An unfolded protein-induced conformational switch activates mammalian IRE1

Karagöz GE[1,*], Acosta-Alvear D[1,2], Nguyen H[3], Lee CP[1,4], Chu F[3], Walter P[1,*]

[1]Department of Biochemistry and Biophysics and Howard Hughes Medical Institute, University of California, San Francisco, CA 94143, USA.
[2]Current address: Department of Molecular, Cellular and Developmental Biology, University of California, Santa Barbara, CA 93106, USA.
[3]Department of Molecular, Cellular, and Biomedical Sciences, University of New Hampshire, Durham, NH 03824, USA.
[4] Current address: BioMarin Pharmaceutical Inc., San Rafael, CA 94901, USA

*: Corresponding author

Abstract

The unfolded protein response (UPR) adjusts the cell's protein folding capacity in the endoplasmic reticulum (ER) according to need. IRE1 is the most conserved UPR sensor in eukaryotic cells. It has remained controversial, however, whether mammalian and yeast IRE1 use a common mechanism for ER stress sensing. Here, we show that similar to yeast, human IRE1α's ER-lumenal domain (hIRE1α LD) binds peptides with a characteristic amino acid bias. Peptides and unfolded proteins bind to hIRE1α LD's MHC-like groove and induce allosteric changes that lead to its oligomerization. Mutation of a hydrophobic patch at the oligomerization interface decoupled peptide binding to hIRE1α LD from its oligomerization, yet retained peptide-induced allosteric coupling within the domain. Importantly, impairing oligomerization of hIRE1α LD abolished IRE1's activity in living cells. Our results provide evidence for a unifying mechanism of IRE1 activation that relies on unfolded protein binding-induced oligomerization.

Introduction

Protein-folding homeostasis is critical for proper cell function. Accordingly, cells evolved surveillance mechanisms to monitor protein-folding status and elicit adaptive responses to adjust protein-folding capacity according to need (Balchin, Hayer-Hartl et al., 2016 , Bukau, Weissman et al., 2006, Walter & Ron, 2011). In the endoplasmic reticulum (ER), where the majority of transmembrane and soluble secretory proteins fold and mature, protein-folding homeostasis is ensured by a network of signaling pathways collectively known as the unfolded protein response (UPR) (Walter & Ron, 2011). In metazoans, perturbations leading to the accumulation of mis- or unfolded proteins in the ER are recognized as "ER stress" by three unique ER-resident UPR sensors, IRE1, PERK and ATF6 (Cox, Shamu et al., 1993, Cox & Walter, 1996, Harding, Zhang et al., 2000, Niwa, Sidrauski et al., 1999, Sidrauski & Walter, 1997 , Tirasophon, Lee et al., 2000, Walter & Ron, 2011, Yoshida, Haze et al., 1998, Yoshida, Matsui et al., 2001). These sensors transmit information about the protein-folding status in the ER and drive gene expression programs that modulate both the protein-folding load and folding capacity of the ER. If ER stress remains unmitigated, the UPR induces pro-apoptotic pathways, thereby placing the network at the center life-or-death decisions that affect the progression of numerous diseases (Bi, Naczki et al., 2005, Feldman, Chauhan et al., 2005, Lin, Li et al., 2007, Lu, Lawrence et al., 2014, Vidal, Figueroa et al., 2012, Walter & Ron, 2011, Zhang & Kaufman, 2008).

IRE1 drives the most conserved branch of the UPR, which exhibits remarkably similar mechanistic aspects shared between yeast and mammals (Aragon, van Anken et al., 2009, Korennykh, Egea et al., 2009, Li, Korennykh et al., 2010). In mammals, IRE1 exists in two isoforms, α and β. IRE1α is ubiquitously expressed, whereas IRE1β expression is restricted to gastrointestinal and respiratory tracts (Bertolotti, Wang et al., 2001, Tsuru, Fujimoto et al., 2013). Both IRE1 orthologs are trans-membrane kinase/nucleases that oligomerize in the ER-membrane in response to ER stress (Aragon et al., 2009, Li et al., 2010). Oligomerization is crucial for IRE1 activation as it allows for *trans*-autophosphorylation and allosteric activation of its endonuclease domain, which for IRE1α then initiates the unconventional splicing of the *XBP1* mRNA (Aragon et al., 2009, Cox et al., 1993, Cox & Walter, 1996, Korennykh et al., 2009, Li et al., 2010, Sidrauski & Walter, 1997, Yoshida et al., 1998, Yoshida et al., 2001). Spliced *XBP1* mRNA encodes the transcription factor XBP1s, which activates the transcription of several target genes involved in restoring ER homeostasis (Acosta-Alvear, Zhou et al., 2007, Lee, Iwakoshi et al., 2003). While the *XBP1* mRNA is the only known splicing target of IRE1, active IRE1 can also cleave ER-localized mRNAs in a process known as regulated IRE1-dependent decay of messenger RNAs (RIDD), which serves to limit the amount of client proteins entering the ER, thus helping alleviate the folding stress (Hollien, Lin et al., 2009, Hollien & Weissman, 2006).

Two alternative models are used to describe how IRE1's lumenal domain senses ER stress: a recent model where unfolded proteins act directly as activating ligands and an earlier model where IRE1 lumenal domain is indirectly activated through dissociation of the ER-chaperone BiP.

The direct activation model emerged from the crystal structure of the core lumenal domain (cLD) from *S. cerevisiae* IRE1 (yIRE1; "y" for yeast), where yIRE1 cLD

dimers join via a 2-fold symmetric interface IF1^{L} ("L" for lumenal). A putative peptide-binding groove that architecturally resembles that of the major histocompatibility complexes (MHCs) extends across this interface (Credle, Finer-Moore et al., 2005). yIRE1 selectively binds a misfolded mutant of carboxypeptidase Y (Gly255Arg, CPY*) *in vivo,* and purified yIRE1 cLD directly interacts with peptides *in vitro*, leading to its oligomerization. Taken together, these observations support the model that direct binding of unfolded proteins in the ER lumen to IRE1 induces its oligomerization leading to IRE1 activation (Gardner & Walter, 2011).

Due to structural differences between human and yeast IRE1 lumenal domains, it is not yet clear if this mechanism is also used by mammalian IRE1. Although the crystal structure of human IRE1α (hIRE1α) cLD displays conserved structural elements in its core, there are several notable differences between the crystal structures of human and yeast IRE1 cLD known to date **(Figure 1)**. First, the helices flanking the groove in yIRE1 cLD are too closely juxtaposed in the human structure to allow formation of the MHC-like groove present in the yeast (Zhou, Liu et al., 2006). Second, the yIRE1 cLD structure displays a second interface, IF2^{L}, which provides contacts for higher order oligomerization, which was experimentally validated to be indispensible for yIRE1 activation *in vivo* **(Figure 1)**. In the yIRE1 cLD, an α-helix–turn region forms an important element in IF2^{L} making contacts with the incomplete β-propeller in the neighboring protomer. Notably, the residues corresponding to the α-helix–turn are not resolved in the hIRE1α cLD crystal structure (aa V307-Y358). Instead, hIRE1α cLD has two other symmetry mates in addition to the dimerization interface, which appear to be crystal lattice contacts that are predicted to be too energetically unstable to form biologically important oligomerization interfaces (Zhou et al., 2006). Indeed, the equivalent of an IF2^{L} cannot form in the depicted hIRE1α cLD structure because of a steric hindrance from a prominent α-helix ("αB helix"; aa V245-I263) that is absent in yIRE1 cLD **(Figure 1)** (Zhou et al., 2006).

The structural differences between IRE1 orthologs were cast to support the indirect model of IRE1 activation in higher eukaryotes (Zhou et al., 2006). This model poses that due to the aforementioned structural differences—rather than direct unfolded protein binding—it is the reversible dissociation of the ER-resident Hsp70-type chaperone BiP from IRE1's lumenal domain the main driving force regulating hIRE1α activity (Zhou et al., 2006). According to this view, titration of BiP to unfolded proteins upon ER stress licenses IRE1 activation (Bertolotti, Zhang et al., 2000, Carrara, Prischi et al., 2015, Oikawa, Kimata et al., 2009, Zhou et al., 2006). In yeast, however, this view has been experimentally refuted (Kimata, Oikawa et al., 2004, Pincus, Chevalier et al., 2010).

Considering the degree of conservation at various features of IRE1 mechanism of action from yeast to mammals, we favor the unifying direct activation model. Such model finds support in the notion that all structures adopted by a protein in a crystal lattice represent a singular snapshot of many possible conformational states. Therefore, it is entirely plausible that human and yeast IRE1 cLD use a common mechanism of activation and that the divergent structures aforementioned represent different states in a spectrum of possible conformational states that the IRE1 cLD from *any species* could assume. In this scenario, the crystal structure of hIRE1α cLD represents a "closed" conformation that can shift towards an "open" state to allow peptide binding in the MHC-like groove that is apparent in the structure of the yeast

ortholog (**Video 1**) (Gardner, Pincus et al., 2013, Gardner & Walter, 2011). As such, this model predicts specific outcomes that can be experimentally tested. Specifically, that (i) human IRE1 cLD can bind to unfolded polypeptides, (ii) unfolded polypeptide binding stabilizes the open conformation of the hIRE1α cLD, and (iii) the open conformation of hIRE1α cLD favors its oligomerization.

Here, we used complementary biochemical and structural approaches to experimentally explore the mechanism of human IRE1α activation. We show that hIRE1α cLD—just like its yeast ortholog—directly binds select peptides with a characteristic amino acid bias. State-of-the-art NMR experiments that probe dynamic conformational states further support an activation mechanism involving peptide binding to the MHC-like groove and stabilizing the open conformation of hIRE1α cLD. Moreover, we provide insights into the mechanism that couples peptide binding and oligomerization to produce active IRE1 oligomers. Importantly, we show by mutational analysis that lumenal domain driven oligomerization is crucial for IRE1 function in mammalian cells. Taken together, our results resolve the discrepancies between existing models of IRE1 activation and supports a model in which unfolded polypeptides can bind and directly activate human IRE1.

Results

The lumenal domain of human IRE1α binds unfolded polypeptides

To test whether, akin to yeast IRE1, mammalian IRE1 also binds unfolded proteins directly, we employed peptide tiling arrays. To identify hIRE1α cLD-binding peptides, we designed tiling arrays utilizing ER-targeted model proteins known to induce the UPR either by overproduction (proinsulin and 8ab protein from SARS-corona virus (Scheuner, Song et al., 2001, Sung, Chao et al., 2009)) or through destabilizing point mutations (myelin protein zero (MPZ)). The peptide arrays were composed by tiling 18-mer peptides that step through the entire protein sequence, shifting by three amino acids between adjacent spots. We incubated the peptide arrays with purified hIRE1α cLD fused N-terminally to maltose-binding protein (MBP) and probed with an anti-MBP antibody. As shown in **Figure 2A** (left panel), MBP-hIRE1 α cLD bound a select subset of peptides on the arrays. To maximize the available sequence space, we analyzed binding of MBP- hIRE1 α cLD to these peptides irrespective of their topological accessibility in the ER lumen. hIRE1α cLD recognized peptide sequences found in both the ER-lumenal and cytosolic domains of MPZ, which we considered together in our analyses to define the chemical properties of cLD peptide recognition. We found that hIRE1 α cLD-binding peptides with the top 10% binding scores were enriched in cysteine, tyrosine, tryptophan, and arginine **(Figure 2B, Figure 2-figure supplement 1A ,** $p< 0.05$**)**. By contrast, aspartate and glutamate were strongly disfavored, together with glutamine, valine, and serine.

IRE1 and the ER-resident chaperone BiP recognize a different subset of peptides

At a first glance, the amino acid preferences displayed by mammalian IRE1 cLD resemble those of the other chaperones including the ER chaperone BiP (Blond-Elguindi, Cwirla et al., 1993, Deuerling, Patzelt et al., 2003, Flynn, Pohl et al., 1991). Like BiP, hIRE1α cLD favored binding to aromatic and positively charged residues (Blond-Elguindi et al., 1993, Otero, Lizak et al., 2010). BiP is a highly abundant

chaperone in the ER lumen, whereas IRE1 is present at orders of magnitude lower levels (Ghaemmaghami, Huh et al., 2003). Therefore, if IRE1 and BiP recognize the same regions of unfolded proteins, the peptide-binding activity of hIRE1α cLD would depend entirely on saturation of BiP by unfolded substrate proteins—a scenario difficult to reconcile with IRE1's task of dynamically sensing ER stress. To address this point, we compared the binding preferences of mammalian BiP (fused to an N-terminal 10x-histidine tag) on the same peptide arrays. We found sequences recognized by both hIRE1α cLD and BiP **(Figure 2A, Figure 2-figure supplement, 1A,B)**. Importantly, however, we also found profound differences. While IRE1 tolerated both prolines and histidines, BiP strongly disfavored these amino acids **(Figure 2B, Figure 2-figure supplement 1A,** $p< 0.05$**)**. Conversely, BiP tolerated serine and threonines, while IRE1 strongly disfavored them. Thus, IRE1 can recognize regions of unfolded proteins to which BiP would not readily bind and *vice versa*, thereby providing a plausible explanation of how IRE1 could recognize unfolded proteins despite of the vast excess of BiP over hIRE1α LD in the ER.

Peptide recognition by hIRE1α cLD relies on distinct biochemical properties

To measure binding affinities of hIRE1α cLD's interaction with peptides in solution, we selected the two peptides with the highest binding scores in the peptide arrays (MPZ- and 8ab-derived peptides, henceforth referred to as "MPZ1" and "8ab1", respectively) and attached fluorophores at their N-termini. Fluorescence anisotropy revealed that hIRE1α cLD bound to MPZ1 with $K_{1/2} = 24 \pm 4.7$ μM and to 8ab1 with $K_{1/2} = 5 \pm 1.7$ μM **(Figure 2C)**. (Note that we used $K_{1/2}$ to denote a measure of affinity because, as we show below, IRE1 cLD exists in solution as an ensemble of different interconverting conformational states and our measurements therefore score several superimposed equilibria. The measured affinities therefore do not reflect true K_d values). These affinities fall within the same order of magnitude of chaperone binding to unfolded proteins, supporting the notion that similar modes of fast transient interactions with unfolded proteins are adopted by both IRE1 and chaperones (Karagoz, Duarte et al., 2014, Marcinowski, Holler et al., 2011, Street, Lavery et al., 2011).

To identify the minimal region in MPZ1 for binding to hIRE1α cLD, we next divided MPZ1-1 into 12, 11 and 9 amino acid long fragments representing its N-terminal (MPZ1-N), middle (MPZ1-M) and C-terminal (MPZ1-C) regions and measured their respective affinities for hIRE1α cLD. hIRE1α cLD bound to MPZ1-N with a similar affinity as the full-length peptide ($K_{1/2} = 16.0 \pm 2.6$ μM, **Figure 2D)**, whereas the other peptide fragments displayed much lower binding affinities ($K_{1/2} = 377 \pm 54$ μM and 572 ± 107 μM, respectively, assuming similar maximum anisotropy values as for the MPZ1-N peptide). We further truncated MPZ1-N by two residues at a time from either its N- or C-terminus. Deleting amino acids from the C-terminus gradually decreased the affinity **(Figure 2E)**. By contrast, deletion of the first two hydrophobic residues from the N-terminus (leucine and isoleucine) abolished its binding to hIRE1α cLD **(Figure 2F)**. These analyses revealed that the minimum peptide length with a comparable binding affinity to the full-length MPZ1 peptide is a 12-mer. This 12-mer peptide matches the chemical properties we found for hIRE1α cLD binding peptides: it is enriched in aromatics, hydrophobic amino acids and arginines, indicating that specific binding contacts play a role in hIRE1α cLD's interaction with unfolded proteins.

hIRE1α cLD binds to unfolded proteins

To validate that peptides are valid surrogates for unfolded proteins, we next tested binding of intact but constitutively unfolded proteins to hIRE1α cLD. Immunoglobulins (IgGs) mature in the ER using a well-characterized folding pathway, wherein the constant region domain of the IgG heavy chain (C_{H1}) remains disordered until it binds to its cognate partner, the constant region domain of the IgG light chain C_L (Feige, Groscurth et al., 2009). We measured the binding affinity of C_{H1} to hIRE1α cLD by thermophoresis, which reports on changes in the hydration shell of a biomolecule upon interaction with a partner in solution (Jerabek-Willemsen, Wienken et al., 2011). By contrast to earlier findings that showed no measurable binding of hIRE1α cLD to C_{H1} under different experimental conditions (Carrara et al., 2015), our experiments showed that hIRE1α cLD interacts with C_{H1} with a $K_{1/2}$ = 29.2 ± 1.2 μM (**Figure 2G)**. To further validate this observation, we measured binding of hIRE1α cLD to another model unfolded protein by fluorescence anisotropy, the folding mutant of staphylococcal nuclease Δ131Δ (Street et al., 2011) . We observed a comparable binding affinity of $K_{1/2}$ = 21.4 ± 2.3 μM (**Figure 2-figure supplement 2)**. Our data thus show that hIRE1α cLD binds to full-length unfolded proteins with similar affinity as peptides, suggesting that these proteins display a distinct single binding site for hIRE1α cLD.

