), K364 and K429 (yellow circles). Both hv-(FAM)-70 (left panel) and hv-(FAM)-90 (right panel) are capable of labeling CHIP. Both probes can be competed with noncovalent ligand in the form of a 10 amino acid peptide from either Hsp72(SGGPTIEEVD, top panel) or Hsp90α (DDTSMEEVD, bottom panel).
Figure 3.2. TPR photoprobes specifically label CHIP in cell lysate. A, A schematic depicting the workflow for labeling TPR co-chaperones in cell lysate overexpressing CHIP-FLAG. Specific labeling events will be annotated based on the ability for a given labeling event to be competed non-covalently. In this way promiscuous probes can be used to characterize the specificity of candidate ligands. B, hv-(FAM)-70 (lanes 1,2) hv-(FAM)-90 (lanes 3,4), or an equimolar mixture (lanes 5,6) were used to label HEK293T CHIP-FLAG cell lysates (10min hv exposure, lanes 2,4,5). Unexposed controls were included (lanes 1,3,5) in addition to conditions containing excess competitor (lane 7: Hsp72, lane 8: Hsp90α, lane 9 Hsp72+Hsp90α). Labeling was visualized after transfer to a nitrocellulose membrane (top panel). A western blot for CHIP-FLAG identified the CHIP band and detected a low level of crosslinking induced dimerization (middle panel). An overlay suggested that CHIP-FLAG was being labeled by both probes, and that labeling was competed by both competitors. C, A FLAG immunoprecipitation confirmed that CHIP-FLAG was specifically labeled, and that labeling could be competed with excess Hsp72 peptide. D, In-gel fluorescence was used to rank order the affinity of Hsp90α (DDTSMEEVD), Hsp72 (SGGPTIEEVD), tauD421, (SSTGSIDMVD), and CHIPOpt (Ac-LWWPD). Competed bands are denoted with an asterisk (red). CHIPOpt was capable of competing with photoprobe at 1µM.
Figure S3.1. TPR photoprobe labeling time course. CHIP (top panels) or HOP (bottom panels) was incubated at a 1:1 ratio with hv-(FAM)-70 (left panels) or hv-(FAM)-90 (right panels) and exposed to 365 nm light for various amounts of time ranging from 0-10 min. Labeling was visualized by in-gel fluorescence. Coomassie stain was used to control for protein loading.
Figure S3.2. Biotin-TPR photoprobe competition binding to CHIP. A, Biotin functionalized photoprobes corresponding to Hsp70 (red) and Hsp90 (blue) tail sequences. B, Both the Hsp70 and Hsp90 probes are capable of competing with hv-(FAM)-70 for CHIP binding with pKi of 5.8 and 5.1 respectively.
Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:
I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

[Signature]
Author Signature

[Date]
08/30/2019
Date