Large Physiochemical Modeling and Analysis of the Mammalian Unfolded Protein Response

Anirikh Chakrabarti[1], **Lina Aboulmouna**[2], **and Jeffrey Varner**[1,*]

[1]Cornell University, School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, NY 14853, USA
[2]Purdue University, School of Chemical Engineering, West Lafayette, IN 47907, USA
[*]jdv27@cornell.edu

ABSTRACT

Cellular stressors routinely lead to malfunctions in the folding state of critical proteins, thereby playing a vital role in disease states like cancer, diabetes and cardiovascular ailments. Newly synthesized proteins must fold and assemble into unique three-dimensional structures, in order to become functionally active. Cells monitor protein folding by an inbuilt quality-control system involving both the Endoplasmic Reticulum (ER) and the Golgi apparatus. Incorrectly folded proteins are tagged for degradation via ER associated degradation (ERAD) or sent back through a refolding cycle. However, accumulation of incorrectly folded proteins can trigger a cascade of events, termed the Unfolded Protein Response (UPR). UPR leads to re-establishing the cellular homeostasis. In this study, we developed a family of mechanistic models of eukaryotic UPR, which was composed of a system of ordinary differential equations. The objective of this study was to assemble a series of molecular modules describing different aspects of UPR and subsequently analyze these models for fragility and robustness. Kinetic parameters for these models were estimated by comparing simulations with experimental data. Using POETs and a cross-validation scheme, we developed an ensemble of models, consistent with literature data. Model analysis highlighted the presence of a sequential order for firing of the ER stress transducers in UPR. The counteracting effects of the ER stress transducers converge at feedback regulation of the molecular chaperone BiP. The three main regulators of feedback, ATF4, cleaved ATF6 and XBP1s share the load ensuring the phased response of UPR. However these regulators add on to the overall fragility of the system and allow scope for manipulation of UPR as seen by sensitivity/robustness analysis. Downstream effects of UPR include a balance between cell survival and cell death based on the magnitude of the stress. Model analysis suggested that the cell-death axis was relatively robust owing to redundant routes e.g., APAF-1 dependent and APAF-1 independent routes of apoptosis. However the cell-survival axis was relatively susceptible to perturbations owing to multiple levels of regulation, thereby highlighting importance of key proteins like CHOP.

Introduction

Protein folding is strategically important to cellular function in all organisms. In eukaryotes, secreted, membrane-bound and organelle-targeted proteins are typically processed and folded in the endoplasmic reticulum (ER)[1–3]. Intracellular perturbations caused by a variety of stressors disturb the specialized environment of the ER leading to the accumulation of misfolded or unfolded proteins[4,5]. Shifts in folding capacity have been associated with diseases such as cancer, diabetes and cardiovascular disorders[4]. Physiological processes such as aging can also influence protein folding. Normally, cells ensure proper protein folding using a combination of molecular chaperones, foldases and lectins[1]. However, when proper folding can not be restored, unfolded or misfolded proteins are targeted to ER Associated Degradation (ERAD) pathways for processing[3]. If unfolded or misfolded proteins continue to accumulate, eukaryotes induce the unfolded protein response (UPR). In mammalian cells, UPR is a complex signaling program mediated by three ER transmembrane receptors: activating transcription factor 6 (ATF6), inositol requiring kinase 1 (IRE1) and double-stranded RNA-activated protein kinase (PKR)-like endoplasmic reticulum kinase (PERK). UPR performs three functions, adaptation, alarm and apoptosis. During adaptation, the UPR tries to reestablish folding homeostasis by inducing the expression of chaperones that enhance protein folding. Simultaneously, translation is globally attenuated to reduce the ER folding load while the degradation of unfolded proteins is increased. If these steps fail, the UPR induces a cellular alarm and apoptosis program. The alarm phase involves several signal transduction events, ultimately leading to the removal of the translational block and the down-regulation of the expression and activity of pro-survival factors such as the B-cell lymphoma 2 (Bcl2) protein. After the alarm phase, cells can undergo apoptosis, although ER stress can also initiate autophagy[6–12]. Thus, ER folding homeostasis strongly influences mammalian physiology[5].

In this study we used physiochemical computer modeling as a tool to study the mammalian unfolded protein response (UPR). Physiochemical models which describe system-level responses can prioritize experimental directions, generate testable hypothesis, and perhaps identify and validate potential therapeutic targets[13]. For mammalian UPR, physiochemical modeling

might tell us the critical components of each branch, and ultimately what must be manipulated to get a desired network response, for example, enhanced death or survival. However, there are critical limitations with physiochemical models. Foremost amongst these is the large number of unknown model parameters and uncertainty in the model structure. It is typically impossible to uniquely identify model parameters, even with extensive training data and perfect models[14]. Alternatively, ensemble approaches, which use uncertain model families, have emerged as a promising strategy in systems biology and other fields like climate prediction[15–19]. Their central value has been the ability to quantify simulation uncertainty and to experimentally constrain model predictions. For example, Gutenkunst *et al.* showed that predictions were possible using signal transduction model ensembles, despite sometimes only order of magnitude parameter estimates[20]. More recently, Luan *et al.* predicted patient response to therapeutic intervention using an ensemble of human coagulation models where approximately 40% of the model parameters had a coefficient of variation (CV) greater than 0.5[21]. Model ensembles have also been used with parameter dependent analysis techniques, e.g., sensitivity analysis to robustly estimate important signaling network features. For example, Tasseff *et al.* characterized emergent behavior between androgen and growth factor signaling in prostate cancer cell-lines using an ensemble of models and sensitivity analysis[22]. Thus, despite uncertainty, ensembles of physiochemical models can be used to understand qualitative properties of complex biochemical networks.

We developed a population of physiochemical models describing the adaptation, alarm, and apoptosis phases of mammalian UPR. Traditionally, it has been hypothesized that there is a sequential order for firing of the ER stress transducers in UPR. The PERK and ATF6 branches are thought to be activated before IRE1[23] and largely promote ER adaptation to misfolding, while IRE1 is the last one to fire as it has a dual role, transmitting both survival and pro-apoptotic signals. However our modeling analysis suggests that these three branches fire simultaneously with varying rates and the state of the cell in terms of adaptation, alarm or apoptosis is a result of counteracting effects of these three prongs of UPR signaling. The counteracting effects is further substantiated by simulated knockout studies, wherein knockout of one ER stress transducer led to enhancement of the other branches of UPR, e.g, if we want more protein production even with ER stress, PERK KO might suggest a viable route that leads to no translational block. However the IRE1 branch will amplify as will the ATF6 branch ultimately leading to death (less compared to WT). So a viable strategy could be to regulate CHOP, resulting in reduced death by regulation of Bcl2. One common playground for the three ER stress transducers is the feedback regulation of BiP expression. Model analysis substantiated the significance of BiP feedback in the overall robustness of the system. Removal of any one branch of BiP feedback led to increased sensitivity of the other branches. So there was load/responsibility sharing within the system. Interestingly, removal of all nodes of BiP feedback increased the overall robustness of the system. So while BiP feedback is crucial in terms of allowing the cell to adapt to small perturbations, it also makes the system more fragile and susceptible to manipulations. Model analysis further highlighted the presence of redundant routes of regulation of apoptosis such as APAF-1 dependent/independent strategies. While manipulation of pro-death axis of the system (cleaved caspases) was relatively robust to perturbations, manipulation of the pro-survival axis (Bcl2) was relatively feasible via CHOP dependent/independent strategies. Overall UPR was seen to be robust to perturbations (simulated knockout/overexpression studies), thereby highlighting the redundancy and crosstalk within the three branches of UPR. However, BiP regulation at the transcriptional level via intermediates of the ER-stress transducer signaling cascades (e.g. ATF4, XBP1s) was seen to be the key player in regulation of UPR. The UPR model is available in SBML in the supplemental materials.