To test whether multiple binding sites would increase the affinity for hIRE1α cLD, we synthesized a peptide consisting of two MPZ1-N tandem repeats separated by a 5-amino acid spacer (MPZ1-N-2X). Intriguingly, MPZ1-N-2X bound to hIRE1α cLD with an order of magnitude higher affinity ($K_{1/2}$ = 0.456 ± 0.07 μM) compared to MPZ1 peptide (**Figure 2H)**. As we show below, the increased apparent affinity is due to avidity of hIRE1α cLD to the peptide.

hIRE1α cLD is structurally dynamic

To capture evidence for structural rearrangements in hIRE1α cLD predicted by a switch-mechanism that oscillates between inactive closed and active open conformations as we suggest in the Introduction, we employed nuclear magnetic resonance (NMR) spectroscopy. NMR spectroscopy reveals structural information at the atomic level for dynamic protein complexes and is well suited to study structural changes in hIRE1α cLD upon its interaction with peptides and unfolded proteins. The hIRE1α cLD dimer is ~80 kDa and thus is well above the size limit for conventional NMR approaches. We therefore used methyl transverse relaxation optimized spectroscopy (methyl-TROSY), a specific NMR method that allows to extract structural information from large proteins after selective isotopic labeling of isoleucine methyl groups with carbon-13 (^{13}C) (Tugarinov, Hwang et al., 2004, Tugarinov, Sprangers et al., 2007). hIRE1α cLD has 12 isoleucines per monomer, which are evenly distributed throughout the protein (**Figure 3A)**. In hIRE1α cLD's methyl-TROSY spectra, we resolved 7 peaks corresponding to isoleucines **(Figure 3B)**, which then served as sensors of peptide binding and accompanying conformational changes. All isoleucine peaks in hIRE1α cLD's NMR spectrum displayed broad line widths **(Figure 3B)**, which is indicative of chemical exchange resulting from hIRE1α cLD sampling multiple conformational states at the conditions

of the NMR experiments. These data revealed that hIRE1α cLD is dynamic in solution.

To assign the resolved peaks to specific amino acids in the hIRE1α cLD sequence, we mutated each isoleucine to leucine, alanine or valine and monitored the disappearance of each resolved peak in methyl-TROSY spectra of the mutant proteins. This approach allowed us to assign 6 isoleucine peaks unambiguously **(Figure 3C, D, Figure 3-figure supplement 1,2,3**). To further increase the number of NMR visible probes in hIRE1α cLD, we mutated Leu186 and Thr159 to isoleucines (**Figure 3E, F, Figure 3-figure supplement 1C, 1F)**. Leu186 lies in an amphipathic unstructured loop surrounding the putative groove in hIRE1α cLD. The Leu186Ile peak displayed high signal intensity consistent with a dynamic and flexible position (**Figure 3F, Figure 3-figure supplement 1F).** By contrast, Thr159 lies at the β-sheet floor in hIRE1α cLD structure where its side chain faces towards the MHC-like groove and, as expected, the Thr159Ile substitution resulted in a low intensity peak (**Figure 3-figure supplement 1C)**.
We further enhanced the coverage of hIRE1α cLD with NMR-visible probes in complementary experiments in which we labeled threonine side chains with ^{13}C at their γ_2 methyl groups **(Figure 3-figure supplement 4)**. There are 33 threonine residues in hIRE1α cLD, 24 of which were detected by the NMR experiments. While we did not assign threonine peaks in hIRE1α cLD spectrum due to high spectral crowding, they provided an additional "fingerprint" reporting on peptide binding-induced changes in hIRE1α cLD.

Peptide binding stabilizes the open conformation of hIRE1α cLD

Next, we used methyl-TROSY experiment to monitor changes in the environment of isoleucines and threonines in hIRE1α cLD upon peptide binding. A largely overlapping subset of isoleucine and threonine peaks shifted when hIRE1α cLD bound to the peptides MPZ1 or 8ab1, indicating a change in a localized environment upon peptide binding **(Figure 4A,B, Figure 4-figure supplement 1A-C)**. Yet, a subset of isoleucine and threonine peaks displayed peptide specific changes. The chemical shifts displayed by the isoleucine peaks were not very large yet reproducible upon binding of different peptides allowing us to probe peptide induced changes in hIRE1α cLD. By contrast, the threonine peaks displayed larger chemical shifts, which is expected from their higher solvent exposure rendering them more sensitive to binding events **(Figure 4B, Figure 4-figure supplement 1B, 1C)**. Mapping the chemical shift perturbations of the isoleucine peaks upon peptide binding on the hIRE1α cLD structure (**Figure 4A-E, Figure 4-figure supplement 1B, Figure 4-figure supplement 2**) revealed that the isoleucine resonances that shifted most significantly lie on the floor of the central β-sheet (marked by Ile124, Ile128, Thr159Ile) (**Figure 4D, 4E)**. Among these isoleucines, only the side chain of Thr159Ile faces towards the MHC-like groove. We noted that in comparison to the isoleucines 124 and 128, Thr159Ile peak displayed a larger shift upon peptide binding (**Figure 4C, 4E, Figure 4-figure supplement 1D**). In addition to the central β-sheet floor, the αB helix that lies at the ends of hIRE1α cLD dimer (marked by Ile263), and the β-sandwich connecting the β-sheet floor to the αB helix (marked by Ile52) were affected, albeit to a lesser extent. By contrast, the unstructured loop extending from the MHC-like groove (marked by Ile186) was only slightly affected and the

isoleucines positioned in the flexible region that are not resolved in the crystal structure (marked by Ile326 and Ile334) did not shift (**Figure 4D, E**).

Importantly, binding of the unfolded protein C_{H1} shifted the same peaks in the hIRE1α cLD spectra as the short peptides suggesting peptides and unfolded protein chains interact with hIRE1α cLD in a similar way (**Figure 4F, Figure 4-figure supplement 2)**. Taken together, these results indicate that peptide as well as unfolded protein binding populate a distinct conformational state of hRE1 cLD, consistent with a peptide-induced closed-to-open conformational transition. Moreover, the results are consistent with a model in which peptide binding induces conformational changes that propagate from the MHC-like groove via the β-sandwich to affect the regions involved in oligomerization.

Peptide binding maps to the MHC-like groove in hIRE1α cLD

To map the peptide-binding site in hIRE1α cLD with higher precision, we employed paramagnetic relaxation enhancement (PRE) experiments (Gaponenko, Howarth et al., 2000, Gillespie & Shortle, 1997) using MPZ1 modified with a nitroxide spin label, 3-(2-Iodoacetamido)-PROXYL, at cysteine residue, Cys5 (**Figure 5A, B)**. The unpaired electron in the spin label broadens (in a range of 1 to 2.5 nm) or entirely erases (distances < 1 nm) NMR signals in its vicinity in a distance dependent manner (Gottstein, Reckel et al., 2012). Binding of the spin label attached peptide to hIRE1α cLD would result in a decrease in the intensity of isoleucine peaks depending on their relative distance to the peptide-binding site. Therefore, we analyzed the changes in the intensities of all isoleucine signals upon binding of MPZ1-proxyl peptide to hIRE1α cLD (**Figure 5B, Figure 5-figure supplement 1A-C)**. Binding of MPZ1-proxyl to hIRE1α cLD erased the otherwise very strong signal of Leu186Ile and broadened that of Ile124 (**Figure 5B-D and Figure 5-figure supplement 1A-C)**. Importantly, Ile128 and Ile263 signals, which shifted upon MPZ1 binding as discussed above (**Figure 4E**), broadened to a lesser extent, suggesting that these residues lie further from the peptide-binding site (**Figure 5-figure supplement 1C, 1D**). Their resonances therefore shifted due to peptide-induced distant conformational rearrangements. Displaying the normalized PRE effect on hIRE1α cLD structure revealed that MPZ1-proxyl binding mapped to the center of the MHC-like groove, suggesting that peptides bind to MHC-like groove and induce distant conformational changes in hIRE1α cLD (**Figure 5D)**.

Peptide binding induces oligomerization of hIRE1α cLD

To test whether the distant conformational changes in hIRE1α cLD monitored by the NMR experiments are due to peptide binding induced oligomerization, we employed analytical ultracentrifugation (AUC) sedimentation velocity experiments to assess the oligomeric status of hIRE1α cLD in the absence and presence of peptides. At the concentration range used at NMR experiments (25-75 μM), hIRE1α cLD was found as a mixture of various oligomeric states, where the main peaks corresponded to dimers and tetramers **(**with higher amount of tetramers formed at higher concentrations, see **Figure 6A)**. Notably, binding of MPZ1-N to hIRE1α cLD at the NMR concentrations sharpened the tetramer peak and induced formation of larger oligomeric species in these experiments **(Figure 6A)**. The peptide concentration used

in these experiments does not saturate hIRE1α cLD molecules based on a determined $K_{1/2}$ of 16.0 ± 2.6 μM, therefore only a small population of hIRE1α cLD formed higher oligomers (depicted as the pink area) **(Figure 6A)**.

To assess hIRE1α cLD's oligomeric status at varying hIRE1α cLD concentrations, we performed size exclusion chromatography and found that hIRE1α cLD eluted at earlier fractions in a concentration-dependent manner **(Figure 6-figure supplement 1A)**. AUC data confirmed these findings and showed that at concentrations close to its dimerization constant of 2.5 μM, hIRE1α cLD sediment as a single peak with a sedimentation coefficient corresponding to a mixture of monomers and dimers (**Figure 6B**, **Figure 6-figure supplement 1B)**. In this concentration regime (from 1 to 5 μM), the hIRE1α cLD peak progressively shifted to higher sedimentation values with increasing hIRE1α cLD concentration **(Figure 6-figure supplement 1B)**. Peptide binding to hIRE1α cLD shifted the hIRE1α cLD population to even higher sedimentation values **(Figure 6B, blue trace)**, indicating that under these conditions peptide binding stabilized hIRE1α cLD dimers and lead to the formation of oligomers.

Oligomerization leads to global conformational changes in hIRE1α cLD

As hIRE1α cLD populated distinct oligomeric states in a concentration-dependent manner, we next compared the conformational state of hIRE1α cLD at 5 μM (no higher-order oligomer formation detected by AUC) to 50 μM **(**based on **Figure 6C**, approximately 60 % higher-order oligomer formation) by NMR spectroscopy to probe for the structural differences assumed by these two distinct states **(Figure 6C, D)**. In these experiments, we relied on the high sensitivity of selective isoleucine labeling strategy, which could readily detect hIRE1α cLD signals at concentrations as low as 5 μM (**Figure 6-figure supplement 2A, B**).

Notably and similar to effects observed upon peptide binding, oligomerization changed the environment of the αB helix (marked by Ile263) and the β-sandwich connecting the β-sheet floor to the αB helix (marked by Ile52) that both lie at the tips of hIRE1α cLD dimers **(Figure 6D, E, Figure 6-figure supplement 2C)**. These data suggest that these isoleucines are part of the oligomerization interface and/or that their conformational rearrangements are coupled to the formation of the interface. Moreover, NMR experiments showed chemical shifts in the isoleucines on the beta sheet floor of the groove (marked by Ile124 and Ile128) upon formation of higher oligomers (**Figure 6E**). These coupled, global conformational differences observed by NMR strongly underscore the notion that oligomeric hIRE1α cLD adopts an active conformation and displays higher affinity for unfolded protein ligands. To address this notion, we set out to experimentally determine the oligomerization interface and then impair the oligomerization of hIRE1α cLD by mutation.

Identifying the oligomerization interface of hIRE1α cLD

We employed a chemical cross-linking strategy coupled to mass spectrometry to experimentally determine residues that map to the oligomerization interface in hIRE1α cLD. To this end, we cross-linked hIRE1α cLD in the presence and absence of peptides by a homobifunctional cross-linker, BS3 (bis(sulfosuccinimidyl)suberate), which crosslinks primary amines mainly present in lysine side chains. Denaturing SDS-PAGE analysis of hIRE1α cLD after cross-linking revealed that cross-linking

captured oligomeric hIRE1α cLD **(Figure 7A, Figure 7-figure supplement 1)**. We separately isolated the bands corresponding to hIRE1α cLD monomers, dimers and higher oligomers from the gel and analyzed peptides by mass spectrometry. We identified cross-linked peptides by accurate mass measurement of both candidate peptides and their fragment ions (Chu, Baker et al., 2010, Trnka, Baker et al., 2014). In comparative analyses, we separated intra- from inter-molecular cross-links by focusing on peaks that were present only in the covalent dimers and higher oligomers (Wu, Peisley et al., 2013, Zeng-Elmore, Gao et al., 2014). These analyses revealed 5 abundant cross-links between lysines 120•120, 53•347, 53•349, 53•351 and 265•351 (**Figure 7A, Table 1**).

Previous studies of BS3-cross-linked proteins with known crystal structures established that the distance between the αC atoms of cross-linked lysines is less than 28 Å for most cross-links but can be up to 33Å for a few cases due to local protein flexibility (Leitner, Walzthoeni et al., 2010), in agreement with the additive lengths of the cross-linker itself plus twice the length of the lysine side chain. The Lys120•120 cross-link maps to hIRE1α cLD's dimerization interface ($IF1^L$), whereas the four other cross-links are compatible with being positioned at hIRE1α cLD oligomerization interface, $IF2^L$. The cross-links Lys53•347, Lys53•349, Lys53•351 and Lys265•351 each involve one lysine residue (Lys53 and Lys263) that is close to the isoleucines (Ile52 and Ile263) that shifted upon hIRE1α cLD oligomerization **(Figure 6D, E)**, suggesting that they report on the formation of hIRE1α cLD's putative oligomerization interface $IF2^L$. Lys347, Lys349 and Lys351 are located in a region that was not resolved in hIRE1α cLD crystal structure, suggesting that these regions are contributing to the formation of the oligomerization interface in hIRE1α cLD.

We next threaded the sequence of human IRE1 cLD into the yeast crystal structure of the oligomeric state, which fulfilled the distance restraints imposed by the cross-links (**Figure 7B, C)**. This structural model predicted an extensive interface formed by hIRE1α cLD oligomers that involves residues from parts of hIRE1 cLD that are not resolved in the crystal structure, as well as the incomplete β-propeller involved in the formation of the oligomerization interface in yeast Ire1 cLD (**Figure 7C)**.

We used the predictive power of the structural model (hIRE1 cLD threaded into the yeast structure) to identify a patch of four hydrophobic residues WLLI (aa 359-362) suggested to contribute to the hIRE1α cLD oligomerization interface IF2L **(Figure 7C, Figure 7-figure supplement 2)**. Assuming that these residues would be critical for oligomerization, we mutated them ($WLLI^{359\text{-}362}$ to $GSGS^{359\text{-}362}$; "$IF2^L$ mutant") and assessed whether the hIRE1α cLD $IF2^L$ mutant formed oligomers by AUC sedimentation velocity analysis. The experiments revealed that, at a concentration (50 μM) where wild type hIRE1α cLD readily forms oligomers, the hIRE1α cLD $IF2^L$ mutant sediment as a single dimeric peak, showing that the mutation prevents hIRE1α cLD oligomerization **(Figure 7D)**.

The peptide-induced allosteric switch remains intact in $IF2^L$ mutant hIRE1α cLD

To assess whether hIRE1α cLD $IF2^L$ mutant is functional, we tested peptide binding by fluorescent anisotropy experiments. The $IF2^L$ mutant bound MPZ1-N and MPZ1-

N-2X peptide at similar affinities to the wild type protein (with $K_{1/2}$ = 5.4 ± 1.4 μM and $K_{1/2}$ = 0.95 ± 0.4 μM, respectively) (**Figure 7E, Figure 7-figure supplement 3A)**. These results indicated that hIRE1α cLD dimer is the functional unit for peptide binding and that hIRE1α cLD oligomers do not display a higher affinity conformation. Moreover, they also showed that the avidity effect that resulted in higher affinity binding of MPZ1-N-2X peptide to hIRE1α cLD does not require formation of higher hIRE1α cLD oligomers. AUC data confirmed these analyses and showed that binding of MPZ1-N-2X to hIRE1α cLD $IF2^L$ stabilized dimer formation but did not lead to formation of oligomers bridged by MPZ1-N-2X peptide (**Figure 7E, Figure 7-figure supplement 3B, 3C**).

The hIRE1α cLD $IF2^L$ mutant therefore enabled us to decouple peptide induced allosteric communication from the formation of oligomers, both of which could have contributed to the shift of the isoleucine peaks in the NMR experiments. To address this notion, we repeated the NMR experiments with the $IF2^L$ mutant **(Figure 7F, Figure 7-figure supplement 4A-C)**. Similar to WT hIRE1α cLD, MPZ1-N peptide binding to hIRE1α cLD $IF2^L$ mutant shifted isoleucines in the β-sheet floor (marked by Ile124 and Ile128) **(Figure 7F, 7G, Figure 7-figure supplement 4B, 4C)**. Importantly, isoleucine peaks (Ile52 and Ile263) close to the oligomerization interfaces also shifted upon peptide binding to the hIRE1α cLD $IF2^L$ mutant. Thus peptide binding-induced conformational changes in isoleucines distant to the peptide binding site persisted in the hIRE1α cLD $IF2^L$ mutant. Interestingly, MPZ1-N-2X binding shifted isoleucine peaks in the same direction and to a similar extent as binding of MPZ1-N, indicating that hIRE1α cLD $IF2^L$ binds to the same site in these peptides **(Figure 7G, Figure 7-figure supplement 4B, 4C)**. These data suggest that the increased affinity of MPZ1-N-2X is due to a decreased rate of dissociation of the peptide.

