J Ind Microbiol Biotechnol (2008) 35:1659–1668
DOI 10.1007/s10295-008-0430-x

ORIGINAL PAPER

Proteomics analysis of "Rovabio™ Excel", a secreted protein cocktail from the filamentous fungus *Penicillium funiculosum* grown under industrial process fermentation

Olivier Guais · Gisèle Borderies · Carole Pichereaux · Marc Maestracci · Virginie Neugnot · Michel Rossignol · Jean Marie François

Received: 10 June 2008 / Accepted: 30 July 2008 / Published online: 12 August 2008

Abstract MS/MS techniques are well customized now for proteomic analysis, even for non-sequenced organisms, since peptide sequences obtained by these methods can be matched with those found in databases from closely related sequenced organisms. We used this approach to characterize the protein content of the "Rovabio™ Excel", an enzymatic cocktail produced by *Penicillium funiculosum* that is used as feed additive in animal nutrition. Protein separation by bi-dimensional electrophoresis yielded more than 100 spots, from which 37 proteins were unambiguously assigned from peptide sequences. By one-dimensional SDS-gel electrophoresis, 34 proteins were identified among which 8 were not found in the 2-DE analysis. A third method, termed 'peptidic shotgun', which consists in a direct treatment of the cocktail by trypsin followed by separation of the peptides on two-dimensional liquid chromatography, resulted in the identification of two additional proteins not found by the two other methods. Altogether, more than 50 proteins, among which several glycosylhydrolytic, hemicellulolytic and proteolytic enzymes, were identified by combining three separation methods in this enzymatic cocktail. This work confirmed the power of proteome analysis to explore the genome expression of a non-sequenced fungus by taking advantage of sequences from phylogenetically related filamentous fungi and pave the way for further functional analysis of *P. funiculosum*.

Keywords Filamentous fungi · *Penicillium* · Proteome · Secretome · Glycosyl hydrolase

O. Guais · J. M. François
Université de Toulouse, INSA, UPS, INP,
135 Avenue de Rangueil, 31077 Toulouse, France

O. Guais · M. Maestracci · V. Neugnot
Cinabio-Adisseo France S.A.S,
135 Avenue de Rangueil, 31077 Toulouse, France

G. Borderies · C. Pichereaux · M. Rossignol
Plateforme Proteomique Génopole Toulouse Midi-Pyrénées,
IFR40-IPBS CNRS, 205 Route de Narbonne,
31400 Toulouse, France

J. M. François
INRA, UMR792 Ingénierie des Systèmes Biologiques et
des Procédés, 31400 Toulouse, France

J. M. François
CNRS, UMR5504, 31400 Toulouse, France

J. M. François (✉)
INSA, LISBP, 135 Avenue de Rangueil, 31400 Toulouse, France
e-mail: fran_jm@insa-toulouse.fr

Introduction

Filamentous fungi comprise an important class of organisms of biotechnological applications and significant commercial relevance. They are used to produce a wide variety of products ranging from human therapeutics (i.e. antibacterial and antifungal agents) to specialty chemicals (e.g. hydrolytic enzymes) [1]. In this respect, the soil deuteromycete *Penicillium funiculosum* is a remarkable cell factory for protein production, since it can grow efficiently under large-scale industrial conditions to secrete biologically active compounds including several hydrolytic enzymes [2]. This remarkable property has been further improved by mutagenic selection of a derivative of *P. funiculosum* IMI 123457 for higher capacity to secrete xylanases and cellulases under submerged cultures. This optimized strain is exploited now for the production of an enzymatic complex cocktail termed "Rovabio™ Excel."

This product is used as feed additive in animal nutrition for enhancing digestibility of the feed materials that are composed of complex carbohydrates (cellulose, hemicellulose, arabinoxylan, arabinogalactan, etc.), and hence to improve the animal performance and health. Although several enzymatic activities, including some carbohydrolases, proteases, peroxydases are known to be present, the true efficacy of the cocktail is likely depending of the combination of enzymatic activity present in the protein mixture [3].

As a first step towards investigating the enzymatic complexity in the nutritional value of Rovabio™ Excel, we aimed at characterizing the protein content in this cocktail by state-of-the art proteomic technologies. Although the genome of this filamentous fungus is not sequenced yet, a proteomic analysis has proven to be feasible when it is based on MS/MS de novo sequencing of individual peptides rather than peptides mass fingerprinting, followed by identification of the proteins from genome sequences of phylogenetically related filamentous fungi [1]. Considering that the 2-dimensional gel electrophoresis (2-DE) was not sufficient for protein identification, we complemented this analysis by a 1-dimensional separation (SDS-Page analysis) and a peptide shotgun method from which the peptides are directly generated by trypsin treatment of the enzymatic cocktail. More than 50 proteins were identified, showing that the analysis of a complex mixture was better covered by the combination of three preparative protein separation methods.

Materials and methods

Fungal strain and process condition for the Rovabio™ Excel production

The *Penicillium funiculosum* 8/403 strain used for protein production is derived from strain IMI 123457 of the International Mycology Institute collection (Egham, UK) by UV mutagenesis and selection for high production of xylanases and cellulases. The fungus was grown under industrial conditions in a liquid medium, which contains cellulose, corn steep liquor, calcium carbonate and ammonium sulphate. The pH of the medium was maintained between pH 3.0 to 6.0 by addition of HCl or sodium hydroxide and the temperature of incubation was adjusted between 27 and 36 °C. At the end of the fermentation time, fungal cells and solid particles were eliminated by centrifugation. The supernatant was concentrated by ultrafiltration. This crude preparation of enzymes secreted from *Penicillium funiculosum* cultivation was then subjected to specific process packaging to yield the "Rovabio" product. This product which was studied, in this paper, is currently sold by Adisseo as 'Rovabio™ Excel".

