Article:
Kaneva, I.N., Longworth, J., Sudbery, P.E. et al. (1 more author) (2018) Quantitative Proteomic Analysis in Candida albicans using SILAC-based Mass Spectrometry. Proteomics, 18 (5-6). 1700278. ISSN 1615-9853

https://doi.org/10.1002/pmic.201700278

Quantitative proteomic analysis in *Candida albicans* using SILAC-based mass spectrometry

Iliyana N. Kaneva[1,2], Joseph Longworth[1], Peter E. Sudbery*[2], Mark J. Dickman*[1]

[1]ChELSI Institute, Department of Chemical and Biological Engineering, University of Sheffield, Mappin Street, Sheffield, S1 3JD, United Kingdom.

[2] Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, UK

*corresponding authors

Keywords: SILAC; quantitative proteomics; *Candida albicans;* native SILAC

Abstract

Stable isotope labelling by amino acids in cell culture (SILAC) in conjunction with mass spectrometry analysis is a sensitive and reliable technique for quantifying relative differences in protein abundance and post-translational modifications between cell populations. We have developed and utilised SILAC-MS workflows for quantitative proteomics in the fungal pathogen *Candida albicans*. Arginine metabolism provides important cues for escaping host defences during pathogenesis, which limits the use of auxotrophs in *Candida* research. Our strategy eliminates the need for engineering arginine auxotrophs for SILAC experiments and allows the use of *ARG4* as selectable marker during strain construction. Cells that were auxotrophic for lysine were successfully labelled with both lysine and arginine stable isotopes. We found that prototrophic *C. albicans* preferentially uses exogenous arginine and downregulates internal production, which allowed it to achieve high incorporation rates. However, similar to other yeast, *C. albicans* was able to metabolise heavy arginine to heavy proline, which compromised the accuracy of protein quantification. A computational method was developed to correct for the incorporation of heavy proline. In addition, we utilised the developed SILAC labelling in *Candida albicans* for the global quantitative proteomic analysis of a strain expressing a phosphatase-dead mutant Cdc14^{PD}.

Significance of the study

The fungus *Candida albicans* is commonly found on the skin and mucosal surfaces of healthy people, but in susceptible individuals, it causes infections ranging from superficial to life threatening. In this study we have developed and utilised SILAC-MS workflows for quantitative proteomics in the fungal pathogen *Candida albicans*. In *C. albicans*, arginine is commonly used as a selectable marker in strain engineering. Arginine metabolism provides important cues for escaping host defences during pathogenesis, which limits the use of auxotrophs in *Candida* research. In this study we have utilised a strategy that eliminates the need for engineering arginine auxotrophs for SILAC experiments and allows the use of *ARG4* as selectable marker during strain construction in *Candida albicans*. We believe the application of nSILAC in Candida offers a powerful tool for quantitative proteomic studies in the fungal pathogen *Candida albicans*. This opens up new possibilities for studying protein interactions, post-translational modifications and whole proteome changes in response to different factors that occur during the lifecycle of this fungus.

Introduction

Stable isotope labelling by amino acids in cell culture (SILAC) is a metabolic labelling technique that has been widely used for accurate quantitative mass spectrometry (MS) studies [1, 2]. In principle, two cell populations are grown in the presence of either light (naturally occurring) or heavy isotope enriched amino acids, such as lysine and arginine, and incorporate them into their proteomes. Cells from both cultures are harvested and mixed together, so that downstream sample processing does not introduce quantitative errors. Subsequent MS analysis of the peptide digests facilitates global protein characterization and quantification. The light and heavy peptides form “SILAC pairs” in the survey spectrum due to the mass difference of the isotopes. The labels allow the origin of peptides to be traced back to the starting culture and thus SILAC is used to detect relative differences in protein abundance or post-translational modifications between samples. The isotopes have no effect on cell metabolism, so any differences found by SILAC-MS can be attributed to the experimental conditions or genetic make-up of the cells.

In SILAC experiments, it is imperative that cells use the exogenous amino acids present in the media and do not synthesise them on their own. For this reason, studies have typically employed auxotrophic mutants, where genes involved in lysine and/or arginine biosynthesis have been deleted. However, recently it has been demonstrated that the proteomes of prototrophic yeast and bacteria can also be metabolically labelled with heavy lysine [4]. In these organisms, the supplemental lysine in the media inhibited its own production and achieved similar incorporation rates to auxotrophic strains. Importantly, this was shown to occur only in exponentially growing cells. In stationary phase, cells resumed biosynthesis of lysine and the labelling efficiency steadily decreased [5]. Nevertheless, this approach known as native SILAC (nSILAC) is a valuable alternative to traditional SILAC methods, especially in investigations where generation of auxotrophic strains is not possible or not practical.

