Deciphering the molecular mechanisms of Yih1/IMPACT: Identification of important Yih1 amino acids for inhibition of the General Amino Acid Control response.

A thesis presented in partial fulfilment of the requirements for the degree of Master of Science

in Biological Science

at Massey University, Albany New Zealand.

Bianca Joanne-Margo Sansom

2022

Abstract

It is crucial that eukaryotes adjust their cellular gene expression profile in response to dynamic cellular changes. Failure to do this can be detrimental to the cell. Gcn2 is a protein kinase that is central to a signalling pathway that decreases general protein synthesis and increases transcription of specific stress response genes in response to amino acid starvation in eukaryotes. Gcn2 needs to be tightly regulated to maintain cellular homeostasis. Gcn2 needs to directly bind Gcn1 to be activated. Yeast IMPACT homolog 1 (Yih1), or IMPACT (Mammals), is an inhibitor of Gcn2. It does this by competing with Gcn2 for Gcn1 binding. However, Yih1 only inhibits Gcn2 under certain circumstances. This implies that Yih1 itself must be regulated. Furthermore, this means that Yih1 resides in the cell in an inactive form or an active form that is primed for Gcn2 inhibition. The precise interactions between Yih1 and Gcn1 are not known yet, nor is it known how intramolecular interactions determine when Yih1 is active or inactive. Given the importance of IMPACT in neuronal development, it is crucial to understand the molecular mechanisms of Yih1/IMPACT-mediated Gcn2 inhibition. In this research, I aimed to identify relevant Yih1 amino acids that are involved in Gcn2 inhibition, using Baker's yeast (*S. cerevisiae*) as a model organism. At present, the only known important amino acids for Yih1 function are D102 and E106 together. I have utilised and established a yeast system that enabled us to score the ability of Yih1 to inhibit Gcn2. By overexpressing Yih1 with substitutions to predicted important amino acids, I utilised this yeast system to determine whether Gcn2 activation was affected. I found that D102 alone is crucial for Gcn1 binding. I have determined that Yih1 interdomain interaction are ionic and have provided experimental evidence of specific charged amino acids involved in these interactions. I have also shed light on the potential role of posttranslational modifications in modulating Yih1-mediated Gcn2 inhibition. This research helps to provide more insight into the regulation of Gcn2 by Yih1/IMPACT and brings us a step closer to understanding diseases associated with Gcn2 dysregulation, like Alzheimer's disease.

Acknowledgements

I would like to take this opportunity to sincerely thank everybody who has contributed to this work.

I would firstly like to thank my supervisor Associate Professor Dr. Evelyn Sattlegger for her support, knowledge, and feedback during my time under her supervision. I am grateful for the opportunity that she gave me to pursue my passion for science under her supervision.

I would also like to thank Trevor Loo for his expertise and time with Mass Spectrometry, as well as Dr. Stefan Harjes for his help with protein purification and chromatography. I would further like to thank Dr. Anja Schiemann for her assistance in lab research during my times on quick trips to Palmerston North.

A huge thank you to my fellow Sattlegger Lab colleagues. A special thank you to Susanne Gottfried, Reuben Anderson and Aditi Ghuge for their help, support, and knowledge in the lab over the course of this thesis. I have thoroughly enjoyed my time with you all and am appreciative for your guidance.

A special thanks to my parents Mike and Lynn Sansom for their support and encouragement throughout my thesis as well as outside of university. I also owe a huge thank you to my sister, Gabriela Sansom, who has encouraged and challenged me throughout my whole Master's degree. All of your support has enabled me to study and pursue what I am passionate about.

Contents

Abstract i

Acknowledgements ii

List of Tables ix

List of Figures x

List of Abbreviations xii

1.1 Introduction 1

1.1.1 The eIF2α kinase Gcn2 is involved in the Integrated Stress Response in mammals 1

1.1.2 Translation initiation in eukaryotes 1

1.1.3 GCN4 Translational control: Regulatory roles of the uORFs 3

1.2. The working model of Gcn2 activation 6

1.2.1 Alternative models of Gcn2 activation 7

1.3 The structure of Gcn2 9

1.4 Gcn1 and its direct interaction with Gcn2 is essential for Gcn2 activation 11

1.4.1 Gcn1 structure and the binding determinants for the Gcn1-Gcn2 interaction 11

1.5 Yeast IMPACT homolog 1 (Yih1)/IMPACT is negative regulator of Gcn2 when overexpressed 13

1.5.1 The Yih1 RWD domain mediates Gcn1 binding 13

1.5.2 Yih1/IMPACT has an RWD domain and an Ancient domain 14

1.5.3 Yih1-mediated Gcn2 inhibition must be regulated 17

1.6 Yih1/IMPACT-protein interactions: Towards the function of Yih1/IMPACT 18

1.6.1 Yih1/IMPACT-Gcn1 interaction impairs the GAAC/ISR response 18

1.6.2 Mapping the Yih1 binding site of Gcn1 18

1.6.3 Yih1 amino acids D102 and E106 are important for Gcn1 binding 18

1.6.4 Structural modelling of Yih1: Insight into Yih1-mediated Gcn2 inhibition 19

1.6.5 Yih1 interacts with Actin at native levels 19

1.6.6 IMPACT, the mammalian homolog of Yih1, also interacts with actin 20

1.6.7 The IMPACT/Yih1-Actin interaction may prevent the inhibition of Gcn2 20

1.6.8 Yih1-Ribosome interaction: Potential roles in translation regulation 21
1.6.9 Yih1-Ribosome interaction is independent of Gcn1 22
1.6.10 IMPACT also interacts with the Ribosome 22
1.6.11 Yih1/IMPACT interacts with cell-cycle dependent proteins Cdc28/CDK1 22
1.6.12 Yih1 regions for Gcn1, Actin, and Ribosomal Binding overlap 23
1.7 Proposed mechanism of Yih1/IMPACT-mediated Gcn2 inhibition 25
1.8. Biological implications of Yih1/IMPACT 29
1.8.1 The Gcn2-Yih1/IMPACT axis has a major role in mammalian neurophysiology and neural development 29
1.8.2 IMPACT is implicated in neurodevelopmental disorders and diseases 29
1.8.3 IMPACT may modulate cancer cell survival 29
1.8.4 Physiological implication of the Yih1/IMPACT-Gcn2 axis 30
1.9. Hypothesis and Aims 31
2. Materials and Methods 33
2.1.1 Yeast strains and plasmids used 33
2.1.2 Media used 35
2.1.3 Media supplements 36
2.1.4 Drug stock solution used in this study 37
2.1.5 Antibiotics used in this study 37
2.1.6 Growth conditions 37
2.1.7 Permanent storage of Yeast cultures 38
2.1.8 Permanent storage of *E. coli* cultures 38
2.1.9 Primers used in this study 38
2.1.10 Plasmid DNA isolation from Yeast 38
2.1.11 Plasmid DNA isolation from *E. coli* 39
2.1.12 Determining DNA concentration and purity of samples 39
2.1.13 Agarose gel electrophoresis 40

2.2.1 *Escherichia coli* (*E. coli*) Transformation 40
2.2.2 Yeast Transformation 41
2.2.3 Measuring Optical Density of Cultures 43
2.2.4 Semi-quantitative growth assay (SQGA) 43
2.2.5 Polymerase Chain Reaction 43
2.2.6 Colony PCR 45
2.2.7 Fusion PCR 45
2.3 Preparation of Whole Cell Extract for Western Blotting 46
2.3.1 Formaldehyde crosslinking 46
2.3.2 Chemical cell lysis 46
2.4. Western Blot 46
2.4.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) 46
2.4.2 Transfer to nitrocellulose membrane 47
2.4.3 Western blot analysis 48
2.5 Harvesting cells for Yih1 purification 50
2.5 Mechanical cell lysis for generation of whole cell extract 50
2.6 Size Exclusion Chromatography for purification of yeast Yih1 51
2.7 Fast Protein Liquid Chromatography 51
2.8 In-Gel Chymotrypsin Digest 52
2.8.1 Preparation of Gel for Digest 52
2.8.2 In Gel Proteolytic Digest using Chymotrypsin 52
2.9 Tandem Mass Spectrometry (MS/MS) 54
2.10 Site directed mutagenesis 54
2.11 Next Generation Sequencing (NGS) 55
2.12 Statistical analysis 55
3. Results 56
Grouping of Yih1/IMPACT Amino Acids 56

Overexpressing Yih1 for molecular insight on the proteins function 58
3.1 The yeast system used for scoring Yih1-mediated Gcn2 inhibition 59
3.1.2 Plasmid map of Yih1/IMPACT used in this study 60
3.1.3 Semi Quantitative Growth Assay for scoring Yih1-mediated Gcn2 inhibition..... 61
3.1.4 Quantitative analysis of Semi Quantitative Growth Assays..... 62
3.1.5 Overexpression of Yih1 or IMPACT in yeast causes 3AT sensitivity..... 64
3.1.6 Determining expression levels of Yih1/IMPACT 67
3.1.7 Yih1/IMPACT overexpression reduces eIF2 phosphorylation levels..... 69
3.2 Discussion 72
3.2.1 Verification of yeast system for scoring Yih1-mediated Gcn2 inhibition..... 72
3.3 Group 1: Amino acids involved in Gcn1 binding 73
3.3.1 Cells with a Yih1 D102A substitution retain growth under starvation conditions.. 74
3.3.2 Verifying the correct substitution in Yih1 variants 77
3.3.3 Verification of Yih1 protein expression levels..... 80
3.3.4 eIF2α phosphorylation is differentially affected by Yih1 mutants..... 83
3.4 Discussion 86
3.4.1 D102 is important for the Yih1-Gcn1 interaction 86
3.4.2 Potential for salt bridge formation underlying Yih1-Gcn1 interaction 87
3.4.3 E106 is not relevant for the Yih1-Gcn1 interaction or Yih1 interdomain interactions 87
3.4.5 Remaining undetermined amino acids that are involved in the Yih1-Gcn1 interaction 88
3.4.6 Future research towards determining all amino acids involved in the Yih1-Gcn1 interaction 88
3.5 Examining the evolutionary conservation of Yih1/IMPACT family..... 89
3.5.1 Scoring the Yih1/IMPACT alignment using the PAM250 scoring matrix 90
3.6 Discussion 92

3.6.1 Comparative analysis of our Yih1/IMPACT sequence alignment with previous alignments 92
3.7 Group 2: RWD domain amino acids involved in interdomain interactions 93
3.7.1 Substituting E87 and D90 to a different charge enhances Yih1 potency 94
3.7.2 Verifying the correct substitution of Yih1 E87 and D90 variants 97
3.7.3 Determining expression levels of Yih1 Group 2 mutants 100
3.8 Discussion 102
3.8.1 The Yih1 conformation influences the potency of Yih1-mediated Gcn2 inhibition 102
3.8.2 E87 and D90 are the major players in interdomain interactions 105
3.8.3 Potential role of E87 in Yih1 interdomain interactions/Yih1 function 105
3.8.4 Changing the charge of E87 and D90 does not completely inhibit Gcn2 activation 106
3.8.5 Future research 107
3.9 Group 3: Ancient domain amino acids involved in interdomain interactions 107
3.9.2 Changing the charge of K173, R175 and K176 enhances Yih1 potency 109
3.9.3 A R175 substitution to Alanine or Glutamic Acid does not influence Yih1 potency 111
3.9.4 Verifying the correct substitution of Yih1 Group 3 mutations 112
4.0 Discussion 115
4.1.1 Ancient domain residues create a positively charged interface for the RWD domain 115
4.1.1 R175 alone is not important for interdomain interactions 115
4.1.2 Additional residues may be involved in contacting the RWD domain 116
4.1.3 Future research 117
4.2 Group 4: Potential role of Post-Translational Modification in Yih1 function 119
4.2.1 Introduction of R144 substitutions by Site directed mutagenesis 120
4.2.2 Substituting Arg144 to Ala or Lys impairs cell growth 123

4.2.3 Verifying the correct R144 substitutions in Yih1 variants 124
4.2.4 Purification of Yeast GST-Yih1 by Size Exclusion Chromatography 126
4.2.5 Arg144 is not dimethylated in Yeast, its native environment 130
4.3 Discussion 135
4.3.1 A non-methylatable version of Yih1 has enhanced potency in Gcn2 inhibition... 135
4.3.2 No evidence of Yih1 R144 dimethylation was detected in yeast. 136
4.3.3 Biological analysis of the significance of additional Yih1 modifications by Mass Spectrometry 137
4.3.4 Potential role of Arg-144 in Yih1-mediated Gcn2 regulation 138
4.4 General Discussion 140
4.4.1 Our yeast system is sufficient for identifying important amino acids in Yih1 function 140
4.4.2 Specific Yih1 residues are involved in Gcn1 binding 140
4.4.3 Ionic interactions are involved interdomain Yih1 interactions 141
4.4.4 Shedding light on additional regulatory mechanisms that modulate Yih1-mediated Gcn2 inhibition 144
4.5 Conclusion and Prospective studies 145
4.6 References 147
4.7 Supplementary Material 155
S1: IMPACT and Yih1 158
S2: Group 1: Amino acids involved in Gcn1 binding 161
S3: Group 2: RWD domain amino acids involved in interdomain interactions 171
S4: Group 3: Ancient domain amino acids involved in interdomain interactions 181
S5: Group 4: Potential role of Post-Translational Modification in Yih1 function 189

List of Tables

Table 1: Yeast strains used in this study 33
Table 2: Plasmids used in this study 34
Table 3: Composition of Synthetic Defined (SD) medium 35
Table 4: Composition of Yeast Peptone Dextrose (YPD) medium 35
Table 5: Composition of Luria Bertani media 36
Table 6: Components of ILV amino acid stock mixture 37
Table 7: Composition of 3AT used in this study 37
Table 8: Composition of Ampicillin used in this study 37
Table 9: Primers used in this study 38
Table 10: Composition of SOB medium (pH 7) 41
Table 11: Standard steps of PCR 44
Table 12: A standard PCR reaction mix 44
Table 13: Thermocycler protocol for Colony PCR 45
Table 14: Master mix for Colony PCR 45
Table 15: Primary antibodies used in this work 49
Table 16: Secondary antibodies used in this work 49
Table 17: Components of SEC Buffer (pH 7) 51
Table 18: Colloidal stock recipe (250mL) 52
Table 19: Components of Digestion solution in this work 54
Table 20: Grouping of Yih1 amino acids in this study 58
Table 21: Purification of GST-Yih1 from *S. cerevisiae* 127
Table 22: Posttranslational modifications identified in Yih1 by MS/MS 131

List of Figures

Figure 1: Overview of translation initiation in Eukaryotes. 3
Figure 2: Translational control of *GCN4* mRNA. 5
Figure 3: The current working model of Gcn2 activation by uncharged tRNAs. 7
Figure 4: Schematic overview of the Gcn2 structure. 10
Figure 5: Schematic overview of the Gcn1 structure. 12
Figure 6: Yih1/1IMPACT is an inhibitor of Gcn2. 14
Figure 7: Yih1/IMPACT protein structure. 16
Figure 8: The minimal binding regions of Yih1 for Gcn1, the ribosome and actin. 24
Figure 9: The current working model of Yih1/IMPACT mediated Gcn2 inhibition. 27
Figure 10: Overview of the site directed mutagenesis procedure. 55
Figure 11: Yih1 is a competitive inhibitor of Gcn2. 59
Figure 12: Plasmid map of Yih1/IMPACT protein constructs used in this study. 60
Figure 13: Experimental procedure for a Semi-Quantitative Growth Assay (SQGA). 62
Figure 14: Representative quantitative growth scoring system of yeast cells. 63
Figure 15: IMPACT or Yih1 overexpression impairs yeast cell growth under starved conditions. 65
Figure 16: Expression levels of Yih1/IMPACT used in this study. 67
Figure 17: Yih1/IMPACT overexpression impairs eIF2α phosphorylation. 69
Figure 18: Yih1 Amino Acid D102 is important for inhibiting Gcn2 75
Figure 19: Sequencing analysis of Yih1 Group 1 mutants. 78
Figure 20: Expression levels of Yih1 Group 1 mutants 80
Figure 21: eIF2 phosphorylation levels are affected by Yih1 with Group 1 substitutions. 83
Figure 22: Multiple sequence alignment of Yih1/IMPACT proteins. 89
Figure 23: SQGA of Yih1 with either an E87K substitution or a E87 and D90 double substitution. 95
Figure 24: Verifying the presence of Yih1 Group 2 mutations by Sequencing 98
Figure 25: Expression levels of Yih1 Group 2 mutants. 100
Figure 26: Structural model of the surface of Yih1. 103
Figure 27: E87 and R175 are closely positioned at the Yih1 domain interface. 106
Figure 28: Testing the ability of Yih1 with Group 3 substitutions to affect cell growth in starved conditions. 110
Figure 29: Sequence verification of Yih1 Group 3 mutations 113
Figure 30: K173 and E106 are positioned close to each other in Yih1. 118

Figure 31: Introduction of R144 substitutions into Yih1 by Site-directed mutagenesis..... 121
Figure 32: R144 substitution to Alanine or Lysine impairs cell growth in starved conditions. 124
Figure 33: Sequence verification of GST-Yih1 Arg144 mutants. 125
Figure 34: Chromatogram of Size Exclusion Chromatography of GST-Yih1. 128
Figure 35: SDS-page analysis of pure GST-Yih1 fractions from Size Exclusion Chromatography 129
Figure 36: Preparation of GST-Yih1 for Tandem Mass Spectrometry (MS/MS) 130
Figure 37: Analysis of Mass Spectrometry data of GST-Yih1 134
Figure 38: Position of Arg144 (R144) in the Ancient domain of Yih1 135
Figure 39: Superimposition of the Yih1/IMPACT RWD domain. 141

List of Abbreviations

The following abbreviations were used in addition to chemical symbols from the Periodic Table of Elements and the International System of Units (SI)

Abbreviation	Meaning
3AT	3-Amino-1, 2, 4-triaxole
AA	Amino Acid
AD	Ancient Domain
Amp	Ampicillin
APS	Ammonium Persulfate
ATF4	Activating transcription factor 4
BiFC	Bimolecular fluorescence complementation
BSA	Bovine Serum Albumin
BP	Base pair
CDK1	Cyclin Dependent Kinase 1 (Mammalian)
Cdc28	Cyclin-dependent kinase (Yeast)
DNA	Deoxyribonucleic acid
Co-IP	Co-Immunoprecipitation
kDa	Kilo Dalton
DMSO	Dimethylsulfoxide
E. coli	*Escherichia coli*
EDTA	Ethylene Diamine Tetra Acetic Acid
eEF1	Eukaryotic Elongation Factor 1
eEF3	Eukaryotic Elongation Factor 3
eIF2	Eukaryotic Initiation Factor 2
eIF2α	Eukaryotic Initiation Factor 2 alpha subunit
eIF2α-P	Eukaryotic Initiation Factor 2 phosphorylated alpha subunit
eIF2B	Guanine nucleotide exchange factor
EtBr	Ethidium Bromide
GAAC	General Amino Acid Control
Gal	Galactose
Gcn1	General control non-derepressible 1
Gcn2	General control non-derepressible 2

Gcn4	General control non-derepressible 4
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GST	Glutathione S- Transferase
GSH	Glutathione Reduced
GTP	Guanosine triphosphate
Glc	Glucose
ddH20	Double Distilled water
I	Isoleucine
IMPACT	Imprinted and Ancient protein
ISR	Integrated Stress Response
L	Leucine
LB	Luria-Bertani
M	Marker (2log DNA ladder)
mRNA	Messenger ribonucleic acid
MW	Molecular Weight
dNTP	Deoxy Nucleotide Tri Phosphate
OD	Optical density
ORF	Open Reading Frame
uORF	Upstream open reading frames
PAGE	Polyacrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PEG	Polyethylene glycol
Pgk1	3-Phosphoglycerate kinase
Psi	Pounds per square inch
MilliQ	Double de-ionised water
RNA	Ribonucleic Acid
RNase	Ribonuclease
dsRNA	Double Stranded RNA
tRNA	Transfer RNA
rpm	Rotations per minute
RT	Room Temperature
RWDBD	RWD binding domain

RWD	RING Finger WD-repeat containing, DEAD helicase
SD	Synthetic Dextrose
SDS	Sodium Dodecyl Sulphate
SDWILV	SD medium supplemented with Tryptophan, Isoleucine, Leucine and Valine
ss	Single strand
SOC	Super Optimal Broth with Catabolic Repression
SOB	Super Optimal Broth
SQGA	Semi Quantitative Growth Assay
TAE	Tris-Acetate EDTA
TBS	Tris-Buffered Saline
TBS-T	TBS-Tween
TC	Tertiary Complex
TEMED	N,N,N',N'-Tetramethylethylenediamine
TE	Tris EDTA buffer
tRNA	Transfer RNA
RNA	Ribonucleic Acid
$tRNA_i^{Met}$	Initiator methionyl tRNA
$tRNA^{deacyl}$	Uncharged tRNA
TRP1	Phosphoribosyl Anthranilate Isomerase
V	Valine
W	Tryptophan
Yih1	Yeast IMPACT homologue-1
YPD	Yeast Peptone Dextrose
YPG	Yeast Peptone Glycerol
YNB	Yeast Nitrogenous Base

1.1 Introduction

1.1.1 The eIF2α kinase Gcn2 is involved in the Integrated Stress Response in mammals

It is vital that eukaryotes adjust their cellular gene expression profile in response to dynamic environmental or intracellular changes/stress, such as nutrient limitation (Masson, 2019; Sattlegger & Hinnebusch, 2000; Silva, Sattlegger, & Castilho, 2016). Failure to do so can lead to cellular dysregulation and can further manifest as pathological conditions in mammals. Central to this regulatory response is phosphorylation of the eukaryotic translation initiation factor-2 (eIF2) at the α-subunit, which is catalysed by eIF2α kinases. In mammals, there are 4 different eIF2α kinases that are each activated by different stress signals to help the cell appropriately respond; PERK (PKR-like endoplasmic reticulum kinase), PKR (RNA-dependent protein kinase), HRI (Heme regulator inhibitor), and Gcn2 (General Control Non-Derepressible-2) (Roffé, Hajj, Azevedo, Alves, & Castilho, 2013). PERK is activated by the accumulation of unfolded proteins in the endoplasmic reticulum, PKR is activated by dsRNA (double stranded RNA) during viral infection, and HRI is activated by heme deficiency or oxidative stress (Roffé et al., 2013). Additionally, Gcn2 is a highly conserved eIF2α kinase that is present in all eukaryotes, from yeast (*Saccharomyces cerevisiae*) to mammals (Castilho, Shanmugam, Silva, Ramesh, Himme & Sattlegger, 2014). Gcn2 was first discovered in Baker's yeast, where it is the only eIF2α kinase (Castilho et al., 2014). Because of this, and its conservation among eukaryotes, a large amount of research on the molecular mechanisms and regulation of Gcn2 has been conducted in yeast. The activating signal for Gcn2 is amino acid starvation, which is detected in the cell by accumulation of uncharged tRNAs (Castilho et al., 2014). Additionally, Gcn2 is also activated by other stress factors such as UV irradiation and viral replication (Castilho et al., 2014).

1.1.2 Translation initiation in eukaryotes

The substrate for the eIF2α kinase Gcn2 is eIF2. eIF2 has a fundamental involvement in translation initiation (Hinnebusch, 2005; Hinnebusch & Natarajan, 2002). eIF2 delivers charged initiator methionyl-tRNA ($tRNA_i^{Met}$) to the 40S ribosomal subunit in a GTP-bound ternary complex (TC), forming a 43S initiation complex for translation initiation (Hinnebusch, 2005). Subsequently, the 43S initiation complex binds to the 5'-capped end of *GCN4* mRNA to form the 48S complex (Hinnebusch, 2005). This 48S complex then scans *GCN4* mRNA in a 5' to 3' direction until it recognises the AUG start codon by codon-anticodon recognition (Castilho et al., 2014; Hinnebusch, 1997; Hinnebusch, 2005; Hinnebusch & Natarajan, 2002).

Subsequently, the GTP of the ternary complex is hydrolysed to GDP, resulting in the release of eIF2-GDP from the ribosomal complex (Hinnebusch, 2005). The 60S ribosomal subunit is then recruited to the 48S complex to form the 80S ribosomal complex for commencement of *GCN4* mRNA translation (Hinnebusch, 2005). For the translation initiation process to be repeated, the guanine nucleotide exchange factor eIF2B is required to recycle eIF2 back into a GTP-bound form (Castilho et al., 2014; Hinnebusch, 2005). However, during starved cellular conditions, phosphorylation of eIF2α by Gcn2 causes eIF2α to become an inhibitor of its guanine nucleotide exchange factor eIF2B (Castilho et al., 2014; Hinnebusch, 2005). This results in reduced TC formation, meaning that $tRNA_i^{Met}$ cannot be transported to the ribosome for translation initiation (Castilho et al., 2014; Hinnebusch, 2005). Furthermore, uncharged tRNAs ($tRNA^{deacyl}$) accumulate in the cell due to the lack of TC (Castilho et al., 2014; Hinnebusch, 2005). As a result, this phosphorylation event causes a global reduction in cellular protein synthesis levels. Simultaneously, there is an upregulation of expression of genes encoding proteins that are involved in amino acid biosynthesis and transport (Hinnebusch, 2005). This specific cellular response, aimed at recovery from the initial stress signal, involves the specific translational regulation of *GCN4* (in yeast)/*ATF4* (in mammals). *GCN4* codes for a transcription factor that activates the transcription of genes involved in amino acid biosynthesis and transport (Hinnebusch, 2005). This is central to the General Amino Acid Control (GAAC) response in yeast and the Integrated Stress Response (ISR) in mammals.

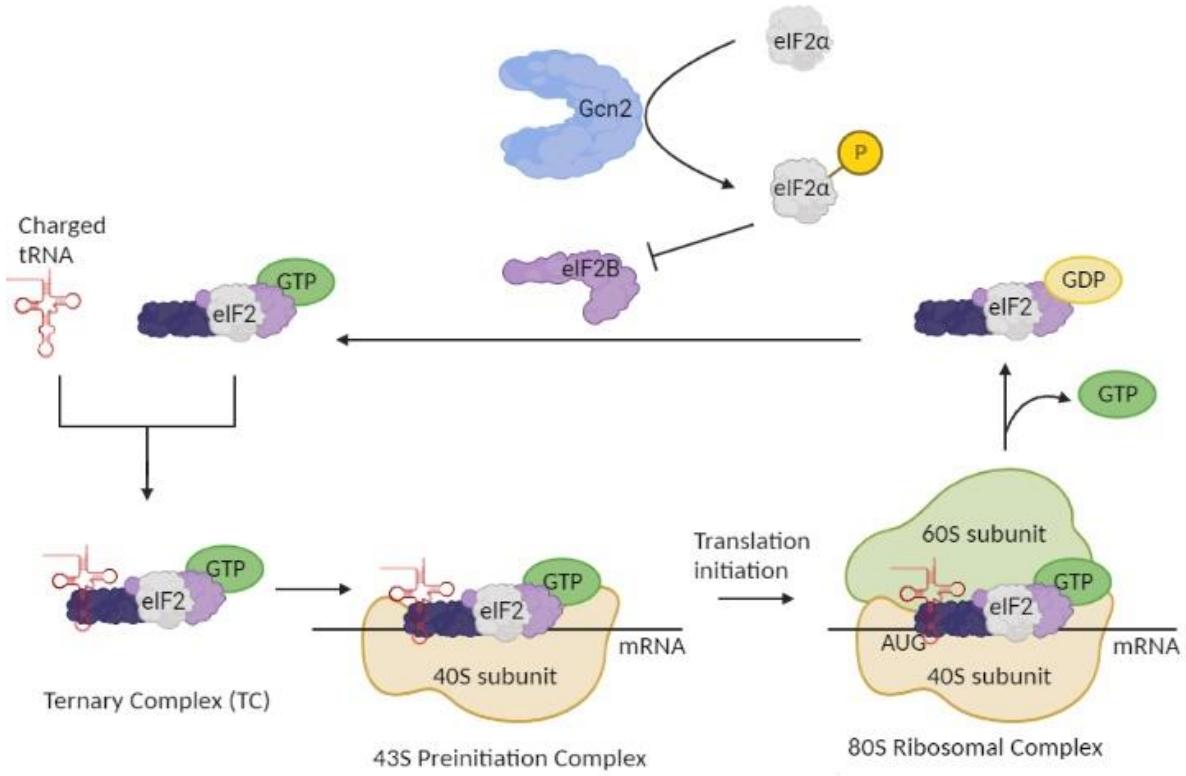


Figure 1: Overview of translation initiation in Eukaryotes.

See text for details. Figure generated in BioRender.

1.1.3 GCN4 Translational control: Regulatory roles of the uORFs

The corresponding reduction in protein synthesis and upregulation of specific genes, in order the counteract the initial stress signal, is a result of the translational regulation of *GCN4* mRNA. This specific response is mediated by four upstream open reading frames (uORFs) in the mRNA of *GCN4* (Hinnebusch, 1997). In this circumstance, uORFs act as negative regulatory elements as their translation impairs translation of *GCN4* mRNA (Hinnebusch, 1997; Hinnebusch, 2005). As mentioned, the 43S preinitiation complex binds *GCN4* mRNA to form the 48S ribosomal complex, where it scans in a 5' to 3' direction for an AUG start codon. The 48S complex will firstly encounter the uORF1 start codon, positioned upstream of the *GCN4* open reading frame, where translation is initiated as described previously (Fig. 2). As uORF1 has a weak stop codon, approximately half of the 40S ribosomal subunits remain bound to the *GCN4* mRNA after uORF1 translation to continue scanning downstream, while the other approximate half of 40S ribosomes dissociate off the *GCN4* mRNA, thus are unable to re-initiate translation (Fig. 2) (Hinnebusch, 1997; Hinnebusch, 2005). In non-starved cells, where the concentration of TC is high, the majority of 40S ribosomes that remained attached to the mRNA will rebind TC before reaching the uORFs further downstream (Fig. 2). Therefore,

translation will be re-initiated at uORFs 2, 3 and 4, which are downstream of uORF1 and have stronger stop codons (Fig. 2) (Hinnebusch, 2005). The uORFs 2 and 3 are functionally redundant. The result of this translational response is that GCN4 translation will be repressed due to the inhibitory effect of the uORFs (Fig. 2). However, in starved cells, the TC levels are low as eIF2α phosphorylation by Gcn2 causes eIF2α to become an inhibitor of its guanine nucleotide exchange factor eIF2B, meaning that GDP (of the eIF2-GDP complex which falls off the ribosome) cannot be recycled back to GTP to become part of the initiation complex for translation (Hinnebusch, 2005). Therefore, the majority of 40S ribosomes that remain attached to the mRNA after uORF1 translation will fail to re-bind TC before reaching the downstream uORFs, such as uORF4. This means that translation of uORFs 2,3 and 4 will be low. However, as the distance between uORF4 and the *GCN4* start codon is long, most of these 40S ribosomes will successfully re-bind TC before reaching the *GCN4* start codon, despite there being a lower concentration of TC (Fig. 2). The result of this is that translation is successfully initiated at the *GCN4* start codon, allowing *GCN4* to be derepressed. The mammalian counterpart to *GCN4* is *ATF4*, which is translationally regulated in a similar manner to the described above. This will not be explained in detail as our studies are conducted in yeast cells.

This change in cellular gene expression profile; a reduction protein synthesis and increase in expression of amino acid biosynthesis proteins, are the basis of the GAAC response in yeast or the ISR in mammals. This signalling pathway enables cells to appropriately respond to and cope with the initial stress signal and restore homeostasis. Hence, Gcn2 is fundamentally involved in the maintenance of normal cell function and for adequate adaptation to stress through changes in gene expression (Masson, 2019). Gcn2 is implicated in many biological processes such as immune system modulation, amino acid homeostasis, nutrient sensing, and neural functions such as memory formation (Castilho et al., 2014; Pereira, Sattlegger, Jiang, Longo, Jaqueta, Hinnebusch, Wek, Mello, Castilho, 2005). Furthermore, various diseases and disorders have been associated with Gcn2, such as cancer and Alzheimer's disease (Castilho et al., 2014). Therefore, it is crucial to understand the molecular mechanisms of Gcn2 in order to understanding normal and abnormal cell function, diseases and potential disease treatment (Castilho et al., 2014). Considering the central role of Gcn2 in cellular processes and translation, it is crucial that the function of Gcn2 is tightly regulated.

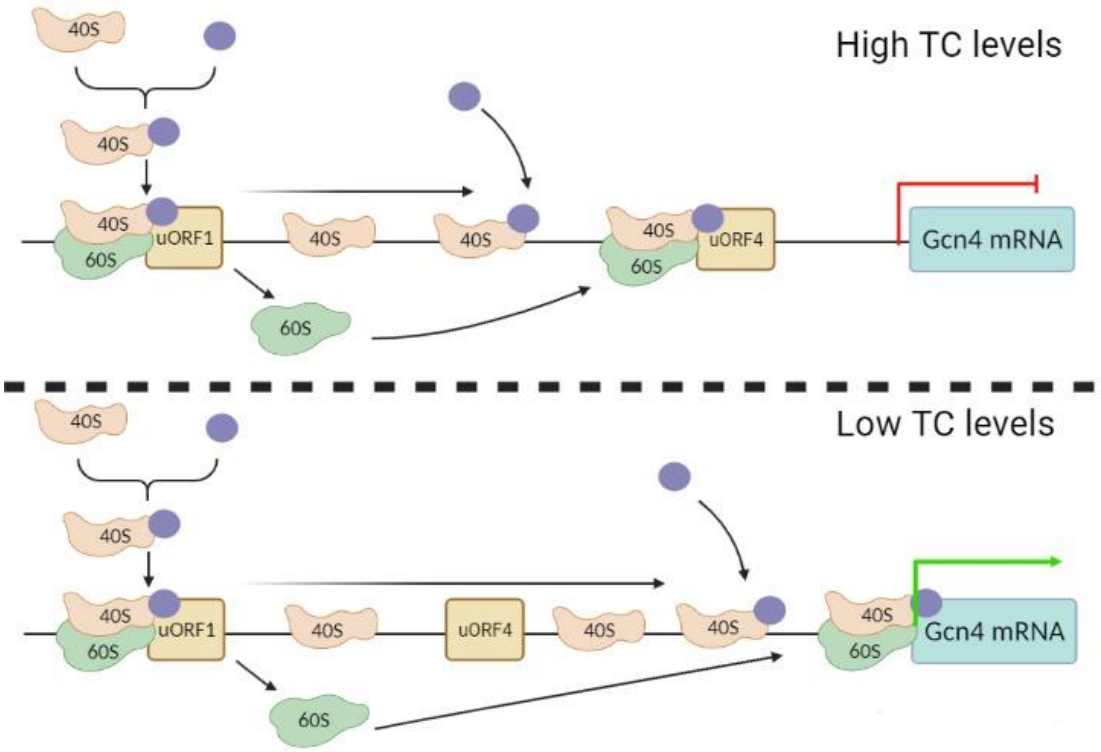


Figure 2: Translational control of *GCN4* mRNA.

Under non-starved conditions, *GCN4* mRNA translation is repressed. This is because high TC levels mean that the majority of 40S ribosomes that are attached to the Gcn4 mRNA will rebind TC before reaching the downstream uORFs. Therefore, translation will be re-initiated at the uORFs downstream of uORF1, at uORFs 2,3 and 4 (Hinnebusch, 2005). This means that *GCN4* translation will be repressed due to the inhibitory effect of the uORFs. The uORFs 2 and 3 are functionally redundant, therefore only uORFs 1 and 4 are shown in this Figure (Top panel). In starved cells, low TC levels means that the majority of 40S ribosomes that remain attached to the mRNA after uORF1 translation will fail to re-bind TC before reaching the downstream uORFs, such as uORF4. This means that translation of uORFs 2,3 and 4 will be low. However, as the distance between uORF4 and the *GCN4* start codon is long, most of these 40S ribosomes will successfully re-bind TC before reaching the *GCN4* start codon, despite there being a lower concentration of TC. The result of this is that translation is successfully initiated at the GCN4 start codon, allowing *GCN4* to be translated. TC: Ternary Complex. Figure generated in BioRender.

1.2. The working model of Gcn2 activation

Gcn1, a coactivator protein of Gcn2, is absolutely essential for Gcn2 activation. Furthermore, Gcn1 and Gcn2 both directly bind to the ribosome, and the formation of a Gcn1-Gcn2-ribosomal complex is also essential for Gcn2 activation (Sattlegger & Hinnebusch, 2005). Additionally, the protein Gcn20 also exists in this complex, however, is not essential for Gcn2 activation. The activating signal for Gcn2 is amino acid starvation, which is communicated to the cell through the accumulation of uncharged tRNAs. Furthermore, depletion of only a single amino acid is required for Gcn2 activation (Castilho et al., 2014)

Research has been carried out to understand how Gcn2 detects this cellular condition and becomes activated. Under normal physiological conditions, Gcn2 exists in an inactive state which is maintained by auto-inhibitory molecular interactions (Castilho et al., 2014). Upon deprivation of one or more amino acids, uncharged tRNA ($tRNA^{deacyl}$) accumulates intracellularly and enters the ribosomal A-site (Castilho et al., 2014). From here, it is proposed that $tRNA^{deacyl}$ is transferred to the Histidyl-tRNA synthetases-like domain (HisRS-like domain) of Gcn2 for Gcn2 to become activated (Zhu, Sobolev, & Wek, 1996; Castilho et al., 2014). Allosteric rearrangements upon uncharged tRNA binding then releases Gcn2 inhibition, enabling it to phosphorylate its substrate eIF2α at Ser-51 (see Fig. 3) (Castilho et al., 2014). Although Gcn1 is essential for Gcn2 activation, the exact role of Gcn1 in this process is ambiguous. In this model, it is suggested that Gcn1 may act as a scaffold protein to hold Gcn2 in the correct conformation for receiving the starvation signal (Castilho et al., 2014). Alternatively, Gcn1 may function in facilitating the transfer of uncharged tRNAs from the ribosomal A-site to Gcn2 (Castilho et al., 2014). This is supported by findings based on the structural homology of a Gcn1 domain (eEF3-like domain) to fungal translation elongation factor 3 (eEF3) (Castilho et al., 2014; Marton, Crouch, & Hinnebusch, 1993; Rakesh, Krishnan, Sattlegger, & Srinivasan, 2017). EF3 is essential for the transfer and release of uncharged tRNA from the ribosomal E-site, suggesting Gcn1 may play a similar role in facilitating transfer of the starvation signal to Gcn2 (Triana, Nierhaus, Ziehler, & Chakraburtty, 1993).

The working model of Gcn2 activation, whereby uncharged tRNA activates Gcn2, is supported experimentally. Firstly, Gcn2 contains a domain homologous to histidyl-tRNA synthetases (HisRS-like domain) in its protein structure, similarly to the bacterial protein RelA (Castilho et al., 2014). In bacteria, RelA binds uncharged tRNAs upon amino acid starvation (Castilho et al., 2014). Therefore, it was proposed that this domain in Gcn2 functions in a similar manner

by binding uncharged tRNA being transferred from the ribosomal A-site during amino acid starvation. Furthermore, mutations to a conserved region of the HisRS-like domain of Gcn2 abrogates the ability for Gcn2 to bind uncharged tRNA and abolishes Gcn2 activation, signalled by reduced eIF2 phosphorylation levels (Wek, Zhu, & Wek, 1995).

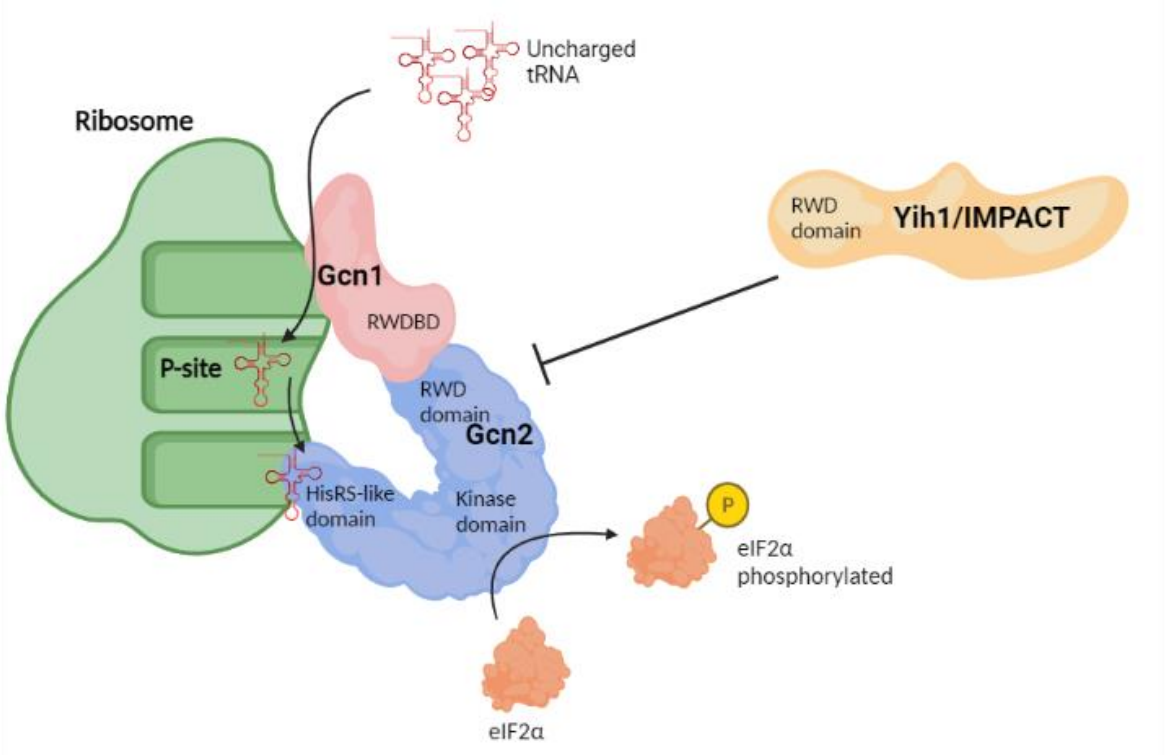


Figure 3: The current working model of Gcn2 activation by uncharged tRNAs.

See text for details. Figure generated in BioRender.

1.2.1 Alternative models of Gcn2 activation

More recent research has shed light on alternative potential mechanisms of Gcn2 activation that may be working in conjunction or independently of uncharged tRNA being the direct activator of Gcn2 (Harding et al., 2019; Ishimura et al., 2016; Inglis et al., 2019). In mammalian cells, Gcn2 activation has been observed upon the induction of ribosome stalling, and this observation was independent of uncharged tRNA levels (Ishimura et al., 2016). These findings challenge the proposed model by suggesting that ribosome stalling, or the ribosomal translational state, may play an important role in Gcn2 activation (Ishimura et al., 2016). Further supporting the role of the ribosomal translational state in Gcn2 activation, potent Gcn2 activation was stimulated by ribosomes and the ribosomal P-stalk in comparison to uncharged tRNAs (Inglis et al., 2019). From this, an alternative model was suggested whereby amino acid starvation stimulates ribosome stalling (Ishimura et al., 2016; Inglis et al., 2019). This would allow P-stalk regions that normally associate with elongation factors during translation to become

available to for Gcn2 activation (Inglis et al., 2019). Currently, it is not known whether this mechanism is working in conjunction with uncharged tRNAs, and/or whether the starvation state of the cell affects the interaction between Gcn2 and the ribosomal P-stalk (Inglis et al., 2019). Further research is required to understand the complete mechanism of how Gcn2 becomes activated.

Regulatory proteins have been identified that positively or negatively impact Gcn2 activation, such as the positive regulatory protein Gcn1. Importantly for this study, Yih1/IMPACT is a negative regulatory protein of Gcn2 which will be discussed further.

1.3 The structure of Gcn2

Gcn2 is a protein kinase that is comprised of 5 functional domains (Castilho et al., 2014). From the N-terminus to the C-terminus, Gcn2 contains an RWD domain (from its presence in RING finger proteins, WD-repeat-containing proteins, and yeast DEAD-like helicases), a pseudokinase (ΨPK) domain, a protein kinase (PK) domain, a histidyl-tRNA synthetase-related domain (HisRS-like domain) and a C-terminal domain (CTD) (see Fig. 4) (Castilho et al., 2014; Masson, 2019). The Gcn2 RWD domain directly interacts with Gcn1 and is essential for Gcn2 activation (Kubota, Sakaki, & Ito, 2000; Sattlegger & Hinnebusch, 2000). For this research, the RWD domain of Gcn2 is of particular interest, and will be discussed deeper, as the regulatory protein Yih1/IMPACT also contains an RWD domain in its protein structure. The ΨPK domain is a positive regulatory domain that allosterically interacts with the PK domain upon Gcn2 activation (Lageix et al., 2014). The PK domain catalyses eIF2α phosphorylation at the Serine-51 residue (Ser-51) upon amino acid starvation, and also becomes auto-phosphorylated during Gcn2 activation (Lageix et al., 2014). The HisRS-like domain is involved in binding $tRNA^{deacyl}$ (Wek et al., 1995). The CTD is involved in $tRNA^{deacyl}$ and ribosomal binding (Wek et al., 1995). Intramolecular auto-inhibitory interactions between Gcn2 domains are central to the regulation of Gcn2 activation (Masson, 2019).

NMR spectroscopy has structurally defined the mouse Gcn2 RWD domain as an αβ-sandwich fold of two layers; one layer having a four-stranded β-sheet and the other having three α-helices with the topological arrangement of αββββαα (Nameki et al., 2004). Characteristic of all RWD-containing proteins, the Gcn2 RWD domain contains an invariant YPXXXP motif (Nameki et al., 2004). The YPXXXP motif is involved in a stable triple β-turn which is surrounded by a core internal hydrogen bond network stabilising the RWD domain (Nameki et al., 2004).

Figure 4: Schematic overview of the Gcn2 structure.

Gcn2 is composed of an RWD domain, Pseudokinase domain (Ψ PK), Protein Kinase domain, HisRS-like domain, and a C-terminal domain (CTD) Binding sites for Gcn1, uncharged tRNA, and the ribosome are indicated by grey arrows.

1.4 Gcn1 and its direct interaction with Gcn2 is essential for Gcn2 activation

A network of proteins has been identified to regulate Gcn2 activity, such as Gcn1 and IMPACT/yeast IMPACT homologue-1 (Yih1), along with other proteins (Castilho et al., 2014; Silva et al., 2016). Importantly, the direct interaction between Gcn2 and its positive regulator Gcn1 is absolutely essential for Gcn2 activation (Castilho et al., 2014; Kubota et al., 2000; Marton et al., 1993; Marton et al., 1997; Sattlegger & Hinnebusch, 2000). Multiple lines of independent research have established the direct Gcn1-Gcn2 interaction both by a protein overlay assay (Sattlegger & Hinnebusch, 2000) and yeast two-hybrid assay (Kubota et al., 2000). Moreover, the direct Gcn1-Gcn2 interaction has been confirmed *in vivo* by a GST-pulldown assay using purified proteins (Sattlegger & Hinnebusch, 2000).

The Gcn1-Gcn2 interaction is mediated by the Gcn2 RWD domain (from its presence in RING finger proteins, WD-repeat-containing proteins, and yeast DEAD-like helicases), which contacts Gcn1 at a distinct C-terminal region notated as the RWD-binding domain (RWDBD) (Kubota et al., 2000; Rakesh et al., 2017; Sattlegger et al., 2004; Sattlegger & Hinnebusch, 2000). The RWDBD domain encompasses a region spanning amino acids 2052-2428 (Sattlegger & Hinnebusch, 2000) or 2048–2383 (Kubota et al., 2000). Moreover, the Gcn1-Gcn2 interaction is mediated by a single amino acid (R2259) situated within the Gcn1 RWDBD (Sattlegger & Hinnebusch, 2000). A single R2259A substitution mutation abolishes the Gcn1-Gcn2 interaction and impairs the GAAC response (Sattlegger & Hinnebusch, 2000). Furthermore, under amino acid starved conditions, overexpression of either the Gcn2 RWD domain or the Gcn1 RWDBD domain impairs yeast cell growth and the GAAC response, and this phenotype is supressed by overexpressing full length Gcn2 or Gcn1 (Garcia-Barrio, Dong, Ufano, & Hinnebusch, 2000; Kubota et al., 2000; Kubota et al., 2001; Sattlegger & Hinnebusch, 2000). These collective findings strongly support the statement that that the direct Gcn1-Gcn2 interaction is essential for Gcn2 activation.

1.4.1 Gcn1 structure and the binding determinants for the Gcn1-Gcn2 interaction

Gcn1 is a large protein that is conserved among all eukaryotes and resides in the cell cytoplasm (Castilho et al., 2014). Gcn1 is the essential binding partner of Gcn2, as the direct Gcn1-Gcn2 interaction is required for Gcn2 activation (Castilho et al., 2014; Sattlegger & Hinnebusch, 2000). Furthermore, a region of Gcn1, referred to as the eEF3-like region (amino acids 1330-1641), is comprised of numerous HEAT repeats homologous to the fungal translation elongation factor-3 (eEF-3), which is involved in transfer of tRNA from the ribosome during translation (Castilho et al., 2014; Rakesh et al., 2017). Furthermore, Gcn1 interacts directly

with translating ribosomes and this interaction requires Gcn1 amino acids 1-2052 (Sattlegger & Hinnebusch, 2000). It is suggested that Gcn1 may act as a scaffold protein to position Gcn2 for receiving the starvation signal, or it may mediate the transfer for the starvation signal, uncharged tRNA, to Gcn2 (Castilho et al., 2014; Sattlegger & Hinnebusch, 2000). Gcn20, an additional yeast protein, binds Gcn1 at the eEF3-like region (Castilho et al., 2014). However, Gcn20 is not essential for the Gcn1-Gcn2 interaction or Gcn2 activation, and a human homologue has not yet been identified (Castilho et al., 2014). As it has been mentioned, Gcn2 and Yih1 both compete for Gcn1 binding through their shared RWD domains. The region of Gcn1 where these proteins bind is referred to as the RWD-binding domain (RWDBD) (Rakesh et al., 2017; Sattlegger & Hinnebusch, 2000). Furthermore, the highly conserved amino acid residue Arg-2259, which is part of a fully conserved ITGPLIR motif of the RWD domain, is essential for RWD-domain binding (Castilho et al., 2014; Rakesh et al., 2017; Sattlegger & Hinnebusch, 2000).

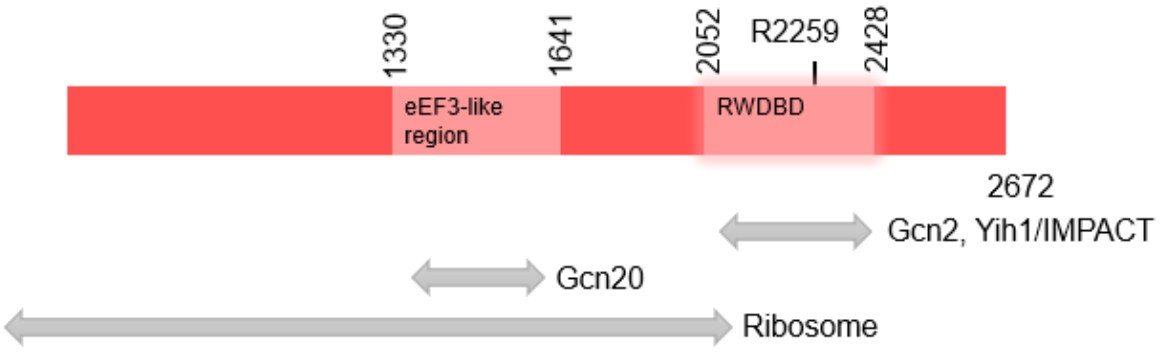


Figure 5: Schematic overview of the Gcn1 structure.

Binding regions of Gcn2, Yih1/IMPACT, Gcn20 and the Ribosome to Gcn1 are indicated by grey arrows. The RWDBD spans amino acids 2052-2428 (Sattlegger & Hinnebusch, 2000) or 2048-2383 (Kubtoa et al., 2000). Figure modified from Castilho et al. (2014).

1.5 Yeast IMPACT homolog 1 (Yih1)/IMPACT is negative regulator of Gcn2 when overexpressed

Yih1 (Yeast IMPACT homolog 1), the yeast homolog of the mammalian protein IMPACT (imprinted and ancient protein) is a negative regulator of Gcn2 (Sattlegger et al., 2011). IMPACT is a paternally imprinted gene that was firstly discovered in mice, however, the function of this protein was not identified (Hagiwara et al., 1997). Overexpression of Yih1 (in yeast) or IMPACT (mammals) impairs the GAAC response, represented by decreased eIF2α phosphorylation and impaired cell growth, indicative of impaired Gcn2 activation (Pereira et al., 2005; Roffé et al., 2013; Sattlegger et al., 2011). Further studies to decipher the role of Yih1/IMPACT have demonstrated that overexpressed Yih1/IMPACT inhibits Gcn2 by directly binding and sequestering Gcn1, the essential binding partner of Gcn2 (Sattlegger et al., 2004; Sattlegger et al., 2011; Waller, Lee, & Sattlegger, 2012; Roffé et al., 2013). In this way, experimental studies support the proposal that Yih1/IMPACT reduces Gcn2-Gcn1 complex formation, causing Gcn2 inhibition (Sattlegger et al., 2004; Sattlegger et al., 2011). As a result, the GAAC response, otherwise known as the Integrated Stress Response in mammals, is impaired.

1.5.1 The Yih1 RWD domain mediates Gcn1 binding

As it was previously discussed, the Gcn2-Gcn1 interaction is critical for Gcn2 activation upon amino acid starved conditions. Like Gcn2, Yih1 and its mammalian homolog IMPACT have an RWD domain in their protein structure that directly interacts with Gcn1 at the same region where Gcn2 binds, called the RWDBD domain (Fig. 6) (Castilho et al., 2014; Pereira et al., 2005; Sattlegger et al., 2004; Sattlegger et al., 2011; Waller et al., 2012). It is implied that through the RWD domain, Yih1/IMPACT competes with Gcn2 for Gcn1 binding, impairing the Gcn2-Gcn1 interaction and consequently preventing Gcn2 activation and the starvation response (Fig. 6) (Pereira et al., 2005; Roffé et al., 2013; Sattlegger et al., 2004; Sattlegger et al., 2011). Supporting this, multiple studies have demonstrated that overexpression of the Yih1 RWD domain alone is sufficient for inhibition of the GAAC response (Kubota et al., 2000; Sattlegger et al., 2011).

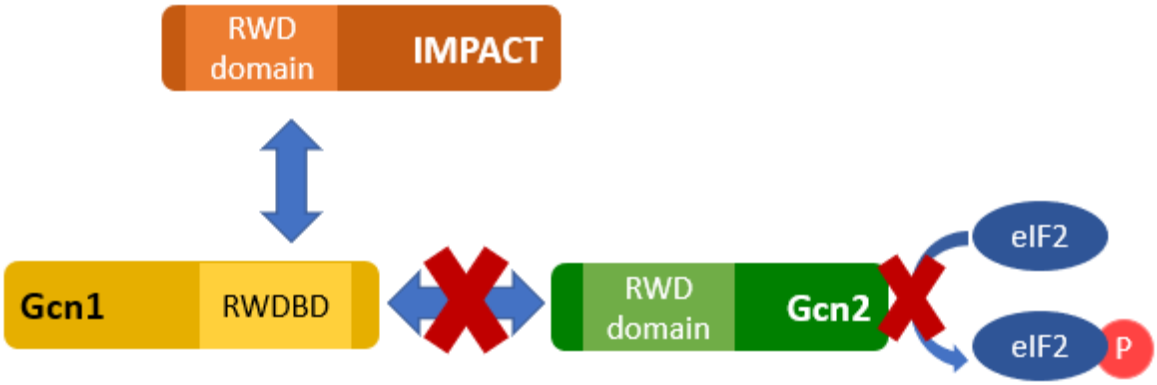


Figure 6: Yih1/1IMPACT is an inhibitor of Gcn2.

Yih1 and Gcn2 both bind to Gcn1 at its RWDBD. Yih1 competes with Gcn2 for Gcn1 binding through its RWD domain. Consequently, the Gcn1-Gcn2 interaction is impaired, meaning Gcn2 cannot be activated and cannot catalyse phosphorylation of its substrate eIF2α at Ser-51. The general amino acid control response is impaired as a result.

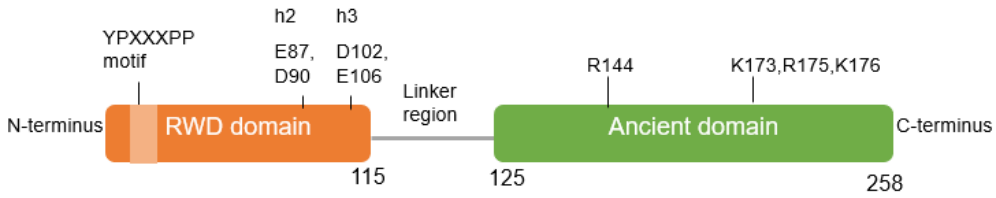
1.5.2 Yih1/IMPACT has an RWD domain and an Ancient domain

Yih1 is a protein with a length of 257 amino acids (Sattlegger et al., 2004; Sattlegger et al., 2011). It is composed of two independent protein domains: an RWD domain and an 'Ancient domain'. These domains are linked by an unstructured flexible linker region (Fig. 7) (Harjes et al., 2021; Nameki et al., 2004; Rakesh et al., 2017; Sattlegger et al., 2011). The RWD domain spans amino acids 1-115 and is situated towards the N-terminal region of Yih1. Structurally, the Yih1 RWD domain is an αβ-sandwich fold consisting of a β-sheet (with 4 β-strands) and 3 α-helices (notated h1, h2 and h3), which is similar to other RWD domain-containing proteins, such as Gcn2 (Sattlegger et al., 2011). A highly conserved YPXXXP motif present in helix 1 (h1) of the RWD domain forms a triple β-turn that is involved in a hydrogen bond network that is important for structural stability of Yih1 (Sattlegger et al., 2011; Nameki et al., 2004). These residues also undergo interactions with the fully conserved Gly-96 residue between helices 2 and 3 (Sattlegger et al., 2011). Overall, the N-terminal region of the Yih1 RWD domain is important for maintenance of the Yih1 protein structure (Sattlegger et al., 2011). A linker region comprised of 26 amino acids, with highly charged residues, connects the RWD domain and the ancient domain (Harjes et al., 2021; Sattlegger et al., 2011). This linker region may provide flexibility for changes in orientation between the two Yih1 domains (Harjes et al., 2021; Sattlegger et al., 2011).

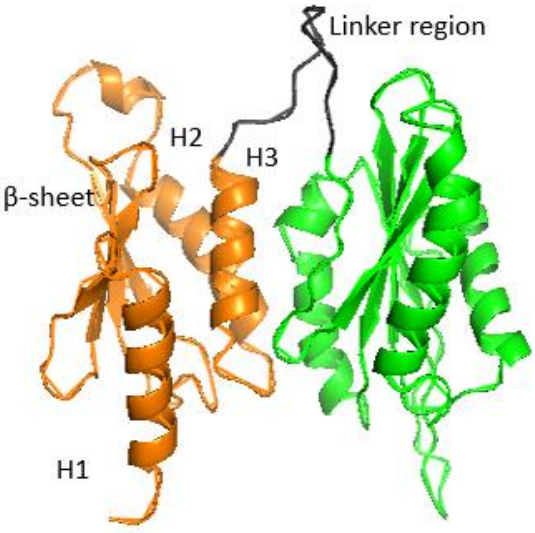
The C-terminal region of Yih1 comprises of the Ancient domain, which is highly conserved among all kingdoms of life, and was therefore named the Ancient domain (Hagiwara et al.,

1997; Sattlegger et al., 2011). The Yih1 Ancient domain has homology to the prokaryotic protein YigZ, which currently has an unknown function (Hagiwara et al., 1997; Harjes et al., 2021; Sattlegger et al., 2011). Structural modelling of the Yih1 Ancient domain, based off the YigZ structure, indicates that it structurally exists as a five-stranded β-sheet "sandwiched" together by one α-helix on one side and two α-helices on the other side of the β-sheet (Sattlegger et al., 2011). Highly conserved residues are present in loop regions of the ancient domain in both Yih1 and YigZ (Harjes et al., 2021; Sattlegger et al., 2011). Although this protein domain has an unknown function in general, as well as in Yih1 function, its high conservation points towards the importance of this domain in the proteins structure and/or function.

A



B



C

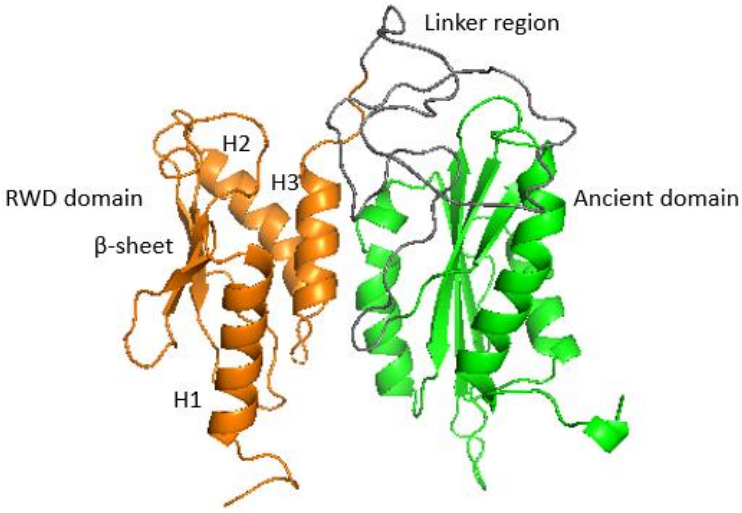


Figure 7: Yih1/IMPACT protein structure.

A: Schematic overview of Yih1 protein structure. The RWD domain is coloured in orange. The ancient domain is coloured in green. The important amino acid residues and motifs are labelled. B: Structural ensemble of Yih1 created with PyMOL Molecular Graphics System, Version 1.8. The RWD domain, with helices and β-sheet labelled, are represented in orange. The linker region which connects the two domains is represented in grey. The ancient domain

is represented in green. Important protein structural characteristics are labelled. C: Structural ensemble of IMPACT created with PyMOL Molecular Graphics System, Version 1.8. RWD domain is represented in orange. The linker region which connects the two domains is represented in grey. The ancient domain is represented in green. Important protein structural characteristics are labelled.

1.5.3 Yih1-mediated Gcn2 inhibition must be regulated

A *YIH1* deletion does not result in increased Gcn2 activity (Sattlegger et al., 2004). Furthermore, Yih1/IMPACT expressed at native levels does not impair the GAAC response (Sattlegger et al., 2004). Therefore, rather than acting as a general regulator, it is suggested that Yih1 only inhibits Gcn2 under certain circumstances (Sattlegger et al., 2004; Sattlegger et al., 2011). This implies that Yih1 must be mediated by additional factors such as intraprotein or interprotein interactions and/or by its binding partners (Sattlegger et al., 2004; Sattlegger et al., 2011). Furthermore, this also raises the possibility of intramolecular interactions that influence the ability of Yih1 to contact its binding partners and regulate Gcn2.

1.6 Yih1/IMPACT-protein interactions: Towards the function of Yih1/IMPACT

1.6.1 Yih1/IMPACT-Gcn1 interaction impairs the GAAC/ISR response

Yih1 directly binds Gcn1, and the direct Yih1-Gcn1 interaction has been supported experimentally by a yeast two-hybrid assay and co-precipitation experiments (Kubota et al., 2000; Sattlegger et al., 2004; Sattlegger et al., 2011). The Yih1-Gcn1 interaction was associated with an impaired GAAC response, indicating that Gcn2 was impaired by the Yih1-Gcn1 interaction (Sattlegger et al., 2004; Sattlegger et al., 2011). The direct IMPACT-Gcn1 interaction has also been experimentally supported, and this also associated with an impaired GAAC response (Pereira et al., 2004; Waller et al., 2012). Therefore, the Gcn1-binding activity of Yih1/IMPACT family proteins is conserved from yeast to mammals.