Two-dimensional gel electrophoresis

Sample preparation

The Rovabio™ product was dialysed in a 3.5 kDa cut-off dialysis bag (Slide-A-lyzer dialysis cassettes from Pierce, UK) overnight at 4 °C against 1,000 volumes of 5 mM Tris buffer, pH 7.5. Precipitation of proteins was carried by mixing one volume of supernatant with one volume of phenol saturated with Tris buffer (0.01 M, pH 8). After 10 min at RT and centrifugation at $3{,}000\times g$ for 5 min, the upper phase was removed and the organic phase was washed with one volume of Tris 10 mM pH 8, followed by the removal of the aqueous phase and addition of ammonium acetate at a concentration of 0.1 M to the phenolic phase. The proteins were finally precipitated by adding 4 volumes of cold methanol for 2 h at 20 °C. The precipitate was collected by centrifugation at $10{,}000\times g$ for 15 min, washed two times with 1 ml of an ice-cold 80% (v/v) methanol/water solution. The protein pellet was brought to dryness by centrifugation using a speed vacuum apparatus.

Electrophoresis

Protein samples (around 150 µg) were loaded by passive in-gel rehydration at 20 °C in 150 µL of a DeStreak rehydration solution (Amersham) complemented with 2% IPG buffer pH 3–10, 0.3% DTT for immobilized pH gradient gel (IPG) strips (7 cm length) in the linear pH range 3–10 (Bio-Rad, Hercules, CA, USA). About 250 µl of the same rehydration solution were used for immobilized pH gradient gel (IPG) strips (11 cm length) in the linear pH range 3–6. Isolectrofocusing (IEF) was performed using a MultiphorII system (GE Healthcare) in a stepwise fashion (all steps were done in gradient voltage, with the upper limits of amperage and power of 1 mA and 5 W, respectively. Steps: (1) 50 V, 1 min; (2) 50 V, 1 h; (3) 200 V, 30 min; (4) 200 V, 2 h; (5) 1,000 V, 30 min; (6) 1,000 V, 1 h 30; (7) 3,500 V, 8 h; (8) 3,500 V, 15 h) at 20 °C for a total of 55 kVh. Prior to SDS-PAGE, the IPG strips were incubated for 15 min in 0.1 M Tris pH 7.5, 6 M urea, 30% v/v glycerol, 2% w/v SDS, 50 mM DTT and traces of bromophenol blue and then for 15 min, with the same buffer except that DTT was replaced by 250 mM iodoacetamide. The equilibrated IPG strips were transferred onto a 12% polyacrylamide slab gels (20 × 20 cm) without stacking gel and overlaid with 0.5% low-melting agarose. The second dimension was run at constant current (25 mA). All gels were fixed overnight in 50% methanol and stained with the Colloidal coomassie blue G-250 (Serva Electrophoresis). Images of gels were captured with the image scanner (GE Healthcare). Image analysis was operated with Image master 2D soft (GE healthcare). Settings were customized in

order to improve spots detection. A triplicate of 2-DE gels homemade using three protein preparations from the same Rovabio™ Excel extract allowed us to validate the reliability of the method. All the spots detected were the same in the three gels and a standard pattern could be defined.

In-gel digestion

The stained spots in 2-DE were excised with a 2 mm inner diameter Pasteur pipette. Each gel piece was washed twice with 100 µL of a 25 mM ammonium bicarbonate, 50% acetonitrile (ACN) solution for 10 min. After drying in vacuo, gel pieces were rehydrated with 20 µL of modified trypsin (purchased from Promega; sequencing grade) at 10 ng/µL in 25 mM sodium bicarbonate pH 8 and digested overnight at 37 °C. The peptides were extracted as described previously [4]. Finally, the peptide mixtures were solubilized in 15 µL ACN 5%, formic acid 0.1% prior to LC-MS/MS analysis.

Protein and peptide shotgun methods

SDS-PAGE and protein shotgun

The Rovabio™ Excel protein complex was cleaned up by fractionation on ion exchange chromatography prior to apply the protein shotgun method. To this end, the protein cocktail was dialysed overnight against 20 mM NaAcetate pH 5, centrifuged for 15 min at 10,000×*g*, resuspended at 0.1 mg/ml in the NaAcetate buffer and loaded on a SP-Sepharose column (Hi-Trap; Amersham Biosciences) equilibrated with the same buffer. In the case of anion exchange chromatography, the proteins were precipitated by dropping the pH to 1.0 with dropwise addition of concentrated HCl at 4 °C. The suspension was centrifuged for 15 min at 10,000×*g* (4 °C). The pellet was solubilized overnight at 4 °C in 20 mM NaAcetate pH 5. After a second centrifugation, the supernatant was loaded on a Q-Sepharose column (Hi-Trap; Amersham Biosciences) equilibrated with the same acetate buffer, pH 5.0. Elution of protein was carried by a linear NaCl gradient (0–1 M) made in 20 mM Na acetate buffer, pH 5.0 at a flow rate of 0.4 ml/min. Fractions of 0.4 ml were collected and those containing proteins were pooled and precipitated by adding four volumes of cold acetone (−20 °C) for 2 h. After centrifugation for 15 min at 10,000×*g*, the pellet was suspended in 250 mM TrisHCl pH 6.8 containing 500 mM DTT, 10% SDS, 0.005% bromophenol blue and 50% glycerol and boiled for 5 min at 100 °C. A sample corresponding to 50 µg total protein was transferred on 12% SDS-gels [4]. The electrophoresis was carried out using a Bio-Rad mini protean II system at constant voltage of 200 V. The gels were fixed overnight in 50% methanol and stained with the Coomassie blue G250 (Serva electrophoresis). Migration lanes were manually sliced into 15 pieces of the same size. The slices were washed, digested as described previously for in-gel digestion.