The possibility of performing nSILAC with arginine prototrophs has not been explored yet. The use of heavy arginine for metabolic labelling has been more problematic because many organisms, including yeast, use it as a precursor for synthesis of other amino acids, mainly proline [6, 7]. As a result, each proline-containing peptide from the heavy-labelled strain forms a cluster of monoisotopic peaks on the mass spectrum, with only the highest peak intensity used in quantitation. This reduces the intensity of the heavy peptides and skews the calculated peptide ratio towards the light version. To solve this issue, several strategies have been proposed to either correct peptide intensity post-quantitation, or prevent the conversion from happening in the first place. The rate of arginine-to-proline conversion may be modulated by carefully titrating the availability of both amino acids in the culture medium [8-11]. An excess of arginine and limited proline are likely to stimulate the conversion. On the other hand, shortage of arginine may lead to incomplete labelling, reduced growth rate and disruptions in metabolic processes. However, high concentrations of proline are not advised either, because it can also be converted to arginine and, again, compromise peptide quantitation. A more effective strategy has been described in *Schizosaccharomyces pombe*, where deletion of several arginase genes prevented the conversion of arginine to other amino acids [12, 13]. In addition the use of $^{15}N_4$-labelled arginine in the light medium and $^{13}C_6{}^{15}N_4$-labelled arginine in the heavy medium has also been used, allowing an internal correction to be applied at the quantitation step [14]. Lastly, several bioinformatic approaches have been developed to correct the intensity of proline-containing peptides [15-17].

Another issue related to the use of heavy arginine is that some species can metabolise it to different analogues. The most common version of heavy arginine used for metabolic labelling is Arg10. In fission yeast, Arg10 is converted not only to proline, but also to Arg7, resulting in additional lower molecular weight “pre-peaks” in the isotopic clusters of arginine-containing peptides [13]. This compromises the sensitivity of MS and significantly reduces the number of identified peptides, and limits peptide quantitation.

The fungus *Candida albicans* is commonly found on the skin and mucosal surfaces of healthy people, but in susceptible individuals, it causes infections ranging from superficial to life threatening [18]. The transition from commensal to pathogenic state is a multifactorial event that is still not fully understood. Important strategies that allow *C. albicans* to adapt to a wide range of conditions include biofilm formation, morphogenetic plasticity and secretion of adhesins and hydrolases [19, 20]. A better understanding of the molecular pathways governing these events is urgently needed in order to develop more effective treatments for *C. albicans* infections.

Although metabolic labelling is ideally performed in conjunction with essential amino acids, this is not always possible. In *C. albicans*, arginine is commonly used as a selectable marker in strain engineering. Therefore, nSILAC in conjunction with arginine prototrophs would be a valuable alternative to traditional SILAC approaches. This is particularly relevant to studies investigating *C. albicans* host defence and morphological switching. Arginine biosynthesis is strongly upregulated in the early stages of cell phagocytosis by macrophages [21]. In this scenario, arginine is converted to urea, and the latter is degraded to CO_2, which induces yeast-to-hypha transition and allows cells to escape the attack [22]. Arginine deficient strains cannot make the hyphal switch inside macrophages, and therefore prototrophs are required for studies investigating these processes.

In this study, we have developed and optimised SILAC labelling of *C. albicans* proteome for high throughput quantitative proteomics. We show that labelling with heavy arginine does not require the use of auxotrophic mutants, but necessitates additional correction of peptide quantitation in order to account for heavy proline conversion. Furthermore, as proof of principle, we have carried out the first global quantitative proteomic analysis using SILAC in *C. albicans* looking at proteome changes that occur in response to overexpressing an inactivated phosphatase $Cdc14^{PD}$ involved in cell separation, mitotic exit and morphogenesis.

