

# Matching the proteome to the genome: the microbody of penicillin-producing *Penicillium chrysogenum* cells

Jan A. K. W. Kiel · Marco A. van den Berg ·  
Fabrizia Fusetti · Bert Poolman ·  
Roel A. L. Bovenberg · Marten Veenhuis ·  
Ida J. van der Klei

Received: 23 December 2008 / Revised: 5 January 2009 / Accepted: 5 January 2009 / Published online: 21 January 2009  
© The Author(s) 2009. This article is published with open access at Springerlink.com

**Abstract** In the filamentous fungus *Penicillium chrysogenum*, microbodies are essential for penicillin biosynthesis. To better understand the role of these organelles in antibiotics production, we determined the matrix enzyme contents of *P. chrysogenum* microbodies. Using a novel in silico approach, we first obtained a catalogue of 200 *P. chrysogenum* proteins with putative microbody targeting signals (PTSs). This included two orthologs of proteins involved in cephalosporin biosynthesis, which we demonstrate to be bona fide microbody matrix constituents. Subsequently, we performed a proteomics based inventory of *P. chrysogenum* microbody matrix proteins using nano-LC-MS/MS analysis. We identified 89 microbody proteins,

79 with a PTS, including the two known microbody-borne penicillin biosynthesis enzymes, isopenicillin N:acyl CoA acyltransferase and phenylacetyl-CoA ligase. Comparative analysis revealed that 69 out of 79 PTS proteins identified experimentally were in the reference list. A prominent microbody protein was identified as a novel fumarate reductase-cytochrome b5 fusion protein, which contains an internal PTS2 between the two functional domains. We show that this protein indeed localizes to *P. chrysogenum* microbodies.

**Keywords**  $\beta$ -Lactam antibiotics · Filamentous fungus · In silico analysis · Proteomics · Peroxisome

**Electronic supplementary material** The online version of this article (doi:10.1007/s10142-009-0110-6) contains supplementary material, which is available to authorized users.

J. A. K. W. Kiel (✉) · M. Veenhuis · I. J. van der Klei  
Molecular Cell Biology, Groningen Biomolecular Sciences  
and Biotechnology Institute (GBB), University of Groningen,  
P.O. Box 14, 9750 AA Haren, The Netherlands  
e-mail: J.A.K.W.Kiel@RUG.nl

J. A. K. W. Kiel · M. Veenhuis · I. J. van der Klei  
Kluyver Centre for Genomics of Industrial Fermentation,  
Julianalaan 67,  
2628 BC Delft, The Netherlands

M. A. van den Berg · R. A. L. Bovenberg  
DSM Anti-Infectives, DSM Gist,  
P.O. Box 425, 2600 AK Delft, The Netherlands

F. Fusetti · B. Poolman  
Department of Biochemistry, Groningen Biomolecular  
Sciences and Biotechnology Institute (GBB) & Zernike Institute  
for Advanced Materials (ZIAM), University of Groningen,  
Nijenborgh 4,  
9747 AG Groningen, The Netherlands

## Introduction

$\beta$ -Lactam antibiotics like penicillin have had a profound impact on human health and belong to one of the largest-selling classes of drugs worldwide (Bruggink and Roy 2001). Since the discovery of penicillin by Alexander Fleming, strain improvement procedures using the filamentous fungus *Penicillium chrysogenum* have enhanced the production of this antibiotic at least a 1,000-fold. It was observed that *P. chrysogenum* strains with superior penicillin production capabilities not only had increased levels of the penicillin biosynthetic enzymes but also an enhanced number and volume fraction of microbodies in the cell (Müller et al. 1991). Furthermore, microbodies were seen to proliferate in *P. chrysogenum* cells cultivated on media that induce penicillin production, which was also observed with the less potent penicillin producer *Aspergillus nidulans* (Valenciano et al. 1998). The importance of microbodies in penicillin production was further emphasized by the finding that an increase in the number of microbodies by genetic

manipulation of the level of the peroxin Pex11p enhanced antibiotic production by *P. chrysogenum* (Kiel et al. 2005).

Microbodies are important organelles that are present in all eukaryotic cells. These single membrane-bound organelles contain a protein-rich matrix consisting of enzymes involved in highly diverse metabolic pathways (Veenhuis and Harder 1991; van den Bosch et al. 1992). Microbodies are crucial organelles in plants (Schumann et al. 2003), *Trypanosomes* (Guerra-Giraldez et al. 2002) as well as in mammals, as demonstrated by the discovery of inherited human metabolic disorders, that are associated with defects in microbody biogenesis or function (Gould and Valle 2000). Microbody matrix enzymes are nuclear-encoded proteins that are synthesized on cytosolic polysomes and become post-translationally imported into their target organelle (Lazarow and Fujiki 1985). To allow correct sorting, these enzymes contain specific microbody (peroxisomal) targeting signals (designated PTSs), which are recognized by specific receptor proteins in the cytosol. The most abundant signal, PTS1, located at the carboxy terminus of microbody enzymes, has the generally accepted consensus sequence -[SCA]-[KRH]-[LM]-COOH (Subramani 1993). A second PTS, designated PTS2, identified in only a subset of microbody matrix enzymes, is located within the N-terminal 50 amino acids (Rachubinski and Subramani 1995). Recently, Petriv et al. (2004) have refined the consensus sequence of this nonapeptide to -[RK]-[LVIQ]-x<sub>(2)</sub>-[LVIHQ]-[LSGAK]-x-[HQ]-[LAF]-, which should describe all known variants of PTS2.

The importance of microbodies in penicillin production is underscored by the identification of a PTS1-related sequence on key enzymes of penicillin biosynthesis, isopenicillin N:acyl CoA acyltransferase (IAT; -ARL-COOH; Barredo et al. 1989) and two phenylacetyl-CoA ligases (PclA, -SKI-COOH and PhlB, -AKL-COOH; Lamas-Maceiras et al. 2006; Wang et al. 2007). Indeed, IAT and PclA have been shown to be located in the microbody matrix (Müller et al. 1991; Gidijala et al. 2007; W.H. Meijer et al. unpublished data). Furthermore, a mutant strain of *P. chrysogenum* in which the PTS1 of IAT was removed was unable to synthesize penicillin (Müller et al. 1992). Unfortunately, despite the availability of multiple

genome sequences, detailed knowledge of the microbody proteome in filamentous fungi is still limited. Although identification of proteins with a PTS in fungal proteomes is feasible, the available consensus sequences for both PTS1 and PTS2 are biased because they have been created using mainly mammalian and baker's yeast proteins.

Our goal was obtaining a detailed understanding of the proteins involved in microbody biogenesis, stability, and function in the  $\beta$ -lactam producer *P. chrysogenum*. In previous studies, we have identified the *P. chrysogenum* proteins involved in microbody biogenesis and turn-over (Kiel et al. 2006; Meijer et al. 2007). In the current study, we investigated the matrix contents of microbodies in penicillin-producing *P. chrysogenum* hyphae by mass spectrometry (MS)-based proteomics. Moreover, to provide a reference set for the proteomics approach, we have identified all putative PTS1 and PTS2 proteins encoded by the *P. chrysogenum* genome by in silico analysis. Our analyses indicate that microbodies in *P. chrysogenum* are not specialized structures solely devoted to penicillin biosynthesis, but are in fact multi-purpose organelles. Furthermore, during penicillin production, *P. chrysogenum* microbodies only contain a subset of the available 200 PTS proteins.

## Experimental procedures

### Strains and media

The fungal strains used in this study are indicated in Table 1. *P. chrysogenum* cells were grown on (1) YGG medium containing 0.8% KCl, 1.6% glucose, 0.67% yeast nitrogen base (Difco), 0.15% citric acid, 0.6% K<sub>2</sub>HPO<sub>4</sub>, 0.2% yeast extract pH 6.2 supplemented with penicillin and streptomycin (Gibco) or (2) penicillin production medium (Hillenga et al. 1994) supplemented with 0.05% phenoxyacetic acid. *P. chrysogenum* transformants were selected on media with acetamide as sole nitrogen source as described (Cantoral et al. 1987). Fluoroacetamide plates (Royer et al. 1999) were used to select recombinant spores that had lost the *A. nidulans amdS* gene.

**Table 1** *P. chrysogenum* strains used in this study

Strain	Genotype/characteristics	Reference
DS17690	High penicillin-producing strain	Harris et al. (2006)
DsRed.SKL (AmdS <sup>+</sup> )	DS17690 with integrated P <sub>pcbC</sub> -DsRed-Tl.SKL-T <sub>penDE</sub> cassette; (AmdS <sup>+</sup> )	This study
DsRed.SKL	Fluoroacetamide resistant derivative of DsRed.SKL (AmdS <sup>-</sup> )	This study
DsRed.SKL eGFP.Pc12g03090	DsRed.SKL producing eGFP.Pc12g03090 fusion protein	This study
DsRed.SKL eGFP.Pc20g13500	DsRed.SKL producing eGFP.Pc20g13500 fusion protein	This study
DsRed.SKL eGFP.Pc22g13680	DsRed.SKL producing eGFP.Pc22g13680 fusion protein	This study

*Escherichia coli* DH5 $\alpha$  was used for plasmid constructions and was grown on LB-medium (Sambrook et al. 1989) supplemented with the appropriate antibiotics.

## DNA procedures

Oligonucleotide primers and plasmids used in this study are indicated in Tables 2 and 3, respectively. Standard recombinant DNA manipulations (Sambrook et al. 1989) were used throughout this study. Preparation and transformation of *P. chrysogenum* protoplasts was performed as described by Cantoral et al. (1987). Polymerase chain reactions were carried out with *Pwo* polymerase or Expand high-fidelity polymerase (Roche). Cloned PCR fragments were sequenced to confirm the correct DNA sequence. For analysis of *P. chrysogenum* DNA and deduced amino acid sequences, the Clone Manager 5 program (Scientific and Educational Software, Durham, USA) was used. Homology searches were performed using either a stand-alone version of the BLAST program (Altschul et al. 1990) or the BLAST algorithm of the National Center for Biotechnological Information (Altschul et al. 1997). Alignments of amino acid sequences were constructed using the Clustal\_X program (Thompson et al. 1997) and displayed using the GeneDoc program (<http://www.psc.edu/biomed/genedoc>). Motif searches were performed using the Genome2D program (Baerends et al. 2004). Graphics showing the distribution and amino acid composition of consensus sequences were generated using Weblogo (Crooks et al. 2004; <http://weblogo.berkeley.edu/>). Secondary structure prediction was performed using JPRED3 (<http://www.compbio.dundee.ac.uk/~www-jpred/>).

## Construction of plasmids

A *P. chrysogenum* DS17690 derivative producing the red fluorescent microbody marker protein DsRed.SKL was constructed as follows: First, a 743 bp *Bam*HI–*Sma*I fragment containing the *DsRed-T1.SKL* gene was isolated from plasmid pHIPZ4-DsRed-T1.SKL and cloned between the *Bam*HI and *Sma*I sites of pGBRH2. This resulted in plasmid pBBK-001. From this plasmid, a 2.2 kb *Not*I fragment containing the  $P_{pcbC}$ -*DsRed-T1.SKL*- $T_{penDE}$  ex-

pression cassette was isolated and co-transformed into *P. chrysogenum* DS17690 protoplasts with a 6 kb *Not*I–*Spe*I fragment from plasmid pNiGANi containing an *A. nidulans amdS* expression cassette flanked by two identical *P. chrysogenum niaD* fragments. Transformants able to grow on plates with acetamide as sole nitrogen source were selected and analyzed for the presence of red fluorescent microbodies. From the resulting strain (DsRed.SKL (AmdS<sup>+</sup>)), a derivative that had lost its *amdS* cassette was isolated by fluoroacetamide selection (Royer et al. 1999), resulting in strain DsRed.SKL.