IRE1 lumenal domain-driven oligomerization is crucial for IRE1 function in mammalian cells

To test the importance of oligomerization for hIRE1α function *in vivo*, we generated cell lines that stably express hIRE1α $IF2^L$ mutant as the only form of hIRE1α. To this end, we introduced the hIRE1α $IF2^L$ mutant into mouse embryonic fibroblasts (MEFs) deficient for both isoforms of IRE1 ($IRE1\alpha^{-/-}/IRE1\beta^{-/-}$). In addition, we attached a GFP tag to IRE1's cytoplasmic flexible linker retaining its function as published previously for HEK293 cells (Li et al., 2010). In parallel, we introduced hIRE1α-GFP to $IRE1\alpha^{-/-}/IRE1\beta^{-/-}$ MEFs to compare hIRE1α activity at similar conditions. In these cell lines, we controlled hIRE1α expression via a doxycycline-inducible promoter. In the absence of doxycycline, cells expressed low levels of hIRE1α due to the leakiness of the promoter. In those conditions, the expression level of the hIRE1α-GFP-$IF2^L$ mutant was similar to hIRE1α-GFP and to the level of endogenous IRE1α from wild-type MEFs, as assessed by Western blot analysis **(Figure 8A, B).**

We next monitored the *XBP1* mRNA splicing activity of IRE1 in $IRE1\alpha^{-/-}/IRE1\beta^{-/-}$ MEFs harboring hIRE1α-GFP or hIRE1α-GFP- $IF2^L$ (**Figure 8C**). We found that unlike hIRE1α-GFP, hIRE1α-GFP- $IF2^L$ mutant did not splice *XBP1* mRNA after induction of ER stress by tunicamycin, a chemical stressor that impairs ER-folding homeostasis by inhibiting N-linked glycosylation (**Figure 8B, Figure 8-figure supplement 1**) (Heifetz, Keenan et al., 1979). IRE1's RNase activity is preceded by

the autophosphorylation of its kinase domain, which can be monitored by a phospho-specific antibody. Western blot analysis showed no signal corresponding to phospho-IRE1 in the IRE1α$^{-/-}$/IRE1β$^{-/-}$ cells expressing hIRE1α- IF2^{L}-GFP-IF2^{L}, by contrast to the same cells reconstituted with wild type hIRE1α-GFP, or in contrast to wild type MEFs, in which we detected phosphorylation of the endogenous protein **(Figure 8B, Figure 8-figure supplement 2**). Lastly, confocal microscopy revealed that under ER stress conditions where hIRE1α-GFP readily formed foci (>70%, n=88, **Figure 8D, Figure 8-figure supplement 3A**), reflecting its assembly into active oligomers, the hIRE1α-GFP-IF2^{L} mutant failed to do so (**Figure 8D, Figure 8-figure supplement 3B, 3C)**. These data confirmed that cLD-mediated oligomerization is crucial for IRE1 function in cells.

Discussion

To date, the mechanism by which mammalian IRE1 senses ER stress has remained controversial. Here, we provide evidence that activation of human IRE1 occurs via direct recognition of unfolded proteins and that the mechanism of ER stress sensing is conserved from yeast to mammals. This conclusion is based on six independent lines of evidence. First, we found that hIRE1α cLD binds peptides with a characteristic amino acid bias. Second, NMR spectroscopy suggested that peptides bind to hIRE1α cLD's MHC-like groove and induce a conformational change including the distant α B helix. In this way, occupation of the peptide-binding groove is allosterically communicated, which, we propose, culminates in the formation of a functional oligomerization interface corresponding to IF2^{L} in yIRE1 cLD. Third, binding of minimal-length peptides induces formation of hIRE1α cLD oligomers as assessed by AUC analyses, further supporting this notion. Fourth, cross-linking experiments captured the oligomerization interfaces, which allowed identification of a functionally crucial hydrophobic patch at IF2^{L}. Fifth, mutation of this patch uncoupled peptide binding from oligomerization but retained the allosteric coupling within the domain. Sixth, impairing the oligomerization of hIRE1α cLD abolished IRE1's activity in living cells, attesting to the physiological relevance of the activation mechanism proposed here.

Taken together, our data converge on a model (**Figure 9**) in which unfolded protein-binding activates a switch in IRE1's cLD, leading to rearrangements that render it compatible with the formation of IF2^{L} and therefore stabilizing an active oligomeric conformation (**Video 1**). cLD-mediated oligomerization on the lumenal side of the ER, in turn, would juxtapose hIRE1α's cytosolic kinase domains in the face-to-face confirmation allowing its *trans*-autophosphorylation, followed by stacking of its RNase domains in back-to-back orientation. These conformational rearrangements then lead to RNase activation, and thus allowing information flow across the ER membrane. Interestingly, our data show that impairment of lumenal domain oligomerization diminished IRE1's both RNase and kinase activities in cells.

Currently due to lack of biochemical and structural understanding of IRE1's interaction with the ER-resident chaperone BiP, its role in regulating IRE1 activity remains unknown. Although it is clear that BiP is released from IRE1 upon ER stress (Bertolotti et al., 2000), current models proposing BiP as the primary regulator of IRE1 activity do not address how the formation of active IRE1 oligomers would form (Carrara et al., 2015, Oikawa et al., 2009, Zhou et al., 2006). By contrast, our data

indicate that peptide-binding is important for lumenal domain-driven IRE1 oligomerization, leading to its activation. We therefore consider it most plausible that BiP binding modulates the response via tuning IRE1's oligomerization equilibrium, similar to what was shown for the yeast counterpart (Pincus et al., 2010). In this way, BiP binding would buffer IRE1 activity at the early stages of the ER stress when the chaperones are not overwhelmed by the unfolded protein load, and during the deactivation phase, when the protein folding homeostasis is achieved. In this scenario, unfolded protein accumulation exerts synergistic effects on IRE1 activation, simultaneously freeing more IRE1 from BiP upon ER stress and inducing IRE1's oligomerization/activation through their direct binding to the sensor (Pincus et al., 2010).

Despite these profound similarities in the salient features of ER stress sensing and processing, yeast and human IRE1 cLD display some distinct oligomerization properties. Whereas yIRE1 cLD precipitously assembles into larger oligomers at concentrations that exceed its dimerization constant (Gardner & Walter, 2011), hIRE1α cLD forms discrete dimers, which in a concentration-dependent manner gradually assemble into tetramers. hIRE1α cLD oligomers are in a dynamic equilibrium of different states, apparent from our size exclusion chromatography and AUC analyses and hIRE1α cLD forms even larger oligomers when bound to peptides. These observations are consistent with the model that the αB helix, which may hinder formation of hIRE1α oligomers as previously suggested (Zhou et al., 2006) participates in conformational changes that release its block on oligomerization. At higher hIRE1α cLD concentrations, the conformational equilibrium of the αB helix is shifted towards the active state. Peptide binding allosterically releases this inhibition and stabilizes the active hIRE1α oligomers. We anticipate that the effect of peptide binding-induced oligomerization would be more pronounced under physiological conditions, where hIRE1α is tethered to ER-membrane with diffusion limited to two dimensions.

We speculate that the conformational change in the αB helix allows the incomplete β-propeller to form contacts with the residues from the flexible region, which is not resolved in the crystal structure (V307-Y358) forming the oligomerization interface in hIRE1α cLD. In this conformation, αB helix may provide additional contact sites contributing to the oligomerization interface. Interestingly, one of the symmetry mates captured by hIRE1α cLD crystal structure shows contacts of the αB helix with the hydrophobic stretch (^{359}WLLI^{362}), which we show to be important for oligomerization. We anticipate that in addition to this hydrophobic stretch, additional contacts contributed by these flexible parts may further facilitate oligomer formation.

hIRE1α cLD's groove is enriched in aromatic residues and displays a negatively charged surface. In this way, the amino acids lining the groove chemically complement hIRE1α cLD binding peptides identified in our study, which are enriched in aromatics and arginines. In the crystal structure of hIRE1α cLD in the "closed" conformation, the α-helices forming the MHC-like groove are close together and mask the residues on the β-sheet floor. When these helices are moved approximately 6 Å apart from one another, the groove deepens and exposes more hydrophobicity mostly contributed by newly exposed aromatic residues. Thus, opening the groove exposes surface chemistry that is conducive to IRE1 binding peptides. Our data

support a model in which widening of the groove is allosterically coupled to the formation of the $IF2^L$-like oligomerization interfaces.

We showed that a 12-mer peptide is the shortest derivative of MPZ1 peptide that binds hIRE1α cLD with undiminished affinity when compared to the original 21-mer peptide, indicating that a 12-mer provides maximal contact with cognate interfaces in hIRE1α cLD groove. It is plausible that similar to MHC molecules, select amino acids in unfolded polypeptides act as "anchor residues" providing contact sites for hIRE1 α cLD binding (2010, Fremont, Matsumura et al., 1992, Matsumura, Fremont et al., 1992, Wilson & Fremont, 1993). Notably, assuming an extended peptide backbone with an average length of 3.4 Å per peptide bond, a 12-mer peptide can fit without constraints into the 39 Å-long groove in the structural model presented here. This notion suggests that the groove ensures preferential binding of fully exposed, unfolded 39Å-stretch of a polypeptide chain. This recognition principle is therefore similar to that of Hsp70-type chaperones, where the structural constraints in the cavity of the substrate-binding domain allow interaction with the substrates only in their extended, unfolded conformation, although Hsp70 only binds a 7 amino acids stretch (Rudiger, Buchberger et al., 1997, Rudiger, Germeroth et al., 1997).

Supporting the notion of mechanistic similarities in unfolded protein recognition between chaperone proteins and IRE1, hIRE1α cLD and the ER-resident chaperone BiP bind partially overlapping as well as distinct sets of peptides tested in our peptide arrays, as previously shown for the orthologous yeast proteins (Gardner & Walter, 2011). Importantly, the presence of distinct hIRE1α cLD binding peptides liberates IRE1 from an otherwise inevitable failure to compete with highly abundant BiP for binding sites in unfolded proteins. hIRE1α cLD's affinity for peptides measured here varied between 5 and 30 μM, which is within the same order of magnitude but at the lower range of those reported for most chaperones (Karagoz et al., 2014, Marcinowski et al., 2011, Street et al., 2011). For example, hIRE1α cLD binds the IgG's C_H1 unfolded domain with ~30 μM affinity whereas BiP was shown to bind the same protein with ~7 μM (Marcinowski et al., 2011). We surmise that this difference has been selected in evolution to set the threshold for unfolded protein recognition slightly higher for the UPR sensors when compared to that of chaperones so that the UPR is not triggered until a critical concentration of unfolded proteins accumulates. Moreover, our data with the MPZ1-N-2X peptide suggested that IRE1 could display higher affinity for select polypeptides that present more than a single IRE1 binding site.

IRE1 dysfunction contributes to the development of numerous diseases, including cancer (such as multiple myeloma (Mimura, Fulciniti et al., 2012)), metabolic disorders (such as obesity and diabetes (Fonseca, Burcin et al., 2009, Hotamisligil, 2010)) and neurodegenerative diseases (such as amyotrophic lateral sclerosis and Hungtinton's disease (Hetz, Thielen et al., 2009, Matus, Nassif et al., 2009, Vidal et al., 2012)). Depending on the disease context, IRE1 makes life or death decisions in response to altered ER function manifested in these pathological conditions (Walter P. & D., 2011). Our data showing that unfolded proteins stabilize a distinct IRE1 conformation suggest novel approaches to manipulate IRE1 pharmacologically. For example, it will be promising to design or screen for small molecule modulators that lock IRE1's groove in the open or closed conformation based on the chemical signature of the IRE1 binding peptides identified here. Such compounds could act as

agonists or antagonists of IRE1 activity. As such it should be possible to develop new classes of pharmaceuticals to induce or inhibit the IRE1 branch of the UPR, driving the desired IRE1 output depending on the disease context.

Acknowledgments

We thank the Walter lab members for insightful comments on the manuscript. Purified Δ131Δ protein was a generous gift from Daniel Elnatan and David Agard. We thank Johannes Buchner for C_{H1} expression plasmid and Linda Hendershot for hamster BiP expression plasmid. We are thankful to Avi Ashkenazi and David Lawrence for the anti-phosho-IRE1 antibody and pGpHUSH.puro vector. We thank the MIT Biopolymers Institute for synthesizing the peptide arrays. This work was supported by NSF CLF #1307367 (to FC). PW is an Investigator of the Howard Hughes Medical Institute.

Competing Financial Interests

The authors declare no competing financial interests.

Table 1. Crosslinked peptides in hIre1cLD unique to dimer and oligomers

			Band number on the SDS-PAGE (with peptides)							Band number on the SDS-PAGE (without peptides)				
Xlink 1	**Xlink 2**		**5**	**4**	**3**	**2**	**1 (monomer)**	**Xlink 1**	**Xlink 2**	**5**	**4**	**3**	**2**	**1 (monomer)**
K53	K347	AVE (%)	4.1%	3.2%	3.7%	4.9%	0.0%	K53	K347	3.3%	4.3%	4.4%	3.4%	0.0%
		STDEV(%)	0.4%	1.9%	1.0%	1.0%	0.0%			0.3%	1.9%	2.1%	2.5%	0.0%
K53	K349	AVE (%)	1.4%	1.0%	1.4%	1.5%	0.0%	K53	K349	1.6%	1.4%	1.6%	1.6%	0.0%
		STDEV(%)	0.2%	0.4%	0.3%	0.8%	0.0%			0.5%	0.5%	0.9%	0.5%	0.0%
K53	K351	AVE (%)	4.0%	2.4%	3.2%	3.4%	0.0%	K53	K351	2.3%	3.2%	2.6%	2.1%	0.0%
		STDEV(%)	1.1%	2.4%	0.8%	0.3%	0.0%			0.4%	1.0%	1.7%	2.3%	0.0%
K121	K121	AVE (%)	3.7%	2.9%	1.5%	0.8%	0.0%	K121	K121	2.6%	2.1%	2.1%	0.3%	0.0%
		STDEV(%)	0.9%	1.4%	0.0%	0.3%	0.0%			0.2%	0.8%	1.5%	0.1%	0.0%
K351	K265	AVE (%)	1.2%	0.8%	1.0%	0.4%	0.0%	K351	K265	0.6%	0.9%	0.9%	0.4%	0.0%
		STDEV(%)	0.1%	0.6%	0.5%	0.3%	0.0%			0.3%	0.2%	0.4%	0.3%	0.0%
K53	K265	AVE (%)	0.5%	0.4%	0.2%	0.2%	0.0%	K53	K265	0.3%	0.4%	0.3%	0.1%	0.0%
		STDEV(%)	0.1%	0.1%	0.0%	0.0%	0.0%			0.1%	0.3%	0.1%	0.1%	0.0%

720
721
722

Figure Legends

Figure 1 Human and yeast IRE1 cLD's crystal structures display distinct features.

The αB helix in hIRE1α cLD structure (pdb: 2hz6) and the helix-turn region and the incomplete β-propeller in yIRE1 cLD (pdb: 2be1) structure are indicated with arrows. The interfaces $IF1^L$ and $IF2^L$ and the unresolved dynamic region (aaV307-Y358) in hIRE1α cLD crystal structure are depicted with dashed lines. The distance between the helices surrounding the groove in yIre1 and hIRE1α cLD is depicted with red arrows.

Figure 2 hIRE1α cLD binds peptides and unfolded proteins

A. Peptide arrays tiled with 18mer peptides derived from proinsulin, myelin protein zero (MPZ), 8ab are probed with MBP-hIRE1α cLD (on the left) or His_{10}-BiP (on the right).

B. Comparison of the amino acid preferences of MBP-hIRE1α cLD (blue) and His_{10}-BiP (gray). The peptide arrays were quantified using Max Quant. The binding intensity in each spot was normalized to max signal intensity in the peptide array. The peptides with the top 10% binding scores were selected and the occurrence of each amino acid in these top-binding peptides was normalized to their total abundance in the arrays. The normalized occurrences are plotted in $\log_2$ scale. Blue stars depict the amino acids that are significantly enriched or depleted in hIRE1α cLD binders ($p < 0.05$), whereas red stars depict differences in binding preferences of hIRE1α cLD and BiP ($p < 0.05$).

C. hIRE1α cLD binds peptides derived from proteins MPZ, "MPZ1" (in blue) (peptides F16-F17 in Fig.1a, sequence: `LIRYCWLRRQAALQRRISAME`) and 8ab, "8ab1" (in green) (peptide H20 in Fig.1a, sequence: `WLCALGKVLPFHRWHTMV`) with a $K_{1/2}$ of 24 ± 4.7 μM and 5 ± 1.7 μM, respectively, determined by fluorescence anisotropy measurements.

D. Fluorescence anisotropy measurements show that N-terminal 12mer derivative of MPZ1 peptide, "MPZ1-N" binds to hIRE1α cLD with a similar affinity as the full-length peptide, with a $K_{1/2=}$ 16.0 ± 2.6 μM. The binding curves of N-terminal (MPZ1-N, sequence: `LIRYCWLRRQAA`), Middle (MPZ1-M, sequence: `WLRRQAALQRR`) and C-terminal (MPZ1-C, sequence: `LQRRISAME`) fragments are shown in black, dark gray, and light gray, respectively.

E. The binding affinity of C-terminal truncations of MPZ1-N for hIRE1α cLD was measured by fluorescence anisotropy. The binding curves for the truncated peptides are shown in different shades of gray.

F. Fluorescence anisotropy measurements with N-terminal truncations of the MPZ1-N peptide are shown in different shades of gray.

G. IRE1 cLD binds to unfolded C_{H1} domain of IgG1 with a $K_{1/2}$ of 29.2 ± 1.2 μM determined by microscale thermophoresis measurements.

H. Fluorescence anisotropy measurements show that MPZ1-N-2X peptide where MPZ1-N peptide sequence is repeated twice in the peptide binds tighter to hIRE1α cLD (peptide sequence: **`LIRYAWLRRQAA`**`LQRR`**`LIRYAWLRRQAA`**).

<u>Figure 2- figure supplement 1</u>. hIRE1α cLD shows preference for arginines and aromatic residues.

