Cloning and Characterization of a Candidate Gene from the Medicinal Plant *Catharanthus roseus* Through Transient Expression in Mesophyll Protoplasts

Patrícia Duarte et al.

*IBMC – Instituto de Biologia Molecular Celular, Universidade do Porto, Portugal*

1. Introduction

*Catharanthus roseus* (L.) G. Don. accumulates in the leaves the dimeric terpenoid indole alkaloids (TIAs) vinblastine and vincristine, which were the first natural anticancer products to be clinically used, and are still among the most valuable agents used in cancer chemotherapy. The great pharmacological importance of the TIAs, associated with the low abundance of the anticancer alkaloids in the plant, stimulated intense research on the TIA pathway, and *C. roseus* has now become one of the most extensively studied medicinal plants. About 130 TIAs have already been isolated from *C. roseus*, and it has been shown that the biosynthesis of vinblastine is highly complex, involving at least 30 steps from the amino-acid tryptophan and the monoterpenoid geraniol (Loyola-Vargas et al., 2007; van der Heijden et al., 2004). All the TIAs of *C. roseus* derive from the common precursor strictosidine, after which the TIA pathway splits into several branches including a short one leading to ajmalicine and serpentine (used as an antihypertensive and as a sedative respectively), and two long branches leading to vindoline and catharanthine - the leaf abundant monomeric precursors of vinblastine and vincristine (Loyola-Vargas et al., 2007; van der Heijden et al., 2004). In our lab, we have performed the characterization of a key biosynthetic step of the anticancer TIAs - the biosynthesis of the first dimeric TIA, α-3',4'-anhydrovinblastine (AVLB), from vindoline and catharanthine. We identified a leaf class III peroxidase (Prx) with AVLB synthase activity and we purified and characterized this enzyme, which was named *Catharanthus roseus* peroxidase 1, CroPrx1 (Bakalovic et al., 2006; Peroxibase, http://peroxibase.toulouse.inra.fr index.php). We have further shown the localization of CroPrx1 in the same subcellular compartment where alkaloids accumulate, the vacuole, through biochemical and molecular methodologies,
and we performed the cloning and molecular characterization of CroPrx1 (Costa et al., 2008; Duarte et al., 2010; Sottomayor et al., 1996; Sottomayor et al., 1998). In the course of this work, we have discovered, cloned and characterized another Prx gene from C. roseus, CroPrx3, which will be addressed in this chapter.

Class III peroxidases (Prxs; EC 1.11.1.7) are typical of plants and form large multigenic families. They are multifunctional enzymes that catalyze the oxidation of small molecules at the expense of H$_2$O$_2$, and are capable of recognizing a broad range of substrates. Prxs show an extraordinary diversity, with the presence of a high number of isoenzymes in a single plant, and have mainly been implicated in key processes determining the architecture and defense properties of the cell wall. They are also thought to play a role in biotic and abiotic stress resistance, in secondary metabolism, and in hydrogen peroxide scavenging and production (Ferreres et al., 2011; Fry, 2004; Passardi et al., 2005; Ros Barceló et al., 2004; Sottomayor et al., 2004). Regarding subcellular localization, Prxs are either vacuolar or extracellular. They are targeted to the secretory pathway by an N-terminal signal peptide (SP) that directs the protein to the endoplasmic reticulum (ER), from where they follow the default pathway to the cell wall, or they are sorted to the vacuole if an additional sorting signal exists (Costa et al., 2008). When compared to any characterized mature plant Prx, all vacuolar Prxs show the presence of a C-terminal extension (CTE) in their deduced amino-acid sequence (Welinder et al., 2002). Previous results obtained in our laboratory revealed that the CTE of CroPrx1 is both necessary and sufficient to target the sGFP reporter to the central vacuole, both in C. roseus bombarded cells and in Arabidopsis protoplasts, and must therefore constitute the Prx vacuolar sorting signal (Costa et al., 2008; Duarte et al., 2010). In contrast with the high number of studies on cell wall Prxs, much less is known about their vacuolar counterparts (Costa et al., 2008; Welinder et al., 2002).

On the other hand, although much is known about the C. roseus TIA pathway and its regulation, so far, enzyme/gene characterization is still lacking for many biosynthetic steps, no effective master switch of the pathway has been identified, and the membrane transport mechanisms of TIAs are basically uncharacterized (Sottomayor et al., 2004; van der Heijden et al., 2004; Verpoorte et al., 2007). Recently, omic approaches and in silico data-mining of EST libraries are being intensively used to enable the identification of candidate genes implicated in TIA metabolism (Carqueijeiro et al., 2010; Liscombe et al., 2010; Murata et al., 2008; Rischer et al., 2006). This task has now been made easier due to the release of the first assemblies of C. roseus transcriptomes by the consortium Medicinal Plant Genomics Resource, together with data for 10 more medicinal plants (http://medicinalplantgenomics.msu.edu/index.shtml). Therefore, the use of molecular cloning strategies to retrieve the cDNAs of C. roseus candidate genes and the development of tools to characterize these genes is of great importance.