Results

Formulation of the UPR network architecture.

The UPR network described the ER folding cycle, ER-associated degradation (ERAD), ER-stress transducer (PERK, IRE1 and ATF6) signaling cascades and stress-induced caspase activation (Fig. **??**). The network consisted of 636 protein or mRNA species interconnected by 1090 interactions (Fig. **??** Inset). Connectivity was formulated from a comprehensive review of the primary literature[1–5,23–29], and from on-line databases; String-8[30], NetworKIN[31] and TRANSFAC. Model connectivity was not specific to a single cell-line; rather, it was a canonical representation of the pathways involved in monitoring and controlling the folding capacity of a generic well-mixed ER compartment. UPR induction was modeled as the release of BiP from the ER stress transducers, PERK, IRE1α, and ATF6 leading initially to adaptation of the folding cycle and then, subsequently, to alarm and apoptosis. The adaption phase of UPR was marked by general translation attenuation, selective transcriptional programs for key species like bZIP transcription factor ATF4[32], cellular inhibitor of apoptosis (cIAP)[33], molecular chaperones e.g., BiP[34], and enhanced clearance of accumulated proteins via ERAD. The alarm and apoptosis phases were mediated by the induction of CHOP[35], regulation of Bcl2, Bcl2-antagonist of cell death (BAD)[36] and (TNF) receptor associated factor 2 (TRAF2)[23,37–39] activation (Fig. **??**). Model connectivity is available in SBML format from the supplemental materials. Complete details of the interactions in the model are also enlisted in the supplementary materials.

The population of UPR models recapitulated adaptation, alarm, and apoptotic events across multiple cell-lines and timescales.

Despite a significant identifiability challenge, the multiobjective POETs algorithm generated a predictive UPR model population. The three phases of UPR were modeled using mass action kinetics within an ordinary differential equation (ODE) framework. While ODEs and mass-action kinetics are common methods of modeling biological pathways[40–42], this modeling strategy resulted in a large number of unknown model parameters. These parameters were not uniquely identifiable (data not shown). Instead, we estimated an experimentally constrained population of parameters using multiobjective optimization. A population of the 1726 unknown model parameters (1090 kinetic parameters and 636 initial conditions) was estimated from 33 dynamic and steady state data-sets taken from literature (Table **??**). The residual between model simulations and each of the experimental constraints was simultaneously minimized using the multiobjective POETs algorithm[43]. A leave-eight-out cross-validation strategy was used to independently estimate the training and prediction error over the 33 data sets; we estimated four different model families, where eight of the 33 objectives were reserved for validation and 25 were used for model training of each family. Further details on model and training in supplementary materials. POETs generated model families that predicted approximately 94% of the objective functions with a significantly higher likelihood than a random control. However, the specific value of any given parameter was likely not well described (data not shown). The coefficient of variation (CV) for the model parameters ranged from 0.5 - 1.6, where approximately 65% of the parameters were constrained with a CV ≤ 1.0 (supplementary materials Fig. S**??**). The most constrained parameters involved a wide-array of functions e.g., regulation of PERK, eIF2α, ATF4, Calcineurin, BiP, CHOP and ATF6 signaling. However, the least constrained parameters involved JNK and apoptosis interactions. POETs identified Pareto fronts between several objectives, e.g., O13$\times$O14, O25$\times$O29, O11$\times$O29, and O27$\times$O2, in the training data (Fig. **??**). Strong Pareto fronts suggested an inability to simultaneously model different aspects of the training data. However, fronts could also result from experimental artifacts, e.g., variation between cell-lines, time-scale differences, or from functional relationships in the data. Globally, adaptation and alarm phase training constraints conflicted with those involving apoptosis. For example, objectives involving caspase-7 or caspase-9 activity conflicted with phosphorylated eIF2α levels. Phosphorylation of eIF2α by activated PERK attenuates translation, which decreases the ER folding load. Thus, eIF2α phosphorylation is a key early adaptive event in UPR. On the other hand, caspase-9 is a stress-induced death marker activated only after UPR has failed to restore folding homeostasis. Conflicts between these early and late phase markers suggested the UPR time-scale was perhaps cell-line or perturbation dependent. Similarly, negative feedback may lead to conflicting objectives as well (Fig. **??** and supplementary information).

We trained the UPR ensemble using data from UPR initiation events (e.g., PERK activation) and the downstream activation of proteins involved in apoptosis (BiP, caspase-3 and caspase-7). ER stress induced by exposure to thapsigargin (Tg), a non-competitive inhibitor of SERCA Ca^{2+} transporters, leads to the dissociation of BiP from the ER-stress transducers. For these initial proof-of-concept simulations, we assumed that the action of Tg and other stress-inducing agents such as dithiothreitol (DTT) was identical, i.e., induction of BiP dissociation. Within 20 minutes after Tg exposure, PERK was activated and transmitted adaptation signals downstream. The population of UPR models recapitulated the timescale of PERK phosphorylation (Fig. **??**A) as well as its downstream signaling activity, for example, the phosphorylation of eIF2α (Fig. **??**H). The nuclear fraction of ATF4 increased from approximately zero (untreated cells) to a maximum value 4 hrs after Tg exposure. While the model ensemble generally predicted the correct trend, there was significant error in the early time points for ATF4 (Fig. **??**G). The phosphorylation of eIF2α by PERK is required for ATF4 activation. Interestingly, when we compared model simulations of p-eIF2α levels following Tg (1μM) exposure in mouse embryonic fibroblasts (MEFs) with measurements (O15[44]), the model correctly captured the appropriate behavior. To test the functionality of the ATF6 branch of the UPR model, we compared simulations with measurements of cleaved ATF6 in tunicamycin-treated MEFs[45]. ER stress is known to lead to the release of BiP from ATF6. Cleaved ATF6 is then translocated to the nucleus where it up-regulates gene expression[46,47]. Simulations of cleaved ATF6 levels following UPR initiation were consistent with measurements (Fig. **??**C). The formation of p-PERK or p-IRE1 initiates a complex series of events that operate on both short and long time-scales. Signals from the ER stress transducers converge downstream to regulate BiP transcription[48–51]. Model ensemble recapitulated correct trends of BiP mRNA with maximum levels $\sim$ 8 hrs, similar to what was seen in experiments done on HEK293 cells (Fig. **??**D)[52]. One of the long-term outcomes of PERK/IRE1 activation is apoptotic cell death. In the proof-of-concept ER-model the link between UPR and apoptosis occurred through the action of eIF2α, the dual role of the ATF4 transcription factor and caspase-12 activation by IRE1-TRAF2 signaling axis. We constrained model parameters associated with the activation of the cell-death program using measurements of pro/caspase-7 levels, pro/caspase-9 levels, pro/caspase-3 levels, pro/caspase-12 levels and PARP cleavage mediated by executioner caspases following treatment with 0.5μM Tg[53]. These experiments were performed in Sak2 cells that lack Apaf-1 protein expression[53]. Thus, the data allowed us to include a non-Apaf-1 mediated stress-induced caspase activation pathway into the model. The population of models recapitulated caspase-3 (Fig. **??**K) and caspase-9 (Fig. **??**J), as well as cleaved Parp levels (Fig. **??**L) following exposure to ER stress-inducers. Interestingly, while PERK activation occurred on the timescale of minutes, initiator and executioner caspase activation occurred over 36 hrs. Thus, the population of UPR models

captured complex signaling events occurring across multiple timescales.

