Biotechnology Reports 38 (2023) e00801

Contents lists available at ScienceDirect

Biotechnology Reports

journal homepage: www.elsevier.com/locate/btre

Research Article

Molecular background of HAC1-driven improvement in the secretion of recombinant protein in *Yarrowia lipolytica* based on comparative transcriptomics

Paulina Korpys-Woźniak, Ewelina Celińska *

Department of Biotechnology and Food Microbiology, Poznan University of Life Sciences, ul. Wojska Polskiego 48, 60-637 Poznań, Poland

ARTICLE INFO

Keywords:
Secretion
Heterologous protein
Yeast
Endoplasmic reticulum (ER) stress
Recombinant protein production

ABSTRACT

While the unfolded protein response (UPR) and its major regulator – transcription factor Hac1 are well-conserved across *Eukarya*, species-specific variations are repeatedly reported. Here we investigated molecular mechanisms by which co-over-expression of *HAC1* improves secretion of a recombinant protein (r-Prot) in *Yarrowia lipolytica*, using comparative transcriptomics. Co-over-expression of *HAC1* caused an >2-fold increase in secreted r-Prot, but its intracellular levels were decreased. The unconventional splicing rate of the *HAC1* mRNA was counted through transcript sequencing. Multiple biological processes were affected in the *HAC1*-and-r-Prot co-over-expressing strain, including ribosome biogenesis, nuclear and mitochondrial events, cell cycle arrest, attenuation of gene expression by RNA polymerase III and II, as well as modulation of proteolysis and RNA metabolism; but whether the *HAC1* co-over-expression/induction was the actual causative agent for these changes, was not always clear. We settled that the expression of the "conventional" *HAC1* targets (*KAR2* and *PDI1*) is not affected by its over-expression.

1. Introduction

Research on yeast protein synthesis and secretion is relevant to basic and applied studies. Knowledge of the molecular background of this biological process provides direct indications for its improvement through molecular and bioprocessing approaches. In our previous studies, we investigated the global transcriptome of *Yarrowia lipolytica* upon the over-synthesis of several biochemically different recombinant (secretory) proteins (r(*s*)-Prots) [1]. Both general and protein type-specific molecular responses were noted. Knowledge gained from that study prompted the design of genetic engineering strategies enhancing the production of rs-Prots in this species [2]. Genes that were significantly deregulated due to r(*s*)-Prot over-synthesis were used as "secretory helpers" assisting the synthesis and secretion of rs-Prot. In both those previous studies, we paid particular attention to the expression pattern of a key gene involved in the modulation of protein synthesis and secretion, namely *HAC1* (*YALI0B12716g*; hereafter all genes names are abbreviated – without *YALI0* prefix) and its regulome.

HAC1 encodes a transcription factor mediating unfolded protein response (UPR), which is triggered by the accumulation of incorrectly folded polypeptides in the over-loaded endoplasmic reticulum (ER) [3–7]. The biological sense of UPR is to restore cellular homeostasis by debottlenecking the synthesis-folding-secretion pathway. Hac1 mediates the deregulation of hundreds of genes involved in polypeptide formation, folding, and maturation, but also in lipid synthesis, membrane expansion, and many others [4,8]. It is thus one of the most frequently targeted genes to improve the production of rs-Prots in yeasts [4,7,9–15], recently reviewed in [16]. *HAC1* homologs were described in different yeast species, and details of their operation in *Saccharomyces cerevisiae* and several other species are well-known [4,5,7,9,17]. While the Hac1-mediated UPR pathway is phylogenetically conserved in *Eukarya*, some variation, even amongst different yeast species, occurs [4,10,18–20]. The variation relates to the structure of the *HAC1* transcript, its expression, and splicing pattern, as well as the sequence of DNA binding motif, and the Hac1-driven regulome [4,10,16,18]. Hence, it becomes obvious that for the successful engineering of rs-Prot production in a given yeast species, more insight into the specificities of the UPR operation is needed.

In our previous studies on rs-Prot synthesis and secretion in *Y. lipolytica,* we found that *HAC1* up-regulation was dependent on biochemical characteristics of r(*s*)-Prot synthesized, and the rate of its synthesis [1]. High-level over-synthesis of a small secretory fluorescent

* Corresponding author.
E-mail address: ewelina.celinska@up.poznan.pl (E. Celińska).

https://doi.org/10.1016/j.btre.2023.e00801
Received 10 April 2023; Accepted 7 May 2023
Available online 8 May 2023

protein (sc-YFP) or highly-disulfide-bonded amylase did not lead to the induction of *HAC1* expression. In contrast, the formation of "difficult-to-synthesize" highly-glycosylated secretory glucoamylase (TlG) and intracellular YFP (in-YFP) triggered *HAC1* up-regulation and initiation of a spectrum of downstream effects. Notably, it was surprising to see that the intracellular r-Prot (in-YFP) over-synthesis triggered the up-regulation of *HAC1* and the following downstream events, which are typically associated with ER compartment. But the transcriptomics profiles induced by the rs-Prot (TlG) and r-Prot (in-YFP) over-expression were highly correlated. In addition, the amounts of synthesized r(*s*)-Prots were generally lower; which was associated with the induction of massive vacuolar proteolysis. Such an observation could incorrectly suggest that up-regulation of *HAC1* did not bring any improvement in the rs-Prots synthesis in *Y. lipolytica*. However, in the following study, we evidenced that co-over-expression of *HAC1* with a secretory YFP (*SCYFP* gene) [2] exerted a positive effect on the rs-Prot productivity, increasing it by 30%.

It was surprising to see that the *HAC1* up-regulation, either caused by the accumulation of incorrectly folded polypeptides [1] or by over-expression from a synthetic construct [2], did not induce expression of its known direct downstream targets, *i.a.* protein disulfide isomerase *PDI1* (*E03036g*) and a multifunctional chaperone *KAR2* (*E13706g*). Previous studies conducted with *Pichia pastoris* and *S. cerevisiae* [5,11,18,21] reported that up-regulation of *HAC1* was accompanied by enhanced expression of these downstream targets. Moreover, upregulation of *KAR2* gene expression was also seen previously in *Y. lipolytica* upon overexpression of a spliced variant of *HAC1*, as assayed by real-time qPCR analysis [19]. We previously speculated that in the case of *ylHAC1* co-over-expression, the outcome was a consequence of using a native, non-spliced variant of the *HAC1* gene [2], which accumulated without being fully processed. For being translated, *HAC1* requires unconventional splicing by dimerized protein Ire1 (*A14839g*), which is a sensor of unfolded polypeptides and a nuclease. Under balanced conditions, Ire1 is bound to Kar2 and is not available for the *HAC1* transcript processing. Accumulation of incorrectly folded polypeptides causes Kar2 dissociation from its interaction partner and releases Ire1 for mRNA splicing. The key role of nucleotide exchange factor Sls1 (*E32703g*) in the Kar2-Ire1 interaction in *Y. lipolytica* was evidenced [22–24] and later observed in the context of the global transcriptome [1]. Nonetheless, the lack of transcriptional response from *KAR2* and *PDI1* was observed in both cases – (i) when the native *HAC1* gene was synthetically over-expressed and thus could accumulate due to insufficient processing, and (ii) when it was natively up-regulated and presumably fully processed due to complete onset of UPR. Hence, the question of the specificity of *HAC1* operation in *Y. lipolytica* remained open.

This study aimed to further characterize the secretory pathway in *Y. lipolytica* with a particular focus on *HAC1* operation. To this end, we conducted comparative transcriptomics analysis of steady-state-maintained *Y. lipolytica* cells: (i) over-expressing native *HAC1* and sc-YFP reporter protein from synthetic constructs, (ii) over-synthetizing sc-YFP at high levels, and (iii) subjected to (cytosolic) UPR and up-regulation of *HAC1* triggered by over-synthesis of intracellular reporter protein (in-YFP). Using such an approach we were able to dissect the effects triggered by different conditions (*HAC1* / high-level synthesis of rs-Prot) and evaluate their practical consequences in terms of rs-Prot production.

2. Materials and methods

2.1. Yeast strains

All the strains analyzed in this study were constructed previously [1,2,25]. The strains are listed in Supplementary Table S1. *Y. lipolytica HAC1_scYFP* (ECY212.3) strain was constructed on a platform of Po1f strain (*MatA, leu2–270, ura3–302, xpr2–322, axp-2*) by co-over-expressing: (i) a secretory yellow fluorescence protein (*SCYFP*), and (ii) a native bZIP transcription factor *HAC1* (*B12716g*) (Supplementary Table S2). The genes were cloned in JMP62-URA3ex and JMP62-LEU2ex vectors (Supplementary Fig. S1), respectively, according to a previously described methodology [26]. *Y. lipolytica* strains GGY251 and GGY249 are Po1f-derivates co-over-expressing a YFP reporter in secreted (sc-YFP) and intracellular form (in-YFP), respectively [25]. For secretion, the YFP-encoding gene was transcriptionally fused with a signal sequence SP1, native for Exo-1,3-beta-glucanase (*B03564g*) [27,28]. The reference strain, GGY231, is a prototrophic version of Po1f (not over-expressing *YFP*). The *scYFP, inYFP,* and the reference strain were constructed via the Golden Gate strategy [29–31]. All the strains were maintained in glycerol stocks at -80 °C, and revived by plating on YPD agar ([$g\ L^{-1}$]: yeast extract (BTL, Poland), 10; bactopeptone (BTL), 20; glucose (POCh, Poland), 20; agar (BTL, Poland), 15) and incubation at 30 °C. Visualization of the strains' growth on agar plates and microscopic images are provided in Supplementary Fig. S2.

2.2. Chemostat cultivation

Data on growth, r(*s*)-Prot synthesis, and carbon utilization by *scYFP* and *inYFP* strains were obtained in the previous study [1]. Bioreactor cultivations of the *HAC1_scYFP* strain were executed according to a previously described methodology [1,25]. Briefly, pre-cultures were developed in shake flasks in 30 mL YPG_{20} medium ([$g\ L^{-1}$]: yeast extract, 10; bactopeptone, 20; glycerol (POCh, Poland), 20), at 28 °C, agitation at 250 rpm for 22 h. Continuous cultures were conducted in 0.5 L stirred tank bioreactors Multifors 2 (Infors HT, Bottmingen Basel, Switzerland), in 0.3 L of YPG_{100} medium ([$g\ L^{-1}$]: yeast extract, 10; bactopeptone, 20; glycerol, 100), inoculated at 10%. The conditions were monitored and controlled at the following set points: temperature, 28 °C; pH, 5.5 (correction by: 20% NaOH, 10% H_2SO_4); airflow, 2.0 vvm; pO_2 was kept at 20% by setting a cascade with stirring (100 to 1200 rpm). Foaming was minimized by the automatic addition of a defoaming agent AntiFoam 204 (Merck/Sigma-Aldrich; USA). The initial batch mode culture was continued until full consumption of glycerol. Feeding was initiated at a dilution rate [D] of 0.20 [h^{-1}]. A steady state was reached after at least 6 residence times [τ]. The cultures were conducted in biological duplicates.

2.3. Biomass concentration measurements

Yeast biomass concentration was analyzed by measuring optical density at 600 nm (OD 600 nm). Prior to the measurement the biomass was washed twice and diluted in a sterile saline solution (0.85% NaCl). OD600 measurements were done using flat-bottomed 96-well microplates (Corning; Sigma-Aldrich, USA) in Tecan Infinite M200 automatic plate reader (TECAN, Austria). All the measurements were conducted in technical duplicates of each biological replicate.

