Overproduction and purification of Haa1 protein envisaging functional and structural studies of this transcription factor involved in yeast adaptive response to acetic acid

Ricardo Aguilar Andrade Ribeiro

Abstract

The transcription factor Haa1 is the main regulator of yeast genomic response to acetic acid stress, regulating, directly or indirectly, the transcription of 80% of the acetic acid-activated genes. Using error-prone PCR, a more tolerant Haa1 mutant (S135F), with a serine residue exchange to a phenylalanine residue in position 135, was selected in a previous work. This Haa1 mutation leads to increased transcriptional activation of Haa1-target genes, presumably due to the effect of altered $Haa1_{S135F}$ structure in binding interactions with the target genes promoter binding sites. In this work, it was attempted the expression of the Haa1 protein using the *Pichia pastoris* host system. Given that this strategy was not successful, the *E.coli*-pET expression system was used to overproduce the his-tagged-Haa1 and $Haa1_{S135F}$ proteins, followed by their purification using Ion Metal Affinity Chromatography and subsequent polishing by Fast Protein Liquid Chromatography.

In order to get some insight into the complexes formed between the Haa1 proteins and the promoter region of one of their target genes, *TPO3*, Atomic Force Microscopy was used. Haa1 and $Haa1_{S135F}$ proteins were found to be bind at 1/3 of the *TPO3* promoter region, where the Haa1-binding site maps. Preliminary results appear to suggest the existence of differences in the height of Haa1 proteins bound to DNA, presumable related with their binding.

Keywords: Transcription factor, Haa1, *Saccharomyces cerevisiae*, Acetic acid response, Recombinant protein expression, Atomic Force Microscopy

Introduction

All different living cell types have the ability to respond to changes in environmental conditions, in particular to environmental insults. Among the different genes, whose transcription is activated in response to environmental stress, are those related to carbohydrate metabolism, fatty acid metabolism, maintenance of the cellular redox potential, solutes transport, protein folding and degradation, vacuolar and mitochondrial functions, intracellular signalling, detoxification of reactive oxygen species (ROS), cell wall modulation, autophagy and secretion, DNA-damage repair mechanisms, and others. [1 2]

Man has been using applying natural preservatives for centuries, even without knowing how they protect food from spoiling. The resistance of spoilage yeasts to weak acids poses a serious problem regarding weak acid usefulness as preservative agents, bringing major economic losses. [3] At low pH of the growth medium, depending of the concentration present, weak acids may have a strong effect on microbial growth and viability. In low pH cultures, the undissociated form of the weak acid will enter in yeast cell by simple diffusion across the lipid bilayer. Therefore, the uncharged acid will diffuse across the plasma membrane to the cytoplasm more easily than the charged form. The charged form is retained if a specific permease is not present to facilitate diffusion or active transport. Once the acid passes through the plasma membrane to the cytoplasm in the undissociated form, it will dissociates in the cell as a consequence of the near neutral intracellular pH of the cytosol, generating a proton and the acid anion (H^+, $XCOO^-$, respectively). The charged anion accumulates in the cytosol because it will not readily diffuse through the cellular lipid bilayer of the cell. [4 5 6] Concerning the physiological action of acetic acid in yeast, When low pH cultures of *S.cerevisiae* are exposed to acetic acid, the undissociated form is of sufficiently low molecular weight to enter in the cell by facilitated diffusion through the open plasma membrane glycerol channel Fps1p, and/or by passive diffusion across the plasma membrane. Once inside the cell, acetic acid will dissociate leading to an increase of the intracellular pool of acetate anions, triggering a multifaceted adaptation stress response (figure 1). [13 15 20 35 8] In order to counteract acidification of cytosol and dissipation of H^+ gradient, plasma membrane H^+-ATPase Pma1 and the vacuolar H^+-ATPase (V-ATPase) activity increases (figure 1). [9 10 11]

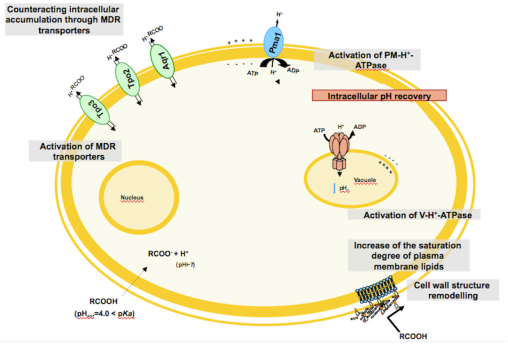

Figure 1 - Mechanistic model for the adaptive yeast response to weak acid stress. The low pH of the external environments favours the entrance of the undissociated form of the weak acid (RCOOH) into the cell. Once in the cytosol, the weak acid dissociates, leading to $RCOO^-$ accumulation in the cytosol. To counteract intracellular acidification H^+-ATPases present in plasma membrane and vacuolar membranes are activated. Cell wall remodelling contributes to reduce the diffusion rate of undissociated weak acids and reduce weak acid-induced plasma membrane damage. Adapted from Mira *et al.* (2010) [15]

Haa1 transcription is considered the main regulator of yeast adaptive genomic response to acetic acid, since approximately 80% of the acetic acid activated genes are regulated, indirectly or directly, by this transcription factor. [12] The protective role exerted by *HAA1* is observed for acetic propionic and lactic acids [13] Moreover, several genes from the Haa1-regulon are themselves important determinants of tolerance to acetic acid. [13 14]

The Haa1p-depedent genes encode proteins clustered in several functional classes such as those related with multidrug resistance transporters (42%), cell wall (4.6%), lipid metabolism (6.2%), regulation of carbohydrate metabolism (12.3%), protein folding (4.6%), carbohydrate metabolism (9.2%), amino acid metabolism (3.1%), nucleic acid processing (4.6%) and transcription factors (9.2%), which illustrates the very important role of Haa1 at the different physiological levels. [12 15] Among the acetic acid-responsive genes, *TPO2* and *TPO3, AQR1*, which encode three Multidrug Resistance transporters (MDR) belonging to the Major Facilitator Superfamily (MFS), were described to exert a role in protection against acetic acid, apparently by mediating the active expulsion of acetate from the cell interior. [13]

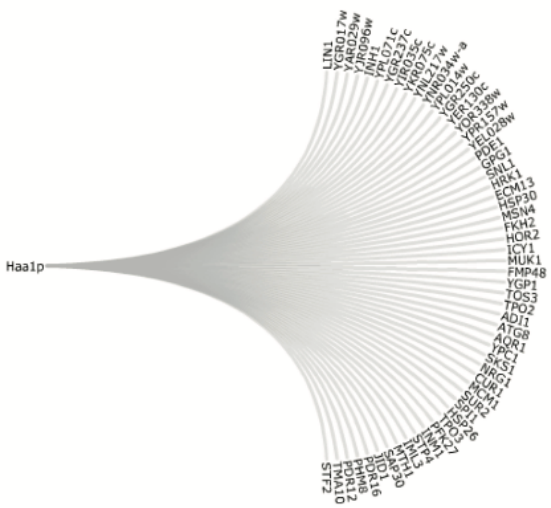

Figure 2 - Acetic acid-induced genes presumably directly regulated by Haa1 transcription factor, having the minimal functional Haa1-responsive element [5'- (G/C)(A/C)GG(G/C)G-3')] in their promoter region. Mira *et al.* 2011 [3]

The mapping of the promoter region of *TPO3* allowed the identification of a functional binding site for Haa1 [5'-(G/C)(A/C)GG(G/C)G-3'], designated as Haa1-responsive element (HRE), in which it was found that Haa1p binds to *TPO3* promoter region, proving that *TPO3* is a Haa1-direct target gene. [3]. A list of genes presumed to be directly regulated by Haa1 is indicated in figure 3. Additionally, among the genes that are regulated by Haa1 transcription factor, the expression of *SAP30* and *HRK1* genes are the most relevant upon acetic acid stress resistance. *HRK1* encodes a protein kinase belonging to a subfamily of protein kinases implicated in activation of the plasma membrane H^+-ATPase Pma1 in response to glucose metabolism. [12 16] Acetic acid challenged Δ*hrk* and Δ*haa1* mutant yeast cells showed increased accumulation of acetic acid. This may lead to speculate that the protein kinase is involved in the Hrk1 regulation by phosphorylation of plasma membrane acetate transporters. [10] Another Hrk1p target is Gph1p, a cytosolic enzyme involved in the catabolism of glycogen, whose expression seems to be protective against acetic acid stress.

