Martin-Luther-Universität Halle-Wittenberg
Fachbereich Biochemie/Biotechnologie

cDNA library construction towards the identification of the ipecac alkaloid biosynthesis genes in tissue cultures of *Psychotria ipecacuanha*: Preliminary characterization of the first isolated enzyme, ipecoside β-D-glucosidase

Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)

vorgelegt der
Mathematisch-Naturwissenschaftlich-Technischen Fakultät
(mathematisch-naturwissenschaftlicher Bereich)
der Martin-Luther-Universität Halle-Wittenberg

von Herrn Alfonso Lara Quesada
geb. am 23.03.1974 in Costa Rica

Gutachter 1. Prof. Dr. T.M. Kutchan
2. Prof. Dr. J. Stöckigt
3. Prof. Dr. B. Dräger

Halle (Saale), 24. Mai 2007

urn:nbn:de:gbv:3-000011846
[http://nbn-resolving.de/urn/resolver.pl?urn=nbn%3Ade%3Agbv%3A3-000011846]

Content

1. Introduction **1**

1.1. *Psychotria ipecacuanha* 1

1.2. Tissue culture and main alkaloid contains 2

1.3. Alkaloids 3

1.4. Monoterpenoid indole alkaloid biosynthesis 5

1.5. Monoterpenoid isoquinoline (ipecac) alkaloid biosynthesis 6

1.6. *β*-Glucosidases 9
1.6.1. Glycoside hydrolases 9
1.6.2. *β*-D-Glucosidases 10

1.7. The molecular biology in the secondary metabolism 12
1.7.1. Homologous sequences and degenerate primers 12
1.7.2. Expressed sequence tags (EST) 14

1.8. Aim of the Research 15

2. Materials and methods **16**

2.1. Materials 16

2.1.1. Organisms 16
2.1.1.1. Plants 16
2.1.1.2. Bacterial Strains 16
2.1.2. Nucleic acids and nucleotides 17
2.1.2.1. Plasmids 17
2.1.2.1.1. Cloning vectors 17
2.1.2.1.2. Expression Vector 17
2.1.2.2. Oligonucleotides 18
2.1.2.3. Nucleotides 19
2.1.3. Biological Preparations 19
2.1.4. Chemicals 20
2.1.5. Materials and reagents 21
2.1.6. Instruments 22
2.1.7. Software 23

2.2. Plant Tissue Culture Methods 24
2.2.1. Growth conditions and Media 24
2.2.2. Shoots Micropropagation 24
2.2.3. Roots culture 24
2.2.4. Methyl jasmonate induced cultures 25

2.3. Microbiological methods 26

2.3.1. Preparation of media and agar plates 26
2.3.2. Transformation with thermal shock 26

2.4. Electrophoresis gel 26
2.4.1. DNA agarose gel 26
2.4.2. RNA agarose gel 27
2.4.3. Protein polyacrylamide gel electrophoresis (PAGE) 27

2.5. Isolation of nucleic acids 27
2.5.1. DNA mini-preparations 27
2.5.2. Plasmid miniprep in 96^{th} microtiter plate 28
2.5.3. DNA extraction from an agarose gel 28
2.5.4. Purification of PCR products 28
2.5.5. Total RNA extraction 29
2.5.6 Isolation of poly $(A)^+$ RNA 29

2.6. Determination of concentrations 29
2.6.1. Determination of nucleic acid concentrations with UV spectrophotometers 29
2.6.2. Determination of the protein concentration 29

2.7. Polymerase Chain Reaction (PCR) 30
2.7.1. Standard PCR 30
2.7.2. Reverse Transcriptase reaction PCR (RT-PCR) 30

2.8. cDNA library construction 30

2.9. Degenerated primer method 31

2.10. Rapid Amplification of cDNA Ends (RACE) 31
2.10.1. Construction of the cDNA 31

2.11. Sequencing of DNA 31

2.12. Cloning techniques 33
2.12.1. Restriction reactions 33
2.12.2. Ligation 33

2.13. Protein expression 34
2.13.1. Expression of recombinant Protein in *E. coli* 34

2.14. Enzymatic reactions 35
2.14.1. Activity measurement in E. coli expressed recombinant enzyme 35

2.15. Chemical substances analysis 35
2.15.1 Alkaloid extraction from plant tissues 35
2.15.2. Alkaloid extraction from liquid medium 36

2.16. Chromatographic methods 36
2.16.1 High Performance Liquid Chromatography (HPLC) 36
2.16.2 Liquid Chromatography – Mass Spectrometry (LC-MS, TOF) 37
2.16.3. LC-ESI-MS/MSSI-MS/MS 37

3. Results **38**

3.1. Plant tissue culture 38
3.1.1. *Psychotria ipecacuanha* shoots micropropagation 38
3.1.2. *Psychotria ipecacuanha* root culture 38

3.2. Quantitative analysis of the alkaloid content in tissue cultures 41
3.2.1. Identification of the main ipecac alkaloids 41
3.2.2. Additional ipecac alkaloids in *P. ipecacuanha* 42
3.2.3. Elicitation of tissue cultures with methyl jasmonate 43

3.3. Cloning of alkaloid biosynthesis genes from *P. ipecacuanha* 46
3.3.1. cDNA library construction 46
3.3.2. Expressed Sequence Tags (EST) analysis 47
3.3.3. Complete gene sequence analysis 48
3.3.3.1. Strictosidine synthase-like gene clone 48
3.3.3.2. Isolation of a strictosidine synthase-like gene by degenerate primers method 51
3.3.4. Ipecoside glucosidase clone 52
3.3.4.1. Analysis of the first clone 52
3.3.4.2. Rapid Amplification of cDNA End (5'-RACE) 52
3.3.4.3. Full-length clone sequence analysis 53

3.4. Enzymatic activity analysis of the strictosidine synthase-like gene clone 58
3.4.1. Heterologus expression and purification 58
3.4.2. Enzymatic activity test 59

3.5. Enzymatic characterization of the of ipecoside glucosidase gene clone (Ipe-Gluc) 61
3.5.1 Heterologus expression and purification 61
3.5.2. Enzymatic activity test 63
3.5.2.1. Determination of pH and temperature optimum 63
3.5.2.2. Determination of the enzyme specificity and stability 65
3.5.3. Mass spectrometric analysis 68

4. Discussion **72**

4.1. Tissue culture and alkaloid contain 72

4.2. cDNA library construction 74

4.3. cDNA identification and analysis 75

4.4. Strictosidine synthase-like cDNA (ips*tr*-like) 76

4.4.1. Sequence analysis 76
4.4.2. Isolation approach by the degenerate primer method 78

4.5. Ipecoside glucosidase (Ipe-Gluc) 79
4.5.1. Isolated clones 79
4.5.2. Sequence analysis 79

4.5.3. Recombinant enzyme 80
4.5.4. Preliminary enzyme characterization 81

5. Summary **83**

6. References **86**

7. Acknowlegments **95**

Abbreviations

%	percent
µ	micro
µl	micro liter
A	Ampere
aa	amino acid
AMP	Adenosine monophosphate
APS	Ammonium persulfate
ATP	Adenosine triphosphate
AU	Absorption unit (s)
BAP	Benzyl amino purine
BLAST	Basic Local Alignment Search Tool
bp	base pair
BSA	Bovine serum albumin
°C	grad Celsius
cDNA	Complementary DNA
cm	centimeter
cpm	counts per minute
ctrl	controls
Da	Daltons
dATP	desoxyadenosintriphosphate, desoxyadenine
dCTP	desoxycytidinetriphosphate, desoxycytosine
ddNTP(s)	didesoxynucleosidetriphosphate, didesoxynucleotide
dGTP	desoxyguanosinetriphosphate, desoxyguanine
dH_2O	distillated water
dhc	dehydrocodeine
DIS	Deacetylipecoside
DIIS	Deacetylisoipecoside

DNA	desoxyribonucleic acid
dNTP	2´-desoxynucleoside-5´-triphosphate
dTTP	desoxythymidintriphosphate, desoxythymine
dw	dry weight
EDTA	Ethylenediamine tetracetate
e.g.	exampli gratia – for example
et al.	*et alii* – and others
ESI-MS	Electro Spray Ionization-Mass Spectrometer
EST	Expressed Sequence Tag
EtOH	ethanol
g	gram
GSP	Gene Specific Primer
h	hour
HCl	Chlorhydric acid
HEB	His-Tag elution buffer
HLB	His-Tag lysis buffer
HPLC	High Performance Liquid Chromatography
HWB	His-Tag wash buffer
IAA	Indole acetic acid
Ipe-Gluc	Ipecoside glucosidase
IPTG	Isopropyl-1-thio-β-D-galaktopyranoside
JA	Jasmonic acid
kDa	kilodaltons
l	liter
LC-MS	Liquid Chromatography-Mass spectrometer
lux	lumen . m^{-2}
m	milli / meter
M	Molar
MeJA	Methyl jasmonate

mg	milligram
min	minute
MOPS	3-(N-morpholino)-propansulfone acid
mRNA	messenger RNA
MS	Murashige and Skoog media
MT	Methyltransferase
nt	nucleotide
NAA	Naphthalene acetic acid
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NCBI	National Center for Biotechnology Information
ORF	Open Reading Frame
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase Chain Reaction
pH	pondus hydrogenii
RACE	Rapid Amplification of cDNA Ends
RG	Raucaffricine glucosidase
RGM	Root growth media
RIM	Root induction media
RMM	Root multiplication media
RNA	Riboncleic acid
RNase	Ribonuclease
rpm	revolutions per minute
RT	Room Temperature
RT-PCR	Reverse Transcription-Polymerase Chain Reaction
SDS	Sodium dodecyl sulfate
sec	second (s)
SG	Strictosidine glucosidase
STR	Strictosidine synthase

syn	synonym
t	time
TAE	Tris/Acetat/EDTA
TDC	Tryptophan decarboxylase
TE	Tris/EDTA
TEMED	N, N, N', N'-Tetramethylethylendiamine
TIA	Terpenoid Indole Alkaloids
t_R	Retention time
Tris	Tris-(hydroxymethyle)-aminomethane
U	Unit(s)
UB	Urea buffer
UV	ultraviolet
V	volt
(v/v)	volume per volume
(w/v)	weight per volume

Abbreviations for amino acids:

Chemical name	**One letter code**	**Three letter code**
Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartic acid	D	Asp
Cysteine	C	Cys
Glutamic acid	E	Glu
Glutamine	Q	Gln
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile

Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

1. Introduction

1.1. *Psychotria ipecacuanha*

The plants of *Psychotria ipecacuanha* (Brot.) Stokes (Rubiaceae) are perennial subshrubs (20 - 40 cm), the natural habitat of which is the understory in tropical hot humid forest of Central and South America. The roots of *P. ipecacuanha* are reddish-brown; consist of twisted fragments of around 15 cm long, no thicker than 6 mm and annulated (fig. 1.1.-A). The best known active principles in this plant are the isoquinoline alkaloids emetine and cephaeline (fig. 1.1.-B), localized principally in the cortex of the root and rhizome, but also many others alkaloids are isolated from different organs of this and related plant species (Bruneton, 1999). The so called ipecac alkaloids are used nowadays mainly in medicine as emetics (an agent that induces vomiting), expectorants and anti-amoebicides (Yoshimatsu & Shimomura, 1991). The use of this plant as a traditional herbal medicinal resource for several illnesses began somewhere with the native Brazilians. Commercial harvesting of this species began at the 16th Century and European traders establish an export system to their home countries, where the drug was mostly used against dysentery (Evans, 1989; Oliveira and Martins, 2002). The protein-inhibiting effects of emetine or its somewhat safer form, dihydroemetine, and other ipecac alkaloids may account for the ability of the plant to inhibit growth or kill several types of microorganisms including parasites like amoeba (Wright and Phillipson, 1990; Schmeller and Wink, 1998). Other important compounds are psychotrine and *O*-methylpsychotrine that appear to be selective inhibitors of human immunodeficiency virus-1 reverse transcriptase (Tan *et al.*, 1991).

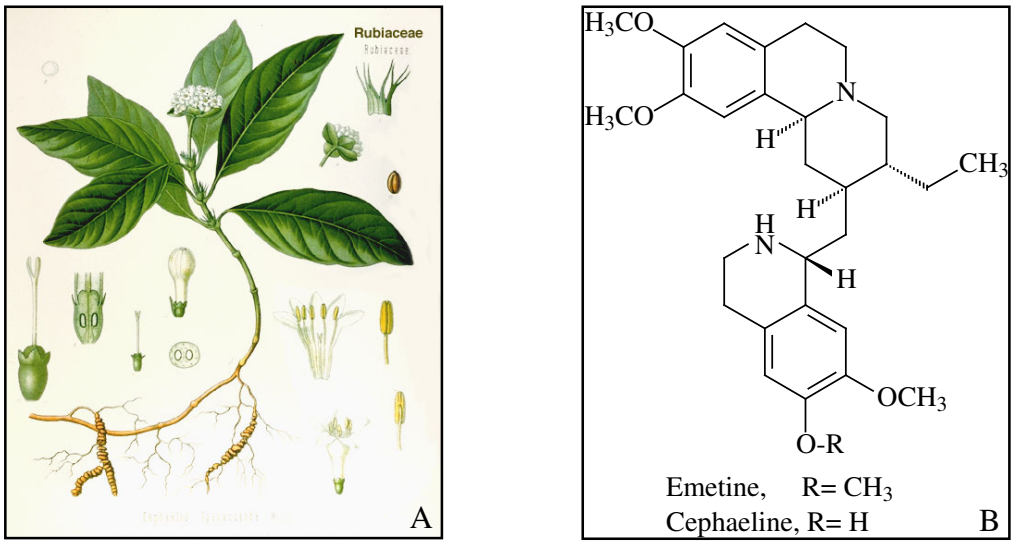

Figure 1.1. A- *Psychotria ipecacuanha* plant illustration showing flowers, roots, rhizome and other organs. B- Molecular structure of the principal terpenoid isoquinoline (ipecac) alkaloids.

The species occurs naturally in three discrete geographic regions: 1- Central America (e.g. Panama and Costa Rica) and northern part of South America (Colombia), 2- Southwestern part of the Brazilian Amazon (Rodonia and Mato Grosso); and 3- Atlantic rain forest along the Brazilian coast. Works on germplasm collection and conservation studies were initiated in Brazil (Skorupa and Assis, 1998), and a study of the genetic erosion of *P. ipecacuanha* was done considering different factors of risk for the *in situ* and *ex situ* conservation on its genetic resources (Oliveira and Martins, 2002). In the same direction, the recent reproductive study on meiotic behavior and pollen viability found that although this species is known to propagate by vegetative multiplication, the sexual reproduction seems to be as important as vegetative propagation (Souza *et al.*, 2006). An interesting description of how *P. ipecacuanha* was identified in Brazil, brought to Europe and later on introduced in India for cultivation can be read in Lloyd (1897). Also a brief story about the different names acquired in the time after its introduction in Europe and the identification of the first alkaloids are nicely told (Lloyd, 1897).

Although the relevance of this plant and the use of its roots to produce the ipecac syrup since hundred years, studies are still lacking for a properly identification and a description of the phenotypic variations of the species. Boer and Thulin (2005) tried to clarify the differences between *Cephaelis* (syn = *Psychotria*) *ipecacuanha* and *Cephaelis acuminata*. Moreover, the synonyms *Psychotria, Cephaelis*, *Callicoca*, *Uragoga* and *Carapichea* have been used; *Carapichea* is the name used for the genus after the most recent revision of the group (Andersson, 2002). For practical reasons, in this work the species will be referred to with the synonym *Psychotria ipecacuanha*.

1.2. Tissue culture and main alkaloid contains

The *in vitro* culture of *P. ipecacuanha* has been largely studied and used for several propose. Some of the most important studies are on clonal propagation by adventitious shoots, root culture, cell suspension culture (Ikeda *et al.*, 1988; Jha *et al.*, 1991; Yoshimatsu and Shimomura, 1991a), somatic embryogenesis (Rout *et al.*, 2000) and *in vitro* alkaloid production (Jha *et al.*, 1988; Teshima *et al.*, 1988; Yoshimatsu and Shimomura, 1991b) and are summarized by Yoshimatsu and Shimomura (1993). The amount of the ipecac alkaloids emetine and cephaeline related to the tissue cultures was shown to be relatively stable. In liquid media, cultured roots were found to accumulate between 0.16 - 0.51 % dry weight (dw) emetine and 0.62 - 1.32 % dw cephaeline. In adventitious shoot cultures without roots, amounts were estimated between 0.04 - 0.07 % dw emetine and 0.4 - 0.5 % dw cephaeline.

Callus cultures generated from hypocotyl explants show the best growth rate on B5 solid medium and the highest alkaloid production (0.93 % dw emetine and 0.346 % dw cephaeline) was found on SH solid medium, both with the same growth regulator supplementation (8 mg/l IBA, 4 mg/l IAA and 4 mg/l NAA) (Jha *et al.*, 1988). However, Teshima *et al.* (1988) found only traces of the alkaloids in calli generated from leaf segments on MS medium with 1 mg/l 2,4-D. Adventitious roots generated from leaves and cultured in MS liquid medium with 3.0 mg/l IAA produced 0.412 and 1.32 % dw of emetine and cephaeline respectively. Root culture supplemented with 1 mg/l NAA also yielded 0.366 and 1.043 % dw of emetine and cephaeline, respectively (Teshima *et al.*, 1988). The production of alkaloids in shoots was, in most of the cases, higher in parental plants (0.82 and 2.16 % dw of emetine and cephaeline, respectively) than in the *in vitro* plants (0.07 and 0.5 % dw of emetine and cephaeline, respectively) from different tested culture conditions (Yoshimatsu and Shimomura, 1991b). This characteristic seems to be generalized for the different plant organs including roots. Also the alkaloid concentrations can be directly related to the maturity of the plant (Yoshimatsu and Shimomura, 1993).

Other works on *P. ipecacuanha* showed the effect of exogenous feeding of callus cultures on alkaloid production. Cephaeline concentration in callus was increased in media supplemented with L-phenylalanine and was doubled with exogenously provided shikimic acid. The concentration of emetine did not show an effect in the presence of any of the proposed biosynthetic precursors (Veeresham *et al.*, 1994). The most recent work on *P. ipecacuanha* tissue culture reports the transformation of *in vitro* petiole segments with *Agrobacterium rhizogenes*. The transformation was confirmed by Southern blot analysis; the slow growth of the transformed roots on a half strength salt MS media without growth regulators was also observed as a characteristic of transformed roots. Regenerated plants from transformed roots showed differences compared to untransformed plants that included vigorous root growth. A brief description of the alkaloid production is mentioned with respect to light conditions, growth regulators and transformed roots (Yoshimatsu *et al.*, 2003).

1.3. Alkaloids

Alkaloids are considered secondary metabolism products and, in general, consist of low molecular weight nitrogen-containing organic compounds, usually with a heterocyclic structure. About 20 % of plant species accumulate alkaloids, which are mostly derived, through the decarboxylation of amino acid precursors (*i.e.* histidine, tyrosine, tryptophan,

lysine and ornithine) to produce their amines, but also are derived from anthranilic acid and nicotinic acid (Wink, 1999). By coupling the amines to different chemical partners, plants can produce a reduced number of the central key intermediates which can then serve as, in some cases, particularly versatile sources for diverse kinds of alkaloids (*i.e.* strictosidine for isoquinoline and norcoclaurine for benzylisoquinoline alkaloids) (De Luca and Laflamme, 2001). Alkaloids are of interest because of their biological activities including medicinal relevance (Kutchan, 1995) and also an important ecological function in the plant defense against pathogens and herbivores (Baldwin, 1999). Nowadays, more than 16,000 alkaloid structures have been elucidated. Several studies on alkaloid-producing plants suggest that the biosynthesis and accumulation are highly regulated process included cell-, tissue-, development- and environment-specific controls (Kutchan, 2005; De Luca and St Pierre, 2000).

An important objective in the research of the different secondary metabolites, like alkaloids, is the uncovering of their biosynthetic pathways and the corresponding induction signals involved in the mechanism control (Kutchan, 1998). In cell suspension cultures of *C. roseus* treated with a fungal elicitor, a rapid transient induction of tryptophan decarboxylase and strictosidine synthase enzyme activities was detected, followed by the accumulation of indole alkaloids (Eiler *et al*., 1987) and later on a transcriptional control was suggested for these events due to the transient appearance of the mRNAs for both enzymes (Roewer *et al*., 1992). In addition, a rapid and transient accumulation of jasmonic acid (JA) and methyl jasmonate (MeJa) was found in cell suspension cultures of *Rauvolfia canescens* and *Eschscholtzia californica* treated with yeast elicitor. Moreover, the induction of *R. canescens* with 250 µM of methyl jasmonate after 96 h increased the concentration of the alkaloid raucaffricine from 7 to 203 mg/l medium. These results demonstrated the role of JA and MeJa in the intracellular signal transduction pathway leading from the elicitor to the accumulation of secondary compounds (Gundlach *et al.*, 1992). More recently it was also shown that MeJa to promote tabersonine biosynthesis in hairy root cultures of *C. roseus* (Rodriguez et al., 2003). At the molecular level, the transcriptional regulation of terpenoid indole alkaloid (TIA) biosynthesis genes under the effect of MeJa has been also investigated. The interaction of jasmonate- and elicitor-responsive element (JERE) with two jasmonic acid-responsive transcription factors called octadecanoid-responsive *Catharanthus* AP2-domain proteins (ORCAs) activate the expression of the strictosidine synthase gene, which produce the common precursor of a wide range of different TIAs (Memelink *et al.*, 2001).

1.4. Monoterpenoid indole alkaloid biosynthesis

The monoterpenoid indole alkaloid-producing plants are a source of many pharmaceutical drugs. Some of the most relevant alkaloids are the antineoplastics vinblastine and vincristine, the antihypertensives ajmalicine and ajmaline, the antimalarial quinine. The first specific enzyme of monoterpenoid indole alkaloid biosynthesis (fig. 1.2) that was identified was strictosidine synthase (STR) (Kutchan, 1989), found in *Rauvolfia serpentina*, which condenses tryptamine and secologanin to form the first intermediate 3α(*S*)-strictosidine (Stöckigt and Zenk, 1977). Before the first characterization of the strictosidine synthase cDNA *str*, activity of the enzyme was known in different *Catharanthus* species, *Rauvolfia* species, both from the Apocynaceae family, and also in Rubiacea, Solanaceae and Leguminosae species (Treimer and Zenk, 1979). The STR-encoding cDNA was also isolated and characterized from *Catharanthus roseus* (McKnight *et al.*, 1990). Recently, the crystallization of STR (Ma *et al.*, 2004) and its crystallization with the substrate tryptamine (Koepke *et al.*, 2005) for the structure and future reaction mechanisms determination have been achieved. The model presents a novel six-bladed β-propeller fold in plants (Ma *et al.*, 2006). A future understanding of the binding and catalytic processes open the possibility to decipher similar mechanisms in the likely related enzymes deacetylipecoside and deacetylisoipecoside synthases.

Secologanin, a key component in the generation of monoterpenoid indole and ipecac alkaloids, is formed via a hydroxylation of geraniol by the geraniol 10-hydroxylase, which leads to the generation of 10-hydroxygeraniol. Although the biosynthetic pathway has not yet been completely elucidated, the formation of secologanin is finally achieved by the cytochrome p450, secologanin synthase (Irmler *et al*., 2000). Tryptamine is obtained by the decarboxylation of L-tryptophan by the tryptophan decarboxylase enzyme (TDC) (De Luca *et al.*, 1989). Secologanin and tryptamine are combined by a Pictet-Spengler condensation to form the intermediate 3α(*S*)-strictosidine, which is then deglycosylated by strictosidine β-D-glucosidase (SG) to yield an unstable dialdehyde (fig. 1.2). By an undetermined mechanism, the unstable structure is channeled to one of several possible skeletons, which are the key compounds for the formation of the major alkaloid types (Hemscheidt and Zenk, 1980; De Luca and Laflamme, 2001). Beginning with strictosidine synthase, the biosynthesis of ajmaline in *R. serpentina* requires at least ten different enzymes (Kutchan, 1998).

In some plants species of the genus *Psychotria* (Rubiaceae), all from the Brazil region, terpenoid indole alkaloids have also been found (Pasquali *et al.*, 2006). *N*-β-D-glucopyranosyl vincosamide was found in shoots of *P. leiocarpa* (Henriques *et al*., 2004); the TIA lyaloside, the major alkaloid in *P. suterrella*, shows toxicity and induces convulsions in rodents (De santos *et al.*, 2001); psychollatine (formerly known as umbellatine) in *P. umbellta* is a TIA structurally related to serotonin with anxiolytic (drug prescribed for the treatment of symptoms of anxiety) and opioid-like analgesic effects (Both *et al.*, 2005); and brachycerine from leaves of *P. brachyceras* has antioxidant and anti-inflammatory effects (Kerber *et al.*, 2001). Psychollatine and brachycerine have been characterized as a new class of TIAs because they are produced from tryptamine and a nonsecologanin terpenoid moiety related to loganin. Therefore, the authors do not expect to find STR in these plants but a STR-like form capable of synthesizing such alkaloids (Pasquali *et al.*, 2006). It is important to mention that *Psychotria* species form one of the largest genera of angiosperms (1000 to 1650 species world wide). It is taxonomically complex due to the unwieldy number of species and difficulties in subgrouping (Kerber *et al.*, 2001). Related to this high taxonomic variety could be the high heterogeneity even in their secondary metabolism.