2.4. Fluorometry

Fluorescence (FL) of intracellular (in-YFP) and secreted (sc-YFP) fraction of the YFP reporter protein in the *HAC1_scYFP* cells was analyzed in flat-bottomed 96-well microplates (Corning; Sigma-Aldrich) in Tecan Infinite M200 automatic plate reader at the wavelengths settings (excitation/emission) 495/527 nm. The FL of sc-YFP was measured in 200 µl of biomass-devoid supernatant. FL of in-YFP was measured in 200 µl of pre-washed culture biomass, re-suspended, and diluted in the saline solution. FL measurements were normalized vs background fluorescence: either biomass of the control strain for in-YFP or YPG medium for sc-YFP. FL results were expressed as: (i) in-YFP – raw FL readout of the washed biomass in FL units [RFU – Relative Fluorescence Unit], (ii) sp_in-YFP - specific FL of the washed biomass in [RFU OD600 nm^{-1}], (iii) sc-YFP – raw FL readout of the supernatant in [RFU], (iv) sp_sc-YFP - specific FL of the supernatant in [RFU OD600 nm^{-1}]. All the

measurements were conducted in technical duplicates of each biological replicate. Additionally, the FL results were randomly verified through fluorescence microscopy observations (Zeiss, Axiovert 200).

2.5. High-performance liquid chromatography

The concentration of glycerol and typical products of *Y. lipolytica* metabolism (mannitol, erythritol, citric acid, and α-ketoglutaric acid) was measured by High-Performance Liquid Chromatography (HPLC), as described previously [25]. Briefly, the analysis was executed using Agilent Technologies 1200 series (Agilent Technologies, Santa Clara, USA), equipped with a refractive-index detector (G1362A) and a Rezex ROA-Organic Acid *H*+ column (Phenomenex, Torrance, USA), and 0.005 N H2SO4 as eluent at a flow rate of 0.6 [mL min^{-1}], at 40 °C.

2.6. Protein isolation and Protease activity assay

2.6.1. Total protein isolation and BCA assay

Total protein extraction was conducted according to the previously described protocol [32]. Briefly, cellular pellets were resuspended in ice-cold breaking buffer (0.1 M sodium phosphate buffer pH 7.4, 5 μM DTT, 5% GLY). The cells were disrupted by glass beads (Sigma-Aldrich, St. Louis, USA) under repeated cycles (5x) of mixing at 30 strokes s^{-1} for 30 s in a MixerMill MM400 (Retsch GmbH, Haan, Germany) and incubation on ice for 1 min. The cellular debris was then separated by centrifugation (24,652 xg, 4 °C, 10 min; Eppendorf 5430 R; Eppendorf, Hamburg, Germany). The cellular extracts were subjected to buffer exchange by ultrafiltration through a 3 kDa MWCO point cellulose membrane (Amicon Ultra; Merck, Darmstadt, Germany).

Total protein concentration in the obtained extracts was determined using a commercial BCA (Bicinchoninic Acid) Protein Assay Kit (Merck) in a microtiter plate format. A standard curve was prepared using a BSA standard solution. The reaction was conducted according to the manufacturer's protocol. Reading of the assay colorimetric was conducted in transparent flat-bottomed 96-well microplates (Corning; Sigma-Aldrich) using a Multiskan SkyHigh Plate Reader - UV/Vis Spectrophotometer at 562 nm wavelength (ThermoFisher Scientific, Waltham, USA).

The assay was conducted in two independent runs and a technical triplicate.

2.6.2. Protease activity assay with FTC-casein

Proteolytic activity in the total protein extracts was examined using a commercial Pierce™ Fluorescent Protease Assay Kit (ThermoFisher Scientific), relying on FTC-casein substrate digestion and its fluorescence. The protein extracts were subjected to the reaction in a microtiter plate format. A standard curve was prepared using a standardized solution of trypsin. Fluorescence reading was conducted in white opaque plates, as recommended by the kit's manufacturer, using a Tecan Infinite M200 automatic plate reader at the wavelength settings (excitation/emission) 485/538 nm.

The assay was conducted in two independent runs and a technical triplicate.

2.6.3. Statistical analysis

Initial data processing and their graphical representation were done using Microsoft Excel and Data Analysis plugin. The statistical importance of the differences between compared sets of data was analyzed using two-way analysis of variance (ANOVA) and Tukey's HSD multiple comparison tests using the STATISTICA software (StatSoft, Inc., Tulsa, USA). Datasets were initially tested by Kruskal-Wallis and Levene's tests to verify ANOVA's presumptions. The results were considered statistically significant at a p-value < 0.05. The results were expressed as mean ± standard deviation (±SD) of the replicates.

2.7. RNA extraction and NGS sequencing

Steady-state-maintained yeast cells were harvested and processed according to a Bead-Beat Total RNA Mini Kit protocol (A&A Biotechnology, Gdynia, Poland) to extract total RNA. The quality of RNA was verified through agarose gel electrophoresis [33] and spectrophotometry (NanoDrop, Thermo Fisher Scientific). Library construction and NGS RNA sequencing were executed using NovaSeq 6000 (Illumina) apparatus at Genomed S.A. (Warsaw, Poland). Raw sequencing reads are available on the NCBI Sequence Read Archive under the BioProject number PRJNA869113 (see Data Availability section).

2.7.1. RNAseq data analysis

2.7.1.1. Raw data processing and data filtration. Cutadapt [34] was used for NGS adapters removal. Initial filtering was done using the *q25* quality parameter and *m15* the minimum length of the reads, followed by re-filtering to remove reads <20 bp. Quality control was conducted using FastQC (2010; www.bioinformatics.babraham.ac.uk). Post-QC data were aligned and mapped to the *Y. lipolytica* CLIB122 reference genome (GenBank assembly identifier: GCA_000002525.2) using Hisat2 [35]. Reading pairs were counted using the HTseq program [36]. Primary Gene Ontology (GO) annotation was conducted using the reference genome template (GCF_000002525.2; manually curated). The final results were processed in R (www.rproject.org) using the DESeq2 package [37]. All Differentially Expressed Genes (DEGs) met the requirement of FDR ≤ 0.05 (adjusted p-value) tested *via the* Wald test. Only DEGs meeting all those criteria were considered in further analyses. Gene name and function annotation was manually curated by cross-referencing several databases – Panther, UniProt, GRYC, and NCBI and blasting (blastp) the sequences against the database's collections. The most relevant matches were chosen to finally describe a given gene. A complete list of all DEGs is given in Supplementary Table S3. Relative gene expression level is expressed as fold change (FC) over the control prototrophic strain.

2.7.1.2. Enrichment / overrepresentation tests. Statistical overrepresentation test and Enrichment Test were conducted using PANTHER v14.0 online tool and database [38,39]. The initial GO terms annotation by DESeq2 and the reference genome was followed by manual curation using PANTHER resources and tools. The biological function was assigned to each DEG according to PANTER's GO Ontology database DOI: 10.5281/zenodo.4437524 Released 2021- 01–01. Overrepresentation of specific biological processes, molecular functions, and cellular components in a subset of DEGs was tested using GO complete annotation data sets, Fisher's exact test, and False Discovery Rate for multiple testing. Only results with $p < 0.05$ were considered significant and displayed.

2.7.1.3. Comparative analysis through Venn diagrams. Prior to the Venn diagram analysis, DEGs predicted by DESeq2 were filtered for $\log_2$-fold change >1.0; only such DEGs were considered in this analysis. Similarities and dissimilarities in DEGs subsets were searched using InteractiVenn online tool; graphics were adjusted using Inkscape (https://inkscape.org/release/inkscape-1.0.2/).

2.7.1.4. Comparative analysis – heat-map generation. Heat-maps of DEGs in the studied strains were prepared using Morpheus online tool (https://software.broadinstitute. org/morpheus).

2.7.1.5. Quantitation of spliced-to-unspliced HAC1 transcript ratio. As the presence of an intron site was not described in the reference genome used, the intron existence and location within the *B12716g* locus had to be identified de novo. The detection and visualization of the potential splice sites were conducted using the SGSeq package [40] in the R

programming language. Mapping of exons and introns in the *B12716g* gene was conducted in the IGV program [41]. Coverage depth within the gene fragments (intron and exons) was used as an indicator of actual splicing events. Coverage was calculated in fragments: intron (NC_006068.1:1702623–1702651), exons (NC_006068.1:1702612–1702622, NC_006068.1:1702652–1702662). The average coverage depth was calculated by SamTools [42].

2.8. Basic bioinformatics

DNA and protein sequences were retrieved from NCBI's GenBank, Nucleotide or Protein databases (https://www.ncbi.nlm.nih.gov/), UniProt (https://www.uniprot.org/), or GRYC (http://gryc.inra.fr/). Sequence analysis and primer design were done in Benchling (https://www.benchling.com/). Multiple sequence alignments were done using BioEdit [43] or MultAlin [44].

3. Results and discussion

The principal aim of this study was to investigate in more detail the *HAC1* gene operation in *Y. lipolytica* and its effect on r(*s*)-Prot synthesis, as well as to verify the previous unexpected observations on its direct (co-)regulome. To this end, three *Y. lipolytica* strains over-synthetizing YFP reporter protein were investigated: (i) strain *scYFP* expressing *HAC1* at the level of control strain, despite high-level over-synthesis of the secretory reporter protein [1], (ii) strain *inYFP* over-synthetizing intracellular YFP, with elevated *HAC1* expression level by 1.23-fold over the control (assayed by RNAseq) [1], (iii) strain *HAC1_scYFP* with synthetically induced *HAC1* expression by more than 6-fold over the prototrophic control and ~ 30% higher synthesis of YFP over the control, seen previously in batch cultures [2]). All the strains over-synthesized YFP reporter, to enable assessment of practical consequences of the *HAC1* overexpression on rs-Prot formation. We used the same model protein in all the strains, to eliminate protein-specific effects, seen previously [25]. Transcriptomes of *inYFP* and *scYFP* strains were obtained and analyzed previously [1]. Here they are used as references, enabling a more thorough characterization of the *HAC1* transcriptome.