Peptide shotgun

The Rovabio™ product was dialysed at 4 °C on a vivaspin (Sartorius) at a 10 kDa cut off. About 500 µg from this protein concentrate was boiled for 10 min in 250 mM sodium bicarbonate pH 8, prior to add 25 µg of trypsin (Promega). After overnight digestion at 37 °C, the resulting peptide mixture was lyophilized, resuspended in 250 µL of 80% ACN/0.1% formic acid solution and then kept on ice for 1 h to eliminate undigested proteins. After centrifugation for 15 min at 10,000×*g*, the supernatant was diluted three times with deionized water and loaded on a Supelcolsil™ LC-SCX cation exchange column (25-cm-long, 3-mm-inner diameter, 5 µm beads diameter) equilibrated in 25% ACN and 0.1% formic acid. Elution was performed with by a stepwise gradient of 1 M Ammonium acetate (5, 15, 25 and 100%) in 25% ACN and 0.1%, formic acid at a flow rate of 0.4 ml/min. Fractions of 800 µL were collected and freeze-dried. The dried residues were solubilized in 15 µL of 5%, ACN and 0.1% formic acid.

Nano LC-MS/MS analysis and search in databases

The proteins were identified by LC-MS/MS analysis using a Q-TRAP mass spectrometer (Applied Biosystems/MDS Sciex, Foster City, CA, USA) connected to a nano-HPLC system (LC-Packings). Peptide mixtures achieved by digestion of proteins with trypsin enzyme and prepared as described in the previous paragraph were fractionated by nanoflow liquid chromatography using a 10 cm-long × 75 µm-inner diameter C18 capillary column and analyzed on-line by electrospray ionization tandem mass spectrometry. The bound peptides were eluted from the C18 column at a 200 nl/min flow rate with a 60-min linear gradient of 2–60% ACN in 0.1% formic acid followed by a 5-min linear gradient of 60–90% ACN in 0.1% formic acid. Analyst software (version 1.4) was used to collect mass spectrometry data, and the mass range for the MS survey scan was *m/z* 400–1,400. The MS/MS data of the two most intense ions were collected sequentially following each MS scan using the dynamic exclusion parameter where a specific ion was sequenced only twice and was excluded from the list for 45 s. The peak lists were generated using the MASCOT script included in the Analyst software (Applied Biosystems/MDS Sciex).

The acquired data were compared with the NCBI genome database restricted to all sequenced fungal genomes using the MASCOT software (Matrix Science,

London, UK). Searches were performed with a mass tolerance of 0.5 Da in MS mode and 0.3 Da in MS/MS mode. Variable modifications selected for searching included carbamidomethylation of cysteine and oxidation of methionine. Mascot uses a probability-based "Mowse score" to evaluate data obtained from tandem mass spectra. Mowse scores were reported as $-10 \times \log_{10}(P)$ where P is the probability that the observed match between the experimental data and the database sequence is a random event. Mowse scores greater than 39 were considered significant ($P < 0.05$). This automatic search was completed by a manual interpretation of mass spectrometry data of the most intense spectra from MS/MS data non-attributed by MASCOT. MS/MS spectra were subjected to de novo sequence analysis using Bioanalyst 1.4 software (Applied Biosystems/MDS Sciex), and the generated peptide sequences were used for homology searching using available on-line tools (MS BLAST, European Molecular Biology Laboratory) as described by Waridel et al. [5] with the default search settings and the non redundant NCBI database limited to fungal protein entries.

Miscellaneous

The concentration of protein was determined by the Bradford assay (Bio-rad) according to the manufacturer instructions using BSA as a standard.

Results and discussion

Characterization of the Rovabio™ Excel by 2-DE analysis

The secreted proteins present in the Rovabio™ Excel product were separated by a two dimensional gel electrophoresis (2-DE) directly from the bulk extract since a dialysis of the cocktail with a cut off of 5 kDa prior to gel electrophoresis did not improve the separation and the quality of the gels. In the first dimension, we initially used a pH range of 3–10 and noticed that the vast majority of proteins migrated in a pH range between 3 and 5.5, indicating that most of the proteins present in the Rovabio™ Excel had a pI in this pH range. Such a profile was actually very similar to that obtained for other fungi cultivated under solid-state fermentation [6] or on complex carbohydrates [7]. Therefore, for subsequent experiments, pH 3–6 IPG strips were used for the first dimension, followed by a 12% SDS-PAGE for the second dimension. A typical 2 DE-gel made with 100 µg of crude preparation of Rovabio™ Excel is shown in Fig. 1a. This separation method resulted in the isolation of more than 100 distinct spots, in a highly reproducible manner, even using samples of Rovabio™ Excel that had been left at room temperature for 2 weeks, or warmed for 4 h at 37 °C before 2-DE analysis (data not shown).