### 1.1 Scientific problem

We first became interested in C. roseus due to the high importance credited to the coupling reaction leading to the first dimeric TIA, AVLB, due to its regulatory importance and potential application for the semisynthetic production of the anticancer alkaloids. We have succeeded to identify, characterize and clone a leaf vacuolar Prx with AVLB synthase activity, CroPrx1 (Costa et al., 2008; Duarte et al., 2010; Sottomayor et al., 1996; Sottomayor et al., 1998), and while seeking to characterize the full genomic sequence of
CroPrx1, we have found a different highly homologous Prx gene, CroPrx3. This gene codes for two alternative splicing transcripts/proteins, and given the high similarity of CroPrx3-a and CroPrx3-b with CroPrx1, it was hypothesized that they could be involved in TIA metabolism, since for instance the conversion of ajmalicine into serpentine has also been suggested to be mediated by an uncharacterized Prx (Blom et al., 1991; Sottomayor et al., 2004). Important for the implication of a Prx in TIA metabolism, is its subcellular localization in the vacuole, where the TIA substrates and products accumulate. Considering all this, we decided to clone the CroPrx3 cDNAs and to investigate the subcellular sorting of the CroPrx3 proteins using GFP fusions and their transient expression in protoplasts. As such, we decided to also develop a methodology for the isolation and transformation of C. roseus mesophyll protoplasts, since the use of an homologous system is far more reliable.

2. Materials and methods

2.1 Outline of the experimental approach

The experimental approach used to clone and characterize CroPrx3 involved three main steps: i) cloning of CroPrx3, ii) construction of GFP-CroPrx3 fusions, and iii) transient expression in C. roseus mesophyll protoplasts, as outlined below.

i. Cloning of CroPrx3 involved the characterization of CroPrx1 positive clones from a C. roseus genomic library by a strategy combining subcloning of restriction fragments, PCR, sequencing, and assembly of sequence data. This enabled the in silico generation of a new Prx gene highly similar to CroPrx1, CroPrx3, which coded for two alternative splicing transcripts, CroPrx3-a and CroPrx3-b. Since RT-PCR revealed that both CroPrx3 transcripts were highly expressed in the roots, RT-PCR using root RNA and specific primers designed to cover the full coding sequences, enabled to retrieve CroPrx3s full cDNAs.

ii. In order to study the subcellular sorting of CroPrx3s, which has functional relevance, a number of GFP-CroPrx3-b fusions were designed and generated, including putative sorting signals correctly cloned at the N- and C-terminus of GFP. Concerning the choice of FP and the rules for design of FP fusions to investigate subcellular localization of candidate gene products, please see Duarte et al. (2010).

iii. For transient expression of the GFP-CroPrx3 fusions in C. roseus cells, a fast and efficient technique for isolation of mesophyll protoplasts was developed and a procedure for their PEG-mediated transformation was optimized resulting in high transformation yields. The localization of fluorescence was monitored using a confocal microscope.

2.2 Plant material and growth conditions

Plants of Catharanthus roseus (L.) G. Don cv. Little Bright Eye were grown at 25°C, in a growth chamber, under a 16 h photoperiod, using white fluorescent light with a maximum intensity of 70 µmol m⁻² s⁻¹. Seeds were acquired from AustraHort (Australia) and voucher specimens are deposited at the Herbarium of the Department of Biology of the Faculty of Sciences of the University of Porto (PO 61912). Callus cultures of the C. roseus line MP183L were kindly provided by Johan Memelink from the University of Leiden, The Netherlands, and were grown as described by Pasquali et al. (1992).
2.3 Cloning of CroPrx3

2.3.1 Isolation and sequencing of CroPrx3

The CroPrx3 gene was found while screening a genomic library for CroPrx1 (Genbank accession number AM236087). Screening was carried out using a C. roseus genomic library kindly provided by Johan Memelink from the University of Leiden, The Netherlands, which was produced using C. roseus genomic DNA partially digested with Sau3AI and cloned in λGEM11 (Promega) BamHI arms (Goddijn et al., 1994). Around 600,000 primary phage plaques grown on Escherichia coli host strain KW251 were screened by hybridization with a mixture of two radiolabelled probes specific for CroPrx1, using standard plaque lift methods (Sambrook et al., 1989). After three rounds of screening, four distinct positive clones were encountered corresponding to genomic inserts of 15 to 20 kbp, and two clones, λ8 and λ17, were characterized further. These clones were not accessible to direct sequencing due to their big size, therefore they were digested with EcoRI and the restriction fragments were analyzed by Southern blotting using the same probes as above (Sambrook et al., 1989). The positive and/or common restriction fragments of the two clones were subcloned either in the plasmid pGEM4 (Promega) or in pBluescript SK+ (Stratagene), and were sequenced using first the vectors universal primers, followed by subsequently designed internal primers, until coverage of the entire sequences was obtained. From the first data acquired, it became clear that the gene present in the clones was not CroPrx1 but a highly similar Prx new gene, which was named CroPrx3 (Bakalovic et al., 2006; Peroxibase, http://peroxibase.toulouse.inra.fr index.php). Sequence data of overlapping regions obtained from both strands was assembled and analyzed using PDRAW 32 version3, to produce the complete gene sequence of CroPrx3 (Genbank accession number AM937226).