1.6.2 Mapping the Yih1 binding site of Gcn1

To further understand the interaction between Yih1/IMPACT and Gcn1 better, the region of Gcn1 where Yih1/IMPACT and Gcn2 directly bind was mapped. Yih1/IMPACT and Gcn2 both directly interact with Gcn1 at the Gcn1 RWDBD (Sattlegger et al., 2004; Sattlegger et al., 2011). Furthermore, binding of Yih1/IMPACT and Gcn2 to the RWDBD is achieved through their respective RWD domains (see Fig. 6) (Sattlegger et al., 2004; Sattlegger et al., 2011). Like Gcn2, the Yih1/IMPACT-Gcn1 interaction is also mediated by a single amino acid residue, Arg2259 (R2259), in a highly conserved segment of the Gcn1 RWDBD (Sattlegger et al., 2004). This was supported experimentally as a R2259A substitution abolishes the Yih1-Gcn1 interaction (Sattlegger et al., 2004). Therefore, Yih1/IMPACT and Gcn2 share key essential binding determinants of Gcn1, supporting the implication that there is competition between these two proteins for Gcn1 binding. It is strongly supported that through their respective RWD domains, Yih1/IMPACT competes with Gcn2 for Gcn1 binding (Fig. 6) (Cambiaghi et al., 2014; Pereira et al., 2005; Sattlegger et al., 2011; Silva et al., 2016).

1.6.3 Yih1 amino acids D102 and E106 are important for Gcn1 binding

The important Gcn1-binding determinants of Yih1 are suggested to be within amino acids 74-114, in which helix 2 (h2) and helix 3 (h3) are located (Fig. 7) (Sattlegger et al., 2011). This was based on the finding that while a Yih1 fragment consisting of the RWD domain and part of the ancient domain (amino acids 1-132) was sufficient for the Yih1-Gcn1 interaction, Yih1 amino acids 68–171 and 68–258 can also bind Gcn1 (Sattlegger et al., 2011). In saying this, the minimal Yih1 region for Gcn1 binding was mapped to amino acids 68-171, in which h2 and h3 of the Yih1 RWD domain and part of the ancient domain are located (Fig. 8) (Sattlegger et al., 2011). Furthermore, a Yih1 fragment truncated of N-terminal RWD amino acids 1-67 binds Gcn1 with greater affinity than full length Yih1, suggesting that Yih1 amino acids 1-67

restricts the accessibility of Yih1 to Gcn1 (Sattlegger et al., 2011). This is supported by finding addressing the structural organisation of RWD domains, as the conserved YPXXXP motif, important for the stability of protein structure, is within amino acids 1-67 (Sattlegger et al., 2011).

It is important to understand the binding determinants of the Yih1-Gcn1 interaction in more detail to understand how Yih1 inhibits Gcn2. By analysing the Yih1 structural model for similar residues of Gcn2 that are involved in Gcn1 binding, Yih1 amino acids Asp-102 (D102) and Glu-106 (E106), two surface-exposed negatively charged residues situated in H3 of the Yih1 RWD domain have been identified (Fig. 7) (Sattlegger et al., 2011). D102 and E106 occupy similar positions and charges to Gcn2 amino acids Glu-125 and Glu-136, respectively. These Gcn2 amino acids are involved in Gcn1 binding, suggesting that Yih1 D102 and E106 may be involved in Gcn1 binding (Castilho et al., 2014; Sattlegger et al., 2011). Substitution of both D102 and E106 to alanine (*GST-YIH1*H3*) results in an impaired GAAC response along with reduced levels of Gcn1 sequestration upon starvation (Sattlegger et al., 2011). Accordingly, these findings imply that D102 and E106, situated in helix H3 of the Yih1 RWD domain, are important for Gcn1 binding and for impairing Gcn2 activation *in vivo* (Sattlegger et al., 2011).

1.6.4 Structural modelling of Yih1: Insight into Yih1-mediated Gcn2 inhibition

A recently published study has determined the structure of Yih1, enabling a more comprehensive understanding of the potential mechanism by which Yih1 inhibits Gcn2 (Harjes et al., 2021). Biophysical techniques demonstrated that Yih1 exists as a tight compact structure that is maintained by interdomain interactions (Harjes et al., 2021). The amino acid residues that are important for Gcn1 binding are buried away within the Yih1 protein structure (Harjes et al., 2021). Therefore, Yih1 must undergo a conformational change in order to be primed for the Yih1-Gcn1 interaction and therefore for Gcn2 inhibition (Harjes et al., 2021). It was further implied that the Yih1 domains are maintained in a 'closed' position by interdomain electrostatic interactions formed by amino acid residues identified at the interface of the two domains (Harjes et al., 2021). Together with the fact that Yih1 is not a general regulator of Gcn2, it is strongly suggested that Yih1 must undergo conformational rearrangements, from a closed 'inactive' form to an open 'active' form in order to be able to inhibit Gcn2 (Harjes et al., 2021; Sattlegger et al., 2004; Sattlegger et al., 2011).

1.6.5 Yih1 interacts with actin at native levels

In addition to Gcn1, Yih1 interacts with other proteins that may modulate its ability to inhibit Gcn2. Actin is a crucial protein in eukaryotes as it polymerises to form the cellular actin

cytoskeleton, which is a dynamic structure that undergoes constant assembly and reassembly with respect to the cellular state (Lee & Dominguez, 2010; Ramesh et al., 2021; Waller et al., 2012). Yih1 interacts with monomeric actin (G-actin), and this has been experimentally supported by independent lines of research (Sattlegger et al., 2004; Sattlegger et al., 2011; Waller et al., 2012). The Yih1-actin interaction was first isolated by co-precipitation assays, whereby native purified FLAG-tagged Yih1 was found in a 1:1 complex with actin (*ACT1*) (Sattlegger et al., 2004). Furthermore, this interaction between endogenous Yih1 and actin was independent of Gcn1 (Sattlegger et al., 2004). It has not yet been investigated whether Yih1 binds to F-actin (Filamentous actin) as well as G-actin.

1.6.6 IMPACT, the mammalian homolog of Yih1, also interacts with actin

As IMPACT binds Gcn1 and is a negative regulator of Gcn2 in mammalian cell lines, analogous to the activity of Yih1 in yeast, it was proposed that IMPACT may also interact with actin (Pereira et al., 2005; Waller et al., 2012). In yeast strains overexpressing IMPACT with Yih1 deleted (*yih1Δ*) or Yih1 and Gcn1 both deleted (*yih1Δgcn1Δ*), IMPACT coprecipitates actin (Waller et al., 2012). Therefore, IMPACT also interacts with actin, and that the IMPACT-actin interaction is independent of Gcn1, rather than being an indirect interaction that occurs through IMPACT-Gcn1 binding (Waller et al., 2012). Furthermore, the conserved functional homology between Yih1 and IMPACT suggests that this interaction has an important role in the function of Yih1/IMPACT.

1.6.7 The IMPACT/Yih1-Actin interaction may prevent the inhibition of Gcn2

As it was mentioned previously, Yih1 is not a general negative regulator of the GAAC response at endogenous levels (Sattlegger et al., 2004). Endogenous Yih1 exists in a complex with G-actin (Sattlegger et al., 2004; Silva et al., 2015; Silva et al., 2016). Furthermore, a yeast strain with reduced actin levels (*act1Δ/ACT1*) overexpressing Yih1 experienced an impaired GAAC response, indicative of Gcn2 inhibition (Sattlegger et al., 2004). Moreover, deletion of Yih1 in the *act1Δ/ACT1* strain suppressed the impairment of the GAAC response (Sattlegger et al., 2004). This implies that the Yih1-actin interaction mediates Gcn2 activation and therefore the GAAC response (Sattlegger et al., 2004). In addition, no morphological differences to actin were observed when Yih1 was either deleted or overexpressed, revoking the possibility that Yih1 associates with actin to change the actin morphology, such as its polymerisation state (Sattlegger et al., 2004).

Overall, experimental data from protein purification and genetic techniques and implies that endogenous Yih1 exists bound to G-actin, and that the release of Yih1 from actin allows Yih1

to impair Gcn2 activation by competing for Gcn1 binding (Sattlegger et al., 2004; Sattlegger et al., 2011; Silva et al., 2016). Therefore, the Yih1-actin interaction may prevent the inhibition of Gcn2 activation by Yih1 (Gasparski et al., 2022; Sattlegger et al., 2011; Silva et al., 2016).

It is possible to speculate that the Yih1-actin interaction occurs under "normal" physiological conditions to prevent Gcn2 inhibition when it is not necessary (Harjes et al., 2021; Sattlegger et al., 2004; Sattlegger et al., 2011; Silva et al., 2015; Silva et al., 2016). In this way, there is possible spatial and temporal regulation of Gcn2 by Yih1. Therefore, a potential mechanism whereby Yih1 is involved in spatial and/or temporal crosstalk between translation and the cellular actin cytoskeleton may be speculated, and this is experimentally supported (Harjes et al., 2021; Sattlegger et al., 2004). Furthermore, considering that the actin binding activity of Yih1/IMPACT is evolutionarily conserved among eukaryotes, this suggests that the Yih1/IMPACT-actin interaction may have an important role in Yih1/IMPACT functionality (Waller et al., 2012).

1.6.8 Yih1-Ribosome interaction: Potential roles in translation regulation

Gcn1 and Gcn2 both reside on the ribosome in a complex, allowing for quick detection of the starvation signal and therefore efficient and fast activation or inhibition of Gcn2 (Waller et al., 2012). This is crucial as Gcn2 as involved in essential cellular processes such as translational control and regulation of protein synthesis. Considering this, it was proposed that Yih1 may be somehow involved in mediating the interaction between these proteins and the ribosome. Accordingly, it has been found that Yih1 also interacts with the ribosome (Waller et al., 2012). The Yih1-ribosomal interaction was first demonstrated by a co-sedimentation assay. In this type of assay, exponentially growing cells are treated with formaldehyde to induce crosslinking, and then velocity sedimentation is performed through a sucrose gradient. This allows for real-time detection of protein-protein and/or protein-nucleic acid interactions in physiological conditions (Waller et al., 2012). The fractions are then subjected to immunoblotting to detect, in this case, which fraction Yih1 was co-sedimented. In this study, researchers detected Yih1 in the heavy fractions where polyribosomes are expected, suggesting that Yih1 co-sedimented with the polyribosomes (Waller et al., 2012). Further investigations using co-precipitation, to analyse whether Yih1 was interacting with ribosomes and not just a protein of similar size, confirmed that Yih1 was directly interacting with polyribosomes, specifically with Rpl39 (large ribosomal protein 39) (Waller et al., 2012).

1.6.9 Yih1-Ribosome interaction is independent of Gcn1

As Gcn1 binds to the ribosome, the possibility was raised that Gcn1 could be bridging the interaction between Yih1 and the ribosome (Waller et al., 2012). To investigate this, the co-sedimentation assay was repeated using a Gcn1 deletion strain (*gcn1Δ*) overexpressing Yih1, and then using a Yih1 and Gcn1 double deletion strain (*yih1Δgcn1Δ*). The latter strain was used to exclude the possibility of endogenous Yih1 binding the ribosome. In both circumstances, Yih1 was detected co-sedimented in heavy fractions with polyribosomes (Waller et al., 2012). Importantly, this strongly suggests that the interaction between Yih1 and the ribosome is independent of Gcn1 (Waller et al., 2012). Furthermore, it demonstrates that both overexpressed and endogenous Yih1 physically associate with polyribosomes.

1.6.10 IMPACT also interacts with the Ribosome

Yih1 and IMPACT seem to have functional homologies, as they both are capable of inhibiting Gcn2 activation when overexpressed (Pereira et al., 2005; Sattlegger et al., 2004). As Yih1/IMPACT both interact with Gcn1 and actin, it was relevant to test whether IMPACT is also capable of binding the ribosome. Co-sedimentation assays using GST-IMPACT showed that IMPACT was also discovered co-sedimented in heavier fractions with yeast polyribosomes (Waller et al., 2012). To further decipher whether IMPACT was physically binding to the ribosome, authors then carried out a co-precipitation experiment, similar to that done with Yih1, in both a *yih1Δgcn1Δ* double deletion strain, and a *yih1Δ* strain. It was found that IMPACT does indeed co-precipitate in a complex with yeast ribosomal complexes, specifically Rpl39 and Rps22, and this interaction was also independent of Gcn1 (Waller et al., 2012).

The finding that Yih1 and IMPACT are both capable of binding the ribosome suggests that the ribosome binding activity of Yih1/IMPACT family proteins is evolutionarily conserved. Therefore, the Yih1/IMPACT-Ribosomal interaction may be an important component of the function of Yih1/IMPACT in Gcn2 inhibition.

1.6.11 Yih1/IMPACT interacts with cell-cycle dependent proteins Cdc28/CDK1

Pulldown experiments and BiFC (Bimolecular fluorescence complementation) techniques show that native Yih1 physically interacts with Cdc28, a protein that is essential for cell cycle progression in budding yeast (Silva et al., 2015). Furthermore, cells with a *YIH1* deletion experience a delay in cell cycle progression, particularly in G2/M cell cycle phases (Silva et al., 2015). This delay was reverted in wild type cells expressing Yih1 (Silva et al., 2015). IMPACT was also shown to physically interact with CDK1, the mammalian homologue of

Cdc28 (Silva et al., 2015). This suggests that this interaction is evolutionarily conserved and therefore may underlie a crucial physiological role of Yih1/IMPACT in the cell cycle.

1.6.12 Yih1 regions for Gcn1, Actin, and Ribosomal Binding overlap

Yih1 has various binding partners, such as Gcn1, actin, ribosomal proteins, and Cdc28. To gain more insight into how these proteins may modulate the function of Yih1 in Gcn2 inhibition, the minimal regions for binding of these proteins to Yih1 have been mapped. The minimal region of Yih1 for actin binding was mapped to amino acids 68-258, harbouring part of the RWD domain and the ancient domain (Fig. 8) (Sattlegger et al., 2011; Waller et al., 2012). Furthermore, as is true for the Yih1-Gcn1 interaction, N-terminal amino acids 1-67 appear to hinder the Yih1-actin interaction (Sattlegger et al., 2011). This is consistent with the finding that amino acids 1-67 (N terminal RWD domain) are important for maintaining the structure of Yih1 (Sattlegger et al., 2011). Therefore, the actin and Gcn1 binding sites on Yih1 overlap by amino acids 68-171 (Fig. 8). In addition, the Yih1-actin interaction is independent of Gcn1 (Sattlegger et al., 2004). Yih1 with mutations to amino acids proposed to be involved in Gcn1 binding, D102 and E106, experiences reduced actin binding (Sattlegger et at., 2011). On the other hand, Yih1 with substitutions to E87 and D90, amino acids proposed to be involved in interdomain interactions, experiences increased actin binding (Sattlegger et al., 2011). Together, these results imply that Gcn1 and actin have different binding determinants on Yih1.

The minimal Yih1 region for ribosome binding was mapped to amino acids 68-171, and this interaction is independent of Gcn1 (Waller et al., 2012). As this is also the minimal Yih1 region required for Gcn1 binding and actin binding, it is strongly implied that these proteins must have different binding determinants (Sattlegger et al., 2011). In conclusion, these interactions seem to occur independently of each other as their binding sites overlap.

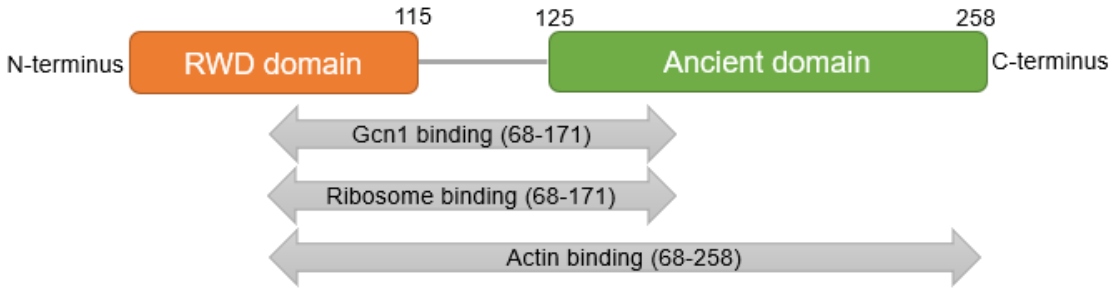


Figure 8: The minimal binding regions of Yih1 for Gcn1, the ribosome and actin.

See text for details.

1.7 Proposed mechanism of Yih1/IMPACT-mediated Gcn2 inhibition

As mentioned previously, Yih1 is not a general regulator of Gcn2, as deletion of *YIH1* does not result in increased Gcn2 activity (Ramesh et al., 2021; Sattlegger et al., 2011). Furthermore, as shown in Fig. 8, the Yih1 binding sites for Gcn1, actin, and the ribosome overlap. However, these interactions seem to be mediated by different binding determinants as the interaction with Yih1 and its binding partners are independent. Regulatory mechanisms must be in place to ensure that Yih1 does not constantly inhibit Gcn2, but rather only under certain cellular conditions where/when appropriate (Akram et al., 2020; Harjes et al., 2021; Ramesh et al., 2021; Sattlegger et al., 2011).

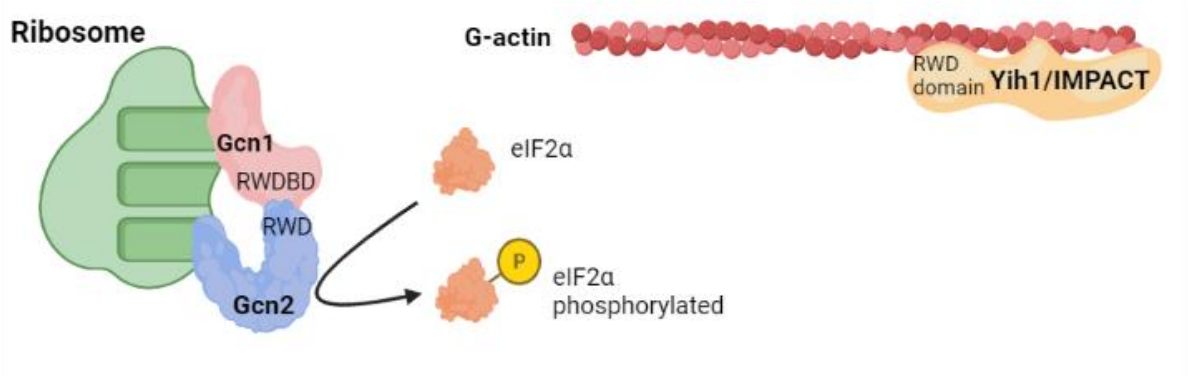
Based on the experimental evidence, a model has been proposed to address how Yih1 impairs Gcn2 activation. This current working model suggests that under certain cellular conditions, Yih1 exists in a complex with G-actin, keeping it inactive and thereby allowing Gcn2 activation (Gasparski et al., 2022; Sattlegger et al., 2011; Silva et al., 2016). Under these conditions, active Gcn2 can phosphorylate eIF2α, which results in a general reduction in protein synthesis and specific upregulation of certain transcription factors. An unknown stimulus/signal may then result in the release of Yih1 from actin allowing Yih1 to shuttle to the ribosome, where Gcn1 and Gcn2 both reside (Marton et al., 1997; Sattlegger & Hinnebusch, 2000). This would enable close physical proximity for Yih1 to compete with Gcn2 for Gcn1 binding and therefore inhibiting Gcn2. Therefore, this working model suggests that Yih1 may shuttle between actin and the ribosome for modulation of Gcn2 activity. This would allow for spatial and temporal modulation of protein synthesis, which is crucial for the regulation of cellular processes such as cell differentiation (Akram et al., 2020; Sattlegger et al., 2011; Silva et al., 2016; Waller et al., 2012).

In addition to this model, Yih1 must exist within the cell in an ‘active’ form and in an ‘inactive’ form. Evidence suggests that Yih1 must undergo a conformational change to become “opened” up and primed for Gcn2 inhibition. One significant finding implying this is that the important residues for Gcn1 binding are buried at the Yih1 domain interface (Harjes et al., 2021; Sattlegger et al., 2011). Therefore, the conformation of Yih1 may modulate its ability to inhibit Gcn2. However, the connection between the Yih1 conformational change and the mechanism of Yih1-mediated Gcn2 inhibition is still unclear. It is implied that Yih1 is held in a “closed” conformation by interdomain interactions (Fig, 9C, left panel) (Harjes et al., 2021; Waller et al., 2012). Upon disruption of these interactions (by currently unknown signals), Yih1 can

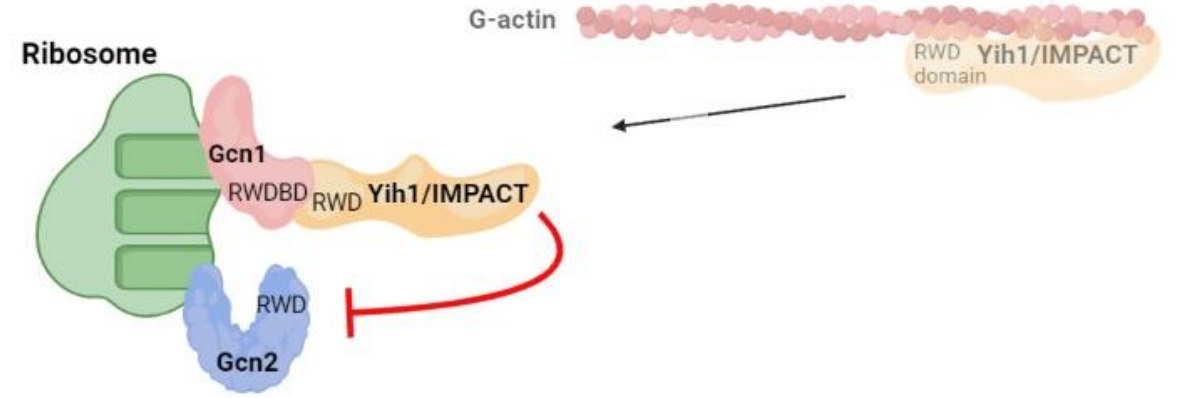
“open” up, exposing the important Gcn1 binding determinants and thereby binding Gcn1 and inhibiting Gcn2 (Fig. 9C, right panel) (Harjes et al., 2021; Waller et al., 2012).

The Yih1-ribosome interaction can occur independently of Gcn1 (Waller et al., 2012). Therefore, it is unknown whether Yih1 is bound to the ribosome while it inhibits Gcn1, or whether Yih1 only interacts with Gcn1, which resides on the ribosome (Waller et al., 2012). Furthermore, the conditions and/or signal that results in the release of Yih1 from actin is unclear. It is possible to speculate that the polymerisation state of actin controls this (Gasparski et al., 2022). It is also possible that the cellular starvation state influences the localisation of Yih1.

A



B



C

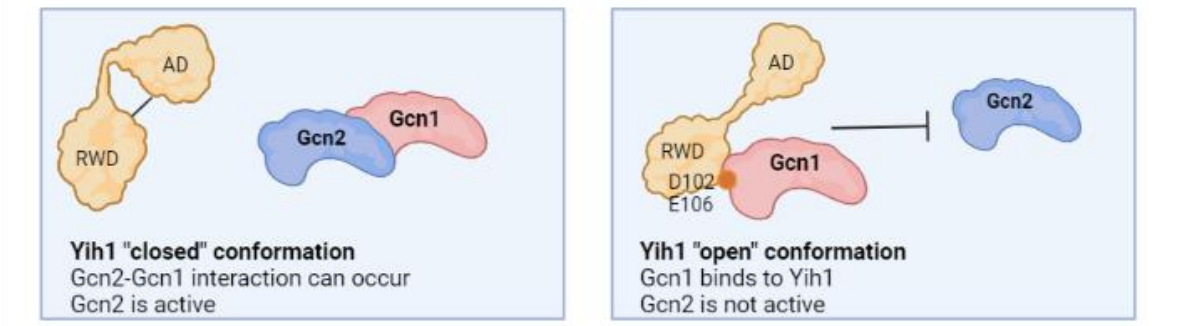


Figure 9: The current working model of Yih1/IMPACT mediated Gcn2 inhibition.

A: Yih1 resides in a complex with G-actin, which prevents it from inhibiting Gcn2. Active Gcn2 is therefore able to catalyse phosphorylation of eIF2α, resulting in a general reduction in protein synthesis and specific upregulation of certain transcription factors. B: Yih1 is released from G-actin, where it can compete with Gcn2 for Gcn1 binding and inhibit Gcn2 activation. Yih1 either binds to the ribosome while it inhibits Gcn2 or binds Gcn1 alone. In either case, Gcn2 activation is impair, leading to reduced eIF2α phosphorylation. C: Yih1 is held in a "closed" conformation by interdomain interactions (left panel). Upon disruption of

these interactions (by currently unknown signals), Yih1 can “open” up, exposing the important Gcn1 binding determinants, allowing Yih1 to bind Gcn1 and therefore inhibit Gcn2. The connection between the Yih1 conformational change and the mechanism of Yih1-mediated Gcn2 regulation is unclear. Figure is not drawn to scale. Figure was generated in BioRender programme.

1.8. Biological implications of Yih1/IMPACT

1.8.1 The Gcn2-Yih1/IMPACT axis has a major role in mammalian neurophysiology and neural development

IMPACT is expressed at high levels in neurons, particularly in hypothalamic neurons (Pereira et al., 2005). This is further emphasized by findings that the highest IMPACT mRNA levels are found in the mouse brain (Hagiwara et al., 1997). An inverse correlation has been established between IMPACT expression and Gcn2 activity in mammalian neurons (Pereira et al., 2005). High expression levels of IMPACT correlate with reduced Gcn2 activity and subsequently eIF2 phosphorylation levels in neurons during neuronal differentiation (Bittencourt et al., 2008; Pereira et al., 2005; Roffé et al., 2013). Furthermore, IMPACT stimulates neurite outgrowth/neuritogenesis, while Gcn2 inhibits this process (Roffé et al., 2013). Based off these findings, an implication of IMPACT in these neural developmental processes has been established. It is also likely that the role of IMPACT in these processes is through modulation of Gcn2 activity. Hypothetically, in brain areas that require high levels of translation, such as during neuronal differentiation, IMPACT inhibits Gcn2 activity, thereby allowing high levels of general translation/protein synthesis (Roffé et al., 2013). However, in regions where this is not required, Gcn2 is activated meaning that general translation is impaired (Roffé et al., 2013). As IMPACT also interacts with actin, it is possible that crosstalk between the actin cytoskeleton in the brain and protein synthesis is occurring in these areas. The interaction between IMPACT and actin is potentially involved in spatiotemporal modulation of protein synthesis for processes such as neurite outgrowth and neural differentiation (Silva et al., 2016; Waller et al., 2012).

1.8.2 IMPACT is implicated in neurodevelopmental disorders and diseases

Because of the involvement of IMPACT in neurological development, Yih1/IMPACT is strongly implicated in neurodevelopmental disorders and diseases, such as Alzheimer's disease (Bittencourt et al., 2008; Pereira et al., 2005; Roffé et al., 2013; Sattlegger et al., 2011). Because of the relevance of IMPACT in mammals, understanding the function and molecular mechanisms of Yih1/IMPACT will allow insight into treatment and/or prevention of these neurodevelopmental disorders and diseases.

1.8.3 IMPACT may modulate cancer cell survival

Yih1 and IMPACT interact with Cdc28 and CDK1, respectively (Silva et al., 2015). These proteins are centrally involved in cell cycle progression in eukaryotes. It has been predicted that Yih1/IMPACT are involved in modulation of the cell cycle and in cell proliferation, presumably through their activity with Cdc28/CDK1 (Silva et al., 2015; Tomek et al., 2018).

Therefore, the role of IMPACT may have implications in cancer cell survival by enabling cancer cells to survive in different stress environments, such as amino acid starvation (Silva et al., 2015; Tomek et al., 2018). Accordingly, IMPACT expression is high in a range of mammalian cancer cells (Tomek et al., 2018). Furthermore, this is associated with reduced activity of stress response pathways and a specific enhanced activity of pathways involved in translational regulation, such as the mTOR pathway (Tomek et al., 2018). Overall, IMPACT has a central role in cell cycle progression, which is further implicated in cancer cell survival and proliferation.

1.8.4 Physiological implication of the Yih1/IMPACT-Gcn2 axis

The important biological and physiological implications of Yih1/IMPACT are also emphasized in Gcn2 regulation. Gcn2 is centrally involved in normal and abnormal cell function, along with the maintenance of eukaryotic cell homeostasis during stress (Masson, 2019). Furthermore, Gcn2 is implicated in many biological processes such as immune system modulation, amino acid homeostasis, and neural functions such as memory formation (Castilho et al., 2014; Pereira et al., 2005). Additionally, various diseases and disorders have also been associated with Gcn2, such as cancer (Castilho et al., 2014). Understanding the regulation of Gcn2 by Yih1/IMPACT may provide insight for a therapeutic target for pharmacological drug development for treating or preventing diseases and disorders that are associated with Gcn2 (Castilho et al., 2014).

1.9. Hypothesis and Aims

Yih1/IMPACT only inhibits Gcn2 under certain circumstances rather than acting as a general inhibitor (Bittencourt et al., 2008; Pereira et al., 2005; Sattlegger et al., 2011). This implies that Yih1 itself must be regulated regarding its activity in Gcn2 inhibition. Furthermore, this means that Yih1/IMPACT can reside in the cell in an inactive form or in an active form that is primed for Gcn2 inhibition. However, the precise molecular interactions between Yih1 and Gcn1 are not known yet, nor is it known how intramolecular interactions determine when Yih1 is in its inactive or active form. To gain a proper understanding of how Yih1 inhibits Gcn2, we need to know in detail the molecular mechanisms involved in Yih1/IMPACT function. Our study contributes to understanding the molecular mechanisms behind the regulation of Gcn2 activation by Yih1/IMPACT. This work aims to identify the relevant amino acids of Yih1 that are involved in inhibition of Gcn2 and the starvation response, using Baker's yeast (*S. cerevisiae*) as a model organism. The following hypotheses were proposed;

Hypothesis 1: Specific Yih1 amino acids are involved in Gcn1 binding to modulate Gcn2 inhibition.

Hypothesis 2: Specific Yih1 amino acids are involved in interdomain interactions which affect the ability of Yih1 to modulate Gcn2 activity.

Hypothesis 3: Additional regulatory mechanisms are in place that modulate the ability of Yih1 to inhibit Gcn2.

To test these hypotheses and meet our aim of gaining a better understanding of the function of Yih1 in the context of Gcn2 inhibition, the objectives are;

1. Establish a yeast system to score for Yih1-mediated Gcn2 inhibition

   As mentioned, yeast cell growth is an indicator of Gcn2 activation levels. In this way, yeast cell growth is inversely correlated to Yih1 activation levels. We can therefore utilise indicators of Gcn2 activation; yeast cell growth and eIF2 phosphorylation, to determine the activation levels of Yih1. We aim to establish a yeast system to score for the ability of Yih1 to inhibit Gcn2.

2. Identify Yih1 amino acid(s) that are relevant for Gcn1 binding

   Identifying relevant Yih1 amino acids for Gcn1 binding is important as the Yih1-Gcn1 interaction is essential for Gcn2 inhibition. To test this hypothesis, the effect of

overexpressed Yih1 with substitutions to the proposed residues involved in Gcn1 binding on Gcn2 activation will be assessed using our established yeast system.

3. Identify Yih1 amino acids that are relevant for interdomain interactions

In the cell, Yih1 exists in a tight compact structure, with the Gcn1 binding determinants buried at the protein interface. Therefore, Yih1 must undergo conformational changes to become primed for Gcn1 binding. However, it is still unclear how intramolecular interactions determine when Yih1 is in its inactive or active form. To test this hypothesis the effect of overexpressed Yih1, with substitutions to proposed residues involved in interdomain interactions, on Gcn2 activation will be analysed using our established yeast system.

4. Identify Yih1 amino acids potentially involved in additional means of regulation

The potential involvement of additional regulatory mechanisms that may modulate Yih1 activity, such as Arginine methylation, has been implied by recent findings (Harjes et al., 2021). However, the connection between this potential modulation and Gcn2 inhibition is still unclear. We hypothesise that additional regulatory mechanisms are in place that modulate the ability of Yih1 to inhibit Gcn2. To test this hypothesis, the effect of overexpressed Yih1 with substitutions to residues proposed to be involved in protein modulation, on Gcn2 activation will be analysed using our established yeast system.

2. Materials and Methods

2.1.1 Yeast strains and plasmids used

Yeast strains and plasmids that were used in this study are described in Table 2.1 and Table 2.2.

Table 1: Yeast strains used in this study

Strain	**Genotype**	**Source**
H1511	MATα, *ura3-52, trp1-63, leu2-3, leu2-112, GAL2*$^{+}$	Foiani et al., (1991)
H2256	*gcn1Δ*	Sattlegger & Hinnebusch, 2000
ESY11001b	*yih1Δ*	Sattlegger et al. (2011)

Table 2: Plasmids used in this study

Plasmids	**Gene**	**Selectable marker**	**Vector**	**Source**
pES128-9	GST alone	*Amp^R, URA3*	pEG(KT)	Sattlegger & Hinnebusch, 2000
pES187-B1	GST-Yih1 (2-258)	*Amp^R, URA3*	pES128-9	Sattlegger et al., 2004
pES234-6-2	GST-IMPACT	*Amp^R, URA3*	pES128-9	Pereira et al., 2004
pES332-3-1	GST-YIH1(2-258)-E87A,D90A,D102A,E106A	*Amp^R, URA3*	pES128-9	Sattlegger et al. (2011)
	GST-Yih1(2-258)-D102A	*Amp^R, URA3*	pES128-9	This study
	GST-Yih1(2-258)-E106A	*Amp^R, URA3*	pES128-9	This study
	GST-Yih1(2-258)-K173A,R175A,K176A	*Amp^R, URA3*	pES128-9	This study
	Gst-Yih1(2-258)-K173E,R175E,K176E	*Amp^R, URA3*	pES128-9	This study
	GST-Yih1(2-258)-R175A	*Amp^R, URA3*	pES128-9	This study
	GST-Yih1(2-258)-R175E	*Amp^R, URA3*	pES128-9	This study
pES330-5-3	GST-Yih1(2-258)-ED-87,90-AA	*Amp^R, URA3*	pES128-9	This study
	GST-Yih1(2-258)-ED-87,90-RR	*Amp^R, URA3*	pES128-9	This study
	GST-Yih1(2-258)-E-87-K	*Amp^R, URA3*	pES128-9	This study
	GST-Yih1(2-258)-ED-87,90-KK	*Amp^R, URA3*	pES128-9	This study
	GST-Yih1(2-258)-R144A			This study
	GST-Yih1(2-258)-R144K			This study

2.1.2 Media used

Unless otherwise stated, all media was prepared using double de-ionised water (MilliQ) and made up to 1L. Media was sterilised by either autoclaving at 121°C at 15psi or filter sterilized through a 0.2 µm cellulose acetate filter (Sartorius Stedim Biotech).

Synthetic Defined (SD) medium

Table 3: Composition of Synthetic Defined (SD) medium

Components	**Amount**
Yeast nitrogen base (YNB) (FORMEDIUM™)	1.9g
Ammonium sulphate (FORMEDIUM™)	5g
Agar FORMEDIUM™ **	20g
Double de-ionised water (MilliQ)	950mL

**Agar for solid media only

YNB (1.9g) and Ammonium Sulfate (5g) (and Agar for solid media) was made up to 950mL with H20 and sterilised by autoclaving. 50mL of 40% glucose, 20mL of ILV stock solution, and 10mL of Tryptophan stock solution was added in sterile conditions. This is referred to as SDWILV + glucose in this thesis.

Yeast Peptone Dextrose (YPD) medium

Table 4: Composition of Yeast Peptone Dextrose (YPD) medium

Components	**Amount**
Yeast extract (FORMEDIUM™)	10g
Peptone (FORMEDIUM™)	20g
Agar FORMEDIUM™ **	20g
Double de-ionised water (MilliQ)	950mL

**Agar for solid media only

Yeast extract (10g) and Peptone (20g) (and Agar for solid media) was dissolved in 950mL MilliQ water and sterilised by autoclaving. Just before use, 50mL of 40% glucose (2% w/v) was added to the media.

Luria Bertani media

Table 5: Composition of Luria Bertani media

Components	Amount
Bacto tryptone (FORMEDIUM^{TM})	10g
Yeast extract (FORMEDIUM^{TM})	5g
Sodium chloride (Ajax)	10g
Agar FORMEDIUM^{TM} **	20g
Double de-ionised water (MilliQ)	950mL

**Agar for solid media only

Bacto tryptone (10g), Yeast extract (5g), NaCl (10g) (and 20g Agar for solid media) was dissolved in up to 1L MilliQ water. Just before use, 1mL Ampicillin (50µg/mL) was added to the medium.

Liquid media

Liquid media was dissolved and sterilized by autoclaving (121°C and 15 psi for 20 minutes). Media supplements were added under sterile conditions after liquid media had cooled down to approximately 60°C. Media was stored at room temperature.

Solid media

Solid media was prepared by adding agar to the media, which was then sterilized by autoclaving (121°C and 15 psi for 20 minutes). Once cooled to approximately 60°C, media supplements were added and poured (distributed) into to petri dish plates. Plates were left to solidify at room temperature overnight (at least), and then were stored at 4°C.

2.1.3 Media supplements

Glucose/Galactose stock solutions

40% weight by volume Glucose or Galactose was dissolved in water and sterilised by autoclaving at 121°C at 15 psi and stored at room temperature.

Amino acid stock solutions

Amino acid isoleucine, leucine, valine mixture solution (ILV) was prepared using double de-ionised water (MilliQ), sterilised by autoclaving at 121°C at 15 psi, and stored at room temperature. Tryptophan (W) (0.8g/100mL) was filter sterilized through a 0.2 µm cellulose acetate filter (Sartorius Stedim Biotech), wrapped in aluminium foil, and stored at 4°C. ILV (and W) amino acid mixture was added to media under sterile conditions, and this media supplement is referred to as WILV.

Table 6: Components of ILV amino acid stock mixture

Amino acid	Amount
Isoleucine	0.656g
Leucine	2.62g
Valine	0.589g
Double de-ionised water (MilliQ)	Up to 200mL

2.1.4 Drug stock solution used in this study

Table 7: Composition of 3AT used in this study

Drug	Solvent	Final concentration (µg/100mL)
3-Amino-1,2,4-triazole (3AT) (FORMEDIUM™)	Water	16.816 g/100mL

2.1.5 Antibiotics used in this study

Table 8: Composition of Ampicillin used in this study

Antibiotic	Solvent	Final concentration (µg/mL)
Ampicillin (FORMEDIUM™)	Water	50µg/mL

2.1.6 Growth conditions

Growth of *Saccharomyces cerevisiae*

For growth on solid media, yeast cultures were streaked onto SD or YPD media and grown at 30°C. For growth in liquid medium, yeast cultures were inoculated in SD or YPD media and grown at 30°C, spinning 160rpm. Yeast cultures were then stored at 4°C.

Growth of *Escherichia coli*

For growth on solid media, bacterial cultures were streaked onto LB media supplemented with Ampicillin and grown at 37°C. For growth in liquid medium, bacterial cultures were inoculated in LB medium supplemented with Ampicillin and growth at 37°C, shaking 160rpm. Bacterial cultures were then stored at 4°C.

2.1.7 Permanent storage of Yeast cultures

Yeast strains were streaked on solid SD media supplemented with ILV, W, and glucose (SDWILV + glucose), then grown at 30°C for 2-3 days. Cells were then transferred, using a toothpick, to 2mL tubes of sterile 30% glycerol, mixed by vortexing, and stored at -80°C.

30% Glycerol solution

15mL of 100% Glycerol was added to 30mL Water, then distributed into 2mL PERM tubes and autoclaved

2.1.8 Permanent storage of *E. coli* cultures

For permanent storage of *E. coli* cultures, bacterial cultures were grown overnight at 37°C, then 500uL was transferred to 2mL tubes of sterile 100% glycerol. Tubes were mixed by vortexing and stored at 80°C.

2.1.9 Primers used in this study

Table 9: Primers used in this study

Primer name	**Sequence (5' to 3')**
ES48	CTGGCAAGCCACGTTTGGT
ES3035	CTCGAGCTCAAGCTTAATTCA
ES3035-B	ACGATCTGCCTCGCGCGT
ES400-48	ACTTCTTTGGCTAAGCGC
ES400-49	CTAGACCTACTGAAGACCGAC
ES3227	GGAGGCCGTCGAGGCCA
ES3228	AATGAAATCCGATTTCAAGTTCT
ES4141-s-R144A-Yih1	CCATTACTGATGCTGGGTCGACTTTCATGG CCTTTGC
ES4142-r-R144A-Yih1	TCGACCCAGCATCAGTAATGGGGTCCGAC
ES4143-s-R144K-Yih1	ATTACTGATAAGGGCTCGACTTTCATGGCC
ES4144-r-R144K-Yih1	GTCGAGCCCTTATCAGTAATGGGGTCCGA C

2.1.10 Plasmid DNA isolation from Yeast

Yeast plasmids were isolated using Zymoprep Yeast Plasmid Miniprep I, which is based off the alkaline lysis method. Yeast cultures were grown for 2-3 days in 4mL SDWILV and glucose. 1.5mL of fully grown yeast culture was then transferred to a 1.5mL tube and spun by

centrifugation at 1000rpm for 2 minutes. The supernatant was then discarded, leaving the cell pellet. 150µL of Digestion Buffer was then added to the tube, along with 2µL Zymolase™. The pellet was then resuspended by flicking and vortexing. The sample was incubated for 15-60 minutes at 37°C. Following incubation, 150µL of Lysis Buffer was added to the tube and mixed well by vortexing. 150µL of Neutralising Buffer was then added to the tube and mixed again by vortexing. Samples were centrifuged at maximum speed (150,000rpm) for 2 minutes. The supernatant was transferred to new tubes to which 400µL of isopropanol was added and then tubes were mixed well. Samples were centrifuged at maximum speed (150000 rpm) for 8 minutes, generating a small pellet of Plasmid DNA. Supernatant was removed and the plasmid pellet was resuspended in 35µL TE buffer (pH 8.0), then stored at -20°C.

2.1.11 Plasmid DNA isolation from *E. coli*

Plasmids were isolated from bacterial strains using the alkaline lysis method based on the ZR (Zymo Research) Plasmid Miniprep kit. 4mL of bacterial cultures were grown overnight in liquid LB media with ampicillin (37°C, shaking). Bacterial cultures were centrifuged for 15 minutes at 4500 rpm (4°C), generating cell pellets. The supernatant was discarded and then 200µL of Resuspension Buffer (containing RNase) was added to the tube and mixed by vortexing until cultures were dissolved in solution. This mixture was then transferred to 1.5mL tubes. 200µL of Lysis Buffer was then added to the tubes and samples were mixed by inverting 2-4 times, or until solution became homogenous. 400µL of Neutralizing Buffer was then added and mixed gently until the sample appeared light yellow, indicating that the neutralization was complete. The lysate was incubated at RT for 1-2 minutes, and subsequently centrifuged for 2 minutes at 14,000 rpm to separate the cell debris from supernatant. The supernatant was transferred to a Zymo-Spin™ IIN column in a Collection Tube, which was then centrifuged for 30 seconds, and flow through discarded. 200µL of Endo-Wash Buffer was added to the column and centrifuged for 30 seconds. 400µL of Plasmid Wash Buffer (containing ethanol) was then added to the column and centrifuged for a further minute. The column was then transferred to a new 1.5mL tube, to which 30µL of DNA Elution buffer was added to the column. The column was then centrifuged for 30 seconds, eluting the isolated Plasmid DNA. The pellet was re-suspended in 40 µl of TE buffer (0.1M Tris HCl pH 8.0, 10mM EDTA) and stored at -20°C.

2.1.12 Determining DNA concentration and purity of samples

A Nanodrop™ spectrophotometer was used to determine the concentration and purity of DNA samples. 1uL of MilliQ H_20 was used as a calibration reference for the DNA samples. 1µL of

DNA sample was loaded onto the Nanodrop and concentration and purity of samples were determined.

2.1.13 Agarose gel electrophoresis

Agarose gel electrophoresis is a molecular technique that was used to separate DNA fragments by size. This was utilised to confirm the presence of DNA fragments before sequencing. A 1-%2% TAE agarose gel was prepared by dissolving 1g-2g of agarose in 100mL of 1X TAE buffer by heating. 10µL of Ethidium Bromide (EtBr) was then added. The gel was mixed and poured into a casting tray, with a comb to create wells, and left to cool and solidify for 20-30 minutes. 5µL of isolated plasmid DNA was added to 1µL of 6X loading dye, mixed thoroughly, and then loaded into the wells of the agarose gel. 2.5µL of a 2-log DNA ladder (0.1-10kb, New England Biolabs) was loaded as a control standard to determine molecular size. The gel was then run at a constant voltage (100V) for 30 minutes. DNA bands were visualized using a UV transilluminator, and images were captured and quantified using BioRad software.

50X Tris Acetate EDTA (TAE) buffer (1L)

Tris	276g
Acetic acid	57.1mL
EDTA	18.6g
Water	Up to IL

6X loading dye

Bromophenol blue	0.25mg
Glycerol	3mL
Water	7mL

2.2.1 *Escherichia coli* (*E. coli*) Transformation

Competent *E. coli* cells (stored at -80°C) in a 1.5mL tube were thawed on ice for no longer than 5 minutes. 3uL of plasmid DNA was added to the tube and mixed softly with pipette, followed by incubation on ice for 30 minutes. The cells were then subject to heat shock in a 42°C water bath for 90 seconds, then placed directly back on ice for 2 minutes. Cells were then regenerated by adding 500-700µL SOB medium and incubated for 45-60 minutes at 37°C, 160rpm. Cells

were then transferred, using glass beads, to solid LB agar plates with ampicillin, and incubated overnight at 37°C.

SOB medium (pH 7)

Table 10: Composition of SOB medium (pH 7)

Components	**Amount**
Tryptone (FORMEDIUM^{TM})	20 g
Yeast extract (FORMEDIUM^{TM})	5 g
NaCl	0.5 g
KCl (250mM)	10mL
$MgCl_2$ (2M)	5mL
Water	Up to 1L

SOB medium is prepared by dissolving solutes in 950mL water. Following this, 10mL of 250mM KCl is added, followed by 2M $MgCl_2$ and pH of the mixture is adjusted to pH 7 using NaOH. The volume is then adjusted to 1L by adding water.

2.2.2 Yeast Transformation

Making competent yeast cells

Yeast cells were grown overnight in a glass tube containing 3mL YPD liquid media, rolling at 30°C. 1mL of the overnight cultures were then inoculated in 50mL YPD liquid media in a 250mL flask, and grown at 30°C shaking (160rpm) until reaching an OD_{600} of 0.8. Subsequently, cultures were transferred to a falcon tube and spun for 5 minutes at 4000rpm (4°C). The supernatant was discarded leaving the cell pellet remaining, which was resuspended in 8mL Solution 1, then incubated at 30°C for 30 minutes shaking, producing competent yeast cells (Andrews, Boone, Davis, & Fields, 2016).

.

Transformation of competent yeast cells

Competent yeast cells were spun for 5 minutes at 4000rpm (4°C), then the supernatant was discarded, and the cell pellet was resuspended in 500uL Solution 1. Herring sperm DNA (10mg/mL) was heated at 100°C for 10 minutes to denature the DNA, then cooled on ice for 5 minutes, producing single stranded DNA. 5µL of this Herring sperm DNA, 5µL of plasmid

DNA and 100µL of competent yeast cultures, was added to an Eppendorf tube and pipetted up and down, then incubated for 30 minutes at 30°C. 700µL of Solution 2 was then added to the Eppendorf tube and the culture was incubated for a further 45 minutes at 30°C, shaking (160rpm). Cultures were then subject to heat shock at 42°C for 10 minutes. Following heat shock, cultures were cooled on ice for 5 minutes then spun at 1000rpm for 5 minutes, then the supernatant was discarded. The cell pellet was resuspended with 80µL liquid SDWILV + glucose and distributed on solid SDWILV + glucose plates using sterile glass beads. Plates were incubated at 30°C for 2-4 days (Andrews et al., 2016).

Yeast transformation solutions

Solution 1

Components	**Amount**
10X TE (pH 7.4)	1mL
1M Lithium acetate	1mL
Water	8mL

Solution 2

Components	**Amount**
44% PEG (in 1X TE)	9mL
1M Lithium acetate	1mL

10X TE

1.211g Tris and 372.24mg EDTA were dissolved in 80mL water, and pH was adjusted to 7.4 with HCl. The volume was then made up to 100mL with water. The solution was then autoclaved at 121 psi.

1M Lithium Acetate

10.2g Lithium acetate with dissolved in 10mL 10X TE, and then made up to 100mL with water. The solution was then filter sterilized.

44% PEG (in 1X TE)

44g PEG (Sigma P4338) was dissolved in 10mL 10X TE, and then made up to 100mL with water. The solution was then autoclaved

2.2.3 Measuring Optical Density of Cultures

Optical measurement of cultures was measured at 600nm (OD_{600nm}) on an Eppendorf BioPhotometer Plus.

2.2.4 Semi-quantitative growth assay (SQGA)

Yih1 mutants were constructed using site-directed mutagenesis prior to this study (unless stated). Yih1 mutant proteins were tagged with a GST-fusion protein for detection and placed under the control of a galactose-inducible promoter. For the SQGA, previously transformed colonies of yeast H1511 (or ESY11001B) strains expressing either GST-tagged Yih1 with different Yih1 mutations (as indicated), GST alone, GST-Yih1 (wild type), or with *GCN1* deleted (*gcn1Δ*) were picked from solid agar plates and inoculated in 4mL SDWILV liquid medium containing glucose. Cultures were grown at 30°C, rotating at 150rpm until saturation. A 10-fold serial dilution (1 to 1:10000) was performed for each culture and 5µL of each dilution was transferred onto solid agar plates containing SDWILV with either glucose, galactose, or galactose and 3-Amino-1,2,4-triazole (3AT: 30mM, 60mM, 90mM, 120mM, 150mM). 3AT is a drug that causes histidine starvation, thus was used to induce amino acid starvation (Sattlegger et al., 2011). On the control plates, glucose was used as a carbon source to test the growth of strains under non-starved conditions. As Yih1 proteins are under control of a galactose-inducible promoter, galactose was added to control plates and 3AT-supplemented plates to induce overexpression of Yih1 proteins. Plates were stored at 30°C. Colony growth was observed over 13-14 days by scanning twice a day (for the first 3 days), and then once every sequential day. Plates were scanned using MicroTek Scanning systems. Two replicate plates were used as a control measure.

2.2.5 Polymerase Chain Reaction

Polymerase Chain Research is an in vitro technique used in molecular biology to amplify a specific DNA fragment from genomic or plasmid DNA. A small oligomer (primer) of approximately 20-25 nucleotide sequences long (forward and reverse), as well as a heat-stable

polymerase, are used to amplify the fragment of interest. Standard PCR reactions were carried out in an Applied Biosystems® Veriti® 96-Well Thermal Cycler. The PCR process can be separated into 3 main steps:

1. Denaturation: DNA is heated to approximately 95°C, resulting in separation of double stranded DNA into single strands. This is preparing the template for annealing of the primers.
2. Annealing: The template is cooled to 55°C to allow the specific primers to anneal to the single-stranded DNA.
3. Extension: A heat-stable polymerase uses the primer sequence as a template to amplify the DNA fragment of interest. This is carried out at 72°C.

Table 11: Standard steps of PCR

Step	**Temperature**	**Duration**	**Cycles**
Initial denaturation	95°C	3 minutes	1
Denaturation	95°C	30 seconds	35 (x3)
Annealing	55°C	15 seconds	
Extension	72°C	1 minute/kB	
Final Extension	72°C	1 minute/kB	1

The three PCR steps are repeated cyclically 30-35 times resulting in amplification of the DNA fragment of interest. The standard PCR reaction volume was 20µL and was prepared as follows;

Table 12: A standard PCR reaction mix

Component	**Volume**
PCR-grade water	Up to 20µL
10X *Kapa/Taq* buffer	3 µL
50mM $MgSo_4$	0.4µL
10 mM dNTP Mix	0.4µL
10 µM Forward primer	0.6 µL
10 µM Reverse primer	0.6 µL
5 U/µl *KAPA/*Taq DNA polymerase	0.2 µL
Template DNA	0.4µL

2.2.6 Colony PCR

Colony PCR is a type of PCR that allows us to amplify a gene of interest from whole yeast cells rather than pure plasmid DNA. For this, a small amount of fresh yeast cells was added to the PCR reaction as the template instead of plasmid DNA. Before the PCR reaction, cells are lysed by heating in 0.2M NaOH for 8 minutes (100^{o}C). Then, a typical cycling protocol for colony PCR is described as follows;

Table 13: Thermocycler protocol for Colony PCR

Step	**Temperature in °C**	**Duration**	**Cycles**
Initial denaturation	94^{o}C	5 minutes	1
Denaturation	94 ^{o}C	30 seconds	35
Annealing	55^{o}C	30 seconds	
Extension	72 ^{o}C	60 seconds	
Final Extension	72 ^{o}C	60 seconds	1
Hold	4°C	∞	

Table 14: Master mix for Colony PCR

Component	Volume
PCR-grade water	Up to 20µL
10X *kapa/Taq* buffer	2 µL
50mM $MgSo_4$ (only if needed)	1.2 µL
10 mM dNTP Mix	0.4µL
10 µM Forward primer	0.8 µL
10 µM Reverse primer	0.8 µL
5 U/µl *kapa* Taq DNA polymerase	0.08 µL
Fresh yeast cells	1µL

2.2.7 Fusion PCR

Fusion PCR is another type of PCR that is utilised to produce a fusion DNA fragment from different DNA fragments. This is achieved through overlap extension and does not require the

addition of restriction sites or DNA ligase. Instead, a DNA fragment is used as a template and an overlapping primer sequence is utilised to generate a full length PCR product. The PCR thermocycler protocol is the same as in Table 3.3, except the annealing temperature is 69°C.

2.3 Preparation of Whole Cell Extract for Western Blotting

Yeast cultures were grown to saturation in 4mL SDWILV + 2% Glucose (30°C, 160rpm). Subsequently, overnight cultures were inoculated in 25mL SD medium + 2% Galactose (with a starting OD_{600} 0.05-0.1) and grown to exponential phase (OD_{600} 0.6-0.8) (30°C, 150rpm). Just before harvesting, cultures were treated with 3AT (30mM) for an exposure time of 30min (30°C, 160rpm) to induce amino acid starvation conditions. This latter step was disregarded for the analysis of expression levels.

2.3.1 Formaldehyde crosslinking

Cultures were crosslinked by transferring the culture to a falcon tube with 10g ice chips and formaldehyde (1% final concentration) and left on ice for 1h, inverting every 15 minutes. Glycine (0.25M final concentration) was added to quench excess formaldehyde. Cell pellets were then harvested by centrifugation for 5min at 4500rpm (4°C). Cell pellets were either stored at -80°C or used for chemical lysis immediately.

2.3.2 Chemical cell lysis

Harvested cells were resuspended in 1mL MilliQ water and pelleted by centrifugation. For chemical cell lysis, the cell pellet was incubated in NaOH (200µL of 0.1M NaOH) for 5 minutes, followed by centrifugation for 5 minutes at RT (4000rpm). The supernatant was discarded, and the remaining sample was used for western blot analysis.

2.4. Western Blot

2.4.1 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

Whole cell extract (WCE) of yeast cells were prepared as described above. Equal amounts of whole cell extract were resolved, and proteins were separated using SDS-PAGE on a 4-20% gradient gel. 4% and 20% acrylamide solutions were prepared and stored at 4°C until use. Two clean plastic gel plates were sealed together around the sides using 1% Agarose and 2% Agarose, then left at RT to solidify. The gel caste was secured in a rack and a gradient mixer was used to pour a 4-20% gradient gel. 7.5mL of 4% Acrylamide solution was added into left chamber of gradient mixer, while 7.5mL of 20% Acrylamide solution was added into the right chamber. While mixing with a magnetic stirrer, 75uL Ammonium persulfate (APS) and 7.5uL

Tetramethyl-ethylene-diamine (TEMED) were added to each chamber containing acrylamide solution. Subsequently, the acrylamide mixture was appropriately poured between the plates and a loading comb was inserted. The gel was left to solidify at RT for no less than 1 hour and then used for SDS-page or stored wrapped in wet paper towels at 4°C until use. The solidified gel was placed into a gel chamber and the comb was removed. The gel chamber was then filled with cold running buffer. Samples for SDS-page were prepared by adding 2x Protein loading buffer, being heated at 80°C for 8 minutes, vortexed, and then loaded onto the gel. See Blue Pre-stained protein ladder (Thermo Fischer) was also loaded onto the gel. Gel electrophoresis was carried out at (250V, 100mA, 1h).

	4% Acrylamide solution	**20% Acrylamide solution**
MilliQ water	75.3mL	47mL
40% Acrylamide	12mL	100mL
Tris (pH 8.8)	30mL	50mL
SDS (10%)	1.2mL	2mL

Running buffer (2.5L)	**Amount**
10% SDS	25mL
10X Tris	250mL
MilliQ water	2225mL

2x Protein loading buffer (100mL)	**Amount**
Tris HCl (pH 6.8)	25mL
SDS (4%)	4g
Glycerol (20%)	20mL
Bromophenol blue (0.1%)	0.1g
β-Mercaptoethanol	

2.4.2 Transfer to nitrocellulose membrane

After SDS-page, proteins were transferred from the gel onto a nitrocellulose membrane by a wet transfer. A wet "sandwich" was composed while submerged in a container filled with cold Transfer buffer. Firstly, the gel was placed in a pack of a sponge and 2x whatman paper (soaked

in cold transfer buffer). The nitrocellulose membrane was then placed on top of the gel followed by 2x whatman paper and another sponge. The sandwich was secured together and placed into the transfer unit in the appropriate orientation (the gel facing the negative electrode and then membrane facing the positive electrode). Cold transfer buffer was used to fill up the transfer unit and an ice pack was also placed into the transfer unit. The transfer unit was placed in a box of ice-cold water and the transfer was performed (100V, 1Amp, 1 hour 20 minutes).

Transfer buffer (1L)	**Amount**
Methanol	200mL
MilliQ water	700mL
Tris/glycine (10x)	100mL

10X Tris Glycine	**Amount**
Tris	151.5 g
Glycine	720g
Water	Up to 5L

2.4.3 Western blot analysis

Following the transfer, the membrane was stained using Ponceau S (1% w/v in 1% acidic acid) (20 minutes at RT), and subsequently destained by rinsing multiple times with a 5% acetic acid solution. The membrane was then rinsed with TBS-T and incubated in blocking solution using 3% BSA (bovine serum albumin) in TBS-T for eIF2α-P and GST or using 5% fat-free milk in TBS-T solution for Pgk-1. For analysis of expression levels, samples were then subject to immunoblotting using antibodies against GST (1:2000) and then the housekeeping gene Phosphoglycerate kinase (Pgk1) (1:2000) as a control. For analysis of eIF2α phosphorylation, antibodies to specifically detect phosphorylated eIF2α at the Serine-51 residue (eIF2α-P) (1:1000) and Pgk1 (control) were used (Table 5). The membrane was incubated with the appropriate primary antibody overnight at 4°C, then washed with TBS-T (3 times: 5, 10, and 10 minutes, respectively). The membrane was then incubated with the appropriate secondary antibody (Table 6) for 1 hour then washed with TBS-T (4 times: 5, 10, 10 and 15 minutes, respectively). Chemiluminescent solution (Solution A and Solution B) was added onto the membrane and the signals were visualized using ChemiDocTM Image System (BioRad) and analysed using Image Lab software and Microsoft Excel. Images used for analysis, and

statistical analyses of data, are provided in *Supplementary Material (Figures S1.2, S1.3, S2.3, S2.4, S3.2)*.

Table 15: Primary antibodies used in this work

Primary antibody	**Dilution**	**Secondary antibody**	**Source**
Pgk1	1:5000	Anti-mouse	Invitrogen
α-GST	1:2000	Anti-rabbit	SantaCruz
eIF2-P	1:1000	Anti-rabbit	Invitrogen

Table 16: Secondary antibodies used in this work

Secondary antibody	**Dilution**	**Source**
Anti-mouse	1:50,000	Pierce
Anti-rabbit	1:100,000	Pierce

Blocking solutions used in this study

- 5% Milk powder in TBS-T
    - 2.5g of Milk powder was dissolved in 50mL TBS-T
- 3% BSA in TBS-T
    - 1.5g of BSA (Bovine Serum Albumin) was dissolved in 50mL TBS-T

5% Acetic acid (1L)	**Amount**
Acetic acid (glacial)	50mL
Water	950mL

0.8% Ponceau stain	**Amount**
Ponceau (Helena)	0.8g
Acetic acid (5%)	100mL

10X Tris Buffered Saline (TBS) (1L)	Amount
NaCl	80g
Tris base	24g
HCl	Up to pH 7.4

TBS-T (1L)	Amount
10X TBS	100mL
MilliQ water	900mL
Tween-20	1mL

2.5 Harvesting cells for Yih1 purification

Cells were grown on a large scale. Yeast cells containing GST-tagged Yih1 were grown in 300mL SDWILV + Galactose (2%) to an OD_{600} of 0.8-1. The culture was then transferred to 300mL bottles and fresh ice chips were used to fill the bottle (approximately 75g ice chips). All steps from here were carried out on ice or at 4°C. Cultures were spun at 4200rpm for 5 minutes, leaving a cell pellet, and the supernatant was discarded. The cell pellet was then resuspended in 50mL ice cold water and spun at 4000rpm for 5 minutes. The supernatant was discarded and the cell pellet was transferred to round bottom tubes using 10mL ice cold water. Cultures were spun for a further 5 minutes at 4200rpm and the supernatant was discarded, leaving the cell pellet. Tubes were tapped upside down on a paper towel to remove any residual water. Cell pellets were stored at -80°C until further use.

2.5.1 Mechanical cell lysis for generation of whole cell extract

Cell pellets were thawed on a mixture of ice and water. Pellets were dissolved in 200uL ice cold breaking buffer. In this instance, we used SEC buffer (10-20mM) and TCEP (an EDTA free protease inhibitor) (see Table 17). 700µL acid-washed glass beads were then added to the mixture and cells were lysed by 10 repeated cycles of vortexing at maximum speed for 30 seconds followed by 30 seconds standing on ice. The cell lysate was then separated from the glass beads by centrifugation at 4000rpm for 5 minutes. The supernatant was then transferred to new 1.5mL tubes and centrifuged for a further 10 minutes at 14,000rpm to ensure all cell lysate was extracted. From here, the cell lysate was transferred to a new 1.5mL tube again.

Table 17: Components of SEC Buffer (pH 7)

Components	**Amount**
NaH_2PO_4 (10mM)	1.2g/L
EDTA (2mM)	0.58g/L
TCEP (protease inhibitor) (2mM)	0.5733g/L
Water	Up to 1L

SEC buffer was prepared by dissolving NaH_2PO_4 (10mM), EDTA (2mM), and TCEP (2mM) in approximately 800mL MilliQ H_2O. pH was then adjusted to 7.0 using NaH_2PO_4 (500mM), and then MilliQ H_2O was added up to 1L.