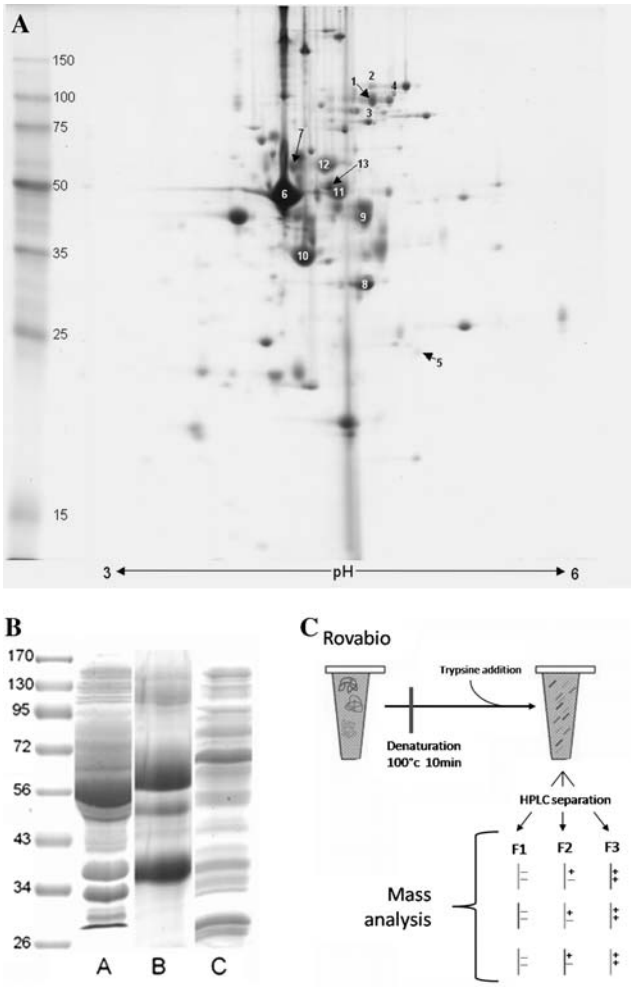

Fig. 1 Representation of the three different techniques used to separate the protein content of the Rovabio™ Excel. **a** A two dimensional gel electrophoresis of proteins of Rovabio stained with Coomassie brilliant blue G250. About 100 µg of proteins of a Rovabio™ Excel sample were loaded. Electrofocusing was realized on 11 cm pH 3–6 strips, proteins were then separated using a 12% gel. A total of 100 spots, including indicated spots, were excised and subjected to in gel trypsin digestion prior to the nanoLC-MS/MS analysis. Ladder weights are in kDa. Number on the spots referred to proteins that corresponded to enzymes previously known to be present by enzymatic assays, **b** SDS-PAGE 12% of 50 µg of protein from crude extract of Rovabio™ Excel (**a**); 50 µg of Rovabio™ proteins eluted at 1 M NaCl S-sepharose column (**b**); and 50 µg of proteins eluted at 3 M NaCl Q-sepharose column (**c**). The staining was realized with Coomassie brilliant blue R250 and ladder weights are in kDa. **c** A heat denaturated extract of Rovabio™ Excel was digested by trypsin. The peptides were separated based on their charge on an ionic exchange chromatography column before analysis by tandem LC-MS/MS

All unique spots were excised from the 2-DE gels, digested with trypsin and subjected to LC-MS/MS analysis. The spectrometric data were computerized using MASCOT. This algorithm compares experimental fragmentation

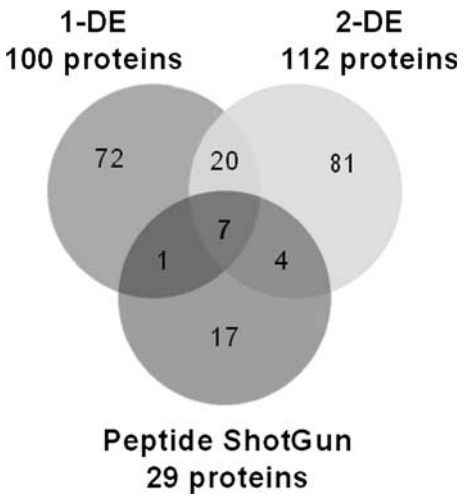

Fig. 2 Overlapping of protein identification by the three methodological approaches for proteomics analysis by LC-MS/MS. Proteins were identified by homology search from fungal genome and NCBI nonredundant protein databases from one or more unique peptides hits. Unknown/hypothetical proteins were included

spectra with calculated fragment ion mass values, obtained by applying trypsin cleavage rules to the entries in a sequence database. Identification relies on a good correlation between experimental and calculated data. For this reason, MASCOT requires exact matches. Since there is very little genomic sequence data available for *Penicillium funiculosum*, the strategy was to search for exact peptides match in a database incorporating all known fungal proteins present in the NCBI non redundant protein database. From a total of >100 spots excised and in trypsin-digested gel, we obtained more than 300 peptides sequences. From this high number of sequences, 112 proteins were identified based on one or more unique peptide hits (Fig. 2). However, this number was reduced to 37 after comparison of the identified proteins. In fact, our search using peptides match to protein sequences allowed identifying, in several cases, proteins that belong to the same family, while they were present under different name in different fungal species. For this reason, this kind of hits was compiled as single entries, assuming that they were only one protein of this family in *Penicillium funiculosum*. Moreover, this list evicted protein hits that have hypothetical or yet uncharacterized functions. Finally, 9 proteins were unambiguously assigned to enzymatic activities that were known to be highly present in the Rovabio™ Excel cocktail. As expected, protein bearing these activities were identified by many peptides, resulting in high peptides coverage (Tables 1, 2, 3, 4). In addition, most of these abundant proteins were identified by peptide sequences in more than one spot. This was particularly the case for cellobiohydrolase and xylanases. This repetitive identification of proteins in different spots could be explained by the existence of protein isoforms or by a partial proteolysis of intact proteins that has occurred during the fermentation process and/or during the process packaging step. In favor of the latter possibility, it was previously shown that the presence of protease inhibitors in the *Aspergillus flavus* mycelia during the extraction procedure reduced the presence of multiple spots in the identification of the secreted proteins [7]. Alternatively, isoforms of cellobiohydrolase have been reported [8]. Thus, additional work, including acquisition of the *P. funiculosum* genome sequence, is required to support either one or the other of the two propositions. Nonetheless, the proteomic data generated through this work will be a good basis for further exploitation of functional genomic of filamentous fungi. We also found that a relative large number of spots did not produce significant hits in the database queries. This lack of identity could be explained by our strategy to use homologous databases from closely related fungi. Indeed, the Mascot software allows matching each fragmentation spectrum with a theoretical spectrum of peptidic sequences present within the databases used. Absence of the sequence in these databases or even a single amino acid modification in the peptide sequence would fail to find this protein.