Other relevant Haa1 target*SAP30* encodes a subunit of the Rpd3L histone deacetylase complex, playing a pivotal role in the regulation of transcription. Elimination of *SAP30* gene led to a strong susceptibility phenotype to acetic acid. [10] *HSP30* encodes a hydrophobic plasma membrane protein whose expression is activated by heat shock, ethanol and weak acid stress. [17] The transcription of *HSP30* gene is highly induced upon acetic acid stress in a Haa1-dependent manner. [10] Cell wall proteins encoded by *SPI1* and *YGP1* genes also have a protective role against the diffusional entry of the acid. [13] *SPI1* is involved in the decrease of cell porosity, whose transcription is activated in dependence of Haa1 transcription factor in the presence of acetic acid. [18] *YGP1* is a cell wall-related secretory glycoprotein. [13 19 20]

Recombinant protein production requires the selection of an appropriate host/expression system, dependent of the characteristics of the protein and intended application. [21] One of the widely used *P.pastoris* commercial protein expression systems is PichiaPink™ Expression system (Invitrogen) [22] *P.pastoris*, unlike bacterial expression systems, is a eukaryotic system that has the ability to perform the post-translational modifications occurring in *S.cerevisiae* and in more complex eukaryotes: correct folding and assembly, disulphide bond formation, O- and N- linked glycosylation of signal sequences. [23 24] Furthermore, the choice to use inducible an *AOX1* promoter that is tightly regulated holds advantage for heterologous gene expression, since it is strongly repressed by glycerol or glucose and, upon depletion of these carbon sources, the promoter is de-repressed and fully induced in the presence of methanol, allowing to uncouple the phase of growth from the protein production phase. In addition, *P.pastoris* has the ability to secrete properly folded, pos-translationally processed and functional recombinant proteins in the supernatant together with the fact that *P.pastoris* secrete low levels of native proteins, meaning that the secreted heterologous protein constitutes the vast majority of the total protein in the culture supernatant, convenient for recovery and purification steps.[25 26] Although *P.pastoris* has been successfully used to produce various recombinant

heterologous proteins, the variability in efficiency in protein secretion is problematic. In fact, translocation determined by the secretion signal peptide, processing and folding in the Endoplasmic Reticulum (ER) and Golgi, and secretion out of the cell are reported as physiological limitations that compromise protein production. [27 28 29 30 31] Concerning protein synthesis and translocation, the *S.cerevisiae* α mating sequence used in this work, fused with the recombinant protein, mediates the entry in the secretory pathway and is one of the most commonly secretion signal peptides used to drive secretion in *P.pastoris.*
Another system widely used for protein overproduction consists in using *E.coli* has the host cell, although some problems may arise when expressing a heterologous protein like *S.cerevisiae* Haa1 transcription factor, namely the lack of post-translational modifications. In pET expression system (Novagen), one of the most widely and efficient systems to express heterologous protein in *E.coli*, the gene to be overexpressed is under the control of T7 promoter, which is strongly induced by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to the growth medium. [32] IPTG induces transcription of the T7 RNA polymerase gene, already present in the genome of the bacteria used for overexpression, which then activates the T7 promoter and allows the expression of the recombinant gene. The pET vector also allows a coding sequence for a C-terminal Histidine-tag fused to the recombinant protein, facilitating protein purification and detection. Histidine-tagged proteins have a high selective affinity for Ni^{2+} and several other metal ions that can be immobilized on chromatographic media using chelating ligands. Another purification technique used in this work is Size Exclusion Chromatography (SEC). SEC separates molecules according to differences in size as they pass through a SEC medium packed in a column.

AFM is a promising technique for the advance of science in many areas of biological research. A core component of the equipment is a micro-machined probe mounted at the end of a cantilever. A voltage-driven piezoelectric (PZT) transducer enables spatial positioning of the cantilever probe in the *x, y* and *z* directions with precision, converting a electrical stimuli in mechanical motion. The AFM head contains an optical deflection system based on a laser diode and a photodetector. The principle of AFM operation is based on the scanning probe in the *x-y* plane over the sample surface. Moreover, the interaction force between the probe and the sample surface is monitored by measuring the deflection (vertical bending) of the cantilever, detected by a laser beam and reflected to a photodiode. By recording every pixel scanned in the *x-y* plane it is possible to reconstruct the surface topography of the sample. [33 34 35] Currently, AFM has been used in biological research to determine the function of membrane proteins, to explore and quantify interactions between DNA and ligands such as peptides or proteins or even to analyse the biomechanical and biochemical properties of cell surfaces at a nanometer resolution. [36 37 38] Focusing on yeast studies, AFM can be used as a complementary tool to tackle some challenges as to visualise what is occurring at the cell surface and how a stress can have impact on the biophysical properties of the cell wall or other cell components or study proteins present in cell surface that have a pivotal role in adhesion and microbial infection. [39 40]

Materials and Methods

Strains, growth conditions and plasmids

The methylotrophic *Pichia pastoris* PichiaPink™ strain (*ade2*, *prb1*, *pep4*) was used for plasmid maintenance and to express the recombinant protein, using Pichia Pink™ Expression System from Invitrogen
For overexpression of the recombinant protein Histidine-tagged Haa1 and the derived mutant using pET expression system (Novagen),*Escherichia coli BL21-CodonPlus (DE3)-RIL* (*E. coli* B F– *ompT hsdS*(rB– mB–) *dcm*+ Tet r *gal* λ(DE3) *end A* Hte [*argU ileY leuW* Cam^r]) (Stratagene) was chosen.
The *Escherichia coli XL1-Blue* strain (*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F´ *proAB lacI*q*Z_M15* Tn*10* (Tet)]) was the bacterial strain used for plasmid maintenance and general molecular cloning procedures. For obtain suitable transformants, *Pichia* cells were cultured in Yeast Peptone Dextrose (YPD) medium, containing 20 g glucose (Merck), 20 g bactopeptone (Difco) and 10 g yeast extract (Difco) and selected on Pichia Adenine Dropout™ (PAD) agar lacking adenine. To express the recombinant protein, *Pichia* cells were cultured in Buffered Glycerol-complex (BMGY) medium (1%(m/v) yeast extract, 2%(m/v) peptone, 100mM potassium phosphate pH 6.0, 1.34%(m/v) YNB, 0.00004%(m/v)biotin and 1%(v/v) glycerol) and Buffered Methanol-complex (BMMY) medium (1%(m/v) yeast extract, 2%(m/v) peptone, 100mM potassium phosphate pH 6.0, 1.34%(m/v) YNB, 0.00004%(m/v)biotin and 1%(v/v) methanol). BMGY and BMMY media were both prepared from stock solutions Yeast Nitrogen Base with ammonium sulphate (YNB) without amino acids (13.4%(m/v) YNB), 0.02%(m/v) biotin, 5%(v/v) methanol, 10%(v/v) glycerol and 1M potassium phosphate buffer pH 6.0. *E.coli* cells were cultured in rich solid or liquid Luria Bertani (LB) growth medium. For *E.coli* strain used for plasmid maintenance and cloning procedures, LB medium was supplemented with ampicillin (150mg/L), while for *E.coli* strain used for overexpression of the recombinant proteins, LB medium was supplemented with ampicillin (150mg/L) and chloramphenicol (30mg/L) to maintain selection during growth.

