MACQUARIE
University

Engineering Biosensors for the Detection of Unfolded Protein Response in *Saccharomyces cerevisiae*

Kai Peng

A thesis submitted in partial fulfilment of the degree of Master of Research

Supervisor: Dr. Thomas Williams

Department of chemistry and biomolecular sciences

10th October 2017

OGTR ethics approval reference number: 5201401059

Declaration of Authorship and Of Thesis Submission History

This is to certify that I, Kai Peng, being a candidate for the degree of Master of Research, have not submitted this work to any other university or educational institution for the purposes of attaining a higher degree. I also certify that this work is my own, and was not taken, in whole or in part, from un-attributed sources.

Signature of Candidate
Date: 9/10/2017

Full Name & Signature of Witness
Date: 10-10-2017

Table of Contents

Abbreviations

bp	base pair
DTT	Dithiothreitol
ER	Endoplasmic reticulum
ERAD	ER associated degradation
eroGFP	ER-targeted redox sensitive GFP
FITC	Fluorescein
FSC	Forward scatter
GFP	Green fluorescent protein
ORF	open reading frame
SC	Synthetic complete
sfGFP	Super-fold GFP
Tm	Tunicamycin
URA	Uracil
UPR	Unfolded protein response
UTR	Untranslated region

Abstract

The unfolded protein response (UPR) is activated when the protein folding capacity of the endoplasmic reticulum (ER) is overwhelmed by the accumulation of unfolded proteins inside the ER lumen. In *S. cerevisiae,* the UPR facilitates the up- and down-regulations of a collection of UPR target genes via the *IRE1/HAC1* dependent signalling pathway to bring the ER back to homeostasis. In the application of *S. cerevisiae* as a cell factory for producing heterologous proteins, low levels of target protein secretion are often associated with the up-regulation of UPR.

In this study we constructed nine UPR biosensors using the putative promoter sequences of nine UPR genes including *DER1, ERO1, EUG1, HAC1, KAR2, PDI1, PMT2, SEC12* and *OST2* as sensors, and green fluorescent protein as reporter. We have evaluated these biosensors in their capacities to detect the activation and different levels of UPR, and found that the sensors incorporating the *DER1* and *ERO1* putative promoter sequences gave the best responses across a range of UPR induction levels. This study highlighted the prospect of providing a real-time and high-throughput detection of unfolded protein induced stress and its concomitant influences on protein secretion.

1. Introduction

1.1 *S. cerevisiae* as a cell factory for heterologous protein production

Since the onset of the genetic engineering in the early 1970s, the application of microbial cells as industrial-scale production platforms has rapidly expanded (Cohen et al., 1973). Today, hundreds of chemicals and proteins are produced using recombinant microbial cells (Kuhad et al., 2011; Ferrer-Miralles and Villaverde, 2013; Rosano and Ceccarelli, 2014; Wang et al., 2017). The baker's yeast *S. cerevisiae* is a popular choice as a microbial cell factory, with a long standing fermentation history, industrial robustness and amenability to genetic modification through the extensive knowledge and well-developed engineering toolkits. Yeast's ability to perform the required post-translational modifications to produce correctly folded eukaryotic proteins makes it ideal for the commercial production of biopharmaceuticals, with examples including human coagulation factor, human albumin, anticoagulant, hirudin, hormones like insulin and human growth hormone and a range of recombinant vaccines (Hou et al., 2012; Martı´nez et al., 2012; Walsh, 2014; Williams et al., 2016; Wang et al., 2017). As with most fungi, producing proteins in yeast has an additional benefit: the ability to efficiently secrete proteins into the extracellular space, greatly reducing the cost of downstream purification. Although many commercial proteins are produced in yeast, the number is relatively low compared with that of bacteria, as the yield of many heterologous proteins failed to reach economically viable titres in *S. cerevisiae* (Idiris et al., 2010; Hou et al., 2012).

Several factors have been reported to limit the secretory protein yields by *S. cerevisiae*, such as the proteolytic degradation (Idiris et al., 2010). But most studies suggested that the bottlenecks along the secretory pathway are the underlying cause for the poor protein secretion (Valkonen et al., 2003; Gasser et al., 2008; Idiris et al., 2010). This is not a trivial problem, as the yeast secretory pathway is a highly regulated and dynamic system, engaging almost a third of the 6,000 or so genes within the *S .cerevisiae* genome. These protein-specific bottlenecks in the secretory pathway are the major reasons for the slow development of super-secreting yeast strains. The ambitious attempts to identify these bottlenecks and counteract with rational engineering approaches frequently revealed that no single genetic alteration could relieve the bottlenecks for all the heterologous proteins (Kroukamp et al. 2017). Although successes in engineering are not always straight forward or intuitive, the increasing knowledge of the secretion pathway serves as a solid foundation for future endeavours.

1.2 The secretory pathway in *S. cerevisiae*

The secretory pathway is the protein transportation system involving the shuttling of newly synthesised proteins via specialised membrane vesicles from the ER to other internal compartments like the Golgi apparatus, vacuole and endosome, as well as for expulsion of these protein cargos into the environment by the fusion of secretion vesicles to the cell membrane (Rapoport 2007). For secretory proteins to travel along the secretory pathway, their nascent polypeptides are first translocated into ER and submitted to the processing and subsequent quality control steps e.g. proper folding, disulphide bond formation, glycosylation and in some cases assembly into multi-subunit complexes (Hou et al., 2012). The protein quality control mechanism in the ER ensures that only properly folded proteins are transported further along the secretory pathway; while the terminally misfolded proteins are removed from the ER and degraded via the ER associated degradation (ERAD) pathway (Travers et al., 2000). Further processing is carried out in the Golgi, which may include the cleavage of the pro-sequences and the extension of both *O*- and *N*-linked glycans. The Golgi apparatus also exercises post-ER quality control, where proteins not suitable for secretion are sorted and directed to endosomes or vacuoles for degradation (Arvan et al., 2002). Properly folded, modified and matured proteins are further packed by the post-Golgi sorting mechanism to be secreted to the extracellular space.

1.3 The *IRE1/HAC1* dependent unfolded protein response (UPR) signalling pathway

The ER plays a critical role in the secretory pathway, as the primary location for protein folding and glycosylation. Although the ER has tightly regulated mechanisms in place to maintain the protein processing homeostasis, its protein processing capacity can be overwhelmed by the protein processing workload under various situations like redox balance change in the ER, glucose starvation, disruption of cellular calcium balance, inhibition of glycosylation, inability of protein export, overexpression of heterologous proteins or viral infections, resulting in the accumulation of unfolded proteins in the ER lumen and causing ER stress (Valkonen et al., 2003; Tang et al., 2015). When this happens, it not only impairs the cell's ability to process organelle-targeted proteins, but also abolishes the protein secretion and growth by limiting the transportation of lipids and proteins to the sites of membrane expansion.

In order to relieve the ER stress, the UPR pathway is activated. The UPR is a conserved protein quality control mechanism in eukaryotes that regulates the UPR target genes to restore the

homeostasis between the protein processing workload and the processing capacity in the ER (Mori 2009). An extensive list of the UPR regulated genes can be found in the studies by Travers et al. (2000) and Kimata et al. (2006). In *S. cerevisiae*, the UPR pathway is less complexed than in metazoans (Walter and Ron 2011), with only one UPR transducer (Ire1p) and one UPR transcription factor (Hac1p). *S. cerevisiae* mutants lacking a functional UPR pathway are hypersensitive to chemicals and cultivation conditions that induce ER stress (Chawla et al. 2011; Rubio et al. 2011).

The UPR pathway relies on the coordination between *HAC1* and *IRE1*. As the gene responsible for activating the transcription of UPR target genes, *HAC1* is first transcribed into immature *HAC1* mRNA containing a 252 base pairs intron flanked by two exons. This intron forms a long-range base pairing loop with the 5'-untranslated region (UTR) of the immature *HAC1* mRNA (Fig. 1), inhibiting its translation when the UPR is not activated (Rüegsegger et al., 2001).

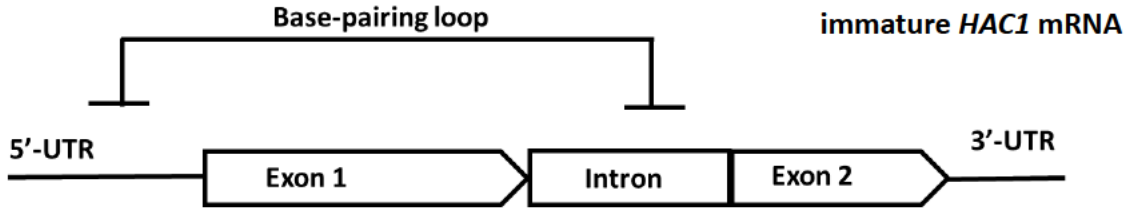


Figure 1. Schematic of the immature *HAC1* mRNA. A long-range base-pairing loop between the intron and 5'-UTR prevents its translation without UPR activation. The RNase activity of Ire1p complex formed upon UPR activation removes the intron, allowing its translation into Hac1p.

IRE1 is the other critical player in the UPR, as the gene encoding Ire1p, the ER stress sensor. Ire1p is an ER single transmembrane protein with a hydrophobic unfolded protein binding domain on the ER luminal side, plus one kinase domain and one RNase domain both exposed on the cytosolic side (Walter and Ron, 2011). The workflow of the UPR pathway is illustrated in Figure 2. When the protein processing in ER is in homeostasis, Ire1p exist as monomers. Upon the protein processing capability reaching its limit in ER, the binding of accumulated unfolded proteins to Ire1p triggers the oligomerization of Ire1p's luminal domains, followed by the oligomerization of the cytosolic domains (van Anken et al., 2014). In this way, Ire1p assemble into higher order cluster (Korennykh et al., 2009) and its cytosolic endonuclease activity becomes activated. The unspliced immature *HAC1* mRNA dock onto the cytosolic linker of the Ire1p complex (Aragon et al., 2009) and the *HAC1* intron is spliced out by Ire1p's RNAse activity through an unconventional spliceosome-

independent mechanism, producing mature *HAC1* mRNA available for translation (Ogawa et al., 2004). As the only known UPR transcription factor in *S. cerevisiae*, Hac1p then up or down regulates a collection of genes termed UPR genes through recognizing and binding to UPR element (UPRE), which is a cis-acting consensus sequence upstream of many, but not all UPR genes (Mori et al., 1998). Three consensus UPREs are currently known, termed UPRE-1, 2 and 3 respectively with consensus sequences as: CANCNTG (Mori et al., 1998), T(C/T)ACGTGT(C/T)(A/C) and AGGACAAC (Patil et al., 2004). Besides Hac1p, later studies found Gcn4p, a bZip transcription activator responsible for inducing amino acid biosynthesis under amino acid starvation, and its activator Gcn2p are also compulsory for the regulation of the UPR genes. Deletion of the *GCN4* gene, while keeping *HAC1* intact also abolishes the UPR function (Herzog et al., 2013).

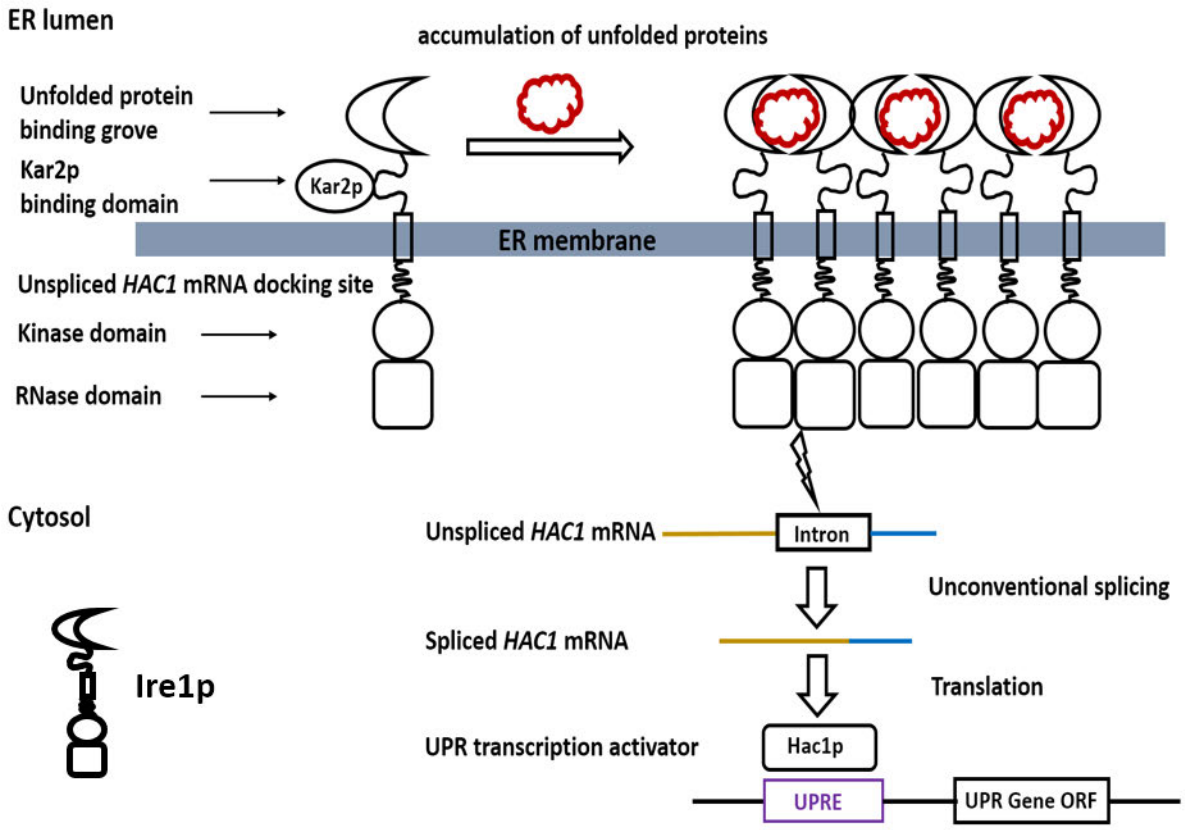


Figure 2. Mechanism of the UPR signalling pathway in *S. cerevisiae*. In the unstressed situation, Ire1p exist as monomer associated with Kar2p. Upon sensing the accumulation of unfolded proteins, Ire1p undergo oligomerization to form high order complex. Kar2p dissociate with Ire1p in this process. The oligomerization activates the endonuclease activity on Ire1p's RNase domain, which then specifically removes the intron from the *HAC1* mRNA to produce mature *HAC1* mRNA. The spliced mature *HAC1* mRNA can then be translated into Hac1p that subsequently up or down regulates the UPR genes, by binding with UPREs in the promoters of these genes. Imaged adapted from van Anken et al., 2014.

Besides the accumulation of unfolded proteins, later studies found that other events could also activate the *IRE1/HAC1* signalling pathway. One is inositol depletion or starvation which leads to the aggregation of Ire1p possibly through the disruption or abnormality of ER membrane (Volmer and Ron, 2015; Promlek et al., 2011). Another unfolded protein independent trigger of UPR recently confirmed is ethanol stress (6-8% ethanol), although the molecular mechanism of which is still under investigation (Navarro-Tapia et al., 2017).

1.4 The function of the UPR

About 400 UPR genes have been identified up to date, counting for around 6% of all the genes in *S. cerevisiae* genome. The effects of the UPR gene products cover a wide range of cellular responses and pathways, including: 1) increasing the ER volume by upregulating lipid and inositol synthesis to accommodate more proteins; 2) increasing the ER translocation machinery to improve protein import into ER; 3) increasing the protein processing capacity inside the ER lumen and maintain the ER redox balance by enhancing the production of various ER resident proteins, for example the ER chaperones Kar2p to assist protein folding, Pdi1p and Ero1p for disulphide bond formation and Pmt1p to improve glycosylation rates; 4) up-regulating the vesicle trafficking capacity by enhancing COPII vesicle formation; 5) activating some components of the ERAD pathway like Der1p, to degrade terminally misfolded proteins; 6) down-regulating the expression of some secretory proteins to temporarily free up the secretion pathway (Travers et al., 2000; Otte and Barlowe, 2004; Kimata et al., 2006; Hou et al., 2012). Around half the UPR genes contain at least one of the already known UPREs in their promoters (Patil et al., 2004), while the consensus sequences for Hac1p binding to the remaining UPR genes are yet to be identified.

