Planta (2005) 221: 690–704
DOI 10.1007/s00425-004-1474-4

ORIGINAL ARTICLE

**Denise I. Jacobs · Marco Gaspari · Jan van der Greef
Robert van der Heijden · Robert Verpoorte**

Proteome analysis of the medicinal plant *Catharanthus roseus*

Received: 17 May 2004 / Accepted: 3 December 2004 / Published online: 29 January 2005

Abstract A proteomic approach is undertaken aiming at the identification of novel proteins involved in the alkaloid biosynthesis of *Catharanthus roseus*. The *C. roseus* cell suspension culture A11 accumulates the terpenoid indole alkaloids strictosidine, ajmalicine and vindolinine. Cells were grown for 21 days, and alkaloid accumulation was monitored during this period. After a rapid increase between day 3 and day 6, the alkaloid content reached a maximum on day 16. Systematic analysis of the proteome was performed by two-dimensional polyacrylamide gel electrophoresis. After day 3, the proteome started to change with an increasing number of protein spots. On day 13, the proteome changed back to roughly the same as at the start of the growth cycle. 88 protein spots were selected for identification by mass spectrometry (MALDI-MS/MS). Of these, 58 were identified, including two isoforms of strictosidine synthase (EC 4.3.3.2), which catalyzes the formation of strictosidine in the alkaloid biosynthesis; tryptophan synthase (EC 4.1.1.28), which is needed for the supply of the alkaloid precursor tryptamine; 12-oxophytodienoate reductase, which is indirectly involved in the alkaloid biosynthesis as it catalyzes the last step in the biosynthesis of the regulator jasmonic acid. Unique sequences were found, which may also relate to unidentified biosynthetic proteins.

Keywords Alkaloid biosynthesis · *Catharanthus* · MALDI-MS/MS · Proteomics · Strictosidine synthase

D. I. Jacobs · R. Verpoorte
Division of Pharmacognosy, Leiden/Amsterdam
Center for Drug Research, Gorlaeus Laboratories,
P.O. Box 9502, 2300 Leiden, The Netherlands

M. Gaspari · J. van der Greef · R. van der Heijden (✉)
Division of Analytical Biosciences, Leiden/Amsterdam
Center for Drug Research, Gorlaeus Laboratories,
P.O. Box 9502, 2300 Leiden, The Netherlands
E-mail: heijden@lacdr.leidenuniv.nl
Tel.: +31-71-5274320
Fax: +31-71-5274277

Abbreviations IPG: Immobilized pH gradient · 2D-PAGE: Two-dimensional polyacrylamide gel electrophoresis · MALDI-TOF: Matrix assisted laser desorption ionization time of flight · MS: Mass spectrometry

Introduction

Catharanthus roseus (Madagascar periwinkle) is the source of the antitumor drugs vinblastine and vincristine. These terpenoid indole alkaloids are minor products in the leaves of the plants, making their isolation elaborate and costly. The production of natural products by means of biotechnological methods like large-scale culture of plant cells and metabolic engineering may provide a good alternative (Verpoorte et al. 1999). The low yield of the alkaloids in the plant combined with their high market price made *C. roseus* an important model system for plant biotechnology and secondary metabolism studies. Up to now, a number of enzymes involved in the biosynthetic pathway of alkaloids (cf. Fig. 1) has been isolated and characterized and several genes were cloned (Verpoorte et al. 1997; van der Heijden et al. (2004). Furthermore, transcription factors have been identified in *C. roseus*, which regulate the expression of genes coding for enzymes participating in alkaloid biosynthesis (Van der Fits and Memelink 2000).

Plant cell cultures supply biomass under controlled conditions and are therefore a good system for comparative proteome and metabolome studies. Previously, it was shown that *C. roseus* cell line A11 accumulates the alkaloid precursor secologanin (Dagnino et al. 1996; Contin et al. 1998). This accumulation is in contrast to other reported *C. roseus* cell lines. Generally, the availability of secologanin is limiting the alkaloid accumulation (Moreno et al. 1993; Whitmer et al. 1998).

Secologanin accumulation of cell line A11 depends on the inoculum density of the culture at the start of a growth cycle (Contin et al. 1999). At low inoculum

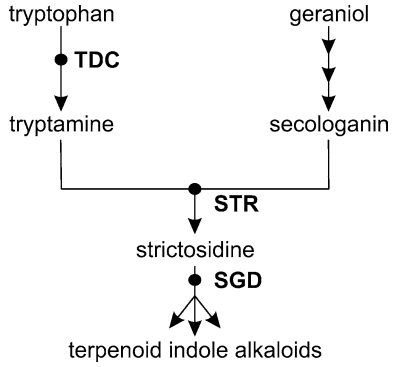

Fig. 1 Schematic representation of the biosynthesis of terpenoid indole alkaloids. The indicated enzymes are *STR* strictosidine synthase, *SGD* strictosidine β-glucosidase, *TDC* tryptophan decarboxylase

density (40 g fresh weight (FW)/l) tryptamine accumulated, while secologanin accumulated at higher inoculum density (≥160 g FW/l). The maximum content of secologanin was observed between day 10 and 20 after inoculation. The cell line was also found to accumulate the alkaloids ajmalicine, tabersonine, vindolinine, catharanthine, strictosidine and serpentine. The accumulation of alkaloids reached a maximum at an inoculum density of 160 g FW/l. Secologanin was still available; obviously tryptamine was limiting for alkaloid biosynthesis.

In order to find novel proteins involved in the alkaloid biosynthesis of *C. roseus*, a proteomic approach was now undertaken. Classic approaches for studying the biosynthesis of secondary metabolites were aimed at identification of the enzymes one by one through analysis of intermediates, followed by development of specific enzyme test methods and subsequent purification of the specific enzymes. This approach is time-consuming and often hampered due to the chemical instability of intermediates, which makes the development of an appropriate activity assay difficult. Moreover, many proteins (e.g. transport or regulatory proteins) have an activity that cannot be assayed. These proteins are easily overlooked by the classic approaches.

Proteomics has already been used for many different applications in plant sciences, including the study of proteins of biosynthetic pathways leading to secondary metabolites (reviewed by Jacobs et al. 2000). The aim of this study was to demonstrate that a proteomic approach for the identification of proteins involved in alkaloid biosynthesis is feasible, and to find novel sequences, keeping in mind that novel proteins involved in alkaloid biosynthesis will be absent from the current protein- and gene-sequence databases. As a first approach, based on our previous biochemical studies, we used the *C. roseus* cell line A11 grown for 21 days. On days 0, 3, 6, 9, 13, and 20, cells were harvested and analyzed for their iridoid and alkaloid content. In parallel, protein composition was analyzed by two-dimensional polyacrylamide gel electrophoresis (2D-PAGE). Several proteins showing interesting expression profiles (see Results) were identified by tandem mass spectrometry on a quadrupole time-of-flight mass spectrometer equipped with a MALDI (matrix-assisted laser desorption ionization) source and the measured masses were used for database searches. The identified proteins are discussed with respect to their expression profiles during the growth cycle and with respect to their putative role in alkaloid accumulation within *C. roseus* cells.

Materials and methods

Cell suspension cultures

Cell suspension cultures of *C. roseus* (cell line A11; Contin et al. 1998, 1999) were grown in medium containing B5 salts (Gamborg et al. 1968), 0.1 g/l myo-inositol, 10 mg/l thiamine-diHCl, 1 mg/l pyridoxine-HCl, 1 mg/l nicotinic acid, 1.86 mg/l naphthalene acetic acid and 20 g/l sucrose at pH 5.8 ± 0.05. For maintaining the cell cultures, they were subcultured every 3 weeks by a twofold dilution with fresh medium in 250-ml flasks containing 100 ml of medium. For the growth experiments, eight to nine grams of biomass, filtered over a sterile glass filter, was added per 250-ml flask containing 50 ml of the same medium. The cultures were kept on a gyratory shaker at 100 rpm in a climate room at 25 ± 1°C and 1,600 lux for 24 h/day.

The growth experiment was performed in duplicate. The cells were harvested at the days indicated by filtration over a glass filter and washed with water. They were immediately frozen in liquid nitrogen and stored at −80°C until further use. For analysis by 2D-PAGE, the cells from the duplicate growth experiments were mixed.

Alkaloid determination

One gram FW cells was homogenized with 1 ml of 20% trichloroacetic acid (TCA) using an UltraTurrax and centrifuged for 5 min at 16,000 *g*. The supernatant was analyzed for the presence of strictosidine, ajmalicine, and vindolinine by HPLC (van der Heijden et al. 1987). The injection volume of the supernatant was 100 µl. Alkaloids were separated on a µBondapak-Phenyl 300×3.9 mm column (Waters) with an eluent containing 50 mM sodium dihydrogen phosphate buffer (adjusted to pH 3.9 with phosphoric acid)–acetonitrile–2-methoxyethanol (80:15:5, by volume) at a flow rate of 2 ml/min. The alkaloids were identified by their retention time and UV spectra, which were recorded with a photodiode array detector (990, Waters). The peak areas at 240 nm were used to quantify each alkaloid.