Overexpression of *HAA1* in *Pichia pastoris*

Following isolation of *HAA1* from *S.cerevisiae* S288C, the product was cloned into pPINK™ vector producing the recombinant pPINK-HC(+)-$HAA1_{WT}$ plasmid. The fidelity of the cloned sequence was evaluated by DNA sequencing. To prepare electrocompetent cells, PichiaPink™ (from Invitrogen included in the Pichia Pink™ Expression System Kit) strain 4 (*ade2, prb1, pep4*) cells were grown at 30°C for 3 days and a single colony was used to pre-inoculate 20ml of YPD media in a sterile100ml flask at 30 °C at 250 rpm for 2 days. The starter culture was used to inoculate 100 ml of YPD media in a 1L flask in order to start a growth with 0.2 OD_{600nm}, at 30 °C at 250 rpm during 1 day. When the inoculum reached 1.3-1.5 OD_{600nm}, cell suspension was centrifuged at 2000g at 4°C during 5 minutes. Pellet was ressuspended in 20 ml YPD with 4ml HEPES buffer 1M and 250 µl DTT, followed by incubation at 30°C during 15 minutes without agitation. Cell suspension were bringing to 100 ml with cold bidistillate sterile water and centrifuged at 1500g at 4°C for 5 minutes. Pellet was washed by centrifugation at 1500G at 4°C

during 5 minutes in 50 ml cold bidistillate sterile water and ressuspended in 300 µl cold sorbitol 1M. Electroporation was performed according to the instrument manufacturer's instructions for yeast.
The attempt of overproducing the recombinant Haa1 protein was performed according to the User Manual Pichia Pink™ Expression System Kit (Invitrogen). The production of the recombinant his-tagged Haa1 protein was evaluated through Western blot.

Overexpression of $Haa1_{WT}$ and the derived mutant $Haa1_{S135F}$ in *E.coli*

The overexpression of Haa1 and the derived mutant harbouring the mutation S135F was carried out using E. *coli BL21-CodonPlus (DE3)-RIL* (Agilent Technologies) as host cell, using pET expression system (Novagen) The two constructs harbouring $HAA1_{WT}$ coding sequence and the derived mutant $HAA1_{S135F}$ coding sequence were already available in lab (Silvia Henriques, unpublished results) and The conditions for overexpression were performed according to Mira *et al.* (2008). The production of the recombinant Haa1 and $Haa1_{S135F}$ his-tagged proteins was evaluated through SDS-PAGE and Western blot.

Purification by Ion Metal Affinity Chromatography and Fast Protein Liquid Chromatography of $Haa1_{WT}$ and the derived mutant $Haa1_{S135F}$

To each millilitre of cell culture volume, 50 µl of lysis buffer (10 mM imidazole, 10mM Tris, 0.625 mM NaCl) and 4 µl of protease inhibitor (100mM PMFS) were added and the cell suspension was kept in ice. Cell suspensions were sonicated to promote cell disruption using Sonifier® 250 sonicator, which were subjected to 4 sonication cycles for 30 seconds interspersed with 30 seconds on ice, using Duty cycle of 50% and an output control of 4.5. Cell free extracts were centrifuged at 25.000g for 30 minutes at 4°C and supernatant was collected and frozen at -80°C until further use. Since Haa1 and the derived mutant included a C– terminal 6×His tag, the purification was achieved by Ion Metal Affinity Chromatography (IMAC), using a nickel coated column (HisTrap ® HP). The column was first washed with water and equilibrated with buffer A (10 mM imidazole, 10mM Tris, 0.625 mMNaCl). The supernatant was loaded on the column after being clarified by centrifugation at 25.000g for 30 minutes at 4°C, and filtered through a Millipore™ syringe filter. The column was again washed with buffer A, and with 5% buffer B (500 mM imidazole, 10mM Tris, 0.625 mM NaCl) Protein elution was carried out using a stepwise increasing gradient of imidazole that ranged from 10mM to 500mM. Under these conditions, the recombinant protein $Haa1_{WT}$ and the derived mutant started to elute when imidazole concentration was above 150mM.
The purity of each collected protein fraction was assessed by SDS-PAGE with GelCode™ Blue Safe Stain (ThermoFisher Scientific) and Western Blot analysis. The purest fractions were mixed and concentrated using Amicon® Ultra-15 Centrifugal Filters 10kDa cutoff (Merck Millipore), according to the manufacturer's instructions. The concentrated fractions of the recombinant protein Haa1 and the derived mutant were subjected to Fast Protein Liquid Chromatography (FPLC), using a 10/300 GL column (Healthcare®) containing Superdex® 75 matrix. The column was equilibrated with 10mM Tris-HCl, 200 mM NaCl, pH 7.4, at a constant flow rate of 1.2 ml/min. After sample injection, elution was performed with the same buffer used to equilibrate the column. The purest fractions were mixed and concentrated and further analysed by SDS-PAGE with GelCode™ Blue Safe Stain (Thermo Fisher Scientific).

Western blot

This method was performed to evaluate the purity of the eluted fractions (supernatant and cell free extracts) of the his-tagged $Haa1_{WT}$ protein overproduced in *P.pastoris* and the purity of the eluted fractions of the his-tagged Haa1 and $Haa1_{S135F}$ proteins purified by Ni^{2+} -based Ion Metal Affinity Chromatography, overproduced in *E.coli*. Denatured samples were run on a 8%polyacrilamide gel and transbloted onto nitrocellulose membranes. Membranes were incubated for 1 hour with 15mL of blocking solution (80 g/l NaCl, 2 g/l KCl, 6.1 g/l Na_2HPO_4 and 2 g/l KH_2PO_4, pH 7.4, 0.05 %v/v Twin-20, 5.0 %w/v non-fat dry milk). The membrane was then incubated with 10 µl of his-probe mouse monoclonal IgG_1 (Santa Cruz Biotechnology) and left at 4°C *overnight* with gentle agitation. After incubation, membranes were washed 3 times, each time 10 minutes with 15ml PBST buffer (80 g/l NaCl, 2 g/l KCl, 6.1 g/l Na_2HPO_4 and 2 g/l KH_2PO_4, pH 7.4, 0.05 %v/v Twin-20) and then incubated 1 hour with 8µl of secondary antibody goat anti-mouse IgG (Santa Cruz Biotechnology). After that time, were washed 3 times with 20ml PBST(10 minutes each wash). The signal was detected by chemiluminescence using Western Lighting® Ultra kit(Perquin Elmas)

SDS-PAGE

This method was performed to evaluate the elution profile of the his-tagged Haa1 and $Haa1_{S135F}$ proteins purified by Ni^{2+} -based Ion Metal Affinity Chromatography, overproduced in *E.coli.* For SDS-PAGE followed by protein staining, 12 µl of each protein eluted fraction were supplemented with 3ul of 5× SDS sample buffer without reducing agent, and boiled for 1 minute at 100°C. The mixture was loaded on a 8% non-reducing SDS polyacrylamide gel. PageRuler®plus PreStained Protein Ladder (Thermo Fisher Scientific) was used as ladder. The electrophoresis was run at 150 Volts for approximately 1.3 hours. The gel was stained with GelCode™ Blue Safe Stain (ThermoFisher Scientific.