Plasmids containing  $P_{pcbC}$ -*eGFP*. *Pc12g03090*- $T_{penDE}$ ,  $P_{pcbC}$ -*eGFP*.*Pc20g13500*- $T_{penDE}$ , and  $P_{pcbC}$ -*eGFP*.*Pc22g13680*- $T_{penDE}$  expression cassettes encoding the fluorescent hybrid proteins eGFP.Pc12g03090, eGFP.Pc20g13500, and eGFP.Pc22g13680, respectively, were constructed using multisite Gateway technology (Invitrogen). First, we created plasmid pGBRH2-eGFP containing a  $P_{pcbC}$ -*eGFP*- $T_{penDE}$  expression cassette by cloning a 734 bp *Hind*III–*Sma*I fragment from pRBG14, containing the *eGFP* gene, between the *Hind*III and *Sma*I sites of pGBRH2. From pGBRH2.eGFP, a 1.75 kb DNA fragment comprising  $P_{pcbC}$  and *eGFP* lacking a stop codon was amplified with primers  $P_{pcbC}$ Attb4f and XFPAttb3r. This fragment was recombined into the pDONR<sup>TM</sup> P4-P1R vector, thereby creating pENTR41/ $P_{pcbC}$ -eGFP. In addition, by PCR using the primer combinations Pc12g03090-Fw + Pc12g03090-Rv, Pc20g13500-Fw2 + Pc20g13500-Rv, and Pc22g13680-Fw + Pc22g13680-Rv, we amplified the 2.0 kb *Pc12g03090*, 2.1 kb *Pc20g13500*, and 1.2 kb *Pc22g13680* coding sequences, respectively, using either *P. chrysogenum* DS17690 genomic DNA (Pc20g13500 and Pc22g13680) or DNA from a cDNA library (Pc12g03090; see Kiel et al. 2000) as templates. These fragments were recombined into vector pDONR<sup>TM</sup> 221 resulting in plasmids pENTR221/Pc12g03090, pENTR221/Pc20g13500, and pENTR221/Pc22g13680, respectively. Finally, the resulting plasmids were recombined with plasmid pENTR23/HIS8. $T_{penDE}$  containing the *penDE* terminator and vector pDEST<sup>TM</sup> R4-R3/AMDS thereby creating pEXP/eGFP.Pc12g03090, pEXP/eGFP.Pc20g13500, and pEXP/eGFP.Pc22g13680. For localization of the fluorescent fusion proteins encoded by the plasmids in *P. chrysogenum*, the plasmids were linearized

**Table 2** Primers used in this study

$P_{pcbC}$ Attb4f	5' GGGGACAACCTTTGTATAGAAAAGTTGTCCTTATACTGGGCCTGCTGCATTGG 3'
XFPAttb3r	5' GGGGACTGCTTTTTGTACAAAAGCTGCTTGTACAGCTCGTCCATGCCG 3'
Pc12g03090-Fw	5' GGGGACAAGTTTGTACAAAAAAGCAGGCTTTATGGCTGCACCTCAAGTAATTGTGGTC 3'
Pc12g03090-Rv	5' GGGGACCACTTTGTACAAGAAAGCTGGGTTTCTACAGCTCCAAGCTGGGGGTCTG 3'
Pc20g13500-Fw2	5' GGGGACAAGTTTGTACAAAAAAGCAGGCTTGATGGATAGGTACGTGGCCAAACAG 3'
Pc20g13500-Rv	5' GGGGACCACTTTGTACAAGAAAGCTGGGTTTCCACAGCCGAACACTCCCACTCGAC 3'
Pc22g13680-Fw	5' GGGGACAAGTTTGTACAAAAAAGCAGGCTTTATGTCACCAATTGAAAGGAATCAAAGTC 3'
Pc22g13680-Rv	5' GGGGACCACTTTGTACAAGAAAGCTGGGTTTCAAATCTTTGACTTCCCATCTATCGC 3'

**Table 3** Plasmids used in this study

pGBRH2	pBluescript vector with <i>pcbC</i> promoter and <i>penDE</i> terminator	Kiel et al. (2005)
pGBRH2-eGFP	pGBRH2 with the <i>eGFP</i> gene	This study
pRBG14	plasmid with <i>eGFP</i> gene in polylinker	Lab collection
pHIPZ4-DsRed-T1.SKL	Plasmid containing <i>DsRed-T1.SKL</i> gene	Monastyrska et al. (2005)
pBBK-001	pGBRH2 with DsRed-T1.SKL	This study
pNiGaNi	Plasmid with <i>niaD</i> - <i>P<sub>gpdA</sub></i> - <i>An.amdS</i> - <i>T<sub>pcbC</sub></i> - <i>niaD</i> cassette	Lab collection
pDONR™ P4-P1R	Multisite Gateway vector	Invitrogen
pENTR41/ <i>P<sub>pcbC</sub></i> -eGFP	pDONR™ P4-P1R with <i>pcbC</i> promoter and <i>eGFP</i> gene lacking a stop codon	This study
pDONR™ 221	Multisite Gateway vector	Invitrogen
pENTR221/ <i>Pc12g03090</i>	pDONR221 with <i>Pc12g03090</i> CDS	This study
pENTR221/ <i>Pc20g13500</i>	pDONR221 with <i>Pc20g13500</i> CDS	This study
pENTR221/ <i>Pc22g13680</i>	pDONR221 with <i>Pc22g13680</i> CDS	This study
pENTR23/ <i>HIS8</i> . <i>T<sub>penDE</sub></i>	pDONR P2R-P3 containing His <sub>8</sub> tag and <i>penDE</i> terminator	W.H. Meijer et al. unpublished
pDEST™ R4-R3/AMDS	pDEST™ R4-R3 containing the <i>P<sub>gpdA</sub></i> - <i>An.amdS</i> - <i>T<sub>penDE</sub></i> cassette	W.H. Meijer et al., unpublished
pEXP/eGFP. <i>Pc12g03090</i>	pDEST™ R4-R3/AMDS with <i>P<sub>pcbC</sub></i> - <i>eGFP.Pc12g03090</i> - <i>T<sub>penDE</sub></i> cassette	This study
pEXP/eGFP. <i>Pc20g13500</i>	pDEST™ R4-R3/AMDS with <i>P<sub>pcbC</sub></i> - <i>eGFP.Pc20g13500</i> - <i>T<sub>penDE</sub></i> cassette	This study
pEXP/eGFP. <i>Pc22g13680</i>	pDEST™ R4-R3/AMDS with <i>P<sub>pcbC</sub></i> - <i>eGFP.Pc22g13680</i> - <i>T<sub>penDE</sub></i> cassette	This study

CDS coding sequence

with *AatIII* (pEXP/eGFP.*Pc12g03090*) or *DraI* (pEXP/*Pc20g13500* and pEXP/*Pc22g13680*) and transformed into protoplasts of *P. chrysogenum* DsRed.SK.L. Transformants able to grow on plates with acetamide as sole nitrogen source were selected and analyzed for the presence of eGFP fluorescence. This resulted in strains DsRed.SK.L eGFP.*Pc12g03090*, DsRed.SK.L eGFP.*Pc20g13500*, and DsRed.SK.L eGFP.*Pc22g13680*.

#### In silico analysis

Identification of PTS1 proteins encoded by the *P. chrysogenum* genome was performed in five steps. During step 1, we collected the amino acid sequences of 106 known microbody proteins (mainly from *Homo sapiens* and *Saccharomyces cerevisiae*—see Supplementary Table 1) and used these as queries in local BlastP searches against the *P. chrysogenum* protein database (van den Berg et al. 2008; consisting of 13,670 proteins) to identify homologous protein sequences. The expect threshold was set to  $e^{-004}$ . The resulting protein sequences were then manually analyzed for the presence of a sequence that conformed to the PTS1-related consensus 1: [SAPGTCLIVMHEDNQWYF]-[KRHNQ]-[LIVMF]-COOH. This resulted in the identification of 125 proteins. For step 2, we reasoned that the proteins identified in step 1 only represented a subset of proteins that have the identified PTS1-related sequences because PTS1 proteins with previously unidentified functions will have been missed in the initial BlastP searches. Therefore, all identified PTS1-related sequences were used as queries in a motif search on the *P. chrysogenum* protein database using Genome2D (Baerends et al. 2004). This resulted in identification of 329 proteins. For step 3, we assumed that a bona fide microbody protein

should also have a PTS1 in related species. Thus, all identified protein sequences were used in Blast searches to identify orthologs in filamentous ascomycetes (for a list of organisms used in this analysis, see Supplementary Table 2). Of the resulting hit lists, the C-terminal 3 amino acids of each identified protein were manually compared to the following more strict PTS1 consensus sequences: -[SA]-[RKHN]-[LI]-COOH, -[TGPCVEMNHW]-[KR]-L-COOH, and -A-R-F-COOH. When at least two filamentous ascomycetes other than *P. chrysogenum* the C terminus of the ortholog conformed to one of the strict consensus sequences, the *P. chrysogenum* protein was judged to be a bona fide PTS1 protein. When no true ortholog could be identified (or only in a single fungal species), paralogs with the highest similarity were also scored. Step 4: During the analysis described in step 3, we occasionally identified proteins from filamentous ascomycetes related to *P. chrysogenum* with a C-terminal sequence that conformed to the original PTS1-related consensus 1, but that was not represented in the set of 125 *P. chrysogenum* proteins initially identified. This might indicate that our PTS1 analysis had not been exhaustive. Therefore, we repeated step 2 of our analysis with these newly identified PTS1-related sequences and obtained an additional set of 82 proteins that were also further analyzed as described above. Step 5: Finally, after omitting a few false positives, i.e., orthologs of proteins that were published as being structural proteins, transcription factors, secreted proteins, or integral membrane spanning proteins, we obtained a list of 190 putative PTS1 proteins.

In a similar manner, we identified PTS2 containing proteins in *P. chrysogenum*. We started with the same set of Blast results as described in step 1 for PTS1 proteins (see above). These were analyzed for the presence of an N-

terminal sequence that conformed to the PTS2-related consensus 1: M-X<sub>(0,40)</sub>-[RK]-[LVIMQF]-X<sub>(5)</sub>-[HQKRN]-[LVIMAFY]-. Preliminary secondary structure analysis of known PTS2 sequences suggested that this sequence has the capability to form an  $\alpha$ -helical structure. As a result, during step 1, proline residues in the nonapeptide were disallowed. This resulted in a total of 38 proteins. In step 2, we used all the identified PTS2 sequences as queries to screen for *P. chrysogenum* proteins with PTS2 motifs based on the PTS2-related consensus 1 (see above) using Genome2D, again disallowing proteins with prolines in the nonapeptide. This resulted in the identification of 203 proteins. Subsequently, we analyzed whether the putative PTS2 signals were conserved in other filamentous ascomycetes. Orthologs were identified by Blast analyses, and the resulting protein sequences were analyzed for the presence of the strict PTS2 consensus: M-X<sub>(0,40)</sub>-[RK]-[LVIQ]-X<sub>(2)</sub>-[LVIHQ]-[LSGAKT]-X-[HQ]-[LAFIV]-. When in at least two filamentous ascomycetes other than *P. chrysogenum* the N terminus of the ortholog conformed to the strict consensus, the *P. chrysogenum* protein was judged to be a bona fide PTS2 protein. As described above, during this analysis, we sometimes identified proteins from filamentous ascomycetes related to *P. chrysogenum* with an N-terminal PTS2 sequence that conformed to the original PTS2-related consensus 1, but that was not present as a putative PTS2 in the set of 38 *P. chrysogenum* proteins initially identified. Again, we repeated step 2 of our analysis with these newly identified PTS2-related sequences and obtained an additional set of 16 proteins that were analyzed according to step 3. Finally, after omitting a few structural proteins, we obtained a list of ten putative PTS2 proteins.

#### General biochemical techniques

Crude extracts of *P. chrysogenum* cells were prepared as follows: after cultivation, one volume of culture was mixed with an equal volume of 25% trichloroacetic acid (TCA), and the mixture was subsequently placed overnight at  $-20^{\circ}\text{C}$ . After thawing on ice, the cells were pelleted by centrifugation (5 min, 14,000 rpm, eppendorf centrifuge,  $4^{\circ}\text{C}$ ), washed twice with 80% acetone ( $-20^{\circ}\text{C}$ ), and air-dried. The resulting pellet was suspended in 250  $\mu\text{l}$  1% SDS/0.1 M NaOH. Subsequently, an equal volume of 2 $\times$  SDS-PAA gel loading buffer was added, and the samples were boiled for 5 min at  $100^{\circ}\text{C}$ .

Protein concentrations were determined using either the Bio-Rad Protein Assay system or the RC/DC Protein Assay system (Bio-Rad) using bovine serum albumin as standard. SDS-PAGE and Western blotting were performed by established procedures. Western blots were decorated with polyclonal antibodies against selected *P. chrysogenum* proteins.