Signal flow analysis of UPR highlighted modes of crosstalk and redundancy in BiP and Bcl2 regulation.

Traditionally, it has been hypothesized that there is a sequential order for firing of the ER stress transducers in UPR. The PERK and ATF6 branches are thought to be activated before IRE1[23] and largely promote ER adaptation to misfolding, while IRE1 is the last one to fire as it has a dual role, transmitting both survival and pro-apoptotic signals. However, our modeling analysis suggests that these three branches fire simultaneously with varying rates and that the state of the cell in terms of adaptation, alarm, or apoptosis is a result of counteracting effects of these three prongs of UPR signaling (Fig. **??** and supplementary materials Fig. S**??**). UPR induction in the model was controlled by manipulation of the generation rate of unfolded or misfolded protein (qP) in the ER compartment. Upon UPR induction, initially ($\leq$ 1hr) the response is damped marking the adaptation phase of UPR. Adaptation is followed by an increase in the activity of the IRE1α, PERK, and ATF6 cascades at $\sim$ 1 hr marking the onset of the alarm phase, which leads to a final steady state $\sim$ 8-10 hrs marking the onset of the commitment or apoptosis phase of UPR (Fig. **??**). This time frame was consistent with the maximum levels of BiP mRNA upon UPR induction[52]. Our analysis was substantiated further by looking at the fluxes at different phases of UPR induction. As compared to P1 (No-UPR Steady State Fig. **??**), we saw that early on at 1 hr after unfolded protein dose there was a marked increase in ATF4 and CHOP regulation, ATF6 signaling along with unfolded protein sensing and degradation. These are hallmarks of the adaptation-alarm phase of the UPR response (Fig. **??**D). The apoptosis phase ($\geq$ 8-10 hrs P2) was marked by increased BiP regulation, enhanced ATF4 transcriptional activity, increased mitochondrial membrane permeability, and increased apoptotic fluxes resulting in cell commitment to apoptosis mediated cell death. Certain modules were seen to reduce like IRE1-TRAF2 signaling, ASK1 activation. However, not much difference was seen in terms of apoptotic fluxes, denoting the commitment of the cell to death. On the contrary, if we reduced the load of proteins in the adaptation-alarm phase (P4), we saw that the cell could recuperate using its ERAD machinery and the regulation of BiP (Fig. **??**E-F).

We further investigated the effect of simulated knockout (KO) and overexpression (OX) of key proteins on the UPR system (supplementary materials Fig. S**??**). Overall signal flow analysis highlighted the extensive amount of crosstalk within the three branches of the UPR network. Taken together, the population of UPR models recapitulated both short- and long timescale behavior following overload of unfolded proteins in a variety of cell types. It also captured the integration between multiple pathways and generated specific and testable hypothesis about the role of network components in signal propagation.

Sensitivity analysis stratified locally and globally important components of the UPR architecture.

First-order sensitivity coefficients were computed, time-averaged (over approximately 8 hrs of simulated time following ER stress), and rank-ordered for the 1090 model parameters and 636 species under normal and UPR induced conditions. Five parameter sets were selected from the parent ensemble based on rightful representation of the diversity for the sensitivity calculations (further details in the supplementary materials). We selected only a few diverse parameter sets for the sensitivity analysis, because of the computational cost of computing all 1090 coefficients over several hours of simulated time.

Sensitivity analysis conducted over discrete two hour time windows revealed the time evolution of the importance of UPR network modules (Fig. **??**). Comparison of the 0 - 2 hrs time window with itself (top panel, first column of Fig. **??**), supported the earlier results that infrastructure components were globally critical followed by ERAD species. These species remained important in all time windows. On the other hand, during the initial 0 - 2 hrs window, ER stress transduction pathway components were robust. Comparison of the 0 - 2 hrs time window with later time points (working down the first column, Fig. **??**), showed the increasing importance of different modules as a function of time. For example, components of the PERK and IRE1 modules were more important in the 2 - 4 hrs window compared to the earlier time points, while alarm and apoptotic phase species were more important in the 6 - 8 hrs window compared to the earlier time points. Specifically, signal integration via the transcriptional activity of ATF6, ATF4, and XBP1s along with the role of RCAN1 and cIAP in apoptosis were significantly more important at 6-8 hrs as compared to 0-2 hrs time window. This is consistent with the dominant role of the negative feedback via the transcriptional regulation of BiP in UPR. Interestingly, the majority of species rankings were similar after 6 hrs (bottom row, Fig. **??** and supplementary materials Fig. S**??**). This analysis supports the structural analysis results, wherein the signal integrators like XBP1s and transcriptional regulation of BiP via ATF6 and ATF4 are the key regulators of the phased response of UPR.

To further investigate the role of positive BiP regulation via ATF6, ATF4 and XBP1s, we conducted sensitivity analysis upon knocking out the feedback branches of BiP for the nominal parameter set (supplementary materials Fig. S**??**). This highlights the essential role and relevance of targeting the BiP feedback in manipulating UPR and specific importance of ATF4 amongst the three feedback branches.

Robustness analysis predicted the phenotypic consequence of structural perturbations to the UPR network.

We calculated the direction of unfolded protein load induced concentration shifts for 636 markers following single parameter knockouts (Edge KO), single gene knockouts (GKO), and single gene overexpression (GOX) to the UPR network for the nominal parameter set (Fig. **??**). Robustness coefficients were used to quantify the effect of structural perturbations on network markers. Coefficients with values > 1 (< 1) indicated a marker increased (decreased) compared to the basal state, while a value ~ 1 indicated approximately no change following a perturbation. Phenotypic behavior of the UPR models were analyzed as a result of perturbation using downstream markers like Caspase 3 (marker for cell death), Bcl2 (marker for cell-survival) and other key signal integrators of the signals from ER stress transducers e.g., XBP1s, CHOP and ATF4.

Overall, we found that in the survival-death phenotypic plane, the pro-death phenotype (marked by robustness coefficients $>$ 1 for Caspase 3) was seen to be relatively robust to structural (both GKO, GOX and Edge KO) perturbations. Few perturbations lead to increased Caspase 3 levels, e.g., overexpression of Procaspases 9/3 (Fig **??** A, Table **??**). This robustness of the apoptotic marker caspase 3, can be attributed to the redundant sources of cell death (e.g., APAF-1 dependent and APAF-1 independent strategies). Interestingly, manipulation of the pro-survival axis via regulation of Bcl2 was possible (Fig **??**A, Table **??**). For example, KO of PERK/ATF4 signaling components led to increased Bcl2 marker levels (Fig **??**A). This behavior is attributed to the dominant role of PERK-ATF4 mediated regulation of CHOP which downstream leads to down-regulation of Bcl2 levels (seen earlier in signal analysis results, supplementary information).

Robustness analysis allowed us to investigate complex network properties like redundancy and crosstalk. For example, the direct correlation between ATF4 and CHOP was further noted in the ATF4-CHOP phenotypic plane. Any perturbations affecting ATF4 affected CHOP levels in the same manner (reduction in ATF4 levels lead to reduced levels of CHOP) (Fig. **??**A-B). However, owing to redundant sources of CHOP regulation (e.g., via XBP1s), effect on CHOP was damped in relation to significant changes in ATF4 levels. In the XBP1s-CHOP plane, we see at lower levels of XBP1s and CHOP, there is a direct relation between XBP1s levels and CHOP levels. However, there exists very few strategies of having both high XBP1s levels and CHOP levels indicating that higher XBP1s doesn't necessarily mean higher CHOP levels. To further investigate the implications of the feedback regulation of BiP via ATF4/ATF6/XBP1s, we simulated KOs of these components over the entire ensemble (data not shown). Upon knockout of BiP feedback, BiP regulation was found to be very strong resulting in drastic reductions in BiP levels and ultimately a stronger and faster UPR response (supplementary materials Fig. S**??**).