Proteomic analysis of Rovabio™ Excel complex using 1-DE

As illustrated in Fig. 1a, the protein complexity of the Rovabio™ Excel was qualitatively well resolved by a 2D-gel electrophoresis analysis and allowed already a good estimation of the enzymes present in this cocktail. While

Table 1 Peptide identification matching of *P. funiculosum* proteins and/or genes specifically by 2-DE method

Proteins	Peptides	Mowse score	No. spot	Accession no.[a]
Endoglucanase C	R.LAVDPNDNSILYFGAR.S	92	11	Q8TFP1_ASPNG
		95	18	
Beta-glucosidase	K.TLHELYLWPFADAVR.A	40	9	Q5B5S8_EMENI
Beta-galactosidase	R.GPLNEGGLYAER.Q	94	17	Q2UCU3_ASPOR/BAE60622
	R.IMLFSGEVHPFR.L	48		
Feruloyl esterase B	K.YVISTYGADSSK.V	53	16	CAC14144/Q9HE18

[a] The dot present in the peptidic sequence indicates the proteolytic clivage that occurred. Sequences without dots are deduced from de novo sequencing

Table 2 Peptide identification matching of *P. funiculosum* proteins and/or genes specifically by 1-DE method

Proteins	Peptides	Mowse score	Fraction	Accession no.[a]
Alpha-L-rhamnosidase	R.AYQGYFHSNDELLNR.I	50	B	Q4WJF8_ASPFU
Beta-galactosidase	R.IMLFSGEVHPFR.L	68	B	Q8TFE6_9EURO
Beta-galactosidase	R.GPLNEGGLYAER.Q	70	B	AAC60538
Beta-glucosidase	K.AVQFVSQLTLAEK.V	60	A, C	Q8TGI8_TALEM
	R.GKGVDVQLGPVAGPLGR.S	86		
Carboxypeptidase	K.IYVTGESYAGR.Y	60	B, C	Q4WI07_ASPFU
Feruloyl esterase B	K.YVISTYGADSSK.V	73	A	CAC14144
Endo-1,4-xylanase precursor	R.VGGTVTTANHFAAWK.A	82	C, A	Q9HFH0_PENFN
	R.VNQPSIEGTSTFNQYWSVR.T			
Endoglucanase C	R.LAVDPNDNSILYFGAR.S	105	A, C	Q8TFP1_ASPNG
Ferulic acid esterase A precursor	R.QYTLTLPSNYNPNK.A	68	A, B, C	Q8WZI7_PENFN/CAC85738
Glutaminase A	K.KYGVPLDTR.H	46	C	Q4WWT9_ASPFU

[a] The dot present in the peptidic sequence indicates the proteolytic clivage that occurred. Sequences without dots are deduced from de novo sequencing

Table 3 Peptide identification matching of *P. funiculosum* proteins and/or genes specifically by shotgun method

Proteins	Peptides	Mowse score	Accession no.[a]
Cellulase	R.QALLSETGGGNVQ.S	83	S68153
	R.QALLSETGGGNVQS.C	64	
Endo-1,4-xylanase B precursor	R.AEAINYNQDY.I	55	Q8J0K5_PENFN
	R.TSGTVTVQNHF.N	68	
	R.AEAINYNQDYIA.S	45	
	M.AVESWSGSGSGQI.S	62	
	M.AVESWSGSGSGQISL.S	84	
	R.AEAINYNQDYIASGAN.V	49	
	Q.AGMNLGTLNY.Q	38	
Endo-1,4-beta-glucanase	L.NLNQAAIDGIR.E	59	BAD72778
Endoglucanase A	T.NSGAYAVLDPH.N	59	Q96WR0_ASPKA
Endoglucanase II	AAGATSQYIFAEGN	68	O59951_ASPAC
			Q8J135_ASPAC
Rhamnogalacturonase	S.NGEVYNMAIR.G	60	Q2U7A4_ASPOR

[a] The dot present in the peptidic sequence indicates the proteolytic clivage that occurred. Sequences without dots are deduced from de novo sequencing

the majority of the proteins were separated in pH range between 3 and 6, it is not excluded that highly acidic and basic proteins and/or low abundant protein due to the relatively low sensitivity of the staining dye have escaped to this separation procedure. To circumvent this problem, nanoLC-MS/MS can be performed from proteins that are separated on a single SDS-PAGE electrophoresis (1-DE) analysis. In this case, the 1-DE gel is cut in several small pieces (in this case, 15 slices were produced from a 5 cm gel length), the bulk of proteins for each slice are in-gel digested and the resulting peptides are subject to mass spectrometer analysis. However, the weakness of this method is the low separation of the proteins on a 1-DE, which displaces the resolution of the analysis at the level of the mass spectrometer. To get partially around this low resolution, a Rovabio™ Excel sample was subjected to ions exchange chromatography fractionation prior to SDS-PAGE electrophoresis. When 1 mg of the cocktail was loaded on a SP-Sepharose (cation exchange chromatography), less than 5% of the total proteins were retained on the column, which were eventually eluted in the presence of 1 M NaCl. On the contrary, more than 90% of proteins were retained on a Q-Sepharose column (anion exchange chromatography), and not more than 10% of these proteins were eluted with a 3 M NaCl concentration. This unexpected data could be explained by the fact that the more abundant proteins in the cocktail are endo-glucanases, xylanases and cellobiohydrolases. These protein harbor two physico-chemical properties that may account for this tight binding on this type of cationic column; namely a pI ranging between 4.2 and 4.7 and a carbohydrate-binding module for retention on sugar polymers. This low elution was slightly improved to 20% when the Rovabio™ Excel sample was treated at pH 1 prior to load on the anion exchange chromatography. The 1-DE

