Published in final edited form as:

J Proteome Res. 2011 November 4; 10(11): 5251–5259. doi:10.1021/pr200551e.

A SILAC compatible strain of *Pichia pastoris* for expression of isotopically labeled protein standards and quantitative proteomics

Ryan J. Austin¹, Rolf E. Kuestner¹, Debbie K. Chang¹, Knut R. Madden³, and Daniel B. Martin^{1,2,*}

¹Institute for Systems Biology, 401 Terry Ave. N., Seattle, WA 98109

²Seattle Cancer Care Alliance, 825 Eastlake Avenue E., Seattle, WA 98109

³BioGrammatics Inc., 2120-D Las Palmas Dr., Carlsbad, CA 92011

Abstract

The methylotrophic yeast *Pichia pastoris* is a powerful eukaryotic platform for the production of heterologous protein. Recent publication of the *P. pastoris* genome has facilitated strain development toward biopharmaceutical and environmental science applications, and has advanced the organism as a model system for the study of peroxisome biogenesis and methanol metabolism. Here we report the development of a *P. pastoris* arg-/lys- auxotrophic strain compatible with SILAC (stable isotope labeling by amino acids in cell culture) proteomic studies, which is capable of generating large quantities of isotopically labeled protein for mass spectrometry-based biomarker measurements. We demonstrate the utility of this strain to produce high purity human serum albumin uniformly labeled with isotopically heavy arginine and lysine. In addition, we demonstrate the first quantitative proteomic analysis of methanol metabolism in *P. pastoris*, reporting new evidence for a malate-aspartate NADH shuttle mechanism in the organism. This strain will be a useful model organism for the study of metabolism and peroxisome generation.

Keywords

Pichia pastoris; Absolute SILAC; SILAC; quantitative mass spectrometry; methanol assimilation; methanol dissimilation; peroxisome biogenesis

Introduction

Pichia pastoris is the most widely used yeast platform for heterologous protein expression. The methylotroph was developed as an expression system by Phillips Petroleum and the Salk Institute Biotechnology/Industrial Associates (SIBIA, La Jolla, CA), and offers the advantages of 1) tightly controlled methanol inducible transgene expression under the alcohol oxidase (AOX1) promoter, and 2) respiratory growth at high cell density¹. The

*Dr. Daniel Martin can be reached at the Seattle Cancer Care Alliance. dmartin@seattlecca.org; phone: (206) 288-1284; fax: (206) 732-2187.

Supporting Information

Supplemental materials include detailed proteomic information for SILAC datasets; detailed methods for the construction of the *P. pastoris* auxotroph with PCR confirmation of ARG4 and LYS2 gene disruption (Fig S1, S2); analysis of RBP4 expression in *Pichia* auxotroph (Fig S3); trypsin digestion of *Pichia* SILAC lysates (Fig S4); blast alignment of *S. cerevisiae* aspartate-malate carrier protein (Fig S5); table of H/L protein ratios at 24 hours growth on methanol (Table S1) and 48 hours growth (Table S2); GO-enrichment for genes up-regulated on methanol (Table S3); and GO-enrichment for genes down-regulated on methanol (Table S4).

molecular genetics of *P. pastoris* are analogous to *S. cerevisiae* allowing for simple integration of transgenes into the yeast genome via homologous recombination and a variety of *P. pastoris* strains and selectable markers have been characterized ², with expression kits commercially available for intracellular and secreted protein production (*Invitrogen* and *BioGrammatics Inc.*). Importantly, *P. pastoris* is capable of performing many of the posttranslational modifications (PTM's) found in higher eukaryotes, such as folding, disulfide bond formation, and glycosylation. To date, over 500 functional proteins have been reported to be produced in *P. pastoris* ³, with yields typically much greater than achieved by insect and mammalian cell cultures. While many eukaryotic expression systems require complex media reagents, *P. pastoris* can be cultured to high densities on a minimal defined media allowing for inexpensive scale-up and purification of secreted protein directly from the culture. High bioactivities of human proteins expressed in *P. pastoris* have recently made the yeast a popular platform for generation of monoclonal antibodies and other protein therapeutics ³, with reports demonstrating the potential of further glyco-engineering strains for biopharmaceutical applications ⁴. One additional application for which *P. pastoris* may be particularly well suited, is the recombinant production of stable isotope labeled proteins for NMR and mass spectrometry (MS)-based measurements ⁵.

MS-quantification of absolute protein biomarker levels in tissues, blood, and other biofluids is demonstrating increasing utility in medical diagnoses ⁶. Presently, many clinically relevant biomarkers are measured by quantitative immunoassay, a technology that has the advantage of high sensitivity and relatively low cost where robust immunoreagents are not available. However, the increased sensitivities of targeted MS are presenting MS as a viable alternative to immuno-based methods, particularly in the research setting ^{7,8}. Absolute SILAC, an MS-technique developed by Matthias Mann and coworkers ⁹, achieves accurate quantification of proteins in complex biological mixtures by spiking full-length isotopically labeled protein standards into samples before co-purification and MS-analysis. Incorporation of ¹³C¹⁵N-labeled heavy arginine and lysine residues into protein standards results in predictable mass shifts in the tryptic peptides of these standards, which are detectable by MS. This differential across multiple peptides allows for robust absolute quantification of endogenous protein. The technique, however, is limited by the need to purify a full-length protein standard uniformly labeled with heavy arginine and lysine that is highly similar to the target biomarker. Absolute SILAC standards have previously been expressed in *E. coli* and *S. cerevisiae* strains auxotrophic for arginine and lysine ^{9,10}, allowing complete and specific incorporation of isotopic label provided in the growth media. The capability of *P. pastoris* to perform many eukaryotic PTM's and the established profile of functional protein therapeutics expressed in the organism make the yeast an attractive platform for the production of Absolute SILAC standards. Towards this application, we have engineered a *P. pastoris* strain that is auxotrophic for arginine and lysine. We have tested the secreted expression of isotopically labeled standards in this strain using the protein human serum albumin (HSA) and have established the fidelity of label incorporation to approach 100% by high mass accuracy tandem MS (MS/MS).