1.5 The interplay between UPR and ERAD

Both as protein quality control mechanisms, the UPR and ERAD actively interact to maintain the ER homeostasis. While UPR increases the strength of ERAD by up regulating many genes involved in ERAD like *DER1*, *HRD1* and *DOA4*, ERAD removes terminally misfolded proteins that could not be rescued by UPR (Mori et al. 2009). Interestingly, *S. cerevisiae* cells remains viable and grow in the absence of either the UPR or ERAD, as long as the other function remains intact and without facing extreme or prolonged ER stress (Travers et al., 2000). Without ERAD, UPR is constitutively activated at low levels to compensate for this loss in protein quality control, but it does not lead to

detectable growth impairment (Zhou and Schekman, 1999). Similarly, ERAD on its own is sufficient to eliminate the misfolded proteins and protect the cell from impaired growth as it was shown that the level of misfolded proteins in unchallenged UPR knockout mutants are relatively low under the unstressed conditions (Valkonen et al., 2003).

1.6 The correlation between UPR and heterologous protein production in *S. cerevisiae*

As mentioned earlier, *S. cerevisiae* does not always secrete heterologous proteins at high levels, mainly due to the bottlenecks at several steps along the secretory pathway. Of these bottlenecks, the protein folding and quality control in the ER is frequently found to be the rate limiting steps for protein secretion (Valkonen et al., 2003; Gasser et al., 2008). Heterologous protein production in yeast are often facilitated by introducing the genes of interest at high copy numbers or under the transcriptional regulation of strong promoters to increase the transcription and potentially protein production. Considering how protein modification and folding proceed in the ER, it is not surprising that the overexpression of heterologous proteins frequently causes the accumulation of unfolded proteins in the ER, consequently inducing the UPR (Liu et al., 2013). This often results in reduced protein secretion titres, due to the longer retention times in the ER and elevated protein degradation (Mattanovich et al., 2004; Rakestraw and Wittrup et al., 2005; Xu et al., 2005; Young et al. 2011). The cellular response to heterologously produced proteins was elegantly demonstrated by Gasser et al. (2008), observing significant increases in the expression levels of many known UPR responsive genes in antibody fragments secreting strains. By augmenting the protein folding capacity though the overexpression of an ER chaperone Pdi1p, they not only enhanced the level of the secreted antibody fragment, but also reduced UPR. Whether the enhanced product secretion was due to the attenuation of the UPR or augmentation from the higher levels of Pdi1p was unclear from this study, though. Another study by Ilmen et al. (2011) supported the former argument. The authors evaluated the secretion of 24 heterologous cellobiohydrolases (CBH) produced by *S. cerevisiae* and found the highest level of CBH secretion were generally correlated with the low levels of UPR.

The impact of the UPR on protein secretion and the potential for engineering this response to increase protein secretion, becomes more evident from the review by Kroukamp et al. (in press), indicating that approximately 25% of rational design approaches to increase protein secretion involved the modulation of ER chaperones and the UPR genes. These rational strategies are often

limited by the availability of knowledge on the specific bottlenecks and are often only applicable to only a small range of heterologous proteins due to the protein-specific requirements for efficient secretion. By monitoring the UPR signalling pathway, we can potentially get a glance of the protein folding stress in the ER and cell's potential to secrete heterologous proteins. To evaluate the UPR in real-time, biosensors for UPR activation could serve as a useful alternative to the traditional RNA studies. These biosensors could be developed into high throughput tools to assess a strain's ability to produce a certain protein at a high level, or alternatively evaluate the 'secretability' of different heterologous proteins produced in *S. cerevisiae*.

1.7 The detection of the UPR

Classically, the UPR is monitored by assaying the levels of the unspliced and spliced mRNA of the *HAC1* transcript (Travers et al., 2000; Kimata et al., 2006). This is a laborious and low throughput process, only able to provide insight on the ER physiology several hours after the fact. More recently, transcriptomic approaches have allowed a more comprehensive view of the gene regulation during the UPR, but is also unable to provide a real-time feedback. To facilitate the real-time insights into the activation of the UPR, *in vivo* UPR biosensors were devised. The current UPR sensors can be classified into two categories, based on their targets of detection: 1) ER environment sensors that monitor the protein folding environment inside the ER lumen and report the ER environment change leading to or caused by the accumulation of unfolded proteins; 2) UPR signalling sensors that make use of the UPR transcriptional activation, brought about by Hac1p or the unconventional splicing mechanism of *HAC1* mRNA.

The first ER environment UPR sensor was developed by Merksamer et al. (2008). In order to track the ER redox environment change during the UPR activation, the authors fused an ER-targeted redox-sensitive GFP (eroGFP) with a C-terminal HDEL ER retention signal (Fig. 3A). The excitation spectra of the eroGFP shift depending on the redox status of its environment. Thus, the ratio of eroGFP excitable at different wavelengths provides an estimate of the relative redox potential of the ER protein folding environment, which is critical to the oxidative folding of many proteins. A more reducing ER lumen indicate more unfolded proteins struggling with disulphide bond formation and isomerization, thus inducing a higher level of UPR that can be detected by the eroGFP sensor.

Another intriguing ER environment-UPR sensor was developed by Lajoie et al. in 2012, where Kar2p was fused with super-fold GFP (Fig. 3B). Kar2p assists protein folding inside the ER lumen by binding with the hydrophobic patches exposed on protein surface, or tagging a terminally misfolded protein for degradation. The more unfolded proteins accumulating in ER, the more Kar2p will be bound to them and have reduced diffusional mobility. The mobility change of Kar2p-sfGFP cause changes to the photo bleaching time of the sfGFP that can be measured using fluorescence microscope. In this way, a correlation between the diffusional mobility of Kar2p-sfGFP and the level of the unfolded protein load inside the ER could be established.

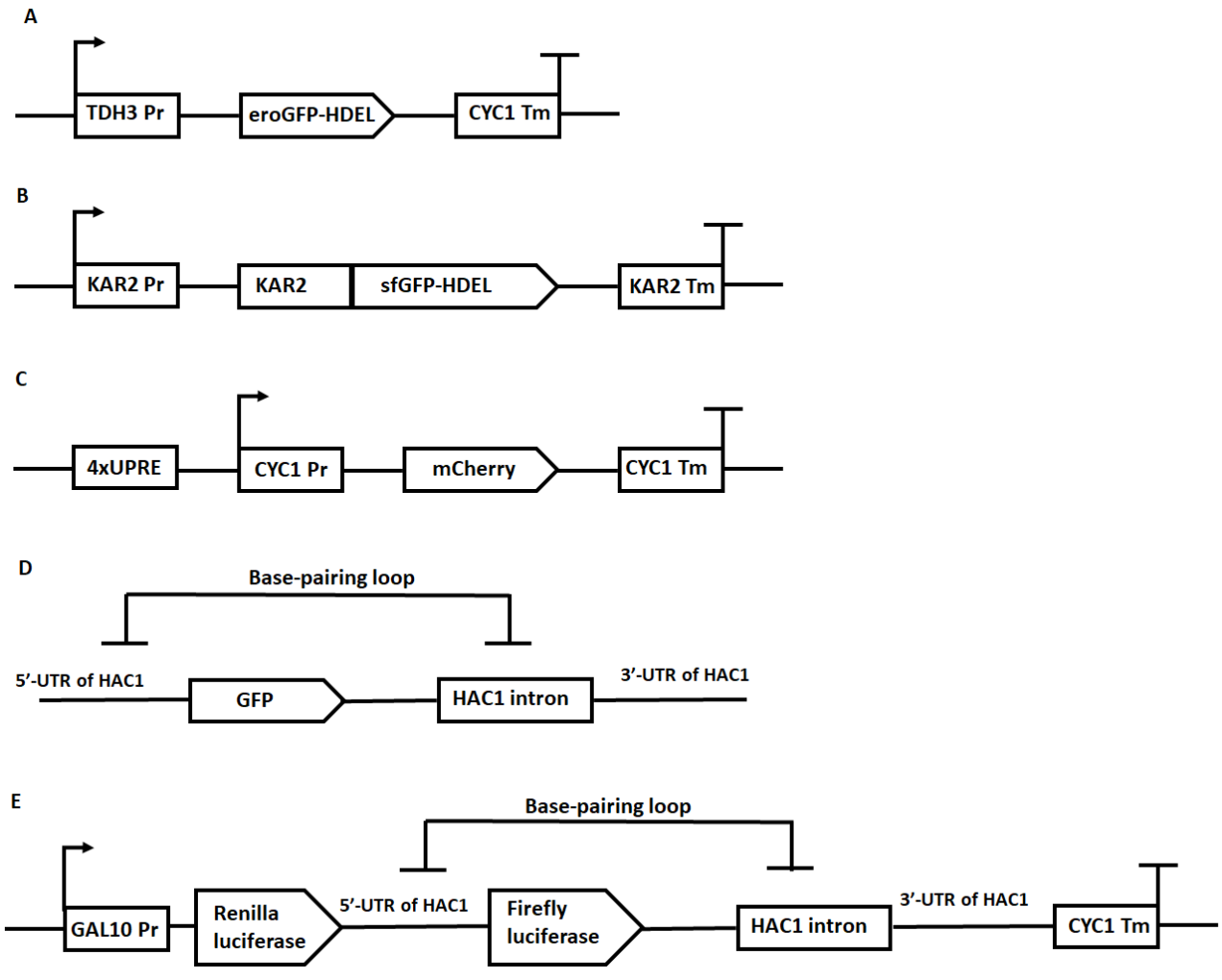


Figure 3. Schematics of previously reported UPR sensors. Pr: promoter; Tm: terminator. **A)** eroGFP ER redox sensor. **B)** Kar2-sfGFP diffusional mobility sensor. **C)** 4xUPRE-mCherry UPR transcription sensor. **D)** *HAC1* mRNA splicing reporter. **E)** Dual-luciferase UPR splicing reporter.

The most commonly used UPR signalling sensor relies on UPRE to detect Hac1p, the UPR transcription activator and reports via the subsequent expression of the downstream reporter protein. The earliest version of such sensor was implemented by Mori et al. (1996), where a 22 bp sequence containing the UPRE-1 of *KAR2* were used as the Hac1p detecting sequence. Later as the consensus UPRE was defined, the sensing sequence was further optimized to the 7bp UPRE-1 of *KAR2.* Merksamer et al. (2008) designed another UPR signalling sensor consisting of four UPRE-1 in tandem upstream of a *CYC1* minimal promoter, followed by a mCherry reporter protein (Fig. 3C). Since then, several studies on yeast UPR had engaged this sensor to detect the UPR activation (Pincus et al., 2010; Rubio et al., 2011).

Besides UPR transcription, another UPR signalling sensor makes use of the unconventional splicing of *HAC1* mRNA. One example replaced the first exon in the unspliced *HAC1* mRNA sequence with GFP sequence (Fig. 3D, Aragon et al., 2009). The resulting mRNA behaves like that of *HAC1* before splicing: in the unstressed situation, the translation of the reporter GFP is inhibited by a long-range base-pairing loop formed between the *HAC1* intron and the 5'-UTR. The activation of the UPR and the subsequent removal of the intron via the unconventional splicing enables the translation of the GFP reporter.

Another more intricate splicing reporter adopted a dual luciferases reporting system (Fang et al., 2015) as depicted in Fig. 3E. The long-range loop between the 5'-UTR of *HAC*1 mRNA and the *HAC1* intron represses the translation of the Firefly luciferase while allowing the translation of the Renilla luciferase. The RNAse activity of the Ire1p complex upon UPR activation splices out the inhibitory loop and enables the translation of the Firefly luciferase. Therefore, the UPR splicing activities can be probed by the ratiometric values of the two luciferases.

1.8 The aim of this study

The real-time reporting capabilities of biosensors for *in vivo* processes has been widely exploited in both applied biotechnology and fundamental approaches. Several biosensors for the detection of the UPR have been constructed for the elucidation of the mechanisms by which the ER stress is initiated, and have expanded our fundamental understanding of the UPR. Up to date, the UPR biosensor design mainly focused on the detection of the 'on and off' status of the UPR, or to monitor the changes in the ER stress in response to the genetic alterations in the response pathway. Here, we aimed to design, build and evaluate UPR biosensors with characteristics

suitable for the systematic detection of the protein secretion stress. We evaluated a selection of non-conventional putative UPR gene promoters on their sensitivities and dynamic ranges in response to ER stress. Several factors that reduced the sensitivity of the UPR sensors was also identified in this study, providing vital knowledge for further sensor optimisation.

2. Methods and materials

2.1 Strains and media

The *S. cerevisiae* strains used in this study are listed in Table 1. The CEN.PK2-1C *hac1Δ* strain was constructed by isolating the *hac1* locus of the BY4741 *hac1Δ* strain, which contained the open reading frame of the *HAC1* gene being replaced by the *kanMX4* cassette. Primers HAC1D_F and HAC1D_R (Table 2) were used to amplify the *hac1* locus from the BY4741 *Δhac1* strain, and the PCR product subsequently transformed into the CEN.PK2-1C strain. Putative CEN.PK2-1C *hac1* cells were selected on YPD agar (1% yeast extract (Sigma-Aldrich), 2% peptone (Sigma-Aldrich), 2% glucose and 2% bacteriological agar (Sigma-Aldrich)) plates supplemented with 300 µg/mL G418-sulfate.

To construct the CEN.PK2-1C strains with the integrated *DER1* and *ERO1* sensors, the regions on the pRS416-DER-GFP and pRS416-ERO-GFP (Table 1) plasmids, comprising of the *URA3* gene, *DER1*/*ERO1* putative promoter and GFP sequence were PCR amplified using the primers URA3_F and URA3_R (Table 2). The PCR products were used to transform CEN.PK2-1 and inserted into the *ura3* locus through homologous recombination to restore the inactive *ura3* allele. Putative CEN.PK2-1C strains with the integrated *DER1* and *ERO1* sensors were selected on SC^{-ura} agar plates containing 0.68% yeast nitrogen base without amino acids (Sigma-Aldrich), 0.192% yeast synthetic drop-out medium supplements without uracil (Sigma-Aldrich), 2% glucose and 2% bacteriological agar.

CEN.PK2-1C derivative strains carrying the pRS416-UPR promoter-GFP sensors were selected on SC^{-ura} agar plates and routinely cultured in SC^{-ura} media to maintain selection for the plasmids. For flow cytometry analysis, all strains were cultivated in buffered double strength SC^{-ura} media (1.36% yeast nitrogen base without amino acids, 0.384% yeast synthetic drop-out medium supplements without uracil, 2% glucose and 2% succinic acid). The pH was adjusted to 6.0 using NaOH.

2.2 *E. coli* and Yeast transformation

For *E. coli* transformations, plasmid DNA was added to chemically competent cells and incubated on ice for 30min, then heat shocked at 42 °C for 45 seconds and spread onto selective lysogeny broth (LB) plates: 1% tryptone (Sigma-Aldrich), 0.5% yeast extract (Sigma-Aldrich), 1% NaCl, 2%

agar supplemented with 100 µg/mL ampicillin. Putative *E. coli* colonies were routinely cultivated in LB media with 100 µg/mL ampicillin overnight, before used for plasmids extraction.

For yeast transformation, the lithium acetate/single-stranded DNA protocol described by Gietz et al. 2007 was used. In each reaction, approximately $1x10^8$ cells from overnight cultures were harvested (roughly equivalent to 1mL of culture with $OD_{600} = 3$) and resuspended in 48 µl 50% (w/v) PEG 3350, 7.2 µl 1M LiOAc, 10 µl single-stranded carrier DNA (2mg/mL), 6.8 µl MiliQ water plus up to 1 µg of DNA to be transferred. Cells were subsequently heat shocked at 42°C for 2 hours, with the exception for CEN.PK2-1C *Δhac1*, which were heat-sensitive and heat shocked for 1 hour instead. The cells were then harvested and selected on agar plates described previously.

Table 1. *S. cerevisiae* strains and plasmids used in this study.