2.6 Size Exclusion Chromatography for purification of yeast Yih1

All steps here were carried out at 4°C. As the Yih1 protein construct used in this study has a GST (Glutathione-S-Transferase) tag, the purification technique used here exploited the affinity that GST has for glutathione, its substrate molecule. The plastic gel filtration column was prepared by loading approximately 2mL of Glutathione Agarose (Protino® Glutathione Agarose 4B) into the column and washed using 10mL of freshly prepared SEC buffer (see Table 17) 2 times to remove ethanol contained in the Glutathione agarose. To immobilize GST-Yih1, yeast protein lysate (approximately 5mL; generated as described in Materials and Methods Sections 2.5 and 2.5.1) consisting of overexpressed GST-Yih1, was loaded directly onto the column. When the protein lysate is loaded directly onto the column, the GST would bind to its substrate molecule glutathione and therefore immobilise GST-Yih1 in the affinity resin gel matrix. The flowthrough, containing non-specific proteins, was collected in a falcon tube, and stored on ice. Subsequently, the column was washed a further 3 times with SEC buffer. Glutathione reduced (GSH) was then prepared (112mg GSH in 15mL SEC buffer- total concentration of 25mM) and loaded directly onto the column to elute GST-Yih1, the target protein, and flow through was collected. This process was repeated two times subsequently using 1mL of Glutathione reduced to elute any remaining GST-Yih1. Aliquots were frozen at -20°C. A 5µL aliquot used for separation by SDS-PAGE gel electrophoresis, and a 100µL aliquot was used for Fast Protein Liquid Chromatography.

2.7 Fast Protein Liquid Chromatography

Fast Protein Liquid Chromatography (FPLC) was carried out by Dr Stefan Harjes using a 75kDa FPLC size exclusion column on a ÄKTA Explorer System at 4°C.

2.8 In-Gel Chymotrypsin Digest

2.8.1 Preparation of Gel for Digest

Purified GST-Yih1 was resolved on a 4-20% SDS-page gel by gel electrophoresis. The gel was then fixed with 40% methanol/MilliQ water and shaken for 20 minutes. The fixative was removed, and the gel was covered with MilliQ water, then shaken for 10 minutes. The gel was then stained with Colloidal Coomassie Blue stain by removing MilliQ water and covering the gel with freshly diluted Colloidal Coomassie Blue (4 parts Colloidal Coomassie Blue stock to 1 part Methanol) and shaken overnight. The gel was then destained with MilliQ water, and this process was repeated until the background had faded.

Table 18: Colloidal Coomassie Blue stock recipe (250mL)

Component	**Amount**
Phosphoric acid (85%)	3mL
Coomassie G-250	0.25g
Ammonium sulfate	25g
MilliQ H20	Up to 250mL

2.8.2 In Gel Proteolytic Digest using Chymotrypsin

The SDS-page gel was placed on a clean glass plate on top of a light box. The desired protein band (approximately 1mm in width) was excised using a clean scalpel blade and then diced into very fine pieces and placed into an Eppendorf Protein LoBind microcentrifuge tube. The gel pieces were then destained by washing with 300uL of freshly prepared 1x ABC (Ammonium bicarbonate)/50% methanol at 45°C until the gel slice became colourless. Gel pieces were then dehydrated by addition of 300uL 80% MeCN (Acetonitrile) for 1 minute, or until gel pieces appeared shrunken and opaque with a colourless supernatant. The supernatant was discarded, and gel pieces were further dried using a centrifugal evaporator (SpeedVac Thermo Scientific Savant SPD131DDA) for 10 minutes.

1. Reduction and Alkylation (Use LC-MS grade water, acid and solvents from this point on)

Dried gel pieces were reduced with 30µL of freshly prepared Reducing Solution and incubating for 1hr at 45°C. The supernatant was removed, and gel pieces were washed with 100µL 1X ABC. The liquid was removed, and gel pieces were dehydrated with 300µL 80% MeCN, which was then removed. Gel pieces were then dried using the centrifugal evaporator (approximately 5 minutes).

Gel pieces were then alkylated with 30µL of freshly prepared Alkylation Solution and incubated for 30 minutes at room temperature in a dark room. The supernatant was removed, and gel pieces were washed with 100µL 1X ABC. The supernatant was removed, and gel pieces were dehydrated with 300µL of 80% MeCN. This process was repeated and then the liquid was removed, and gel pieces were dried by centrifugal evaporation until gel pieces were bone dry.

2. Digestion

Approximately 30µL of Digestion solution (prepared as shown below) was added and the sample was incubated on ice for 10 minutes, to swell the gel pieces. After swelling, excess digestion solution was removed and 30µL 1X ABC was added to keep the gel pieces moist during incubation. Sample was incubated with Digestion solution (containing Chymotrypsinogen A) at 37°C overnight.

3. Extraction

Tubes were centrifuged and sonicated in an ultrasonic bath (2 minutes), followed by centrifugation for a further 2 minutes. The supernatant of the digestion was collected into a fresh LoBind tube. 50uL of 5% formic acid/40% MeCN was added, followed by sonication for 2 minutes and centrifugation for 2 minutes. The sample was spun in a centrifugal evaporator to reduce the volume to 30µL, ensuring the sample did not dry completely. From here, samples could be stored at -80°C or spun by centrifugation at full speed for 15 minutes (4°C) and then transferred into HPLC vial inserts carefully.

1X ABC solution

50 mM Ammonium bicarbonate (0.04 g/10 mL) pH>7.9.

Reducing Solution

Reducing solution was prepared by dissolving DTT (Dithiothreitol) (0.00154 g/mL) in H_2O to a final concentration of 10 mM.

Alkylation solution

Alkylation solution was prepared by dissolving Iodoacetamide (0.0037 g/mL) in H_2O to a concentration of 20 mM.

Digestion solution

Chymotrypsinogen A was used for protease digestions in this work and was activated by Trypsin before use by adding 2µL of SOL-u trypsin to 1 mg/ml of chymotrypsinogen A in 1x ABC. This solution was then incubated at 37°C for 2 hours. Solution was dissolved 20x with 1x ABC and then added to digestion solution (see Table 19).

Table 19: Components of Digestion solution in this work

Component	**Amount (200uL total)**
1X ABC	196uL
CaCl (1mM)	0.2uL
Chymotrypsinogen A (as prepared above) (Accession number P00766)	4uL

2.9 Tandem Mass Spectrometry (MS/MS)

Tandem Mass Spectrometry (MS/MS) of purified GST-Yih1 fragments was carried out by Trevor Loo due to time constraints limiting access to equipment at the time. Therefore, these time restraints restricted that ability for a detailed understanding of the methodology of Tandem Mass Spectrometry. Credit is given to Trevor Loo for Mass Spectrometry.

2.10 Site directed mutagenesis

Site-directed mutagenesis is achieved based on the principle that a circularised double-stranded plasmid can be amplified using complementary primers designed specifically carrying the required mutation flanked by sequences matching those flanking the desired mutagenesis site, resulting in introduction of the desired mutation.

Yih1 proteins with selected mutations were produced by site-directed mutagenesis. Initial PCR reactions were set up using the plasmid pES187-b1 as a template, as well as mutagenesis and flanking primers that amplified two halves of the same selected Yih1 fragment with overlapping regions containing the inserted mutation. The following PCR reaction uses the resulting PCR fragments (depicted as PCR A and PCR B, Fig. 10) as the template DNA, along with flanking primers that annealed upstream and downstream of the Yih1 fragment. This results in full-amplified Yih1 fragments containing the desired mutations. PCR fragments were then visualised by agarose gel electrophoresis against a 2-long DNA ladder of known sizes to determine fragment size. The Yih1 fragments were transformed into yeast (as described in Materials and Methods, Section 2.2.2).

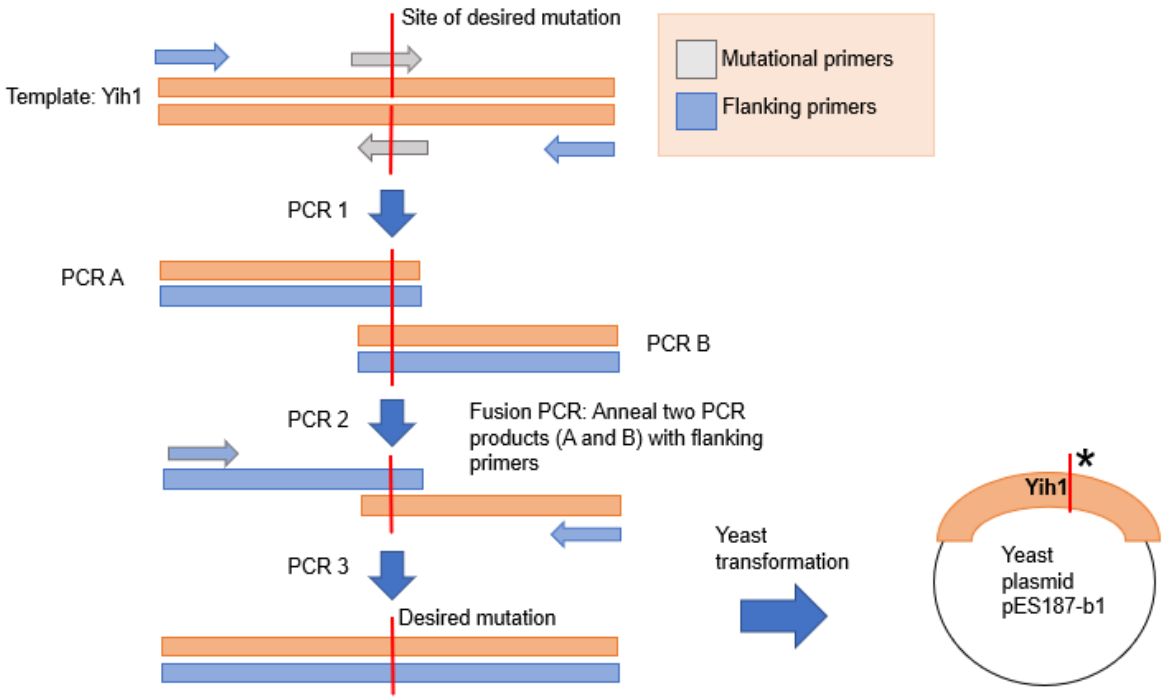


Figure 10: Overview of the site directed mutagenesis procedure.

See text for details.

2.11 Next Generation Sequencing (NGS)

Plasmid DNA was commercially sequenced by Macrogen's Next-generation sequencing service. Next-generation sequencing is a DNA sequencing tool that is used to determine the nucleotide or amino acid sequence from an entire genome/ plasmid DNA with the help of oligomers. Samples for sequencing were prepared according to the requirements of Macrogen. The purity and concentration of the DNA sample was determined through nanodrop. The oligomers were sent according to the requirements of Macrogen. After the generation of sequence, the sequence was aligned through Nucleotide BLAST (Basic Local Alignment Search Tool), and the alignment is searched to determine the sequence and/or any of the mutations were introduced.

2.12 Statistical analysis

Statistical analyses were carried out using Microsoft Excel. Two data values were used for statistical analyses; hence, more samples are necessary for more representative data. Student T-test (Paired two-sample T-tests), mean, standard deviation and standard error values are provided in Supplementary Material.

3. Results

In this study, we aimed to shed light on the molecular mechanisms behind how Yih1 carries out the inhibition of Gcn2. As alluded to previously, under physiological conditions, the Gcn1-Gcn2 interaction occurs even in the presence of Yih1 (Sattlegger et al., 2004). This suggests that Yih1 must exist in an inactive form and an active form that is primed for Gcn2 inhibition (Harjes et al., 2021). It is not yet known what cue shifts Yih1 from the inactive to the active form and vice versa. However, previous research has demonstrated that overexpression of Yih1/IMPACT in yeast is a suitable way to render Yih1 capable of Gcn2 inhibition (Sattlegger et al., 2004). When overexpressed, Yih1/IMPACT impairs the GAAC response by competing with Gcn2 for Gcn1 binding, thereby disrupting the Gcn1-Gcn2 protein-protein interaction (Sattlegger et al., 2004). A possible mechanism is that overexpressed Yih1/IMPACT may be primed for Gcn1 binding because of the excessive amount of protein in the cell increasing the affinity between Yih1 and its binding partner Gcn1. In this way, Gcn2 inhibition is encouraged, allowing us to examine the molecular function of Yih1/IMPACT in this process in molecular detail.

Based on the recently resolved Yih1 structure, amino acids have been predicted to be involved in Yih1 function by the means of different mechanisms, such as targets for modification, inter-domain interactions, and protein interactions (Harjes et al., 2021). Some of these are computationally predicted amino acids. The aim of this study is to test/verify, in the living cell, the predicted function of these amino acids. This will give us insight into the function of Yih1 in protein-protein interactions and how it might get primed for Gcn1 binding. We have analysed the Yih1 structure, together with previous literature, and have predicted Yih1/IMPACT amino acids that may function in Gcn2 inhibition (Table 20). The first step of this study will be to categorise these amino acids into groups of their predicted function.

Grouping of Yih1/IMPACT Amino Acids

The aim of this study was to further unveil the molecular mechanisms behind Yih1-mediated Gcn2 regulation. Because of this, we interrogated earlier publications on the Yih1 structure and amino acid sequence to find amino acids potentially relevant for Yih1 function (Sattlegger et al., 2011; Harjes et al., 2021). The amino acids were assigned into groups based on their proposed role in the function of Yih1 (Table 20). Group 1 is comprised of Yih1/IMPACT amino acids that are proposed to be involved in the direct interaction between Yih1/IMPACT and its binding partner Gcn1. These amino acids have been previously published (Sattlegger et al., 2011). Group 2 is comprised of Yih1 amino acids that are proposed to be involved in

interdomain Yih1 interactions (Table 20). These amino acids residues; E87 and D90, are positioned in the RWD domain (Harjes et al., 2021; Sattlegger et al. 2011). E87 and D90 have also previously been published by Sattlegger et al., 2011. Group 3 is comprised of amino acids, K173, R175 and K176, which are also proposed to mediate the RWD-ancient domain (interdomain) interaction and instead are positioned in the Yih1 ancient domain. It is suggested that amino acid residues in Group 2 and Group 3 (Table 20) may form the interface between the two Yih1 domains. These amino acids therefore may be involving in mediated the orientation and conformation of Yih1 (Harjes et al., 2021).

In addition to these groups, we have also taken interest in a recent finding showing that when Yih1 is expressed in *E. coli*, the ancient domain residue R144 is dimethylated (Harjes et al., 2021). This amino acid is fully conserved among Yih1/IMPACT family of proteins. Group 4 is composed of the amino acid residue R144, in which it is predicted that this residue may be involved in the function of Yih1.

Table 20: Grouping of Yih1 amino acids in this study

Group	Proposed role	Amino acid(s)	Mutations	Source
1	Gcn1 binding (Yih1-Gcn1 interaction)	D102 (Yih1) E106 (Yih1)	D102A E106A EDDE-87,90,102,106-AAAA	Sattlegger et al. (2011)
2	Interdomain interactions (RWD domain)	E87 D90	E87K ED-87,90-AA ED-87,90-KK ED-87,90-RR	This work Sattlegger et al. (2011) This work This work
3	Interdomain interactions (Ancient domain)	K173 R175 K176	R175A R175E KRK-173,175,176-AAA KRK-173,175,176-EEE	This work This work This work This work
4	Yih1 Methylation/Post Translational Modification	R144	R144A R144K	This work This work

Overexpressing Yih1 for molecular insight on the proteins function

As mentioned prior, Yih1 is not a general inhibitor of Gcn2 under native cellular conditions (Sattlegger et al., 2004). This is because the Gcn1-Gcn2 interaction occurs even in the presence of Yih1 (Sattlegger et al., 2004). From an experimental approach, this means that it is difficult to score whether Yih1 inhibits Gcn2 under physiological conditions. As the aim of this thesis is to decipher the molecular mechanisms behind Yih1-mediated Gcn2 inhibition, we utilised overexpression of Yih1 in yeast to drive the interaction between Yih1 and Gcn1. Yih1 overexpression has been experimentally shown to be functional in inhibiting Gcn2 (Sattlegger et al., 2004; Sattlegger et al., 2011). By encouraging the Yih1-Gcn1 interaction, this allows us to get a detailed molecular understanding of how Yih1 impairs Gcn2.

To test the proposed functions of these amino acids in the living cell, amino acids that are predicted to be involved in this process were substituted, and it was tested whether the mutated

Yih1 was still capable of inhibiting Gcn2, as well as whether it became more or less potent in Gcn2 inhibition. The rationale/approach used in this study is; if overexpressed Yih1 can impair the Gcn1-Gcn2 interaction, then Gcn2 activation will be disrupted, meaning cells will be unable to cope with starvation. This will impair their ability to grow. However, if a Yih1 mutant is less able to impair the Gcn1-Gcn2 interaction, then Gcn2 activation is restored. This will not impair the yeast cells ability to grow. If a Yih1 mutation enhances Gcn1 binding, then Gcn2 activation will be impaired more strongly. This means that the ability of yeast cells to grow will be even more impaired. Therefore, the degree of Gcn2 activity inversely correlates to the ability of Yih1 to disrupt the Gcn1-Gcn2 interaction (Fig. 11). It also gives us information on how strongly Yih1 is primed to bind its essential binding partner, Gcn1. Based off this rationale, we can determine whether a certain amino acid is important for the function of Yih1 in Gcn2 inhibition.

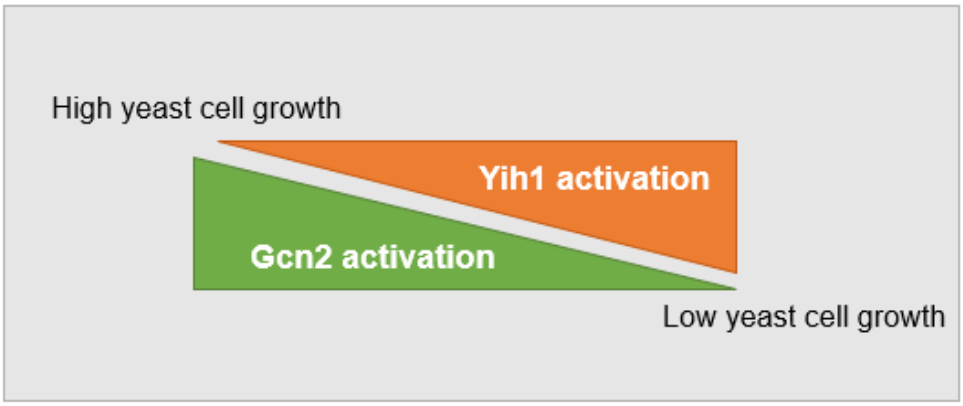


Figure 11: Yih1 is a competitive inhibitor of Gcn2.

The degree of Gcn2 activity inversely correlates to the ability of Yih1 to disrupt the Gcn1-Gcn2 interaction. High Gcn2 activity is correlated with high levels of yeast cell growth, whereas high levels of Yih1 activation impairs yeast cell growth through inhibition of Gcn1. See text for details.

3.1 The yeast system used for scoring Yih1-mediated Gcn2 inhibition

Gcn2 activation can be easily scored in the living cell by analysing yeast cell growth under starved conditions. When Gcn2 can be activated, cells can overcome starvation and grow on starvation medium. However, if cells cannot activate Gcn2, then they are not able to grow on starvation medium. The stronger Gcn2 can be activated, the better the cells can grow on starvation medium. In this thesis, Yih1/IMPACT and mutated Yih1 will be overexpressed in yeast cells, and their ability to inhibit Gcn2 will be determined by analysing yeast cell growth in starved conditions. Next, an overview is given on the plasmids used to overexpress Yih1/IMPACT in yeast (Fig. 12).

3.1.2 Plasmid map of Yih1/IMPACT used in this study

Yih1/IMPACT overexpression in yeast was accomplished by introducing into yeast a plasmid harbouring the *Yih1/IMPACT* open reading frame (ORF). These plasmids had been published and described previously (Sattlegger et al., 2000; Sattlegger et al., 2004; Pereira et al., 2005). *Yih1* and *IMPACT* ORFs are fused at the C-terminus with a Glutathione-S-transferase tag. Furthermore, both *Yih1* and *IMPACT* are under the control of a Galactose Inducible Promotor (GAL1), allowing their expression to be induced in the presence of Galactose. These plasmids also contain the *URA3* gene to complement the *URA3* deletion in the yeast cell, allowing the plasmids to be maintained in the cell. In this study, these plasmids were used to generate mutated versions of Yih1 and IMPACT. These plasmids were transformed into yeast cells to then determine the effect of Yih1/IMPACT to inhibit Gcn2.

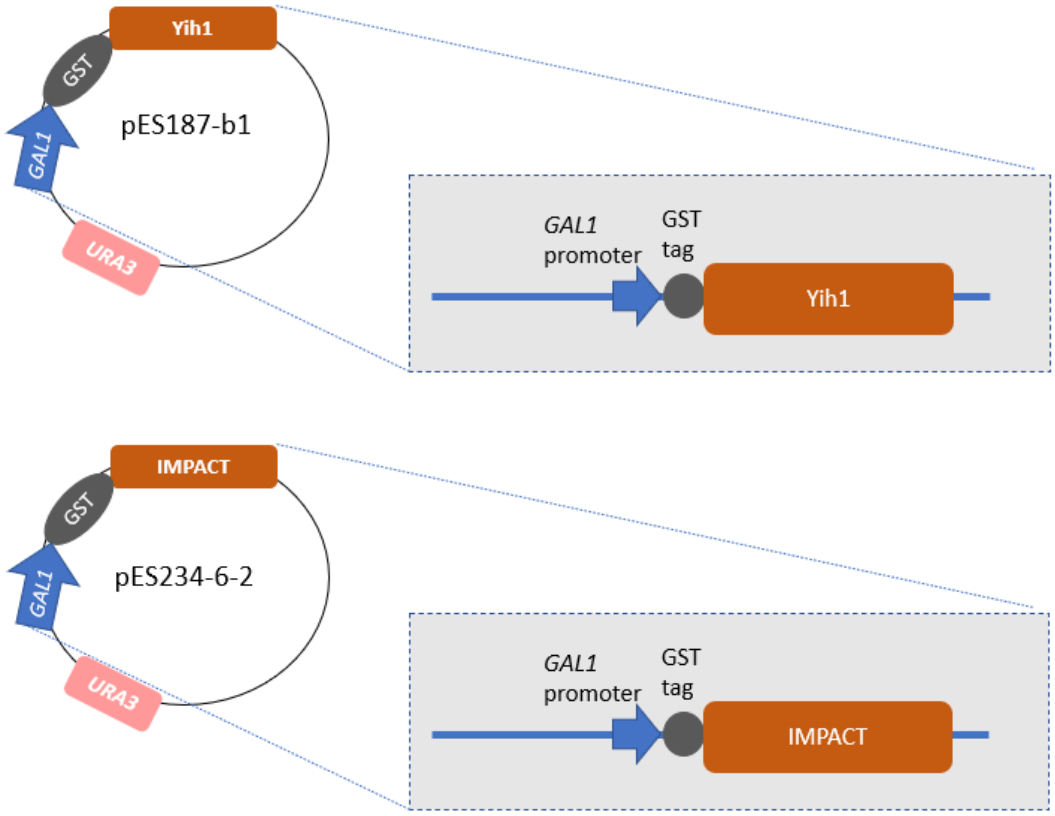


Figure 12: Plasmid map of Yih1/IMPACT protein constructs used in this study.

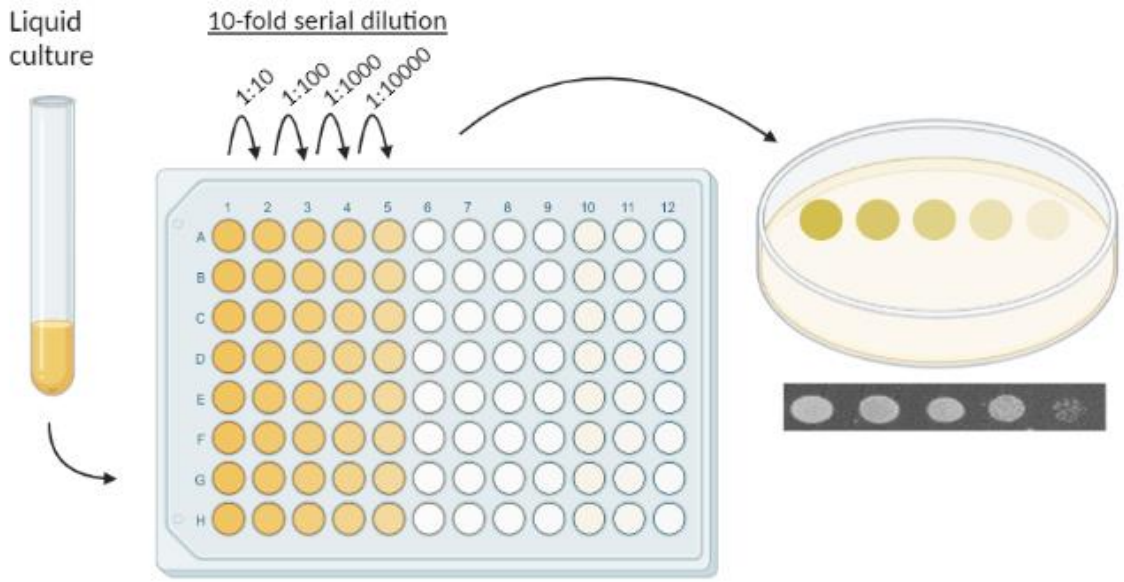
Yih1 and *IMPACT* were inserted into the vector pES128-9 to generate plasmids pES234-6-2, and pES187-b1, respectively. *Yih1/IMPACT* is tagged with a Glutathione-S transferase (GST) tag located upstream of their ORFs. *Yih1* and *IMPACT* are under the control of a Galactose inducible promotor (*GAL1*), to induce their overexpression. Additionally, the vector has a *URA3* selectable marker to complement the *URA3* deletion in the yeast cell, thereby allowing the plasmids to be maintained in the cell. Plasmids were transformed into yeast strains and used for further experiments in this study.

3.1.3 Semi Quantitative Growth Assay for scoring Yih1-mediated Gcn2 inhibition

As alluded to previously, the degree of Gcn2 activity directly correlates with the ability of yeast cells to grow on starvation media. In this study, we performed a Semi-Quantitative Growth Assay (SQGA), which allowed us to gauge the yeast growth rate and hence the level of Gcn2 activity in living cells. To do this, colonies of yeast strains overexpressing GST-tagged Yih1/IMPACT proteins were grown to saturation. Each saturated culture was subject to a 10-fold serial dilution, and then an aliquot was transferred onto solid medium containing either Glucose only, Galactose only, or Galactose and 3-Amino-1,2,4-triazole (3AT). 3AT is a drug that induces amino acid starvation conditions in yeast cells by causing histidine starvation (Sattlegger et al., 2011). On the control plates, glucose was used as a carbon source to test the growth of these strains under conditions where no proteins are overexpressed. As the expression of Yih1/IMPACT proteins are under control of a galactose-inducible promoter, additional control plates contained galactose to test whether the overexpression of Yih1/IMPACT affects cell growth. To test for the effect of overexpressed Yih1 on Gcn2 activity, cultured were plated on plates supplemented contain galactose and 3AT. Different concentrations of 3AT (30mM, 60mM, 90mM, 120mM, 150mM) was used to score for high and low levels of 3AT sensitivity. These 3AT concentrations have been sufficient for generating a phenotype in previous literature (Sattlegger & Hinnebusch, 2004). Colony growth was observed at 30°C over 13-14 days.

Yeast cell growth on solid medium containing 3AT is an indicator of Gcn2 activation (Sattlegger et al., 2011). If Yih1 is functionally able to disrupt Gcn2 activation, and consequently impair cell growth, these strains are 3AT sensitive. This is also referred to as a general control non-functional (Gcn^-) phenotype as strains that cannot activate Gcn2 are not able to overcome amino acid starvation, thus cannot grow (Sattlegger & Hinnebusch, 2000). On the other hand, if Yih1 is impaired in inhibiting Gcn2 then growth will be maintained, meaning that these strains are 3AT resistant.

A



B

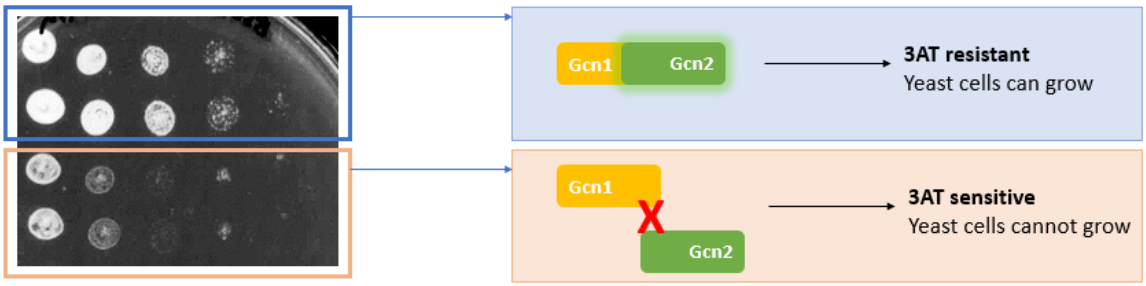


Figure 13: Experimental procedure for a Semi-Quantitative Growth Assay (SQGA).

A, Fully grown yeast cultures were subject to a 10-fold serial dilution (1:10, 1:100, 1:1000, 1:10000) in a 96-well microtiter plate, and then 5uL of each dilution was transferred onto solid media (as shown in the diagram). Cell growth is then analysed over a 14-day period. A representative image of a serial dilution of yeast cultures is shown below the depicted agar plate. Figure generated using BioRender. B, Cells that can grow on 3AT are '3AT resistant', whereas cells that cannot grow on 3AT are '3AT sensitive'. Panels shown in B are cell growth levels on a microtiter plate with 60mM 3AT concentration.

3.1.4 Quantitative analysis of Semi Quantitative Growth Assays

For a quantitative understanding of the growth differences between strains, a growth score value was calculated and assigned for each strain. The growth score was based on how well yeast cells could grow on 3AT medium. Firstly, for each dilution of a strain, the score was calculated as follows; Full growth (100%)= 2, 75% growth= 1.5, 50% growth= 1, 25% growth= 0.5, and No growth= 0 (as shown in Fig. 14). As 5 dilutions are used for each strain, the total

score was calculated out of 10 (5 dilutions multiplied by 2). Growth scores were calculated for all Yih1 proteins and other strains (GST alone, GST-Yih1, and *gcn1*Δ) used in the growth assays, as shown in Supplementary Material (Fig. S1.1). For the control plates with Galactose only (no 3AT), growth scores were calculated for two duplicate plates and the average was calculated. For the plates containing different concentrations of 3AT, one representative plate for each concentration were used for quantification and the average score was calculated. Five different 3AT concentrations were analysed (30mM, 60mM, 90mM, 120mM, and 150mM). Once the total growth scores were calculated for the growth of strains on 3AT plates, this value was then divided by the calculated value of the control plate containing Galactose only. This was to compensate for any growth defects that were unrelated to the amino acid starvation response. The total growth scores of strains on 3AT plates were also divided by the growth scores of wild type strains (GST alone or GST-Yih1) to compare the effect of mutating specific amino acids on Yih1 function. Normalised growth scores were then plotted in a bar graph, and raw data was collated in Supplementary Material (Figs S1.1, S2.2, S3.1. S4.1, S5.1).

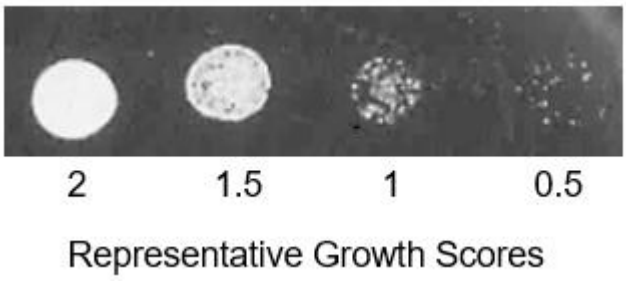
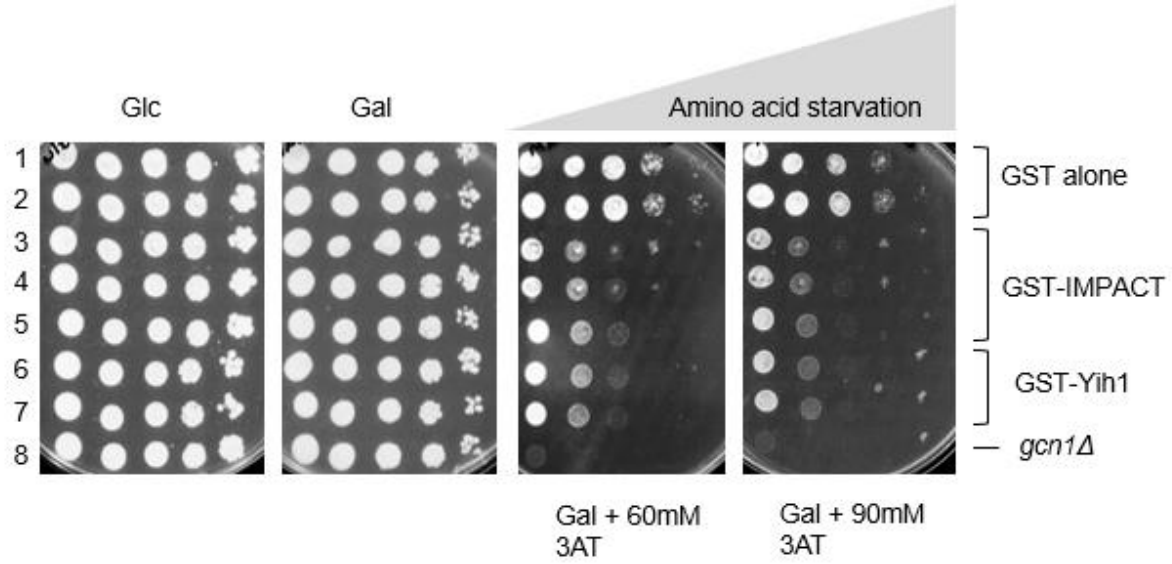


Figure 14: Representative quantitative growth scoring system of yeast cells.

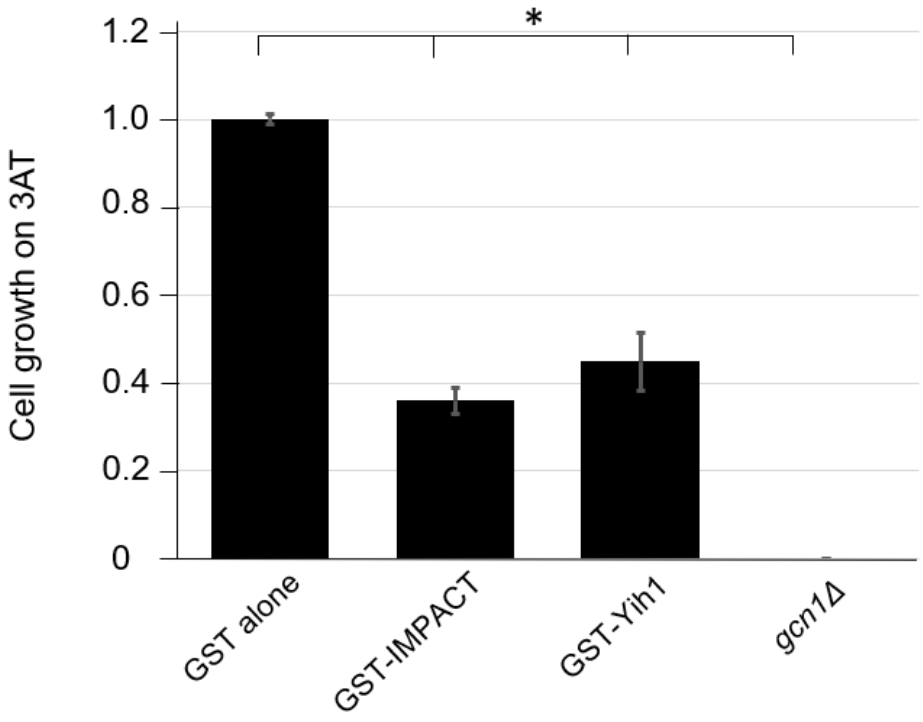
2= Full growth, 1.5= 75% growth, 1= 50% growth and 0.5= 25% growth.

3.1.5 Overexpression of Yih1 or IMPACT in yeast causes 3AT sensitivity

A



B



C

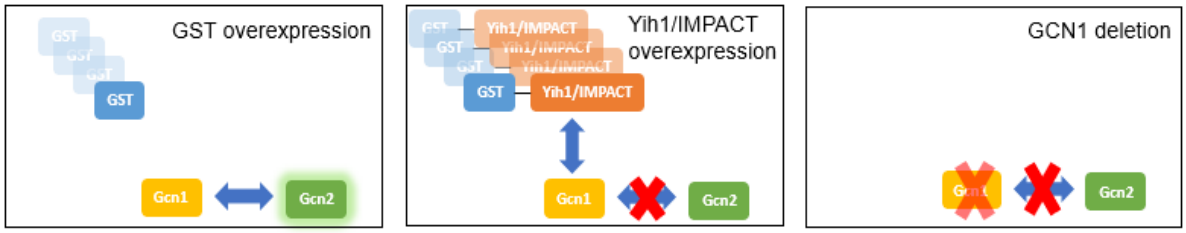


Figure 15: IMPACT or Yih1 overexpression impairs yeast cell growth under starved conditions.

A, 10-fold serial dilutions of overnight cultures from a YIH1Δ strain (ES11001B) expressing either GST alone, GST-IMPACT, or GST-Yih1 under a galactose-inducible promoter from plasmids pES128-9, pES234-6-2, and pES187-b1, respectively, were prepared (as described in Materials and Methods). A *GCN1* deletion strain (H2256) was also used as a control (*gcn1Δ*). 5uL of each dilution was transferred onto SD plates containing either Glucose or Galactose alone, as well as Galactose and different concentrations of 3AT (as described in Materials and Methods). The ability of yeast strains to grow on starvation medium (+3AT) was then observed over 14 days (Supplementary Material, Fig. S1.1). B, Bar graph of growth scores for each strain in (A). Quantitation of these results were carried out as described in Section 3.1.4. Full calculations are supplied in Supplementary Material, Fig. S1.1. P values greater than 0.05 are indicated with '*'. C, Left panel: GST overexpression does not affect the Gcn1-Gcn2 interaction, allowing Gcn2 to be activating and cells to grow in starved conditions. Middle panel: Yih1/IMPACT overexpression impairs that Gcn1-Gcn2 interaction by binding Gcn1, meaning Gcn2 is not activated. Right panel: Gcn2 cannot be activated without the Gcn1-Gcn2 interaction.

Yih1 and IMPACT overexpression has been shown previously to impair yeast cell growth in starved conditions, implying that Gcn2 activation was affected (Pereira et al., 2005). Because of this, yeast cell growth can be analysed as a quantitative measure of Gcn2 activity. We can utilise this even further to understand the molecular mechanisms of Yih1/IMPACT by examining the inverse correlation between Gcn2 activation and Yih1/IMPACT overexpression. The aim here was to first use these plasmids to test/verify the yeast system that is going to be used in this study.

As described in Section 1.5, IMPACT is the mammalian counterpart of Yih1 and previous research has shown that Yih1 and IMPACT are functionally homologous, as they both inhibit Gcn2 (Sattlegger et al., 2004; Pereira et al., 2005; Sattlegger et al., 2011). To test if the Yih1/IMPACT protein constructs used in this study were able to elicit the same impairment of Gcn2 activation, a SQGA was carried out, as described in Section 3.1.3. As depicted in the two left boxes of Fig. 15A, all strains expressing the respective proteins were able to grow at a similar rate on non-starvation media with either Glucose or Galactose, ensuring that none of the strains were experiencing growth defects. As seen in Lanes 1 and 2, strains expressing GST alone were able to grow on starvation medium (Fig. 15A). Furthermore, a yeast strain with a

GCN1 deletion was unable to grow on starvation medium (Fig. 15A, Lane 8). This is expected as the Gcn1-Gcn2 interaction is essential for cells to overcome amino acid starvation and grow. As shown in Lanes 3-5, overexpression of IMPACT severely impaired cell growth, meaning that these cells demonstrated a 3AT sensitivity phenotype (Fig. 15A). This 3AT sensitivity was also demonstrated by yeast strains overexpressing Yih1, indicating that Yih1 and IMPACT can both inhibit Gcn2 when overexpressed (Lanes 6 and 7, Fig. 15A). Furthermore, the differences in growth between GST alone and IMPACT or Yih1 were statistically significant ($P \leq 0.05$) (Fig 15B). Once confirming that Yih1 and IMPACT both inhibit Gcn2, we then aimed to determine the expression levels of these proteins, to validate their expression within the cell and also test for any differences in expression levels between strains.

3.1.6 Determining expression levels of Yih1/IMPACT

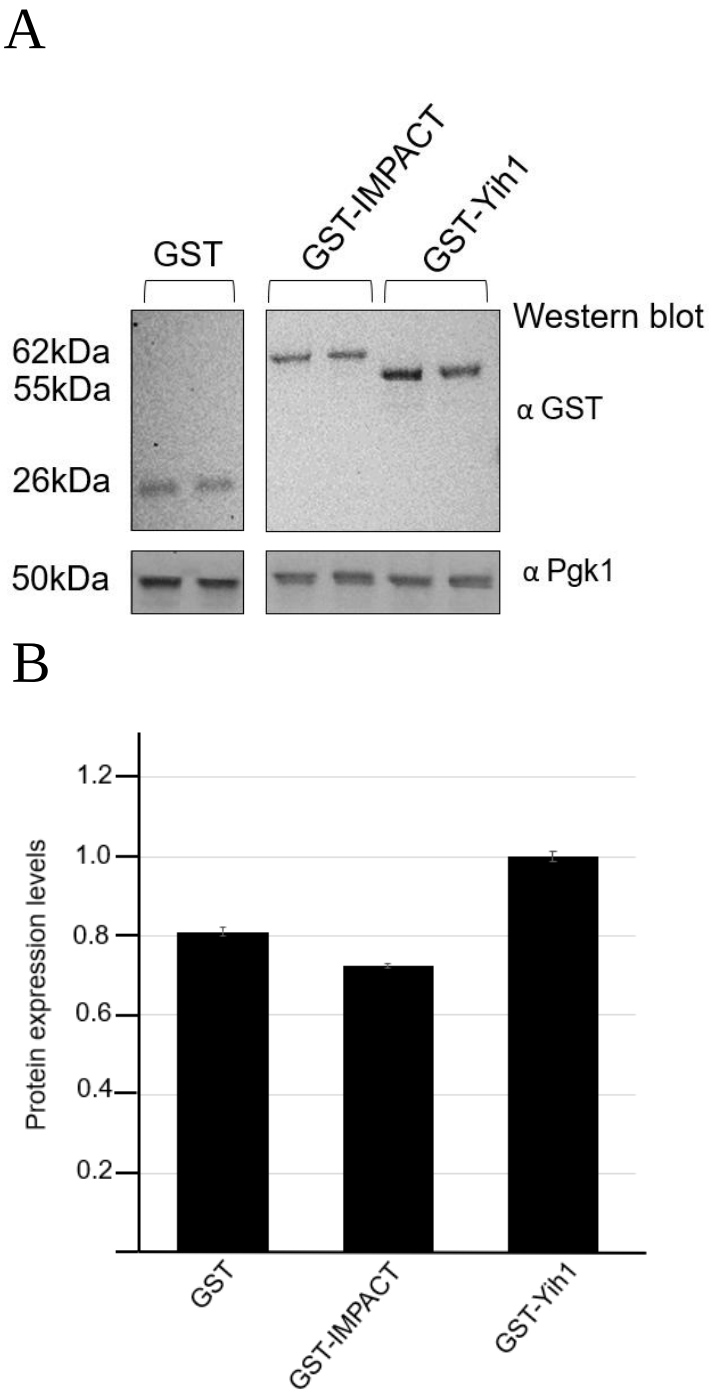


Figure 16: Expression levels of Yih1/IMPACT used in this study.

A: Yeast cells (ESY11001B background) expressing GST alone, GST-Yih1, or GST-IMPACT, under control of a galactose-inducible promoter were grown to exponential phase in SDWILV + galactose. Whole cell extracts were subject to immunoblotting using antibodies against GST (top two panels) and Pgk1 (bottom two panels) for normalization. Proteins were detected using chemiluminescence and at least two samples were analysed. Top and bottom panels, probing for GST and Pgk1 respectively, are both cut/cropped from the same membrane. Furthermore, top panels and bottom panels probing for GST and Pgk1

are representative of the same exposure time (non-modified images of membrane supplied in Supplementary Material, Fig. S1.2A). B: Expression levels were quantified by measuring band intensity using NIH Image J software and normalised to Yih1 wild type (WT). An image with a lower exposure to this Figure was used for quantification of band intensity measurements (Supplementary Material, Fig. S1.2A). Error bars are shown as black blunt ended arrows. Each data bar is representative of an average expression level calculated from two replicate samples (Supplementary Material, Fig. S1.2).

After determining the 3AT sensitivity phenotype of strains overexpressing Yih1 or IMPACT, it was important to test the expression levels of these proteins. This was to ensure that any differences in expression levels did not account for the observed growth differences in Fig. 15. As mentioned in Section 3.1.2, the Yih1 and IMPACT protein constructs are tagged with a GST fusion protein. Therefore, a suitable way to compare the expression levels between these proteins is by carrying out a western blotting analysis using an anti GST antibody. To do this, overnight cultures of yeast strains expressing either GST only, GST-Yih1 or GST-IMPACT were inoculated in flasks containing liquid medium (as described in *Materials & Methods*) at 30°C, 160rpm. Strains were grown to an OD_{600} of approximately 0.8, and then subjected to formaldehyde cross-linking and subsequently cells were harvested. Whole cell extracts were generated and then used for SDS-PAGE and western blot analysis, as described in Materials and Methods, Section 2.4). The western membrane was probed with an anti GST antibody, to detect GST. The membrane was also probed with an anti-Phosphoglycerate kinase-1 antibody (anti-Pgk1). Pgk1 is a housekeeping gene and was therefore used as a control measure to ensure that equal amounts of protein samples were loaded onto the gel. As seen in Fig. 16, for yeast strains overexpressing GST, a band was detected at approximately 26kDa, the predicted molecular weight of GST. Furthermore, strains overexpressing GST-Yih1 produced a band at approximately 55kDa, which is the predicted molecular weight of GST-tagged Yih1. A band at approximately 62kDa was also detected in samples overexpressing GST-IMPACT, the predicted molecular weight of GST-IMPACT (Fig. 16A). This western blot was then subject to quantification using ImageLab (Supplementary Material, Fig. S1.2B). In ImageLab, a value is prescribed to each western blot band signal which represents the intensity of the band. For each band, this quantified signal for GST is divided by that of Pgk1, and then normalised to the levels of GST-IMPACT. A T-test was also carried out to determine any significant differences in expression levels between samples. As shown in Figure 16B, GST-Yih1 was slightly more overexpressed than GST only and GST-IMPACT, however the values were not significantly different ($P \geq 0.05$). Furthermore, a T test showed that there were no significant differences in expression levels between GST alone, GST-IMPACT and GST-Yih1 (Supplementary Material, Fig. S1.2C).

3.1.7 Yih1/IMPACT overexpression reduces eIF2 phosphorylation levels

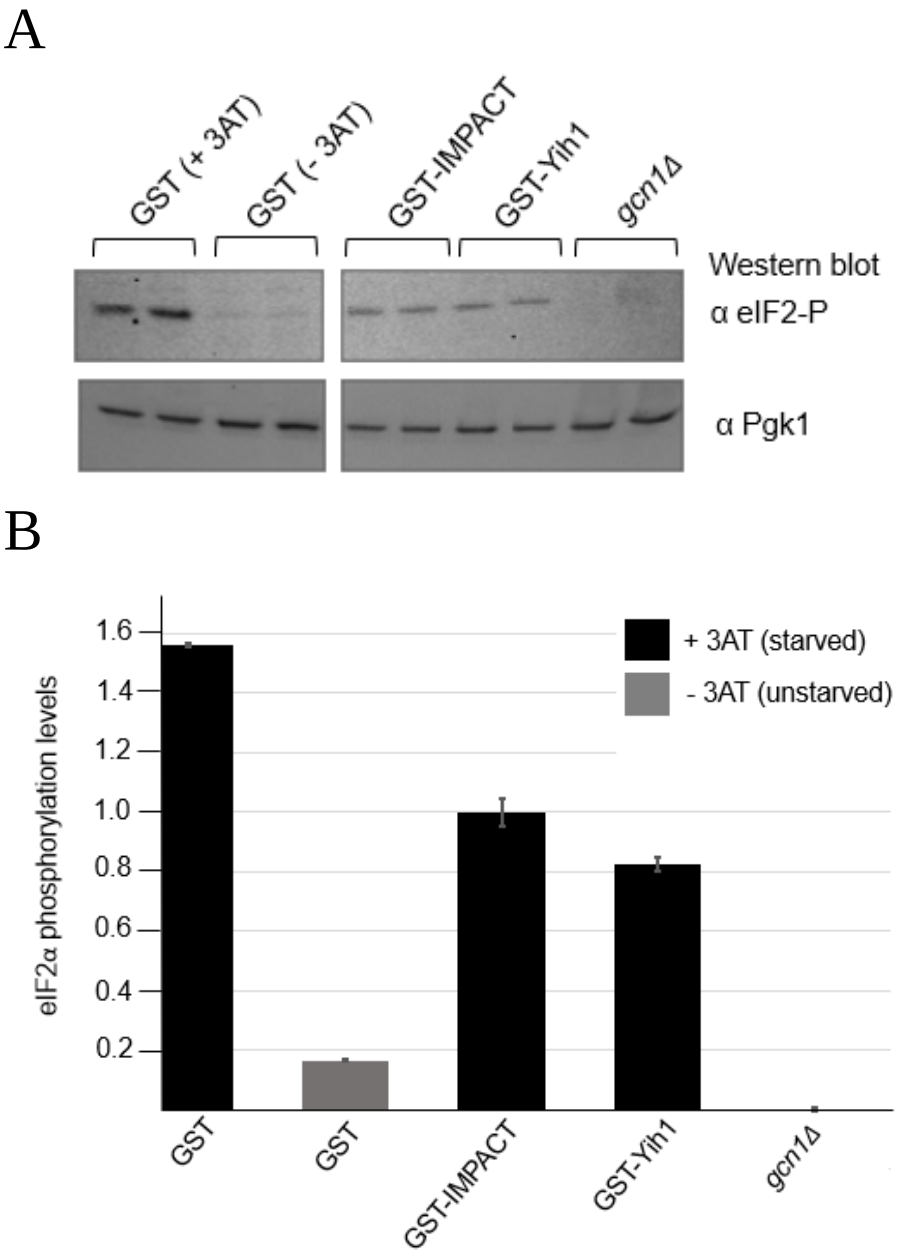


Figure 17: Yih1/IMPACT overexpression impairs eIF2α phosphorylation.

A, Whole cell extracts of GST alone, GST-IMPACT, GST-Yih1 and a *gcn1*Δ strain were subject to western blot analysis using an anti-eIF2α-P antibody to detect cellular eIF2α phosphorylation levels. Membranes were also probed with an anti-Pgk1 antibody, to detect Pgk1 to determine whether loading amounts were equal. The top and bottom panels, probing for eIF2p and Pgk1 respectively, are both appropriately cut/cropped from the same western membrane. Furthermore, top panels probing for eIF2p are representative of the same exposure time. Bottom panels probing for Pgk1 are also representative of the same exposure time. (non-modified images of membrane supplied in Supplementary Material, Fig. S1.3A). B, Band intensities from (A) were quantified using ImageJ software, and then normalised to GST-IMPACT (Supplementary Material, Fig. S1.3B).

After determining that overexpressed Yih1 and IMPACT both resulted in increased 3AT sensitivity, we next wanted to test whether this phenotype was caused by Gcn2 inhibition. A suitable way to do this is by measuring phosphorylation levels of eIF2α, the substrate for Gcn2. As Gcn2 is the only eIF2α kinase in yeast, levels of eIF2α phosphorylation levels are directly proportional to Gcn2 activation levels.

To measure eIF2α phosphorylation levels, yeast cultures expressing the indicated proteins (Fig. 17) were inoculated in liquid medium and grown to an OD_{600} of 0.8, as similarly described in the section above (Section 3.1.6). To induce amino acid starvation, 3AT was added at OD_{600} 0.6 for an exposure time of 30 minutes, and then at OD600 0.8 cells were subjected to crosslinking by treating with formaldehyde. Following crosslinking, cells were harvested and whole cell extract was generated, as described in *Materials & Methods*, Section 2.3. Proteins were then separated by SDS page gel electrophoresis and samples were subject to western blot analysis.

The membrane was probed with an anti eIF2α-P (eIF2α phosphorylated) antibody to determine levels of eIF2α phosphorylation. The membrane was also probed for Pgk1 using an anti Pgk1 antibody as a control. As mentioned previously, probing for Pgk1 enabled us to determine whether equal amounts of protein extract were loaded onto the gel. As shown in Fig. 17, under non-starved conditions (-3AT), levels of eIF2α phosphorylation were lower when compared to amino acid starved conditions. This was expected as Gcn2 is activated by amino acid starvation, meaning we would expect an increase in phosphorylation of its substrate, eIF2a. Furthermore, no eIF2α phosphorylation was detected in *gcn1Δ strains*, in line with the requirement of Gcn1 for Gcn2 activation, meaning that Gcn2 cannot be activated in a strain deleted for GCN1 (Fig. 17). As shown in Fig. 17A, and quantified in Fig. 17B, eIF2α phosphorylation levels were reduced by almost half in strains overexpressing GST-Yih1 under starved conditions. Furthermore, eIF2 phosphorylation levels were also reduced in strains overexpressing GST-IMPACT. These results are in line with previous findings that Yih1/IMPACT overexpression reduces eIF2α phosphorylation, indicated Gcn2 inhibition by Yih1/IMPACT. This result indicated that the 3AT sensitivity phenotype produced in Fig. 15 was due to Gcn2 inhibition, as demonstrated by reduced eIF2α phosphorylation levels (Fig. 17). A T-test was carried out, and no significant differences were calculated between GST alone (+3AT) and GST-IMPACT

or GST-Yih1 (Supplementary Material, Fig S1.3). However, sample size was minimal and a more comprehensive T-test would be enabled by a larger sample size.

3.2 Discussion

3.2.1 Verification of yeast system for scoring Yih1-mediated Gcn2 inhibition

Although previously published, yeast cell growth is inversely correlated to the degree of Gcn2 activation and can therefore be utilised to gauge the inhibition of Gcn2 by Yih1. We wanted to ensure that our yeast system, together with overexpression techniques, were compatible with scoring for Yih1/IMPACT-mediated Gcn2 inhibition. Together with the overexpression techniques, SQGA and western blot analysis for testing expression levels and eIF2α phosphorylation, we found consistently strong differences in phenotypes between strains overexpressing GST alone as compared to Yih1 or IMPACT. This indicated that our yeast system worked and was sufficient for understanding the inhibition of Gcn2 by Yih1/IMPACT. We next aimed to apply this yeast system to assessing the predicted amino acid groups of Yih1/IMPACT on their involvement in Gcn2 inhibition.

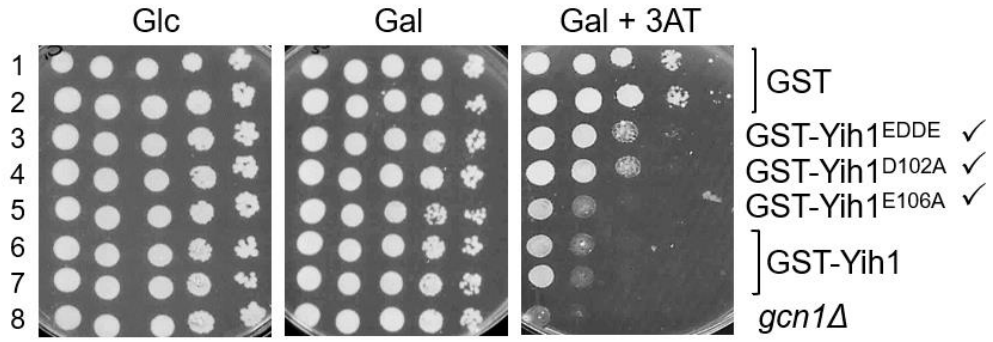
3.3 Group 1: Amino acids involved in Gcn1 binding

As it was previously published, yeast cells overexpressing Yih1 with a D102A and E106A double substitution retained the ability to grow under starved conditions, suggesting that Gcn2 could be activated (Sattlegger et al., 2011). This growth phenotype was associated with reduced Gcn1 binding to Yih1 (Sattlegger et al., 2011). This suggests that both or one of these amino acids (D102 and/or E106) are involved in Gcn1 binding, and therefore for Yih1-mediated Gcn2 inhibition. In this section, we aimed to determine whether one or both of these amino acids, D102 or E106, are necessary for mediating Gcn2 inhibition. The precise knowledge of which amino acid mediates the interaction with Gcn1 is critical for deciphering the molecular mechanisms involved in Yih1-mediated Gcn2 regulation

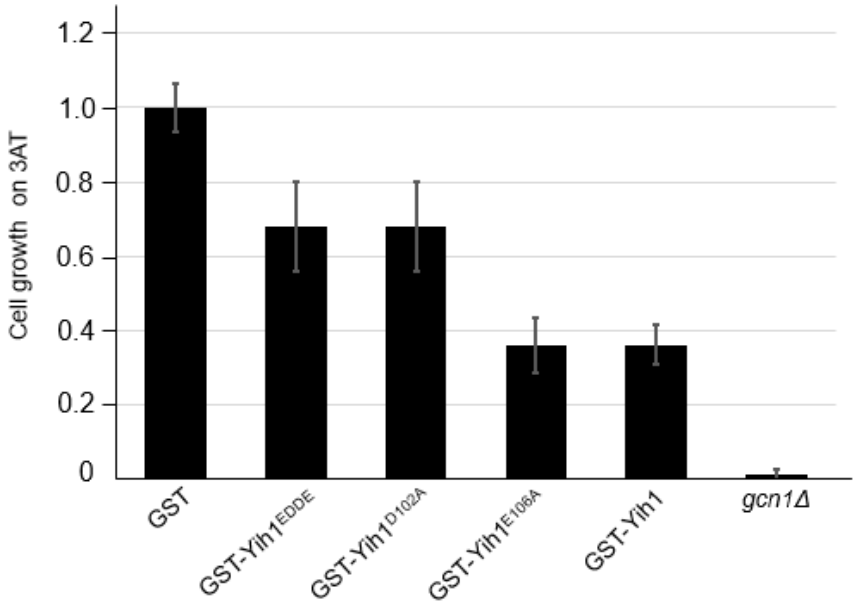
For this, plasmids containing GST-tagged Yih1 proteins with D102 or E106 substituted to Alanine (GST-Yih1^{D102A} or GST-Yih1^{E106A}) as well as an EDDE-87,90,102,106-AAA substitution (GST-Yih1^{EDDE}) were made available by Sattlegger Lab. We then utilised our established yeast system to firstly determine if mutating one of these amino acids impacts the ability of yeast cells to grow on starvation media using a SQGA.

3.3.1 Cells with a Yih1 D102A substitution retain growth under starvation conditions

A



B



C

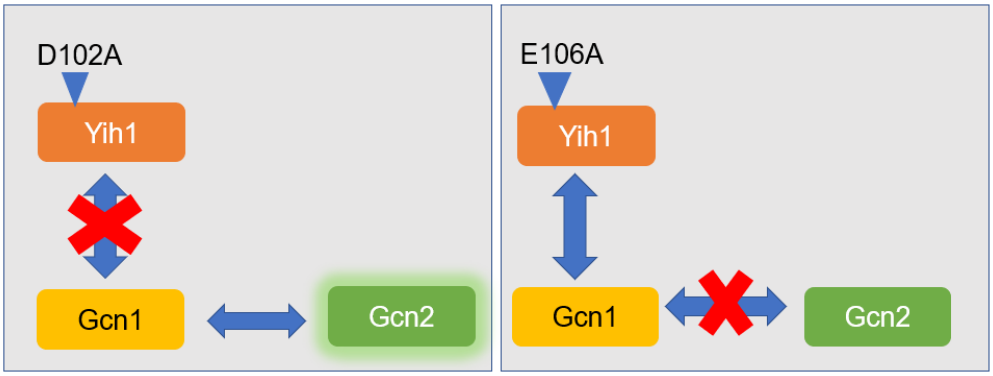


Figure 18: Yih1 Amino Acid D102 is important for inhibiting Gcn2

A: Previously transformed colonies of ESY11001B strains expressing either GST-tagged Yih1 with different mutations to the RWD domain (EDDE-86,90,102,106-AAAA, D-102-A, or E-106-A), GST alone, GST-Yih1 (wild type), or *gcn1Δ* (H2259 strain) were grown to saturation as described (Materials and Methods). The growth abilities of each strain were compared to that of strains overexpressing GST-alone of wild-type Yih1. Lack of growth on 3AT indicates that Gcn2 is unable to become activated. Growth on 3AT indicates Gcn2 activation, meaning cells can grow. For each plate, left to right: 10fold serial dilutions (1 to 1:10000). Figure is representative result of 3 independent experiments. B: Quantitative representation of the semi-quantitative growth assay shown in A. Each bar is representative of an average of two independent strains. Growth was scored using a quantitative system, described in Section 3.1.4. Values were then normalised that of the strain overexpressing GST alone. Error bars are shown as dark grey blunt end arrows. C: (Left) If Yih1 can not bind to Gcn1, then Gcn1-Gcn2 complex formation will occur, and subsequently Gcn2 can be activated. This means that eIF2 will become phosphorylated. (Right) If Yih1 can interact with Gcn1, then Gcn1-Gcn2 complex formation will be impaired, meaning Gcn2 activation is impaired. Thus, Gcn2 will not be able to phosphorylate its substrate eIF2α. Figure is not drawn to scale. ✓= sequence confirmed. × = sequence incorrect.

Based off the rationale (explained in Section 3.1.3), yeast strains expressing either GST alone, GST-Yih1 (wild type), *gcn1Δ*, or GST-tagged Yih1 proteins with mutations (as indicated in Fig. 3.3.1) under control of a galactose-inducible promoter were subject to a SQGA as described (Materials and Methods, Section 2.2.4). As seen in Fig. 18A, all strains demonstrated full growth on control plates containing glucose or galactose without 3AT treatment. These observations ensured that overexpression of the respective proteins did not affect an essential cellular function, which would have manifested itself by impaired growth under non-starved conditions. Furthermore, a yeast strain overexpressing GST alone, as a control, was able to grow on 3AT medium (3AT resistance phenotype) (Fig. 18A rows 1 and 2). This growth phenotype implied that the Gcn1-Gcn2 interaction was not impaired, meaning that Gcn2 could be activated, allowing cells to grow in starved conditions. As a negative control measure, a *gcn1Δ* strain presented severely impaired growth on amino acid starved plates (3AT sensitivity phenotype) in comparison to the strain overexpressing GST alone (Fig. 18A, row 8). This result is consistent with the expectation that Gcn2 activation requires its positive regulator (binding partner) Gcn1, so deletion of Gcn1 means that cells cannot grow in amino acid starved conditions. Furthermore, yeast cells overexpressing GST-Yih1 (wild type) demonstrated impaired growth (3AT sensitivity) upon 3AT treatment in comparison to the control GST-alone strain (Fig. 18A, rows 6 and 7). Although this impaired growth phenotype of GST-Yih1 (wild type) strains was not as severe as that of the *gcn1Δ* strain, suggesting that Gcn2 was still slightly active, it was still more sensitive to 3AT than GST alone (Fig. 18A) This result, consistent with the inhibition of Gcn2 activation by Yih1 (wild type), is supported by published work whereby

overexpressed Yih1 impairs the general amino acid control response (Sattlegger et al., 2011; Sattlegger & Hinnebusch, 2004).