1.5. Monoterpenoid isoquinoline (ipecac) alkaloid biosynthesis

The ipecac alkaloids are well known to be used against parasites (Muhammad *et al.*, 2003), as emetic (vomiting inductor), against drug absorption (Teshima *et al.*, 1990), as expectorant and because their inhibitor effect on the RNA, DNA and protein synthesis (Grollman, 1968). Tan *et al.* (1991) show the inhibitory activity effect of *O*-methylpsychotrine sulfate heptahydrate and psychotrine dihydrogen oxalate against the human immunodeficiency virus type 1 reverse transcriptase (HIV-1 RT) and HIV-2 RT. An extensive review of the biological activity of ipecac alkaloids is summarized and briefly described in Fujii and Ohba (1983) and expanded upon in Fujii and Ohba (1998).

Ipecac alkaloids are found in the plant families Rubiaceae, Alangiaceae and Icacinaceae (Evans, 1989). *Psychotria ipecacuanha* and *Alangium lamarckii* are two commonly used plants for the extraction and study of ipecac alkaloids. From *Psychotria spp.*, approximately twenty ipecac alkaloids have been isolated: emetine, cephaeline, psychotrine, *O*-methylpsychotrine, emetamine, protoemetine, ipecoside, 6-*O*-methylipecoside, 7-*O*-methylipecoside, ipecoside acid, *trans*-cephaeloside, 6-*O*-methyl-*trans*-cephaeloside, 6-*O*-methyl-*cis*-cephaeloside, *cis*-cephaeloside, neoipecoside, 7-*O*-methylneoipecoside,

3-4-dehydroneoipecoside (possibly an artifact), alangiside, demethylalangiside and 3-*O*-demethyl-2-*O*-methylalangiside. Emetine, cephaeline, psychotrine, alangiside, demethylalangiside and some other bases closely related to the ipecac alkaloids were also found in *Alangium lamarckii* (Fujii and Ohba, 1998). Five additional, to the above mentioned, ipecac alkaloids were also isolated from *Psychotria* (*acuminata* ?) roots by Itoh *et al.* (1999). Recently, 7-*O*-demethylisocephaeline was isolated from stem bark of *P. klugii* plants (Muhammad *et al.*, 2003).

Although the relative similarities in the genesis of the first intermediates and homologous steps in the biosynthesis of terpenoid indole and ipecac alkaloids, the high medical relevance of compounds like vinblastine, vincristine and ajmaline intensified the research to uncover the processes in the production of the TIAs rather than the ipecac alkaloids. Nevertheless, efforts have been made also to understand the metabolic pathway leading to the production of the compounds emetine and cephaeline.

The first concrete proposal for the biosynthetic pathway of the ipecac alkaloids cephaeline and emetine as well as the glucosides ipecoside and alangiside was made in the work of Nagakura *et al.* (1978). The condensation of dopamine and secologanin in a Pictet-Spengler manner, through possibly two different enzymes, would form the intermediates deacetylipecoside and deacetylisoipecoside (Battersby *et al.*, 1969). The epimer with the *β*-configuration (deacetylipecoside) would be inactivated by acetylation in *Psychotria* species to ipecoside or by a lactam formation in *Alangium* species to alangiside. The *α*-configuration (deacetylisoipecoside) would be the intermediate for the formation, probably through protoemetine, to several terpenoid isoquinoline alkaloids including cephaeline and emetine (Nagakura *et al.*, 1978) (fig. 1.2). The work of De-Eknamkul *et al.* (1997) demonstrated the condensation of dopamine and secologanin by two different enzymes. The condensation of dopamine and secologanin produced both epimers (*R*) and (*S*) in presence of enzyme crude extracts from *Alangium* leaves. Through the partial purification of the enzyme, only the *R*-epimer of demethylalangiside (stable form of deacetylipecoside) was detected. This indicated that after the partial enzyme purification, the preparation contained only the activity of deacetylipecoside synthase and it is presumed the instability of deacetylisoipecoside synthase in the purification preparation.

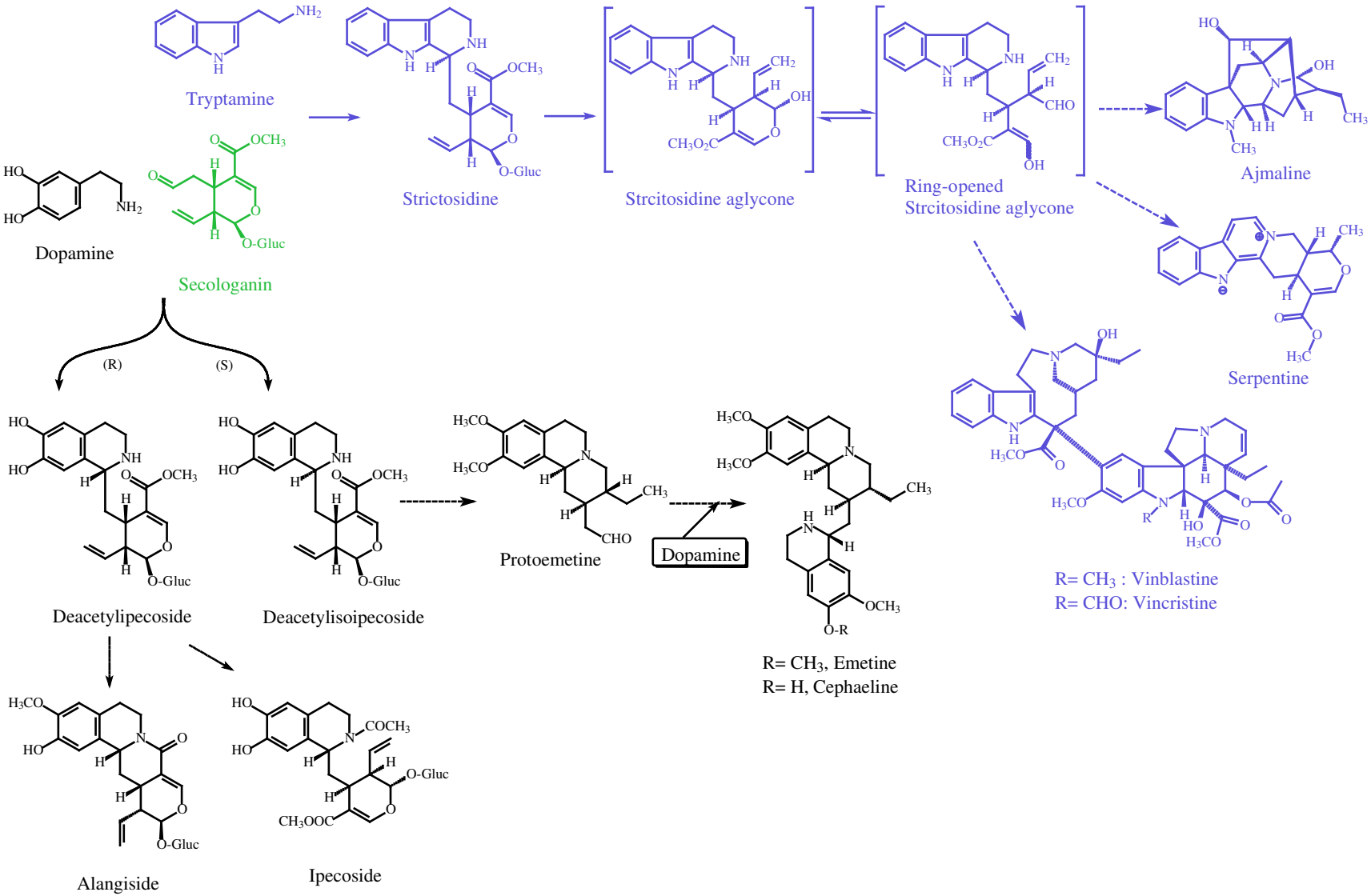

Figure 1.2. Biosynthetic pathways of monoterpenoid indole alkaloids (blue) and ipecac alkaloids. Secologanin (green) take part, with tryptamine, in the condensation of the intermediate strictosidine and, with dopamine, of the intermediates deacetylipecoside and deacetylisoipecoside.

1.6. β-Glucosidases

1.6.1. Glycoside hydrolases

Enzyme classification is based on the type of reaction that the enzyme catalyzes and on its substrate specificity (EC classification). This classification does not necessarily reflect structural features of the enzyme. Since many enzymes, especially glycoside hydrolases show broad substrate specificity there is a need for a classification based on amino acid sequence similarities. Glycoside hydrolases that catalyze the hydrolysis of *O*-glycosyl linkages between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety are classified into 108 different families (at November 2006 in http://afmb.cnrs-mrs.fr/CAZY/) based on their sequence similarities. Many enzymes displaying identical substrate specificity are found in different families displaying totally unrelated three-dimensional folds. Because the fold of proteins is better conserved than their sequences, some of the families can then be grouped in one of the 14 formed 'clans', each one named with a letter (from GH-A to GH-N). (Davies and Henrissat, 1995).

Enzymatic hydrolysis of the glycosidic bond requires two critical residues: a proton donor and nucleophile/base. The two major mechanisms for this hydrolysis lead to an overall retention or an inversion of the anomeric configuration. In both cases the position of the proton donor is identical, however, with inverting enzymes the nucleophile/base is more distant and a water molecule is accommodated between the two catalytic residues resulting in a product with opposite stereochemistry to the substrate (Davies and Henrissat, 1995).

The glycoside hydrolase family 1, which share the GH-A clan with the families 2, 5, 10, 17, 26, 30, 35, 39, 42, 50, 51, 53, 59, 72, 79 and 86, is characterized to present a $(\beta/\alpha)_8$ barrel structure fold, hydrolysis via overall retention mechanism, a glutamic acid (Glu) as catalytic nucleophile/base and a Glu as catalytic proton donor (Coutinho and Henrissat, 1999). In the glycoside hydrolase family 1, because of the sequence similarity, are included enzymes like β-glucosidase (*e.i.*, amygdalase, cellobiase) [EC 3.2.1.21], strictosidine β-glucosidase [EC 3.2.1.105] from *C. roseus* (Geerlings *et al.*, 2000) and from *R. serpentina* (Gerasimenko *et al.*, 2002), raucaffricine β-glucosidase [EC 3.2.1.125] from *R. serpentina* (Warzecha *et al.*, 2000) and many other glucosidases and glycosidases (Coutinho and Henrissat, 1999).

1.6.2. *β*-D-Glucosidases

β-Glucosidases are some of the oldest enzymes known. As early as 1837, Wöhler and Liebig (1837) reported on *β*-glucosidase (emulsin) action in almond on the substrate amygdalin, a cyanogenic glucoside found in almond. *β*-glucosidases occur widely among all organisms and most of them share some common features: they are exoglucolases with a subunit size of 55 to 65 kDa and acidic pH optima (pH 4.0 - 6.2) that require *β*-glucosides as substrates (Esen, 1993). In plants, *β*-glucosidases have been implicated in numerous different plant functions, such as growth regulation by the release of active plant hormones (Campos *et al*., 1992; Brzobohatý *et al*., 1993; Falk and Rask, 1995; Dietz *et al*., 2000), lignification (Hösel *el al*., 1982; Dharmawardhana *et al*., 1999) and plant defense (Poulton, 1990). Cyanogenic glucosides, for example, are *O*-*β*-glucosides from which hydrogen cyanide (HCN) is released upon tissue damage. The sugar moiety is enzymatically cleaved by an *O*-*β*-glucosidase and the corresponding aglycone is formed. The aglycone is then, spontaneously or enzymatically by an *α*-hydroxynitrile lyase, converted to carbonyl compounds and HCN (Poulton, 1990).

If a catalytic residue of a member of one family is known, one can immediately trace the presence of this residue in the other members (Henrissat, 1991). For the *Agrobacterium* *β*-glucosidase (family 1), the glutamic acid Glu-358 has been shown to be directly involved in glycosidic bond cleavage by acting as a nucleophile (Withers et al., 1990). Nine different mutations generated at the corresponding glutamate of a *β*-glucosidase from *A. faecalis* resulted in a decrease in activity of at least 10,000-fold (Trimbur *et al.*, 1992). Multiple sequence alignments of *β*-glucosidases in the family 1 members show that the same glutamate residue is absolutely conserved in all catalytically competent members of the group (Henrissat, 1991) and the 5 residues around (YITENG) (Keresztessy *et al.*, 2001) are also highly conserved. Another important residue in this conserved region is glycine, a replacement by serine or cysteine result in a severe reduction in activity of the enzyme. Therefore, it has been proposed that once the stereochemical mechanism (*i.e*. catalytic mechanisms) is established for one member of a family, it may be safely extended to other members of that family (Gebler *et al*., 1992). The conserved region DxxRxxY seems to be also relevant for the glucosidase enzyme activity. A substitution of aspartic acid produced the greatest activity reduction depending on the substitute used. However of the activity reduction, the DxxRxxY motif has not a critical role in catalysis (Trimbur *et al*, 1992).

Strictosidine *β*-D-glucosidase (Geerlings *et al.*, 2000; Gerasimenko *et al*., 2002) as well as raucaffricine *β*-D-glucosidase (Warzecha *et al.*, 1999; Warzecha *et al.*, 2000) present in their sequences both motifs, and the residue glutamate for the proton donor and nucleophile are also conserved. Crystallization and the preliminary X-ray analysis of strictosidine *β*-glucosidase have been achieved recently (Barleben *et al.*, 2005), but studies about the activity of these motifs still are not reported. Nevertheless, the presence also of the second motif (DxxRxxY) in both sequences suggests a relevance of these conserved regions for the catalytic activity in both *β*-D-glucosidases.

The crystallization and enzyme structure has been determined for the cyanogenic *β*-glucosidase linamarase from white clover (Barrett *et al*., 1995) and the *β*-glucosidase in maize (ZMGlu1) (Czjzek *et al*., 2000; Czjzek *et al*., 2001). For ZMGlu1, the aglycone moiety of the substrate is sandwiched between the residues W-378 on one side and F-198, F-205, and F-466 on the other. Thus, specific conformations of these four hydrophobic amino acids and the shape of the aglycone-binding site determine the aglycone recognition and substrate specificity in Glu1. All residues but W-378 are variable among *β*-glucosidases that differ in substrate specificity, supporting the conclusion that these sites are the basis of aglycone recognition and binding (Czjzek *et al*., 2000). Other studies about substrate specificity show that the mode of aglycone binding in enzyme substrate complexes of dhurrin glucosidase (Dhr1) differs from the homologous maize enzyme. Specifically, the data suggest that Asn-259, Phe-261, and Ser-462, located in the aglycone-binding site of *S. bicolor* Dhr1, are crucial for aglycone recognition and binding. The residue at position 462 presents a high divergence throughout the whole family 1 *β*-glucosidases in *Arabidopsis*. This variability may reflect the diversity of substrates hydrolyzed by these enzymes. This is also true for the residue at position 463, equally located in the aglycone binding pocket, which is as variable as position 462 and therefore most likely highly important for the substrate specificity of these enzymes (Verdoucq *et al.*, 2004).

1.7. The molecular biology in the secondary metabolism

1.7.1. Homologous sequences and degenerate primers.

In the last twenty years, the investigations on plant secondary metabolism have employed the newest molecular techniques with the aim of accelerating the discovery of genes for a specific reaction or even clarify complete biosynthetic pathways. Kutchan in 1998 recognize already the success of a decade of molecular genetics to the alkaloid field and made a review of the works on the principal alkaloids groups (Kutchan, 1998). The success in finding new genes is linked to the appropriate strategies and the correct use of the molecular methods. Selection of oligonucleotide primers, for example, is of high relevance to the polymerase chain reaction (PCR), RT-PCR, oligo-hybridization and DNA sequencing. Proper primer design is actually one of the most important factor/steps in successful DNA sequencing (Abd-Elsalam, 2003). Several parameters including the length of the primer, % GC content and the 3'-end sequence need be optimized for successful PCR. Help for the design of appropriate specific as well as degenerate primers can be found in numerous web-base resources (Abd-Elsalam, 2003; McPherson *et al.*, 1991).

From the extensive list of works done on alkaloid biosynthesis research, some of the recent works done with primers based on conserved amino acid sequences in plant methyltransferases (MT) for the isolation of homologous enzymes involved in the biosynthesis of the alkaloid berberine in *Thalicum tuberosum*. The initial obtained cDNAs allowed the identification of *mt* clones with a broad specificity including phenylpropanoids and alkaloids (Frick and Kutchan, 1999). Similarly, (*R,S*)-reticuline 7-OMT and (*R,S*)-norcoclaurine 6-OMT cDNA were amplified from *Papaver somniferum* by RT-PCR using primers based on internal amino acid sequences of previously isolated proteins with similar sequence to plant OMT (Ounaroon *et al.*, 2003) and in the terpenoid indole alkaloids also strictosidine and raucaffricine glucosidases were isolated with degenerate primers from *Rauvolfia serpentina* and *Catharanthus roseus* (Warzecha *et al.*, 2000; Geerlings *et al.*, 2000; Gerasimenko *et al.*, 2002).

Strictosidine synthase STR1, responsible for the stereospecific condensation of tryptamine with secologanin (fig. 1.2), was discovered in 1977 (Stöckigt and Zenk, 1977); the cDNA was cloned ten years later using primers based on internal amino acid sequences of the protein from *Rauvolfia serpentina* (Kutchan, 1988). STR1 was also identified in a *Catharanthus*

roseus cDNA library by a labeled oligonucleotide derived from the sequence of *R. serpentina* STR1. From 45,000 clones screened, only one clone hybridized to the labeled probe and when heterologously expressed resulted in active recombinant STR1 enzyme (McKnight *et al.*, 1990). Based on the molecular methods, *str*-like genes have been found in complex alkaloid producing and also in complex alkaloid-free plants like *Arabidopsis thaliana* and soybean, and even in animals. Hypotheses on the origin and evolution have been proposed without enzymatic characterization of the gene products in the alkaloid free species (Kutchan, 1993; De Luca and Laflamme, 2001, Facchini *et al.*, 2004).

Highly degenerate primers, based on a single region of peptide sequence data, may also be used in rapid amplification on cDNA ends (RACE) PCR (Frohman *et al.*, 1988). This allows the efficient amplification of the 5'- and 3'-ends of a target cDNA and, if required, the subsequent generation of a full-length cDNA by virtue of the overlapping sequences between the two original primers (McPherson *et al.*, 1991). Raucaffricine glucosidase (RG) is an enzyme that catalyzes the formation of vomilenine in the biosynthesis of the alkaloid ajmaline (fig. 1.3). Although the enzyme is present a maxim activity with the substrate raucaffricine, it accepts also strictosidine to a small degree (Warzecha *et al.*, 2000). For the molecular cloning and later functional bacterial expression of *rg*, two degenerate oligonucleotides derived from conserved regions in other glucosidases were used in combination with those derived from RG protein fragments that were previously sequenced (Warzecha *et al.*, 1999). Degenerate polymerase chain reaction primers also were designed based on the homology between different plant β-glucosidases and used to identify and isolate a strictosidine glucosidase (*sg*) clone from a *C. roseus* cDNA library (Geerlings *et al.*, 2000). From *R. serpentina,* a cDNA encoding SG was successfully amplified with primers designed on the basis of comparison of cDNA sequences of *sg* from *C. roseus* and *rg* from *R. serpentina*. Both SG enzymes were heterologously expressed and the activity was shown for several substrates, strictosidine having the highest relative activity (Geerlings *et al.*, 2000; Gerasimenko *et al.*, 2002).

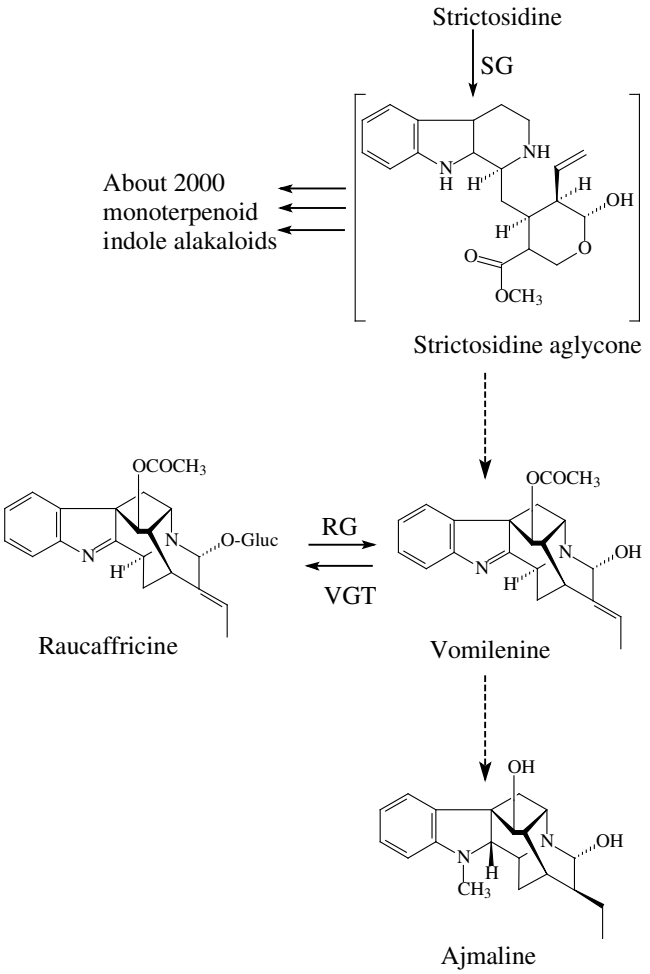

Figure 1.3. Abbreviated biosynthetic pathway to the ajmaline alkaloid. Some enzymes involved in the pathway found so far are showed: SG – strictosidine glucosidase, RG – raucaffricine glucosidase and VGT – vomilenine gucosyl transferase.

1.7.2. Expressed sequence tags (EST)

Large-scale single-pass sequencing of cDNA clones randomly picked from libraries has proven to be a powerful approach to discover genes (Adams *et al.*, 1991; Bonaldo *et al.*, 1996). With the construction of a cDNA-library and the partial sequencing of randomly chosen clones, expressed sequence tags (EST), is expected to scan the expression of mRNAs of a determined tissue under specific physiological and environmental conditions for a plant species (Bonaldo *et al.*, 1996; NCBI, 2004). Also combined methods like macroarray and expressed sequence tags (EST) allowed identification of a specific *O*-methyltransferase involved in benzylisoquinoline biosynthesis (Ziegler *et al.*, 2005) and a reductase (salutaridine reductase) specific to morphine biosynthesis in *P. somniferum* (Ziegler *et al.*, 2006). Recently, two cDNA libraries were generated from RNA isolated from the base part of young leaves and from root tips of *C. roseus* and their expressed sequence tags were

analyzed. The development of microarray based transcriptome analysis and gene discovery are aims in the near future work (Murata *et al.*, 2006). A study of Shukla *et al.* (2006) used the differential gene expression in the two main tissues (leaf and root) of *C. roseus* for the comparative analysis of transcriptomes between both tissues.

1.8. Aim of the Research

In spite of the medicinal importance of the ipecac alkaloids (Fujii and Ohba, 1998) and their use throughout hundreds of years, little is known about of their biosynthesis. Investigations have been mostly directed towards the *in vitro* culture and chemical characterization of tissues (Yoshimatsu and Shimomura, 1993).

Even though investigations have demonstrated the presence of two enzymes involved in the first step of the metabolic pathway (Nagakura *et al.*, 1978 ; De-Eknamkul *et al.* 1997), so far no genes have been isolated.

This work had the aims to make, first, an overview of the alkaloids present in *in vitro* cultures of *Psychotria ipecacuanha* and to determinate the effect of the well known elicitor methyl jasmonate on the alkaloid production of leaves and roots of *in vitro* cultures. The main objective was to generate the first cDNA library from *in vitro* root culture from *Psychotria ipecacuanha*. An expressed sequence tag (EST) analysis of the cDNA library would facilitate identification of cDNAs potentially involved in the alkaloid biosynthesis. The cloning of full-length cDNAs, functional heterologous expression and characterization of the recombinant enzymes would begin to unravel the pathways in the biosynthesis of ipecac and alangiside alkaloids at the molecular level.

2. Materials and methods

2.1. Materials

2.1.1. Organisms
2.1.1.1. Plants

Psychotria ipecacuanha in vitro plants (fig. 2.1. -A) were obtained from the Tropical Plant Biotechnology Laboratory of the Technologic Institute of Costa Rica in San Carlos, region where the plants are also cultivated in field. This plant belongs to the Rubiaceae family and it is found in tropical and subtropical forest in regions of Latin-America (fig. 2.1. -B).

Figure 2.1. A- *In vitro* plants of *Psychotria ipecacuanha* in a MS solid media. B- Original location, in Costa Rica, where the plant material was obtained for the establishment of the *in vitro* plants.

2.1.1.2. Bacteria strains

Escherichia coli strains used:

DH5α (Clontech, California): F^-, *deo*R, *end*A1, *gyr*A96, *hsd*R17, ($r_k^- m_k^+$), *rec*A1, *rel*A1, *sup*E44, Φ80*lac*ZΔM15, *thi*-1, Δ(*lac*ZYA-*arg*FV)U169.

TOP10 (Invitrogen, Karlsruhe):
F^-, *mcr*A, Δ(*mrr-hsd*RMS-*mcr*BC), Δ*lac*X74, Φ80*lac*ZΔM15, *deo*R, *rec*A1, *end*A1, *gal*K, *nup*G, *ara*D139, Δ(*ara-leu*)7697, *rps*L, (Str^R), *gal*U.