#### Isolation of microbodies

*P. chrysogenum* hyphae were harvested by filtration and washed with KC buffer (0.8 M KCL 10 mM citrate pH 6.2). Hyphae were protoplasted for 3 h at  $25^{\circ}\text{C}$  in KC buffer supplemented with 20 mg/ml lysing enzymes (Sigma L1412). Subsequently, protoplasts were separated from unprotoplasted material by filtration over glass wool, cooled to  $4^{\circ}\text{C}$ , and then washed successively with (1) KC buffer containing 1 mM PMSF, (2) a mixture of 50% KC buffer and 50% sorbitol buffer (1.2 M sorbitol in buffer B: 5 mM MES pH 5.5, 1 mM MgCl<sub>2</sub>, 1 mM EDTA) with 1 mM PMSF, and (3) sorbitol buffer with 1 mM PMSF. The protoplasts were then taken up in ice-cold sorbitol buffer supplemented with protease inhibitors (Complete, Roche +1 mM PMSF), lysed by 2 $\times$ 10 strokes of a potter and subjected to differential centrifugation (successively 10 min at 4,000 $\times g$ , 10 min at 6,000 $\times g$  and 20 min at 30,000 $\times g$  at  $4^{\circ}\text{C}$ ). The 30,000 $\times g$  pellet containing mainly mitochondria and microbodies was carefully suspended in ice-cold 40% sucrose in buffer B and loaded on discontinuous sucrose gradients (Douma et al. 1985). After centrifugation for 2.5 h at 32,000 $\times g$  and  $4^{\circ}\text{C}$ , approximately 20 fractions of 1.5 ml were taken from each gradient. Gradient fractions were either directly analyzed for sucrose concentrations and cytochrome *c* oxidase activities (Douma et al. 1985), lysed to obtain the matrix fraction (see below) or TCA-precipitated for Western blot analysis.

#### Lysis of microbody fractions

Purified microbodies were lysed by mixing sucrose gradient fractions with an equal volume of ice-cold 0.1 M Tris-HCl pH 8.0. After ultracentrifugation for 15 min at 200,000 $\times g$  at  $4^{\circ}\text{C}$ , the membrane pellet was resuspended in one volume of ice-cold 0.1 M Tris-HCl pH 8.0 and the mixture was again centrifuged. Supernatants were pooled and the proteins were TCA precipitated. Finally, proteins were dissolved in SDS-PAA gel loading buffer at a concentration of 2 mg/ml.

#### Separation of microbody matrix proteins by 1-D SDS/PAGE and in-gel tryptic digestion

For mass spectrometry analysis, 60  $\mu\text{g}$  of microbody matrix proteins were separated in duplo on 7.5% or 12.5% PAA minigels. After electrophoresis, PAA gels were stained with Bio Safe Coomassie<sup>R</sup> Blue G250 (Bio-Rad), followed by extensive destaining with water. Entire 7.5% gels (protein range >25 kDa) were cut in 32 gel slices. From 12.5% gels, only the lower portion with proteins <30 kDa was cut into 14 gel pieces. The 46 excised gel fractions were washed at room temperature with water, 30% acetonitrile in 25 mM

$\text{NH}_4\text{HCO}_3$ , and 50% acetonitrile in 25 mM  $\text{NH}_4\text{HCO}_3$  until they were completely destained. After drying under vacuum, proteins in the gel slices were reduced and alkylated by incubation in 10 mM DTT in 100 mM  $\text{NH}_4\text{HCO}_3$  for 30 min at 56°C followed by treatment with 55 mM iodoacetamide in 100 mM  $\text{NH}_4\text{HCO}_3$  for 30 min at room temperature. The gel slices were then washed with 50% acetonitrile in 25 mM  $\text{NH}_4\text{HCO}_3$ , dried under vacuum, and subsequently rehydrated in 20  $\mu\text{l}$  trypsin solution (Promega V5111; 10 ng/ $\mu\text{l}$  in 25 mM  $\text{NH}_4\text{HCO}_3$ ), followed by overnight incubation at 37°C. The tryptic peptides were recovered from the gel slices by three subsequent extractions with 75% acetonitrile 1.25% aqueous formic acid. The extracted peptides were pooled and concentrated under vacuum and redissolved in 50  $\mu\text{l}$  5% acetonitrile in 0.1% trifluoroacetic acid (TFA).

#### Liquid chromatography-mass spectrometry

MS analysis was performed using a MALDI-TOF/TOF 4800 Proteomics Analyzer (Applied Biosystems, Foster City, CA, USA). After initial MS analysis of all 46 tryptic peptide mixtures, 20  $\mu\text{l}$  of each fraction was taken, and 32 pools were prepared based on the estimated protein content of each fraction. The peptide mixtures were then separated on a C18 capillary column (75  $\mu\text{m}\times 150$  mm, 3  $\mu\text{m}$  particle size Zorbax 300SB-C18 HPLC column, Agilent Technologies, Amstelveen, The Netherlands) mounted on an Ettan MDLC nanoflow liquid chromatography system (Amersham Biosciences, Freiburg, Germany). The elution buffers used in the gradient were buffer A (0.07% aqueous TFA) and buffer B (80% acetonitrile, 0.05% aqueous TFA). A gradient from 5–50% B was performed in 90 min, at a flow rate of 200 nL/min. Column effluent was mixed 1:4 v/v with a matrix solution of 2.3 mg/ml  $\alpha$ -cyano-4-hydroxycinnamic acid (LaserBio Labs, Sophia-Antipolis, France) containing 2 nM angiotensin II and 4 nM adrenocorticotrophin hormone 18–36 (Sigma, Zwijndrecht, The Netherlands) as internal standards for mass calibration. Fractions of 16-s width were spotted on blank MALDI targets with a Probot system (LC Packings, Amsterdam, the Netherlands). MS data acquisition was performed in positive ion mode. Peptides with signal-to-noise level higher than 40 were selected for MS/MS analysis.

#### Database searching

Protein identification was carried out using Mascot (version 1.9.05, Matrix Science, London, UK) and X! Tandem ([www.thegpm.org](http://www.thegpm.org); version 2006.04.01.2) searching against the *P. chrysogenum* protein sequence database (van den Berg et al. 2008). To monitor the false positive rate, the protein database was combined with a reverse version of the same database.

Peptide mass tolerance was set to 0.2 Da, allowing for one missed trypsin cleavage. Oxidation of methionine and iodoacetamide derivation of cysteine was specified as variable modifications. The result files from each gel slice were combined and analyzed using the program Scaffold (version Scaffold-01\_07\_00, Proteome Software, Portland, OR, USA) to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the Peptide Prophet algorithm (Keller et al. 2002). Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least two identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii et al. 2003).

#### Ultrastructural analysis

*P. chrysogenum* cells, organellar pellets, and sucrose gradient fractions were fixed and prepared for electron microscopy as described previously (Waterham et al. 1994). For localization of fluorescent fusion proteins in *P. chrysogenum*, conidiospores of *P. chrysogenum* DsRed.SKL eGFP.Pc12g03090, DsRed.SKL eGFP.Pc20g13500, and DsRed.SKL eGFP.Pc22g13680 were germinated in penicillin production media supplemented with phenoxyacetic acid. After 24–48 h of cultivation, hyphae were visualized using a Zeiss LSM510 confocal laser scanning microscope (Zeiss Netherlands b.v. Weesp, The Netherlands) equipped with a Zeiss Plan—Apochromatic 63 $\times$  NA 1.4 objective. eGFP fluorescence was analyzed by excitation with a 488 nm Ar/Kr laser with 2% output, and fluorescence was detected by a 500–530 nm Photo Multiplier Tube. DsRed fluorescence was analyzed by excitation with a 545 nm He/Ne laser with 20% output, and fluorescence was detected by a 565–650 nm Photo Multiplier Tube.

#### Results

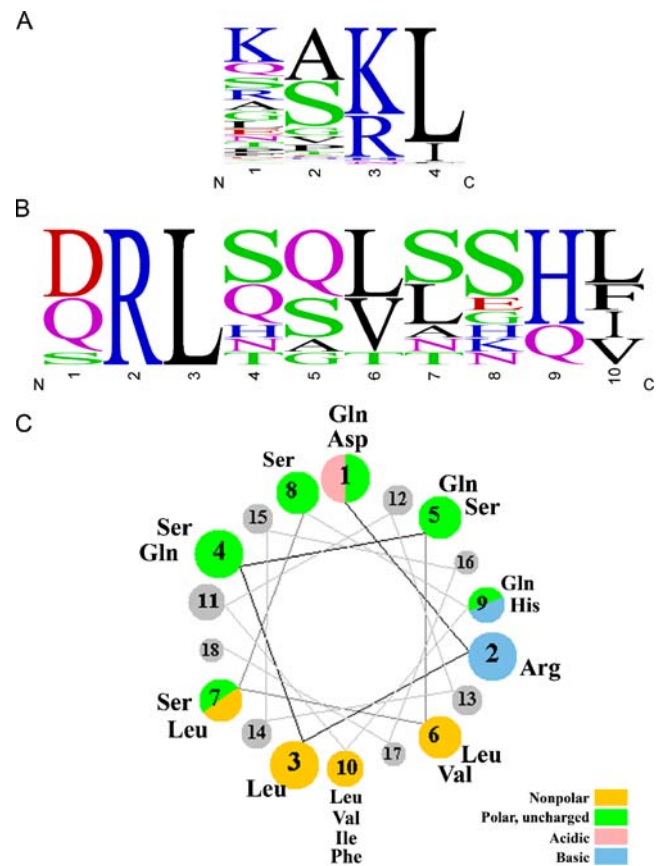
In silico identification of *P. chrysogenum* proteins with putative microbody targeting signals

To facilitate an adequate comparison between *P. chrysogenum* microbody matrix proteins identified by a proteomics approach and those actually encoded by the fungal genome, we first prepared a reference list containing all *P. chrysogenum* proteins with a putative microbody targeting signal (PTS). An initial Genome2D search for such proteins in the *P. chrysogenum* genome (encoding 13,670 proteins) using general consensus sequences (Subramani 1993; Petriv et al. 2004) revealed a set of 150 PTS1 and 25 PTS2 proteins (data

not shown). However, the published PTS consensus sequences are mainly based on mammalian and baker's yeast data, which may not fully apply to filamentous fungi. We observed that the resulting data sets lacked certain known microbody proteins (e.g., PclA). Moreover, they also contained a significant number of small proteins that showed no similarity to proteins in databases, but nevertheless carried a PTS consensus sequence. These findings may significantly limit the use of the data sets as reference lists. Therefore, we decided to follow a more rational approach to identify *P. chrysogenum* PTS proteins. This approach was based on three assumptions namely: (1) *P. chrysogenum* orthologs/paralogs of known human/yeast microbody-borne proteins are candidate PTS proteins; (2) the *P. chrysogenum* PTS consensus sequences resemble, but are not necessarily identical to, those observed in baker's yeast and man; and (3) orthologs of a bona fide *P. chrysogenum* PTS protein also have a PTS in related filamentous ascomycetes (for details see the "Experimental procedures" section).

