Elucidation of the Metabolic Fate of Glucose in the Filamentous Fungus Trichoderma reesei Using Expressed Sequence Tag (EST) Analysis and cDNA Microarrays*

Received for publication, August 9, 2001, and in revised form, November 15, 2001
Published, JBC Papers in Press, February 1, 2002, DOI 10.1074/jbc.M107651200

Felipe S. Chambergo‡, Eric D. Bonaccorsi‡, Ari J. S. Ferreira‡, Augusto S. P. Ramos‡, José Ribamar Ferreira Júnior‡, José Abrahão-Neto‡, João P. Simon Farah‡, and Hamza El-Dorry‡

From the Departments of ‡Biochemistry and ¶Chemistry, Institute of Chemistry, University of São Paulo, Avenida Prof. Lineu Prestes 748, São Paulo SP 05508–900, Brazil

Despite the intense interest in the metabolic regulation and evolution of the ATP-producing pathways, the long standing question of why most multicellular microorganisms metabolize glucose by respiration rather than fermentation remains unanswered. One such microorganism is the cellulolytic fungus Trichoderma reesei (Hypocreaceae). Using EST analysis and cDNA microarrays, we find that in T. reesei expression of the genes encoding the enzymes of the tricarboxylic acid cycle and the proteins of the electron transport chain is programmed in a way that favors the oxidation of pyruvate via the tricarboxylic acid cycle rather than its reduction to ethanol by fermentation. Moreover, the results indicate that acetaldehyde may be channeled into acetate rather than ethanol, thus preventing the regeneration of NAD⁺, a pivotal product required for anaerobic metabolism. The studies also point out that the regulatory machinery controlled by glucose was most probably the target of evolutionary pressure that directed the flow of metabolites into respiratory metabolism rather than fermentation. This finding has significant implications for the development of metabolically engineered cellulolytic microorganisms for fuel production from cellulose biomass.

Evolution has produced a diverse array of metabolic pathways and regulatory mechanisms that reflect the adaptation of an immense variety of microorganisms to different environments and nutritional requirements. A prominent example is the metabolism of glucose, the primary and preferred fuel for eukaryotic microorganisms. Although glucose is metabolized by a highly conserved series of connected enzymatic reactions, the mechanisms that regulate its fate and the properties of the ATP-producing pathways have been subjected to selection pressure during evolution. Aerobic (respiration) and anaerobic (fermentation) pathways are used by microorganisms to obtain energy from glucose, in the form of ATP. These pathways allow organisms to produce ATP at different rates and with different efficiencies; respiration proceeds at a lower rate and with a high yield, whereas fermentation operates at higher rates but with lower yield. Selection pressure imposed by energy limitation and the high ATP yield of respiration has been implicated in facilitating the evolutionary transition from unicellular to undifferentiated multicellular organisms (1).

Unicellular microorganisms, such as the yeast Saccharomyces cerevisiae, use both pathways depending on the metabolic state of the cell, whereas multicellular microorganisms, such as filamentous fungi, preferentially use respiration (2). Mucor racemosus, a dimorphic fungus that can grow either in a unicellular (yeast-like) or a multicellular (mycelial) form, also uses both; the unicellular form exploits fermentation, whereas the multicellular form is capable of respiration (3–5).

S. cerevisiae preferentially ferments glucose, even in the presence of oxygen, producing ethanol and CO₂ by anaerobic metabolism. Only after exhaustion of the available glucose is respiration activated, and the yeast cells then use the ethanol as a carbon and energy source for aerobic metabolism. The switch from anaerobic to aerobic metabolism, referred to as the diauxic shift, has stimulated a profusion of research on metabolic regulation in S. cerevisiae and other eukaryotic microorganisms (6–9). However, several basic questions remain unanswered. For example, why do eukaryotic microorganisms other than S. cerevisiae preferentially obtain energy from glucose by respiration and, therefore, do not undergo a diauxic shift? What factor(s) determine(s) this difference and at what molecular level has selection operated? The answers to these questions will have a critical impact not only on our basic knowledge of the metabolic regulation of glucose utilization and its evolution but also on the potential use of eukaryotic microorganisms for metabolic engineering and the production of useful compounds.