Materials and methods

Generation of strains – All strains were derived from the *Candida albicans* strain MDL04 (*lys2::CmLEU2/lys2::CdHIS1 arg4Δ/arg4Δ leu2Δ/leu2Δ his1Δ/his1Δ ura3Δ::imm434/ura3Δ::imm434 iro1Δ::imm434/iro1Δ::imm434*), which was obtained from Carol Munro from The University of Aberdeen. *CDC14* was fused to *MYC-URA3* using a PCR cassette as previously described [23]. *CDC14-MYC* was amplified from genomic DNA by PCR using the primers 5'-GGCGGGCTCGAGGGCTTTCCTTTCCTTTGCTATG-3' and 5'-GGCGGGTCTAGACTAATTTGTGAGTTTAGTATACATGC-3' and cloned into pRSC3 plasmid via ligation with BamHI and XbaI and using the linker oligonucleotides GATCCCTCCCAGAAC and TCGAGTTCTGGGAGG. This resulted in plasmid pINK1 containing *CDC14-MYC-URA3*, in which the point mutation CDC14C275S was inserted by PCR amplification using the primers 5'-GCAGTACATTCTAAAGCAGGGTTAGG-3' and 5'-CCTAACCCTGCTTTAGAATGTACTGC-3'. *Cdc14C275S-MYC-URA3* was excised and transformed into MDL04 as above. A regulatable promoter sequence *ARG4-MET3* was fused to *cdc14C275S* as described by Gola, et al. using the primers 5'-AAATGTATATAACGAAGATGACTATCATCAATGGTCCGGTTAGTAAAGCGAACAAGCTTTATAAAAATAGTTA TGCTGAACGTACCATGAAGCTTCGTACGCTGCAGGTC-3' and 5'-AAAGGTAGAACAATCAATTTGAAGTAGATTTTCCCAACATACTTTTAAGAAACTCTATAAGAGGCACATGAAC CAGTGAACTATGCATGTTTTCTGGGGAGGGTATTTAC-3' [24]. Successful transformation was confirmed by PCR, DNA sequencing and Western blot. In the article, this strain is written as $Cdc14^{PD}$.

SILAC media and growth conditions –MET3 promoter-repressing media containing: 6.7 g/L Difco yeast nitrogen base without amino acids, 20 g/L D-glucose, 670 mg/L Formedium™ complete supplement mixture drop-out: -arginine, -lysine, -methionine, 60.6 mg/L methionine, 373 mg/L cysteine and 80 mg/L uridine. MET3 promoter-inducing media contained the same ingredients, except no methionine and no cysteine were added. In addition, light medium contained 80 mg/L light arginine (Arg0: $^{12}C_6$, $^{14}N_4$) and 80 mg/L light lysine (Lys0: $^{12}C_6$, $^{14}N_2$), both supplied by Sigma

Aldrich. In addition, heavy medium contained 100 mg/L heavy arginine (Arg10: $^{13}C_6$, $^{15}N_4$) and 100 mg/L heavy lysine (Lys8: $^{13}C_6$, $^{15}N_2$), both supplied by Cambridge Isotope Laboratories, Inc. Cells were routinely grown overnight at 30 °C, 200 rpm in 50 ml of either light or heavy MET3 promoter-repressing media. In the morning, each culture was re-inoculated into 0.5 L light or heavy MET3 promoter-inducing media at OD_{595}=0.25 and cells were allowed to grow for approximately 4 hrs, at 30 °C, 200 rpm until they reached OD_{595}=0.7. At this point cells were harvested by centrifugation, frozen in liquid nitrogen and stored at -80 °C.

Sample preparation and fractionation – Cell pellets were thawed on ice and re-suspended in 5 ml cold lysis buffer (20 mM HEPES, 150 mM NaCl, EDTA-free protease inhibitors (Roche), 1mM PMSF, pH 7.4). For SILAC experiments, equal amount of light and heavy cells were mixed together and broken in a high pressure cell disrupter (Constant Systems Ltd.) at 35 psi, 4 °C. Lysates were cleared of cell debris by centrifugation and proteins were denatured by heating and separated by SDS-PAGE. Polyacrylamide gels were stained with Coomassie InstantBlue and all protein bands were excised and digested with trypsin as previously described [25]. Briefly, protein bands were excised and destained by washing with 50 mM ammonium bicarbonate (ABC) and acetonitrile (ACN). Dehydrated gel bands were soaked in digestion buffer (12.5 ng/µl trypsin in 50 mM ABC) overnight and peptides extracted by incubating with ACN, 5% formic acid.