Sample preparation for 2D-PAGE

Samples for 2D-PAGE were prepared according to Jacobs et al. (2001). One gram of frozen cells was

ground with a mortar and pestle, which were pre-cooled with liquid nitrogen. In brief, 10 ml of cold (−20°C) 10% TCA in acetone containing 0.07% β-mercaptoethanol was directly added to the sample (Granier 1988). The sample was kept at −20°C for 2 h to allow for complete precipitation. After centrifugation (15 min, 3,000 *g*), the samples were washed thrice with 10 ml acetone (−20°C) containing 0.07% β-mercaptoethanol to remove the TCA. The precipitate was lyophilized for 1 h and then solubilized in 2 ml of urea-buffer (9.5 M urea, 2% CHAPS, 0.5% CA, 65 mM DTT). This mixture was repeatedly vortexed, and it was centrifuged (5 min, 16,000 *g*) after 1 h. The supernatant was recovered and stored in aliquots at −80°C. The protein concentration was determined by the method of Peterson (1977).

2D-PAGE and gel analysis

The 2D-PAGE was performed according to Görg et al. (2000). About 100 µg of protein was loaded on an 18-cm IPG-strip with pH gradient between 4 and 7 (IPG; immobilized pH gradient, Amersham Biosciences) by in-gel rehydration at 30 V for 11 h. Isoelectric focusing (IEF) was performed at 20°C for 75 kVh on an IPGphor (Amersham Biosciences). Prior to SDS-PAGE, the IPG-strips were equilibrated for 2×10 min in 6 M urea, 30% glycerol, 2% SDS in 0.05 M Tris-HCl pH 8.6 containing 1% DTT for the first equilibration step and 4% iodoacetamide for the second equilibration step. SDS-PAGE (12% T, 2.6% C, Laemmli buffer system (Laemmli 1970) was performed at 20°C with the ISO-DALT system (Amersham Biosciences). For calibration, low-molecular weight marker proteins (Amersham Biosciences) were applied on the gel via a small piece of filter paper. The gels were stained with the fluorescent Sypro Ruby protein gel stain according to the manufacturer (Bio-Rad). Fluorescence image acquisition was performed on a Typhoon 9400 fluorescence scanner (Amersham Biosciences). For Sypro Ruby detection, the instrument was equipped with a 532 nm laser and a 610 nm band-pass emission filter. The gel images were analyzed with the software packages ImageMaster 2D Elite v3.01 and ImageMaster Database v3.11 (Amersham Biosciences). Spot detection was performed using the following parameters: sensitivity 9660, operator size 27, noise factor 5, background 0. Spot detection and matching as performed by ImageMaster were checked and only edited when necessary. For background subtraction, the mode of non-spot option was used with a factor of 45. The spot intensities of individual spots were normalized to the total intensity of all spots on the gel. Of all samples, five replicate gels were made. The gel with the most protein spots was selected as reference gel. Only spots that were detected on at least four out of five gels were involved in further analysis.

In-gel trypsin digestion and mass spectrometry

The protein spots selected for identification by mass spectrometry were manually excised from the gels while they were visualized on an UV transilluminator at 280 nm. In-gel digestion was performed as described by Shevchenko et al. (1996). The gel pieces were washed twice with 100 mM ammonium bicarbonate and 50% acetonitrile, reduced with 10 mM DTT, followed by alkylation with 55 mM iodoacetamide. The gel pieces were washed twice again and dried in a SpeedVac. For proteolytic digestion, the gel pieces were rehydrated in digestion buffer containing 50 mM ammonium bicarbonate, 5 mM calcium chloride and 25 ng/µl trypsin (sequencing grade, Promega). After overnight incubation at 37°C, the peptides were extracted with acetonitrile and 5% formic acid. These extracts were pooled and dried in a SpeedVac. The lyophilized digest was redissolved in 10 µl of 0.1% (v/v) trifluoroacetic acid (TFA), and purified with ZipTip µC18 pipette tips (Millipore), using the procedure recommended by the manufacturer, except for the elution step. Peptides were eluted directly from the ZipTip onto the MALDI target in 1–1.5 µl matrix solution. Matrix solution was 15 mg/ml 2,5-dihydroxybenzoic acid (DHB) in 0.1% TFA/acetonitrile 1:1 (v/v) for MALDI-TOF experiments and 40 mg/ml DHB in 0.1% TFA/acetonitrile 1:1 (v/v) for oMALDI-QSTAR experiments.

The MALDI-TOF peptide mass spectra were acquired in reflector and delayed extraction mode on a BIFLEX III mass spectrometer (Bruker), using the maximum acceleration voltage (19 kV). Typically, a mass spectrum was the average of 100–200 laser shots. All further processing, including spectrum calibration based on trypsin autodigest, was performed using the software package XMASS 5.0 (Bruker).

The MALDI-MS/MS data were acquired on an oMALDI-QSTAR Pulsar quadrupole time-of-flight mass spectrometer (Applied Biosystems/MDS Sciex), which was permanently run in enhanced mode. Full MS spectra were continuously acquired for 0.5–1 min (m/z range 600–3,000). Major peaks in the spectra were selected for MS/MS experiments. Collision energy was manually optimized to achieve the most informative MS/MS data and acquisition proceeded for 1–3 min per spectrum. MS/MS data could be acquired for 3–10 peptides per spot, the number of peptides being limited by sensitivity and/or by sample depletion on the MALDI target. Mass calibration was achieved externally using a peptide calibration standard with mass range 1,000–4,000 Da (Bruker). Whenever oMALDI-QSTAR single MS peptide fingerprinting data were used, they were searched as described for MALDI-TOF data, but with a mass accuracy of 30 ppm. For de novo sequencing of tryptic peptides, interpretation of tandem mass spectra was achieved manually using Bioanalyst version 1.1 (Applied Biosystems/MDS Sciex).

Database search

The MALDI-TOF peptide fingerprinting data were queried against all Viridiplantae entries in the NCBInr (non-redundant) database using the MASCOT search engine of Matrix Science Ltd. (http://www.matrixscience.com; Perkins et al. 1999). Search parameters were: trypsin cleavage; allow up to one missed cleavage; no restriction on protein mass; peptide mass tolerance 75 ppm; quantitative modification: carbamidomethyl (C); nonquantitative modification: oxidation (M).

For MALDI-MS/MS data acquired on the oMALDI-QSTAR instrument, database search was performed using the MS/MS ion search available on the MASCOT search engine. Data were queried against all Viridiplantae entries in the NCBInr database. Search parameters were the same, except for: peptide mass tolerance 100 ppm; fragment ion mass tolerance 0.2 Da.

Results

Alkaloid accumulation during a growth cycle

Figure 2 shows the accumulation of alkaloids during the growth cycle of *C. roseus* cell line A11. The major alkaloids that accumulated in the cells were vindolinine, ajmalicine and strictosidine. In contrast to the earlier results with this cell line (Contin et al. 1999), no accumulation of the iridoid secologanin was found. This change in behavior of the cell line was probably due to a change of the environment (the cell line had to be moved to another climate room) or due to variability of the cell line itself.

A few days after inoculation, a fast increase in alkaloid accumulation was observed. After day 6, the amount was still increasing and reaches a maximum on day 16. After an increase in biomass the first days, the accumulation of alkaloids starts to increase.

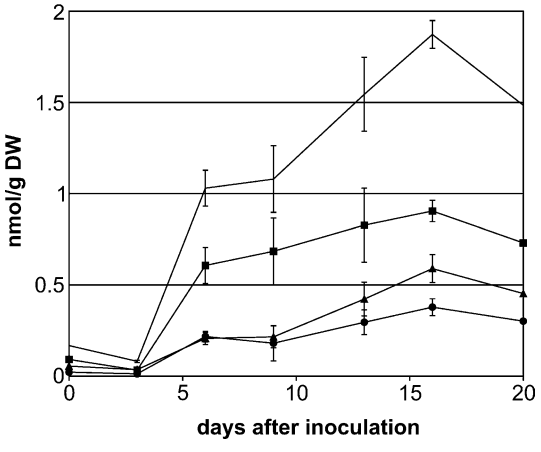

Fig. 2 Alkaloid accumulation during the growth cycle of *Catharanthus roseus* cell line A11. The mean values of strictosidine (*filled circle*), ajmalicine (*filled triangle*), vindolinine (*filled square*) and total (*no symbol*) are given. The *error bars* represent the two measured values

Changing protein patterns during growth

The 2D-gels of proteins extracted from *C. roseus* cells harvested on various days after inoculation are shown in Fig. 3. The 2D-gels from days 0, 16 and 20 were largely the same as on day 3 and 13, respectively, and were therefore omitted. Most changes occurred between day 3 and 13; in this period, alkaloid accumulation also took place (Fig. 2). After day 3, the protein pattern started to change. The most obvious changes occurred around day 9, when most protein spots appeared new on the 2D-gels, especially protein spots with a M_w of above 43 kDa and a pI between 6 and 7. At the end of the growth cycle, the protein pattern changed back to roughly the same as the beginning of the cycle.

