In *Proceedings of the VI International Plant Peroxidase Symposium* (eds. Acosta M, Rodríguez-López JN & Pedreño MA) (pp 152-158). University of Murcia & University of A Coruña (2002).

# CLONING OF A PEROXIDASE ENZYME INVOLVED IN THE BIOSYNTHESIS OF PHARMACEUTICALLY ACTIVE TERPENOID INDOLE ALKALOIDS IN *CATHARANTHUS ROSEUS* (L.) G. DON

Hilliou, F.[1a], Costa, M.[1a], Almeida, I.[2], Lopes Cardoso, I.[2], Leech, M.[1], Ros Barceló, A.[3] & Sottomayor, M.[2]*

[1]John Innes Centre, Norwich, UK; [2]Institute for Molecular and Cell Biology, Porto University, Portugal; [3]Department of Plant Biology (Plant Physiology), Murcia University, Spain. [a] These authors contributed equally to this work. * Author for correspondence: IBMC, Rua do Campo Alegre, 823, 4150-180 PORTO, Portugal, email: msottoma@ibmc.up.pt

## INTRODUCTION

*Catharanthus roseus* is a medicinal plant that produces in the leaves the anticancer drugs vinblastine (VLB) and vincristine (VCR). VLB and VCR, commonly known as the *Vinca* alkaloids, were the first natural anticancer agents to be clinically used, and are widely applied as an indispensable part of many curative and coadjuvant regimens in cancer chemotherapy (1). Due to the pharmaceutical relevance and demand of these alkaloids, and the low levels present in *C. roseus*, the monoterpenoid indole alkaloid biosynthetic pathway has been intensively studied in this plant (2,3,4,5,6). Production in cell cultures as an alternative source of alkaloids has also been intensively investigated, but the pathway is poorly expressed in this system and the production of the dimeric alkaloids has never been achieved in cell cultures (5,7).

The biosynthetic pathway of vinblastine is highly complex, involving more than twenty enzymatic steps, nine of which are now well characterized at the enzyme and gene level. Recently, regulatory genes of the initial part of the pathway (ORCAs) have been cloned, in what consists a highly promising strategy for the manipulation of the pathway (5,6). However, there are still a number of steps of the pathway that are not characterized, even at the chemical intermediates level, what may hinder the successful manipulation of the dimeric alkaloid levels.

The dimerizing step, thought to be particularly relevant from a regulatory point of view, consists in the coupling of the monomeric precursors catharanthine and vindoline to yield α-3',4'-anhydrovinblastine (AVLB) (Figure 1), which is the direct precursor of VLB and VCR. The search for the enzyme catalysing this reaction led to the finding of a basic peroxidase present in vacuoles of *C. roseus* mesophyll cells that shows AVLB synthase activity – CRPRX1 (8). CRPRX1 was the single peroxidase isoenzyme detected in leaf extracts using the sensitive substrate 4-methoxy-α-naphthol (8,9). The AVLB synthase/peroxidase was purified from *C. roseus* leaves, and the enzyme was shown to be a high spin ferric heme protein belonging to the plant peroxidase superfamily (class III peroxidases) (10). The interaction of peroxidase with the two substrates catharanthine and vindoline to yield AVLB was also investigated and a radical propagated mechanism was proposed (11).

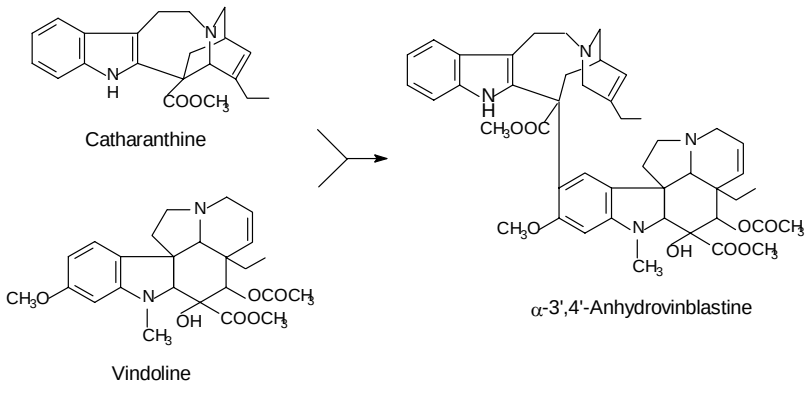


**Figure 1**. Coupling of vindoline and catharanthine to yield a-3',4'-anhydrovinblastine.

To further characterise the basic peroxidase of *C. roseus* leaves involved in indole alkaloid biosynthesis, we have isolated a cDNA clone, by means of a PCR strategy using degenerate primers designed from a partial amino acid sequence of the purified protein.

