Investigation of Malic Acid Production in *Aspergillus oryzae* under Nitrogen Starvation Conditions

Christoph Knuf,^a Intawat Nookaew,^a Stephen H. Brown,^b Michael McCulloch,^b Alan Berry,^b Jens Nielsen^a

Department of Chemical and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden^a; Novozymes, Inc., Davis, California, USA^b

Malic acid has great potential for replacing petrochemical building blocks in the future. For this application, high yields, rates, and titers are essential in order to sustain a viable biotechnological production process. Natural high-capacity malic acid producers like the malic acid producer *Aspergillus flavus* have so far been disqualified because of special growth requirements or the production of mycotoxins. As *A. oryzae* is a very close relative or even an ecotype of *A. flavus*, it is likely that its high malic acid production capabilities with a generally regarded as safe (GRAS) status may be combined with already existing large-scale fermentation experience. In order to verify the malic acid production potential, two wild-type strains, NRRL3485 and NRRL3488, were compared in shake flasks. As NRRL3488 showed a volumetric production rate twice as high as that of NRRL3485, this strain was selected for further investigation of the influence of two different nitrogen sources on malic acid secretion. The cultivation in lab-scale fermentors resulted in a higher final titer, 30.27 ± 1.05 g liter⁻¹, using peptone than the one of 22.27 ± 0.46 g liter⁻¹ obtained when ammonium was used. Through transcriptome analysis, a binding site similar to the one of the *Saccharomyces cerevisiae* yeast transcription factor Msn2/4 was identified in the upstream regions of glycolytic genes and the cytosolic malic acid production pathway from pyruvate via oxaloacetate to malate, which suggests that malic acid production is a stress response. Furthermore, the pyruvate carboxylase reaction was identified as a target for metabolic engineering, after it was confirmed to be transcriptionally regulated through the correlation of intracellular fluxes and transcriptional changes.

Malic acid belongs to the group of C₄ dicarboxylic acids, which are structurally similar to maleic acid and maleic anhydride, which represent key building blocks in the chemical industry. The C₄ dicarboxylic acids may therefore replace petrochemically derived compounds in the future, when increased oil and gas prices favor the production of renewable chemicals from biomass. The C₄ acids of interest, malic, succinic, and fumaric acids, are intermediates of the tricarboxylic acid (TCA) cycle and are naturally produced by many organisms. The first patent on malic acid production was filed in 1960 (1). The inventors selected an *Aspergillus flavus* strain to be the best natural producer and optimized the fermentation process for this organism, resulting in final titers of 58.4 g liter⁻¹ after 9 days of fermentation from minimal medium identical to MAF3 medium (see below) containing 0.2% ammonium sulfate and 100 g liter⁻¹ glucose. This represents a yield of 0.78 mol mol⁻¹ on glucose and a productivity of 0.27 g liter⁻¹ h⁻¹. Furthermore, they investigated the effect of the nitrogen source, including, among others, peptone and ammonium sulfate, and reported final titers of 32.6 g liter⁻¹ and 30.4 g liter⁻¹, respectively, after 7 days of fermentation from 100 g liter⁻¹ glucose. The same strain was used in the late 1980s and early 1990s for further investigation of the metabolism toward malic acid production. It was reported that *de novo* enzyme synthesis during nitrogen starvation resulted in an increase of malate synthesis, as malate dehydrogenase activity increased and fumarase activity changed only slightly (2). In shake flasks and fermentors, the molar yield on glucose was 0.68 mol mol⁻¹ after 8 days and 0.57 mol mol⁻¹ after 6 days (2). Further ¹³C nuclear magnetic resonance analysis of the produced malic acid showed that the majority of the acid was produced via the reductive TCA cycle branch, from pyruvate via oxaloacetate to malate (3). For *Aspergillus flavus*, the localization of pyruvate carboxylase was shown to be cytosolic (4), whereas isoenzymes could be found in the cytosol and mitochondrion in the case of the closely related organism *A. oryzae*. Isoenzymes of malate dehydro-

genase were shown to be active in both the cytosol and mitochondrion compartments for *Aspergillus flavus* (3). After optimization of the fermentation process, yields of 1.26 mol mol⁻¹ on glucose and a productivity of 0.59 g liter⁻¹ h⁻¹ were achieved in fermentors after 190 h of fermentation (5).

Though high yields and titers could be achieved using *A. flavus*, aflatoxin production by this organism disqualifies it for industrial malic acid production. Consequently, several attempts have been made to produce malic acid using highly engineered model organisms, like *Saccharomyces cerevisiae* and *Escherichia coli*. An *S. cerevisiae* strain overexpressing pyruvate carboxylase, malate dehydrogenase, and a malate exporter and carrying a pyruvate decarboxylase deletion reached malate yields of 0.42 mol mol⁻¹ on glucose at a productivity of 0.19 g liter⁻¹ h⁻¹ (6). Engineered *E. coli* strains could also reach very high molar yields and high productivities, e.g., 1.42 mol mol⁻¹ and 0.47 g liter⁻¹ h⁻¹ (7), in an anaerobic two-stage fermentation or 0.74 mol mol⁻¹ and 0.74 g liter⁻¹ h⁻¹ (8). The first strain was developed from a strain already engineered for succinic acid production and carried 11 gene deletions in total. In the latter strain, the ATP generation during the malic acid production process was changed by overexpressing the *Mannheimia succiniciproducens* phosphoenolpyruvate carboxykinase. The yields and productivities obtained with these engineered strains are similar to those obtained with the *A. flavus* wild-type

Received 4 May 2013 Accepted 21 July 2013

Published ahead of print 26 July 2013

Address correspondence to Jens Nielsen, nielsenj@chalmers.se.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/AEM.01445-13>.