In addition to its utility for heterologous protein expression, our auxotrophic *Pichia* strain expands the available model organisms in which quantitative shotgun proteomics can be performed using SILAC (stable isotope labeling by amino acids in cell culture) ¹¹. In a typical SILAC workflow, two populations of differentially stimulated cells are grown up separately on media containing either light or heavy arginine and lysine amino-acids. Incorporation of ¹³C¹⁵N-labeled heavy amino-acids into cellular proteomes results in a predictable mass shift in the tryptic peptides of these proteins that is detectable by mass spectrometry and this differential allows for the relative proteome quantification of heavy versus light samples ¹². Basing our MS-analysis on the recently published *P. pastoris* genome ¹³, we have performed a SILAC study of methanol growth in our auxotrophic strain.

P. pastoris is a model organism for the investigation of methanol metabolism¹⁴, and as expected we observe the key enzymes in methanol metabolism to be dramatically up-regulated on methanol growth media. We also find novel evidence for a malate-aspartate NADH shuttle mechanism, which may function in the methanol dissimilation pathway. The work represents the first quantitative proteomic analysis of methanol metabolism in *P. pastoris* and offers a template for investigation of other model behaviors in the organism, including peroxisome biogenesis¹⁵, pexophagy¹⁶, and vesicle secretion studies¹⁷.

Experimental Procedures

P. pastoris auxotroph BG08 *arg4 lys2*

An *arg*-/*lys*- auxotrophic strain of *P. pastoris* (*BG08 arg4 lys2*) was generated by sequential knockout of the ARG4 and LYS2 genes. Briefly, electrocompetent BG08 cells provided by BioGrammatics Inc., (Carlsbad, CA) were electroporated with 2.1 µg linearized ARG4 knockout vector (dmrk22) and plated on 200 µg/mL nourseothricin. Selected nourseothricin resistant colonies were screened by replica plating on minimal media ± arginine to identify auxotrophs and these colonies were assessed for specific disruption of ARG4 by specific PCR amplification. Clone ARG- 1A (*BG08 arg4*) was electroporated with 2.0 µg linearized LYS2 knockout vector (dmrk23), selected on 200 µg/mL hygromycin, and screened on *arg*+/*lys*-, *arg*+/*lys*+ minimal media. The disruption of LYS2 in *BG08 arg4 lys2* was confirmed by PCR. Sequences of knockout vectors; sequences of the disrupted ARG4 and LYS2 loci, and elaborated methods are included in the supplemental materials.

Heterologous protein expression in *BG08 arg4 lys2*

Secreted HSA was expressed under control of the alcohol oxidase promoter (AOX1) in *BG08 arg4 lys2* using the expression vector pJAZ-HSA (*BioGrammatics*). This vector incorporates the HSA ORF with its native leader sequence targeting the protein for secretion. Briefly, electrocompetent *BG08 arg4 lys2* cells were electroporated with 2.0 µg linearized pJAZ-HSA and plated on 100 µg/mL zeocin. Individual transformants were grown up as 2 mL cultures and induced to secrete recombinant HSA on methanol-based media. A 2 mL culture of BMG media [100 mM potassium phosphate (pH 6.0); 1.34% YNB (w/v); 16 nM d-biotin; 1% glycerol (v/v); and add-back aminoacids: Asp (80 mg/L), Iso (50 mg/L), Leu (100 mg/L), Met (20 mg/L), Phe (50 mg/L), Thr (200 mg/L), Trp (50 mg/L), Tyr (50 mg/L), Val (140 mg/L), Pro (200 mg/L), Glu (80 mg/L), Adenine (20 mg/L); isotopically light or heavy Arg (50 mg/L) and Lys (50 mg/L) (*Isotech*)] containing glycerol as its carbon source was inoculated from a single colony and incubated overnight at 250 rpm, 30 °C. Cells were spun down at 2,000 rcf and exchanged into 2 mL methanol-based BMM media [100 mM potassium phosphate (pH 6); 1.34% YNB (w/v); add-back amino acids; 16 nM d-biotin; 0.5% MeOH (v/v)] to induce recombinant protein expression for 60 hrs. Cultures were supplemented with additional methanol [100 µl 10% methanol (v/v)] at 24, 36, and 48 hr time points following induction. At 60 hrs, cultures were spun down at 3,000 rcf to remove cells. Secreted HSA in the culture media was concentrated by Centricon Spin-Filter YM-10 (*Amicon*) prior to MS-analysis. Similar *P. pastoris* expression vectors incorporating an α mating factor leader sequence were used to express secreted RBP4 and hGH in the auxotroph strain following similar protocols (pJAZαMF-RBP4 and pJAZαMF-hGH, *BioGrammatics, Inc.*).

Methanol/glycerol SILAC in *BG08 arg4 lys2*

2 mL cultures of *BG08 arg4 lys2* were inoculated in either heavy or light BMG media [100 mM potassium phosphate (pH 6); 1.34% YNB (w/v); 16 nM d-biotin; 1% glycerol (v/v); and add-back amino-acids] and grown 24 hrs at 250 rpm, 30 °C. Heavy cultures were switched into BMM media [100 mM potassium phosphate (pH 6); 1.34% YNB (w/v); add-back

amino acids; 16 nM d-biotin; 0.5% MeOH (v/v)] and grown to 24 and 48 hr time points with methanol supplementation. Cells were spun down at 3,000 rcf, washed and resuspended in Breaking Buffer [50 mM sodium phosphate, pH 7.4; 1 mM PMSF; 1 mM EDTA] to an OD₆₀₀ = 50. Light and heavy samples were mixed 1:1 (normalized by OD₆₀₀) and mechanically lysed using acid treated beads (*Scientific Industries Inc.*) vortexed for 30 minutes at 4 °C. The soluble lysate was separated from beads and cellular debris by centrifugation (14,000 rcf, 10 minutes). Lysates were quantified by Bradford assay and 25 µg samples aliquoted for subsequent trypsinization and MS analysis.