Global analysis of the gene knockout robustness coefficients by clustering to minimize the variance, provided systems-level insight into the UPR network. As in Fig. **??**C, dendrogram of the single GKOs revealed clusters with inherent functional relationships. For example, the most distinct separation was between unfolded protein sensing and IRE1/PERK signal initiation from the rest of the knockouts. PERK/ATF4 branch as seen earlier, plays a dominant role in the regulation of BiP and CHOP upon the onset of UPR. Similarly, IRE1/TRAF2 signaling axis is valuable to the apoptosis module. Another interesting functional module was that of CHOP, involving p38MAPK which leads to down-regulation of Bcl2 levels which considerably affects the apoptosis module. We computed the magnitude of the orthogonal components of the single GKO coefficients (Fig. **??**C Inset). The orthogonal component was used to establish the uniqueness of the knockout. All the single GKOs were found to have orthogonal components greater than one (> 1) with a 95% confidence. KO of BiP, Procaspase 9, IRE1, PERK and TRAF2 were some of the knockouts which produced the most unique effects. This is in accord to the critical role of BiP in initiating and further regulating the time scale of progression and the ultimate result of the UPR response. So PERK via ATF4 plays a key role in regulation of BiP, is rightly seen to have a major effect as compared to other GKOs. Similarly, regulation of the apoptosis branch via IRE1-TRAF2 and Procaspase 9 rightly were seen to have the most effect as compared to other GKOs.

Discussion

Proteins requiring post-translational modifications such as N-linked glycosylation or disulfide bond formation are processed in the endoplasmic reticulum (ER). A diverse array of cellular stresses can lead to dysfunction of the ER, and ultimately to an imbalance between protein-folding capacity and protein-folding load. Unfolded or misfolded proteins are tagged for degradation via ER associated degradation (ERAD) or sent back through the folding cycle. Continued accumulation of incorrectly folded proteins can also trigger the Unfolded Protein Response (UPR). In this study, we formulated a mechanistic model of the cellular response to stress (protein overload) and studied the core regulatory aspects and downstream effects of UPR induction. PRKR-like ER kinase (PERK), inositol-requiring kinase 1 (IRE1) and activating transcription factor 6 (ATF6) were modeled as the key UPR initiators. While UPR has been extensively studied[1–5,23–29], a detailed mathematical model to investigate the complexities involved is lacking. The UPR network architecture is based on extensive review of the literature (details of connectivity in supplementary materials)[1–5,23–29]. Mass balance equations describing 636 species interconnected by 1090 interactions were formulated using mass-action kinetics within an ordinary differential equation (ODE) framework. Four model populations were estimated using multi-objective optimization (33 objective functions) in conjunction with a leave-eight out cross-validation strategy using POETs[43]. These model populations were then analyzed using population-based sensitivity and

robustness analysis. Overall we identified sources of network crosstalk and redundancy within the UPR module which could be the reason for emergent properties and be the links for aberrations in cellular adaptation to stress.

A key finding of our study was that the overall outcome of UPR was as a result of simultaneous firing and competition between signaling mediated by the three ER-stress transducers: PERK, IRE1, and ATF6. This is in contrast to the traditional belief that PERK and ATF6 branches are activated before IRE1[23]. So what we hypothesize is that instead of a sequential ordering of these branches, the state of the cell in terms of adaptation, alarm, or apoptosis is a result of counteracting effects of these three prongs of UPR signaling. The counteracting/competing effects were further substantiated in simulated knockout studies, wherein knockout of one ER stress transducer led to enhancement of the other branches of UPR. Signal transduction architectures frequently contain redundancy, feedback, and crosstalk. These topological features ensure signal propagation is adaptable, efficient, and robust. However, they also make reprogramming signal flow challenging. This was highlighted remarkably in the case of UPR. Signals from the three ER-stress transducers converged at the level of up-regulation of BiP. This is the key junction which regulates the three stages (onset and time) of adaptation, alarm, and apoptosis. In this regard, regulators of this feedback cleaved ATF6, ATF4, and XBP1s and were seen as highly sensitive components of UPR. Interestingly these components put-together increased the overall fragility of the system and presented a greater scope of manipulation of the UPR response. This was substantiated by sensitivity analysis upon KO of the feedback loops, where we saw increased stability of the UPR module. When these feedback components were knocked out individually, the system overall remained stable thanks to increased activity and load sharing via the other feedback branches. Amongst the three components of feedback, we identified ATF4 as the key load bearer/regulator. This was substantiated by signal flow, robustness and sensitivity analysis. This is really interesting as ATF4 protein has shown to be present in greater levels in cancer compared to normal tissue, and it is up-regulated by signals of the tumor microenvironment such as hypoxia/anoxia, oxidative stress, and ER stress[55]. So any aberrations in regulation of ATF4 could potentially serve as a specific target in cancer therapy. As a target, ATF4 is attractive because it is also potentially involved in angiogenesis and adaptation of cancer cells to hypoxia/anoxia, which are major problems in cancer progression[55].

Downstream effects of UPR range from cellular adaptation/survival (low stress) to the cell committing to apoptosis mediated death (high stress). Our modeling analysis suggested that the cell-death phenotype (marked by increased levels of Caspase 3 as compared to WT) was relatively robust. This robustness could be attributed to redundant routes of APAF-1 dependent and APAF-1 independent routes of apoptosis. This claim is supported by experimental evidence in Sak2 cells[53] and as seen by our simulated knockout studies where we identified two distinct populations representing clear distinctions in APAF-1 dependent and independent routes of apoptosis. Interestingly, manipulation of the pro-survival phenotype (marked by increased levels of Bcl2 as compared to WT) was feasible. The most effective route was via manipulation of the PERK/ATF4/CHOP branch. This was substantiated by simulated CHOP KO experiments over the entire ensemble, wherein we identified two distinct populations within the ensembles. One with a strong effect of CHOP mediated down-regulation of Bcl2 (marked by $\sim$ 10 fold increase in Bcl2 levels) and the other with very little effect of CHOP on Bcl2 levels. This complex network behavior could be attributed to other conflicting means of regulation of Bcl2 levels. Rightfully so, induction of CHOP is involved in the development of various diseases and several therapeutic interventions[56]. For instance, suppression of CHOP by RNA interference, decoy oligodeoxynucleotides or drug inhibitors have a significant therapeutic potential to modulate type I diabetes and brain ischemia. On the other hand, overexpression of CHOP may represent a new class of anticancer therapy. Since induction of BiP has been observed in a variety of tumor cells, overexpression of CHOP directed by the BiP promoter may be used as a highly specific therapy for cancer[56]. Model analysis also highlighted the essential role of RCAN1 and IRE1-TRAF2 routes of apoptosis. ATF6 induces regulation of calcineurin 1 (RCAN1) expression[27]. RCAN1 sequesters calcineurin[27], a calcium activated protein-phosphotase B, that dephosphorylates Bcl2-antagonist of cell death (BAD) at S75 or S99[36]. This leads to sequestering of Bcl2 by BAD, which inhibits its downstream anti-apoptotic activity[36]. Recently, a number of ATF6 homologs have been identified, e.g., OASIS, CREBH, LUMAN/CREB3, CREB4, and BBF2H7 that are processed in a similar way as ATF6, yet their function remains unknown[57]. Thus, ER-stress induced ATF6 signaling may be responsible for additional undiscovered functionality.