Noteworthy, the up-regulation of *HAC1* expression observed in the *inYFP* strain [1] cannot result from a canonical UPR induction, since in-YFP protein is not targeted for secretion and hence does not enter ER compartment. On the other hand, the *HAC1* up-regulation in the *inYFP* strain was not an isolated phenomenon caused by in-YFP over-synthesis but was indeed accompanied by a number of other gene deregulations seen also in the other strain, over-synthetizing secretory glucoamylase TlG [1], facing actual UPR. Their transcriptomes shared a high level of similarity, suggesting a similar biological response to the r(*s*)-Prot over-synthesis. It is possible that this *HAC1* upregulation in the *inYFP* strain results from an indirect action of "cytosolic UPR" stress [13] governed by Hsf1, or some other cross-talk mechanism. Such a tight cross-talk between "cytosolic-" and ER-UPR is well described in plant species [45], for which the implication of bZIP60 and bZIP28/17 transcription factors (40% similarity of the former to Hac1, based on blastp) was studied in detail. The here elicited transcriptome of the *HAC1_scYFP* strain shares similarities with both *S. cerevisiae*'s transcriptomes: induced by the "ER-specific responder" Hac1 [46], and by the "cytosol-specific responder" Hsf1 [13]. For example, genes involved in fatty acid metabolism (*MGA2/D24101g* FC: 1.18), heme biosynthesis (*HEM12/C01716g* FC: 0.61, *HEM13/E29975g* FC: 0.78, *HEM15/F19470g* FC: 0.33), cell wall biogenesis (*EXG2/F05390g* FC: 1.29, *PKC1/F09746g* FC: 0.85), or phospholipid biosynthesis (*OPI3/E12441g* FC: 1.27) share the same expression pattern with *S. cerevisiae*'s *HAC1*-induced regulome [46]. On the other hand, the other genes deregulated in the *HAC1_scYFP* strain, playing role in ubiquitination and proteolysis (*UBI4/F18403g* FC: 1.27, *PIB1/D15510g* FC: 0.78), carbohydrate metabolism (*UGP1/A02310g* FC: 1.16), signal transduction (*YDR247W/B08558g/* FC: 0.83), energy generation (*PGK1/D12400g* FC: 1.19), or chaperones (*SSA8* (*SSA4*)*/D22352g* FC: 2.17, *HSP12/D20526g* FC: 1.41), were previously identified as *HSF1*-specific [13]. Such a strong overlap of here observed transcriptome of the *HAC1_scYFP* strain indicates that indeed the transcriptome of the *inYFP* strain could develop based on the accumulation of the in-YFP in the cytoplasm, which prompted *HSF1* up-regulation and concomitant *HAC1* induction (by an unknown mechanism) followed by activation of its downstream specific targets. What supports this hypothesis is that the canonical UPR transcriptome of the *TlG* strain shares the majority of the abovementioned marker genes with *HAC1*-specific transcriptome, and only one gene with the *HSF1*-specific transcriptome from *S. cerevisiae* (assayed previously [1], analyzed here).

3.1. HAC1 prevents excessive accumulation of r(s)-Prots

The strains analyzed in this study were cultured in continuous mode, and at the moment of transcriptome profiling, they were maintained in steady-state (for *HAC1_scYFP*: Supplementary Fig. S3; for *inYFP* and *scYFP* strains: [1]). The reporter YFP protein was synthesized at the levels presented in Table 1. Interestingly, co-over-expression of *HAC1* caused a nearly 7-fold drop in the retained fraction of YFP but promoted its secretion, suggesting that the multidirectional activity of Hac1 was "used" for both decreasing YFP synthesis and enhancing its secretion. The level of extracellular FL generated by *HAC1_scYFP* was nearly 2.5-fold higher over the *scYFP* strain in terms of total amounts [kRFU L^{-1}]. While the overall amounts of rs-Prots upon *HAC1* co-overexpression were lower than in the more efficient strain, *scYFP,* the latter encountered an undesired phenomenon of high YFP retention within the cell.

3.2. HAC1's unconventional splicing

While high-level over-expression of *HAC1* in the *HAC1_scYFP* strain was confirmed earlier [2], here we quantified the fraction of transcripts that were spliced to generate an active form of Hac1 [4,18–20,47]. First, it was necessary to identify and map the putative splicing site de novo, as it turned out that the *HAC1* gene from the reference genome (GCF_000002525.2) lacked this feature; even though the sequence of the intron was previously described [19,20]. The splicing sites were identified *de novo* based on the transcripts (cDNA) sequencing results and their coverage. The 29 bp fragment with the decreased coverage level visualized in Supplementary Fig. S4, corresponds to the previously described intron ([19,20]; cagtgatgactgtcgcaactactgaccag). This result provides evidence from the RNA level, that the previously identified "most likely" splicing sites [20] are operable.

Despite the high over-representation of *HAC1* transcripts (Fig. 1.A), the counts of a spliced form were comparable to the prototrophic control level, and the UPR-less *scYFP* strain (Fig. 1.B). It is commonly accepted that for effective operation, *HAC1* transcript requires unconventional splicing, to remove the translation-inhibiting short intron sequence [4, 18,47]. The results presented in (Fig. 1.A.B) suggest that it is another mechanism, rather than accumulation of the substrate (i.e. *HAC1* mRNA), that promotes the formation of an effective = spliced *HAC1* transcript in *Y. lipolytica*. As clearly demonstrated previously, accumulation of *XBP1* mRNA (a metazoan homolog of yeast *HAC1* [48]; for sequence analysis see Supplementary Fig. S5) was a prerequisite for the following Ire1-mediated splicing reaction, to overcome the degradation of protein and to produce significant amounts of the spliced form of *XBP1* [49]. Based on the here observed patterns of the intron-less transcript accumulation (Fig. 1.B) and distribution of the YFP protein (Table 1), it can be concluded that the high accumulation of YFP inside the cell (the highest in *scYFP* strain) does not promote the rate of the *HAC1* splicing (the lowest in the *scYFP* and *HAC1_scYFP* strains). The "easy-to-fold" YFP did not elicit the necessary signal of the hydrophobic domain exposure, even when accumulated at a high level.

Table 1
Amounts of YFP synthesized by Y. lipolytica strains under steady-state conditions.

Fraction Fraction	*HAC1* expression level by RNAseq [fold over control]	biomass in-YFP [kRFU L^{-1}]	supernatant sc-YFP [kRFU L^{-1}]	biomass sp.in-YFP [kRFU OD_{600}^{-1}]	supernatant sp.sc-YFP [kRFU OD_{600}^{-1}]	Reference
inYFP	1.23	130 809.60 ± 5 201.20		42 726.32 ± 3635.71		[1]
scYFP	1.0	812 108.30 ± 5 715.62	24 785.28 ± 187.33	257 051.30 ± 23 576.03	7 899.84 ± 1 260.26	
HAC1_scYFP	21.0	118 922.90 ± 3 224.39	60 123.13 ± 847.64	18 736.20 ± 923.66	9 415.35 ± 343.21	This study

The previous detailed biochemical studies [22] evidenced the key role of the Sls1 nucleotide exchange factor in regulating the interaction between Ire1 and Kar2 in *Y. lipolytica*; that ultimately leads to *HAC1* splicing. As evidenced previously at the protein level, the intracellular levels of Sls1 are induced by the accumulation of unfolded proteins [23, 24,50]. In the present study, the *SLS1* transcript levels in the *HAC1_scYFP* and *inYFP* strains were not significantly different from the control, but in the UPR-less strain (*scYFP*) its expression was decreased (see Fig. 2), which is consistent with the results obtained on the protein level [22]. On the other hand, previous studies showed that *SLS1* co-overexpression caused a significant enhancement in the rs-Prot synthesis under permissive temperature [2]. The level of improvement was comparable to the one triggered by co-overexpression of *PDI1, CNE1*, and *HAC1* [2]. The key question arose from results depicted in Fig. 1.A.B is about the mechanism underlying the observed improvement in the rs-YFP secretion, operating despite the comparable amounts of the intron-less, effective *HAC1* transcript in the *HAC1_scYFP* and the control strains. One possible explanation is that, as evidenced for *XBP1* [49], both forms – spliced and unspliced *HAC1* can modulate cellular UPR activity in yeast via direct binding to the complementary cis-elements; but the spliced variant is much more efficient. At this moment, it is only speculation that requires further in-depth studies.

3.3. *Genes highly responsive to HAC1 co-over-expression and r(s)-Prot over-synthesis – dissection of the effects*

Co-over-expression of *SCYFP* and *HAC1* genes contributed to significant up-/ down-regulation of 316 / 237 genes over the prototrophic control strain (Supplementary Fig. S6; Supplementary Table S3). Amongst the most up-regulated DEGs (fold change > 2) we identified *i. a. F15411g* encoding Zn-Fe transporter (up-regulated by > 18-fold), other genes involved in transmembrane transportation of Fe (*TAF1 D05401g, D26081g, FRE2 B10846g*), phosphate (*PHO89 E23859g*), MFS-type transporters (*C00825g*), Zn-regulated transporter (*F21659g*) and other membrane proteins putatively involved in transmembrane transportation (*A07623g, E12881g*). In addition, two ribonucleases (*RNH1 C09658g, E12331g*) and reverse transcriptase domain-bearing protein (*C19888g*), were identified amongst the most up-regulated genes upon *HAC1+SCYFP* co-over-expression, highlighting intensive RNA degradation. Strikingly, all these genes were also found to be up-regulated in the *inYFP* strain with natively induced *HAC1* expression (Fig. 3.A). However, *F15411g, TAF1, D26081g, FRE2, PHO89, C00825g*, and *E12881g* were all previously identified in the "up-regulome" of the high-level synthesis strains (*scYFP* and *SoA*), having a basic level of *HAC1* expression [1], but decreased *SLS1* expression. On the other hand, transporters *F21659g, A07623g*, as well as RNA metabolism-involved genes *RNH1, E12331g,* and *C19888g* were up-regulated solely when *HAC1* expression level was enhanced / *SLS1* transcription was not inhibited (Fig. 1.A; Fig. 2).

Amongst the genes up-regulated in the *HAC1_scYFP* strain, but to a lesser extent, we found aquaporin (*F01210g*), biofilm regulator (*E31757g*), phosphatase *PHO2* (*C19866g*), DeoR-type transcription factor (*E19030g*), Dauer up regulated-related putative ER-resident thioltransferase (*F22187g*), alternative oxidase *AOX1* (*D09933g*) and catalase *CTT1* (*E34749g*) (Fig. 3.A). All these genes were previously found up-regulated in response to high-level over-synthesis of rs-Prot in at least one of the analyzed strains (*scYFP* or *SoA*) [1]; so their up-regulation in the *HAC1_scYFP* strain results rather from sc-YFP over-synthesis than *HAC1* up-regulation (to be discussed hereafter).

Within the most down-regulated genes in the *HAC1_scYFP* strain, we identified a transcription factor *D01353g* of an unknown specific role, followed by reverse-transcriptase *A07238g* (showing the opposite pattern of expression to *C19888g* having the same function), and *URC5* involved in uracil catabolism (*E35156g*) (Supplementary Table S3). The same transcription factor *D01353g* was found amongst the most down-regulated DEGs in both the *Y. lipolytica* strains over-synthetizing rs-Prot

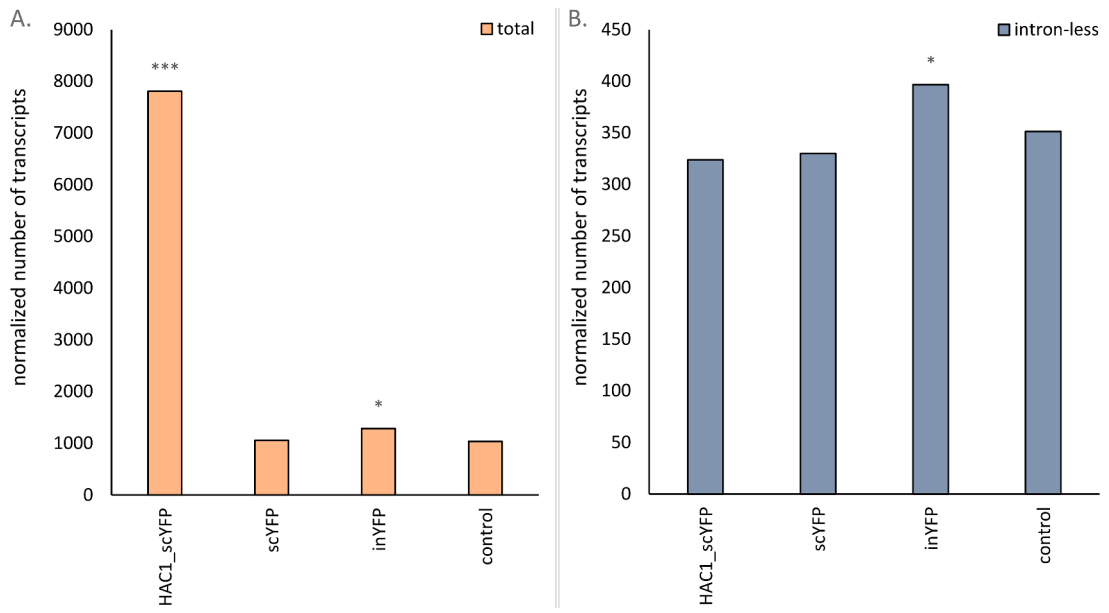

Fig. 1. Quantitation of *HAC1* transcripts total number and number of transcripts devoid of an intron in the analyzed strains' transcriptomes. The *HAC1* transcripts sequenced and quantified by RNAseq analysis were mapped for the occurrence of exons and introns. Total denotes a normalized number of all transcripts based on exon counts. Intron-less denotes a normalized number of transcripts that were devoid of an intron. *, *** signifies statistical significance at $p < 0.05$ and 0.001.