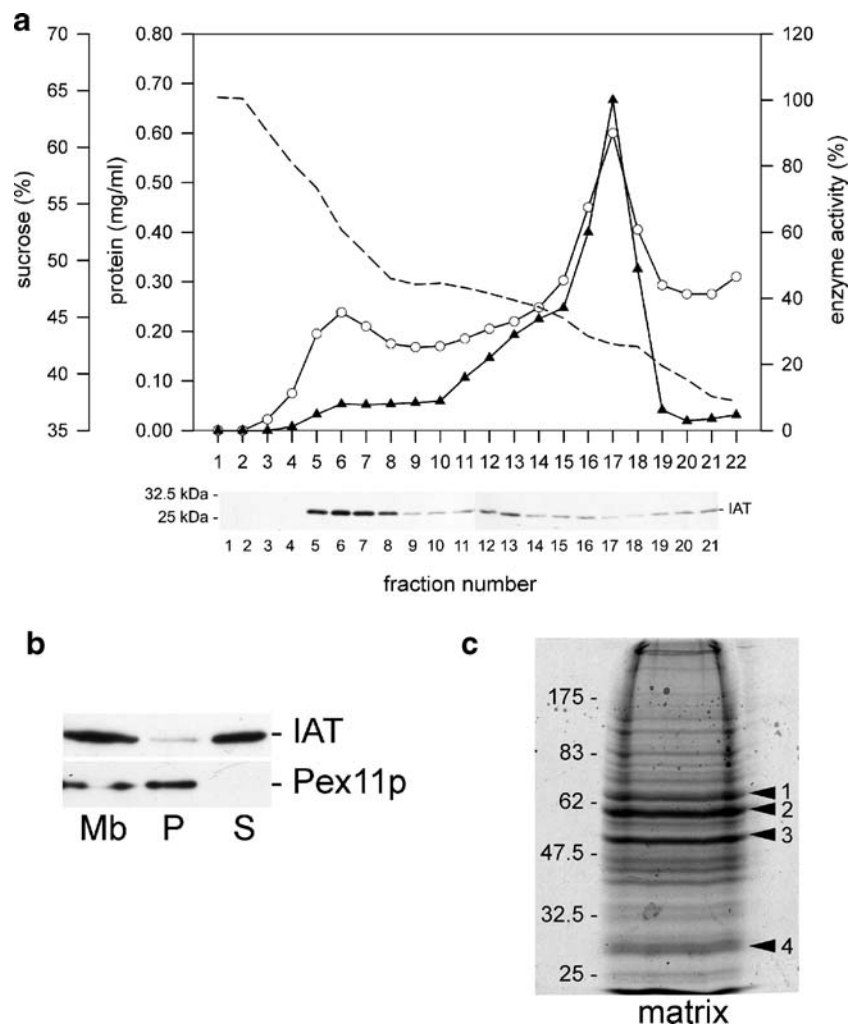
The results of this approach are detailed in Supplementary Table 3. In total, 190 putative PTS1 proteins were identified as well as ten putative PTS2 proteins. Analysis of the sorting signals of these proteins revealed some remarkable characteristics. For PTS1, 45 different signals were identified in *P. chrysogenum*. The most frequent signals were -AKL-COOH (37), -SKL-COOH (35), -ARL-COOH (20), -SRL-COOH (11), -SKI-COOH (nine), and -GKL-COOH (eight) constituting 63% of the total number of PTS1 proteins, with leucine being the preferred C-terminal amino acid (in 83% of the cases). Furthermore, in 25% of the identified signals, a lysine residue directly preceded the tripeptide (cf. Maynard and Berg 2007). A Weblogo graphic of the *P. chrysogenum* PTS1 consensus is presented in Fig. 1a. Combined, the data suggest that *P. chrysogenum* indeed utilizes an organism-specific set of PTS1 signals. For PTS2, almost all signals were found within the first 20 N-terminal amino acids of the proteins. The *P. chrysogenum* PTS2 contains a fully conserved arginine–leucine sequence. In addition, also the amino acid directly preceding the nonapeptide appears to be conserved in *P. chrysogenum* PTS2 proteins (being mostly aspartate or glutamine). A Weblogo graphic of the resulting decapeptide is depicted in Fig. 1b. Secondary structure prediction using the Jpred3 programme indicated that the *P. chrysogenum* PTS2 signal has the capacity to fold into an  $\alpha$ -helix. A helical wheel representation of the decapeptide based on the most abundant residues observed in putative *P. chrysogenum* PTS2 sequences is shown in Fig. 1c and indicates that the PTS2 has the potential to form a three-faced helix with a basic, an aliphatic, and a polar side.

By manual annotation, we have classified the identified PTS proteins according to their possible function (see Supplementary Table 3). This revealed the presence of



**Fig. 1** PTS signals of *P. chrysogenum* proteins. **a** Weblogo graphic showing the distribution and amino acid composition of the *P. chrysogenum* PTS1 consensus sequence. For the graphic, the C-terminal four amino acids of 190 putative *P. chrysogenum* PTS1 proteins were used. For clarity, residue frequencies rather than information contents are indicated. Residues with similar properties are indicated in the same color. Polar amino acids (G, S, T, Y, C, Q, N) are green, basic amino acids (K, R, H) blue, acidic amino acids (D, E) red, and hydrophobic amino acids (A, V, L, I, P, W, F, M) are black. The data indicate that the most frequently used PTS1 sequences in *P. chrysogenum* are -AKL-COOH, -SKL-COOH, -ARL-COOH, and -SRL-COOH. **b** Weblogo graphic showing the distribution and amino acid composition of the *P. chrysogenum* PTS2 consensus sequence. For the graphic, the amino acid sequences of ten putative *P. chrysogenum* PTS2 decapeptides were used. For clarity, residue frequencies rather than information contents are indicated. The color code is identical to that in **a**. **c** Helical wheel representation of the *P. chrysogenum* PTS2 sequence. For creating the helical wheel, the PTS2 consensus sequence -[QD]-R-L-[QS]-[QS]-[VL]-[SL]-S-[QH]-[LVIF]- was used that comprises the most frequently occurring amino acids in *P. chrysogenum* PTS2 sequences. The color code is indicated. The *P. chrysogenum* PTS2 has the capacity to form an  $\alpha$ -helix with three faces that possess different properties, basic (blue), polar (green, red), and aliphatic (yellow). For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article

various paralogs (e.g., multiple D-amino acid oxidases and copper amine oxidases), which might have different substrate specificities or might be solely produced under specific growth conditions. In some cases, genes encoding PTS



**Fig. 2** Isolation of *P. chrysogenum* microbodies. **a** Sucrose density gradient centrifugation of an organellar pellet fraction prepared from *P. chrysogenum* hyphae cultured for 48 h on penicillin production media supplemented with phenoxyacetic acid. Sucrose concentrations (dashed line), protein concentrations (open circles), and the distribution of the activities of the mitochondrial marker protein cytochrome *c* oxidase (closed triangles) are indicated. Enzyme activities are expressed as percentages of the specific activities of the peak fractions, which were arbitrarily set to 100%. The Western blot shows the distribution of the microbody protein IAT in the fractions of the gradients. Equal portions of each fraction were loaded per lane. IAT is present in high-density microbody fractions, well separated from the mitochondria. **b** Separation of microbody matrix enzymes from membrane components. Microbody peak fractions from sucrose gradients (*Mb*) were lysed by treatment with 0.1 M Tris-HCl pH 8.0 and the matrix (*S*) and membrane (*P*) components were separated by

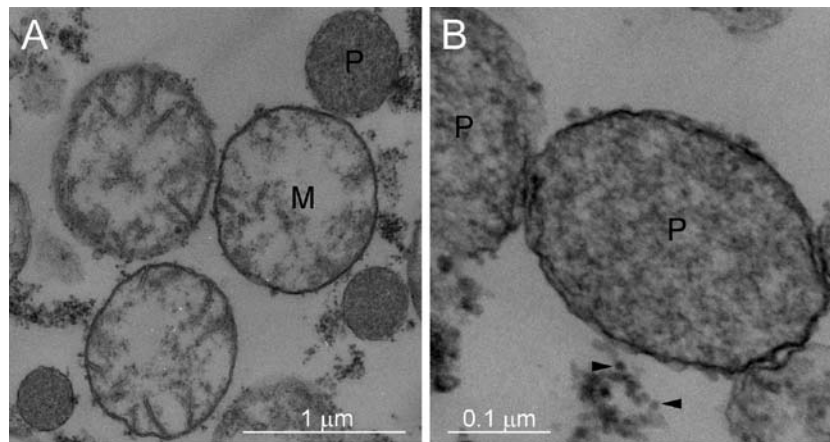
ultracentrifugation, TCA-precipitated, and prepared for Western blotting using equal volume fractions per lane. The blot was detected with specific antibodies against IAT and Pex11p. Almost all IAT protein is present in the microbody matrix fraction, while Pex11p was only found in the membrane fraction. **c** The matrix contents of *P. chrysogenum* microbodies. Microbody matrix proteins were separated by 1-D SDS/PAGE and stained with Bio Safe Coomassie<sup>R</sup> Blue G250. Protein bands were excised and prepared for MS analysis. Four prominent protein bands are indicated with arrowheads. *Band 1* was found to represent a novel fumarate reductase-cytochrome *b5* fusion protein (Pc12g03090), *band 2* was identified as isocitrate lyase (ICL1, Pc22g03660), *band 3* contained a mixture of three proteins with equal Mr, translation elongation factor 1- $\alpha$  (Pc13g02940),  $\alpha$ -tubulin (Pc22g21650), and  $\beta$ -tubulin (Pc13g13470), while *band 4* contained IAT protein (Pc21g21370). Marker bands are indicated in kilodalton to the left

proteins appear to be present in clusters in the *P. chrysogenum* genome (indicated in Supplementary Table 4). Striking examples of this are as follows: the cluster Pc13g14410 (a putative palmitoyl-CoA oxidase)/Pc13g14420 (a putative acyl-CoA synthetase)/Pc13g14430 (a putative enoyl-CoA hydratase/isomerase) and the cluster Pc20g15650 (a putative acetyl-CoA acetyltransferase)/Pc20g15660 (a putative sterol carrier protein-2)/

Pc20g15670 (glutathione-dependent formaldehyde-activating enzyme). The proposed function of the proteins encoded by each cluster suggests that they may belong to the same biochemical pathway.

We also identified a few novel PTS1 proteins that might be involved in secondary metabolism. Next to the expected penicillin biosynthesis enzymes IAT, PclA, and PhIB, we also identified five other PclA-related proteins that are





**Fig. 3** Electron microscopical analysis of purified *P. chrysogenum* microbodies. **a** A 30,000 $\times$ g organellar pellet fraction prepared from *P. chrysogenum* DS17690 hyphae cultured for 48 h on penicillin production media supplemented with phenoxyacetic acid contained

predominantly mitochondria (*M*) and microbodies (*P*). **b** Microbodies (*P*) in a purified microbody peak fraction from a sucrose gradient. The fraction also contained ribosomes (*arrowheads*) that were associated with the organelles

annotated as coumarate-CoA ligases (Pc06g01160, Pc13g12270, Pc21g20650, Pc21g22010, and Pc22g24780). In addition, we found three PTS1 proteins (Pc20g13500, Pc22g13680, and Pc22g24630) that might be involved in the early steps of cephalosporin formation. This is remarkable since *P. chrysogenum* does not synthesize this  $\beta$ -lactam antibiotic. Pc20g13500 and Pc22g13680 are orthologs of *Acremonium chrysogenum* CefD1 and CefD2, respectively (Ullan et al. 2002), while Pc22g24630 represents a small (truncated) protein that also shows weak similarity to CefD2. In the filamentous fungus *A. chrysogenum*, CefD1 and CefD2 are involved in cephalosporin biosynthesis at the stage of isopenicillin N epimerization, a prerequisite for subsequent expansion of the  $\beta$ -lactam ring (for review, see Brakhage 1998).

As expected, several of the identified PTS proteins represent enzymes related to the beta-oxidation of fatty acids, including 14 acyl-CoA synthetases (two with a PTS2), three PTS1-containing acyl-CoA oxidases, and three PTS2-containing 3-keto-acyl thiolases (and one lacking a PTS). Remarkably, we also identified nine PTS1-containing acyl-CoA dehydrogenases, suggesting the presence of an acyl-CoA dehydrogenation step in microbodies similar to that required for beta-oxidation of fatty acids in mammalian mitochondria (cf. Shen and Burger 2009).

We observed that the *P. chrysogenum* ortholog of the aspartate transaminase Aat2, Pc22g19440, has both a putative PTS2 that is conserved in a number of other filamentous fungi, as well as a not conserved PTS1 (-AKL-COOH). Other filamentous fungi and yeast species, including *S. cerevisiae* (Verleur et al. 1997), contain an Aat2 protein with either a PTS1 or a PTS2 (Supplementary Fig. S1). Remarkably, in the PTS2-containing Aat2 orthologs, this signal is located in an extended N terminus and is

immediately followed by a conserved Met residue. This might suggest differential transcription of the Pc22g19440 gene, with the second Met residue representing an alternative startcodon for a possible cytosolic form of the protein, a situation that has been observed for an Aat2 homolog in certain plants (cf. Gebhardt et al. 1998).

Another remarkable finding is that protein Pc21g23730, a putative feruloyl-CoA synthetase, contains a PTS2 in *P. chrysogenum* and in the related *Aspergilli*, but a PTS1 in other filamentous ascomycetes (e.g., *Neurospora crassa*).