To carry out a comprehensive investigation aimed at understanding these differences, we have established an EST data base for the filamentous fungus Trichoderma reesei. Using the complementary DNA microarray technology we analyzed the gene expression profile during glucose exhaustion and compared it to the temporal program of gene expression accompanying the metabolic shift from fermentation to respiration in S. cerevisiae (6). The fungus T. reesei was chosen for this study because its natural habitats and nutritional requirements are very different from those of S. cerevisiae. Although relatively high concentrations of sugars prevail in the natural habitats of S. cerevisiae, the ubiquitous soil inhabitant T. reesei (10) has

* This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo Grants Genoma-FAPESP 97/13461–1, FAPESP 97/5267–0, and PADCT-CNpq 62.0533/98–6. The costs of publication of this article were defrayed in part by the payment of page charges.

† Present address: Dept. of Biochemical and Pharmaceutical Technology, Pharmaceutical Sciences School, University of São Paulo, São Paulo SP 05508–900, Brazil.

‡ To whom correspondence should be addressed. Tel.: 55-11-3091-3848; Fax: 55-11-3091-3848; E-mail address: dorry@iq.usp.br.
adapted to a nutrient-poor environment in which it exploits extracellular hydrolases, such as cellulase, to obtain glucose from polysaccharides (11). In addition, the genus *Trichoderma* includes species of economic importance. Enzymes produced by *Trichoderma* are used in the textile, food, and paper industries (12–14). Moreover, strains of *Trichoderma* that produce chitinolytic enzymes are mycoparasitic and can be used as biocontrol agents against plant-pathogenic fungi (15).

This report is based on the sequences of the 5’ ends of 2835 randomly selected cDNA clones that corresponded to 1151 unique transcripts. The complete sequence of the mitochondrial genome of *T. reesei* is also presented. Putative functions were assigned to 36.0% of these transcripts, unknown proteins represent 3.0%, whereas 61% of the ESTs showed no significant similarity to any other sequence in the data base, indicating that these sequences are specific to filamentous fungi and/or *T. reesei*. We also show that patterns of glucose-dependent regulation of gene transcription in *S. cerevisiae* and *T. reesei* differ in regard to critical genes whose products control the direction of flow of primary metabolites. Although the expression of genes for proteins involved in the tricarboxylic acid cycle and in mitochondrial respiration is repressed strongly in *S. cerevisiae* in the presence of glucose, in *T. reesei* these genes remain active under these conditions. Thus, in *T. reesei*, in contrast to *S. cerevisiae*, aerobic metabolism prevails in the presence of high levels of glucose.

**EXPERIMENTAL PROCEDURES**

*Media, Growth Conditions, and Metabolyte Analysis—* *T. reesei*, strain QM 9414, was obtained from the American Type Culture Collection (ATCC 26921). A 0.5-liter inoculum (containing 10^7 spores/ml) was added to a 14-liter fermentation vessel containing 10 liters of culture medium (16) supplemented with glucose at a final concentration of 100 mM. The culture was maintained at 28 °C with constant agitation and aeration. Aliquots of the culture were withdrawn, as indicated, and mycelium was collected by filtration and frozen in liquid nitrogen.

Glucose concentration in the culture supernatants was measured using a SERA-PAK kit (Bayer). Ethanol in the culture supernatants and acetate were measured enzymatically using the TC acetic acid and TC ethanol kits obtained from Roche Molecular Biochemicals.