On day 9, the most protein spots were detected on the 2D-gels; i.e., 905 protein spots on average based on five gels (Table 1) and a maximum of 988. For gel analysis, the 2D-gel with the most protein spots (988) was selected as a reference gel (Fig. 3c). All the gels in this experiment were matched against this reference gel.

Due to experimental variations, not all protein spots were detected on each gel of the same sample. To average out this experimental variation, five gels were made for each sample. Growth curves have been given in detail for *C. roseus* cell cultures by Contin et al. (1999); and on day 9, most cells are in the late exponential phase or are about to enter the stationary phase. Of the protein spots detected on the five gels of day 9, 95% could be matched to the reference gel. Although the gels of the other days contained less protein spots, only 83% of the protein spots on day 6 and 13, and 65% of the protein spots on day 3 could be matched to the reference gel (Table 1). Besides the many new protein spots that appeared on day 9, also many protein spots disappeared. These are the protein spots detected on day 3, 6 or 13, but not on day 9. For example, after comparing day 6 and day 9, approximately 75 protein spots disappeared and approximately 300 protein spots appeared.

Even though the most obvious changes on the 2D-gels occurred around day 9, the most interesting changes with respect to alkaloid biosynthesis are those between day 3 and 6. As is shown in Fig. 2, alkaloid accumulation started between day 3 and day 6 and the amount of alkaloids in the cells continued to increase until day 16. Therefore, the most interesting protein spots with respect to alkaloid accumulation are those, which appeared on day 6 and did not disappear completely later on. A selection of protein spots was made of spots showing a pattern in agreement with the alkaloid accumulation. With the ImageMaster Database software, criteria were chosen to make two selections of protein

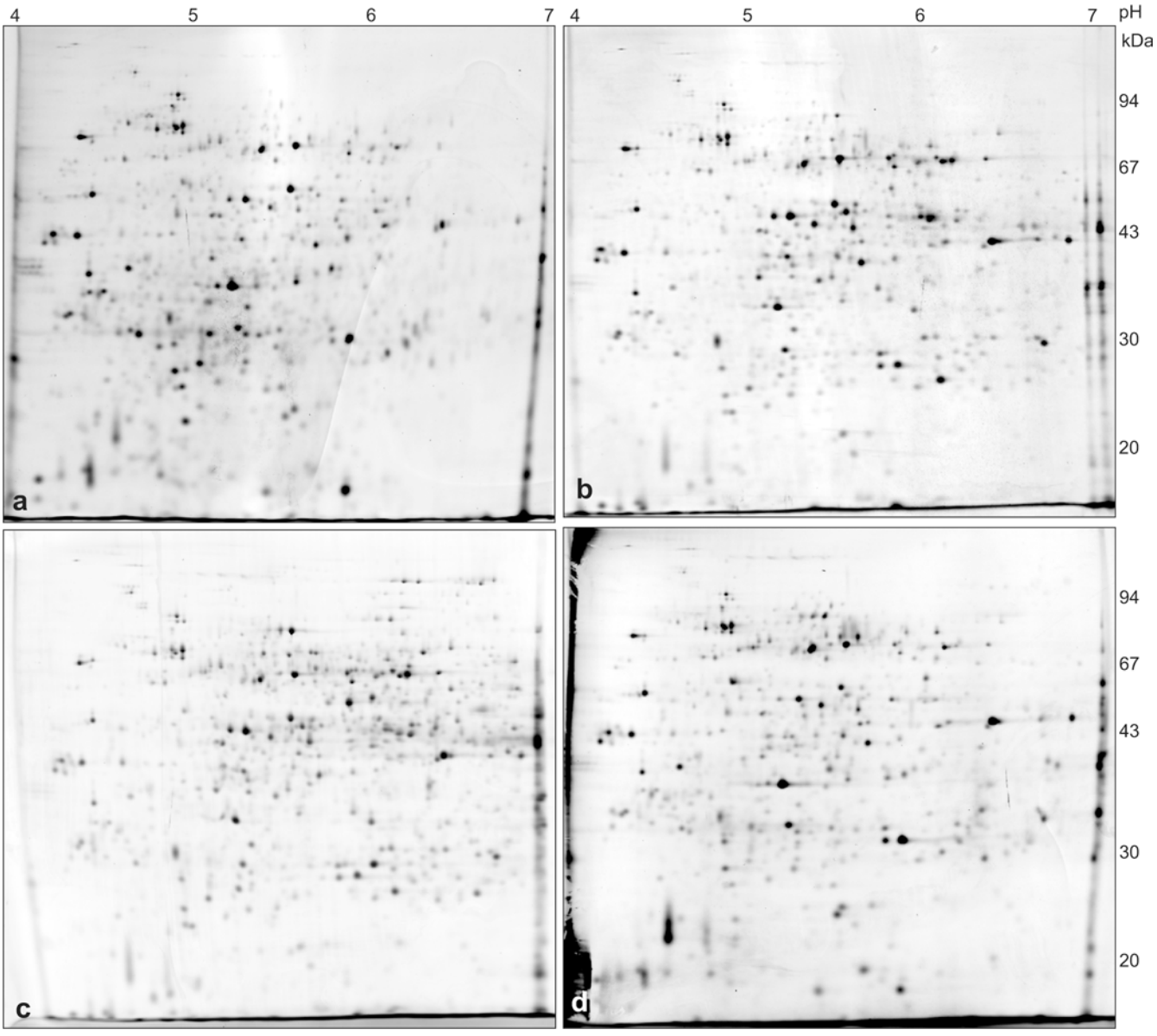

Fig. 3 The 2D-gels of proteins from *Catharanthus roseus* cell line A11 harvested on day 3 (**a**), 6 (**b**), 9 (**c**) and 13 (**d**)

spots. In the first selection (S1), protein spots were selected that (1) were detected on 2D-gels of day 6; (2) had a normalized intensity on day 6 of at least twice the intensity on day 3; (3) and did not decrease more than twice on day 9 and day 13. In the second selection (S2), protein spots were selected that (1) were detected on 2D-gels of day 6; (2) had a normalized intensity on day 6 of at least twice the intensity on day 3; (3) were detected on 2D-gels of day 9 and day 13 (although they could be very small) and (4) did not belong to S1. Only protein spots detected on at least four out of five gels were considered in these selections. Selection S1 resulted in a total of 56 spots and selection S2 resulted in an additional 55 protein spots. These protein spots are indicated in Fig. 4.

It is most likely that proteins involved in alkaloid biosynthesis are found in selection S1, because on day 13, the amount of alkaloids in the cell-suspension culture is still increasing, thus the biosynthetic proteins should still be present. However, it cannot be excluded that only the activity of one or a few enzymes is needed at this stage and that other proteins involved in alkaloid biosynthesis have disappeared already and thus belong to selection S2.

In total, 88 protein spots were selected for identification by mass spectrometry. From the above-mentioned selections S1 and S2, the 18 and 15 protein spots with the largest intensities were selected. Also 25 protein spots that were strongly induced on day 9 were selected (selection T); because on this day, many changes occur in the cell culture on protein level (Fig. 3). Moreover, 30 protein spots were selected that were always detected on 2D-gels of *C. roseus* in quite large amounts (selection O).

Table 1 Number of protein spots detected and matched on day 3, 6, 9, and 13 of the growth cycle of *Catharanthus roseus* cell line A11

Days after inoculation	Average number of spots on a gel[a]	Average number of matches[a]
3	560 (±16)	365 (65%)
6	660 (±51)	550 (83%)
9	905 (±61)	860 (95%)
13	410 (±26)	340 (83%)

[a]Average of five gels. The gels are matched against the reference gel of day 9 containing the most spots (988). The values for the standard deviation and the percentages of matched spots are given within brackets

Protein identification

The term ''protein identification'' will be used in the text because of its commonly accepted meaning, even though it would be more appropriate to speak about ''gene product identification'', being this what is actually achieved by MS and MS/MS analysis of gel-separated proteins (Rappsilber and Mann 2002).