Strain/Plasmid name	Relevant genotype	Source
Yeast strains		
CEN.PK2-1C	*MATa his3 Δ1 leu2-3_112 ura3-52 trp1-289*	Etian and Kotter, 2007
BY4741 *Δhac1*	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 hac1::kanMX4*	Brachmann et al., 1999
CEN.PK2-1C *Δhac1*	*MATa his3 Δ1 leu2-3_112 ura3-52 trp1-289 hac1::kanMX4*	This study
CEN.PK DER-GFP (Integrated *DER1* sensor)	*MATa his3 Δ1 leu2-3_112 trp1-289 URA3:: DER1pro-GFP*	This study
CEN.PK ERO-GFP (Integrated *ERO1* sensor)	*MATa his3 Δ1 leu2-3_112 trp1-289 URA3:: ERO1pro-GFP*	This study
Plasmids		
pJET1.2	*AmpR*	Genbank EF694056
pRS416	*ARS CEN6 URA3 AmpR*	Genbank U03450
pRS416-PGK-GFP	*ARS CEN6 URA3 AmpR* $PGK1_{pro}$*-GFP-*$CYC1_{ter}$	Kroukamp et al. unpublished
pRS416-DER-GFP	*ARS CEN6 URA3 AmpR* $DER1_{pro}$*-GFP-*$CYC1_{ter}$	This study
pRS416-ERO-GFP	*ARS CEN6 URA3 AmpR* $ERO1_{pro}$*-GFP-*$CYC1_{ter}$	This study
pRS416-EUG-GFP	*ARS CEN6 URA3 AmpR* $EUG1_{pro}$*-GFP-*$CYC1_{ter}$	This study
pRS416-HAC-GFP	*ARS CEN6 URA3 AmpR* $HAC1_{pro}$*-GFP-*$CYC1_{ter}$	This study
pRS416-KAR-GFP	*ARS CEN6 URA3 AmpR* $KAR2_{pro}$*-GFP-*$CYC1_{ter}$	This study
pRS416-PDI-GFP	*ARS CEN6 URA3 AmpR* $PDI1_{pro}$*-GFP-*$CYC1_{ter}$	This study
pRS416-PMT-GFP	*ARS CEN6 URA3 AmpR* $PMT1_{pro}$*-GFP-*$CYC1_{ter}$	This study
pRS416-SEC12-GFP	*ARS CEN6 URA3 AmpR* $SEC12_{pro}$*-GFP-*$CYC1_{ter}$	This study
pRS416-OST2-GFP	*ARS CEN6 URA3 AmpR* $OST2_{pro}$*-GFP-*$CYC1_{ter}$	This study
pRS416-GAL1-GFP	*ARS CEN6 URA3 AmpR* $GAL1_{pro}$*-GFP-*$CYC1_{ter}$	This study

2.3 DNA purification

Plasmid DNA was extracted from overnight *E. coli* cultures using the Monarch® Miniprep kit (NEB) as recommended by the supplier. Yeast genomic DNA extractions for PCR, were prepared using a modified protocol described by Lööke et al. (2011). Yeast cells from overnight cultures were lysed by incubating them with 200 mM lithium acetate and 1% sodium dodecyl sulphate at 70°C for 15 minutes. The cell lysate was precipitated with 100% ethanol. The pellet was subsequently washed with 70% ethanol and dried before dissolving in MiliQ water. DNA restriction fragments and PCR products were routinely extracted from agarose gels using the Zymoclean Gel DNA Recovery Kit (Zymo Research), according to the manufacturer's specifications.

Table 2. PCR primers used in this study.

Primer	Sequence	Fragment amplified
DER1_F	GATCCCGAATTC AATTTACCGACGCAC	*DER1* 250 bp putative promoter
DER1_R	GATCCCTTAATTAA ATTGCTTGGGCTT	
ERO1_F	GATCCCGAATTC GTCCTTTGGTTTCTC	*ERO1* 250 bp putative promoter
ERO1_R	GATCCCTTAATTAA GTTTTACCTGCAC	
EUG1_F	GATCCCGAATTC CTTTCCTGGTTAGTC	*EUG1* 250 bp putative promoter
EUG1_R	GATCCCTTAATTAA GGTATTATATGAG	
HAC1_F	GATCCCGAATTC TATACATCGTACCGA	*HAC1* 250 bp putative promoter
HAC1_R	GATCCCTTAATTAA AGTGGCGGTTGTT	
KAR2_F	GATCCCGAATTC ATTCATCTAATTTTG	*KAR2* 250 bp putative promoter
KAR2_R	GATCCCTTAATTAA GGTATGTTTGATA	
PDI1_F	GATCCCGAATTC TACCCTGTCGGGCGG	*PDI1* 250 bp putative promoter
PDI1_R	GATCCCTTAATTAA AACGGGATAGATG	
PMT1_F	GATCCCGAATTC CTTGGAACCAGAAGA	*PMT1* 250 bp putative promoter
PMT1_R	GATCCCTTAATTAA GACCGTGCTTGTA	
SEC12_F	GATCCCGAATTC TGTAATGTACTGGTC	*SEC12* 250 bp putative promoter
SEC12_R	GATCCCTTAATTAA AGTGGCTCGATTG	
SEC62_F	GATCCCGAATTC AACTCCTAAATCTGA	*SEC62* 250 bp putative promoter
SEC62_R	GATCCCTTAATTAA GTTCTCCTGTGTT	
OST2_F	GATCCCGAATTC CTGGAAATAACGAGT	*OST2* 250 bp putative promoter
OST2_R	GATCCCTTAATTAA ATTAGTCATGTAT	
GFP_R	TTATTTGTACAATTCATCCATACC	For verifying the UPR sensor plasmids
HAC1D_F	GAGTCGCAAATATACATCGTACCGAGTG	*hac1* locus of the BY4741 Δ*hac1* strain
HAC1D_R	CACACACTAACCGGAGACAGAACAG	
URA3_F	ATGTCGAAAGCTACATATAAGGAACGT GCTGCTACTCTTAACTATGCGGCATCAGAC	For preparing the integrated sensor constructs
URA3_R	TTAGTTTTGCTGGCCGCATCTTCTCAAATA TGCTTCCCAGCGCGCGCAATTAACCCT	

2.4 UPR biosensor construction and verification

We built nine UPR sensors using the putative UPR promoters of nine UPR up-regulated genes: *DER1, ERO1, EUG1, HAC1, KAR2, PDI1, PMT1, SEC12* and *OST2* (Table 3). The nine putative promoters were amplified by PCR from the genomic DNA extracted from the *S. cerevisiae* S288C strain, using primers for the respective putative UPR promoters listed in Table 2. The putative UPR promoter sequences were devised as the 250 base pairs upstream of the start codon of the respective UPR genes; 80bp longer than the predicted *KAR2* promoter sequence (Fig. 4A). These sequences all contained motives representing known UPRE elements. The motives were searched using the Geneious software.

Table 3. The functions of the UPR regulated genes used in this study

UPR gene	Function category	Protein
DER1	ERAD	ER membrane exporter of misfolded polypeptide
ERO1	Protein folding	ER thiol oxidase needed for protein disulphide bond formation
EUG1	Protein folding	ER protein disulphide bond isomerase
HAC1	UPR signalling	UPR transcription factor
KAR2	Protein folding/ translocation	ER chaperone
PDI1	Protein folding	ER protein disulphide bond isomerase
PMT1	Glycosylation	ER protein O-mannosyltransferase
SEC12	Vesicle transport	Guanine nucleotide exchange factor for COPII vesicle formation
OST2	Glycosylation	Subunit of ER lumen oligosaccharyltransferase complex

The *Eco*RI (GAATTC) and *Pac*I (TTAATTAA) restriction sites were incorporated into the 5'-end overhangs of the forward and reverse primers, respectively. The PCR products of the putative promoters were separated by agarose gel electrophoresis, isolated and cloned into the pJET1.2 cloning vector using the CloneJET PCR Cloning Kit (ThermoFisher) as recommended by the manufacturer and subsequently transformed into the chemically competent *E. coli* DH5α cells.

The UPR sensor plasmids were constructed by liberating the putative promoter sequences from the respective pJET2.1 plasmids with *Eco*RI and *Pac*I restriction digestion and the subsequent ligation of these fragments into the *Eco*RI and *Pac*I sites of the pRS416-PGK-GFP plasmids (Table 2) provided by Kroukamp et al. (unpublished), replacing the original *PGK1* promoter. A graphical illustration of the biosensor construction is shown in Fig. 4. In these constructs, the gene coding for green fluorescent protein (GFP) was used as reporter. The *GFP* orf was placed immediately downstream of the putative UPR promoters. Transcription of the *GFP* gene was terminated using the *S. cerevisiae CYC1* terminator sequence. These constructs would allow for the direct

measurement of the promoter activation in response to UPR stress, though increased GFP fluorescence. The pRS416 plasmids containing the respective putative UPR promoter sequences were verified by PCR amplification (Fig. 5) using the forward primers DER1_F, ERO1_F, EUG_1F, HAC1_F, KAR2_F, PDI1_F, PMT1_F, SEC12_F, OST2_F and the reserve primer GFP_R (Table 2), that amplified the section containing both the putative primers and the GFP reporter. The presence and integrity of the putative promoter sequences were also verified by Sanger sequencing (Macrogen, South Korea).

The chromosome integration versions of the *DER1* and *ERO1* sensors were obtained as described in the strains and media section. In this study we will refer to the specific UPR biosensors as the '*UPR GENE* sensor' for convenience, e.g. the *KAR2* sensor for the sensor driving GFP expression using the putative *KAR2* promoter sequence.

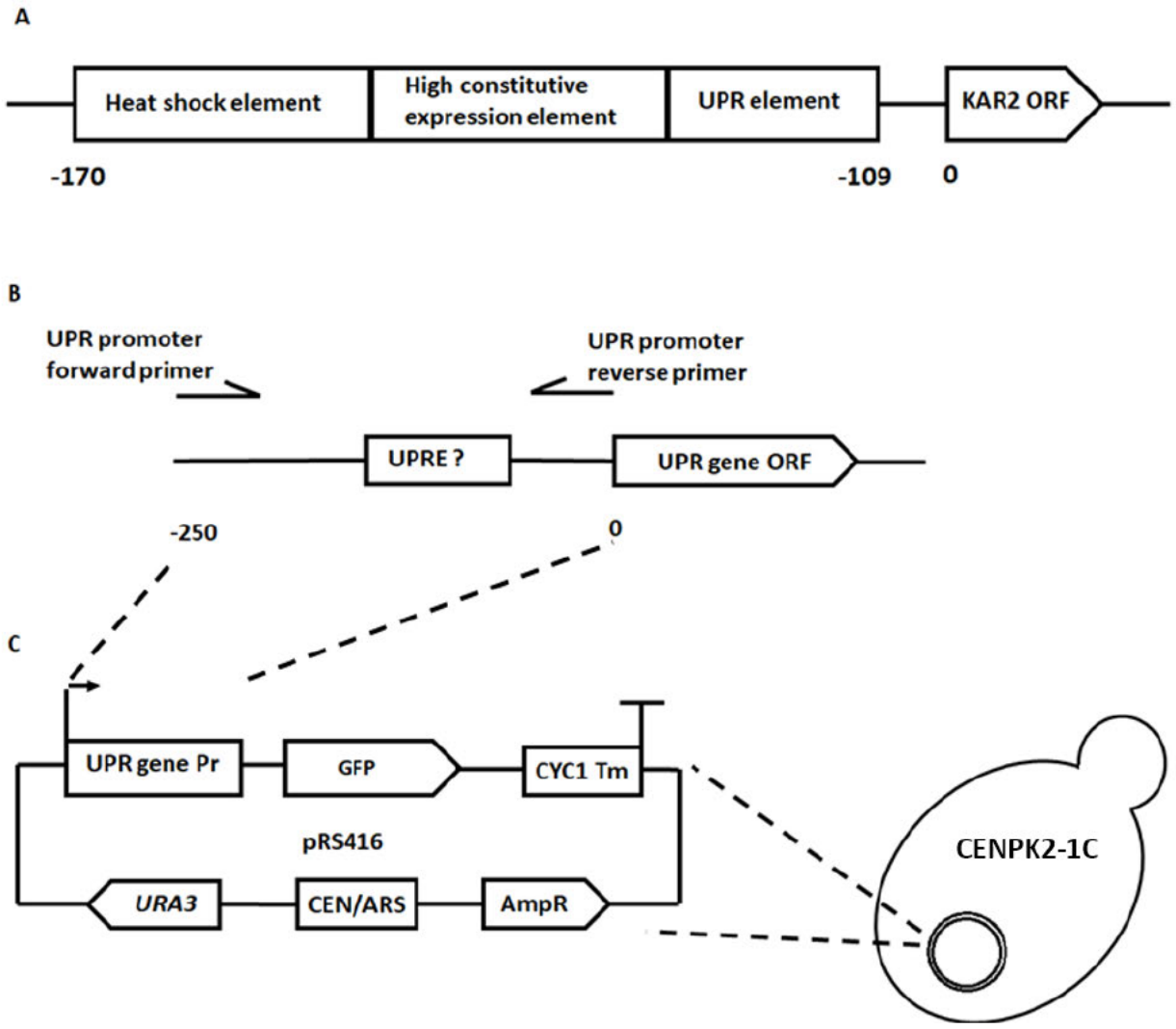


Figure 4. Schematics of **A)** *KAR2* promoter elements of *S. cerevisiae* (Mori et al., 1992), showing the location of UPRE and supporting the UPR responsiveness by the putative UPR promoters. **B)** the 250 bp region upstream UPR gene initiation codon was amplified as putative UPR promoter. **C)** UPR sensor plasmids were constructed using the putative promoter and transformed into CEN.PK2-1C strain.

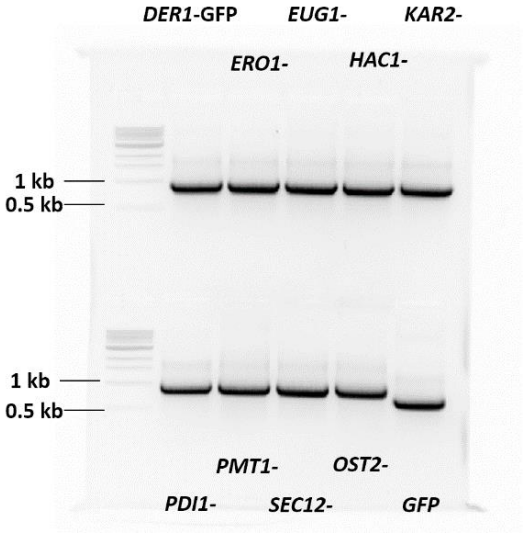


Figure 5. Gel image of the PCR products of the pRS416-UPR promoter-GFP sensors. For each sensor, the sequence containing the putative promoter and the GFP reporter was PCR amplified. The 900 bp bands indicated the successful linking between the putative promoter and GFP. The last lane at bottom was the PCR products of the GFP gene only, as a comparison.

2.5 UPR biosensor evaluation

Overnight pre-cultures were prepared for each of the UPR sensor containing strains. The pre-cultures were used to inoculate 30mL double strength buffered SC^{-URA} selective media to OD_{600}=0.15 at T_0.

For tests using DTT as a chemical UPR inducer, appropriate volumes of a 1 M DTT stock were added into each flask to obtain the final DTT concentrations of 0, 0.5, 1, 2 and 3 mM. All samples were prepared in triplicate. The cultures were incubated at 30°C with shaking at 200 rpm. Samples were collected at 30 minute intervals for 8 hours. The OD_{600} of each sample was determined using PHERAstar spectrophotometer (BMG Labtech) and the relative GFP fluorescence was determined with a Beckman Coutler CytoFlex S flow cytometer. The recommended gains based on the calibration beads were used for fluorescence signal acquisition in the flow cytometry analysis. The FITC channel (525nm) was monitored to collect GFP fluorescence signals. The flow rates were set between 6-30 µl/min and 10,000 events were recorded for each sample.

To evaluate the sensors' response to Tm, appropriate volumes of a 1 mg/mL Tm stock were added to each flask containing to obtain the final Tm concentrations of 0, 0.2, 0.5, 1 and 2 µg/mL. The subsequent incubation, sampling, cell densities and fluorescent signals reading were done in the same way as for the DTT tests, except the incubation time was capped at 4 hours after induction and samples were taken hourly, except for the *DER1* and *ERO1* sensors that were sampled every 30 minutes. The CEN.PK2-1C *Δhac1* strains containing the *DER1* and *ERO1* sensors, and the CEN.PK2-1C strains with the integrated *DER1* and *ERO1* sensors were prepared and analysed the same way as in the Tm test, except only one Tm concentration of 0.5 µg/mL was used.

3. Results

3.1 UPR sensor screening

An effective UPR sensors should have 1) good resolution power or sensitivity to distinguish between distinct levels of UPR; 2) a wide dynamic range where the sensor remain responsive; 3) a moderate level of signal strength; 4) and ideally have a close-to-linear response to ER stress levels. In this study we have evaluated the performances of the nine UPR sensors by the fluorescence signals produced in response to different ER stress levels. The subsequent results would provide insight into the best performing candidates from the nine UPR sensors, which would be the focus of our future development and optimisation.

3.1.1 Sensor evaluation using dithiothreitol (DTT) as the UPR inducer

With the exception of the *HAC1* and *KAR2* promoters, there have been no previous studies on the *in vivo* response of the putative UPR promoters used in this study. Therefore, in order to obtain a preliminary understanding of the responses of these UPR sensors we built, we first challenged the cells carrying different UPR sensors with DTT, a chemical inducer of UPR. DTT reversibly unfolds the polypeptides in the ER lumen, causing the accumulation of unfolded protein and triggering the UPR.