2.3.2 RT-PCR analysis of C. roseus organs

RNA was isolated using the hot phenol-LiCl procedure previously described by van Slogteren et al. (1983) and Menke et al. (1996) from various organs of five month old C. roseus plants, from 7-day old C. roseus seedlings, and from callus cultures. The different tissues were ground under liquid nitrogen and total RNA was isolated by a 5 min extraction with two volumes of hot phenol buffer (80 ºC, 1:1 mixture of phenol containing 0.1 % hydroxyquinoline with 100 mM LiCl, 10 mM EDTA, 1 % SDS, 100 mM Tris pH 9.0) and one volume of chloroform. The mixture was microcentrifuged 30 min at 5000 rpm at room temperature (RT), and the aqueous phase was mixed with one volume of chloroform, shaken for 5 min and centrifuged as before. The aqueous phase volume was measured, mixed with one-third of its volume of 8 M LiCl (2 M final concentration) and the RNA was precipitated overnight at 4 ºC. The RNA was then collected by microcentrifugation at 10000 rpm for 30 minutes at 4 ºC. The pellet was washed twice with 70 % ethanol and dried under vacuum. Alternatively, RNA can be extracted with the RNeasy Plant Mini Kit (QIAGEN). The dried RNA was dissolved in water, microcentrifuged at 10000 rpm for 10 minutes, at 4 ºC, to precipitate impurities (polysaccharides), the RNase inhibitor RNAsin (Promega) was added, and the RNA was stored at -20 ºC. Integrity and equal loading of RNA was verified by ethidium bromide staining of ribosomal RNA bands after agarose electrophoresis. Quantification of the RNA samples was also performed spectrophotometrically.

RNA was treated with DNase I (Promega) and RT-PCR was performed as follows. Reverse transcription polymerase chain reactions were performed using the Promega Reverse Transcription System (Promega), using a poly(dT)12-18 as a first strand primer. PCR
reactions to detect the presence of transcripts from *CroPrx1*, *CroPrx3-a* and *CroPrx3-b* were performed with DFS-Taq DNA Polymerase (Bioron) for 35 cycles, using the primer pairs described in Table 1. These were cross-tested for specificity using cloned DNA from each gene as template for PCR reactions. The amount of cDNA included in the PCR reactions of each set of experiments was estimated after standardization using the reference gene *Rps9* encoding the 40 S ribosomal protein S9 (Genbank accession number AJ749993; Menke et al., 1999). The RT-PCR products were subjected to electrophoresis in 1.5% agarose gels and amplification products were visualized by staining with ethidium bromide. Band intensity was assessed both visually and with the help of the image analysis software Kodak DC120 Gel Electrophoresis Analysis System.

<table>
<thead>
<tr>
<th>Amplification product</th>
<th>Primer name</th>
<th>Primer sequence 5'-3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>CroPrx3-a fragment</td>
<td>forward</td>
<td>CTCATCTGCTCTTCTATTTGG</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>CAGCACAAAGACATGGCTGA</td>
</tr>
<tr>
<td>CroPrx3-b fragment</td>
<td>forward</td>
<td>GCTGGCTGGAAGGAACGAG</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>CAGCACAAAGACATGGCTGA</td>
</tr>
<tr>
<td>Rps9 fragment</td>
<td>forward</td>
<td>GAGCGTTTGGATGCTGAGTT</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>TCATCTCCATCACCACCAGA</td>
</tr>
</tbody>
</table>

Table 1. Sequence of the primers used in the RT-PCR analysis of *C. roseus* organs for the presence of *CroPrx3-a* and *CroPrx3-b*.

### 2.3.3 Isolation of the complete cDNAs of *CroPrx3-a* and *CroPrx3-b*

Total RNA was extracted from *C. roseus* roots as described in the previous section. Reverse transcription was performed using the Transcriptor Reverse Transcriptase (Roche) using a poly(dT)12-18 as a first strand primer. The generated cDNA was then PCR amplified using 5'UTR specific primers for the two alternative splicing transcripts and a common 3'UTR reverse primer (Table 2, Fig. 3). PCRs were carried out using 0.5 µg of cDNA as template and Pfu DNA Polymerase (Promega). The generated PCR products were cloned into the vector pGEM-Teasy (Promega) and the clones obtained were sequenced to verify the inclusion of the complete coding regions of *CroPrx3-a* and *CroPrx3-b*.

<table>
<thead>
<tr>
<th>Amplification product</th>
<th>Primer</th>
<th>Primer sequence 5'-3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>CroPrx3-a</td>
<td>forward</td>
<td>GGAACATATGGCTTTGATTGC</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>CAGCACAAAGACATGGCTGA</td>
</tr>
<tr>
<td>CroPrx3-b</td>
<td>forward</td>
<td>GCTGGCTGGAAGGAACGAG</td>
</tr>
<tr>
<td></td>
<td>reverse</td>
<td>CAGCACAAAGACATGGCTGA</td>
</tr>
</tbody>
</table>

Table 2. Sequence of the primers used for the isolation of the *CroPrx3-a* and *CroPrx3-b* complete cDNAs.