Given the structural and parametric uncertainty associated with the current version of the UPR model, we were still able to extract interesting insight into the complex inner workings and present falsifiable hypothesis regarding manipulating the UPR program. While we did an extensive search of the literature to formulate the model, we were missing certain key structural aspects of UPR which could provide a more comprehensive analysis for further generations of the model. One key missing aspect is the negative regulation of the three ER-stress transducers. Given PERK's central role in translation attenuation, cells have evolved multiple axes to regulate PERK activity[58]. Additionally, little is known about ATF6. XBP!u, the unspliced form of XBP1, has been implicated as a negative regulator for ATF6[64]. IRE1α activity is regulated by several proteins, specifically members of the Bcl2 protein family Bax and Bak[70]. More detailed netowrk considerations for future model generations that may lend more relevant insight into manipulating UPR is found in supplementary material.

Methods

Formulation and solution of the model equations.

The unfolded protein response model was formulated as a set of coupled ordinary differential equations (ODEs):

$$\frac{d\mathbf{x}}{dt} = \mathbf{S} \cdot \mathbf{r}(\mathbf{x}, \mathbf{p}) \qquad \mathbf{x}(t_o) = \mathbf{x}_o \tag{1}$$

The symbol $\mathbf{S}$ denotes the stoichiometric matrix (636×1090). The quantity $\mathbf{x}$ denotes the concentration vector of proteins or protein complexes (636×1). The term $\mathbf{r}(\mathbf{x}, \mathbf{p})$ denotes the vector of reaction rates (1090×1). Each row in $\mathbf{S}$ described a protein or protein-protein complex, while each column described the stoichiometry of network interactions. Thus, the (i, j) element of $\mathbf{S}$, denoted by σ_{ij}, described how protein i was involved in rate j. If $\sigma_{ij} < 0$, then protein i was consumed in r_j. Conversely, if $\sigma_{ij} > 0$, protein i was produced by r_j. Lastly, if $\sigma_{ij} = 0$, there was no protein i in rate j. All of these interactions were obtained from the literature.

We assumed mass-action kinetics for each interaction in the network. The rate expression for interaction q was given by:

$$r_q(\mathbf{x}, k_q) = k_q \prod_{j \in \{\mathbf{R}_q\}} x_j^{-\sigma_{jq}} \tag{2}$$

The set $\{\mathbf{R}_q\}$ denotes reactants for reaction q while σ_{jq} denotes the stoichiometric coefficient (element of the matrix $\mathbf{S}$) governing species j in reaction q. All reversible interactions were split into two irreversible steps. The mass-action formulation, while expanding the dimension of the UPR model, regularized the mathematical structure; this allowed automatic generation of the model code using UNIVERSAL and regularized the unknown model parameters (parameters were one of only three types, association, dissociation or catalytic rate constants). UNIVERSAL, an open source Java code generator, generates multiple code types from text and SBML inputs. UNIVERSAL is freely available as a Google Code project (http://code.google.com/p/universal-code-generator/). Thus, although mass-action kinetics increased the number of parameters and species, they reduced the complexity of model analysis. In this study, we considered well-mixed nuclear, cytosolic and extracellular compartments. The model equations were solved using the LSODE routine in OCTAVE (v 3.1.0; www.octave.org) on an Apple workstation (Apple, Cupertino, CA; OS X v10.6.4).

Unfolded protein response conditions were simulated by running the model to steady state and then providing a dose of proteins. The steady-state was estimated numerically by repeatedly solving the model equations and estimating the difference between subsequent time points:

$$\|\mathbf{x}(t + \Delta t) - \mathbf{x}(t)\|_2 \leq \gamma \tag{3}$$

The quantities $\mathbf{x}(t)$ and $\mathbf{x}(t + \Delta t)$ denote the simulated concentration vector at time t and $t + \Delta t$, respectively. The L_2 vector-norm was used as the distance metric. We used $\Delta t = 1$ s and $\gamma = 0.001$ for all simulations.

Estimation and cross-validation of a population of models using Pareto Optimal Ensemble Techniques (POETs).

POETs is a multiobjective optimization strategy which integrates several local search strategies e.g., Simulated Annealing (SA) or Pattern Search (PS) with a Pareto-rank-based fitness assignment[43]. Denote a candidate parameter set at iteration $i+1$ as $\mathbf{k}_{i+1}$. The squared error for $\mathbf{k}_{i+1}$ for training set j was defined as:

$$E_j(\mathbf{k}) = \sum_{i=1}^{\mathscr{T}_j} \left(\hat{\mathscr{M}}_{ij} - \hat{y}_{ij}(\mathbf{k}) \right)^2 \tag{4}$$

The symbol $\hat{\mathscr{M}}_{ij}$ denotes scaled experimental observations (from training set j) while the symbol $\hat{y}_{ij}$ denotes the scaled simulation output (from training set j). The quantity i denotes the sampled time-index and $\mathscr{T}_j$ denotes the number of time points for experiment j. The read-out from the training immunoblots was band intensity where we assumed intensity was only loosely proportional to concentration. Suppose we have the intensity for species x at time $i = \{t_1, t_2, .., t_n\}$ in condition j. The scaled measurement would then be given by:

$$\hat{\mathscr{M}}_{ij} = \frac{\mathscr{M}_{ij} - \min_i \mathscr{M}_{ij}}{\max_i \mathscr{M}_{ij} - \min_i \mathscr{M}_{ij}} \tag{5}$$

Under this scaling, the lowest intensity band equaled zero while the highest intensity band equaled one. A similar scaling was defined for the simulation output.

We computed the Pareto rank of $\mathbf{k}_{i+1}$ by comparing the simulation error at iteration $i+1$ against the simulation archive $\mathbf{K}_i$. We used the Fonseca and Fleming ranking scheme[76]:

$$rank\left(\mathbf{k}_{i+1} \mid \mathbf{K}_i\right) = p \tag{6}$$

where p denotes the number of parameter sets that dominate parameter set $\mathbf{k}_{i+1}$. Parameter sets on or near the optimal trade-off surface have small rank. Sets with increasing rank are progressively further away from the optimal trade-off surface. The parameter set $\mathbf{k}_{i+1}$ was accepted or rejected by the SA with probability $\mathscr{P}\left(\mathbf{k}_{i+1}\right)$:

$$\mathscr{P}\left(\mathbf{k}_{i+1}\right) \equiv \exp\left\{-rank\left(\mathbf{k}_{i+1} \mid \mathbf{K}_i\right)/T\right\} \tag{7}$$

where T is the computational annealing temperature. The initial temperature $T_o = n/log(2)$, where n is user defined ($n = 4$ for this study). The final temperature was $T_f = 0.1$. The annealing temperature was discretized into 10 quanta between T_o and T_f and adjusted according to the schedule $T_k = \beta^k T_0$ where β was defined as:

$$\beta = \left(\frac{T_f}{T_o}\right)^{1/10} \tag{8}$$

The epoch-counter k was incremented after the addition of 100 members to the ensemble. Thus, as the ensemble grew, the likelihood of accepting parameter sets with a large Pareto rank decreased. To generate parameter diversity, we randomly perturbed each parameter by $\leq \pm 25\%$. We performed a local pattern-search every q steps to minimize the residual for a single randomly selected objective. The local pattern-search algorithm has been described previously[77,78]. The parameter ensemble used in the simulation and sensitivity studies was generated from the low-rank parameter sets in $\mathbf{K}_i$.