TOP10F' (Invitrogen, Karlsruhe): *mcr*A, Δ(*mcr*BC-*hsd*RMS-*mrr*), *end*A1, *rec*A1, *rel*A1, *gyr*A96, Φ80*lac*ZΔM15, *deo*R, *nup*G, *ara*D139, F{*lac*I^{q}, *Tn*10(Tet^{r})}, *gal*U, Δ*lac*X74, *gal*K, Δ(*ara-leu*)7697.

XL1-Blue MRF' (Stratagene, California): *(mcrA)183 .(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac [F 'proAB lacI^{q} Z .M15 Tn 10 (Tet^{r})].*

Rosetta 2(DE3) (Novagen, USA): *F– ompT hsdSB(rB– mB–) gal dcm (DE3) pRARE23 (CamR).*

2.1.2. Nucleic acids and nucleotides

2.1.2.1. Plasmids
2.1.2.1.1. Cloning vectors

pCR®2.1 (Invitrogen, Karlsruhe)
This is a 3.9 kb lineal cloning vector that allows a quick one-step cloning strategy for the direct insertion of a PCR product into the vector. *Taq* polymerase has a nontemplate-dependent activity that adds a single deoxyadenosine (A) to the 3'-ends of PCR products. The linearized vector has single 3'-deoxythymidine (T) residues. This allows PCR inserts to ligate efficiently with the vector.

pGEM-T Easy vector (Promega, Mannheim)
This 3.0 kb lineal cloning vector has a 3'-thymidine overhang that makes possible the ligation of *Taq* DNA polymerase amplified PCR products. General specifications for this vector were T7 and SP6 transcription as well as promoter position, MCS, lacO, a lacZ start codon, an α-lactamase sequence as well as pUC/M13 forward and reverse sequencing start point.

2.1.2.1.2. Expression vector

pET100/D-TOPO (Invitrogen, Karlsruhe):
This 5.7 kb vector contains a TOPO cloning site that presents a GTGG antisense 5'-sticky-end overhang and a 3'-blunt-end. Other specifications are a T7 promotor and a transcription-termination position as well as a T7r-primer position, RBS, a lac-operator (lacO), a V5-epitop, a c-terminus 6x His-tag, a pBR322 replication origin and ampicillin resistance.

2.1.2.2. Oligonucleotides

Synthetic oligonucleotides were obtained from MWG-Biotech AG, Ebersberg and Biomers.net, Ulm. These oligonucleotides were used in polymerase chain reactions (PCR), for cloning, sequencing and RT-PCR.

Strictosidine synthase specific oligonucleotides **Calculated Tm**

str1-5'	5' ATG GCC AAA CTT TCT GAT TC 3'	**53.2 °C**
str1-5'-2	5' CGA GTT ATC AAG TAC GAA GGA 3'	**58.4 °C**
str1-3'	5' GCC ATG GAA CAG GGT TC 3'	**55.2 °C**

Degenerated oligonucleotides (from Biomers.net, Ulm)

STR III	5' TTY AAG TGG YT5 TAC GC 3'	**50.4 °C**
STR III b	5' TTY AAG TGG YT5 TAT GC 3'	**47.1 °C**
STR IV	5' TTA ATV AAR TAY GAC CC 3'	**45.0 °C**
STR IV b	5' TTA ATV AAR TAY GAT CC 3'	**41.9 °C**
STR VII-r	5' ACT TSW ARA ATR TTN CCG A 3'	**54.7 °C**
STR VII b-r	5' ACT TSW ARA ATR TTN CCA A 3'	**51.8 °C**
STR VIII-r	5' RTG CTC TTG AAY YTG CTC 3'	**53.7 °C**
STR VIII b-r	5' RTG CTC TTG AAY YTG TTC 3'	**50.5 °C**

Cloning oligonucleotides

3-GlucSpecif	5' CCA ATC TTC AAT TGT CAA AAG GGA CGC 3'	**63.4 °C**
5-GlucSpecif	5' GGC ACG AGG TGG AAG CAA TGG AG 3'	**66.0 °C**
5-Gluc3+adap	5' CAC CAC AAC AAC AAT GTC TAG TG 3'	**58.9 °C**
5-GlucSpefull	5' GTG CAA TCC CTC AAC AAC AAC AAC AAT G 3'	**61.9 °C**
DIS_3	5' AGA GGA GGA ACA TAC ACC 3'	**53.7 °C**
DIS_5	5' CAC CAT GCC TGC ATG 3'	**54.3 °C**
Glu_3	5' GTC AAA AGG GAC GCT TTC 3'	**53.7 °C**
Glu_5	5' CAC CGC ACG AGG TG 3'	**48.0 °C**
GSP1_Gluco	5' GAA CAG AGT TAC ATA GGG TAC 3' Did not work	**55.9 °C**
GSP2_Gluco	5' CCT GCG TTT ATA CTC CCA CCA GGC AAT ATC 3'	**70.8 °C**
GSPI_Gluc	5' GCC ATC ATA TTG ATC TTG TAA GGC TTG GGG 3'	**66.8 °C**
GSPII_Gluc	5' CTC CCA CCA GGC AAT ATC CGT GGC CAT G 3'	**71.0 °C**
GSPI_Glu2race	5' GCA GAG CTC CGC GAA GTC ACG GAA GTC G 3'	**72.4 °C**
GSPII_Glu2	5' CCA GGC AAT ATC CGT GGC CAT GAT ATG GAG 3'	**69.5 °C**

Sequencing oligonucleotides:

Gluc-621	5’ GAG TAT TTG CTA GCA CTG CAG 3’	**57.9 °C**
Ph0212_60_Glu_C	5’ GCA TGA CAT AAA GGA GAA CTA C 3’	**56.5 °C**
ph0212_60_Glu_D	5’ GGA TGA GAA TAA ATG AAA GCG TC 3’	**57.1 °C**
ph0212_60_Gluco	5’ CTG TTC CAT TGG GAT GTT C 3’	**54.5 °C**
Ph0212_60_Gluco_B	5’ ACT TGT GAC TCA GTG GAT G 3’	**55.2 °C**
Ph0212_Gluco_BII	5’ ACT GAT GGC AAT TTC TAT ACC AC 3’	**57.1 °C**
Ph0412_62_Syn_C	5’ AGT GGA GGA GAA GGA TG 3’	**52.8 °C**
ph0412_62_Syn_D	5’ CTG CAG TTT CAA TGT TCA AGA AG 3’	**57.1 °C**
ph0412_62_Synth	5’ ACA GGT CAG GCT TAT GTG 3’	**53.7 °C**
Ph0412_62_Synth_B	5’ TCC TAA TGG TGT CTC CAT G 3’	**55.2 °C**
Ph0412_Synth_CII	5’ GAG ATT TGG TTA TCC TGA TGT G 3’	**56.5 °C**
Str_Synth_am1020	5’ CCA CTA CGT GAT GGA AAT TG 3’	**55.3 °C**

General oligonucleotides

M13 rev	5’ CAG GAA ACA GCT ATG ACC 3’	**53.7 °C**
M13 uni	5’ TGT AAA ACG ACG GCC AGT 3’	**53.7 °C**
SP6	5’ GAT TTA GGT GAC ACT ATA GAA TAC 3’	**55.5°C**
T3	5’ GCT CGA AAT TAA CCC TCA CTA AAG 3’	**59.3 °C**
T7	5’ GAA TTG TAA TAC GAC TCA CTA TAG 3’	**55.9 °C**
T7 TOPO	5’ T AAT ACG ACT CAC TAT AGG 3’	**53.2 °C**
5’TriplEx 5’long	5’ CAA GCT CCG AGA TCT GGA CGA GC 3’	**65.8 °C**
3’TriplEx 3’long	5’ ATA CGA CTC ACT ATA GGG CGA ATT GGCC 3’	**64.5 °C**
T7_REV	5’ TAG TTA TTG CTC AGC GGT GG 3’	**57.3 °C**
dT20VN	5’-$(T)_{20}$ VN-3’	
T7pZI-1	5’ AGT AAT ACG ACT CAC TAT AGG 3’	**54.0 °C**

Y= C or T, 5= Inosin, , R= A or G, W= A or T, S= G or C V= A, C or G, N= A, C, G or T

2.1.2.3. Nucleotides

dATP, dTTP, dCTP, dGTP	Promega (Mannheim)
1 kb base pair ladder	New England Biolabs
O'GeneRuler™ 100bp DNA Ladder Plus	Fermentas (St. Leon-Rot)

2.1.3. Biological preparations

Enzymes:

BD Advantage™ 2 PCR Enzyme System	BD Biosciences
*Bam*HI	New England Biolabs
*Eco*RI	New England Biolabs
Lysozyme	Sigma-Aldrich (Taufkirchen)
M-MLV Reverse Transcriptase, RNase H(-)	Promega (Mannheim)
Pfu-Polymerase	Promega (Mannheim)

T4-DNA Ligase	Roche (Boehringer Mannheim)
Taq-Polymerase	Promega (Mannheim)

Proteins:

Protein Molecular Weight Marker	Fermentas (St. Leon-Rot)
Bovine serum albumin (BSA)	Roth (Karlsruhe)
Yeast extracts	Difco (Detroit)

Antibiotics:

Ampicillin 50 mg/ml	Sigma-Aldrich (Taufkirchen)
Chloramphenicol 30 mg/ml	Sigma-Aldrich (Taufkirchen)
Carbenicillin 50 mg/ml DMSO	Duchefa Biochemie (Netherlands)
Kanamycin 50 mg/ml	Roth (Karlsruhe)
Tetracyclin 5 mg/ml	Roche (Boehringer Mannheim)

2.1.4. Chemicals

[α-^{32}P]-dATP, 3000 Ci/mmol	Biomedicals, ICN
Agar	Serva (Heidelberg)
Ammonium persulfate	Serva (Heidelberg)
6-Benzylaminopurine (BAP)	Duchefa Biochemie (Netherlands)
Bromophenol blue	Sigma-Aldrich (Taufkirchen)
Chloroform	Roth (Karlsruhe)
Coomassie Brilliant Blue G-250	Serva (Heidelberg)
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich (Taufkirchen)
Dopamine	Sigma-Aldrich (Taufkirchen)
DTT (Dithiothreitol)	Roth (Karlsruhe)
EDTA (Ethylendiamintetraacetic acid)	Roth (Karlsruhe)
Acetic acid	Roth (Karlsruhe)
Ethanol	Merck (Darmstadt)
Ethidium bromide	Sigma-Aldrich (Taufkirchen)
Formaldehyde	Merck (Darmstadt)
Formamide	Fluka (Taufkirchen)
Glucose, D(+)	Merck (Darmstadt)
Glycerol	Roth (Karlsruhe)
Glycine	Roth (Karlsruhe)
HCl	Roth (Karlsruhe)
Imidazol	Roth (Karlsruhe)
IPTG (Isopropyl-1-thio-α-galaktopyranosid)	Fluka (Taufkirchen)
Isopropanol (2-Propanol)	Merck (Darmstadt)
LiCl	Merck (Darmstadt)
MES (2-(N-Morpholino) ethansulfon acid)	Serva (Heidelberg)
Methanol	Merck (Darmstadt)
$MgCl_2$	Roth (Karlsruhe)
$MgSO_4$	Sigma-Aldrich (Taufkirchen)
3-(N-morpholino)-2-propansulfon acid (MOPS)	Roth (Karlsruhe)
NaCl	Roth (Karlsruhe)
NaOH	Roth (Karlsruhe)

a- Naphthalene acetic acid (NAA)	Duchefa Biochemie (Netherlands)
Nicotinic acid	Roth (Karlsruhe)
p-Nitrophenil-ß-D-glucopyranoside	CHEMAPOL
Secologanin	Sigma-Aldrich (Taufkirchen)
Sodium acetate	Merck (Darmstadt)
Sodium cianoborohydride	Sigma-Aldrich (Taufkirchen)
Sodium citrate	Roth (Karlsruhe)
Orange G	Sigma-Aldrich (Taufkirchen)
Peptone	Difco (Detroit)
Phenol-Chloroform	Roth (Karlsruhe)
Polyacrilamine Gel 30	Roth (Karlsruhe)
PVP (Polyvinylpyrrolidon)	Sigma-Aldrich (Taufkirchen)
Pyridoxine-hydrochloride	Roth (Karlsruhe)
Saccarose	Roth (Karlsruhe)
SDS (Sodium lauryl sulphate)	Roth (Karlsruhe)
Sephadex G-50 Superfine	Amersham Pharmacia
Sorbitol	Roth (Karlsruhe)
N,N,N',N'-tetramethylethylen-diamine(TEMED)	Roth (Karlsruhe)
thiamine- hydrochloride	Roth (Karlsruhe)
Tris (Tris-(hydroxymethyl)-aminomethane)	Roth (Karlsruhe)
Tween 20	Roth (Karlsruhe)
Urea	Merck (Darmstadt)
5-brom-4-chlor-indolyl-ß̃-D-galactopyranoside	Roth (Karlsruhe)
ß-Mercaptoethanol	Roth (Karlsruhe)

Ipecoside, arbutin, raucaffricine, strictosidine and vincoside lactam substrates were kindly provided by Professor Joachim Stöckigt, Pharmacy Institute, University of Mainz, Germany. Apigenin, galloyl, kaempferol, naringenin, salicin and quercetin substrates were obtained kindly from Dr. Thomas Vogt, Leibniz Institute of Plant Biochemistry, Halle, Germany.

2.1.5. Materials and reagents

Kits:

BD SMART™ RACE cDNA Amplification Kit	BD Biosciences (Erembodegem)
Big Dye Terminators Version 1.1	Applied Biosystems (Lincoln)
Gene Racer™ Kit Version F	Invitrogen (Karlsruhe)
M-MLV Reverse Transcriptase (H-)	Invitrogen (Karlsruhe)
Marathon cDNA amplification kit	Clontech (California)
Megaprime™ DNA Labelling System	GE Healthcare (München)
Oligotex® mRNA Mini Kit	Qiagen (Hilden)
pGEM®-T Easy Vektor System I	Promega (Mannheim)
pET 100/D TOPO Expression kit	Invitrogen (Karlsruhe)
QIAprep® Spin Miniprep Kit	Qiagen (Hilden)
QIAquick® Gel Extraction Kit	Qiagen (Hilden)
QIAquick® PCR Purification Kit	Qiagen (Hilden)

SuperScript First-Strand Synthesis System	Invitrogen (Karlsruhe)
ZAP Express® cDNA Synthesis Kit	Stratagene (California)
ZAP Express® cDNA Gigapack® III Gold Cloning Kit	Stratagene (California)

Others:	
Biodyne Membrane	Pall
Filter paper GB 004 Gel blotting paper	Schleicher and Schuell
PD-10 Desalting column	Amersham Biosciences (Freiburg)
Phosphor-image screen	Molecular Dynamics (München)
ProbeQuant™ G-50 Micro Columns	Amersham Biosciences (Freiburg)

2.1.6. Instruments

Autoclave:	Varioklav vapor sterilisator Typ 250T (H+T)
Centrifuge:	Centrifuge 5810R and 5415D and C (Eppendorf, Hamburg), 4K10 and 3K18 (Sigma-Aldrich (Taufkirchen)), Sorvall RC 26 Plus (DuPont).
Electrophoretic:	vertical und horizontal gel apparatus (Biometra) electrophoresis power supply Microcomputer Model E455 (Consort) Gene Genius Bio Imaging Systems (Syngene)
LC/MS:	1100 Series (Agilent Waldbronn, Germany); MS-TOF (Applied Biosystems Lincoln, USA); Turbulon Spray source (PE-Sciex, Concord, ON, Canada)
HPLC:	1100 Series (Agilent)
Measure Instruments:	Storm Phosphorimager (Molecular Dynamics, München), Scintillation counter (Beckman LS 6000 TA), UV Spectrophotometer (Pharmacia Biotech)
PCR-Machine:	Thermal Cycler GenAmp PCR System 9700 (PE Applied Biosystems, Lincoln), PTC 200 Peltier Thermal Cycler (MJ Research), Mastercycler grathent (Eppendorf, Hamburg)
pH-meter:	Inolab (WTW) pH Level 1.
Radioactivity measure:	Storage Phosphor Screen (Molecular Dynamics, München) Storm 860 (Molecular Dynamics, München) Image Eraser (Molecular Dynamics, München) Multi Purpose Scintillation Counter LS 6500 (Beckman Coulter)

Sequencing instrument:	ABI 310 and ABI 3100 Avant Genetic Analyzer (Applied Biosystems, Lincoln)
Power suppliers:	Standard Power Pack P25 (Biometra), E455 (Consort), PHERO-stab 500(Biotech-Fischer)
Sterile chamber:	HERAsafe (Heraeus) Varioklav (H+P Labortechnik AG)
Centrifuge:	Centrifuge 5415 C (Eppendorf, Hamburg), Centrifuge 5403 (Eppendorf, Hamburg), Centrifuge 5810 R (Eppendorf, Hamburg) Picofuge (Stratagene, California) Sorvall RC 26 Plus (Du Pont)
Others:	DU 640 Spectrophotometer (Beckman) Weighing machine Model OC 210-A (Omnilab) Weighing machine Model MC 1 (Sartorius) Hybridising oven Model 7601 (GFL) Cold water bad Model K15 (Haake) Water bad Model 13 A (Julabo) Orbital shaker Model RCT basic (IKA Werke) SPD Speed Vac (Thermo Savant) Thermo mixer 5436 (Eppendorf, Hamburg) UV Stratalinker 1800 (Stratagene, California) Vortex-Genie 2 (Scientific Industries) Bio Imaging System (Syngene) Digital Graphic Printer UP-D895 (Sony) Milli-Qplus machine (Millipore)

2.1.7. Software

DNA and protein sequence data were processed using the program package of DNASTAR Inc. from Lasergene (2003) and Basic Local Alignment Search Tool (BLAST) (Altschul *et al.*, 1990) from the National Center for Biotechnology Information. BLAST from the Munich Information Center for Protein Sequences (MIPS) - *Arabidopsis thaliana* Database (MAtDB) was also used to visualize the distribution of the clone homologies. The deduced amino acid sequence was scanned for the occurrence of conserved patterns using the PROSITE database Hofmann *et al.*, (1999). The prediction of transmembrane helices was carried out using the servers HMMTOP Tusnady and Simon (1998), TMHMM Sonnhammer *et al.*, (1998) and SOSUI from the Tokyo University of Agriculture and Technology. Signal peptides that predict subcellular localization were analyzed by PSORT server Nakai and Kanehisa (1992).

2.2. Plant tissue culture methods

2.2.1. Growth conditions and media

- *Psychotria* medium: Murashige and Skoog (1962) media (MS) with 3% sugar, 10 ml / l vitamins solution and 100 mg / l inositol. The pH was adjusted to 5.7 before autoclaving 20 min 120 °C
- Vitamin solution: in a litter dH_2O are dissolved; 0.5 mg nicotinic acid, 0.5 mg pyridoxine hydrochloride and 0.1 mg thiamine hydrochloride.
- Solid medium: For a solid medium 10 g/l of agar were added

Psychotria ipecacuanha in vitro shoots and roots were cultivated separately. Shoots were cultivated in flasks containing *Psychotria* solid medium, figure 2.1. -A, under sterile conditions. After sub-cultivation in the corresponding medium, the shoot containing flasks were placed in growth chambers under a 16 h light with intensity of 750 lux and at 26±1 °C. Roots were cultivated in 300 ml conical flasks containing 50 ml of *Psychotria* liquid medium in darkness, at 26±1 °C and on a orbital shaker at 90 rpm as standard conditions.

2.2.2. Shoot micropropagation

Micropropagation media: *Psychotria* medium supplemented with 3mg/l 6-benzylaminopurine (BAP) and 0.01 mg / l α-naphthalene acetic acid (NAA)

Stem sections with 1 to 3 nodes were sub-cultivated every 6 to 8 weeks in freshly prepared *Psychotria* solid medium with growth under the above mentioned conditions for the maintenance and multiplication of the plants.

2.2.3. Root culture

- Root induction medium (RIM): *Psychotria* medium supplemented with 2mg/l NAA
- Root multiplication media (RMM): *Psychotria* medium supplemented with 1 mg/l NAA
- Root growth media (RGM): *Psychotria* medium supplemented with 0 to 0.5 mg/l NAA

Complete single *in vitro* plants were sub-cultured on solid RIM under the shoot growth conditions. Plants were kept under the same conditions and medium until roots were harvested, around 8 to 10 weeks later. Root tips, 0.8 to 1.0 cm long, were excised from plants and separately cultivated on solid *Psychotria* medium supplemented with various concentrations of plant growth regulators. Well-developed roots were cultivated in liquid *Psychotria* medium under root growth

conditions. Sub-cultivation of the roots was done every 3 to 4 weeks. Once the roots were multiplied, they were sub-cultivated in RGM under the same root growth conditions.

2.2.4. Methyl jasmonate induced cultures

Effects of methyl jasmonate (MeJa) were tested on the production of alkaloids in tissue cultures of *P. ipecacuanha*. Once the tissues were subcultivated in a freshly prepared medium, 10 µl of 9.2 mM MeJa per 1 ml of media were added to the root liquid medium or onto the *Psychotria* solid media. The time points of analysis for the induced and control roots were: 6, 12, 18, 24, 30, 36, 48, 60, 72 and 120 hours. For induced and control leaves, the time points were: 24, 48, 96, 192 and 240 hours. Two to three repetitions were made for each treatment. Media and growth conditions were the same indicated above without growth regulator supplementation.

2.3. Microbiological methods

2.3.1. Preparation of media and agar plates

- Luria Bertani medium (LB): 1% (w/v) peptone / tryptone, 0.5% (w/v) yeast extract and 1% (w/v) NaCl dissolved in dH2O. The pH was adjusted with NaOH to pH 7.0
- SOC medium: 2% (w/v) peptone / tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, pH 7.0. After autoclaving, the medium was supplemented with 10 mM MgCl2·6H2O, 10 mM MgSO4·7H2O and 0.4% glucose, all from filtered sterile stock solutions.

For solid LB medium in Petri dishes, 1.5% (w/v) agar was added. SOC medium was used for the transformation in E. coli (Maniatis *et al*., 1982).

2.3.2. Transformation by thermal shock

A 200 µl aliquot heat shock competent cells stored at -80 °C was thawed slowly on ice. Plasmid or ligation reaction was added. The preparation was mixed by gentle agitation and incubated on ice for 30 min. After the incubation, the cells were heat shocked in a 42 °C water bad for 30 s and immediately transferred to ice for 2 min. SOC medium pre-warmed to 37 °C was added to the cell preparation and incubated at 37 °C in a orbital shaker at 200 rpm for 1 h. Various sized aliquots were spread over 37 °C pre-warmed LB agar Petri dishes containing the corresponding antibiotic for selection. Finally, Petri dishes with the transformed cells were incubated overnight at 37 °C. Cell colonies were picked and grown overnight in a tube containing 2-5 ml LB liquid media with antibiotic for further work and glycerol stock preparation.

2.4. Electrophoresis gel

2.4.1. DNA agarose gel

- 50X TAE: 242g Tris-Base, 57.1 ml Acetic acid, 100 ml 0.5 M EDTA pH 8.0, H2O up to 1 litter
- 5X Loading buffer: Glycerol (50% w/v), Orange G (0.001% w/v)

The agarose electrophoresis gel allows the separation and identification of DNA because of the size difference among the fragments. For fragments of 0.5-25 kb size, 1X TAE buffer with 1% (w/v) agarose was heated near to boiling, cooled down to 50 °C, ethidium bromide (0.4 µg/ml) was added and then was poured in a horizontal gel chamber. Samples were mixed with the loading buffer, loaded in the gel and run at 100 V. The size of the fragments was determined by comparison with a DNA ladder marker.

2.4.2. RNA agarose gel

- 10X FA buffer: 200 mM MOPS, 50 mM NaAc, 10 mM EDTA and H2O up to 1 litter, pH 7.0
- 1X running buffer: 1x FA buffer, 0.22 M formaldehyde and H2O up to 1 litter
- Loading buffer: 0.1M EDTA (pH 8.0), 0.25% bromphenol blue, 0.25% xylencyanol, 50% glycerol and 100µg/ml ethidium bromide.
- Denaturing buffer 100µl: 13 µl 10X MOPS, 23µl 37% formaldehyde, 64 µl formamide and 20 µl 10X loading buffer

For the analysis of RNA, a formaldehyde agarose gel was used. A volume of 1X FA buffer was heated with agarose 1.2% (w/v), the gel was cooled down to 50 °C, then formaldehyde was added (0.22 M final concentration) and ethidium bromide (0.4 µg/ml); the solution was then poured into a horizontal gel chamber to cool. The samples were prepared with RNA loading buffer and denatured at 65 °C for 5 min, cooled down on ice, loaded on the gel and run at 60-80 V in 1X running buffer.

2.4.3. Protein polyacrylamide gel electrophoresis (PAGE)

Detection of the expressed protein was made by SDS polyacrylamide gel electrophoresis according to protocols in Sambrook *et al.* (1989). Proteins were stained in the gel with Coomassie Brilliant Blue solution.