**cDNA Library—** Total cellular RNA was extracted from glycerol-grown *T. reesei* cultures by the guanidium isothiocyanate procedure (17), and poly(A)^+ RNA was purified using oligo(dT) chromatography. A unidirectional cDNA library was constructed in the Uni-ZAP XR vector. In vivo excision of pBluescript plasmids was performed in *Escherichia coli* SOLR (Stratagene). To assess the quality of the library, the ratio of recombinants to non-recombinants and the average size of the cDNA inserts were determined by PCR analysis of the DNA from 96 individual clones.

**DNA Sequencing—** Mitochondrial DNA was isolated by cesium chloride/bisbenzimide density gradient centrifugation (18). Shotgun libraries were constructed from sheared mitochondrial DNA cloned into pUC18. Plasmid DNA from individual colonies was prepared with the Concert rapid plasmid miniprep system (Invitrogen), and DNA sequencing reactions were performed using the BigDye terminator cycle sequencing kit (PerkinElmer Life Sciences) and the M13 reverse and M13 (~20) primers (Stratagene). For ESTs, single-pass sequences of the 5’ ends of cDNAs were performed. Samples were loaded on an ABI 377 DNA sequencer (PerkinElmer Life Sciences) for automated sequence analysis.

**Computational Analysis—** Sequences were edited for each EST using the program phred/phrap/cor (19–21). Only ESTs with a minimum length of 150 bases and a phred quality value of at least 20 were considered for further analysis. Edited sequences were translated and used as query sequences to search the GenBank non-redundant protein data base by using the program BLASTX (22) at the National Center for Biotechnology Information (NCBI). Scores ≥80 were considered to be significant, and the top-scoring genes were used to group the transcripts into functional clusters. For computational and graphical analysis of the microarrays data we employed the Cluster and TreeView programs (23). Mitochondrial DNA consensus was generated by phred+phrap+consed (19–21). ORFs were predicted by ORF finder (NCBI) and searched against the GenBank non-redundant protein data base using the program BLASTP (22). tRNAs and rRNAs were located using tRNAscan-SE v.1.11 program (24) and BLASTN, respectively (25).

**Microarray Analysis—** Inserts were amplified by PCR in a 96-well format using M13 reverse and M13 (~20) primers (Stratagene). PCR products were then purified in a 96-well filtration plate using the Millipore MultiScreen Assay System. Each PCR product was verified by agarose gel electrophoresis and was considered correct if the amplified product resulted in a single band. These DNA were spotted on glass slides and hybridized with fluorescently labeled cDNA prepared by reverse transcription in the presence of Cy5 or Cy5-labeled deoxyuridine triphosphate (Glass fluorescent kit; CLONTECH). cDNA prepared from cells harvested at 83 mM glucose was labeled with Cy3 (reference sample), and those prepared at each later time were labeled with the Cy5 fluor. Hybridization, image analysis, and integration were performed with a GeneTac hybridization station, GeneTAC biochip analyzer, and GeneTAC integrator 3.0.1 (Genomic Solutions, www.genomicsolutions.com).

**RESULTS**

**cDNA Library Analysis—** A unidirectional cDNA library was constructed from mycelia of *T. reesei*, grown on glycerol as the sole carbon source as described under “Experimental Procedures.” The library was named TrEST-A, and we randomly selected 4320 clones for sequencing. PCR analysis of 96 individual clones revealed that 99% produced an amplification product with an average size of 1.2 kb. We obtained 2835 ESTs with a minimum length of 150 bases and a Phred quality value of at least 20. Of the 2835 ESTs, 808 sequences remained as singletons, and 2027 sequences formed 343 clusters. Therefore, this analysis shows that we obtained the partial sequences of 1151 expressed genes of *T. reesei*. The clusters ranged in size from 2 (177 clusters) to 90 (1 cluster) sequences.