LC-MS –Samples re-suspended in 0.1% TFA were analysed on an Ultimate 3000 online nano liquid chromatography system with PepMap300 C18 trapping column (ThermoFisher), coupled to Q Exactive HF hybrid Quadrupole Orbitrap (ThermoFisher). Peptides were eluted onto a 50 cm x 75 µm Easy-spray PepMap C18 analytical column. During analysis the column temperature was maintained at 40 °C. Peptides were eluted at a flow rate of 300 nL/min using a gradient of 3% to 40% over 60m. Analytical buffers were composed of 0.1% Formic acid (FA) and either 3% ACN (Buffer A) or 80% ACN (Buffer B). 0.1% TFA and 3% ACN (Loading buffer). Spray voltage was set to 2.1 kV, S-lens RF level at 60, and heated capillary at 250 °C. MS1 scan resolution was set to 120 000 at m/z 200 and MS2

189 resolution was set to 15,000 at m/z 200. Full scan target was $3*10^6$ with a maximum fill time of 100
190 ms. Mass range was set to 375 - 1500. Target value for fragment scans was set at $5*10^4$, and
191 intensity threshold was kept at $5*10^4$. Isolation width was set at 1.2. A fixed first mass of 100 was
192 used. Normalized collision energy was set at 28. Peptide match was set to Preferred, and isotope
193 exclusion was on. All data was acquired in profile mode using positive polarity.

194 *Data analysis* – Raw files were processed by MaxQuant v.1.5.2.8 using Andromeda search engine
195 [26]. All files from a single experiment were processed together and peptide searches were
196 performed against the *Candida albicans* protein database SC5314 Assembly 21 downloaded from
197 candidagenome.org. Arg10 and Lys8 were specified as heavy labels and up to 3 arginine or lysine
198 labelled amino acids per peptide were allowed. Methionine oxidation and Pro6 were set as variable
199 modifications and up to 6 per peptide were allowed. Specific digestion with trypsin was chosen with
200 no more than 2 missed cleavages allowed. Re-quantify was selected to identify missing SILAC pairs.
201 The peptide tolerance was 4.5 ppm and intensity threshold was set to 500. Peptides were allowed to
202 be 8-25 amino acids long and up to 4600 Da. False discovery rates (FDR) for peptide spectral
203 matches (PSM) and proteins was limited to 0.01. Ratios were then corrected using in house
204 constructed shiny application (Shiny-Pro6Correction) written in R to determine the median isotopic
205 ratio for evidence peptides lacking proline and comparing it to the median isotopic ratio of evidence
206 peptides containing a single proline without isotopic labelling [27]. Using this difference ratio as a
207 correction factor, the observed isotopic ratio of all peptides is corrected based on their proline
208 content (corrected peptide ratio = peptide ratio*(1+correction factor)*number of prolines in the
209 peptide). The median corrected isotopic ratio for all evidence peptides within a protein group is then
210 reported as the corrected protein isotopic ratio with a normalised correction reporting the
211 adjustment giving a median ratio of 1. Shiny-Pro6Correction is implemented entirely in the R
212 language. It can be hosted/launched by any system with R installed, including Windows, Mac OS and
213 most Linux distributions. Information technology administrators can also host Shiny-
214 Pro6Correction from a remote server, in which case users need only have a Web browser installed.

Shiny-Pro6Correction is provided free of charge under MIT open-source license through GitHub at https://josephlongworth.github.io/Shiny-Pro6Correction/. Statistical analysis of the data was done in Perseus, where proteins significantly enriched in either heavy or light isotopes were identified using the Significance B test based on Benjamini-Hochberg FDR 0.05 [26]. The efficiency of heavy amino acid incorporation was calculated using the formula: Incorporation = (H/L)/(H/L+1)×100, where H/L is the normalised ratio of heavy-to-light peptide intensities generated by the MaxQuant software. Incorporation was calculated for individual peptides first, and then the average of all peptides. The mass spectrometry data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD007825 [28].

Results and Discussion

Strain construction and SILAC media formulation.

We constructed a *C. albicans* strain expressing an inactive phosphatase Cdc14^{PD} that forms stable interactions with its substrates, but is not able to dephosphorylate them. In the course of engineering the strain, it was necessary to use the *ARG4* gene as a selectable marker, which renders the cells prototrophic for arginine and auxotrophic for lysine. The mutation did not affect the growth rate or the phenotype of *C. albicans* yeast at least for the duration of SILAC experiments, indicating that Cdc14 is a haplosufficient allele. Cdc14^{PD} was also able to form true hyphae for at least 3 hours after promoter induction. Cdc14^{PD} was put under the control of a regulatable *MET3* promoter, while the *ARG4* gene was controlled by a separate promoter.