Polyacrylamide gels	**Resolving gel, 12%**		**Stacking gel, 4%**	
H_2O	1.5	ml	1.5	ml
1.5 M Tris-HCl pH 8.8	1.25	ml	0.625	ml
20% SDS (w/v)	50	µl	0.025	µl
Acrylamide / bis-acrylamide 30%	2.0	ml	312	µl
10% Ammonium persulfate	40	µl	20	µl
TEMED	10	µl	5	µl
Total volume	5	ml	2.5	ml

2.5. Isolation of nucleic acids

2.5.1. DNA mini-preparations

DNA mini-preparation consists of the isolation of plasmids from bacteria cell cultures. Five ml of LB medium with the corresponding antibiotic were inoculated with the selected bacteria and cultured at 37 °C overnight and 180 rpm in an orbital shaker. Purification of the plasmid was

achieved from 3 ml of overnight culture using the QIAprep Spin Miniprep kit following the micro centrifuge method according to the manufacturer's protocol.

2.5.2. Plasmid miniprep in 96-well microtiter plate

- P1 solution (keep at 4°C): 50 mM Tris-Cl (pH 8.0), 10 mM EDTA, 100 µg/ml RNAse A.
- P2 solution: 0.2N NaOH, 1% SDS
- P3 solution 3M K-acetate (pH 5.5)
- P4 solution (fresh preparation) 6.1M KI

Bacteria were inoculated into 1.0 ml 2X LB liquid media with the corresponding antibiotic in each cell from a 96 deep well block and grown 24 h at 37°C in an orbital shaker, 300 rpm. Afterwards, the block was centrifuged 5 min at 1500 g. The pellets were re-suspended in 80 µl of P1 solution by vortex. Cells were lysed in 80 µl of P2 solution with brief mixing and incubation at room temperature for 2 min. P2 solution was neutralized with 80 µl of P3 solution and incubated for 2-5 min. The block was centrifuged at 16.000 x g for 15 min. Supernatants were taken carefully from the block and added to a Multiscreen Flat Bottom (MFB) plate prepared with 150 µl of P4 solution and mixed, centrifuged at 1000 x g for 5 min, twice washed with 200 µl 80% EtOH and centrifuged at 1000 x g for 5 min and then an additional centrifugation for 10 min to remove residual EtOH. Finally, the plasmids were eluted with 60 µl of TE or sterile H2O in a microtiter plate, incubated 2-5 min, and centrifuged at 1000 x g for 5 min. The microtiter plate was sealed with Parafilm for storage.

2.5.3. DNA extraction from an agarose gel

DNA fragments from endonuclease digestion, ligations or PCR products were isolated and purified from agarose electrophoresis gels with a QIAquick gel extraction kit. After electrophoretic separation, the fragments, visible under UV light in presence of ethidium bromide, were cut out directly from the gel and handled according to the manufacturer's protocol.

2.5.4. Purification of PCR products

The purification of the PCR products from the remaining reaction components was achieved with QIAquick PCR purification kit according to the manufacturer's protocol.

2.5.5. Total RNA extraction

- Extraction buffer: 4M guanidine thiocyanate, 50mM sodium citrate, 0.5% sarcosyl, 0.1 M *β*-mercaptoethanol (added just before use). The buffer was mixed 1:1 with Roti-aqua-phenol

Total RNA isolation was performed by the guanidine thiocyanate method described by Chomczynski *et al.* (1987). Plant tissue was frozen with liquid nitrogen, ground with a mortar and mixed with 7.0 ml phenol buffer per 500 mg ground tissue. The mixture was incubated 5 min at room temperature; afterwards 2 ml of chloroform were added and vigorously mixed by vortexing. The RNA contained in the supernatant was taken into a new tube and precipitated with isopropanol at RT. The pellet was rinsed twice with 75% ethanol. Finally, the ethanol was completely evaporated before the RNA was re-suspended in 50-100 µl of sterile distilled water. The quality and quantity of RNA were calculated with a spectrophotometer, according to Sambrook *et al.* (1989).

2.5.6 Isolation of poly (A)+ RNA

Poly (A)+ RNA (mRNA) was isolated from total RNA using an OligotexTM mRNA mini kit according to the manufacturer's protocol. The poly-A tail of the mRNA was hybridized under high-salt concentration to a dTTP oligomer that was coupled to a solid-phase matrix. After washing, the hybridizing mRNA was released from the matrix by lowering the ionic strength.

2.6. Determination of concentrations

2.6.1. Determination of nucleic acid concentrations with UV spectrophotometers

The concentration and quality of nucleic acids was estimated according to Sambrook *et al*. (1989).

2.6.2. Determination of the protein concentration

The protein concentration was estimated with the method of Bradford (1976) and the Bio-Rad protein assay solution (Bio-Rad).

2.7. Polymerase Chain Reaction (PCR)
2.7.1. Standard PCR

A standard PCR mixture consists of:

DNA Template	100 - 200 ng	
20 µM primer A	1 µl	
20 µM primer B	1 µl	
2.5 mM dNTP mix	1 µl	
25 mM MgCl2	1 µl	
DNA polymerase	1.5 U	
10 x PCR buffer	5 µl ddH2O	up to 50 µl

General PCR-Cycle:

Denaturalization	94 °C	2-5 min	
Denaturalization	94 °C	30-45 s	25-30
Annealing	Tm + 5°C	30-45 s	
Extension	72 °C	1-5 min	
Final extension	72 °C	5-7 min	
Stop	4 °C		

Taq-polymerase (Thermophilus aquaticus) or Pfu-polymerase (Pyrococcus furiosus) was used as thermo stabile DNA-polymerases. The annealing temperatures, time and number of cycles were dependent on the melting temperature of the primers used.

2.7.2. Reverse Transcriptase reaction PCR (RT-PCR)

For the transcription of mRNA to cDNA of a specific clone it was used the M-MLV reverse transcriptase, RNase H minus, point mutant from Promega and according to the manufacturer's protocol.

2.8. cDNA library construction

Phage storage buffer (PSB): 8 mM MgSO4, 50 mM Tris-HCl pH 7.5, 100 mM NaCl
DMSO buffer: PSB with 20% (v/v) DMSO

A cDNA library was constructed using the ZAP Express® cDNA Synthesis Kit and ZAP Express® cDNA Gigapack® III Gold Cloning Kit according to the manufacturer's protocol. Isolated mRNA (2.93 µg) from in vitro roots was used as template for cDNA synthesis. For the

filtration step it was used the ProbeQuantTM G-50 Micro Columns according to the manufacturer's protocol. The storage of the cDNA bank was done in PBS at –20°C and in DMSO buffer at –80°C. The library was amplified with a Marathon cDNA amplification kit and the primers 5'TriplEx 5'long and 3'TriplEx 3'long according to the manufacturer's protocol.

2.9. Degenerated primer method

Through a RT-PCR the degenerated primer method is supposed to isolate partially a homologous gene from poly $(A)^+$ RNA using primers which come from conserved regions of proteins with a near related function among them and with the proposed protein goal. The conserved regions should allow the design of 17-20 long nucleotide primers with a similar melting temperature, the primers should be of low degeneracy, the 3'-terminus must have a C or G nucleotide and not be degenerated and both primers, 5'-forward and 3'-forward, should be designed under the same considerations. The designed degenerate primers used for the isolation DIS and DIIS genes are listed in the section (2.1.2.2). The mRNA used was obtained from leaves of *in vitro* cultured plants.

2.10. Rapid Amplification of cDNA Ends (RACE)
2.10.1. Construction of the cDNA

After the construction of the cDNA library, an additional step was required to obtain the full-length clone of the ipecoside glucosidase. With the BD SMART™ RACE cDNA amplification kit and the gene specific primers (GSP) it was possible to generate a full-length clone of ipecoside glucosidase performed according to manufacture's protocol.

2.11. Sequencing of DNA

DNA sequencing was achieved by the didesoxyribonucleotide method (Sanger *et al.*, 1977). For this, didesoxyribonucleotides marked with a fluorescent dye was detected by a laser. The DNA sequencing was obtained with the protocol Big Dye Terminators version 1.1. The purification of the sequencing reaction products was achieved with Sephadex G-50 Superfine on a MultiScreen 96-well plate. The purified sequencing products were analysed on an ABI 310 or ABI 3100 Genetic Analyzer. The sequencing primers used for the ipecoside glucosidase clone were ph0212_60_Gluco, Ph0212_60_Gluco_B, Ph0212_Gluco_BII, Ph0212_60_Glu_C, ph0212_

60_Glu_D, and Gluc-621, and for the strictosidine synthase-like clone, the sequencing primers used were ph0412_62_Synth, Ph0412_62_Synth_B, Ph0412_62_Syn_C, Ph0412_Synth_CII, ph0412_62_Syn_D, and Str_Synth_am1020.

Reaction components:

1-5 µl Plasmid-DNA (500 ng)
1 µl Primer (10 pmol)
4 µl Big Dye Mix V 1.1.
up to10 µl ddH_2O
Cycle:

Denaturalization	96 °C	10 s	
Annealing	50 °C	5s	} 25
cycles			
Extension	60 °C	4 min	
Reaction stops	4 °C		

2.12. Cloning techniques
2.12.1. Restriction reactions

The standard restriction endonuclease digestion was set up as given below. The enzymatic reaction was placed at 37°C for at least 2 hours.

DNA	1-10 µg
Restriction enzyme	1 U/ µg DNA
10x reaction buffer	5 µl
ddH_2O	up to 50 µl

2.12.2. Ligation

- Standard ligation components: 50 ng of fresh PCR product, 1 µl 10X ligation buffer, 25 ng of vector, 1 µl DNA ligase and H_2O up to 10 µl.

The ligation reaction was incubated at 14 °C overnight. The recommended molar proportion of the vector : insert should be 1:3. The amount of the insert was calculated as follows

$$\frac{\text{ng vector x kb insert}}{\text{kb vector}} \times \text{molar proportion insert: vector} = \text{ng insert}$$

For ligation of the newly generated clones in the pET100/D-TOPO expression vector, an additional amplification with 5'-adaptor containing primers was required. For ipecoside glucosidase, the primers 5-Gluc3+adap and Glu_3 were used and for the strictosidine synthase-like clone the primers DIS_5 and DIS_3 were used. The ligation with a pET 100/D TOPO Expression kit and the amplification with a BD Advantage™ 2 PCR enzyme system were achieved according to the manufacturer's protocols.

2.13. Protein expression
2.13.1. Expression of recombinant protein in *E. coli*

- His-tag Wash Buffer (HWB): 50 mM Tris-HCl pH 7.0, 500 mM NaCl, 2.5 mM imidazole, 10% glycerol, 10 mM β-mercaptoethanol
- His-tag Lysis Buffer (HLB): HWB plus, 1% Tween 20, 750 µg/ml lysozyme
- His-tag Elution Buffer (HEB): 50 mM Tris-HCl pH 7.0, 500 mM NaCl, 250 mM imidazole, 10% glycerol, 10 mM β-mercaptoethanol
- TALON Metal Affinity Resin (BD Biosciences)
- TALON 2 ml disposable gravity columns Column (BD Biosciences)
- PD-10 Column Sephadex G-25 M (GE Healthcare)
- Ipecoside glucosidase storage buffer: 20mM tris (pH 8.0), 20mM NaCl, 1mM EDTA, 20% glycerine
- Strictosidine synthase-like protein storage buffer: 10 mM tricine (pH 7.5), 3mM EDTA, 10 mM NACl, 10% glycerine

Escherichia coli Rosetta 2 strain was transformed with the pET100/D-TOPO vector containing the clone of interest and selected on Luria–Bertani (LB) medium supplemented with 50 mg/l carbenicillin disodium, 34 mg/l chloramphenicol and 1% glucose. For purification of the recombinant protein, 5 ml of overnight freshly grown bacterial culture were inoculated into 2.0 l of the above mentioned LB medium and incubated at 28 °C in an orbital shaker at 180 rpm. IPTG was added to a final concentration of 1.0 mM in the medium when the culture reached an $OD_{600} = 0.6$. Cells were incubated further at 18 °C until the $OD_{600} = 1.6$, then the cells were collected by centrifugation for 15 min at 10,000 x g, at 4 °C. The cell pellet from the 2.0 l culture was resuspended in 40 ml HLB and incubated on ice for 40 - 60 min. From this step on, except in specific cases, the protein was maintained at 4 °C. Cells were lysed by sonication at the end of the incubation; finally the crude extract was centrifuged 15 min at 10,000 x g. Aliquots were taken from each preparation step and frozen. These samples were analysed by SDS-PAGE to observe the presence of the protein in induced and un-induced cultures, crude extracts and extracted cultures, pellet and supernatant extractions. The protein of interest was purified from the crude protein extract with a Talon Metal Affinity Resin according to the manufacture's protocol. The protein extract was incubated with 1 ml of His-Tag Talon resin 2 h on an orbital shaker. The Talon resin was collected by centrifugation 2 min at 700 x g. The protein-bound resin was twice resuspended in 20 ml HWB, incubated 10 min on an orbital shaker and collected by centrifugation 2 min at 700 g. The Talon resin was resuspended in 5 ml HWB, loaded onto a Talon affinity column and rinsed with 3 ml HWB. The protein bond to the Talon resin was then eluted with 3.5 ml of HEB and collected in seven 0.5 ml fractions. The fractions with the major protein concentration were put together up to a volume of 2.5 ml and loaded onto a desalting PD-10

column equilibrated with the corresponding storage buffer (ipecoside glucosidase storage buffer or strictosidine synthase-like protein storage buffer). The protein was eluted with 3.0 ml of the storage buffer, divided into 2 x 1.5 ml volumes and stored at 4 °C or -20 °C. The protein concentration was determined according to Bradford (1976).

2.14. Enzymatic reactions

2.14.1. Activity measurement in *E. coli* expressed recombinant enzyme

Estimation of the pH optimum for ipecoside glucosidase was achieved with reaction preparations with 0.03 µg enzyme per 100 µl of reaction and 100 µM ipecoside incubated 30 min at 30 °C in 0.1 M citrate-phosphate buffer at pH 3.5, 4.0, 4.5, 5.0, 5.5 and 6.0. The temperature optimum was determined under the same conditions mentioned above with a pH of 3.5 and the temperatures 20, 30, 40 and 50 °C. In order to intercept putative precursors of cathenamine formed immediately after hydrolysis of strictosidine, the enzymatic reaction was carried out in presence of reducing agent $NaBH_3CN$ which was expected to reduce aldehyde groups and thus prevent it from further conversions (Stöckigt and Zenk, 1977; Winkel-Shirley, 2001; Stöckigt, 1995). The reactions were terminated by addition of 50 µl MeOH. After centrifugation (17,000 x g, 5 min) the supernatants were analyzed by HPLC to determine the specificity of the reaction.

Strictosidine synthase-like enzyme reactions were tested with the buffers citrate-phosphate, potassium phosphate, sodium phosphate and tricine at different pHs varying from 4 to 8.

2.15. Chemical substances analysis

2.15.1 Alkaloid extraction from plant tissues

For the extraction of alkaloids from *in vitro* cultured leaves and roots, between 300 and 800 mg tissue were used. First, the tissue was frozen and powdered in liquid nitrogen with a mortar and pestle, 5 ml 70% ethanol were added and the suspension mixed. A sample of 1 ml was taken, 60 µg of dehydrocodeine (dhc) was added as internal standard and the pH was adjusted to 8.0 - 9.0 with $NaHCO_3$. The preparation was vortexed for 2 min and centrifuged at 17,000 x g in a table centrifuge for 5 min. The upper phase was transferred to a new tube and the volume was evaporated 300 µl with the rotavapor. Finally, the alkaloids were extracted two times with 600µl of ethyl acetate by vortexing and centrifugation. The ethyl acetate was completely evaporated with the rotavapor and then alkaloids were resuspended in 1 ml 70% ethanol.

2.15.2. Alkaloid extraction from liquid medium

Liquid medium in which roots were grown for (1 - 30 days) was filtered before the alkaloids extraction treatment. Samples of 50 ml of liquid medium were extracted directly with 40 ml of ethyl acetate and finally re-suspended in 500 µl of 70% ethanol. As internal standard, 60 µg of dehydrocodeine (dhc) were added.

2.16. Chromatographic methods

2.16.1 High Performance Liquid Chromatography (HPLC)

Mobile phases:

A1 - 98% H_2O, 2.0% HCN, 0.01% H_3PO_4	B1 - 90% HCN, 2.0% H_2O, 0.01% H_3PO_4
A2 - 98% H_2O, 2.0% HCN, 0.2% HOAc	B2 - 90% HCN, 2.0% H_2O, 0.2% HOAc
A3 - 98% H_2O, 2.0% HCN, 0.2% HCOOH	B3 - 90% HCN, 2.0% H_2O, 0.2% HCOOH

General conditions for HPLC analysis were a LiChrospher 60 RP-select B (250- 4, 5µm) column, an injection volume of 5 - 15 µl, a flow rate of 1.0 ml/min and the wavelength detection at 210, 255 and 285 nm. For each substance or group of related substances, the appropriate solvent system was used. ALK 46 method used the solvents A1 and B1 with the following gradient for B1, from 0 to 25 min 0% to 46%, from 25 to 26 min to 100%, 26 to 33 min 100%, 33 to 35 min to 0%, 35 to 40 min 0%. FORM 37 method used the solvents A2 and B2 with the following gradient for B2, from 0 to 20 min 0% to 37%, from 20 to 21 min to 100%, 21 to 24 min 100%, 25 to 29 min to 0%. STRIC-1 method used the solvents A3 and B3 with the following gradient for B3, from 0 to 1 min 15% to 25%, from 1 to 7.5 min to 40%, 7,5 to 10 min 40%, 10 to 10.5 min to 85%, 10.5 to 15 min 85%, 15 to 15.5 min to 15%, 15.5 to 20 min to 15%. The method ALK 46 was applied for general analysis of samples from alkaloid extraction from plant tissues and for the emetine and cephaeline standards. It was also used for analysis of the enzymatic reactions of ipecoside glucosidase. The method STRIC-1 was used for general analysis of non-polar alkaloids of the enzymatic reactions of ipecoside glucosidase. The method FORM 37 was used for the analysis of enzymatic reactions of the strictosidine synthase-like protein and also for general preparation of samples for further enzymatic reactions and LC-MS analysis.

2.16.2 Liquid Chromatography – Mass Spectrometry (LC-MS, TOF)

Electrospray ionization - mass spectrometry (ESI-MS) measurements and LC separations were carried out on a Mariner TOF mass spectrometer (Applied Biosystems, Lincoln, USA) equipped with a Turbulon Spray source (PE-Sciex, Concord, ON, Canada) using an LC1100 series system of Agilent (Waldbronn, Germany), adapted to flow rates at 0.2mL/min. Samples were injected (2µL) on a Superspher 60 RP-select B column (125x2 mm, 5µm). The following LC conditions were used: solvent A, MeCN : H_2O (2:98) and solvent B, MeCN : H_2O (98:2), containing 0.2% formic acid in both cases. The gradient increased from 0% to 46% B in 25min, to 90% in 1 min and held at 90% for 7min, post time was 5min. The TOF mass spectrometer was operated in the positive ion mode, nebulizer gas (N_2) flow was 0.5 L/min, curtain gas (N_2) flow 1.5 l / min and heater gas (N_2) 7 L/min. The spray tip potential of the ion source was 5.5 kV, heater temperature was 350^{o}C, nozzle potential was 200 V, quadrupole temperature was 144^{o}C, and detector voltage 1.95 kV. The other settings varied depending on tuning.

2.16.3. LC-ESI-MS/MSSI-MS/MS

The positive ion ESI mass spectra of strictosidine and cathenamine were obtained from a Finnigan MAT TSQ Quantum Ultra AM system equipped with a hot ESI source (HESI, electrospray voltage 3.0 kV, sheath gas: nitrogen; vaporizer temperature: 50 ^{o}C; capillary temperature: 250 ^{o}C; The MS system is coupled with a Surveyor Plus micro-HPLC (Thermo Electron), equipped with a Ultrasep ES RP18E-column (5 µm, 1x100 mm, SepServ). For HPLC, a gradient system was used starting from H_2O:CH_3CN 85:15 (each of them containing 0.2% HOAc) to 10:90 within 15 min; flow rate 50 µl min^{-1}. The collision-induced dissociation (CID) mass spectra of compound during the HPLC run were recorded with a collision energy of 25 eV for the $[M+H]^+$-ions at *m/z* 531 (strictosidine) and 351 (cathenamine), respectively (collision gas:argon; collision pressure: 1.5 mTorr). The multiple reaction monitoring (MRM) measurements were carried out by using the following transitions (scan time 0.2 sec, peak width 0.7): strictosidine: *m/z* 531 → *m/z* 514, *m/z* 531 → *m/z* 352; cathenamine: *m/z* 351 → *m/z* 249, *m/z* 351 → *m/z* 170.

3.1. Plant tissue culture

3.1.1. *Psychotria ipecacuanha* shoots micropropagation

Psychotria ipecacuanha in vitro plants were cultivated and multiplied successfully on micropropagation media (2.2.2.) (fig. 3.1.A). For sub-cultivation, stem sections with 1 to 3 nodes were used. From 100 *in vitro* plants brought from Costa Rica, it was possible to generate more than 4000 *in vitro* plants in one year. During the multiplication period, plant material was taken and used for diverse experiments and preparations like the induction and cultivation of roots, material for DNA and RNA extractions, alkaloids extraction and for maintenance of stock material.

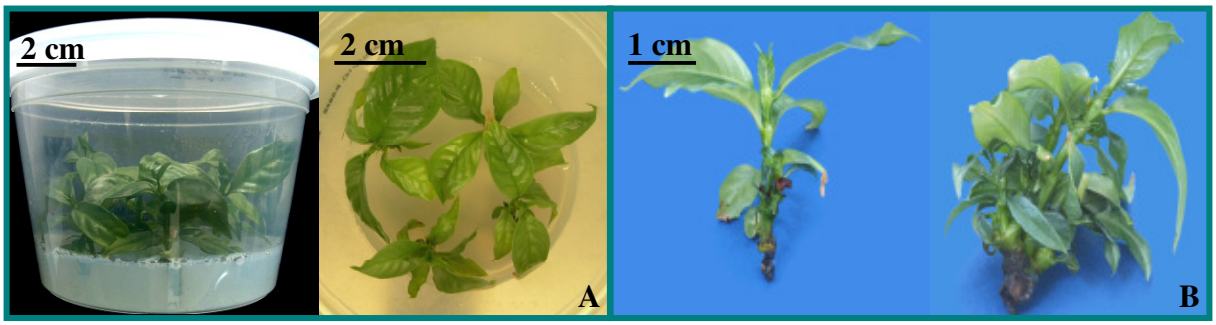

Figure 3.1.A- *In vitro* plants of *Psychotria ipecacuanha* cultivated on micropropagation media. B- Typical single plant and plant with several newly generated shoots.

3.1.2. *Psychotria ipecacuanha* root culture

Psychotria ipecacuanha plants micropropagated by sub-cultivation of stems did not generate roots without an external induction. For the generation of the new roots, plants with a well-developed aerial part (fig. 3.1 B) were placed on root induction medium (RIM) (2.2.3.). After 10 weeks, it was observed 1 cm roots growing from a dark brown callus at the base of the plants (fig. 3.2.).

Figure 3.2. Plant with several shoots showing new induced roots. The roots grew around callus from the media submerged stem.

The roots were separated from the plants and grown on the same medium, freshly prepared. Once the roots were growing and developing stably, around 2 weeks later, the effect of supplementation with α-naphthalene acetic acid (NAA) or 6-benzylaminopurine (BAP) in the *Psychotria* medium with the concentrations 0.5 µg/l and 1.0 µg/l each and in combination with 2.0 µg / l of NAA and 0.02µg / l of BAP on root development was tested.

An evaluation of the root development was done 1, 4 and 8 weeks after the subcultivation in media containing the different treatments. Roots cultured without growth regulators continued growing, but new roots were not generated (fig. 3.3 A-C). Benzylaminopurine, in concentrations of 1.0 µg/l (fig. 3.3 D-F) or lower (data not showed), did not have a major effect on root production and growth. α-Naphthalene acetic acid (0.5 - 1.0 µg/l) induced a major production of roots (fig. 3.3 G-L), nevertheless the elongation and development of the new roots was also reduced and, in some cases, the roots grew with callus characteristics, uncontrolled cell growth and fragile tissue (fig. 3.3.-I). A better development of roots was observed when induced roots were subcultivated on a growth regulator-free *Psychotria* medium. A combination of 2.0 µg/l of NAA and 0.02µg/l of BAP increased the number of roots, but the effect on the elongation and growth was stronger and the root tips became dark and unhealthy (data not show).