(sc-YFP and SoA), studied previously [1]. A deeper search for the *D01353g* function revealed that it bears some similarities with the repressor of RNA polymerase II. So it seems that its down-regulation releases RNA polymerase II from its repressive activity. Inhibition of uracil catabolism was observed in both the *HAC1*-over-expressing strains and in the *scYFP* strain (Fig. 3.B), as highlighted by the down-regulation of multiple DEGs in the latter [1].

Transcriptomes of *HAC1_scYFP* and *inYFP* strains shared down-regulation of RNA endonuclease (*A12683g*) that specifically degrades the RNA of RNA-DNA hybrids (Fig. 3.B), Fe-O2 oxidoreductase (*D07282g*), and nicotinic acid plasma membrane permease (*F21109g*). These genes were significantly down-regulated solely in the strains with

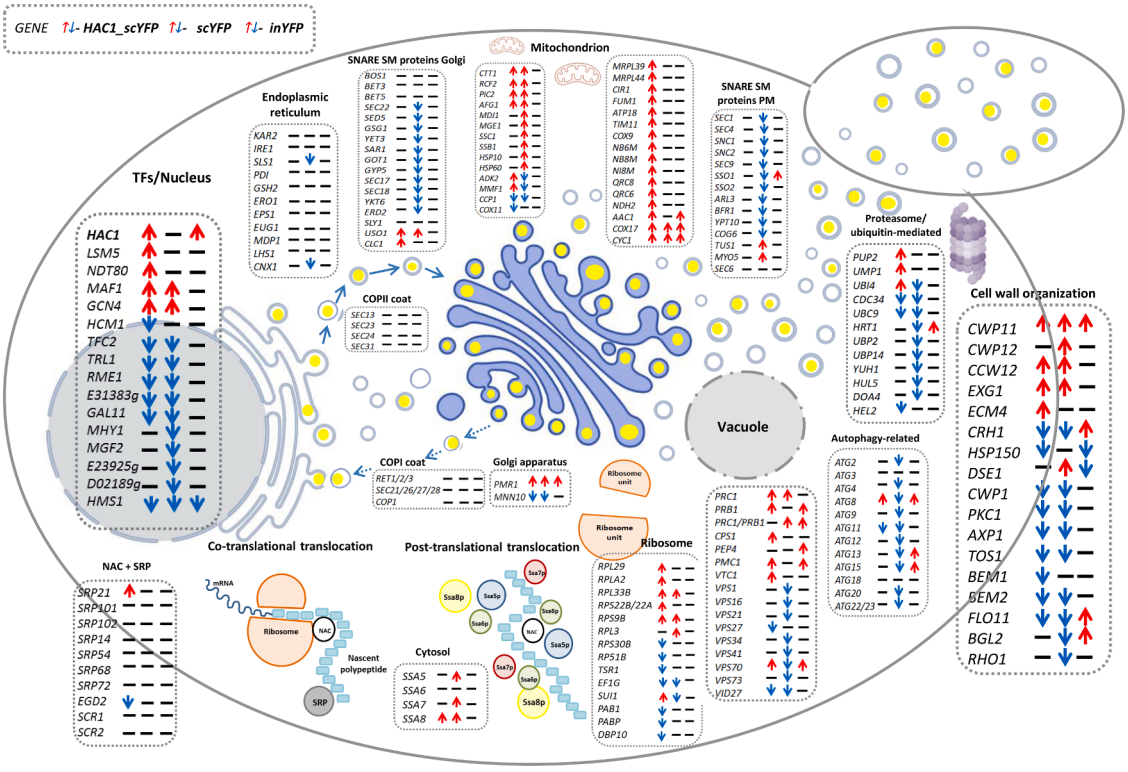

Fig. 2. Overview of expression patterns of selected genes of relevance to r(*s*)-Prot synthesis and secretion. Red arrows (↑) significant up-regulation over the control strain; blue arrows (↓) down-regulation over the control strain; black dash (–) no significant change in expression level over the control strain. Legend indicates the order of strains: left column: *HAC1_ scYFP*; middle column: *scYFP*; right column: *inYFP*. Numerical values of estimated Fold Change (FC) and a full list of genes depicted in this Fig., with their specific functions, is given in Supplementary Table S4.

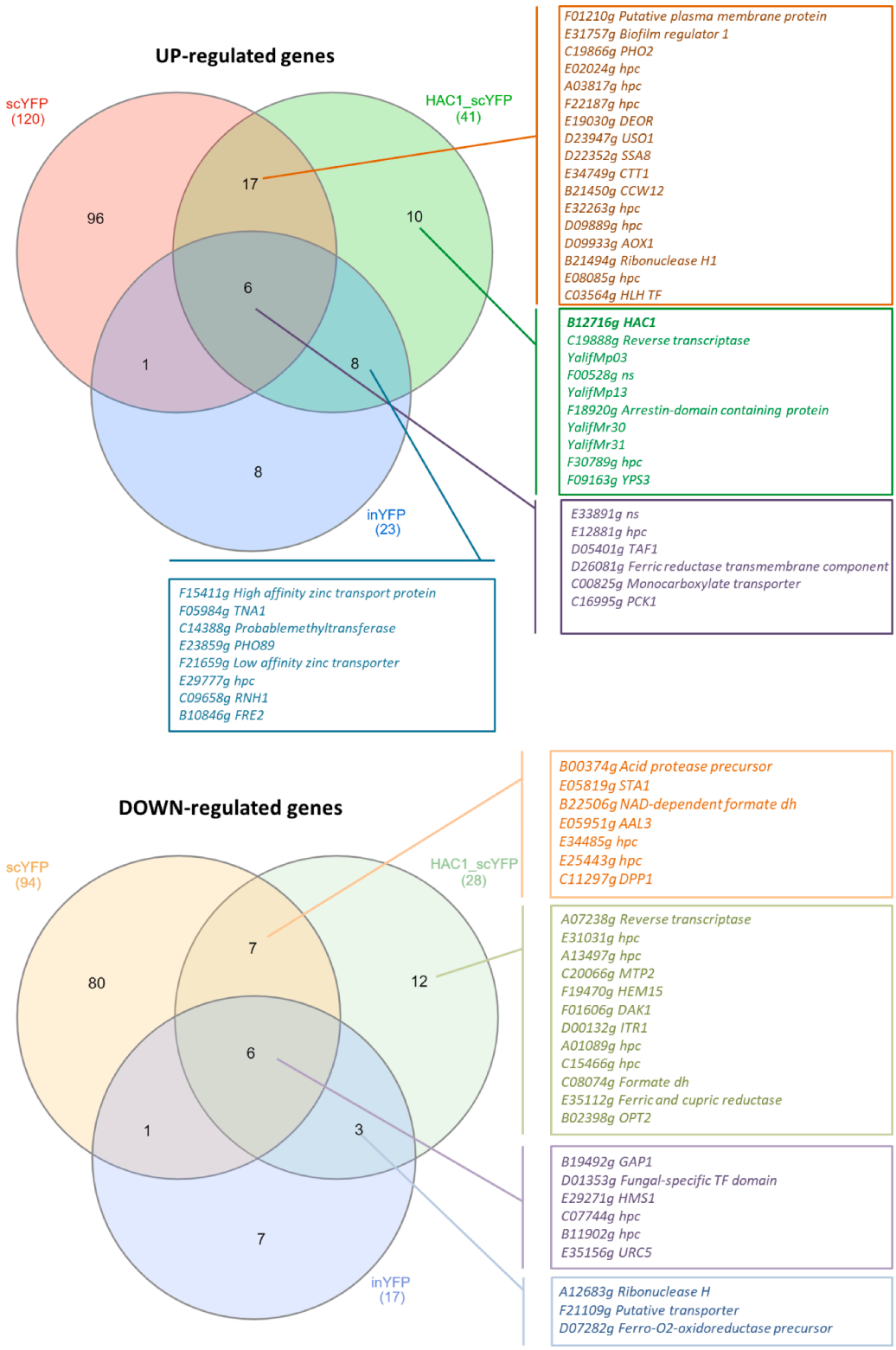

Fig. 3. Venn Diagrams analysis illustrating shared and dissimilar gene sets for individual strains. The numbers indicate the number of shared or strain-specific genes, up-/downregulated in *HAC1_scYFP, scYFP,* and *inYFP* strains; (A) up-regulated DEGs or (B) down-regulated DEGs. Venn diagrams analysis was done with the DEGs lists restricted to $\log_2$-fold change>1 and *p*-value < 0.05. YALI_code of the gene and a short description is given for selected genes. Abbreviations: TF – Transcription Factor, dh – dehydrogenase, hpc – hypothetical protein conserved in the *Yarrowia* clade, ns – no similarity.

an elevated level of *HAC1* expression (Fig. 1.A). On the other hand, there was a group of genes that were commonly down-regulated in the strains over-synthetizing sc-/in-YFP, like *AAL3* (*E05951g*) activating acyl chains, *CUP9*-related transcription factor regulating RNA polymerase II (*E29271g*), but also *GAP1* general amino acid permeases (*B19492g*), and flocculin *STA1* (*E05819g*) (Fig. 3.B; not all the genes are shown in Venn diagrams, as they did not always meet 2-fold change criterion for all the strains).