A. Comparison of the amino acid preferences of MBP-hIRE1α cLD (blue) and His_{10}-BiP (gray) with the amino acid composition of all peptides displayed on the array (total, black). The frequency of each amino acid present in peptides with top 10% binding score is shown for hIRE1α cLD and BiP. The experimental error is calculated from three experimental replicates. Blue stars depict the amino acids that are significantly enriched or depleted in hIRE1α cLD binders ($p < 0.05$),

whereas red stars show the amino acids that differ between hIRE1α cLD and BiP binders ($p < 0.05$).

B. Comparison of MBP-hIRE1α cLD (blue) and His_{10}-BiP (gray) binding to peptides derived from MPZ, proinsulin and 8ab. The contribution of each amino acid to MBP-hIRE1α cLD (blue) and His_{10}-BiP (gray) binding was calculated by averaging the intensity of 6 neighboring spots that contain that amino acid. The topology of the proteins are shown at the bottom of each graph (SP: signal peptide, TM: transmembrane domain). Precursor peptides of insulin are depicted in the graph. Red bars indicate the location of the MPZ1 and 8ab1 peptides in the protein sequence.

Figure 2- figure supplement 2

hIRE1α cLD binds the folding mutant of Staphylococcal nuclease Δ131Δ with 21.4 ± 2.3 μM affinity determined by fluorescence anisotropy.

Figure 3 NMR spectroscopy reveals dynamic nature of hIRE1α cLD

A. Isoleucines serving as probes in the NMR experiments are evenly distributed throughout hIRE1α cLD. hIRE1α cLD structural model is shown in gray, with space-filling isoleucine side chains shown in red. The structural model of hIRE1α cLD was generated by I-Tasser webserver using hIRE1α cLD crystal structure (pdb:2hz6) as a template to visualize the loops that are not resolved in the crystal structure (Roy, Kucukural et al., 2010, Zhang, 2008). The dimerization interface $IF1^{L}$ of hIRE1α cLD is depicted with a dashed line.

B. Methyl-TROSY spectrum of hIRE1α cLD with selective ^{13}C labeling at δ_1- methyl group of isoleucines resolves 7 peaks, indicated by red dots.

C. Ile128 is highlighted as red spheres on hIRE1α cLD structural model. hIRE1α cLD is shown in gray, with isoleucine side chains are depicted as grayspace-fillings.

D. Assignment strategy for isoleucines in hIRE1α cLD. The WT hIRE1α cLD spectrum (in black) is overlaid with the spectrum of Ile128Leu mutant (depicted in red). The signal that disappeared in the mutant spectrum corresponds to Ile128 peak and is depicted with a circle.

E. The side chain of Leu186 is highlighted as red spheres on hIRE1 α cLD structural model.

F. The WT hIRE1α cLD spectrum (in black) is overlaid with the spectrum of Leu186Ile mutant (depicted in red). The signal that appeared in the mutant spectrum that corresponds Leu186Ile peak is depicted with a circle.

Figure 3- figure supplement 1. Assignment strategy of isoleucines in hIRE1α cLD spectrum.

A-F. The WT hIRE1α cLD (in black) is overlaid with isoleucine mutants (depicted in red). The mutants Ile52Leu **(A)**, Ile124Leu **(B)**, resulted in disappearance of corresponding residues from the NMR spectrum. By contrast, we did not observe peaks disappear in the hIRE1α cLD Ile362Leu spectrum **(Figure 3-Figure Supplement 2F).** The triple Ile326/334/362Val mutation assigned disappearing peaks to Ile326 and Ile334 (**D)**. The Leu186Ile (**E)** and Ile263Ala (**F)** mutants were made independently in the context of the Ile326/334/362Val triple mutation to overcome spectral crowding and unambiguously assign these residues in the resulting quadruple mutants. Thr159Ile (**C)** Leu186Ile **(E)** mutations were made to increase the number of Ile probes.

Figure 3- figure supplement 2

A-F. No peaks disappeared in the spectrum of Ile100Leu (**A)**, Ile115Leu (**B)**, Ile163Ala (**C)** Ile219Ala (**D)**, Ile243Ala (**E)** and Ile362Leu (**F)** mutants. This is likely due to signal overlap, to low signal intensity of these residues, and/or to a combination of the two. These experiments confirm the unambiguity of the assignment of the other residues.

Figure 3- figure supplement 3

The assigned isoleucines in wild type hIRE1α cLD are depicted on the spectrum

Figure 3- figure supplement 4

Methyl-TROSY spectrum of hIRE1α cLD with selective ^{13}C labeling at γ_2 methyl group of threonines resolves 24 residues. The threonine peaks are colored as red dots.

Figure 4 Peptide binding induces conformational changes in hIRE1α cLD

A. Close-up of the isoleucine peaks in the methyl-TROSY spectrum of hIRE1α cLD Leu186Ile alone (black, 50 μM) and of hIRE1α cLD bound to MPZ1 (1:1 molar ratio) (upper panel, blue), or to 8ab1 peptide (lower panel, green, 1:1 molar ratio) shows that peptide binding shifts select peaks. The shift of each peak is indicated with brackets consistent with the color code in **Figure 4-figure supplement 4**, where yellow color indicates chemical shift perturbation values $\Delta v > 0.005$, orange, $\Delta v > 0.010$ and red $\Delta v > 0.020$. The identities of isoleucine peaks are indicated on top of each peak.

B. Close-up of the threonine peaks in the methyl-TROSY spectrum of hIRE1α cLD Leu186Ile alone (black, 50 μM) and of hIRE1α cLD bound to MPZ1 (1:1 molar ratio) (upper panel, blue), or to 8ab1 peptide (lower panel, green, 1:1 molar ratio) shows that peptide binding shifts select peaks upon binding of peptides. The chemical shift of each peak is indicated as in **Figure 4A** based on the chemical shift perturbations calculated in **Figure 4-figure supplement 4**.

C. Close-up of select isoleucine peaks in the methyl-TROSY spectrum of hIRE1α cLD T159I mutant alone (black, 25 μM) and in the spectrum of hIRE1α cLD bound to MPZ1-N peptide (1:1 molar ratio) (blue). The chemical shift of each peak is indicated with brackets consistent with the color code in **Figure 4-figure supplement 6**.

D. Important structural regions of hIRE1 cLD are depicted on the structural model of hIRE1α cLD by arrows. The red dashed-lines indicate the dimerization interface $IF1^L$ of hIRE1α cLD , whereas the black box shows the β-sheet floor of the MHC-like groove.

E. The isoleucine peaks shifting upon MPZ1 binding are mapped into the hIRE1α cLD structural model based on their combined chemical shift perturbation values as shown in **Figure 4-figure supplement 4 and 6**. The red spheres indicate isoleucine peaks with significant shifts ($\Delta v > 0.020$), orange; moderate shifts $\Delta v > 0.010$, and yellow spheres show isoleucines that shift slightly upon peptide binding, $\Delta v > 0.005$. The isoleucine peaks that do not change significantly ($\Delta v < 0.05$) are depicted in blue.

F. Close-up of select isoleucine peaks in the methyl-TROSY spectrum of hIRE1α cLD triple mutant Ile326/334/362Leu alone (black, 50 μM) overlaid with the spectrum of hIRE1α cLD when bound to C_{H1} domain (1:1 molar ratio) (red). The shifts are indicated with brackets consistent with color coding in **Figure 4B** and **Figure 4-figure supplement 7**.

Figure 4- figure supplement 1

A. Methyl-TROSY spectrum of hIRE1α cLD Leu186Ile at 50 μM alone (right), hIRE1α cLD Leu186Ile bound to MPZ1 peptide, at 50 μM (1:1 molar ratio,

middle), hIRE1α cLD Leu186Ile bound to 8ab1 peptide, at 50 μM (1:1 molar ratio, left).

B. Chemical shift perturbation analysis of isoleucine signals upon MPZ1 (blue) and 8ab1 (green) peptide binding to hIRE1α cLD (based on the spectra in Fig. 4a). The combined chemical shift perturbation of the isoleucines are calculated as $\Delta V= ((0.25\Delta vC)^2 + (\Delta vH)^2)^{1/2}$. The yellow line indicates $\Delta v > 0.005$, orange line $\Delta v > 0.010$ and red line $\Delta v > 0.020$. '*' indicates isoleucine peaks that split.Chemical shift perturbation analysis of threonine signals upon MPZ1 (blue) and 8ab1 (green) peptide binding to hIRE1α cLD (based on the spectra in **Fig. 4b**). The combined chemical shift perturbation of the isoleucines are calculated as $\Delta V= ((0.25\Delta vC)^2 + (\Delta vH)^2)^{1/2}$. The yellow line indicates $\Delta v > 0.005$, orange line $\Delta v > 0.010$ and red line $\Delta v > 0.020$. '*' indicates isoleucine peaks that split.

C. Chemical shift perturbation analysis of isoleucine signals upon MPZ1 (blue) peptide binding to hIRE1α cLD T159I mutant (based on the spectrum in Fig. 4c). The yellow line indicates $\Delta v > 0.005$, orange line $\Delta v > 0.010$ and red line $\Delta v > 0.020$.

Figure 4- figure supplement 2

Chemical shift perturbation analysis of isoleucine signals upon binding of C_{H1} domain binding to hIRE1α cLD I326-334-362L triple mutant (based on the spectrum in **Fig. 4f**). The yellow line indicates $\Delta v > 0.005$, orange line $\Delta v > 0.010$ and red line $\Delta v > 0.020$. '*' indicates isoleucine peaks that split.

Figure 5 Peptide binding maps to the center of MHC-like groove

A. Schematic representation of spin label attached MPZ1 peptide.

B. Comparison of the methyl-TROSY spectra of hIRE1α cLD Leu186Ile in the absence (black, 75 μM) and presence of spin-labeled MPZ1 peptide (red, 1:1 molar ratio).

C. The normalized PRE effect on isoleucine peaks upon binding of spin-labeled peptide. The intensity of isoleucine peaks upon MPZ1-proxyl binding is divided by their intensity in the reference spectrum (I_{PRE}/I_0) (**Figure 5-figure supplement 3**) and further normalized to their surface exposed area to exclude possible contributions from non-specific interactions with the spin label attached peptide (Clore & Iwahara, 2009) (see Methods).

D. The normalized PRE effect is mapped on the structural model of hIRE1α cLD. The isoleucine peaks in hIRE1α cLD that are broadened upon peptide binding are depicted with a color gradient from red to green as space filling side-chains in the hIRE1α cLD structural model based on decreasing degree of broadening using normalized PRE effect in **Figure 5C**.

Figure 5- figure supplement 1

A. Comparison of the methyl-TROSY spectra of hIRE1α cLD L186I/I326-334-362L quadruple mutant in the absence (black, 50 μM) and presence of spin-labeled MPZ1 peptide (red in 1:1 molar ratio).

B. The intensity of isoleucine peaks upon MPZ1-proxyl binding to L186I/I326-334-362L quadruple mutant is divided to their intensity in the reference spectrum (I_{PRE}/I_0).

C. The intensity of isoleucine peaks upon MPZ1-proxyl binding to L186I mutant is divided to their intensity in the reference spectrum (I_{PRE}/I_0).

Figure 6 hIRE1α cLD forms dynamic oligomers

A. AUC sedimentation velocity measurements of hIRE1α cLD alone (75 μM) (gray line) versus hIRE1α cLD with 75 μM MPZ1-N peptide (blue line). Pink region indicates larger hIRE1α cLD oligomers formed upon peptide binding.

B. AUC sedimentation velocity measurements of hIRE1α cLD alone (5 μM) (gray line) versus hIRE1α cLD with 50 μM MPZ1 peptide (blue line), pink region indicates the shift in the AUC profile upon peptide binding.
C. AUC sedimentation velocity measurements of hIRE1α cLD at 5, 25 and 75 μM are shown in different shades of gray.
D. Close-up of isoleucine peaks in the methyl-TROSY spectrum of hIRE1α cLD at 5 μM (black) overlaid with the spectrum of hIRE1α cLD at 50 μM (red).
E. The isoleucine peaks shifting upon oligomerization are mapped into the hIRE1α cLD structure based on the chemical shift perturbation values shown in **Figure 6-figure supplement 5.** The red spheres indicate isoleucine peaks that display most significant shifts ($\Delta v > 0.020$), orange; moderate shifts $\Delta v > 0.010$, and yellow; slight shifts $\Delta v > 0.005$. The isoleucine peaks that do not change significantly ($\Delta v < 0.005$) are depicted in blue.

Figure 6- figure supplement 1

A. Size exclusion chromatography with increasing concentrations of hIRE1α cLD shows that it elutes in earlier fractions in a concentration dependent manner (500 μM, 250 μM, 125 μM, 62.5 μM of hIRE1α cLD is loaded on the column). The arrows show the expected sizes of the most represented oligomeric species of hIRE1α cLD.
B. Comparison of sedimentation profiles of hIRE1α cLD at 1, 2.5 and 5 μM (shades of gray) by AUC sedimentation velocity experiments.

Figure 6- figure supplement 2

A. Methyl-TROSY spectrum of hIRE1α cLD Leu186Ile at 5 μM.
B. Methyl-TROSY spectrum of hIRE1α cLD Leu186Ile at 50 μM.
C. Chemical shift perturbation analysis of isoleucine signals upon oligomerization of hIRE1α cLD (based on the spectra in Fig. 6d). The combined chemical shift perturbation of the isoleucines are calculated as $\Delta V = ((0.25\Delta vC)^2 + (\Delta vH)^2)^{1/2}$. The yellow line indicates $\Delta v > 0.005$, orange line $\Delta v > 0.010$ and red line $\Delta v > 0.020$.

Figure 7 Cross-linking coupled to mass spectrometry identified the oligomerization interface of hIRE1α cLD

A. Tandem mass spectrometry (MS) profile of the peptide crosslinked at Lys53 and Lys351. Extracted Ion chromatography (XIC) of the peptide peak in monomeric versus oligomeric hIRE1α shows its absence in cross-linker treated monomer proteins.
B. Mapping cross-link sites on the structural model of hIRE1α cLD by threading on the oligomeric yeast crystal structure. Each monomer is colored as shades of gray. The Lys53, Lys265 are shown as orange and blue spheres, respectively, and Lys347,Lys349, Lys351 are shown as red spheres. The dashed lines indicate the cross-links between the lysines.
C. The amino acids forming the oligomerization interface are shown as spheres and colored by red (indicating oxygens), blue (indicating nitrogens) and white (indicating carbons). The Trp359 that is mutated in the hIRE1α $IF2^L$ mutant is colored as green and pink in different protomers.
D. ^{359}WLLI^{362}-GSGS mutation (hIRE1α cLD $IF2^L$ mutant) impairs hIRE1α cLD oligomerization determined by AUC sedimentation velocity experiments. Gray line depicts wild type hIRE1 α cLD (25 μM) and red shows the hIRE1α cLD $IF2^L$ mutant (50 μM).
E. hIRE1α cLD $IF2^L$ mutant binds MPZ1-N-2X peptide (red curve) at similar affinity $K_{1/2}$=0.95 ± 0.4 as wild type hIRE1α cLD (black curve)($K_{1/2}$ = 0.456 ± 0.7 μM) determined by fluorescence anisotropy measurements.

F. Close-up of isoleucine peaks in the methyl-TROSY spectrum of hIRE1α cLD $IF2^L$ mutant alone (black, 50 μM) overlaid with the spectrum of hIRE1α cLD bound to MPZ1-N-2X (upper panel, red, 1:1 molar ratio).

G. The isoleucine peaks shifting upon peptide binding to hIRE1α cLD $IF2^L$ mutant are mapped on hIRE1α cLD structural model based on the chemical shift perturbations calculated in **Figure 7-figure supplement 4**) consistent with the color code in **Figure 4E.**

Figure 7- figure supplement 1

SDS-PAGE analysis of BS3 (1 mM) cross-linked hIRE1α cLD (20 μM). After incubation with the BS3 cross-linker at the indicated times, the reaction is stopped with addition of Tris-HCl and the proteins are analyzed with the SDS-PAGE after coomassie blue staining.

Figure 7- figure supplement 2

The hydrophobic stretch ^{359}WLLI^{362} that is mutated in hIRE1α cLD $IF2^L$ mutant is shown on the structural model of hIRE1α cLD based on yeast crystal structure. The mutated amino acids are shown as spheres, where red spheres depict oxygens, blue, nitrogens and pink and green spheres depict carbons in different protomers.

Figure 7- figure supplement 3

A. hIRE1α cLD $IF2^L$ mutant binds MPZ1-N peptide ($K_{1/2}$ = 5.4 ± 1.4 μM) at similar affinity to wild type hIRE1α cLD ($K_{1/2=}$ 16.0 ± 2.6 μM) determined by fluorescence anisotropy measurements .

B. AUC sedimentation velocity measurements of hIRE1α cLD IF^L mutant alone (50 μM) (gray line) versus hIRE1α cLD with 12.5 μM MPZ1-N-2X peptide (blue line).

C. AUC sedimentation velocity measurements of hIRE1α cLD IF^L mutant alone (5 μM) (gray line) versus hIRE1α cLD with 5 μM MPZ1-N-2X peptide (blue line).

Figure 7- figure supplement 4

A. Methyl-TROSY spectrum of hIRE1α cLD $IF2^L$ at 50 μM.

B. Methyl-TROSY spectrum of hIRE1α cLD $IF2^L$ bound to MPZ1-N peptide, at 50 μM (1:1 molar ratio).

C. Chemical shift perturbation analysis of MPZ1-N peptide binding to hIRE1α cLD IF^{2L} mutant. The combined chemical shift perturbations of the isoleucine resonances were calculated as $\Delta v = ((0.25\Delta vC)^2 + (\Delta vH)^2)^{1/2}$.