We first examined if the evaluated DTT concentrations (0.5, 1, 2 and 3 mM) had any impact on the yeast growth rate, since DTT was known to be growth-inhibitory (Lee et al., 2008). Except for the cells with the *OST2* sensor, DTT concentration of 2 mM was able to cause significant reduction in the growth rates compared to that of unstressed cells (Fig. 6). As a representative of the typical response to DTT, the *DER1* sensor containing cells achieved a growth rate of 0.402 h^{-1} when unstressed, but had significantly lower growth rates of 0.298 ($p=7.6x10^{-3}$) and 0.175 ($p= 4.9x10^{-4}$) in the presence of 2mM and 3mM DTT, respectively (Fig. 6). DTT concentrations below 1mM did not cause any detectable decrease in growth rates ($p>0.05$) of the cells carrying the *DER1, ERO1, KAR2, PDI1, PMT1, SEC12* and *GAL1* sensors (Fig. 6).

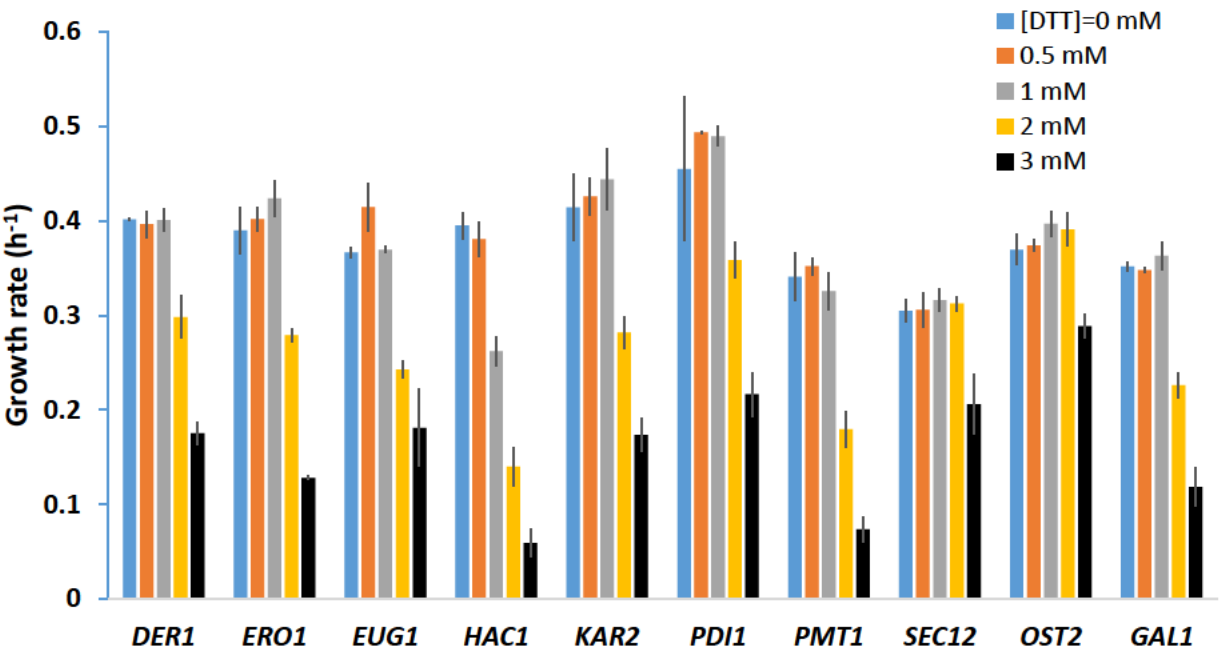


Figure 6. Growth rates of the cells carrying the UPR sensors at different DTT concentrations. The growth rates were calculated based on the cell densities incubated at 30˚C and shaking at 200 rpm over 8 hours. Three biological replicates were sampled. Error bars represent standard deviation.

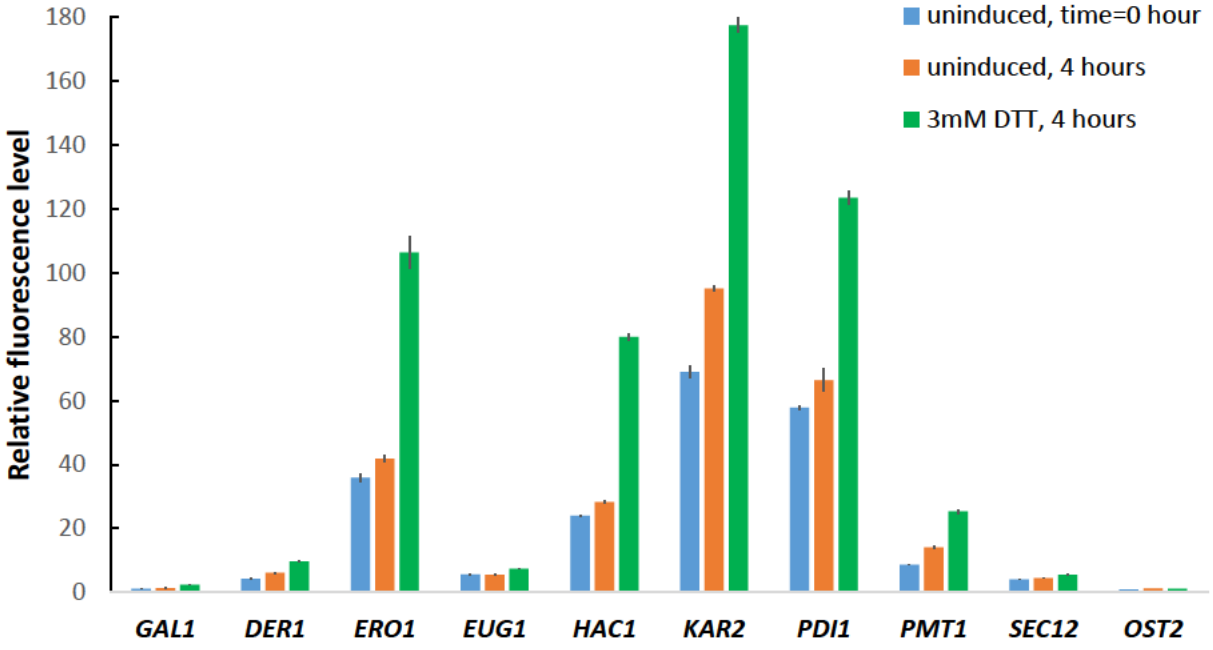


Figure 7. Comparison of the relative fluorescence signal levels of the UPR sensors induced by DTT. For each sensor, the uninduced basal signal levels at 0 and 4 hours, and the signals induced by 3mM DTT after 4 hours were displayed. The arbitrary fluorescent levels were normalised against the lowest fluorescent level across (uninduced *OST2* sensor, 0 hours). Samples were incubated at 30˚C and shaking at 200 rpm. Three biological replicates were sampled. Error bars represent the standard deviation from the mean value.

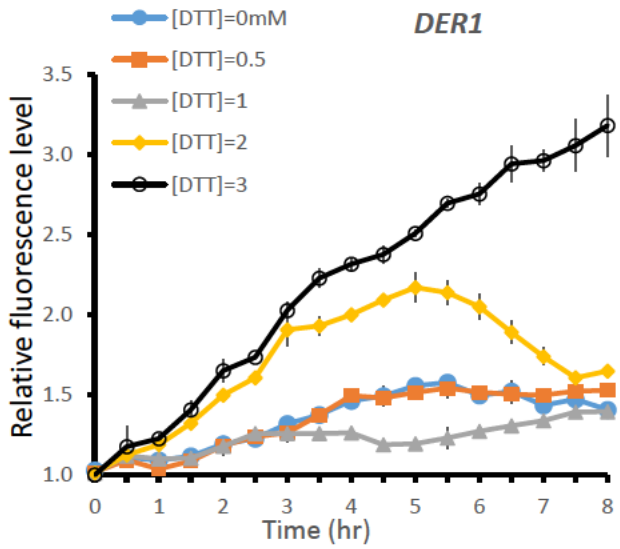
All the sensors had a basal level of GFP expression. With the exception of the *OST2* sensor, all the other sensors were responsive to DTT (Fig. 7 and Fig. 8). This might be due to a loss of essential UPREs or other promoter elements critical to core functionality. Therefore the *OST2* sensor was discarded from further test and discussion. This on the other hand reflected our limited knowledge about regulatory elements of UPR genes, even though the *S. cerevisiae* genome had been well-annotated.

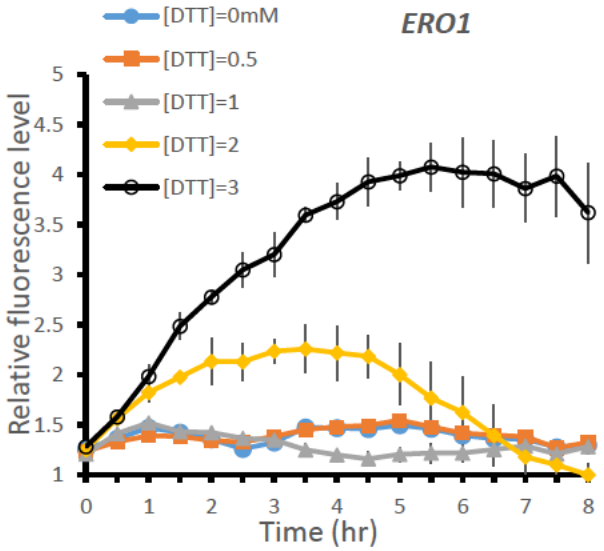
The basal fluorescence signals of the sensor in unstressed cells were all above the corresponding value of the *GAL1*-GFP control plasmid (Fig. 7). Only the *ERO1* and *HAC1* sensors had a relatively stable basal (uninduced) signal over 8 hours (Fig. 8); while the *DER1* ($p=4x10^{-4}$), *KAR2* (p=0.001), *PDI1* (p=0.027), *PMT1* (p=0.001) and *SEC12* (p=0.006) sensors displayed increased basal signal levels by 4 hours after induction (Fig. 8). The *EUG1* sensor displayed a decrease in basal fluorescence over time (Fig. 8).

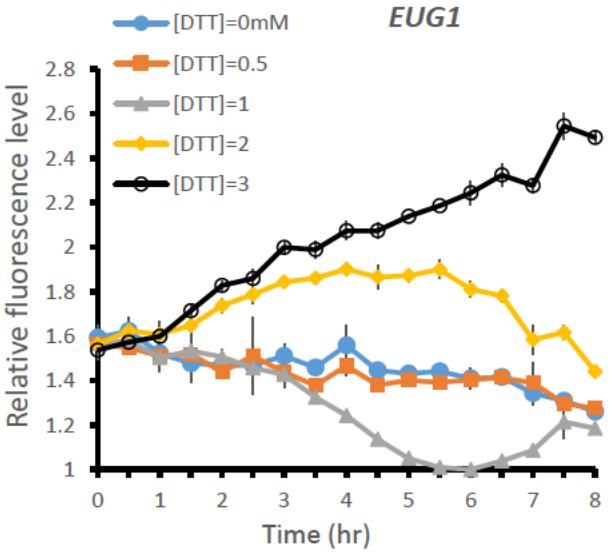
When the cells faced low ER stress represented by 0.5mM DTT, the fluorescence levels remained unchanged from that of the uninduced state (Fig. 8). Considering the unchanged growth rates at the same concentration (Fig. 7), it might indicate that DTT concentrations below this level could not cause substantial ER stress to trigger differentiated signals from the UPR sensors; or alternatively, the sensors were not sensitive enough to detect such low stress.

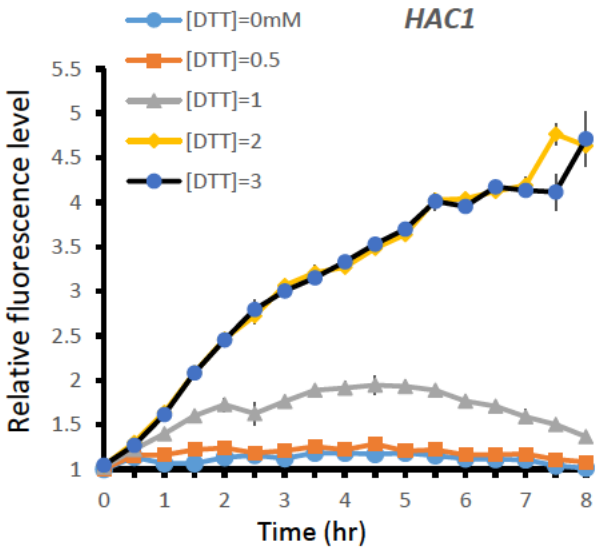
The fluorescence responses of the UPR sensors started to display more heterogeneity at 1mM DTT. The responses of the *PMT1* and *SEC12* sensors did not show observable deviations from the uninduced signal levels (Fig. 8). At 1 mM DTT, the *DER1, ERO1, EUG1, KAR2* and *PDI1* sensors signal levels fluctuated over time after induction, displaying fluorescence values below that of the unstressed cells at 4 hours after induction ($p<0.05$, Fig. 8). The *HAC1* sensor was the only sensor that had a significantly higher fluorescence levels at 1mM DTT compared to the uninduced cells (Fig. 8).

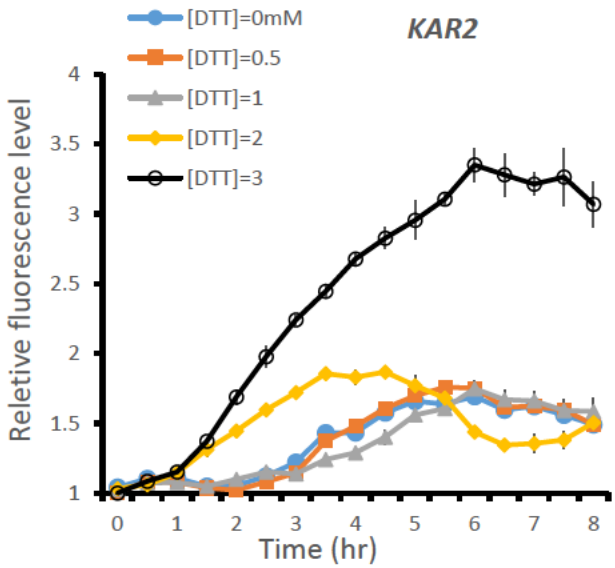
At DTT levels of 2 mM, the signal levels of the *DER1*, *ERO1*, *EUG1*, *KAR2* and *PDI1* sensors initially rose above the basal levels until 4-5 hours before declining (Fig. 8). The signal levels of the *ERO1*, *KAR2* and *PDI1* sensors even dropped below that of the uninduced cells (Fig. 8). Only the *PMT1* and *SEC12* sensors had kept their fluorescence levels at 2 mM DTT steadily rising and residing between the uninduced and 3 mM DTT curves over time (Fig. 8).