Copyright © 2013, American Society for Microbiology. All Rights Reserved.
doi:10.1128/AEM.01445-13

strain. Comparative genomics of *A. flavus* and *A. oryzae* suggest that these are very close relatives or even ecotypes of the same species (9), which suggests that they have similar malic acid production capabilities. This leads to the question of whether *A. oryzae* strains are suitable for malic acid production, as well as what impact the nitrogen source has on malic acid production capacity.

In this study, we present *A. oryzae* as a cell factory for the production of malic acid which combines high malic acid production capabilities and production security using a class 1 organism, which would be preferred for industrial production. With the introduction of systems biology tools (10, 11) and the availability of the whole-genome sequence (12), high-throughput analysis has become possible. By using the *A. oryzae* genome-scale metabolic model (GEM) in combination with microarrays for transcriptome analysis, we investigated the malic acid production mechanisms and predicted metabolic engineering targets to further increase malic acid production yields and productivities to industrial targets.

MATERIALS AND METHODS

Strains. Wild-type *Aspergillus oryzae* strains NRRL3485 and NRRL3488 were initially compared with each other. Strains NRRL3485 (DSM1862) and NRRL3488 were obtained from the German Collection of Microorganisms and Cell Cultures (Deutsche Sammlung von Mikroorganismen und Zellkulturen [DSMZ]) and the Agricultural Research Service (ARS) culture collection (Northern Regional Research Laboratory [NRRL]), respectively. The physiological characterization and transcriptome analysis were performed using NRRL3488. Spores of both strains were stored at -80°C in 20% glycerol solution.

Media. The compositions of the spore propagation medium (Cove-N-Gly) and the preculture medium (G2-GLY) were described before (13). In this study, the preculture medium also contained 100 g/liter CaCO_3 . *A. oryzae* malic acid fermentation (MAF3) medium consisted of 50 g liter $^{-1}$ glucose monohydrate, 0.15 g liter $^{-1}$ KH_2PO_4 , 0.15 g liter $^{-1}$ K_2HPO_4 , 0.10 g liter $^{-1}$ $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.10 g liter $^{-1}$ $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0.005 g liter $^{-1}$ $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0.005 g liter $^{-1}$ NaCl, 100.0 g liter $^{-1}$ CaCO_3 , and 1 ml liter $^{-1}$ pluronic acid (PE6100; BASF). The nitrogen sources used were either 1.4 g liter $^{-1}$ $(\text{NH}_4)_2\text{SO}_4$ or 6 g liter $^{-1}$ Bacto peptone.

Preparation of inoculum. Cove-N-Gly plates were inoculated with a suspension of spores from the -80°C stock of each strain. The plates were incubated at 30°C for 7 days. The spores were subsequently harvested by addition of 10 ml of 0.01% Tween 80 solution. The spore suspension was used to inoculate the preculture medium (G2-GLY) with a final concentration of 6×10^9 spores liter $^{-1}$. The preculture was incubated with shaking at 250 rpm in a 500-ml shake flask without baffles for 24 h at 30°C . Thereafter, the main cultures for malic acid production in fermentors or shake flasks were inoculated with preculture broth using 10% of the final working volume.

Batch cultivations. For evaluation of the malic acid production capabilities, batch cultivations were carried out in shake flasks and fermentors. In the case of shake flask cultivations, the strains were incubated in 100 ml MAF3 medium supplied with peptone as the nitrogen source in 500-ml Erlenmeyer flasks with agitation at 250 rpm in an orbital shaker. The cultivations in bioreactors were performed in 2.7-liter Applikon bioreactors (Applikon, Schiedam, The Netherlands) with 2-liter working volumes. Reactors were equipped with two Rushton six-blade disc impellers, and the temperature was maintained at 34°C . The pH was buffered by the calcium carbonate in the medium. The temperature, agitation, gassing, pH, and composition of the off gas were monitored and/or controlled using a DasGip monitoring and control system (DasGip, Jülich, Germany). The stirrer speed was set to 950 rpm, and the aeration rate was 1 volume of gas per volume of fermentation broth per minute (vvm). The concentrations of oxygen and carbon dioxide in the exhaust gas

were analyzed with DasGip fedbatch-pro gas analysis systems with the off-gas analyzer GA4, based on zirconium dioxide and a two-beam infrared sensor.

Sampling. One fraction of the sample withdrawn from the fermentors was stored at -20°C for subsequent analysis of extracellular metabolites. For quantification of cell mass, a known sample volume was treated with 2 N HCl in order to dissolve the calcium carbonate. The treated broth was centrifuged, and the pellet was washed once. After that, the wet biomass pellet was redissolved in distilled water and poured onto a preweighed aluminum dish. Aluminum dishes were kept at 90°C for 24 h in order to evaporate the water and subsequently stored in a desiccator until weighing.

Metabolite analysis. The concentrations of sugars and metabolites in the culture medium were determined by high-pressure liquid chromatography (HPLC). Organic acids were measured using a Synergi 4- μm Hydro-RP 80-Å HPLC column (Phenomenex Ltd., Aschaffenburg, Germany), together with a Dionex Ultimate 3000 system and a photodiode array detector (Dionex, Sunnyvale, CA) at a wavelength of 210 nm. The eluent consisted of 145 mM phosphoric acid with 10% methanol at pH 3. Elution was performed isocratically at a flow rate of 0.7 ml/min and a temperature of 20°C .

Glucose was determined using an Aminex HPX-87H column from Bio-Rad (Bio-Rad, Sundbyberg, Sweden). The assay was performed at 45°C at a flow rate of 0.6 ml/min with a 5 mM sulfuric acid solution for isocratic elution.