Atomic force microscopy: visualization of $Haa1_{WT}$ and derived mutated $Haa1_{S135F}$ bound to *TPO3* promoter region

For preparation of *TPO3* promoter region, a 1000bp DNA fragment containing the Haa1 binding site (5′- GAGGGG-3′) was amplified from *S.cerevisiae* S288C by PCR. Each amplification by PCR was done containing 0.5 µl of genomic DNA, 5 µl of Phusion Buffer 10X (Thermo Fisher Scientific), 0.5 µl of 10mM dNTPs, 0.5 µl of each 0.1 mM primer solution, 2 µl of 50mM $MgCl_2$ (Thermo Fisher Scientific) solution, 0.75 µl of 100% DMSO (Thermo Fisher Scientific), 15 µl of RNAfree bidistillate de-ionized water and 0.25 µl of Phusion polymerase (Thermo Fisher Scientific). PCR amplification cycling parameters were the following: preincubation at 98°C for 1 minute, 35 cycles of denaturation at 98°C for 10 seconds, followed by annealing step

at 62°C for 20 seconds and an extension step at 72°C for 1 minute and 10 seconds. A final extension step was carried out at 72°C for 7 minutes. Primers used to amplify *TPO3* promoter region were 5'AGCAACATAATTGACTGACCC 3' and 5'CATTTCTGTTTATCTTTGGCTAG 3'. DNA amplification product was confirmed in a 0.8% agarose gel electrophoresis. The PCR product was gel-purified by using Zymno clean Gel DNA Recovery Kit (Zymno Research) and quantified using ImageJ software, by determination of agarose band intensity. For AFM imaging, the solutions containing DNA and Haa1 protein or the derived mutant were prepared in a ratio of 1:1 and 1:5 (DNA:protein). 1nM of dsDNA was either incubated with 2nM or 10nM of protein for 30 minutes at room temperature. Control solutions were also prepared, containing 1nM of dsDNA or 10nM of Haa1 recombinant protein. After incubation, 5mM of $MgCL_2$ solution, 10mM of HEPES (pH 6.5) were added.10 µl of the solution was adsorved on a freshly cleaved mica for 15 minutes. The drop was rinsed 7 times with 200 µl of RNAfree bidistillate de-ionized water and dried on air for approximately 10 minutes.To obtain image of the DNA-protein complexes, PicoSPM system (Molecular Imaging, Inc.) was operated in tapping mode was used for these experiments. High resolution Cr-Au cantilevers were employed for imaging in air. Cantilever oscilation frequency was tuned to the resonance frequency of the cantilever (160 KHz), with a Force Constant of 5N/m. The 512×512 resolution images were obtained with a scan rate between 0.7-1Hz and a scan range between 1-5µm.

Results

Overexpression of *HAA1* in *Pichia pastoris*

The expression levels of *HAA1* were assessed by Western blot, using an anti-His_6 tag specific antibody to confirm the presence of his-tagged Haa1 within the induced-time course samples, either in cell free extracts or in the supernatant, since it was expected that the recombinant protein would be secreted due to the presence of *S.cerevisiae* α mating sequence linked to the recombinant protein. The results are shown in figure 3 and 4. Concerning the supernatant analysis during cultivation, It was expected to see a clear signal detected by immunoassay corresponding to Haa1 protein with a molecular with approximately 80 KDa, the predicted molecular weight of the His-labelled Haa1 protein with the fused TEV cleavage sequence, assuming that the *S.cerevisiae* α mating sequence would be cleaved transversing secretion pathway. No bands with approximately 80 KDa, corresponding to the estimated size of the His-labelled Haa1 protein fused with the TEV cleavage sequence were detected in the supernatant samples obtained during the cultivation, neither at 25°C nor at 30°C induction temperatures (lanes 3-10 in Figure 3; lanes 3-7 in figure 4). Regarding the cell free extracts analysed during cultivation, contrary to our expectations, the results show two specific bands that were detected by immunoassay with approximately 70 KDa and approximately 100KDa. Results also show unspecific multiple bands with different molecular weights, apparently. The band detected by immunoassay with approximately 100KDa may presumably corresponds to the His-labelled Haa1 protein with the fused TEV cleavage sequence and with the *S.cerevisiae* α mating sequence that remained retained intracellularly, in which the α mating sequence wasn't properly processed and cleaved. The reason for the presence of a band detected by immunoassay with approximately 70 KDa remains unclear. However, it is likely that this may correspond to proteolytic degradation of the recombinant protein. Since the attempt of overproduce and secrete to the supernatant (convenient for recombinant protein recovery and purification) the his-tagged Haa1 protein using *P.pastoris* was not achieved, overproduction of the Haa1 using *E.coli* as host cell was performed.

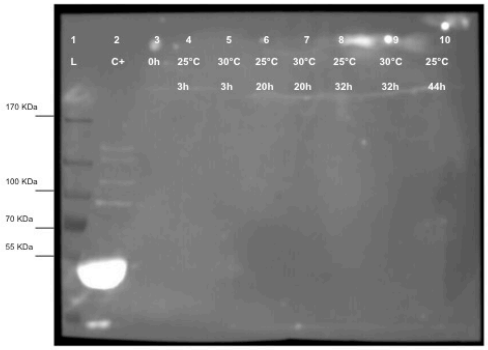

Figure 3 - Western blot analysis of the His-tag Haa1 overexpression in *P.pastoris*, using *AOX1* methanol inducible promoter. Haa1 was detected using anti-His_6 tag specific antibody. Lane 3-10 - *P.pastoris* culture supernatants, collected at different time points during cultivation after methanol induction, at either 25°C or 30°C. Lane 2 – Recombinant BceC protein containing a C-terminal $His6_6$ tag used as a positive control. . Lane 1 - PageRuler®plus PreStained (Thermo Fisher Scientific) Protein Ladder

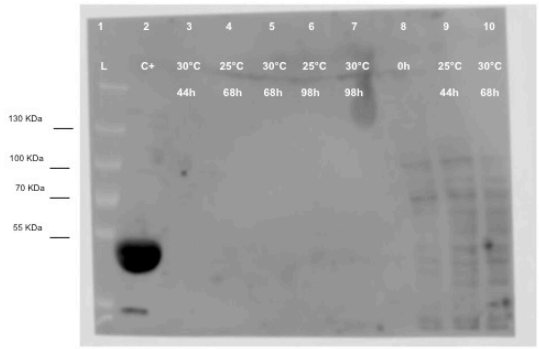

Figure 4 - Western blot analysis of the His-tag Haa1 overexpression in *P.pastoris*, using *AOX1* methanol inducible promoter. Haa1 was detected using anti-His_6 tag specific antibody. . Lane 3-7 - *P.pastoris* culture supernatants, collected at different time points during cultivation after methanol induction, at either 25°C or 30°C. Lane 8-10 - *P.pastoris* cell free extracts collected at different time points during cultivation after methanol induction, at either 25°C or 30°C. Lane 2 – Recombinant BceC protein containing a C-terminal $His6_6$ tag used as a positive control.Lane 1 - PageRuler®plus PreStained (Thermo Fisher Scientific) Protein Ladder