Figure 8 Lumenal domain driven oligomerization is crucial for IRE1 function

A. Western blot analyses show the levels of hIRE1α-GFP and hIRE1α-GFP-$IF2^L$ mutant proteins stably expressed in IRE1$\alpha^{-/-}$/IRE1$\beta^{-/-}$ MEFs in response to various doxycycline concentrations. hIRE1α is detected by anti-IRE1 antibody and GAPDH is probed as the loading control. The lower panel shows Western blot analysis comparing IRE1 levels in hIRE1α-GFP and hIRE1α-GFP-$IF2^L$ mutant cell lines in the absence of doxycycline side by side.

B. Unlike hIRE1α-GFP, the hIRE1α-GFP-$IF2^L$ mutant does not splice *XBP1* mRNA after induction of ER stress by the chemical ER stressor tunicamycin (5 μg/ml). *XBP1* mRNA splicing is determined by semi quantitative PCR. The spliced and unspliced forms of *XBP1* mRNA are indicated. Splicing assays in are conducted in IRE1$\alpha^{-/-}$/IRE1$\beta^{-/-}$ MEFs reconstituted with hIRE1α-GFP or the hIRE1α-GFP-$IF2^L$ mutant in the absence of doxycycline.

C. Western blot analyses of hIRE1α-GFP and hIRE1α-GFP- $IF2^L$ mutant reconstituted in IRE1$\alpha^{-/-}$/IRE1$\beta^{-/-}$ MEFs and MEFs isolated from wild type mice are

probed with anti-IRE1 and anti-phospho-IRE1 antibody. The cells are treated with 5 µg/ml tunicamycin for inducing ER stress.

D. Confocal microscopy images of IRE1$\alpha^{-/-}$/IRE1$\beta^{-/-}$ MEFs reconstituted with hIRE1α-GFP-IF2^{L} mutant and hIRE1α-GFP after 4hrs of chemically induced ER stress by tunicamycin (5 µg/ml).

Figure 8- figure supplement 1

Western blot analyses of cell lysates collected at various times after ER stress induction from IRE1$\alpha^{-/-}$/IRE1$\beta^{-/-}$ MEFs reconstituted with hIRE1α-GFP and hIRE1α-GFP-IF2^{L} ("L" for lumenal) mutant are probed with anti-phospho IRE1 antibody (upper panel), whereas GAPDH is the loading control (lower panel).

Figure 8- figure supplement 2

Unlike hIRE1α-GFP, the hIRE1α-GFP-IF2^{L} mutant does not splice *XBP1* mRNA after induction of ER stress by tunicamycin (Tm, 5 µg/ml) at various time points and 2hrs after thapsigargin (Tg, 100 nM) treatment. The spliced and unspliced forms of *XBP1* mRNA are indicated. Splicing assays in are conducted in IRE1$\alpha^{-/-}$/IRE1$\beta^{-/-}$ MEFs reconstituted with hIRE1α-GFP or the hIRE1α-GFP-IF2^{L} mutant in the absence of doxycycline.

Figure 8- Figure Supplement 3

A. Confocal microscopy images of IRE1$\alpha^{-/-}$/IRE1$\beta^{-/-}$ MEFs reconstituted with hIRE1α-GFP. Around > 70% of wild type cells form foci upon ER stress induction (n=88). The cells are treated for 4 hrs by tunicamycin to induce ER stress before fixation and fluorescence imaging of hIRE1α-GFP (green) and DNA (blue, DAPI)

B. Confocal microscopy images of IRE1$\alpha^{-/-}$/IRE1$\beta^{-/-}$ MEFs reconstituted with hIRE1-GFP-IF2^{L} mutant. The cells are treated for 4 hrs by tunicamycin to induce ER stress before fixation and fluorescence imaging of hIRE1-GFP-IF2^{L} mutant (green) and DNA (blue, DAPI).

C. Live cell imaging with IRE1$\alpha^{-/-}$/IRE1$\beta^{-/-}$ MEFs reconstituted with hIRE1α-IF2^{L}-GFP. Imaging was performed after 4 hrs of ER stress induction (between 4-5 hrs of ER stress) by tunicamycin.

Figure 9 Model of human IRE1 activation

Apo-hIRE1α LD dimers are found in equilibrium between closed and open conformations (Step 1) (note that for simplicity IRE1's cytoplasmic kinase/RNase domains are not displayed in the model). Upon ER stress, unfolded proteins accumulating in the ER lumen bind hIRE1α LD. Unfolded protein binding stabilizes hIRE1α LD in the open conformation and induces a conformational change in the αB helix and the neighboring structural elements (Steps 2). This conformational change releases the block on oligomerization, thus leading to active hIRE1α oligomers by allowing the formation of an IF2^{L}-like interface in hIRE1α LD (Step 3). Oligomerization driven by hIRE1α LD subsequently activates its kinase and RNase domains. When protein-folding homeostasis is achieved, the dynamic hIRE1α LD oligomers re-adopt the inactive hIRE1α LD conformation (Step 4).

Table 1. List of cross-linked peptides in hIRE1α cLD detected by mass spectrometry and their relative abundance.

Video 1. The model displaying the transition of hIRE1α cLD from the "closed" to "open" state. The crystal structure of hIRE1α cLD is used to represent (pdb: 2hz6) the closed and the hIRE1α cLD structural model based on the yeast crystal structure (pdb: 2be1) represents the open state. The movie is generated using Pymol.

Materials and Methods

Reagents

Synthetic peptides were ordered from Elim Biosciences and GenScript at > 95% purity.

Protein Purification

To express MBP-hIRE1α cLD (aa 24-389), human IRE1α cDNA sequences were cloned into a pMalC2p vector to create a hIRE1α cLD fused on its N-terminus to MBP. To express His_{10}-hIRE1α cLD, hIRE1α cLD was cloned into pet16b(+) vector containing a FactorXa protease cleavage site. Additionally, His_{10}-hIRE1α cLD and IRE1 LD coding sequences were cloned into pet47b(+) vector with a preScission protease cleavage site. Hamster BiP with an N-terminal His_{10}-tag was cloned into pet16b(+) vector, which was modified to introduce a preScission protease site C-terminal to the His_{10}-tag. For expression of the proteins, the plasmid of interest was transformed into *Escherichia coli* strain BL21DE3* RIPL (Agilent Technologies) or Rosetta2 cells (Novagen). Cells were grown in Luria Broth at 37º C until OD_{600} = 0.6. Protein expression was induced with 0.3 mM IPTG, and cells were grown at 21º C overnight. For selective labeling, cells were grown according to published protocols (Tugarinov & Kay, 2004). Briefly, cells were grown at minimal media in D_2O supplemented with deuterated glucose as the primary carbon source. For purification, cells were resuspended in Lysis Buffer (50 mM HEPES pH 7.2, 400 mM NaCl, 4 mM dithiothreitol (DTT)(or 5 mM β-mercaptoethanol, if a nickel column was used)) and were lysed in an Avestin EmulsiFlex-C3 cell disruptor at 16,000 psi. The supernatant was collected after centrifugation for 40 min at 30,000xg. MBP-IRE1 cLD constructs were purified on an MBP-amylose resin (New England Biolabs) and eluted with 10 mM amylose in Elution Buffer (50 mM HEPES pH 7.2, 150 mM NaCl, 4 mM DTT) after washing the column with 20 column volumes of Lysis Buffer. The eluate was then diluted with 50 mM HEPES (pH 7.2) buffer to 50 mM NaCl and applied to a MonoQ ion exchange column and eluted with a linear gradient from 50 mM to 1 M NaCl. The protein was further purified on a Superdex 200 10/300 gel filtration column equilibrated with Buffer A (25 mM HEPES pH 7.2, 150 mM NaCl, 2 mM *tris*(2carboxyethyl)phosphine (TCEP). The initial purification of His_6- and His_{10}-hIRE1α cLD and His_{10}-BiP constructs were performed on a His-TRAP column (GE Healthcare), where the protein was eluted with gradient from 20 mM to 500 mM imidazole. The eluate was purified on a MonoQ column, before the His_6-tag (pet47b+) or His_{10}-tag (pet16b+) were removed by either PreScission protease (GE Healthcare, 1 unit of enzyme for 100 µg of protein) or FactorXa (NEB, 1 µg of FactorXa per 100 µg of protein), respectively. The tag removal was performed at 4º C overnight after the protein concentration was adjusted to 1 mg/mL. C_{H1} domain of IgG was purified under reducing conditions as described (Feige et al., 2009). Protein concentrations were determined using extinction coefficient at 280 nm predicted by the Expasy ProtParam tool (http://web.expasy.org/protparam/).

Peptide arrays

Peptide arrays were purchased from the MIT Biopolymers Laboratory. The tiling

arrays were composed of 18-mer peptides that were tiled along the CPY*, MPZ, insulin, lysozyme and PTIP sequences with a 3 amino acid shift between adjacent spots. In the mutational arrays, peptides were synthesized to systematically mutate each amino acid in the core region of the CPY*-derived peptide. The arrays were incubated in 100% methanol for 10 minutes, then in Binding Buffer (50 mM HEPES pH 7.2, 150 mM NaCl, 0.02% Tween-20, 2 mM DTT) three times for 10 min each. For BiP experiments, ADP and $MgCl_2$ were added to the binding buffer to final concentrations of 1 mM and 5 mM, respectively. The arrays were then incubated for 1 h at room temperature with 500 nM MBP-hIRE1α cLD or His_{10}-BiP and washed again three times with 10 min incubation in between the washes in Binding Buffer to remove any unbound protein. Using a semi-dry transfer apparatus, the bound protein was transferred to a nitrocellulose membrane and detected with anti-MBP antiserum (NEB) or anti-His_6 antibody (Abcam). The contribution of each amino acid to hIRE1α cLD and BiP binding was calculated as described previously (Gardner & Walter, 2011). The peptide arrays were quantified using Max Quant. The binding intensity in each spot was normalized to max signal intensity in the peptide array. The peptides with the top 10% binding scores were selected and the occurrence of each amino acid in these top-binding peptides was calculated. This value is normalized to their abundance in the arrays (Fig. 2a). To calculate experimental error, the amino acid occurrences of top binders were calculated for independent replicates. The statistical significance ($p < 0.05$) is determined using non-paired t-test by the Prism software (Fig. Supp. 2b).

Fluorescence Anisotropy

For fluorescence anisotropy measurements, MPZ1 peptide attached to 5-carboxyfluorescein (5-FAM) at its C-terminus was obtained at >95% purity from ELIM Biopharmaceuticals. For the remaining peptides (8ab1, MPZ1-N, MPZ1-M, MPZ1-C and MPZ1-N) derivatives were synthesized with 5-FAM attached to their N-terminus by GenScript at >95% purity. Binding affinities of hIRE1α cLD or IRE1 mutants to FAM-labeled peptides were measured by the change in fluorescence anisotropy on a Spectramax-M5 plate reader with excitation at 485 nm and emission at 525 nm with increasing concentrations of hIRE1α cLD. Fluorescently labeled peptides were used in a concentration range of 50-100 nM. The reaction volume of each data point was 20 µL and the measurements were performed in 384-well, black flat-bottomed plates after incubation of peptide with hIRE1α cLD or its mutants for 30 min at 25° C. Binding curves were fitted using Prism Software (GraphPad) using the following equation: $F_{bound} = r_{free} + (r_{max} - r_{free})/(1+10_{((LogK^{1/2}-X)\cdot n_H)})$, where F_{bound} is the fraction of peptide bound, r_{max} and r_{free} are the anisotropy values at maximum and minimum plateaus, respectively. n_H is the Hill coefficient and x is the concentration of the protein in log scale. Curvefitting is performed with minimal constraints to obtain $K_{1/2}$ values with high R_2 values. However, as this equation does not take into account the equilibria between hIRE1 cLD dimers/oligomers, these apperant $K_{1/2}$ values do not reflect the dissociation constant.

Microscale Thermophoresis Experiments (MST)

MST experiments were performed with a Monolith NT.115 instrument (NanoTemper Technologies, Germany). All experiments were done with the following buffer: 25 mM HEPES pH 7.2, 150 mM NaCl, 1 mM TCEP, 0.025% Tween-20. hIRE1α cLD

was labeled using the Monolith NT Protein labeling Kit Red-Maleimide. Labeled protein was used in the measurements at a concentration of 50 nM. It was mixed with equal volumes unlabeled interaction partner in two-fold serial dilutions. Hydrophilic-treated capillaries (NanoTemper Technologies) were used for all the measurements. All experiments were performed at 50 % LED power and 40-60-80% IR-laser at 25°C.

AUC Sedimentation Velocity Experiments

Sedimentation velocity experiments were carried out in a Beckman Optima XL-A analytical centrifuge at 40,000xg at 20º C with An-60 Ti rotor. All experiments were performed in buffer containing 25 mM HEPES pH 7.2, 150 mM NaCl, 2 mM DTT. Samples (400 µL) and reference buffer (410 µL) were loaded into AUC cells for each experiment. Samples of hIRE1α cLD at 5 µM were scanned at 280 nm, whereas hIRE1α cLD at concentrations higher than 25 µM were scanned at 290 nm to prevent detector saturation at high protein concentrations. Data analysis was performed using the SEDFIT software employing the c(s) method with time invariant and radial invariant noise fitting (Schuck, 2000). Buffer viscosity was calculated by Sednterp.

NMR experiments

NMR experiments were performed on an 800 MHz Bruker AVANCE-I spectrometer with a TXI Cryoprobe equipped with an actively shielded Z-gradient at 298.0 K. Samples were buffer-exchanged into 25 mM phosphate buffer pH 7.2, 150 mM NaCl and 2 mM DTT in 100% D_2O on Vivaspin columns (Millipore). The concentration of WT hIRE1α cLD and hIRE1α cLD mutants varied from 25 - 400 µM in 250 µL volume. Samples were placed in a Shigemi advanced NMR microtube. For peptide and unfolded protein binding experiments, the peptides were dissolved in the same buffer at high concentrations (1-2 mM) and titrated in 1:0.5, 1:1 and 1:2 molar ratios. Two-dimensional [^{13}C, ^{1}H]-HMQC methyl correlation experiments on $^{13}CH_3$–Ile hIRE1α cLD were acquired with 86* and 768* complex points in the ^{13}C and ^{1}H dimensions, respectively. All spectra were processed with TOPSPIN 3.2 and analyzed with Sparky.

Attachment of the spin label to MPZ1 peptide and PRE experiments

MPZ1 peptide at 200 µM was labeled with 3-(2-iodoacetamido)-proxyl (Sigma) at the single cysteine, Cys5 in 25 mM phosphate buffer pH 7.2, 150 mM NaCl in the presence of 2 mM spin-label at 4° C for 8 h. The labeled peptide was then dialyzed in a Slide-a-Lyzer dialysis cassette (Thermo Fisher Scientific) with 2 kDa cut-off to remove the excess spin-label and to exchange the buffer to deuterated buffer (25 mM phosphate buffer pH 7.2, 150 mM NaCl, 2mM DTT) for NMR experiments. Control samples used in the reference experiments contained (1-oxyl-2,2,5,5-tetramethylpyrroline-3-methyl) methanethiosulfonate spin-label that was treated the same way as the proxyl-labeled peptide. Wild type hIRE1α cLD and quadruple mutant hIRE1α cLD (Leu186Ile, Ile326/334/362Val) and single mutant Leu186Ile were used in PRE experiments at 75 µM and 100 µM protein concentration respectively, in the presence and absence of equimolar concentrations of MPZ1-proxyl peptide. We normalized the PRE effect with the surface exposed area displayed by that isoleucine to exclude possible contributions from non-specific

interactions with the spin label attached peptide (Clore & Iwahara, 2009). The normalized PRE values are calculated as follows, the solvent accessible surface area for isoleucines are calculated using the "GETAREA" webserver (http://curie.utmb.edu/getarea.html, (Fraczkiewicz & Braun, 1998)) based on hIRE1 α cLD crystal structure. The maximum solvent accessible surface by these isoleucines is normalized to 1 and the normalized values are multiplied with the PRE effect. The PRE effect is calculated by dividing the intensity of isoleucine signals in the control experiments with the isoleucine signals in the presence of MPZ1-proxyl peptide.

Cross-linking Experiments

10 μM, 20 μM and 50 μM hIRE1α cLD was incubated with 500 μM and 1mM BS3 cross-linker for 15 and 30 min at room temperature. Same reaction was performed for hIRE1 cLD pre-bound to 50 μM MPZ1-N for 30 min on ice. The reaction was stopped with the addition of 1M Tris-HCl at pH 8.0 at end concentration of 50 mM Tris-HCl, and incubated for 10 min at room temperature before running the SDS-PAGE gel.

LC-MS/MS analysis and cross-linked peptide identification

Cross-linked products were in-gel digested and analyzed by LC-MS and LC-MS-MS as described previously (Wu et al., 2013, Zeng-Elmore et al., 2014). Briefly, 1μl aliquot of the digestion mixture was injected into an Dionex Ultimate 3000 RSLCnano UHPLC system (Dionex Corporation, Sunnyvale, CA), and separated by a 75 μm × 25 cm PepMap RSLC column (100 Å, 2 μm) at a flow rate of ~450 nl/min. The eluant was connected directly to a nanoelectrospray ionization source of an LTQ Orbitrap XL mass spectrometer (Thermo Scientific, Waltham, MA). LC-MS data were acquired in an information-dependent acquisition mode, cycling between a MS scan (m/z 315-2,000) acquired in the Orbitrap, followed by low-energy CID analysis on 3 most intense multiply charged precursors acquired in the linear ion trap.

Cross-linked peptides were identified using an integrated module in Protein Prospector, based on a bioinformatic strategy described previously (Chu et al., 2010, Trnka et al., 2014). The score of a cross-linked peptide was based on number and types of fragment ions identified, as well as the sequence and charge state of the cross-linked peptide. Only results where the score difference is greater than 0 (i.e. the cross-linked peptide match was better than a single peptide match alone) are considered. Tandem MS spectra of cross-linked peptides were manually inspected to ensure data quality. With the threshold of peptide score and expectation value for oligomer-only cross-linked peptides, no decoy match was returned.