Table 4 Proteins identified in common by 2-DE, protein shotgun and peptide shotgun. A dash at the edge of a sequence indicates the beginning or the end of the protein sequence

Proteins	Peptides	Accession no.	No. spot
Xylanase/cellobiohydrolase precursor	R.LGVTDFYGSGK.T[a,b]	Q8WZJ4_PENFN	6-7
	R.LNFVTGSNVGSR.T[a,b]		
	R.YYVQNGVVIPQPSSK.I[a,b]		
	K.TVESQSGSSYVVFSDIK.V[a,b]		
	K.TVDTTKPFTVVTQFVTDDGTSSGSLSEIR.R[b]		
	R.TYLMADNTHY.Q[c]		
	T.QFVTDDGTSSGSLSEIR.R[c]		
	F.VTDDGTSSGSLSEIR.R[c]		
	N.SGAITLDANWR.W[c]		
	K.ISGISGNVIN.S[c]		
	K.NMGSALEAGM.V[c]		
	Y.FVTMDADGGVSK.Y[c]		
	L.YFVTMDADGGVSK.Y[c]		
Endo-1,4-xylanase D precursor	K.EHINGVVTHYK.G[a]	Q5ZNB1_PENFN	8-9
	K.LYYNDYNIEYAGAK.A[a,b]		
	K.AAGLLYFGTAVDNPDLSDSK.Y[a,b]		
	K.YLVNLETADFGQITPANAMK.W[a,b]		
	K.ANMAAFTALGVDVAITELDIR.M[b]		
	K.GQCYAWDVVNEALNEDGTYR.Q[b]		
	R.IDGVGLQSHFIVGQTPSLATQK.A[a,b]		
	K.WQPTEPSQGSYTFTQGDQIASLAK.S[a,b]		
	R.QNVFYQHIGEAYIPIAFAAAAAADPNAK.L[b]		
	R.MTLPDTSAL.Q[c]		
	SALQTQQ[c]		
	K.TVDTTKPF.T[c]		
	N.DYNIEYAGAK.A[c]		
Cellobiohydrolase II	K.AYIDSIVAQLK.A[a,b]	O93837_9ASCO	10-11-12
	K.AAEIPSFVWLDTAAK.V[a,b]		
	K.VPTMGTYLANIEAANK.A[a,b]		
	K.AGASPPIAGIFVVYDLPDR.D[a,b]		
	K.AYPDVHTILIIEPDSLANMVTNLSTAK.C[b]		
	R.GLATNVANYNAW.S[c]		
	R.GLATNVANYN.A[c]		
	L.ATNVANYNAW.S[c]		
	A.SNGEYTVANNGVANYK.A[c]		
	N.GEYTVANNGVANYK.A[c]		
	L.LTSNGWPNAH.F[c]		
	R.GLATNVANY.N[c]		
	L.LTSNGWPNAH.F[c]		
	Q.LLTNANPALV.-[c]		
	VQLLTNANPALV[c]		
	F.VVYDLPDR.D[c]		
	A.ASGNPFSGY.Q[c]		
	Y.IDSIVAQLK.A[c]		
	M.GTYLANIEAANK.A[c]		

Table 4 continued

Proteins	Peptides	Accession no.	No. spot
Alpha-galactosidase	R.VTGDINATWDR.I[a,b]	O13295_PENPU	13
	K.EYTAQLEAHDVAVLK.V[a,b]		
	R.AHFALWAIMK.S[a]		
	K.NADLIAFNQDPVIGKPALPYK.Q[a,b]		
	LILPDDVGR[c]		

[a] Sequences obtained from 2-DE

[b] sequences obtained from 1-DE

[c] sequences obtained from peptide shot-gun

The dot present in the peptidic sequence indicates where the proteolytic clivage occurred. Sequences without dots are deduced from de novo sequencing

gel pattern obtained from crude Rovabio™ Excel and purified preparations from ions exchange chromatography is reported in Fig. 1b. The separation of proteins was relatively poor, except with the fractions eluted from the Q-sepharose column (Fig. 1b) that showed many peptide bands ranging from high (>100 kDa) to low (<30 kDa) molecular mass.

Mass spectrometer analysis of protein content in these three preparations after their separation on a 1-DE gel yielded about 100 candidate proteins (Fig. 2). The use of the three fractions was necessary because many proteins were found only in one of them (data not shown). Table 2 summarized the list of unambiguously assigned proteins from 1-DE analysis in the Rovabio™ Excel, taking into account that they were identified by at least two unique peptides sequences that exactly matched known proteins of the fungal nonredundant protein databases.