For the purpose of SILAC experiments, Cdc14^{PD} cells were grown in heavy medium ($^{8}K^{10}R$) and wild type cells were grown in light medium. In order to keep the growth conditions of both strains as similar as possible, both media were prepared to be either *MET3*-inducing or *MET3*-represing, although wild type cells did not have an ectopic MET3 promoter. The optimal concentration of lysine was empirically determined to be 0.44 mM. At these concentrations, cells grew at the same rate in both media (Supplementary figure 1). The same concentration of arginine was used in media formulation. The results show that neither the mutant allele CDC14^{PD}, nor the heavy isotope-containing media affected the growth of *C. albicans*. In addition, alternative media (including hyphae inducing media) could also be suitable for SILAC, as long as they do not contain the light versions of amino acids used for labelling.

Efficiency of SILAC metabolic labelling.

Cdc14^{PD} cells were grown overnight in heavy medium allowing approximately 10 cycles of division. Soluble protein extracts were prepared and fractionated by SDS-PAGE, and a single fraction was

analysed by MS to determine the labelling efficiency (see supplementary figures 2 and 3). 63.6% of the peptides identified contained lysine (but not arginine) with a Lys8 incorporation of 95.43% (see supplementary figure 3). Further analysis of the heavy labelling of *C. albicans* was performed by examining individual MS spectra of lysine containing peptides (see figure 1). The results show that the vast majority of peptides displayed a single isotope cluster corresponding to the heavy labelled peptide. Signals corresponding to the light peptides were not observed or comparable to noise levels.

Analysis of individual MS spectra of arginine containing peptides show that unlabelled peptides had significantly lower signal intensities than heavy peptides, typically below 10% (see figure 2). In addition, the global peptide analysis revealed that the heavy arginine labelling efficiency was 89.54%. These results demonstrate that *C. albicans* preferentially use amino acids from the media even when they are not essential for survival. Although the incorporation efficiency of Arg10 was not as high as that of Lys8, it was sufficient for the purpose of SILAC experiments.

Almost a third of all peptides identified (28.1%) contained arginine but not lysine. If cells were not labelled with Arg10, these peptides would not be included for protein quantification, resulting in significant loss of data. Therefore, we achieved a more accurate protein quantification by using both arginine and lysine incorporation during SILAC labelling. Since two thirds of the peptides contain lysine rather than arginine, the overall isotope incorporation was 93.87%. Thus, our strategy of double labelling in lys^- arg^+ cells will generate improved quantitative mass spectrometry data compared to using Lys8 only as a single label. Furthermore, these results suggest that *C. albicans* is a suitable organism for nSILAC.

Correction of arginine to proline conversion.

Arginine conversion has been well characterised in *S. cerevisiae* [29], and this process is likely to occur in *C. albicans*, which has orthologues of all enzymes involved in the reaction. Cellular metabolism of Arg10 may potentially lead to biosynthesis of heavy proline, in particular Pro6 and additional heavy arginine, namely Arg7. We examined the heavy labelled sample for the presence of Pro6 and Arg7 and found that proline conversion occurred at a rate of 18.9% (see supplementary figure 4), but Arg7 was not detected. Arginine to proline conversion was also analysed by examining individual MS spectra of proline containing peptides (see figure 3A/B). The results clearly demonstrate the additional isotope clusters present in the spectra of proline containing peptides due to the incorporation of heavy proline. The incorporation of heavy proline will result in an underestimation of the heavy:light ratio using automated software analysis such as MaxQuant. Therefore, we developed a computational method to correct for the heavy arginine to proline conversion by measuring the change in observed peptide evidence isotopic ratio, stratified by proline content. The heavy isotope underestimation was determined based on the peptides containing a single proline residue. Peptides with one potential conversion provide the superior measurement of the error observed given their greater abundance. This correction factor could then be applied to all proline-containing peptides based on their respective proline content.

The isotopic ratio distribution for peptides depending on their proline content is shown in figure 3C/D pre and post correction. Peptides were sorted according to the number of proline residues comparing the median H/L ratio. Peptides containing a single proline were determined to have a 35.7% lower H/L ratio than peptides with no prolines (see figure 3C). Therefore, a correction factor of 0.357 per proline was applied to the affected peptides. After correction, peptide ratios were normalised so that the median H/L ratio was adjusted to 1. Peptides in all groups displayed similar H/L ratios regardless of their proline count (see figure 3C). The alignment of the distributions post correction indicate the true differential in the experiment. It should be noted that whilst Pro6

containing peptides were not utilised in the determination of protein quantifications, they were included in the original search as their observations are valid to support protein identifications.