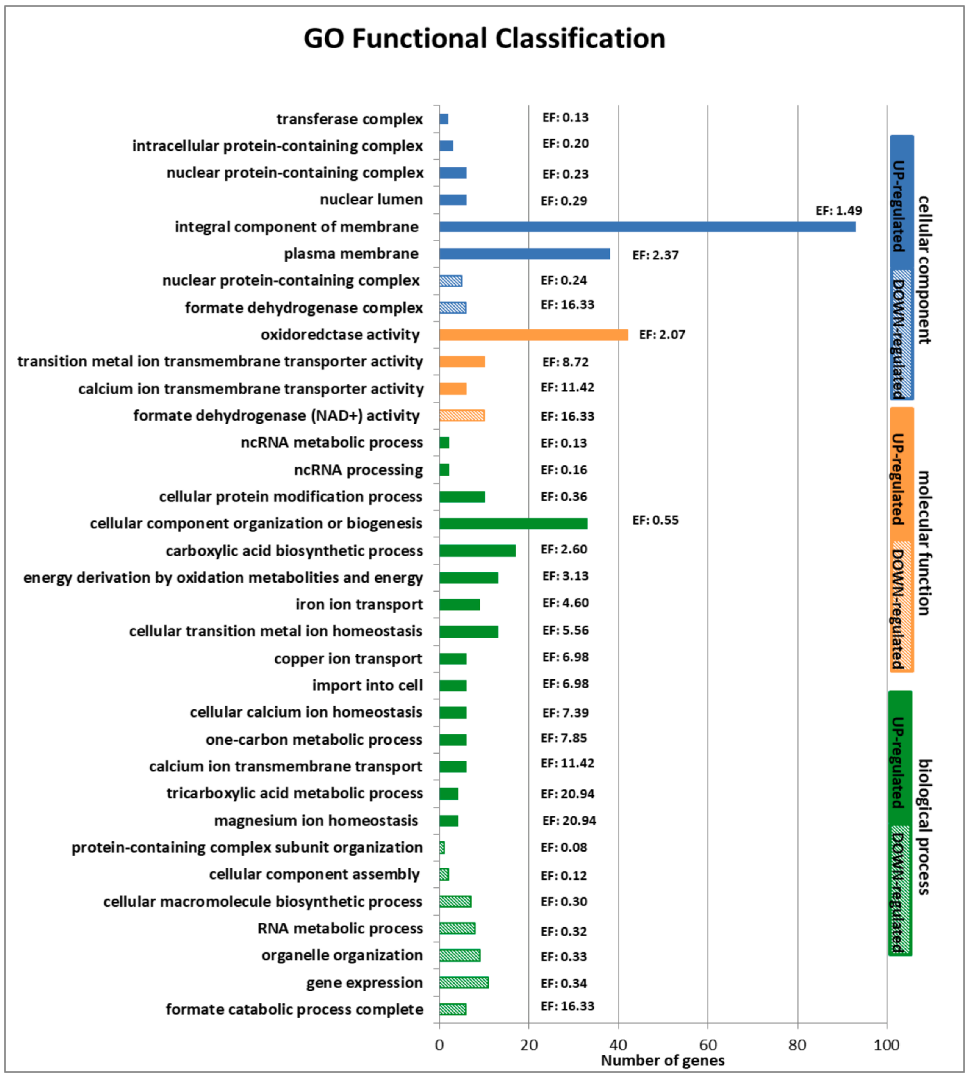

Fig. 4. Functional classification and over-representation test of DEGs in *HAC1-scYFP* strain transcriptome. GO term statistical over-representation test was conducted using the Panther tool on complete DEGs lists. The vertical coordinates are the significantly enriched GO terms, and the horizontal coordinates are the numbers of the up-regulated (full) and down-regulated (dashes) genes in these GO terms. Blue - cellular component; orange - molecular function; green - biological process. Only GO terms enriched at a significance level of $p < 0.05$ are shown. EF – Enrichment fold.

From the above it is clear that the transcriptome of the *HAC1_scYFP* strain bears several unique biological features, but also shares the others with one of the compared strains – *inYFP* or *scYFP* (Figs. 2, 3.A.B). It implies that in the *HAC1_scYFP* strain, the deregulation in the gene expression could equally well result from sc-YFP over-synthesis, the metabolic burden caused by the co-over-expression of the two genes, or over-representation of *HAC1* transcript (and notably - concomitant constant level of *SLS1*, which is an important finding of this study).

For example, several membrane transporters, genes involved in uracil catabolism, and RNA degradation were deregulated solely upon *HAC1* over-representation (*HAC1_scYFP* and *inYFP*). The former observation well corresponds with the results of functional classification followed by an overrepresentation test for that strain (Fig. 4) showing that the cellular component "plasma membrane" and molecular function "transporter" were the most affected (vs the control strain) upon *HAC1* and *SCYFP* co-over-expression. Inhibition of uracil catabolism concomitant with up-regulation of ribonucleases, which decompose RNA to single nucleotides, suggests that RNA components are not degraded but intensively re-used. These biological processes could be linked with the high disproportion between *HAC1* transcript in spliced and un-spliced form (Fig. 1.A.B) seen in the *HAC1_scYFP* strain; and to a lesser extent, the *HAC1* mRNA turnover in the *inYFP* strain.

Genes involved in oxidative stress response (Fig. 3.A.B), cell cycle progression (Fig. 5), and several RNA polymerase II and III repressors (Figs. 2, 3.A.B, 4) were deregulated in response to a rs-Prot over-synthesis, rather than to the specific action of operable Hac1. Likewise, based on DEGs profiles illustrated in Fig. 2, it seems that nucleus-localized genes (mainly transcription factors) were more responsive to the rs-Prot over-synthesis rather than to the increased occupancy of the *HAC1* transcript; inferred based on the high similarity of *HAC1_scYFP* and *scYFP* strains profiles.

Finally, the global transcriptomes comparison suggests that the unique up-regulation of multiple genes localized to mitochondria and involved in cellular energetics in the *HAC1_scYFP* is a consequence of the co-over-expression of two heterologous genes, and the high-energy-consuming secretion process (Table 1). In this regard, the high-level synthesis and secretion strain, *scYFP*, shows partial similarity.

Still, we are aware that the abovementioned correlations between unique DEGs indicated above and generated phenotypes may equally well indicate a direct functional interaction, or be a consequence of some other not identified phenomena.

3.4. (Co)-Regulome of HAC1

Further in relation to the specificity of *HAC1* operation in *Y. lipolytica*, of key importance for us was to investigate the "direct regulome" of *HAC1*, including ER-resident foldases and chaperones (e.g. *KAR2, PDI1, etc.*), since we previously observed that they are insensitive to changes in *HAC1* expression level in *Y. lipolytica* [1,2]. That observation stayed in contrast to what has been earlier seen in *S. cerevisiae* and

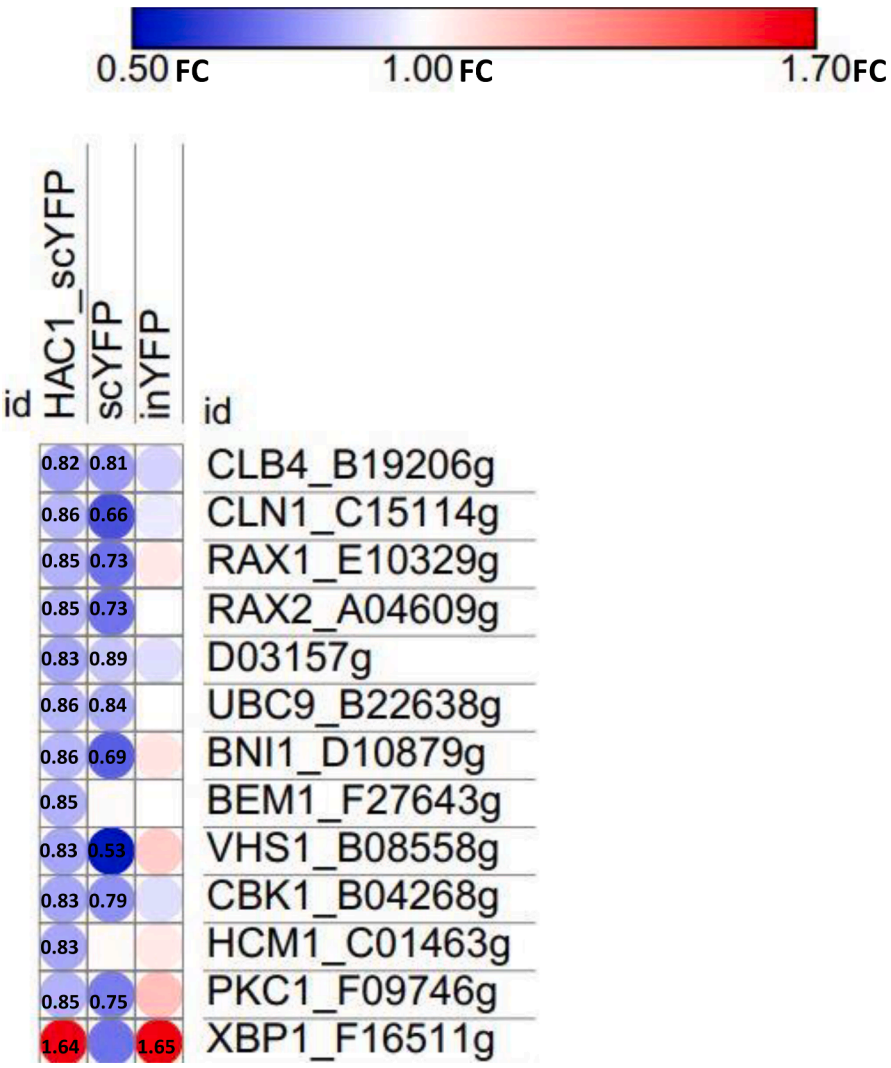

Fig. 5. Expression patterns of DEGs involved in cell cycle progression in HAC1_scYFP, scYFP, and inYFP strains. YALI_ code and a short description of a given gene are shown. Assigned names and functions were manually curated for all the ambiguous hits by blasting the sequence against UniProt and NCBI non-redundant protein database, and inferring the most accurate match. Red (up-regulation), blue (down-regulation). Numbers indicate estimated Fold Change (FC). A full list of genes depicted in this Fig., with their specific functions, is given in Supplementary Table S5.

P. pastoris [5,11,18,21], or even for *Y. lipolytica* [19] (assayed by RT-qPCR). The current transcriptomics analysis indicates that none of the ER foldases nor chaperones exhibited any changes in transcription level upon *HAC1* synthetic over-expression or its native induction in the *inYFP* strain (Fig. 2; Supplementary Table S4). Their expression level was not affected either when *HAC1* was synthetically co-overexpressed (assayed here by RNAseq or previously by RT-qPCR [2]), or when the spliced *HAC1* variant was in abundance in the *inYFP* strain (assayed previously by RNAseq [1]). This observation stays in agreement, with our previous observation, and supports the statement that their transcription is not correlated with the *HAC1* expression level in *Y. lipolytica*.

Amongst the genes known to be implicated in *HAC1*-driven UPR in yeast is *XBP1*, which in the present study showed an expression pattern highly corresponding to *HAC1* (Fig. 5). As described for *S. cerevisiae, XBP1* encodes a transcriptional repressor binding to promoter sequences of cyclin genes, *CYS3, SMF2*, and *CLN1* hence is also claimed to be involved in the G1/S transition of the mitotic cell cycle [51,52]. Homonimic but not homologous *XBP1* (X-box binding protein 1; see alignments in Supplementary Fig. S5) is a widely studied metazoan transcriptional repressor, undergoing unconventional mRNA splicing under conditions of ER stress [48,49,53]. Noteworthy, metazoan *XBP1* is a homolog of yeast *HAC1*, while yeast *XBP1* is a separate gene, operable in yeast (at least in *S. cerevisiae* and *Y. lipolytica*; please see Supplementary Fig. S5). Both transcription factors Hac1 and Xbp1 operate in the yeast cells, and both undergo unconventional splicing by Ire1. Upon ER stress, *XBP1* mRNA is spliced by Ire1 (as *HAC1*, discussed above), thereby generating functional *XBP1* transcript, which is translocated to the nucleus to initiate transcriptional programs that regulate a subset of UPR- and non-UPR-associated genes. In yeast, Xbp1 is known to repress 15% of all genes as cells transition to quiescence (G1 arrest). It is specifically interesting since *XBP1* was also claimed to be implicated in UPR. It is also known that *XBP1* expression is induced by different stress conditions. A protein encoded by the *ylXBP1* gene (*F16511g*) shows significant similarity to the *S. cerevisiae*'s homolog (YIL101C1; Supplementary Fig. S5). Based on its unique up-regulation in *HAC1_scYFP* and *inYFP* it can be speculated that its expression is associated with the *HAC1* transcription / (cytosolic) UPR induction in *Y. lipolytica*. However, this statement requires further investigation.