We simultaneously calculated training and prediction error during the parameter estimation procedure using leave-eight-out cross-validation[79]. The complete set of training data (33 objectives) was subdivided into four bins; in each bin 25 data sets were reserved for training while eight were reserved for prediction. In the first bin $DS_1 \ldots DS_8$ were used for validation while $DS_9 \ldots DS_{33}$ were used for training. In the second bin $DS_9 \ldots DS_{16}$ were used for validation while $DS_1 \ldots DS_8$ $DS_{17} \ldots DS_{33}$ were used for training, etc. Thus, we formulated four ensembles from which we evenly selected parameter sets for the *parent* ensemble (Fig. **??**). While cross-validation required that we generate additional model populations, we trained and tested against all the data sets.

Sensitivity and robustness analysis of the population of EHR models.

Sensitivity coefficients were calculated as shown previously[43] using five models selected from the ensemble (red points, supplementary materials Fig. S**??**). The resulting sensitivity coefficients were scaled and time-averaged (Trapezoid rule):

$$\mathscr{N}_{ij} \equiv \frac{1}{T}\int_0^T dt \cdot \left|\alpha_{ij}\left(t\right) s_{ij}(t)\right| \tag{9}$$

where T denotes the final simulation time and $\alpha_{ij} = 1$. The time-averaged sensitivity coefficients were then organized into an array for each ensemble member:

$$\mathscr{N}^{(\varepsilon)} = \begin{pmatrix} \mathscr{N}_{11}^{(\varepsilon)} & \mathscr{N}_{12}^{(\varepsilon)} & \cdots & \mathscr{N}_{1j}^{(\varepsilon)} & \cdots & \mathscr{N}_{1P}^{(\varepsilon)} \\ \mathscr{N}_{21}^{(\varepsilon)} & \mathscr{N}_{22}^{(\varepsilon)} & \cdots & \mathscr{N}_{2j}^{(\varepsilon)} & \cdots & \mathscr{N}_{2P}^{(\varepsilon)} \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathscr{N}_{M1}^{(\varepsilon)} & \mathscr{N}_{M2}^{(\varepsilon)} & \cdots & \mathscr{N}_{Mj}^{(\varepsilon)} & \cdots & \mathscr{N}_{MP}^{(\varepsilon)} \end{pmatrix} \qquad \varepsilon = 1,2,\ldots,N_\varepsilon \tag{10}$$

where ε denotes the index of the ensemble member, P denotes the number of parameters, N_ε denotes the number of ensemble samples and M denotes the number of model species. To estimate the relative fragility or robustness of species and reactions in the network, we decomposed the $\mathscr{N}^{(\varepsilon)}$ matrix using Singular Value Decomposition (SVD):

$$\mathscr{N}^{(\varepsilon)} = \mathbf{U}^{(\varepsilon)}\mathbf{\Sigma}^{(\varepsilon)}\mathbf{V}^{T,(\varepsilon)} \tag{11}$$

Coefficients of the left (right) singular vectors corresponding to largest β singular values of $\mathscr{N}^{(\varepsilon)}$ were rank-ordered to estimate important species (reaction) combinations. Only coefficients with magnitude greater than a threshold ($\delta = 0.1$) were considered. The fraction of the β vectors in which a reaction or species index occurred was used to rank its importance.

Robustness coefficients of the form:

$$\alpha\left(i,j,t_o,t_f\right) = \left(\int_{t_o}^{t_f} x_i\left(t\right)dt\right)^{-1}\left(\int_{t_o}^{t_f} x_i^{(j)}\left(t\right)dt\right) \tag{12}$$

were calculated to understand the robustness of the network. The robustness coefficient $\alpha\left(i,j,t_o,t_f\right)$ is the ratio of the integrated concentration of a network marker in the presence (numerator) and absence (denominator) of structural or operational perturbation. The quantities t_0 and t_f denote the initial and final simulation time respectively, while i and j denote the indices for the marker and the perturbation respectively. If $\alpha\left(i,j,t_o,t_f\right) > 1$, then the perturbation *increased* the marker concentration. Conversely, if $\alpha\left(i,j,t_o,t_f\right) \ll 1$ the perturbation *decreased* the marker concentration. Lastly, if $\alpha\left(i,j,t_o,t_f\right) \sim 1$ the perturbation did not influence the marker concentration.

Species clustering and dendrogram.

A dendrogram was derived by considering each of the knockouts(over-expressions) as variables and the average log of robustness coefficient (LRC) for each of the species as observations. We used the Euclidean norm in LRC space as the distance metric. The linkage function (objective function for identifying variable clusters) was the inner squared distance (minimum variance algorithm). The Statistical Toolbox of Matlab (The Mathworks, Natick, MA) was used to generate the distances, linkages and the final dendrogram.

Identification of distinguishable species.

Robustness coefficients were used to rank-order knockout(overexpression) experiments in terms of the greatest unique responses and identify species which were linearly distinguishable. The response of the knockout(overexpression) was measured in terms of the robustness coefficients. The LRC had desirable linear properties, such that no response (no change in trajectories from wild-type) returns a value of zero and similar negative and positive responses have different directions but similar magnitudes. We considered the unique component of the response to be the orthogonal component in LRC space and the magnitude of the response to be the Euclidean norm. The orthogonal components and there magnitude were identified for each parameter set in the ensemble by first choosing the knockout(overexpression) with the greatest magnitude, x_1 and placing it in the empty set $\mathscr{V}$. The knockout(overexpression) x_1 defines the orthogonal directions in the LRC space. We then calculated the orthogonal components for all remaining knockouts(overexpressions) relative to x_1, and added the knockout(overexpression) species with the greatest orthogonal magnitude to set $\mathscr{V}$. In general the components of all remaining x_i orthogonal to set $\mathscr{V}$ were calculated and the largest was moved into set $\mathscr{V}$. This process was continued until all knockout(overexpression) species, x_i were added to set $\mathscr{V}$. Mathematically two species were considered distinguishable if and only if they were linearly independent (the orthogonal components were non-zero). We considered a threshold value of one or five and performed a student t-test (Matlab Statistical Toolbox, The Mathworks, Natick, MA) to identify which species had orthogonal components above the threshold with a 95% confidence over the ensemble.

References

1. Naidoo, N. Er and aging-protein folding and the er stress response. *Ageing Res Rev* **8**, 150–9 (2009).
2. Ron, D. Translational control in the endoplasmic reticulum stress response. *J Clin Invest* **110**, 1383–8 (2002).
3. Kaufman, R. J. *et al.* The unfolded protein response in nutrient sensing and differentiation. *Nat Rev Mol Cell Biol* **3**, 411–21 (2002).
4. Ellgaard, L. & Helenius, A. Quality control in the endoplasmic reticulum. *Nat Rev Mol Cell Biol* **4**, 181–91 (2003).
5. Fonseca, S. G., Burcin, M., Gromada, J. & Urano, F. Endoplasmic reticulum stress in beta-cells and development of diabetes. *Curr Opin Pharmacol* **9**, 763–70 (2009).
6. Ogata, M. *et al.* Autophagy is activated for cell survival after endoplasmic reticulum stress. *Mol Cell Biol* **26**, 9220–31 (2006).
7. Yorimitsu, T., Nair, U., Yang, Z. & Klionsky, D. J. Endoplasmic reticulum stress triggers autophagy. *J Biol Chem* **281**, 30299–304 (2006).
8. Bernales, S., McDonald, K. L. & Walter, P. Autophagy counterbalances endoplasmic reticulum expansion during the unfolded protein response. *PLoS Biol* **4**, e423 (2006).
9. Kamimoto, T. *et al.* Intracellular inclusions containing mutant alpha1-antitrypsin z are propagated in the absence of autophagic activity. *J Biol Chem* **281**, 4467–76 (2006).
10. Høyer-Hansen, M. *et al.* Control of macroautophagy by calcium, calmodulin-dependent kinase kinase-beta, and bcl-2. *Mol Cell* **25**, 193–205 (2007).
11. Kouroku, Y. *et al.* Er stress (perk/eif2alpha phosphorylation) mediates the polyglutamine-induced lc3 conversion, an essential step for autophagy formation. *Cell Death Differ* **14**, 230–9 (2007).