Enzyme assay. Enzyme assays were performed with samples from shake flasks. A 50-ml sample of medium was poured into a strainer. The mycelia were washed with distilled water until most CaCO_3 was removed. The washed mycelia were subsequently collected in a Falcon tube and stored at -20°C until further use. For cell disruption, the frozen pellet was placed in a prechilled mortar and ground into an even, smooth paste. The paste was mixed into a smooth suspension after addition of 400 μl of phosphate-buffered saline buffer (pH 7.4). The crude enzyme extract was separated from the debris by centrifugation at $18,000 \times g$ and cleaned up using a Micro Bio spin column and SSC ($1 \times \text{SSC}$ is 0.15 M NaCl plus 0.015 M sodium citrate) buffer (Bio-Rad). The clean flowthrough was immediately used for protein quantification and enzyme assay. The assay mixture for malate dehydrogenase contained 100 mM Tris, pH 8.0, 10 mM NaHCO_3 , 5 mM MgCl_2 , 0.667 mM oxaloacetic acid, 0.2 mM NADH. The reaction was started with the addition of the cell extract. The malate dehydrogenase activity was measured spectrophotometrically by monitoring NADH oxidation at 340 nm. The pyruvate carboxylase activity was measured indirectly by coupling the assay to the malate dehydrogenase reaction. The assay mixture for pyruvate carboxylase contained 100 mM Tris, pH 8.0, 10 mM NaHCO_3 , 5 mM MgCl_2 , 1 mM pyruvate, 0.2 mM NADH, concentrated malate dehydrogenase, and 1 mM ATP. The assays were performed at 30°C with freshly prepared extracts. One unit of enzyme converts 1 μmol of substrate and coenzyme per minute, which, in this case, is 1 μmol of NADH to NAD^+ per minute. The total protein concentration was measured with a NanoDrop 2000 apparatus (Thermo Scientific).

RNA extraction. Samples for RNA extraction were briefly treated with HCl in order to solubilize the CaCO_3 . The samples were subsequently filtered through Miracloth (Calbiochem, San Diego, CA), and the filter cake was shock frozen in liquid nitrogen and stored at -80°C . For RNA extraction, the frozen mycelium was ground to a powder in a prechilled mortar with a prechilled pestle. RNA was extracted from the mycelium powder using an RNeasy plant minikit (Qiagen) according to the protocol described by the manufacturer. The quality of the extracted RNA was checked with a 2100 Bioanalyzer apparatus (Agilent Technologies, Inc., Santa Clara, CA).

Microarray processing. Details concerning the microarray manufacturing and design, biotin-labeled cDNA and microarray processing, and microarray data acquisition can be found in a previously published report (13).

Microarray analysis. The microarrays were analyzed using the R program, version 2.15.1 (14), and the R package Piano (15) for microarray analysis. This package combines several other packages, as described below. The Affy package (16) was used to load the CEL files. Those were preprocessed using the Plier package (17) and applying cubic spline normalization. For further statistical pairwise analysis to determine significantly differentially expressed genes, the Limma package (18) was utilized for moderated Student's *t* test. The standard errors within each gene were moderated using empirical Bayesian statistics, and multiple testing was adjusted for by applying the Benjamini-Hochberg method (19). The *P* values and fold changes for each gene so obtained were further analyzed using the reporter algorithm (20) for gene ontology (GO) term analysis and reporter metabolites. For each gene set, distinct directional *P* values were calculated. The distinct directional *P* value is the outcome of two one-tailed *t* tests for either up- or downregulation, taking the transcriptional changes of all genes in the gene set into account. The case for the most significant distinct directional *P* value was chosen. Hence, the distinct directional *P* value, other than the directional *P* value, shows the overall significance of up- or downregulation of genes in a given gene set and not only the significance of the up- or downregulated genes in the gene set. For the reporter metabolite analysis, the topology was inferred from the *A. oryzae* GEM. The topology for the GO term analysis was downloaded from the AspGD *Aspergillus* genome database (21).

In order to identify transcriptionally regulated reactions, a correlation between expression and flux changes was searched for. For this, the method of random sampling (22) was applied to generate 3,292 possible flux distributions for each of the conditions and constraints considered according to the measured external fluxes. Loops were avoided by setting the default bounds for the metabolic fluxes to infinity and negative infinity instead of 1,000 and $-1,000$, as is usually done. In order to obtain significance scores for the flux change in each reaction, the first sample of condition A (exponential phase) was compared with all the samples of condition B (stationary phase). The number of times that condition A had a higher or lower value for each reaction was subsequently computed. Then, the same procedure was repeated for the second sample of condition A, and this procedure was continued until the last sample of condition A was analyzed. The probability score for a flux increase or decrease is reflected by the fraction of $3,292 \times 3,292$ comparisons in which condition A had a higher or lower value than condition B. This probability score was subsequently compared with the probability scores of transcriptional changes. A correlation of the significance scores of both flux and transcript hints toward a transcriptional regulation of this reaction.

RESULTS

Physiological investigation. An initial screening of wild-type *A. oryzae* strains showed broad malic acid production capabilities. To illustrate this, the two wild-type *A. oryzae* strains NRRL3485 and NRRL3488, which were located at the lower and upper ends of the malic acid production spectrum, respectively, were cultivated over 78 h in shake flasks containing MAF3 medium with 6 g liter⁻¹ peptone as the nitrogen source for initial biomass production. Strain NRRL3485 secreted malic acid at a volumetric rate of 0.299 ± 0.011 g liter⁻¹ h⁻¹, whereas NRRL3488 produced malic acid at almost twice that rate of 0.563 ± 0.020 g liter⁻¹ h⁻¹. The final titers were 23.12 ± 0 g liter⁻¹ and 38.86 ± 2.80 g liter⁻¹ for NRRL3485 and NRRL3488, respectively.