Next, the growth phenotypes of strains overexpressing Yih1 with different mutations to the RWD domain (as described in Fig. 18) were analysed. Besides D102 and E106, additional RWD domain amino acid residues that have also been proposed to be involved in the Yih1-Gcn1 interaction are Glu-86 (E86) and Asp-90 (D90) (Harjes et al., 2021; Sattlegger et al., 2011). Therefore, in addition to the single D102A and E106A Yih1 mutant, a strain overexpressing Yih1 with alanine substitution mutations to E86, D90, D102, and E106 (GST-Yih1^{EDDE}) was also tested for growth under starved conditions. As seen in Fig. 18, yeast cells overexpressing GST-Yih1^{EDDE} retained their ability to grow on starvation medium with respect to strains overexpressing GST-Yih1 (wild type), despite being slightly more 3AT resistant in comparison to the strain overexpressing GST-alone (Fig. 18A and B). This finding is consistent with the suggestion that mutation of all, or one, of these amino acids in the Yih1 RWD domain leads to an impaired Yih1 function. Therefore, it is implied that either one, some, or all these amino acid residues are involved in the Yih1-Gcn1 interaction. To test whether one of these amino acids are sufficient for the Yih1-Gcn1 interaction, growth of strains overexpressing either a Yih1 containing a D102A (GST-Yih1^{D102A}) or E106A (GST-Yih1^{E106A}) substitution was analysed. Importantly, as seen in Fig. 18, yeast cells overexpressing GST-Yih1^{D102A} were able to grow under starvation conditions, therefore conferring 3AT resistance. Although this strain did not grow as well as the strains overexpressing GST-alone, it grew better than the wild type strain overexpressing Yih1 (Fig. 18A, row 4). This result suggests that the Yih1 D102A substitution impaired the Yih1-Gcn1 interaction, thereby enabling the Gcn1-Gcn2 interaction and subsequent Gcn2 activation to occur. Furthermore, the growth of yeast cells overexpressing GST-Yih1^{D102A} was comparable to that demonstrated by the GST-Yih1^{EDDE} mutant, suggesting that D102 is responsible for this growth phenotype (Fig. 18). Yeast strains overexpressing a GST-Yih1^{E106A} mutant demonstrated impaired growth under starvation conditions (3AT sensitivity), comparable to that demonstrated by wild type Yih1 (Fig. 18). This finding implies that mutation of Yih1 E106 did not impair the ability for Yih1 to bind Gcn1 and interrupt the Gcn1-Gcn2 interaction. Therefore, Gcn2 activation was impaired by Yih1, so consequently cell growth was reduced.

Results from the SQGA imply that D102 is important for Yih1 to inhibit Gcn2, as a GST-Yih1^{D102A} mutant lost its ability to inhibit Gcn2. To further verify this finding, it was important

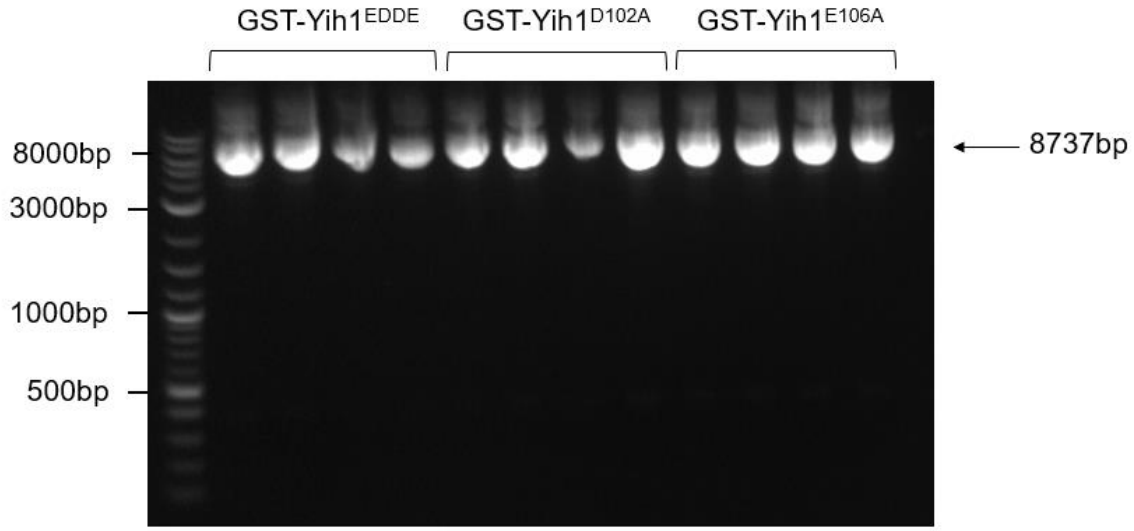
to ensure that Yih1 mutant proteins were being stably expressed, and more directly test whether Gcn2 activation was being affected by expression of these Yih1 proteins.

3.3.2 Verifying the correct substitution in Yih1 variants

Site directed mutagenesis by a PCR-based approach allows for the insertion of desired mutations into a protein sequence (Materials and Methods, Section 2.10). However, it is also possible that unexpected mutations are inserted throughout this process, thereby confounding results produced by yeast cells expressing these proteins. Therefore, after determining the phenotypes of yeast strains overexpressing GST-Yih1 with either a D102A, E106A, or EDDE mutation, we next wanted to check that the desired mutations were correctly inserted into Yih1 by Site directed mutagenesis, which was carried out prior to this study. This was to ensure than the phenotypes gathered in Fig. 18 were representative of overexpressing these Yih1 mutants.

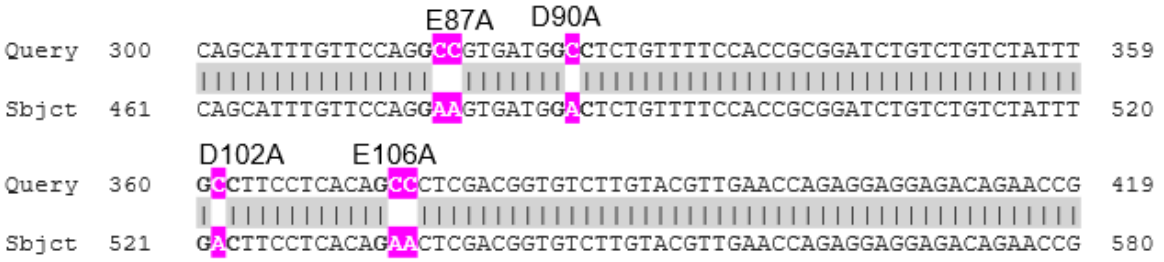
To do this, overnight cultures of yeast strains expressing either GST-Yih1^{EDDE}, GST-Yih1^{D102A}, or GST-Yih1^{E106A} were grown to saturation in liquid SDWILV medium. The plasmid DNA was then isolated from yeast using the Zymoprep Yeast Plasmid Miniprep I Kit, based off the alkaline lysis method (Materials and Methods). Following this, the isolated yeast plasmids were then transformed into the *E. coli DH5α* strain, for further plasmid isolation out of *E. coli*, for sequencing. Presence of the isolated plasmids were confirmed by agarose gel electrophoresis, where an expected band of approximately 8737bp was produced; the size of the Yih1 pES187-B1 plasmid (Fig. 19A). The plasmid DNA of mutated Yih1 proteins were then sent for commercial sequencing (by the Macrogen DNA sequencing service), along with primers ES3035-B and ES48, which cover the area of the whole Yih1 sequence. The plasmid DNA sequence was aligned with the sequence of yeast Yih1 (UniProtKB—P25637) through BLAST (basic local alignment search tool), to compare the sequence of Yih1 with the mutant Yih1 proteins. This enabled us to determine whether the expected mutations had been correctly inserted, as well as infer whether any unexpected mutations had been incorrected introduced during Site directed mutagenesis.

A



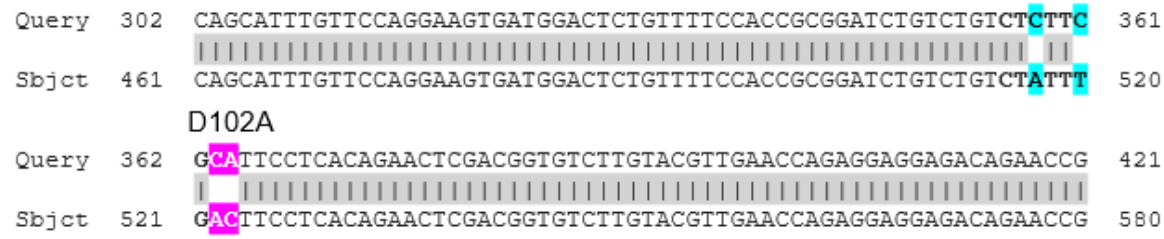
B

GST-Yih1^{EDDE}



C

GST-Yih1^{D102A}



D

GST-Yih1^{E106A}

Figure 19: Sequencing analysis of Yih1 Group 1 mutants.

A: Plasmid DNA samples of Yih1 with the indicated mutations ('A' and 'B' colonies/transformants) were isolated from E. coli and resolved on a 1% TAE agarose gel, along with a 2log DNA ladder marker to determine the size of samples. Duplicates were loaded for each colony (A,B). B-D, Nucleotide BLAST alignments; sequence on the top is the query sequence (Yih1 with the indicated mutation), and the bottom sequence is the wild type Yih1 sequence used for comparison/alignment (UniProtKB—Yih1 P25637): the area of interest is shown here, with full alignments shown in Supplementary Material, Fig S2.1. B: Nucleotide BLAST alignment for $Yih1^{EDDE}$ A and B colonies. C, Nucleotide BLAST alignment for $Yih1^{D102A}$ A and B colonies. D, Nucleotide BLAST alignment for $Yih1^{E106A}$ A and B colonies. Magenta colour indicates mutations that are expected. Turquoise colour indicates unexpected silent mutations. Red colour indicates unexpected mutations.

As seen in Fig. 19, sequencing analysis of the Yih1 mutant proteins by nucleotide BLAST alignment comparison with wild type Yih1 indicated that all expected mutations were correctly inserted into the plasmids containing GST-$Yih1^{D102A}$, GST-$Yih1^{E106A}$ and GST-$Yih1^{EDDE}$. As shown in Fig. 19B, no undesired mutations were introduced into the plasmid containing GST-$Yih1^{EDDE}$ during Site directed mutagenesis, suggesting that the phenotype that was produced by this strain represented the effects of the EDDE mutations alone. Analysis of the alignments of the sequences of Yih1 with single Group 1 mutations showed that undesired mutations were inserted into the sequences of GST-$Yih1^{D102A}$ and GST-$Yih1^{E106A}$ (Turquoise highlighted nucleotides). In saying this, these mutations were silent mutations, and do not change the amino acid sequence of the protein, therefore these silent mutations were unlikely to have confounded the results gathered.

After confirming that the correct desired mutations were present, we then wanted to further utilise our yeast system by testing the expression level of these Yih1 mutant proteins. This was to ensure that these proteins are being expressed within the living cells and allows us to gauge the expression levels of these proteins.

3.3.3 Verification of Yih1 protein expression levels

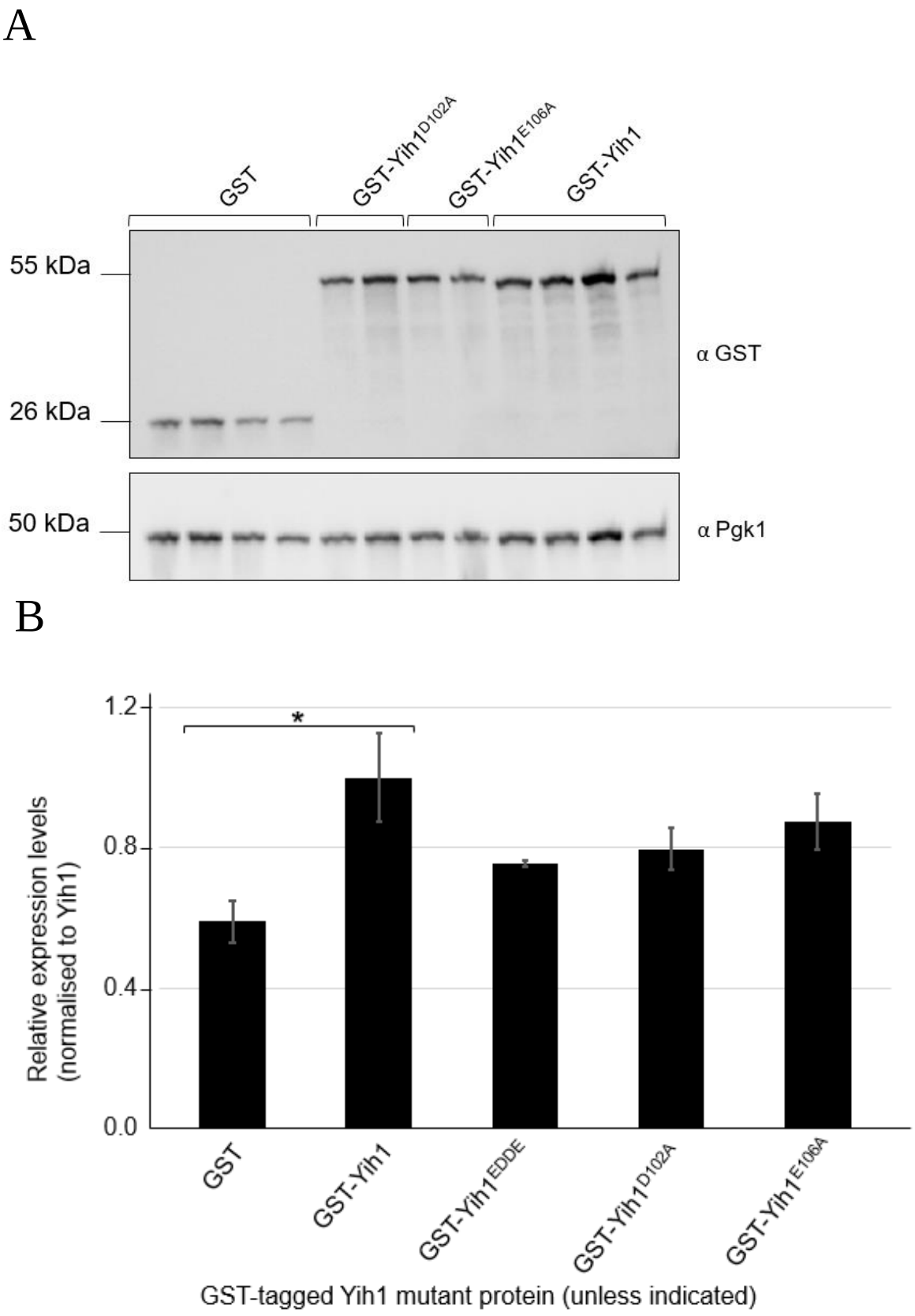


Figure 20: Expression levels of Yih1 Group 1 mutants

A: Yeast cells (ESY11001B background) expressing the indicated GST-tagged Yih1 proteins (or GST alone as a control) under control of a galactose-inducible promoter were grown to

exponential phase in SDWILV liquid medium with galactose as a carbon source. Whole cell extracts were subject to immunoblotting using antibodies against GST and Pgk1 for normalization. The bottom row of bands represents levels of the control housekeeping protein Pgk1. Proteins were detected using chemiluminescence and at least two samples were analysed. B: Expression levels were quantified by band intensity measurements using NIH Image J software and normalised to Yih1 wild type (WT). An image with a lower exposure to this Figure was used for quantification of band intensity measurements (Supplementary Material, Fig. S4). Error bars are shown as black blunt ended arrows. Each data bar is representative of an average expression level calculated from two replicate samples (Supplementary Material, Fig. S2.3). Statistically significant data ($P \leq 0.05$) is indicated by a black line with a '*' on top.

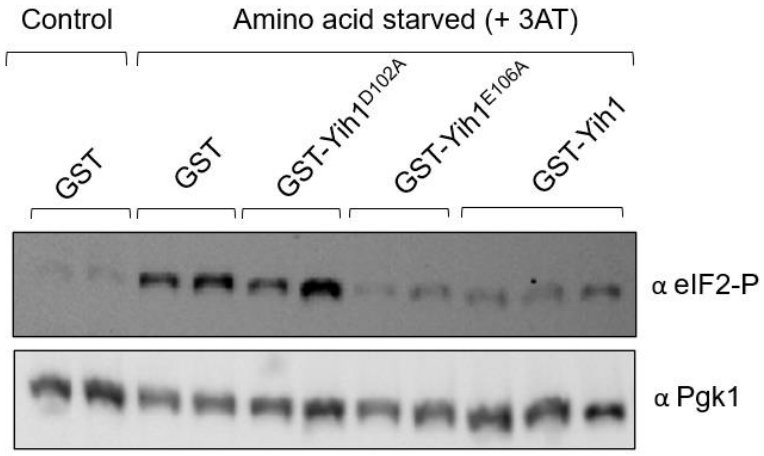
Mutation of a protein can destabilise the structure, making it more susceptible to protein degradation. Before continuing with the studies, it was important to verify that the mutations to the Yih1 protein structure did not affect the stability of the Yih1 protein. This was to ensure that the Yih1 mutant proteins were able to be expressed stably inside the yeast living cell. As the Yih1 proteins are fused with a GST tag, expression levels of GST were analysed in yeast wild-type ESY11001B strains comprising either GST-alone, GST-Yih1 (wild type) or GST-Yih1 (harbouring the Group 1 mutations indicated in Fig. 18). For this, cells were grown to exponential phase (OD_{600} 0.6-0.8) and subsequently harvested for generation of whole cell extract. Equal amounts of whole cell extract were resolved by SDS-PAGE, and samples were then subject to immunoblotting (western blot) using antibodies against GST (as described in Materials and Methods, Section 2.4). The immunoblot was also probed for the housekeeping gene Pgk1 as a loading control. This ruled out the possibility of unequal loading accounting for any differences in expression levels.

As it can be seen in Fig. 20, proteins were separated according to their molecular weight. A molecular weight marker of known protein sizes was loaded onto the gel to determine size. As shown by the bands in lanes labelled 'GST', GST alone has a lower molecular weight (approximately 26kDa) than GST-tagged Yih1 proteins (Fig. 20A and B). All GST-Yih1 proteins showed a distinctive band at approximately 55kDa, the molecular weight of GST-Yih1 (26kDa + 29kDa) suggesting that they were stably expressed in the yeast ESY11001B strain (Fig. 20). In order to quantitatively compare expression levels, the band signal intensities were quantified using Image J software, and for each sample the signal intensity was normalised to Yih1. The signal intensity of the band correlates to the expression level of the protein. The values were plotted in a bar graph (Fig. 20B).

We found that mutant proteins GST-Yih1^{EDDE} and GST-Yih1^{D102A} are expressed at slightly lower levels in comparison to wild type Yih1 (Fig. 20B). It is important to consider this difference in expression levels when analysing their effect on Gcn2 activity, as the growth phenotypes demonstrated by these mutant proteins may be even stronger than it is shown in Section 3.3.1. Upon confirming that the Yih1 proteins were able to be overexpressed in ESY11001B yeast cells, further evidence was required to test the effect of these mutant Yih1 proteins on Gcn2 activation.

3.3.4 eIF2α phosphorylation is differentially affected by Yih1 mutants

A



B

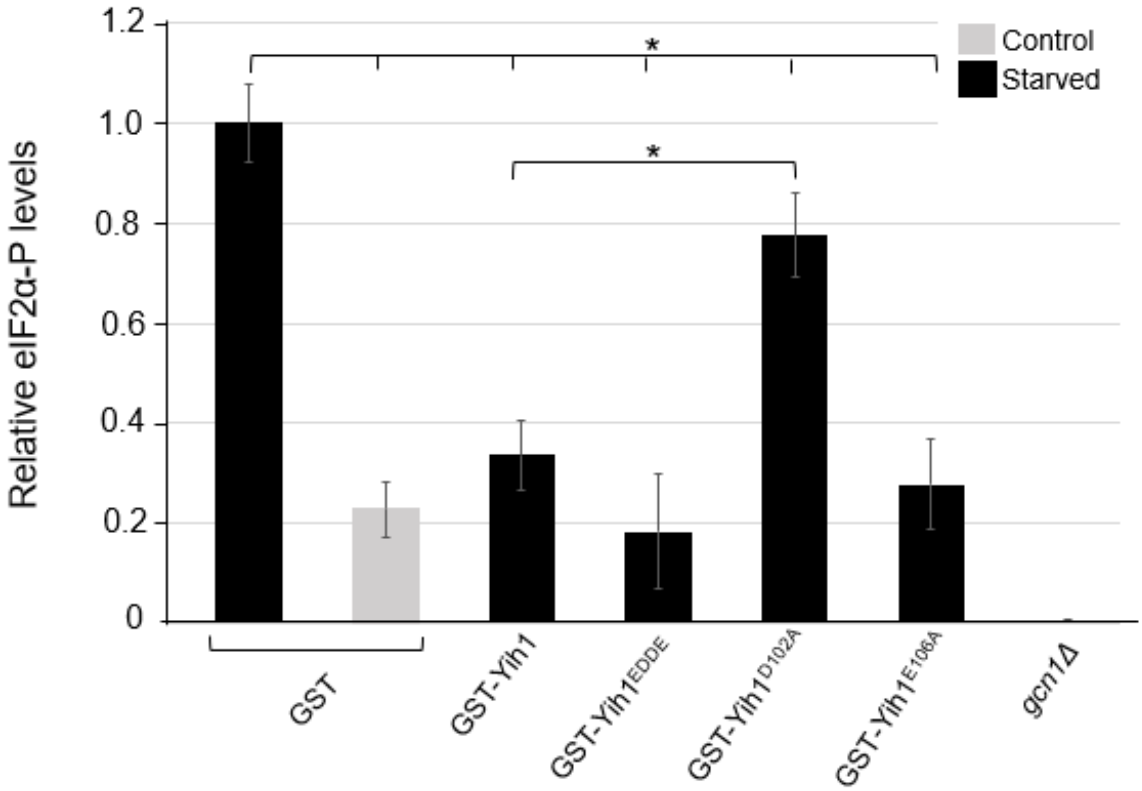


Figure 21: eIF2 phosphorylation levels are affected by Yih1 with Group 1 substitutions.

A: Whole cell extract of ESY11001B strains harbouring GST alone (control), Yih1 wild type (WT) (control), or Yih1 mutant proteins as described in Fig. 3.3.1 were prepared and subject

to immunoblotting (Materials and Methods). Antibodies against Pgk1 and eIF2 phosphorylated at Serine-51 of the α-subunit were used. Proteins were detected using chemiluminescence. At least two replicate samples were used for each data point. B: Quantification of eIF2 phosphorylation levels in (A). Data is an average from at least two replicate samples. Data was quantified by measuring band intensity and analysed using NIH Image J software and Microsoft Excel. An image with a lower intensity/exposure to this Figure was used for quantification of measurements (Supplementary Material, Fig. S2.4). The standard error is indicated by the black blunt-end bars. Data are normalised to Yih1 wild type. Statistically significant data ($P \leq 0.05$) is indicated by a black line with a '*' on top.

If Yih1 impairs the Gcn1-Gcn2 interaction, the eIF2α kinase Gcn2 cannot be activated, meaning it is unable to phosphorylate its substrate eIF2α. However, if Yih1 does not impair the Gcn1-Gcn2 interaction, then Gcn2 can be activated and is able to phosphorylate eIF2α. As Gcn2 is the only eIF2α kinase in yeast, eIF2α phosphorylation levels correlates with the levels of Gcn2 activation within the cell. Utilising this, yeast strains expressing different mutations to the Yih1 RWD domain (as described in Fig. 21) were subjected to immunoblotting in order to analyse eIF2α phosphorylation levels (as described in Materials and Methods, Section 2.4). Antibodies were used to detect eIF2α phosphorylated at the α-subunit at Ser-51 (eIF2α-P). As well as this, antibodies to detect Pgk1 were also used to normalise the eIF2 phosphorylation levels that were detected in each strain.

In accordance with previous findings, strains overexpressing GST-alone displayed strong levels of eIF2α-P that were detected upon starvation (Pereira et al., 2005; Sattlegger et al., 2011) (Fig. 21). This finding is consistent with the expectation that Gcn2 was activated and able to phosphorylate its substrate, eIF2α. Consistent with previous findings, overexpressing GST-Yih1 (wild type) lead to a dramatic reduction in eIF2α-P levels in comparison to the control GST-alone strain (Sattlegger et al., 2004; Pereira et al., 2005; Sattlegger et al., 2011) (Fig. 21). This result is in accordance with the implication that Yih1 is a negative regulator of Gcn2 that impairs the Gcn1-Gcn2 interaction, meaning that Gcn2 was inhibited and therefore could not phosphorylate eIF2α.

Interestingly, yeast cells overexpressing GST-Yih1^{D102A} experienced detectable eIF2α phosphorylation levels comparable to that of GST-alone upon starvation conditions (Fig. 21A and B). These eIF2α phosphorylation levels were significantly higher than that of GST-Yih1 (wild type) ($P \leq 0.05$). This result is consistent with the suggestion that a D102A mutation to Yih1 impaired its function in inhibiting Gcn2. Therefore, Gcn2 activation could proceed and subsequent eIF2α phosphorylation could occur. eIF2α phosphorylation levels demonstrated by

strains overexpressing GST-Yih1^{EDDE} were comparable that detected by strains overexpressing GST-Yih1^{D102A}, further supporting the responsibility of D102 for this phenotype. By contrast, strains overexpressing GST-Yih1^{E106A} resulted in a relatively reduced level of eIF2α-P detection in comparison to GST-alone (Fig. 21A and B). These eIF2α phosphorylation levels were similar to that of cells overexpressing GST-Yih1 (wild type). Furthermore, there was no significant difference in eIF2α phosphorylation between GST-Yih1^{E106A} and wild type GST-Yih1. This result implies that a E106A mutation to Yih1 did not impair its ability to inhibit Gcn2.

Student T-tests were performed to determine whether the observed differences in eIF2α phosphorylation were significant (Supplementary Materials, Fig. S2.2C*)*. For all T-tests, the null hypothesis stated that the difference between data is significant. As shown in Fig. 21B, there was a significant difference in eIF2α phosphorylation levels between GST-Yih1^{D102A} and wild type GST-Yih1. Furthermore, there was no significant difference in eIF2α phosphorylation levels between GST-Yih1^{E106A} and wild type GST-Yih1 (Fig. 21B).

Results from this western blot assay reinforces the importance of D102 of the Yih1 RWD domain in the inhibition of Gcn2. Together, results from the SQGA and western blot suggests that between D102 and E106, D102 of the Yih1 RWD domain is important for the Yih1-Gcn1 interaction, and therefore for Gcn2 inhibition.

3.4 Discussion

3.4.1 D102 is important for the Yih1-Gcn1 interaction

Yih1 and Gcn2 proteins both bind to Gcn1 by their RWD domains (Sattlegger et al., 2011). By superimposing the structurally modelled RWD domains of both Yih1 and Gcn2, D102 was found to occupy the same position as E125 in the Gcn2 RWD domain (Nameki et al., 2004; Sattlegger et al., 2011; Sattlegger & Hinnebusch, 2000). This Gcn2 amino acid is involved in the Gcn1-Gcn2 interaction (Nameki et al., 2004). Based on the structural model, it was therefore proposed that D102 may be important for the Yih1-Gcn1 interaction (Sattlegger et al., 2011).

We hypothesised that specific Yih1 amino acids are involved in Gcn1 binding to modulate Gcn2 inhibition. As mentioned prior, previous research has demonstrated that amino acids D102 and E106 are important for Gcn1-binding and for the Yih1-Gcn1 interaction (Sattlegger et al., 2011). Despite these findings, it remains to be determined whether both or just one of these amino acids are relevant for Yih1 to bind Gcn1 and inhibit Gcn2. Using our established yeast system, this work provides evidence that Asp-102 (D102) is important for Yih1 to inhibit Gcn2. This conclusion arose from the observation that overexpression of Yih1 D102A showed yeast cell growth and increased eIF2α-P levels, in comparison to wild type Yih1, in starved conditions (Fig. 18 and Fig. 21). Based on our yeast system, we predicated that if D102 is important for binding Gcn1, we would expect to see a phenotype consistent with Gcn2 activation upon substituting D102 to Alanine. This is because a D102A substitution would impair the ability of Yih1 to bind Gcn1, thereby enabling the Gcn1-Gcn2 interaction and subsequent Gcn2 activation to occur. We saw a phenotype consistent with these predictions, strongly suggesting that Yih1 D102 is important for Gcn1 binding.

An alternative explanation for the observed phenotypes could be that the mutations hindered the stability of Yih1, which may have reduced protein levels in the cell. Protein-protein interactions depend on the concentration of interaction partners, and high concentrations can drive these interactions. We found that GST-Yih1^{D102A} and GST-Yih1^{EDDE} were slightly less overexpressed than wild type GST-Yih1 (Fig. 20). To validate the expression levels of these Yih1 mutant proteins, attention was brought to findings from the determination of Yih1 fragment expression levels in Sattlegger et al. (2011). In the respective work, it was found that a Yih1 fragment spanning amino acids 2-171 was overexpressed at lower levels relative to wild type Yih1, however its expression level was still relevant for generating a phenotype (Sattlegger et al., 2011). Hence, the comparable difference in expression levels between GST-Yih1 (wild

type) and GST-Yih1^{D102A} or GST-Yih1^{EDDE} suggests that their expression levels are sufficient to represent the phenotype of these Yih1 mutants. Therefore, this phenomenon was not due to reduced expression levels.

3.4.2 Potential for salt bridge formation underlying Yih1-Gcn1 interaction

The highly conserved Gcn1 residue R2259 is essential for the Yih1-Gcn1 interaction (Sattlegger et al., 2004). R2259 and the area surrounding this residue have highly positively charged side chains (Rakesh et al., 2017). Our finding that D102, a negatively charged residue, is important for the Gcn1 binding points towards the suggestion that ionic interactions are involved in stabilising the Yih1-Gcn1 interaction.

3.4.3 E106 is not relevant for the Yih1-Gcn1 interaction or Yih1 interdomain interactions

Previous publications demonstrated that Yih1 with double D102A and E106A substitutions (GST-Yih1^{*H3}) elicited an impaired ability of inhibit Gcn2, which was associated with reduced Gcn1 binding (Sattlegger et al., 2011). If Yih1 E106 was involved in Gcn1 binding, then the same phenotype would be expected to be elicited by a single E106A substitution. However, in comparison to GST-Yih1^{D102A}, GST-Yih1^{E106A} was still able to impair Gcn2 activation as well as wild type Yih1 (Fig. 18 and Fig. 21). This was indicated by reduced yeast cell growth on starvation medium and an associated reduction in eIF2α-P levels (Fig. 18 and Fig. 21). Furthermore, there was no significant difference in eIF2α phosphorylation levels between GST-Yih1^{E106A} and GST-Yih1 (Supplementary Material, Fig. S2.4C). Together, this suggests that E106 is not required for Gcn1 binding.

Structural modelling of Yih1 showed E106 to be positioned close to the positively charged residue K173 (Fig. 30) (Harjes et al., 2021). Therefore, it was suggested that E106 may be involved in ionic interactions between the two Yih1 domains (Harjes et al., 2021). We found that overexpressing GST-Yih1^{E106A} did not influence the potency of Yih1 in comparison to wild type Yih1 (Fig. 18 and Fig. 21). If E106 is involved in interdomain interactions, a possible explanation for GST-Yih1^{E106A} retaining its ability to inhibit Gcn2 would be that by substiuting E106 to Alanine, thereby neutralising the side chain charge, the Yih1 interdomain interactions are weakened. This would expose the Gcn1-binding determinant D102, meaning that the Yih1-Gcn1 interaction could still occur. However, it can be argued that if E106 was involved in interdomain interactions, an E106A substitution should elicit a 3AT sensitivity phenotype and reduced eIF2α phosphorylation levels, as we would expect Yih1 to become more “open” and therefore more potent in Gcn2 inhibition. As we did not see a phenotype for E106A mutants, it suggests that E106 is not relevant for Gcn1 binding or for interdomain Yih1 interactions.

3.4.5 Remaining undetermined amino acids that are involved in the Yih1-Gcn1 interaction

The growth phenotype of cells overexpressing a Yih1 D102A mutant was not completely reverted (Fig. 18). This suggests that Yih1 D102A still has some affinity to Gcn1. Therefore, there must be remaining Yih1 amino acids in addition to D102 that are also involved Gcn1 binding. In order to decipher, in detail, the molecular mechanisms behind Yih1-mediated Gcn2 inhibition, we need to establish the involvement of remaining Yih1 amino acids in Gcn1 binding.

Now that we know one amino acid that is involved in Gcn1 binding, we can target this area of Yih1 to find other amino acids that are involved in this interaction. We know that the Gcn1 binding region is predicted to be within amino acids 74-114, suggesting that the important amino acids for Gcn1 binding are within this region (Sattlegger et al., 2011). Furthermore, the fully conserved Gcn1 residue Arg-2259 is essential for Yih1 binding (Sattlegger et al., 2004). Considering the charge relevance of Yih1 D102 and Gcn1 R2259, it could be speculated that the interaction between Yih1 and Gcn1 may be mediated by electrostatic interactions. By assessing the sequence of Yih1 within these parameters, potential candidates may be identified for future mutagenesis studies.

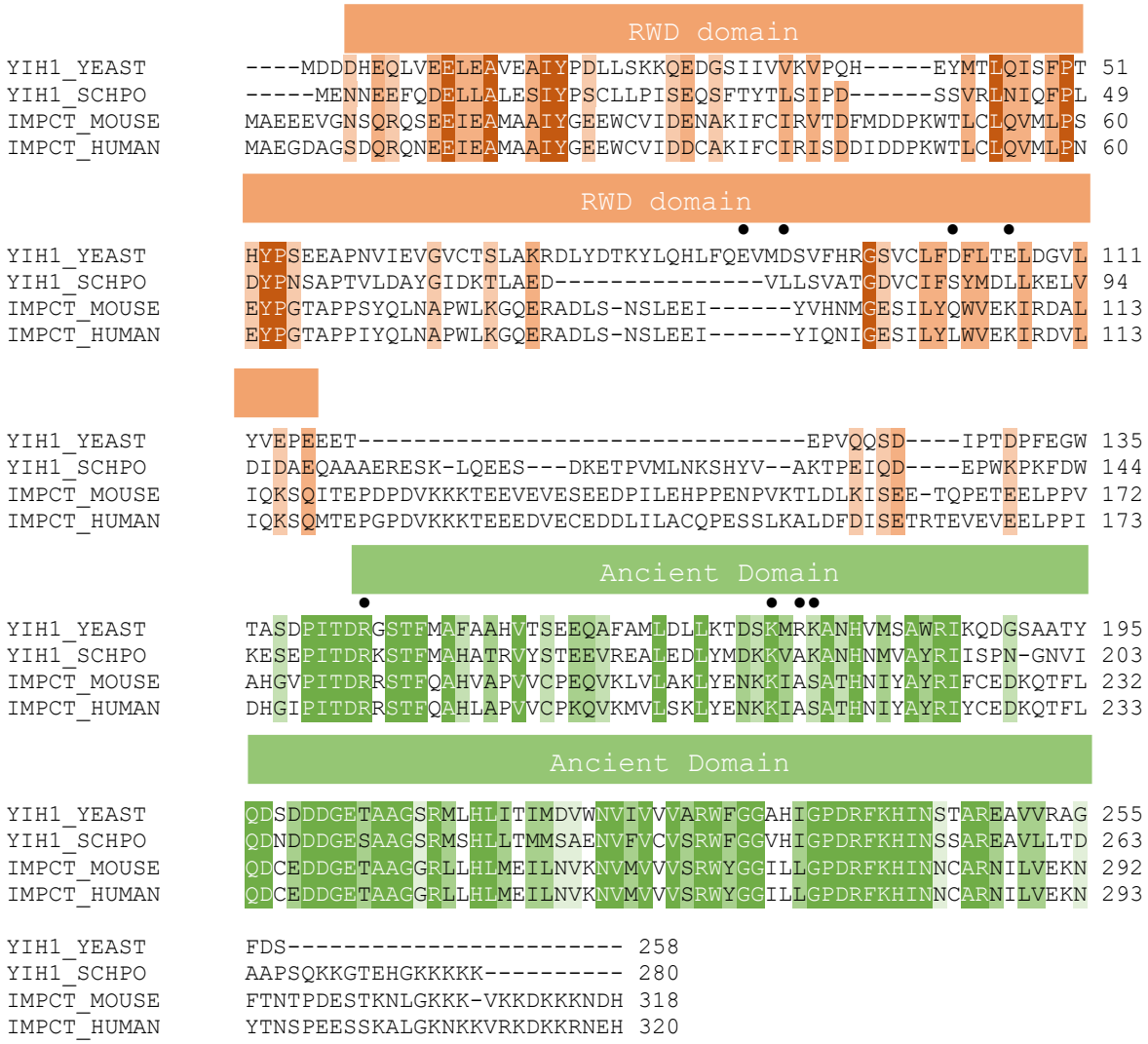
3.4.6 Future research towards determining all amino acids involved in the Yih1-Gcn1 interaction

In this section, we aimed to decipher the molecular mechanisms that underly the Yih1-Gcn1 interaction, and thereby are involved in Gcn2 inhibition. The current experiments were used as an indication for Yih1-Gcn1 binding, but future research is important to provide final evidence of the importance of D102 in Gcn1 binding. It is crucial that we understand the Gcn1 binding determinants of Yih1 as the Yih1-Gcn1 interaction is essential for Gcn2 inhibition (Sattlegger et al., 2011; Sattlegger & Hinnebusch, 2004). Future research involving physical interaction assays such as GST-pulldown assays to determine whether a Yih1 D102A substitution abolishes the Yih1-Gcn1 interaction would be insightful. This would provide direct evidence of the role of D102 in Gcn1 binding.

Future mutagenesis and screening using our established yeast system (Section 3.1) will also be insightful to identify other amino acids involved in Gcn1 binding. Upon identifying Yih1 D102 as an important amino acid for Gcn2 inhibition and for the Yih1-Gcn1 interaction, we next aimed to assess the evolutionary conservation of these amino acid residues in the Yih1/IMPACT family by a Multiple Sequence Alignment.

3.5 Examining the evolutionary conservation of Yih1/IMPACT family

A



B

```
Percent Identity  Matrix - created by Clustal2.1


  1: sp|O55091|IMPCT_MOUSE  100.00   82.39   33.07   30.32
  2: sp|Q9P2X3|IMPCT_HUMAN   82.39  100.00   31.08   29.14
  3: sp|P25637|YIH1_YEAST    33.07   31.08  100.00   41.67
  4: sp|O13997|YIH1_SCHPO    30.32   29.14   41.67  100.00
```

Figure 22: Multiple sequence alignment of Yih1/IMPACT proteins.
A: Yih1/IMPACT protein sequences from different organisms were gathered from the UniProt database, and a sequence alignment was carried out using Clustal Omega, available on the EMBL-EBI website. The RWD domain and the ancient domain are highlighted in orange and green, respectively. Black dots represent amino acids that were analysed in this

study for their involvement in Yih1-mediated Gcn2 regulation. Conservation of amino acid residues between organisms was scored using the Gonnet PAM 250 Scoring Matrix. Residues highlighted the darkest colour shade (with white writing) are fully conserved residues, while the slightly lighter highlighter shade indicates conservation between residues of strongly similar properties (>0.5 score on Gonnet PAM 250 matrix). The lightest highlighter shade indicates conservation between residues of weakly similar properties (<0.5 score on Gonnet PAM 250 matrix). B: Percent Identity Matrix was also calculated from the sequence alignment, which gives a value/score that represents the sequence similarly between two alignments. *Saccharomyces cerevisiae* (Yeast) (Yih1: P25637), *Schizosaccharomyces pombe (SCHPO) (Yih1:* O13997), *Mus musculus (Mouse) (IMPACT:* O55091), *Homo sapiens (Human) (IMPACT:* Q9P2X3).

To analyse the evolutionary relationship between the protein sequences of Yih1/IMPACT from different species, a Multiple Sequence Alignment (MSA) was carried out. Protein sequence alignments are beneficial for studying evolutionary homologies which can give an insight into important amino acids in the proteins structure or function (Tiwary, 2021). If an amino acid residue is evolutionarily conserved in the protein family, it suggests that this residue is important for the proteins overall structure or function. Protein sequences of Yih1/IMPACT from various species were gathered through the UniProt database, to which a Clustal Omega alignment was carried out on the EMBL-EBI website (Fig. 22).

3.5.1 Scoring the Yih1/IMPACT alignment using the PAM250 scoring matrix

The protein sequences of Yih1/IMPACT from different species were scored using a PAM250 scoring matrix (Percent/Point Accepted Mutations, where "Accepted" means that the mutation has been adopted by the sequence) (Dayhoff et al., 1978). PAM is an alignment scoring matrix that assigns a score taking into consideration the physical and chemical relatedness of amino acid substitutions as well as the evolutionary probability for these substitutions to occur (Dayhoff et al., 1978; Tiwary, 2021). The PAM250 matrix assumes that approximately 250 mutations occur per. 100 amino acids (Dayhoff et al., 1978). Gap penalties are also employed, which are negative penalty scores given to the overall alignment score because of a gap opening in the alignment (Nozaki & Bellgard, 2005). This results in a score that reflects the evolutionary conservation of amino acid residues. This score is represented in our MSA by the lightness/darkness of the highlighter colours orange (for RWD domain) or green (for Ancient domain) (Fig. 22A). Amino acid residues that are fully conserved across all species are highlighted the darkest tint of colour. Residues that have conservation between groups of highly similar physiochemical properties, scoring >0.5 on the PAM250 matrix, are coloured the subsequently lighter shade of colour. Residues that have conservation between groups of

weakly similar physiochemical properties, scoring <0.5 on the PAM250 matrix, are coloured the lightest shade of colour.

Percent Identity Matrix was also calculated from the sequence alignment, which gives a value that represents the overall sequence similarly between two protein alignments. This is useful for determining the evolutionary divergence of a protein as well as scoring confidence of the alignment. Protein sequence identity scores of 30% or over is considered a high sequence similarity and is therefore associated with a high significance level (Feng et al., 1985). Between Yih1 yeast and IMPACT mouse there is 33.07% sequence similarity. Between Yih1 yeast and IMPACT human there is 31.08% sequence similarity. The highest sequence similarity is between IMPACT mouse and IMPACT human, which was a score of 82.39% (Fig. 22B).

3.6 Discussion

3.6.1 Comparative analysis of our Yih1/IMPACT sequence alignment with previous alignments

The chosen scoring matrix used for comparison of sequences can strongly influence the outcome of the sequence alignment. This is because different alignment scoring matrices have different parameters. Therefore, our alignment was compared to a previous alignment carried out of Yih1/IMPACT family proteins (Sattlegger et al., 2011). As it can be seen in Fig. 22 and consistently with previous findings, the RWD domain has low sequence conservation (Harjes et al., 2021; Kubota et al., 2000; Sattlegger et al., 2011). Despite this, Yih1 and Gcn2 RWD domains have a structural similarity when superimposed (Sattlegger et al., 2011).

There are characteristic amino acid residues in our alignment as well as previous alignments that are fully conserved among RWD domain-containing proteins (Nameki et al., 2004; Sattlegger et al., 2011). Such fully conserved amino acid residues, scoring >0.5 on the PAM250 matrix, include the Glutamic acid and Glycine residue (E11 and G96 of yeast Yih1, respectively), as well as the YPXXXP motif (Fig. 22). All these residues are involved in the structural stability and conformational determination of the RWD domain (Sattlegger et al., 2011.). Based off the importance of these residues on the protein structure and function, is therefore plausible for these residues to be fully conserved among RWD domain-containing proteins. We next decided to analyse the conservation of the Yih1 ancient domain.

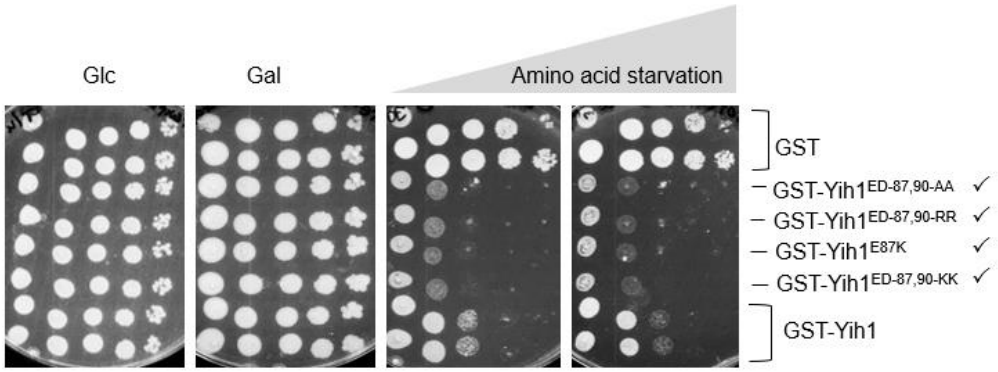
As shown in our alignment, and comparable to previous alignments, the Ancient Domain has higher sequence conservation in comparison to the RWD domain (Fig. 22) (Sattlegger et al., 2011). There are also characteristic amino acids of the Ancient Domain that are fully conserved in our alignment and in the alignment carried out by Sattlegger et al. (2011). These include the DDGE motif and the GPDRF(R/K)XIX motif (Fig. 22A) (Sattlegger et al., 2011). These fully conserved amino acid motifs are located in the loop regions of the ancient domain (Sattlegger et al., 2011). Although the Ancient domain is relatively highly conserved, its function remains unknown (Sattlegger et al., 2011). The high conservation of the Ancient domain strongly supports the notion that this domain must have an important role in the function of Yih1, such as being an interface for different binding partners (Harjes et al., 2021).

3.7 Group 2: RWD domain amino acids involved in interdomain interactions

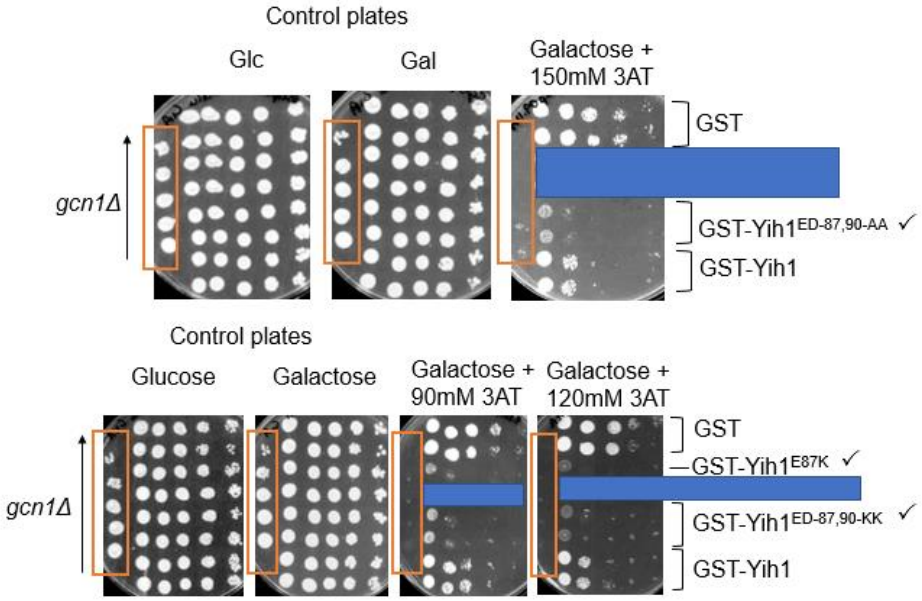
Interestingly, yeast cells overexpressing Yih1 with substitutions to both E87 and D90 (GST-Yih1*^{H2}) showed an enhanced 3AT sensitivity phenotype, which was also associated with increased Gcn1 and actin binding (Sattlegger et al., 2011). One explanation for this would be that substituting E87 and D90 with alanine may have disrupted the interdomain interactions, thereby 'opening up' Yih1 and allowing it to become more accessible to Gcn1 and actin (Sattlegger et al., 2011). Hence, we speculated that amino acids E87 and D90 may be involved in interdomain Yih1 interactions. To further clarify the role of these amino acids in Yih1-mediated Gcn2 inhibition, we first carried out a SQGA.

3.7.1 Substituting E87 and D90 to a different charge enhances Yih1 potency

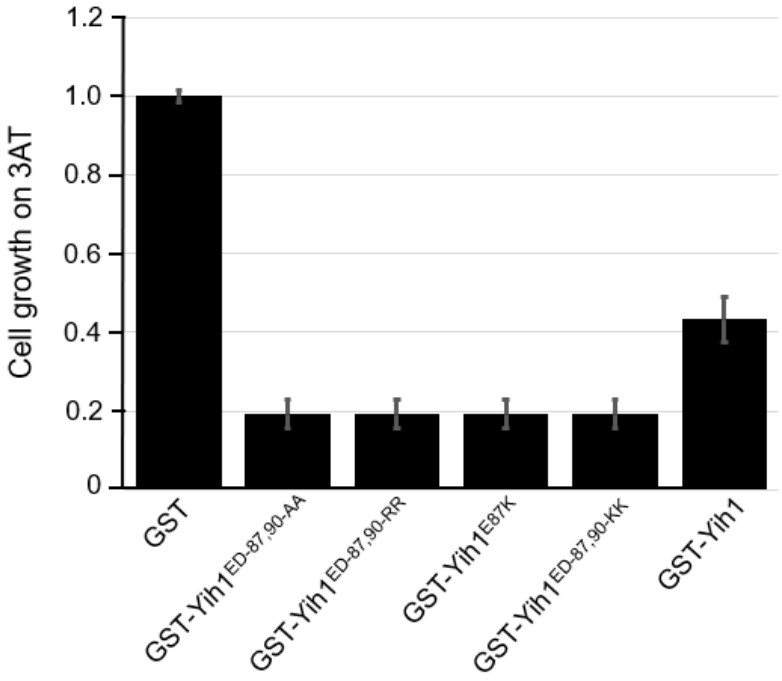
A



B



C



D

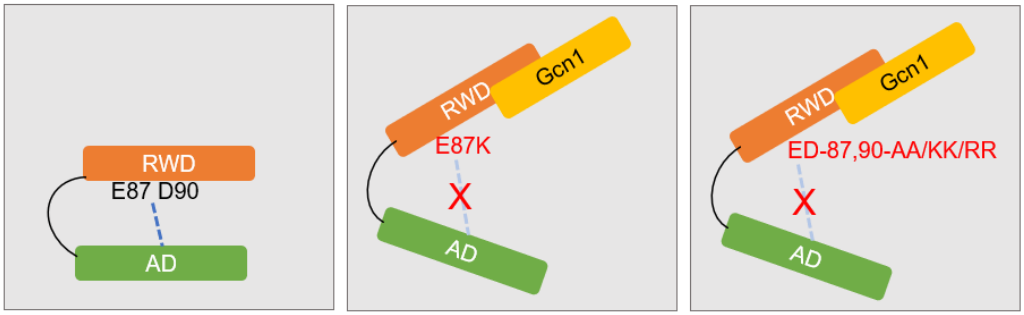


Figure 23: SQGA of Yih1 with either an E87K substitution or a E87 and D90 double substitution.

A: 5uL of 10-fold serially diluted overnight yeast cultures expressing the indicated Yih1 mutations were transferred to solid SD medium supplemented with either 2% glucose or galactose, and then galactose with increasing 3AT concentrations to induce amino acid starvation. The 3AT concentrations of the two right panels are 60mM and 90mM from left to right. The Figure is representative of two growth assays using two independent colonies/transformants. B: 5uL of 10-fold serially diluted overnight yeast cultures expressing the indicated Yih1 mutations were transferred to solid SD medium supplemented with either 2% glucose or galactose, and then galactose with increasing 3AT concentrations to induce amino acid starvation. Blue boxes represent samples that were incorrect in sequence and have been blocked out for simplicity. Full images without blue boxes are supplied in Supplementary Material, Fig. S3.4. Orange boxes highlight the growth phenotype of *gcn1Δ* strains, which represent full Gcn2 inhibition. C: Growth of yeast strains in the SQGA (from A) was scored using a quantitative system as described in Section 3.1.4. Each bar is representative of an average growth score of two independent strains. Values were

normalised to that of the strain overexpressing GST alone. Error bars are shown as dark grey blunt end arrows. D: E87 and/or D90 may be involved in Yih1 interdomain interactions. Substituting E87 alone, or E87 and D90 to a neutral or positive charge may 'open up' the Yih1 structure by disrupting interdomain interactions. RWD= RWD domain, AD= Ancient domain. E87= Glutamic acid-87, D90= Aspartic acid-90. Glc: 2% Glucose, Gal: 2% Galactose. ✓ = sequence confirmed. × = sequence incorrect

We wanted to investigate the potential involvement of the Group 2 amino acids in Yih1 function. Yeast strains expressing GST-Yih1 with a single substitution to E87 (E87K), or double substitutions to both E87 and D90, were made available to me by Sattlegger Lab. E87 and D90 were double substituted to either Alanine (A) (as published previously by Sattlegger et al., 2011), Arginine (R) or Lysine (K), to analyse the effect of changing the amino acid charge on yeast cell growth in starved conditions (Fig. 23). The ability for these Yih1 mutant proteins to impair Gcn2 activation, based off our established yeast system (as described in Section 3.1) was analysed by a SQGA.

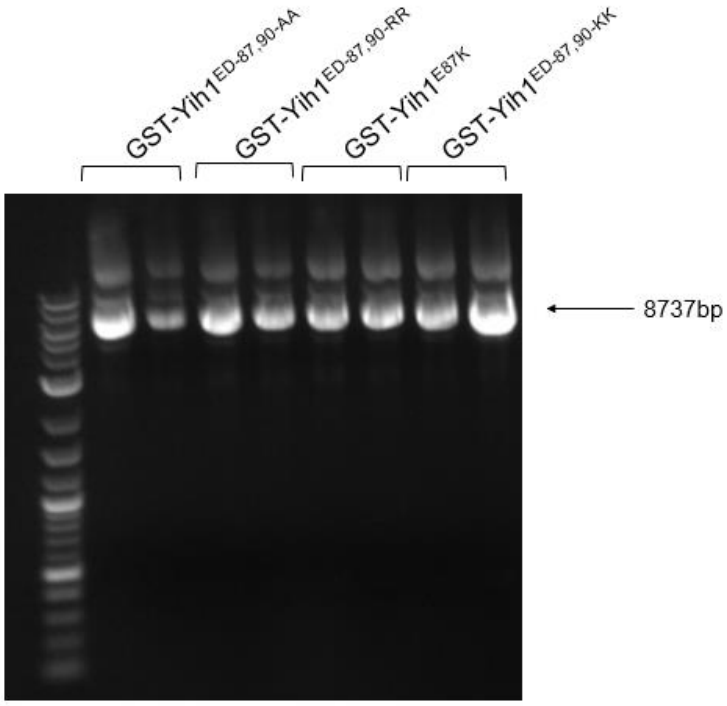
As shown in Fig. 23, all yeast strains expressing the indicated GST-tagged Yih1 mutant proteins demonstrated full growth on SD medium with either Glucose or Galactose, indicating that no growth defects were present in these yeast strains. Expectedly, strains overexpressing GST alone were resistant to amino acid starvation (3AT resistance), in line with the expectation of Gcn2 activation by amino acid starvation in these strains (Fig. 23). In comparison, strains overexpressing GST-Yih1 (wild type) displayed 3AT sensitivity, demonstrated by reduced growth on 3AT medium, as compared to strains overexpressing GST alone, implying that Yih1 was able to impair Gcn2 in these conditions, leading to reduced cell growth (Fig. 23) (Sattlegger et al., 2011). Furthermore, in a *gcn1Δ* strain, we can see that yeast cell growth on 3AT was completely abolished, in line with the requirement of Gcn2 for Gcn1 to be activated (Fig. 23B)

In line with previous publications, GST-Yih1$^{ED\text{-}87,90\text{-}AA}$ demonstrated an enhanced 3AT sensitivity phenotype in comparison to wild type Yih1 (Fig. 23). This suggests that this mutation enhanced Yih1 potency (Fig. 23) (Sattlegger et al., 2011). When comparing between the Yih1 double substitution strains, when E87 and D90 were both mutated to either Alanine, Lysine or Arginine, cell growth was equally impaired between the Yih1 mutants (Fig. 23). Furthermore, Yih1 with a single E87K substitution was sufficient for generating an enhanced 3AT sensitivity phenotype in comparison to wild type Yih1 (Fig. 23). This suggests that when overexpressed, Yih1 proteins with the indicated H2 mutations did not lose their ability to inhibit Gcn2, but rather that these mutations may have enhanced Yih1 potency in Gcn2 inhibition.

Once determining the effect of mutating these Yih1 amino acids on cell growth, and therefore on Gcn2 activation, we then aimed to verify the presence of these substitutions in Yih1.

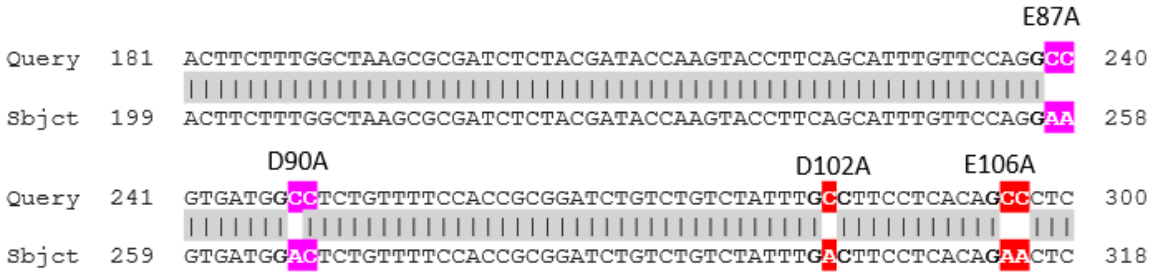
3.7.2 Verifying the correct substitution of Yih1 E87 and D90 variants

A



B

GST-Yih1$^{ED-87,90-AA}$ (A)



C

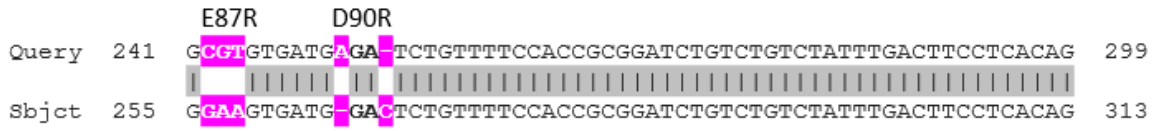
GST-Yih1$^{ED-87,90-AA}$ (B)

D

GST-Yih1$^{ED-87,90-RR}$ (A)

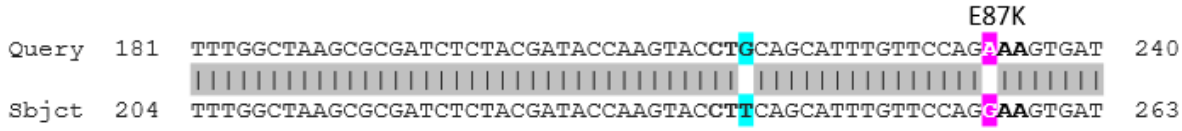
E

GST-Yih1$^{ED\text{-}87,90\text{-}RR}$ (B)



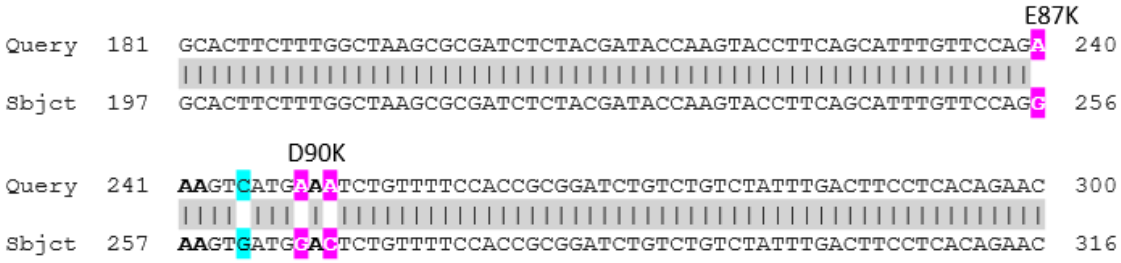
F

GST-Yih1^{E87K} (A)



G

GST-Yih1$^{ED\text{-}87,90\text{-}KK}$ (A)



H

GST-Yih1$^{ED\text{-}87,90\text{-}KK}$ (B)

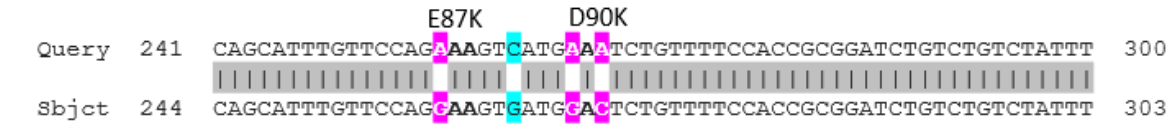


Figure 24: Verifying the presence of Yih1 Group 2 mutations by Sequencing

A: Plasmid DNA of yeast cells expressing the indicated Group 2 mutations was resolved by 1% agarose gel electrophoresis and then sent in for commercial sequencing as described in Materials and Methods. B,C: Nucleotide BLAST alignment for Yih1$^{ED\text{-}87,90\text{-}AA}$ A and B colonies. D,E: Nucleotide BLAST alignment for Yih1$^{ED\text{-}87,90\text{-}RR}$ A and B colonies. F: Nucleotide BLAST alignment for Yih1^{E87A} A colony. G,H: Nucleotide BLAST alignment for Yih1$^{ED\text{-}87,90\text{-}KK}$ A and B colonies. Query sequence: Yih1 with the indicated E87 and/or D90 mutation. Sbjct (Subject) sequence: Wild type Yih1 sequence used for comparison/alignment (UniProtKB—Yih1 P25637). The area of interest is shown here, with full alignments shown in Supplementary Material (Section S3). Magenta colour

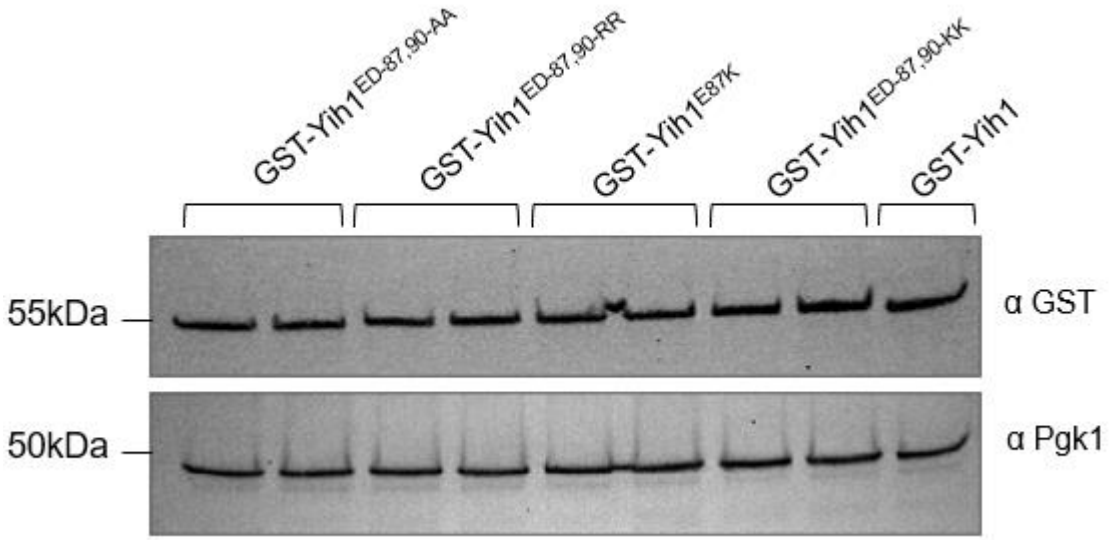
indicates mutations that are expected. Turquoise colour indicates unexpected silent mutations. Red colour indicates unexpected mutations.