Mass spectrometry analysis of *P. chrysogenum* microbody matrix proteins

*P. chrysogenum* microbody fractions were purified by differential centrifugation of lysed protoplasts, followed by sucrose density centrifugation of organellar pellets to separate microbodies from other organelles like mitochondria. The results of this approach are shown in Fig. 2a. Enzyme analysis and Western blotting demonstrated that microbodies, characterized by the presence of IAT (fractions 5 to 8), were well separated from mitochondria, characterized by cytochrome *c* oxidase activity (fraction 15 to 19). This was confirmed by electron microscopical analysis using an organellar pellet (Fig. 3a) as well as the microbody peak fraction of a sucrose gradient (Fig. 3b). Subsequently, we used the microbody peak fractions to separate microbody matrix and membrane components by osmotic lysis (Fig. 2b).

For proteomics characterization, microbody matrix protein fractions were separated on 1-D SDS/PAA gels (Fig. 2c). The gels were cut into small strips and incubated with trypsin, and the eluted peptides were dried. Subsequently, each peptide fraction was analyzed by nanoliquid chromatography tandem mass spectrometry (LC-MS/MS).

**Table 4** *P. chrysogenum* microbody proteins identified by LC-MS/MS analysis

ORF number	Protein	Protein coverage %	Mw (kDa)	PTS1	PTS2
<b>Secondary metabolism</b>					
Pc21g21370	Acyl-coenzyme A/isopenicillin N acyltransferase	58	40	-nARL*	
Pc22g14900	Phenylacetyl-CoA ligase PclA	11	63	-gSKI*	
Pc22g20270	Phenylacetyl-CoA ligase PhlB	13	63	-kAKL*	
Pc22g20380	Putative acetyltransferase SidF	7	51	-nAKL*	
<b>Beta oxidation-related</b>					
Pc12g01170	Putative enoyl-CoA hydratase/isomerase family protein	13	30	-kAKL*	
Pc12g08530	Putative long-chain acyl-CoA dehydrogenase	9	48	-nAKI*	
Pc13g01890	Putative acyl-CoA synthetase	24	63		6/sRLsgllgHF
Pc13g02710	Putative beta-oxidation protein/dehydratase	56	36	-kSKL*	
Pc13g05130	Putative acetyl-CoA synthetase	27	75	-kSKL*	
Pc13g05940	Trifunctional enzyme Fox-2	48	97	-kAKL*	
Pc13g12930	Putative 3-ketoacyl-CoA thiolase	51	44		4/dRLsslslHL
Pc13g15840	Putative enoyl-CoA hydratase/isomerase family protein	22	31	-fEKL*	
Pc15g00410	Putative 3-ketoacyl-CoA thiolase	36	44		6/qRLnsvanQL
Pc20g01800	Putative palmitoyl-CoA oxidase	8	78	-kSKL*	
Pc20g12470	Delta3-cis-delta2-trans-enoil-CoA isomerase Eci1	12	31	-rHKL*	
Pc20g15010	Carnitine <i>O</i> -acetyltransferase	31	73	-kPKL*	
Pc21g04680	Putative 3-ketoacyl-CoA thiolase	10	43		
Pc21g17590	Putative acyl-CoA dehydrogenase	73	49	-kSHL*	
Pc21g19000	Putative acyl-CoA dehydrogenase	7	48	-qSKV*	
Pc21g20710	Putative acyl-CoA dehydrogenase	21	57		
Pc21g21810	Putative levodione reductase/short chain dehydrogenase	68	33		7/qRLqalskQL
Pc22g03680	Putative NADPH-dependent beta-ketoacyl reductase	58	28	-sAKL*	
Pc22g06820	3-ketoacyl-CoA thiolase Pthik	40	43		6/qRLsqvssHF
Pc22g15030	Peroxisomal 2,4-dienoyl-CoA reductase Sps19	57	33	-kSKL*	
Pc22g17070	Putative sterol carrier protein-2	43	49	-IARF*	
Pc22g19950	Long-chain-fatty-acid-CoA synthetase	4	76		
Pc22g20810	Putative enoyl-CoA hydratase	13	32	[-dSKL*]	
Pc22g22390	Putative enoyl-CoA hydratase/isomerase family protein	11	34	-pSKL*	
Pc22g24780	Putative 4-coumarate-CoA ligase	8	61	-kAKL*	
<b>Glyoxylate cycle-related</b>					
Pc12g04750	Malate dehydrogenase	72	34		
Pc12g05400	Citrate synthase CitA	24	52	-gAKL*	
Pc20g03610	NADP-dependent isocitrate dehydrogenase IcdA	29	54	-kARL*	
Pc20g13550	Malate synthase AcuE	30	61	-aAKL*	
Pc22g03660	Isocitrate lyase AcuD	70	60		222/nRLvairaQA
<b>N-metabolism</b>					
Pc12g09740	Putative fructosyl amine oxidase	15	46	-qAKL*	
Pc13g10660	Putative fructosyl amino acid oxidase	29	48		
Pc16g14240	Putative amidase	42	63	-sAKL*	
Pc20g09430	Ureidoglycolate hydrolase Dal3	28	28	-pAKL*	
Pc21g03120	Putative D-amino acid oxidase	28	38		
Pc21g04560	Putative copper amine oxidase	10	77		
Pc21g05590	Putative D-amino acid oxidase	19	40	-rAKL*	
Pc22g00060	Putative fructosyl amine oxidase	5	49	-aAKL*	
Pc22g19440	Aspartate transaminase Aat2	42	48	(-sAKL*)	9/dRLsslslSHIM
Pc22g20960	Urate oxidase Uaz	60	34	-kAKL*	
<b>Antioxidants</b>					
Pc12g13740	Catalase/oxidase KatG	32	82	-rPRL*	
Pc13g11580	Putative epoxide hydrolase	14	29	-kARL*	
Pc16g07060	Putative catalase	12	55	-sSRL*	
Pc22g21240	Catalase CatC	16	57	-pARL*	
Pc22g24640	Putative peroxiredoxin	19	18		

**Table 4** (continued)

ORF number	Protein	Protein coverage %	Mw (kDa)	PTS1	PTS2
Proteases/hydrolases/esterases					
Pc06g01040	Putative acyl-CoA thioesterase	29	40	[-ISKI*]	
Pc12g15100	Putative acyl-CoA thioesterase	20	40	[-aANL*]	
Pc15g00630	Putative thioesterase	22	19		
Pc21g01650	Putative thioesterase	12	21	-eAKL*	
Pc21g04900	Putative thioesterase family protein	14	25	(-dSKL*)	
Other					
Pc12g03090	Fumarate reductase Osm1 with cytochrome B5 domain	54	67		478/qRLgqisHI
Pc12g14280	Putative flavocytochrome b2 protein/FMN-dependent dehydrogenase	27	55	-kPKL*	
Pc13g03610	Putative YcaC related amidohydrolase	13	27	-fSKI*	
Pc13g06370	Protein weakly similar to ethanolamine utilization protein	30	17	-kAKL*	
Pc13g07430	Putative dioxygenase	15	36	-aSKL*	
Pc13g10320	Putative short chain dehydrogenase	33	31		5/dRLtqvneHL
Pc13g12080	Putative protein with alpha/beta hydrolase fold	7	40	-aARL*	
Pc14g00990	Putative GMC-family oxidoreductase	6	61		
Pc14g01270	Putative fumarylacetoacetate hydrolase family protein	35	31	-sSKL*	
Pc16g03120	Conserved hypothetical protein	23	31	[-ISKM*}	
Pc16g07070	Formate oxidase Fod1	26	66	-sSKL*	
Pc16g03770	Protein with cupin domain	25	31	(-ISRL*)	
Pc16g08670	C-8 sterol isomerase Erg-1	12	26	-qLKL*	
Pc16g14230	Putative NADPH-dependent beta-ketoacyl reductase	37	29	-kGRL*	
Pc18g01220	Fructose-bisphosphate aldolase Fba1	46	39	-aGQL*	
Pc18g02760	Putative aldehyde dehydrogenase	10	54	-gTKL*	
Pc18g03000	Aldose 1-epimerase	14	35	-kSRL*	
Pc20g03330	Glucose-6-phosphate 1-dehydrogenase GsdA	7	58	-pNRL*	
Pc20g04300	Putative flavin-containing monooxygenase	47	54	-tARL*	
Pc20g08300	Conserved hypothetical protein	21	39	-eSKI*	
Pc20g09970	Putative short chain dehydrogenase	26	34	[-rAKI*]	
Pc20g13690	Putative CoA-binding protein	18	16	-sQRL*	
Pc21g00970	Woronin body protein Hex1	56	60	-gSRL*	
Pc21g02970	Putative 2-polyprenyl-6-methoxyphenol hydroxylase	7	71	-qSRL*	
Pc21g03400	Triose-phosphate-isomerase TpiA	27	27	-nARI*	
Pc21g04910	Conserved hypothetical protein	19	37	-mEKL*	
Pc21g08790	Putative NADP-dependent oxidoreductase	9	38	-kAKL*	
Pc21g11670	Putative NADPH/quinone reductase/Zn-dependent oxidoreductase	46	36	-kSKL*	
Pc22g10030	Betaine-aldehyde dehydrogenase	20	53	-fAKL*	
Pc22g11470	Putative flavocytochrome b2 protein/FMN-dependent dehydrogenase	13	55	-rARL*	
Pc22g13390	Putative NADPH/quinone reductase/Zn-dependent oxidoreductase	59	39	-gSKL*	
Pc22g14270	Putative 2-nitropropane dioxygenase family oxidoreductase	17	36	-sSKL*	
Pc22g19490	Putative 2-nitropropane dioxygenase family oxidoreductase	44	38	-vSKL*	
Pc22g22110	Glycerol 3-phosphate dehydrogenase GfdA	16	48	(-aARL*)	
Pc22g25570	Putative aminoglycoside phosphotransferase	14	40	-eGKL*	

The asterisk represents the carboxy terminal COOH group of the proteins. Hyphens indicate that the PTS is not conserved in orthologs in other filamentous ascomycetes. Square brackets indicate that the protein sequence in the *P. chrysogenum* database was incorrect and needed correction

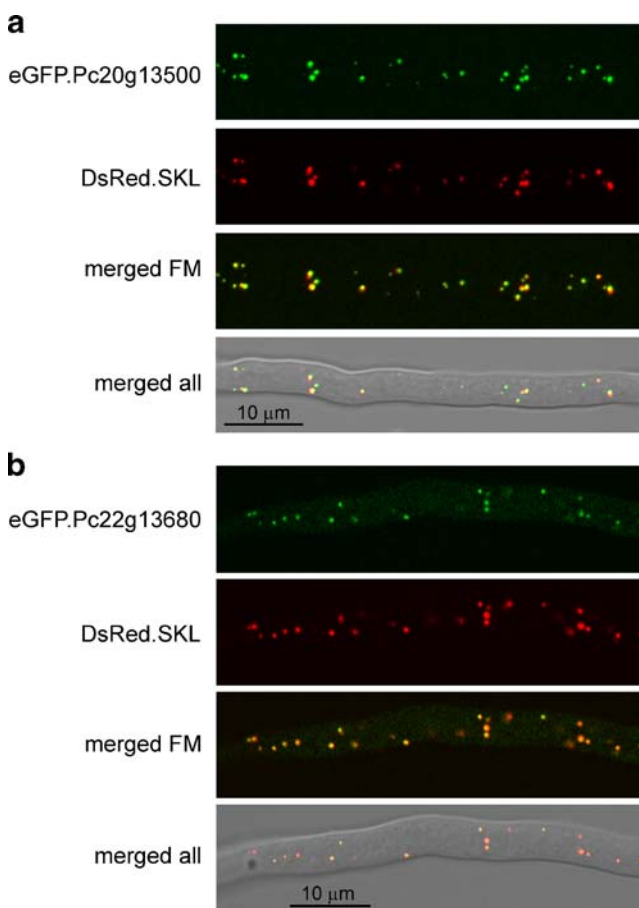
The obtained 35,000 MS/MS spectra were searched against the *P. chrysogenum* protein database (13,670 sequences) using stringent matching criteria. This resulted in the confident identification of 546 proteins (Supplementary Table 5). Following manual annotation, the identified proteins were divided in different subclasses, based on

their putative subcellular localization. A large group of proteins (79) showed a PTS (Table 4). Indeed, among the 50 best scoring proteins, 30 were putative microbody matrix proteins. These probably represent the most prominent proteins in the sample, which is reflected by a high protein coverage score. This shows that the analyzed