Proteomic Analysis / Mass Spectrometry

HSA and normalized SILAC samples were prepared for MS analysis in similar fashion with proteins denatured, reduced, alkylated, and trypsinized using an SDS/Urea-based denaturing protocol¹⁸. Detergent was removed from the tryptic digest by multiple cation exchanges over an Oasis MCX extraction cartridge (*Waters Inc.*) and subsequently desalted by UltraMicroSpin Vydac C18 silica column following the manufacture's specifications (*Nest Group, Inc.*). Tryptic peptides were dissolved in loading buffer [0.1% formic acid; 1% acetonitrile; 98.9% water], trapped on a fused silica fritted capillary precolumn packed with 2 cm reverse-phase Magic C18Aq RP spherical silica (75 µm ID, 5 µm, 200 Å; *Michrom Bioresources*), and separated over a 15 cm reverse-phase Magic C18Aq RP analytical column (50 µm ID, 5 µm, 100 Å). The gradient program was a 60 min linear gradient from 2-35% acetonitrile at a flow rate of 0.35 µl/min (*Agilent 1100 Series LC system*). Nanospray ESI MS/MS analysis was performed using a *Thermo Scientific* LTQ Orbitrap. MS/MS was acquired over a range of 50-2000 m/z using a 60 sec dynamic exclusion time and a 35.0 V collision voltage.

MS spectra were converted to universal mzXML file format by ReAdW version 4.3.1¹⁹ and searched against protein database FASTA files for human (NCBI) or *P. pastoris* strains GS115 and DSMZ 70382 (Uniprot). Searches were performed using X!Tandem²⁰ with the following parameters: tolerable tryptic termini = 1; identifications based on b and y ions; parent mass tolerance = 3.00; daughter ion mass tolerance = 0.50; fixed modifications include carboxyamidomethylation of cysteine (57.02); variable modifications include oxidation of methionine (15.99), SILAC heavy arginine: ¹³C6-¹⁵N4 (10.01), SILAC heavy lysine: ¹³C6-¹⁵N2 8.01, SILAC heavy proline ¹³C5-¹⁵N (6.01). MS/MS peptide assignments were validated by PeptideProphet²¹ and protein assignments validated by Protein Prophet²². Statistical analysis and error rate information are included in the supplemental materials. Quantitative SILAC ratios for proteins were determined using XPRESS software²³, available in the current TPP distribution²⁴. Precursor ion elution profiles of heavy vs. light peptides were determined with a mass tolerance of 0.05. The area under the curve (AUC) was used to determine a SILAC ratio for each peptide. Outlier AUC ratios were identified by RelEx least squares regression fit of AUC profiles²⁵ and removed to generate a parsimonious dataset. The uncertainties of SILAC ratios were determined for each protein expression level for which multiple peptide measurements were available. Finally, protein SILAC ratios for the 24 and 48 hour datasets were normalized about a median H/L ratio = 1 for subsequent analyses.

For SILAC experiments 24 and 48 hr sample time points gave 259 and 212 high-confidence protein IDs (probability > 0.9), respectively. Of these identifications, 186 proteins were redundant to both data sets. The 48 hour data set was taken as our standard for functional analyses (Table S2). *P. pastoris* protein ID's were mapped to *S. cerevisiae* protein orthologs using InParanoid 7²⁶, to generate an input list for Gene-GO enrichment analysis by DAVID²⁷. Twenty-three up-regulated genes (log H/L > 0.5) were assessed for molecular function enrichment over a background of 200 identified genes. Protein function associated with peroxisome, aspartate binding-site, and microbody GO-terms exhibited the greatest

fold-enrichment (Table S3). A converse analysis of 33 down-regulated genes ($\log H/L < -0.5$) showed high enrichment for ribosomal RNA processing (Table S4).

Results and Discussion

Design of *P. pastoris* auxotroph BG08 *arg4 lys2*

We have created the first SILAC-compatible strain of *P. pastoris* for recombinant expression of protein standards labeled with $^{13}C^{15}N$ heavy arginine and lysine. Specific labeling of these residues ensures that all tryptic peptides of the protein standard, with the exception of the carboxy-terminal peptide, will contain a basic amino-acid with a known isotopic shift detectable by LC-MS/MS. In selecting a reference strain of *P. pastoris*, we considered those strains for which genomic information is available, including GS115²⁸, DSMZ 70382¹³, X33 (*Integrated Genomics Inc.*) and BG08 (*BioGrammatics Inc.*); BG08 was selected because of the open source availability of this strain for research purposes. The auxotrophic *P. pastoris* strain (*BG08 arg4 lys2*) was created by replacing internal codons of the ARG4 and LYS2 genes with drug resistant genes using established yeast genetics²⁹. Knockout vector schematics are shown in figure 1. Gene substitution of ARG4 (arginosuccinate lyase) and LYS2 (α -amino adipate reductase) was confirmed by PCR amplification of the genomic DNA. As expected, *BG08 arg4 lys2* showed no growth on minimal media lacking either arginine or lysine and a crude lysate exhibited complete incorporation of $^{13}C^{15}N$ labeled arginine and lysine when grown on heavy SILAC media, as determined by MS/MS (data not shown). Growth characteristics of the auxotroph were similar to the parent Bg08 strain ("wild type") on glycerol and methanol based media with an electroporation efficiency of 1.5×10^5 cfu/ μ g, conducive to molecular cloning work. The possible metabolic conversion of arginine to proline in the *BG08 arg4 lys2* strain was a concern, as this phenomena has previously been reported in the yeast *S. pombe* and can result in non-uniform protein labeling³⁰. However, no evidence for any arginine to proline conversion was found by searching MS/MS data for incorporation of heavy proline.