Using BLASTX (22) and a stringency score ≥80, the total number of ESTs that could be assigned a cellular role on the basis of sequence similarity to proteins with known function was 348. The remaining ESTs are either unclassified (52 sequences), show similarity to sequences of unknown function (19 sequences), or have no significant similarity to any protein sequences in the data bases (no matches, 732 sequences). Those ESTs that encode putative protein sequences that show similarity to products in the NCBI non-redundant data base were classified into functional groups (Fig. 1). The functional groups presented in Fig. 1 are principally based on the classification developed at the Institute for Genomic Research (TIGR, Rockville, MD; available at www.tigr.org/dacs/tiger-scripts/egad_scripts/role_report.spl) (25). The complete list of ESTs classified into functional groups is available through the Internet (trichoderma.iq.usp.br/TrEST.html). Most of the known transcripts belong to groups related to housekeeping genes such as those involved in metabolism (14%), protein synthesis (9%), and RNA synthesis (2%). The high percentage of the ESTs that showed no hits (61%) most probably reflect the absence in the data base of a completely sequenced genome related to filamentous fungi. In fact the data base contains only 98 sequences from *T. reesei*. Therefore, our data represent an increase of more than 10-fold in the number of *T. reesei* expressed genes in the data base.

**Gene Expression Analysis during Glucose Exhaustion—** Homology searches using the sequenced cDNAs against the GenBank data base revealed that sufficient coverage had been achieved to allow a comprehensive study of the gene expression pattern during glucose exhaustion. Using complementary DNA microarray technology we analyzed the expression of the available set of *T. reesei* genes after attachment to glass slides. We compared transcript populations from cells harvested when glucose reached 83 mM to those expressed at various times as the glucose level declined (Fig. 2). Fluorescently labeled cDNA

---

1 The abbreviation used is: EST, expressed sequence tag.
was prepared from mRNA isolated from cells at 83 mM glucose in the presence of Cy3 (green)-labeled dUTP and from mRNA obtained at 46, 11, 1, and 0 mM glucose in the presence of Cy5 (red)-labeled dUTP. The labeled cDNAs were mixed and hybridized, in duplicate, to the microarrays.

The first striking finding was that many genes coding for enzymes of the tricarboxylic acid cycle were not repressed in glucose-rich medium, and those that were, citrate synthase and α-ketoglutarate dehydrogenase, were only partially repressed. We identified 14 genes that are expressed at ≥2-fold higher levels as glucose is depleted from the growth medium (Fig. 3). An overall view of the changes in the expression of genes involved in the metabolic pathways controlling the utilization of glucose in T. reesei is shown in Fig. 4. For purposes of comparison we also present the same portions of the metabolic pathways of S. cerevisiae as reported previously by DeRisi et al. (6).

The abundances of transcripts encoding enzymes of the glycolytic pathway in T. reesei were either unaffected or, in some cases, decreased slightly upon glucose exhaustion. The gene encoding enolase, however, is highly expressed in the presence of glucose-rich medium and is markedly repressed on depletion of the sugar. Thus, as in S. cerevisiae, up-regulation of the glycolytic transcripts in the presence of glucose will increase the flow of metabolites through the glycolytic pathway to yield pyruvate. The flow of metabolites in this direction is facilitated by the fact that two enzymes involved in the first steps of the pentose phosphate pathway, glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase, are expressed only at relatively low levels in the presence of glucose (Fig. 3).

However, a striking difference between the two microorganisms concerns the fate of pyruvate, as is evident from the pattern of expression of the transcripts of the genes for tricarboxylic acid cycle enzymes seen upon depletion of glucose.
Whereas high concentrations of glucose strongly repress genes encoding enzymes of the tricarboxylic acid cycle in *S. cerevisiae*, the corresponding transcripts in *T. reesei* behave quite differently. We have identified the genes for 5 of the 8 successive reaction steps in the tricarboxylic acid cycle. Glucose partially represses expression of the genes for citrate synthase and α-ketoglutarate dehydrogenase, whereas those for isocitrate dehydrogenase, succinate dehydrogenase, and malate dehydrogenase are unaffected (Fig. 4). If expression of these enzymes are controlled mainly at the transcriptional level, then pyruvate will be oxidized because of the higher level of expression of the tricarboxylic acid cycle mRNAs when glucose concentration is high. In contrast, in *S. cerevisiae*, pyruvate is...
channeled to acetaldehyde under these conditions, as a result of the strong repression of transcription of the genes for enzymes of the tricarboxylic acid cycle by glucose.