The faster-producing strain, NRRL3488, was selected for further evaluation of malic acid production during nitrogen starvation in well-controlled lab-scale bioreactors. The medium contained either the complex nitrogen source peptone or the nitrogen-containing salt ammonium sulfate, which is more suitable for use in industrial large-scale processes, as it is cheaper and simplifies downstream processing. The batch cultivations were performed in quadruplicate experiments for

each nitrogen source. The results are presented in Fig. 1, which illustrates the profiles of the averaged biomass, the concentrations of the extracellular metabolites of malate, succinate, and citrate, and glucose concentrations for the peptone (Fig. 1A) and ammonium (Fig. 1B) conditions. The maximum specific growth rates, the rates of malic acid volumetric production, the rate of glucose consumption, and the molar yields of the organic acids in the exponential and stationary phases for each nitrogen source are summarized in Table 1.

Comparison of the growth curves of NRRL3488 revealed similar behaviors on peptone and ammonium. The strain showed exponential growth for the first 12 to 15 h of fermentation with a slightly elevated maximum specific growth rate and prolonged exponential phase on peptone, resulting in a biomass concentration of about 5 g liter⁻¹ on peptone in comparison to one of 4 g liter⁻¹ on ammonium at the end of the exponential phase. At this point, nitrogen limitation was confirmed by measurement of ammonium concentrations in the medium in case of ammonium fermentation. Thereafter, an adaptation phase, during which the biomass increased under both conditions, but to a higher extent on peptone, bridged over to a stationary phase, in which the biomass stayed constant on peptone and ammonium at about 8.5 g liter⁻¹ and 6 g liter⁻¹, respectively.

During the exponential growth phase, only malate secretion could be detected in significant amounts under both conditions. On peptone, small amounts of succinate and citrate were detected, and on ammonium, marginal succinate and citrate secretion was detected. During the stationary phase, malate, citrate, succinate, and pyruvate were detected in the fermentation broth under both conditions. The yields of malate on a glucose base were tripled during the stationary phase, resulting in final titers of malate of 30.27 ± 1.05 g liter⁻¹ in the peptone case and 22.27 ± 0.46 g liter⁻¹ for the ammonium cultivations. This substantial secretion of organic acids, mainly malate, raised the question about the underlying regulation and mechanisms of the switch between exponential growth, with biomass production as the objective function, and the stationary phase, in which half of the metabolized glucose is directly converted to malic acid and secreted into the fermentation broth. In order to further investigate this issue, comparative transcriptome analysis was performed.

Comparative transcriptome analysis. In order to understand the key processes in the shift toward malic acid production and the response of the cell to nitrogen deprivation, results of transcription analysis of samples taken in mid-exponential phase (6 h) and stationary phase (30 h) from triplicate cultivations were compared pairwise for each nitrogen source using Student's *t*-test statistics. For the statistical analysis of the transcriptional changes, Student's *t* test with an adjusted-*P*-value cutoff of 0.005 was applied. In total, the expression of 3,766 genes changed significantly under the peptone condition; of these, 946 genes had lower levels of expression and 2,820 had higher levels of expression in response to nitrogen starvation (30 h versus 6 h). Among the 4,540 differentially expressed genes under the ammonium condition, 28% (1,306 genes) showed a lower level of expression in response to nitrogen starvation. For further investigation, reporter metabolite analysis and gene ontology (GO) term analysis were performed using the reporter function of the R package Piano.

Reporter feature analysis of the transcriptome data. For the reporter metabolite analysis, the gene metabolite network was inferred from an updated version of the published *A. oryzae* model

FIG 1 Extracellular metabolite concentrations during bioreactor cultivations of NRRL3488 in MAF medium containing 6 g liter⁻¹ peptone (A) or 1.4 g liter⁻¹ ammonium sulfate (B) in 2.7-liter Applikon bioreactors with a working volume of 2 liters controlled by a DasGip control system. The results shown are averages and standard errors of 4 reactors. Mal, malate; Cit, citrate; Suc, succinate; Gluc, glucose; DW, dry weight.

(10), which is available in the BioMet tool box (23). The reporter metabolites with a distinct directional *P* value of less than 0.001 in either of the two comparisons are shown in Fig. 2B. A full list of the results is provided in Data Set S2 in the supplemental material. Among those 59 metabolites around which the most significant transcriptional changes occurred, there were only 8 metabolites with a significant upregulation of the correlated genes. It was not surprising to find both intra- and extracellular ammonia among the upregulated reporter metabolites in nitrogen-starved cells. Allantoate and urate, which are intermediates or end products of purine metabolism, were found, as were two metabolites of glutathione metabolism and the end product of glutathione metabolism, glutathione. The last identified metabolite was UDP, which is involved in the polymerization of glycogen. Among the residual group of 51 metabolites associated with repressed expression, 8 metabolites that are involved in amino acid synthesis and 8 that

are involved in the TCA cycle were found, and 10 were energy or reduction equivalents.

Another method for correlating gene expression and biological processes is GO term analysis. For this analysis, only the ontology of biological processes was used as input for the reporter features algorithm (24). A heat map displaying the most significantly changed GO terms that were characterized by a distinct directional *P* value of less than 0.001 under at least one condition is represented in Fig. 2A. A full list of the results is provided in Data Set S3 in the supplemental material. Among the 15 induced processes, significant changes in piecemeal microautophagy of the nucleus, the purine base catabolic process, protein ubiquitination, or conidium formation were found. The list of GO terms associated with transcriptional repression reflects the same general trend of reduced cellular viability shown by the induction-associated GO terms. In association with the scarcity of available nitro-

TABLE 1 Physiological data FOR NRRL3488 grown in malic acid fermentation medium with either peptone or ammonium as the nitrogen source^a