Overexpression in *E.coli* and purification of $Haa1_{WT}$ and the derived mutant $Haa1_{S135F}$

The overproduction of the recombinant his-tagged Haa1 and $Haa1_{S135F}$ proteins was carried out using *E.coli* as host cell, using pET expression system (Novagen). The genes to be expressed from the recombinant plasmids have a fused C-terminal 6×His-tag, allowing the desired protein purification by Ion Metal Affinity Chromatography (IMAC) Regarding Haa1 purification by IMAC, from the analysis of the elution profile of the different fractions it is possible to see that, although some protein was eluted at imidazole concentrations of 120mM (lane 5 figure 5 A) and B)), most of the recombinant protein started to be eluted at concentrations of 130mM imidazole and above (lane 6-17 figure 5 A) and B). Furthermore, co-purification of *E.coli* native contaminant proteins exhibiting affinity for Ni^{2+} is evident, particularly with imidazole elution concentrations between 120 mM and 150mM (Figure 5 A and B). To avoid the more contaminated fractions to assure higher purity, only the Haa1 protein fractions eluted with a imidazole concentration of approximately 160mM and above (lane 9-17 figure 5 A) and B). A

similar methodology was used to purify the derived Haa1 mutant by IMAC, in which the elution profile of the derived Haa1 mutant was similar to the elution profile of $Haa1_{WT}$. (data not shown)

Previous results obtained in our lab suggested that Haa1 behaves as a dimer in solution and its elution volume was established using the column, used for the FPLC polishing step. For that reason, the purified fractions of Haa1 (78KDa) were collected at an apparent elution volume equivalent to a 156 KDa molecule. However, by comparing the protein extracts of Haa1 protein before and after the FPLC polishing step, no major improvements in purification were found since protein contaminants with a lower molecular weight than the one expected for Haa1 protein (78KDa) were still visible after this step (Figure 6). Also, the amount of the recovered protein after the FPLC polishing step was lower (Figure 6). A similar methodology was used for the derived Haa1 mutant by performing a FPLC polishing step with similar results (data not shown).

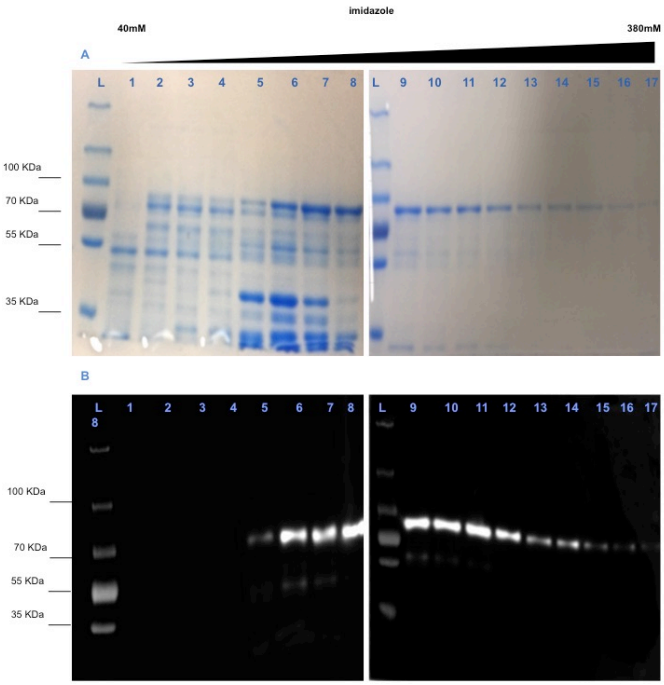

Figure 5 - Analysis of the elution profile of the recombinant his-tagged Haa1 protein, purified by Ni^{2+} -based Ion Metal Affinity Chromatography. A) SDS-PAGE analysis of the elution profile with an increased gradient of imidazole. Gel was stained with GelCode™ Blue Safe Stain (Thermo Fisher Scientific) B) Western blot analysis of the elution profile with an increased gradient of imidazole, using Anti-His_6 tag specific antibody to confirm the presence of Haa1 (with a molecular weight of 78 kDa). The fractions collected at imidazole concentrations above 160mM (lane 9 and above) were gathered and concentrated. L) - PageRuler®plus PreStained Protein Ladder (ThermoFisher Scientific) used as ladder.

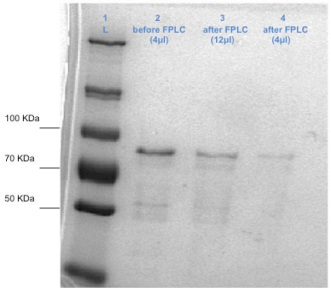

Figure 6 - An example of the various SDS-PAGE gels for separation of Haa1 protein fractions purified by Ion Metal Affinity Chromatography. Lane 2 - Purified Haa1 solution without FPLC polishing step. Lane 3 and 4 – Different volumes of the purified Haa1 solution after FPLC polishing step. L) - PageRuler®plus PreStained Protein Ladder (Thermo Fisher Scientific) used as ladder.

Atomic force microscopy: visualization of $Haa1_{WT}$ and derived mutated $Haa1_{S135F}$ bound to *TPO3* promoter DNA region

Protein:DNA complexes were selected based on a end-to-end distance of the DNA molecule with approximately 300nm of length expected for *TPO3* promoter region and the location of Haa1 bound to *TPO3* promoter region, approximately at 1/3 of the DNA chain, concomitant with the location of Haa1-responsive element between nucleotides -690 and -590 of the *TPO3* promoter.[3]

AFM images indicating the formation of $Haa1_{WT}$:DNA complexes is shown in figure 7 A) and B). In figure 7 A) the length of the DNA fragment has approximately 300nm and in figure 7 B) the DNA fragment has approximately 400nm. A globular shape presumably corresponding to the Haa1 protein is found attached to the DNA in both cases, at 1/3 of the DNA chain, at 100nm position approximately. The measured height of the attached protein to the DNA chain is approximately between 2.2nm and 3nm (Figure 7 A) and B)). Similar images indicating the formation of complexes between $Haa1_{S135F}$ and *TPO3* promoter region by AFM imaging are also shown in Figure 7 C) and D). In both, the length of the DNA fragment corresponding to *TPO3* promoter region is approximately 300nm. A globular shape corresponding to the Haa1 derived mutant is found in both images. In figure 7 C), the protein seems to be attached at 1/3 of the DNA chain, at 100nm position approximately. However, in figure 7 D) the protein seems to be attached at 1/2 of the DNA chain approximately. Also, uneven measured height variations in some regions of the DNA:protein complex are shown in figure 7 D), possibly due to the presence of holes in the mica surface near and/or above the protein:DNA complex, corresponding to black spots visible in the image. Moreover, by comparing with the $Haa1_{WT}$:DNA complexes, height measurements of the attached protein that presumably corresponds to $Haa1_{S135F}$ are lower, being approximately between 0.8nm and 1.1nm (Figure 7 C) and D)).