A second difference between the two species concerns the fate of the acetaldehyde formed by the decarboxylation of pyruvate by pyruvate decarboxylase, which is up-regulated in both microorganisms in the presence of glucose. In *S. cerevisiae*, the acetaldehyde formed is reduced to ethanol by NADH in a reaction catalyzed by alcohol dehydrogenase and is not converted to acetic acid. In contrast, in *T. reesei* we have identified two paralogous genes for aldehyde dehydrogenase, the ALD2 transcript is strongly repressed by glucose, but ALD1 is not affected. If both enzymes have comparable specificity, then acetaldehyde will be converted to acetate in *T. reesei* even in the presence of glucose. To address this question, we measured the concentration of ethanol and acetate after the addition of a high concentration of glucose to *T. reesei* culture grown in the presence of glycerol. The results show that, in contrast to *S. cerevisiae*, ethanol concentration did not change, whereas the concentration of acetate increased upon the addition of glucose (Fig. 5). This result indicates that, although one of the two paralogous genes for aldehyde dehydrogenase is repressed by glucose, acetate will be produced in the presence of a high concentration of glucose.

Upon exhaustion of glucose, the activation of the gene encoding acetyl-coenzyme A synthase will allow the entry of acetate, produced via the pyruvate bypass route, to replenish the tricarboxylic acid cycle. Under these conditions, the genes encoding phosphoenolpyruvate carboxykinase will also be activated, allowing the tricarboxylic acid cycle intermediates to flow via oxaloacetate to fuel the gluconeogenic pathway (Fig. 4).

We validated the expression of selected genes involved in the metabolic pathways controlling the utilization of glucose by Northern analysis. The results are presented in Fig. 6 and were found to be in agreement with the results obtained from the microarrays.

**Mitochondrial Activity**—Aerobic metabolism requires the expression of proteins involved in mitochondrial activity and the flow of electrons and protons through the complex of respiratory chain proteins that are encoded by mitochondrial and nuclear genes. Therefore, to determine whether transcripts encoded by the mitochondrial genome of *T. reesei* are subject to strong repression by glucose, as in *S. cerevisiae*, the complete sequence of the mitochondrial genome of the filamentous fungus was determined. The 42,130-bp circular mitochondrial DNA encodes 15 polypeptides, 2 rRNAs, and 25 tRNAs, all of which are transcribed from the same DNA strand (Fig. 7A). We have measured the abundance of several RNAs transcribed from different parts of the mitochondrial genome. The probes used are indicated in Fig. 7A and included three synthetic oligonucleotides complementary to gene sequences for cytochrome *c* oxidase polypeptide III (COX III) and for subunits 2 and 5 of NADH ubiquinone oxidoreductase (ND2, ND5). Synthetic oligonucleotides complementary to these genes were used as probes. C, Northern blot showing the effect of glucose concentration on the level of transcripts of the nuclear genes coding for COX V and COX VI. The actin transcript (ACT) is included as a control (34).

![Map of *T. reesei* mtDNA and effect of glucose on the expression of mitochondrial and nuclear transcripts coding for mitochondrial proteins](Image)

Fig. 7. Map of *T. reesei* mtDNA and effect of glucose on the expression of mitochondrial and nuclear transcripts coding for mitochondrial proteins. A, the unique direction of transcription is indicated by the arrow above the map; exons are presented as filled boxes and introns as open boxes. Genes that were used for Northern analysis are marked in red. B, Northern blot showing the effect of glucose concentration on the level of the transcripts (marked in red in A) of the genes for cytochrome *c* oxidase subunit III (COX III) and subunits 2 and 5 of NADH-ubiquinone oxidoreductase (ND2, ND5). Synthetic oligonucleotides complementary to these genes were used as probes. C, Northern blot showing the effect of glucose concentration on the level of transcripts of the nuclear genes coding for COX V and COX VI. The actin transcript (ACT) is included as a control (34).
DISCUSSION