Due to the rather poor genome and protein sequence information currently available on *C. roseus*, identification relied almost exclusively on homology to proteins belonging to other plant species. Initially, the efficacy of MALDI-TOF peptide fingerprinting for identification was tested on a number of protein spots. Because proteins of interest were likely not to be present in the database, simple peptide mass fingerprinting required a highly homologous protein from another species to be available in the database. Furthermore, it has already

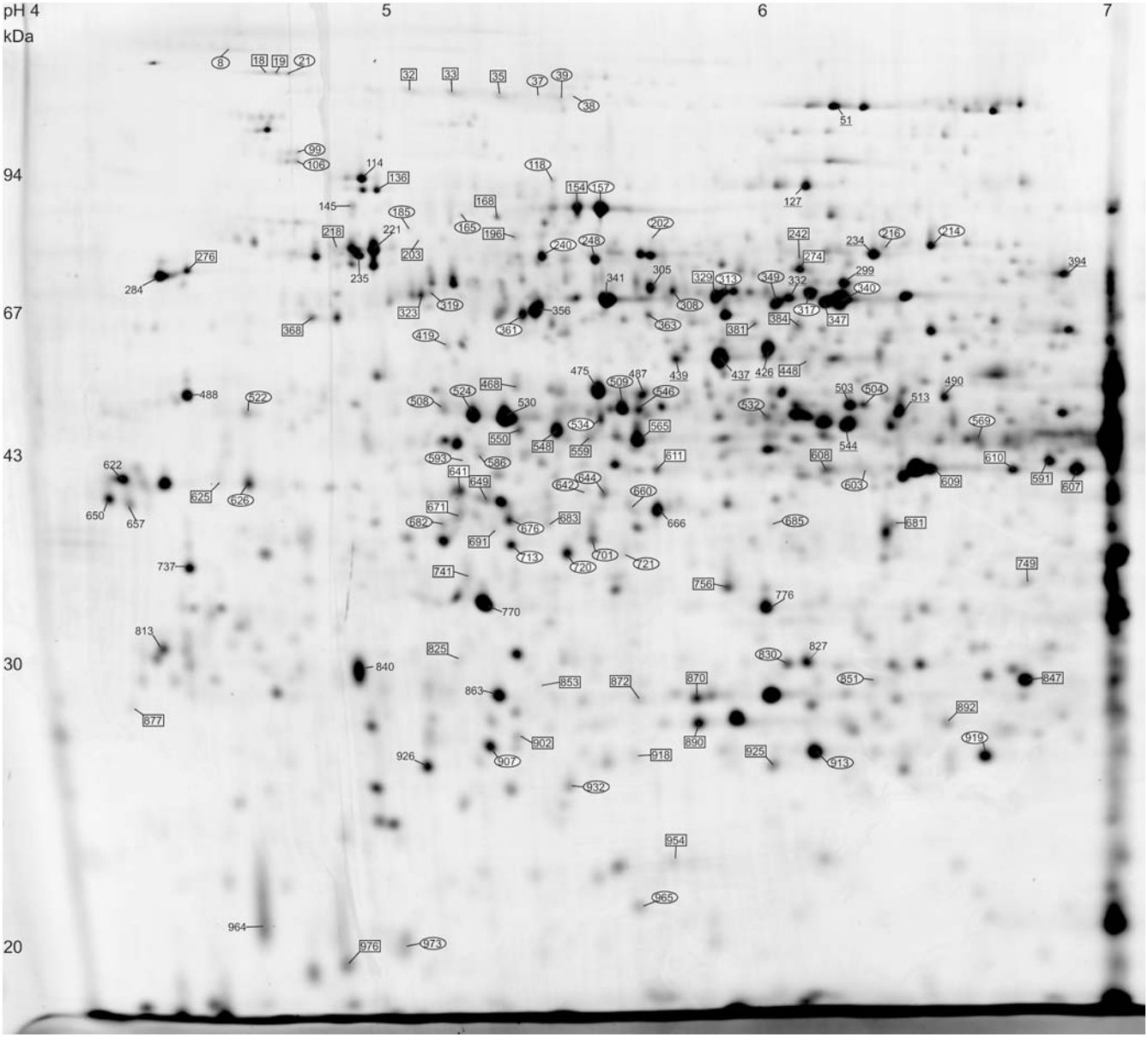

Fig. 4 Reference 2D-gel of *Catharanthus roseus* cell line A11 (day 9 of the growth cycle). Indicated are the protein spots belonging to selection S1 (*square boxes*) and S2 (*oval boxes*), of which the *white*

Table 2 Identified protiens from 2D-gels of *Catharanthus roseus*

Spot no.	Selection	Protein name	Accession no.	M_w (exp)[a]	pI(exp)	M_w(theor)[a]	pI(theor)	Organism	Matched sequences[b])
51	T	Aconitase	gi-15233349Q42560	102.3	6.22	98.2	5.98	*Arabidopsis thaliana*	ILLESAIR, DFNSYGSR, IDKLPYSIR, QVEIPFKPAR, FYSLPALNDPR
114	O	HSP70 (DnaK)	gi-7441880 O50036	81.9	4.92	76.1	5.19	*Spinacia oleracea* (spinach)	IVDWLAASFKR, FAAEEISAQVLR, QFAAEEISAQVLR, AVVTVPAYFNDSQR, LSFSDLDEVILVGGSTR
127	T	Transketolase 1	gi-7433617O78327	80.2	6.14	80.1	6.16	*Capsicum annuum* (bell pepper)	NPYWFNR, DGNTGYDEIR, VTTTIGFGSPNK, ALPTYTP1ESPADATR
145	O	HSP70	gi-7441856 Q39641	75.4	4.89	75.4	5.15	*Cucumis sativus* (cucumber)	IVDWLAASFKR, FAAEEISAQVLR, QFAAEEISAQVLR, AVVTVPAYFNDSQR, LSFSDLDEVILVGGSTR, QAVVNPENTFFSVK
157	S2	HSP68/HSP70 like	gi-300265/ gi-15235740	74.7	5.57	71.1	4.97	*Lycopersicon peruvianum* (peruvian tomato)/*Arabidopsis thaliana*	IAGLDVQR, HLNISLTR, VQEVVSEVFGK, EVDEVLLDGGMTR, AVITVPAYFNDAQR, VNPDEAVAMGAALQGGILR
214	S2	Dihydroxyacid dehydratase	gi-15229546Q9LIR4	67.5	6.49	64.9	5.85	*Arabidopsis thaliana*	SVGLELTLDDFQK
221	O	Disulfide isomerase precursor	gi-11133775 Q43116/gi-15223975	66.7	4.95	53.0	4.91	*Ricinus communis* (castor bean)/ *Arabidopsis thaliana*	IQGFPTLK, SEPIPEVNNEPVK, TKEDIISFIEK
234	T	Phosphoglycerate dehydrogenase-like protein	gi-15235282 O49485	65.6	6.33	63.3	6.16	*Arabidopsis thaliana*	ILNDETFAK, ISLCDALIVR, LAVQLVAGGSGVK, AIGVELVGFEVAISTA, VASILGDENVNVSFMSVGR
235	O	Chaperone 60 α-subunit	gi-1351030 P21239	65.4	4.91	57.0	4.78	*Brassica napus* (rape)	*APGFGER, GVLNVVAVK, DIIPILEK, LLVEFENAR, LADAVGLTLGPR, GYISPQFVTNPEK, ELSETDSVYDSEK, TNDSAGDGTTTASVLAR, AIELPDAMENAGAALIR*
240	S2	Chaperone 60 β-subunit	gi-2506277P08927	65.2	5.41	62.9	5.85	*Pisum sativum* (pea)	APGFGER,VVEGMQFDR, LADLVGVTLGPK, IAALKAPGFGER, DLVNILEDAIR, VVAAGANPVLITR, GYISPYFVTDSEK, EVELEDPVENIGAK, AAVEEGIVVGGGCTLLR, DTTTIVGDGSTQEAVNKR, TNDLAGDGTTTSVVLAQGLIAEGVK