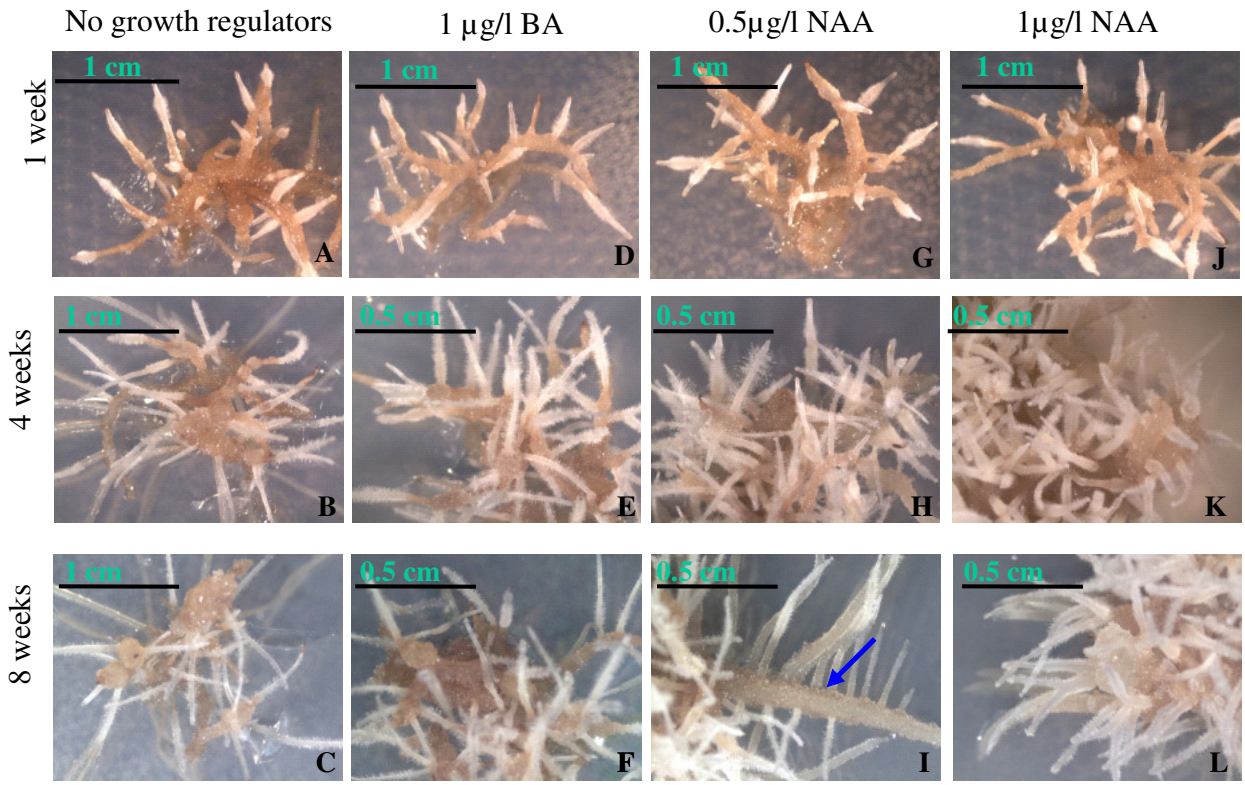

Figure 3.3. Root cultures on a *Psychotria* medium containing different growth regulators and concentrations after 1, 4 and 8 weeks. Arrow indicates a root with callus characteristics.

Solid media grown roots formed clusters with calli in the middle and new roots on the periphery. Each cluster was subcultivated then in conical glassware (fig. 3.4) with root multiplication media (RMM) and root growth media (RGM) (2.2.3.).

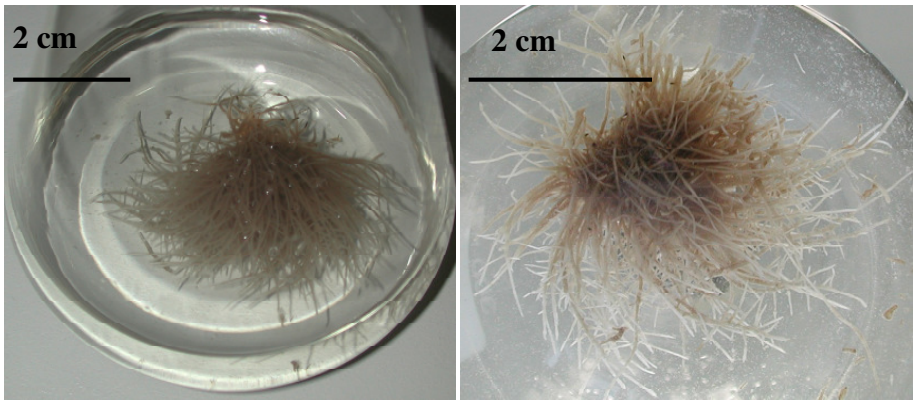

Figure 3.4. *P. ipecacuanha* roots culture in RGM (without growth regulators) 8 weeks after subcultivation from the solid media culture.

In the middle of the cluster, the roots were connected by a light brown fragile callus. Therefore, the roots were separated easy from the clusters without injury. This characteristic allows the continued multiplication and analysis of root without growth alterations due to manipulation.

3.2. Qualitative analysis of the alkaloid content in tissue cultures
3.2.1. Identification of the main ipecac alkaloids

The extraction of ipecac alkaloids from leaves of *in vitro* plants and root cultures of *P. ipecacuanha* was achieved successfully according to the protocol (2.15.1). Cephaeline and emetine were identified by comparison with standards and fragmentation analysis with the method STRIC-1 (2.17.1) by LC/ESI-Selected Reaction Monitoring (fig. 3.5).

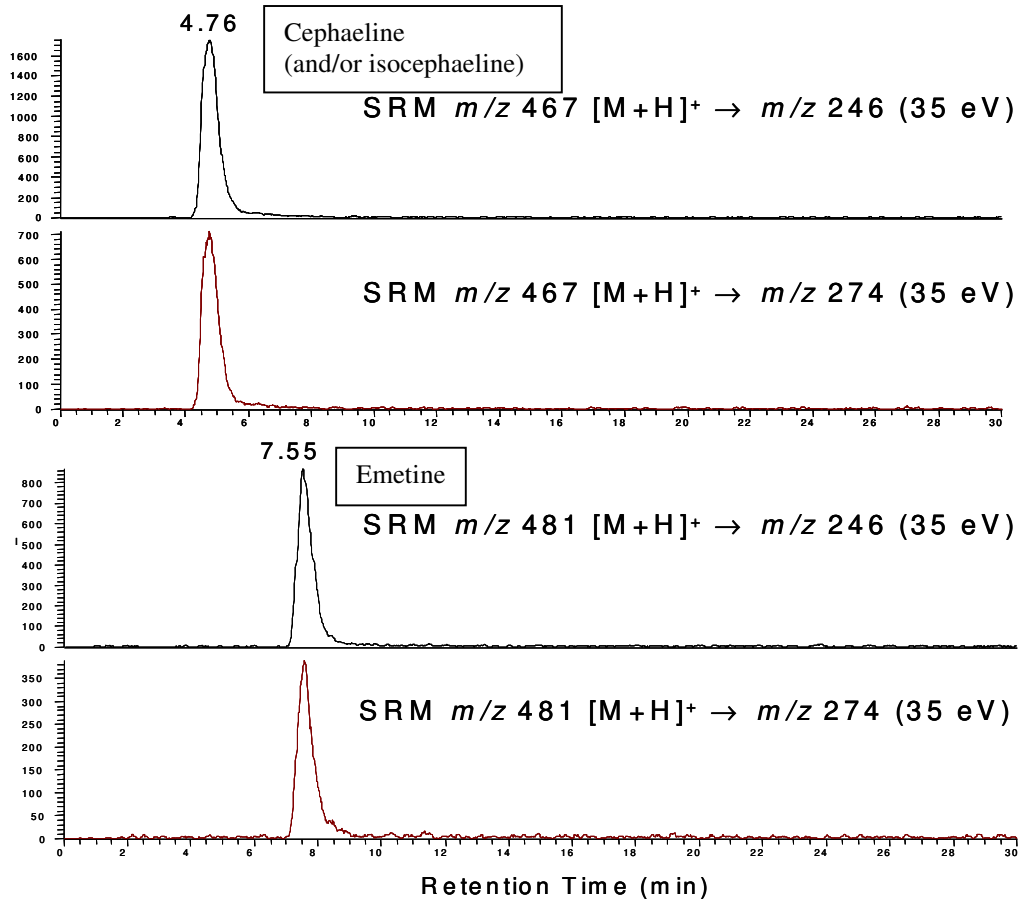

Figure 3.5 Identification of the alkaloids emetine and cephaeline from tissue culture extraction by typical ion fragmentation in LC/ESI-SRM.

The detection of cephaeline $[M+H]^+$, *m/z* 467, t_R 4.76 min and emetine $[M+H]^+$, *m/z* 481, t_R 7.55 min was corroborated in both cases with the fragment at *m/z* 274 and the fragment at *m/z* 246 (fig. 3.6). The specific stereochemistry of cephaeline or/and isocephaeline was not established.

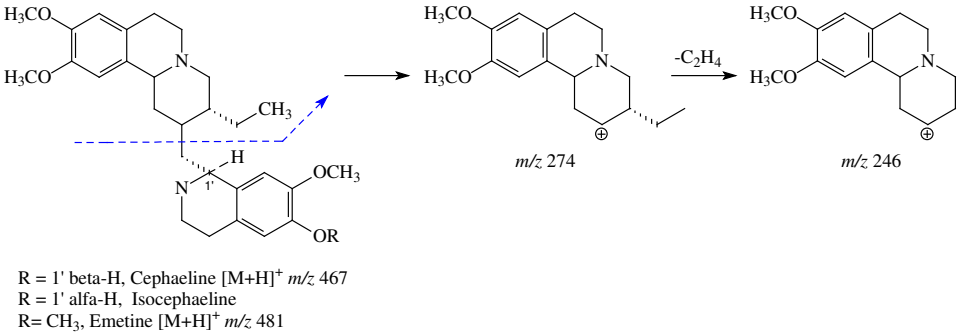

Figure 3.6 Structures of mass spectral fragmentation of emetine and cephaeline determined by LC/ESI-Selected Reaction Monitoring.

3.2.2. Additional ipecac alkaloids in *P. ipecacuanha*

In the general screening of the alkaloid extractions by LC-MS/TOF and the method ALK-40, the known intermediates for the ipecac alkaloid pathway were also identified: ipecoside $[M+H]^+$, *m/z* 566, t_R 17.7 min; *O*-methylipecoside $[M+H]^+$, *m/z* 580, t_R 20.7 min; protoemetine $[M+H]^+$, *m/z* 318, t_R 17.5 min; protoemetinol $[M+H]^+$, *m/z* 320, t_R 17.0 min; emetamine $[M+H]^+$, *m/z* 477, t_R 22.7 min; psychotrine $[M+H]^+$, *m/z* 465, t_R 14.0 min and *O*-methylpsychotrine $[M+H]^+$, *m/z* 479, t_R 22.0 min (fig. 3.7).

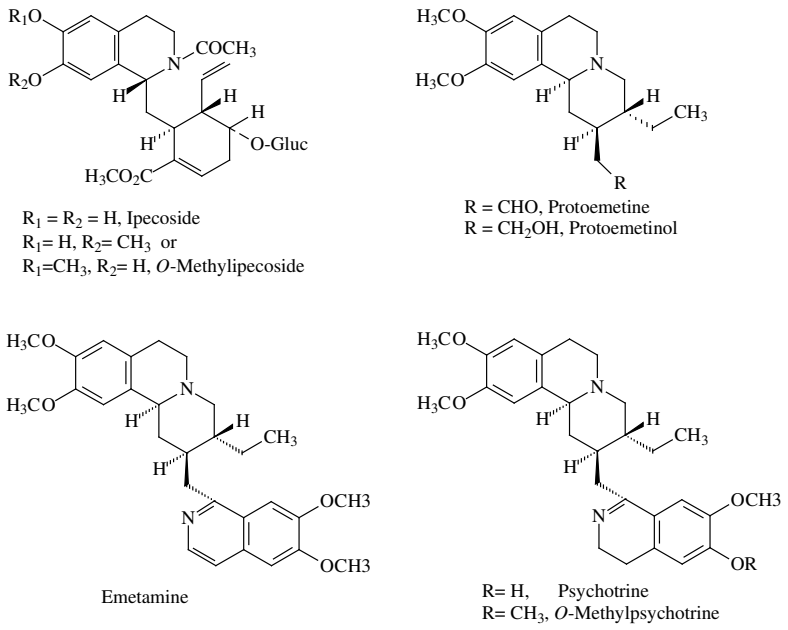

Figure 3.7 Structure of ipecac alkaloids detected in extracts of tissue cultures of *P. ipecacuanha* by LC-MS/TOF.

Applying the LC-MS/TOF technique, the MS spectra showed additional intense signals of the sodium adducts of ipecoside $[M+Na]^+$, *m/z* 588 (fig. 3.8) and *O*-methylipecoside $[M+Na]^+$ *m/z* 602 (fig. 3.9). A comparison of the key ion fragments indicated the presence of the

sodium adduct of *O*-methylipecoside by an increase of 14 mass units with respect to the sodium adduct of ipecoside (schemes in fig. 3.8 and 3.9).

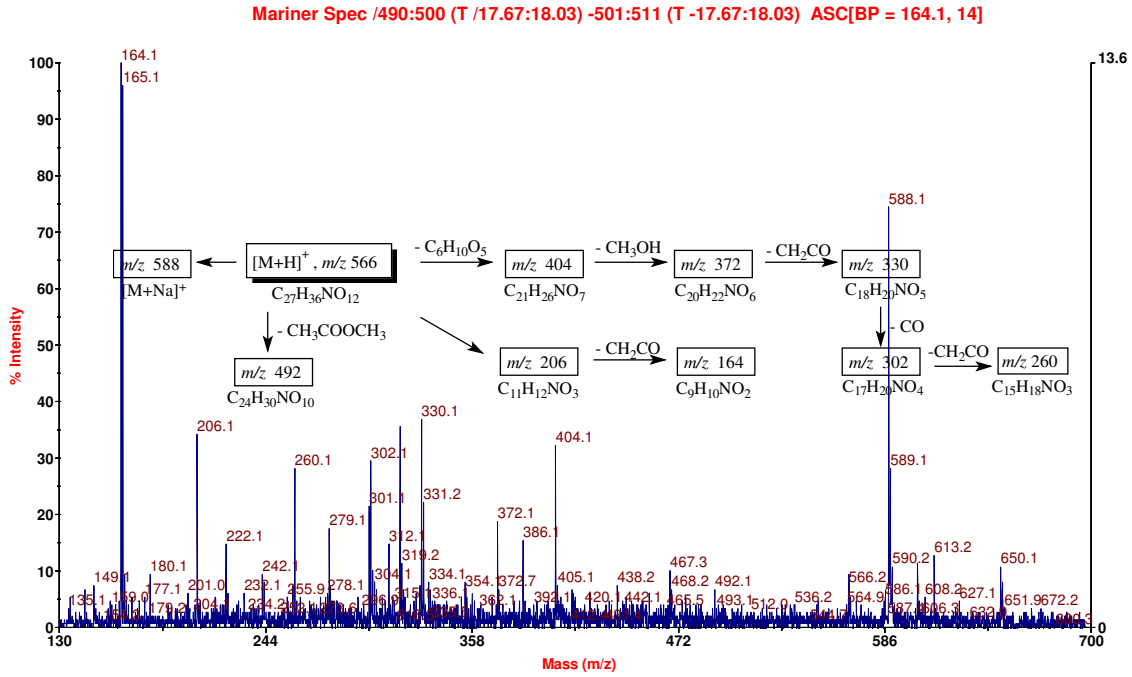

Figure 3.8. ESI-MS spectra and key ion fractionation scheme of the sodium adduct of ipecoside obtained by LC-MS/TOF.

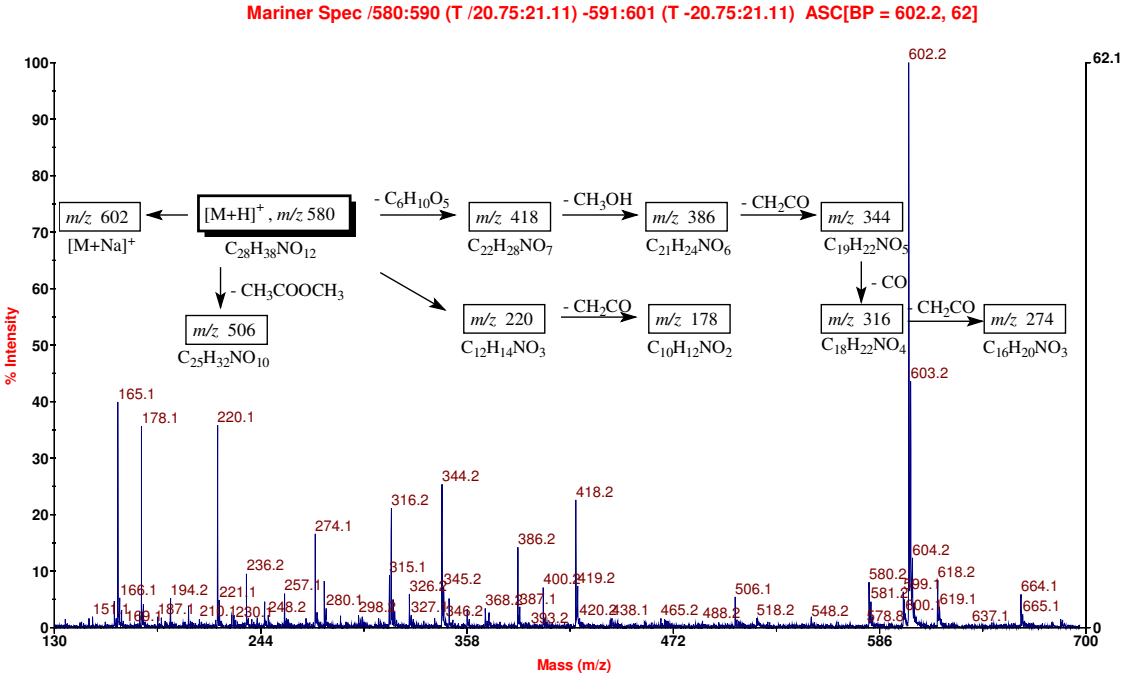

Figure 3.9. ESI-MS spectra and key ion fractionation scheme of the sodium adduct of *O*-methylipecoside obtained by LC-MS/TOF.

3.2.3. Elicitation of tissue cultures with methyl jasmonate

The concentration of cephaeline and emetine was estimated for leaves and roots under *in vitro* culture conditions for *P. ipecacuanha* based on standard compounds and the internal standard

dihydrocodeine (dhc) according to the section (2.2) and determined by LC/ESI-SRM. From alkaloid extractions of *in vitro* plant leaves, the concentration of cephaeline was estimated to be 150 - 300 µg/100 mg fresh material weight (fw) and emetine to be 120 - 230 µg/100 mg fw (ctrl in fig. 3.10-A). From alkaloid extractions of root cultures, the concentration of cephaeline was 30 - 70 µg / 100 mg fw and emetine 2 - 3.5 µg/100 mg fw (ctrl in fig. 3.11-A). The effect of 100 µM methyl jasmonate added to tissue cultures on the accumulation of cephaeline, emetine and the other detected ipecac alkaloids over the time was analysed. A clear induction effect on the alkaloid production could not be inferred from this study for either leaves (fig. 3.10-A) or roots (fig. 3.10-B). However, it was notable that at 12 and 48 hours there was a transient increase in the concentration of cephaeline in both leaves and roots. Emetine presented also a slight concentration increase at 48 hours in leaves; while in roots the concentration was lower.

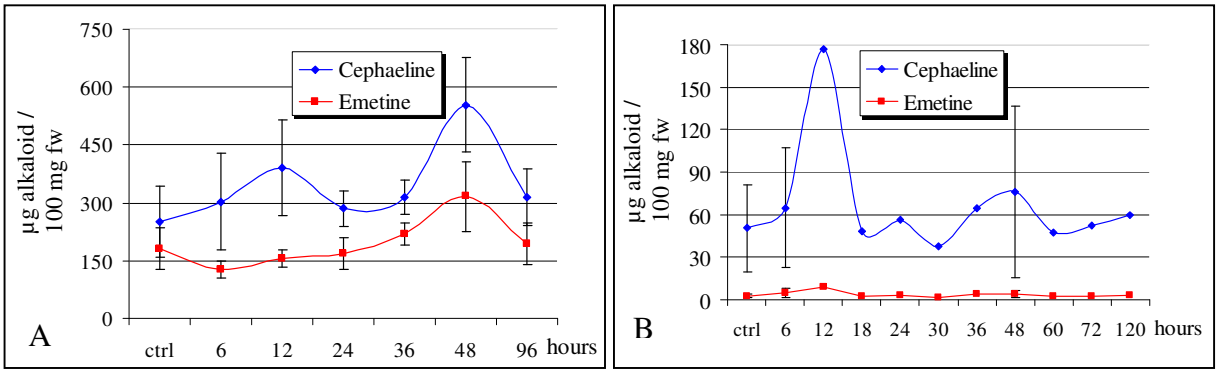

Figure 3.10. Effect induction of tissue cultures with 100 µM methyl jasmonate on the concentrations of cephaeline and emetine at various times after addition of elicitor. A- leaves (std dev of n =3) and B- roots (n =1 and 2). The estimations are given in µg alkaloid per 100 mg of fresh tissue weight (fw); ctrl - not induced tissues.

For the remainder of the identified ipecac alkaloids, it was not possible to establish a precise quantitation due to a lack of standards. In general, an overall induction of alkaloid accumulation was not observed in leaves (fig. 3.11-A) or roots (fig. 3.11-B).

In summary, the chemical analysis showed the presence of the typical ipecac alkaloids in *P. ipecacuanha* tissue cultures. Cephaeline and emetine were clearly identified and quantified for each organ, roots and shoots, cultivated separately. With the exception of ipecoside which was found only in roots, other identified ipecac alkaloids were found in both roots and leaves. No overall induction effects on the production of alkaloids in tissue cultures in presence of methyl jasmonate were observed. However, a transient increase in the alkaloid accumulation at 48 hours was observed.

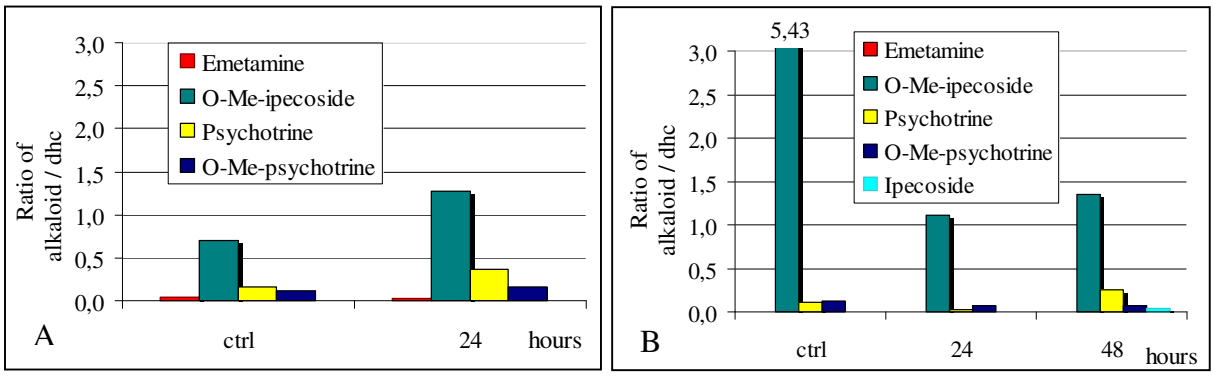

Figure 3.11. Effect on the relative concentration of various ipecac alkaloids after induction of tissue cultures with 100 µM methyl jasmonate. The values were determined by the ration of the integrated area of the extracted molecular ion for the various alkaloids compared to that of the internal standard dihydrocodeine (dhc). A- leaves and B- roots. (n =1); ctrl - not induced tissues.

3.3. Cloning of alkaloid biosynthesis genes from *P. ipecacuanha*.
3.3.1. cDNA library construction

A complementary DNA (cDNA) library was produced (2.8.) using *P. ipecacuanha in vitro* roots under the conditions described in 2.2.5. From 400 µg of extracted total RNA per 1.0 g fresh weight root tissue culture (3.1.2.) 2.93 µg mRNA were isolated (2.5.6.). The synthesis of the first-strand cDNA was done using the mRNA. The synthesized radioactive first and second-strand cDNAs were visualized before and after a filtration step by a Phosphor imager screen detector and a Phosphor imager Storm 860 apparatus (fig. 3.12-A). The cDNA transcripts were fractionated and also visualized (fig. 3.12-B). Fractions from 8 to 12 were combined and used to generate the library. Around 40 ng of cDNA from the pooled fractions, 2.0 µl of 20 ng/µl (fig. 3.12-C), were ligated to the ZAP Express vector for further analysis.

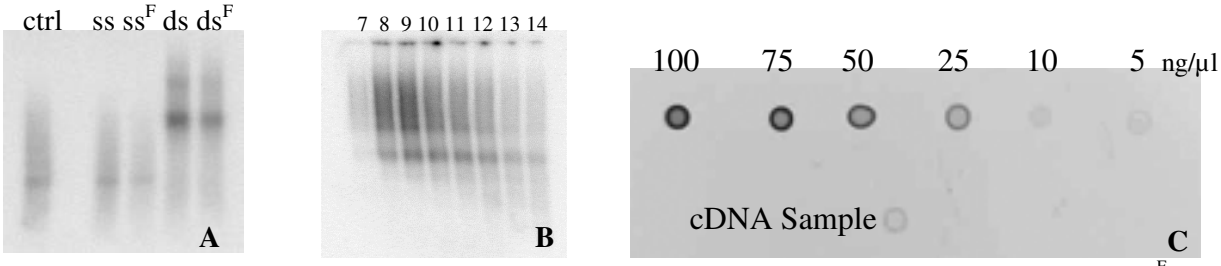

Figure 3.12. A-Single (ss) and double strand (ds) synthesized cDNA before and after the filtration step (^{F}), as control [α-^{32}P]ATP annealed mRNA is also shown (ctrl). B-Fractions of the radioactively labelled cDNA to estimate quality and length. C-Estimation of the pooled cDNA concentration by comparing with different Lamda/DNA EcoRI + HIN III concentrations.