Recombinant expression of isotopically labeled HSA

Protein quantification by Absolute SILAC requires that a pure calibrated heavy protein standard be spiked into a sample at a concentration similar to the target protein being measured. The permissible concentration difference between target and standard is theoretically limited by the dynamic range of MS-detection (100 to 1000-fold), however impurities in the isotopic standard can narrow this window considerably. To evaluate the potential of our auxotrophic strain to produce protein standards uniformly labeled at high-fidelity, recombinant expression of secreted human serum albumin (HSA) was tested. HSA levels in urine are diagnostic/prognostic for renal disease and cardiovascular morbidity and the biomarker has recently been quantified by LC-MS/MS in human studies³¹. The expression construct pJAZ-HSA coding for a secreted HSA under the control of the host's inducible alcohol oxidase (AOX1) promoter was transformed into *BG08 arg4 lys2*. Heterologous expression of HSA was induced by switching the growth culture from a glycerol to methanol based media, which resulted in a dramatic increase in relative expression of the recombinant protein (Fig 2a). A combination of strong induction and the low baseline secretome with expression regulated by the *P. pastoris* AOX1 promoter allowed for relatively pure recombinant HSA to be harvested directly from the culture media. A recent report demonstrated that recombinant HSA (67 kDa) is degraded in *P. pastoris* by the aspartic protease yapsin, to generate a 43 kDa amino-terminal HSA fragment³². A similar fragment is evident with HSA expressions in *Bg08 arg- lys-*, although nominal when cultures are induced in a potassium buffered media (pH 6.0, < 2% by image densitometry). The expression level of HSA on heavy media in shakeflask cultures was measured to be greater than 70 mg/L by both Bradford and BCA protein quantification

methods. High mass accuracy MS/MS of isotopically labeled HSA identified 31 unique peptide sequences accounting for 55% of the full-length protein sequence (Fig 2b). Twenty-two of these peptides gave quantifiable light to heavy SILAC ratios (L/H), which show the fraction of light amino acid incorporation in HSA to be less than 1% (Fig 2c, Table 1), consistent with the purity of isotopic amino-acids employed (>98% $^{13}\text{C}^{15}\text{N}$). Secreted expression of uniformly labeled HSA isotopic standard in the *P. pastoris* BG08 arg4 lys2 demonstrates the utility of this strain for production of MS biomarker standards. HSA is one of many potential plasma and urine biomarker proteins that, when spiked into bio-specimens, will provide quality controls by producing multi-peptide quantification standards on digestion. And with inexpensive production scale up, heavy HSA also presents a simple MS standard for normalizing run-to-run variability and batch controls for quantitative analysis. Furthermore, the expression of additional plasma protein standards including retinol binding protein 4 (RBP4) and human growth hormone (hGH) with the *Pichia* auxotroph have been performed and demonstrate secretion yield and purity comparable to HSA (Fig S3). Both RBP4 and hGH are present at low concentrations in the serum and have previously been quantified by MS *in situ*^{33, 34}. Like HSA they can also now serve as internal controls for proteomic studies of serum.

Quantitative proteomic analysis of methanol metabolism in *P. pastoris*

P. pastoris is an important organism in industry and basic science. A systems-based understanding of *P. pastoris* could expand the range of applications for the organism, with potential value in protein therapeutics^{4, 35} and biofuels³⁶. Beyond these uses, *P. pastoris* is a valuable model system for the study of methanol metabolism¹⁴, peroxisome biogenesis¹⁵, pexophagy¹⁶, and vesicle secretion¹⁷. Recently, systems level studies of these processes have been accelerated by publication of the *P. pastoris* genome^{28, 37}, which has facilitated metabolomic³⁸, transcriptomic³⁹, and proteomic investigations^{37, 40-42}. Quantitative analysis of the *P. pastoris* proteome has, however, not yet been reported.

To investigate the utility of the BG08 arg4 lys2 strain for proteomic analysis of *P. pastoris*, a SILAC-based study of methanol metabolism in the organism was performed. BG08 arg4 lys2 cells were cultured to saturation on glycerol-based media supplemented with either light or heavy arginine and lysine amino acids. Light cells were harvested at time point zero, while heavy cells were transferred to methanol-based media with heavy arginine and lysine and further incubated for 24 or 48 hours. Light and heavy samples were normalized and mixed prior to protein sample preparation for MS. The resulting ratio of heavy to light protein in the mixed samples corresponds to the relative expression of proteins on methanol versus glycerol at time point zero. Two hundred forty four of 259 proteins identified at the 24 hour time point and 197 of 212 proteins identified at the 48 hour time point gave quantifiable heavy:light (H/L) SILAC ratios. Of these proteins, 186 were identified with high confidence in both samples. Complete tables of protein IDs and SILAC ratios at 24 and 48 hours are provided in the supplemental materials (Tables S1, S2).

As expected, the proteins associated with peroxisome function and methanol metabolism are up-regulated on methanol media. A plot of H/L ratios at 48 hours shows the distribution of *P. pastoris* protein levels in methanol versus a glycerol baseline (Fig 3a). A statistical analysis of protein gene ontology (GO)-enrichment shows those proteins up-regulated on methanol ($\log_{10} \text{H/L} > 0.5$) to be enriched for peroxisomal and aspartate-binding molecular functions while down-regulated proteins ($\log_{10} \text{H/L} < -0.5$) are enriched for RNA-processing and ribosomal synthesis functions (Tables S3, S4). Proteins exhibiting H/L ratios within a half-log of the mean ($-0.5 < \log_{10} \text{H/L} < 0.5$) comprised 79% and 66% of quantified IDs at 24 and 48 hours, respectively. Methanol metabolism in *P. pastoris* involves two primary pathways: 1) methanol assimilation to cellular carbon, and 2) methanol dissimilation into carbon dioxide with commensurate production of energy (Fig 3b). Key to

the assimilation pathway are the peroxisomal enzymes alcohol oxidase (uniprot: C4R917; gene ID: AOX) and dihydroxyacetone synthase (C0LQF4; DAS), which are both known to be dramatically up-regulated in cells grown on methanol 2. The principal isoforms of these two enzymes, corresponding to the AOX1 and DAS1 genes, show the highest up-regulation (70-fold), consistent with the known promoter function of these two genes. The secondary isoforms, DAS2 and AOX2, are also induced on methanol, along with pyruvate carboxylase (C4R339), which participates in the trafficking of alcohol oxidase to the peroxisome⁴³. Two subsequent enzymatic steps in the assimilation of methanol are performed outside of the peroxisome by dihydroxyacetone kinase (C4R5Q6; DAK) and fructose 1,6 biphosphate aldolase (C4QWS2), both of which are strongly up-regulated at 48 hours. Together, four of the five proteins most up-regulated in our experiment are known to function in the methanol assimilation pathway.