## MATERIAL AND METHODS

*Plant material*
Plants of *Catharanthus roseus* (L.) G. Don (cv. Little Bright Eyes) were grown at 25ºC in a growth chamber under a 16/8 h white fluorescent light photoperiod.

*Enzyme purification*
Enzyme purification was carried out using a 6 step protocol involving protein precipitation simultaneous with homogenization in 75% (v/v) acetone at -20°C, followed by ammonium sulfate precipitation and 4 chromatographic steps, as described in (10).

*Amino acid sequence analysis and tryptic digestion*
Amino acid sequencing was carried out at the Protein Sequencing Facility of the John Innes Center, Nitrogen Fixation Laboratory, by Dr. M. J. Naldrett. Briefly, 200 pmol of purified protein were analysed by SDS-PAGE, electroblotted onto a polyvinylidene difluoride membrane (PVDF, Millipore) and stained with Coomassie Blue R250. The membrane piece carrying the protein was used directly for automatic Edman degradation of the N-terminus on a Procise protein sequencer (model 491, Applied Biosystems). The heme was removed to 200 pmol of pure protein and the apoenzyme was used for trypsin digestion as described in (12). Peptides resulting from digestion were separated (12) and sequenced by automatic Edman degradation as referred above.

*Construction of cDNA library*
Total RNA was isolated from the first pair of leaves of 60 cm high flowering plants of *C. roseus.* Poly(A)$^+$ mRNA was isolated using a PolyATtrack mRNA kit from Promega. The cDNA was synthesized and inserted into λ-ZAP using the ZAP-cDNA$^{TM}$ kit from Stratagene.

*Isolation of a Full-length cDNA clone coding for CrPrx1*
Semi-nested PCR amplification of the cDNA library using primers AVLB1 (TAY CCI AAY ATH GAY CCI ACI ATG)* + T7 (GTA ATA CGA CTC ACT ATA GGG C) for 1$^{st}$ reaction, followed by primers AVLB2 (AAY ATH GAY CCI ACI ATG GAY C)* + T7 for 2$^{nd}$ reaction, produced a band of 250 bp containing a part of the sequence of the AVLB primers. Semi-nested PCR amplification of the cDNA library using primers AVLB3 (GTY TGR TCC ATI GTI GGR TC)* + Rev (GGA AAC AGC TAT GAC CAT G) for 1$^{st}$ reaction, followed by primers AVLB4 (TCC ATI GTI GGR TCD ATR TTI GG)* + Rev for 2$^{nd}$ reaction, produced a band of 650 bp containing a part of the sequence of the AVLB primers. The deduced amino acid sequence of this band contained the sequence of one of peptides resulting from trypsin digestion of CRPRX1.

The design of new primers and two rounds of PCR reactions enabled the amplification of two segments of the cDNA including sequences of the sequenced peptidic fragments. These products were cloned into pBluescriptSK II+ (Sratagene), sequenced to confirm their identity, and used to screen the cDNA library. The three rounds screening of approximately 750,000 plaques enabled the selection of 13 positive clones, and the sequence of *CrPrx1* was reconstructed from the overlapping sequences obtained from clones 3, 5, 7 and 11, to produce an incomplete cDNA sequence. The complete sequence was obtained by Rapid amplification of cDNA ends (RACE system: GibcoBRL). *I = inosine; Y = C + T; R = A + G; D = G + A + T; H = A + T + C.

*Southern and Northern Blot analysis*
Genomic DNA was extracted from leaves of *C. roseus* using the CTAB method (13). RNA was isolated using Trizol Reagent (Gibco BRL) from pools of roots, stems, petals, reproductive organs and leaves from several mature plants of *C. roseus*. Digested DNAs (10 μg of genomic DNA digested with BglII, ClaI, EcoRI or HincII) were electrophoresed in 1% agarose gels in Tris-borate/EDTA buffer and transferred to Hybond-N membranes by capillary blotting with 0.5 M NaOH. 15 μg total RNA from each sample was electrophoresed in 1.5% agarose/formaldehyde gels as described in (14), and transferred to Hybond-N membranes membranes by capillary blotting with 50 mM sodium phosphate and 5 mM EDTA pH 6.5. Cross-linked blots were (pre)-hybridised in 50% deionised formamide, 5×SSPE and 5% SDS at 42°C, washed with 0.1×SSPE and 0.5% SDS at 65°C, and exposed at -80°C to Fuji-RX films mounted on Kyokko intensifying screens.