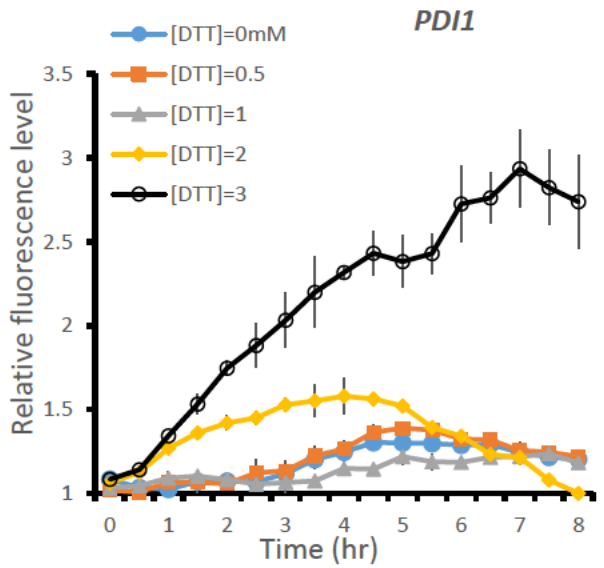

DER1
[DTT]=0mM
[DTT]=0.5
[DTT]=1
[DTT]=2
[DTT]=3
Relative fluorescence level
Time (hr)


ERO1
[DTT]=0mM
[DTT]=0.5
[DTT]=1
[DTT]=2
[DTT]=3
Relative fluorescence level
Time (hr)


EUG1
[DTT]=0mM
[DTT]=0.5
[DTT]=1
[DTT]=2
[DTT]=3
Relative fluorescence level
Time (hr)


HAC1
[DTT]=0mM
[DTT]=0.5
[DTT]=1
[DTT]=2
[DTT]=3
Relative fluorescence level
Time (hr)


KAR2
[DTT]=0mM
[DTT]=0.5
[DTT]=1
[DTT]=2
[DTT]=3
Reletive fluorescence level
Time (hr)


PDI1
[DTT]=0mM
[DTT]=0.5
[DTT]=1
[DTT]=2
[DTT]=3
Relative fluorescence level
Time (hr)

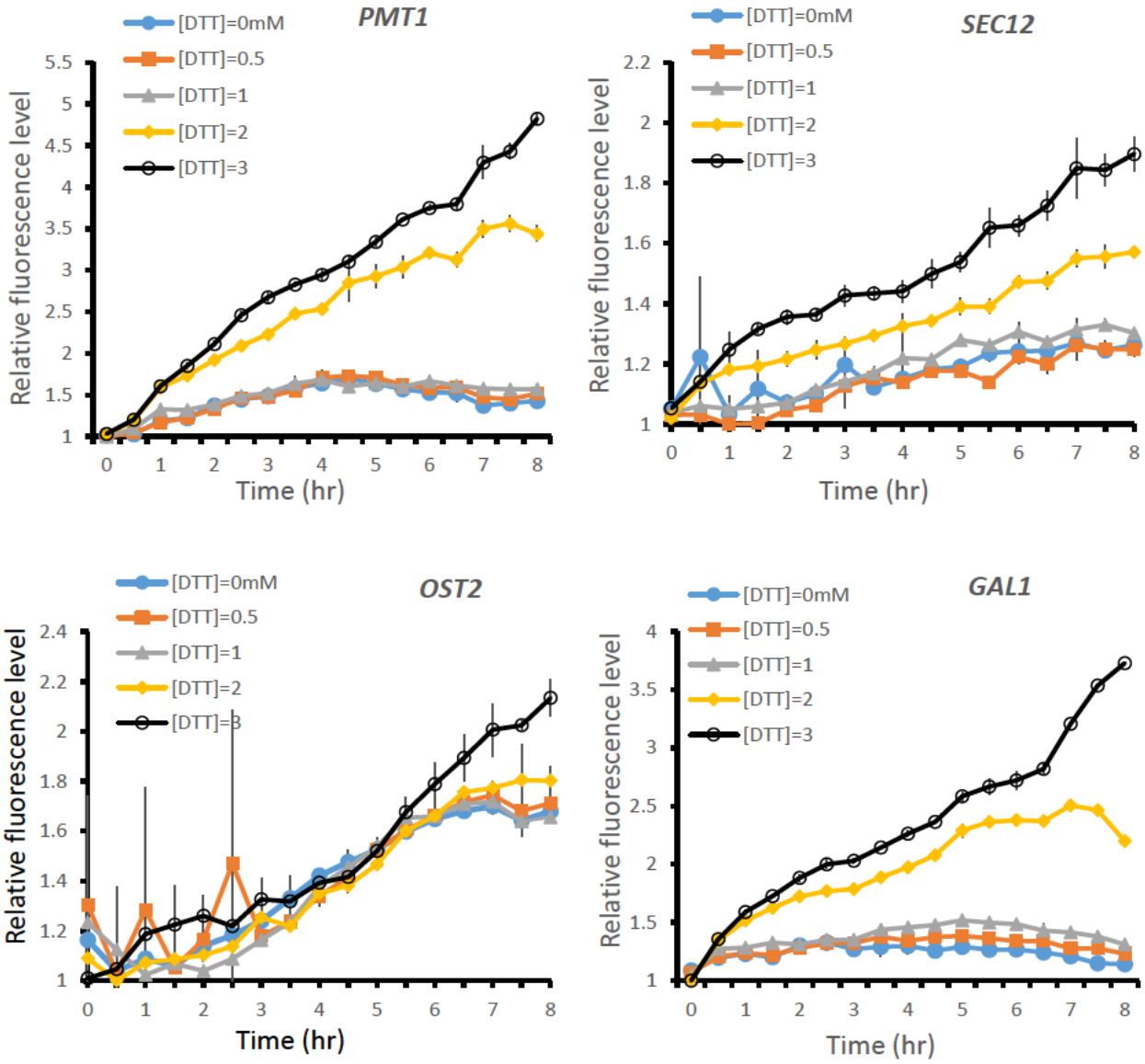


Figure 8. The relative fluorescence signal levels of the UPR sensors over 8 hours. The fluorescence levels were recorded every half an hour since induction. For each sensor, the arbitrary fluorescence levels were normalised against the lowest recorded fluorescence value by the same sensor. Samples were incubated at 30°C and shaking at 200 rpm. Three biological replicates were sampled. Error bars represent the standard deviations.

The similar rise-and-drop pattern in signal levels were also observed at 3 mM DTT with *ERO1, KAR2* and *PDI1* sensors between 5-7 hours since induction; while those of *DER1*, *EUG1*, *PMT1* and *SEC12* showed no such tendency (Fig. 8). Nevertheless, the fluorescence levels at 3 mM DTT still remained the highest for all the evaluated DTT concentrations, except for the *HAC1* sensor (Fig. 8).

The *HAC1* sensor displayed no significant differences in signal levels between 2 mM and 3 mM DTT induction (Fig. 8), and was unique in showing an 'on/off' switch like behaviour in response to ER

stress. This switch-like behaviour of the *HAC1* sensor and its undesired upper detection limit did not fit well with our criteria for a good UPR sensor. The *HAC1* sensor was subsequently excluded from the final list of UPR sensor candidates.

Interestingly the *GAL1* control sensor responded to 2 mM and 3 mM DTT by issuing elevated fluorescence levels above those with 0-1 mM DTT (Fig. 8). This promoter was chosen to provide a reference to the fluorescence signal levels, since no UPRE elements were found and were not expected to respond to ER stress. Nonetheless, the GFP expression from this sensor remained low throughout the evaluated induction times, even at relatively high DTT concentrations (Fig. 7).

3.1.2 Sensor evaluation using tunicamycin (Tm) as the UPR inducer

After obtaining the preliminary knowledge on the sensors' response to DTT induced stress, we decided to determine the compatibility and responsiveness of these sensors to another UPR inducer, tunicamycin (Tm). Tm induces UPR via a different mechanism from DTT by inhibiting the *N*-linked glycosylation in the ER thus causing the accumulation of unfolded proteins, and only affects newly synthesized proteins; while DTT affects both new and the existing proteins. We designed a Tm concentration gradient of 0, 0.2, 0.5, 1.0 and 2.0 µg/mL and performed the GFP fluorescence determinations similar to the DTT test. From the DTT tests, we achieved sufficient resolution of the signals by 4 hours, so the incubation time was reduced to 4 hours after induction with Tm.

Examining the Tm fluorescence results, it was observed that the *EUG1* sensor had a reduction in fluorescence at Tm concentrations below 0.5 µg/mL, though the sensor still responded to 1 and 2 µg/mL Tm through increased fluorescence levels (Fig. 9). This unexpected reduction in signal levels when induced by low concentration of Tm, excluded the *EUG1* sensor from our potential candidates for the UPR sensors.

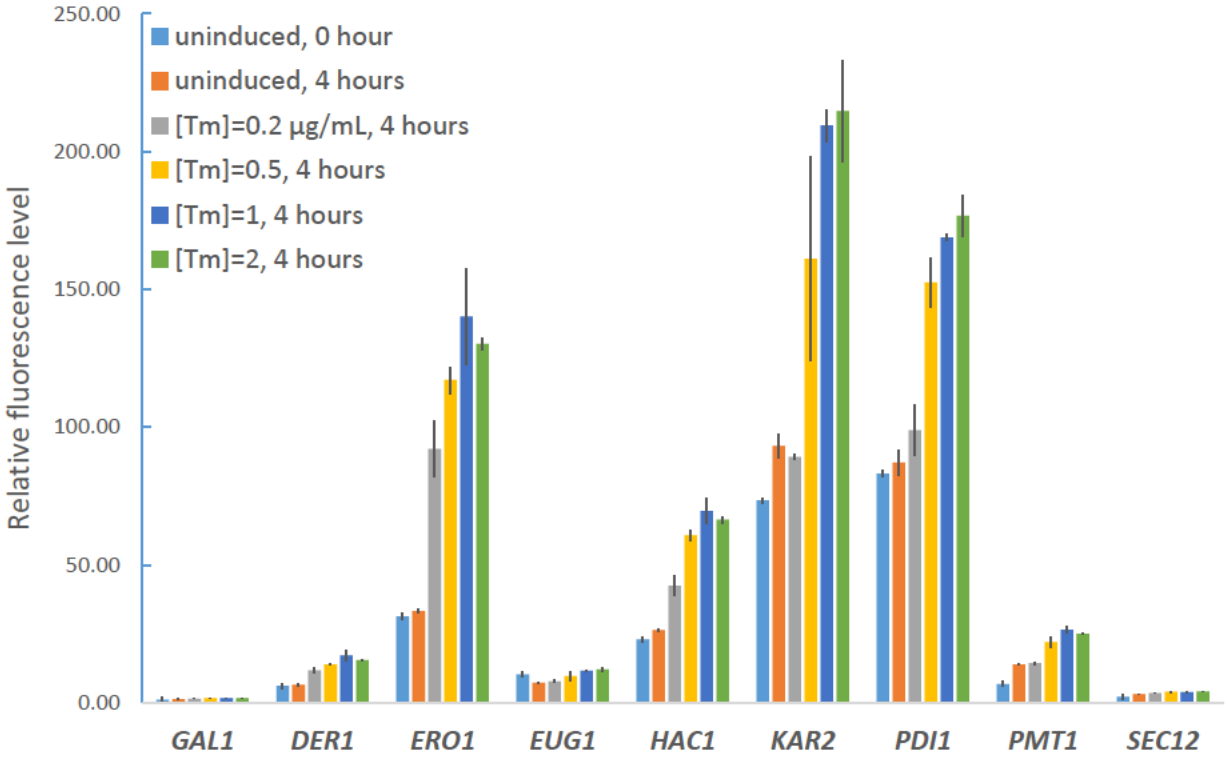


Figure 9. The relative fluorescence signal levels of the UPR sensors induced by Tm. For each sensor, the uninduced basal signal levels at 0 and 4 hours, and the signals induced by 0.2, 0.5, 1 and 2 µg/mL Tm after 4 hours were displayed. The arbitrary fluorescence levels were normalised against the lowest recorded fluorescence value across all the sensors (uninduced *GAL1*, 0 hours). Samples were incubated at 30°C and shaking at 200 rpm. Three biological replicates were sampled. Error bars represent the standard deviation from the mean value.

After considering the lowest concentrations of DTT and Tm that could be resolved with each sensor, we decided to proceed with the fine gradient UPR inductions using the *DER1* and *ERO1* sensors. Both sensors were able to resolve ER stress induced by Tm concentrations as low as 0.2 µg/mL and had the most gradual response (best linear response range) to the increase of chemical UPR inducers (Fig. 9). The *DER1* sensor was chosen as a low fluorescence sensor and the *ERO1* as a high fluorescence sensor.

3.2 Dynamic and linear range determination using a fine Tm gradient

After deciding on *DER1* and *ERO1* sensors as our final candidates with favourable biosensor characteristics, we evaluated their responses against a fine gradient of Tm: 0, 0.05, 0.10, 0.25, 0.50, 0.75, 1.00, 1.25, 1.50 and 1.75 µg/mL, in order to investigate their resolution power to separate ER stress levels close to each other. From the preliminary Tm tests, we already

discovered that fluorescent levels at 2 µg/mL Tm were indistinguishable from 1 µg/mL Tm samples, resulting in a non-linear response. So we reduced the maximal Tm concentrations to 1.75 µg/mL to better probe the linear response range the two sensors.

We examined the fluorescent signal responses of the *DER1* and *ERO1* sensors across the Tm concentration gradients at 4 hours since UPR induction, to determine the dynamic ranges of both sensors (Fig. 10). At Tm concentrations below 0.1 µg/mL, the *DER1* sensor levels did not show significantly different signals from the uninduced isogenic strain ($p>0.296$). The *ERO1* sensor was more sensitive within this range of Tm and had a higher fluorescence signal at 0.1 µg/mL than 0.05 µg/mL Tm ($p=0.026$), but could not separate 0.05 µg/mL Tm from the uninduced with confidence ($p=0.494$).

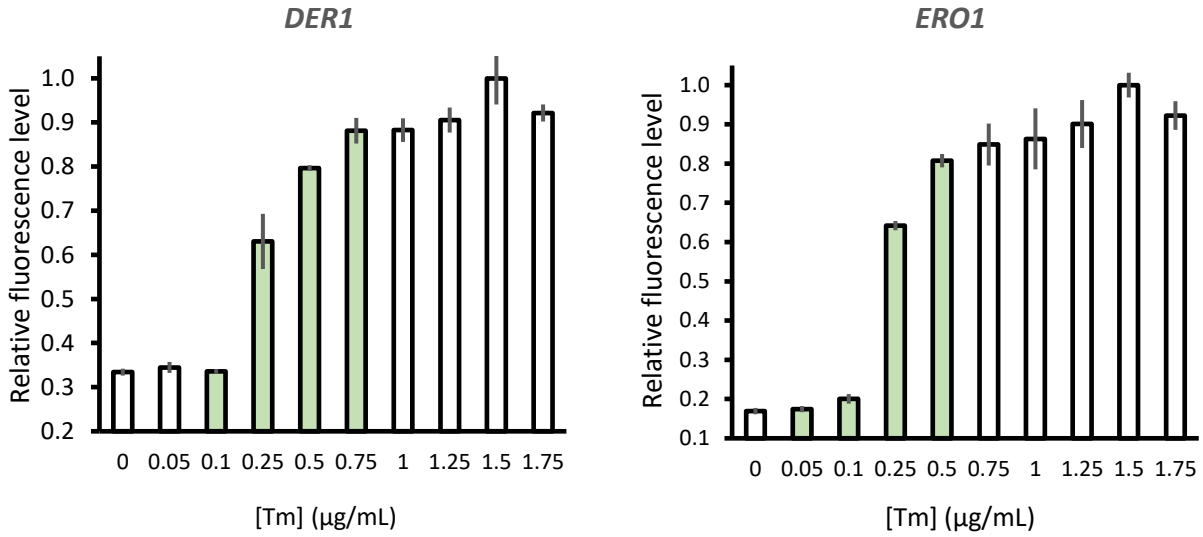


Figure 10. The relative fluorescence signal levels of the *DER1* and *ERO1* UPR sensors after 4 hours since UPR induction. UPR was induced by a Tm concentrations gradient of 0.05, 0.1, 0.25, 0.5, 0.75, 1, 1.25, 1.5 and 1.75 µg/mL Tm. The arbitrary fluorescent levels were normalised against the highest recorded fluorescence value by the same sensor. The dynamic range of each sensor was highlighted in green. Samples were incubated at 30°C and shaking at 200 rpm. Three biological replicates were sampled. Error bars represent standard deviation.

The *DER1* sensor had a linear response range between 0.10, 0.25, 0.50 and 0.75 µg/mL Tm ($p=0.008$, 0.022, 0.019) by issuing higher signals stepwise along the Tm gradient (Fig. 10 *DER1*). The *ERO1* sensor had a similar linear response range, but was sensitive to lower Tm concentrations of 0.05, 0.10, 0.25 and 0.50 µg/mL ($p=0.026$, 0.001, 0.001, Fig. 10 *ERO1*). A plateau in the signal was observed between 0.75 and 1.25 µg/mL Tm for the *DER1* sensor and between 0.5 and 1.25 µg/mL Tm for the *ERO1* sensor. This irresponsive plateau slowly led to the peaks of signals at 1.5

μg/mL Tm that had a significant increase from the signal levels at 0.75 μg/mL ($p<0.05$). Similar to the previous Tm test results, when the Tm concentration reached the maximal along the concentration series of 1.75 μg/mL, the signal levels of the *ERO1* sensor dropped again from that at 1.5 μg/mL Tm ($p=0.034$). However, the *ERO1* sensor did not experience significant signal drop from 1.5 μg/mL to 1.75 μg/mL Tm ($p=0.081$).

We determined the preliminary dynamic range of response was slightly better for the *ERO1* sensor with a shift greater than 60% of the total fluorescence between the start and end of the response (Fig. 10). *DER1* had a dynamic range of 55% of its total maximum fluorescence (Fig. 10).

3.3 Biosensor test in *HAC1* deleted strain

DTT and Tm might also induce other cellular stress or responses other than the intended the UPR. Due to the limited knowledge about the structures of the *DER1* and *ERO1* promoters, it was unclear if the putative promoters we used contained regulatory elements that could respond to non-UPR stimuli. To this end, we evaluated the two sensors in the CEN.PK *Δhac1* strains, where the *IRE1/HAC1* dependent UPR signalling pathway had been disconnected, compared to the native CEN.PK strains.

The first noticed phenomenon was when unstressed the sensors' fluorescence levels in the *Δhac1* strain after 4 hours became significantly higher than in the wild type (Fig. 11): 2.2 folds as much reported by the *DER1* sensor ($p=1x10^{-5}$) and 1.9 folds by *ERO1* sensor ($p=1x10^{-3}$).