274	S1	Probable methylmalonate semialdehyde dehydrogenase	gi-7431455	63.0	6.13	56.9	5.99	*Oryza sativa* (rice)	AGVQFYTQIK,ASFAGDLNFYGK
276	S1	26S proteasome regulatory subunit S5A	gi-7484606	62.7	4.43	42.8	4.50	*Mesembryanthenum crystallinum* (common iceplant)	IIVFAGSPVK
284	O	Calreticulin	gi-11131745 P93508	61.7	4.36	45.5	4.36	*Ricinus communis*	FEDGWENR, TLVFQFSVK, YVGIELWQVK, GSLYSDWDLLPPK, FYAISAEFPEFSNK
299	T	Mitochondrial processing peptidase, β-chain	gi-12802327Q9AXQ2	60.6	6.25	58.8	6.56	*Cucumis melo* (oriental melon)	EQTTYYAK, EPAIFTGSEVR
313	S2	Alanine aminotransferase, putative	gi-5892836EST277651	59.2	5.92	58.5[c]	5.62[c]	*Lycopersicon esculentum* (tomato)	SETQGLFSADAIER, ALVVINPGNPTGQVLGEANQR
317	S2	ATP synthase α-chain	gi-6561638 Q9T775	58.9	6.15	45.3	7.82	*Euptelea polyandra*	LELAQYR, AVDSLVPIGR, VVDALGVPIDGR
323	S1	Tubulin β-chain	gi-100551P258319	58.6	5.08	43.3	4.60	*Avena sativa* (oat)	LAVNLIPFPR, YLTASAVFR, VSEQFTAMFR, *IREEYP, MMLTFSVFPSPK, ALIDSCFRTLK, FPGQLNSDLRK, KLAVNLIPFPR, EVDEQMINVQNK,*
329	S1	Aminopeptidase	gi-2492530 Q42876	58.3	5.89	54.2	6.33	*Lycopersicon esculentum*	FQNPLLQK, PGAAITAALFLK, LTFDSGGYNIK, GLTFDSGGYNIK
332	T	Enolase	gi-1169534P42896	58.2	6.09	47.9	5.56	*Ricinus communis*	VNQIGSVTESIEAVR, AAVPSGASTGIYEALELR
340	S2	Enolase	gi-1169534P42896	57.8	6.24	47.9	5.56	*Ricinus communis*	YNQLLR, AGWGVMASHR, YDLNFKEEK, AGALVKGIPLYK, VNQIGWVTESIEAVR, AAVPSGASTGIYEALELR, EEELGAEAVYAGAKFR
341	O	ATP synthase β-chain	P17614	57.7	5.59	54.1	5.13	*Nicotiana plumbaginifolia* (curled-leaved tobacco)	*VVDLLAPYQR, GGKIGLFGGAGVGK, AHGGFSVFAGVGER, VLNTGSPITVPVGR, FTQANSEVSALLGR, LVLEVAQHLGENVVR, CALVYGQMNEPPGAR, QISELGIYPAVDPLDSTSR, EAPAFVEQATEQQILVTGIK, IPSAVGYQPTLALTLGGLQER*
347	S1	Aldehyde dehydrogenase (NAD+)-like protein	gi-15228319Q9SU63	57.3	6.19	58.6	7.11	*Arabidopsis thaliana*	EEIFGPVQSILK, FADLVEK, TFPTLDPR
356	O	Mitochondrial processing peptidase, α II chain precursor	gi-2117355 P29677	56.4	5.40	54.7	5.71	*Solanum tuberosum* (potato)	SAILMNLESR, EPMIYTYDALK, NPAFLDWEVIEQLEK
361	S2	α-tubulin	gi-11967906Q9FRU1	55.5	5.36	49.8	4.94	*Nicotiana tabacum* (tobacco)	LISQVISSLTASLR, SLDIERPTYTNLNR, AVFVDLEPTVIDEVR
394	T	Tryptophan synthase β-chain	Q9FFW8	53.3	6.49	55.7	6.79	*Arabidopsis thaliana*	DLPEEVLDVYK, FGQQMER

Table 2 (Contd.)

Spot no.	Selection	Protein name	Accession no.	M_w (exp)[a]	pI(exp)	M_w(theor)[a]	pI(theor)	Organism	Matched sequences[b])
426	T	S-Adenosyl methionine synthase	gi-14600072 Q94FA4	50.9	6.04	43.2	5.67	*Brassica juncea* (indian mustard)	SGAYIVR, TIFHLNPSGR, FVIGGPHGDAGLTGR, TNMVMVFGEITTK
437	T	S-Adenosyl methionine synthase	gi-3024127 Q96552	49.5	5.90	43.0	5.51	*Catharanthus roseus*	SGAYIVR, SIVASGLAR, IPDKEILKTAAYGHFGR, TIFHLNPSGR, FVIGGPHGDAGLTGR, TQVTVEYYNDNGAMVPIR
439	T	S-Adenosyl methionine synthase	gi-14600072 Q94FA4	49.5	5.78	43.2	5.67	*Brassica juncea*	SGAYIVR, IPDKEILK, TAAYGHFGR, TIFHLNPSGR, FVIGGPHGDAGLTGR, TNMVMVFGEITTK
475	O	Actin 104	gi-3219764 P93375	45.7	5.57	37.1	5.57	*Nicotiana tabacum*	*DLYGNIVLSGGSTMFPGIADR, GYMFTTSAEREIVR, VAPEEHPVLLTEAPLNPK, VVAPPERK, AGFAGDDAPR, IKVVAPPER, IKVVAPPERK, GYMFTTSAER, AVFPSIVGRPR, DAYVGDEAQSKR, IWHHTFYNELR*
487	O	Actin	gi-9965319	45.3	5.69	41.7	5.31	*Setaria italica* (foxtail millet)	AGFAGDDAPR, GYSFTTTAER, AVFPSIVGRPR, SYELPDGQVI TIGAER
488	O	Ankyrin-repeat protein HBP-1	gi-13310811	45.1	4.43	37.3	*4.45*	*Nicotiana tabacum*	NTALHYAAGYGR, YWNDQETPKK, TALHFACGYGEVK
490	T	Aspartate transaminase	gi-629503 P46248	45.1	6.53	44.2	6.98	*Arabidopsis thaliana*	LAAALIER, VPWSEYR, VATIQGLSGTGSLR
503	T	Probable 12-oxophytodienoate reductase	gi-7449901 O23777	44.1	6.26	41.9	5.90	*Catharanthus roseus*	FQLSHR, GLTPLYR, ADLVVYGR, LFLANPDLPK, RFELNAPLNK, GTFMAAGGYDREDGIK
513	T	Alcohol dehydrogenase	gi-113379 P14674	43.7	6.40	41.1	5.92	*Solanum tuberosum*	GQNPVFPR, FGVTEFVNPK, GTFFGNYKPR, THPMNFLNER, GSSVAIFGLGGVGLAAAEGAR, IIGVDLNAPR
534	S2	Rubisco large chain	gi-1173916Q37173	42.7	5.57	49.0	6.15	*Cyanastrum cordifolium*	DTDILAAFR,DLAKEGNEIIR, *IIGFDNVR, EVTLGFVDILR*
544	T	Glutamine synthase pr-2	gi-121345 P04771	42.1	6.26	39.2	5.40	*Phaseolus vulgaris* (kidney bean)	YILER, GYFEDR, IFKDPFR, SPLSDLINLNSESTEK
565	S1	Aspartate semialdehyde dehydrogenase	gi-20856224	40.6	5.67	40.7	6.52	*Arabidopsis thaliana*	VDNSSAFR, IDDR
569	S2	Fructose 1,6-bisphosphate aldolase (fragment)	gi-14484932Q94FU4	40.5	6.62	38.6[c]	6.48[c]	*Lycopersicon esculentum*	AQAAFLTR, VAPEVIAEYTVR
609	S1	UM53TE05	gi-24072217EST14383740	38.0	6.49			*Populus balsamifera* (western balsam poplar)	GTPFDFTTEKK
610	S1	Putative malate dehydrogenase	gi-15219721P93819	38.0	6.72	35.6	6.11	*Arabidopsis thaliana*	ALGQISER,EFAPSIPAK, VLVVANPANTNALILK, *VLVTGAAGQIGYALVPMIAR, MELVDAAFPLLK*