Approximately 30,000 primary clones were generated as estimated from plaque titration. Plaques were picked and stored in PSB at -20°C and in 20% DMSO at -80°C according to section 2.8. Amplified cDNAs were visualized and selected according to length after analysis by agarose gel electrophoresis; those between 500 and 2500 nt long (fig. 3.13) were sequenced.

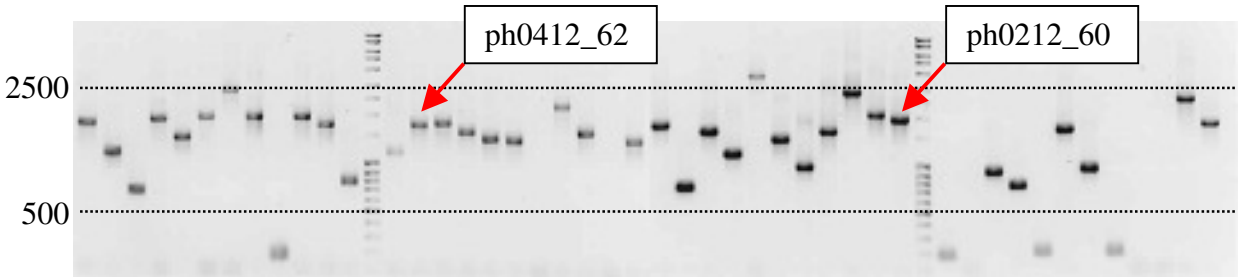

Figure 3.13. Agarose electrophoresis gel of selected amplified cDNAs showing the clones identified as ph0412_62 and ph0212_60.

3.3.2. Expressed Sequence Tags (EST) analysis

ESTs were obtained from a total of 1050 single-pass sequenced cDNAs (2.11.). The average length of the sequences read was 324 nucleotides. The analysis of the sequences in the software SeqMan from DNASTAR generated 694 different contigs or groups of cDNAs with identical sequences. The program blastx Basic Local Alignment Search Tool from the NCBI, which compares the nucleotide query sequence translated in all reading frames against the protein sequence database, allowed identifying homologies between the partially sequenced clones (ESTs) and the sequences registered in the NCBI data base. Especially interesting were the clone sequences ph0412_62, related to a putative strictosidine synthase-like gene from *A. thaliana*, and ph0212_60, related to strictosidine *β*-glucosidase from *Catharanthus roseus* and raucaffricine-*O*-*β*-D-glucosidase from *Rauvolfia serpentina*, both clones are indicated by red arrows in figure 3.13. Homology values of ph0412_62, ph0212_60 and other ESTs are show on the table 3.1.

Table 3.1. Names of potential homologues for some of the *P. ipecacuanha* EST sequences and the corresponding relation values. ESTs with presumed direct relation to alkaloid metabolism are separated by a double line.

Name of homologue protein	Expect value	Identity %	Rate of Identity	Similarity %
1,4-benzoquinone reductase-like protein [*A. thaliana*]	$2E^{-46}$	87.4	(90/103)	94.2
ABC-type transport protein [*A. thaliana*]	$5E^{-58}$	93.7	(74/79)	97.5
ADP-glucose pyrophosphorylase [*Ipomoea batatas*]	$3E^{-42}$	79.3	(88/111)	85.6
Cysteine protease [*Ipomoea batatas*]	$7E^{-55}$	78.5	(95/121)	88.4
Cytochrome P450 [*Catharanthus roseus*]	$1E^{-14}$	63,2	(36/57)	80.7
Cytochrome P450-like protein [*A. thaliana*]	$1E^{-24}$	55,2	(58/105)	72.4
Putative cytochrome P450 [*Oryza sativa*]	$4E^{-26}$	54,2	(64/118)	66.9
F-box family protein (FKF1)/adagio 3 (ADO3) [*A. thaliana*]	$1E^{-78}$	91.2	(134/147)	95.2
Flavonoid *O*-methyltransferase [*Catharanthus roseus*]	$2E^{-34}$	57,9	(73/126)	77.0
NADP dependent malic enzyme [*Phaseolus vulgaris*]	$4E^{-58}$	94.8	(109/115)	98.3
Transcription factor [*Nicotiana plumbaginifolia*]	$1E^{-37}$	94.0	(79/84)	97.6
Transcription factor JERF1 [*Lycopersicon esculentum*]	$1E^{-13}$	73.3	(33/45)	82.2
10-hydroxygeraniol oxidoreductase [*Catharanthus roseus*]	$8E^{-28}$	73,4	(58/79)	77.2
Cytochrome P450 (CYPLXXII) (*Secologanin synthase*)	$8E^{-25}$	49.6	(58/117)	67.5
Putative strictosidine synthase-like [*A. thaliana*]	$3E^{-51}$	78,2	(93/119)	86,6
Strictosidine *β*-glucosidase [*Catharanthus roseus*]	$2E^{-67}$	64,4	(112/174)	81.6
Raucaffricine-*O*-*β*-D-glucosidase [*Rauvolfia serpentina*]	$2E^{-34}$	67,0	(65/97)	80.4

An examination in the distribution of the homologies obtained for all generated sequences with BLAST from the Munich Information Center for Protein Sequences (MIPS)-*Arabidopsis thaliana* Database (MAtDB) showed that 54% of the ESTs present showed homology to a known protein, 15% had homology to proteins of unknown function and 34% were not homologous to any sequence in the database (fig. 3.14-A). From those ESTs with

homology to known functional proteins, 41% were potentially involved in metabolism and the remaining 59% were classified in other cellular functions such as development (4%), interactions (18%), cell cycle and protein synthesis (29%) and others (8%) (fig. 3.14-B). Within the ESTs related to metabolism, most of the sequences had a general function (78%), while the remainder were related to secondary metabolism, including alkaloid biosynthesis (1%) (fig. 3.14-C).

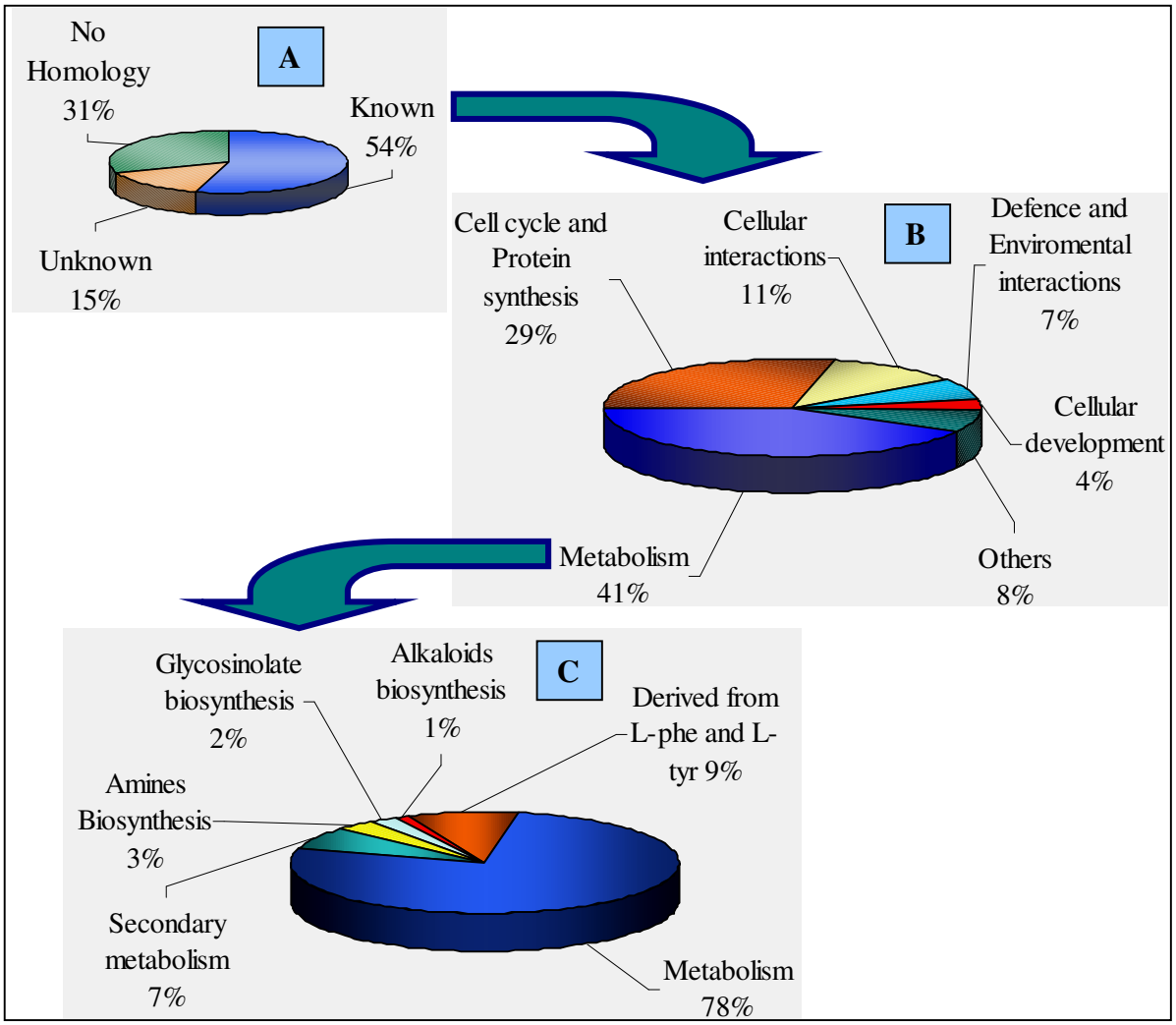

Figure 3.14 Classification of the ESTs according to the matching homology in the MatDB. A- Hits of the ESTs to registered databank sequences with known and unknown function. B- Distribution of the ESTs related to a general cellular activity. C- Overview of the distribution of metabolism-related sequences.

3.3.3. Complete gene sequence analysis

3.3.3.1. Strictosidine synthase-like gene clone (ipSTR-like)

The strictosidine synthase-like gene clone was amplified by the PCR method (2.7.1.) with the Pfu polymerase and the primers DIS_5 and DIS_3. It was subsequently ligated (2.11.2.) into the pET100/D-TOPO® expression vector and transformed into TOP10 cells from Invitrogen (2.3.2.). The complete sequenced of the insert was achieved with the primers T3,

ph0412_62_Synth, Ph0412_62_Synth_B, Ph0412_62_Syn_C, Ph0412_Synth_CII, ph0412_62_Syn_D and Str_Synth_am1020 by the primer walking method (2.11.).

The sequence analysis of strictosidine synthase-like gene clone by EditSeq software program from DNASTAR (fig. 3.15) identified a 1769 bp long clone with a 1056 bp long open reading frame (ORF) and a theoretical protein estimated at 351 amino acids with a molecular weight of 39,652 Daltons.

```
1      CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACGCGGGATAG
57    CAACGTCGAGTGTATCTCCCAAAAAAAACATCCACGAGGGGTGACTCGTGAACTGTC
114   GGACAATGACGCCGGCCGGTTTTCTYGCCGGGKTTTTTTCTTCTGTTGGCGCTTTAT
171   TGTGGCTTGGACCCTTTTAAGCACAGTGCGATTAYTGWTKWTCCTRAKTWTGTGGCT
228   CTTAGGGTGGAGATGCCTGCATGGAGTGAAATCCCAGTTGAGAAAGACACTCAGAAT
                M  P  A  W  S  E  I  P  V  E  K  D  T  Q  N
285   TTGCTTCAGAAATCGGAGATTAAGTTTTTGAATCAGGTTCAGGGTCCTGAGAGTATG
      L  L  Q  K  S  E  I  K  F  L  N  Q  V  Q  G  P  E  S  M
342   GCCTTTGATCCTCAGGGACGTGGGCCTTATACGGGGGTTGCTGATGGCAGGGTCCTG
      A  F  D  P  Q  G  R  G  P  Y  T  G  V  A  D  G  R  V  L
399   TTCTGGGATGGGGAGAAATGGAAAGATTTTGCTTATACTTCGGCTAACAGGTCAGGC
      F  W  D  G  E  K  W  K  D  F  A  Y  T  S  A  N  R  S  G
456   TTATGTGACCCTAAACCATCTCCCTTGCTGTACTTGAAGAATGAGCATATTTGTGGA
      L  C  D  P  K  P  S  P  L  L  Y  L  K  N  E  H  I  C  G
513   CGGCCTCTGGGTTTAAGATTTGACAAGAAAACAGGTGATTTATACATTGCAGATGCA
      R  P  L  G  L  R  F  D  K  K  T  G  D  L  Y  I  A  D  A
570   TACTTCGGGTTGATGAAGGTTGGACCTGAAGGTGGGTTAGCAACACCATTGGTAACT
      Y  F  G  L  M  K  V  G  P  E  G  G  L  A  T  P  L  V  T
627   GAGGCTGAAGGAGTGCCACTTCGATTCACTAATGACCTGGACTTCGATGATGAAGGA
      E  A  E  G  V  P  L  R  F  T  N  D  L  D  F  D  D  E  G
684   AACATTTATTTCACTGATAGCAGCACTAAATATCAACGAAGGAACTTTATGCAGTTA
      N  I  Y  F  T  D  S  S  T  K  Y  Q  R  R  N  F  M  Q  L
741   GTTTTCTCTGGAGATGATAGTGGACGGCTTTTGAAACATAATCCTGAAACCAAAGAA
      V  F  S  G  D  D  S  G  R  L  L  K  H  N  P  E  T  K  E
798   ACCACTGTTCTTGTTAGGGGTCTCCAATTTCCTAATGGTGTCTCCATGAGCAAGGAT
      T  T  V  L  V  R  G  L  Q  F  P  N  G  V  S  M  S  K  D
855   CGGTCATTCTTTGTCTTCTGTGAGGGTTCAATTGGCAGGTTGCGCAAGTACTGGTTG
      R  S  F  F  V  F  C  E  G  S  I  G  R  L  R  K  Y  W  L
912   AAAGGCGAAAAAGCAGGGACCTCAGAAGTAATGGCTGTCCTGCCAGGCTTTCCTGAC
      K  G  E  K  A  G  T  S  E  V  M  A  V  L  P  G  F  P  D
969   AACATCAGAACAAATGAAAAAGGTGAATTTTGGGTTGCGATTCATGCTCGCCGTACC
      N  I  R  T  N  E  K  G  E  F  W  V  A  I  H  A  R  R  T
1026  CTTTATGCTTATATGTGTGCTTTGTACCCAAGAGTTCGATTGTTCTTGTTGAAGCTT
      L  Y  A  Y  M  C  A  L  Y  P  R  V  R  L  F  L  L  K  L
1083  CCAATACCAATAAAGATCCACTACGTGATGGAAATTGGAGGCCGGCCTCATGCAGTT
      P  I  P  I  K  I  H  Y  V  M  E  I  G  G  R  P  H  A  V
1140  GTAGTGAAGTATAGCCCAGAAGGTAAGGTTTTACAAATATTGGAGGATAGACAAGGA
      V  V  K  Y  S  P  E  G  K  V  L  Q  I  L  E  D  R  Q  G
1197  AAAGTTGTTAAAGCAGTGAGTGAAGTGGAGGAGAAGGATGGGAaGCTTTGGAtGGGC
      K  V  V  K  A  V  S  E  V  E  E  K  D  G  K  L  W  M  G
1254  AGCGTTTTGATGCCTTTTGTCGCAGTTTACACATTAGAATAAAGGGGGAAAAAAACC
      S  V  L  M  P  F  V  A  V  Y  T  L  E  *
1311  ATTTCCTTACTCTTTGGTGTATGTTCCTCCTCTCCCCTTGGTGTGCCGGGGGTCCTA
1368  AATGCTCGCCGCAAGTCATAGATGAAGGGAATTCAGGCTTTTGGCTTTTTATGAGCA
1425  TAGCTGTATGGATTGTCTACTTCCGACCTTTTTACTTCATTCCCCTGCATGATGTTT
```

```
1482  CACCAACATAGCTATTTTATGGTTCGAACTGAGATTTGGTTATCCTGATGTGCTTTA
1539  ACATTCTATGAAGTATCTTCCTTTTTAGATCTTTGGAATGCGAGTTTGTTTTGAGTC
1596  ACGAGCTTATACCAAGTCAGAAGATGTCTCTGCAGTTTCAATGTTCAAGAAGATAAA
1653  ATCTACGGATGTCGAGGTGGCAGGGAAACAAATGAGTGACTCAAAGGATGGATGTGT
1710  ATAATTTGTGGGATTTTATTTACATTGTGATGAATAAGAAAAGAATTGGCAGTAAAA
1767  AAA
```

Figure 3.15. Nucleotide sequence of the strictosidine synthase-like gene clone and the conceptual translation to the amino acid sequence. Green letters are nucleotides not clearly identified.

An analysis of the open reading frame of the strictosidine synthase-like gene clone with the program blastx from the NCBI showed homology to 3 well-characterized strictosidine synthase proteins. The highest homology found was for a protein isolated from *Ophiorrhiza pumila* (Rubiaceae) with 40% identity (94 of 231 aa), then, *Rauvolfia serpentina* (Apocynaceae) with 39% identity (90 of 230 aa) and *Catharanthus roseus* (Apocynaceae) with 37% identity (86 of 232 aa). A direct comparison of the amino acid sequences are shown in the figure 3.16.

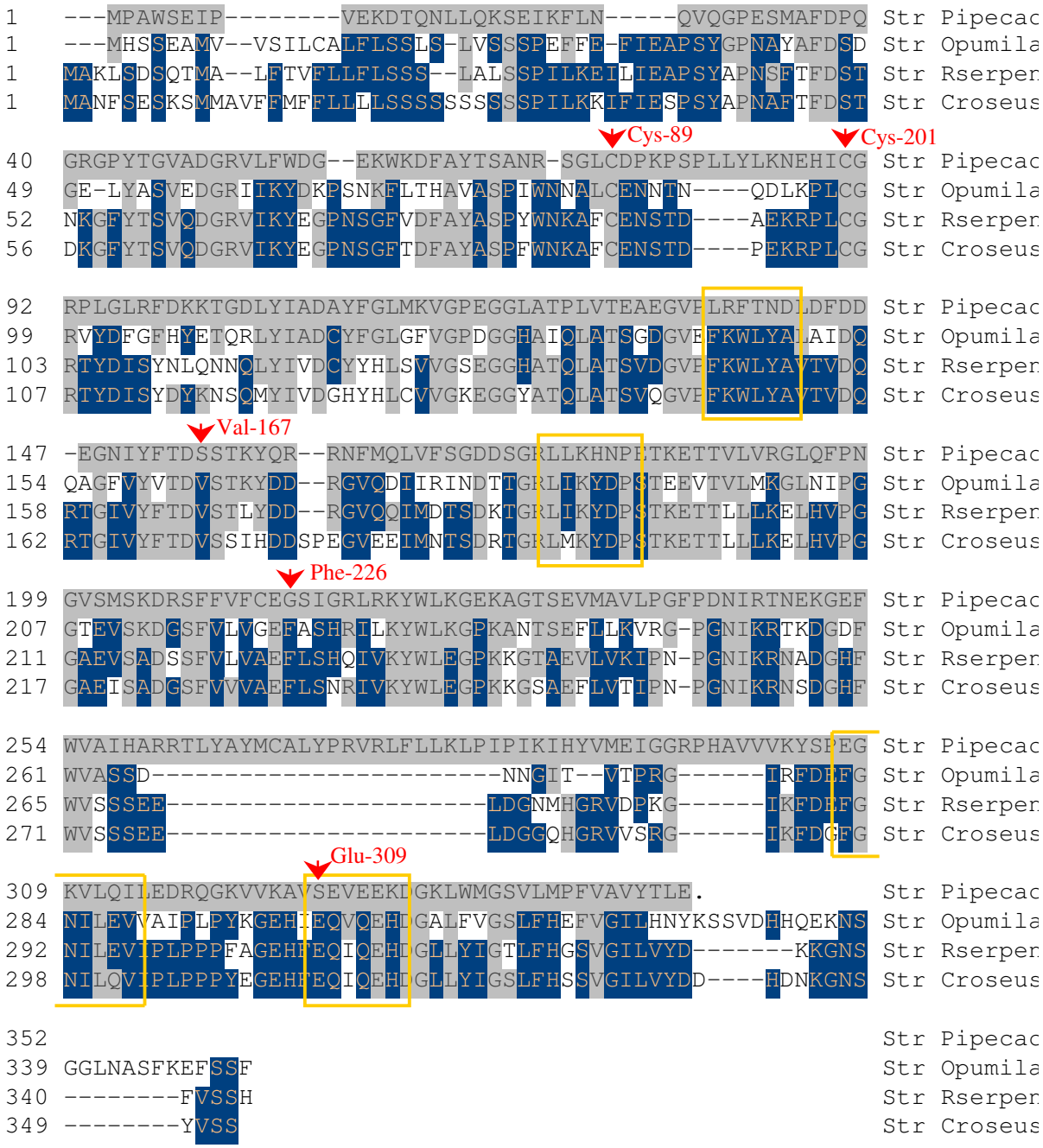

Figure 3.16. Amino acid sequence alignment of the strictosidine synthase-like of *P. ipecacuanha* (ipSTR-like) and characterized STR1 from *O. pumila*, *R. serpentina* and *C. roseus* (NCBI database). The amino acid sequence of ipSTR-like and its consensus with the STR1 sequences are highlighted in grey. The consensus sequences only among the characterized STR1 are highlighted in blue. The yellow frames were the conserved regions selected for the degenerate primers designing. Arrow heads show some residue positions with relevance in STR1.

3.3.3.2. Isolation of a strictosidine synthase-like gene by degenerate primers method

Prior to the construction of the cDNA library, isolation of a strictosidine synthase-like gene from *P. ipecacuanha* leaves poly $(A)^+$ RNA was attempted using the degenerate primer method and RT-PCR (2.7.2.). For the design of the degenerate primers, conserved amino acid sequences of strictosidine synthase from *C. roseus*, *O. pumila* and *R. serpentina* were

analyzed. Under the degenerate primer design conditions (2.9.), only 4 conserved regions were considered appropriate, indicated by yellow frames in the figure 3.16. The selected amino acid sequences, from 5' to 3', with the respective primer names were FKWLYA (STR III and STR III b) and L(I/M)KYDP (STR IV and STR IV b) for the 5'-forward primers and FGNIL(E/Q)V (STR VII-r and STR VII b-r) and EQ(I/V)QEH (STR VIII-r and STR VIII b-r) for the 3'-forward primers. The nucleotide sequences of the genes for the compared proteins were the final criteria to define the degenerated oligonucleotide sequences listed in the section 2.1.2.2. A strictosidine synthase-like gene could not be isolated by the degenerate primer method. From the 4 conserved amino acid regions chosen, the region "FKWLYA" had no consensus at all with the clone ph0412_62, the region "L(I/M)KYDP" matched only the first, third and last amino acid, the region "FGNIL(E/Q)V" mismatched 5 of the 7 amino acids and "EQ(I/V)QEH" mismatched 4 of 6 amino acids; in the last two regions, the mismatch included the first and last amino acid for each selected region (fig. 3.16).

3.3.4. Ipecoside glucosidase clone

3.3.4.1. Analysis of the first clone

The first ipecoside glucosidase clone (1451 bp) was sequenced with the primers T3, ph0212_60_Gluco, Ph0212_60_Gluco_B, ph0212_Gluco_BII, Ph0212_60_Glu_C and ph0212_60_Glu_D by the primer walking method (2.11.). An analysis with the program blastx from the NCBI suggested a missing fragment of approximately 200 bp at the 5' end compared with the closely related protein raucaffricine *O-β*-glucosidase from *R. serpentina* with 53% identity (256 of 482 aa). In addition, the amino acid sequence contained a stop codon at position 624.

3.3.4.2. Rapid Amplification of cDNA End (5'-RACE)

A full-length cDNA derived from the ipecoside glucosidase transcript was obtained in 3 steps; first, the fragment known was isolated again (2.7.2.) to corroborate the unexpected stop codon, then a 5'-RACE protocol (2.10.) was developed to identify the missing 5'-end and finally the full-length clone was generated by RT-PCR (2.7.2.). In all amplifications and re-amplifications, Pfu polymerase was used as well as an appropriate Tm for each primer pair.
A second partial cDNA of the glucosidase-like gene (Glucosidase_2) was obtained from leaf mRNA using the specific primers 5-GlucSpecif and 3-GlucSpecif and the primer dT20VN. A re-amplification with the primers Glu_5 and Glu_3 made it possible to ligate the clone into

the pET 100 TOPO vector and later on to transform the plasmid into TOP 10 cells (fig. 3.17-A). The Glucosidase_2 clone sequence, nevertheless, presented 5 amino acid mismatches with the first clone and a glutamine (Q) instead of the stop codon at position 624. These differences are indicated by yellow and red highlighted amino acids, respectively, in figure 3.18 and table 3.2. A 5'-RACE amplification was then carried out to generate the missing 5'-end. The cDNA was amplified with the primers GSPI_Glu2 and GSPII_Glu2. A 529 bp fragment was sequenced, corresponding to the length from position 327 of the primer GSPI_Glu2 in the Glucosidase_2 clone, plus the approximately 200 bp missing fragment (fig. 3.17-B). A preliminary analysis with the SeqMan program showed a corresponding partial consensus of the generated fragment with the Glucosidase_2 clone and also the sequence of both fragments combined was homologous to a complete glucosidase gene from the NCBI data bank (data not shown).