*Alkaloid quantification and IEF study of CRPRX1 in C. roseus leaf extracts*
Quantification of leaf alkaloids and detection of CRPRX1 activity by isoelectric focusing of protein extracts, followed by staining of peroxidase activity with 4-methoxy-α-naphthol (4-MN), was performed according to (8), (9) and (10).

## RESULTS AND DISCUSSION

The basic peroxidase present in *C. roseus* leaves was purified to homogeneity and was submitted to automatic Edman degradation. The N-terminus was highly blocked but it was possible, nevertheless, to obtain the sequence of the first 11 amino acids – PPTVSGLSYTF. The purified protein was digested with trypsin and the N-terminal sequence of 4 purified peptides was determined. The sequence of the fragment IYPNIDPTMDQT was found to have the lower identity with other plant peroxidases and, since it was also encoded by triplets of low degeneracy, it was therefore chosen for the design of degenerated oligonucleotide primers.

The consecutive selection of identified amplification products and design of new, specific primers, enabled the cloning of two fragments of *CrPrx1* which were used for screening of the c-DNA library. This produced an incomplete cDNA sequence that was extended by Rapid amplification of cDNA ends (RACE) to obtain a full-length cDNA clone of 1383 bp.

The deduced amino acid sequence of the complete cDNA corresponds to a polypeptide with 363 amino acids that includes the sequence of the tryptic digestion fragments of the pure protein and thus corresponds to the basic peroxidase isoenzyme present in *C. roseus* leaves – CRPRX1. The main results of analysis of the amino acid sequence are shown in Table 4.

**Table 4** – Prosequences and consensus sequences detected in CRPRX1.

| Putative (P) and identified (I) sites | Localization in protein - sequence |
|---|---|
| N-terminal Signal Peptide – ER (I) | -34 to -1 |
| C-terminal Propeptide (P) | 305 - 329 |
| Peroxidase proximal heme-ligand signature with invariant H (I) | 168 - DVVALSGG**H**TI |
| Peroxidase active site signature with invariant H and R (I) | 38 - AAGLL**R**LHF**H**DC |
| Conserved Cysteines – 8 (I) and disulfide bridges – 4 (P) | 16 - 97, 49 - 54, , 103 – 298, 183 - 210 |
| N- glycosylation sites – 6 (P) | 75 - NLTL, 151 - NTSA, 165 – NATD[a], 217 - NRTF, 266 - NQTL, 297 – NCSL |
| Glycosaminoglycan attachment sites – 1 (P) | 65 - SGPG |
| Protein kinase C phosporylation sites – 3 (P) | 77 – TLR, 252 – TDR, 299 - SLR |
| Casein kinase II phosphorylation sites – 6 (P) | 102 – SCSD, 137 – TRAD, 185 – SFDE, 219 – TFLD, 246 – SDQD, 316 - SVVE |
| N-myristoylation sites – 4 (P) | 6 – GLSYTF, 53 – GCDSSV, 132 – GLNFAT, 242 – GLFTSD |
| Amidation sites – 1 (P) | 127 - LGRR |

[a] N165 corresponds to a zero yield during amino acid sequencing of one of the pure protein tryptic fragments, confirming glycosylation.
Identification and prediction of specific sequences was made using The PROSITE database.

Alignment studies and comparison with consensus peroxidase sequences show that CRPRX1 contains all conserved and highly conserved residues typical of class III peroxidases (Table 4). The localization of the N-terminal sequence determined for the purified protein enables the identification of a signal peptide with 34 amino acids corresponding to the ER signal peptide present in all plant peroxidases. Alignment with previously characterized plant peroxidases and the fact that all known mature plant peroxidases terminate near the same site, 6 to 8 amino acid residues after the last cysteine (15), indicate the presence in CRPRX1of a C-terminal propeptide with 23 to 25 amino acids.

BLAST analysis of the CRPRX1 deduced amino acid sequence revealed high homology with the protein coded by a recently cloned tobacco peroxidase cDNA, NTPER9-6 (AY032674; (16)), with a basic peroxidase from azuki bean induced by ethylene, AZ42 (D11337; (17)), and with a spinach peroxidase, SOPRXR3 (Y10464; (18)). The closest Arabidopsis class III peroxidase is ATP4a / ATPRXR6 / ATP12 (AC016163), and the barley seed basic peroxidase BP1 (M73234; (19)) also shares significant homology with CRPRX1. All the homologous proteins possess a C-terminal propeptide, identified for BP1, and putative for the others, which is thought to be responsible for the targeting of these peroxidases to the vacuole. This is in agreement with the previously shown vacuolar localization of CRPRX1 (8). The phylogenetic relation of CRPRX1 with the above mentioned peroxidases from other species is represented in figure 2.