Nitrogen source	Final titer of malate (g liter ⁻¹)	Phase ^b	$\mu_{\max}^c$ (h ⁻¹)	r_{malate}^d (mmol liter ⁻¹ h ⁻¹)	r_s^e (mmol liter ⁻¹ h ⁻¹)	Yield on glucose (mol mol ⁻¹)			
						Citrate	Malate	Succinate	Pyruvate
Peptone	30.27 ± 1.05	Exp.	0.23 ± 0.01	4.22 ± 0.25	8.10 ± 0.81	ND ^f	0.33 ± 0.05	ND	ND
		Stat.		6.61 ± 0.57	6.13 ± 0.34	0.03 ± 0.01	0.98 ± 0.13	0.14 ± 0.03	0.02 ± 0.01
Ammonium	22.27 ± 0.46	Exp.	0.21 ± 0.05	1.59 ± 0.15	3.71 ± 0.44	0.01 ± 0.01	0.34 ± 0.06	0.07 ± 0.02	ND
		Stat.		4.36 ± 0.14	3.92 ± 0.10	0.07 ± 0.01	1.09 ± 0.05	0.20 ± 0.01	0.03 ± 0.00

^a The numbers stated are the means of four individual bioreactors ± standard errors.

^b Exp., exponential growth phase; Stat., stationary phase.

^c $\mu_{\max}$, maximum specific growth rate.

^d r_{malate} , specific malate production rate.

^e r_s , substrate consumption rate.

^f ND, not determined.

gen, several biosynthetic processes for amino acid synthesis seemed to be repressed. Consistent with the low availability of amino acids, GO terms associated with protein synthesis, starting from translation/translational elongation and continuing to pro-

tein folding and intracellular protein transport, were correlated with transcriptional repression. As the availability of nitrogen limited cell growth, the largest sink of energy in the cell was absent. The cells seemed to respond to this with transcriptional repression

FIG 2 Heat map of overrepresented biological process GO terms (A) and reporter metabolites (B) in the comparison of the nitrogen starvation phase to exponential growth phase. GO terms and reporter metabolites with *P* values of at most 0.005 under either of the conditions with peptone or ammonium are shown. BP, biological process; SSU, small subunit; CoA, coenzyme A; SRP, signal recognition particle; Golgi, Golgi apparatus; ER, endoplasmic reticulum; acp, acyl carrier protein.

FIG 3 Schematic drawing of the reactions related to malic acid secretion, including the significance and direction of transcriptional changes of the genes encoding the enzymes catalyzing those steps. The reaction arrows show the direction of the reaction stated in the GEM of *A. oryzae*. Double-headed arrows indicate reactions assumed to be reversible. The numbers next to the reaction arrows correlate with the reaction numbers in the model. The arrowed fields indicate the significance of transcriptional change in the comparison of the stationary phase to the exponential phase, with the ammonium condition indicated in the left set of boxes and the peptone condition indicated in the right set of boxes. The darker that the shading is, the more significant that the change is. The direction of transcriptional change is indicated by the direction of the arrows: up arrows, upregulation in stationary phase; down arrows, downregulation in stationary phase. Asterisks, data for the most significantly changed genes of an enzyme complex.

of energy-supplying processes, such as aerobic respiration and processes connected to that, e.g., mitochondrial electron transport or ATP synthesis-coupled proton transport.

Taken together, the reporter metabolite and GO term analyses point toward a transcriptional response of the cell to the nitrogen starvation stress that leads to a degradation of cellular components, degradation of nitrogen-containing compounds in order to recycle the nitrogen, and reduced nitrogen consumption in protein production. However, the findings so far do not directly point toward an explanation for the high level of malic acid secretion. To shed light on the malic acid production pathway, the *P* values of the reactions related to it were plotted onto a metabolic map, including the pathway from glucose to malic acid, taking the di-

rection of fold change into account (Fig. 3). From this map it becomes obvious that glycolysis seems to be upregulated at the transcriptional level. Except for the glucose-6-phosphate isomerase, fructose bisphosphate aldolase, and pyruvate kinase steps, for the other reactions from glucose to malic acid, at least one gene involved in those reactions was more highly expressed in response to nitrogen starvation under both conditions. On the other hand, the general trend for the expression of genes involved in the TCA cycle was for these to be downregulated, which is in agreement with the results of the GO term analysis regarding the reduced ATP generation during oxidative phosphorylation. In order to find possible transcription factors that regulate the overexpression of the genes involved in the malic acid production pathway, the

upstream sequences of the genes that showed a significant induction during the starvation phase were investigated for conserved sequences. The upstream sequences of the genes of interest were retrieved from the web service regulatory sequence analysis tools (RSAT) using the retrieve sequence function, and subsequently, the oligonucleotide analysis tool (25) of the motif discovery functions. Two conserved motifs were discovered, kwCCCCTCCy and kbbCACCGGTGvwm (where k is G or T, w is A or T, y is C or T, b is C, G, or T, and v is A, C, or G). The conserved 6-oligonucleotide sequence contained in the first motif (CCCCTC) has an occurrence P value of $6.6\text{E}-06$. By passing this pattern on to the pattern-matching tool of the web service interface YEASTRACT (26), we identified this to be the binding site for the *Saccharomyces cerevisiae* yeast transcription factor Msn2/4, which is the transcriptional activator of the multistress response in *S. cerevisiae* (27).