611	S1	NAD-malate dehydrogenase	gi-5123836Q9XQP4	38.0	5.73	33.9	5.56	*Nicotiana tabacum*	ANTFVAQKK, ALDALKPELK, LFGVTTLDVVR, IQNAGTEVVEAK, KLFGVTTLDVVR, DDLFNINANIVK
622	O	Cytochroom c oxidase subunit 6b	gi-15219886 Q9S7L9	37.2	4.42	21.2	4.29	*Arabidopsis thaliana*	YVEYHR, FPTTNQTR, LETAPADFR
641	S1	Putative fructokinase	gi-15221364	36.3	5.18	37.6	5.51	*Arabidopsis thaliana*	TALAFVTLR, APGGAPANVAIAVSR
650	O	Cysteine protease	gi-11994462	35.7	4.20	40.3	4.77	*Arabidopsis thaliana*	VVTIDGYEDVPR
657	O	Cysteine protease	gi-11994462	35.3	4.26	40.3	4.77	*Arabidopsis thaliana*	VVTIDGYEDVPR
666	O	Cysteine synthase	gi-3127890	34.9	5.73	27.9	5.89	*Cicer arietinum* (chickpea)	RIILR, YLSSVLFESVR, LIGVVFPSFVER, AFGAELVLTDPAK, YLSSVLFESVRK
671	S1	Strictosidine synthase	gi-18222P18417	34.5	5.18	35.7	5.12	*Catharanthus roseus*	GFYTSVQDGR, TYDISYDYK
676	S2	Strictosidine synthase	gi-18222P18417	34.4	5.32	35.7	5.12	*Catharanthus roseus*	GFYTSVQDGR, WLYAVTVDQR
713	S2	20S proteasome α-6 subunit	gi-22947842	32.5	5.32	29.9	5.07	*Nicotiana benthamiana*	DALFAIR, QGSAAIGLR, ENLLKDALFAIR, *LFQVEYAMEAVK, THVILACVNK, VDDHIGVAIAGLTADGR*
737	O	Elongation factor 1-β	gi-6166140 Q40680	31.1	4.42	24.7	4.36	*Oryza sativa*	LVPVGYGIK, LDEYLLTR, KLDEYLLTR
770	O	Oxygen evolving enhancer protein	gi-131384 P14226	29.1	5.23	26.5	5.08	*Pisum sativum*	*TYLEVK, GSSFLDPK, VPFLFTIK, LTFDEIQSK, RLTFDEIQSK, IQGVWYAQLES, LCLEPTSFTVK, KLCLEPTSFTVK, DGIDYAAVTVQLPGGER*
813	O	Nascent polypeptide associated complex α-chain	gi-6752882 Q9M612	27.0	4.32	22.4	4.32	*Pinus taeda* (loblolly pine)	IEDLSSQLQTQAAEQFK, NILFVQSNPDVFK
827	T	Acetone-cyanohydrin lyase (putative)	Q8S8S9	27.2	6.17	29.8	4.97	*Arabidopsis thaliana*	FTEEGFGSVPR, VNLAASGL
863	O	Triose phosphate isomerase	gi-1174745 P46225	25.1	5.26	27.3	5.01	*Secale cereale* (rye)	FFVGGNWK, QEDIDGFLVGGASLK, IIYGGSVDGGNCAELAK
907	S2	Ascorbate peroxidase	gi-804973Q42661	23.5	5.22	27.4	5.32	*Capsicum annuum*	TGGPFGTMR, EGLLQLPSDK, ALLSDPAFRPLVEK
919	S2	Manganese superoxide dismutase (fragment)	gi-11967889 Q9G2P0	23.2	6.63	23.0[c]	5.89[c]	*Prunus persica*(peach)	NVRPDYLK, LVVETTANQDPLVTK
926	O	Ascorbate peroxidase	gi-804973Q42661	22.8	5.04	27.4	5.32	*Capsicum annuum*	TGGPFGTMR, ALLSDPAFRPLVEK
964	O	H-protein	gi-737595 P16048	19.8	4.54	13.9	4.65	*Pisum sativum*	YAPSHEWVK, VKPSSPAELESLMGPK

[a] (exp) represents the experimental value and (theor) the theoretical value calculated with th Calculate pI/M_w tool on the ExPASy server

[b] Sequences in bold are matched using Mascot MS/MS Ion Search, sequences in italic are matched using Mascot Peptide Mass Fingerprint, underlined sequences are obtained by de novo sequencing

[c] Theoretical M_w and pI are based on the nearest homologue found using BLAST at NCBI. 313: gi-14334825, *Arabidopsis thaliana*; 569: gi-10800924, *Persea americana* (avocado); 919: gi-464775, *Hevea brasiliensis*

been demonstrated that high sequence homology to a database entry might not be sufficient to guarantee identification, because homology has to remain conserved at the tryptic digest level (Clauser et al. 1999). In fact, as expected, results were successful just in a limited number of cases.

Evaluation of MALDI-TOF analyses demonstrated that MS/MS sequence information from analyzed tryptic peptides was necessary to analyze the *C. roseus* protein spots. In fact, MS/MS ion search raised the percentage of identification considerably. The orthogonal acceleration MALDI quadrupole time-of-flight instrument provided excellent sensitivity and both MS fingerprinting and MS/MS data by performing one simple experiment. The only drawback was the fact that MS/MS had to be performed on singly charged ions (because sample ionization was achieved by MALDI), which is generally less informative than MS/MS performed on doubly charged ions. Nevertheless, good quality MS/MS data were obtained for the large majority of proteins analyzed and, when needed, partial or full de novo sequencing could be achieved for many fragmented peptides.

The MALDI-MS/MS data acquired on the QSTAR MS instrument were used for searching databases adopting a sequential strategy. The first step was to use only the 3–10 peptides for which MS/MS information was available, because they represented the most reliable and informative part of the total MS data acquired. Database search was performed using the MS/MS ion search available on the MASCOT search engine, querying against all Viridiplantae entries in the NCBInr (non-redundant) database. If a significant score was obtained for one or more peptides, error tolerant search was performed on the sequence of the identified protein in order to look for additional homologous peptides containing a modification, e.g., an amino acid substitution. If MS/MS ion search against NCBInr failed, data were re-searched against all Viridiplantae entries of the pdbEST-others database. When the second MS/MS ion search also failed, or when the score of MS/MS identification was at the borderline significance/random event, internally calibrated peptide mass fingerprinting data were also used as an alternative strategy or as additional confirmation. When all database-related strategies failed, de novo sequencing was performed on fragmented peptides in order to perform homology searches. Only MS/MS data providing sequence strings of at least six amino acids were considered for the purpose.

By using the database searches, 56 protein spots could be identified (Table 2, Fig. 4). The amino acid sequences obtained by de novo sequencing of seven additional protein spots are given in Table 3. The identification of the remaining spots failed in eight cases because of poor MS data, while in 17 cases the database search failed and the obtained MS data were not sufficient for de novo sequencing.

The amino acid sequences obtained by de novo sequencing were used for searching the non-redundant database from EMBL with the MSBLAST search engine at the EMBL server (http://dove.embl-heidelberg.de/Blast2/msblast.html; Shevchenko et al. 2001). Most of the sequences did not show any homology with a protein from the databases. However, two additional protein spots were identified. Spot 394 was identified as tryptophan synthase β-chain, as the two sequences were largely homologue and a significant identification score was obtained (Table 4), and spot 827 was identified as acetone-cyanohydrin lyase, also based on high homology of two sequences. The third sequence of spot 827 did not result in a significant hit. Of spot 305, the sequence VFASPLAR is homologue to a putative dihydrolipoamide acetyltransferase from *Arabidposis thaliana*. However, this is not sufficient for a significant identification score. The sequences obtained from the other spots are unique sequences, as they did not show any homology with proteins from the databases. Of particular interest are the isoforms 524 and 530. These are large protein spots on the 2D-gels and gave very good MS/MS spectra and long sequence strings, but absolutely no significant hits were obtained. These or some of these sequences showing no significant homology, may be interesting with respect to the alkaloid biosynthesis. To clarify the function of these proteins, however, the amino acid sequences will be needed to generate PCR primers for the isolation of the coding genes.

Discussion

The two isoforms of strictosidine synthase (671(S1), 676(S2)) are among the most interesting identified proteins from this study, since their involvement in the alkaloid biosynthesis. Strictosidine synthase catalyzes the condensation of tryptamine and secologanin to form strictosidine (Fig. 1), the common precursor of all terpenoid indole alkaloids. Strictosidine synthase is a glycosylated enzyme, of which six isoforms have been detected in *C. roseus* (De Waal et al. 1995). The identified isoforms meet the selection criteria S1 and S2 and thus exhibit a pattern similar to that of alkaloid accumulation in the cell culture (Fig. 5). Especially the increase between day 3 and day 6 is evident. The identification of strictosidine synthase shows that proteome analysis, by a combination of 2D-PAGE and mass spectrometry, is capable of identifying proteins from secondary metabolism in plants. Furthermore, it proves that the experimental setup and the selection of the criteria for S1 and S2 were correctly chosen. In total, approximately 10% of the protein spots detected on the 2D-gels of *C. roseus* cell line A11 meet these criteria. So far, 24 of them could be identified. It is assumed that by proteome analysis (2D-PAGE combined with mass spectrometry), only high abundant proteins can be identified. It is also assumed that proteins involved in secondary metabolism are low abundant proteins. As