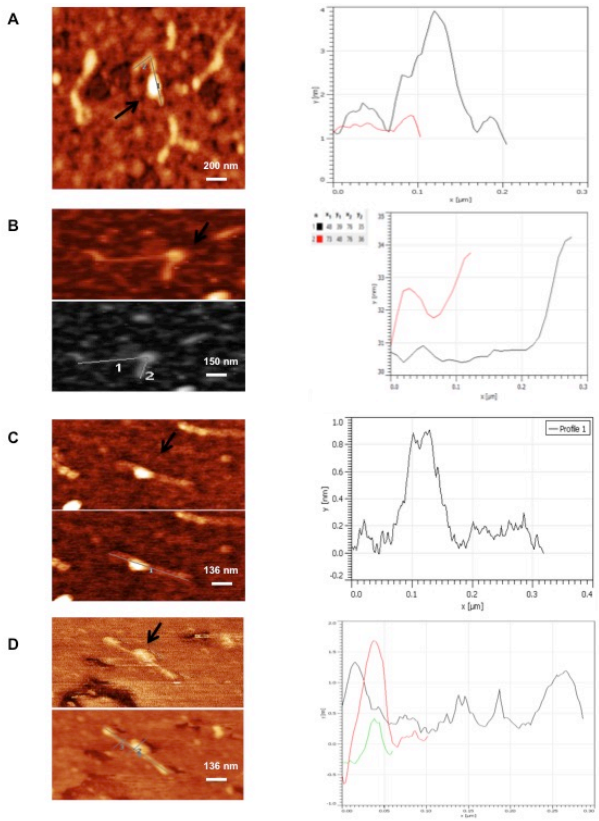

Figure 7 - Atomic Force Microscopy Imaging and processing for analysis using Gwyddion software 2.45. Protein:DNA complexes were selected based on the end-to-end distance of the DNA molecule with approximately 300nm of length expected for *TPO3* promoter region and the location of Haa1 bound to *TPO3* promoter region, approximately at 1/3 of the DNA chain, concomitant with the location of Haa1-responsive element. A and B) - AFM image of $Haa1_{WT}$:DNA complex at a 1:5 molar coupling stoichiometry and the respective measurement based on a end-to-end distance of the DNA, on the right. The putative $Haa1_{WT}$ bound to DNA has 2.2nm of height in A and 3nm of height in B. The scan size corresponds to 2μm^2 in A and 5 μm^2 in B. C and D) - AFM image of $Haa1_{S135F}$:DNA complex at a 1:5 molar coupling stoichiometry and the respective measurement based on the end-to-end distance of the DNA, on the right. The putative $Haa1_{S135F}$ bound to DNA has 0.8nm of height in C and 1.1nm of height in D. The scan size corresponds to 1μm^2. The protein:DNA complexes are indicated with a black arrow.

Discussion

The aim of the present work was to overproduce and purify the *S.cerevisiae* Haa1 transcription factor and S135F mutated form that is known to confer increased tolerance to acetic acid. The final objective to obtain a suitable amount of the protein for further functional and structural studies of this transcription factor involved in the adaptive response of yeast to acetic acid. To fulfil this objective, the first experimental strategy was to overexpress the *HAA1* gene using the commercial PichiaPink™ Expression system using *Pichia pastoris* as cell host. The attempt to overproducing and secrete the Haa1 transcription factor was not achieved It was expected that the recombinant Haa1 protein would be secreted and detected in the culture supernatant, as a 80KDa protein detected by Western blot, corresponding to the His-labelled Haa1 protein with the fused TEV cleavage sequence, this assuming that the fused *S.cerevisiae* α mating sequence would be cleaved through the secretion pathway. However, results indicate that the recombinant Haa1 transcription factor was produced and retained intracellularly and that a fraction of the retained recombinant protein was possibly degradated. Several studies suggest that one of the most critical bottlenecks that influence the productivity of an expression system is the recombinant protein translocation and secretion out of the cell. [27 28 29 30 31] The secretion limiting step of the recombinant Haa1 protein may be associated with the processing of the secretion peptide signal (the *S.cerevisiae* α mating sequence fused to the Haa1 protein, used in this work), since processing at the α mating sequence *Kex2* site is often incomplete or it does not occur, leaving a N-terminus with various portions of the α mating sequence still attached and compromising secretion out of the cell. [41 42 43 44 45] Additionally, the N-terminus of the recombinant Haa1 protein may fold on a structure that prevents processing enzymes from reaching their processing site. [23] Moreover, the fraction of the retained recombinant Haa1 protein that seems degradated may be related to the activation of Endoplasmic Reticulum Associated Protein Degradation (ERAD) pathway, since it was reported that overproduction of recombinant proteins leads to endoplasmic reticulum (ER) overloading and consequent accumulation of misfolded or unfolded proteins, triggering the Unfolded Protein Response (UPR) pathway. [46 47 48 49] As consequence, the UPR pathway triggers the ERAD pathway, in which unfolded or misfolded proteins are subjected to proteosomal degradation. [46 47 48 49] Additionally, the recombinant Haa1 may impose a significant metabolic burden during the overproduction of the recombinant Haa1 protein. Since Haa1 is a yeast transcription factor and *P.pastoris* also have a *HAA1* homologue gene, this may also contribute to the decreased expression levels observed in this study. To overcome these problems in protein overproduction some strategies could be applied. It is reported that a co-feeding of glycerol and methanol over a certain time, whereby the glycerol feeding is decreased slowly and methanol feeding is increased, can be beneficial for *P.pastoris* cells to adapt better for protein overproduction. [50 51] Also, the Haa1 transcription factor is not secreted in their native host. For this reason it could be beneficial to overproduce the recombinant Haa1 intracellularly, without the secretion leader signal peptide, to mimic a more natural environment similar to its production in their native host.

Since the attempt to overproduce extracellularly the $Haa1_{WT}$ transcription factor in *P.pastoris* was unsuccessful, the overexpression of Haa1 and the derived mutant harbouring the S135F mutation was carried out using *E.coli* as host cell using pET expression system. Both recombinant Haa1 and the derived mutant his-tagged proteins were successfully overexpressed. However, the amount of recovered protein was low, and no major improvements in purification were found using Fast Protein Liquid Chromatography (FPLC) polishing step after performing Ion Metal Affinity Chromatography (IMAC). In order to obtain more pure protein fractions and to assure higher amounts of recovered protein, polishing step must be optimized after IMAC purification. For instance, the use of heparin chromatography could be a suitable strategy. The structure and negative charge of heparin enable it to mimic DNA in its overall binding properties and it can confer higher selectivity in purification of the transcription factor Haa1 [52] It is reported that heparin is able to bind to a range of biomolecules, including transcription factors. [52 53] Remarkably, the very preliminary experiments carried out using Atomic Force Microscopy (AFM) allowed the visualization of the complexes structures, in which Haa1 transcription factor is apparently bound to the *TPO3* promoter region (1000bp upstream the *TPO3* start codon), with approximately 300nm of length. Apparently, similar complex structures were also visualized between the Haa1 derived mutant and *TPO3* promoter region. Interestingly, by focusing on the position distribution of the protein on the DNA sequence, it is possible to hypothesise a