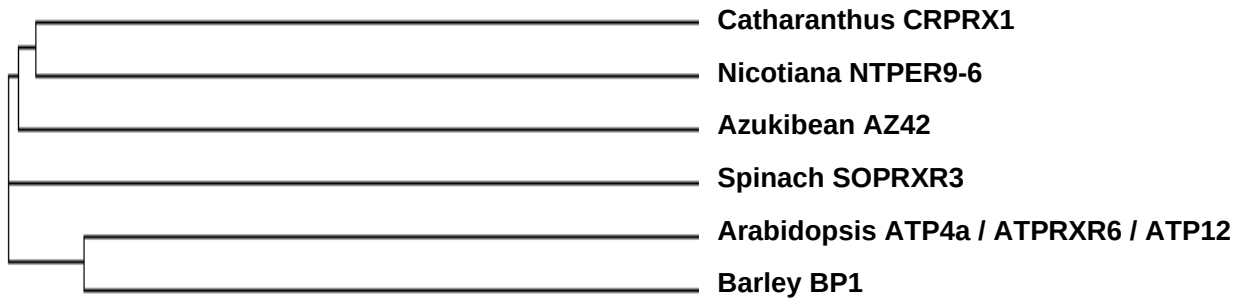


**Figure 2.** Phylogenetic distances between CRPRX1 and putative orthologous peroxidases of other species.

Southern blot analysis (data not shown) revealed that the *CrPrx1* gene is most likely present as a single copy gene in *C. roseus* genome, similarly to what is shown in (20) for the putative paralogous gene of Arabidopsis *AtP4a / AtPrxr6 / AtP12*. In fact, *AtP12* is one of the two single isolated genes in the phylogenetic tree of Arabidopsis 74 peroxidase genes constructed by Tognolly *et al.* (20), in contrast with the other 71 genes grouped in 5 clusters.

The expression of the *CrPrx1* mRNA was observed early during plant development (Figure 3A) being expressed in cotyledons, hypocotyls and roots in 12 days old seedlings. In mature flowering plants, *CrPrx1* transcripts were detected in all aerial organs of the plant (data not shown). The basic peroxidase of *C. roseus* was also expressed during all stages of leaf development (Figure 3**B**).

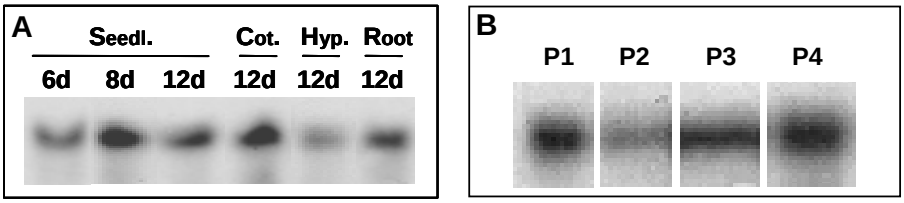


**Figure 3.** Expression of CrPRX1 in *C. roseus* plants. **A** – During development of seedlings (Seedl. – seedlings; Cot. – cotyledons). **B** – During development of leaves in mature plants (P1- fist pair of leaves).

In order to find out if the presence of the CRPRX1 protein in mesophyll cells was related with the capacity of cells to produce AVLB, we studied the alkaloid profiles and CRPRX1 levels during leaf development. For this, 4 plants of *C. roseus* with 16 weeks were analysed on a single leaf basis, in order to avoid that variations between plants masked variations between the different leaf pairs of one plant – this is particularly relevant for alkaloids, since different plants may have very different alkaloid levels. For each leaf pair, one of the leaves was used for alkaloid quantification and the opposite leaf for protein extraction and analyses of peroxidase activity in IEF gels (Figure 4). The plants used had 5 pairs of leaves in which the 5$^{th}$ pair already showed the first signs of senescence.