Identification of transcriptionally regulated fluxes. In the next step of the analysis of malic acid production under conditions of nitrogen starvation, we identified transcriptionally regulated reactions by correlating the changes in flux and the transcription of genes encoding the enzymatic steps. Those transcriptionally regulated reactions are possible targets for metabolic engineering, as a simple overexpression of the gene should directly positively influence the flux of the corresponding reaction. By using a random sampling approach, the average and standard deviation for each flux were calculated for the reactions in the *A. oryzae* GEM, using the measured exchange fluxes as constraints (the GEM in XLSX format, including constraints, can be found in Data Set S4 in the supplemental material) (22). For the calculation of exchange fluxes, the physiological data for the ammonium fermentation were used. During the starvation phase, the biomass exceeded the theoretical biomass. As the biomass concentration has a big influence on the specific rates and it is inherently difficult to determine the biomass concentration exactly from filamentous fungal fermentations, a theoretical value was calculated from the amount of nitrogen supplied using elemental biomass compositions from previous studies. Very accurate measurements for *A. oryzae* have been obtained by Pedersen et al. (28), but the samples were taken from chemostats with a maximum growth rate of 0.17 h^{-1} . In order to get a better estimate for unlimited growth, the elemental biomass composition for *A. niger* for unlimited growth was used, and a maximum biomass concentration of 4.19 g liter^{-1} was calculated (29). As this value correlates well with the biomass concentration at the end of the exponential growth phase, the specific rates during the stationary phase were calculated using this biomass concentration for better comparison of the rates during exponential and stationary phases. After obtaining the flux distribution for each of the two conditions, the significance of the change in each reaction rate could be calculated. By correlating the P values of the transcriptional and flux changes for each reaction, genes that have direct transcriptional control over the flux of the reaction could be identified (22). In total, expression and theoretical flux were significantly positively correlated for only two genes, encoding the cytosolic pyruvate carboxylase (AspGD accession number AO090023000801 [http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090023000801&organism=A_oryzae_RIB40]) and lactate dehydrogenase (AspGD accession number AO090023000577 [http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090023000577&organism=A_oryzae_RIB40]).

A list with negatively correlated reactions can be found in Data Set S5 in the supplemental material. As those transcriptionally

regulated genes are prone to be regulated directly by transcription factors, the same procedure described above was followed to find regulatory sequences. The pattern matching resulted in the conserved consensus sequence vwTCAATTGAwb. The possible recognition pattern CAATTG had an occurrence P value of $4.9\text{E}-10$. The search for the yeast transcription factor known to bind to this sequence did not return any hits.

In order to investigate the feasibility of the overexpression of pyruvate carboxylase, the enzyme activities of pyruvate carboxylase and malate dehydrogenase were measured in the exponential (6 h) and stationary (48 h) phases on MAF medium with ammonium and a glucose concentration of 100 g liter^{-1} in shake flasks. The activity of pyruvate carboxylase increased slightly from 0.024 ± 0.004 to 0.033 ± 0.007 units mg^{-1} total protein. In the malate dehydrogenase case, the activity decreased slightly from 4.848 ± 0.828 to 4.304 ± 0.358 units mg^{-1} total protein.

The changes in activities were expected to differ more significantly. The reasoning for this expectation was based on the fold changes in expression and because the enzymes were identified to be transcriptionally regulated. Therefore, the possibility of ubiquitination and a resulting faster protein degradation were considered. The ubiquitination sites of the genes involved in the malic acid production pathway, *pyc* (AspGD accession number AO090023000801 [http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090023000801&organism=A_oryzae_RIB40]), *mdhA* (AspGD accession number AO090701000013 [http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090701000013&organism=A_oryzae_RIB40]), and *mae3* (AspGD accession number AO090023000318 [http://www.aspergillusgenome.org/cgi-bin/locus.pl?locus=AO090023000318&organism=A_oryzae_RIB40])), were predicted using the UbPred program (30). For *mdhA* and *mae3*, ubiquitination sites were predicted with low confidence, whereas two sites were found in *pyc* with high confidence.

DISCUSSION

An initial test of two wild-type *A. oryzae* strains led to the selection of NRRL3488 for further investigation. NRRL3488 was capable of producing malic acid in shake flasks with peptone as the nitrogen source at a volumetric rate of $0.563 \pm 0.020\text{ g liter}^{-1}\text{ h}^{-1}$. This productivity was almost twice as high as the one of NRRL3485, indicating a large variation in malic acid production among *A. oryzae* wild-type strains. Therefore, further investigation of high malic acid production in lab-scale bioreactors using either peptone or ammonium as the nitrogen source was conducted with NRRL3488. The productivity during the stationary phase was 34% higher in the case of the peptone cultivations than in the case of the ammonium cultivations; nevertheless, the yields did not differ significantly. This might be due to the same metabolic efficiency but different biomasses built up from the amount of nitrogen available in the two cultivations.

NRRL3488 showed high final titers, productivities, and yields during cultivation for 47.5 h on ammonium salt medium. Though the transient yield of NRRL3488 (1.09 mol mol^{-1}) was slightly lower than that previously reported for *A. flavus* (1.26 mol mol^{-1}), the volumetric productivities were almost equal: $0.59\text{ g liter}^{-1}\text{ h}^{-1}$ in the case of *A. flavus* and $0.58\text{ g liter}^{-1}\text{ h}^{-1}$ in the case of NRRL3488 (5). The productivity and yield were more than twice as high in NRRL3488 than in a highly engineered *S. cerevisiae* strain (6) during the acid production phase. In comparison with engineered *E. coli* strains, strain *E. coli* XZ658 (7) showed a higher

molar yield, but the productivity of NRRL3488 was about 23.4% higher. On the other hand, *E. coli* WGS-10 (8) showed a lower yield but a 21.1% higher productivity than NRRL3488.

In order to investigate the regulatory mechanisms of this massive secretion of malic acid by a wild-type *A. oryzae* strain, transcription analysis using DNA microarrays was used. The general trend indicated by GO term and reporter metabolite analyses comparing the stationary phase with the exponential growth phase showed that the cells respond to nitrogen starvation by recycling nitrogen by degradation of proteins and other nitrogen-containing cellular compounds and reducing protein synthesis.