The expression of proteins involved in methanol dissimilation, the second primary pathway of methanol metabolism (for energy production), are also up-regulated; all cytoplasmic enzymes in the pathway show quantifiable induction. These include: S-hydroxymethyl dehydrogenase (C4R6A5; FLD), S-formyl glutathione hydrolase (C1PHG5; FGH), and formate dehydrogenase (C1PHG6; FDH). Several additional proteins identified as highly up-regulated are likely to function in the production of energy from methanol, including the protein inhibitor of ATP hydrolysis (C4QWR9), which enhances ATP generation in the mitochondria; and nicotinamidase (C4R4B7, C4QUZ5), which functions in recycling of NAD⁺ via the salvage pathway. Of further interest are the proteins aspartate aminotransferase (C4QWE4, C4R862) and malate dehydrogenase (C4R911, C4R024), which potentially participate in the generation of ATP by a malate-aspartate NADH shuttle (MAS) mechanism.

Oxidation of formaldehyde to CO₂ in the dissimilation pathway generates cytoplasmic NADH that must be transported into the mitochondria to generate ATP via electron transport. In *S. cerevisiae* and higher eukaryotes this is accomplished using a high-yield MAS mechanism reliant upon 4 enzymes: mitochondrial and cytosolic malate dehydrogenases, aspartate aminotransferases, the aspartate/glutamate carrier (AGC) and the α -ketoglutarate/malate carrier (OGC)⁴⁴. Evidence for MAS function in *S. cerevisiae* comes from functional studies of the mitochondrial aspartate-glutamate transporter (Agc1p), which was identified through sequence homology of the carboxy-terminal domain to human AGC and functionally characterized through reconstitution in phospholipid vesicles⁴⁵. To our knowledge, no evidence for a malate-aspartate NADH shuttle mechanism in *P. pastoris* has previously been reported, although the production of cytosolic NADH in the methylotroph suggests the advantage of such a mechanism. While an aspartate/glutamate carrier protein has not presently been identified in *P. pastoris*, a candidate protein, C4R4J8, belongs to the appropriate subclass of calcium-binding mitochondrial carrier proteins and shares 51% protein sequence identity with the Agc1p carboxy-terminal domain (Fig S5).

In addition to the primary products of methanol metabolism, hydrogen peroxide and reactive oxygen species are byproducts of methanol breakdown in the peroxisome. The highly efficient enzyme catalase decomposes hydrogen peroxide to oxygen. We observe it to be significantly up-regulated in *Pichia* grown on methanol along with other proteins that likely function to protect the peroxisome from oxidative stress, including two superoxide dismutases (C4R8X7, C4QXC7), membrane localized stress response protein (C4R8G2) and cytosolic glutathione reductase (C1PHG3) (Table S2). The glutathione peroxidase PMP20 (C1PHG2), known to detoxify reactive oxygen species in the peroxisome⁴⁶, was not identified at the 48 time point, but was identified and observed to be up-regulated 8-fold at 24 hours. Separately, two mitochondrial enzymes, mitochondrial aldehyde dehydrogenase (C4R6P6) and mitochondrial peroxiredoxin (C4QX37), are significantly up-regulated on

methanol, consistent with a bimodal distribution of these enzymes between mitochondria and peroxisomes¹⁵. Among the list of proteins highly up-regulated on methanol, is proteinase B (C4QYT0) (Fig 2a). This enzyme presumably functions in the cellular stress response to methanol, however, we observe the counterpart, proteinase A (C4R6G8), which is required to activate proteinase B, to be down-regulated 7-fold. The balance of these vacuolar proteases, along with carboxypeptidase, is thought to be related to the processes of micro/macro-pexophagy in *P. pastoris*¹⁶.

Beyond static time point analyses, SILAC measurements may be expanded to quantify the time course of protein expression in organisms. A time course of *P. pastoris* on methanol at 0, 24, and 48 hours, illustrates the yeast's metabolic dynamics in finer detail (Fig 4). We observe, for instance, the level of dihydroxyacetone synthase increasing faster than that of alcohol oxidase. This relationship had previously been predicted as a mechanism for *P. pastoris* to prevent high peroxisomal formaldehyde concentrations¹⁵. More general trends are also apparent in the symmetry of the time course data. Beginning at the zero time point when the culture media is switched from glycerol to methanol, the methanol metabolic enzymes show a precipitous up-regulation that is shadowed by gradual decline of glycolytic enzymes necessary for the metabolism of glycerol.

The results clearly demonstrate the utility of SILAC in *Pichia* to study the proteomics of a model behavior with a temporal resolution that is akin to transcriptomic qRT-PCR. This work offers a proteomic map of methanol metabolism in *P. pastoris* and presents novel findings pursuable in the organism including a putative MAC function in the methanol dissimilation pathway.

Conclusion

Many in the field expect that the next decade will see proteomics make inroads in identifying biomarkers useful for clinical conditions of all types. Heavy protein markers such as HSA, RBP4, and hGH demonstrated here will serve as crucial standardization tools in this endeavor. *P. pastoris* is ideally suited in this role given its known success in producing secreted proteins. The arg- lys- auxotrophic strain of *Pichia* provides a novel route to obtaining isotopically heavy proteins quickly and at a reasonable cost. The *Pichia*-produced whole protein standards are less vulnerable to quantification errors resulting from the use of single peptide standards⁴⁷, where adsorption of the peptide during preparation can cause unexpected variations in the standard and unexpected PTM's can unpredictably affect the level of unmodified endogenous peptide. Furthermore, the *P. pastoris* system demonstrated is archival, distributable, and easily manipulated owing to the organism's history of commercialization (the parent strain, auxotroph, *BG08 arg4 lys2*, and recombinant pJAZ-HSA, pJAZ-RBP4, pJAZ-hGH strains are available from BioGrammatics).