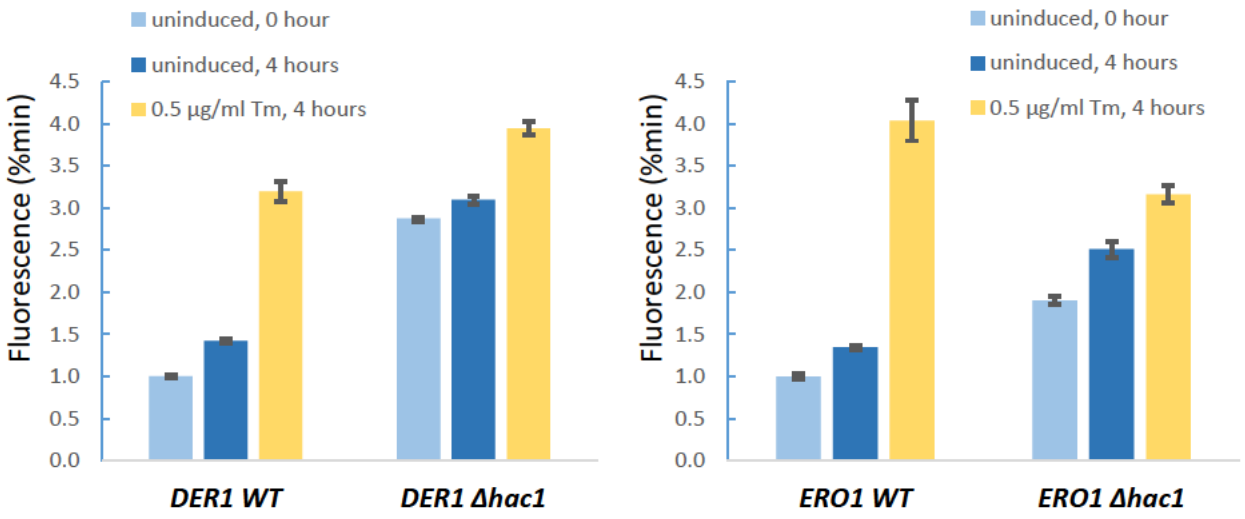


Figure 11. The relative fluorescence signal levels of the *DER1* and *ERO1* sensors in *Δhac1* strains. For each sensor, the uninduced basal signal levels at 0 and 4 hours, and the signals induced by 0.5 µg/mL Tm after 4 hours were displayed. The arbitrary fluorescence levels were normalised against the lowest recorded fluorescence value by the same sensor (uninduced *DER1* and *ERO1*, 0 hours). Samples were incubated at 30°C and shaking at 200 rpm. Three biological replicates were sampled. Error bars represent standard deviation.

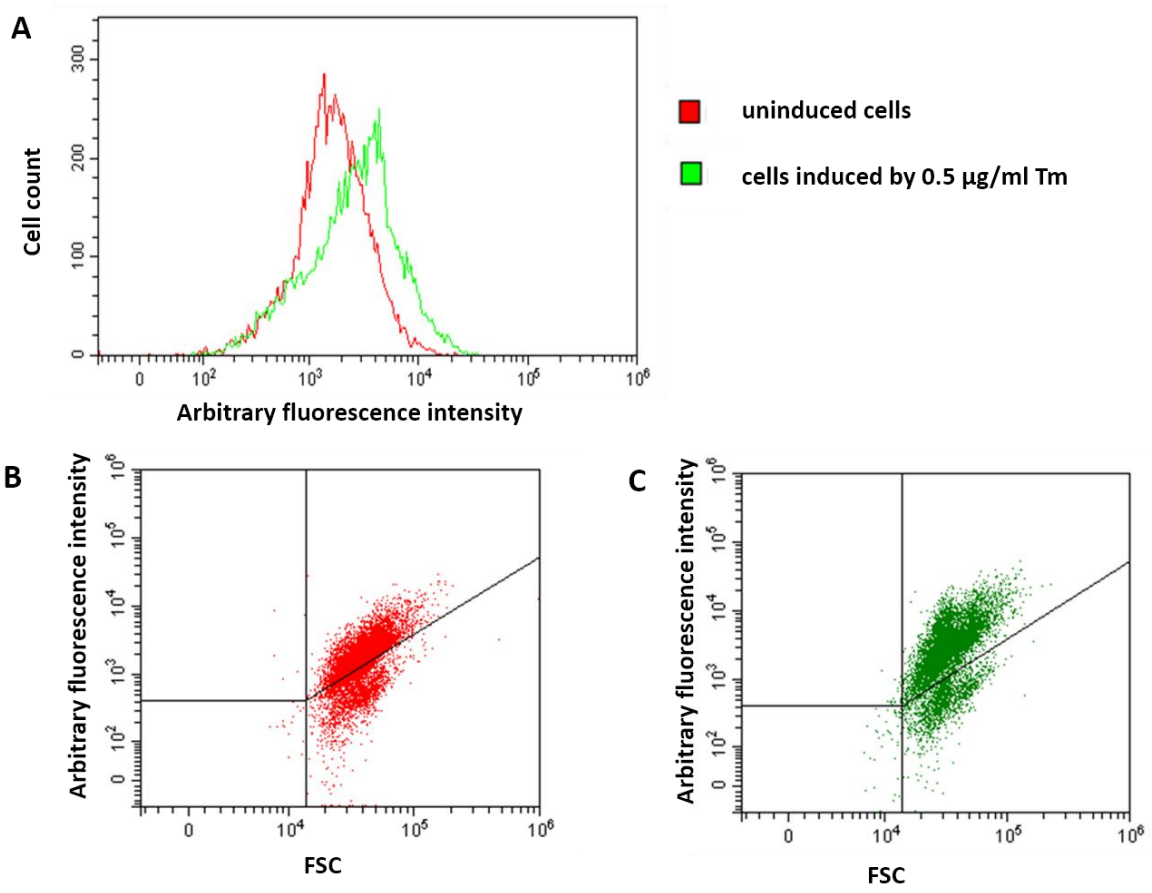
We then compared the magnitudes of signal increments between CEN.PK and CEN.PK *Δhac1* strains when induced with 0.5 µg/mL Tm, in order to determine if the responses were actually caused by UPR. In the Tm stressed wild type CEN.PK cells with the *DER1* sensor, the signals levels increased to 3.2 times after 4 hours. Within the same timeframe, the induced signal amplification dropped to 1.4 times in the *Δhac1* strain ($p=1x10^{-4}$), demonstrating a major loss of UPR sensor sensitivity (Fig. 11). The similar pattern was confirmed by the *ERO1* sensor as well (Fig. 11), where the signal amplification decreased to 1.7 folds in the *Δhac1* strain from 4.0 folds in wild type ($p=9.5x10^{-4}$). Hereby, we confirmed the *DER1* and *ERO1* sensors' responses to UPR inducers were mainly due to the UPR pathway activity. One point worth noticing was, although the response magnitudes of the *DER1* and *ERO1* sensors to Tm had been significantly lowered by *HAC1* deletion, they were still higher than the rise in the basal signal levels in the unstressed *Δhac1* cells over the same period (Fig. 11), which were 1.1 folds for the *DER1* sensor ($p=1.4x10^{-4}$) and 1.3 folds for the *ERO1* sensor (p=0.006).

Another interesting finding was in the induced *Δhac1* strains, the *DER1* sensor reported a higher level signal than in the induced wild type strain ($p=1.4x10^{-3}$, Fig. 11), while the *ERO1* sensor issued a lower signal (p=0.014, Fig. 11).

3.4 Test of chromosome integrated sensors

While we preformed the previous tests, we also noticed that except the *OST2* sensor and the *GAL1* control sensor, cells harbouring the rest of the UPR sensors displayed discrete population distributions on the fluorescence level-forward scatter (FITC-FSC) plots from the flow cytometry, as shown in Fig. 12 B and C. Because the distribution of FSC was continuous, we reasonably suspected the discrete distribution was caused by discontinuity in fluorescence levels, i.e. GFP levels.

One possible cause was the pRS416 plasmids used to carry the UPR sensor cassette. Though as a centromeric plasmid, the pRS416-UPR promoter-GFP is usually maintained as a single copy per yeast cell, we hypothesised that there existed a sub-population of cells with varying copy numbers of pRS416, e.g. two copies or loss of plasmids; or alternatively, the promoters had a bimodal expression characteristic. To test the above theories, we decided to rule out the possible instability associated with the pRS416 plasmids. We thus performed chromosomal integration of the *DER1* and *ERO1* sensor sequences in to the CEN.PK2-1C strain. The fluorescence profile of the cells with plasmid based and chromosome integrated *DER1* and *ERO1* sensors were subsequently investigated.



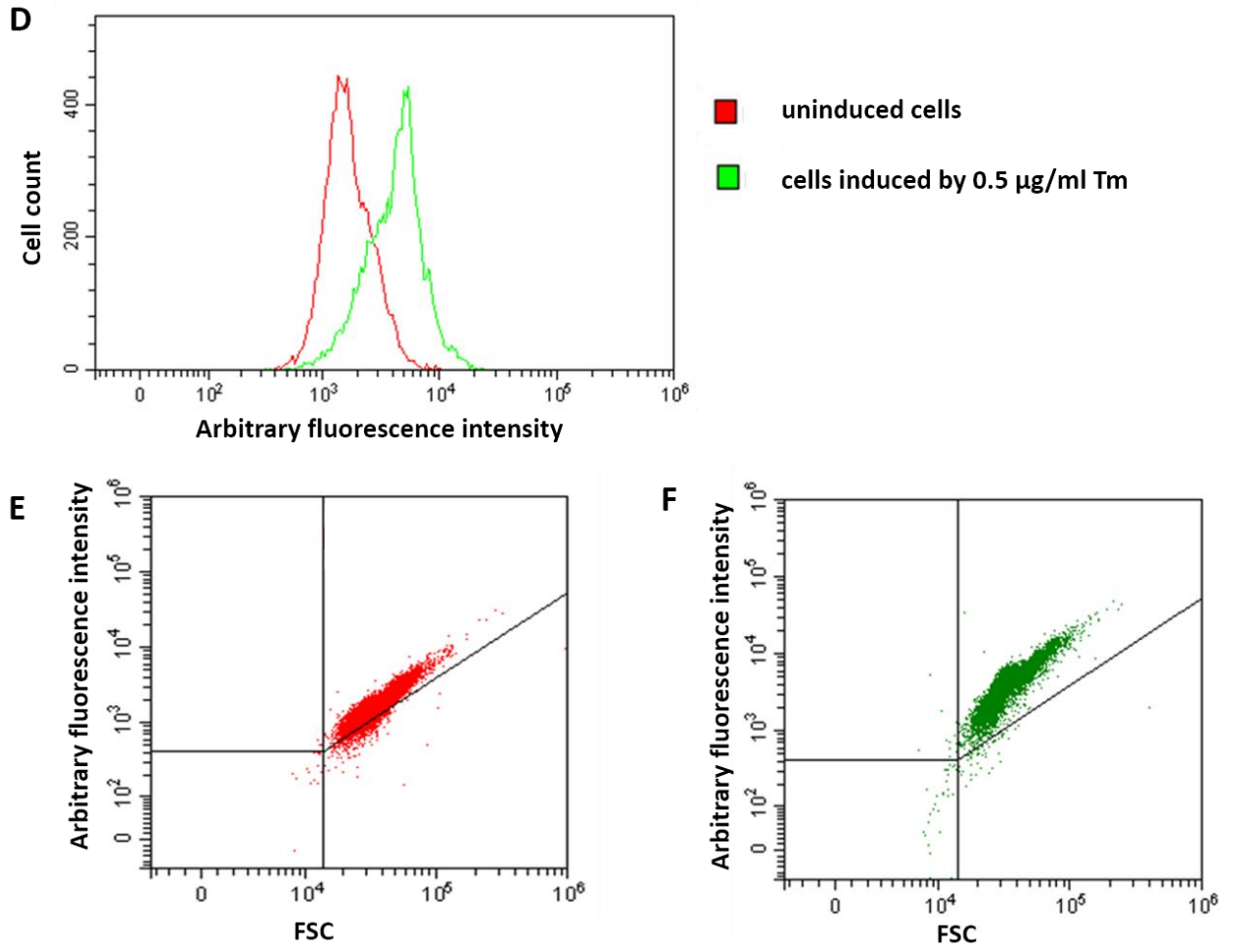


Figure 12. Distribution of the cells carrying the plasmid based *DER1* sensor (A-C) and chromosome integrated *DER1* sensor (D-F). UPR was induced by 0.5 µg/mL Tm. Samples were incubated at 30°C and shaking at 200 rpm over 4 hours. Ten thousand events were displayed. **A)** and **D)** histograms of fluorescence distributions of the uninduced cells and induced cells. **B)** and **E)** FITC-FSC plots of uninduced cells. **C)** and **F)** FITC-FSC plots of induced cells.

As previously seen, the cells with the plasmid based *DER1* sensor displayed two subpopulations with discrete fluorescence levels under both unstressed and induced conditions (Fig. 12 B, C). These subpopulations with overall lower fluorescence levels represented either a non-fluorescent population or background noise, considering the only two sensors that did not have this phenomenon (the *OST2* and *GAL1* sensors, Fig. 7) also had the lowest fluorescent levels. However, such low-fluorescent subpopulations disappeared in strains with the chromosome integrated version of the *DER1* sensor, and the cells formed a single uniform population (Fig. 12 E, F). The chromosome integration also reduced the deviation in population fluorescence, thus the induced cells could be better separated from the unstressed, illustrated by the sharper peaks on the FITC histograms (Fig. 12 A, D).

However for the cells with the integrated *ERO1* sensor, the population on the FITC-FSC plot had enlarged variations in fluorescence levels compared with the population carrying the plasmid-based sensors (Fig. S1). In this case, the chromosome integration actually amplified the noise and made the population fluorescence more dispersed. This might be caused by a problem that occurred during the construction of this strain, which was not apparent during the confirmation procedures and will be a subject for further investigation.

4. Discussion

4.1 GFP signal reduction

As shown in the DTT tests, a reduction in signal levels were observed with the *DER1*, *ERO1*, *EUG1*, *KAR2* and *PDI1* sensors at 2mM DTT and *ERO1, KAR2* and *PDI1* sensors at 3mM DTT. Depending on the GFP turnover rate, the fluorescence curves in Fig. 8 depicted the net accumulative effect of GFP overtime rather than the real-time expression, so the fluorescence levels were not supposed to reduce without GFP being removed. GFP has a half-life of about 8 hours (Pincus et al., 2010), making the *in vivo* auto-degradation of the GFP an unlikely explanation for this reduction. The GFP reporter used in this study did not have secretion or ER retention signal sequences, thus remained in cytosol after synthesis. Therefore, it was possible the yeast cytosolic degradation mechanism had been activated by the production of GFP to handle the unwanted proteins in the cytosol. There was thus a competition between the GFP degradation rate and its production. The accumulative fluorescence levels would drop when the degradation of GFP overtook its production. The initial rise in GFP signal before the drop occurred might also have suggested the cytosolic degradation gained activity later than GFP accumulation had started.

Several observations supported the GFP cytosolic degradation hypothesis: for the same sensor e.g. the *KAR2* sensor (Fig. 8), the reduction in GFP signals occurred later when facing 3mM DTT (from 8 hours, $p=0.048$) than at 2mM DTT (from 5.5 hours, $p=0.005$); as the higher the GFP production level, the longer it took for the degradation to catch up with the production. To the contrary, sensors with relatively low levels of GFP production like the *DER1* sensor (Fig. 7), also displayed signal reduction at 2mM DTT after 5.5 hours (Fig. 8). This might suggest that the activation of ERAD is influencing cytosolic protein degradation as well, without the obvious accumulation of undesirable cytosolic proteins.

It is unclear if the Tm treated cells also displayed this phenomenon of signal reduction, as the incubation was stopped at 4 hours. Following-up studies would provide insight whether these reduction in fluorescence levels is aperiodic or cyclic of nature.

4.2 Variations in the basal signal levels

As shown in Fig. 7 and Fig. 9, the putative UPR promoters used in this study maintained basal level expression without the UPR induction. This was in line with the fact that the UPR genes

maintained basal level expression in order to provide protein processing and quality control functions under unstressed conditions. The level of basal fluorescence varied in orders of magnitude between the sensors. The *ERO1*, *KAR2* and *PDI1* sensors had high basal fluorescence, at least 40 times higher than that of the *GAL1* control sensor at 4 hours; while the *DER1*, *EUG1* and *SEC12* sensors had lower levels of basal florescence, which were only about 6 times that of the *GAL1* control sensor (Fig. 7). This reflected the varied activity levels of the respective UPR genes in their native cellular roles. Kar2p has a high constitutive expression as an essential ER chaperone. Ero1p and Pdi1p are vital to the formation and isomerization of peptide disulphide bond, thus also maintained a high constitutive expression. Der1p as an ER membrane exporter of misfolded proteins in the ERAD pathway, kept relatively low level of expression, due to the limited presence of the terminally misfolded proteins when unstressed (Valkonen et al., 2003). Upon UPR induction, the sensors with high basal fluorescence generally had much greater induced signal levels compared to the sensors with lower basal fluorescence (Fig. 7).