The metabolic function responsible for this might be an increased ubiquitination of proteins. Two strong ubiquitination sites were predicted for pyruvate carboxylase but not for other proteins directly involved in malic acid production. Nevertheless, this potential ubiquitination might explain the need for a more significant upregulation of *pyc* than malate dehydrogenase gene expression and still result in enzyme activity slightly higher than that during the exponential growth phase. As the *in vitro* specific activity of pyruvate carboxylase is 2 orders of magnitude lower than that of malate dehydrogenase, it seems to be the rate-limiting step in the malic acid production pathway.

As nitrogen depletion hampers cellular growth, the generation of ATP via oxidative phosphorylation and the production of reduction equivalents are not needed anymore. This could also be seen from the GO terms, which contained genes with significant repression, and reporter metabolites, around which a general repression of genes was detected.

The regulation of energy metabolism might be the reason for malic acid secretion. During the stationary phase, the cells are not able to respire glucose to carbon dioxide, as this would result in the production of large amounts of ATP that the cells cannot use due to a lack of growth. The advantage of malic acid production from glucose through the reductive cytosolic TCA branch from pyruvate via oxaloacetate is that it is completely balanced in terms of ATP and NADH production from glycolysis to pyruvate and consumption in the subsequent steps to malate. This pathway enables *A. oryzae* to produce large amounts of malic acid uncoupled from growth, which makes it a very interesting trait for future industrial applications. From an ecological and evolutionary point of view, direct glucose conversion to malic acid makes sense, as (i) glucose consumption can continue, reducing the availability for competing microorganisms, like many fast-growing bacteria; (ii) microbial growth is oppressed by low pH, whereas aspergilli are tolerant to acidic pHs (*A. oryzae*, for example, shows optimal growth over a broad pH range from 3 to 7 [31]); and (iii) sequence analysis of *A. oryzae* showed that it has the largest expansion of secretory hydrolytic enzymes that work at low pH, in comparison to the expansion for *A. nidulans* and *A. fumigatus* (12). This might help the cell to survive in an environment with only complex nitrogen sources available. In order to achieve acidification of the medium up to the lower end of the growth optimum, malic acid production is very efficient in its acidifying potential, with pK_a s at 3.46 and 5.10.

The switch between ATP generation and malic acid secretion might be regulated by a transcription factor binding to the same recognition pattern as the yeast transcription factor Msn2/4. This transcriptional activator is known to positively affect transcription of genes under stressed conditions, such as conditions of osmotic, temperature, and nitrogen stress in budding yeast (32).

Taken together, we present a natural malate-producing microorganism that does not have to be disqualified for commercial large-scale production because of either the production of toxic by-products or a dependence on complex medium. The yield and productivity of the organism are comparable to those of highly engineered *E. coli* strains and exceed those of *S. cerevisiae*, leaving space for additional improvement by metabolic engineering. One target for this might be the overexpression of the pyruvate carboxylase, which was identified to be transcriptionally regulated and seems to be a rate-limiting step in malic acid biosynthesis.

ACKNOWLEDGMENTS

We thank Sergio Bordel for assisting in identifying transcriptionally regulated reactions and Wanwipa Vongsangnak for valuable discussions.

We acknowledge funding for our research activities from the Chalmers Foundation, the Knut and Alice Wallenberg Foundation, the European Research Council (grant no. 247013), and Novozymes, Inc.

REFERENCES

1. Abe S, Akira F, Ken-Ichiro T. November 1962. Method of producing L-malic acid by fermentation. US patent 3,063,910.
2. Peleg Y, Stieglitz B, Goldberg I. 1988. Malic acid accumulation by *Aspergillus flavus*. Appl. Microbiol. Biotechnol. 28:69–75.
3. Peleg Y, Barak A, Scrutton M, Goldberg I. 1989. Malic acid accumulation by *Aspergillus flavus*. Appl. Microbiol. Biotechnol. 30:176–183.
4. Bercovitz A, Peleg Y, Battat E, Rokem JS, Goldberg I. 1990. Localization of pyruvate carboxylase in organic acid-producing *Aspergillus* strains. Appl. Environ. Microbiol. 56:1594–1597.
5. Battat E, Peleg Y, Bercovitz A, Rokem JS, Goldberg I. 1991. Optimization of L-malic acid production by *Aspergillus flavus* in a stirred fermentor. Biotechnol. Bioeng. 37:1108–1116.
6. Zelle RM, de Hulster E, van Winden WA, de Waard P, Dijkema C, Winkler AA, Geertman J-MA, van Dijken JP, Pronk JT, van Maris AJA. 2008. Malic acid production by *Saccharomyces cerevisiae*: engineering of pyruvate carboxylation, oxaloacetate reduction, and malate export. Appl. Environ. Microbiol. 74:2766–2777.
7. Zhang X, Wang X, Shanmugam KT, Ingram LO. 2011. L-Malate production by metabolically engineered *Escherichia coli*. Appl. Environ. Microbiol. 77:427–434.
8. Moon SY, Hong SH, Kim TY, Lee SY. 2008. Metabolic engineering of *Escherichia coli* for the production of malic acid. Biochem. Eng. J. 40:312–320.
9. Payne GA, Nierman WC, Wortman JR, Pritchard BL, Brown D, Dean RA, Bhatnagar D, Cleveland TE, Machida M, Yu J. 2006. Whole genome comparison of *Aspergillus flavus* and *A. oryzae*. Med. Mycol. 44:9–11.
10. Vongsangnak W, Olsen P, Hansen K, Krogsgaard S, Nielsen J. 2008. Improved annotation through genome-scale metabolic modeling of *Aspergillus oryzae*. BMC Genomics 9:245. doi:10.1186/1471-2164-9-245.
11. Andersen MR, Vongsangnak W, Panagiotou G, Salazar MP, Lehmann L, Nielsen J. 2008. A trispecies *Aspergillus* microarray: comparative transcriptomics of three *Aspergillus* species. Proc. Natl. Acad. Sci. U. S. A. 105:4387–4392.
12. Machida M, Asai K, Sano M, Tanaka T, Kumagai T, Terai G, Kusumoto K-I, Arima T, Akita O, Kashiwagi Y, Abe K, Gomi K, Horiuchi H, Kitamoto K, Kobayashi T, Takeuchi M, Denning DW, Galagan JE, Nierman WC, Yu J, Archer DB, Bennett JW, Bhatnagar D, Cleveland TE, Fedorova ND, Gotoh O, Horikawa H, Hosoyama A, Ichinomiya M, Igarashi R, Iwashita K, Juvvadi PR, Kato M, Kato Y, Kin T, Kokubun A, Maeda H, Maeyama N, Maruyama J-I, Nagasaki H, Nakajima T, Oda K, Okada K, Paulsen I, Sakamoto K, Sawano T, Takahashi M, Takase K, Terabayashi Y, Wortman JR, et al. 2005. Genome sequencing and analysis of *Aspergillus oryzae*. Nature 438:1157–1161.
13. Vongsangnak W, Hansen K, Nielsen J. 2011. Integrated analysis of the global transcriptional response to α -amylase over-production by *Aspergillus oryzae*. Biotechnol. Bioeng. 108:1130–1139.
14. The R Development Core Team. 2012. R: a language and environment for statistical computing. The R Foundation for Statistical Computing, Vienna, Austria.
15. Varemo L, Nielsen J, Nookaew I. 2013. Enriching the gene set analysis of

- genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic Acids Res.* 41: 4378–4391.
16. Gautier L, Cope L, Bolstad BM, Irizarry RA. 2004. Affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 20:307–315.
 17. Affymetric Inc, Crispin JM, PICR. plier: implements the Affymetrix PLIER algorithm. R package, version 1.28.0. Affymetrix, Santa Clara, CA.
 18. Smyth GK. 2005. Limma: linear models for microarray data, p 397–420. In Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W (ed), *Bioinformatics and computational biology solutions using {R} and bioconductor*. Springer, New York, NY.
 19. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57:289–300.
 20. Patil KR, Nielsen J. 2005. Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proc. Natl. Acad. Sci. U. S. A.* 102:2685–2689.
 21. Arnaud MB, Cerqueira GC, Inglis DO, Skrzypek MS, Binkley J, Chibucos MC, Crabtree J, Howarth C, Orvis J, Shah P, Wymore F, Binkley G, Miyasato SR, Simison M, Sherlock G, Wortman JR. 2012. The *Aspergillus* Genome Database (AspGD): recent developments in comprehensive multispecies curation, comparative genomics and community resources. *Nucleic Acids Res.* 40:D653–D659.
 22. Bordel S, Agren R, Nielsen J. 2010. Sampling the solution space in genome-scale metabolic networks reveals transcriptional regulation in key enzymes. *PLoS Comput. Biol.* 6:e1000859. doi:10.1371/journal.pcbi.1000859.
 23. Cvijovic M, Olivares-Hernández R, Agren R, Dahr N, Vongsangnak W, Nookaew I, Patil KR, Nielsen J. 2010. BioMet toolbox: genome-wide analysis of metabolism. *Nucleic Acids Res.* 38:W144–W149. doi:10.1093/nar/gkq404.
 24. Oliveira A, Patil K, Nielsen J. 2008. Architecture of transcriptional regulatory circuits is knitted over the topology of bio-molecular interaction networks. *BMC Syst. Biol.* 2:17. doi:10.1186/1752-0509-2-17.
 25. van Helden J, André B, Collado-Vides J. 1998. Extracting regulatory sites from the upstream region of yeast genes by computational analysis of oligonucleotide frequencies. *J. Mol. Biol.* 281:827–842.
 26. Abdulrehman D, Monteiro PT, Teixeira MC, Mira NP, Lourenço AB, dos Santos SC, Cabrito TR, Francisco AP, Madeira SC, Aires RS, Oliveira AL, Sá-Correia I, Freitas AT. 2011. YEASTRACT: providing a programmatic access to curated transcriptional regulatory associations in *Saccharomyces cerevisiae* through a web services interface. *Nucleic Acids Res.* 39:D136–D140. doi:10.1093/nar/gkq964.
 27. Schmitt AP, McEntee K. 1996. Msn2p, a zinc finger DNA-binding protein, is the transcriptional activator of the multistress response in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* 93:5777–5782.
 28. Pedersen H, Carlsen M, Nielsen J. 1999. Identification of enzymes and quantification of metabolic fluxes in the wild type and in a recombinant *Aspergillus oryzae* strain. *Appl. Environ. Microbiol.* 65:11–19.
 29. Stephanopoulos GN, Aristidou AA, Nielsen J. 1998. *Metabolic engineering*. Academic Press, New York, NY.
 30. Radivojac P, Vacic V, Haynes C, Cocklin RR, Mohan A, Heyen JW, Goebel MG, Iakoucheva LM. 2010. Identification, analysis, and prediction of protein ubiquitination sites. *Proteins* 78:365–380.
 31. Carlsen M, Spohr AB, Nielsen J, Villadsen J. 1996. Morphology and physiology of an α -amylase producing strain of *Aspergillus oryzae* during batch cultivations. *Biotechnol. Bioeng.* 49:266–276.
 32. Santhanam A, Hartley A, Düvel K, Broach JR, Garrett S. 2004. PP2A phosphatase activity is required for stress and Tor kinase regulation of yeast stress response factor Msn2p. *Eukaryot. Cell* 3:1261–1271.