After determining the phenotypes of the Yih1 E87 D90 mutant strains, we wanted to check that the desired mutations were correctly inserted into Yih1 to ensure than the phenotypes gathered in Fig. 23 were representative of overexpressing these mutant proteins. As similarly described in Fig. 19, overnight cultures of yeast strains expressing Yih1 with the indicated mutations (Fig. 24) were grown to saturation and isolated from yeast, then subsequently transformed and isolated from *E. coli* for sequencing analysis by Macrogen. These plasmid sequences were aligned with yeast Yih1 (UniProtKB—P25637) through Nucleotide BLAST and then analysed.

As seen in Fig. 24, all desired mutations were verified in each respective yeast strain, as indicated by magenta highlighted nucleotides. The only undesired mutations that changed the protein sequence was found in the A strain of GST-Yih1$^{ED\text{-}87,90\text{-}AA}$, in which a D102A and E106A mutation was also present (Fig. 24B). However, the 'B' strain of GST-Yih1$^{ED\text{-}87,90\text{-}AA}$ was sequence confirmed and did not have any undesired mutations (Fig. 24C). This meant that we could trust this sample for analysing the phenotype. After confirming the presence of the expected Yih1 E87 and/or D90 substitutions, we next wanted to test the expression levels of these proteins inside the living cell.

3.7.3 Determining expression levels of Yih1 Group 2 mutants

A



B

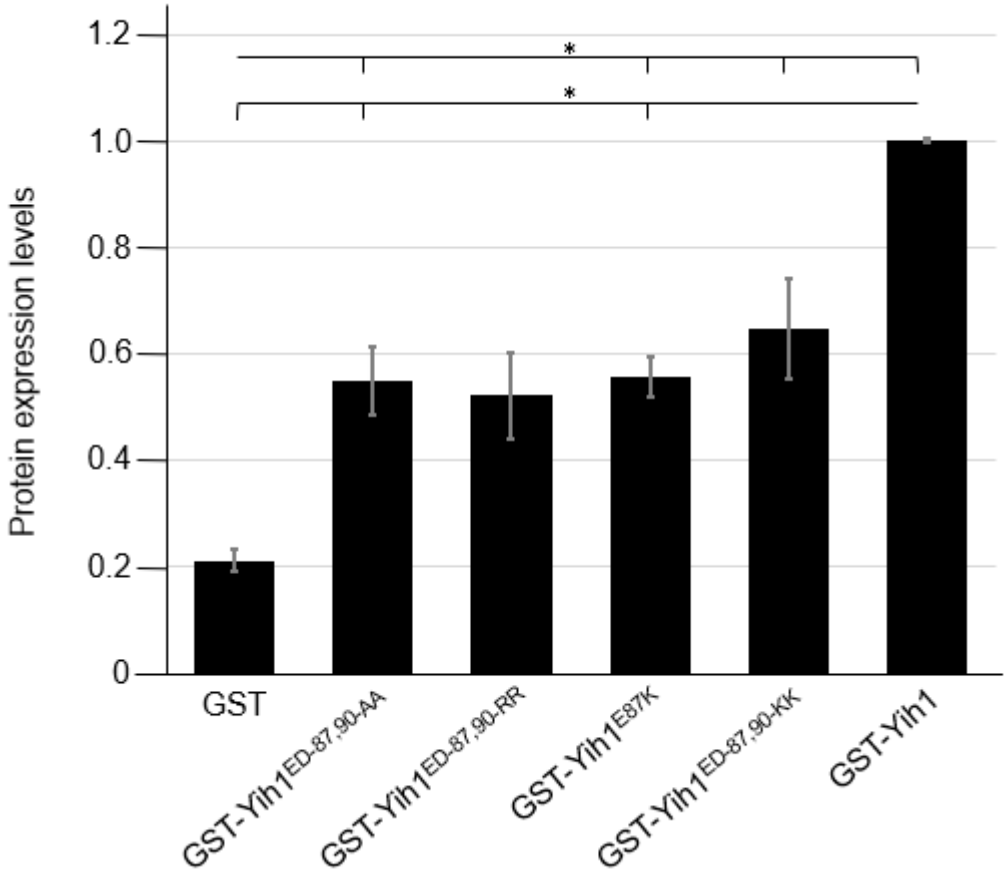


Figure 25: Expression levels of Yih1 Group 2 mutants.

A: ESY11001B (Yih1 deletion) yeast strains expressing the indicated GST-tagged Yih1 proteins (or GST alone) were grown to exponential phase and harvested for generation of whole cell extracts as described in Materials and Methods. Whole cell extracts were subject to immunoblotting using antibodies against GST and Pgk1. The bottom row of bands represents levels of the control housekeeping protein Pgk1. Proteins were detected using chemiluminescence and at least two samples were analysed. B: Expression levels were quantified by band intensity measurements using NIH Image J software and normalised to Yih1 wild type (WT). An image with a lower exposure to this Figure was used for quantification of band intensity measurements (Supplementary Material, Fig. S3.2). Error bars are shown as black blunt ended arrows. Each data bar is representative of an average expression level calculated from two replicate samples (Supplementary Material, Fig. S3.2). Statistically significant data ($P \leq 0.05$) is indicated by a black line with a '*'.

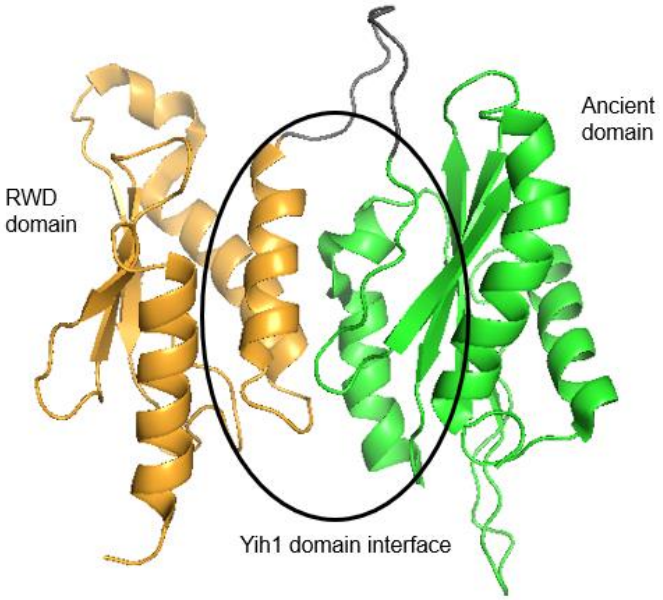
We next wanted to determine the expression levels of these Yih1 variants to ensure that mutating E87 and D90 did not destabilise the Yih1 structure and/or result in Yih1 not being expressed inside the living cell. Yeast cells expressing GST alone, GST-Yih1, or GST-Yih1 with the indicated Group 2 mutations (as described in Fig. 25) were grown to exponential phase (OD_{600} 0.6-0.8) and harvested for generation of whole cell extract. Equal amounts of whole cell extract were resolved by SDS-PAGE, separating proteins according to their molecular weight. Western blot analysis was subsequently carried out using antibodies against GST, as well as for Pgk1, to rule out the possibility of unequal loading accounting for any differences in expression levels.

Upon probing the membrane for GST, a strong band of approximately 55kDa, the expected molecular weight of GST fused to Yih1 (both wild type and mutant variants) was detected, suggesting that they were all sufficiently expressed (Fig. 25A). Furthermore, probing the membrane for Pgk1 resulted in a strong band of approximately 50kDa, the molecular weight of Pgk1 (Fig. 25A). To quantitatively compare expression levels, the signal intensities of the bands were quantified using Image J software, and for each sample the signal intensity was normalised to Yih1 (Supplementary Material, Fig. S3.2C). The signal intensity of the band correlates to the expression level of the protein. The values were plotted in a bar graph (Fig. 25B). As shown in Fig. 25B, the mutant Yih1 variants (GST-Yih1$^{ED\text{-}87,90\text{-}AA}$, GST-Yih1$^{ED\text{-}87,90\text{-}RR}$, GST-Yih1^{E87K} and GST-Yih1$^{ED\text{-}87,90\text{-}KK}$) are all approximately one third (1/3) less overexpressed that wild type Yih1. Despite being less overexpressed than wild type Yih1, these Yih1 variants were still able to produce a strong 3AT sensitivity phenotype in our SQGA, suggesting that these expression levels were sufficient for generating a phenotype.

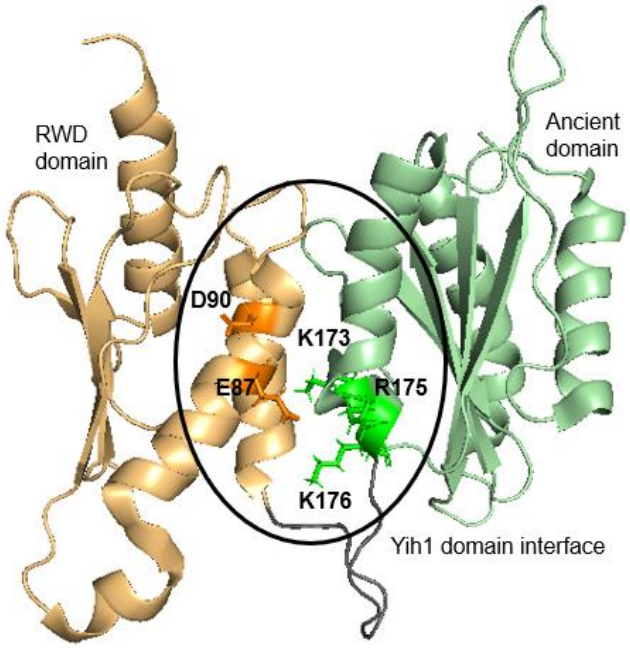
3.8 Discussion

3.8.1 The Yih1 conformation influences the potency of Yih1-mediated Gcn2 inhibition

A



B



C

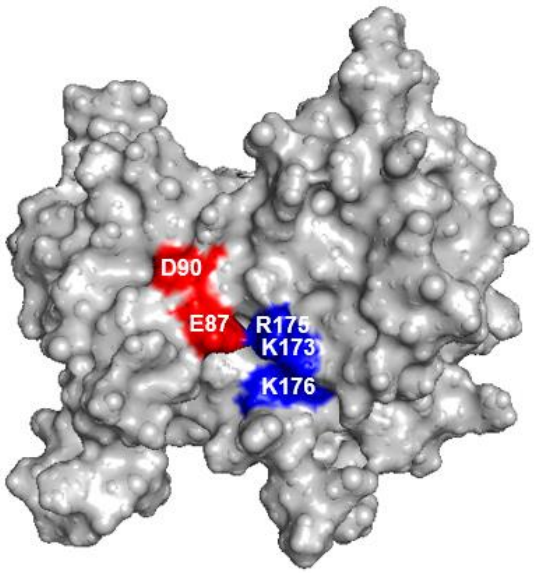


Figure 26: Structural model of the surface of Yih1.

A: Structural model of Yih1 generated in PyMOL™ version 2.3.4. The RWD is coloured in orange, the linker region is coloured in grey and the Ancient domain is coloured in green. The interface between the two Yih1 domains is circled and labelled. B: Group 2 (Orange) residues and Group 3 (Green) residues are situated at the Yih1 domain interface. C: RWD domain residues E87 and D90 (red) (Group 2 amino acids) are negatively charged. Ancient domain residues K173, R175 and K176 (blue) (Group 3 amino acids0 are positively charged. E87 and R175 are positioned close to eachother. Model was constructed in PyMOL™ version 2.3.4

Earlier studies demonstrated that a double substitution of E87 and D90 to Alanine resulted in an increased 3AT sensitivity phenotype when compared to wild type Yih1, suggesting that these mutations enhanced Yih1 potency (Sattlegger et al., 2011). E87 and D90 are negatively charged, and it was proposed that they mediate ionic interactions between the RWD and ancient domain (Harjes et al., 2021; Sattlegger et al., 2011). Furthermore, E87 and D90 are spatially positioned close to positively charged Ancient domain residues, which also supports the notion of interdomain electrostatic interactions (Fig. 26C). Therefore, it was reasoned that substituting E87 and D90 to Alanine would weaken interdomain interactions between the RWD and ancient domain, thereby changing the Yih1 conformation to a more ‘open’ conformation which is more primed for Gcn1 binding and therefore Gcn2 inhibition (Harjes et al., 2021; Sattlegger et al., 2011).

The role of Yih1 interdomain interactions is not yet clear in regard to Yih1-mediated Gcn2 inhibition. To get more insight into this, it is important to experimentally evaluate whether one or both RWD domain amino acids are necessary for mediating interdomain interactions.

Upon determining that Yih1 with E87 and D90 substituted to Alanine enhances Yih1 potency, consistent with previous published research, we next analysed the effect of substituting E87 and D90 to oppositely charged residues Lysine (K) or Arginine (R) (Sattlegger et al., 2011). The rationale behind choosing a Lysine or Arginine substitution was that if the interdomain Yih1 interactions are mediated by ionic interactions, we would expect that substituting these residues to the opposite charge would result in an electrostatic repulsion between the Yih1 domains. This would "open up" the Yih1 structure and therefore expose the Gcn1 binding determinants, resulting in stronger Gcn2 inhibition.

We found that substituting the negative charge of both E87 and D90 to a positive charge increased Yih1 potency during starved conditions, as found for an Alanine substitution, suggesting that Gcn2 inhibition is enhanced in either scenario (Section 3.7.1). Interestingly, there was no difference in Yih1 potency between an Alanine substitution or a Lysine (or Arginine) substitution, which will be further discussed in Section 3.8.2. As it was predicted, a possible explanation for these results is that changing the charge of these amino acids changed the Yih1 structure by electrostatic repulsions weakening the interdomain interactions. Accordingly, this would expose the Gcn1 binding determinants and thereby allow better access of Yih1-binding partners to their respective binding sites. Therefore, our results in the living yeast cell suggests that one or both of E87 and D90 are involved in Yih1 interdomain interactions.

Overexpression of wild type Yih1 lead to reduced yeast cell growth by approximately half the amount of that by overexpressing GST alone (Fig. 23C). This implies that Gcn2 activation was reduced by approximately half the amount when overexpressing GST-Yih1. By comparing the phenotype of cells overexpressing GST-Yih1 to cells with a *gcn1Δ*, we can see that GST-Yih1 does not completely inhibit Gcn2, as there is still some cell growth in starved conditions. However, overexpression of Yih1 with the indicated Group 2 mutations resulted in a further reduced cell growth by approximately 3/4 of that by overexpressing GST alone (Fig. 23C). Furthermore, in the Yih1 mutant strains, Gcn2 activation was reduced by approximately 1/2 more of that by wild type Yih1 (Fig. 23). A plausible reason for this is that wild type Yih1 is in a slightly open conformation, with some interdomain interactions still intact, making the Yih1

potency half of its potential inhibition. However, with the Group 2 mutations, the Yih1 protein structure is forced into even more of an open conformation due to structural relaxation caused by electrostatic repulsions between domains. Accordingly, this would result in stronger Gcn2 inhibition in these Yih1 mutants when compared to wild type Yih1. In other words, it implies that the relative orientation of the RWD domain and ancient domain determines the potency of Yih1.

3.8.2 E87 and D90 are the major players in interdomain interactions

We can see that when mutating E87 and D90 at the same time, the charge, whether it is changed to neutral or the opposite charge, is redundant on the effect of Yih1 potency (Fig. 23). If there are additional RWD domain amino acids other than E87 and D90 involved in ionic interactions, we would expect that mutating these to Lysine or Arginine would cause a stronger 3AT sensitivity phenotype than mutating them to Alanine. This is because it is likely that electrostatic repulsions would disrupt any additional nearby interactions, forcing Yih1 into a more 'open' conformation, resulting in stronger Yih1 potency than that of an Alanine substitution. Consistent with this idea, it also would be expected that mutating E87 and D90 to Alanine would not have as much of an effect on Yih1 potency. This is because the neutral charge of Alanine would not cause strong electrostatic repulsion, and any additional amino acids involved would rescue the interdomain interactions, resulting in a phenotype like that of wild type Yih1 (Fig. 23). As we see a comparable 3AT hypersensitivity phenotype between all E87 D90 mutants, it suggests that in the RWD domain, E87 and D90 are the major players in Yih1 interdomain interactions.

3.8.3 Potential role of E87 in Yih1 interdomain interactions/Yih1 function

As mentioned prior, to get a detailed understanding of the role of interdomain interactions in Yih1 function, it is important to experimentally evaluate whether one or both of E87 or D90 are necessary for mediating interdomain interactions. Interestingly, an enhanced 3AT sensitivity phenotype was demonstrated by yeast cells overexpressing Yih1 with an E87K substitution (Fig. 23). This suggests that substituting this amino acid alone was sufficient for enhancing Yih1 potency. Furthermore, this suggests that E87 also may have an integral role in Yih1 interdomain interactions, and thereby for Yih1-mediated Gcn2 inhibition. E87 is positioned closely to R175, in the ancient domain, in the computational structural model of Yih1 (Fig. 27). In the structural model of Yih1, the distance between E87 and R175 is approximately 4.3 angstroms (Fig. 27) These residues were suggested to contact each other through ionic interactions such as interdomain salt bridges (Harjes et al., 2021).

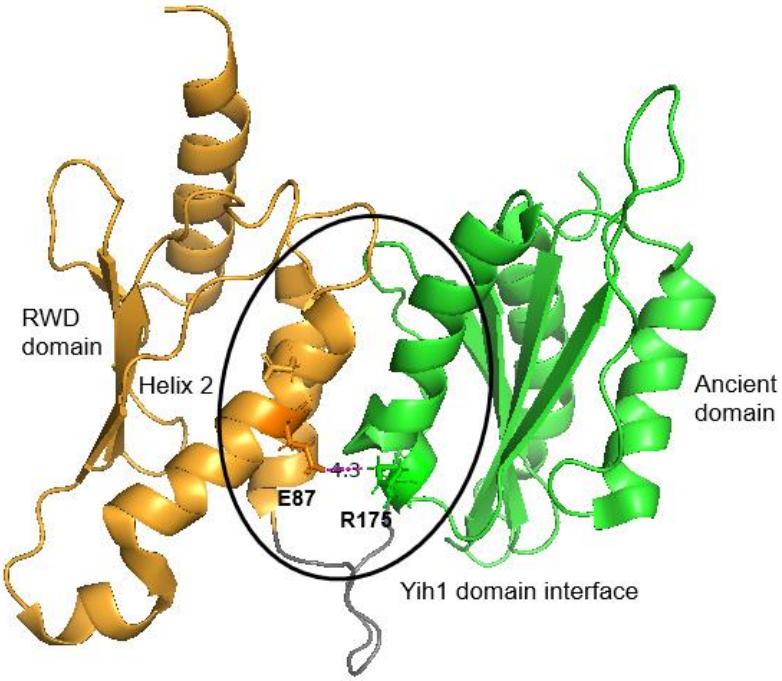


Figure 27: E87 and R175 are closely positioned at the Yih1 domain interface.

Structural modelling of Yih1 with the interface between the RWD and ancient domains circled. Shown is the approximate distance between E87 (Orange) and R175 (Green), which is 4.3 angstroms.

Our findings here points towards a suggestion that changing the charge of E87 alone to a positive charge would cause electrostatic repulsions between the two Yih1 domains, thereby 'opening up' Yih1 and exposing its Gcn1 binding determinants. Therefore, it is possible that in the RWD domain, E87 is sufficient for maintaining interdomain interactions. However, another potential interpretation is that due to the close structural proximity of E87 and D90, the electrostatic repulsions caused by changing E87 to K would disrupt any interdomain interactions involving D90, therefore opening Yih1 up.

3.8.4 Changing the charge of E87 and D90 does not completely inhibit Gcn2 activation

Our work here provides insight into the amino acids and the relevant area of the Yih1 RWD domain that are involved in interdomain interactions. However, as shown in Fig. 23, strains overexpressing Yih1 with E87 and D90 double mutations still show a small level of Gcn2 activation, as shown by a small amount of yeast cell growth on starvation media in undiluted cultures (Fig. 23A right two panels, first columns). Furthermore, the Yih1 double substitution mutants still show slight Gcn2 activation in comparison to a *gcn1Δ* strain, which depicts the phenotype of full Gcn2 inhibition (Fig 23B). Therefore, although Yih1 potency is greatly enhanced by these substitutions, the phenotype is not completely reverted. This means that Yih1 is still able to be slightly active with double E87 and D90 substitutions. This implies that

there may be remaining underlying factors or amino acid interactions that are involved in Yih1 interdomain interactions.

3.8.5 Future research

Our results suggest that ionic interactions involving E87 and D90 are important for maintaining Yih1 in a 'closed' conformation. Furthermore, this implies that the conformation of the RWD domain relative to the ancient domain determines the potency of Yih1. However, some biological questions remain, such as what may trigger such a conformational change of Yih1 under physiological conditions. Furthermore, it is still unclear whether E87 and/or D90 are functionally redundant in Yih1 interdomain interactions. As mentioned, we see a strong 3AT sensitivity phenotype with a E87K mutant, suggesting E87 alone is sufficient for Yih1 interdomain interactions. However, we cannot exclude the possibility that electrostatic repulsions from substituting E87 to K would disrupt interactions involving D90. To address this, analysing the effect of overexpressing Yih1 with a single D90A or D90K would be insightful. This would enable us to determine whether D90 alone is also sufficient for enhancing Yih1 potency.

Due to time constraints, the effect of overexpressing the Yih1 group 2 mutants on cellular eIF2 phosphorylation levels were not analysed. Further research will benefit from this experiment as it will allow more direct insight into how overexpression of these mutants affects Gcn2 activation levels in the cell.

Moreover, as the phenotype was not completely reverted by these Group 2 substitutions, it is possible that other amino acids and/or factors are involved in the interdomain interactions. Further research is necessary to uncover the precise process underlying Yih1 interdomain interactions. Through our studies, we have identified a relevant area of the RWD domain that is strongly implied to be involved in interdomain interactions. We can further apply our established yeast system to screen for other amino acids in this area that potentially mediate interdomain interactions.

Once the molecular mechanisms of Yih1 interdomain interactions are determined, the next steps will involve screening IMPACT for similar contacts. For this, the yeast system that we have established in this thesis can be utilised.

3.9 Group 3: Ancient domain amino acids involved in interdomain interactions

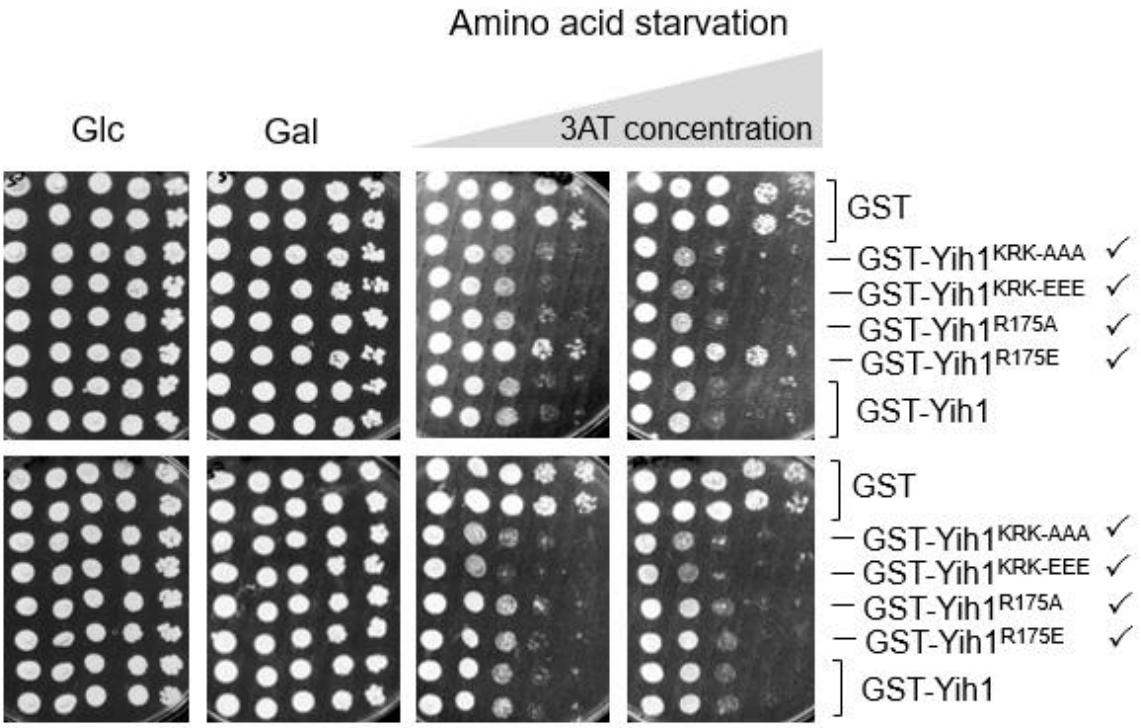
Structural determination of Yih1 through biophysical techniques has revealed that K173, R175, K176, residing in the ancient domain, are positioned close to E87 and E106 in the RWD domain

(Fig. 26) (Harjes et al., 2021). K173, R175 and K176 are positively charged (Harjes et al., 2021). It was therefore suggested that these amino acids may be involved in interdomain salt bridge formation between Yih1 domains.

These suggestions were based off computational predictions (Harjes et al., 2021). Here, we aimed to experimentally test the importance, if any, of these residues in the role of Yih1 in inhibiting Gcn2 using our yeast system. To gain a more comprehensive understanding of how these Yih1 ancient domain residues may be involved, these amino acids were mutated by Site-directed mutagenesis and transformed into a yeast plasmid prior to this study (as described in Materials and Methods, Section 2.2.2) to generate GST-Yih1 with the indicated Group 3 mutations (Fig. 28). We then utilised our yeast system to score the ability for these Yih1 mutants to inhibit Gcn2. Similar to the rationale in Section 3.7 and 3.8, we speculated that if the interaction between Yih1 domains is electrostatic, then changing these positively charged residues to the opposite charge should weaken the interaction between the two domains, exposing the Gcn1-binding determinants of Yih1 and allowing Yih1 to interact with its binding partner Gcn1.

3.9.2 Changing the charge of K173, R175 and K176 enhances Yih1 potency

A



B

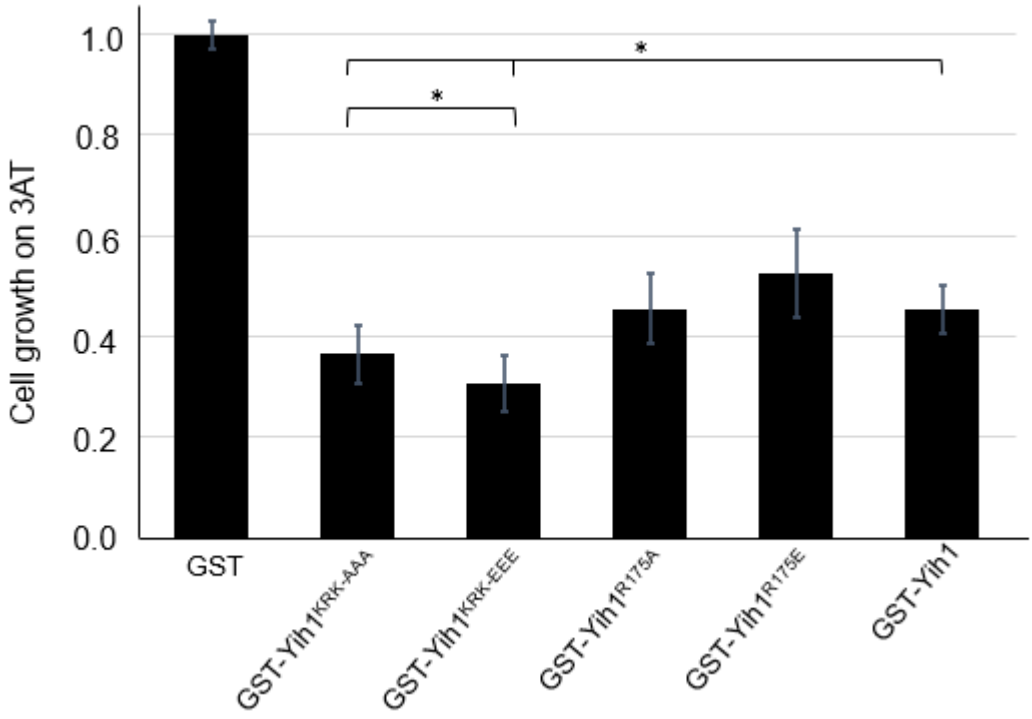


Figure 28: Testing the ability of Yih1 with Group 3 substitutions to affect cell growth in starved conditions.

A, 5uL of 10 fold serial dilutions of fully grown yeast strains (ESY11001B background) expressing either GST-tagged Yih1 with the notated Group 3 amino acid substitutions, GST alone or GST-Yih1 (wild type) were transferred onto solid SD medium with or without 3AT, and the growth of strains was observed every day for 14 days. The growth abilities of each strain expressing a Yih1 mutant were compared to that of strains overexpressing GST-alone and GST-Yih1 (wild type). Two independent transformants were analysed; 'A' strains: top row of panels, 'B' strains: bottom row of panels. Glc: 2% Glucose. Gal: 2% Gal. KRK-AAA: K173A, R175A, K176A. KRK-EEE: K173E, R175E, K176E. B, Quantitative analysis of the growth ability of the indicated strains on starvation medium. Growth scores from A and B strains were normalised and used to give an overall score. Grey bars represent Standard Error. Significant values ($P \leq 0.05$) are indicated by a '*'. ✓ = sequence confirmed. × = sequence incorrect

To assess whether overexpression of the Yih1 Group 3 substitution mutants influenced the ability for Yih1 to impair Gcn2 activation *in vivo*, a SQGA was carried out. For this, previously transformed yeast colonies (A colonies: top row of panels, B colonies: bottom row of panels) of ESY11001B strains expressing either GST-tagged Yih1 with the notated Group 3 amino acid substitutions, GST alone or GST-Yih1 (wild type) were grown to saturation as described (Fig. 28) (Materials and Methods, Section 2.1.6). For each strain, two different transformants (A and B) were analysed in order to determine the reproducibility of results. These yeast strains were then subject to a SQGA as described (Materials and Methods, Section 2.2.4).

As seen in Fig. 28, all strains demonstrated full growth on control plates containing glucose or galactose without 3AT treatment. These observations ensured that overexpression of these proteins did not affect an essential cellular function, which would have manifested itself by impaired growth under non-starved conditions. Furthermore, yeast strains overexpressing GST alone was able to grow on 3AT medium (3AT resistance) (Fig. 28). This 3AT resistances phenotype indicated that Gcn2 could be activated, allowing cells to grow in starved conditions. Furthermore, strains overexpressing GST-Yih1 (wild type) demonstrated reduced yeast cell growth on starvation medium, implying that Gcn2 function was impaired which resulted in impaired cell growth (Fig. 28).

Yeast strains overexpressing GST-Yih1$^{KRK\text{-}AAA}$ showed a stronger 3AT sensitivity phenotype than the strain overexpressing wild type GST-Yih1 (Fig. 28) ($P = \leq 0.05$). This implies that changing the charge of these amino acids to a neutral charge slightly enhances Yih1 potency when compared to wild type Yih1. We speculated that if these residues are involved in

electrostatic interactions, then changing them to a negative charge may enhance Yih1 potency due to electrostatic repulsions. A and B strains of GST-Yih1$^{KRK\text{-}EEE}$ both show stronger 3AT sensitivity in comparison to GST-Yih1 (Fig. 28) (P= ≤ 0.05). In comparison to GST-Yih1$^{KRK\text{-}AAA}$, only GST-Yih1$^{KRK\text{-}EEE}$ B strain, but not the 'A' strain showed a stronger 3AT sensitivity (Fig. 28). The inconsistences between the A and B strains GST-Yih1$^{KRK\text{-}EEE}$ need to be revisited following this study.

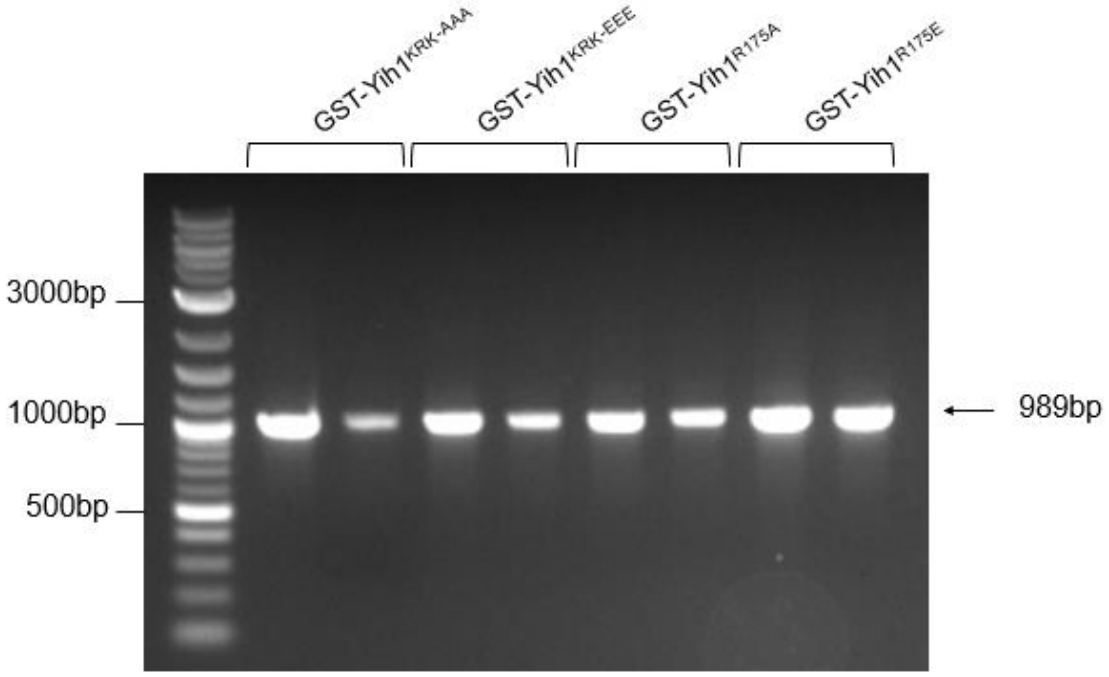
3.9.3 A R175 substitution to Alanine or Glutamic Acid does not influence Yih1 potency

The next step was to investigate whether one or all of these amino acids were responsible for the enhanced 3AT sensitivity phenotype. The growth ability of yeast strains overexpressing GST-Yih1 with a R175A or R175E substitution (GST-Yih1^{R175A}, GST-Yih1^{R175E}) was scored. As shown in Fig. 28, strains overexpressing GST-Yih1^{R175A} consistently showed a growth phenotype like that of wild type Yih1. Furthermore, a T test confirmed that there were no significant differences between GST-Yih1^{R175A} and GST-Yih1 (P value= 0.169). This implies that changing this residue to a neutral charge does not impact the potency of Yih1.

We then wanted to examine whether swapping the charge of R175 to a negative charge impacts the ability of yeast cells to grow in starved conditions. The 'A' strain of GST-Yih1^{R175E} showed slightly enhanced 3AT resistance, similar to GST alone, in comparison to Yih1 (Fig. 28). On the other hand, strain 'B' of GST-Yih1^{R175E} demonstrated a 3AT sensitivity phenotype like that of wild type Yih1 (Fig. 28). Despite these differences between strains, there was no significant differences in potency of GST-Yih1^{R175E} and GST-Yih1 wild type (P value = 0.506). However, the inconsistencies in these results need to be further addressed experimentally.

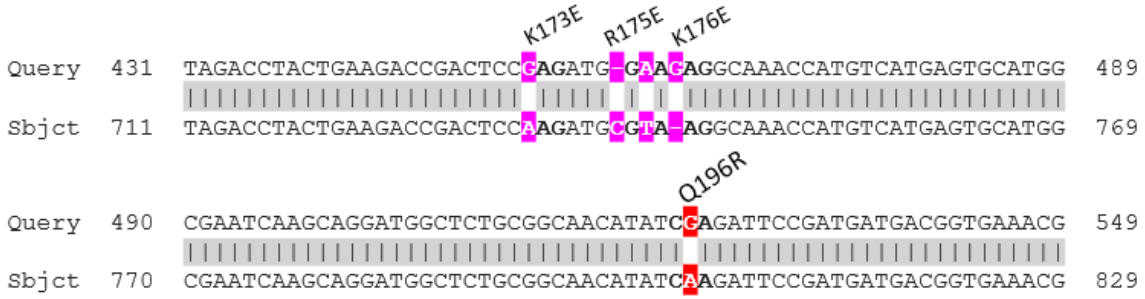
3.9.4 Verifying the correct substitution of Yih1 Group 3 mutations

A

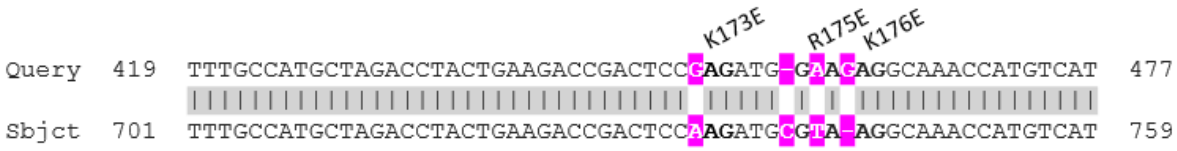


B GST-Yih1$^{KRK-AAA}$ (A and B are the same)

C GST-Yih1$^{KRK-EEE}$ (A)

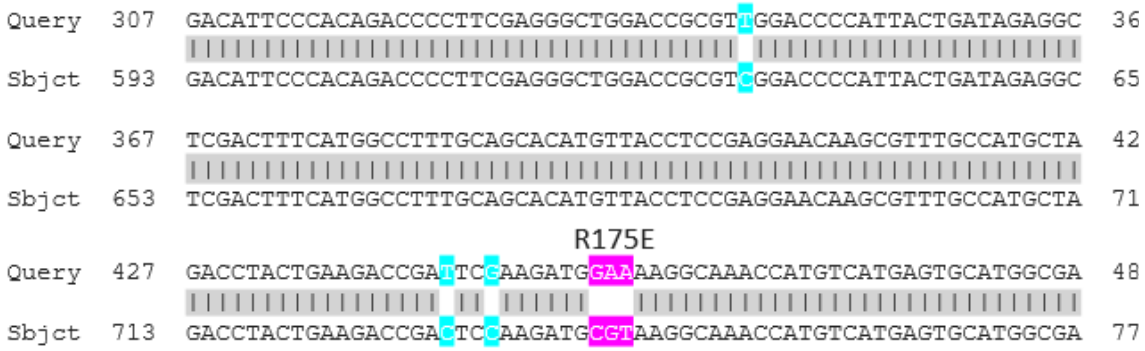


D GST-Yih1$^{KRK-EEE}$ (B)



E GST-Yih1^{R175A} (B)

F GST-Yih1^{R175E} (A)



G GST-Yih1^{R175E} (B)

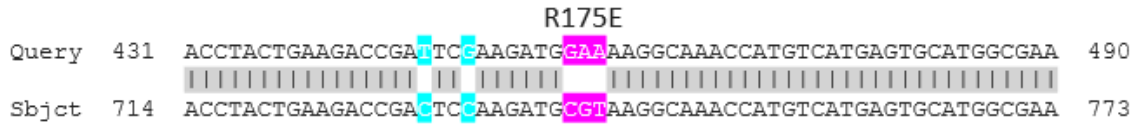


Figure 29: Sequence verification of Yih1 Group 3 mutations

A: Plasmid DNA of the indicated GST-Yih1 mutant proteins were amplified by colony PCR using Taq polymerase and using area primers ES3227 and ES3228 and then prepared for sequencing (Materials and Methods). Two colonies for each strain were analysed. B: Sequencing alignment area of interest for GST-Yih1$^{KRK-AAA}$. Alignment is representative of colonies A and B. C: Sequencing alignment area of interest for GST-Yih1$^{KRK-EEE}$ colony A. D: Sequence alignment area of interest for GST-Yih1$^{KRK-EEE}$ colony B. E: Sequencing alignment area of interest for GST-Yih1^{R175A}. Alignment is representative of colonies A and B. F: Sequencing alignment area of interest for GST-Yih1^{R175E} colony A. G: Sequencing alignment area of interest for GST-Yih1^{R175E} colony B. Magenta highlighted nucleotides are expected mutations. Blue highlighted nucleotides are unexpected silent mutations (do not

change the protein sequence). Red highlighted nucleotides are unexpected mutations which change the protein sequence. Query: Yih1 protein sequence (UniProtKB - P25637). Subject: Yih1 mutant protein sequence. Full BLAST sequencing alignments are attached in Supplementary Material (Section S4, Fig. S4.2). Sequencing order HC00073122.

It is important to sequence verify the Yih1 mutant proteins to confirm that the phenotype produced by these strains are representative of the desired Yih1 mutation. It is also important to analyse whether any undesired mutations were mistakenly inserted during construction of the mutant protein. Yeast strains (ESY11001B) expressing either GST-Yih1$^{KRK\text{-}AAA}$, GST-Yih1$^{KRK\text{-}EEE}$, GST-Yih1^{R175A} or GST-Yih1^{R175E} were grown to saturation at 30°C on solid medium (Materials and Methods, Section 2.1.6). Colony PCR was then carried out using the area primers ES3227 and ES3228, to amplify the corresponding Yih1 region for sequencing. Following this, samples were resolved on a 1% agarose gel to confirm that the expected Yih1 fragment size was amplified by colony PCR (Fig. 29A). These plasmid DNA samples were then sent in for sequencing in compliance with the regulations by Macrogen (Materials and Methods, Section 2.11). Nucleotide BLAST sequencing alignment of the Yih1 mutant protein sequence with wild type yeast Yih1 (UniProtKB—P25637) was carried out to determine whether the desired mutations were correctly inserted.

As shown in Fig. 29B, GST-Yih1$^{KRK\text{-}AAA}$ (both transformants ‘A’ and ‘B’) was sequence confirmed, with no unexpected mutations being present. Sequencing of Yih1$^{KRK\text{-}EEE}$ also confirmed that the expected mutations (K173E, R175E, K176E) were present in both A and B transformants (Fig. 29C and Fig. 29D). However, an unexpected non-target Q196R mutation was also introduced into the sequence of transformant ‘A’ only (Fig. 29C). The GST-Yih1^{R175A} sequence was confirmed with the expected R175A mutation there, and no unexpected mutations (for both A and B strains) (Fig. 29E). Furthermore, GST-Yih1^{R175E} (A and B strains) were both sequence confirmed, with unexpected silent mutations being present (Fig. 29F and Fig. 29G).

4.0 Discussion

4.1.1 Ancient domain residues create a positively charged interface for the RWD domain

D102, the amino acid that is important for the Yih1-Gcn1 interaction, is buried in the interface between the two Yih1 domains (Harjes et al., 2021). Furthermore, Yih1 only impairs Gcn2 activation under certain unknown circumstances. This suggests that Yih1 may undergo a conformational change to become primed to interact with Gcn1. Yih1 ancient domain residues K173, R175 and K176 are positioned close to E87 and D90 in the RWD domain (Fig. 26) (Harjes et al., 2021). These residues were suggested to form the interface between the two Yih1 domains (Fig. 26) (Harjes et al., 2021). Therefore, we targeted the ancient domain residues to determine their importance, if any, in Yih1-mediated Gcn2 inhibition. We found that swapping the charge of K173, R175 and K176 enhances Yih1 potency (Section 3.7.1). Therefore, this suggests that changing the charge of these amino acids influences the ability of Yih1 to inhibit Gcn2.

When comparing between changing the charge to neutral or negative, the 'B' strain of GST-Yih1$^{KRK\text{-}EEE}$ (Fig. 28 bottom panels), but not the A strain (Fig. 28 top panels), showed slightly enhanced Yih1 potency compared to GST-Yih1$^{KRK\text{-}AAA}$. Furthermore, there was a significant difference in Yih1 potency between GST-Yih1$^{KRK\text{-}AAA}$ and GST-Yih1$^{KRK\text{-}EEE}$ ($P = \leq 0.05$). One explanation for this result is the swapping the charge of these positive residues forces Yih1 into a more open conformation due to electrostatic repulsions. This would enable Yih1 to bind more strongly to Gcn1, and thereby inhibit Gcn2 stronger than wild type Yih1.

Sequencing analysis of the triple mutation strains showed that GST-Yih1$^{KRK\text{-}AAA}$ was verified with no mistakes in the sequence. Furthermore, while the 'B' strain of GST-Yih1$^{KRK\text{-}EEE}$ was verified, with no undesired mutations, the A strain had an undesired mutation (Q196R). This may account for the discrepancies in the results between GST-Yih1$^{KRK\text{-}EEE}$ A and B strains. Due to time restraints, the expression levels of these mutant proteins were not tested. Therefore, we cannot exclude the possibility that these proteins are expressed at different levels. The analysis of expression levels needs to be revisited in future research.

4.1.1 R175 alone is not important for interdomain interactions

Understanding whether one or multiple amino acids are responsible for the enhanced 3AT sensitivity of cells overexpressing Yih1$^{KRK\text{-}AAA}$ and Yih1$^{KRK\text{-}EEE}$ mutants is important for a detailed understanding of the molecular mechanisms behind Yih1 function. In our study, R175 was targeted as Yih1 modelling suggested that it is close to E87, with the potential for an electrostatic interaction (Fig. 26) (Harjes et al., 2021). We did not see a significant phenotype

when substituting R175 to Alanine, as cells grew the same as cells overexpressing wild type Yih1 (Fig. 28) (P value= 0.169). This implies that changing the charge of R175 to neutral does not affect the potency of Yih1 in Gcn2 inhibition.

When determining the effect of swapping the charge of R175 to a negative charge (R175E substitution), we saw a slight 3AT resistance phenotype in one yeast strain ('A' strain, Fig. 28). However, the 'B' strain of a R175E mutant showed a similar yeast cell growth rate of that by GST-Yih1 overexpression (Fig. 28). Despite this, there was no significant difference between the growth of GST-Yih1^{R175E} and wild type GST-Yih1 (P value= 0.506). This suggests that swapping the charge of R175 does not affect the potency of Yih1. One interpretation of these results is that R175 on its own is not sufficient for maintaining interdomain interactions. Accordingly, this would mean that mutating R175 only would not disrupt the protein structure and accordingly, Yih1 would stay in the same conformation as wild type GST-Yih1. It is important to consider that the differences in phenotypes between GST-Yih1^{R175E} A and B strains may have been caused by differences in expression levels between the two strains. Furthermore, it is possible that GST-Yih1^{R175A} and GST-Yih1^{R175E} were not expressed as highly as the triple mutation proteins. As mentioned, to exclude this possibility, this urgently needs to be revisited by checking the expression levels of these proteins in the future to address these inconsistencies.

4.1.2 Additional residues may be involved in contacting the RWD domain

The 3AT sensitivity phenotype of the GST-Yih1$^{KRK-AAA}$ and GST-Yih1$^{KRK-EEE}$ were weak, as Yih1 potency was only enhanced by approximately 1/3 (one third) in comparison to wild type Yih1 (Fig. 28A and Fig. 28B). Furthermore, the potency of Yih1 was not as strong as it was when RWD domain amino acids E87 and D90 were substituted (Fig. 23). Moreover, it is also possible to speculate that the interdomain interactions may not be solely due to one exclusive amino acid, but rather the charge that it possesses. As the phenotype of cells overexpressing GST-Yih1 harbouring the Group 3 mutations were not completely reverted, it is possible that other ancient domain amino acids are maintaining the interdomain interactions in Yih1.

Another possible explanation would be that all of the ancient domain residues (K173, R175 and K176) are required for maintaining interdomain interactions between the Yih1 domains. If this is true, we would expect that the other interdomain interactions, mediated by K173 and K176, would remain intact, or compensate for the disruption, when substituting R175 alone.

This would mean that the Yih1 would stay in its “normal” conformation and therefore no enhancement in Yih1 potency would occur in comparison to wild type Yih1.

4.1.3 Future research

Future research is required to completely decipher the amino acids that are involved in maintaining interdomain interactions between the Yih1 domains. Another residue that may give insight into the interdomain interactions is K173 (Harjes et al., 2021). In our alignment, K173 is fully conserved among all species (Fig. 22). This suggests that it has an important role in Yih1/IMPACT function. Furthermore, structural modelling has predicted K173 to interact with E106 of the RWD domain, implying K173 to be involved in interdomain interactions (Harjes et al., 2021). Supporting this, E106 and K173 are spatially arranged in close proximity, being approximately 2.4 angstroms apart (Fig. 30). Further mutagenesis studies, utilising our established yeast system, on K173 alone would be insightful for experimentally determining whether this residue alone is important for in interdomain interactions. As mentioned prior, expression levels of the Yih1 mutants in this section needs to be revisited, as a priority, to account for any phenotypic variance. Furthermore, eIF2 phosphorylation assays of cells overexpressing these Yih1 mutants would be beneficial to determine the effect of substituting these residues on the ability of Yih1 to inhibit Gcn2.

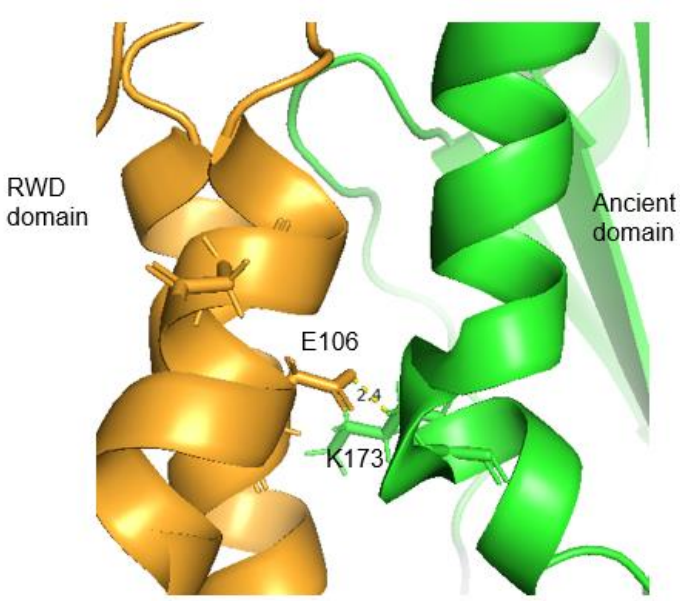


Figure 30: K173 and E106 are positioned close to each other in Yih1.

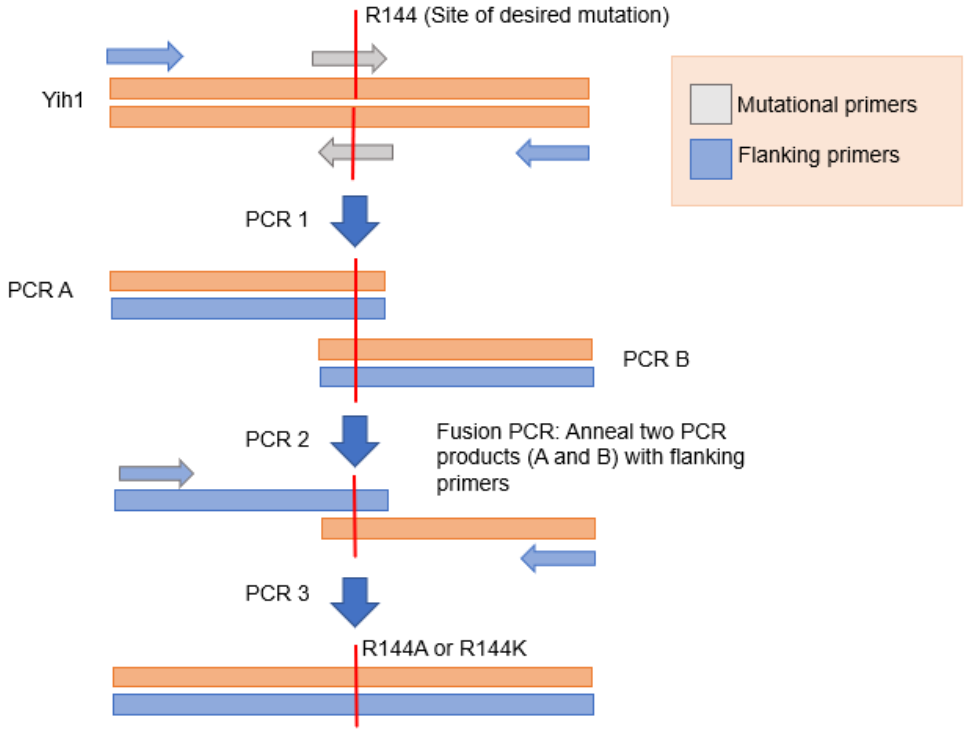
Shown above is a close-up view of the Yih1 domain interface, with the distance between E106 and K173 indicated (approximately 2.4 angstroms distance separation). The RWD domain is coloured in orange, and the Ancient domain is coloured in green.

4.2 Group 4: Potential role of Post-Translational Modification in Yih1 function

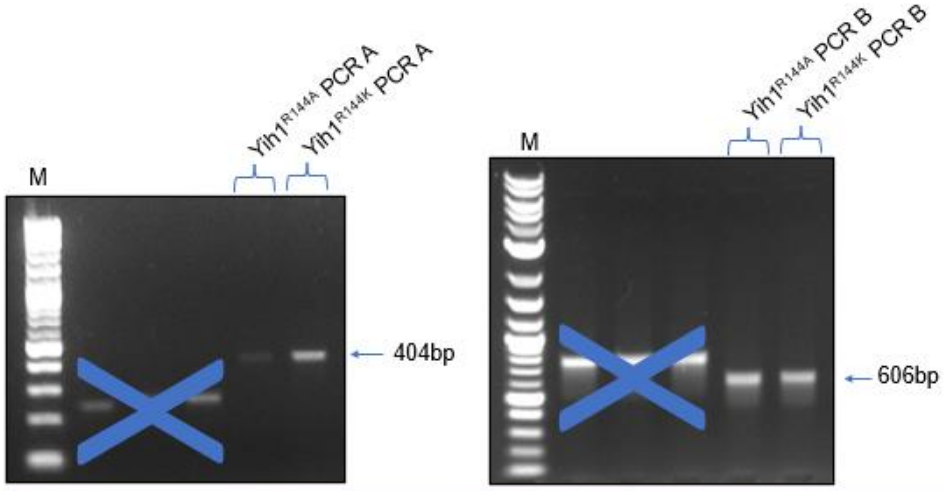
When expressed in *E. coli*, the Yih1 amino acid Arg144 (R144) is dimethylated (Harjes et al., 2021). We wanted to address the involvement of R144 methylation in Yih1-mediated Gcn2 inhibition in its native environment (yeast). Assuming that this methylation also exists in yeast, we carried out site-directed mutagenesis of R144 to evaluate the effect substituting this residue on the function of Yih1 in Gcn2 inhibition using our yeast system. R144 was substituted to Alanine, which would prevent the methylation at the position 144 (Weber et al., 2009). R144 was also substituted to Lysine (R144K), which cannot be methylated, therefore mimicking non-methylatable Yih1 at position 144 (Weber et al., 2009). The Mutagenesis procedure is outlined below.

4.2.1 Introduction of R144 substitutions by Site directed mutagenesis

A

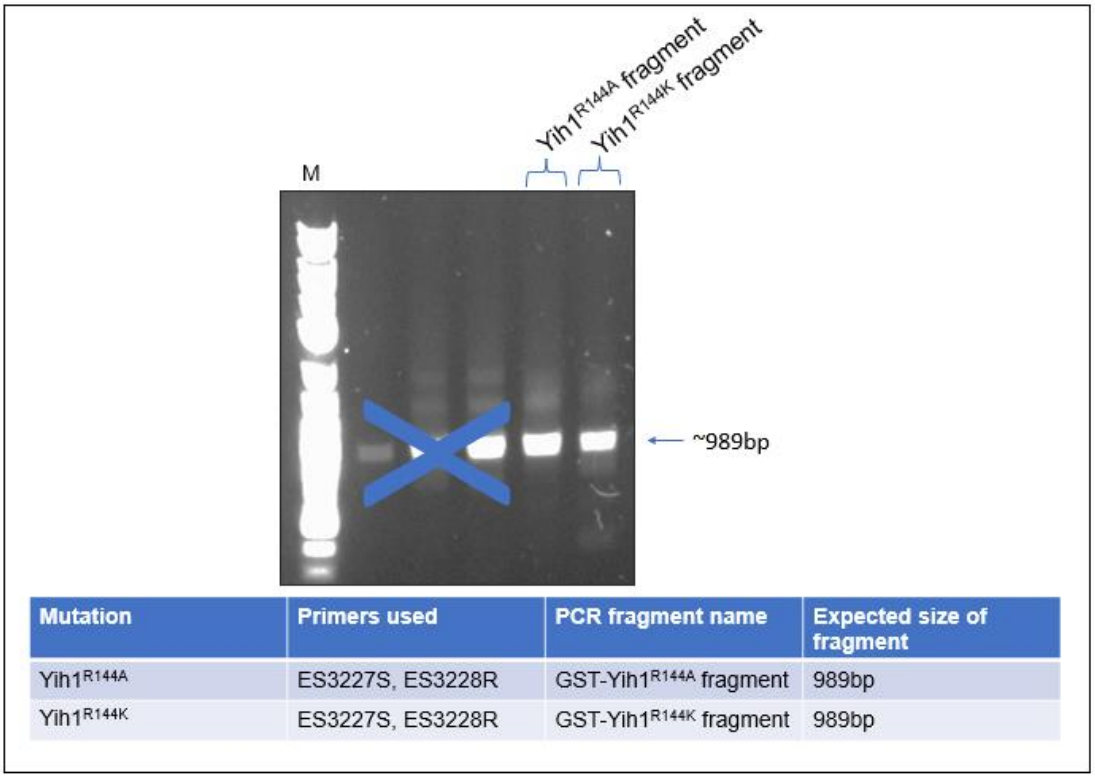


B



Mutation	Primers used	PCR fragment name	Expected size of fragment
Yih1^{R144A}	ES3227S, ES4142R	PCR A	404bp
	ES3228R, ES4141S	PCR B	606bp
Yih1^{R144K}	ES3227S, ES4144R	PCR A	404bp
	ES3228R, ES4143S	PCR B	606bp

C



Mutation	Primers used	PCR fragment name	Expected size of fragment
Yih1^{R144A}	ES3227S, ES3228R	GST-Yih1^{R144A} fragment	989bp
Yih1^{R144K}	ES3227S, ES3228R	GST-Yih1^{R144K} fragment	989bp

Figure 31: Introduction of R144 substitutions into Yih1 by Site-directed mutagenesis.

A: Experimental procedure for Site-directed mutagenesis to introduce a desired mutation into Yih1. See text for details. B: A Yih1 plasmid (pES187-b1) was used as a template for the initial PCR reactions to generate two halves of the same Yih1 fragment with overlapping regions that contain the desired mutation. The PCR fragments (A and B) were resolved by 2% (PCR B)/1% (PCR A) agarose gel electrophoresis along with a 2-log DNA ladder marker of known sizes to determine the fragment size. C: Fusion PCR was carried out using PCR A and PCR B (Fig. 31B) as the template, and ES3227S and ES3228R as the area primers, to generate a complete Yih1 fragment with the desired mutations. The PCR fragment was resolved by 1% agarose gel electrophoresis along with a 2-log DNA ladder marker of known sizes to determine the fragment size. Blue tables indicate the desired Yih1 mutation, primers used for Site-directed mutagenesis, PCR names and the expected fragment size.

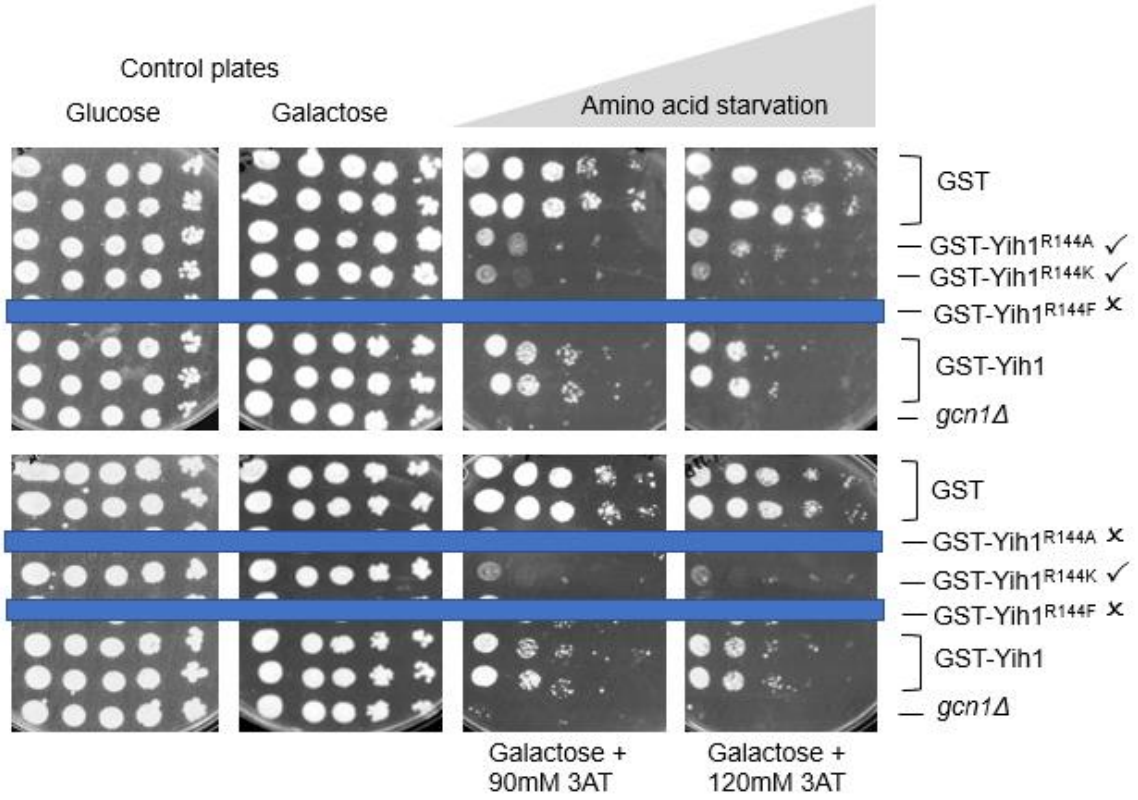
To test whether R144 is involved in Yih1-mediated Gcn2 inhibition, we first substituted R144 to either Alanine (A) or Lysine (K) using Site-directed mutagenesis. For the first PCR reaction, a plasmid containing the Yih1 ORF (pES187-b1) was used as the template. For each substitution mutation to be inserted, in the initial round of PCR, two separate reactions were carried out using the indicated mutational and flanking primer pairs, indicated in Fig. 31. This resulted in amplification of two halves of the same Yih1 fragment with overlapping regions that contain the desired mutation, designated PCR A and PCR B. The ‘A’ PCR fragments were resolved by 1% agarose gel electrophoresis and the fragment size was determined by

comparison with the 2-long DNA ladder (M) (Fig. 31B). The ‘B’ PCR fragments were resolved by 2% agarose gel electrophoresis and the fragment size was determined by comparison with the 2-log DNA ladder (M) (Fig. 31B). In the next PCR reaction, the mutated Yih1 fragment halves (PCR A and PCR B) were used as the template DNA along with the flanking primers ES3227S and ES3228R, which annealed upstream and downstream of the desired Yih1 mutation site (Fig. 31A). The resulting amplicon was a Yih1 fragment containing the desired mutation (R144A or R144K). These fragments were resolved by agarose gel electrophoresis and the fragment size was determined by comparison with the 2-log DNA ladder (M) (Fig. 31C).