### 2.4 Generation of GFP-*CroPrx3* fusions

In order to study the subcellular localization of CroPrx3s, a series of fusion constructs of *CroPrx3-b* regions with GFP were designed and generated, as shown in Fig. 1. The *CroPrx3-b* sequences used in the fusions were amplified by PCR using a proofreading Pfu DNA polymerase (Fermentas) and primers including specific *CroPrx3-b* sequences plus the restriction sites required for directional cloning, including additional nucleotides to
guarantee efficient cleavage close to fragments termini, according to manufacturer indications. The primers are shown in Table 3 and their localization relative to Prxs domains are shown in Fig. 1A. Primers were designed to be in frame with GFP, and the fusion sequences were analyzed with the ExPASy Translate tool (http://expasy.org/tools/dna.html) to confirm the generation of a correct open reading frame.

![Fig. 1](image-url)

**Fig. 1.** Schematic representation of the cloning strategy followed to obtain the GFP-CroPrx3-b constructs. A) Representation of the different regions of CroPrx3-b and the primers used for their amplification. B) Schematic representation of the constructs generated to study the subcellular localization of CroPrx3-b. 35S - cauliflower mosaic virus strong promoter, SP - signal peptide, CTE - C-terminal extension, MP - mature protein, nos T - *Agrobacterium tumefaciens* nopaline synthase terminator. The two first constructs were generated in the plasmid pTH-2, while the remaining were generated in pTH-2BN.

<table>
<thead>
<tr>
<th>CroPrx3-b amplification product</th>
<th>Primer name</th>
<th>Primer sequence 5’-3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP</td>
<td>SP Fwd</td>
<td>AGCC GTCGACAAAAATGGTTTTATGAGTTTCCTTTTTC</td>
</tr>
<tr>
<td></td>
<td>SP Rev</td>
<td>GATGCCATGGTTGTGAGCTTCGATATGG</td>
</tr>
<tr>
<td>CTE</td>
<td>CTE Fwd</td>
<td>CAAGATCTTTCCGGAATGCCCGCAGCGGACTTCCTTT</td>
</tr>
<tr>
<td></td>
<td>CTE Rev</td>
<td>CGCCTCGAGTTAAAAACATGGAACAAGCCAACTTTCG</td>
</tr>
<tr>
<td>MP</td>
<td>MP Fwd</td>
<td>CAAGATCTTTGCCACCTATAGTGGAGCTTCCTTTTC</td>
</tr>
<tr>
<td></td>
<td>MP Rev</td>
<td>CGCCTCGAGTTAGCCATTCGCCAAGCCAACTTTCG</td>
</tr>
<tr>
<td>MP-CTE</td>
<td>MP Fwd</td>
<td>CAAGATCTTTGCCACCTATAGTGGAGCTTCCTTTTC</td>
</tr>
<tr>
<td></td>
<td>CTE Rev</td>
<td>CGCCTCGAGTTAAAAACATGGAACAAGCCAACTTTCG</td>
</tr>
</tbody>
</table>

Table 3. Sequence of the primers used for the generation of the GFP-CroPrx3-b fusion constructs. Engineered restriction sites are underlined. SP - signal peptide, MP - mature protein, CTE - C-terminal extension.
The fusion of CroPrx3-b sequences at the N-terminus coding sequence of GFP were generated using the plasmid pTH-2 corresponding to pUC18 carrying the 35SΩ-sGFP(S65T)-nos construct and an ampicillin/carbenicillin-resistance marker (Niwa et al., 1999). This plasmid may be requested from Yasuo Niwa (niwa@fns1.u-shizuokaken.ac.jp). Whenever fusions at the C-terminus coding sequence were required, the plasmid used was pTH-2BN (Kuijt et al., 2004), which lacks a stop codon in the end of the GFP sequence and may be requested from Johan Memelink (j.memelink@biology.leidenuniv.nl).

The amplified CroPrx3-b signal peptide coding sequence (SP) was cloned in frame in pTH-2 and in pTH-2BN, using SalI and NcoI, to generate 35S::SP-GFP constructs in each of the plasmids. The 35S::SP-GFP construct in pTH-2 (Fig.1B) was used directly for transformation, while the 35S::SP-GFP construct in pTH-2BN was used to further insert at the C-terminus coding sequence of GFP, using BglII and XhoI, the following sequences:

i. the CTE coding sequence of CroPrx3-b to generate the construct 35S::SP-GFP-CTE (Fig.1B)

ii. the mature protein coding sequence of CroPrx3-b (excluding the SP and CTE coding sequences) to generate the construct 35S::SP-GFP-MP (Fig.1B)

iii. the mature protein plus CTE coding sequences of CroPrx3-b to generate the construct 35S::SP-GFP-MP-CTE (Fig.1).