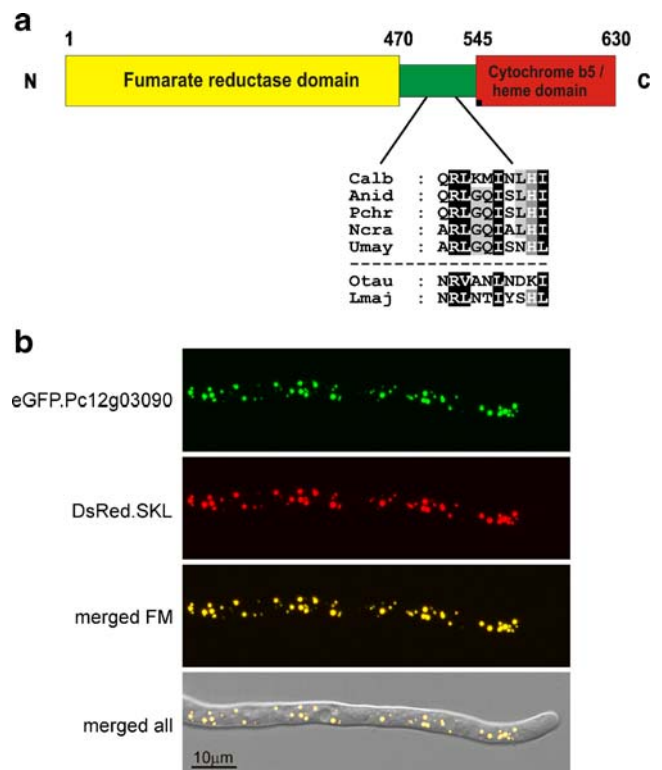
cellular fraction was indeed enriched in microbodies. Furthermore, a consistent fraction of the identified proteome (119 proteins) was represented by proteins involved in translation, mainly ribosomal proteins. Although ribosomal proteins could be considered as contaminants of the microbody fraction, their presence can also be explained by the observation that ribosomes were sometimes seen adjacent to purified microbodies in electron microscopy images (Fig. 4b). Additionally, proteins originating from other subcellular compartments were also identified (Supplementary Table 5), albeit with much lower scores.

Three out of four prominent protein bands visible by Coomassie staining on 1-D PAA gels (arrows in Fig. 2c) were identified as Pc12g03090 (band 1), Pc22g03660

(isocitrate lyase, ICL1; band 2), and Pc21g21370 (IAT; band 4). Remarkably, among these proteins, only the well-studied penicillin biosynthesis protein IAT possesses an obvious PTS. Pc12g03090 is a novel fumarate reductase-cytochrome b5 fusion protein conserved in filamentous fungi and a few yeast species (Supplementary Fig. S2 and Fig. 5a). This protein does not have a PTS at either its N or C termini, but contains a conserved sequence in a loop between the fumarate reductase and cytochrome b5 domains that fits well with the *P. chrysogenum* PTS2 consensus (478/qRLgqislHI, Fig. 5a). Thus, Pc12g03090 may have an internal sorting signal. ICL1 is a key component of the glyoxylate cycle and was demonstrated



**Fig. 4** *P. chrysogenum* orthologs of CefD1 (Pc20g13500) and CefD2 (Pc22g13680) localize to microbodies. *P. chrysogenum* DsRed.SK1 eGFP.Pc20g13500 (a) and DsRed.SK1 eGFP.Pc22g13680 (b) cells were cultivated on penicillin production medium supplemented with phenoxyacetic acid for 30 h and analyzed by confocal laser scanning fluorescence microscopy. In both cases, GFP and DsRed fluorescence co-localized, indicating that both fusion proteins are present in microbodies. DsRed.SK1 eGFP.Pc22g13680 cells also showed a faint cytosolic eGFP fluorescence, indicating that a portion of the eGFP.Pc22g13680 fusion protein is located in the cytosol. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article



**Fig. 5** Pc12g03090 is a novel fumarate reductase-cytochrome b5 fusion protein with a putative internal PTS2 sequence that locates to microbodies in *P. chrysogenum*. **a** Schematic representation of the structure of *P. chrysogenum* Pc12g03090. This protein consists of a fumarate reductase domain, a cytochrome b5 heme binding domain, as well as a linker region containing a putative PTS2. This internal PTS2 is conserved in many organisms, mainly filamentous fungi. An alignment of representative PTS2 sequences is also shown. These sequences were taken from the alignment in Supplementary Fig. S2. Conserved residues are shaded. **b** Localization of Pc12g03090 in *P. chrysogenum*. *P. chrysogenum* DsRed.SK1 eGFP.Pc12g03090 cells were cultivated on penicillin production medium supplemented with phenoxyacetic acid for 2 days and analyzed by confocal laser scanning fluorescence microscopy. The GFP fluorescence completely co-localized with the DsRed fluorescence, indicating that the fusion protein is present in microbodies. For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article

to be a microbody protein in *A. nidulans* (Valenciano et al. 1998). Recently, it was shown that import of *A. nidulans* ICL1 in microbodies depends on the PTS2 receptor Pex7p (Hynes et al. 2008). In all filamentous ascomycetes, including *P. chrysogenum*, ICL1 lacks a recognizable PTS, while in some yeast species and certain basidiomycetes, the protein contains a PTS1 (for alignment, see Supplementary Fig. S3). A screen for a possible internal PTS2 sequence identified the related sequence 222/nRLvairaQA, which deviates significantly from the identified PTS2 consensus (Fig. 1b) and is located in a highly conserved region of ICL1, making its function as a targeting signal rather doubtful.

Next to prominent proteins like ICL1, we identified other glyoxylate cycle enzymes as well as multiple proteins presumed to function in  $\beta$ -oxidation, including all three PTS2 containing 3-keto-acyl thiolases as well as the PTS-lacking paralog Pc21g04680 (Table 4). Remarkably, only a single acyl-CoA oxidase (Pc20g01800) was detected, while four acyl-CoA dehydrogenases were identified, again suggesting that fatty acids might also be metabolized via an acyl-CoA dehydrogenation step in *P. chrysogenum* microbodies.

Among the proteins involved in penicillin biosynthesis, the known microbody-borne enzymes IAT and PclA were both identified. In addition, we found the PclA-related proteins PhlB and Pc22g24780, that is annotated as a 4-coumarate-CoA ligase. Additionally, the enzymes involved in the initial stages of penicillin biosynthesis,  $\delta$ (L- $\alpha$ -aminoadipyl)-L-cysteinyl-D-valine synthetase and isopenicillin N synthase, both present in high concentrations in the cytosol (van der Lende et al. 2002), were also detected.

#### Matching the proteome to the genome

Comparison of the obtained proteomics catalogue with the reference list of the 200 in silico identified PTS proteins revealed that 69 out of 79 PTS proteins identified experimentally were in the reference list. Apparently, only a subset of the available PTS proteins locates to microbodies at penicillin inducing conditions. Five proteins (Pc06g01040, Pc12g15100, Pc16g03120, Pc20g09970, Pc22g20810) lacking a PTS in the protein database were annotated as putative peroxisomal proteins on the basis of their PTS1-containing orthologs in other fungi. Closer analysis revealed that improper intron splicing has caused that the coding sequences of these *P. chrysogenum* proteins were incorrect. After correction, all five proteins were shown to contain a PTS1.

Improper intron splicing may also explain the case of Pc22g22110, encoding glycerol 3-phosphate dehydrogenase GfpA. At first glance, GfpA only has a PTS1 in *P. chrysogenum*. However, closer analysis of *gfpA* gene

sequences in Aspergilli revealed that incorrect intron splicing may have precluded identification of the PTS1 in these organisms.

The other four PTS proteins not present on the reference list were ICL1 and Pc12g03090 that have an internal PTS2 (see above), as well as Pc16g03770, a conserved protein with a Cupin domain, and Pc21g04900, a putative thioesterase. Although in *P. chrysogenum*, Pc16g03770 and Pc21g04900 contain a PTS1, orthologs in other organisms lack a PTS, explaining their absence in the reference list.

#### Subcellular localization of newly identified PTS proteins in *P. chrysogenum*

To confirm the strength of our in silico analysis, we determined the subcellular location of two identified PTS1 proteins in *P. chrysogenum* using GFP-fusion proteins. We fused GFP to the N terminus of both Pc20g13500 and Pc22g13680, the orthologs of *A. chrysogenum* CefD1 and CefD2, respectively, and produced the fusion proteins in a *P. chrysogenum* strain also producing DsRed.SKL. After cultivation of selected transformants on penicillin production media supplemented with phenoxyacetic acid to induce the expression of the fusion genes, hyphae were visualized by confocal laser scanning microscopy (CLSM). In both cases, we observed co-localization of the DsRed and GFP fluorescence, implying that the GFP-fusion proteins are indeed targeted to microbodies (Fig. 4).

Similarly, to confirm the subcellular location of the novel fumarate reductase-cytochrome b5 fusion protein Pc12g03090 identified by mass spectrometry, we expressed a *GFP.Pc12g03090* fusion gene in *P. chrysogenum* DsRed.SKL. CLSM analysis revealed a complete co-localization of the DsRed and GFP fluorescence, implying that Pc12g03090 is indeed a bona fide microbody protein (Fig. 5b).

#### Discussion

We aimed to identify the matrix enzymes present in microbodies of the filamentous fungus *P. chrysogenum* during penicillin-producing conditions and match them with a reference list containing microbody enzymes predicted from genome analysis. By a novel in silico approach, we found 200 *P. chrysogenum* microbody proteins containing a PTS. Additionally, a proteomics study on the contents of purified *P. chrysogenum* microbodies by nano-LC-MS/MS analysis has identified 79 PTS proteins and ten other putative microbody matrix proteins. We found that 69 out of 79 PTS proteins identified experimentally were in the reference list. We have corroborated our analyses by demonstrating the microbody location for three novel microbody proteins using *GFP* gene fusions, namely orthologs of *A. chrysogenum*

CefD1 (Pc20g13500) and CefD2 (Pc22g13680) as well as a novel fumarate reductase-cytochrome b5 fusion protein (Pc12g03090). Our *in silico* and proteomics analyses clearly indicate that microbodies in *P. chrysogenum* are not specialized structures solely devoted to penicillin biosynthesis, but are in fact multi-purpose organelles. Moreover, the identification of hydrogen-peroxide-producing oxidases and hydrogen-peroxide-decomposing catalases in *P. chrysogenum* microbodies implies that these organelles should actually be considered peroxisomes. This contrasts to the situation observed in *N. crassa*, where microbodies lack catalase (Schliebs et al. 2006).

#### *In silico* identification of *P. chrysogenum* microbody proteins

Our *in silico* approach to identify PTS proteins in the *P. chrysogenum* genome was based on the assumptions of (1) conservation of enzyme function, (2) PTS sequence conservation, and (3) conservation of subcellular localization. The first assumption implied that *P. chrysogenum* orthologs of known microbody proteins are candidate PTS proteins. This was based on the notion that certain major microbody-borne biochemical pathways are conserved in various organisms, e.g., the  $\beta$ -oxidation pathway of fatty acids (Poirier et al. 2006). Indeed, this has provided the basis for our reference set of proteins. Of the 125 candidate PTS1 proteins identified this way, 98 (78%) were considered *bona fide* PTS1 proteins. Similarly, seven out of ten PTS2 proteins in the final reference set were present in the initial set of candidate PTS2 proteins. Our second assumption included that the *P. chrysogenum* PTS sequences resembled, but were not necessarily identical to, those observed in baker's yeast and mammals. This was based on the observation that a PTS signal from a heterologous protein can target a marker protein into microbodies in baker's yeast as well as in mammalian cells (see Gould et al. 1990; Purdue and Lazarow 1996). Moreover, the fusion of -SKL-COOH to the C terminus of the normally cytosolic GFP protein results in its sorting to microbodies in all eukaryotes including *P. chrysogenum* (see Kiel et al. 2004).