The Yih1^{R144A} or Yih1^{R144K} fragments were then subject to yeast transformation using a Yih1 deletion strain (ES11001b) containing the pES187-b1 plasmid (1:40 dilution) that was previously cut open using BglII and SpeI restriction sites prior to this study. In pES187-b1, the Yih1 ORF is fused to GST tag at the C terminus. Furthermore, in this plasmid, the expression of the Yih1 ORF is under the control of a Galactose Inducible Promotor (GAL1). This transformation resulted in the generation of plasmid born Yih1 variants (GST-Yih1^{R144A} and GST-Yih1^{R144K}) expressed from the Galactose-inducible promotor. Samples were plated on solid SDWLIV + glucose using glass beads and incubated overnight at 30°C. With these transformants, a Semi Quantitative Growth Assay was carried out, as described in Materials and Methods.

4.2.2 Substituting Arg144 to Ala or Lys impairs cell growth

A



B

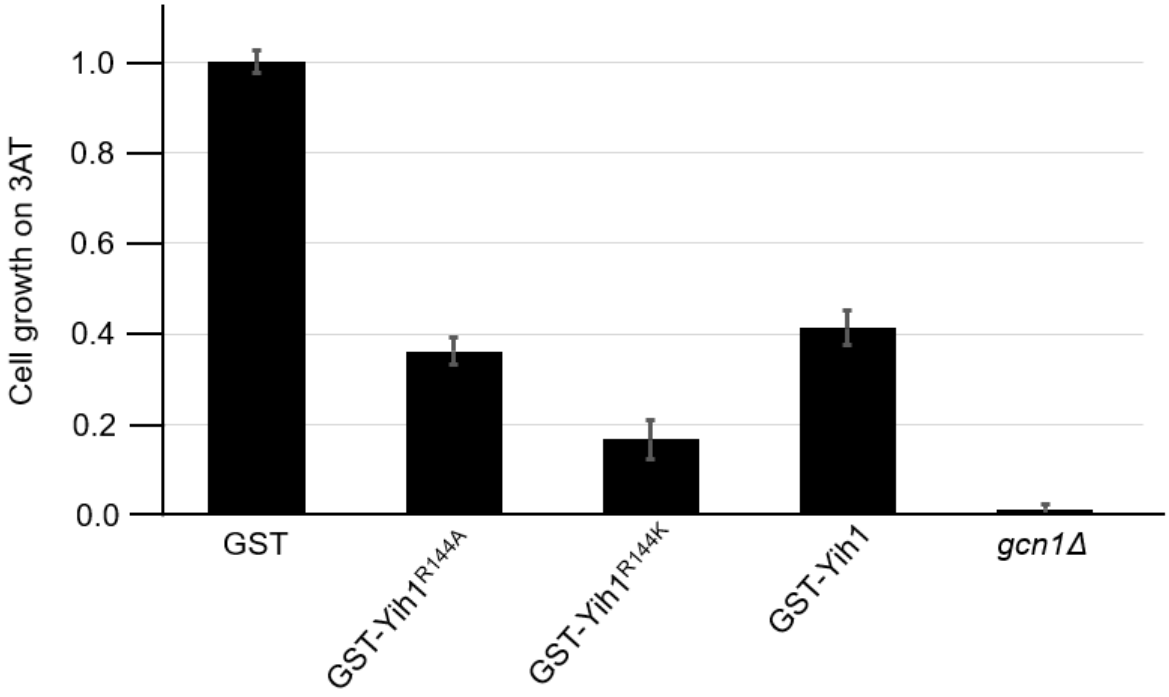


Figure 32: R144 substitution to Alanine or Lysine impairs cell growth in starved conditions.

A: 5uL of 10-fold serially diluted overnight yeast cultures, expressing either GST alone, GST-Yih1, a *gcn1Δ* strain or the indicated Yih1 R144 mutations were transferred into solid SD medium supplemented with either 2% glucose or galactose, and then galactose with increasing 3AT concentrations to induce amino acid starvation. Yeast cell growth of strains overexpressing the indicated proteins was analysed by a SQGA. The top panels are A colonies, and bottom panels are B colonies. 3AT: 3-Amino-1,2,4-triazole. ✓= sequence confirmed. × = sequence incorrect. B: Yeast cell growth from the SQGA (A) was quantified, as described in Section 3.1.4. Complete calculations and data are available in Supplementary Material, Fig. S5.1).

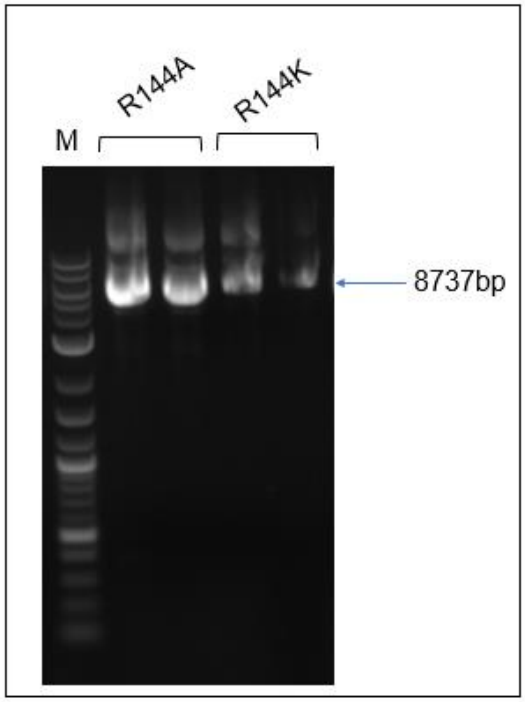
Two colonies from the yeast transformation, designated 'A' and 'B' were grown to saturation at 30°C in liquid medium (Materials and Methods, Section 2.1.6). A SQGA was then carried out as described in Materials and Methods, Section 2.2.4. All yeast strains successfully grew on medium containing either Glucose or Galactose without 3AT, indicating that no growth defects were present (Fig. 32A). Cells overexpressing GST alone demonstrated an enhanced 3AT resistance in comparison to wild type Yih1 (Fig. 32). Furthermore, cells overexpressing GST-Yih1 demonstrated a stronger 3AT sensitivity phenotype when compared to GST alone (Fig. 32). Furthermore, yeast cells expressing a *gcn1*Δ strain demonstrated strong 3AT sensitivity, consistent with the inhibition of Gcn2 as Gcn2 cannot be activated without its binding partner Gcn1 (Fig. 32).

As seen in Fig. 32, strains overexpressing GST-Yih1^{R144A} showed an enhanced 3AT sensitivity phenotype in comparison to wild type Yih1. Moreover, strains overexpressing GST-Yih1^{R144K}, therefore mimicking methylation, showed a much stronger sensitivity to 3AT when compared to that wild type Yih1 (Fig. 32). Furthermore, the 3AT sensitivity phenotype of GST-Yih1^{R144K} was noticeably stronger than that of GST-Yih1^{R144A} (Fig. 32A and Fig. 32B). Therefore, results from this SQGA imply that substituting R144 to Alanine or Lysine influences the ability of Yih1 to inhibit Gcn2.

After determining these different phenotypes produced by Yih1 R144 mutants, and to further address our question of the potential role of R144 methylation in yeast Yih1 function, we next sought clarification that the desired mutations had been correctly inserted into Yih1.

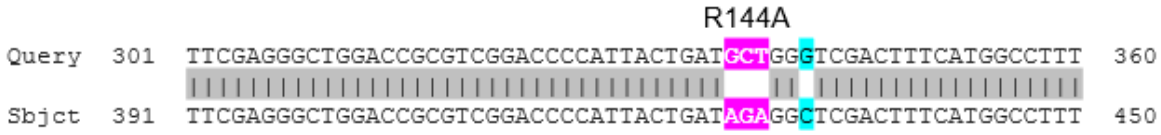
4.2.3 Verifying the correct R144 substitutions in Yih1 variants

A



B

GST-Yih1^{R144A} (A)



C

GST-Yih1^{R144K} (A and B)

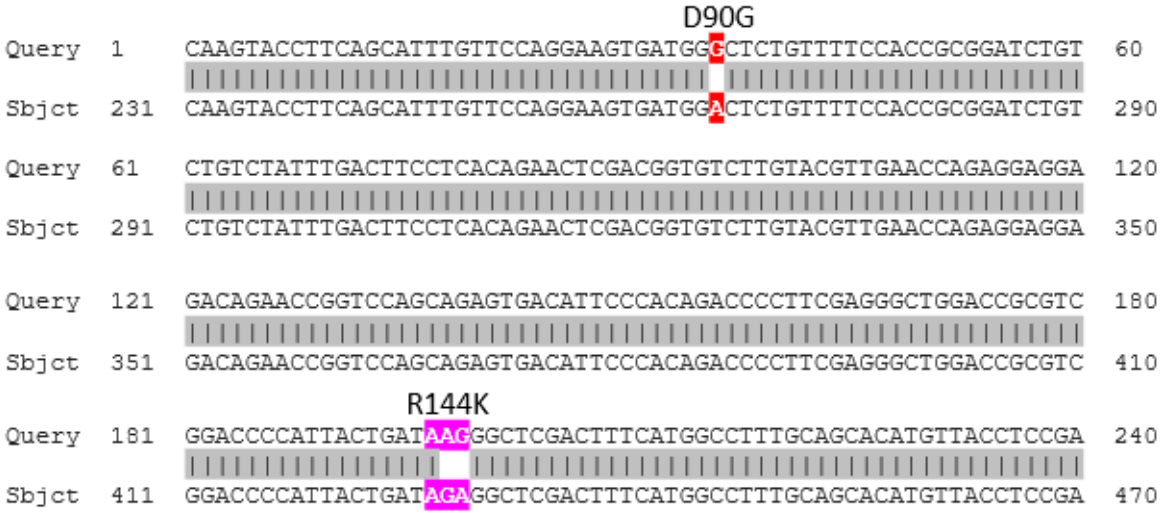


Figure 33: Sequence verification of GST-Yih1 Arg144 mutants.

A: GST-Yih1 containing the indicated R144 substitutions were grown to saturation in SDWILV + glucose and the plasmid DNA was purified as described in Section 4.2.3 text. Pure DNA samples were resolved by 1% agarose gel electrophoresis and then sample DNA concentration and purity of samples were measured using a Nanodrop. DNA samples were then sent in to Macrogen for commercial sequencing. B: Sequencing analysis of GST-Yih1^{R144A} (A). C: Sequencing analysis of GST-Yih1^{R144K} (A and B). Query sequence: Plasmid DNA fragments isolated in (A) containing either GST-Yih1^{R144A} or GST-Yih1^{R144K}. Sbjct (Subject) sequence: Wild type Yih1 sequence used for comparison/alignment (UniProtKB—Yih1 P25637). The area of interest is shown here, with full alignments shown in the Supplementary Material (Fig. S4.2). Magenta colour indicates mutations that are expected. Turquoise colour indicates unexpected silent mutations. Red colour indicates unexpected mutations.

We checked the sequence of our Yih1 R144 mutants in order to verify the presence of the desired mutations. For this, yeast strains expressing GST-Yih1^{R144A} and GST-Yih1^{R144K} were grown to saturation at 30°C, and then the plasmids were isolated from yeast and transformed into *E. coli*. After this, plasmids were further isolated from *E. coli* (Materials and Methods, Section 2.1.11). These samples of plasmid DNA were resolved by 1% agarose gel electrophoresis, along with a 2-log DNA ladder of known sizes, to determine the size of the DNA fragment. Samples were then sent to Macrogen for commercial sequencing. The samples were then subject to Nucleotide BLAST sequence alignment with the wild type Yih1 protein sequence (UniProt - P25637). The presence of the R144A substitution was verified in the A strain, but not the B strain (Fig. 33B). Furthermore, the desired R144K mutation was corrected inserted into the Yih1 protein sequence of GST-Yih1^{R144K} (Fig. 33) However, analysis of the BLAST alignment indicated that an undesired D90G substitution was introduced upstream of R144. This means that we cannot exclude that the 3AT sensitivity phenotype of cells overexpressing GST-Yih1^{R144K} was influenced by the presence of the D90G mutation.

4.2.4 Purification of Yeast GST-Yih1 by Size Exclusion Chromatography

In parallel to studying the effect of a R144 amino acid substitution, we aimed to investigate whether Yih1 is also dimethylated (or methylated) at position 144 in yeast. For this, GST-Yih1 was overexpressed in *Saccharomyces cerevisiae* and subsequently harvested and mechanically lysed in preparation for Size Exclusion Chromatography.

Gel filtration chromatography was firstly carried out to purify GST-Yih1. As the Yih1 protein construct used in this study has a GST (Glutathione-S-Transferase) tag, the purification technique used here exploited the affinity that GST has for glutathione, its substrate molecule.

A gel filtration column was prepared by loading Glutathione Agarose resin (Protino® Glutathione Agarose 4B) onto the column and then washing using SEC buffer, as described in Materials and Methods. To immobilize GST-Yih1, yeast protein lysate, consisting of overexpressed GST-Yih1, was loaded directly onto the column. When the protein lysate is loaded directly onto the column, the GST would bind to its substrate molecule glutathione and therefore immobilise GST-Yih1 in the affinity resin gel matrix. The flowthrough, containing non-specific proteins, was collected in a falcon tube, and stored on ice. Subsequently, the column was washed a further 3 times with SEC buffer (as described in Materials and Methods, Section 2.6). Glutathione reduced (GSH) was then prepared and loaded directly onto the column to elute GST-Yih1, the target protein, and flow through was collected. This process was repeated two times subsequently using 1mL of Glutathione reduced to elute any remaining GST-Yih1. The sample, containing GST-Yih1, was then concentrated using a Centricon concentrator. proteins absorbance (280nm) of the flow through was measure on a Nanodrop™, yielding a concentration of 9.41mg/mL. The yields and protein concentrations are presented in Table 21: Purification of GST-Yih1 from *S. cerevisiae*).

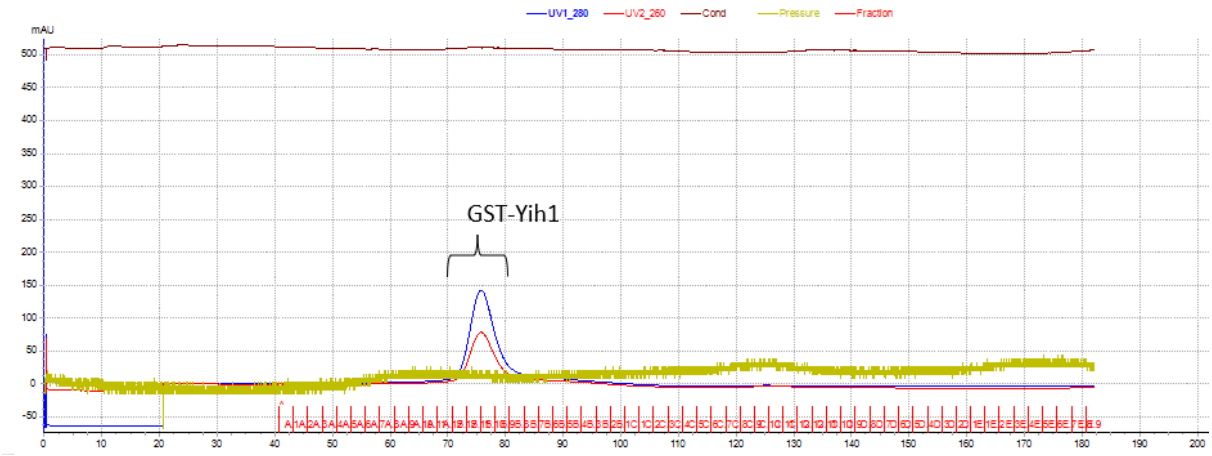
Table 21: Purification of GST-Yih1 from *S. cerevisiae*

Purification Method	**Volume (mL)**	**Protein concentration (mg/mL)**
Yeast Whole Cell Lysate (Flowthrough) from SEC	~20	44.87
Size Exclusion Chromatography (GST-Yih1 eluted with GSH)	0.15	9.41
Fast Protein Liquid Chromatography (GST-Yih1 after FPLC)	10.1	8.06

The eluted GST-Yih1 from Gel filtration chromatography was then subject to Fast Protein Liquid Chromatography (FPLC) for further purification. FPLC was carried out by Dr. Stefan Harjes using a 75kDa FPLC size exclusion column on a ÄKTA Explorer System at 4°C. As shown in Fig. 34, GST-Yih1 elutes a single peak at approximately 75 minutes, which was consistent for a protein the size of GST-Yih1 (55kDa). This peak corresponded to Fractions B10 and B11 (Fig. 34B). Fractions B10 and B11 were collected and concentrated (protein concentration 8.09mg/mL), and then subject to SDS page gel electrophoresis. The gel was stained with Colloidal Coomassie (as described in Materials and Methods, Section 2.8), which

resolved a strong single band of approximately 64kDa (Fig. 35). Additionally, samples from the protein purifications (Table 21) were also subject to western blot analysis and the membrane was probed for GST using anti-GST antibodies, also resulting in a band of approximately 64kDa. (Fig. 35B). Interestingly, the 64kDa molecular weight is 10kDa larger than the expected molecular weight of GST-Yih1 (55kDa). The difference in molecular weight may be attributable to the presence of Posttranslational Modifications in GST-Yih1. This reinforced in requirement for Mass Spectrometry in order to characterise any Posttranslational Modifications present in Yih1.

A



B

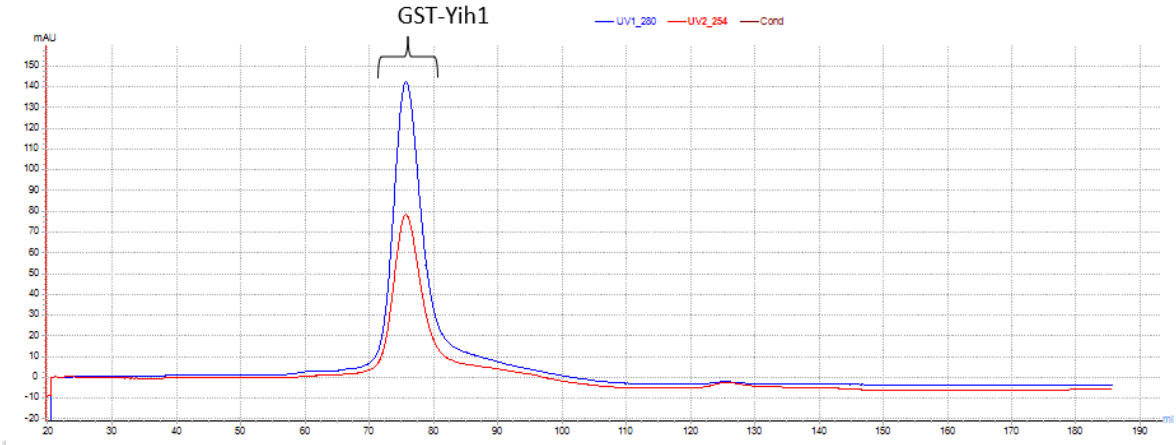


Figure 34: Chromatogram of Size Exclusion Chromatography of GST-Yih1.

Blue line: $Abs_{280nm.}$ Red line: Abs_{250nm}. Fast Protein Liquid Chromatography was performed using a 75kDa size exclusion column on an ÄKTA Explorer System at 4°C.

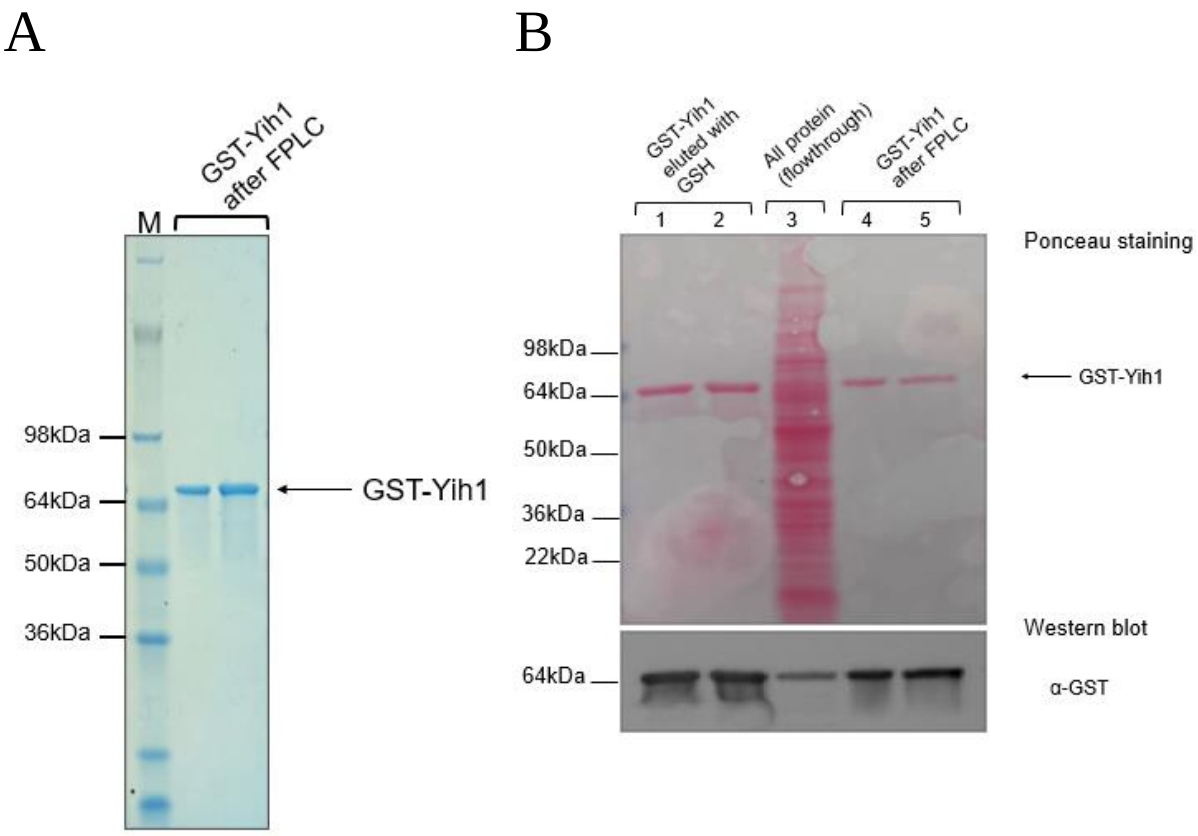


Figure 35: SDS-page analysis of pure GST-Yih1 fractions from Size Exclusion Chromatography

A: SDS-PAGE gel electrophoresis (4-20% acrylamide gel) of the collected fractions B10 and B11 from FPLC. Protein concentration was 8.08mg/mL, and the sample was diluted by 30x (145uL Protein Loading Buffer and 5uL sample): 2uL was loaded in first lane, 4uL was loaded in second lane. The gel was stained with Colloidal Coomassie stain (as described in Materials and Methods, Section 2.8). These bands were excised for In Gel Chymotrypsin Digest in preparation for Mass Spectrometry. M= Marker. B: Western blot analysis of purified GST-Yih1 after Size Exclusion Chromatography (sample concentrations are listed in Table 21). The membrane was stained with Ponceau stain and then probed with a GST antibody.

4.2.5 Arg144 is not dimethylated in Yeast, its native environment

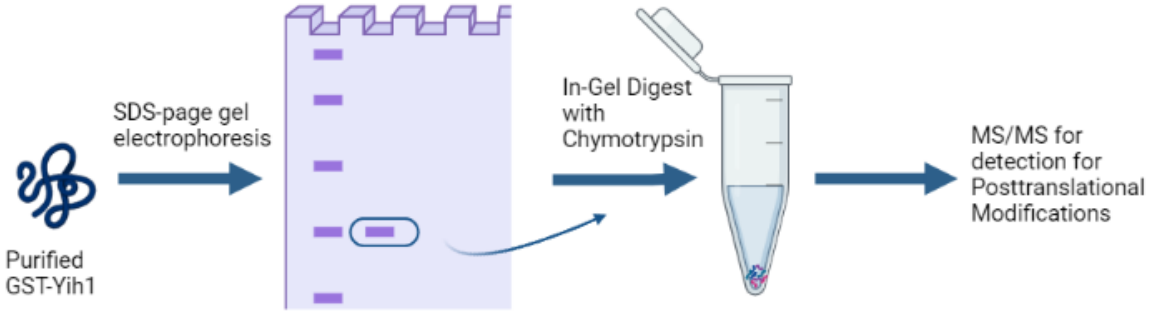


Figure 36: Preparation of GST-Yih1 for Tandem Mass Spectrometry (MS/MS)

GST-Yih1 was purified by SEC and then proteins were separated by SDS-page gel electrophoresis. Bands were stained with Colloidal Coomassie, excised from the gel, and digested with Chymotrypsin. The resulting peptides were subject to Tandem Mass Spectrometry (MS/MS) to detect Posttranslational Modifications.

Following on from SDS-page gel electrophoresis, pure GST-Yih1 fractions from FPLC (Section 4.2.4) were subject to an in-gel Chymotrypsin digest to enzymatically digest the protein into peptides. The resulting mixture of peptides were analysed by Tandem Mass Spectroscopy, and the peptides and/or posttranslational modifications were identified by matching the peptide fragment ion spectra to theoretical spectra based on a protein database. The aim of this was to identify the presence of Posttranslational modifications, such as R144 dimethylation, and this allowed us to do so. Tandem Mass Spectroscopy was carried out and analysed by Trevor Loo due to limited time and availability to equipment. These results will be further discussed.

Sequence coverage by Chymotrypsin was 43% and the measured molecular weight for GST-Yih1 was measured as 55kDa, which is consistent with the expected molecular weight of GST-tagged Yih1 (GST molecular weight: 26kDa + Yih1 molecular weight: 29kDa) (Fig. 37B). The peptide containing Arg144 was found 15 times by Peptide Spectrum Match, and not one was dimethylated or methylated (Fig. 37A and Fig. 37B). Therefore, we found no evidence of Yih1 R144 dimethylation when Yih1 was expressed in yeast. Other modifications that were detected includes Carbamidomethylation of a Cysteine residue (C67) and Oxidation of Methionine residues (as listed in Table 22). Due to the sequence coverage of GST-Yih1 being 43%, we cannot exclude the possibility that other Posttranslational modifications exist in Yih1 that was not covered.

In addition to R144, Lys36 (K36) was also found to be dimethylated when expressed in *E. coli* (Harjes et al., 2021). This sequence area was covered by Chymotrypsin, and no K36 dimethylation was identified in yeast (Fig. 37A). No other Posttranslational modifications were searched for outside of these parameters.

Table 22: Posttranslational modifications identified in Yih1 by MS/MS.

Yih1 residue	Post translational Modification	Delta Mass (Da)
Methionine-43 (M43)	Oxidation	31.9**
Cysteine-67 (C67)	Carbamidomethylation	57*
Methionine-149 (M149)	Oxidation	31.9**
Methionine-164 (M164)	Oxidation	31.9**
Methionine-217 (M217)	Oxidation	31.9**

*Delta Mass (Da) retrieved from Kim, Zhong, & Pandey, 2016

**Delta Mass (Da) retrieved from Raftery, 2014.

A


Coverage
ProteinCard
GST_YIH1
Annotate PTMs reported in Uniprot
Show only PTMs
Include PSMs that are Filtered Out
Coverage: 43.38%
Found Modifications:
C Carbamidomethyl (C)
O Oxidation (M)
Sequence
Modification List
K36
R144

B

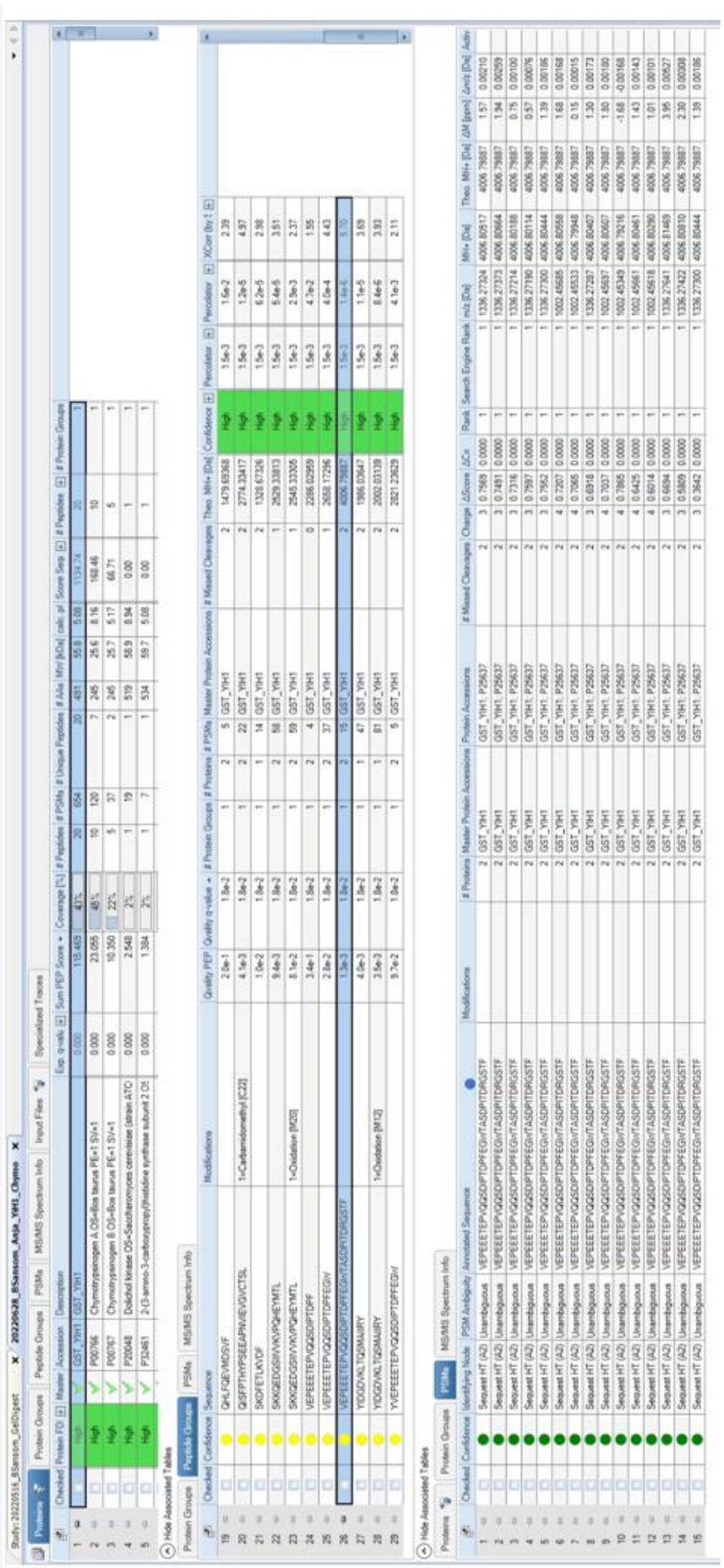


Figure 37: Analysis of Mass Spectrometry data of GST-Yih1

A: Sequence coverage by Chymotrypsin was 43%. The areas covered are highlighted in ‘green’. R144 is highlighted ‘blue’, with no posttranslational modifications detected. B: The measured molecular weight of GST-Yih1 was 55kDa. The peptide containing R144 was found 15 times and no peptide contained dimethylated R144. ● = R144 or K36 position.

4.3 Discussion

4.3.1 A non-methylatable version of Yih1 has enhanced potency in Gcn2 inhibition

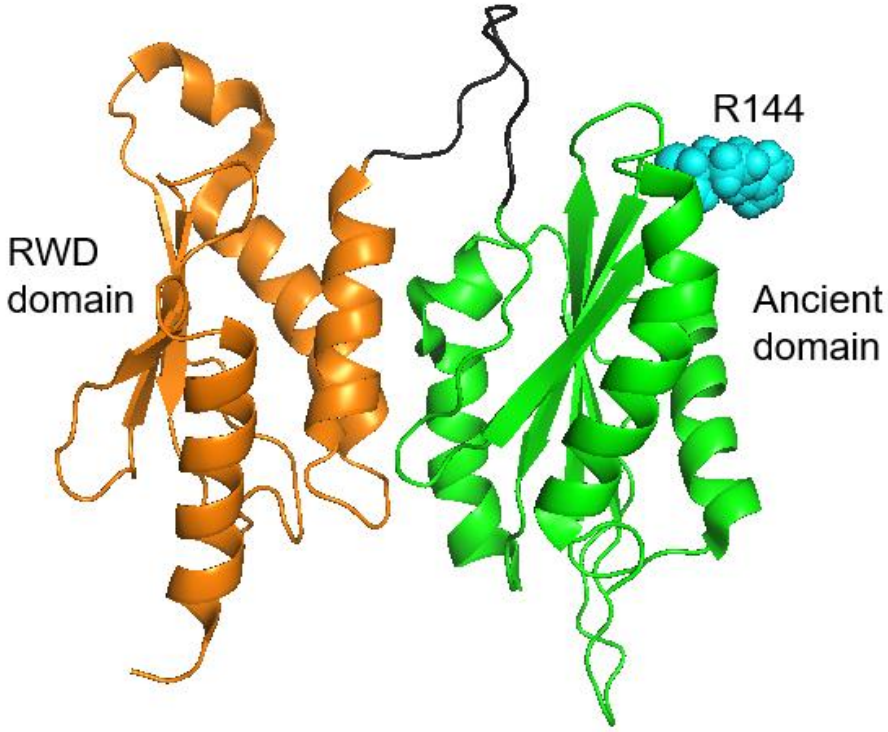


Figure 38: Position of Arg144 (R144) in the Ancient domain of Yih1

The RWD domain is coloured in orange, while the Ancient domain is coloured in green. The linker region is coloured dark grey. Arg144 is labelled in Cyan colour, with the side chain depicted as spheres. Yih1 structural model was constructed using PyMol™ Version 2.3.4.

Post translational modifications, such as arginine methylation, can underlie important mechanisms that modulate protein activity and binding activity (Bedford & Clarke, 2009). Roles of arginine methylation have been recognised to have physiological roles in cellular processes such as in signal transduction, transcriptional control, and protein translocation (Bedford & Clarke, 2009). Recent studies have shed light on an additional mechanism that could be involved in interplay between Yih1 and Gcn2 regulation. R144 was dimethylated when expressed in *E. coli* (Harjes et al., 2021). Authors suggested that this modification may modulate the activity of Yih1 (Harjes et al., 2021).

Methylation is the addition of a methyl group to the guanidinium nitrogen contained in the Arginine structure (Pahlich et al., 2006). In turn, this increases the hydrophobicity of the Arginine side chain, and therefore shields the surface charge that is normally provided by the positively charged Arginine (Guccione, 2019). This posttranslational modification influences

the biochemical properties of the protein, which impacts its function and activity. Our mutagenesis studies here showed that substituting Yih1 R144 to Alanine or Lysine enhances Yih1 potency, shown by stronger 3AT sensitivity in comparison to wild type Yih1 (Fig. 32). Furthermore, Yih1 potency was stronger by GST-Yih1^{R144K} overexpression compared to GST-Yih1^{R144A} overexpression (Fig. 32). By substituting R144 to Alanine, methylation at this site is prevented (Weber et al., 2009). Furthermore, the positive charge that is provided by Arginine is removed. By substituting R144 to Lysine, methylation also cannot occur, therefore mimicking non-methylatable Yih1 at position 144 (Weber et al., 2009). Therefore, our preliminary results suggest that non-methylated Yih1 has a greater potency in Gcn2 inhibition in comparison to normal Yih1. Hence, it is possible to propose that R144 has a role in the function of Yih1 in Gcn2 inhibition.

It cannot be excluded that these mutant proteins were not expressed at the same levels. As time restraints prevented us from testing their expression levels, this urgently needs to be revisited in the future. It will also be insightful to analyse eIF2 phosphorylation levels in yeast cells overexpression GST-Yih1^{R144A} or GST-Yih1^{R144K}, to directly determine their effect on Gcn2 activation.

Another potential interpretation of this result is that methylation at this site may be Arginine specific. To address this potential explanation, we decided to test whether R144 is dimethylated in its native environment yeast.

4.3.2 No evidence of Yih1 R144 dimethylation was detected in yeast.

Mass Spectrometry (MS) has been well established and recognised as a method for characterising Posttranslational modifications of proteins (Pahlich et al., 2006). Yih1 was found to be dimethylated when expressed in *E. coli* (Harjes et al., 2021). Furthermore, we witnessed a phenotype of enhanced Yih1 potency in a non-methylatable version of Yih1 (GST-Yih1^{R144A}). Upon separation of purified GST-Yih1 (by SEC) by SDS-page electrophoresis, a band of approximately 64kDa was detected (Fig. 35). This is higher than the predicted molecular weight of GST-Yih1 (55kDa), suggesting that Posttranslational modifications may be accounting for this difference in molecular weight. To determine whether Yih1 R144 was also dimethylated in its native environment (yeast), we utilised Mass Spectrometry of Chymotrypsin-digested GST-Yih1 peptide fragments. Tandem Mass Spectroscopy was carried out by Trevor Loo. Our MS data here showed no evidence of Yih1 R144 methylation or dimethylation when expressed in yeast (Fig. 37). There are numerous explanations for this result. One possibility is that overexpression of GST-Yih1 may have ‘overloaded’ the yeast cell with high protein levels,

disrupting the ability for R144 to become methylated. Another possibility is that R144 is not constitutively methylated, but rather that this methylation is regulated in a spatial or temporal manner. Accordingly, if this is true, the conditions and the specific physiological stimulus that encourages methylation needs to be established. Furthermore, the role that constitutive methylation has on the function of Yih1 in Gcn2 inhibition would have to be investigated. Lastly, it could be possible that R144 methylation is an artefact of Yih1 expression in *E. coli*.

In addition to R144, MS analysis from earlier studies also detected dimethylation of the K36 residue when Yih1 was expressed in *E. coli* (Harjes et al., 2021). However, there was no evidence implying that double dimethylation was present, suggesting that either R144 or K36 was dimethylated in *E.coli*. We found no evidence of K36 dimethylation in yeast (Fig. 37).

These are preliminary results, and the relatively low sequence coverage of GST-Yih1 by Chymotrypsin means that future research will benefit from repeating the enzymatic digest of GST-Yih1 to increase the efficacy of our MS results.

4.3.3 Biological analysis of the significance of additional Yih1 modifications by Mass Spectrometry

Due to the sequence coverage being 43%, we cannot exclude the possibility that other Posttranslational modifications or methylation sites are present in the protein sequence that was not covered. Despite finding no evidence of R144 dimethylation or methylation, MS identified other Yih1 modifications which will be further discussed. MS analysis, carried out by Trevor Loo, identified the Carbamidomethylation of Yih1 C67 (Table 22, Fig. 37B). Disulfide bonds between cysteine residues contributes to the formation of the tertiary protein structure (Suttapitugsakul et al., 2017). During an in-gel digest in preparation for MS/MS, these disulfide bonds need to broken, which would otherwise make it difficult to analyse the peptide sequence due to protein folding (Boja & Fales, 2001; Kim, Zhong, & Pandey, 2016; Suttapitugsakul et al., 2017). This is addressed by Reduction and Alkylation steps during in-gel digest (Materials and Methods, Section 2.8). In these steps, the sample is treated with a Reducing agent to break disulfide bonds between cysteine residues. Next, alkylation of the sample by the addition of Iodoacetamide (IAA) (in this study) prevents the re-formation of disulfide bonds between cysteine residues by Carbamidomethylation (Boja & Fales, 2001; Kim, Zhong, & Pandey, 2016; Suttapitugsakul et al., 2017). Therefore, according to the literature, Carbamidomethylation of cysteine residues is a common artefact of Mass Spectrometry-based peptide analysis that can occur during an in-gel digest (Boja & Fales, 2001; Kim, Zhong, &

Pandey, 2016; Pahlich et al., 2006; Suttapitugsakul et al., 2017). In saying this, the witnessed C67 Carbamidomethylation was likely an artefact produced during sample preparation, rather than an intentional posttranslational modification (Table 22, Fig. 37B).

MS analysis also identified the Oxidation of numerous Methionine residues of Yih1 (see Table 22 and Fig. 37A and Fig. 37B). Oxidation of residues such as Methionine is often another artefact caused by sample handling during preparation for Mass Spectrometry (Bettinger et al., 2019). This alludes to the ability of Methionine to become oxidised readily due to the presence of sulfur in its side chain (Bettinger et al., 2019; Schöneich, 2005). According to literature, Methionine oxidation is likely a result of spontaneous protein damage that occurs in response to oxidative stress during sample preparation (Bettinger et al., 2019; Schöneich, 2005). This can hinder the ability to identify and characterise sites of biologically significant oxidation sites that occur intentionally (Bettinger et al., 2019, Schöneich, 2005).In our research, the Oxidation of Methionine residues was likely to be an artefact caused by sample preparation. However, as these are preliminary results, future research will be insightful to determine the biological relevance, if any, of these oxidation events.

4.3.4 Potential role of Arg-144 in Yih1-mediated Gcn2 regulation

The fact that we did not detect R144 dimethylation in our MS analysis does not exclude the potential role of R144 in Yih1-mediated Gcn2 regulation. As overexpression of GST-Yih1^{R144A} enhanced the potency of Yih1 in Gcn2 inhibition, this strongly suggests a role of R144 in Yih1 function. Arg144 is positively charged and is situated in a highly conserved region of the ancient domain (Harjes et al., 2021). By analysing the structural model of Yih1, we can see that the R144 side chain protrudes from the protein surface, potentially for recognition by potential binding partners (Fig. 38). Furthermore, R144 itself is fully conserved in the Yih1/IMPACT protein family, implying that it is important for the proteins structure or function (Section 3.5).

As mentioned in Section 1.5.2, the Ancient domain is highly conserved, however the function of this domain remains relatively ambiguous. Earlier structural studies that resolved the crystal structure of YigZ (the Yih1 Ancient domain homologue) has given insight into the potential role of the Yih1 Ancient domain in DNA/RNA binding (Park et al., 2004; Zamora Caballero, 2016). This has been further supported by findings that IMPACT homolog in *Chaetomium thermophilum* (*C. thermophilum*) has nuclease activity (Zamora Caballero, 2016). Furthermore, the site of DNA/RNA binding has been mapped to a highly conserved and hydrophobic Ancient domain region that consists of a deep groove surrounded by positively charged residues,

including R144, that are capable of binding RNA (Harjes et al., 2021; Park et al., 2016; Zamora Caballero, 2016). An earlier study proposed that that R144 dimethylation may modulate the nuclease activity of Yih1/IMPACT and in turn determine when Yih1 is active or inactive in Gcn2 inhibition (Harjes et al., 2021). We cannot exclude the possibility that R144 is not constitutively methylated, and that this may modulate nuclease activity of Yih1. In turn, this may impact the ability of Yih1 to become active in Gcn2 inhibition. In saying this, further research is required to fully elucidate the role of R144 in Yih1-mediated Gcn2 inhibition. Yih1 Arg144 corresponds to the R181 residue in IMPACT (mouse) (Fig. 22). Following our studies, it would be insightful to carry out mutagenesis of R181 to Alanine in IMPACT to determine whether this substitution also enhances IMPACT potency. From here, a more in-depth experiment would involve MS to determine any posttranslational modifications, such as arginine methylation that exists in IMPACT.

4.4 General Discussion

4.4.1 Our yeast system is sufficient for identifying important amino acids in Yih1 function

It is crucial that we understand when and how Gcn2 is inhibited by Yih1. The aim of this research was to identify the relevant amino acids of Yih1 that are involved in impairing Gcn2 activation and the starvation response, using Baker's yeast (*S. cerevisiae*) as a model organism. To meet this aim, we have refined and utilised a yeast system that was first introduced by Sattlegger & Hinnebusch, (2000). This yeast system involves experimental procedures that utilises the phenotypic response of Gcn2 activation/inactivation to screen for important Yih1 residues in Gcn2 inhibition. Through this research, we have demonstrated that this system can be utilised as a screening method for finding more important residues in Yih1 function moving forwards (Section 3.1). This system includes substituting residues predicted to be important Yih1 function, overexpressing these mutant proteins in yeast (to drive the Yih1-Gcn1 interaction) and then examining the effect on Gcn2 activation. This can be determined phenotypically by analysing cell growth and eIF2α phosphorylation levels. Furthermore, this experimental procedure can be applied to a more widespread context, to other protein-protein interactions, given that an indicative phenotype can be analysed.

4.4.2 Specific Yih1 residues are involved in Gcn1 binding

In this work, we hypothesised that specific Yih1 amino acids are involved in Gcn1 binding to modulate Gcn2 inhibition. Our results show that the Yih1 residue D102, but not E106, is important for Gcn1 binding (Section 3.4.1). Computational modelling of the Yih1 structure shows D102 to be buried at the interface between the RWD and ancient domain (Harjes et al., 2021). Therefore, it is strongly implied that a conformational change must occur to expose this important binding determinant for Gcn1 binding. A potential explanation for our result is that mutating D102 to Alanine relaxed the Yih1 structure, exposing D102, the relevant Gcn1 binding determinant. However, while the D102A substitution relaxed the Yih1 structure, it diminished the ability of this Yih1 mutant to bind Gcn1, enabling Gcn2 activation to occur in the presence of overexpressed GST-Yih1^{D102A}.

As mentioned in Section 3.4.5, the phenotype of strains overexpressing GST-Yih1^{D102A} was not completely reverted, suggesting the involvement of other amino acids in Gcn1 binding. Therefore, the immediate next steps from here would be to screen the predicted Gcn1 binding region of Yih1 (amino acids 74-114) for other candidate residues that are involved in binding

Gcn1 (Sattlegger et al., 2011). Another "cherry on top" experiment would be to carry out physical interaction experiments, such as pull-down assays, to determine whether a Yih1 D102A substitution abolishes the Yih1-Gcn1 interaction.

Considering the biological relevance of IMPACT in neuronal function and development, a crucial prospective study following this work would be to screen IMPACT for the relevant Gcn1 binding determinants in mammalian IMPACT (mouse or human). In our MSA, we can see that IMPACT E107 and D111 are located in a similar region of the RWD domain to Yih1 D102 and E106 and are also negatively charged (Fig. 22). Furthermore, Q104 (mouse IMPACT) and L104 (human IMPACT) structurally occupy the same position as Yih1 D102 (see Fig. 39). Despite not being highly conserved, the structural positioning homology indicates that these residues would be interesting candidates to substitute and analyse the effect of their overexpression of Gcn2 activity (Fig. 22, Fig. 39).

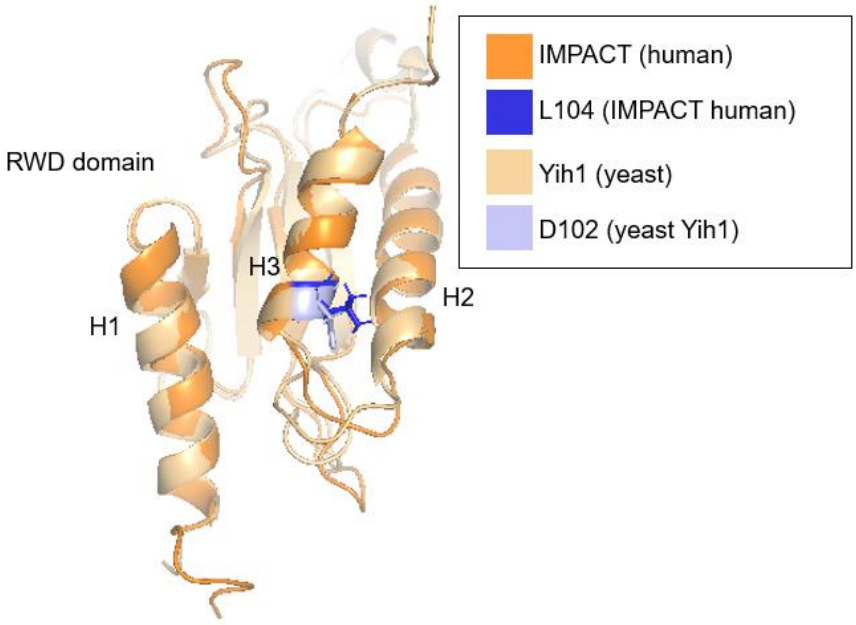


Figure 39: Superimposition of the Yih1/IMPACT RWD domain.

Shown above is the RWD domains of Yih1 (*S. cerevisiae*) and IMPACT (*H. sapiens*) structurally superimposed. D102 and L104, which occupy the same position as yeast D102, is labelled. H= Helix. Figure generated through PyMOL™ Version 2.3.4.

4.4.3 Ionic interactions are involved interdomain Yih1 interactions

In this work, we have also hypothesised that specific Yih1 amino acids are involved in interdomain interactions which affect the ability of Yih1 to modulate Gcn2 activity. As shown in Fig. 26, the interface between the Yih1 domains contains a great proportion of negatively

charged residues of the RWD domain (Group 2) and positively charged residues of the Ancient domain (Group 3), which are all relatively structurally and positionally close to each other (Harjes et al., 2021). Earlier research proposed that ionic interactions may be involved in maintaining the relative conformations of the Yih1 RWD and Ancient domain, and that these interactions may be involved in modulating the activity of Yih1 (Harjes et al., 2021; Sattlegger et al., 2011). However, the connection between these interdomain interactions and how this may determine whether Yih1 is in its inactive or active form in Gcn2 inhibition are unclear.

To examine the role of ionic interactions in this, we substituted Group 2 residues, and Group 3 residues, to the opposite charge. The rationale behind choosing an oppositely charged amino acid is that if Yih1 interdomain interactions are mediated by ionic interactions, then we would expect that upon substitution of these residues to the opposite charge, electrostatic repulsions would cause Yih1 to change into a more open conformation. In turn, this would result in stronger Gcn2 inhibition. We also substituted these residues to Alanine, thereby removing the charge, to determine whether the substitution would relax the Yih1 structure and thereby also result in stronger Gcn2 inhibition. Accordingly, we found that substituting E87 and D90 to the opposite charge, or removing the charge completely, enhances Yih1 potency in Gcn2 inhibition (Section 3.7.1). Furthermore, substituting K173, R175 and K176 together to the opposite charge, or removing the charge, also enhances Yih1 potency (Section 3.9.2). Therefore, through our mutagenesis studies, it appears that removing the charge of these residues enhances Yih1 potency (Section 3.7.1 and Section 3.9.2). Consistently with our speculation, this suggests that these substitutions relax and/or disrupt the interdomain electrostatic interactions, thereby exposing Yih1 to Gcn1 and allowing stronger Gcn2 inhibition. Therefore, we have determined that interdomain interactions between Yih1 are likely to be ionic, and that Yih1 potency is affected by the relative conformation of these domains. We have also provided experimental evidence that identifies specific charged amino acids in each Yih1 domain that may be involved in maintaining these interdomain interactions.

There are considerations that need to be addressed when analysing the ability of Yih1 with substitutions to impair Gcn2 activation. Firstly, the level that a protein is expressed in the cell affects its affinity to its binding partners. Therefore, higher expression levels of a Yih1 mutant generates a stronger ability for Yih1 to bind Gcn1 through mass action. Furthermore, it is possible that mutating specific amino acids of Yih1 alters the protein structure sterically, and therefore its function. This is because different residues have different biochemical properties.

These considerations highlight the importance of measuring expression levels to ensure that the Yih1 proteins are being stably expressed in the yeast cell.

As mentioned prior, mutagenesis can change the protein structure, which can further alter protein function. In this study, we selected Alanine substitutions due to the biochemical properties of alanine being a small, non-bulky residue because of its simple side chain consisting of a methyl group (CH_3) (Cunningham & Wells, 1989). Furthermore, alanine is chemically inert. Together this means that substitution to Alanine does not impose any strong steric effects or hindrances, meaning that native protein structure is preserved and the risk of this mutation on changing the protein structure is minimized (Cunningham & Wells, 1989). This makes Alanine a suitable candidate for mutagenesis to examine the structural and functional activity of Yih1 in the context of this study. As we saw a strong 3AT sensitivity phenotype, consistent with enhanced Yih1 potency, when substituting the charged residues to Alanine, this strongly suggests that these residues are involved in maintaining interdomain interactions. Therefore, we can see that maintaining the correct residue charge is important for maintaining Yih1 interdomain interactions.

It could also be speculated that the interdomain Yih1 interactions may not be solely due to one exclusive amino acid/interaction, but rather the charge that it possesses. A prospective study to test this would be to swap the charges of all Group 2 and Group 3 residues and examine the effect that this has on Yih1 potency using our established yeast system. If the interdomain interactions are primarily based on charge relevance, we would expect that swapping the charge of E87 and D90 and K173, R175 and K176 would result in a phenotype like that of GST-Yih1 wild type. This is because it would be expected that electrostatic interactions would be maintained, meaning that the Yih1 conformation would not change with respect to its wild type conformation. However, potential roadblocks to this could be that introducing a high number of mutations into Yih1 may disrupt the ability of the protein to be expressed or may disrupt the protein structure. This would have to be considered and accounted for accordingly.

Through our studies, we have gained more insight into the interdomain interactions of Yih1, which are likely to involved electrostatic interactions between charged residues. If a conformational change of Yih1 is involved to prime Yih1 for Gcn2 inhibition, it is yet to be determined when this change occurs. It is possible to speculate that dynamic changes in the cellular state, such as starvation, may influence the relative conformation of the RWD domain and the Ancient domain. Furthermore, the conformation of Yih1 may be modulated by its

different binding partners, such as Actin or the ribosome. Together, our preliminary findings contribute to building a map of the amino acids involved in maintaining interdomain interactions between Yih1. They also give us more insight into how disruption of these interactions impacts the ability of Yih1 to inhibit Gcn2.

4.4.4 Shedding light on additional regulatory mechanisms that modulate Yih1-mediated Gcn2 inhibition

Studies have discussed the potential involvement of Arginine methylation in influencing the ability of Yih1 to inhibit Gcn2 (Harjes et al., 2021). This came about upon the finding that R144 was dimethylated when expressed in *E. coli* (Harjes et al., 2021). Authors suggested that R144 dimethylation may determine when Yih1 can become active or inactive in inhibiting Gcn2 (Harjes et al., 2021). In this research, we also aimed to shed light on additional potential mechanisms of Yih1 regulation that may modulate Gcn2 inhibition. We hypothesised that additional regulatory mechanisms are in place that modulate the ability of Yih1 to inhibit Gcn2. Although we found no evidence of R144 dimethylation when Yih1 was expressed in yeast, substituting R144 to Alanine or Lysine strongly enhances Yih1 potency, suggesting a role of this residue in Yih1-mediated Gcn2 inhibition (Section 4.2.2).

The fact that a band of 64kDa (higher than the predicated GST-Yih1 molecular weight of 55kDa) was resolved by SDS-page electrophoresis following GST-Yih1 purification by SEC is important to consider moving forward (Fig 35A). Because of the difference in molecular weight, the possibility of a residue containing a posttranslational modification wasn't covered by Chymotrypsin could be speculated. More research is required to elucidate the involvement of PTMs in Yih1-mediated Gcn2 inhibition, and furthermore to determine the role of R144 in modulating the ability of Yih1 to inhibit Gcn2.

4.5 Conclusion and Prospective studies

Gcn2 is central to the GAAC response, a signalling pathway aimed at changing the cellular gene expression profile in response to environmental or intracellular stress, such as amino acid starvation, in eukaryotes. Therefore, Gcn2 activity needs to be tightly regulated to maintain cellular homeostasis. Gcn1 is an absolutely essential binding partner of Gcn2 (Castilho et al., 2014). Yih1 is an inhibitor of Gcn2, and it carries out this inhibition by competing with Gcn2 for Gcn1 binding (Sattlegger et al., 2004). At present, the only known amino acids that are important for Yih1 function are D102 and E106 together (Sattlegger et al., 2011). Considering the biological implications of the Yih1/IMPACT-Gcn2 modulation axis (Section 1.8), it is crucial to understand the molecular mechanisms behind Yih1-mediated Gcn2 regulation.

The aim of this study was to identify relevant amino acids of Yih1 that are involved in impairing Gcn2 activation and the starvation response. To do this, we have grouped our proposed amino acids into 4 groups; amino acids potentially involved in Gcn1 binding (Group 1), RWD domain amino acids potentially involved in interdomain interactions (Group 2), ancient domain amino acids potentially involved in interdomain interactions (Group 3), and amino acids potentially involved in posttranslational modifications that may modulate Yih1 activity (Group 4). Through this research, utilising our established yeast system, we have found that D102 alone is crucial for Gcn1 binding. We have also determined that interdomain interactions between Yih1 are likely to be ionic, and that the conformation of these domains determines the potency of Yih1. In parallel, we have provided experimental evidence that sheds light on specific amino acids in each Yih1 domain that may be involved in maintaining these interdomain interactions. We have also found that although R144 is not dimethylated in yeast, substitution of this residue to alanine strongly enhances Yih1 potency. This preliminary result sheds light on another potential means of Yih1 regulation.

Specific prospective studies and future directions for our amino acid groups have been discussed in Sections 3.4.6, 3.8.5, 4.1.3, 4.3.4 and 4.4. However, the important overarching experiments moving forwards would involve introducing the equivalent substitutions (from this study) into IMPACT (mammalian) to determine how these substitutions affect Gcn2 activity by analysing yeast cell growth and eIF2α phosphorylation levels in starved conditions. We have shown that IMPACT overexpression in yeast impairs Gcn2 activity, shown by a strong 3AT sensitivity phenotype (Section 3.1.7). Therefore, yeast is a sufficient expression system to overexpress IMPACT with substitutions to predicted important residues and determine their effect on Gcn2 activity.

It is crucial that we understand when and how Gcn2 is inhibited by Yih1/IMPACT. This will bring us a step closer to understanding diseases associated with Gcn2 dysregulation, like Alzheimer's disease and other neuronal disorders. Our work has provided insight into this area by shedding light on important amino acids for Gcn1 binding, interdomain interactions, and potential other modulation mechanisms such as posttranslational modifications.

4.6 References

Akram, Z., Ahmed, I., Mack, H., Kaur, R., Silva, R. C., Castilho, B. A., . . . Munn, A. L. J. C. (2020). Yeast as a model to understand actin-mediated cellular functions in mammals—illustrated with four actin cytoskeleton proteins. *9*(3), 672. doi:10.3390/cells9030672

Andrews, B., Boone, C. M., Davis, T., & Fields, S. (2016). *Budding Yeast: A Laboratory Manual*: Cold Spring Harbor Laboratory Press.

Bedford, M. T., & Clarke, S. G. (2009). Protein Arginine Methylation in Mammals: Who, What, and Why. *Molecular Cell, 33*(1), 1-13. doi:10.1016/j.molcel.2008.12.013

Bettinger, J. Q., Welle, K. A., Hryhorenko, J. R., & Ghaemmaghami, S. J. J. o. p. r. (2019). Quantitative analysis of in vivo methionine oxidation of the human proteome. *19*(2), 624-633. doi:10.1021/acs.jproteome.9b00505

Bittencourt, S., Avedissian, M., Mello, L. E., Pereira, C. M., Delamano, A., & Castilho, B. A. (2008). Distribution of the protein IMPACT, an inhibitor of GCN2, in the mouse, rat, and marmoset brain. *Journal of Comparative Neurology, 507*(5), 1811-1830-1830. doi:10.1002/cne.21652

Boja, E. S., & Fales, H. M. J. A. c. (2001). Overalkylation of a protein digest with iodoacetamide. *73*(15), 3576-3582. doi:10.1021/ac0103423

Cambiaghi, T. D., Pereira, C. M., Shanmugam, R., Bolech, M., Wek, R. C., Sattlegger, E., & Castilho, B. A. (2014). Evolutionarily conserved IMPACT impairs various stress responses that require GCN1 for activating the eIF2 kinase GCN2. *Biochemical and Biophysical Research Communications, 443*(2), 592-597. doi:10.1016/j.bbrc.2013.12.021

Castilho, B. A., Shanmugam, R., Silva, R. C., Ramesh, R., Himme, B. M., & Sattlegger, E. (2014). Keeping the eIF2 alpha kinase Gcn2 in check. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research, 1843*(9), 1948-1968. doi:10.1016/j.bbamcr.2014.04.006

Cunningham, B. C., & Wells, J. A. J. S. (1989). High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *244*(4908), 1081-1085. doi:10.1126/science.2471267

Dayhoff, M., Schwartz, R., Orcutt, B. J. A. o. p. s., & structure. (1978). A model of evolutionary change in proteins. *5*, 345-352. doi:doi:10.1.1.145.4315

Dever, T. E., Chen, J. J., Barber, G. N., Cigan, A. M., Feng, L., Donahue, T. F., . . . Hinnebusch, A. G. (1993). Mammalian eukaryotic initiation factor 2 alpha kinases

functionally substitute for GCN2 protein kinase in the GCN4 translational control mechanism of yeast. *Proc Natl Acad Sci U S A, 90*(10), 4616-4620. doi:10.1073/pnas.90.10.4616

Dong, J. S., Qiu, H. F., Garcia-Barrio, M., Anderson, J., & Hinnebusch, A. G. (2000). Uncharged tRNA activates GCN2 by displacing the protein kinase moiety from a bipartite tRNA-Binding domain. In (Vol. 6, pp. 269-279).