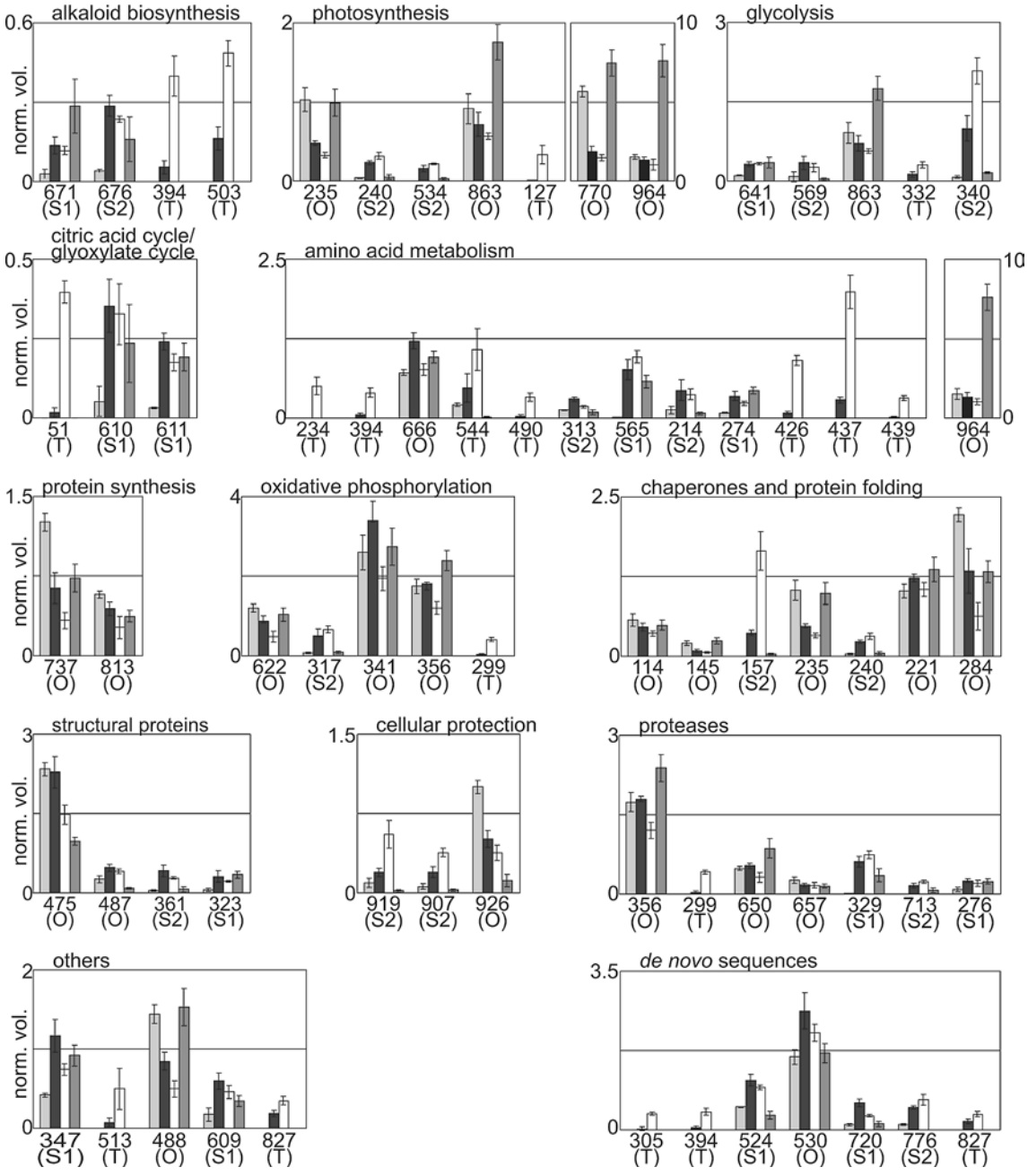

Fig. 5 Expression profiles of the identified proteins from *Catharanthus roseus* grouped in different cellular functions. Normalized spot intensities of day 3 (*light gray*), day 6 (*black*), day 9 (*white*) and day 13 (*dark gray*)

shown by our results, at least one of these assumptions is not correct for strictosidine synthase.

Another protein that is important for the alkaloid biosynthesis is tryptophan synthase (394(T)). Alkaloid producing cells need a sufficient supply of tryptophan, as tryptamine is one of the building blocks. In the *C. roseus* cell line A11, tryptamine used to be the limiting precursor for the biosynthesis of alkaloids (Contin et al. 1998). However, a change in behavior of the cell line was observed as alkaloids accumulated instead of secologanin. Tryptophan synthase is only detected during a short period, in a small quantity on day 6 and in a large quantity on day 9. For the synthesis of tryptophan, serine is required (Ho and Saito 2001). Phosphoglycerate dehydrogenase (234(T)), which could only be detected on day 9, is involved in serine biosynthesis via the 3-phosphoglycerate pathway. It has been proposed that this pathway is of minor significance compared to the glycolate pathway during photosynthetic activity (Ho and Saito 2001). As a result of the oxygenase activity of Rubisco (photorespiration), phosphoglycolate and 3-phosphoglycerate are formed. Phosphoglycolate is converted via the glycolate pathway, of which glyoxylate and glycine are intermediates. H-protein (964(O)) is a

subunit of the glycine decarboxylation complex converting glycine to serine (Douce et al. 2001). H-protein was detected on every day of the growth cycle in a relatively large amount. Probably, on day 9, extra serine was required or there was a decrease in photorespiration, which resulted in an induction of phosphoglycerate dehydrogenase (234(T)). Whether this serine was needed for tryptophan synthesis, cannot be directly deduced. Many of the identified proteins involved in amino acid biosynthesis are strongly induced on day 9, after which they disappear again (selection T). This period is just after the onset of alkaloid accumulation.

Another interesting protein in relation to alkaloid biosynthesis is 12-oxophytodienoate reductase (503(T)), which is involved in jasmonate biosynthesis. Jasmonic acid is essential for signal transduction. Its involvement in stress reactions was demonstrated by induction of defense proteins including proteins involved in the biosynthesis of secondary metabolites. In *C. roseus*, it also induces the expression of genes involved in alkaloid biosynthesis (Memelink et al. 2001). 12-oxophytodienoate reductase is only detected during a short period, in a small quantity on day 6 and in a large quantity on day 9. Since the enzyme might be indirectly involved in the alkaloid biosynthesis, its expression profile may be out-of-phase to that of the biosynthetic proteins.

Most of the proteins identified in this study are housekeeping proteins involved in common cellular processes, e.g., glycolysis, oxidative phosphorylation, protein synthesis and folding. The identified glycolytic enzymes belong to the selections S1 and S2. One of the intermediates of the glycolysis is glyceraldehyde 3-phosphate, which is also an intermediate of many other processes including photosynthesis and the biosynthesis of isopentenyl diphosphate via the MEP-pathway (Rohmer 1999). The alkaloid precursor secologanin is also derived from this pathway (Contin et al 1998). The isomerization of the three-carbon phosphorylated sugars dihydroxyacetone phosphate and glyceraldehyde 3-phosphate is catalyzed by triose phosphate isomerase (863(O)). The amounts detected of this enzyme during the growth cycle are quite constant, except for day 13 when it is increased. Two protein spots of malate dehydrogenase (610, 611(S1)) from the citric acid and glyoxylate cycle also meet the selection criteria of S1. An unexpected result is that aconitase (51(T)) is only detected on day 9 in a large amount. Possibly another isoform or isoenzyme is present in the cells. The two identified subunits from the 26S proteasome (713, 276) belong to the selections S1 and S2 as well. This might indicate that during growth and alkaloid accumulation of the cell culture, there is also an increased protein turnover. However, in contrast, the amounts detected of two cysteine proteases (650, 657(O)) are quite constant.

The identified proteins from the oxidative phosphorylation are detected in rather constant amounts during the growth cycle, as are the proteins involved in protein synthesis and the structural proteins. The proteins involved in protein folding show much variation in behavior, some are quite constant, but others vary considerably in their amounts detected during the different time points of the growth cycle. The two ascorbate peroxidases (907(S2), 926(O)), concerned with cellular protection against oxidative stress, show opposite profiles in amounts detected on the 2D-gels. Seven proteins participating in photosynthesis are identified as well. The presence of these proteins implies that this cell line, which has a green color, could be photosynthetically active. However, it is not essential for plant cell cultures to have photosynthetic activity as long as sugars are available from the medium.