The high and low fluorescence levels may have influences on the dynamic ranges of the UPR sensors. In this study, the *DER1* sensor had an overall lower signal levels than the *ERO1* sensor. Meanwhile, both its upper and lower detection limits were above those of the *ERO1* sensor, i.e. having an upper shift of the dynamic range: 0.1 to 0.75 µg/mL Tm by the *DER1* sensor versus 0.05 to 0.50 µg/mL Tm by the *ERO1* sensor. Therefore, it was speculated that a UPR sensor with relatively lower signal levels was better at resolving higher levels of ER stress and vice versa.

The basal signal levels also had different patterns of fluctuation over time. As mentioned in the DTT test, those of the *ERO1* and *HAC1* sensors remained relatively stable over 8 hours, while the *DER1*, *KAR2*, *PDI1*, *PMT1* and *SEC12* sensors had the basal signals increased over time. One possibility was the activities of some putative UPR promoters were somehow growth dependent thus might have increasing expression overtime. As the growth and cell division necessitated a higher flux of proteins through the secretion pathway for the cell expansion over time, requiring increased protein folding capacity. The *EUG1* sensor was the only one showing reduced basal signal levels over time, e.g. in the DTT test since 7 hours the basal signal levels had significantly dropped below that at 4 hours ($p<0.025$, Fig. 8). The drop was more apparent in the Tm test (Fig. 9). A convincing explanation had been elusive, as it seemed the putative *EUG1* promoter reduced its transcriptional activity over time, while at the same DTT or Tm concentrations the growth rates were still similar to those of other sensors (Fig. 6).

4.3 Biosensor response to *HAC1* deletion

In the *Δhac1* strains, both the *DER1* and *ERO1* sensors had elevated basal signal levels compared with the wild type strains (Fig. 11). This supported the observations of other studies that in the absence of UPR, the ERAD servers as a dynamic substitute for the removal of misfolded proteins in unchallenged environment (Travers et al., 2000). Besides being a UPR target gene, *DER1* is also involved in ERAD pathway; thus the increased activity of ERAD in *Δhac1* strain might have been a contributing factor in the up-regulated response of the *DER1* sensor independent of Hac1p signalling. On the other hand, *ERO1* as a gene responsible for disulphide bond formation has not been reported to be linked to ERAD. This fact was also displayed in our results, with a drop in Tm treated *ERO1* sensor fluorescence in the absence of Hac1p. However, this did not abolish the *ERO1* induction completely. The increased workload of protein processing in the ER after losing UPR as a quality control possibly up-regulated *ERO1* via another signalling pathway, without the involvement of UPR. This finding supported our previous speculation in the *Δhac1* strains tests that the putative promoters of *DER1* and *ERO1* contain non-UPR regulatory elements, which could up-regulate *DER1* and *ERO1* upon other cellular events. In this study we had shown that the elevated non-UPR stress response in the *Δhac1* strain was more effective in inducing the *DER1* sensor than the *ERO1* sensor, which was also in line with the fact that *DER1* is an ERAD gene.

4.4 Biosensor response to non-UPR stress

The previous comparative evaluation of the *DER1* and *ERO1* sensors in wild type and Δ*hac1* deletion strains provided informative insight into the responses of the UPR sensors. After *HAC1* knockout, although the magnitude of the induced response of the *DER1* sensor was significantly lowered to 1.4 folds increase over 4 hours comparing with the 3.2 folds increase in wild type, this response was still significantly above the 1.1 folds increase of the basal signal levels in the *Δhac1* strain. Such residual response above the basal level was also found with the *ERO1* sensor in the *Δhac1* strain (Fig. 11). Therefore, the *DER1* and *ERO1* sensors seemed to be induced by other non-UPR cellular stress or events as well, which could be the residual effects caused by Tm.

Supporting study on this issue was found, where Patil et al. (2004) also reported the *ERO1* promoter had residual up-regulation in the wild type yeast when induced with 2mM DTT after the removal of its UPRE-2. However, the residual up-regulation was not observed in the *Δhac1* strain with UPRE-2 intact *ERO1* promoter, in contrast to what we saw. Patil et al.'s result did not exclude

the possibility that the cryptic regulation element responsible for the residual up-regulation in his study was another UPRE hiding in the *ERO1* promoter. From the results obtained in our study, we could not exclude this possibility either; however, it is likely that this Hac1p independent activation is not just limited to *ERO1*. This might also be linked to the inducible response of the *GAL1* sensor (Fig. 8 and Fig. 9), where *GAL1* had not been reported as a UPR target gene. Non-UPR regulatory elements may also exist in the promoters of other UPR genes. For example, *KAR2* promoter contains heat shock element (Fig. 4A) and might thus exhibit non-UPR regulation when stressed by DTT or Tm, if the two UPR chemical inducers also cause heat shock-like stress.

4.5 The switch-like response of the *HAC1* sensor

Hac1p is also the transcription activator of itself (Ogawa et al., 2004), thus once Hac1p is produced upon UPR induction a positive feedback loop is formed to rapidly amplify its own transcription. Our DTT test showed the *HAC1* sensor was insensitive to the range of ER stress levels equivalent to 2-3mM DTT (Fig. 8). One possible explanation was that the transcription from the *HAC1* putative promoter could be saturated at moderate ER stress (e.g. 2-3mM DTT) due to the positive feedback of its transcription. Together with its half-life of only 2 minutes (Chapman and Walter 1997), the abundance of Hac1p reached maximum level and could not increase further despite any increase in the UPR inducer concentration. In this case, the *HAC1* promoter would behave like an 'on/off switch' and the *HAC1* sensor would not respond to ER stress once the DTT or Tm level reached the saturation threshold.

Another explanation was a step-wise response, where the *HAC1* promoter responded to discrete ranges of ER stress instead of having a gradual response, i.e. our *HAC1* sensor may still respond further provided the DTT concentration was high enough. In our Tm test, the *HAC1* sensor signal levels were similar between 1 and 2 µg/mL Tm by 4 hours (Fig. 9); and these signal level were not higher than those at 2mM or 3mM DTT (Fig. 7). Together with the DTT test, it seemed the 'on/off switch-like' response was more plausible for the *HAC1* sensor. However, the irresponsiveness of the *HAC1* sensor between 1-2 µg/mL Tm might be a result of the more severe cellular stress caused by the high levels of Tm comparing with treatment by 2-3mM DTT, rather than reflecting a real 'on/off switch' behaviour.

The possibility of the 'on/off switch' behaviour by the *HAC1* promoter raised another interesting thought. If the splicing of the *HAC1* mRNA or its translation is saturable, any UPR sensor relying on

the UPR gene transcription mechanism might have its maximal signal level capped by the saturated supply of Hac1p. This suggests that once the UPR inducer level is enough to saturate the supply of Hac1p, the UPR transcription sensor signal would be insensitive to the further increases in the UPR inducer levels. Also at this stage, the strength of UPR matched the magnitude of ER stress by the duration of UPR (Pincus et al., 2010), instead of the Hac1p levels. If this argument is true, the levels of ER stress that were inseparable at certain time points by the UPR sensors, might eventually become separable given longer time. This will be investigated in our follow-up studies.

4.6 Chromosome integration of UPR sensor

During the initial sensor evaluations, several distinct fluorescent cell populations were observed. Although the average fluorescence values corresponded with the level of induction in most cases, having distinct differential responses in individual cells was an undesirable feature. It was unclear if these populations were generated by bimodal expression patterns of the sensor, or due to the inconsistency in the copy number of the 'single-copy' pRS416 centromeric plasmids.

As an example, the *DER1* sensor also displayed distinct fluorescent population distribution (Fig. 12). We eliminated the sensor copy number variability by integrating it as single copy into the yeast genome, and repeated the flow cytometry analysis. The FITC-FSC plot of the plasmid based *DER1* sensor (Fig. 12) displayed discrete populations as previously observed, with about 10% cells residing within a subpopulation with a low fluorescence profile. This subpopulation potentially represented the cells that had lost the pRS416 plasmids, but still persisted in the population without actively growing. Over the 4 hours incubation, the size of this non-fluorescent subpopulation remained relatively stable occupying about 10% of the total population, while the overall cell density kept increasing. With the chromosomal integration of the sensor, we effectively increased the stability of the *DER1* sensor. In both uninduced and induced states, the chromosomal integrated *DER1* sensor displayed uniform fluorescence profile, eliminated the non-fluorescent subpopulation and increased the sensitivity of the sensor (Fig. 12). These results suggested that the distinct fluorescent populations was caused by variable plasmid copy numbers, thus sensor integration is an appropriate strategy to optimise its sensitivity.

4.7 The concentrations of the UPR chemical inducers

In the UPR related studies so far, DTT and Tm have been used as the standard chemical inducers of UPR. A translation between the concentrations of DTT and Tm would be beneficial for future study on UPR, as well as for UPR sensor assessment to have a better control on the inducer dosages. Although there has been no study focused on this issue, from the published work (Table 4) and this one (Fig. 7 and Fig. 9) we could still manage to make a preliminary comparison.

DTT concentrations below 1mM and Tm concentrations below 0.1 μg/mL induce low level ER stress. At such stress levels, generally no obvious grow inhibition were observed (Fig. 6). UPR sensor tests within such ranges were usually difficult to distinguish between the low inducer levels, therefore such low DTT or Tm concentrations were not often used by researchers wishing to ensure a differentiated responses from UPR sensors. Accordingly, we would like to recommend a parallel assay of *HAC1* mRNA, or incorporation of the *HAC1* mRNA splicing reporter when applying low level inducers, to ensure the UPR is actually activated.

Moderate ER stress can be induced by 2-3 mM DTT or 0.2-0.75 μg/mL Tm. Within this range, the cells can still overcome or adapt the ER stress and resume growth overtime, with decreased growth rates. DTT concentration above 4-5mM or Tm concentration over 1 μg/mL represent high stress levels, and the cells could hardly proliferate with such stress. At very high UPR inducer levels, the UPR sensors signal may also reach saturation. The cell growth could not recover as UPR could not compensate the high ER stress anymore.

When the cells were facing moderate or high UPR stress, the growth curves may no longer be good references for the stress levels, as under such conditions the relationship between the growth inhibition effects and the UPR levels may not be linear. For example, in our study the cells carrying the *ERO1* sensor at 3 mM DTT had similar cell density to that at 1.5 μg/mL Tm ($p=0.135$), but the sensor signal levels were much lower at 3mM DTT ($p=2.7\times10^{-4}$). Besides that, high concentrations of UPR inducers add another complication to UPR sensor assessment through the activation of non-UPR specific stress responses (Pincus et al., 2014; Navarro-Tapia et al., 2017), which were also discussed in our previous section. For example, treatment by 8mM DTT was known to upregulate the non-UPR genes like *ATG5*, *ATG7* and *ATG8* in the *Δhac1* or *ΔIRE1* strains (Bernales et al., 2006) that are responsible for autophagy of the overabundant endomembrane system. This could probably affect the ER environment sensors like the eroGFP and Kar2p-sfGFP sensors. Severe ER stress could also impair the cellular functions essential to the UPR sensors, like

transcription, translation and protein degradation. The ability of the UPR sensors to distinguish between high UPR stress and other stress responses could thus be a consideration of the future sensor optimization, as well.

5. Conclusion and future development

This is the first study dedicated to the engineering and assessment of the UPR sensors for detecting protein secretion stress. Here, we assessed nine putative UPR promoters as a first step to construct sensitive real-time biosensors for protein production stress. The *DER1* and *ERO1* sensors had the best dynamic ranges in response to Tm, and will be the focus for future engineering to increase the dynamic ranges and sensitivities to ER stress. The *DER1* sensor resolved the different ER stress levels caused by 0.1 to 0.75 μg/mL Tm, while ERO1 sensor was successful with 0.05 to 0.50 μg/mL Tm. The *HAC1* sensor was unique in exhibiting an 'on/off switch' like behaviour, where the signal level could be possibly saturated by the treatment of 2-3mM DTT and might be useful as a switching mechanism to modulate the final UPR sensing circuit.

The research up to date has typically focused on using the UPR sensors (especially the 4xUPRE-mCherry) to report the on/off behaviour of the UPR, as well as to separate high and low levels of UPR regulation, where the inducer concentrations were designed relatively far apart to guarantee differentiated responses. No study has so far demonstrated the systematic dissection of the UPR and elucidated the cross-activation through other stress response pathways. Table 4 summarized the DTT and Tm levels that could be distinguished by the previously reported UPR sensors. When designing biosensors, it is imperative to consider their resolving power at low inducer levels. This might be limited by the natural action of the UPR, as it is buffered against activation at very low levels of ER stress, through the association between Kar2p and Ire1p (Pincus et al., 2010), which is an important protective mechanism preventing a hypersensitive UPR.

There are several considerations on how our prototype UPR sensors, i.e. the *DER1* and *ERO1* sensors, can be optimized for improved resolution and dynamic ranges. The UPR (Hac1p) sensing sequence ahead the reporter protein is crucial for the design of our sensors, and optimisation of this sequence could effectively improve the sensor strength and sensitivity. The 4xUPRE-mCherry sensor used the UPRE-1 of *KAR2* as the core UPR sensing sequence, while ours relied on single copy putative promoters of *DER1* and *ERO1*. We can optimise the putative promoters by using longer promoter regions or removing the sequences responsible for the basal level expression and non-UPR responsiveness. Synthetic UPR promoters could also be made by combining different UPREs and varying the number of repeats, which might achieve better linear responses and wider dynamic ranges for the UPR sensors. Besides the sensing sequences, the GFP reporter can also be modified through the attachment of a degradation signal. The removal of the overabundant GFP would not only reduce the potential saturating effect of the UPR sensors caused by the

accumulated GFP, but also reduce the competition with the endogenous cellular functions. In addition, by removing GFP the UPR sensors can also be prepared for the next cycle of UPR detection after the previous UPR deactivation.

Table 4. The resolution power of the currently available UPR sensors. The DTT and Tm concentrations that could be distinguished by each sensors were listed.

UPR sensor	[DTT] resolution (mM)	[Tm] resolution (μg/mL)	Source
eroGFP	0, 2	0, 1	Merksamer et al., 2008
	0, 1		Navarro-Tapia et al., 2017
	0, 1, 2		Rubio et al., 2011
Kar2p-sfGFP	0, 5	0, 0.025, 1	Lajoie et al., 2012
4xUPRE-mCherry	0, 2	0, 1	Merksamer et al., 2008
	0, 1		Navarro-Tapia et al., 2016
	0, 5		Lajoie et al., 2012
	0, 1.5, 2.2		Pincus et al., 2010
Splicing reporter		0, 1	Lajoie et al., 2012
	0, 2		Anken et al., 2014
Dual-luciferase	0, 0.1-0.5	0, 0.5, 2	Fang et al., 2015
***DER1* sensor**	0-1, 2, 3	0-0.1, 0.25, 0.5, 0.75-1.25, 1.5	this study
***ERO1* sensor**	0-1, 2, 3	0-0.05, 0.1, 0.25, 0.5-1.25, 1.5	this study

To eliminate the UPR-independent responses of the *DER1* and *ERO1* sensors, we can take advantage of the highly specific splicing of the immature *HAC1* mRNA. Like the splicing reporter and dual-luciferase sensor, we can incorporate the 5'-UTR of *HAC1* mRNA upstream the *DER1/ERO1* promoter and *HAC1* intron within the GFP sequence. The long-range base-pairing similar to that in unspliced *HAC1* mRNA could eliminate the basal level signal; as this system will function as an 'AND-gate', requiring both the induction of the sensor and the UPR-dependent *IRE1*-splicing to produce the signal. Such design also appeared to have no significant competition with the processing of the endogenous *HAC1* mRNA and normal UPR signalling (Jonikas et al., 2009). An alternative strategy is to incorporate an internal control, and use the ratio between the UPR sensor signal and the control signal. The control signal can come from the same sensor construct like the dual-luciferase sensor, or from another constitutively expressed reporter like the *TEF2* promoter-RFP reporter designed by Jonikas et al. (2009), to serve as an internal control for normalising the gene expression in cells with different metabolic states.