Intriguingly, amongst the genes significantly down-regulated in the scYFP-over-synthetizing strains we identified a tRNA ligase *TRL1* involved in tRNA biogenesis (*D14916g*). Strikingly, one of the *TRL1*'s functions is the ligation of exons after unconventional splicing of *HAC1/XBP1* transcript to initiate UPR [54]. Aside from *HAC1* and *IRE1, TRL1* (*RLG1* in *S. cerevisiae*) is considered the key element of UPR response in *S. cerevisiae* [49]. *TRL1*'s down-regulation was observed in all the *Y. lipolytica* strains co-overexpressing any secretory r-Prot (by 14%, 33%, 14%, and 24% for *HAC1_scYFP, scYFP, TlG*, and *SoA* strain, respectively [1]), but not for *inYFP*, which is the only case for which a significant increase in spliced *HAC1* transcript counts was observed (Fig. 1.B). While it does not explain how co-over-expression of *HAC1*

contributed to enhanced rs-Prot secretion, it suggests that Trl1 ligase is indeed implicated in unconventional splicing of *HAC1/XBP1* in *Y. lipolytica,* and is the Rlg1's homolog.

3.5. Biological processes affected by HAC1 and SCYFP co-over-expression

Biological processes are most frequently modulated by the cumulative action of many minor modifications of multiple genes. The most efficient way of finding and characterizing such global changes is functional classification followed by an overrepresentation test, that yields numerical value (enrichment fold, EF) indicating whether the group of DEGs is enriched in a specific biological process, molecular function, or localized to a specific cellular compartment. Therefore, all DEGs identified in the *HAC1_scYFP* strain (also those showing expression fold change < 2, but significantly deregulated) were searched for enrichment in specific biological functions. The results of this analysis are illustrated in Fig. 4. Functional categorization showed that the most affected cellular component was the plasma membrane and its integral components, represented by more than 100 up-regulated DEGs. It could be expected since multiple highly-deregulated DEGs were represented by trans-membrane transporters of various types. No such clear indication could be inferred from the enrichment test on down-regulated cellular components, except from the formate dehydrogenase complex, which also constituted an enriched molecular function and a biological process identified amongst the down-regulated DEGs (see Fig. 4). Oxidoreductase activity and ion transporters significantly up-regulated molecular functions in the *HAC1_scYFP* strain. Up-regulation of ion transporter function well corresponds with enrichment in plasma membrane cellular component. Likewise, many different biological processes related to ions transportation and import into the cell were identified as enriched biological processes up-regulated upon co-over-expression of *HAC1* and *SCYFP*. Energy derivation, tricarboxylic acid cycle (TCA), and one-carbon metabolic processes were also identified as significantly up-regulated biological processes (Fig. 4, Supplementary Table S3). Within the latter group, we found many genes involved in amino acid synthesis and conversion (see Supplementary Table S3). Enrichment in energy derivation and TCA well corresponds with concerted up-regulation of mitochondria-localized genes involved in cellular energetics, which was unique for the *HAC1_scYFP* strain (see Fig. 2; Supplementary Table S4). Intensive rebuilding of cellular structures was highlighted by over 30 up-regulated DEGs assigned to cellular components organization and biogenesis, and over 10 DEGs assigned to cellular components assembly that were down-regulated (Fig. 4). Careful analysis of the DEGs assigned to these categories showed that cellular components organization and biogenesis were specifically represented by genes localized to mitochondria, involved in vesicle transportation and chromatin re-structuring (specific DEGs can be found in Supplementary Table S3).

Changes to the protein synthesis process were hallmarked by the deregulation of several ribosomal subunit proteins, and eukaryotic translation initiation factor *SUI1* (*C07524g*; up-regulated by 23%; Fig. 2; Supplementary Table S4). On the contrary, RNA metabolism and gene expression, as well as macromolecule biosynthesis were all down-regulated biological processes upon *HAC1* and *SCYFP* co-over-expression. Specific DEGs classified to these down-regulated categories were ribosomal proteins *RP51B* (*F14465g*; down-regulated by 18%), *RPS30B* (*F14663g*; down-regulated by 29%), pre-rRNA processing protein *TSR1* (*B08756g*; down-regulated by 13%), elongation factor EF-1-gamma (*C16049g*; down-regulated by 20%), as well as tRNA ligase *TRL1* involved in tRNA biogenesis (*D14916g*; down-regulated by 14%).

3.6. Biological processes of specific interest for rs-Prots production

As inferred from the previous [1] and current transcriptomics data, as well as the amounts of YFP reporter protein distribution (Table 1), three specific biological processes were postulated as Hac1-driven mechanisms for optimization of rs-Prots synthesis in *Y. lipolytica*: (i) rate of r(*s*)-Prot synthesis, (ii) degradation of proteins, and (iii) vesicular transportation across the secretory pathway. Having transcriptomics data, we were able to specifically track these processes in the analyzed strains at the gene expression level. Transcriptomics data for the genes specifically implicated in these biological processes are globally presented in Fig. 2.

3.6.1. Gene expression regulation

A wide set of genes deregulated upon *HAC1* co-over-expression with *SCYFP* was localized to the nucleus (Fig. 2; Supplementary Table S4). Many of the identified DEGs were transcription factors of different scopes. For example, *MAF1* (*F10541g*; up-regulated by 82%), and *TFC2* (*F05104g*; down-regulated by 25%) exert their regulatory action towards RNA polymerase III. Maf1 is a negative regulator of RNA polymerase III, involved in tRNA processing and stability; while Tfc2 is a positive regulator of rRNA expression. The down-regulated transcription factor *E31383g* is involved in the positive regulation of transcription by RNA polymerase II; while transcription factor *GCN4* (*E27742g*), up-regulated in *HAC1_scYFP* and *scYFP* strains (Fig. 2; Supplementary Table S4), is a transcription factor that is responsible for the activation of more than 30 genes required for amino acid and purine biosynthesis in response to their deficiency. The expression pattern of transcription factors *MAF1, TFC2*, and *E31383g* across the analyzed strains suggests that transcription by the RNA polymerase III and II is limited upon high-level rs-Prot synthesis and secretion. This observation well corresponds with the down-regulation of the "gene expression" functional category (Fig. 4). Up-regulation of transcription factor *GCN4* in *HAC1_scYFP* and *scYFP* strains (Fig. 2; Supplementary Table S4) highlights the occurrence of nitrogen starvation in the scYFP over-synthetizing strains. Upon induction of UPR, *GCN4* was up-regulated in *P. pastoris* but down-regulated in *S. cerevisiae* [11]. Gcn4 was shown to be activated upon ER stress and to play an essential role in the UPR together with or downstream of Hac1 [55]. Our data show that *GCN4* expression level is not directly linked to *HAC1* expression in *Y. lipolytica*. It is more associated with amino acid starvation under intensive over-synthesis of rs-Prot. Co-over-expression of *HAC1* and *SCYFP* caused significant changes to the ribosome operation (Fig. 2; Supplementary Table S4). In the face of indicated above attenuation of transcription by RNA polymerases III and II, the down-regulation of multiple translation-associated genes is more understandable (Fig. 2, Supplementary Table S4). Notably, such concerted changes to ribosome-related genes were observed neither in highly-synthetizing *scYFP* nor in spliced *HAC1*-up-regulated *inYFP* strain, so it is plausible that a combination of both genes' co-over-expression contributed to this effect. Altogether these data indicate that upon high-level rs-Prot synthesis gene expression is attenuated and the cells encounter nitrogen starvation.

3.6.2. Protein degradation via different pathways

The second proposed mechanism by which Hac1 executes its protective role in *Y. lipolytica* cells is targeting excessively accumulated, incorrectly folded r(*s*)-Prots to degradation (Fig. 2). We observed that indeed this biological process was up-regulated in *HAC1_scYFP* strain, even though the model protein was "easy-to-fold" (based on previous comparisons [1,25]). Specifically, the main vacuolar protease *PRB1* (*B16500g*; up-regulation by 22%), vacuolar Ca2+ ATPase *PMC1* (*D04873g*; up-regulation by 37%), vacuolar protein sorting factor *VPS70* (*B05258g*; up-regulation by 31%) and essential for autophagy *ATG8* gene (*E02662g*; up-regulation by 22%) were all up-regulated. This observation was contrasted by global down-regulation of the genes involved in any of the protein degradation pathways (proteasome, autophagy / vacuolar) in the *scYFP* strain and the other highly over-synthetizing strain *SoA* [1], with no change or up-regulation in the *HAC1*-up-regulation strains (*HAC1_scYFP, inYFP* and *TlG*), (Fig. 2; Supplementary Table S4) [1]. Amongst the genes specifically up-regulated

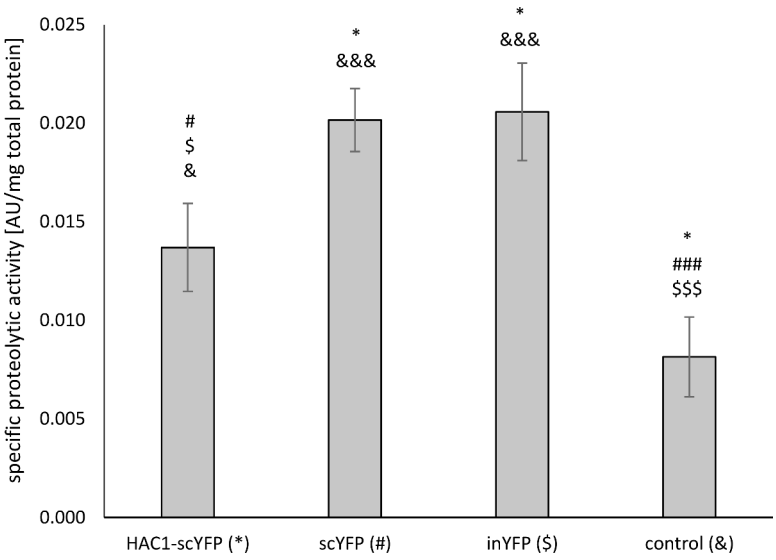

Fig. 6. Total intracellular proteolytic activity of the analyzed strains. Proteolytic activity was determined in total protein extracts. Y axis: specific proteolytic activity equivalent to trypsin activity in standardized solution per a unit of total protein concentration expressed in [AU/mg total protein]. Each strain is marked with a symbol: *, #, $, and &, for *HAC1_scYFP, scYFP, inYFP*, and control, respectively. Top labels indicate the statistical significance of the differences at the level of $p < 0.05$ and 0.001.

in *HAC1*-induced strains was *E02662g* encoding the crucial for autophagy initiation - Atg8. *ATG8* is the key gene involved in autophagy, directing the vacuolar degradation of autophagosomes. Interestingly, it is also known to be involved in the ER-specific autophagy process and is essential for the survival of cells subjected to severe ER stress. Hence, based on such a specific expression pattern, conserved across so many genes, we hypothesized that the *HAC1* expression level is directly linked to the vacuolar sorting and degradation.