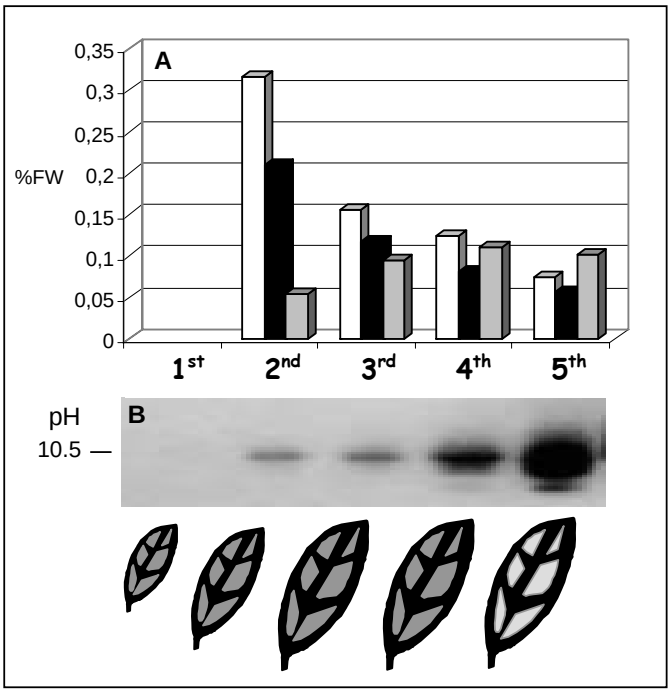


**Figure 4**. Alkaloid levels and CRPRX1 accumulation during leaf development. A – levels of catharanthine (white), vindoline (black) and AVLB (grey). B – IEF showing the band corresponding to CRPRX1 stained with 4-MN.

We observed that, during leaf development, the levels of the monomeric precursors of AVLB, catharanthine and vindoline, decreased progressively, with a simultaneous increase in the product AVLB and in the amount of CRPRX1 activity present. In pre-senescent leaves (5$^{th}$) CRPRX1 continued to increase, while AVLB started to decrease slightly. These results are consistent with the proposed role of CRPRX1 in the biosynthesis of AVLB. In leaves entering senescence, the monomeric precursors seem to become rate limiting for the dimerization reaction, and CRPRX1 must assume a different and relevant function during leaf senescence, as indicated by its high activity in pre-senescent leaves (5$^{th}$ in Figure 4). It was not possible to obtain results for the 1$^{st}$ leaf pair, neither in what concerns alkaloid levels or peroxidase detection. However, the analyses of pools of first leaves confimed the general trend of increasing of CRPRX1 and AVLB levels with leaf age (data not shown).

## CONCLUSIONS

In this work, we have cloned, sequenced and characterized the cDNA corresponding to the single peroxidase isoenzyme present in *Catharanthus roseus* leaves – the basic class III peroxidase CRPRX1 with a pI of 10.5, previously shown to be responsible for the biosynthesis of the indole alkaloid $\alpha$-3',4' – anhydrovinblastine (AVLB) in mesophyll vacuoles (8,10). The predicted amino acid sequence contains all conserved and highly conserved residues typical of class III peroxidases, it includes an ER N-terminal signal peptide, and a C-terminal extension believed to target peroxidases to the vacuole.

CRPRX1 is expressed in most organs of the plant, and is particularly important in leaves where it is the single peroxidase isoenzyme detected (8). Its activity levels in developing leaves correlate well with its assigned function in the biosynthesis of AVLB, but also suggest a function during leaf senescence, indicating its multifunctional character, typical of plant peroxidases.

**ACKNOWLEDGEMENTS**

This work was partially supported by fellowships of the John Innes Center to Frederique Hilliou and Manuela Costa, by grant PI-70/00615/FS/01 of *Fundación Séneca* to A. Ros Barceló, by grant LIS/882/13 (99/2000) of *Instituto de Cooperação Científica e Tecnológica Internacional / The British Council* to Mariana Sottomayor and Mark Leech and by a scholarsip from *Bilateral Agreement Fundação para a Ciência e Tecnologia (PT) / Ministério de Educación y Ciencia (SP)* to Mariana Sottomayor. We wish to thank Dr. M. J. Naldrett for having done such a good job in sequencing such a difficult protein.