Our UPR sensors ultimately aims to offer a high throughput and quantitative measure of the UPR during industrial fermentation, identify heterologous proteins that induce higher ER stress and provide a variable for secretion improvement via attenuating the UPR. In future we'd like to first evaluate the UPR sensors against ER stress induced by the production of heterologous proteins, e.g. cellulase to test if the sensor signals can be related to the secretion levels of heterologous proteins. Next, we plan to assess the UPR sensor against the environmental factors in industrial practice that induce UPR like heat and ethanol stress, and test if the varying UPR levels can be readily distinguished and measured. Ideally, we would like to achieve real-time monitoring of UPR, presenting UPR levels as a relevant and controllable parameter in industrial practice. Comparing with conventional protein-specific strain and expression engineering, UPR regulation offers a different approach to improve recombinant protein secretion by simultaneous regulation of a collection of genes, the products of which increase the overall ER processing and secretory pathway capacities.

Acknowledgements

We would like to thank Dr Heinrich Kroukamp, Dr Niel van Wyk and Prof Ian Paulsen for the valuable scientific discussions in this project.

References

Aragón, T., van Anken, E., Pincus, D., Serafimova, I. M., Korennykh, A. V., Rubio, C. A., & Walter, P. (2009). Messenger RNA targeting to endoplasmic reticulum stress signalling sites. *Nature*, 457(7230), 736–740.

Arvan, P., Zhao, X., Ramos-Castaneda, J., & Chang, A. (2002). Secretory pathway quality control operating in Golgi, plasmalemmal, and endosomal systems. *Traffic*, 3(11), 771–780.

Bernales, S., McDonald, K. L., & Walter, P. (2006). Autophagy counterbalances endoplasmic reticulum expansion during the unfolded protein response. *PLoS Biology*, 4(12), 2311–2324.

Chapman, R. E., & Walter, P. (1997). Translational attenuation mediated by an mRNA intron. *Current Biology* : CB, 7(11), 850–859.

Chawla, A., Chakrabarti, S., Ghosh, G., & Niwa, M. (2011). Attenuation of yeast UPR is essential for survival and is mediated by IRE1 kinase. *Journal of Cell Biology*, 193(1), 41–50.

Cohen, S. N., Chang, A. C. Y., Boyer, H. W., & Helling, R. B. (1973). Construction of biologically functional bacterial plasmids *in vitro*. *Proceedings of the National Academy of Sciences*, 70(11), 3240–3244.

Fang, Z., Kuang, X., Zhang, Y., Shi, P., & Huang, Z. (2015). A novel HAC1-based dual-luciferase reporter vector for detecting endoplasmic reticulum stress and unfolded protein response in yeast *Saccharomyces cerevisiae*. *Plasmid*, 79, 48–53.

Ferrer-Miralles, N., & Villaverde, A. (2013). Bacterial cell factories for recombinant protein production; expanding the catalogue. *Microbial Cell Factories*, 12(1), 113.

Gasser, B., Saloheimo, M., Rinas, U., Dragosits, M., Rodríguez-Carmona, E., Baumann, K., Villaverde, A. (2008). Protein folding and conformational stress in microbial cells producing recombinant proteins: a host comparative overview. *Microbial Cell Factories*, 7(1), 11.

Gietz, R. D., & Schiestl, R. H. (2007). Microtiter plate transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*, 2(1), 5–8.

Herzog, B., Popova, B., Jakobshagen, A., Shahpasandzadeh, H., & Braus, G. H. (2013). Mutual cross talk between the regulators Hac1 of the unfolded protein response and Gcn4 of the general amino acid control of *Saccharomyces cerevisiae*. *Eukaryotic Cell*, 12(8), 1142–1154.

Hou, J., Tyo, K. E. J., Liu, Z., Petranovic, D., & Nielsen, J. (2012). Metabolic engineering of recombinant protein secretion by *Saccharomyces cerevisiae*. *FEMS Yeast Research*. 12, 491-510.

Idiris, A., Tohda, H., Kumagai, H., & Takegawa, K. (2010). Engineering of protein secretion in yeast: Strategies and impact on protein production. *Applied Microbiology and Biotechnology*. 86, 403-417.

Ilmén, M., den Haan, R., Brevnova, E., McBride, J., Wiswall, E., Froehlich, A., Penttilä, M. (2011). High level secretion of cellobiohydrolases by *Saccharomyces cerevisiae*. *Biotechnology for Biofuels*, 4(1), 30.

Jonikas, M. C., Collins, S. R., Denic, V., Oh, E., Quan, E. M., Schmid, V., Schuldiner, M. (2009). Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. *Science,* 323, 1693-1697.

Kimata, Y., Ishiwata-Kimata, Y., Yamada, S., & Kohno, K. (2006). Yeast unfolded protein response pathway regulates expression of genes for anti-oxidative stress and for cell surface proteins. *Genes to Cells*, 11, 59-69.

Korennykh, A. V., Egea, P. F., Korostelev, A. A., Finer-Moore, J., Zhang, C., Shokat, K. M., Walter, P. (2009). The unfolded protein response signals through high-order assembly of Ire1. *Nature,* 457(7230), 687–693.

Kroukamp, H., den Haan, R., van Zyl, J.-H & van Zyl, W. H. (2017). Rational strain engineering interventions to enhance cellulase secretion by *Saccharomyces cerevisiae*. *Biofuels, Bioprod. Bioref.* doi:10.1002/bbb.1824

Kuhad, R. C., Gupta, R., & Singh, A. (2011). Microbial cellulases and their industrialaApplications. *Enzyme Research*, 2011, 1–10.

Lajoie, P., Moir, R. D., Willis, I. M., & Snapp, E. L. (2012). Kar2p availability defines distinct forms of endoplasmic reticulum stress in living cells. *Molecular Biology of the Cell*, 23(5), 955–964.

Lee, K. P. K., Dey, M., Neculai, D., Cao, C., Dever, T. E., & Sicheri, F. (2008). Structure of the dual enzyme Ire1 reveals the basis for catalysis and regulation in nonconventional RNA splicing. *Cell,* 132, 89-100.

Liu, Z. H., Hou, J., Martinez, J. L., Petranovic, D., & Nielsen, J. (2013). Correlation of cell growth and heterologous protein production by *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol*, 97, 8955-8962.

Löoke, M., Kristjuahan, K., & Kristjuhan, A. (2011). Extraction of genomic DNA from yeasts for PCR-based applications. *Biotechniques*, 50(5), 325–328.

Merksamer, P. I., Trusina, A., & Papa, F. R. (2008). Real-time redox measurements during endoplasmic reticulum stress reveal interlinked protein folding functions. *Cell,* 135(5), 933–947.

Martínez, J. L., Liu, L., Petranovic, D., & Nielsen, J. (2012). Pharmaceutical protein production by yeast: Towards production of human blood proteins by microbial fermentation. *Current Opinion in Biotechnology*, 23(6), 965–971.

Mattanovich, D., Gasser, B., Hohenblum, H., & Sauer, M. (2004). Stress in recombinant protein producing yeasts. Journal of Biotechnology, 113, 121–135.

Mori, K., Ogawa, N., Kawahara, T., Yanagi, H., & Yura, T. (1998). Palindrome with spacer of one nucleotide is characteristic of the cis-acting unfolded protein response element in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry*, 273(16), 9912–9920.

Mori, K., Sant, a, Kohno, K., Normington, K., Gething, M. J., & Sambrook, J. F. (1992). A 22 bp cis-acting element is necessary and sufficient for the induction of the yeast *KAR2* (BiP) gene by unfolded proteins. *The EMBO Journal*, 11(7), 2583–2593.

Mori, K. (2009). Signalling pathways in the unfolded protein response: Development from yeast to mammals. *Journal of Biochemistry*, 146(6), 743–750.

Navarro-Tapia, E., Nana, R. K., Querol, A., & Pérez-Torrado, R. (2016). Ethanol cellular defense induce unfolded protein response in yeast. *Frontiers in Microbiology*, 7, 1–12.

Navarro-Tapia, E., Prez-Torrado, R., & Querol, A. (2017). Ethanol effects involve non-canonical unfolded protein response activation in yeast cells. *Frontiers in Microbiology*, 8, 1–12.

Ogawa, N., & Mori, K. (2004). Autoregulation of the *HAC1* gene is required for sustained activation of the yeast unfolded protein response. Genes to Cells, 9, 95-104.

Otte, S., & Barlowe, C. (2004). Sorting signals can direct receptor-mediated export of soluble proteins into COPII vesicles. *Nature Cell Biology,* 6(12), 1189–1194.

Patil, C., & Walter, P. (2001). Intracellular signalling from the endoplasmic reticulum to the nucleus: The unfolded protein response in yeast and mammals. *Current Opinion in Cell Biology*. 13, 349-356.

Patil, C. K., Li, H., & Walter, P. (2004). Gcn4p and novel upstream activating sequences regulate targets of the unfolded protein response. *PLoS Biology*, 2(8), 1208-1223.

Pincus, D., Chevalier, M. W., Aragón, T., van Anken, E., Vidal, S. E., El-Samad, H., & Walter, P. (2010). BiP binding to the ER-stress sensor Ire1 tunes the homeostatic behavior of the unfolded protein response. *PLoS Biology*, 8(7), 1-14.

Promlek, T., Ishiwata-Kimata, Y., Shido, M., Sakuramoto, M., Kohno, K., & Kimata, Y. (2011). Membrane aberrancy and unfolded proteins activate the endoplasmic reticulum stress sensor Ire1 in different ways. *Molecular Biology of the Cell,* 22(18), 3520–3532.

Rakestraw, A., & Wittrup, K. D. (2006). Contrasting secretory processing of simultaneously expressed heterologous proteins in *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering*, 93(5), 896–905.

Rapoport, T. A. (2007). Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature,* 450(7170), 663–669.

Rosano, G. L., & Ceccarelli, E. A. (2014). Recombinant protein expression in *Escherichia coli*: Advances and challenges. *Frontiers in Microbiology*, 5(APR), 1–17.

Rubio, C., Pincus, D., Korennykh, A., Schuck, S., El-Samad, H., & Walter, P. (2011). Homeostatic adaptation to endoplasmic reticulum stress depends on Ire1 kinase activity. *Journal of Cell Biology*, 193(1), 171-184.

Rüegsegger, U., Leber, J. H., & Walter, P. (2001). Block of *HAC1* mRNA translation by long-range base pairing is released by cytoplasmic splicing upon induction of the unfolded protein response. *Cell*, 107(1), 103–114.

Tang, H., Bao, X., Shen, Y., Song, M., Wang, S., Wang, C., & Hou, J. (2015). Engineering protein folding and translocation improves heterologous protein secretion in Saccharomyces cerevisiae. *Biotechnology and Bioengineering*, 112(9), 1872–1882.

Travers, K. J., Patil, C. K., Wodicka, L., Lockhart, D. J., Weissman, J. S., & Walter, P. (2000). Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. *Cell*, 101(3), 249–258.

Valkonen, M., Penttilä, M., & Saloheimo, M. (2003). Effects of inactivation and constitutive expression of the unfolded-protein response pathway on protein production in the yeast *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology*, 69(4), 2065–72.

van Anken, E., Pincus, D., Coyle, S., Aragón, T., Osman, C., Lari, F., Walter, P. (2014). Specificity in endoplasmic reticulum-stress signaling in yeast entails a step-wise engagement of *HAC1* mRNA to clusters of the stress sensor Ire1. *eLife*, 3, 1-17.

Volmer, R., & Ron, D. (2015). Lipid-dependent regulation of the unfolded protein response. *Current opinion in cell biology,* 33, 67-73.

Walsh, G. (2010). Biopharmaceutical benchmarks 2010. *Nature Biotechnology*, 28(9), 917-924.

Walter, P., & Ron, D. (2011). The unfolded protein response: from stress pathway to homeostatic regulation. *Science*, 334(6059), 1081–1086.

Wang, G., Huang, M., & Nielsen, J. (2017). Exploring the potential of *Saccharomyces cerevisiae* for biopharmaceutical protein production. *Current Opinion in Biotechnology*, 48, 77-84.

Williams, T. C., Pretorius, I. S., & Paulsen, I. T. (2016). Synthetic evolution of metabolic productivity using biosensors. *Trends in Biotechnology*. 34(5), 371-381.

Young, C. L., Yuraszeck, T., & Robinson, A. S. (2011). Decreased secretion and unfolded protein response upregulation. *Methods in Enzymology*, 491, 235–260.

Zhou, M., & Schekman, R. (1999). The Engagement of Sec61p in the ER Dislocation Process. *Molecular Cell*, 4(6), 925–934.

Supplementary information

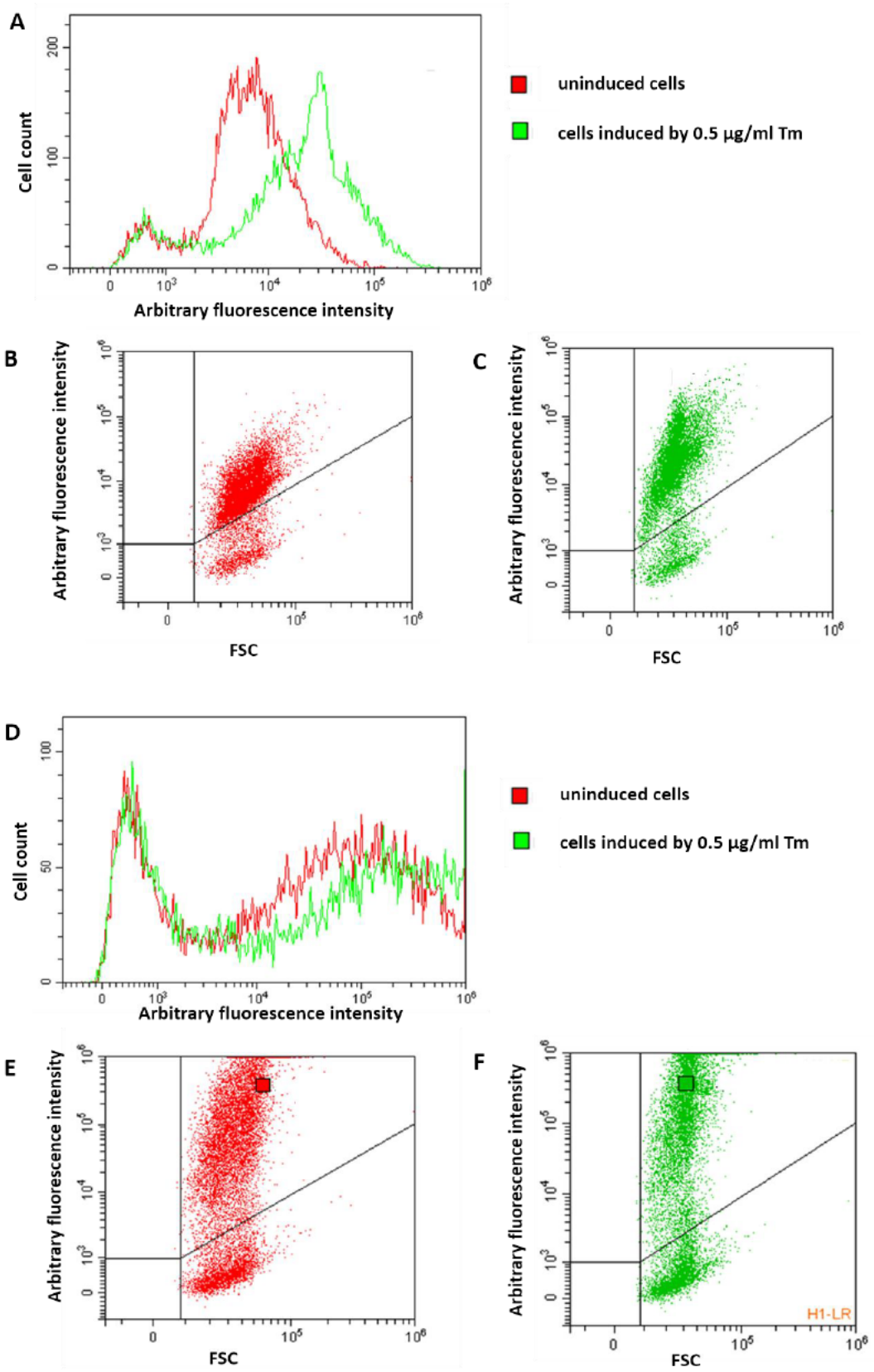


Figure S1. Distribution of the cells carrying the plasmid based *ERO1* sensor (A-C) and chromosome integrated *ERO1* sensor (D-F). UPR was induced by 0.5 µg/mL Tm. Samples were incubated at 30°C and shaking at 200 rpm over 4 hours. Ten thousand events were displayed. **A)** and **D)** histograms of fluorescence distributions of the uninduced cells and induced cells. **B)** and **E)** FITC-FSC plots of the uninduced cells. **C)** and F) FITC-FSC plots of the induced cells.