To verify this statement, we examined the total intracellular proteolytic activity of the strains studied here. Based on the assay results (Fig. 6), any over-synthesis of r(*s*)-Prot led to increased total proteolytic activity. On the other hand, co-over-expression of *HAC1* contributed to a significant decrease of proteolysis vs the other two YFP-over-synthetizing strains, but still it was higher in *HAC1_scYFP* strain than in the control. The biological sense of the enhancement in total proteolytic activity upon over-synthesis of r(*s*)-Prot seems clear and is most probably related to globally enhanced protein turnover. However, such an explanation does not correspond to the overall profile of transcriptomics data, from which it is clear that proteolysis is inhibited in the *scYFP* strain. Careful analysis of the transcriptomics data revealed that many of the down-regulated genes assigned to the functional category "proteolysis" are vacuolar sorting factors, ubiquitination-mediating factors, and autophagy-related factors, which do not possess proteolytic activity *per se*, which was assayed in the test. In other words, it was the process of the protein targeting for degradation, that was decreased in the *scYFP* strain. Specific proteolytic activity was the highest in the strains facing enhanced YFP retention, i.e. *scYFP* and *inYFP* (Table 1, Supplementary Fig. S2). On the other hand, co-over-expression of *HAC1* contributed to a significant decrease of proteolysis vs these two strains, which still was higher than in the control. The decreased proteolytic activity in the *HAC1_scYFP* strain could partially explain the higher amounts of secreted YFP (Table 1).

3.6.3. *Vesicular transportation across the secretory pathway*

The data presented in Table 1 suggest that one of the reasons for increased amounts of the secreted YFP reporter in the *HAC1_scYFP* strain could be intensified transportation through the secretory pathway. From amongst the genes involved in anterograde and retrograde COP-mediated transport, intra- and post-Golgi trafficking, all except two (*USO1, D23947g*, essential for vesicle-mediated ER to Golgi transport; up-regulated by 2.58-fold; *CLC1, E33253g*, vesicle coat protein, clathrin light chain; up-regulated by 18%) remained at the control expression level. At the same time, many genes involved in vesicular transportation were down-regulated in the *scYFP* strain (Fig. 2; Supplementary Table S4), which over-synthetized the rs-Prot at a very high level and secreted it. Looking at the data illustrated in Fig. 2, it was not possible to indicate specific DEGs that could account for such an outcome. The surprising down-regulation of vesicular transportation genes under high-level synthesis and secretion in the *scYFP* strain is not well understood at this moment.

3.6.4. *Cell cycle progression*

Based on the conducted enrichment test (not shown), we identified another functional category that was enriched amongst DEGs. Namely, a set of genes involved in cell cycle progression was significantly down-regulated in response to the co-over-expression of *HAC1* and *SCYFP* (Fig. 5; Supplementary Table S5). A similar, yet not identical, expression pattern was observed for the *scYFP* strain. The expression pattern of the genes involved in cell-cycle progression was characterized by concerted down-regulation of several serine/threonine-protein kinases involved in cell division and morphogenesis (*VHS1, B08558g; CBK1, B04268g; PKC1, F09746g*, G2/M-specific), two cyclins *CLB4* (*B19206g*, G2/M-specific) and *CLN1* (*C15114g*, G1/S-specific), as well as *UBC9* gene (*B22638g*) specifically involved in the degradation of cyclins of S- and M-phases. Furthermore, we observed significant down-regulation of an *HCM1* transcription factor involved in cell cycle-specific transcription (*C01463g*), proteins *RAX1* (*E10329g*) and *RAX2* (*A04609g*), *BNI1* (*D10879g*) and *BEM1* (*F27643g*), all implicated in biological processes like bud site selection, cytokinesis, maintenance of cell polarity; as well as GTP-binding protein *D03157g* involved in the mitotic spindle and division septum assembly. Strikingly, in line with the down-regulation of cell cycle progression, we noted significant up-regulation of the *XBP1* repressor gene (*F16511g*) in both *HAC1_scYFP* and *inYFP* strains, contrasted by its down-regulation in the *scYFP* strain (Fig. 5); its involvement in cell cycle regulation but also in ER stress response was confirmed for *S. cerevisiae* [51–53]. Considering previously [1] and currently observed expression patterns of *HAC1* and *XBP1* across all the strains analyzed (*SoA, scYFP, HAC1_scYFP, inYFP,* and *TlG*) it can be concluded that the expression of the two genes is correlated. Based on the "guilt-by-association" rule, it suggests their functional interaction of yet unknown mechanisms.

4. Conclusions

In conclusion, this study provides insight into several new elements of the secretory pathway in *Y. lipolytica*. A practical outcome of this research is demonstrating that *HAC1* co-overexpression in *Y. lipolytica* leads to enhanced secretion of the correctly folded target protein and its

limited retention. The identified associated molecular mechanisms were massive changes in mitochondria and ribosomes operation, cell cycle arrest, attenuation of gene expression by RNA polymerase III and II, as well as modulation of proteolysis and RNA metabolism; but whether the *HAC1* co-over-expression/induction was the actual causative agent for these changes, was not always clear. On the other hand, it was settled that the expression level of the "conventional" *HAC1* targets, like *KAR2, PDI1,* and other ER-localized foldases, is not affected by increased expression of *HAC1* in *Y. lipolytica*. Moreover, our transcriptomics data suggest the implication of the Xbp1 transcription factor (*F16511g*) and the Trl1 ligase (*D14916g*) in the *HAC1*-driven cellular response and unconventional splicing of *HAC1* mRNA in *Y. lipolytica*, respectively.

Funding

This study was partly financially supported by the Ministry of Science and Higher Education (currently, Ministry of Education and Science) project number: DI 2017 000947, in the part concerning chemostat cultivations, samples processing, and part of RNAseq and bioinformatics. The research was also partly funded by National Science Centre, Poland, grant number 2021/41/B/NZ9/00086; in the part concerning the investigation of the *HAC1* splicing based on transcripts sequencing, and the impact on *HAC1* overexpression on the intracellular proteolysis.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

CRediT authorship contribution statement

Paulina Korpys-Woźniak: Investigation, Methodology, Visualization, Data curation, Writing – original draft, Writing – review & editing. **Ewelina Celińska:** Conceptualization, Supervision, Formal analysis, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data accompanying this research are presented directly in the manuscript, and supplementary materials or are available in Sequence Read Archive, NCBI database under BioProject number: PRJNA869113.

Acknowledgments

Graphical Abstract was prepared with BioRender.com.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.btre.2023.e00801.

References

[1] P. Korpys-Woźniak, E. Celińska, Global transcriptome profiling reveals genes responding to overproduction of a small secretory, a high cysteine- and a high glycosylation-bearing protein in *Yarrowia lipolytica*, Biotechnol. Rep. 31 (2021) e00646, https://doi.org/10.1016/j.btre.2021.e00646.

[2] P. Korpys-Woźniak, P. Kubiak, E. Celińska, Secretory helpers for enhanced production of heterologous proteins in *Yarrowia lipolytica*, Biotechnol. Rep. 32 (2021) e00669, https://doi.org/10.1016/j.btre.2021.e00669.

[3] K. Mori, T. Kawahara, H. Yoshida, H. Yanagi, T. Yura, Signalling from endoplasmic reticulum to nucleus: transcription factor with a basic-leucine zipper motif is required for the unfolded protein-response pathway, Genes Cells 1 (1996) 803–817, https://doi.org/10.1046/j.1365-2443.1996.d01-274.x.

[4] M. Guerfal, S. Ryckaert, P.P. Jacobs, P. Ameloot, K. Van Craenenbroeck, R. Derycke, N. Callewaert, The HAC1 gene from *Pichia pastoris*: characterization and effect of its overexpression on the production of secreted, surface displayed and membrane proteins, Microb. Cell Fact 9 (2010) 1–12, https://doi.org/10.1186/1475-2859-9-49.

[5] A.R. Hesketh, J.I. Castrillo, T. Sawyer, D.B. Archer, S.G. Oliver, Investigating the physiological response of *Pichia (Komagataella) pastoris* GS115 to the heterologous expression of misfolded proteins using chemostat cultures, Appl. Microbiol. Biotechnol. 97 (2013) 9747–9762, https://doi.org/10.1007/s00253-013-5186-1.

[6] D. Mattanovich, B. Gasser, H. Hohenblum, M. Sauer, Stress in recombinant protein producing yeasts, J. Biotechnol. 113 (2004) 121–135, https://doi.org/10.1016/j.jbiotec.2004.04.035.

[7] A. Graf, B. Gasser, M. Dragosits, M. Sauer, G.G. Leparc, T. Tüchler, D.P. Kreil, D. Mattanovich, Novel insights into the unfolded protein response using *Pichia pastoris* specific DNA microarrays, Biomed. Chromatogr. 9 (2008) 1–13, https://doi.org/10.1186/1471-2164-9-390.

[8] K.J. Kauffman, E.M. Pridgen, F.J. Doyle, P.S. Dhurjati, A.S. Robinson, Decreased protein expression and intermittent recoveries in BiP levels result from cellular stress during heterologous protein expression in *Saccharomyces cerevisiae*, Biotechnol. Prog. 18 (2002) 942–950, https://doi.org/10.1021/bp025518g.

[9] K.E.J. Tyo, Z. Liu, D. Petranovic, J. Nielsen, Imbalance of heterologous protein folding and disulfide bond formation rates yields runaway oxidative stress, BMC Biol. 10 (2012), https://doi.org/10.1186/1741-7007-10-16.

[10] M. Hernández-Elvira, F. Torres-Quiroz, A. Escamilla-Ayala, E. Domínguez-Martin, R. Escalante, L. Kawasaki, L. Ongay-Larios, R. Coria, The unfolded protein response pathway in the yeast *kluyveromyces lactis*. A comparative view among yeast species, Cells 7 (2018), https://doi.org/10.3390/cells7080106.

[11] B. Gasser, M. Maurer, J. Rautio, M. Sauer, A. Bhattacharyya, M. Saloheimo, M. Penttilä, D. Mattanovich, Monitoring of transcriptional regulation in *Pichia pastoris* under protein production conditions, Biomed. Chromatogr. 8 (2007) 1–18, https://doi.org/10.1186/1471-2164-8-179.

[12] J. Liu, Q. Han, Q. Cheng, Y. Chen, R. Wang, X. Li, Y. Liu, D. Yan, Efficient expression of human lysozyme through the increased gene dosage and co-expression of transcription factor Hac1p in *Pichia pastoris*, Curr. Microbiol. 77 (2020) 846–854, https://doi.org/10.1007/s00284-019-01872-9.

[13] J. Hou, H. Tang, Z. Liu, T. Österlund, J. Nielsen, D. Petranovic, Management of the endoplasmic reticulum stress by activation of the heat shock response in yeast, FEMS Yeast Res 14 (2014) 481–494, https://doi.org/10.1111/1567-1364.12125.

[14] B. Gasser, M. Sauer, M. Maurer, G. Stadlmayr, D. Mattanovich, Transcriptomics-based identification of novel factors enhancing heterologous protein secretion in yeasts, Appl Environ Microbiol 73 (2007) 6499–6507, https://doi.org/10.1128/AEM.01196-07.

[15] X. Zhao, W. Xie, Y. Lin, X. Lin, S. Zheng, S. Han, Combined strategies for improving the heterologous expression of an alkaline lipase from *Acinetobacter radioresistens* CMC-1 in *Pichia pastoris*, Process Biochem. 48 (2013) 1317–1323, https://doi.org/10.1016/j.procbio.2013.06.021.