The PTS1 signal is recognized by the highly conserved tetratricopeptide repeat domains of the PTS1 receptor protein Pex5p (Gatto et al. 2000). Recently, we demonstrated that *P. chrysogenum* Pex5p could efficiently sort catalase (-SKI-COOH), dihydroxyacetone synthase (-NKL-COOH), and eGFP.SK1 into microbodies in the yeast *Hansenula polymorpha* (Kiel et al. 2004). Thus, the *P. chrysogenum* PTS1 receptor can recognize highly diverse signals, which is consistent with the identification of 45 different putative PTS1 signals. Our final assumption was that a *P. chrysogenum* protein can be considered a *bona fide* PTS protein when an ortholog in (at least) two related filamentous ascomycetes

also carries a PTS. This criterion is based on the high sequence similarity between enzymes in filamentous ascomycetes. This ensures that many small orphan proteins in the *P. chrysogenum* protein database that lack homologs in other species will not contaminate the reference set. It must be noted that protein databases of filamentous fungi may contain many incorrect entries due to improper intron splicing, which often affects correct determination of the N and C termini of proteins. As a result, a PTS may not be conserved in all species hence the non-stringent requirement for identification of only two PTS-containing orthologs. Based on our proteomics inventory, we propose an alternative splicing for at least five *P. chrysogenum* proteins that did not appear to carry a PTS1 sequence. In fact, upon manual annotation, a PTS1 sequence could be identified, thus improving protein sequence annotation. However, we expect that this phenomenon will only add a limited number of additional proteins to the list and will not affect the identified PTS1 consensus significantly.

Next to the PTS proteins involved in known microbody functions (e.g.,  $\beta$ -oxidation, glyoxylate cycle), our reference set also contained some unexpected PTS1 proteins. These included orthologs of the *A. chrysogenum* CefD1 and CefD2 proteins involved in the biosynthesis of cephalosporin, a  $\beta$ -lactam derivative that is not produced by *P. chrysogenum*, but which utilizes the same precursor as penicillin, isopenicillin N (IPN). CefD1 and CefD2 catalyze the first step that commits the pathway in *A. chrysogenum* to the production of cephalosporins, the isomerization of the l- $\alpha$ -amino adipate side chain of IPN to the D-enantiomer to give penicillin N. The subsequent activities of the CefEF and CefG proteins will result in ring expansion of the  $\beta$ -lactam ring and production of cephalosporin C (Brakhage 1998). CefD1 is highly similar to very-long-chain-fatty-acid-CoA synthetases, while CefD2 is highly similar to human racemases (Ullan et al. 2002). Members of both families have been found in mammalian microbodies (Wanders 2004). The identification of *Pc*-CefD1 and *Pc*-CefD2 and the finding that both are located in microbodies (Fig. 4) implies that in *P. chrysogenum*, conversion of IPN into penicillin N can occur in the protective environment of the microbody. However, *P. chrysogenum* does not encode an ortholog of CefEF, precluding formation of cephalosporin C. Of note is that also in *A. chrysogenum*, the CefD1 and CefD2 proteins contain putative PTS1 signals, implying that microbodies may actually be essential for cephalosporin formation in this organism.

Other unexpected proteins identified *in silico* included aldose 1-epimerase (Pc18g03000, -SRL-COOH), fructose-bisphosphate aldolase (Pc18g01220; -GQL-COOH), glucose-6-phosphate 1-dehydrogenase (Pc20g03330; -NRL-COOH), triose-phosphate-isomerase (Pc21g03400; -ARI-COOH), and a pyruvate carboxylase like protein

(Pc22g15180; -GKL-COOH; see Supplementary Table 3). We cannot exclude that a number of these highly produced, putative cytosolic enzymes may represent false positives in our dataset due to the extremely high sequence similarity between orthologs, including their C termini. Nevertheless, we identified peptides of many of these dominant proteins in the microbody matrix fraction (see Supplementary Table 5). This may suggest that these proteins have multiple localizations in the cell, a feature that was observed before in *S. cerevisiae*. When grown on oleate, the *Sc-Aat2* protein (Pc22g19440 in *P. chrysogenum*) is in microbodies, while it is located in the cytosol during growth on glucose (Verleur et al. 1997). In *A. nidulans*, differential transcription of the gene encoding the NADP-dependent isocitrate dehydrogenase (Pc20g03610 in *P. chrysogenum*; -AKL-COOH) results in the synthesis of either a long mitochondrial/microbody form or a short cytosolic/microbody form of the enzyme dependent on the culture conditions (Szewczyk et al. 2001). Additional studies will have to address whether the indicated proteins have multiple locations in *P. chrysogenum*.

Was our *in silico* analysis exhaustive? Obviously, our approach will not resolve proteins lacking a PTS, and also, proteins with an internal PTS2 will be difficult to identify. During our analysis, we have concentrated on the first 50 N-terminal amino acids of proteins to look for putative PTS2 sequences. In most cases, this is sufficient because the extreme N and C termini of a protein are most exposed and thus available for binding a receptor moiety. However, in exceptional cases, an internal domain may also be available for binding. We suspect that in Pc12g03090, a fumarate dehydrogenase-cytochrome b5 fusion protein, the highly conserved PTS2-related sequence present in the region between the two domains represents its targeting information. We have confirmed the microbody location of Pc12g03090 using a GFP-fusion protein (Fig. 5). Nevertheless, it is unclear whether this sorting indeed proceeds via the PTS2 pathway. Regarding other possible false negatives, the outcome of our proteomics study has largely confirmed our *in silico* approach; out of 79 proteins with a PTS at either of the N or C termini, only two putative PTS1 proteins were not in our reference list (Pc16g03770, Pc21g04900) because their orthologs had no PTS. Further analysis will have to confirm that Pc16g03770 and Pc21g04900 are indeed bona fide microbody matrix proteins in *P. chrysogenum*.

#### Proteomics of *P. chrysogenum* microbodies

Our proteomics analysis of purified microbody matrix fractions resulted in a list of 546 proteins of which approximately 16% were putative microbody matrix proteins, despite the fact that microbody fractions from sucrose gradients were highly enriched for IAT and almost devoid of

mitochondria. Nevertheless, prominent matrix components from other organelles (mitochondria, ER, vacuole) were also identified. The major contamination in our microbody preparation is represented by the 119 proteins involved in translation, mostly ribosomal proteins. This number does not include the many potential nuclear-associated proteins involved in RNA splicing and ribosome assembly, which in fact might also be attached to ribosomes. Proteins involved in translation (e.g., translation elongation factor 1- $\alpha$ ) have been observed before in microbody fractions (Schäfer et al. 2001; Kiel et al. 2007). Electron micrographs suggest that ribosomes are sometimes attached to microbodies in purified *P. chrysogenum* microbody fractions (Fig. 4b), which may in fact reflect localized translation of mRNAs encoding microbody proteins analogous to what has been observed for mitochondria (for review, see Gonsalvez et al. 2005).

The dataset of 89 putative microbody proteins identified by mass spectrometry (79 with a PTS and ten without; Table 4) contained multiple components involved in  $\beta$ -oxidation, despite the fact that no fatty acids were present in the cultivation medium.  $\beta$ -oxidation in yeast species such as *S. cerevisiae* occurs solely in microbodies and is primarily devoted to the degradation of extracellularly fed fatty acids generating acetyl-CoA that is used as carbon and energy source for growth (Poirier et al. 2006). Indeed, in baker's yeast, the genes encoding  $\beta$ -oxidation enzymes become highly induced upon growth on oleate (Kal et al. 1999). In higher eukaryotes,  $\beta$ -oxidation is required for the metabolism of fatty acids as well, but is also essential for more specialized functions, e.g., auxin conversion in plants (Zolman et al. 2007; reviewed by Poirier et al. 2006). A remarkable observation is that filamentous fungi like *P. chrysogenum* harbor nine PTS-containing putative acyl-CoA dehydrogenases, a number of which were also identified in the proteomics profiling, as well as isozymes with a mitochondrial signal sequence. Recent data suggest that, like mammals, filamentous fungi are capable of  $\beta$ -oxidation in both microbodies and mitochondria (Maggio-Hall and Keller 2004; Maggio-Hall et al. 2008). Although this would explain the presence of multiple mitochondrial type acyl-CoA dehydrogenases, it does not give clues as to why *P. chrysogenum* would require nine microbody-borne isoenzymes in addition to PTS1-containing acyl-CoA oxidases (cf. Shen and Burger 2009). In theory,  $\beta$ -oxidation via an acyl-CoA dehydrogenase is more advantageous to the cell because the reducing equivalents generated in the dehydrogenation step can be recovered (e.g., in NADH), while they are lost in the form of hydrogen peroxide in case of acyl-CoA oxidase function. Recently, it was demonstrated that in *Arabidopsis thaliana*, auxin conversion requires a  $\beta$ -oxidation-related reaction that involves an acyl-CoA dehydrogenase with a microbody targeting signal (Zolman et al. 2007). In addition to this, it must be noted that *N.*

*crassa* microbodies may not contain catalase to decompose hydrogen peroxide (Schliebs et al. 2006), suggesting that  $\beta$ -oxidation in this organism apparently does not produce high levels of microbody-borne hydrogen peroxide.

Related to this may be the identification of the novel fumarate reductase-cytochrome b5 fusion protein (Pc12g03090) as a prominent component of the microbody matrix in *P. chrysogenum* (Figs. 3c and 5). Generally, fumarate reductases convert fumarate into succinate with concurrent oxidation of NADH to NAD<sup>+</sup>. Thus, the presence of this enzyme in *P. chrysogenum* microbodies may point to a novel way to balance the amount of reducing equivalents in the microbody. In this respect, the cytochrome b5 module of the protein may actually assist by recruiting the reducing equivalents from either free NADH or directly from the FAD molecules that function as co-factors of the acyl CoA dehydrogenases. The PTS2 and the cytochrome b5 domain of Pc12g03090 are mainly conserved in filamentous fungi and not in yeast species like *S. cerevisiae* (cf. Fig. 5a). Interestingly, in baker's yeast peroxisomal NAD<sup>+</sup> is regenerated by a malate shuttle via the conversion of oxaloacetate into malate (van Roermund et al. 1995).

ICL1 is another prominent protein of *P. chrysogenum* microbody matrix fractions. This protein is a bona fide microbody protein in *A. nidulans* (Valenciano et al. 1998), but is cytosolic in baker's yeast (reviewed by Kunze et al. 2006). ICL1 is involved in the glyoxylate cycle and as such fuels the tricarboxylic acid cycle by converting isocitrate into glyoxylate and succinate. Thus, two prominent microbody enzymes of *P. chrysogenum* (ICL1 and Pc12g03090) may produce significant amounts of succinate inside the organelle. This implies the presence of a succinate export pathway from the microbody to the cytosol, possibly followed by import of succinate into mitochondria via the mitochondrial fumarate/succinate transporter (Palmieri et al. 1997). In the mitochondria, succinate may be re-converted into fumarate and transported back via the cytosol to the microbody. Alternatively, succinate may become oxidized further, ultimately yielding CO<sub>2</sub> and water, while the generated reducing equivalents are converted into ATP. This scenario would allow reducing equivalents formed in the microbody to be utilized for energy generation in mitochondria. However, the required fumarate/succinate shuttle in the *P. chrysogenum* microbody membrane has not been identified.

The manuscript describing the *P. chrysogenum* genome sequence (van den Berg et al. 2008) also included an analysis of the transcriptome of *P. chrysogenum* DS17690 cells cultured in a glucose-limited chemostat supplemented with phenylacetic acid, conditions that are thought to mimic penicillin-producing conditions. We have compared the results of our proteomics study with the absolute transcript

levels of the *P. chrysogenum* genes encoding the in silico identified PTS proteins (Supplementary Table 6). The results indicate that our mass spectrometry analysis has identified a significant number of those proteins from which the gene was significantly/highly expressed during chemostat cultivation. Thus, from the 50 *P. chrysogenum* genes in this list with the highest relative expression, we identified 37 gene products. Conversely, from the 50 genes with the lowest hybridization scores, i.e., which were virtually not expressed under the conditions used, only three gene products were identified.

In conclusion, we have identified *P. chrysogenum* microbody matrix enzymes by in silico and proteomics approaches and confirmed the location of three enzymes by fluorescence microscopy. This study forms a solid basis to better understand which microbody proteins are of importance to *P. chrysogenum* during penicillin production.

**Acknowledgments** J.A.K.W.K. was financially supported by the Netherlands Ministry of Economic Affairs and the B-Basic partner organizations ([www.b-basic.nl](http://www.b-basic.nl)) through B-Basic, a public-private NWO-ACTS programme (ACTS = Advanced Chemical Technologies for Sustainability). F.F. and B.P. were supported by the Netherlands Proteomics Centre (<http://www.netherlandsproteomicscentre.nl/>). We thank Wim Huibers, Rolf Kanninga, Wiebe H. Meijer, and Jennifer Srijbosch for expert technical assistance.