Feng, D. F., Johnson, M. S., & Doolittle, R. F. (1985). Aligning amino acid sequences: Comparison of commonly used methods. *Journal of Molecular Evolution, 21*(2), 112-125. doi:10.1007/BF02100085

Ferraz, R. C., Camara, H., De-Souza, E. A., Pinto, S., Pinca, A. P. F., Sato, V. N., . . . Castilho, B. A. (2016). IMPACT is a GCN2 inhibitor that limits lifespan in Caenorhabditis elegans. *BMC Biology, 14*(1) doi:10.1186/s12915-016-0301-2

Foiani, M., Cigan, A. M., Paddon, C. J., Harashima, S., & Hinnebusch, A. G. (1991). GCD2, a translational repressor of the GCN4 gene, has a general function in the initiation of protein synthesis in Saccharomyces cerevisiae. *Mol Cell Biol, 11*(6), 3203-3216. doi:10.1128/mcb.11.6.3203

Fuhrmann, J., Clancy, K. W., & Thompson, P. R. (2015). Chemical Biology of Protein Arginine Modifications in Epigenetic Regulation. *Chemical Reviews, 115*(11), 5413-5461. doi:10.1021/acs.chemrev.5b00003

Gallinetti, J., Harputlugil, E., & Mitchell, J. R. (2013). Amino acid sensing in dietary-restriction-mediated longevity: roles of signal-transducing kinases GCN2 and TOR. *Biochem J, 449*(1), 1-10. doi:10.1042/bj20121098

Garcia-Barrio, M., Dong, J., Cherkasova, V. A., Zhang, X., Zhang, F., Ufano, S., . . . Hinnebusch, A. G. (2002). Serine 577 is phosphorylated and negatively affects the tRNA binding and eIF2alpha kinase activities of GCN2. *The Journal of biological chemistry, 277*(34), 30675-30683. doi:10.1074/jbc.M203187200

Garcia-Barrio, M., Dong, J., Ufano, S., & Hinnebusch, A. G. (2000). Association of GCN1-GCN20 regulatory complex with the N-terminus of eIF2alpha kinase GCN2 is required for GCN2 activation. *The EMBO journal, 19*(8), 1887-1899. doi:10.1093/emboj/19.8.1887

Gasparski, A. N., Mason, D. E., Moissoglu, K., & Mili, S. (2022). Regulation and outcomes of localized RNA translation. *Wiley interdisciplinary reviews. RNA*, e1721. doi:10.1002/wrna.1721

Guccione, E., & Richard, S. (2019). The regulation, functions and clinical relevance of arginine methylation. *Nature Reviews Molecular Cell Biology, 20*(10), 642-657. doi:10.1038/s41580-019-0155-x

Hagiwara, Y., Hirai, M., Nishiyama, K., Kanazawa, I., Ueda, T., Sakaki, Y., & Ito, T. J. P. o. t. N. A. o. S. (1997). Screening for imprinted genes by allelic message display: identification of a paternally expressed gene impact on mouse chromosome 18. *94*(17), 9249-9254. doi:10.1073/pnas.94.17.9249

Harding, H. P., Ordonez, A., Allen, F., Parts, L., Inglis, A. J., Williams, R. L., & Ron, D. J. E. (2019). The ribosomal P-stalk couples amino acid starvation to GCN2 activation in mammalian cells. *8*, e50149. doi:10.7554/eLife.50149

Harjes, E., Jameson, G. B., Tu, Y. H., Burr, N., Loo, T. S., Goroncy, A. K., . . . Göbl, C. J. F. l. (2021). Experimentally based structural model of Yih1 provides insight into its function in controlling the key translational regulator Gcn2. *595*(3), 324-340. doi:10.1002/1873-3468.13990

Hinnebusch, A. G. (2005). Translational Regulation of GCN4 and the General Amino Acid Control of Yeast. *Annual Review of Microbiology, 59*(1), 407-450. doi:10.1146/annurev.micro.59.031805.133833

Hinnebusch, A. G., & Natarajan, K. J. E. c. (2002). Gcn4p, a master regulator of gene expression, is controlled at multiple levels by diverse signals of starvation and stress. *1*(1), 22-32. doi:10.1128/EC.01.1.22-32.2002

Hinnebusch, A. G. J. J. o. B. C. (1997). Translational regulation of yeast GCN4: a window on factors that control initiator-tRNA binding to the ribosome. *272*(35), 21661-21664. doi:10.1074/jbc.272.35.21661

Hotamisligil, G. S., Erbay, E., & Hotamisligil, G. S. (2008). Nutrient sensing and inflammation in metabolic diseases. *Nature Reviews Immunology, 8*(12), 923-934. doi:10.1038/nri2449

Inglis, A. J., Masson, G. R., Shao, S., Perisic, O., McLaughlin, S. H., Hegde, R. S., & Williams, R. L. J. P. o. t. N. A. o. S. (2019). Activation of GCN2 by the ribosomal P-stalk. *116*(11), 4946-4954. doi:10.1073/pnas.1813352116

Ishimura, R., Nagy, G., Dotu, I., Chuang, J. H., & Ackerman, S. L. J. E. (2016). Activation of GCN2 kinase by ribosome stalling links translation elongation with translation initiation. *5*, e14295. doi:10.7554/eLife.14295

Kim, M. S., Zhong, J., & Pandey, A. J. P. (2016). Common errors in mass spectrometry-based analysis of post-translational modifications. *16*(5), 700-714. doi:10.1002/pmic.201500355

Kubota, H., Ota, K., Sakaki, Y., & Ito, T. (2001). Budding yeast GCN1 binds the GI domain to activate the eIF2alpha kinase GCN2. *J Biol Chem, 276*(20), 17591-17596. doi:10.1074/jbc.M011793200

Kubota, H., Sakaki, Y., & Ito, T. (2000). GI domain-mediated association of the eukaryotic initiation factor 2alpha kinase GCN2 with its activator GCN1 is required for general amino acid control in budding yeast. *J Biol Chem, 275*(27), 20243-20246. doi:10.1074/jbc.C000262200

Lageix, S., Dever, T. E., Hinnebusch, A. G., & Rothenburg, S. (2014). Enhanced Interaction between Pseudokinase and Kinase Domains in Gcn2 stimulates eIF2α Phosphorylation in Starved Cells. *PLoS Genetics, 10*(5) doi:10.1371/journal.pgen.1004326

Lageix, S., Zhang, J., Rothenburg, S., & Hinnebusch, A. G. (2015). Interaction between the tRNA-Binding and C-Terminal Domains of Yeast Gcn2 Regulates Kinase Activity In Vivo. *PLoS Genetics, 11*(2), 1-28. doi:10.1371/journal.pgen.1004991

Lee, S. H., Dominguez, R. J. M., & cells. (2010). Regulation of actin cytoskeleton dynamics in cells. *29*(4), 311-325. doi:10.1007/s10059-010-0053-8

Liu, Y., Chen, X., Zhang, Y., Wan, Y., Wu, T., & Song, J. (2013). A mutant screening method by critical annealing temperature-PCR for site-directed mutagenesis. *BMC Biotechnology, 13* doi:10.1186/1472-6750-13-21

Marton, M. J., Crouch, D., & Hinnebusch, A. G. (1993). GCN1, a translational activator of GCN4 in Saccharomyces cerevisiae, is required for phosphorylation of eukaryotic translation initiation factor 2 by protein kinase GCN2. *Mol Cell Biol, 13*(6), 3541-3556. doi:10.1128/mcb.13.6.3541

Marton, M. J., Vazquez de Aldana, C. R., Qiu, H., Chakraburtty, K., & Hinnebusch, A. G. (1997). Evidence that GCN1 and GCN20, translational regulators of GCN4, function on elongating ribosomes in activation of eIF2alpha kinase GCN2. *17*(8), 4474-4489. doi:10.1128/MCB.17.8.4474 %J Molecular and Cellular Biology

Masson, G. R. (2019). Towards a model of GCN2 activation. *Biochemical Society Transactions, 47*(5), 1481-1488. doi:10.1042/BST20190331 %J Biochemical Society Transactions

Munn, A., & Sattlegger, E. J. A. B. (2009). Yeast studies reveal new roles for an ancient skeleton. *40*, 9-13.

Nameki, N., Yoneyama, M., Koshiba, S., Tochio, N., Inoue, M., Seki, E., . . . Yokoyama, S. (2004). Solution structure of the RWD domain of the mouse GCN2 protein. *Protein science : a publication of the Protein Society, 13*(8), 2089-2100. doi:10.1110/ps.04751804

Nozaki, Y., & Bellgard, M. J. B. (2005). Statistical evaluation and comparison of a pairwise alignment algorithm that a priori assigns the number of gaps rather than employing gap penalties. *21*(8), 1421-1428. doi:10.1093/bioinformatics/bti198

Padyana, A. K., Burley, S. K., Roll-Mecak, A., Qiu, H., & Hinnebusch, A. G. (2005). Structural basis for autoinhibition and mutational activation of eukaryotic initiation factor 2α protein kinase GCN2. *Journal of Biological Chemistry, 280*(32), 29289-29299. doi:10.1074/jbc.M504096200

Pahlich, S., Zakaryan, R. P., Gehring, H. J. B. e. B. A.-P., & Proteomics. (2006). Protein arginine methylation: Cellular functions and methods of analysis. *1764*(12), 1890-1903. doi:10.1016/j.bbapap.2006.08.008

Park, F., Gajiwala, K., Eroshkina, G., Furlong, E., He, D., Batiyenko, Y., . . . Bioinformatics. (2004). Crystal structure of YIGZ, a conserved hypothetical protein from Escherichia coli k12 with a novel fold. *55*(3), 775-777. doi:10.1002/prot.20087

Pereira, C. M., Sattlegger, E., Jiang, H. Y., Longo, B. M., Jaqueta, C. B., Hinnebusch, A. G., . . . Castilho, B. A. (2005). IMPACT, a protein preferentially expressed in the mouse brain, binds GCN1 and inhibits GCN2 activation. *J Biol Chem, 280*(31), 28316-28323. doi:10.1074/jbc.M408571200

Qiu, H., Garcia-Barrio, M. T., & Hinnebusch, A. G. (1998). Dimerization by Translation Initiation Factor 2 Kinase GCN2 Is Mediated by Interactions in the C-Terminal Ribosome-Binding Region and the Protein Kinase Domain. *18*(5), 2697-2711. doi:10.1128/MCB.18.5.2697 %J Molecular and Cellular Biology

Qiu, H. F., Hu, C. H., Dong, J. S., & Hinnebusch, A. G. (2002). Mutations that bypass tRNA binding activate the intrinsically defective kinase domain in GCN2. In (Vol. 16, pp. 1271-1280).

Raftery, M. J. (2014). Determination of oxidative protein modifications using mass spectrometry. *Redox Report, 19*(4), 140-147. doi:10.1179/1351000214Y.0000000089

Rakesh, R., Krishnan, R., Sattlegger, E., & Srinivasan, N. (2017). Recognition of a structural domain (RWDBD) in Gcn1 proteins that interacts with the RWD domain containing proteins. *Biology Direct, 12*(1), 12. doi:10.1186/s13062-017-0184-3

Ramesh, R. (2016). *Unraveling the dynamics of protein-protein interactions in the Gcn2 signal transduction pathway : a thesis presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Microbiology and Genetics, Massey University, Albany, New Zealand.* (Doctor of Philosophy (Ph.D.) Doctoral), Massey University, Retrieved from http://hdl.handle.net/10179/10003

Ramesh, R., Dautel, M., Lee, Y., Kim, Y., Storey, K., Gottfried, S., . . . Sattlegger, E. J. F. l. (2021). Asp56 in actin is critical for the full activity of the amino acid starvation-responsive kinase Gcn2. *595*(14), 1886-1901. doi:10.1002/1873-3468.14137

Roffé, M., Hajj, G. N., Azevedo, H. F., Alves, V. S., & Castilho, B. A. (2013). IMPACT is a developmentally regulated protein in neurons that opposes the eukaryotic initiation factor 2α kinase GCN2 in the modulation of neurite outgrowth. *J Biol Chem, 288*(15), 10860-10869. doi:10.1074/jbc.M113.461970

Sattlegger, E., Barbosa, J. A., Moraes, M. C., Martins, R. M., Hinnebusch, A. G., & Castilho, B. A. (2011). Gcn1 and actin binding to Yih1: implications for activation of the eIF2 kinase GCN2. *J Biol Chem, 286*(12), 10341-10355. doi:10.1074/jbc.M110.171587

Sattlegger, E., & Hinnebusch, A. G. (2000). Separate domains in GCN1 for binding protein kinase GCN2 and ribosomes are required for GCN2 activation in amino acid-starved cells. *19*(23), 6622-6633. doi:10.1093/emboj/19.23.6622

Sattlegger, E., & Hinnebusch, A. G. J. J. o. B. C. (2005). Polyribosome binding by GCN1 is required for full activation of eukaryotic translation initiation factor 2α kinase GCN2 during amino acid starvation. *280*(16), 16514-16521. doi:10.1074/jbc.M414566200

Sattlegger, E., Swanson, M. J., Ashcraft, E. A., Jennings, J. L., Fekete, R. A., Link, A. J., & Hinnebusch, A. G. J. J. o. B. C. (2004). YIH1 is an actin-binding protein that inhibits protein kinase GCN2 and impairs general amino acid control when overexpressed. *279*(29), 29952-29962. doi:10.1074/jbc.M404009200

Schöneich, C. (2005). Methionine oxidation by reactive oxygen species: reaction mechanisms and relevance to Alzheimer's disease. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics, 1703*(2), 111-119. doi:10.1016/j.bbapap.2004.09.009

Silje, A., Róbert, Z., & Beáta, G. (2017). Activation of Gcn2 in response to different stresses. *PLoS ONE*(8), e0182143. doi:10.1371/journal.pone.0182143

Silva, R. C., Dautel, M., Di Genova, B. M., Amberg, D. C., Castilho, B. A., & Sattlegger, E. (2015). The Gcn2 Regulator Yih1 Interacts with the Cyclin Dependent Kinase Cdc28 and Promotes Cell Cycle Progression through G2/M in Budding Yeast. *PLoS ONE, 10*(7), 1-29. doi:10.1371/journal.pone.0131070

Silva, R. C., Sattlegger, E., & Castilho, B. A. (2016). Perturbations in actin dynamics reconfigure protein complexes that modulate GCN2 activity and promote an eIF2 response. *129*(24), 4521-4533. doi:10.1242/jcs.194738 %J Journal of Cell Science

Simone, B., Catia M, P., Marcelo, A., Arthur, D., Luiz E, M., & Beatriz A, C. (2008). Distribution of the protein IMPACT, an inhibitor of GCN2, in the mouse, rat, and marmoset brain. *Journal of Comparative Neurology*(5), 1811. doi:10.1002/cne.21652

Srikanth, B., Gillipsie, M., Aman, M., & Nooruddin, K. (2017). Amino Acid Sensing via General Control Nonderepressible-2 Kinase and Immunological Programming. *Frontiers in Immunology* doi:10.3389/fimmu.2017.01719

Suttapitugsakul, S., Xiao, H., Smeekens, J., & Wu, R. J. M. B. (2017). Evaluation and optimization of reduction and alkylation methods to maximize peptide identification with MS-based proteomics. *13*(12), 2574-2582. doi:10.1039/c7mb00393e

Tiwary, B. K. (2022). *Bioinformatics and computational biology : a primer for biologists*: Springer.

Tomek, P., Gore, S. K., Potts, C. L., Print, C. G., Black, M. A., Hallermayr, A., . . . Ching, L.-M. (2018). Imprinted and ancient gene: a potential mediator of cancer cell survival during tryptophan deprivation. *Cell Communication and Signaling, 16*(1), 88. doi:10.1186/s12964-018-0301-7

Triana, F. J., Nierhaus, K. H., Ziehler, J., & Chakraburtty, K. (1993). Defining the function of EF-3, a unique elongation factor in low fungi. In *The translational apparatus* (pp. 327-338): Springer. doi:10.1007/978-1-4615-2407-6_31

Vazquez de Aldana, C. R., Marton, M. J., & Hinnebusch, A. G. (1995). GCN20, a novel ATP binding cassette protein, and GCN1 reside in a complex that mediates activation of the eIF-2 alpha kinase GCN2 in amino acid-starved cells. *14*(13), 3184-3199. doi:10.1002/j.1460-2075.1995.tb07321.x

Waller, T., Lee, S. J., & Sattlegger, E. (2012). Evidence that Yih1 resides in a complex with ribosomes. *Febs j, 279*(10), 1761-1776. doi:10.1111/j.1742-4658.2012.08553.x

Weber, S., Maaß, F., Schuemann, M., Krause, E., Suske, G., & Bauer, U.-M. (2009). PRMT1-mediated arginine methylation of PIAS1 regulates STAT1 signaling. *Genes & Development, 23*(1), 118-132. doi:10.1101/gad.489409

Wek, S. A., Zhu, S., & Wek, R. C. (1995). The histidyl-tRNA synthetase-related sequence in the eIF-2 alpha protein kinase GCN2 interacts with tRNA and is required for activation in response to starvation for different amino acids. *Mol Cell Biol, 15*(8), 4497-4506. doi:10.1128/MCB.15.8.4497

Wu, Q., Schapira, M., Arrowsmith, C. H., & Barsyte-Lovejoy, D. (2021). Protein arginine methylation: from enigmatic functions to therapeutic targeting. *Nature Reviews Drug Discovery, 20*(7), 509-530. doi:10.1038/s41573-021-00159-8

Yuriko, H., Momoki, H., Kazutoshi, N., Ichiro, K., Takayuki, U., Yoshiyuki, S., & Takashi, I. (1997). Screening for Imprinted Genes by Allelic Message Display: Identification of a Paternally Expressed Gene Impact on Mouse Chromosome 18. *Proceedings of the National Academy of Sciences of the United States of America, 94*(17), 9249. doi:10.1073/pnas.94.17.9249

Zamora Caballero, S. (2016). Structural and functional characterization of IMPACT proteins: a novel nuclease family.

Zhu, S., Sobolev, A. Y., & Wek, R. C. J. J. o. B. C. (1996). Histidyl-tRNA synthetase-related sequences in GCN2 protein kinase regulate in vitro phosphorylation of eIF-2. *271*(40), 24989-24994. doi:10.1074/jbc.271.40.24989

Zhu, S., & Wek, R. C. (1998). Ribosome-binding domain of eukaryotic initiation factor-2 kinase GCN2 facilitates translation control. *Journal of Biological Chemistry, 273*(3), 1808-1814. doi:10.1074/jbc.273.3.1808

4.7 Supplementary Material

Table S1: List of reagents used in preparation of solutions and their manufactures.

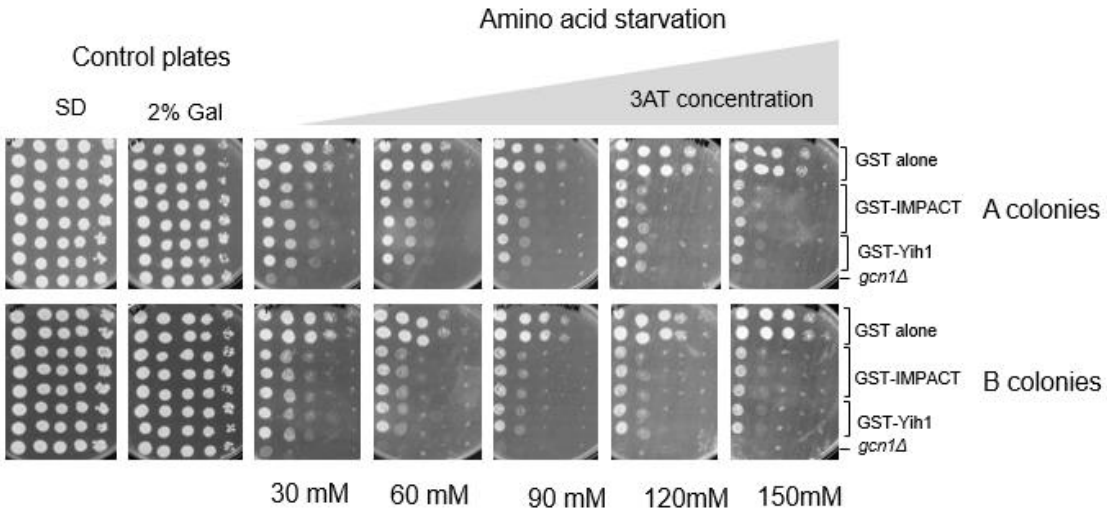
Reagents	**Abbreviation**	**Manufacturer name**	**Country**
Agar		FORMEDIUM™	United Kingdom
Ammonium Persulphate	APS	ACRO ORGANICS	United states of America
Ammonium Bicarbonate	ABC	SIGMA™	United states of America
Acetic Acid		RCI Labscan	Bangkok, Thailand
Acetonitrile (Methyl Cyanide)	MeCN	Fisher Scientific	United States of America
Agarose		FORMEDIUM™	United Kingdom
β-mecaptoethanol	BME	SIGMA™	United states of America
Chymotrypsinogen A (Bos taurus) (P00766)		SIGMA™	United states of America
SOLu-Trypsin (EMS0004)		SIGMA™	United states of America
Colloidal Coomassie (G-250)		SIGMA™	United states of America
Dithiothreitol	DTT	SIGMA™	United states of America
EDTA	Ethylenediamine tetra acetic acid	SIGMA™	United states of America
Formaldehyde		Merck	Germany
Formic Acid			
Lithium Acetate		SIGMA™	United states of America

Poly(ethylene glycol)	PEG P4338	SIGMA™	United states of America
Potassium Chloride	KCl	SIGMA™	United states of America
Magnesium Chloride Hexahydrate	MgCl2	SIGMA™	United states of America
Ampicillin	Amp	FORMEDIUM™	United Kingdom
Ethidium Bromide	EtBr	SIGMA™	United states of America
Ethanol		Ajax Scientific LTD	Canada
Methanol		Ajax Scientific LTD	Canada
Isopropyl β-D-1-thiogalactopyranoside	IPTG	SIGMA™	United States of America
Iodoacetamide	IAA	SIGMA™	United States of America
Yeast Extract		FORMEDIUM™	United Kingdom
Tris Base		FORMEDIUM™	United Kingdom
Tris(2-carboxyethyl)phosphine	TCEP	SIGMA™	United States of America
Tryptone		FORMEDIUM™	United Kingdom
Galactose	Gal	FORMEDIUM™	United Kingdom
Glucose	Glu	FORMEDIUM™	United Kingdom
Glycerol		Fisher Scientific	United States of America
Glutathione Agarose 4B		PROTINO®	Germany

Glutathione Reduced	(GSH)	SIGMA™	United States of America
Yeast Nitrogen Base	YNB	FORMEDIUM™	United Kingdom
Luria-Bertani	LB	Fisher Scientific	United States of America
Peptone		FORMEDIUM™	United Kingdom
Calcium Chloride	CaCl2	SIGMA™	United states of America
Isoleucine		FORMEDIUM™	United Kingdom
leucine		FORMEDIUM™	United Kingdom
Valine		FORMEDIUM™	United Kingdom
Tryptophan		FORMEDIUM™	United Kingdom
3-Amino-1, 2, 4-triaxole	3AT	FORMEDIUM™	United Kingdom
Sodium Chloride	$NaCl_2$	SIGMA™	United states of America
Sodium Hydroxide	NaOH	Ajax UNIVAR	Australia
Monosodium phosphate	NaH_2PO_4	SIGMA™	United States of America

S1: IMPACT and Yih1

A



B

	3AT concentration	Total growth score (out of 10)	Normalisation of 3AT plates with GST	Average normalised total score	Galactose growth score	Normalisation Gal	Normalised 3AT/Normalised Gal)	Average	SD	SE
GST alone		8	1.025641026	1	8.5	1	0.975	1.001	0.031843367	0.01007
	30mM	8	1.025641026		8.5	1	0.975			
		8	1.025641026		8.5	1	0.975			
	60mM	8	1.025641026		8.5	1	0.975			
		8	1.025641026		8.5	1	0.975			
	90mM	8	1.025641026		8.5	1	0.975			
		7.5	0.961538462		8.5	1	1.04			
	120mM	7.5	0.961538462		8.5	1	1.04			
		7.5	0.961538462		8.5	1	1.04			
	150mM	7.5	0.961538462		8.5	1	1.04			
IMPACT		3.5	0.448717949	0.358974359	8.5	1	0.448717949	0.358974	0.11176664	0.028858
		3.5	0.448717949		8.5	1	0.448717949			
	30mM	3.5	0.448717949		8.5	1	0.448717949			
		3.5	0.448717949		8.5	1	0.448717949			
	60mM	3.5	0.448717949		8.5	1	0.448717949			
		3.5	0.448717949		8.5	1	0.448717949			
	90mM	3.5	0.448717949		8.5	1	0.448717949			
		2	0.256410256		8.5	1	0.256410256			
	120mM	2	0.256410256		8.5	1	0.256410256			
		1.5	0.192307692		8.5	1	0.192307692			
	150mM	1.5	0.192307692		8.5	1	0.192307692			
Yih1		6	0.769230769	0.448717949	8.5	1	0.769230769	0.448718	0.206724558	0.065372
	30mM	6	0.769230769		8.5	1	0.769230769			
		4	0.512820513		8.5	1	0.512820513			
	60mM	4	0.512820513		8.5	1	0.512820513			
		4	0.512820513		8.5	1	0.512820513			
	90mM	4	0.512820513		8.5	1	0.512820513			
		2	0.256410256		8.5	1	0.256410256			
	120mM	2	0.256410256		8.5	1	0.256410256			
		1.5	0.192307692		8.5	1	0.192307692			
	150mM	1.5	0.192307692		8.5	1	0.192307692			
gcn1	30mM	0	0	0	8.5	1	0	0	0	0
	60mM	0	0		8.5	1	0			
	90mM	0	0		8.5	1	0			
	120mM	0	0		8.5	1	0			
	150mM	0	0		8.5	1	0			
			Average gst alone growth		Average galactose growth					
			7.8		8.5					

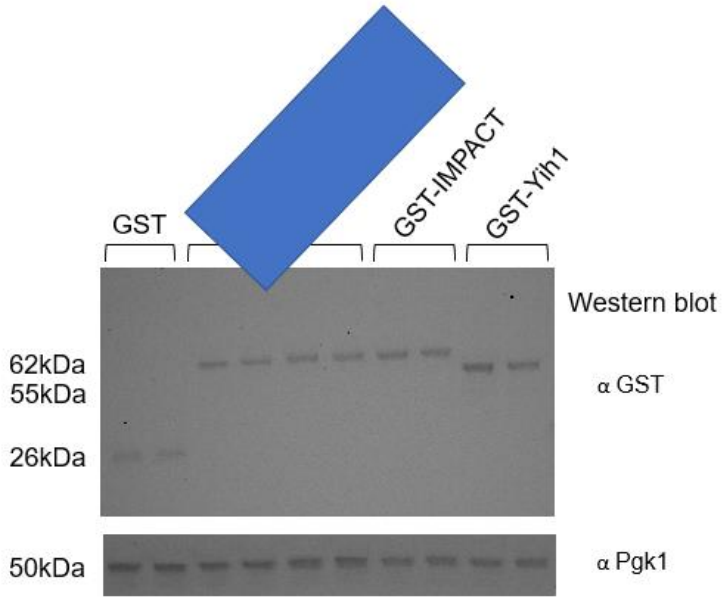
C

	P values			
	GST alone	GST-IMPACT	GST-Yih1	gcn1Δ
GST alone	0	2.43725E-13	6.06E-05	8.6E-15
GST-IMPACT	2.4373E-13	0	0.255199	9.18E-09
GST-Yih1	6.0636E-05	0.255199058	0	0.00011
gcn1Δ	8.595E-15	9.17967E-09	0.00011	0

Figure S1.1: SQGA of yeast cells overexpressing GST alone, GST-Yih1, GST-IMPACT or a gcn1Δ strain.

A: 5µL of ten-fold serially diluted overnight cultures of the yeast strains indicated above were transferred onto solid media with 3AT, or with 2% Glucose or 2% Galactose alone. A representative image showing yeast cell growth at all 3AT concentrations analysed (30mM, 60mM, 90mM, 120mM, 150mM) B: Quantification of SQGA in (A). SD= Standard Deviation, SE= Standard Error. C: Calculated T-test values from data quantified in (B). Values greater than 0.05 ($p \geq 0.05$) are not stastically significant, while values less or equal to 0.05 (≤ 0.05) are statistically significant. For samples with an unequal sample size: a T-test assuming unequal variance was carried out. For samples with equal sample size: a paired T-test two sample for means was carried out.

A



B

	Signal intensities						
	Pgk1	GST	Ratio	Normalisation to Yih1	Average	SD	SE
	3432338	4537674	1.322036	0.675041577	0.81098	0.096369	0.01
	2858908	5243157	1.833972	0.936440012			
	2767414	4629231	1.672764	0.854126111			
GST	2957240	4507677	1.524285	0.778311588			
	2935036	4036797	1.375382	0.702280739	0.725328	0.023047	0.01
IMPACT	2496962	3659688	1.465656	0.748375236			
	2501590	4639113	1.854466	0.946904321	1	0.053096	0.01
Yih1	2020330	4166802	2.062436	1.053095679			

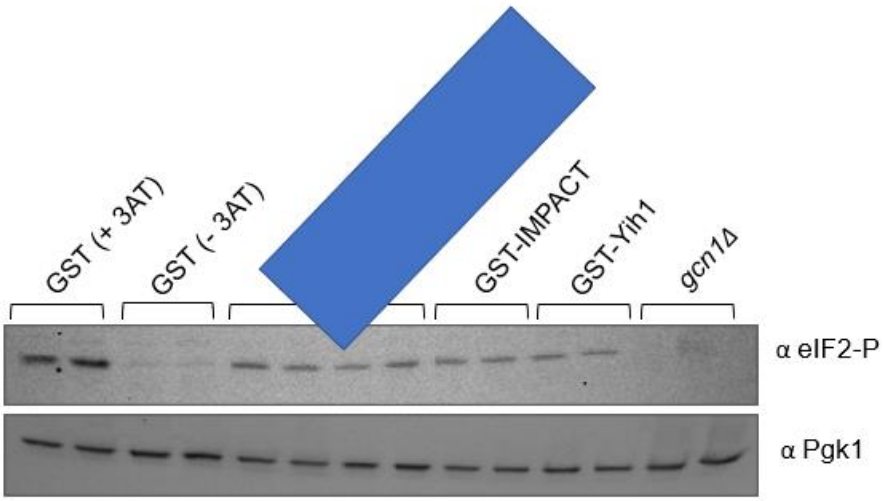
C

	GST	IMPACT	Yih1
GST	0	0.228027	0.09105
IMPACT	0.228027	0	0.069369
Yih1	0.09105	0.069369	0

Figure S1.2: Expression levels of yeast cells overexpressing the indicated GST-tagged proteins.

A: Whole images of western membrane depicted in Fig. 3.1.6, for analysis of expression levels of yeast cells (ESY11001B background) expressing GST alone, GST-Yih1, or GST-IMPACT. The same membrane and exposure times were used for quantification and also for construction of the image panels in Fig. 15. B: Quantification of band signal intensities from Fig. 15. Signal intensity was measured in ImageLab and statistical analysis was carried out in Microsoft Excel. C: Calculated T-test values from data quantified in (B). Values greater than 0.05 ($p \geq 0.05$) are not stastically significant, while values less or equal to 0.05 (≤ 0.05) are statistically significant. For samples with an unequal sample size: a T-test was carried out assuming unequal variance. For samples with equal sample size: a paired T-test two sample for means was carried out.

A



B

	Signal intensity						
	pgk1	eif2p	ratio	Normalised to IMPACT	Average	SD	SE
	571350	310336	0.54316268	1.544477123	1.56437417	0.01989705	0.01407
GST (+3AT)	824625	459446	0.5571575	1.584271215			
	823975	50908	0.06178343	0.175680498	0.1646761	0.0110044	0.00778
GST (-3AT)	738000	39884	0.05404336	0.153671701			
	397525	164502	0.41381548	1.176679778	1	0.17667978	0.12493
GST-IMPACT	463975	134342	0.28954577	0.823320222			
	558950	179712	0.32151713	0.914230426	0.82605591	0.08817452	0.06235
GST-Yih1	505575	131196	0.25949859	0.73788139			
	617275	6604	0.01069864	0.030421451	0.05040807	0.01998662	0.01413
gcn1Δ	680550	16848	0.02475645	0.070394685			

C

	GST (+3AT)	GST (-3AT)	GST-IMPACT	GST-Yih1	*gcn1Δ*
GST (+3AT)	0	0.01405251	0.213374522	0.092528	3.77E-05
GST (-3AT)	0.014052513	0	0.124647424	0.073947	0.168604
GST-IMPACT	0.213374522	0.12464742	0	0.299641	0.13001
GST-Yih1	0.092528321	0.07394674	0.299640529	0	0.088205
gcn1Δ	3.76645E-05	0.1686039	0.130009873	0.088205	0

Figure S1.3: eIF2 phosphorylation levels in yeast cells overexpressing the indicated proteins.

A: Lower expoure image used for quantification of eIF2 phosphorylation levels of yeast cells (ESY11001B background) expressing GST alone, GST-Yih1, or GST-IMPACT, under control of a galactose-inducible promoter, as described in Fig. 16. Proteins were detected using chemiluminescence and at least two samples were analysed. The same membrane and exposure times were used for quantification and for construction of the image panels in Fig. 16.

B: Quantification of data in (A). SD= Standard Deviation. SE= Standard Error

C: Calculated T-test values from data quantified in (B). Values greater than 0.05 ($p \geq 0.05$) are not stastically significant, while values less or equal to 0.05 (≤ 0.05) are statistically significant. For samples with an unequal sample size: a T-test was carried out assuming unequal variance. For samples with equal sample size: a paired T-test two sample for means was carried out.

S2: Group 1: Amino acids involved in Gcn1 binding

A

GST-Yih1^{EDDE} (A)

GAA (E87) to GCC (Ala)
GAC (D90) to GCC (Ala)
GAC (D102) to GCC (Ala)
GAA (E106) to GCC (Ala)

```
Query  60    GATGACGATCACGAACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGAT  119
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  221   GATGACGATCACGAACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGAT  280

Query  120   CTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAA  179
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  281   CTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAA  340

Query  180   TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  239
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  341   TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  400

Query  240   ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  299
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  401   ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  460

Query  300   CAGCATTTGTTCCAGGCCGTGATGGCCTCTGTTTTCCACCGCGGATCTGTCTGTCTATTT  359
             |||||||||||||||| |||||||| ||||||||||||||||||||||||||||||||||
Sbjct  461   CAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTT  520

Query  360   GCCTTCCTCACAGCCCTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  419
             | ||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  521   GACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  580
```

```
Query  420   GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  581   GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  640

Query  480   ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  641   ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  700

Query  540   TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATG  599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  701   TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATG  760

Query  600   AGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGAC  659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  761   AGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGAC  820

Query  660   GGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAAC  719
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  821   GGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAAC  880

Query  720   GTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAA  779
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  881   GTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAA  940

Query  780   CACATCAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCGTAATCATTA  839
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  941   CACATCAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCGTAATCATTA  1000
```

B

GST-Yih1^{EDDE} (B)

GAA (E87) to GCC (Ala)
GAC (D90) to GCC (Ala)
GAC (D102) to GCC (Ala)
GAA (E106) to GCC (Ala)

```
Query  63    GATGACGATCACGAACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGAT  122
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  221   GATGACGATCACGAACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGAT  280

Query  123   CTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAA  182
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  281   CTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAA  340

Query  183   TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  242
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  341   TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  400

Query  243   ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  302
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  401   ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  460

Query  303   CAGCATTTGTTCCAGGCCGTGATGGCCTCTGTTTTCCACCGCGGATCTGTCTGTCTATTT  362
             |||||||||||||||  |||||||  |||||||||||||||||||||||||||||||||a.)b.)
Sbjct  461   CAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTT  520

Query  363   GCCTTCCTCACAGCCCTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  422
             |  ||||||||||  |||||||||||||||||||||||||||||||||||||||||||c.)d.)
Sbjct  521   GACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  580

Query  423   GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  482
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  581   GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  640
```

```
Query  483   ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  542
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  641   ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  700

Query  543   TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATG  602
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  701   TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATG  760

Query  603   AGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGAC  662
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  761   AGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGAC  820

Query  663   GGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAAC  722
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  821   GGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAAC  880

Query  723   GTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAA  782
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  881   GTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAA  940

Query  783   CACATCAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCGTAATCATTA  842
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  941   CACATCAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCGTAATCATTA  1000

Query  843   TGATTATGTCAAGCACCTGTAACCTTGATCACA  875
             ||||||||||||||||||||||||||||||||
Sbjct  1001  TGATTATGTCAAGCACCTGTAACCTTGATCACA  1033
```

C

GST-Yih1^{D102A} (A)

CTA (L100) to CTC: Silent mutation

GAC (D102) to GCA (Alanine)

```
Query  62    GATGACGATCACGAACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGAT  121
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  221   GATGACGATCACGAACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGAT  280

Query  122   CTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAA  181
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  281   CTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAA  340

Query  182   TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  241
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  341   TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  400

Query  242   ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  301
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  401   ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  460

Query  302   CAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTCTTC  361
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  ||
Sbjct  461   CAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTT  520

Query  362   GCATTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  421
             |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  521   GACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  580

Query  422   GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  481
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  581   GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  640

Query  482   ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  541
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  641   ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  700
```

```
Query  542  TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATG  601
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  701  TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATG  760

Query  602  AGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGAC  661
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  761  AGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGAC  820

Query  662  GGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAAC  721
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  821  GGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAAC  880

Query  722  GTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAA  781
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  881  GTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAA  940

Query  782  CACATCAA  789
            ||||||||
Sbjct  941  CACATCAA  948
```

D

GST-Yih1^{D102A} (B)

CTA (L100) to CTC: Silent mutation
TTT (F101) to TCC: Silent mutation
GAC (D102) to GCA (Alanine)

```
Query  58   GATGACGATCACGAACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGAT  117
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  221  GATGACGATCACGAACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGAT  280

Query  118  CTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAA  177
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  281  CTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAA  340

Query  178  TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  237
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  341  TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  400

Query  238  ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  297
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  401  ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  460

Query  298  CAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTCTTC  357
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  a.)b.)
Sbjct  461  CAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTT  520

Query  358  GCATTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  417
            |  ||||||||||||||||||||||||||||||||||||||||||||||||||||||c.)
Sbjct  521  GACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  580

Query  418  GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  477
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  581  GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  640

Query  478  ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  537
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  641  ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  700

Query  538  TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATG  597
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  701  TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATG  760

Query  598  AGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGAC  657
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  761  AGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGAC  820

Query  658  GGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAAC  717
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  821  GGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAAC  880

Query  718  GTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAA  777
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  881  GTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAA  940

Query  778  CACATCAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  828
                ||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  941  CACATCAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  991
```

E

GST-Yih1^{E106A} (A)

GAA (E106) to GCT (Ala)

CTC (L107) to CTA: Silent mutation

```
Query  60   GATGACGATCACGAACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGAT  119
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  221  GATGACGATCACGAACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGAT  280

Query  120  CTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAA  179
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  281  CTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAA  340

Query  180  TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  239
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  341  TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  400

Query  240  ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  299
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  401  ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  460

Query  300  CAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTT  359
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  461  CAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTT  520

Query  360  GACTTCCTCACAGCTCTAGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  419
            |||||||||||||  ||| ||||||||||||||||||||||||||||||||||||||||a.) b.)
Sbjct  521  GACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  580

Query  420  GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  479
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  581  GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  640

Query  480  ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  539
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  641  ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  700

Query  540  TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATG  599
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  701  TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATG  760

Query  600  AGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGAC  659
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  761  AGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGAC  820

Query  660  GGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAAC  719
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  821  GGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAAC  880

Query  720  GTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAA  779
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  881  GTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAA  940

Query  780  CACATCAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  830
            ||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  941  CACATCAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  991
```

F

GST-Yih1^{E106A} (B)

GAA (E106) to GCT (Ala)

CTC (L107) to CTA: Silent mutation

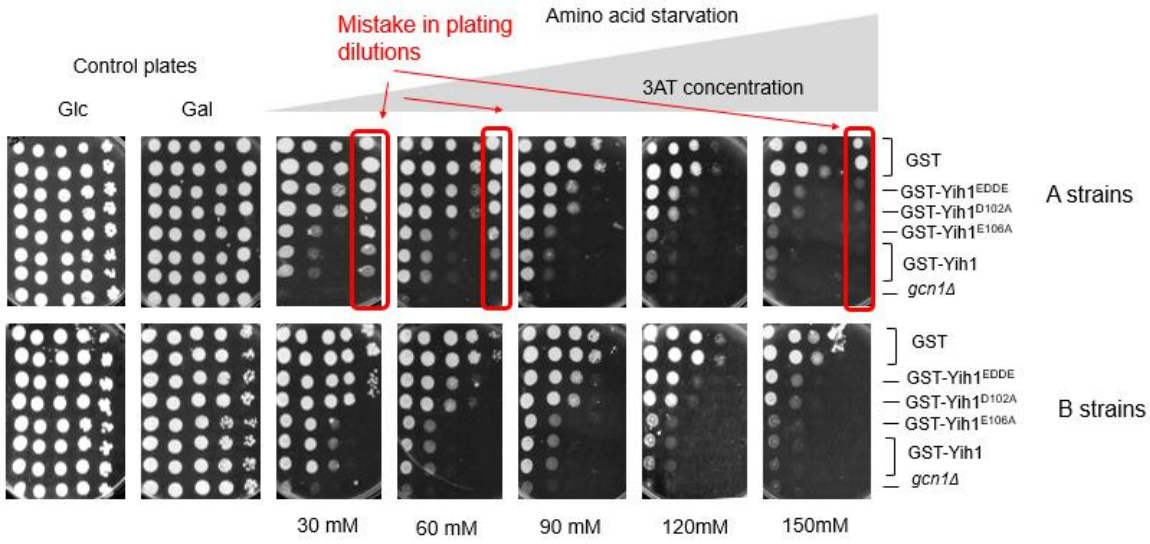
```
Query  60   GATGACGATCACGAACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGAT  119
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  221  GATGACGATCACGAACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGAT  280

Query  120  CTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAA  179
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  281  CTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAA  340

Query  180  TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  239
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  341  TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  400

Query  240  ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  299
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  401  ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  460

Query  300  CAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTT  359
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  461  CAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTT  520

Query  360  GACTTCCTCACAGCTCTAGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  419
            |||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||a.) b.)
Sbjct  521  GACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  580

Query  420  GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  479
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  581  GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  640

Query  480  ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  539
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  641  ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  700

Query  540  TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATG  599
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  701  TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATG  760

Query  600  AGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGAC  659
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  761  AGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGAC  820

Query  660  GGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAAC  719
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  821  GGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAAC  880

Query  720  GTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAA  779
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  881  GTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAA  940

Query  780  CACATCAATTCTACGGCAAGAGAAGCTGTTGTC  812
            |||||||||||||||||||||||||||||||||
Sbjct  941  CACATCAATTCTACGGCAAGAGAAGCTGTTGTC  973
```

Figure S2.1: Whole Nucleotide BLAST sequence alignment of Group 1 Yih1 mutants (Sequencing order H210324-035).

Plasmid DNA was isolated as described in Materials and Methods (Section 2.1.10 and Section 2.1.11) The samples indicated (from A to F) were commercially sequenced using primers 3035-B and ES48. Query: Wild type Yih1 sequence. Sbjct: Plasmid DNA samples

with the indicated substitutions. Magenta colour indicates mutations that are expected. Turquoise colour indicates unexpected silent mutations. Red colour indicates unexpected mutations.

A



B

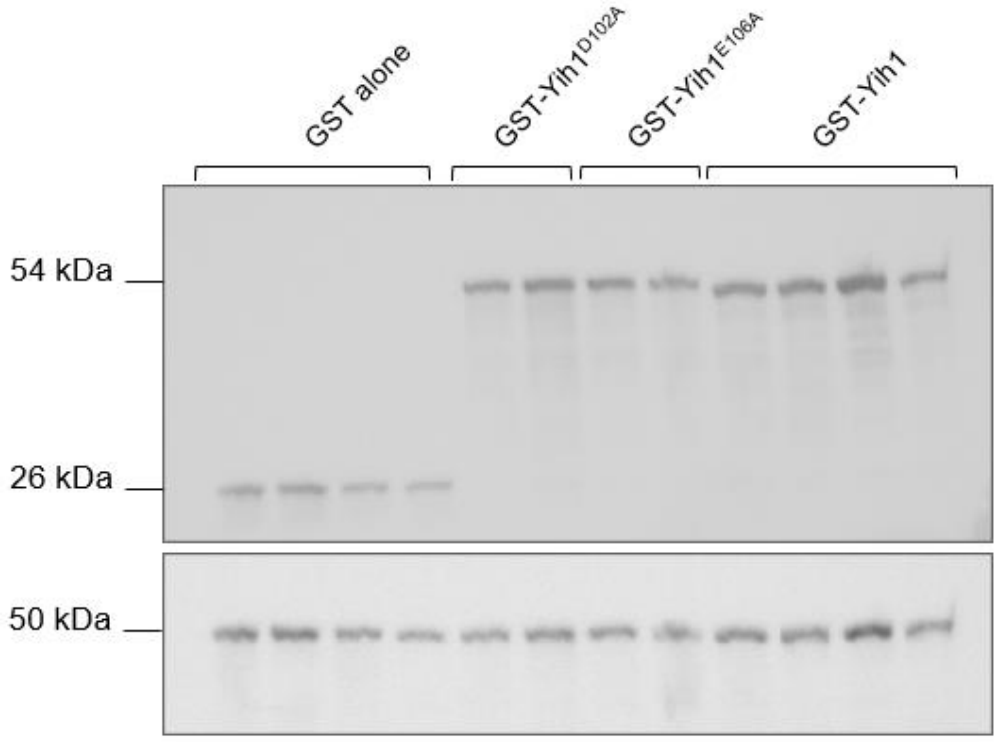
	3AT concentration	Total growth score (out of 10)	Normalisation of 3AT plates with GST	Galactose growth score	Normalisation Gal	Normalised 3AT/Normalised Gal)	Average	SD	SE
GST		9	1.304347826	9	1	1.304347826	1	0.206998	0.065458
	30mM	9	1.304347826	9	1	1.304347826			
		8	1.15942029	9	1	1.15942029			
	60mM	8	1.15942029	9	1	1.15942029			
		6.5	0.942028986	9	1	0.942028986			
	90mM	6.5	0.942028986	9	1	0.942028986			
		6	0.869565217	9	1	0.869565217			
	120mM	6	0.869565217	9	1	0.869565217			
		5	0.724637681	9	1	0.724637681			
	150mM	5	0.724637681	9	1	0.724637681			
GST-Yih1 (edde)	30mM	7.5	1.086956522	9	1	1.086956522	0.68115942	0.269581	0.12056
	60mM	5.5	0.797101449	9	1	0.797101449			
	90mM	5	0.724637681	9	1	0.724637681			
	120mM	3.5	0.507246377	9	1	0.507246377			
	150mM	2	0.289855072	9	1	0.289855072			
D102A	30mM	7.5	1.086956522	9	1	1.086956522	0.68115942	0.269581	0.12056
	60mM	5.5	0.797101449	9	1	0.797101449			
	90mM	5	0.724637681	9	1	0.724637681			
	120mM	3.5	0.507246377	9	1	0.507246377			
	150mM	2	0.289855072	9	1	0.289855072			
E106A	30mM	4.5	0.652173913	9	1	0.652173913	0.362318841	0.165243	0.073899
	60mM	3	0.434782609	9	1	0.434782609			
	90mM	2	0.289855072	9	1	0.289855072			
	120mM	1.5	0.217391304	9	1	0.217391304			
	150mM	1.5	0.217391304	9	1	0.217391304			
Yih1		4.5	0.652173913	9	1	0.652173913	0.362318841	0.165243	0.052254
	30mM	4.5	0.652173913	9	1	0.652173913			
		3	0.434782609	9	1	0.434782609			
	60mM	3	0.434782609	9	1	0.434782609			
		2	0.289855072	9	1	0.289855072			
	90mM	2	0.289855072	9	1	0.289855072			
		1.5	0.217391304	9	1	0.217391304			
	120mM	1.5	0.217391304	9	1	0.217391304			
		1.5	0.217391304	9	1	0.217391304			
	150mM	1.5	0.217391304	9	1	0.217391304			
gcn1Δ	30mM	0	0	9	1	0	0.014492754	0.028986	0.012963
	60mM	0	0	9	1	0			
	90mM	0.5	0.072463768	9	1	0.072463768			
	120mM	0	0	9	1	0			
	150mM	0	0	9	1	0			
			Average gst alone growth	Average galactose growth					
			6.9	9					

C

	GST	GST-Yih1 (EDDE)	GST-Yih1 (D102A)	GST-Yih1 (E106A)	GST-Yih1
GST	0	0.079856079	0.079856079	0.000222368	6.46792E-10
GST-Yih1 (EDDE)	0.079856	0	1	0.00900635	0.080128571
GST-Yih1 (D102A)	0.079856	1	0	0.00900635	0.080128571
GST-Yih1 (E106A)	0.000222	0.00900635	0.00900635	0	1
GST-Yih1	6.47E-10	0.080128571	0.080128571	1	0

Figure S2.2: SQGA of yeast cells overexpressing the indicated proteins.
A: 5µL of ten-fold serially diluted overnight cultures of the yeast strains indicated above were transferred onto solid media with 3AT, or with 2% Glucose or 2% Galactose alone. B: Quantification of SQGA in (A). SD= Standard Deviation, SE= Standard Error. C: Calculated T-test values from data quantified in (B). Values greater than 0.05 ($p \geq 0.05$) are not stastically significant, while values less or equal to 0.05 (≤ 0.05) are statistically significant. For samples with an unequal sample size: a T-test was carried out assuming unequal variance. For samples with equal sample size: a paired T-test two sample for means was carried out.

A



B

	Signal intensities						
	Pgk1	GST	Ratio (GST/Pgk1)	Normalize to Yih1	Average	SD	SE
	2898032	2804472	0.97	0.71	0.5	0.098335	0.040145
	4194174	2646039	0.63	0.46			
	3899636	2903703	0.74	0.55			
	4924496	3132294	0.64	0.47			
	6532238	3548754	0.54	0.40			
GST alone	5819510	3938814	0.68	0.50			
	1914668	3135825	1.64	1.20	1.0	0.242083	0.121042
	2246332	3908850	1.74	1.28			
	2205520	2211759	1.00	0.74			
GST-Yih1	2874244	3067746	1.07	0.78			
	2589738	2609145	1.01	0.74	0.8	0.017387	0.008693
	2442906	2610894	1.07	0.78			
	2319824	2382138	1.03	0.75			
GST-Yih1 (EDDE)	2500742	2538261	1.02	0.75			
	2704422	2790546	1.03	0.76	0.8	0.078378	0.039189
	3431324	3131667	0.91	0.67			
	2970004	3305511	1.11	0.82			
GST-Yih1 (D102A)	3025864	3639405	1.20	0.88			
	2042462	2198262	1.08	0.79	0.9	0.064305	0.032152
	2668702	2995608	1.12	0.82			
	2015254	2283633	1.13	0.83			
GST-Yih1 (E106A)	2461146	3216213	1.31	0.96			
	6838518	1715010	0.25	0.18	0.3	0.099208	0.07015
gcn1Δ	3355286	1748208	0.52	0.38			
				Yih1 (Average ratio)	1.36		

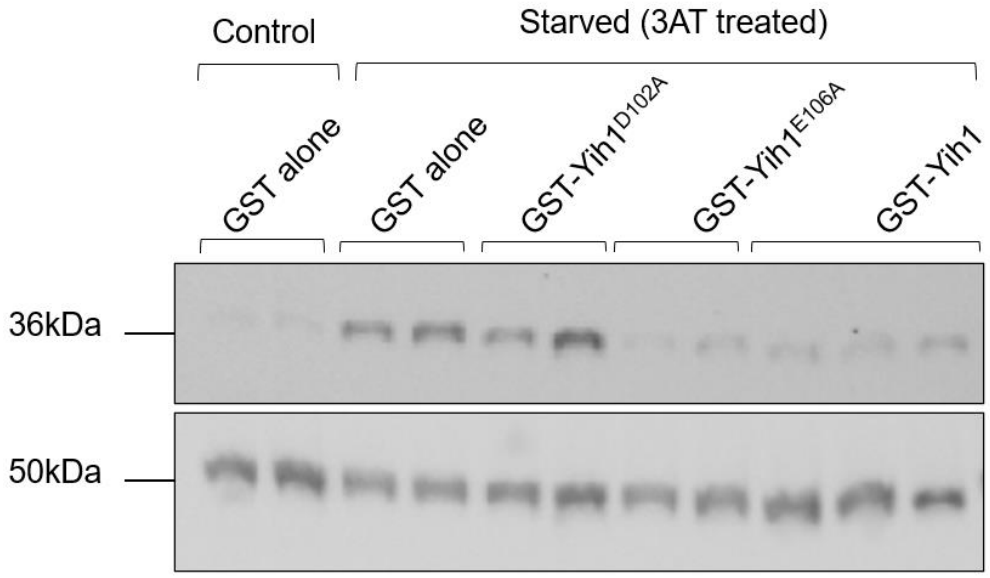
C

	P values from T test				
	GST	GST-Yih1 (EDDE)	GST-Yih1 (D102A)	GST-Yih1 (E106A)	GST-Yih1
GST	0	0.001719277	0.003798166	0.00037606	0.029439
GST-Yih1 (EDDE)	0.001719	0	0.656458127	0.098824846	0.169565
GST-Yih1 (D102A)	0.003798	0.656458127	0	0.110976357	0.314818
GST-Yih1 (E106A)	0.000376	0.098824846	0.110976357	0	0.434747
GST-Yih1	0.029439	0.169564632	0.314818242	0.434747272	0

Figure S2.3: Expression levels of the indicated proteins.

A: Lower exposure time image of the western blotting membrane visualised in Fig. 20, which was used for data quantification and statistical analysis. B: Quantification of band signal intensities from Fig. 20. Signal intensity was measured in ImageLab and statistical analysis was carried out in Microsoft Excel. C: Calculated T-test values from data quantified in (B). Values greater than 0.05 ($p \geq 0.05$) are not stastically significant, while values less or equal to 0.05 (≤ 0.05) are statistically significant. For samples with an unequal sample size: a T-test was carried out assuming unequal variance. For samples with equal sample size: a paired T-test two sample for means was carried out.

A



B

	Signal intensity							
	Pgk1	eIF2-P	eIF2-P w/o background	Ratio (eIF2-P:Pgk1)	Normalisation to GST alone (starved)	Average	SD	SE
GST alone (Unstarved)	6446434	723212		0.112187917	0.16403802	0.310392	0.08613	0.043065
	7182646	1659756		0.23107863	0.337876681			
	4789938	1259748		0.262998811	0.384549473			
	6118342	1485904		0.242860566	0.355103896			
GST alone (Starved)	4323070	3216108		0.743940764	1.087769285	1	0.128728	0.064364
	3979892	3143420		0.789825453	1.154860588			
	4043732	2289532		0.566192814	0.827871226			
	4334622	2755508		0.635697415	0.929498901			
GST-Yih1 (D102A)	4914616	3177776		0.646597008	0.945435979	0.94695	0.187722	0.093861
	4650440	3909836		0.840745392	1.229314292			
	4547536	2183776		0.480210822	0.702150773			
	5454406	3397968		0.622976727	0.910899069			
GST-Yih1 (E106A)	3730194	1211000		0.324647994	0.474691176	0.504423	0.057868	0.028934
	4569614	1810424		0.396187512	0.579294249			
	3959106	1451604		0.366649441	0.536104511			
	4669288	1365504		0.29244373	0.427603006			
GST-Yih1 (EDDE)	6713577	2923150	-365007	-0.054368483	-0.046362932	0.183561	0.226635	0.113317
	7499232	6528238	3240081	0.432055042	0.368436591			
	7960407	2950316	-337841	-0.042440167	-0.036191015			
	7698912	7336078	4047921	0.525778318	0.448359472			
GST-Yih1 (wild type)	5523262	2109688		0.381964136	0.558497229	0.484092		
	5333376	1996316		0.374306256	0.547300091			
	6113820	3120208		0.510353265	0.746224207			
	3854454	1382304		0.358625113	0.524371565			
	6125562	1385804		0.226232956	0.330791473			
	6243324	1625456		0.260351057	0.380678001			
	4959380	1791160		0.361166114	0.528086945			
	4075462	715736		0.175620825	0.256787837			

C

	GST alone (Unstarved)	GST alone (Starved)	GST-Yih1 (D102A)	GST-Yih1 (E106A)	GST-Yih1 (wild type)	GST-Yih1 (EDDE)
GST alone (Unstarved)	0	0.008174273	0.001764222	0.017713514	0.043356741	0.360566705
GST alone (Starved)	0.008174273	0	0.700441696	0.000898293	0.00139365	0.008941349
GST-Yih1 (D102A)	0.001764222	0.700441696	0	0.007966762	0.012470306	0.006600075
GST-Yih1 (E106A)	0.017713514	0.000898293	0.007966762	0	0.757672983	0.104304479
GST-Yih1 (wild type)	0.043356741	0.00139365	0.012470306	0.757672983	0	0.101457434
GST-Yih1 (EDDE)	0.360566705	0.008941349	0.006600075	0.104304479	0.101457434	0

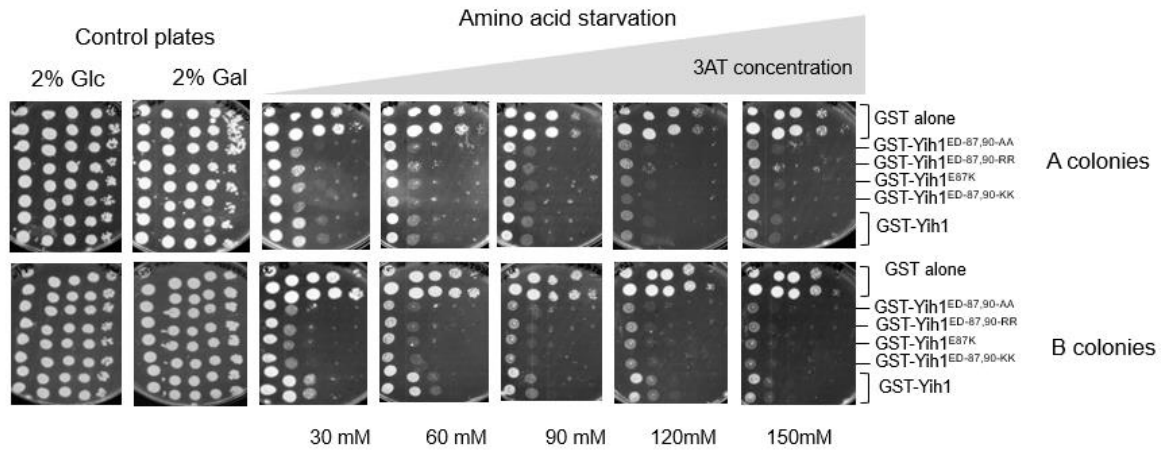
Figure S2.4: eIF2 phosphorylation levels in yeast cells overexpressing the indicated proteins.

A: A lower expoure image used for quantification of eIF2 phosphorylation levels of yeast cells (ESY11001B background) expressing the indicated proteins. Proteins were detected using chemiluminescence and at least two samples were analysed. B: Quantification of data

in (A). C: Calculated P values from T test using data values quantified in (B). Values greater than 0.05 ($p \geq 0.05$) are not stastically significant, while values less or equal to 0.05 (≤ 0.05) are statistically significant.

S3: Group 2: RWD domain amino acids involved in interdomain interactions

A



B

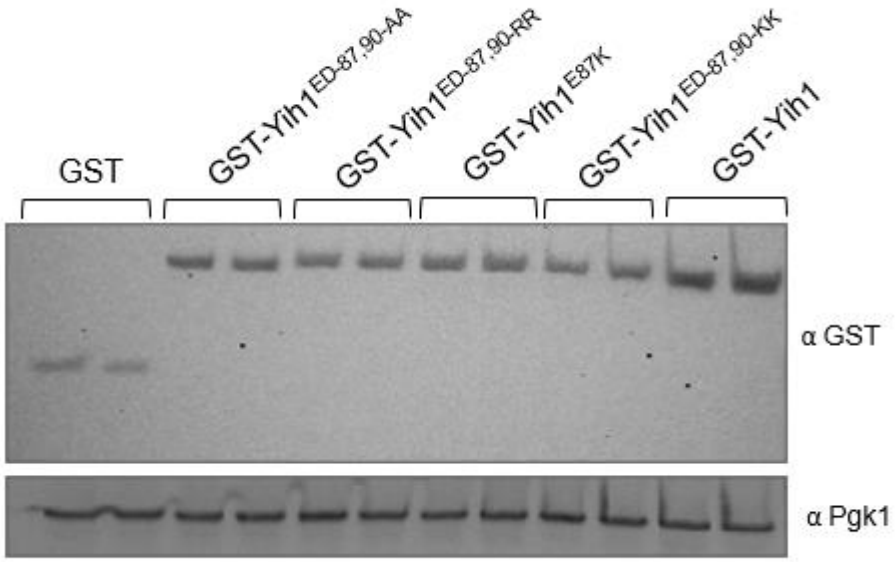
t test						
	GST	Yih1 (ED-87,90-AA)	Yih1 (ED-87,90-RR)	Yih1 (E87K)	Yih1 (ED-87,90-KK)	GST-Yih1
GST	0	8.33648E-06	8.33648E-06	8.33648E-06	8.33648E-06	8.66566E-07
Yih1 (ED-87,90-AA)	8.34E-06	0	0	0	0	0.005485142
Yih1 (ED-87,90-RR)	8.34E-06	0	0	0	0	0.005485142
Yih1 (E87K)	8.34E-06	0	0	0		0.005485142
Yih1 (ED-87,90-KK)	8.34E-06	0	0	0	0	0.005485142
GST-Yih1	8.67E-07	0.005485142	0.005485142	0.005485142	0.005485142	0

C

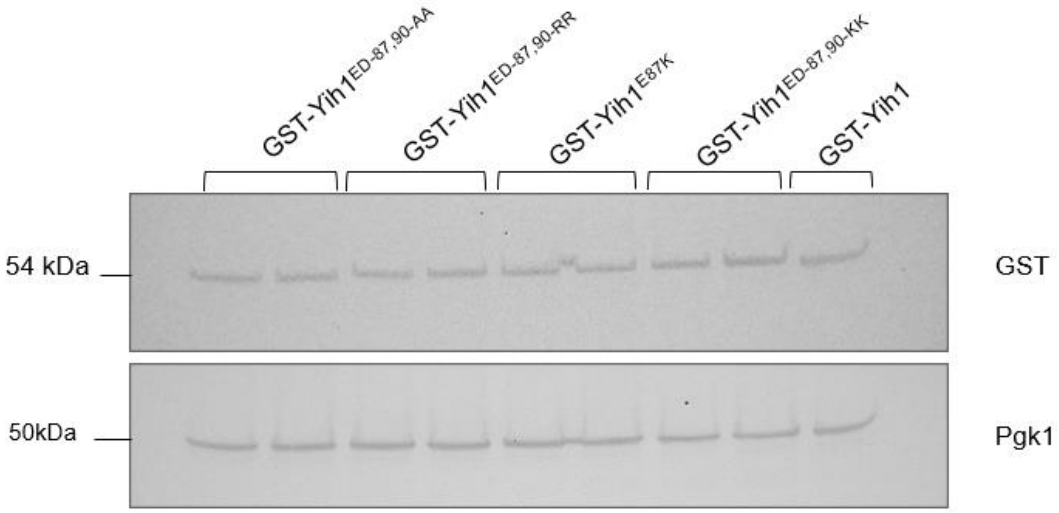
	3AT concentration	Total Growth Score (out of 10)	Normalisation of 3AT plates with GST	Average	Galactose Growth Score	Normalisation Galactose	Normalised 3AT/Normalised Gal	Average	SD	SE
		8	1.01910828	1	9.5	1	1.01910828	1	0.049747	0.015731
	30mM	8.5	1.082802548		9.5	1	1.082802548			
		8	1.01910828		9.5	1	1.01910828			
	60mM	8.5	1.082802548		9.5	1	1.082802548			
		7.5	0.955414013		9.5	1	0.955414013			
	90mM	8	1.01910828		9.5	1	1.01910828			
		7.5	0.955414013		9.5	1	0.955414013			
	120mM	7.5	0.955414013		9.5	1	0.955414013			
		7.5	0.955414013		9.5	1	0.955414013			
GST	150mM	7.5	0.955414013		9.5	1	0.955414013			
	30mM	2.5	0.318471338	0.191083	9.5	1	0.318471338	0.191083	0.080568	0.036031
	60mM	1.5	0.191082803		9.5	1	0.191082803			
	90mM	1.5	0.191082803		9.5	1	0.191082803			
	120mM	1.5	0.191082803		9.5	1	0.191082803			
Yih1 (ED-87,90-AA)	150mM	0.5	0.063694268		9.5	1	0.063694268			
	30mM	2.5	0.318471338	0.191083	9.5	1	0.318471338	0.191083	0.080568	0.036031
	60mM	1.5	0.191082803		9.5	1	0.191082803			
	90mM	1.5	0.191082803		9.5	1	0.191082803			
	120mM	1.5	0.191082803		9.5	1	0.191082803			
Yih1 (ED-87,90-RR)	150mM	0.5	0.063694268		9.5	1	0.063694268			
	30mM	2.5	0.318471338	0.191083	9.5	1	0.318471338	0.191083	0.080568	0.036031
	60mM	1.5	0.191082803		9.5	1	0.191082803			
	90mM	1.5	0.191082803		9.5	1	0.191082803			
	120mM	1.5	0.191082803		9.5	1	0.191082803			
Yih1 (E87K)	150mM	0.5	0.063694268		9.5	1	0.063694268			
	30mM	2.5	0.318471338	0.191083	9.5	1	0.318471338	0.191083	0.080568	0.036031
	60mM	1.5	0.191082803		9.5	1	0.191082803			
	90mM	1.5	0.191082803		9.5	1	0.191082803			
	120mM	1.5	0.191082803		9.5	1	0.191082803			
Yih1 (ED-87,90-KK)	150mM	0.5	0.063694268		9.5	1	0.063694268			
		5.5	0.700636943	0.433121	9.5	1	0.700636943	0.433121	0.181947	0.057537
	30mM	5.5	0.700636943		9.5	1	0.700636943			
		4.5	0.573248408		9.5	1	0.573248408			
	60mM	4.5	0.573248408		9.5	1	0.573248408			
		3	0.382165605		9.5	1	0.382165605			
	90mM	3	0.382165605		9.5	1	0.382165605			
		2.5	0.318471338		9.5	1	0.318471338			
	120mM	2.5	0.318471338		9.5	1	0.318471338			
		1.5	0.191082803		9.5	1	0.191082803			
Yih1	150mM	1.5	0.191082803		9.5	1	0.191082803			
		Average Growth Score (GST)								
		7.85								

Figure S3.1: SQGA of yeast cells overexpressing the indicated proteins.
A: 5µL of ten-fold serially diluted overnight cultures of the yeast strains indicated above were transferred onto solid media with 3AT, or with 2% Glucose or 2% Galactose alone. B: Quantification of SQGA in (A). SD= Standard Deviation, SE= Standard Error. C: Calculated P values from T test using data values quantified in (B). Values greater than 0.05 ($p \geq 0.05$) are not stastically significant, while values less or equal to 0.05 (≤ 0.05) are statistically significant. For samples with an unequal sample size: a T-test was carried out assuming unequal variance. For samples with equal sample size: a paired T-test two sample for means was carried out.