[16] H. Raschmanová, A. Weninger, Z. Knejzlík, K. Melzoch, K. Kovar, Engineering of the unfolded protein response pathway in *Pichia pastoris*: enhancing production of secreted recombinant proteins, Appl. Microbiol. Biotechnol. 105 (2021) 4397–4414, https://doi.org/10.1007/s00253-021-11336-5.

[17] R. Gordân, K.F. Murphy, R.P. McCord, C. Zhu, A. Vedenko, M.L. Bulyk, Curated collection of yeast transcription factor DNA binding specificity data reveals novel structural and gene regulatory insights, Genome Biol. 12 (2011), https://doi.org/10.1186/gb-2011-12-12-r125.

[18] G. Whyteside, R.M. Nor, M.J.C. Alcocer, D.B. Archer, Activation of the unfolded protein response in *Pichia pastoris* requires splicing of a HAC1 mRNA intron and retention of the C-terminal tail of Hac1p, FEBS Lett. 585 (2011) 1037–1041, https://doi.org/10.1016/j.febslet.2011.02.036.

[19] N. Callewaer, W. Vervecken, K. De Pourq, S. Geysens, M. Guerfal, *Yarrowia lipolytica* and *Pichia pastoris* Hac1 nucleic acids, US 8026083 B2, 2011.

[20] M.H. Oh, S.A. Cheon, H.A. Kang, J.Y. Kim, Functional characterization of the unconventional splicing of *Yarrowia lipolytica* HAC1 mRNA induced by unfolded protein response, Yeast 27 (2010) 443–452, https://doi.org/10.1002/yea.1762.

[21] Z. Liu, J. Hou, J.L. Martínez, D. Petranovic, J. Nielsen, Correlation of cell growth and heterologous protein production by *Saccharomyces cerevisiae*, Appl. Microbiol. Biotechnol. 97 (2013) 8955–8962, https://doi.org/10.1007/s00253-013-4715-2.

[22] A. Babour, M. Kabani, A. Boisramé, J.M. Beckerich, Characterization of Ire1 in the yeast *Yarrowia lipolytica* reveals an important role for the Sls1 nucleotide exchange factor in unfolded protein response regulation, Curr. Genet. 53 (2008) 337–346, https://doi.org/10.1007/s00294-008-0190-1.

[23] A. Boisramé, J.M. Beckerich, C. Gaillardin, Sls1p, an endoplasmic reticulum component, is involved in the protein translocation process in the yeast *Yarrowia lipolytica*, J. Biol. Chem. 271 (1996) 11668–11675, https://doi.org/10.1074/jbc.271.20.11668.

[24] A. Boisramé, M. Kabani, J.M. Beckerich, E. Hartmann, C. Gaillardin, Interaction of Kar2p and Sls1p is required for efficient co- translational translocation of secreted proteins in the yeast *Yarrowia lipolytica*, J. Biol. Chem. 273 (1998) 30903–30908, https://doi.org/10.1074/jbc.273.47.30903.

[25] P. Korpys-Woźniak, P. Kubiak, W. Białas, E. Celińska, Impact of overproduced heterologous protein characteristics on physiological response in *Yarrowia lipolytica* steady-state-maintained continuous cultures, Appl. Microbiol. Biotechnol. 104 (2020) 9785–9800, https://doi.org/10.1007/s00253-020-10937-w.
[26] J.M. Nicaud, C. Madzak, P. Van Den Broek, C. Gysler, P. Duboc, P. Niederberger, C. Gaillardin, Protein expression and secretion in the yeast *Yarrowia lipolytica*, FEMS Yeast Res. 2 (2002) 371–379, https://doi.org/10.1016/S1567-1356(02)00082-X.
[27] E. Celińska, M. Borkowska, W. Białas, P. Korpys, J.M. Nicaud, Robust signal peptides for protein secretion in *Yarrowia lipolytica*: identification and characterization of novel secretory tags, Appl. Microbiol. Biotechnol. 102 (2018) 5221–5233, https://doi.org/10.1007/s00253-018-8966-9.
[28] E. Celińska, M. Borkowska, P. Korpys-Woźniak, M. Kubiak, J. Nicaud, P. Kubiak, M. Gorczyca, W. Białas, Optimization of *Yarrowia lipolytica*-based consolidated biocatalyst through synthetic biology approach: transcription units and signal peptides shuffling, Appl. Microbiol. Biotechnol. 104 (2020) 5845–5859, https://doi.org/10.1007/s00253-020-10644-6.
[29] E. Celińska, R. Ledesma-Amaro, M. Larroude, T. Rossignol, C. Pauthenier, J. M. Nicaud, Golden Gate Assembly system dedicated to complex pathway manipulation in *Yarrowia lipolytica*, Microb Biotechnol. 10 (2017) 450–455, https://doi.org/10.1111/1751-7915.12605.
[30] P. Soudier, M. Larroude, E. Celińska, T. Rossignol, J.M.J.-M. Nicaud, Selection of heterologous protein-producing strains in *Yarrowia lipolytica*, in: B. Gasser, D. Mattanovich (Eds.), Methods in Molecular Biology, Springer Science+Business Media, LLC, part of Springer Nature, 2019, pp. 153–168, https://doi.org/10.1007/978-1-4939-9024-5_6.
[31] P. Korpys-Woźniak, M. Kubiak, M. Borkowska, E. Celińska, Construction and Assembly of Standardized Biobricks for Synthetic Pathways Engineering in Yeasts, in: F.Darvishi Harzevili (Ed.), Synthetic Biology of Yeasts: Tools and Applications, Springer International Publishing, Cham, 2022, pp. 27–62, https://doi.org/10.1007/978-3-030-89680-5_2.
[32] M. Kubiak-Szymendera, B. Skupień-Rabian, U. Jankowska, E. Celińska, Hyperosmolarity adversely impacts recombinant protein synthesis by *Yarrowia lipolytica* – molecular background revealed by quantitative proteomics, Appl. Microbiol. Biotechnol. 106 (2021) 349–367, https://doi.org/10.1007/s00253-021-11731-y.
[33] J. Sambrook, D. Russell, Molecular Cloning: A Laboratory Manual, 3rd ed., Cold Spring Harbor Laboratory Press, New York, 2001.
[34] M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, EMBnet J. 17 (2011) 10–12, https://doi.org/10.14806/ej.17.1.200.
[35] D. Kim, J. Paggi, C. Park, C. Bennett, S. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype, Nat. Biotechnol. 37 (2019) 907–915, https://doi.org/10.1038/s41587-019-0201-4.Graph-Based.
[36] S. Anders, P.T. Pyl, W. Huber, HTSeq-A Python framework to work with high-throughput sequencing data, Bioinformatics 31 (2015) 166–169, https://doi.org/10.1093/bioinformatics/btu638.
[37] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (2014) 1–21, https://doi.org/10.1186/s13059-014-0550-8.
[38] H. Mi, A. Muruganujan, X. Huang, D. Ebert, C. Mills, X. Guo, P. Thomas, Protocol update for large-scale genome and gene function analysis with PANTHER Classification System (v.14.0), Nat. Protoc. 14 (2019) 703–721, https://doi.org/10.1038/s41596-019-0128-8.Protocol.
[39] H. Mi, D. Ebert, A. Muruganujan, C. Mills, L.-.P. Albou, T. Mushayamaha, P. D. Thomas, PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API, Nucleic Acids Res. 49 (2021) D394–D403, https://doi.org/10.1093/nar/gkaa1106.
[40] L.D. Goldstein, Y. Cao, G. Pau, M. Lawrence, T.D. Wu, S. Seshagiri, R. Gentleman, Prediction and quantification of splice events from RNA-Seq Data, PLoS ONE 11 (2016), e0156132, https://doi.org/10.1371/journal.pone.0156132.
[41] J.T. Robinson, H. Thorvaldsdóttir, W. Winckler, M. Guttman, E.S. Lander, G. Getz, J.P. Mesirov, Integrative genomics viewer, Nat. Biotechnol. 29 (2011) 24–26, https://doi.org/10.1038/nbt.1754.
[42] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The sequence alignment/map format and SAMtools, Bioinformatics 25 (2009) 2078–2079, https://doi.org/10.1093/bioinformatics/btp352.
[43] T. Hall, Bioedit: a user-friendly biological sequence aligmnet editor and analysis program for Windows 95/98/NT, Nucleic Acids Symp. Ser. 41 (1999) 95–98.
[44] F. Corpet, Multiple sequence alignment with hierarchical clustering, Nucleic Acids Res. 16 (1988) 10881–10890, https://doi.org/10.1093/nar/16.22.10881.
[45] K. Löchli, E. Torbica, M. Haile-Weldeslasie, D. Baku, A. Aziz, D. Bublak, S. Fragkostefanakis, Crosstalk between endoplasmic reticulum and cytosolic unfolded protein response in tomato, Cell Stress Chaperones (2022), https://doi.org/10.1007/s12192-022-01316-7.
[46] K.J. Travers, C.K. Patil, L. Wodicka, D.J. Lockhart, J.S. Weissman, P. Walter, Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation, 2000. https://www.cell.com/cgi/content/full/101/3/249/DC1.
[47] J.S. Cox, P. Walter, A novel mechanism for regulating activity of a transcription factor that controls the unfolded protein response, Cell 87 (1996) 391–404.
[48] H. Yoshida, K. Haze, H. Yanagi, T. Yura, K. Mori, Identification of the cis-acting endoplasmic reticulum stress response element responsible for transcriptional induction of mammalian glucose- regulated proteins: involvement of basic leucine zipper transcription factors, J. Biol. Chem. 273 (1998) 33741–33749, https://doi.org/10.1074/jbc.273.50.33741.
[49] H. Yoshida, T. Matsui, A. Yamamoto, T. Okada, K. Mori, XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to er stress to produce a highly active transcription factor, Cell 107 (2001) 881–891.
[50] M. Kabani, J.-.M. Beckerich, C. Gaillardin, Sls1p stimulates Sec63p-mediated activation of Kar2p in a conformation-dependent manner in the yeast endoplasmic reticulum, Mol. Cell Biol. 20 (2000) 6923–6934, https://doi.org/10.1128/mcb.20.18.6923-6934.2000.
[51] B. Mai, L. Breeden, Xbp1, a stress-induced transcriptional repressor of the *Saccharomyces cerevisiae* Swi4/Mbp1 family, Mol. Cell Biol. 17 (1997) 6491–6501, https://doi.org/10.1128/mcb.17.11.6491.
[52] B. Mai, L. Breeden, CLN1 and its repression by Xbp1 are important for efficient sporulation in budding yeast, Mol. Cell Biol. 20 (2000) 478–487, https://doi.org/10.1128/mcb.20.2.478-487.2000.
[53] T. Iwawaki, M. Tokuda, Function of yeast and amphioxus tRNA ligase in IRE1alpha-dependent XBP1 mRNA splicing, Biochem. Biophys. Res. Commun. 413 (2011) 527–531, https://doi.org/10.1016/j.bbrc.2011.08.129.
[54] K.B. Hooks, S. Griffiths-Jones, Conserved RNA structures in the non-canonical Hac1/Xbp1 intron, RNA Biol. 8 (2011) 552–556, https://doi.org/10.4161/rna.8.4.15396.
[55] C.K. Patil, H. Li, P. Walter, Gcn4p and novel upstream activating sequences regulate targets of the unfolded protein response, PLoS Biol. 2 (2004), https://doi.org/10.1371/journal.pbio.0020246.