The transcript profiles of *T. reesei* during glucose exhaustion presented in this report provide a description of the molecular basis of the shunting of the end product of the glycolytic pathway, pyruvate, into aerobic rather than anaerobic metabolism. They have obvious implications for the long standing question of why some eukaryotic microorganisms, such as *T. reesei*, utilize glucose by aerobic metabolism. First, in the presence of glucose-rich medium, the expression of the genes encoding the enzymes of the tricarboxylic acid cycle will allow the available pyruvate to fuel the tricarboxylic acid cycle. In addition, pyruvate can be converted to acetaldehyde and then to acetate via the pyruvate bypass route. Furthermore, the conversion of acetaldehyde to acetate and not to ethanol as in *S. cerevisiae* precludes the generation of the NAD⁺ required for anaerobic metabolism. Second, analysis of the expression of several mitochondrial and nuclear gene-encoding proteins involved in mitochondrial respiration confirms that *T. reesei* is able to carry out respiration in glucose-rich medium. Regulation of gene transcription by glucose in *S. cerevisiae* and *T. reesei* therefore differs with respect to critical genes, the products of which control the direction of the flow of metabolites. Although the expression of genes involved in the tricarboxylic acid cycle and in mitochondrial respiration is repressed strongly in *S. cerevisiae* in the presence of glucose, in *T. reesei* these genes remain active under these conditions. Thus, aerobic metabolism will prevail in *T. reesei* in the presence of glucose-rich medium. The gene expression profile described for *T. reesei* in this work is most probably also used by other multicellular microorganisms to obtain energy by respiration, rather than fermentation, in the presence of high levels of glucose. The fact that, in *Aspergillus nidulans*, the cytochrome c gene (cyCA) is also not repressed by glucose (28) supports this contention.

In *S. cerevisiae* and *Kluvyomyces lactis*, the main regulatory effect of glucose occurs at the transcriptional level (8, 29). This also seems to be the case in *T. reesei* because the primary effect is on the level of mRNA of critical steps directing the flow of metabolites to aerobic rather than anaerobic metabolism, corresponding well with the fact that *T. reesei* is a preferentially respiratory microorganism. Future analysis of metabolic flux using ¹³C should be of great value to accompany the metabolic gene expression studies reported in our work.

Metabolic engineering of eukaryotic cells for the production of useful compounds represents a formidable challenge. However, the power of this approach is evident from the recent report that an obligate photosynthetic microalga can be converted into a heterotrophic organism by introducing a human gene that encodes a glucose transporter (30). *T. reesei* is capable of hydrolyzing cellulose to glucose (11). The gene expression profile described in this report provides valuable information for the metabolic engineering necessary to turn this preferentially respiratory microorganism into a fermenter, a step that will be required if *T. reesei* is to be used for the efficient production of ethanol fuel from cellulose biomass.

Finally, it is believed that multicellular organisms capable of respiration arose from unicellular fermenters early in the course of evolution, after the concentration of oxygen in the atmosphere began to rise and following the engulfment of aerobic bacteria that evolved into mitochondria (31). It is not surprising, therefore, that the molecular alterations in response to glucose, which determined the switch from fermentation to respiration, are observed in both nuclear and mitochondrial genes coding for mitochondrial enzymes and proteins. Because effectors specific for glucose repression are conserved in yeast and filamentous fungi (32), the crucial alterations most probably occurred in the target sequences for the glucose repressor in the promoters of the relevant genes, which prevented the binding of repressor.

Acknowledgment—We thank Boris Stambuk for valuable discussions.

REFERENCES