In order to test if our method correctly amends the isotope ratios, as a positive control we mixed wild type cells grown in heavy and light SILAC medium in a 1:1 ratio prior to LC MS analysis. Samples were prepared and analysed as previously described (see supplementary figure 5A). The results show that only those peptides with no prolines had a median H/L peptide ratio of 1. In contrast, proline-containing peptides had a significantly lower H/L ratio, which decreased proportionally with the number of proline residues (see supplementary figure 5A). After applying the Shiny-Pro6 correction, all H/L peptide ratios were similar and close to 1, demonstrating that the Shiny-Pro6 method corrects the H/L peptide ratio errors caused by the arginine-to-proline conversion.

We further compared our peptide ratio normalisation method to the MaxQuant normalisation method. The H/L peptide ratios (not corrected for Pro6 and not normalised by MaxQuant) were normalised by both Shiny-Pro6 and MaxQuant (see supplementary figure 5B). The results show the excellent correlation between the different normalisation methods, further demonstrating that Shiny-Pro6 correctly normalises the H/L ratios after correcting for arginine to proline conversion.

Quantitative proteomic analysis of Cdc14^{PD} overexpression in *C. albicans*.

Following the demonstration of the SILAC labelling of *C. albicans* proteome with heavy amino acids (arginine and lysine) and the development of computational methods to correct for the proline conversion, we performed a proof of principle application of SILAC labelling of *C. albicans* and measured the global proteomic changes that occur in response to overexpressing the inactive phosphatase Cdc14^{PD}. Cdc14^{PD} cells were grown in the presence of Lys8 and Arg10, while wild type

cells were grown in media containing only light amino acids. Both strains were grown in heavy MET3-repressing media overnight, and then in heavy MET3-inducing media for 4 hours. Thus, $Cdc14^{PD}$ was overexpressed for only 4 hours. Wild type cells were grown in light media in the same conditions. Cells from both strain were harvested and mixed together prior to cell lysis, protein extraction and GeLC MS analysis. The results are summarised in supplementary Table 1 and Figure 4.

Proteome analysis of *C. albicans* overexpressing an inactive phosphatase $Cdc14^{PD}$ resulted in the identification of 2541 proteins representing 40.1% of the cellular proteome, with 2339 proteins quantified (see supplemental tables 1 and 2). The quantitative SILAC MS analysis revealed that 103 proteins were significantly upregulated (see supplementary table 3), while 28 proteins were significantly downregulated in $Cdc14^{PD}$ compared to the wild type proteome (see figure 4). A gene ontology (GO) analysis of the 131 differentially expressed proteins revealed that almost a third (29.8%) are involved in oxidation-reduction processes and 12.2% take part in monocarboxilic acid metabolic processes (see supplementary figure 6 and supplementary table 4). Oxidoreductases play an important role in regulating CDC14B in mammalian cells, where the phosphatase activity is inhibited by high intracellular concentrations of H_2O_2 [30].Therefore, such changes in the proteome may reflect the presence of $Cdc14^{PD}$ above physiological levels as well as the different amino acid requirements of the mutant strain compared to the wild type strain. Further experiments should be performed to elucidate the relationship between Cdc14 and the proteins presented here and further analyse the interacting partners of Cdc14 in *C. albicans*.

Conclusions

In this study we have for the first time developed and utilised SILAC-MS workflows for quantitative proteomics in the fungal pathogen *Candida albicans*. The fungus *Candida albicans* is commonly found on the skin and mucosal surfaces of healthy people, but in susceptible individuals, it causes infections ranging from superficial to life threatening. In *C. albicans*, arginine is commonly used as a selectable marker in strain engineering. Arginine metabolism provides important cues for escaping host defences during pathogenesis, which limits the use of auxotrophs in *Candida* research. In this study we have utilised a strategy that eliminates the need for engineering arginine auxotrophs for SILAC experiments and allows the use of *ARG4* as selectable marker during strain construction in *Candida albicans*. This opens up new possibilities for studying protein interactions, post-translational modifications and whole proteome changes in response to different factors that occur during the lifecycle of this fungus.

We have developed and used nSILAC in *C. albicans*, although cells were able to synthesise arginine internally, we were able to achieve approximately 90% heavy arginine incorporation, which was slightly lower than that of the essential amino acid lysine. Previously, nSILAC has been applied in baker's and budding yeast using heavy lysine [4]. A protein turnover study done in *S. cerevisiae* showed that the proteome was almost fully labelled in actively growing cells, but cells resumed lysine biosynthesis in stationary phase [5]. In our experiments, cells were grown in heavy medium overnight, when they inevitably reached stationary phase several hours before being re-inoculated into fresh medium. Thus, it is likely that cells were feeding on supplemental heavy arginine in the exponential phase, and switched to production of light arginine in stationary phase. Cells were then allowed to re-enter exponential growth for 4 hours, and likely downregulate the rate of endogenous arginine production. It is proposed that the efficiency of arginine incorporation could be improved by harvesting the cells before they reached stationary phase of growth.