All the generated constructs were sequenced with universal and internal primers in order to certify that all clones were error-free.

2.5 Transient expression in C. roseus mesophyll protoplasts

2.5.1 Isolation of C. roseus mesophyll protoplasts

C. roseus mesophyll protoplasts were obtained using a protocol adapted from Negrutiu et al. (1987), Sottomayor et al. (1996) and Yoo et al. (2007). Approximately 8-10 leaves (~1.5 – 2 g) of adult plants (usually 2nd and 3rd pairs) were cut with frequently renewed scalpel blades into ~1 mm strips excluding the central vein, and immediately transferred, abaxial face down, to a Petri dish with 10 mL of digestion medium composed of 2 % (w/v) cellulase (Onozuka R-10, Duchefa), 0.3 % (w/v) macerozyme (Onozuka R-10, Serva) and 0.1 % pectinase (Sigma) dissolved in MM buffer (0.4 M mannitol and 20 mM Mes, pH 5.6-5.8). The medium was vacuum infiltrated during 15 min, applying slow disruptions of the vacuum every ½ a min. Leaf strips were incubated in the digestion medium for ca. 3 h at 25ºC, in the dark, after which the Petri dishes were placed on an orbital shaker (~60 rpm) for 15 min in the dark and at RT, to help release the protoplasts. The suspension was then filtered through a 100 µm nylon mesh and the filtrate was gently transferred into 15 mL falcon tubes, using sawn-off plastic Pasteur pipettes. The protoplast suspension was centrifuged at 65 g for 5 min at 20 ºC, the supernatant was removed, and the protoplasts were washed twice in MM buffer and once in cold W5 solution (154 mM NaCl, 125 mM CaCl2·2H2O, 5 mM KCl and 2 mM Mes, pH 5.7) – the resuspension of the protoplasts must be performed gently, by flicking the tube after addition of a small volume of medium, and only after that adding the full volume for washing. The last pellets were all resuspended in a minimum volume of W5 and pooled together. Protoplasts were counted using a haemocytometer and were left to rest on ice for 30 min. After this incubation, the protoplasts were pelleted as above and resuspended to a protoplast concentration of 5x10⁶ cells mL⁻¹ using the adequate volume of MMg buffer (0.4 M mannitol, 15 mM MgCl₂ and 4 mM Mes, pH 5.7). At this point, the protoplasts were ready for transformation. The integrity of the isolated protoplasts was
checked by observation under an optical microscope (Olympus) and images were acquired by a coupled Olympus DP 25 Digital Camera and respective software (Cell B, Olympus).

2.5.2 Transformation of *C. roseus* mesophyll protoplasts

Transformation of *C. roseus* mesophyll protoplasts was adapted from Yoo et al. (2007). In all cases, 10 µL of 2 µg µL\(^{-1}\) of ultrapure plasmid DNA were mixed with 100 µL of protoplast suspension with 5x10^5 cells, using 2 mL round bottom eppendorfs (the protoplast suspension was pipetted using a P1000 sawn-off tip). Plasmid DNA used for protoplast transformation was isolated using the QIAGEN Plasmid Midi Kit following the manufacturer’s instructions. One volume (110 µL) of 40% (w/v) PEG, 0.2 M mannitol and 0.1 M CaCl\(_2\).2H\(_2\)O, was slowly added (drop by drop) to the DNA-protoplast mixture, gently flicking the tube after each drop, and the tubes were left to incubate for 15 min at RT. After this incubation time, four volumes of W5 solution were slowly added and the tubes were centrifuged at 600 rpm for 2 min, with acceleration and deceleration set at the minimum (1). The supernatant was removed, the pellet was gently resuspended in 100 µL of W5 solution by flicking the tube, the protoplasts were transferred to 15 mL falcon tubes containing 900 µL of W5, and were incubated in the dark at 25ºC, with the tubes lying in a slight slope, for at least 2 days. Prior to confocal observation, cells were subjected to a 3 h incubation at 35ºC, since this showed to increase GFP expression and/or fluorescence. GFP fluorescence was examined using a Leica SP2 AOBS SE confocal microscope equipped with a scanhead with an argon laser. Visualization of GFP was performed using an excitation wavelength of 488 nm and an emission wavelength window from 506 to 538 nm, and visualization of chloroplast autofluorescence was performed using the same excitation wavelength and an emission wavelength window from 648 to 688 nm.