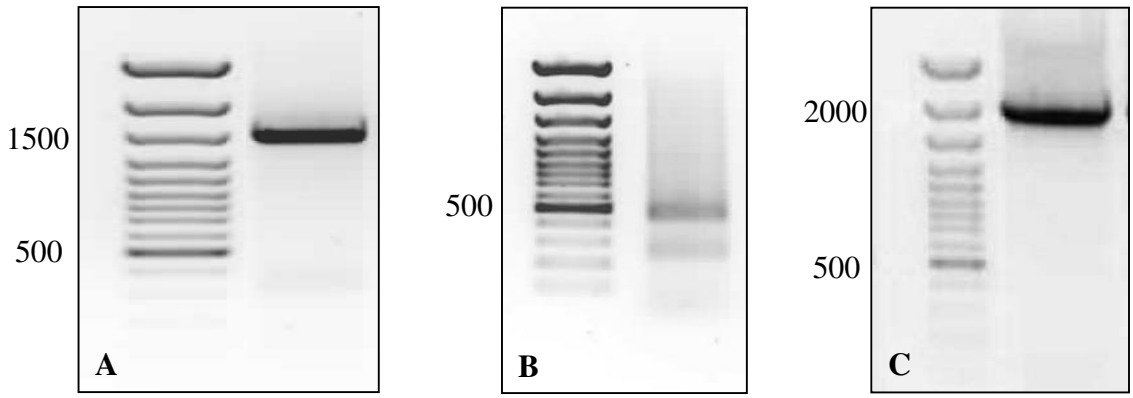

Figure 3.17. RT-PCR amplifications of the glucosidase-like clones. A- Glucosidase_2 (second clone isolated), B- fragment of approximately 500 bp generated by the 5'-RACE method and C- full-length clone of Ipe-Gluc.

3.3.4.3. Full-length clone sequence analysis

The amplification of the full-length cDNA (Ipe-Gluc) was achieved from leaf mRNA using the primers 5-GlucSpefull, designed from the 5'-end fragment, and the primer 3-GlucSpecif. A re-amplification with the primers 5-Gluc3+adap and Glu_3 allowed ligation of the cDNA into the pET 100 TOPO vector, which was subsequently transformed into TOP 10 cells (fig. 3.17-C). The Ipe-Gluc cDNA was sequenced according to section 2.11. and analyzed with the EditSeq program. The full-length clone (fig. 3.18) was 1775 nt long, the ORF consisted of 1632 nt that would form a 543 amino acid protein with a estimated molecular weight of 61807 Daltons.

```
1       CTTTAAGAAGGAGATATACATATGCGGGGTTCTCATCATCATCATCATCATGGT
55   ATGGCTAGCATGACTGGTGGACAGCAAATGGGTCGGGATCTGTACGACGATGACGAT
112  AAGGATCATCCCTTCACCACAACAACAATGTCTAGTGTTTTGCCTACCCCTGTTTTG
                                  M  S  S  V  L  P  T  P  V  L
169  CCTACCCCTGGGAGAAACATCAATCGAGGGCACTTTCCCGATGATTTTATCTTCGGA
     P  T  P  G  R  N  I  N  R  G  H  F  P  D  D  F  I  F  G
226  GCAGGAACATCGTCTTATCAGATTGAAGGTGCTGCCAGAGAAGGCGGTCGAGGGCCC
     A  G  T  S  S  Y  Q  I  E  G  A  A  R  E  G  G  R  G  P
283  AGTATATGGGATACTTTCACACATACGCATCCTGAACTCATACAGGATGGAAGCAAT
     S  I  W  D  T  F  T  H  T  H  P  E  L  I  Q  D  G  S  N
340  GGAGACACGGCCATCAATTCATATAATTTGTACAAGGAAGATATCAAGATAGTGAAG
     G  D  T  A  I  N  S  Y  N  L  Y  K  E  D  I  K  I  V  K
397  CTCATGGGACTTGATGCCTATCGTTTCTCCATATCATGGCCACGGATATTGCCTGGT
     L  M  G  L  D  A  Y  R  F  S  I  S  W  P  R  I  L  P  G
454  GGGAGTATAAACGCAGGCATAAATCAAGAAGGAATTAAGTATTACAACAATCTCATA
     G  S  I  N  A  G  I  N  Q  E  G  I  K  Y  Y  N  N  L  I
511  GATGAGCTTCTGGCCAATGATATCGTACCCTATGTAACTCTGTTCCATTGGGATGTT
     D  E  L  L  A  N  D  I  V  P  Y  V  T  L  F  H  W  D  V
568  CCCCAAGCCTTACAAGATCAATATGATGGCTTYTTAAGTGATAAGATTGTGGACGAC
     P  Q  A  L  Q  D  Q  Y  D  G  F  L  S  D  K  I  V  D  D
625  TTCCGTGACTTCGCGGAGCTCTGCTTTTGGGAATTCGGAGATCGTGTGAAAAATTGG
     F  R  D  F  A  E  L  C  F  W  E  F  G  D  R  V  K  N  W
682  ATTACAATTAATGAGCCAGAGTCCTACAGCAATTTTTTTGGTGTTGCATACGATACG
     I  T  I  N  E  P  E  S  Y  S  N  F  F  G  V  A  Y  D  T
739  CCTCCGAAGGCACACGCCCTCAAAGCTTCTCGTTTATTAGTGCCAACGACAGTGGCT
     P  P  K  A  H  A  L  K  A  S  R  L  L  V  P  T  T  V  A
796  CGACCAAGCAAACCGGTGAGAGTATTTGCTAGCACTGCAGATCCAGGAACTACTACT
     R  P  S  K  P  V  R  V  F  A  S  T  A  D  P  G  T  T  T
853  GCAGATCAGGTATATAAGGTGGGTCACAATCTGCTGCTGGCCCATGCAGCTGCTATT
     A  D  Q  V  Y  K  V  G  H  N  L  L  L  A  H  A  A  A  I
910  CAAGTGTACAGAGACAAGTTTCAGAATACTCAAGAGGGCACATTTGGAATGGCACTT
     Q  V  Y  R  D  K  F  Q  N  T  Q  E  G  T  F  G  M  A  L
967  GTGACTCAGTGGATGAAGCCTTTGAATGAAAACAATCCTGCCGATGTGGAGGCTGCA
     V  T  Q  W  M  K  P  L  N  E  N  N  P  A  D  V  E  A  A
1024 AGCAGGGCTTTTGATTTTAAGTTTGGATGGTTCATGCAGCCTCTGATAACTGGTGAA
     S  R  A  F  D  F  K  F  G  W  F  M  Q  P  L  I  T  G  E
1081 TATCCAAAATCCATGAGACAGTTGCTTGGACCACGGCTTCGTGAATTTACACCAGAC
     Y  P  K  S  M  R  Q  L  L  G  P  R  L  R  E  F  T  P  D
1138 CAAAAGAAGCTGCTCATTGGATCATATGATTATGTTGGGGTGAACTATTATACAGCT
     Q  K  K  L  L  I  G  S  Y  D  Y  V  G  V  N  Y  Y  T  A
1195 ACATATGTATCGAGTGCACAACCACCCCATGATAAAAAGAAAGCAGTTTTTCATACT
     T  Y  V  S  S  A  Q  P  P  H  D  K  K  K  A  V  F  H  T
1252 GATGGCAATTTCTATACCACCGATTCCAAAGATGGTGTACTTATTGGTCCACTGGCT
     D  G  N  F  Y  T  T  D  S  K  D  G  V  L  I  G  P  L  A
1309 GGTCCAGCATGGTTAAACATTGTTCCAGAAGGGATTTATCATGTCTTACAAGACATA
     G  P  A  W  L  N  I  V  P  E  G  I  Y  H  V  L  Q  D  I
1366 AAGGAGAACTATGAGGATCCAGTCATTTATATAACCGAAAATGGTGTTTATGAAGTG
     K  E  N  Y  E  D  P  V  I  Y  I  T  E  N  G  V  Y  E  V
1423 AACGATACAGCGAAGACACTTTCGGAAGCTCGTGTTGATACTACAAGATTACACTAC
     N  D  T  A  K  T  L  S  E  A  R  V  D  T  T  R  L  H  Y
1480 CTCCAGGACCATCTTTCCAAAGTCCTAGAAGCAAGGCATCAGGGAGTAAGAGTTCAG
     L  Q  D  H  L  S  K  V  L  E  A  R  H  Q  G  V  R  V  Q
1537 GGCTACTTAGTGTGGTCATTAATGGATAATTGGGAGCTTAGAGCGGGTTACACATCT
     G  Y  L  V  W  S  L  M  D  N  W  E  L  R  A  G  Y  T  S
1594 CGTTTTGGTCTTATTCACATAGACTACTATAATAATTTTGCGAGATACCCAAAAGAT
     R  F  G  L  I  H  I  D  Y  Y  N  N  F  A  R  Y  P  K  D
1651 TCAGCCATATGGTTCAGGAATGCTTTTCACAAGAGGCTTCGCATCCATGTGAACAAG
```

```
     S  A  I  W  F  R  N  A  F  H  K  R  L  R  I  H  V  N  K
1708 GCCAGACCTCAAGAAGATGATGGAGCATTCGACACCCCCAGGAAAAGGTTAAGAAAG
     A  R  P  Q  E  D  D  G  A  F  D  T  P  R  K  R  L  R  K
1765 TATTAAGATGG
     Y  *
```

Figure 3.18. Nucleotide and predicted amino acid sequences of Ipe-Gluc from *P. ipecacuanha*. Amino acid mismatching of Ipe-Gluc against the two first clones is highlighted grey. Amino acid mismatching and the stop codon of the first clone (ph0212_60) are highlighted yellow and red, respectively. Green amino acids correspond to the 5'-end missing fragment.

Table 3.2. Comparison of the amino acid mismatching among the 3 generated glucosidase clones. The isolated RNA used was obtained from roots (R) and from leaves (L). Colour codes are the same as in figure 3.18.

Position		62	63	64	72	188	192	266	267	268	378	407	492
ph0212_60	R	A	R	G	V	Q	D	D	A	*	N	H	V
Glucosidase_2	L	A	R	G	I	Q	D	N	T	Q	T	Q	V
Ipe-Gluc	L	I	Q	D	I	E	N	N	T	Q	T	Q	I

An analysis of the open reading frame of Ipe-Gluc with the program blastx from the NCBI showed homology to *β*-glucosidase proteins. The highest homology found was to raucaffricine-*O*-*β* -D-glucosidase from *R. serpentina* (Apocynaceae) with 54 % identity (286 of 528 aa). Additional homologies found using the NCBI database and the MegAlign program were to strictosidine glucosidases from *R. serpentina* (Apocynaceae) with 46.8 % similarity and *C. roseus* (Apocynaceae) with 46.6 % similarity. An amino acid sequence comparison of the above mentioned glucosidases and the conserved regions for the glucosidases (glycoside hydrolases) family 1 (Withers *et al*., 1990; Trimbur *et al*., 1992; Keresztessy *et al*., 1994; Henrissat and Davies, 1997) are shown in figure 3.19.

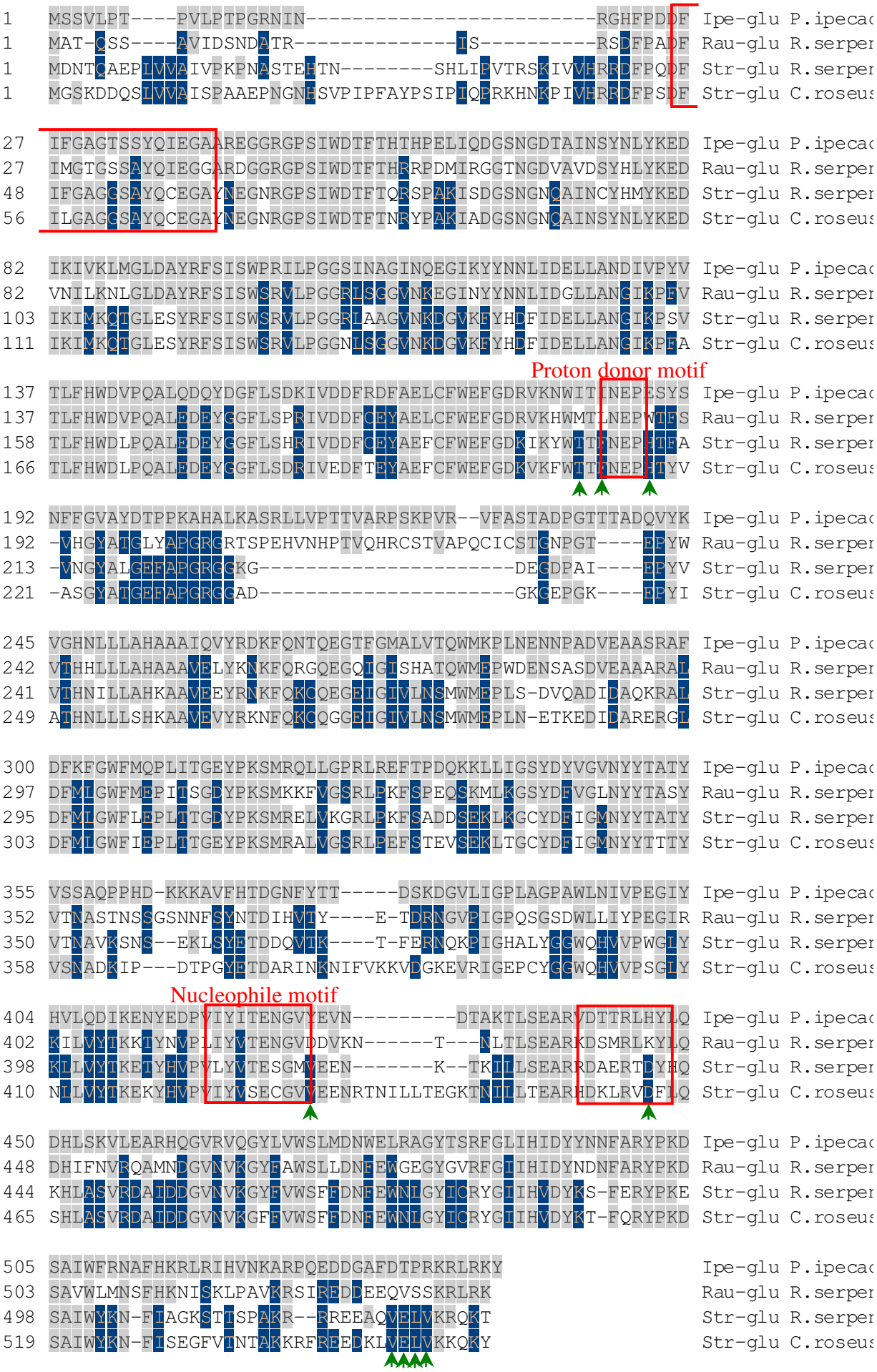
Proton donor motif
Nucleophile motif

Figure 3.19. Amino acid sequence alignment of Ipe-Gluc clone from *P. ipecacuanha*, raucaffricine and strictosidine glucosidases from *R. serpentina* and strictodisine glucosidase from *C. roseus*. The sequence of Ipe-Gluc and the consensus with the characterized glucosidases are highlighted in grey. In blue are highlighted the consensus among the characterized glucosidase sequences. In red frames are identified the conserved motifs in members of the glucosidase family 1. Green arrow heads show some residues conserved in the SGs, but differed in the RG and in the Ipe-Gluc sequence.

3.4. Enzymatic activity analysis of the strictosidine synthase-like cDNA
3.4.1. Heterologus expression and purification

A strictosidine synthase-like cDNA isolated from *P. ipecacuanha* was expected to encode, according to EditSeq program, a 351 amino acid protein with a molecular weight of 39,652 Daltons, an isoelectric point at pH 8.1 and a charge at pH 7.0 of 2.65. Neither a signal peptide indicative of subcellular localization nor trans-membrane helices were detected by the sequence analysis programs PROSITE database, the servers HMMTOP, TMHMM, SOSUI and PSORT given in section (2.1.7.). For enzyme expression, the isolated cDNA was inserted into pET100/D-TOPO vector, which permits addition of a hexahistidine extension at the N-terminus of the insert for affinity purification. The plasmid was then transformed into Rosetta™ 2(DE3) cells. Rosetta™ 2(DE3) is a recombinant *E. coli* bacterial strain that enhances the expression of eukaryotic proteins because it produces higher levels of tRNAs for six codons that are rarely found in common *E. coli* bacteria, but are common in eukaryotic cells. The expression of the clone according to the section (2.14.1.) produced a 43 kDa protein, including the hexahistidine extension and the Xpress™ epitope (which allows an eventual fusion protein by the Anti-Xpress™ Antibodies) regions present in the pET100/D-TOPO vector. Most of the recombinant protein was found in inclusion bodies, which were then dissolved in Urea Buffer (UB) after the first protein extraction with His-Tag Lysis Buffer (HLB). A second protein, around 38 kDa, also isolated was no further investigated. (fig. 3.20). Varying the bacterial culture temperature to either 25 °C or 15 °C did not result in an increase in non-denatured protein.

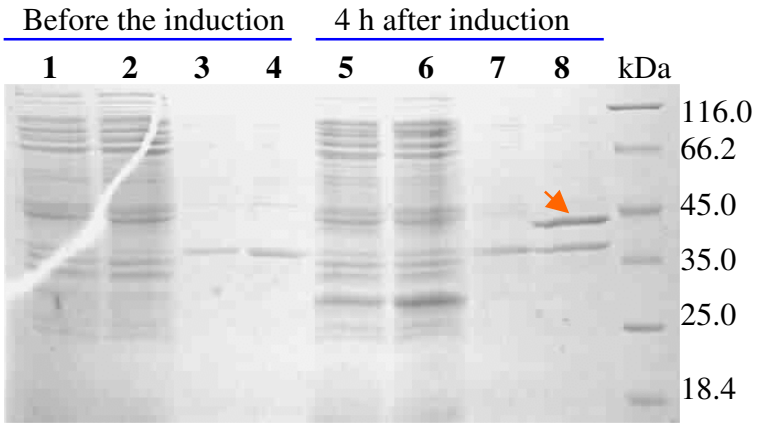

Figure 3.20. SDS-polyacrylamide gel with the recombinant strictosidine synthase-like protein before and after induction with IPTG. Before induction: first lysis extract with HLB (1) and its negative control (2), inclusion bodies dissolved in UB (3) and its negative control (4). After induction: first lysis with HLB (5) and its negative control (6), inclusion bodies dissolved in UB of a negative control (7), the recombinant protein (arrow) extracted

with UB from inclusion bodies (8) and protein molecular mass markers (kDa). The negative controls were the corresponding procedure steps of parallel protein extractions using empty vector-containing bacteria.

Figure 3.21-A shows the purified protein from inclusion bodies dissolved in UB eluted in fractions from the Talon resin column and the desalted protein after a PD10 column. Purified protein was also obtained from first lysis extracted protein with HLB. Protein extractions and purification steps are shown in the figure 3.21-B. A parallel protein extraction and purification of an empty vector containing bacteria was carried out as negative control. From 2 litres of cultured bacteria 0.35 mg purified protein from inclusion bodies dissolved in UB after the PD10 column was obtained with a concentration of 0.10 mg/ml, including the second isolated 38 kDa protein (fig. 3.21-A). The amount of isolated protein from HLB extraction was much lower and the fraction contained a not further studied second protein of 60 kDa (fig. 3.21-B).

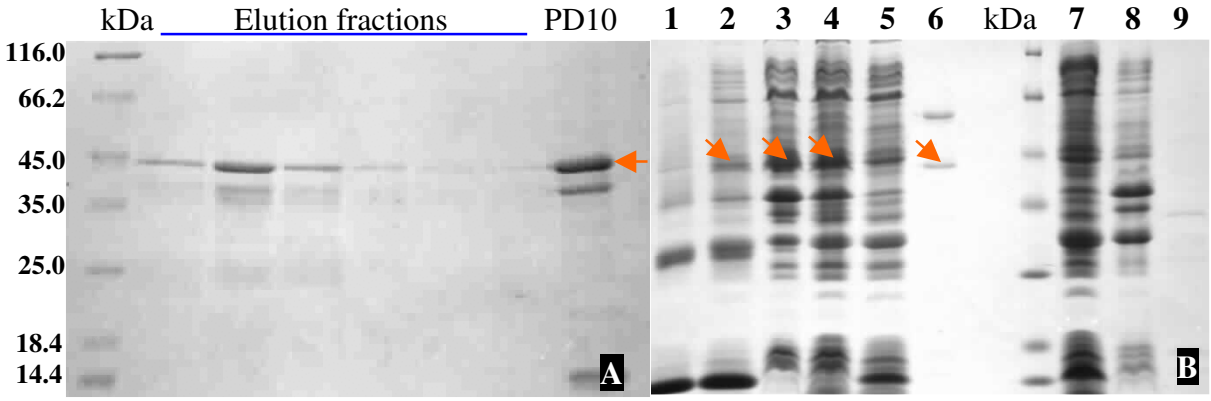

Figure 3.21. Strictosidine synthase-like protein purification. A- Protein elution fractions and the purified protein after the (PD10) column from UB extracted protein. B- Before induction: first lysis extract with HLB (1). After induction: first lysis extract with HLB (2), pellet dissolved in loading buffer (3), pellet dissolved in UB (4), metal affinity no-bound proteins (5), metal affinity purified protein after the PD10 desalting column from HLB extracted protein (6). Empty vector bacteria extractions: first lysis with HLB (7), UB dissolved inclusion bodies (8), metal affinity purification after PD10 desalting column (9). Protein marker (kDa). Arrows show the protein presence.

3.4.2. Enzymatic activity test

Crude protein extractions with HLB and inclusion bodies dissolved in UB as well as purified protein from the same above mentioned extractions were used to test the enzymatic activity of the recombinant strictosidine synthase-like protein. Negative controls were the above described protein preparations boiled 10 min at 95 °C, protein preparations from empty vector containing bacteria as well as protein storage buffer. The basic conditions considered for the enzymatic activity tests were taken from previous works on strictosidine synthase (Treimer and Zenk, 1979) and on deacetylipecoside and deacetylisoipecoside synthases (De-eknamkul *et al.*, 1997), but additional conditions were also assessed (2.15.1.). Under these conditions, it

was not possible to establish an enzymatic activity of the protein since differences were not detected when compared with negative controls. A spontaneous Pictet-Spengler condensation of dopamine and secologanin to form deacetylipecoside (and/or its isoform) $[M+H]^+$, *m/z* 524 t_R:10.5 min (ALK 46 gradient) (fig. 3.22) and deacetylipecoside acid (and/or its isoform) $[M+H]^+$, *m/z* 510, t_R: 9.0 min (ALK 46 gradient) (fig. 3.23) was observed in pHs up to 6.5 in negative controls by HPLC and confirmed by LC-MS/TOF. No differences in the product concentration were detected in the presence of any of the protein preparations compared to the corresponding negative controls.

Activity was also not detected with the potential substrates tryptamine, tryptophan, L-dopa, tyramine or tyrosine in place of dopamine (fig. 3.24).

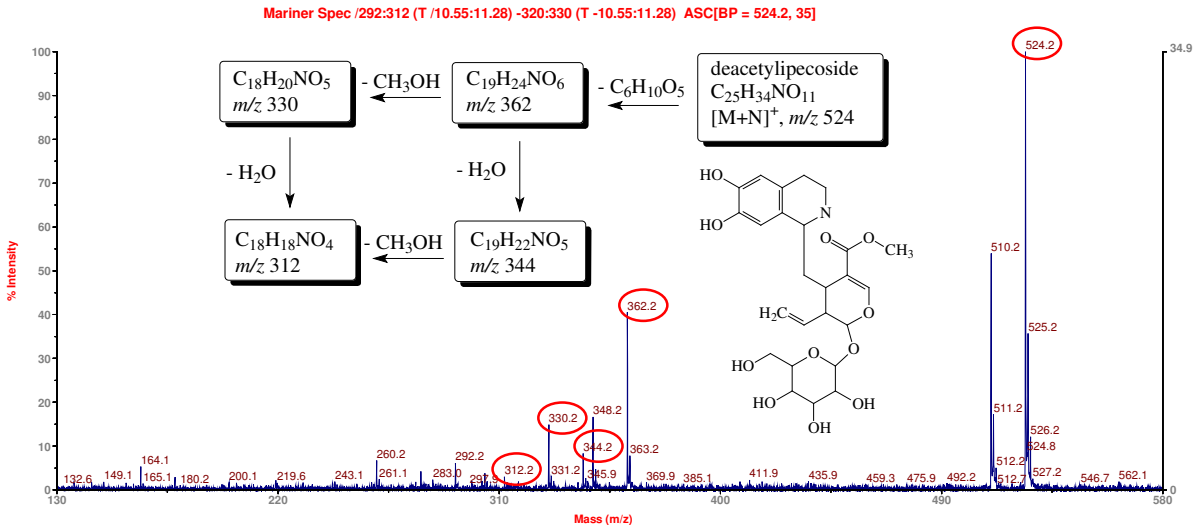

Figure 3.22. ESI-MS spectra and key ion fragmentation of deacetylipecoside (and/or its isoform) condensed from secologanin and dopamine in a control reaction at pH 5.0 obtained by LC-MS/TOF.