A


GST
GST-Yih1ED-87,90-AA
GST-Yih1ED-87,90-RR
GST-Yih1E87K
GST-Yih1ED-87,90-KK
GST-Yih1
α GST
α Pgk1

B


GST-Yih1ED-87,90-AA
GST-Yih1ED-87,90-RR
GST-Yih1E87K
GST-Yih1ED-87,90-KK
GST-Yih1
54 kDa
50kDa
GST
Pgk1

C

	Band Signal Intensity						
	Pgk1	GST	Ratio (GST:Pgk1)	Normalise to Yih1	Average	SD	SE
	402752	204792	0.508481646	0.213208234	0.212430756	0.034734	0.020054
	460665	186228	0.40425906	0.169507318			
GST	339196	205940	0.607141594	0.254576715			
	296244	323764	1.092896396	0.458255499	0.55027176	0.126072	0.063036
	355966	339920	0.954922661	0.400402602			
	431850	64080	0.148384856	0.712410649			
Yih1 (ED-87,90-AA)	462264	60660	0.131223716	0.630018288			
	394446	365064	0.925510716	0.388070065	0.521916541	0.161756	0.080878
	369824	344596	0.931783767	0.390700378			
	499972	54495	0.108996104	0.52330128			
Yih1 (ED-87,90-RR)	429572	70290	0.163627983	0.78559444			
	302484	348740	1.152920485	0.483423822	0.556542339	0.072821	0.036411
	337480	400176	1.185776935	0.497200652			
	508464	61480	0.120913182	0.580516372			
Yih1 (E87K)	487872	67578	0.13851584	0.66502851			
	363454	350616	0.964677786	0.404492962	0.647849979	0.189868	0.094934
	315250	396312	1.257135607	0.527121607			
	482504	78840	0.163397609	0.784488391			
Yih1 (ED-87,90-KK)	368764	67230	0.182311722	0.875296955			
	268918	646268	2.40321585	1.00767729	1	0.006268	0.003619
	263614	623868	2.366596615	0.99232271			
Yih1 (wild type)	348920	72675	0.208285567	1			
			Average Yih1 Ratio				
			2.384906232				

D

	GST	GST-Yih1 (ED-87,90-AA)	GST-Yih1 (ED-87,90-RR)	GST-Yih1 (E87K)	GST-Yih1 (ED-87,90-KK)	GST-Yih1 (wild type)
GST		0.011710749	0.049146379	0.000877298	0.030405516	0.000822873
GST-Yih1 (ED-87,90-AA)	0.011711	0	0.719109166	0.905707772	0.21392982	0.008581256
GST-Yih1 (ED-87,90-RR)	0.049146	0.719109166	0	0.558733808	0.091690604	0.014472588
GST-Yih1 (E87K)	0.000877	0.905707772	0.558733808	0	0.285772482	0.001850027
GST-Yih1 (ED-87,90-KK)	0.030406	0.21392982	0.091690604	0.285772482	0	0.048962962
GST-Yih1 (wild type)	0.000823	0.008581256	0.014472588	0.001850027	0.048962962	0

Figure S3.2: Expression levels of yeast cells overexpressing the indicated GST-tagged proteins.

A, B: Lower exposure images of the western blot membrane pictured in Fig. 25 that were used for quantification. C: Quantification of band signal intensities in (A,B) and statistical analysis. D:Calculated P values from T test using data values quantified in (B). Values greater than 0.05 ($p \geq 0.05$) are not stastically significant, while values less or equal to 0.05 (≤ 0.05) are statistically significant. For samples with an unequal sample size: a T-test was carried out assuming unequal variance. For samples with equal sample size: a paired T-test two sample for means was carried out.

A

GST-Yih1$^{ED\text{-}87,90\text{-}AA}$ (A)

GAA (E87) to GCC (Ala)
GAC (D90) to GCC (Ala)
GAC (D102) to GCC (Ala)
GAA (E106) to GCC (Ala)

```
Query  1    TTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGATCTTCTCTCCAAGAAGCAG  60
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  19   TTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGATCTTCTCTCCAAGAAGCAG  78

Query  61   GAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAATACATGACACTGCAGATC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  79   GAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAATACATGACACTGCAGATC  138

Query  121  TCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTGTCTGC  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  139  TCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTGTCTGC  198

Query  181  ACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCCAGGCC  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  199  ACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCCAGGAA  258

Query  241  GTGATGGCCTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGCCTTCCTCACAGCCCTC  300
            ||||||| ||||||||||||||||||||||||||||||||||| |||||||||| ||| |
Sbjct  259  GTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAGAACTC  318

Query  301  GACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTGACATT  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  319  GACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTGACATT  378

Query  361  CCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCTCGACT  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  379  CCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCTCGACT  438

Query  421  TTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAGACCTA  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  439  TTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAGACCTA  498

Query  481  CTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATGAGTGCATGGCGAATCAAG  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  499  CTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATGAGTGCATGGCGAATCAAG  558

Query  541  CAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCCGCCGGC  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  559  CAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCCGCCGGC  618

Query  601  TCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTGGTGGCC  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  619  TCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTGGTGGCC  678

Query  661  CGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCTACGGCA  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  679  CGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCTACGGCA  738

Query  721  AGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  753
            |||||||||||||||||||||||||||||||
Sbjct  739  AGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  771
```

B

GST-Yih1$^{ED\text{-}87,90\text{-}AA}$ (B)

GAA (E87) to GCC (Ala)

GAC (D90) to GCC (Ala)

```
Query  18   GATGACGATCACGAACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGAT  77
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    GATGACGATCACGAACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGAT  60

Query  78   CTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAA  137
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61   CTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAA  120

Query  138  TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  197
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121  TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  180

Query  198  ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  257
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181  ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  240

Query  258  CAGCATTTGTTCCAGGCCGTGATGGCCTCTGTTTTCCACCGCGGATCTGTCTGTCTATTT  317
```

```
Sbjct  241  CAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTT  300

Query  318  GACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  377

Sbjct  301  GACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  360

Query  378  GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  437

Sbjct  361  GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  420

Query  438  ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  497

Sbjct  421  ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  480

Query  498  TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATG  557

Sbjct  481  TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATG  540

Query  558  AGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGAC  617

Sbjct  541  AGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGAC  600

Query  618  GGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAAC  677

Sbjct  601  GGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAAC  660

Query  678  GTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGT  710

Sbjct  661  GTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGT  693
```

C

GST-Yih1$^{ED\text{-}87,90\text{-}RR}$ (A)

GAA (E87) to CGT (Arg)

GAC (D90) to AGA (Arg)

```
Query  1    CGAGGCCATCTATCCGGATCTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGT  60

Sbjct  42   CGAGGCCATCTATCCGGATCTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGT  101

Query  61   GAAAGTGCCGCAGCATGAATACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTC  120

Sbjct  102  GAAAGTGCCGCAGCATGAATACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTC  161

Query  121  CGAGGAGGCTCCTAATGTCATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCT  180

Sbjct  162  CGAGGAGGCTCCTAATGTCATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCT  221

Query  181  CTACGATACCAAGTACCTTCAGCATTTGTTCCAGCGTGTGATGAGA-TCTGTTTTCCACC  239

Sbjct  222  CTACGATACCAAGTACCTTCAGCATTTGTTCCAGGAAGTGATG-GAC-TCTGTTTTCCACC  280

Query  240  GCGGATCTGTCTGTCTATTTGACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAAC  299

Sbjct  281  GCGGATCTGTCTGTCTATTTGACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAAC  340

Query  300  CAGAGGAGGAGACAGAACCGGTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCT  359

Sbjct  341  CAGAGGAGGAGACAGAACCGGTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCT  400

Query  360  GGACCGCGTCGGACCCCATTACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATG  419

Sbjct  401  GGACCGCGTCGGACCCCATTACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATG  460

Query  420  TTACCTCCGAGGAACAAGCGTTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGC  479

Sbjct  461  TTACCTCCGAGGAACAAGCGTTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGC  520

Query  480  GTAAGGCAAACCATGTCATGAGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACAT  539

Sbjct  521  GTAAGGCAAACCATGTCATGAGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACAT  580

Query  540  ATCAAGATTCCGATGATGACGGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCA  599
```

```
Sbjct  581  ATCAAGATTCCGATGATGACGGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCA  640

Query  600  CCATCATGGATGTGTGGAACGTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACA  659
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  641  CCATCATGGATGTGTGGAACGTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACA  700

Query  660  TAGGTCCCGACCGGTTTAAACACATCAATTCTACGGCAAG  699
            |||||||||||||||||||||||||||||||||||||||
Sbjct  701  TAGGTCCCGACCGGTTTAAACACATCAATTCTACGGCAAG  740
```

D

GST-Yih1$^{ED\text{-}87,90\text{-}RR}$ (B)

GAA (E87) to CGT (Arg)

GAC (D90) to AGA (Arg)

```
Query  1    ACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGATCTTCTCTCCAAGAA  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  15   ACAGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGATCTTCTCTCCAAGAA  74

Query  61   GCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAATACATGACACTGCA  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  75   GCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAATACATGACACTGCA  134

Query  121  GATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTGT  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  135  GATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTGT  194

Query  181  CTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCCA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  195  CTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCCA  254

Query  241  GCGTGTGATGAGA-TCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAG  299
            | |||||||| || ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  255  GGAAGTGATG-GAC TCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAG  313

Query  300  AACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTG  359
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  314  AACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTG  373

Query  360  ACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCT  419
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  374  ACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCT  433

Query  420  CGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAG  479
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  434  CGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAG  493

Query  480  ACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATGAGTGCATGGCGAA  539
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  494  ACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATGAGTGCATGGCGAA  553

Query  540  TCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCCG  599
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  554  TCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCCG  613

Query  600  CCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTGG  659
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  614  CCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTGG  673

Query  660  TGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCTA  719
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  674  TGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCTA  733

Query  720  CGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  757
            |||||||||||||||||||||||||||||||||||||||
Sbjct  734  CGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  771
```

E

GST-Yih1^{E87K} (A)

CTT (L81) to CTG: Silent mutation

GAA (E87) to AAA (Lys)

```
Query  1    CGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGATCTTCTCTCCAAGAAGCAGGAAGA  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  24   CGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGATCTTCTCTCCAAGAAGCAGGAAGA  83

Query  61   CGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAATACATGACACTGCAGATCTCCTT  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  84   CGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAATACATGACACTGCAGATCTCCTT  143

Query  121  CCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTGTCTGCACTTC  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  144  CCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTGTCTGCACTTC  203

Query  181  TTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTGCAGCATTTGTTCCAGAAAGTGAT  240
            |||||||||||||||||||||||||||||||||| ||||||||||||||| ||||||||
Sbjct  204  TTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCCAGGAAGTGAT  263

Query  241  GGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAGAACTCGACGG  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  264  GGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAGAACTCGACGG  323

Query  301  TGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTGACATTCCCAC  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  324  TGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTGACATTCCCAC  383

Query  361  AGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCTCGACTTTCAT  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  384  AGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCTCGACTTTCAT  443

Query  421  GGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAGACCTACTGAA  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  444  GGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAGACCTACTGAA  503

Query  481  GACCGACTCCAAGATGCGTAAGGCAAACCATGTCATGAGTGCATGGCGAATCAAGCAGGA  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  504  GACCGACTCCAAGATGCGTAAGGCAAACCATGTCATGAGTGCATGGCGAATCAAGCAGGA  563

Query  541  TGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCCGCCGGCTCCAG  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  564  TGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCCGCCGGCTCCAG  623

Query  601  AATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTGGTGGCCCGTTG  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  624  AATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTGGTGGCCCGTTG  683

Query  661  GTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCTACGGCAAGAGA  720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  684  GTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCTACGGCAAGAGA  743

Query  721  AGCTGTTGTCAGGGCCGGCTTCGACTCG  748
            |||||||||||||||||||||||||||
Sbjct  744  AGCTGTTGTCAGGGCCGGCTTCGACTCG  771
```

F

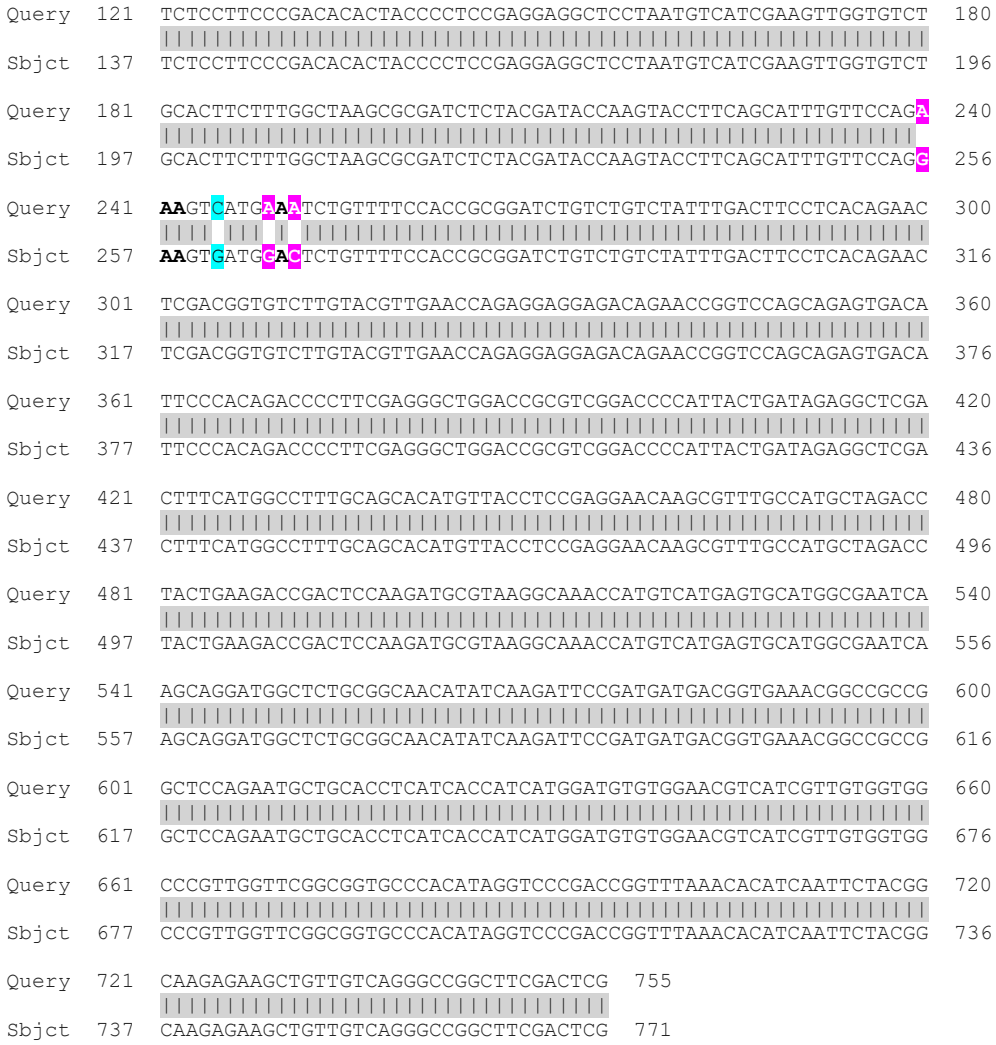
GST-Yih1$^{ED-87,90-KK}$ (A)

GAA (E87) to AAA (Lys)

GTG (V88) to GTC: Silent mutation

GAC (D90) to AAA (Lys)

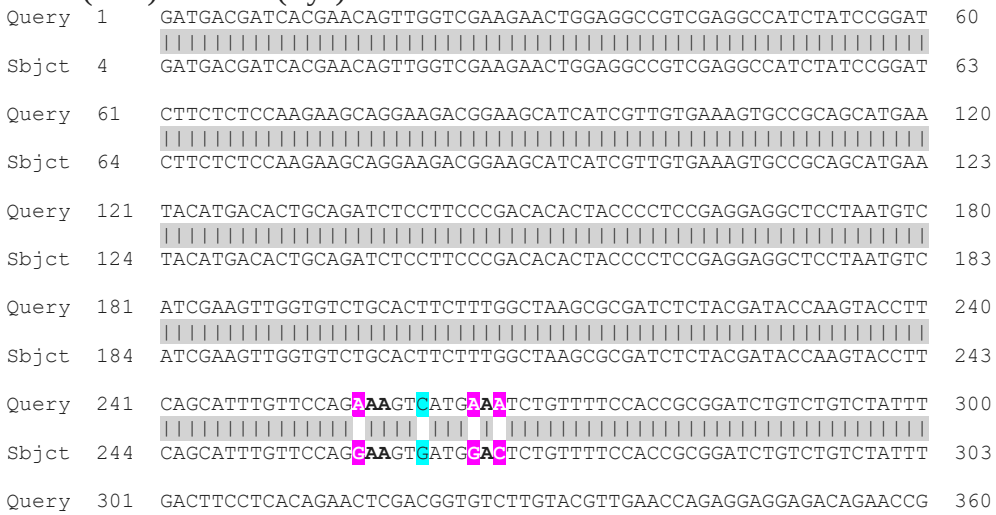
```
Query  1    AGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGATCTTCTCTCCAAGAAGC  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17   AGTTGGTCGAAGAACTGGAGGCCGTCGAGGCCATCTATCCGGATCTTCTCTCCAAGAAGC  76

Query  61   AGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAATACATGACACTGCAGA  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  77   AGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCATGAATACATGACACTGCAGA  136
```



G

GST-Yih1$^{ED\text{-}87,90\text{-}KK}$ (B)

GAA (E87) to AAA (Lys)
GTG (V88) to GTC: Silent mutation
GAC (D90) to AAA (Lys)



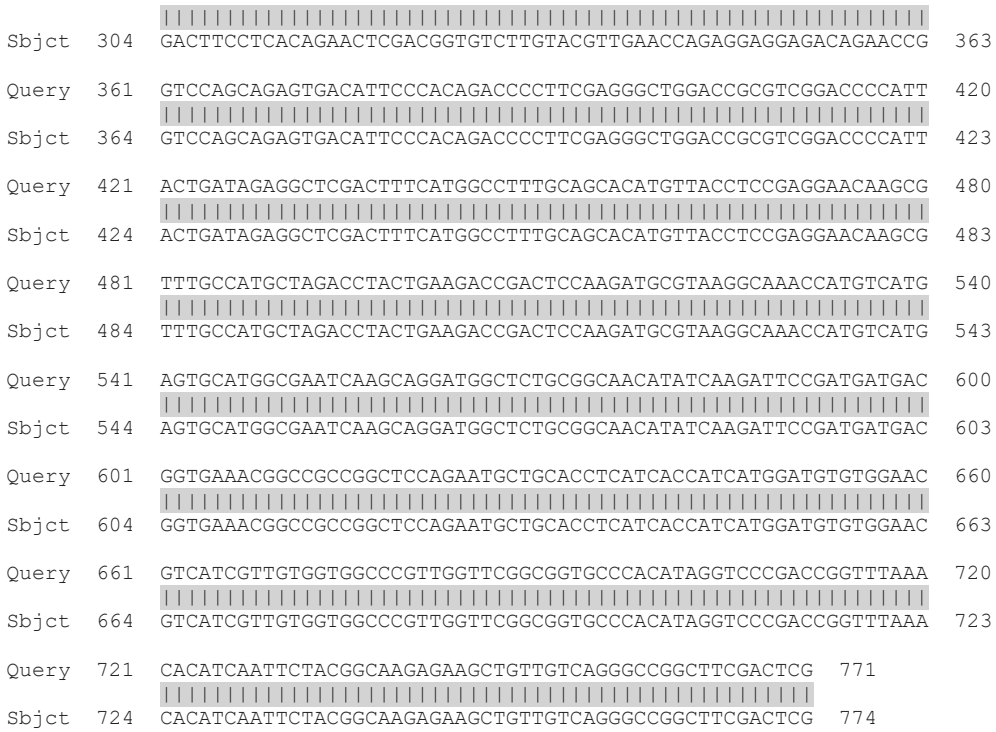


Figure S3.3: Whole Nucleotide BLAST sequence alignment of Group 2 Yih1 mutants
Plasmid DNA was isolated as described in Materials and Methods (Section 2.1.10 and Section 2.1.11). The samples indicated (from A to G) were commercially sequenced using primers 3035-B and ES48. Query: Wild type Yih1 sequence. Sbjct: Plasmid DNA samples with the indicated substitutions. Magenta colour indicates mutations that are expected. Turquoise colour indicates unexpected silent mutations. Red colour indicates unexpected mutations. Sequencing order HC00545595.

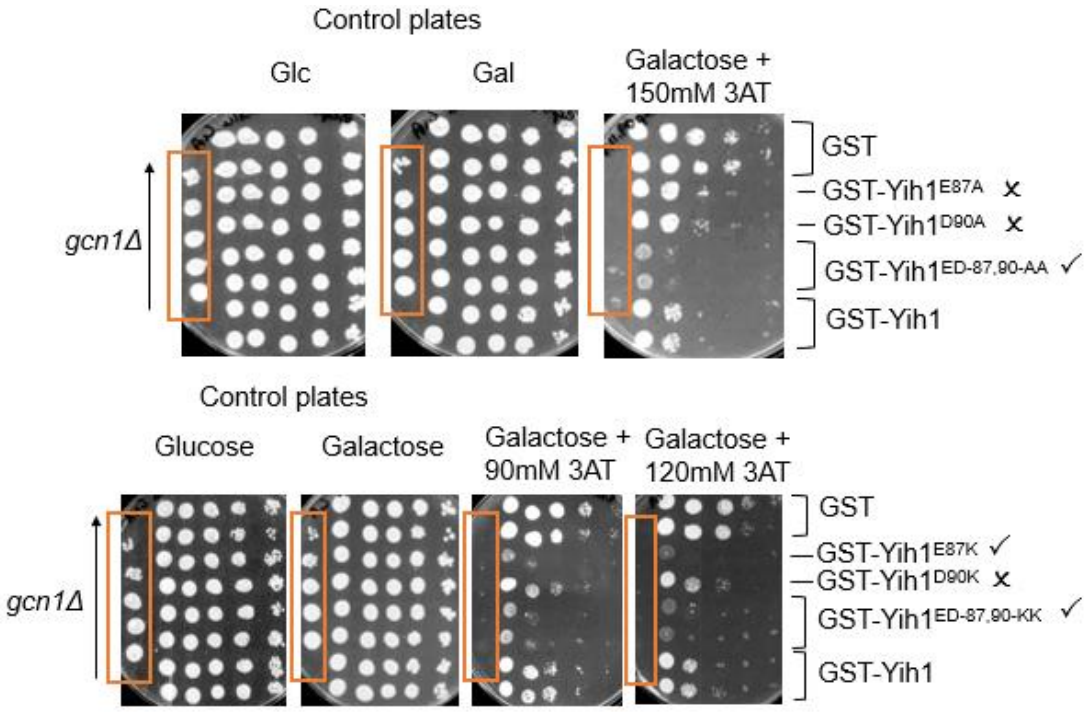
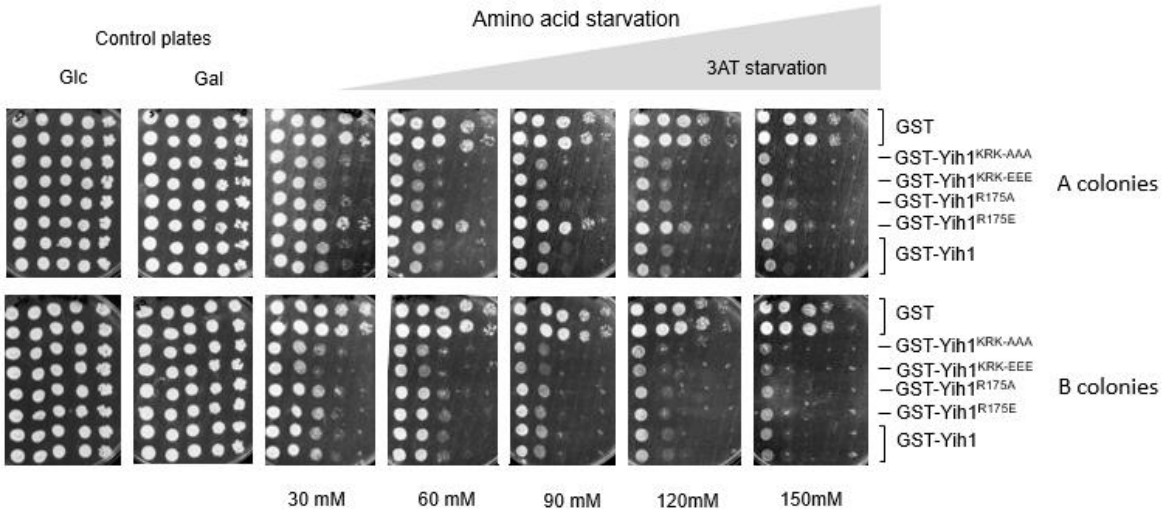


Figure S3.4 SQGA of yeast cells overexpressing the indicated proteins.
5µL of ten-fold serially diluted overnight cultures of the yeast strains indicated above were transferred onto solid media with 3AT, or with 2% Glucose or 2% Galactose alone. GST-Yih1^{E87A}, GST-Yih1^{D90A}, and GST-Yih1^{D90K} have incorrect sequences (the indicated mutations were not inserted). ✓= sequence confirmed. × = sequence incorrect. *gcn1Δ*= a gcn1 deletion strain.

S4: Group 3: Ancient domain amino acids involved in interdomain interactions

A



B

	3AT concentration	Growth score	Normalisation 3AT with GST	Average	Galactose growth score	Normalisation to Galactos	Normalised 3AT/Normalised Gal	Average		SD	SE
		8.5	1.08974359	1	9.5	1.00330033	1.086158907	0.999969		0.1255	0.028063
		8.5	1.08974359		9.5	1.00330033	1.086158907				
		9	1.153846154		9.4375	0.99669967	1.157666836				
	30mM	9	1.153846154		9.4375	0.99669967	1.157666836				
		8.5	1.08974359		9.5	1.00330033	1.086158907				
		8.5	1.08974359		9.5	1.00330033	1.086158907				
		8	1.025641026		9.4375	0.99669967	1.029037188				
	60mM	8	1.025641026		9.4375	0.99669967	1.029037188				
		8.5	1.08974359		9.5	1.00330033	1.086158907				
		8.5	1.08974359		9.5	1.00330033	1.086158907				
		8	1.025641026		9.4375	0.99669967	1.029037188				
	90mM	8	1.025641026		9.4375	0.99669967	1.029037188				
		7	0.897435897		9.5	1.00330033	0.894483806				
		7	0.897435897		9.5	1.00330033	0.894483806				
		8	1.025641026		9.4375	0.99669967	1.029037188				
	120mM	8	1.025641026		9.4375	0.99669967	1.029037188				
		7	0.897435897		9.5	1.00330033	0.894483806				
		7	0.897435897		9.5	1.00330033	0.894483806				
		5.5	0.705128205		9.4375	0.99669967	0.707463067				
GST	150mM	5.5	0.705128205		9.4375	0.99669967	0.707463067				
		5	0.641025641	0.365385	9.5	1.00330033	0.638917004	0.365325		0.183534	0.058038
	30mM	5	0.641025641		9.4375	0.99669967	0.643148242				
		4.5	0.576923077		9.5	1.00330033	0.575025304				
	60mM	3.5	0.448717949		9.4375	0.99669967	0.45020377				
		2	0.256410256		9.5	1.00330033	0.255566802				
	90mM	2	0.256410256		9.4375	0.99669967	0.257259297				
		2	0.256410256		9.5	1.00330033	0.255566802				
	120mM	2	0.256410256		9.4375	0.99669967	0.257259297				
		1.5	0.192307692		9.5	1.00330033	0.191675101				
GST-Yih1 (KRK-AAA)	150mM	1	0.128205128		9.4375	0.99669967	0.128629648				
		5	0.641025641	0.307692	9.5	1.00330033	0.638917004	0.307611		0.171257	0.054156
	30mM	4	0.512820513		9.4375	0.99669967	0.514518594				
		3.5	0.448717949		9.5	1.00330033	0.447241903				
	60mM	3	0.384615385		9.4375	0.99669967	0.385888945				
		2	0.256410256		9.5	1.00330033	0.255566802				
	90mM	2	0.256410256		9.4375	0.99669967	0.257259297				
		1	0.128205128		9.5	1.00330033	0.127783401				
	120mM	1	0.128205128		9.4375	0.99669967	0.128629648				
		1.5	0.192307692		9.5	1.00330033	0.191675101				
GST-Yih1 (KRK-EEE)	150mM	1	0.128205128		9.4375	0.99669967	0.128629648				
		5.5	0.705128205	0.455128	9.5	1.00330033	0.702808704	0.45507		0.217311	0.06872
	30mM	6	0.769230769		9.4375	0.99669967	0.771777891				
		5.5	0.705128205		9.5	1.00330033	0.702808704				
	60mM	5	0.641025641		9.4375	0.99669967	0.643148242				
		3	0.384615385		9.5	1.00330033	0.383350202				
	90mM	3	0.384615385		9.4375	0.99669967	0.385888945				
		2.5	0.320512821		9.5	1.00330033	0.319458502				
	120mM	2	0.256410256		9.4375	0.99669967	0.257259297				
		2	0.256410256		9.5	1.00330033	0.255566802				
GST-Yih1 (R175A)	150mM	1	0.128205128		9.4375	0.99669967	0.128629648				
		8.5	1.08974359	0.525641	9.5	1.00330033	1.086158907	0.525351		0.270986	0.085693
	30mM	6	0.769230769		9.4375	0.99669967	0.771777891				
		5.5	0.705128205		9.5	1.00330033	0.702808704				
	60mM	5	0.641025641		9.4375	0.99669967	0.643148242				
		3	0.384615385		9.5	1.00330033	0.383350202				
	90mM	3	0.384615385		9.4375	0.99669967	0.385888945				
		2.5	0.320512821		9.5	1.00330033	0.319458502				
	120mM	2	0.256410256		9.4375	0.99669967	0.257259297				
		4.5	0.576923077		9.5	1.00330033	0.575025304				
GST-Yih1 (R175E)	150mM	1	0.128205128		9.4375	0.99669967	0.128629648				
		5.5	0.705128205	0.455128	9.5	1.00330033	0.702808704	0.45507		0.217311	0.048592
		5.5	0.705128205		9.5	1.00330033	0.702808704				
		6	0.769230769		9.4375	0.99669967	0.771777891				
	30mM	6	0.769230769		9.4375	0.99669967	0.771777891				
		5.5	0.705128205		9.5	1.00330033	0.702808704				
		5.5	0.705128205		9.5	1.00330033	0.702808704				
		5	0.641025641		9.4375	0.99669967	0.643148242				
	60mM	5	0.641025641		9.4375	0.99669967	0.643148242				
		3	0.384615385		9.5	1.00330033	0.383350202				
		3	0.384615385		9.5	1.00330033	0.383350202				
		3	0.384615385		9.4375	0.99669967	0.385888945				
	90mM	3	0.384615385		9.4375	0.99669967	0.385888945				
		2.5	0.320512821		9.5	1.00330033	0.319458502				
		2.5	0.320512821		9.5	1.00330033	0.319458502				
		2	0.256410256		9.4375	0.99669967	0.257259297				
	120mM	2	0.256410256		9.4375	0.99669967	0.257259297				
		2	0.256410256		9.5	1.00330033	0.255566802				
		2	0.256410256		9.5	1.00330033	0.255566802				
		1	0.128205128		9.4375	0.99669967	0.128629648				
GST-Yih1	150mM	1	0.128205128		9.4375	0.99669967	0.128629648				
			Average GST		Average Galactose						
			7.8		9.46875						

C

	GST	GST-Yih1 (KRK-AAA)	GST-Yih1 (KRK-EEE)	GST-Yih1 (R175A)	GST-Yih1 (R175E)	GST-Yih1
GST	0	3.73286E-07	3.45325E-08	1.45371E-05	0.000398621	7.33E-13
GST-Yih1 (KRK-AAA)	3.73E-07	0	0.018691556	0.001343578	0.007723849	0.007724
GST-Yih1 (KRK-EEE)	3.45E-08	0.018691556	0	0.00062536	0.000550682	0.000551
GST-Yih1 (R175A)	1.45E-05	0.001343578	0.00062536	0	0.169815885	0.169816
GST-Yih1 (R175E)	0.000399	0.007723849	0.000550682	0.169815885	0	0.506131
GST-Yih1	7.33E-13	0.007723849	0.000550682	0.169815885	0.50613088	0

Figure S4.1: SQGA of yeast cells overexpressing the indicated proteins.

A: 5µL of ten-fold serially diluted overnight cultures of the yeast strains indicated above were transferred onto solid media with 3AT, or with 2% Glucose or 2% Galactose alone. Growth of two independent transformants were analysed (colonies A and B). B:

Quantification of SQGA in (A). SD= Standard Deviation, SE= Standard Error. C: Calculated P values from T test using data values quantified in (B). Values greater than 0.05 ($p \geq 0.05$) are not stastically significant, while values less or equal to 0.05 (≤ 0.05) are statistically significant. For samples with an unequal sample size: a T-test was carried out assuming unequal variance. For samples with equal sample size: a paired T-test two sample for means was carried out.

A

GST-Yih1$^{KRK\text{-}AAA}$ (A)

AAG (K173) to GCT (Ala)
CGT (R175) to GCG (Ala)
AAG (K176) to GCC (Ala)

```
Query  123  GTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCAT  182
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  407  GTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCAT  466

Query  183  TTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTC  242
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  467  TTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTC  526

Query  243  CTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAG  302
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  527  CTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAG  586

Query  303  CAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGAT  362
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  587  CAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGAT  646

Query  363  AGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCC  422
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  647  AGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCC  706

Query  423  ATGCTAGACCTACTGAAGACCGACTCCGCTATGGCG-GCCGCAAACCATGTCATGAGTGC  481
            ||||||||||||||||||||||||||   ||| || |||||||||||||||||||||||
Sbjct  707  ATGCTAGACCTACTGAAGACCGACTCCAAGATG-CGTAAGGCAAACCATGTCATGAGTGC  765

Query  482  ATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGA  541
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  766  ATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGA  825

Query  542  AACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCAT  601
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  826  AACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCAT  885

Query  602  CGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACAT  661
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  886  CGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACAT  945

Query  662  CAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  707
            |||||||||||||||||||||||||||||||||||||||||||||
Sbjct  946  CAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  991
```

B

GST-Yih1$^{KRK\text{-}AAA}$ (B)

AAG (K173) to GCT (Ala)
CGT (R175) to GCG (Ala)
AAG (K176) to GCC (Ala)

```
Query  74   AGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTG  133
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  354  AGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTG  413

Query  134  TCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCC  193
```

```
Sbjct  414  TCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCC  473

Query  194  AGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAG  253
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  474  AGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAG  533

Query  254  AACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTG  313
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  534  AACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTG  593

Query  314  ACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCT  373
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  594  ACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCT  653

Query  374  CGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAG  433
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  654  CGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAG  713

Query  434  ACCTACTGAAGACCGACTCCGCTATGCCG-GCCGCAAACCATGTCATGAGTGCATGGCGA  492
            ||||||||||||||||||| ||| || ||||||||||||||||||||||||||||||||
Sbjct  714  ACCTACTGAAGACCGACTCCAAGATG-CGTAAGGCAAACCATGTCATGAGTGCATGGCGA  772

Query  493  ATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCC  552
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  773  ATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCC  832

Query  553  GCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTG  612
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  833  GCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTG  892

Query  613  GTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCT  672
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  893  GTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCT  952

Query  673  ACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  711
            |||||||||||||||||||||||||||||||||||||||
Sbjct  953  ACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  991
```

C

GST-Yih1$^{KRK\text{-}EEE}$ (A)

AAG (K173) to GAG (Glu)
CGT (R175) to GAA (Glu)
AAG (K176) to GAG (Ala)
CAA (Q196) to CGA: Undesired mutation

```
Query  71   TGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTG  130
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  351  TGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTG  410

Query  131  GTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGT  190
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  411  GTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGT  470

Query  191  TCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCA  250
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  471  TCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCA  530

Query  251  CAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGA  310
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  531  CAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGA  590

Query  311  GTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAG  370
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  591  GTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAG  650

Query  371  GCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGC  430
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  651  GCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGC  710
```

```
Query  431  TAGACCTACTGAAGACCGACTCCGAGATG-GAAGAGGCAAACCATGTCATGAGTGCATGG  489
            |||||||||||||||||||||||| |||||| | | ||||||||||||||||||||||||
Sbjct  711  TAGACCTACTGAAGACCGACTCCAAGATGCGTA-AGGCAAACCATGTCATGAGTGCATGG  769

Query  490  CGAATCAAGCAGGATGGCTCTGCGGCAACATATCGAGATTCCGATGATGACGGTGAAACG  549
            |||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Sbjct  770  CGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACG  829

Query  550  GCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTT  609
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  830  GCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTT  889

Query  610  GTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAAT  669
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  890  GTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAAT  949

Query  670  TCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  711
            |||||||||||||||||||||||||||||||||||||||||||
Sbjct  950  TCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  991
```

D

GST-Yih1$^{KRK\text{-}EEE}$ (B)

AAG (K173) to GAG (Glu)
CGT (R175) to GAA (Glu)
AAG (K176) to GAG (Ala)

```
Query  59   TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  118
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  341  TACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTC  400

Query  119  ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  178
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  401  ATCGAAGTTGGTGTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTT  460

Query  179  CAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTT  238
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  461  CAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTT  520

Query  239  GACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  298
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  521  GACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCG  580

Query  299  GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  358
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  581  GTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATT  640

Query  359  ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  418
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  641  ACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCG  700

Query  419  TTTGCCATGCTAGACCTACTGAAGACCGACTCCGAGATG-GAAGAGGCAAACCATGTCAT  477
            |||||||||||||||||||||||||||||||||||||| |||||| | | ||||||||||||
Sbjct  701  TTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTA-AGGCAAACCATGTCAT  759

Query  478  GAGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGA  537
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  760  GAGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGA  819

Query  538  CGGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAA  597
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  820  CGGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAA  879

Query  598  CGTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAA  657
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  880  CGTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAA  939

Query  658  ACACATCAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  709
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  940  ACACATCAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  991
```

E

GST-Yih1^{R175A} (A)

CGT (R175) to GCC (Ala)

```
Query  69   AGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTG  128
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  354  AGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTG  413

Query  129  TCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCC  188
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  414  TCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCC  473

Query  189  AGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAG  248
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  474  AGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAG  533

Query  249  AACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTG  308
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  534  AACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTG  593

Query  309  ACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCT  368
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  594  ACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCT  653

Query  369  CGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAG  428
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  654  CGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAG  713

Query  429  ACCTACTGAAGACCGACTCCAAGATGGCCAAGGCAAACCATGTCATGAGTGCATGGCGAA  488
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||| a.)
Sbjct  714  ACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATGAGTGCATGGCGAA  773

Query  489  TCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCCG  548
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  774  TCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCCG  833

Query  549  CCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTGG  608
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  834  CCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTGG  893

Query  609  TGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCTA  668
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  894  TGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCTA  953

Query  669  CGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  706
            |||||||||||||||||||||||||||||||||||||
Sbjct  954  CGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  991
```

F

GST-Yih1^{R175A} (B)

CGT (R175) to GCC (Ala)

```
Query  74   AGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTG  133
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  354  AGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTG  413

Query  134  TCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCC  193
```

```
Sbjct  414  TCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCC  473

Query  194  AGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAG  253
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  474  AGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAG  533

Query  254  AACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTG  313
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  534  AACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTG  593

Query  314  ACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCT  373
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  594  ACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCT  653

Query  374  CGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAG  433
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  654  CGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAG  713

Query  434  ACCTACTGAAGACCGACTCCAAGATGGCCAAGGCAAACCATGTCATGAGTGCATGGCGAA  493
            |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||| a.)
Sbjct  714  ACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATGAGTGCATGGCGAA  773

Query  494  TCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCCG  553
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  774  TCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCCG  833

Query  554  CCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTGG  613
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  834  CCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTGG  893

Query  614  TGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCTA  673
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  894  TGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCTA  953

Query  674  CGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  711
            |||||||||||||||||||||||||||||||||||||
Sbjct  954  CGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  991
```

G

GST-Yih1^{R175E} (A)

TCG (S138) to TTG: Silent mutation
GAC (D171) to GAT: Silent mutation
TCC (S172) to TCG: Silent mutation
CGT (R175) to GAA (Glu)

```
Query  127  GTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTC  186
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  413  GTCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTC  472

Query  187  CAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACA  246
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  473  CAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACA  532

Query  247  GAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGT  306
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  533  GAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGT  592

Query  307  GACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTTGGACCCCATTACTGATAGAGGC  366
            |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||| b.)
Sbjct  593  GACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGC  652

Query  367  TCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTA  426
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  653  TCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTA  712

Query  427  GACCTACTGAAGACCGATTCGAAGATGGAAAAGGCAAACCATGTCATGAGTGCATGGCGA  486
            |||||||||||||||||  ||  ||||||    |||||||||||||||||||||||||||| c.) d.)
e.)
Sbjct  713  GACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATGAGTGCATGGCGA  772

Query  487  ATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCC  546
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  773  ATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCC  832

Query  547  GCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTG  606
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  833  GCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTG  892

Query  607  GTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCT  666
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  893  GTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCT  952

Query  667  ACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  705
            |||||||||||||||||||||||||||||||||||||||
Sbjct  953  ACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  991
```

H

GST-Yih1^{R175E} (B)

GAC (D171) to GAT: Silent mutation
TCC (S172) to TCG: Silent mutation
CGT (R175) to GAA (Glu)

```
Query  71   AGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTG  130
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  354  AGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTG  413

Query  131  TCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCC  190
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  414  TCTGCACTTCTTTGGCTAAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCC  473

Query  191  AGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAG  250
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  474  AGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAG  533

Query  251  AACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTG  310
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  534  AACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTG  593

Query  311  ACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCT  370
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  594  ACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCT  653

Query  371  CGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAG  430
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  654  CGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAG  713

Query  431  ACCTACTGAAGACCGATTCGAAGATGGAAAAGGCAAACCATGTCATGAGTGCATGGCGAA  490
            |||||||||||||||||  ||  ||||||    |||||||||||||||||||||||||||| a.) b.) c.)
Sbjct  714  ACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAACCATGTCATGAGTGCATGGCGAA  773

Query  491  TCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCCG  550
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  774  TCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCCG  833

Query  551  CCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTGG  610
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  834  CCGGCTCCAGAATGCTGCACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTGG  893

Query  611  TGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCTA  670
```

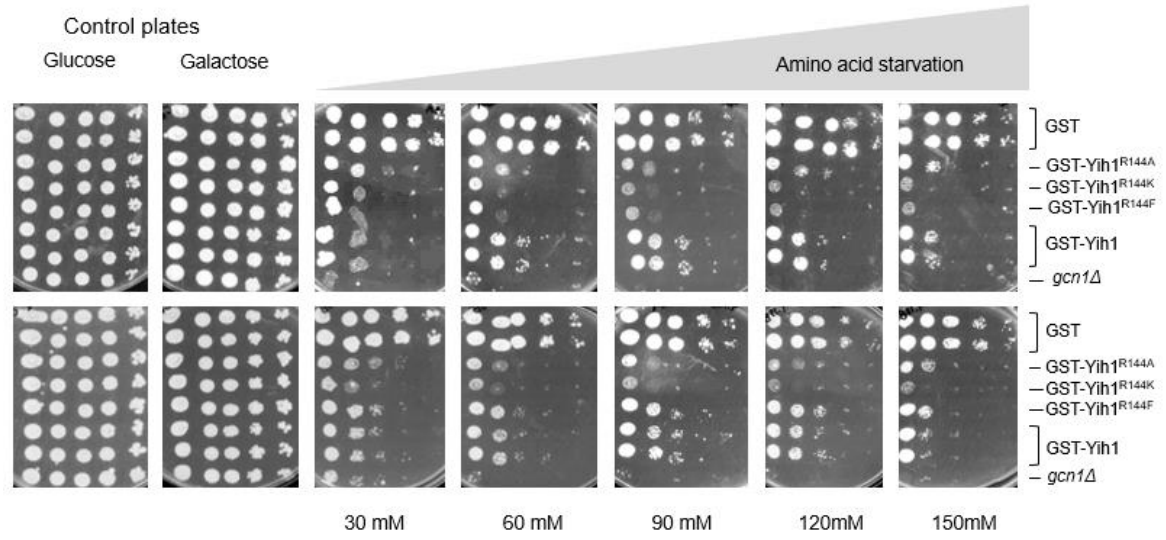
```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  894   TGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCTA  953

Query  671   CGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  708
              |||||||||||||||||||||||||||||||||||||
Sbjct  954   CGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTCG  991
```

Figure S4.2: Whole Nucleotide BLAST sequence alignment of Group 3 Yih1 mutants

Plasmid DNA of the indicated Yih1 proteins were isolated as described in Materials and Methods (Section 2.1.10 and Section 2.1.11). The samples indicated (from A to H) were commercially sequenced using primers 3035-B and ES48. Query: Wild type Yih1 sequence. Sbjct: Plasmid DNA samples with the indicated substitutions. Magenta colour indicates mutations that are expected. Turquoise colour indicates unexpected silent mutations. Red colour indicates unexpected mutations. Sequencing order HC00073122.

S5: Group 4: Potential role of Post-Translational Modification in Yih1 function

A



B

Yeast strain	3AT concentration	Growth score	Normalisation 3AT with GST	Average	Galactose growth score	Normalisation to Galactose	Normalised 3AT/Normalised Galactose	Average	SD	SE
		8	1.03	1.00	9.5	1	1.032258065	1.00	0.077687707	0.024567
	30mM	8.5	1.10		9.5	1	1.096774194			
		8.5	1.10		9.5	1	1.096774194			
	60mM	8.5	1.10		9.5	1	1.096774194			
		7	0.90		9.5	1	0.903225806			
	90mM	7.5	0.97		9.5	1	0.967741935			
		7.5	0.97		9.5	1	0.967741935			
	120mM	8	1.03		9.5	1	1.032258065			
		7	0.90		9.5	1	0.903225806			
GST	150mM	7	0.90		9.5	1	0.903225806			
	30mM	3.5	0.45	0.36	9.5	1	0.451612903	0.36	0.0657938	0.029424
	60mM	2.5	0.32		9.5	1	0.322580645			
	90mM	2	0.26		9.5	1	0.258064516			
	120mM	3	0.39		9.5	1	0.387096774			
GST-Yih1 (R144A)	150mM	3	0.39		9.5	1	0.387096774			
	30mM	2.5	0.32	0.17	9.5	1	0.322580645	0.17	0.0965589	0.043182
	60mM	1.5	0.19		9.5	1	0.193548387			
	90mM	1.5	0.19		9.5	1	0.193548387			
	120mM	0.5	0.06		9.5	1	0.064516129			
GST-Yih1 (R144K)	150mM	0.5	0.06		9.5	1	0.064516129			
GST-Yih1 (R144F)										
		2.5	0.32	0.41	9.5	1	0.322580645	0.41	0.119659593	0.03784
	30mM	2.5	0.32		9.5	1	0.322580645			
		4	0.52		9.5	1	0.516129032			
	60mM	4	0.52		9.5	1	0.516129032			
		4.5	0.58		9.5	1	0.580645161			
	90mM	4.5	0.58		9.5	1	0.580645161			
		2	0.26		9.5	1	0.258064516			
	120mM	2	0.26		9.5	1	0.258064516			
		3	0.39		9.5	1	0.387096774			
GST-Yih1	150mM	3	0.39		9.5	1	0.387096774			
	30mM	0	0.00	0.01	9.5	1	0	0.01	0.025806452	0.011541
	60mM	0.5	0.06		9.5	1	0.064516129			
	90mM	0	0.00		9.5	1	0			
	120mM	0	0.00		9.5	1	0			
gcn1del	150mM	0	0.00		9.5	1	0			
			7.75							
					9.5					

Figure S5.1: SQGA of yeast cells overexpressing the indicated proteins.

A: 5µL of ten-fold serially diluted overnight cultures of the yeast strains indicated above were transferred onto solid media with 3AT, or with 2% Glucose or 2% Galactose alone. Growth of two independent transformants were analysed (colonies A and B) B: Quantification of SQGA in (A). SD= Standard Deviation, SE= Standard Error. Sample size was too small for a sufficient T test.

A

GST-Yih1^{R144A} (A)

AGA (R144) to GCT (Ala)

GGC (G145) to GGG: Silent mutation

```
Query  1    ATCATCGTTGTGAAAGTGCCGCAGCATGAATACATGACACTGCAGATCTCCTTCCCGACA  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  91   ATCATCGTTGTGAAAGTGCCGCAGCATGAATACATGACACTGCAGATCTCCTTCCCGACA  150

Query  61   CACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTGTCTGCACTTCTTTGGCT  120
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  151  CACTACCCCTCCGAGGAGGCTCCTAATGTCATCGAAGTTGGTGTCTGCACTTCTTTGGCT  210

Query  121  AAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCCAGGAAGTGATGGACTCT  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  211  AAGCGCGATCTCTACGATACCAAGTACCTTCAGCATTTGTTCCAGGAAGTGATGGACTCT  270

Query  181  GTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAGAACTCGACGGTGTCTTG  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  271  GTTTTCCACCGCGGATCTGTCTGTCTATTTGACTTCCTCACAGAACTCGACGGTGTCTTG  330

Query  241  TACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTGACATTCCCACAGACCCC  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  331  TACGTTGAACCAGAGGAGGAGACAGAACCGGTCCAGCAGAGTGACATTCCCACAGACCCC  390

Query  301  TTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATGCTGGGTCGACTTTCATGGCCTTT  360
```

```
Sbjct  391  TTCGAGGGCTGGACCGCGTCGGACCCCATTACTGATAGAGGCTCGACTTTCATGGCCTTT  450

Query  361  GCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAGACCTACTGAAGACCGAC  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  451  GCAGCACATGTTACCTCCGAGGAACAAGCGTTTGCCATGCTAGACCTACTGAAGACCGAC  510

Query  421  TCCAAGATGCGTAAGGCAAACCATGTCATGAGTGCATGGCGAATCAAGCAGGATGGCTCT  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  511  TCCAAGATGCGTAAGGCAAACCATGTCATGAGTGCATGGCGAATCAAGCAGGATGGCTCT  570

Query  481  GCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCCGCCGGCTCCAGAATGCTG  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  571  GCGGCAACATATCAAGATTCCGATGATGACGGTGAAACGGCCGCCGGCTCCAGAATGCTG  630

Query  541  CACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTGGTGGCCCGTTGGTTCGGC  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  631  CACCTCATCACCATCATGGATGTGTGGAACGTCATCGTTGTGGTGGCCCGTTGGTTCGGC  690

Query  601  GGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCTACGGCAAGAGAAGCTGTT  660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  691  GGTGCCCACATAGGTCCCGACCGGTTTAAACACATCAATTCTACGGCAAGAGAAGCTGTT  750

Query  661  GTCAGGGCCGGCTTCGACTCG  681
            |||||||||||||||||||||
Sbjct  751  GTCAGGGCCGGCTTCGACTCG  771
```

B

GST-Yih1^{R144A} (B): Mutation not inserted

GAA (E203) to GAG: Silent mutation

```
Query  1    ATACCAAGTACCTTCAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGAT  60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  227  ATACCAAGTACCTTCAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGAT  286

Query  61   CTGTCTGTCTATTTGACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGG  120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  287  CTGTCTGTCTATTTGACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGG  346

Query  121  AGGAGACAGAACCGGTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCG  180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  347  AGGAGACAGAACCGGTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCG  406

Query  181  CGTCGGACCCCATTACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCT  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  407  CGTCGGACCCCATTACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCT  466

Query  241  CCGAGGAACAAGCGTTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGG  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  467  CCGAGGAACAAGCGTTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGG  526

Query  301  CAAACCATGTCATGAGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAG  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  527  CAAACCATGTCATGAGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAG  586

Query  361  ATTCCGATGATGACGGTGAGACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCA  420
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Sbjct  587  ATTCCGATGATGACGGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCA  646

Query  421  TGGATGTGTGGAACGTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTC  480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  647  TGGATGTGTGGAACGTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTC  706

Query  481  CCGACCGGTTTAAACACATCAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCG  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Sbjct  707  CCGACCGGTTTAAACACATCAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCG  766

Query  541  ACTCG  545
            |||||
Sbjct  767  ACTCG  771
```

C

GST-Yih1^{R144K} (A)

GAC (D90) to GGC (Gly)

AGA (R144) to AAG (Lys)

```
Query  1    CAAGTACCTTCAGCATTTGTTCCAGGAAGTGATGGGCTCTGTTTTCCACCGCGGATCTGT  60
            |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Sbjct  231  CAAGTACCTTCAGCATTTGTTCCAGGAAGTGATGGACTCTGTTTTCCACCGCGGATCTGT  290

Query  61   CTGTCTATTTGACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGA  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  291  CTGTCTATTTGACTTCCTCACAGAACTCGACGGTGTCTTGTACGTTGAACCAGAGGAGGA  350

Query  121  GACAGAACCGGTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTC  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  351  GACAGAACCGGTCCAGCAGAGTGACATTCCCACAGACCCCTTCGAGGGCTGGACCGCGTC  410

Query  181  GGACCCCATTACTGATAAGGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGA  240
            |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Sbjct  411  GGACCCCATTACTGATAGAGGCTCGACTTTCATGGCCTTTGCAGCACATGTTACCTCCGA  470

Query  241  GGAACAAGCGTTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAA  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  471  GGAACAAGCGTTTGCCATGCTAGACCTACTGAAGACCGACTCCAAGATGCGTAAGGCAAA  530

Query  301  CCATGTCATGAGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTC  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  531  CCATGTCATGAGTGCATGGCGAATCAAGCAGGATGGCTCTGCGGCAACATATCAAGATTC  590

Query  361  CGATGATGACGGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGA  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  591  CGATGATGACGGTGAAACGGCCGCCGGCTCCAGAATGCTGCACCTCATCACCATCATGGA  650

Query  421  TGTGTGGAACGTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGA  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  651  TGTGTGGAACGTCATCGTTGTGGTGGCCCGTTGGTTCGGCGGTGCCCACATAGGTCCCGA  710

Query  481  CCGGTTTAAACACATCAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTC  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  711  CCGGTTTAAACACATCAATTCTACGGCAAGAGAAGCTGTTGTCAGGGCCGGCTTCGACTC  770

Query  541  G  541
            |
Sbjct  771  G  771
```

D

GST-Yih1^{R144K} (A)

GAC (D90) to GGC (Gly)

AGA (R144) to AAG (Lys)

```
Query  1    GGATCTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCA  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  57   GGATCTTCTCTCCAAGAAGCAGGAAGACGGAAGCATCATCGTTGTGAAAGTGCCGCAGCA  116

Query  61   TGAATACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAA  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  117  TGAATACATGACACTGCAGATCTCCTTCCCGACACACTACCCCTCCGAGGAGGCTCCTAA  176
```

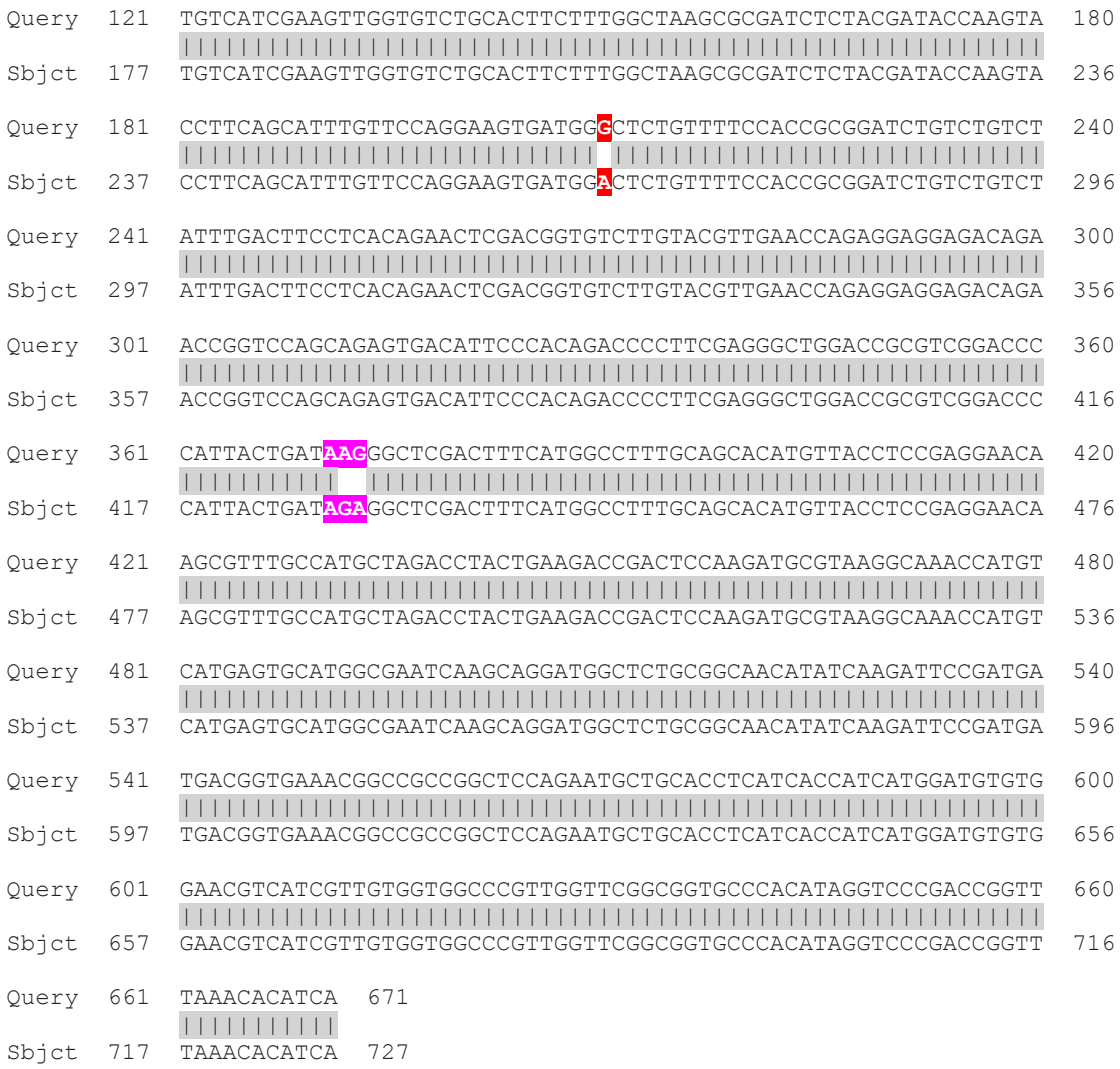


Figure S5.2: Whole Nucleotide BLAST sequence alignment of Group 3 Yih1 mutants

Plasmid DNA of the indicated Yih1 proteins were isolated as described in Materials and Methods (Section 2.1.10 and Section 2.1.11). The samples indicated (from A to D) were commercially sequenced using primers 3035-B and ES48. Query: Wild type Yih1 sequence. Sbjct: Plasmid DNA samples with the indicated substitutions. Magenta colour indicates mutations that are expected. Turquoise colour indicates unexpected silent mutations. Red colour indicates unexpected mutations. Sequencing order HC00488651.

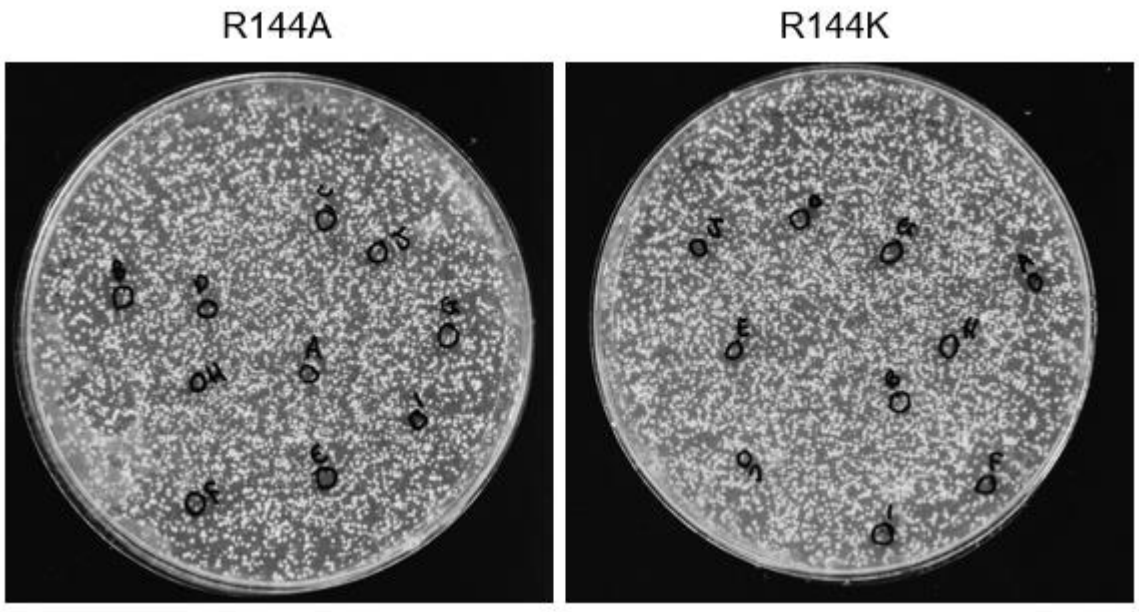


Figure S5.3: Yeast transformation of Yih1^{R144A} or Yih1^{R144K} fragments into pES187-b1 plasmid.

Yih1^{R144A} or Yih1^{R144K} fragments were subject to yeast transformation using a Yih1 deletion strain (ES11001b) containing the pES187-b1 plasmid (1:40 dilution) that was previously cut open using BglII and SpeI restriction sites prior to this study (See details in Materials and Methods). Samples were plated on SDWLIV + glucose using glass beads and incubated overnight at 30°C. With these transformants, two colonies were picked (A and B colonies) and a Semi Quantitative Growth Assay was carried out, as described in Materials and Methods.