12. Fujita, E. *et al.* Two endoplasmic reticulum-associated degradation (erad) systems for the novel variant of the mutant dysferlin: ubiquitin/proteasome erad(i) and autophagy/lysosome erad(ii). *Hum Mol Genet* **16**, 618–29 (2007).

13. Kitano, H. A robustness based approach to systems-oriented drug design. *Nat. Rev. Drug Discov.* **6**, 202 – 210 (2007).

14. Gadkar, K. G., Varner, J. & Doyle, F. J. Model identification of signal transduction networks from data using a state regulator problem. *Syst Biol (Stevenage)* **2**, 17–30 (2005).

15. Battogtokh, D., Asch, D. K., Case, M. E., Arnold, J. & Schuttler, H.-B. An ensemble method for identifying regulatory circuits with special reference to the qa gene cluster of neurospora crassa. *Proc Natl Acad Sci U S A* **99**, 16904–16909 (2002).

16. Kuepfer, L., Peter, M., Sauer, U. & Stelling, J. Ensemble modeling for analysis of cell signaling dynamics. *Nat Biotech* **25**, 1001–1006 (2007).

17. Brown, K. S. & Sethna, J. P. Statistical mechanical approaches to models with many poorly known parameters. *Phys Rev E Stat Nonlin Soft Matter Phys* **68**, 021904 (2003).

18. Palmer, T. *et al.* Representing model uncertainty in weather and climate prediction. *Ann Rev Earth and Planetary Sci* **33**, 163–193 (2005).

19. Song, S. O. & Varner, J. Modeling and analysis of the molecular basis of pain in sensory neurons. *PLoS One* **4**, e6758 (2009).

20. Gutenkunst, R. N. *et al.* Universally sloppy parameter sensitivities in systems biology models. *PLoS Comput Biol* **3**, 1871–78 (2007).

21. Luan, D., Szlam, F., Tanaka, K. A., Barie, P. S. & Varner, J. D. Ensembles of uncertain mathematical models can identify network response to therapeutic interventions. *Mol Biosyst* **6**, 2272–86 (2010).

22. Tasseff, R. *et al.* Analysis of the molecular networks in androgen dependent and independent prostate cancer revealed fragile and robust subsystems. *PLoS One* **5**, e8864 (2010).

23. Szegezdi, E., Logue, S. E., Gorman, A. M. & Samali, A. Mediators of endoplasmic reticulum stress-induced apoptosis. *EMBO Rep* **7**, 880–5 (2006).

24. Schröder, M. & Kaufman, R. J. The mammalian unfolded protein response. *Annu Rev Biochem* **74**, 739–89 (2005).

25. Gotoh, T., Terada, K. & Mori, M. hsp70-dnaj chaperone pairs prevent nitric oxide-mediated apoptosis in raw 264.7 macrophages. *Cell Death Differ* **8**, 357–66 (2001).

26. McCullough, K. D., Martindale, J. L., Klotz, L. O., Aw, T. Y. & Holbrook, N. J. Gadd153 sensitizes cells to endoplasmic reticulum stress by down-regulating bcl2 and perturbing the cellular redox state. *Mol Cell Biol* **21**, 1249–59 (2001).

27. Belmont, P. J. *et al.* Coordination of growth and endoplasmic reticulum stress signaling by regulator of calcineurin 1 (rcan1), a novel atf6-inducible gene. *J Biol Chem* **283**, 14012–21 (2008).

28. Hetz, C. & Glimcher, L. H. Fine-tuning of the unfolded protein response: Assembling the ire1alpha interactome. *Mol Cell* **35**, 551–61 (2009).

29. Urano, F. *et al.* Coupling of stress in the er to activation of jnk protein kinases by transmembrane protein kinase ire1. *Science* **287**, 664–6 (2000).

30. Jensen, L. J. *et al.* String 8–a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res* **37**, D412–6 (2009).

31. Linding, R. *et al.* Systematic discovery of in vivo phosphorylation networks. *Cell* **129**, 1415–26 (2007).

32. Lu, P. D., Harding, H. P. & Ron, D. Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. *J Cell Biol* **167**, 27–33 (2004).

33. Hamanaka, R. B., Bobrovnikova-Marjon, E., Ji, X., Liebhaber, S. A. & Diehl, J. A. Perk-dependent regulation of iap translation during er stress. *Oncogene* **28**, 910–20 (2009).

34. Harding, H. P. *et al.* An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. *Mol Cell* **11**, 619–33 (2003).

35. Ron, D. & Habener, J. F. Chop, a novel developmentally regulated nuclear protein that dimerizes with transcription factors c/ebp and lap and functions as a dominant-negative inhibitor of gene transcription. *Genes Dev* **6**, 439–53 (1992).

36. Wang, H. G. *et al.* Ca2+-induced apoptosis through calcineurin dephosphorylation of bad. *Science* **284**, 339–43 (1999).

37. Lei, K. & Davis, R. J. Jnk phosphorylation of bim-related members of the bcl2 family induces bax-dependent apoptosis. *Proc Natl Acad Sci U S A* **100**, 2432–7 (2003).

38. Putcha, G. V. *et al.* Jnk-mediated bim phosphorylation potentiates bax-dependent apoptosis. *Neuron* **38**, 899–914 (2003).

39. Yamamoto, K., Ichijo, H. & Korsmeyer, S. J. Bcl-2 is phosphorylated and inactivated by an ask1/jun n-terminal protein kinase pathway normally activated at g(2)/m. *Mol Cell Biol* **19**, 8469–78 (1999).

40. Schoeberl, B., Eichler-Jonsson, C., Gilles, E. D. & Müller, G. Computational modeling of the dynamics of the map kinase cascade activated by surface and internalized egf receptors. *Nat Biotechnol* **20**, 370–5 (2002).

41. Chen, W. W. *et al.* Input-output behavior of erbb signaling pathways as revealed by a mass action model trained against dynamic data. *Mol Syst Biol* **5**, 239 (2009).

42. Helmy, M. *et al.* Predicting novel features of toll-like receptor 3 signaling in macrophages. *PLoS One* **4**, e4661 (2009).

43. Song, S. O., Chakrabarti, A. & Varner, J. D. Ensembles of signal transduction models using pareto optimal ensemble techniques (poets). *Biotechnol J* **5**, 768–80 (2010).

44. Yamamoto, K. *et al.* Transcriptional induction of mammalian er quality control proteins is mediated by single or combined action of atf6alpha and xbp1. *Dev Cell* **13**, 365–76 (2007).

45. Lee, K. *et al.* Ire1-mediated unconventional mrna splicing and s2p-mediated atf6 cleavage merge to regulate xbp1 in signaling the unfolded protein response. *Genes Dev* **16**, 452–66 (2002).