Proteomic analysis of Rovabio™ Excel complex using peptide shotgun methods

A third method termed 'Peptide shotgun' also called "DALPC" for direct analysis of large protein complexes technique [9] was applied to the Rovabio™ Excel enzyme cocktail. Briefly, about 1 mg of protein sample was heat denaturated for 10 min and trypsin-digested. After removal of the nondigested proteins, the peptides were separated by an off-line 2D nano liquid chromatography prior to the MS/MS analysis. The first dimension chromatography was performed on a Supelcosil™ LC-SCX column to separate the peptides by their charge. The second dimension occurs on a hydrophobic column that is in line with the injection in the Q-TRAP. This method is qualified as unbiased because all kinds of proteins, acidic, basic, low and high molecular weight should be in theory analyzed. However, we were confronted in a major difficulty to assign sequences to protein identification using MASCOT, because the fragmentation spectra were mostly lacking traditional ions tagging arginine or lysine that resulted from trypsin-digestion and serving as a starting point for ab initio sequencing. The reason of this failure was that many peptides were already present in the Rovabio™ Excel cocktail prior to the trypsin degradation. The presence of these peptides could be due to mild proteolysis by secreted proteases existing in the enzymatic cocktail and/or generated during the formulation and the packaging process leading to the commercial Rovabio™ Excel product. The protein identification was further complicated by the fact that the 'peptide shotgun' analysis is carried out off-line on all fractions collected from Supelcosil column, resulting in peptides from a single protein potentially distributed in several fractions. This configuration severely limits the efficacy of using MASCOT for protein identification, unless grouping all the spectrometric data in a single file. We did not choose this mode of analysis because the size of such a file combined with the need to run MASCOT without digestion rule increased the number of peptides to be analyzed, and would have overcome our computational capabilities. Altogether, this preparative method allowed identifying 29 proteins based on one or more peptide hits (Fig. 2), from which 7 were considered unambiguous, as known to be present in the enzymatic cocktail (Table 3).

Comparison of the three methods

The maximal coverage of proteins present in the Rovabio™ Excel was one of the goals of this study. Therefore, three preparative methods for protein identification have been employed, which together led to the identification of >200 proteins in the enzymatic cocktail based on peptide hits (Fig. 2). This study also showed that the overlapping between the three methods was rather weak, as only seven distinct proteins were identified by the three methods, among which four correspond to abundant enzymes known to be secreted by *P. funiculosum* (Table 4). This result argues that protein content in a biological sample cannot be

covered by a single method. This large number can be explained by the fact that proteins were identified by homology search against all known fungal proteins combined with the annotated and partially annotated fungal genome databases. In addition, protein identification was assigned by exact peptides matching. Therefore, in several cases, peptides sequences obtained by either one of the three methods identified a different protein that was assigned to a different fungal species, but in most cases, these different proteins actually belong to the same functional family and probably bear an identical catalytic function. As for instance, peptides sequences from 2-DE spots were assigned to endoglucanase 2 from *P. janthinellium* and endoglucanase C from *A. niger*, whereas peptides from shotgun analysis identified endoglucanase II from *A. aculeatus*. It is therefore possible that these different peptides actually correspond to the same protein in *P. funiculosum*, because they belong to the same GH5 family (http://www.cazy.org/). Taking into account this remark, we grouped the identified proteins according to their identical family category or catalytic function. Such grouping together leads to of the characterization of at least 50 different enzymatic activities to be present in the Rovabio™ Excel (Tables 4, 5). Therefore, the overlapping between 1-DE and 2-DE was about 40%, whereas it was only 15% with the 'peptide shotgun'. However, even if this method had a much lower resolution than the two previous methods, it complemented the 2-DE and 1-DE analyses, as for instance in this case by identifying a cellulase, and a rhamnogalacturonase that were not found by these two methods (Tables 1, 2, 3). In addition, a great advantage of combining various approaches for a proteomic study on a non-sequenced organism was that the covering rate (with respect to its full sequence size) for a given protein can be dramatically increased, as for instance, for the cellobiohydrolase 093837 and xylanase Q6WJ4, for which the 2-D electrophoresis analysis gave, respectively 19.7% and 16.3% coverage rate, and this coverage increased up to 26 and 21.7% by cumulating the 2-D and 1-D analyses.

Comparison with the secretome from other fungal species

In spite of the fact that the genome of *P. funiculosum* is not sequenced, our proteomic approach successfully identified by sequence homology a larger number of proteins in the Rovabio™ Excel that were originally identified by enzymatic assays. This includes several glycosylhydrolases and lignolytic enzymes involved in degradation of complex carbohydrates, exocellular oxydases (laccase, manganese peroxydase) and some proteases (alkaline and aspartic proteases, carboxypeptidase) (Table 5). The proteomic analysis also identified several cell-wall bound proteins, such as

Table 5 List of putative enzymatic activities identified by 2-DE, 1-DE and peptide shotgun methods present in Rovabio™ Excel from *Penicillium funiculosum* fermentation process

Potential enzymatic activities identified in Rovabio™ Excel		
Protein name	Protein family	Number of matching peptides
Acetyl xylan esterase	CE 5	1
Alkaline protease		1
Alpha-1,2-mannosyltransferase	GT 1	1
Alpha-galactosidase	GH 27	5
Alpha-glucuronidase	GH 67	1
Alpha-L-rhamnosidase	GH 78	1
Alpha-mannosidase	GH 47	1
Alpha-xylosidase	GH 31	1
Arabinofuranosidase	GH 62	8
Aspartic protease		1
Avicelase III		2
Beta-1,3-glucanosyltransferase	GT 3	1
Beta-1,4-xylosidase	GH 3	4
Beta-1,6-glucanase	GH 5	2
Beta-galactosidase	GH 35	10
Beta-glucosidase	GH 3	3
Beta-xylosidase	GH 3	4
Carboxypeptidase		2
Catalase		3
Cellobiohydrolase I	GH 7	2
Cellobiohydrolase II	GH 6	11
Cellulase	GH 5	1
Dextranase	GH 49	1
Endo-1,4-beta-galactanase	GH 53	2
Endo-1,4-xylanase	GH 11	2
Endo-1,4-xylanase B	GH 11	8
Endo-1,4-xylanase D	GH 11	12
Endoglucanase	GH 5	3
Endopygalacturonase	GH 55	1
Exo-beta-D-glucosaminidase	GH 2	2
FAD binding monooxygenase		1
Ferulic acid esterase A	CE 1	1
Feruloyl esterase B	CE 1	1
GDP-mannose 4,6-dehydratase		1
Glucoamylase	GH 15	1
Glutaminase A		1
Laccase		1
Lipase		1
Manganese peroxydase precursor		1
Neutral endopolygalacturonase	GH 28	5
Oxidoreductase		1
Polygalacturonase	GH 28	4
Polyketide synthase		1