Catabolic conversion of arginine to proline is a well-known problem in SILAC experiments using heavy arginine as a label. In *C. albicans*, almost a fifth of all proline residues were synthesised using Arg10 as a precursor. This metabolic conversion can be prevented by titrating concentrations of both arginine and proline in the media. However, this is difficult to achieve in prototrophic organisms because cells need sufficient amount of heavy arginine in order to shut down endogenous production. Engineering an arginase-deficient strain is also possible, but not practical in *C. albicans* as previously described. Therefore, we developed a computational method to correct the protein SILAC (H/L) ratios, which normalises the ratios of proline-containing peptides to those of proline-free peptides. We performed a proof of principle application of SILAC labelling of *C. albicans* and measured the global proteomic changes that occur in response to overexpressing the inactive phosphatase $Cdc14^{PD}$. The analysis resulted in the quantification of 2339 proteins of which 131 proteins were significantly up/down regulated compared to the wild type cells.

Acknowledgements

INK was a student funded by the BBSRC White Rose DTP (BB/J014443/1). MJD and PES acknowledge support from the Biotechnology and Biological Sciences Research Council UK (BB/M012166/1 and BB/J002305/1) and MJD further acknowledges support from Engineering and Physical Sciences Research Council UK.

Figure Legends:

Fig. 1: ESI-MS of ^{8}K-labelled peptides. Analysis of the MS spectra of the tryptic peptides **A)** LYGGAAGEGAGGAGDAK and **B)** MKETAEGFLGTTVK from cells grown on heavy SILAC ($^{8}K^{10}R$) labelled media reveals no significant signal corresponding to the ^{0}K peptide consistent with the high average incorporation efficiency (95.43%) determined from the global analysis of heavy lysine peptides. Blue and red squares show the position of light (^{0}K) and heavy (^{8}K) peptides respectively.

Fig. 2: Incorporation of ^{10}R in prototrophic *C. albicans*. Analysis of the MS spectra of the tryptic peptides **A)** VEIIANDQGNR and **B)** VDEIVLVGGSTR from cells grown on heavy SILAC ($^{8}K^{10}R$) labelled media reveals a small signal corresponding to the ^{0}R peptide in each case consistent with the lower average incorporation efficiency of heavy arginine (89.54%) determined from the global peptides analysis of heavy arginine peptides. Blue and red squares show the position of light (^{0}R) and heavy (^{10}R) peptides respectively.

Fig. 3: Arginine to proline conversion. MS spectra of the tryptic proline-containing peptides **A)** GPLVVYAQDNGIVK and **B)** GHDIPHPITTFDEAGFPDYVLQEVK from *Candida* cells grown on heavy ($^{8}K^{10}R$) and light ($^{0}K^{0}R$) SILAC media mixed in a 1:1 ratio. Blue squares show the position of light (^{0}K) peptides and red squares show the position of the heavy isotope clusters from the ($^{8}K/^{6}P^{8}K$) peptides respectively. **C)** Scatter plots of the peptides H/L ratio versus proline content before and after correction. **D)** A box plot showing a global protein analysis of the H/L ratios before and after correction for heavy arginine to proline conversion.

Fig. 4: Quantitative proteomic analysis in *C. albicans* using SILAC. A) Plot of protein intensities against normalized H/L SILAC ratios of WT *C. albicans* cells grown in light ($^{0}K^{0}R$) media and Cdc14^{PD} *C. albicans* cells grown heavy ($^{8}K^{10}R$) SILAC media. Significant outliers were determined using the Significance B test and are coloured in blue (FDR < 0.05); other proteins are shown in grey. **B)** Histogram showing the normalised heavy to light protein ratio distributions. The frequency represents the number of proteins in each bin.

References

[1] Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., Kuster, B., *Analytical and bioanalytical chemistry* 2007, *389*, 1017-1031.

[2] Zhang, G., Fenyo, D., Neubert, T. A., *Journal of proteome research* 2009, *8*, 1285-1292.