Lentiviral constructs and transduction

The coding sequence of wild type GFP-tagged IRE1 (Li et al., 2010) was amplified by PCR with Phsuion polymerase (NEB) and oligonucleotides with engineered restriction sites. introduced into the Gateway entry vector pSHUTTLE-CMV-TO (kind gift of A. Ashkenazi, Genentech and (Gray, Hoeflich et al., 2007) by PCR and

cloning into cognate KpnI and EcoRI sites. The IRE1-GFP- $IF2^L$ mutant was generated in pSHUTTLE-CMV-TO by site directed mutagenesis of the wild-type sequence. The resulting clones were recombined into pGpHUSH.puro (kind gift of A. Ashkenazi, Genentech and (Gray et al., 2007)), a single lentivirus expression vector that allows the doxycyline-regulatable (Tet^{ON}) expression of a gene-of-interest. VSV-G pseudotyped lentiviral particles were prepared using standard protocols using 293METR packaging cells (kind gift of Brian Ravinovich, formerly at MD Anderson Cancer Center,(Rabinovich, Li et al., 2006)). Viral supernatants were concentrated by filtration (Amicon Ultra centrifugal filter device, 100 kDa MWCO) and used to infect target cells by centrifugal inoculation (spinoculation) at 2000 rpm inn a Beckman GH3.8 rotor outfitted with plate carriers for 90 minutes in presence of 8ug/mL polybrene. The cells were left to recover overnight following infection and were then subjected to puromycin selection as described below.

Cell culture and generation of stable cell lines

IRE1 double-knockout Mouse Embryonic Fibroblasts (MEF) ($IRE1\alpha^{-/-}/IRE1\beta^{-/-}$) and wild-type MEFs are kind gift of D. Ron, University of Cambridge). Cells were not tested for the mycoplasma contamination. Cells were grown in DMEM supplemented with 10% fetal bovine serum, 2 mM L-glutamine, and penicillin/streptomycin. Lentiviral-transduced cells were selected with 6 μg/mL puromycin for 72 hr based on the puromycin concentration defined by the kill curve. Subsequently, a pulse of 25 nM doxycycline was given to induce expression of the GFP-tagged IRE1 transgenes for 10-12 hrs. The following day, the doxycycline was washed out and we selected pseudoclonal cell populations by fluorescent activated cell sorting based on GFP expression for both wild-type (hIRE1α-GFP) and $IF2^L$ mutant (hIRE1α-GFP-$IF2^L$ mutant) forms of IRE1. The cells were selected in a FACS Aria instrument (BD FACSAria3), gating for very narrow GFP expressing populationThis procedure ensures selection of a pseudoclonal population where most cells have similar levels of expression of the transgene of interest while avoiding typical problems associated from monoclonal selection of IRE1-expressing cells; namely an aberrant UPR. The pseudoclonal populations were expanded and frozen as source stocks for experiments.

Live Cell Imaging of hIRE1α -GFP and hIRE1α-GFP-$IF2^L$ mutant

IRE1 double-knockout MEFs ($IRE1\alpha^{-/-}/IRE1\beta^{-/-}$) reconstituted with of hIRE1α -GFP and hIRE1α-GFP-$IF2^L$ mutant were split 2 days before imaging onto ibiTreat dishes (ibidi) at 5×10^4 cells/dish. 25 nM Doxyccline containing medium was added for 10-12 h, withdrawn before imaging and replaced with imaging media consisting of Fluorobrite DMEM (Thermo Scientific), 2.5 % FBS, and 5 mM Hepes at a pH of 7. Cells were imaged at 37 degrees Celsius on a spinning disk confocal with Yokogawa CSUX A1 scan head, Andor iXon EMCCD camera and 40x Plan Apo air Objective NA 0.95 with a 1.5x tube lens for additional magnification giving 60x final or 100X objective. Images were acquired using 488 nm laser at a rate of one frame per 3 minutes with 300 ms exposure time for each time point for an hour after different time points following induction of ER stress by tunicamycin (5 μg/mL) or thapsigargin (100 nM).

Immunofluorescence of hIRE1α -GFP and hIRE1α-GFP-$IF2^L$ mutant

IRE1 double-knockout MEFs ($IRE1\alpha^{-/-}/IRE1\beta^{-/-}$) reconstituted with of hIRE1α -GFP and hIRE1α-GFP-$IF2^{L}$ mutant were grown similar to live cell imaging experiments. After stress induction at various time points, cells were washed 3 times with PBS followed by 3 min fixation with 100% ETOH, followed by 3 times 5 minutes washes with PBS. As these fixation conditions kept GFP intact, they did not require immunostaining of hIRE1α for fluorescence imaging. DAPI staining is performed according to manufacturer's instructions (Thermo Fisher).

Generation of cDNA and semi-quantitative PCR

Cells exposed to DMSO or thapsigargin (100 nM) or tunicamycin (5 µg/ml) were collected in 0.5 ml of TRIzol reagent (Life Technologies) from a 6 well dish and total RNA was extracted following the manufacturer's recommendations. To generate cDNAs, 500 ng of total RNA were reverse transcribed using the SuperScript VILO system (Life Technologies) following the manufacturer's recommendations. The resulting 20 µl reverse transcription reactions were diluted to 10 times to 200 µl with 10 mM Tris– HCl pH 8.2, and 1% of this dilution was used for multiplex semiquantitative PCR. The multiplex PCR was set up using 1 µM of the forward reverese primers, 0.4 units of Taq DNA polymerase (Thermo Scientific), 0.2 mM of each dNTP, and 1.5 mM MgCl2, in a 20 µl reaction using the following buffer system: 75 mM Tris– HCl pH 8.8, 20 mM (NH4)SO4, and 0.01% Tween-20. The oligonucleotide sequences are the following: Hs_XBP1_Fwd: 50 -GGAGTT AAGACAGCGCTTGG-30 ; Hs_XBP1_Rev: 50 -ACTGGGTCCAAGTTG TCCAG-30. PCR products were amplified for 28 cycles and resolved on 3% agarose gels (1:1 mixture of regular and low-melting point agarose) stained with ethidium bromide.

Protein analysis by Western-Blot

Cells were lysed in SDS-PAGE loading buffer (1% SDS, 62.5 mM Tris-HCl pH 6.8, 10% glycerol). Lysates were sonicated and equal amounts were loaded on SDS-PAGE gels (BioRad, Hercules, CA). Proteins were transferred onto nitrocellulose and probed with primary antibodies diluted in Phosphate-buffered saline supplemented with 0.1% Tween 20 and 5% bovine serum albumin at 4 ^{o}C, overnight. The following antibodies were used: anti-IRE1 (1:1000) (14C10, Cell Signaling Technology, Danvers, MA), anti-GAPDH (1:1000) (14C10, Cell Signaling Technology, Danvers, MA and anti-phosho IRE1 antibody(1:500). IRE1 anti-phospho antibody is a kind gift of Avi Ashkenazi's group at Genentech. An HRP-conjugated secondary antibody (Amersham, Piscataway, NJ) was employed to detect immunereactive bands using enhanced chemiluminescence (SuperSignal; Thermo Scientific, Waltham, MA) detected by Li-Cor instrument (Li-Core Biosciences).

References

(2010) Modulators of HSP90. Nature Reviews Drug Discovery 9: 425-425

Acosta-Alvear D, Zhou Y, Blais A, Tsikitis M, Lents NH, Arias C, Lennon CJ, Kluger Y, Dynlacht BD (2007) XBP1 controls diverse cell type- and condition-specific transcriptional regulatory networks. Mol Cell 27: 53-66

1398 Aragon T, van Anken E, Pincus D, Serafimova IM, Korennykh AV, Rubio CA,
1399 Walter P (2009) Messenger RNA targeting to endoplasmic reticulum stress
1400 signalling sites. Nature 457: 736-740

1401 Balchin D, Hayer-Hartl M, Hartl FU (2016) In vivo aspects of protein folding and
1402 quality control. Science 353: aac4354

1403 Bertolotti A, Wang X, Novoa I, Jungreis R, Schlessinger K, Cho JH, West AB, Ron D
1404 (2001) Increased sensitivity to dextran sodium sulfate colitis in IRE1beta-
1405 deficient mice. The Journal of clinical investigation 107: 585-593

1406 Bertolotti A, Zhang Y, Hendershot LM, Harding HP, Ron D (2000) Dynamic
1407 interaction of BiP and ER stress transducers in the unfolded-protein response.
1408 Nature cell biology 2: 326-332

1409 Bi M, Naczki C, Koritzinsky M, Fels D, Blais J, Hu N, Harding H, Novoa I, Varia M,
1410 Raleigh J, Scheuner D, Kaufman RJ, Bell J, Ron D, Wouters BG, Koumenis C (2005)
1411 ER stress-regulated translation increases tolerance to extreme hypoxia and
1412 promotes tumor growth. The EMBO journal 24: 3470-3481

1413 Blond-Elguindi S, Cwirla SE, Dower WJ, Lipshutz RJ, Sprang SR, Sambrook JF,
1414 Gething MJ (1993) Affinity panning of a library of peptides displayed on
1415 bacteriophages reveals the binding specificity of BiP. Cell 75: 717-28

1416 Bukau B, Weissman J, Horwich A (2006) Molecular chaperones and protein
1417 quality control. Cell 125: 443-451

1418 Carrara M, Prischi F, Nowak PR, Kopp MC, Ali MM (2015) Noncanonical binding
1419 of BiP ATPase domain to Ire1 and Perk is dissociated by unfolded protein CH1 to
1420 initiate ER stress signaling. Elife 4

1421 Chu F, Baker PR, Burlingame AL, Chalkley RJ (2010) Finding chimeras: a
1422 bioinformatics strategy for identification of cross-linked peptides. Mol Cell
1423 Proteomics 9: 25-31

1424 Clore GM, Iwahara J (2009) Theory, practice, and applications of paramagnetic
1425 relaxation enhancement for the characterization of transient low-population
1426 states of biological macromolecules and their complexes. Chem Rev 109: 4108-
1427 39

1428 Cox JS, Shamu CE, Walter P (1993) Transcriptional induction of genes encoding
1429 endoplasmic reticulum resident proteins requires a transmembrane protein
1430 kinase. Cell 73: 1197-1206

1431 Cox JS, Walter P (1996) A novel mechanism for regulating activity of a
1432 transcription factor that controls the unfolded protein response. Cell 87: 391-
1433 404

1434 Credle JJ, Finer-Moore JS, Papa FR, Stroud RM, Walter P (2005) On the
1435 mechanism of sensing unfolded protein in the endoplasmic reticulum.

1436 Proceedings of the National Academy of Sciences of the United States of America
1437 102: 18773-18784

1438 Deuerling E, Patzelt H, Vorderwulbecke S, Rauch T, Kramer G, Schaffitzel E, Mogk
1439 A, Schulze-Specking A, Langen H, Bukau B (2003) Trigger Factor and DnaK
1440 possess overlapping substrate pools and binding specificities. Mol Microbiol 47:
1441 1317-28

1442 Feige MJ, Groscurth S, Marcinowski M, Shimizu Y, Kessler H, Hendershot LM,
1443 Buchner J (2009) An unfolded CH1 domain controls the assembly and secretion
1444 of IgG antibodies. Molecular cell 34: 569-579

1445 Feldman DE, Chauhan V, Koong AC (2005) The unfolded protein response: a
1446 novel component of the hypoxic stress response in tumors. Molecular cancer
1447 research : MCR 3: 597-605

1448 Flynn GC, Pohl J, Flocco MT, Rothman JE (1991) Peptide-binding specificity of the
1449 molecular chaperone BiP. Nature 353: 726-30

1450 Fonseca SG, Burcin M, Gromada J, Urano F (2009) Endoplasmic reticulum stress
1451 in beta-cells and development of diabetes. Curr Opin Pharmacol 9: 763-70

1452 Fraczkiewicz R, Braun W (1998) Exact and efficient analytical calculation of the
1453 accessible surface areas and their gradients for macromolecules. J Comput Chem
1454 19: 319-333

1455 Fremont DH, Matsumura M, Stura EA, Peterson PA, Wilson IA (1992) Crystal
1456 structures of two viral peptides in complex with murine MHC class I H-2Kb.
1457 Science 257: 919-27

1458 Gaponenko V, Howarth JW, Columbus L, Gasmi-Seabrook G, Yuan J, Hubbell WL,
1459 Rosevear PR (2000) Protein global fold determination using site-directed spin
1460 and isotope labeling. Protein Sci 9: 302-9

1461 Gardner BM, Pincus D, Gotthardt K, Gallagher CM, Walter P (2013) Endoplasmic
1462 reticulum stress sensing in the unfolded protein response. Cold Spring Harb
1463 Perspect Biol 5: a013169

1464 Gardner BM, Walter P (2011) Unfolded Proteins Are Ire1-Activating Ligands that
1465 Directly Induce the Unfolded Protein Response. Science (New York, NY)

1466 Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O'Shea
1467 EK, Weissman JS (2003) Global analysis of protein expression in yeast. Nature
1468 425: 737-41

1469 Gillespie JR, Shortle D (1997) Characterization of long-range structure in the
1470 denatured state of staphylococcal nuclease. II. Distance restraints from
1471 paramagnetic relaxation and calculation of an ensemble of structures. J Mol Biol
1472 268: 170-84

1473 Gottstein D, Reckel S, Dotsch V, Guntert P (2012) Requirements on paramagnetic
1474 relaxation enhancement data for membrane protein structure determination by
1475 NMR. Structure 20: 1019-27

1476 Gray DC, Hoeflich KP, Peng L, Gu Z, Gogineni A, Murray LJ, Eby M, Kljavin N,
1477 Seshagiri S, Cole MJ, Davis DP (2007) pHUSH: a single vector system for
1478 conditional gene expression. BMC Biotechnol 7: 61

1479 Harding HP, Zhang Y, Bertolotti A, Zeng H, Ron D (2000) Perk is essential for
1480 translational regulation and cell survival during the unfolded protein response.
1481 Molecular cell 5: 897-904

1482 Heifetz A, Keenan RW, Elbein AD (1979) Mechanism of action of tunicamycin on
1483 the UDP-GlcNAc:dolichyl-phosphate Glc-NAc-1-phosphate transferase.
1484 Biochemistry 18: 2186-92

1485 Hetz C, Thielen P, Matus S, Nassif M, Court F, Kiffin R, Martinez G, Cuervo AM,
1486 Brown RH, Glimcher LH (2009) XBP-1 deficiency in the nervous system protects
1487 against amyotrophic lateral sclerosis by increasing autophagy. Genes &
1488 development 23: 2294-2306

1489 Hollien J, Lin JH, Li H, Stevens N, Walter P, Weissman JS (2009) Regulated Ire1-
1490 dependent decay of messenger RNAs in mammalian cells. The Journal of cell
1491 biology 186: 323-331

1492 Hollien J, Weissman JS (2006) Decay of endoplasmic reticulum-localized mRNAs
1493 during the unfolded protein response. Science (New York, NY) 313: 104-107

1494 Hotamisligil GS (2010) Endoplasmic reticulum stress and the inflammatory basis
1495 of metabolic disease. Cell 140: 900-917

1496 Jerabek-Willemsen M, Wienken CJ, Braun D, Baaske P, Duhr S (2011) Molecular
1497 interaction studies using microscale thermophoresis. Assay Drug Dev Technol 9:
1498 342-53

1499 Karagoz GE, Duarte AM, Akoury E, Ippel H, Biernat J, Moran Luengo T, Radli M,
1500 Didenko T, Nordhues BA, Veprintsev DB, Dickey CA, Mandelkow E, Zweckstetter
1501 M, Boelens R, Madl T, Rudiger SG (2014) Hsp90-Tau complex reveals molecular
1502 basis for specificity in chaperone action. Cell 156: 963-74

1503 Kimata Y, Oikawa D, Shimizu Y, Ishiwata-Kimata Y, Kohno K (2004) A role for BiP
1504 as an adjustor for the endoplasmic reticulum stress-sensing protein Ire1. The
1505 Journal of cell biology 167: 445-456

1506 Korennykh AV, Egea PF, Korostelev AA, Finer-Moore J, Zhang C, Shokat KM,
1507 Stroud RM, Walter P (2009) The unfolded protein response signals through high-
1508 order assembly of Ire1. Nature 457: 687-693

1509 Lee AH, Iwakoshi NN, Glimcher LH (2003) XBP-1 regulates a subset of
1510 endoplasmic reticulum resident chaperone genes in the unfolded protein
1511 response. Molecular and cellular biology 23: 7448-7459

1512 Leitner A, Walzthoeni T, Kahraman A, Herzog F, Rinner O, Beck M, Aebersold R
1513 (2010) Probing native protein structures by chemical cross-linking, mass
1514 spectrometry, and bioinformatics. Mol Cell Proteomics 9: 1634-49

1515 Li H, Korennykh AV, Behrman SL, Walter P (2010) Mammalian endoplasmic
1516 reticulum stress sensor IRE1 signals by dynamic clustering. Proceedings of the
1517 National Academy of Sciences of the United States of America 107: 16113-16118

1518 Lin JH, Li H, Yasumura D, Cohen HR, Zhang C, Panning B, Shokat KM, Lavail MM,
1519 Walter P (2007) IRE1 signaling affects cell fate during the unfolded protein
1520 response. Science (New York, NY) 318: 944-949

1521 Lu M, Lawrence DA, Marsters S, Acosta-Alvear D, Kimmig P, Mendez AS, Paton
1522 AW, Paton JC, Walter P, Ashkenazi A (2014) Cell death. Opposing unfolded-
1523 protein-response signals converge on death receptor 5 to control apoptosis.
1524 Science (New York, NY) 345: 98-101