3. Results and discussion

3.1 Cloning of *CroPrx3*

The new Prx gene *CroPrx3* was found while trying to obtain the full genomic sequence of *CroPrx1*, implicated in the biosynthesis of the *C. roseus* anticancer alkaloids. *CroPrx1* positive clones from a *C. roseus* genomic library were obtained through standard screening methods, and were analysed by restriction digestion using an enzyme cutting only once in the previously known coding sequence. This showed that the three smallest bands from clones λ8 and λ17 were similar (data not shown), which was interpreted as the possible presence of the common *CroPrx1* gene in these fragments. This was indeed confirmed by Southern blotting (data not shown), and they were subcloned in two stages. First the λ8 fragments were cloned in pGEM4 with limited success, and in a second stage the λ17 equivalent fragments were successfully cloned in pBluescriptII SK+. These last clones were sequenced using the plasmid universal primers (T3, T7, Fig. 2) and the sequences obtained were used to design internal primers for further rounds of sequencing (IPs, Fig. 2). From the first data acquired it became clear that the gene present in the clones was not *CroPrx1* but a highly similar Prx new gene, which was named *CroPrx3*, using the nomenclature followed by Peroxibase (Bakalovic et al., 2006; http://peroxibase.toulouse.inra.fr/index.php). The whole sequencing data obtained was assembled and analyzed using PDRAW 32 version3, to produce a gene sequence of *CroPrx3* spanning a 5242 bp region (Genbank accession number AM937226), which revealed the gene structure depicted in Fig. 3. *CroPrx3* revealed a striking
structural difference to CroPrx1, due to the presence of two exons 1 in tandem, separated by a spacer sequence, suggesting this is a gene with alternative splicing coding for two proteins with distinct N-terminal regions, CroPrx3-a and CroPrx3-b (manuscript in preparation).

Fig. 2. Schematic representation of the λ17 EcoRI restriction fragments subcloned in the plasmid pBluescriptII SK+ showing the position of the plasmid universal primers (T7,T3) and the newly designed internal primers (IPs) used to obtain the full sequence of the inserts.

Fig. 3. Schematic representation of the CroPrx3 gene with the localization of the primers used to obtain the full length cDNAs of CroPrx3-a and CroPrx3-b. Ex, exon; Int, intron.

In order to investigate the subcellular localization of the proteins codified by CroPrx3, the next step was to isolate and clone the cDNAs of the two transcripts putatively resulting from CroPrx3. For this, it was necessary to use RNA from a tissue/organ where the transcripts were expressed. Therefore, expression analysis by RT-PCR was performed with primers specific for both forms of CroPrx3, revealing that both transcript variants are highly expressed in seedlings, roots and callus cultures, and are generally not expressed in the aerial parts of the plant (Fig. 4).

Fig. 4. RT-PCR analysis of C. roseus organs to assess the presence of the alternative splicing forms of CroPrx3 (CroPrx3-a and CroPrx3-b). Rps9 is commonly used as a house-keeping gene for this species (Menke et al., 1999).
Given the expression results, mRNA extracted from roots was used to isolate the CroPrx3 cDNAs. RT-PCR was performed using 5'UTR specific primers for the two alternative splicing transcripts with a common 3'UTR reverse primer (Table 2, Fig. 3), and enabled the isolation of two complete cDNAs with a 100% sequence match with the two predicted alternative splicing forms. The complete cDNAs obtained for CroPrx3-a and CroPrx3-b were cloned into the vector pGEM-Teasy (Promega).

3.2 Design of GFP-CroPrx3 fusions for the characterization of CroPrx3 subcellular sorting

The subcellular localization of CroPrx3s is highly relevant to investigate their function, since it will indicate which potential substrates are accessible in vivo, namely vacuolar TIAS. Prxs are targeted to the secretory pathway by an N-terminal signal peptide (SP) that directs the protein to the endoplasmic reticulum (ER), from where they follow the default pathway to the cell wall, or they are sorted to the vacuole if an additional sorting signal exists. We have previously shown that CroPrx1 is targeted to the vacuole by a C-terminal extension (CTE) in its amino-acid sequence, which is present in all vacuolar Prxs, but not in cell wall Prxs (Costa et al., 2008; Welinder et al., 2002). CroPrx3-a and CroPrx3-b differ only in the N-terminal regions, with different SPs and N-term of the mature proteins (Fig. 5), and share most of the core polypeptide sequence and the C-terminal sequence.

Fig. 5. Multiple alignment of the N-terminal (A) and C-terminal (B) regions of the C. roseus Prxs, CroPrx3-a, CroPrx3-b and CroPrx1, and PNC1, a peanut Prx localized at the cell wall. The regions highlighted in black correspond to conserved amino acid residues in at least three of the proteins. SPs determined experimentally are underlined, and predicted SPs and CTEs are in italic letters.

Alignment of the predicted CroPrx3s amino acid sequences with CroPrx1, and the well characterized cell wall peanut Prx PNC1 (Genbank accession number M37636) (Fig. 5), shows that both CroPrx3 proteins include a common CTE, and should therefore be sorted to the vacuole. To confirm this, several GFP fusions were designed and generated for CroPrx3-b (Fig.1), in which the coding sequences for the SP, the mature CroPrx3-b protein (MP), and the CTE were cloned in the adequate sides of the GFP coding sequence (Duarte et al., 2010). CroPrx3-b was chosen, since we also purified the protein and obtained MS/MS amino acid sequence data (manuscript in preparation) indicating the precise location of the N-terminal region of the mature CroPrx3-b protein, after cleavage of the SP (Fig. 5A). Therefore, the SP used to fuse at the GFP N-terminus was the sequence in italic in Fig. 5A, plus the subsequent 6 amino acids, to make sure that the cleavage site is correctly recognized. The SP sequence represented for CroPrx3-b in Fig. 5A was predicted using PSORT
The length of the CroPrx3 CTE was deduced from sequence alignment with a well characterized cell wall Prx lacking a vacuolar sorting signal, peanut PCN1 (Fig. 5B), and bearing in mind that all mature Prx proteins terminate four to six residues after the last conserved cysteine (Welinder et al., 2002). In this case, the CTE used included two more amino acid residues than the predicted CTE sequence (Fig. 5B).