Table 5 continued

Potential enzymatic activities identified in Rovabio™ Excel		
Protein name	Protein family	Number of matching peptides
Rhamnogalacturonase	GH 28	1
Steroid monooxygenase		1
Swollenin		3
Type I phosphodiesterase/ Nucleotide pyrophosphatase		1
Ubiquitin-conjugating enzyme		1
Xylanase/cellobiohydrolase	GH 7	10

α-1,2-mannosyltransferase, β-1,3-glucanosyltranferase, endoglucanase, as well as proteins that are not known to be secreted such as actin and some transcription factors. The presence of these proteins may likely arise from lysed cells at the end of the fermentation process. Overall, the protein content is largely comparable to what has been reported previously for other fungal species cultivated on glass wool [10], on rutin [7] or on cellulose [11]. The content in *P. funiculosum* seems perhaps richer than that of *A. niger* although the mode of cultivation and the method of protein identification were different. For the latter filamentous fungus, the growth was on wheat bran, the culture was carried out under submerged and solid-state conditions and, only 29 proteins were identified using peptides mass fingerprinting (MALDI-TOF/MS) instead of peptides amino acid sequences [6].

Concluding remarks

The aim of this study was to obtain a maximal coverage of the protein present in a commercially available enzymatic cocktail generated from *P. funiculosum* under industrial process fermentation. Using a combination of three preparative protein methods coupled to high resolution nano LC-MS/MS techniques, a total of 50 enzymatic activities were unambiguously identified, whereas only 37 were assigned using the classical 2-DE method. This work shows that the combination of multiple approaches in proteomic study maximizes the chance of protein identification, mainly when the genome of the organism under study is not sequenced. The availability of complete genome sequences from many fungal species has been an important factor in this study. However, the relative ambiguity about identification and functional classification of secreted proteins, the relative high number of peptides sequences without protein assignment together with the enormous biotechnological potential of *P. funiculosum* strongly argues for the need to sequence the genome of this filamentous fungus in a near future.

Acknowledgments This work was supported by CINABio-Adisseo (Antony, France) and by a Region Midi Pyrénées grant No. 06001324 to J.M.F. We thank Dr Jean Luc Parrou for critical suggestions during this work.

References

1. Kim Y, Nandakumar MP, Marten MR (2007) Proteomics of filamentous fungi. Trends Biotechnol 25:395–400. doi:10.1016/j.tibtech.2007.07.008
2. Belshaw NJ, Haigh NP, Fish NM, Archer DB, Alcocer MJ (2002) Use of a Histone H4 promoter to drive the expression of homologous and heterologous proteins by *Penicillium funiculosum*. Appl Microbiol Biotechnol 60:455–460. doi:10.1007/s00253-002-1141-2
3. Olukosi OA, Sands JS, Adeola O (2007) Supplementation of carbohydrases or phytase individually or in combination to diets for weanling and growing-finishing pigs. J Anim Sci 85:1702–1711. doi:10.2527/jas.2006-709
4. Borderies G, Jamet E, Lafitte C, Rossignol M, Jauneau A, Boudart G et al (2003) Proteomics of loosely bound cell wall proteins of *Arabidopsis thaliana* cell suspension cultures: a critical analysis. Electrophoresis 24:3421–3432. doi:10.1002/elps.200305608
5. Waridel P, Frank A, Thomas H, Surendranath V, Sunyaev S, Pevzner P et al (2007) Sequence similarity-driven proteomics in organisms with unknown genomes by LC-MS/MS and automated de novo sequencing. Proteomics 7:2318–2329. doi:10.1002/pmic.200700003
6. Oda K, Kakizono D, Yamada O, Iefuji H, Akita O, Iwashita K (2006) Proteomic analysis of extracellular proteins from *Aspergillus oryzae* grown under submerged and solid-state culture conditions. Appl Environ Microbiol 72:3448–3457. doi:10.1128/AEM.72.5.3448-3457.2006
7. Medina ML, Haynes PA, Breci L, Francisco WA (2005) Analysis of secreted proteins from *Aspergillus flavus*. Proteomics 5:3153–3161. doi:10.1002/pmic.200401136
8. Phalip V, Delalande F, Carapito C, Goubet F, Hatsch D, Leize-Wagner E et al (2005) Diversity of the exoproteome of *Fusarium graminearum* grown on plant cell wall. Curr Genet 48:366–379. doi:10.1007/s00294-005-0040-3
9. Link AJ, Eng J, Schieltz DM, Carmack E, Mize GJ, Morris DR et al (1999) Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol 17:676–682. doi:10.1038/10890
10. Zorn H, Peters T, Nimtz M, Berger RG (2005) The secretome of *Pleurotus sapidus*. Proteomics 5:4832–4838. doi:10.1002/pmic.200500015
11. Wymelenberg AV, Sabat G, Martinez D, Rajangam AS, Teeri TT, Gaskell J et al (2005) The Phanerochaete Chrysosporium secretome: database predictions and initial mass spectrometry peptide identifications in cellulose-grown medium. J Biotechnol 118:17–34. doi:10.1016/j.jbiotec.2005.03.010