1525 Marcinowski M, Holler M, Feige MJ, Baerend D, Lamb DC, Buchner J (2011)
1526 Substrate discrimination of the chaperone BiP by autonomous and cochaperone-
1527 regulated conformational transitions. Nature structural & molecular biology 18:
1528 150-158

1529 Matsumura M, Fremont DH, Peterson PA, Wilson IA (1992) Emerging principles
1530 for the recognition of peptide antigens by MHC class I molecules. Science 257:
1531 927-34

1532 Matus S, Nassif M, Glimcher LH, Hetz C (2009) XBP-1 deficiency in the nervous
1533 system reveals a homeostatic switch to activate autophagy. Autophagy 5: 1226-
1534 1228

1535 Mimura N, Fulciniti M, Gorgun G, Tai YT, Cirstea D, Santo L, Hu Y, Fabre C, Minami
1536 J, Ohguchi H, Kiziltepe T, Ikeda H, Kawano Y, French M, Blumenthal M, Tam V,
1537 Kertesz NL, Malyankar UM, Hokenson M, Pham T et al. (2012) Blockade of XBP1
1538 splicing by inhibition of IRE1alpha is a promising therapeutic option in multiple
1539 myeloma. Blood 119: 5772-5781

1540 Niwa M, Sidrauski C, Kaufman RJ, Walter P (1999) A role for presenilin-1 in
1541 nuclear accumulation of Ire1 fragments and induction of the mammalian
1542 unfolded protein response. Cell 99: 691-702

1543 Oikawa D, Kimata Y, Kohno K, Iwawaki T (2009) Activation of mammalian
1544 IRE1alpha upon ER stress depends on dissociation of BiP rather than on direct
1545 interaction with unfolded proteins. Experimental cell research 315: 2496-2504

1546 Otero JH, Lizak B, Hendershot LM (2010) Life and death of a BiP substrate.
1547 Seminars in cell & developmental biology 21: 472-478

1548 Pincus D, Chevalier MW, Aragon T, van Anken E, Vidal SE, El-Samad H, Walter P
1549 (2010) BiP binding to the ER-stress sensor Ire1 tunes the homeostatic behavior
1550 of the unfolded protein response. PLoS biology 8: e1000415

1551 Rabinovich B, Li J, Wolfson M, Lawrence W, Beers C, Chalupny J, Hurren R,
1552 Greenfield B, Miller R, Cosman D (2006) NKG2D splice variants: a reexamination
1553 of adaptor molecule associations. Immunogenetics 58: 81-8

1554 Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated
1555 protein structure and function prediction. Nat Protoc 5: 725-38

1556 Rudiger S, Buchberger A, Bukau B (1997) Interaction of Hsp70 chaperones with
1557 substrates. Nature structural biology 4: 342-349

1558 Rudiger S, Germeroth L, Schneider-Mergener J, Bukau B (1997) Substrate
1559 specificity of the DnaK chaperone determined by screening cellulose-bound
1560 peptide libraries. The EMBO journal 16: 1501-1507

1561 Scheuner D, Song B, McEwen E, Liu C, Laybutt R, Gillespie P, Saunders T, Bonner-
1562 Weir S, Kaufman RJ (2001) Translational control is required for the unfolded
1563 protein response and in vivo glucose homeostasis. Molecular cell 7: 1165-1176

1564 Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation
1565 velocity ultracentrifugation and lamm equation modeling. Biophys J 78: 1606-19

1566 Sidrauski C, Walter P (1997) The transmembrane kinase Ire1p is a site-specific
1567 endonuclease that initiates mRNA splicing in the unfolded protein response. Cell
1568 90: 1031-1039

1569 Street TO, Lavery LA, Agard DA (2011) Substrate binding drives large-scale
1570 conformational changes in the Hsp90 molecular chaperone. Mol Cell 42: 96-105

1571 Sung SC, Chao CY, Jeng KS, Yang JY, Lai MM (2009) The 8ab protein of SARS-CoV
1572 is a luminal ER membrane-associated protein and induces the activation of ATF6.
1573 Virology 387: 402-13

1574 Tirasophon W, Lee K, Callaghan B, Welihinda A, Kaufman RJ (2000) The
1575 endoribonuclease activity of mammalian IRE1 autoregulates its mRNA and is
1576 required for the unfolded protein response. Genes Dev 14: 2725-36

1577 Trnka MJ, Baker PR, Robinson PJ, Burlingame AL, Chalkley RJ (2014) Matching
1578 cross-linked peptide spectra: only as good as the worse identification. Mol Cell
1579 Proteomics 13: 420-34

1580 Tsuru A, Fujimoto N, Takahashi S, Saito M, Nakamura D, Iwano M, Iwawaki T,
1581 Kadokura H, Ron D, Kohno K (2013) Negative feedback by IRE1beta optimizes
1582 mucin production in goblet cells. Proc Natl Acad Sci U S A 110: 2864-9

1583 Tugarinov V, Hwang PM, Kay LE (2004) Nuclear magnetic resonance
1584 spectroscopy of high-molecular-weight proteins. Annual Review of Biochemistry
1585 73: 107-146

1586 Tugarinov V, Kay LE (2004) An isotope labeling strategy for methyl TROSY
1587 spectroscopy. Journal of Biomolecular NMR 28: 165-172

1588 Tugarinov V, Sprangers R, Kay LE (2007) Probing side-chain dynamics in the
1589 proteasome by relaxation violated coherence transfer NMR spectroscopy.
1590 Journal of the American Chemical Society 129: 1743-1750

1591 Vidal RL, Figueroa A, Court FA, Thielen P, Molina C, Wirth C, Caballero B, Kiffin R,
1592 Segura-Aguilar J, Cuervo AM, Glimcher LH, Hetz C (2012) Targeting the UPR
1593 transcription factor XBP1 protects against Huntington's disease through the
1594 regulation of FoxO1 and autophagy. Human molecular genetics 21: 2245-2262

1595 Walter P, Ron D (2011) The unfolded protein response: from stress pathway to
1596 homeostatic regulation. Science (New York, NY) 334: 1081-1086

1597 Walter P., D. R (2011) The unfolded protein response: from stress pathway to
1598 homeostatic regulation. Science 334: 1081-6

1599 Wilson IA, Fremont DH (1993) Structural analysis of MHC class I molecules with
1600 bound peptide antigens. Semin Immunol 5: 75-80

1601 Wu B, Peisley A, Richards C, Yao H, Zeng X, Lin C, Chu F, Walz T, Hur S (2013)
1602 Structural basis for dsRNA recognition, filament formation, and antiviral signal
1603 activation by MDA5. Cell 152: 276-89

1604 Yoshida H, Haze K, Yanagi H, Yura T, Mori K (1998) Identification of the cis-
1605 acting endoplasmic reticulum stress response element responsible for
1606 transcriptional induction of mammalian glucose-regulated proteins. Involvement
1607 of basic leucine zipper transcription factors. The Journal of biological chemistry
1608 273: 33741-33749

1609 Yoshida H, Matsui T, Yamamoto A, Okada T, Mori K (2001) XBP1 mRNA is
1610 induced by ATF6 and spliced by IRE1 in response to ER stress to produce a
1611 highly active transcription factor. Cell 107: 881-891

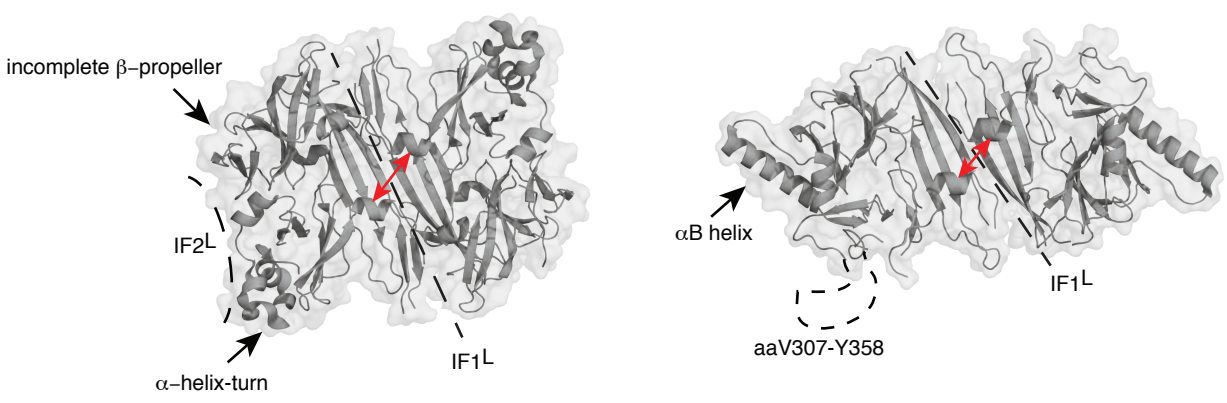
1612 Zeng-Elmore X, Gao XZ, Pellarin R, Schneidman-Duhovny D, Zhang XJ, Kozacka
1613 KA, Tang Y, Sali A, Chalkley RJ, Cote RH, Chu F (2014) Molecular architecture of
1614 photoreceptor phosphodiesterase elucidated by chemical cross-linking and
1615 integrative modeling. J Mol Biol 426: 3713-28

1616 Zhang K, Kaufman RJ (2008) From endoplasmic-reticulum stress to the
1617 inflammatory response. Nature 454: 455-462

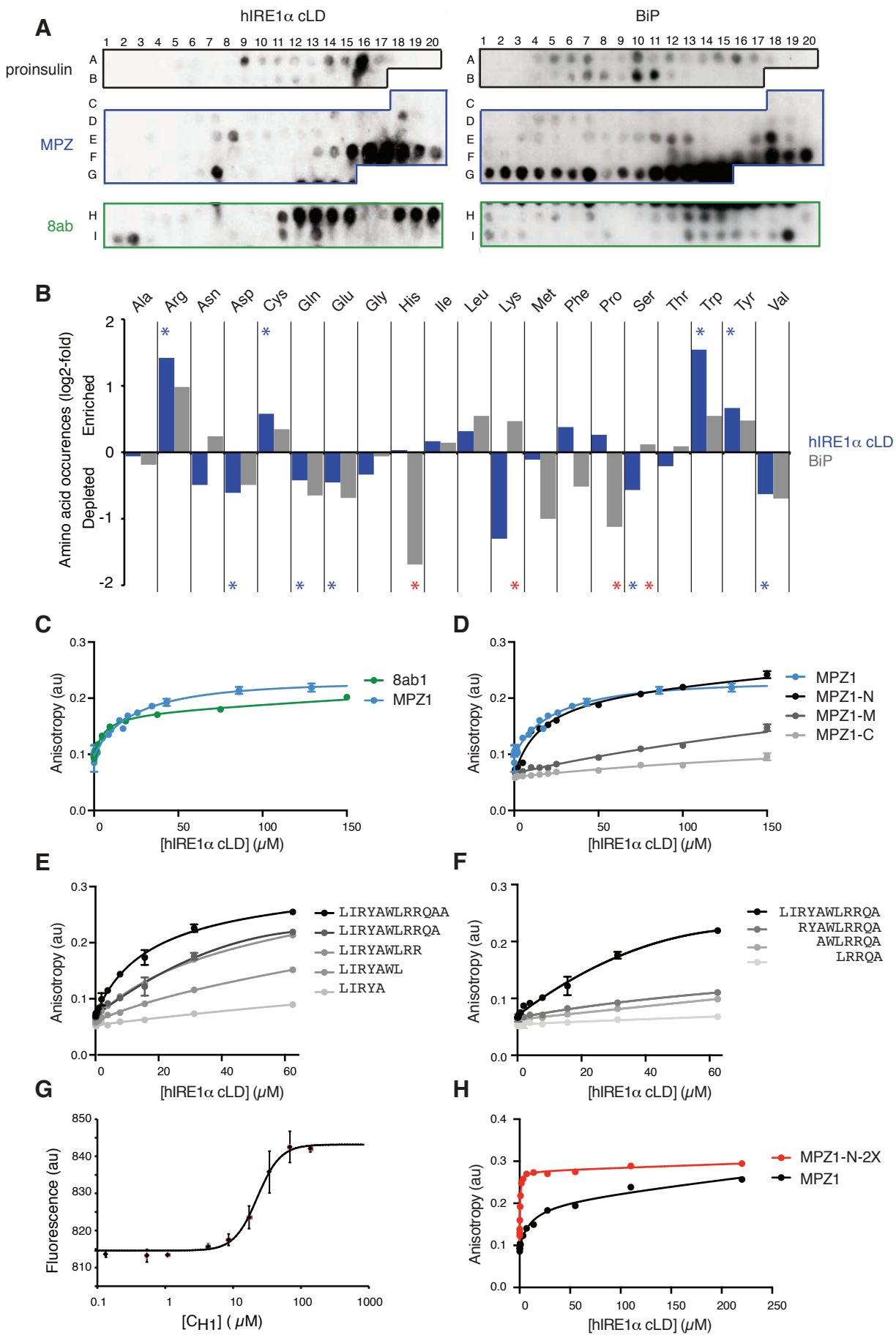
1618 Zhang Y (2008) I-TASSER server for protein 3D structure prediction. BMC
1619 Bioinformatics 9: 40

1620 Zhou J, Liu CY, Back SH, Clark RL, Peisach D, Xu Z, Kaufman RJ (2006) The crystal
1621 structure of human IRE1 luminal domain reveals a conserved dimerization
1622 interface required for activation of the unfolded protein response. Proceedings of
1623 the National Academy of Sciences of the United States of America 103: 14343-
1624 14348
1625