[3] Chen, X., Wei, S., Ji, Y., Guo, X., Yang, F., *Proteomics* 2015, *15*, 3175-3192.

[4] Frohlich, F., Christiano, R., Walther, T. C., *Mol Cell Proteomics* 2013, *12*, 1995-2005.

[5] Martin-Perez, M., Villen, J., *Analytical chemistry* 2015, *87*, 4008-4014.

[6] Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., Mann, M., *Mol Cell Proteomics* 2002, *1*, 376-386.

[7] Gruhler, A., Olsen, J. V., Mohammed, S., Mortensen, P., Færgeman, N. J., Mann, M., Jensen, O. N., *Mol Cell Proteomics* 2005, *4*, 310-327.

[8] Lossner, C., Warnken, U., Pscherer, A., Schnolzer, M., *Analytical biochemistry* 2011, *412*, 123-125.

[9] Blagoev, B., Mann, M., *Methods (San Diego, Calif* 2006, *40*, 243-250.

[10] Schmidt, F., Strozynski, M., Salus, S. S., Nilsen, H., Thiede, B., *Rapid Commun Mass Spectrom* 2007, *21*, 3919-3926.

[11] Bendall, S. C., Hughes, C., Stewart, M. H., Doble, B., Bhatia, M., Lajoie, G. A., *Mol Cell Proteomics* 2008, *7*, 1587-1597.

[12] Bicho, C. C., de Lima Alves, F., Chen, Z. A., Rappsilber, J., Sawin, K. E., *Mol Cell Proteomics* 2010, *9*, 1567-1577.

[13] Borek, W. E., Zou, J., Rappsilber, J., Sawin, K. E., *PloS one* 2015, *10*, e0129548.

[14] Van Hoof, D., Pinkse, M. W., Oostwaard, D. W., Mummery, C. L., Heck, A. J., Krijgsveld, J., *Nature methods* 2007, *4*, 677-678.

[15] Park, S. K., Liao, L., Kim, J. Y., Yates, J. R., 3rd, *Nature methods* 2009, *6*, 184-185.

[16] van Breukelen, B., van den Toorn, H. W., Drugan, M. M., Heck, A. J., *Bioinformatics (Oxford, England)* 2009, *25*, 1472-1473.

[17] Pasculescu, A., Schoof, E. M., Creixell, P., Zheng, Y., Olhovsky, M., Tian, R., So, J., Vanderlaan, R. D., Pawson, T., Linding, R., Colwill, K., *Journal of proteomics* 2014, *100*, 167-173.

[18] Kibbler, C. C., Seaton, S., Barnes, R. A., Gransden, W. R., Holliman, R. E., Johnson, E. M., Perry, J. D., Sullivan, D. J., Wilson, J. A., *The Journal of hospital infection* 2003, *54*, 18-24.

[19] Nobile, C. J., Johnson, A. D., *Annual review of microbiology* 2015, *69*, 71-92.

[20] Schaller, M., Borelli, C., Korting, H. C., Hube, B., *Mycoses* 2005, *48*, 365-377.

[21] Lorenz, M. C., Bender, J. A., Fink, G. R., *Eukaryotic cell* 2004, *3*, 1076-1087.

[22] Ghosh, S., Navarathna, D. H., Roberts, D. D., Cooper, J. T., Atkin, A. L., Petro, T. M., Nickerson, K. W., *Infection and immunity* 2009, *77*, 1596-1605.

[23] Lavoie, H., Sellam, A., Askew, C., Nantel, A., Whiteway, M., *BMC genomics* 2008, *9*, 578.

[24] Gola, S., Martin, R., Walther, A., Dunkler, A., Wendland, J., *Yeast (Chichester, England)* 2003, *20*, 1339-1347.

[25] Pandey, A., Andersen, J. S., Mann, M., *Sci STKE* 2000, *2000*, pl1.

[26] Cox, J., Mann, M., *Nature biotechnology* 2008, *26*, 1367-1372.

[27] Chang, W., Cheng, J., Allaire, J. J., Xie, Y., McPherson, J., 2017.

[28] Vizcaino, J. A., Csordas, A., Del-Toro, N., Dianes, J. A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q. W., Wang, R., Hermjakob, H., *Nucleic acids research* 2016, *44*, 11033.

[29] Ong, S. E., Kratchmarova, I., Mann, M., *Journal of proteome research* 2003, *2*, 173-181.

[30] Lim, J. M., Lee, K. S., Woo, H. A., Kang, D., Rhee, S. G., *The Journal of cell biology* 2015, *210*, 23-33.