## REFERENCES

(1) Levêque, D., Wilhm, J. and Jehl, F. 1996. Pharmacology of *Catharanthus* alkaloids. B. Cancer 83: 176-186.
(2) Meijer, A.H., Verpoorte, R. and Hoge, J.H.C. 1993. Regulation of enzymes and genes involved in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. J. Plant Res. 3: 145-165.
(3) Sottomayor, M., DiCosmo, F. and Ros Barceló, A. 1999. The biogenesis of indole alkaloids. In Hemantarajan, A. (ed). Advances in Plant Physiology, Vol 2. Scientific Publishers (India), Jodhpur, pp 1-41.
(4) Facchini, P.J. 2001. Alkaloid biosynthesis in plants: biochemistry, cell biology, molecular regulation and metabolic engineering applications. Annu. Rev. Plant Physiol. Plant Mol. Biol. 52: 29-66.
(5) Hilliou, F., Van der Fits, L. and Memelink, J. 2001. Molecular regulation of monoterpenoid indole alkaloid biosynthesis. In: Romeo, J.T., Saunders, J.A. and Matthews, B.F. (eds). Regulation of Phytochemicals by Molecular Techniques. Recent Advances in Phytochemistry, vol 35. Elsevier Science, pp 275-95.
(6) Memelink, J., Verpoorte, R., Kijne, J.W. 2001. ORCAnization of jasmonate-responsive gene expression in alkaloid metabolism. Trends Plant Sci. 6: 212-219.
(7) Moreno, P.R.H., Van der Heijden, R. and Verpoorte, R. 1995. Cell and tissue cultures of *Catharanthus roseus*: a literature survey II. Updating from 1988 to 1993. Plant Cell Tiss. Org. Cult. 42: 1-25.
(8) Sottomayor, M., de Pinto, M.C., Salema, R., DiCosmo, F., Pedreño, M.A. and Ros Barcelo, A. 1996. The vacuolar localization of a basic peroxidase isoenzyme responsible for the synthesis of α-3',4'-anhydrovinblastine in *Catharanthus roseus* (L.) G. Don leaves. Plant Cell Environm. 19: 761-767.
(9) Ferrer, M.A., Calderón, A.A., Muñoz, R. and Ros Barceló, A. 1990. 4-Methoxy-α-naftol as a specific substrate for kinetic, zymographic and cytochemical studies on plant peroxidase activities. Phytochem. Anal. 1: 63-69.
(10) Sottomayor, M., Lopéz-Serrano, M., DiCosmo, F. and Ros Barceló, A. 1998. Purification and characterization of α-3',4'-anhydrovinblastine synthase (peroxidase-like) from *Catharanthus roseus* (L.) G. Don. FEBS Lett. 428: 299-303.
(11) Sottomayor, M., DiCosmo, F. and Ros Barceló, A. 1997. On the fate of catharanthine and vindoline during the peroxidase-mediated enzymatic synthesis of α-3',4'-anhydrovinblastine. Enz. Microb. Technol. 21: 543-549.
(12) Kjaersgard, I.V.H., Jespersen, H.M., Rasmussen, S.K. and Welinder, K.G. 1997. Expression and RT-PCR analysis of two peroxidases from *Arabidopsis thaliana* belonging to a novel evolutionary branch of plant peroxidases. Plant Mol. Biol. 33: 699-708.
(13) Doyle, J.J. and Doyle, J.N., 1991. Isolation of plant DNA from fresh tissue. FOCUS 12: 13-15.
(14) Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Smith, J.A., Seidman, J.G. and Struhl, K., 1987. Current protocols in Molecular Biology. New York: Greene Publishing Associates.
(15) Johansson, A., Rasmussen, S.K., Harthill, J.E. and Welinder K.G. 1992. cDNA, amino acid and carbohydrate sequence of barley seed-specific peroxidase BP1. Plant Mol. Biol. 18: 1151-1161.
(16) Dorokhov,Y.L., Ravin,N.V., Zamchuk,L.A., Frolova,O.Y., Skurat,E.V., Makinen,K., Atabekov,J.G. 2001. Nucleotide sequence of tobacco peroxidase gene. Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases. Accession number AY032674 / AAK52084.
(17) Ishige, F., Mori, H., Yamazaki,K. and Imaseki, H. 1993. Identification of a basic glycoprotein induced by ethylene in primary leaves of azuki bean as a cationic peroxidase. Plant Physiol. 101:193-199.
(18) Simon, P. 1993. Diveristy and conservation of plant peroxidases. Plant Peroxidase Newsletter 1: 2-5.
(19) Rasmussen, S.K., Welinder, K.G. and Hejgaard, J. 1991. cDNA cloning, characterization and expression of an endosperm-specific barley peroxidase. Plant Mol. Biol. 16: 317-327.
(20) Tognolli, M., Penel, C., Greppin, H. And Simon, P. 2002. Analysis and expression of the class III peroxidase large gene family in *Arabidopsis thaliana*. Gene 288: 129-138.