Of many proteins identified in this study, more unidentified isoforms or isoenzymes can be present on the gel and in the cells. Therefore, caution has to be taken with the interpretation of the expression profiles. This is reflected by the expression profiles of different identified subunits from the same protein complex. One could expect that these subunits would show the same behavior during the growth cycle. However, for the subunits from tubulin and mitochondrial protein peptidase this is not the case. The amount of β-tubulin (323(S1)) detected on day 13 is much higher than the amount of α-tubulin (361(S2)); however, in the functional protein they occur in a fixed 1:1 ratio. The β-chain of mitochondrial protein peptidase (299(T)) is only detected in a large amount on day 9, while the αII subunit (356(O)) is constantly present in a large amount. Yet, a similar behavior of the two subunits would be expected. This means that probably more unidentified isoforms of the subunits are present. This could be the case for other proteins as well.

The main limitation for the identification of proteins from *C. roseus* by mass spectrometric analysis is the lack of sequence data of both, genes and proteins, from *C. roseus* and from plants in general. For example, the SWISS-PROT database (November 2004) contains only 27 protein entries from *C. roseus*. It is thus necessary for identification of proteins from 2D-gels to obtain sequence data and not rely solely on peptide masses. Such analyses can be done by fragmentation of peptides measured by mass spectrometry or by Edman degradation, as is used for the identification of protein spots

Table 3 Protein sequences obtained by de novo sequencing using MS/MS data of protein spots from *Catharanthus roseus*

Spot no.	Sequences[a]
305 (T)	**WNDLLVLK**, VFASPLAR, **WVEEFLASR**, **AVPYPQR**, **PLQYVLY**
394 (T)	TM(MT)FGQQMER, DLPEEVLDVYK
524 (S1), 530 (O)	**NDLHLS**, **DDVEGYGTSR**, **NSNCDNQD**, **QLDYNVTYGK**, **NQHLSGLTNLLK**
720 (S1)	VVNAVYTR
776 (S2)	**TLPLYLR**, **FY(YF)KPQSK**
827 (T)	**(DV, VD, LT, TL)PEPFQR**, FTEEGFGSVPR, VNLAASGL

[a]Note that by MS/MS L and I cannot be distinguished. Unique sequences are given in bold

Table 4 Homology of sequenced peptides of protein spots from *Catharanthus roseus* resulting from MSBLAST search at the EMBL server

Spot no./protein name	Sequences showing homology	
305	VFASPLAR	
	VFASPLAR	
Dihydrolipoamide acetyltransferase (putative) (Q9SLL0, *Arabidopsis thaliana*)	224 VFASPLAR 231	
394	DLPEEVLDVYK	FGQQMER
	D + PEEVL + + YK	QQME
Tryptophan synthase β-chain (Q9FFW8, *A. thaliana*)	119 DIPEEVLEIYK 129	299 CIQQMEN 304
827	FTEEGFGSVPR	VNLAASGL
	F + EG + GSVPR	V + LAASG +
Acetone–cyanohydrin lyase (putative) (Q8S8S9, *A. thaliana*)	195 FSDEGYGSVPR 205	40 VDLAASGI 47

from opium poppy (Decker et al. 2000). They are needed for *C. roseus* as a model plant for studies of the secondary metabolism.

In total, 88 protein spots from *C. roseus* were subjected to mass spectrometric analysis for identification. Of these, 56 could be identified by searching the databases with the obtained m/z-values. Of eight additional protein spots, amino acid sequence strings were obtained by de novo sequencing, resulting in two extra identifications using MSBLAST search. The other sequences did not show any homology with proteins present in the databases, being therefore of particular interest. These proteins are candidates for further investigations concerning their function. The fact that we identified other proteins that possibly might be involved in alkaloid biosynthesis, their identification will be important. De novo sequencing with a large sequence coverage may be helpful to obtain enough data for homology searching, gene cloning and, eventually, metabolic engineering.

Note added in proof: The protein spot 827 (see Tables 3, 4) could be identified as Protein S, a protein from C. roseus associated with monoterpenoid biosynthesis, recently cloned by Lemenager et al. (entry name Q5XLS1). The second and third peptide, as mentioned in Table 3, showed complete identity.

Acknowledgements The authors would like to thank the Van Leersumfonds (the Netherlands) to make the purchase of an IPGphor and MALDI-sample plates possible.

References

Clauser KR, Baker P, Burlingame AL (1999) Role of accurate mass measurement (+/− 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. Anal Chem 71:2871–2882

Contin A, van der Heijden R, Lefeber AWM, Verpoorte R (1998) The iridoid glucoside secologanin is derived from the novel triose phosphate/pyruvate pathway in a *Catharanthus roseus* cell culture. FEBS Lett 434:413–416

Contin A, van der Heijden R, ten Hoopen HJG, Verpoorte R (1999) The inoculum size triggers tryptamine or secologanin biosynthesis in *Catharanthus roseus* cell cultures. Plant Sci 139:205–211

Dagnino D, Schripsema J, Verpoorte R (1996) Analysis of several iridoid and indole precursors of terpenoid indole alkaloids with a single HPLC run. Planta Med 62:278–280

De Waal A, Meijer AH, Verpoorte R (1995) Strictosidine synthase from *Catharanthus roseus*: purification and characterization of multiple isoforms. Biochem J 306:571–580

Decker G, Wanner G, Zenk MH, Lottspeich F (2000) Characterization of proteins in the latex of of the opium poppy (*Papaver somniferum*) using two-dimensional electrophoresis and microsequencing. Electrophoresis 21:3500–3516

Douce R, Bourguignon J, Neuburger M, Rébeillé R (2001) The glycine decarboxylase system: a fascinating complex. Trends Plant Sci 6:167–176

Gamborg OL, Miller RA, Ojima K (1968) Nutrient requirements of suspension cultures of soybean root cells. Exp Cell Res 50:151–158

Görg A, Obermaier C, Boguth G, Harder A, Scheibe B, Wildgruber R, Weiss W (2000) The current state of two-dimensional electrophoresis with immobilized pH gradients. Electrophoresis 21:1037–1053

Granier F (1988) Extraction of plant proteins for two-dimensional electrophoresis. Electrophoresis 9:712–718

Ho C-L, Saito K (2001) Molecular biology of the plastidic phosphorylated serine biosynthetic pathway in *Arabidopsis thaliana*. Amino Acids 20:243–259

Jacobs DI, van der Heijden R, Verpoorte R (2000) Proteomics in plant biotechnology and secondary metabolism research. Phytochem Anal 11:277–287

Jacobs DI, van Rijssen MS, van der Heijden R, Verpoorte R (2001) Sequential solubilization of proteins precipitated with trichloroacetic acid in acetone from cultured *Catharanthus roseus* cells yields 52% more spots after two-dimensional electrophoresis. Proteomics 1:1345–1350

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 277:680–685

Memelink J, Verpoorte R, Kijne JW (2001) ORCAnization of jasmonate-responsive gene expression in alkaloid metabolism. Trends Plant Sci 6:212–219

Moreno PRH, van der Heijden R, Verpoorte R (1993) Effect of terpenoid precursor feeding and elicitation on formation of indole alkaloids in cell suspension cultures of *Catharanthus roseus*. Plant Cell Rep 12:702–705

Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551–3567

Peterson G L (1977) A simplification of the protein assay method of Lowry et al which is more generally applicable. Anal Biochem 83:346–356

Rappsilber J, Mann M (2002) What does it mean to identify a protein in proteomics? Trends Biochem Sci 27:74–78

Rohmer M (1999) The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. Nat Prod Rep 16:565–574

Shevchenko A, Wilm M, Vorm O, Mann M (1996) Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Anal Chem 68:850–858

Shevchenko A, Sunyaev S, Loboda A, Shevchenko A, Bork P, Ens W, Standing K (2001) Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching. Anal Chem 73:1917–1926

Van der Fits L, Memelink J (2000) ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. Science 289:295–297

Van der Heijden R, Lamping PJ, Out PP, Wijnsma R, Verpoorte R (1987) High performance liquid chromatographic determination of indole alkaloids in a cell suspension culture of *Tabernaemontana divaricata*. J Chromatogr 396:287–295

Van der Heijden R, Jacobs DI, Snoeijer W, Hellerd D, Verpoorte R (2004) The *Catharanthus alkaloids*: pharmacognosy and biotechnology. Curr Med Chem 11:1241–1253

Verpoorte R, van der Heijden R, Moreno PRH (1997) Biosynthesis of terpenoid indole alkaloids in *Catharanthus roseus* cells. In: Cordell GA (ed) The alkaloids, vol 49. Academic, San Diego, pp 221–299

Verpoorte R, van der Heijden R, ten Hoopen HJG, Memelink J (1999) Metabolic engineering of plant secondary metabolite pathways for the production of fine chemicals. Biotechnol Lett 21:467–479

Whitmer S, Canel C, Hallard D, Goncalves C, Verpoorte R (1998) Influence of precursor availability on alkaloid accumulation by transgenic cell lines of *Catharanthus roseus*. Plant Physiol 116:853–857