3.3 Subcellular sorting of GFP-CroPrx3 fusions in C. roseus mesophyll protoplasts

In order to assess the subcellular localization of CroPrx3, a method for transient expression of the GFP-CroPrx3 fusions in C. roseus mesophyll protoplasts was developed. To obtain C. roseus mesophyll protoplasts, we adapted the general method described by Negrutiu et al. (1987) for Arabidopsis, using an enzyme composition adapted from a previous method used in our lab for C. roseus mesophyll (Sottomayor et al., 1996), and adapting the last steps of the procedure to the subsequent PEG mediated transformation method from Yoo et al. (2007). This enabled to obtain a high yield of healthy and pure protoplasts that, under the light microscope, showed no apparent membrane damage or disintegration of the internal structure (Fig. 6). This methodology usually enabled the retrieval of a total of 5x10^7 protoplasts from a Petri dish prepared as described above, which is enough for up to 100 independent transformations – we usually prepare this amount, use the protoplasts for around 6 transformations per researcher, and use the rest for vacuole isolation or biochemical determinations, but the protoplast isolation method can be scaled down.

The method used for transient expression of the GFP-CroPrx3 fusions in C. roseus mesophyll protoplasts was adapted from Yoo et al. (2007) with just a few modifications, since the method reported by these authors for Arabidopsis revealed to work quite well when it was applied directly to C. roseus, indicating that its transference to further species may also be simple. The main modifications that were introduced were a 25 fold increase in protoplast concentration, enabling a significant increase in transformation yield, and an incubation at 35 °C during the 3 h preceding observation, to boost GFP expression and/or fluorescence. The improved method yielded transformation rates of up to 70 - 80% when transforming with the cytosolic GFP construct. It was evident that both the yield and GFP fluorescence intensity decreased with the size of the GFP fusion construct/protein.
For the investigation of the vacuolar sorting of CroPrx3-b by transient expression, the following constructs were used (Fig. 1B), with the following predicted subcellular localizations:

i. GFP → cytosol
ii. SP-GFP → ER, Golgi, extracellular space as final destination
iii. SP-GFP-CTE → ER, Golgi, vacuole as final destination
iv. SP-GFP-MP → ER, Golgi, extracellular space as final destination
v. SP-GFP-MP-CTE → ER, Golgi, vacuole as final destination

The localization of fluorescence was monitored from 24 to 72 h after transformation and the results obtained are shown in Fig. 7. In fact, the time-course of the transient expression and GFP peak accumulation/fluorescence depended on the construct, possibly as a consequence of the different protein paths and destination compartments involved. Thus, cytosolic GFP already appeared in a few protoplasts 24 h after transformation with a strong fluorescence signal, with the number of fluorescent protoplasts continuing to increase until at least 72 h after transformation (Fig. 7A). All the remaining secretory GFPs (ii to v, above) were not visible at 24 h, with ER and/or Golgi labeling being well recognized at 48 h (Fig. 7B to E). For the construct SP-GFP-MP, it was observed that fluorescence had completely disappeared 72 h after transformation (data not shown), possibly indicating a more transient expression for this construct than for SP-GFP, which putatively codes for the same final subcellular destination. On the other hand, this disappearance of fluorescence may be interpreted as an indication that the GFP synthesized during that period in fact had the extracellular space as final destination, since in this localization the dilution effect prevents fluorescence detection, while in a contained compartment the fluorescence of GFP would be expected to be observed for a longer time. For the constructs including the CTE, at 48 h, not only the ER/Golgi were labeled, but also the vacuole presented a clear green fluorescence, indicating this compartment was the final destination (Fig. 7C and E). At 72 h, fluorescence was mainly observed in the vacuole, and had practically disappeared from the ER/Golgi (Fig. 7F). The vacuolar fluorescence signal was never very strong, possibly due to the dilution effect in such a big compartment, and to GFP instability in the vacuole, as a consequence of the acid pH and/or vacuolar proteases. In fact, according to Tamura et al. (2003), GFP is degraded by vacuolar proteases if plant cells are under light conditions, a fact prevented if cells are maintained in the dark. Likewise, we can only observe vacuolar fluorescence in *C. roseus* protoplasts if they are kept in the dark.