specific recognition site for this protein in the *TPO3* promoter region. In fact, the protein:DNA complexes observed mostly appears to be located at 1/3 of the DNA chain, consistent with the location of the Haa1-responsive element (HRE) between nucleotides -690 and -590 of *TPO3* promoter.[3] Another interesting indication was obtained by comparing of the results from height measurements between the putative $Haa1_{WT}$:DNA complexes and the putative $Haa1_{S135F}$:DNA complexes, since the differences in height protein preliminary measurements between the Haa1 and the derived mutant S135F bound to DNA appear to suggest a conformational change upon binding. Proteins and DNA can undergo conformational changes in order to form functional complexes to carry out the required function. [54 55 56] Co-relations, using AFM approach, between conformational changes and the average protein height measurement of bound and unbound forms, were already been reported.[57] The Haa1 derived mutant S135F harbours a single point mutation by substitution of a serine for a phenylalanine in 404 nucleotide position, near the DNA binding domain, and leading to an increase in acetic acid tolerance. Single point mutations, such as the S135F mutation in the Haa1 protein, may be crucial for local conformational changes that may impact protein function. [58] Hydrophobic and aromatic residues were found to be critical for trans-activation domain target contacts, which are essential for efficient transactivation activity. [59 60 61 62 63 64] Moreover, the S135F substitution may be also affecting the trans activation potential of Haa1, for instance by improving the recruitment of RNA transcription machinery. [65] Additionally, nonpolar residues with an aromatic nature such as phenylalanine can contribute to protein function by forming hydrophobic cores that are relevant for global protein, geometrical compatibility and efficient packing, folding and stability. [55 66 67 68] Given that the Haa1 protein is a multi-phosphorylated protein (Sugiyama *et al,. 2014,* phosphoGRID), the elimination of a serine residue that can be prone to phosphorylation and its impact during the binding and/or in the stability of Haa1 protein:DNA complexes remains to be determined. Also, additional experiments to increase the number of Haa1 protein:DNA complexes found must be carried out in order to ensure reproducibility of the results obtained from AFM.

Bibliographic references

1. Ruis, H. & Schüller, C. Stress signaling in yeast. *BioEssays* 17, (1995).
2. Takagi, H., Kitagaki, H. & Brewing, I. *Stress Biology of Yeasts and Fungi*. (Springer Berlin Heidelberg, 2015).
3. Henriques, F. *et al.* Identification of a DNA-binding site for the transcription factor Haa1 , required for *Saccharomyces cerevisiae* response to acetic acid stress. 39, 6896–6907 (2011).
4. Mollapour, M. & Piper, P. W. Weak Organic Acid Resistance of Spoilage Yeasts. *Elsevier Ltd* 143–155 (2008).
5. Allen, I., Laskin, S. . & Geoffrey, M. *Advances in Applied Microbiology 77*. (Elsevier Inc., 2011).
6. Fleet, G. Spoilage yeasts. *Crit. Rev. Biotechnol* 12, 1–44 (1992).
7. Mollapour, M. & Piper, P. W. Hog1p MAP kinase determines acetic acid resistance in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 6, 1274–1280 (2006).
8. Mollapour, M. & Piper, P. Hog1 mitogen-activated protein kinase phosphorylation targets the yeast Fps1 aquaglyceroporin for endocytosis, thereby rendering cells resistant to acetic acid. *Mol Cell Biol* 27, 6446–6456 (2007).
9. Stratford, M. *et al.* International Journal of Food Microbiology Extreme resistance to weak-acid preservatives in the spoilage yeast *Zygosaccharomyces bailii* ☆. *Int. J. Food Microbiol.* 166, 126–134 (2013).
10. Mira, N. P., Teixeira, M. C. & Sa, I. Adaptive Response and Tolerance to Weak Acids in *Saccharomyces cerevisiae* : A Genome-Wide View. *OMICS* 14, 525–540 (2010).
11. Makrantoni, V., Dennison, P., Stark, M. & Coote, P. a novel role for the yeast protein kinase Dbf2p in vacuolar H+-ATPase function and sorbic acid stress tolerance. *Microbiology* 12, 4016–4026 (2007).
12. Mira, N. P., Becker, J. D. & Sa, I. Genomic Expression Program Involving the Haa1p-Regulon in *Saccharomyces cerevisiae* Response to Acetic Acid. *Omi. A J. Integr. Biol.* 14, (2010).
13. Fernandes, A., Mira, N., Vargas, R., Canelhas, I. & Sa-Correia, I. Saccharomyces cerevisiae adaptation to weak acids involves the transcription factor Haa1p and Haa1p-regulated genes. *Biochem Biophys Res Commun* 337, 95–103 (2005).
14. Mira, N. P., Palma, M., Guerreiro, J. F. & Sá-correia, I. Genome-wide identification of *Saccharomyces cerevisiae* genes required for tolerance to acetic acid. *Microb. Cell Fact.* 9, 1–13 (2010).
15. Mira, N. P., Cacho Teixeira, M. & Sá-Correia, I. Adaptive Response and Tolerance to Weak Acids.pdf. *Omics* 14, (2010).
16. Yeheskely-Hayon, D. *et al.* The Roles of the Catalytic and Noncatalytic Activities of Rpd3L and Rpd3S in the Regulation of Gene Transcription in Yeast. *PLoS One* 8, (2013).
17. Piper, P. W. et al. Hsp30, the integral plasma membrane heat shock protein of *Saccharomyces cerevisiae*, is a stress-inducible regulator of plasma membrane ATPase. *Cell Stress Chaperones* (1997).
18. Simões, T., Mira, N., Fernandes, A. & Sá-correia, I. The SPI1 gene, encoding a glycosylphosphatidylinositol-anchored cell wall protein, plays a prominent role in the development of yeast resistance to lipophilic weak-acid food preservatives. *Appl Env. Microbiol* 18, 7168–7165 (2006).
19. Destruelle, M., Holzer, H. & DJ, K. Identification and characterization of a novel yeast gene: the YGP1 gene product is a highly glycosylated secreted protein that is synthesized in response to nutrient limitation. *Mol Cell Biol.* 4, 2740–2754 (1994).
20. Minetaka, S. *et al.* Nuclear localization of Haa1, which is linked to its phosphorylation status, mediates lactic acid tolerance in Saccharomyces cerevisiae. *Appl. Environ. Microbiol* (2014). doi:10.1128/AEM.04241-13
21. Burgess, R. R. & Deutscher, M. P. *Methods in Enzymology : Guide to protein purification*. (2009).
22. Technologies, L. *PichiaPink ™ Expression System Guide*. (2014).
23. Cregg, J. M. *Pichia Protocols*. 389, (Humana Press, 2007).
24. Lin-Cereghino, G. ., Leung, W. . & Lin-Cereghino, J. *Expression of protein in Pichia pastoris.* (Methods Express. Scion Publishing Limited; Oxfordshire, 2007).
25. Cregg JM, Vedvick TS, R. W. Recent advances in the expression of foreign genes in *Pichia pastoris.* BioTechnology 11: 905–910. *BioTechnology* 905–910 (1993).
26. Puxbaum, V., Gasser, B. & Mattanovich, D. The bud tip is the cellular hot spot of protein secretion in yeasts. *Appl Microbiol Biotechnol* 100, 8159–8168 (2016).
27. Briand, L. *et al.* Optimization of the production of a honeybee odorant-binding protein by *Pichia pastoris. Protein Expr. Purif.* 15, 362–369 (1999).
28. Koganesawa, N. *et al.* Construction of an expression system of insect lysozyme lacking thermal stability: the effect of selection of signal sequence on level of expression in the *Pichia pastoris* expression system. *Protein Eng* 14, 705–710 (2001).
29. Kowalski, J., Parekh, R. N., Mao, J. & Wittrup, K. D. Protein folding stability can determine the efficiency of escape from endoplasmic reticulum quality control. *Biol. Chem.* 273, 19453–19458. (1998).
30. Agaphonov, M. O., Romanova, N. V., Trushkina, P. M., Smirnov, V. N. & Ter-Avanesyan, M. D. Aggregation and retention of human urokinase type plasminogen activator in the yeast endoplasmic reticulum. BMC Mol. Biol. 3, 15. *BMC Mol. Biol.* 3, (2002).
31. Mattanovich, D., Gasser, B., Hohenblum, H. & Sauer, M. Stress in recombinant protein producing yeasts. 113, 121–135 (2004).
32. *pET System Manual Novagen pET System Manual Novagen*. (2003).
33. Hinterdorfer, P. & Oijen, A. *Handbook of Single-Molecule Biophsics*. (Springer Science+Business Media, 2009).