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Altschul SF, Gish W, Miller W et al (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
- Altschul SF, Madden TL, Schäffer AA et al (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
- Baerends RJ, Smits WK, de Jong A et al (2004) Genome2D: a visualization tool for the rapid analysis of bacterial transcriptome data. *Genome Biol* 5:R37
- Barredo JL, Van Solingen P, Diez B et al (1989) Cloning and characterization of the acyl-coenzyme A: 6-aminopenicillanic acid-acyltransferase gene of *Penicillium chrysogenum*. *Gene* 83:291–300
- Brakhage AA (1998) Molecular regulation of beta-lactam biosynthesis in filamentous fungi. *Microbiol Mol Biol Rev* 62:547–585
- Bruggink A, Roy PD (2001) Industrial synthesis of semisynthetic antibiotics. In: Bruggink A (ed) *Synthesis of B-lactam antibiotics, chemistry, biocatalysis and process integration*. Kluwer, Dordrecht, pp 12–55
- Cantoral JM, Diez B, Barredo JL et al (1987) High frequency of transformation of *Penicillium chrysogenum*. *Bio/Technol* 5:494–497



- Crooks GE, Hon G, Chandonia JM et al (2004) WebLogo: a sequence logo generator. *Genome Res* 14:1188–1190
- Douma AC, Veenhuis M, de Koning W et al (1985) Dihydroxy-acetone synthase is localized in the peroxisomal matrix of methanol grown *Hansenula polymorpha*. *Arch Microbiol* 143:237–243
- Gatto GJ Jr, Geisbrecht BV, Gould SJ et al (2000) Peroxisomal targeting signal-1 recognition by the TPR domains of human PEX5. *Nat Struct Biol* 7:1091–1095
- Gebhardt JS, Wadsworth GJ, Matthews BF (1998) Characterization of a single soybean cDNA encoding cytosolic and glyoxysomal isozymes of aspartate aminotransferase. *Plant Mol Biol* 37:99–108
- Gidijala L, van der Klei IJ, Veenhuis M et al (2007) Reprogramming *Hansenula polymorpha* for penicillin production: expression of the *Penicillium chrysogenum* *pcl* gene. *FEMS Yeast Res* 7:1160–1167
- Gonsalvez GB, Urbinati CR, Long RM (2005) RNA localization in yeast: moving towards a mechanism. *Biol Cell* 97:75–86
- Gould SJ, Valle D (2000) Peroxisome biogenesis disorders: genetics and cell biology. *Trends Genet* 16:340–345
- Gould SJ, Keller GA, Schneider M et al (1990) Peroxisomal protein import is conserved between yeast, plants, insects and mammals. *EMBO J* 9:85–90
- Guerra-Giraldez C, Quijada L, Clayton CE (2002) Compartmentation of enzymes in a microbody, the glycosome, is essential in *Trypanosoma brucei*. *J Cell Sci* 115:2651–2658
- Harris DM, Diderich JA, van der Krogt ZA et al (2006) Enzymic analysis of NADPH metabolism in beta-lactam-producing *Penicillium chrysogenum*: presence of a mitochondrial NADPH dehydrogenase. *Metab Eng* 8:91–101
- Hillenga DJ, Versantvoort HJ, Driessen AJ et al (1994) Structural and functional properties of plasma membranes from the filamentous fungus *Penicillium chrysogenum*. *Eur J Biochem* 224:581–587
- Hynes MJ, Murray SL, Khew GS et al (2008) Genetic analysis of the role of peroxisomes in the utilization of acetate and fatty acids in *Aspergillus nidulans*. *Genetics* 178:1355–1369
- Kal AJ, van Zonneveld AJ, Benes V et al (1999) Dynamics of gene expression revealed by comparison of serial analysis of gene expression transcript profiles from yeast grown on two different carbon sources. *Mol Biol Cell* 10:1859–1872
- Keller A, Nesvizhskii AI, Kolker E et al (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* 74:5383–5392
- Kiel JA, Hilbrands RE, Bovenberg RA et al (2000) Isolation of *Penicillium chrysogenum* *PEX1* and *PEX6* encoding AAA proteins involved in peroxisome biogenesis. *Appl Microbiol Biotechnol* 54:238–242
- Kiel JA, van den Berg MA, Bovenberg RA et al (2004) *Penicillium chrysogenum* Pex5p mediates differential sorting of PTS1 proteins to microbodies of the methylotrophic yeast *Hansenula polymorpha*. *Fungal Genet Biol* 41:708–720
- Kiel JA, van der Klei IJ, van den Berg MA et al (2005) Overproduction of a single protein, Pc-Pex11p, results in 2-fold enhanced penicillin production by *Penicillium chrysogenum*. *Fungal Genet Biol* 42:154–164
- Kiel JA, Veenhuis M, van der Klei IJ (2006) *PEX* genes in fungal genomes: common, rare or redundant. *Traffic* 7:1291–1303
- Kiel JA, Titorenko VI, van der Klei IJ et al (2007) Overproduction of translation elongation factor 1-alpha (eEF1A) suppresses the peroxisome biogenesis defect in a *Hansenula polymorpha* *pex3* mutant via translational read-through. *FEMS Yeast Res* 7:1114–1125
- Kunze M, Pracharoenwattana I, Smith SM et al (2006) A central role for the peroxisomal membrane in glyoxylate cycle function. *Biochim Biophys Acta* 1763:1441–1452
- Lamas-Maceiras M, Vaca I, Rodríguez E et al (2006) Amplification and disruption of the phenylacetyl-CoA ligase gene of *Penicillium chrysogenum* encoding an aryl-capping enzyme that supplies phenylacetic acid to the isopenicillin N-acyltransferase. *Biochem J* 395:147–155
- Lazarow PB, Fujiki Y (1985) Biogenesis of peroxisomes. *Annu Rev Cell Biol* 1:489–530
- Maggio-Hall LA, Keller NP (2004) Mitochondrial beta-oxidation in *Aspergillus nidulans*. *Mol Microbiol* 54:1173–1185
- Maggio-Hall LA, Lyne P, Wolff JA et al (2008) A single acyl-CoA dehydrogenase is required for catabolism of isoleucine, valine and short-chain fatty acids in *Aspergillus nidulans*. *Fungal Genet Biol* 45:180–189
- Maynard EL, Berg JM (2007) Quantitative analysis of peroxisomal targeting signal type-1 binding to wild-type and pathogenic mutants of Pex5p supports an affinity threshold for peroxisomal protein targeting. *J Mol Biol* 368:1259–1266
- Meijer WH, van der Klei IJ, Veenhuis M et al (2007) *ATG* genes involved in non-selective autophagy are conserved from yeast to man, but the selective Cvt and pexophagy pathways also require organism-specific genes. *Autophagy* 3:106–116
- Monastyrska I, van der Heide M, Krikken AM et al (2005) Atg8 is essential for macropexophagy in *Hansenula polymorpha*. *Traffic* 6:66–74
- Müller WH, Van der Krift TP, Krouwer AJJ et al (1991) Localization of the pathway of the penicillin biosynthesis in *Penicillium chrysogenum*. *EMBO J* 10:489–495
- Müller WH, Bovenberg RA, Groothuis MH et al (1992) Involvement of microbodies in penicillin biosynthesis. *Biochim Biophys Acta* 1116:210–213
- Nesvizhskii AI, Keller A, Kolker E et al (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* 75:4646–4658
- Palmieri L, Lasorsa FM, De Palma A et al (1997) Identification of the yeast *ACR1* gene product as a succinate-fumarate transporter essential for growth on ethanol or acetate. *FEBS Lett* 417:114–118
- Petriv OI, Tang L, Titorenko VI et al (2004) A new definition for the consensus sequence of the peroxisome targeting signal type 2. *J Mol Biol* 341:119–134
- Poirier Y, Antonenkov VD, Glumoff T et al (2006) Peroxisomal beta-oxidation—a metabolic pathway with multiple functions. *Biochim Biophys Acta* 1763:1413–1426
- Purdue PE, Lazarow PB (1996) Targeting of human catalase to peroxisomes is dependent upon a novel COOH-terminal peroxisomal targeting sequence. *J Cell Biol* 134:849–862
- Rachubinski RA, Subramani S (1995) How proteins penetrate peroxisomes. *Cell* 83:525–528
- Royer JC, Christianson LM, Yoder WT et al (1999) Deletion of the trichodiene synthase gene of *Fusarium venenatum*: two systems for repeated gene deletions. *Fungal Genet Biol* 28:68–78
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning, a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor
- Schäfer H, Nau K, Sickmann A et al (2001) Identification of peroxisomal membrane proteins of *Saccharomyces cerevisiae* by mass spectrometry. *Electrophoresis* 22:2955–2968
- Schliebs W, Würtz C, Kunau WH et al (2006) A eukaryote without catalase-containing microbodies: *Neurospora crassa* exhibits a unique cellular distribution of its four catalases. *Eukaryot Cell* 5:1490–1502
- Schumann U, Wanner G, Veenhuis M et al (2003) *AthPEX10*, a nuclear gene essential for peroxisome and storage organelle formation during *Arabidopsis* embryogenesis. *Proc Natl Acad Sci U S A* 100:9626–9631
- Shen YQ, Burger G (2009) Plasticity of a key metabolic pathway in fungi. *Funct Integr Genomics*. doi:10.1007/s10142-008-0095-6, Sep 16

- Subramani S (1993) Protein import into peroxisomes and biogenesis of the organelle. *Annu Rev Cell Biol* 9:445–478
- Szewczyk E, Andrianopoulos A, Davis MA et al (2001) A single gene produces mitochondrial, cytoplasmic, and peroxisomal NADP-dependent isocitrate dehydrogenase in *Aspergillus nidulans*. *J Biol Chem* 276:37722–37729
- Thompson JD, Gibson TJ, Plewniak F et al (1997) The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 24:4876–4882
- Ullan RV, Casqueiro J, Banuelos O et al (2002) A novel epimerization system in fungal secondary metabolism involved in the conversion of isopenicillin N into penicillin N in *Acremonium chrysogenum*. *J Biol Chem* 277:46216–46225
- Valenciano S, De Lucas JR, Van der Klei IJ et al (1998) Characterization of *Aspergillus nidulans* peroxisomes by immunoelectron microscopy. *Arch Microbiol* 170:370–376
- van den Berg MA, Albang R, Albermann K et al (2008) Genome sequencing and analysis of the filamentous fungus *Penicillium chrysogenum*. *Nature Biotech* 26:1161–1168
- van den Bosch H, Schutgens RB, Wanders RJ et al (1992) Biochemistry of peroxisomes. *Annu Rev Biochem* 61:157–197
- van der Lende TR, van de Kamp M, Berg M et al (2002) Delta-(L-alpha-Amino adipyl)-L-cysteiny-D-valine synthetase, that mediates the first committed step in penicillin biosynthesis, is a cytosolic enzyme. *Fungal Genet Biol* 37:49–55
- van Roermund CW, Elgersma Y, Singh N et al (1995) The membrane of peroxisomes in *Saccharomyces cerevisiae* is impermeable to NAD(H) and acetyl-CoA under in vivo conditions. *EMBO J* 14:3480–3486
- Veenhuis M, Harder W (1991) Microbodies. In: Rose A (ed) *The yeasts*, vol. 4. Academic, New York, pp 601–653
- Verleur N, Elgersma Y, Van Roermund CW et al (1997) Cytosolic aspartate aminotransferase encoded by the *AAT2* gene is targeted to the peroxisomes in oleate-grown *Saccharomyces cerevisiae*. *Eur J Biochem* 247:972–980
- Wanders RJ (2004) Peroxisomes, lipid metabolism, and peroxisomal disorders. *Mol Genet Metab* 83:16–27
- Wang FQ, Liu J, Dai M et al (2007) Molecular cloning and functional identification of a novel phenylacetyl-CoA ligase gene from *Penicillium chrysogenum*. *Biochem Biophys Res Commun* 360:453–458
- Waterham HR, Titorenko VI, Haima P et al (1994) The *Hansenula polymorpha PER1* gene is essential for peroxisome biogenesis and encodes a peroxisomal matrix protein with both carboxy- and amino-terminal targeting signals. *J Cell Biol* 127:737–749
- Zolman BK, Nyberg M, Bartel B (2007) IBR3, a novel peroxisomal acyl-CoA dehydrogenase-like protein required for indole-3-butyric acid response. *Plant Mol Biol* 64:59–72