The *BG08 arg4 lys2* strain also provides an excellent tool for quantitative proteomic analysis of methanol metabolism¹⁴, peroxisome biogenesis¹⁵, pexophagy¹⁶, and vesicle secretion models in *P. pastoris*¹⁷. In combination with transcriptomic studies, SILAC brings the power of systems biology to model studies, as well as application driven efforts, in *P. pastoris* including glycoengineering⁴, unnatural protein expression³⁵, and biofuel processing³⁶. The findings with regard to energy metabolism stage a starting point for further investigation of a MAS mechanism in *P. pastoris* and have bearing on the understanding of disordered metabolism in humans, including diabetes and the metabolic syndrome⁴⁸; both diseases expected to be huge national cost drivers in the coming decades.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH/NIGMS grant P50 GM076547. We would like to thank Dr. Ramsey Saleem for helpful criticism on the manuscript and Dr. Patrick May who provided bioinformatics support for InParanoid 7 gene list conversions.

Abbreviations

AGC	aspartate/glutamate carrier
AOX	alcohol oxidase
arg	arginine
ARG4	arginosuccinate lyase
ATP	adenosine triphosphate
BCA	bicinchoninic acid assay
BMG	buffered minimal glycerol
BMM	buffered minimal methanol
DAK	dihydroxyacetone kinase
DAS	dihydroxyacetone synthase
DAVID	database for annotation visualization and integrated discovery
FDH	formate dehydrogenase
FGH	S-formyl glutathione hydrolase
FLD	S-hydroxymethyl dehydrogenase
GO	gene ontology
hGH	human growth hormone
HSA	human serum albumin
LC	liquid chromatography
lys	lysine
LYS2	aminoadipate semialdehyde dehydrogenase
MAS	malate-aspartate NADH shuttle
MCX	multiple cation exchange
MRM	multiple reaction monitoring
MS	mass spectrometry
MS/MS	tandem mass spectrometry
NAD⁺/NADH	nicotinamide adenine dinucleotide
NMR	nuclear magnetic resonance
OGC	α -ketoglutarate/malate carrier
qRT-PCR	quantitative reverse transcription polymerase chain reaction

RBP4	retinol binding protein 4
RP	reverse phase
SIBIA	Salk Institute Biotech/Industrial Associates
SILAC	stable isotope labeling by amino acids in cell culture
PTM	post-translational modification
TPP	trans-proteomic pipeline
YNB	yeast nitrogen base

References

1. Cregg JM, Vedvick TS, Raschke WC. Recent advances in the expression of foreign genes in *Pichia pastoris*. *Biotechnology (N Y)*. 1993; 11(8):905–10. [PubMed: 7763913]
2. Cereghino JL, Cregg JM. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol Rev*. 2000; 24(1):45–66. [PubMed: 10640598]
3. Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast*. 2005; 22(4):249–70. [PubMed: 15704221]
4. Jacobs PP, Geysens S, Vervecken W, Contreras R, Callewaert N. Engineering complex-type N-glycosylation in *Pichia pastoris* using GlycoSwitch technology. *Nat Protoc*. 2009; 4(1):58–70. [PubMed: 19131957]
5. Whittaker JW. Selective isotopic labeling of recombinant proteins using amino acid auxotroph strains. *Methods Mol Biol*. 2007; 389:175–88. [PubMed: 17951643]
6. Vrana JA, Gamez JD, Madden BJ, Theis JD, Bergen HR 3rd, Dogan A. Classification of amyloidosis by laser microdissection and mass spectrometry-based proteomic analysis in clinical biopsy specimens. *Blood*. 2009; 114(24):4957–9. [PubMed: 19797517]
7. Elschnerbroich S, Kislinger T. Targeted proteomics by selected reaction monitoring mass spectrometry: applications to systems biology and biomarker discovery. *Mol Biosyst*. 7(2):292–303. [PubMed: 20976349]
8. Surinova S, Schiess R, Huttenhain R, Cerciello F, Wollscheid B, Aebersold R. On the development of plasma protein biomarkers. *J Proteome Res*. 10(1):5–16. [PubMed: 21142170]
9. Hanke S, Besir H, Oesterhelt D, Mann M. Absolute SILAC for accurate quantitation of proteins in complex mixtures down to the attomole level. *J Proteome Res*. 2008; 7(3):1118–30. [PubMed: 18271523]
10. Austin RJ, Smidansky H, Holstein C, Chang D, Martin D. Proteomic Analysis of the Androgen Receptor via MS-compatible Purification of Biotinylated Protein on Streptavidin Resin. manuscript in submission. 2011
11. Ong SE, Blagoev B, Kratichmarova I, Kristensen DB, Steen H, Pandey A, Mann M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics*. 2002; 1(5):376–86. [PubMed: 12118079]
12. Mann M. Functional and quantitative proteomics using SILAC. *Nat Rev Mol Cell Biol*. 2006; 7(12):952–8. [PubMed: 17139335]
13. Mattanovich D, Callewaert N, Rouze P, Lin YC, Graf A, Redl A, Tiels P, Gasser B, De Schutter K. Open access to sequence: browsing the *Pichia pastoris* genome. *Microb Cell Fact*. 2009; 8:53. [PubMed: 19835590]
14. Hartner FS, Glieder A. Regulation of methanol utilisation pathway genes in yeasts. *Microb Cell Fact*. 2006; 5:39. [PubMed: 17169150]
15. van der Klei IJ, Yurimoto H, Sakai Y, Veenhuis M. The significance of peroxisomes in methanol metabolism in methylotrophic yeast. *Biochim Biophys Acta*. 2006; 1763(12):1453–62. [PubMed: 17023065]