34. Jalili, N. & Laxminarayana, K. A review of atomic force microscopy imaging systems : application to molecular metrology and biological sciences. 14, 907–945 (2004).
35. Lyubchenko, Y. L., Shlyakhtenko, L. S. & Gall, A. A. *Methods in Molecular Biology, DNA -Protein Interactions*. 543, (Humana Press).
36. Morris, V. J., Kirby, A. R. & Gunning, A. P. *Atomic Force Microscopy for Biologists*. (Imperial College Press, 2009).
37. Chang, K., Chiang, Y., Yang, C. & Liou, J. Atomic force microscopy in biology and biomedicine. *Tzu Chi Med. J.* 24, 162–169 (2012).
38. Alsteens, D. & Dufre, Y. F. Atomic force microscopy – looking at mechanosensors on the cell surface. (2009). doi:10.1242/jcs.106005
39. Francois, J. M. *et al.* Use of atomic force microscopy (AFM) to explore cell wall properties and response to stress in the yeast *Saccharomyces cerevisiae*. *Curr Genet* 59, 187–196 (2013).
40. Schiavone, M., Formosa-dague, C., Elsztein, C. & Teste, M. An Atomic Force Microscopy study of yeast response to ethanol stress : Evidence for a role of the plasma membrane in the nanomechanical properties of the cell walls. *Am. Soc. Microbiol*. 33, (2016).
41. Waters, M., Evans, E. & Blobel, G. Prepro-alpha-factor has a cleavable signal sequence. *J Biol Chem* 263, 6209–6214 (1988).
42. Julius, D., Brake, A., Blair, L., Kunisawa, R. & J, T. Isolation of the putative structural gene for the lysine–arginine-cleaving endopepti- dase required for processing of yeast prepro-alpha-factor. *Cell* 37, 1075–1089 (1984).
43. Brake, A. *et al.* Alpha- factor-directed synthesis and secretion of mature foreign proteins in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 81, 4642–4646 (1984).
44. Kurjan, J. & Herskowitz, I. Structure of a yeast pheromone gene (MF alpha): a putative alpha-factor precursor contains four tandem copies of mature alpha-factor. *Cell* 30, 933–43 (1982).
45. Hegde, R. & Bernstein, H. The surprising complexity of signal sequences. *Trends Biochem Sci.* 31, 563–71 (2006).
46. Hoseki, J., Ushioda, R. & Nagata, K. Mechanism and components of endoplasmic reticulum-associated degradation. *J Biochem* 147, 19–25 (2010).
47. Kohno, K. Stress-sensing mechanisms in the unfolded protein response: similarities and differences between yeast and mammals. *J Biochem* 147, 27–33 (2010).
48. Malhotra, J. & Kaufman, R. The endoplasmic reticulum and the unfolded protein response. *Semin Cell Dev Biol* 18, 716–731 (2007).
49. Yu, X., Sha, C., Guo, Y., Xiao, R. & Y, X. High-level expression and characterization of a chimeric lipase from *Rhizopus oryzae* for biodiesel production. *Biotechnol Biofuels* 6, (2013).
50. Minning, S. *et al.* Optimization of the high-level production of Rhizopus oryzae lipase in *Pichia pastoris*. *J Biotechnol* 1, 59–70 (2001).
51. Zhang, W. & Meagher, M. M. Fermentation Strategies for Recombinant Protein Expression in the Methylotrophic Yeast *Pichia pastoris*. *Biotechnol. Bioprocess Eng.* 5, 275–287 (2000).
52. Posch, A. *Methods in Molecular Biology 2D PAGE: Sample Preparation and Fractionation*. 1, (Humana Press, 2008).
53. Hausner, W. & Thomm, M. Purification and Characterization of a General Transcription Factor, aTFB, from the Archaeon *Methanococcus thermolithotrophicus*. *J. Biol. Chem.* 268, 24047–52 (1993).
54. Andrabi, M., Mizuguchi, K. & Ahmad, S. Conformational changes in DNA-binding proteins: Relationships with precomplex features and contributions to specificity and stability Munazah. *Wiley Period. Inc.* 82, 841–857 (2013).
55. Dickerson, R. DNA bending: the prevalence of kinkiness and the virtues of normality. *Nucleic Acids Res* 26, 1906–1926 (1998).
56. Matsumoto, A. & Olson, W. Sequence-dependent motions of DNA: a normal mode analysis at the base-pair level. *Biophys J* 83, 22–41 (2002).
57. Trajkovic, S., Zhang, X. & Sylvia, D. Atomic Force Microscopy Study of Conformational Change of Immobilized Calmodulin. *Natl. Institutes Heal.* 27, 10793–10799 (2012).
58. Schaefer, C. & Rost, B. Predict impact of single amino acid change upon protein structure. *BMC Genomics* 13, S4 (2012).
59. Landrieu, I., Verger, A., Baert, J., Rucktooa, P. & Ferreira, E. Characterization of ERM transactivation domain binding to the ACID / PTOV domain of the Mediator subunit MED25. *Nucleic Acids Res.* 1–12 (2015). doi:10.1093/nar/gkv650
60. Brzovic, P. S. *et al.* The acidic transcription activator gcn4 binds the mediator subunit gal11/med15 using a simple protein interface forming a fuzzy complex. *Mol. Cell* 44, 942–953 (2011).
61. Chabot, P. R. *et al.* Structural and functional characterization of a complex between the acidic transactivation domain of EBNA2 and the Tfb1/p62 subunit of TFIIH. *PLoS Pathog.* (2014).
62. Jennings, B. H., Pickles, L. M., Wainwright, S. M., Roe, S. M. & Pearl, L.H. and Ish-Horowicz, D. Molecular recognition of transcriptional repressor motifs by the WD domain of the Groucho/TLE corepressor. *Mol. Cell,* 22, (2006).
63. Langlois, C., , Mas, C., Di Lello, P., Jenkins, L. M. & Legault, P. and Omichinski, J. G. NMR structure of the complex between the Tfb1 subunit of TFIIH and the activation domain of VP16: structural similarities between VP16 and p53. *J. Am. Chem. Soc.,* 130, 10596–10604. (2008).
64. Sullivan, S. M. *et al*. Mutational analysis of a transcriptional activation region of the VP16 protein of herpes simplex virus. *Nucleic Acids Res* 26, 4487–4496 (1998).
65. Swinnen, S. *et al.* Improvement of yeast tolerance to acetic acid through Haa1 transcription factor engineering: towards the underlying mechanisms (submitted).
66. Ahmad, S., Keskin, O., Sarai, A. & Nussinov, R. Protein – DNA interactions : structural , thermodynamic and clustering patterns of conserved residues in DNA-binding proteins. 36, 5922–5932 (2008).
67. Keskin, O., Ma, B., Nussinov, R. & Aviv, T. Hot Regions in Protein – Protein Interactions : The Organization and Contribution of Structurally Conserved Hot Spot Residues. *J. Mol. Biol.* 345, 1281–1294 (2005).
68. Bhardwaj, N. & Gerstein, M. Relating protein conformational changes to packing efficiency and disorder. 18, 1230–1240 (2009).