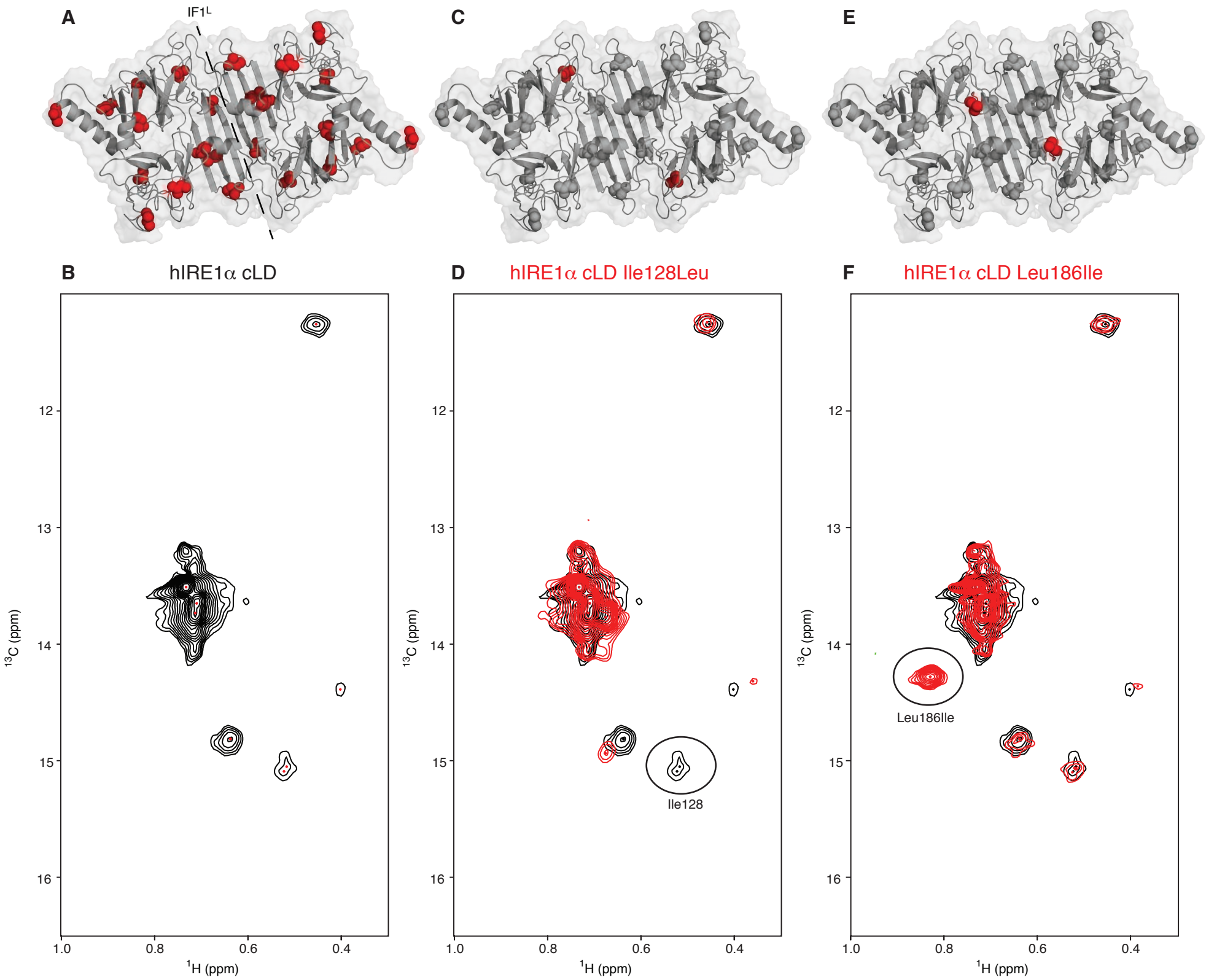
Karagöz et al., Figure 1



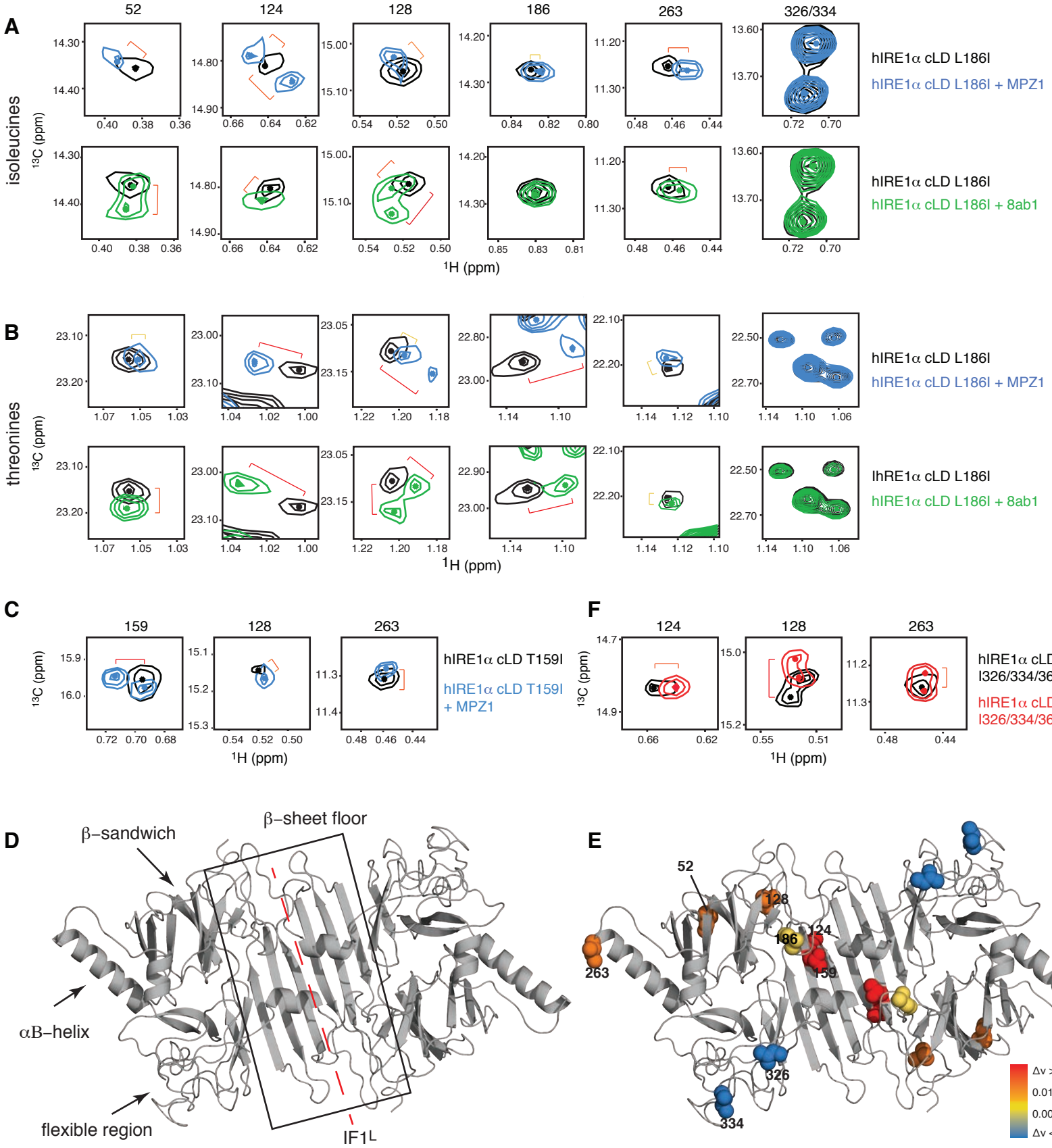
Karagöz et al., Figure 2



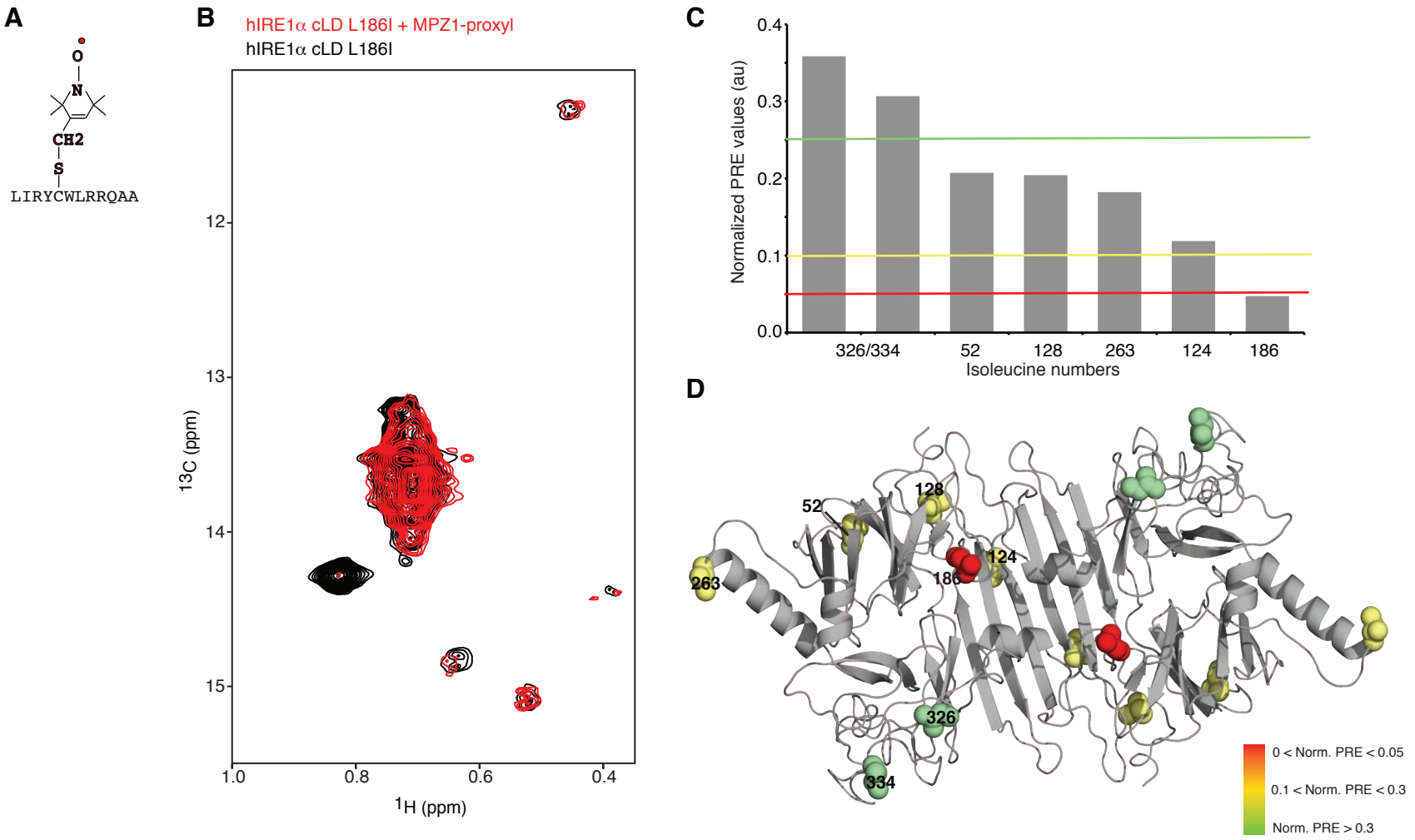
Karagöz et al., Figure 3



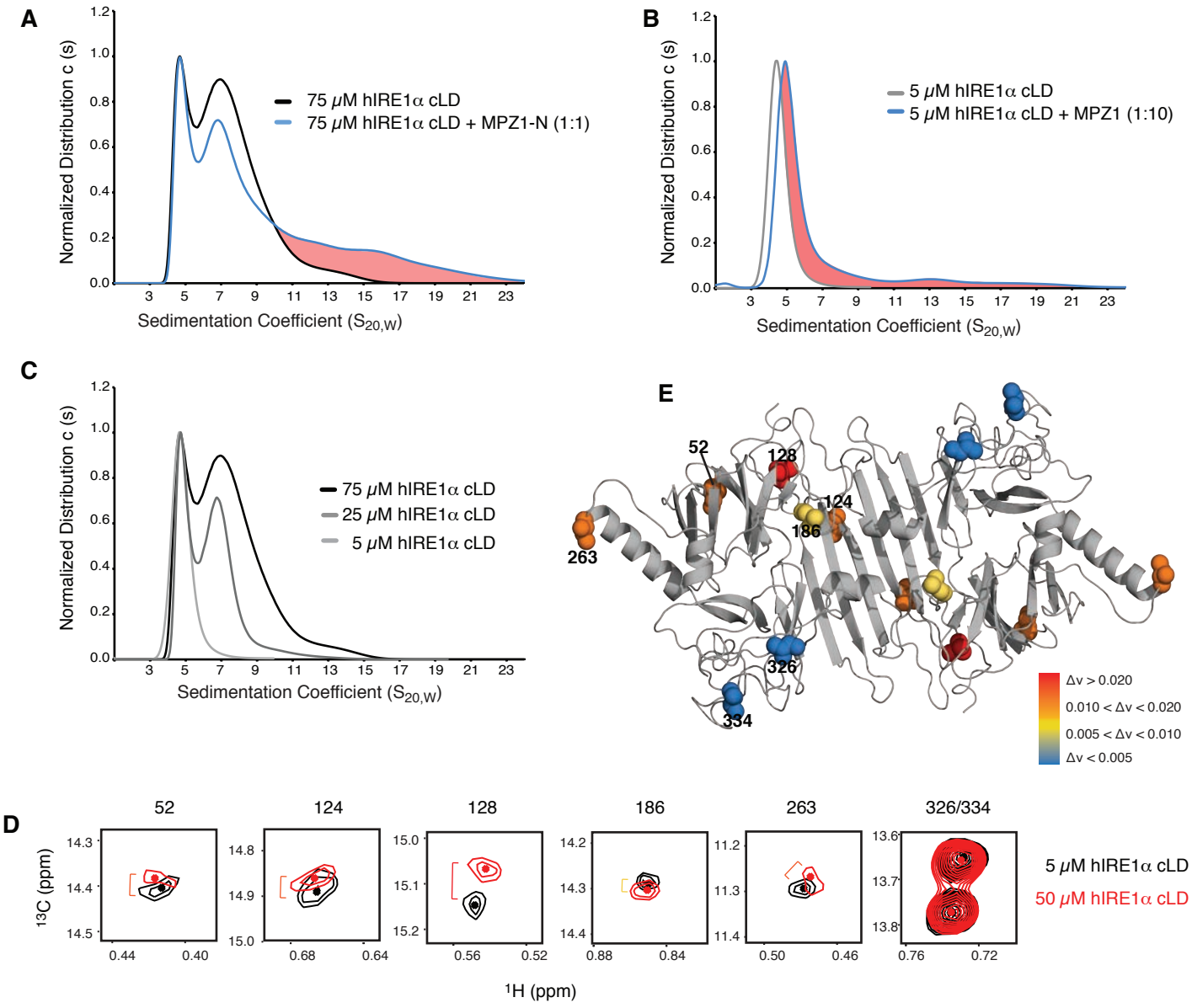
Karagöz et al., Figure 4



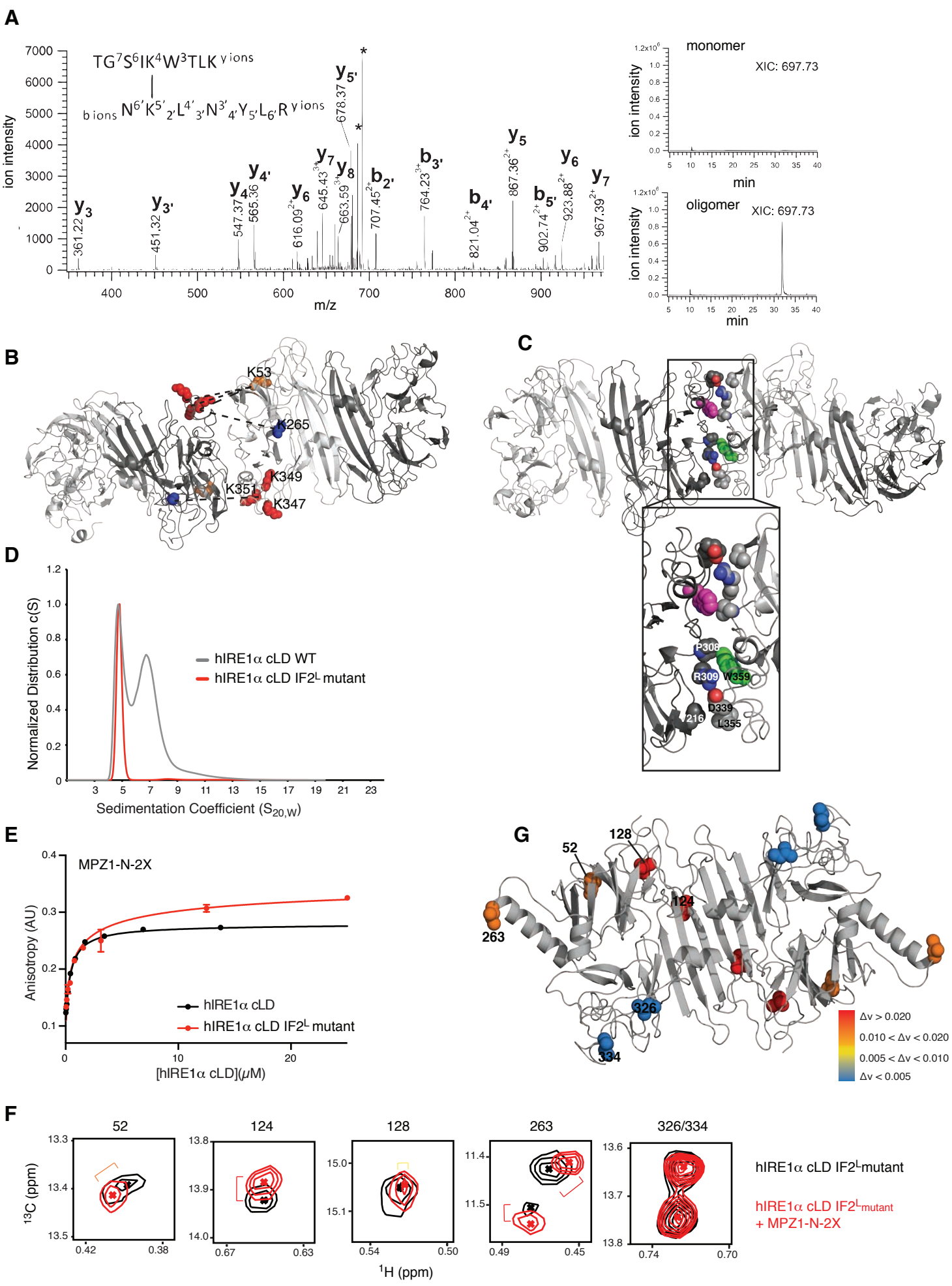
Karagöz et al., Figure 5



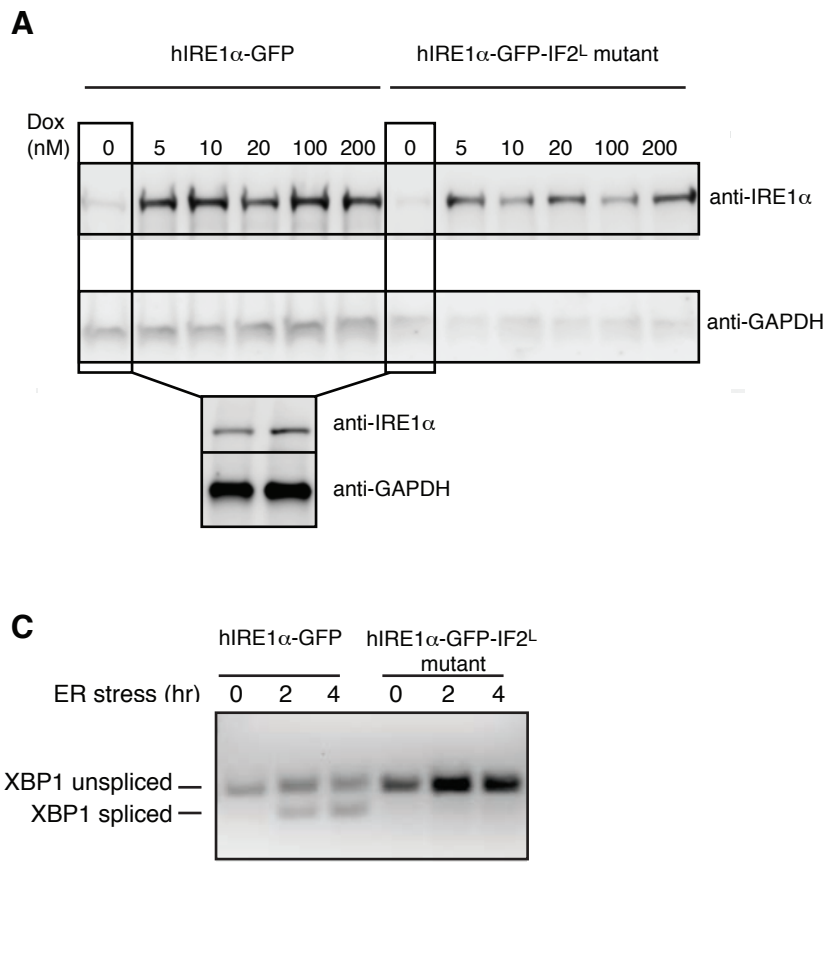
Karagöz et al., Figure 6



Karagöz et al., Figure 7



Karagöz et al., Figure 8



Karagöz et al., Figure 9