16. Dunn WA Jr, Cregg JM, Kiel JA, van der Klei IJ, Oku M, Sakai Y, Sibirny AA, Stasyk OV, Veenhuis M. Pexophagy: the selective autophagy of peroxisomes. *Autophagy*. 2005; 1(2):75–83. [PubMed: 16874024]
17. Payne WE, Kaiser CA, Bevis BJ, Soderholm J, Fu D, Sears IB, Glick BS. Isolation of *Pichia pastoris* genes involved in ER-to-Golgi transport. *Yeast*. 2000; 16(11):979–93. [PubMed: 10923020]
18. Tian Y, Zhou Y, Elliott S, Aebersold R, Zhang H. Solid-phase extraction of N-linked glycopeptides. *Nat Protoc*. 2007; 2(2):334–9. [PubMed: 17406594]
19. ReAdW. http://www.ionsource.com/functional_reviews/readw/t2x_update_readw.htm
20. Craig R, Beavis RC. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics*. 2004; 20(9):1466–7. [PubMed: 14976030]
21. Keller A, Eng J, Zhang N, Li XJ, Aebersold R. A uniform proteomics MS/MS analysis platform utilizing open XML file formats. *Mol Syst Biol*. 2005; 1:20050017.
22. Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem*. 2003; 75(17):4646–58. [PubMed: 14632076]
23. Han DK, Eng J, Zhou H, Aebersold R. Quantitative profiling of differentiation-induced microsomal proteins using isotope-coded affinity tags and mass spectrometry. *Nat Biotechnol*. 2001; 19(10):946–51. [PubMed: 11581660]
24. Trans-Proteomic Pipeline. <http://tools.proteomecenter.org/wiki/index.php?title=Software:TPP>
25. MacCoss MJ, Wu CC, Liu H, Sadygov R, Yates JR 3rd. A correlation algorithm for the automated quantitative analysis of shotgun proteomics data. *Anal Chem*. 2003; 75(24):6912–21. [PubMed: 14670053]
26. Ostlund G, Schmitt T, Forslund K, Kostler T, Messina DN, Roopra S, Frings O, Sonnhammer EL. InParanoid 7: new algorithms and tools for eukaryotic orthology analysis. *Nucleic Acids Res*. 38(Database issue):D196–203. [PubMed: 19892828]
27. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009; 4(1):44–57. [PubMed: 19131956]
28. De Schutter K, Lin YC, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, Rouze P, Van de Peer Y, Callewaert N. Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nat Biotechnol*. 2009; 27(6):561–6. [PubMed: 19465926]
29. Cregg JM, Tolstorukov I, Kusari A, Sunga J, Madden K, Chappell T. Expression in the yeast *Pichia pastoris*. *Methods Enzymol*. 2009; 463:169–89. [PubMed: 19892173]
30. Bicho CC, de Lima Alves F, Chen ZA, Rappsilber J, Sawin KE. A genetic engineering solution to the “arginine conversion problem” in stable isotope labeling by amino acids in cell culture (SILAC). *Mol Cell Proteomics*. 9(7):1567–77. [PubMed: 20460254]
31. Seegmiller JC, Barnidge DR, Burns BE, Larson TS, Lieske JC, Kumar R. Quantification of urinary albumin by using protein cleavage and LC-MS/MS. *Clin Chem*. 2009; 55(6):1100–7. [PubMed: 19325011]
32. Yao XQ, Zhao HL, Xue C, Zhang W, Xiong XH, Wang ZW, Li XY, Liu ZM. Degradation of HSA-AX15(R13K) when expressed in *Pichia pastoris* can be reduced via the disruption of YPS1 gene in this yeast. *J Biotechnol*. 2009; 139(2):131–6. [PubMed: 18977401]
33. Hergenroeder G, Redell JB, Moore AN, Dubinsky WP, Funk RT, Crommett J, Clifton GL, Levine R, Valadka A, Dash PK. Identification of serum biomarkers in brain-injured adults: potential for predicting elevated intracranial pressure. *J Neurotrauma*. 2008; 25(2):79–93. [PubMed: 18260791]
34. Wu SL, Amato H, Biringer R, Choudhary G, Shieh P, Hancock WS. Targeted proteomics of low-level proteins in human plasma by LC/MSn: using human growth hormone as a model system. *J Proteome Res*. 2002; 1(5):459–65. [PubMed: 12645918]
35. Young TS, Ahmad I, Brock A, Schultz PG. Expanding the genetic repertoire of the methylotrophic yeast *Pichia pastoris*. *Biochemistry*. 2009; 48(12):2643–53. [PubMed: 19265424]
36. Mousdale, DM. *Biofuels: biotechnology, chemistry, and sustainable development*. 1st ed. CRC Press; 2008.
37. Mattanovich D, Graf A, Stadlmann J, Dragosits M, Redl A, Maurer M, Kleinheinz M, Sauer M, Altmann F, Gasser B. Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. *Microb Cell Fact*. 2009; 8:29. [PubMed: 19490607]