The SP of CroPrx3-b clearly determined the entrance of GFP in the secretory system, with fluorescence observed in the ER and/or Golgi (Fig. 7B), and therefore corresponds to an ER signal peptide, as generally assumed for Prxs. The addition of the CTE of CroPrx3-b to the secretory SP-GFP was enough to sort the fluorescent fusion to the central vacuole, which could also be observed for the fusion including the MP-CTE on the GFP C-terminus (Fig. 7C, E and F). On the other hand, deletion of the CTE from this latter construct prevented the fusion protein to be sorted to the vacuole, since fluorescence was observed only in ER/Golgi-like patterns (Fig. 7D) and disappeared at 72 h (data not shown). These results show that CroPrx3-b is localized in the vacuole and that the vacuolar sorting information for this protein is indeed localized in its CTE. The vacuolar localization of CroPrx3 means that the vacuolar accumulated TIAs are potential substrates for this enzyme.
Fig. 7. Transient transformation of \textit{C. roseus} mesophyll protoplasts with \textit{GFP-CroPrx3} fusions. The schematic representation of each construct used for transformation is depicted.
right above the respective set of confocal images. A) GFP fluorescence pattern observed for the transformation with the control construct 35S::sGFP. GFP accumulates in the cytosol and the nucleus. B) GFP fluorescence pattern observed for the transformation with the construct 35S::SP-sGFP. GFP fluorescence is observed in the ER and possibly Golgi indicating sorting to the secretory pathway. C) GFP fluorescence pattern observed for the transformation with the construct 35S::SP-sGFP-CTE. The presence of the CTE of CrPrx3-b targets secretory GFP to the central vacuole. D) GFP fluorescence pattern observed for the transformation with the construct 35S::SP-sGFP-MP. GFP fluorescence is observed in the ER and possibly Golgi indicating sorting to the secretory pathway but no fluorescence is observed in the vacuole indicating that the absence of the CrPrx3-b CTE impairs the vacuolar sorting of GFP. E and F) GFP fluorescence pattern observed for the transformation with the construct 35S::SP-sGFP-MP-CTE. Addition of the CTE to the CrPrx3-b MP successfully targets secretory GFP to the central vacuole. E1-E3, 48 h after transformation. F1-F3, 72 h after transformation. Images on the left column – GFP channel. Images on the middle column – red channel showing chloroplast autofluorescence. Images on the right column – merged images of the two channels. Bars = 10 µm.

4. Conclusions

Here, we have cloned a new Prx gene from the medicinal plant C. roseus, CroPrx3, and we have characterized its subcellular sorting through transient expression of GFP fusions in C. roseus mesophyll protoplasts, using a newly developed method for this important medicinal plant. Initial cloning of CroPrx3 was achieved through a classical approach involving screening of a genomic library. Currently, for the isolation of new genomic sequences we have been using, and would suggest, PCR based strategies such as inverse PCR (IPCR) (Costa et al., 2008) and genome walking (Gutiérrez et al., 2009).

The transient expression of GFP-CroPrx3 fusions indicated a vacuolar localization for this enzyme, and a vacuolar sorting function for the CroPrx3 CTE, since the presence of this region was sufficient and necessary for the accumulation of green fluorescence in the central vacuole of C. roseus mesophyll protoplasts. The investigation of the subcellular localization of the proteins codified by candidate genes can thus be efficiently investigated by a strategy involving a careful design of fusions with fluorescent proteins (FPs), followed by transient expression (Duarte et al., 2010).

In fact, a highly useful tool for the characterization of candidate genes, now constantly arising from omic and in silico approaches, is the use of transient expression methodologies to investigate subcellular localization, protein-protein interactions, and to perform diverse functional assays like enzymatic activities, transport, etc. These methodologies are particularly relevant if they have been developed for the homologous organism as done in this work. Here, we have developed an easy and fast method for transient expression in the important medicinal plant C. roseus, contributing with a significant breakthrough for its future research. Moreover, the transference and adaptation of the developed methodology to other species of interest should be possible and encouraged.

5. Acknowledgment

The authors would like to thank the following institutions for their financial support: i) Programa Operacional Ciência e Inovação (POCI 2010) financed by the European Union and
Cloning and Characterization of a Candidate Gene from the Medicinal Plant *Catharanthus roseus* Through Transient Expression in Mesophyll Protoplasts

Fundação para a Ciência e Tecnologia (FCT) from the Portuguese Government; ii) the FCT grants PRAXIS/P/BIA/10267/1998 and POCTI/BIO/38369/2001; iii) an FCT grant to Patricia Duarte (SRFH/BPD/20669/2004), iv) a Scientific Mecenate award from Grupo Jerónimo Martins SA. The authors would also like to thank Yasuo Niwa (University of Shizuoka, Japan) for the pTH-2 plasmid, and Johan Memelink (University of Leiden, The Netherlands) for the *C. roseus* callus culture line MP183L, for the *C. roseus* genomic library, and for the pTH-2BN plasmid.

6. References


The development of molecular cloning technology in the early 1970s created a revolution in the biological and biomedical sciences that extends to this day. The contributions in this book provide the reader with a perspective on how pervasive the applications of molecular cloning have become. The contributions are organized in sections based on application, and range from cancer biology and immunology to plant and evolutionary biology. The chapters also cover a wide range of technical approaches, such as positional cloning and cutting edge tools for recombinant protein expression. This book should appeal to many researchers, who should find its information useful for advancing their fields.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