46. Haze, K., Yoshida, H., Yanagi, H., Yura, T. & Mori, K. Mammalian transcription factor atf6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. *Mol Biol Cell* **10**, 3787–99 (1999).

47. Hai, T. W., Liu, F., Coukos, W. J. & Green, M. R. Transcription factor atf cdna clones: an extensive family of leucine zipper proteins able to selectively form dna-binding heterodimers. *Genes Dev* **3**, 2083–90 (1989).

48. Malhotra, J. D. & Kaufman, R. J. The endoplasmic reticulum and the unfolded protein response. *Semin Cell Dev Biol* **18**, 716–31 (2007).

49. Rao, R. V. & Bredesen, D. E. Misfolded proteins, endoplasmic reticulum stress and neurodegeneration. *Curr Opin Cell Biol* **16**, 653–62 (2004).

50. Kokame, K., Kato, H. & Miyata, T. Identification of erse-ii, a new cis-acting element responsible for the atf6-dependent mammalian unfolded protein response. *J Biol Chem* **276**, 9199–205 (2001).

51. Yoshida, H. *et al.* Atf6 activated by proteolysis binds in the presence of nf-y (cbf) directly to the cis-acting element responsible for the mammalian unfolded protein response. *Mol Cell Biol* **20**, 6755–67 (2000).

52. Lin, J. *et al.* IRE1 signaling affects cell fate during the unfolded protein response. *Science* **318**, 944 (2007).

53. Rao, R. *et al.* Coupling Endoplasmic Reticulum Stress to the Cell Death Program AN Apaf-1-INDEPENDENT INTRINSIC PATHWAY. *J. Biol. Chem.* **277**, 21836–21842 (2002).

54. Ma, Y. & Hendershot, L. M. Delineation of a negative feedback regulatory loop that controls protein translation during endoplasmic reticulum stress. *J Biol Chem* **278**, 34864–73 (2003).

55. Ameri, K. & Harris, A. L. Activating transcription factor 4. *Int J Biochem Cell Biol* **40**, 14–21 (2008).

56. Oyadomari, S. & Mori, M. Roles of CHOP/GADD153 in endoplasmic reticulum stress. *Cell Death & Differentiation* **11**, 381–389 (2003).

57. Ron, D. & Walter, P. Signal integration in the endoplasmic reticulum unfolded protein response. *Nat Rev Mol Cell Biol* **8**, 519–29 (2007).

58. Lee, T. G., Tomita, J., Hovanessian, A. G. & Katze, M. G. Purification and partial characterization of a cellular inhibitor of the interferon-induced protein kinase of mr 68,000 from influenza virus-infected cells. *Proc Natl Acad Sci U S A* **87**, 6208–12 (1990).

59. Yan, W. *et al.* Control of perk eif2alpha kinase activity by the endoplasmic reticulum stress-induced molecular chaperone p58ipk. *Proc Natl Acad Sci U S A* **99**, 15920–5 (2002).

60. van Huizen, R., Martindale, J. L., Gorospe, M. & Holbrook, N. J. P58ipk, a novel endoplasmic reticulum stress-inducible protein and potential negative regulator of eif2alpha signaling. *J Biol Chem* **278**, 15558–64 (2003).

61. Novoa, I., Zeng, H., Harding, H. P. & Ron, D. Feedback inhibition of the unfolded protein response by gadd34-mediated dephosphorylation of eif2alpha. *J Cell Biol* **153**, 1011–22 (2001).

62. Zhan, Q. *et al.* The gadd and myd genes define a novel set of mammalian genes encoding acidic proteins that synergistically suppress cell growth. *Mol Cell Biol* **14**, 2361–71 (1994).

63. Brush, M. H., Weiser, D. C. & Shenolikar, S. Growth arrest and dna damage-inducible protein gadd34 targets protein phosphatase 1 alpha to the endoplasmic reticulum and promotes dephosphorylation of the alpha subunit of eukaryotic translation initiation factor 2. *Mol Cell Biol* **23**, 1292–303 (2003).

64. Yoshida, H., Uemura, A. & Mori, K. pxbp1(u), a negative regulator of the unfolded protein response activator pxbp1(s), targets atf6 but not atf4 in proteasome-mediated degradation. *Cell Struct Funct* **34**, 1–10 (2009).

65. Yoshida, H., Oku, M., Suzuki, M. & Mori, K. pxbp1(u) encoded in xbp1 pre-mrna negatively regulates unfolded protein response activator pxbp1(s) in mammalian er stress response. *J Cell Biol* **172**, 565–75 (2006).

66. Tirosh, B., Iwakoshi, N. N., Glimcher, L. H. & Ploegh, H. L. Rapid turnover of unspliced xbp-1 as a factor that modulates the unfolded protein response. *J Biol Chem* **281**, 5852–60 (2006).

67. Gu, F. *et al.* Protein-tyrosine phosphatase 1b potentiates ire1 signaling during endoplasmic reticulum stress. *J Biol Chem* **279**, 49689–93 (2004).

68. Luo, D. *et al.* Aip1 is critical in transducing ire1-mediated endoplasmic reticulum stress response. *J Biol Chem* **283**, 11905–12 (2008).

69. Zhang, H. *et al.* Aip1/dab2ip, a novel member of the ras-gap family, transduces traf2-induced ask1-jnk activation. *J Biol Chem* **279**, 44955–65 (2004).

70. Hetz, C. *et al.* Proapoptotic bax and bak modulate the unfolded protein response by a direct interaction with ire1alpha. *Science* **312**, 572–6 (2006).

71. Xu, Q. & Reed, J. C. Bax inhibitor-1, a mammalian apoptosis suppressor identified by functional screening in yeast. *Mol Cell* **1**, 337–46 (1998).

72. Bailly-Maitre, B. *et al.* Cytoprotective gene bi-1 is required for intrinsic protection from endoplasmic reticulum stress and ischemia-reperfusion injury. *Proc Natl Acad Sci U S A* **103**, 2809–14 (2006).

73. Lisbona, F. *et al.* Bax inhibitor-1 is a negative regulator of the er stress sensor ire1alpha. *Mol Cell* **33**, 679–91 (2009).

74. Marcu, M. G. *et al.* Heat shock protein 90 modulates the unfolded protein response by stabilizing ire1alpha. *Mol Cell Biol* **22**, 8506–13 (2002).

75. Gupta, S. *et al.* Hsp72 protects cells from er stress-induced apoptosis via enhancement of ire1alpha-xbp1 signaling through a physical interaction. *PLoS Biol* **8**, e1000410 (2010).

76. Fonseca, C., Fleming, P. *et al.* Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization. In *Proceedings of the fifth international conference on genetic algorithms*, 416–423 (Citeseer, 1993).

77. Gadkar, K., Doyle III, F., Crowley, T. & Varner, J. Cybernetic model predictive control of a continuous bioreactor with cell recycle. *Biotechnology progress* **19**, 1487–1497 (2003).

78. Varner, J. Large-scale prediction of phenotype: concept. *Biotechnology and bioengineering* **69**, 664–678 (2000).

79. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *International joint Conference on artificial intelligence*, vol. 14, 1137–1145 (Citeseer, 1995).

Acknowledgements

The project described was supported by Award Number #U54CA143876 from the National Cancer Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health. In addition, special thanks to Gaurav Chakravorty and Mansi Singhal for their generous support.

Author contributions statement

Additional information

To include, in this order: **Accession codes** (where applicable); **Competing financial interests** (mandatory statement).

The corresponding author is responsible for submitting a competing financial interests statement on behalf of all authors of the paper. This statement must be included in the submitted article file.