38. Chung BK, Selvarasu S, Andrea C, Ryu J, Lee H, Ahn J, Lee H, Lee DY. Genome-scale metabolic reconstruction and in silico analysis of methylotrophic yeast *Pichia pastoris* for strain improvement. *Microb Cell Fact*. 9:50. [PubMed: 20594333]
39. Graf A, Gasser B, Dragosits M, Sauer M, Leparo GG, Tuchler T, Kreil DP, Mattanovich D. Novel insights into the unfolded protein response using *Pichia pastoris* specific DNA microarrays. *BMC Genomics*. 2008; 9:390. [PubMed: 18713468]
40. Dragosits M, Stadlmann J, Albiol J, Baumann K, Maurer M, Gasser B, Sauer M, Altmann F, Ferrer P, Mattanovich D. The effect of temperature on the proteome of recombinant *Pichia pastoris*. *J Proteome Res*. 2009; 8(3):1380–92. [PubMed: 19216534]
41. Dragosits M, Stadlmann J, Graf A, Gasser B, Maurer M, Sauer M, Kreil DP, Altmann F, Mattanovich D. The response to unfolded protein is involved in osmotolerance of *Pichia pastoris*. *BMC Genomics*. 2010; 11:207. [PubMed: 20346137]
42. Bridges HR, Fearnley IM, Hirst J. The subunit composition of mitochondrial NADH:ubiquinone oxidoreductase (complex I) from *Pichia pastoris*. *Mol Cell Proteomics*. 9(10):2318–26. [PubMed: 20610779]
43. Ozimek P, van Dijk R, Latchev K, Gancedo C, Wang DY, van der Klei IJ, Veenhuis M. Pyruvate carboxylase is an essential protein in the assembly of yeast peroxisomal oligomeric alcohol oxidase. *Mol Biol Cell*. 2003; 14(2):786–97. [PubMed: 12589070]
44. Palmieri F, Agrimi G, Blanco E, Castegna A, Di Noia MA, Iacobazzi V, Lasorsa FM, Marobbio CM, Palmieri L, Scarcia P, Todisco S, Voza A, Walker J. Identification of mitochondria l carriers in *Saccharomyces cerevisiae* by transport assay of reconstituted recombinant proteins. *Biochim Biophys Acta*. 2006; 1757(9-10):1249–62. [PubMed: 16844075]
45. Caverio S, Voza A, del Arco A, Palmieri L, Villa A, Blanco E, Runswick MJ, Walker JE, Cerdan S, Palmieri F. Satrustegui, J., Identification and metabolic role of the mitochondrial aspartate-glutamate transporter in *Saccharomyces cerevisiae*. *Mol Microbiol*. 2003; 50(4):1257–69. [PubMed: 14622413]
46. Yano T, Takigami E, Yurimoto H, Sakai Y. Yap1-regulated glutathione redox system curtails accumulation of formaldehyde and reactive oxygen species in methanol metabolism of *Pichia pastoris*. *Eukaryot Cell*. 2009; 8(4):540–9. [PubMed: 19252120]
47. Huttenhain R, Malmstrom J, Picotti P, Aebersold R. Perspectives of targeted mass spectrometry for protein biomarker verification. *Curr Opin Chem Biol*. 2009; 13(5-6):518–25. [PubMed: 19818677]
48. Singh I. Biochemistry of peroxisomes in health and disease. *Mol Cell Biochem*. 1997; 167(1-2):1–29. [PubMed: 9059978]

Figure 1.

Generation of *P. pastoris arg- lys-* auxotroph. (A) ARG4 and LYS2 knockout vectors. The NTC (nourseothricin acetyltransferase) and HPT (hygromycin phosphotransferase) disruption vectors were used to knockout arginosuccinate lyase (ARG4) and α -aminoadipate reductase (LYS2) by homologous recombination. (B) PCR-based confirmation of genetic disruption: At left, rk88/rk6 primers amplified a 2488 amplicon generated in the arg-knockout vs. a 1222 amplicon in the 'wild type' *BG08*. At right, rk166/rk167 primers amplify a 1952 amplicon generated in lys- knockout vs. no amplicon in the wild type.

Figure 2.

Heavy HSA expression and MS analysis. (A) Recombinant HSA is secreted into the culture medium at a yield of 50 $\mu\text{g/ml}$ detectable by Coomassie stained SDS-PAGE. A minor 43kD degradation product of the full length HSA (67kD) is also evident, consistent with previous reports of protein cleavage in *P. pastoris*³². The absence of other secreted proteins in the media is apparent in both methanol induced and uninduced samples. (B) MS-analysis of trypsinized HSA covers 55% of the protein sequence. Of 31 non-redundant peptides, 22 peptides show quantifiable L/H ratios by MS. (C) Sample chromatogram overlay for HSA peptide LVTDLDK with the heavy peptide AUC (gray) vs. light peptide AUC (slate).

Figure 3.

P. pastoris protein profiles at 48 hours growth on methanol supplemented with heavy arginine and lysine. (A) Normalized plot of heavy:light ratios for 197 quantified proteins. Proteins up-regulated 8-fold or greater are highlighted (slate bars) and tabulated at right. The majority of up-regulated proteins are involved in methanol metabolism. (B) Schematic of methanol metabolism in *P. pastoris*. 1) alcohol oxidase, 2) catalase, 3) S-hydroxymethyl dehydrogenase, 4) S-formyl glutathione hydrolase, 5) formate dehydrogenase, 6) dihydroxyacetone synthase, 7) dihydroxyacetone kinase, 8) fructose 1,6-bisphosphate aldolase, 9) fructose 1,6-bisphosphatase, 10) glutathione reductase, 11) glutathione peroxidase.

Figure 4.

Time course of individual protein ratios on methanol media. Precipitous up-regulation of methanol metabolic enzymes corresponds with the gradual down-regulation of glycolytic enzymes involved in glycerol metabolism. AOX1 (alcohol oxidase 1), CTA (catalase), DAS1 (dihydroxyacetone synthase 1); DAK (dihydroxyacetone kinase), FBA (fructose 1,6-bisphosphate aldolase); FLD (S-hydroxymethyl dehydrogenase), FGH (S-formyl glutathione hydrolase), FDH (formate dehydrogenase); GAPDH (glyceraldehydes-3-phosphate dehydrogenase), PGK (phosphoglycerate kinase), PGM (phosphoglycerate mutase), ENO (enolase), PK (pyruvate kinase).

Table 1
HSA: quantifiable unique peptides

Peptide (light)	M/Z ^a	L/H ^b
AACLLPK	386.7	0.008
AAFTECCQAADK	686.3	0.007
ADDKETCFAEEGK	500.5	0.010
AVMDDFAAFVEK	679.8	0.006
CCTESLVNR	569.7	0.008
CCAAADPHECYAK	768.3	0.014
ETYGEMADCCAK	717.8	0.032
FQNALLVR	480.7	0.000
KLVAASQAALGL	571.4	0.004
KQTALVELVK	564.9	0.000
KVPQVSTPTLVEVSR	820.4	0.000
KYLYEIAR	528.3	0.005
LVNEVTEFAK	575.3	0.010
LVTDLTK	395.2	0.009
QNCLEFEQLGEYK	829.4	0.009
QEPERNECFLQHKDDNPNLPR	873.7	0.007
RPCFSALEVDETYVPK	637.6	0.001
TCVADESAENCDK	749.8	0.008
VFDEFKPLVEEPQNLIK	682.4	0.003
VPQVSTPTLVEVSR	756.4	0.029
YICENQDSISSK	722.3	0.008
YLYEIAR	464.2	0.000

^aM/Z: mass to charge ratio.

^bL/H: light to heavy ratio