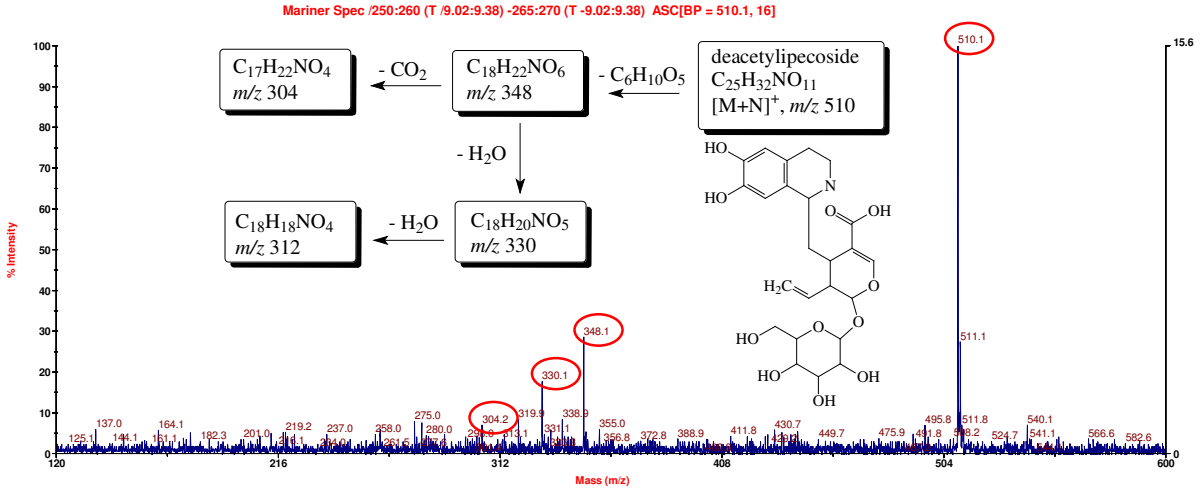

Figure 3.23. ESI-MS spectra and key ion fragmentation of deacetylipecoside acid (and/or its isoform) condensed from secologanin and dopamine in a control reaction at pH 5.0 obtained by LC-MS/TOF.

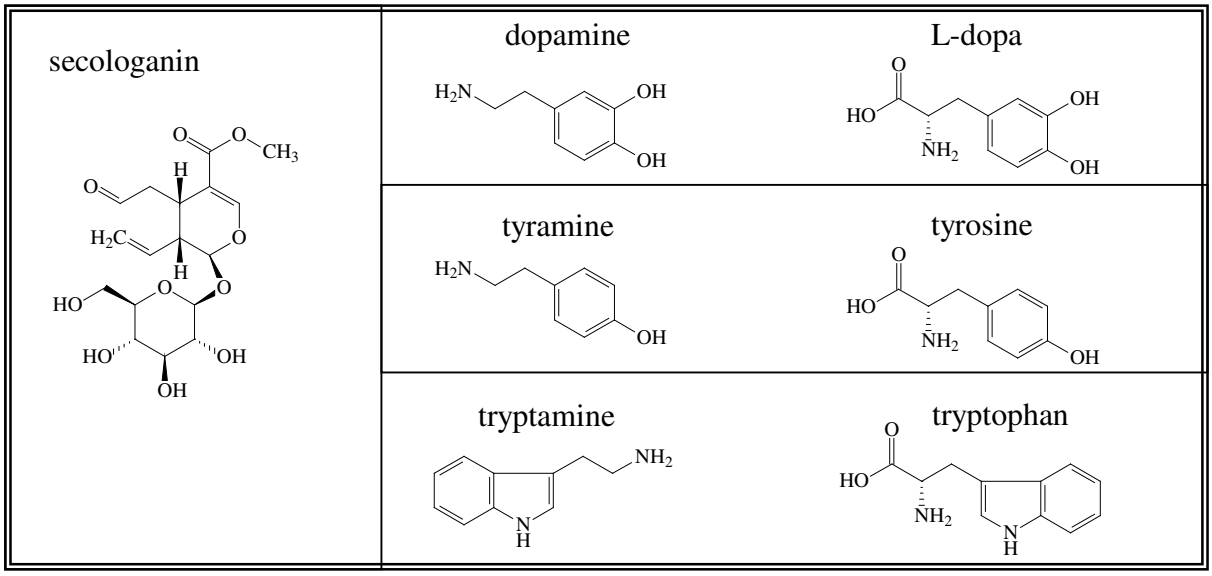

Figure 3.24. Tested substrates with the isolated recombinant strictosidine synthase-like protein. The structures were obtained from www.genome.jp.

3.5. Enzymatic characterization of the of ipecoside glucosidase gene clone (Ipe-Gluc)
3.5.1 Heterologus expression and purification

The ipecoside glucosidase cDNA was expected to encode, according to EditSeq program, a 543 amino acid protein with an estimated molecular weight of 61807 Daltons, an isoelectric point at pH 6.4 and a charge at pH 7.0 of -4.43. Neither a signal peptide for subcellular localization nor transmembrane helices could be identified with the sequence analysis programs PROSITE database, the servers HMMTOP, TMHMM, SOSUI and PSORT given in section (2.1.7.). The isolated cDNA was inserted into the pET100/D-TOPO vector and transformed into Rosetta™ 2(DE3) cells. Expression of the cDNA (2.14.1.) produced a protein of 64 kDa molecular mass, including the hexahistidine extension and the Xpress™ epitope (which allows an eventual fusion protein by the Anti-Xpress™ Antibodies) regions present in the pET100/D-TOPO vector. Similar to the strictosidine synthase-like clone, most of the Ipe-Gluc protein was found in the inclusion bodies dissolved in UB (fig. 3.25). Variations in the bacteria culture temperature did not increase the extraction of non-denatured protein into HLB.

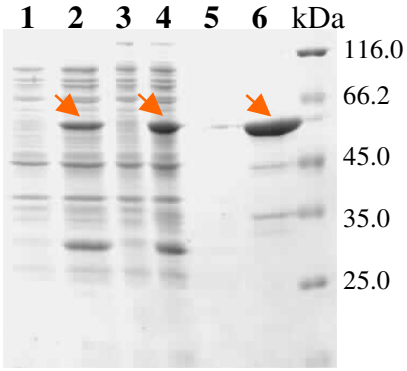

Figure 3.25. SDS-polyacrylamide gel of the recombinant ipecoside glucosidase enzyme. First lysis extraction with HLB before (1) and after (2) the induction, pellet dissolved in loading buffer before (3) and after (4) the induction, UB dissolved inclusion bodies before (5) and after (6) the induction and the protein molecular mass markers (kDa). Arrows show the recombinant protein.

From 2 litres of cultured bacteria 3.2 mg purified protein from inclusion bodies dissolved in UB after the PD10 column was obtained with a concentration of 0.92 mg/ml. A total amount of 0.53 mg desalted protein from a first lysis extraction with HLB was purified with a concentration of 0.15 mg/ml. Figure 3.26 shows selected protein extraction steps of Ipe-Gluc including a final purification from a first lysis protein extraction with HLB.

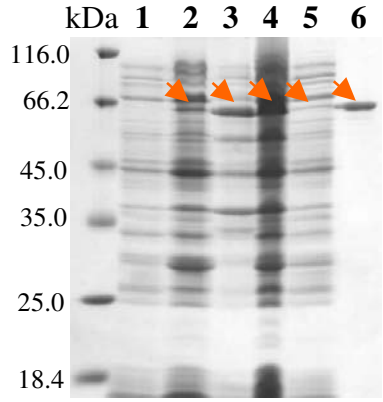

Figure 3.26. Ipecoside glucosidase protein purification. Before induction: First lysis extraction with HLB (1). After induction: First lysis extraction with HLB (2), UB dissolved inclusion bodies (3), pellet dissolved in loading buffer (4), metal affinity unbound proteins (5), metal affinity purified protein after desalting on a PD10 column (6) and protein marker (kDa). Arrows indicate the position of recombinant protein.

3.5.2. Enzymatic activity test

The enzymatic activity of recombinant ipecoside glucosidase was tested with the same kind of protein preparations, including the negative controls, mentioned in (3.4.2.). The enzymatic activity test conditions are described in (2.15.1.). The substrates ipecoside, strictosidine, raucaffricine, vincoside and arbutin were tested due to the amino acid homology found in (3.3.4.3.). Enzymatic activity was detected for the substrate ipecoside (identification details in 3.5.3.). In one minute at 30°C, pH 5.0, 0.6 µg of metal affinity purified protein was able to reduce the concentration of the substrate ipecoside to 50 % compared with the no-enzyme control. Also, around 20% of the ipecoside was not detected in presence of inactive enzyme after 30 min (fig. 3.27).

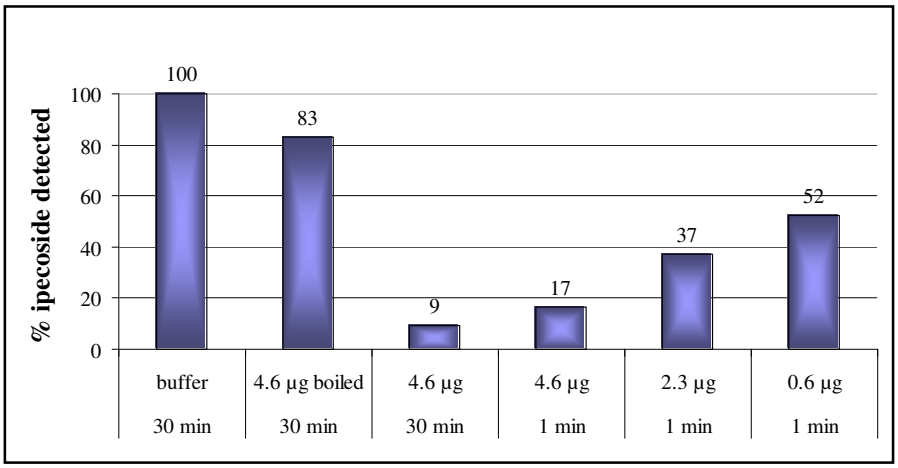

Figure 3.27. Preliminary test of ipecoside glucosidase enzyme activity at 30 °C, pH 5.0 and 100 µM ipecoside in a 100 µl reaction (n= 1) with HPLC detection. Reaction times of 30 min and 1 min were tested in the presence of various concentrations of inactive enzyme (boiled) and active enzyme as well as a no enzyme control (buffer).

3.5.2.1. Determination of pH and temperature optimum

The pH dependency was determined according to (2.15.1.) and compared with a negative control (without enzyme). Figure 3.28-A shows the highest concentration of ipecoside aglycon detected at pH 3.5, which then progressively decreased until pH 6.0. At the pH 3.5, the sum of the molar concentration of ipecoside and ipecoside aglycon was approximately 80%; at pH 5.5 the sum of the concentrations was only 50 % suggesting that 20-50% of the alkaloid added was not accounted for (Fig. 3.28.-A).

The temperature assessment showed no notable differences between 20 °C and 50 °C; nevertheless the highest concentration of ipecoside aglycone and the lowest concentration of ipecoside were detected at 40 °C (fig. 3.28-B).

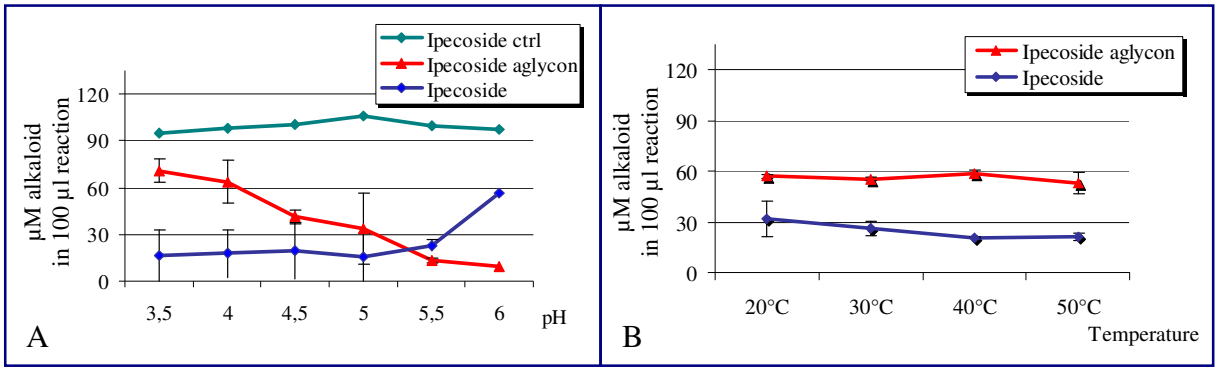

Figure 3.28. HPLC detection of: pH- Effect of pH on the enzyme activity with the substrate ipecoside after 2 min reaction, at 30 °C (n= 3). The negative control (green) corresponds to reaction preparations without enzyme after 5 min (n= 1). Temperature- Effect of temperature on enzyme activity with the substrate ipecoside after 2 min reaction, pH 3.5 and (n=2). 0.6 µg enzyme was used per 100 µl reaction.

3.5.2.2. Determination of the enzyme specificity and stability

The enzyme ipecoside glucosidase was tested under the optimum conditions for the substrate ipecoside, and with the substrates strictosidine, raucaffricine, vincoside lactam, arbutin, apigenin 7-*O*-*β*-glucoside, 7-*O*-*β*-D-glucosyl kaempferol, naringenin 7-*O*-*β*-D-glucoside, galloyl, salicin and 7-*O*-*β*-D-glucosyl quercetin, all *β*-glucosides (fig. 3.29). Enzymatic products were detected by LC-MS and LC-MS-MS analysis when strictosidine was added as substrate. Low concentrations of cathenamine (identification details in 3.5.3), the subsequent reduced product of strictosidine aglycone were detected. Cathenamine could not be detected by simple HPLC analysis, nevertheless the concentration of strictosidine was reduced in presence of the active enzyme (fig. 3.30). The concentration of the substrate strictosidine was also, in some cases, reduced in the presence of inactive enzyme (boiled 10 min at 95 °C), but the exact cause of this effect was not determined.

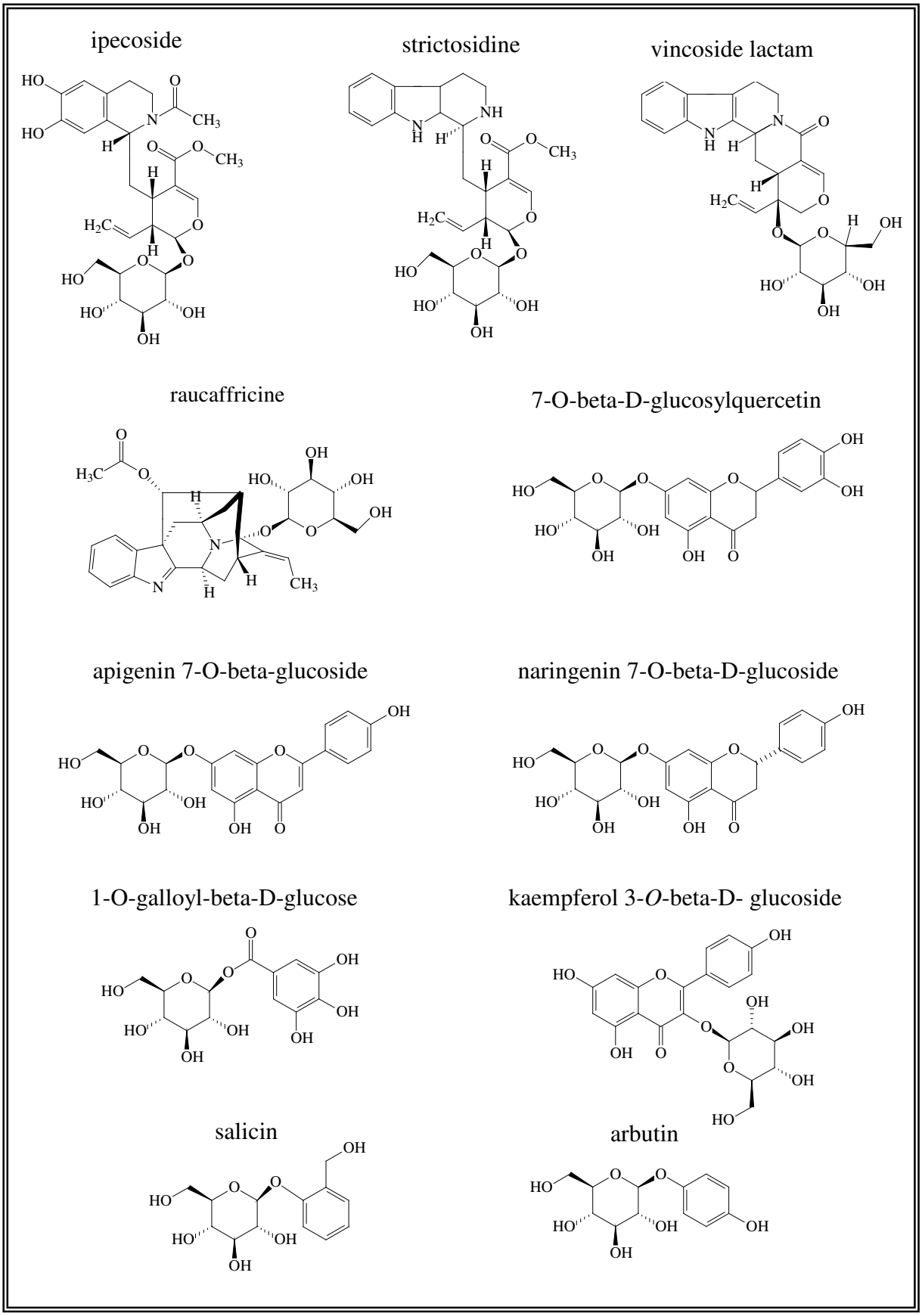

Figure 3.29. Substrates tested with recombinant ipecoside glucosidase protein. The structures were obtained from www.genome.jp.

In the substrate raucaffricine, low amounts of methyl-raucaffricine impurity were detected by LC-MS analysis. The $NaBH_3CN$, added to all assays to maintain the same conditions as with the substrate strictosidine, produced a reduction of both compounds to 1,2-dihydro-racauffricine and 1,2-dihydro-1-methyl raucaffricine respectively (identification details in 3.5.3). A small reduction in the concentration of the reduced substrate mixture was observed at pH 5.5 (fig. 3.30), but enzymatic products were not detected in any of the test reactions.
The remainder of the beta-glucoside substrates tested did not show differences compared with the corresponding controls and products were not detected. Some of the tested substrates in presence of active enzyme are showed in the figure 3.30.

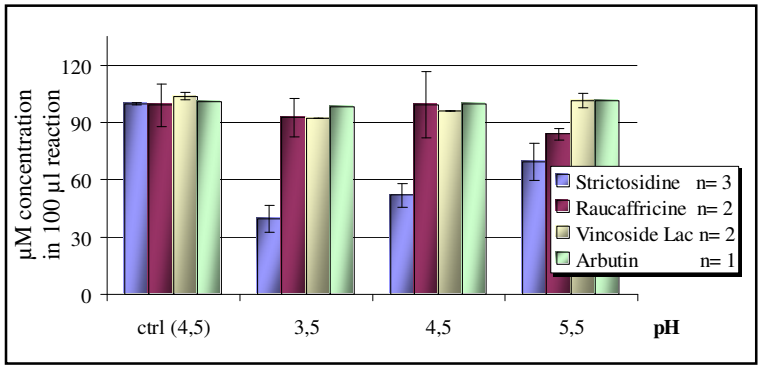

Figure 3.30. Concentration of the tested substrates with 1 µg active enzyme after 60 min reaction and 30 °C, detected by HPLC. The negative control corespond to inactived enzyme with 10 min at 95 °C.

Purified enzymes at different times and stored at different temperatures were tested to evaluate the stability of the enzyme activity (fig 3.31). The enzyme activity showed to be estable, even, more than a month after the extraction and stored at 4 °C. The test was achieved only for the metal affinity purified protein and with the substrate ipecoside.

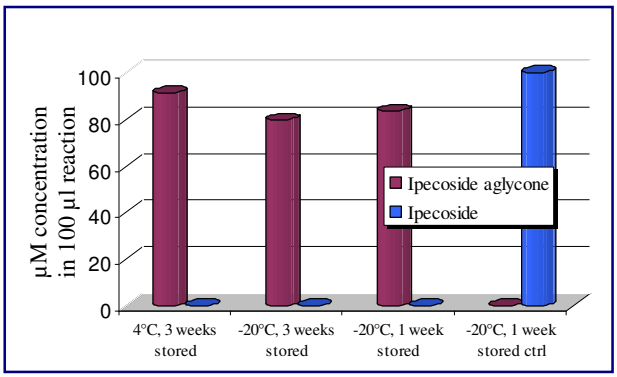

Figure 3.31. Effect of different storage conditions on the enzyme activity with the substrate ipecoside (n=1) detected by HPLC. After 30 min at pH 3.5, 30°C in 100µl reaction. Aliquots of the same purified enzyme were stored at 4°C and –20°C for a period of 1-3 weeks. The control corresponds to an empty vector bacterial protein purification.

3.5.3. Mass spectrometer analysis

The quantification of the substrates and products of reaction were detected by HPLC analysis according to (2.17.1.). The detection and identification of detection of substrates and products in low concentrations were analysed by LC-MS and MS-MS methods described in (2.17.2.) and (2.17.3.)

The analysis of the substrate raucaffricine (*m/z* 513) demonstrated the presence also of methyl raucaffricine (*m/z* 527), nevertheless the analysis of reaction preparations, with and without enzyme, detected the 1,2-dihydroracauffricine $[M+H]^+$, *m/z* 515 and 1,2-dihydro-1-methylraucaffricine $[M+H]^+$, *m/z* 529 (fig. 3.32). A reduction factor present in the reaction preparations to stop additional rections was sodium cyanoborohydride ($NaBH_3CN$).

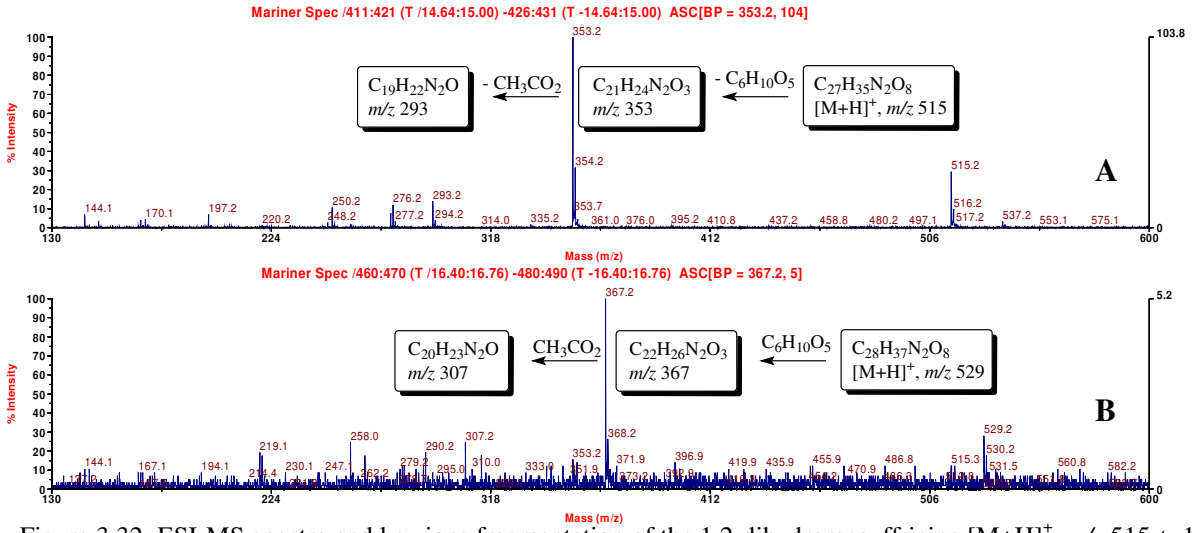

Figure 3.32. ESI-MS spectra and key ions fragmentation of the 1,2-dihydroracauffricine $[M+H]^+$, *m/z* 515 t_R 14.6 min (A) and 1,2-dihydro-1-methylraucaffricine $[M+H]^+$, *m/z* 529 t_R 16.4 min (B) obtained by LC-MS/TOF.

Under optimum conditions for the ipecoside glucosidase enzyme, but without enzyme, only the substrate ipecoside (t_R 14.2 min) was detected and identified. In presence of active enzyme, only the aglycone (t_R 16.1 min) form was found (fig. 3.33).The key ion fragment of both molecules were clear identified including the adducds $[M+Na]^+$, *m/z* 588 (fig. 3.34) for ipecoside and $[M+Na]^+$, *m/z* 426 for ipecoside aglycone (fig. 3.35).

Additionally, an *O*-methylipecoside (t_R 16.4 min) was found included with the substrate ipecoside. Similar to pecoside, the aglycone form (t_R 18.6 min) was detected only in presence of active enzyme (fig. 3.36). The key ion fragment of both molecules were clear identified including the adducts $[M+Na]^+$, *m/z* 602 (fig. 3.37) for ipecoside and $[M+Na]^+$, *m/z* 440 for ipecoside aglycone (fig. 3.38).

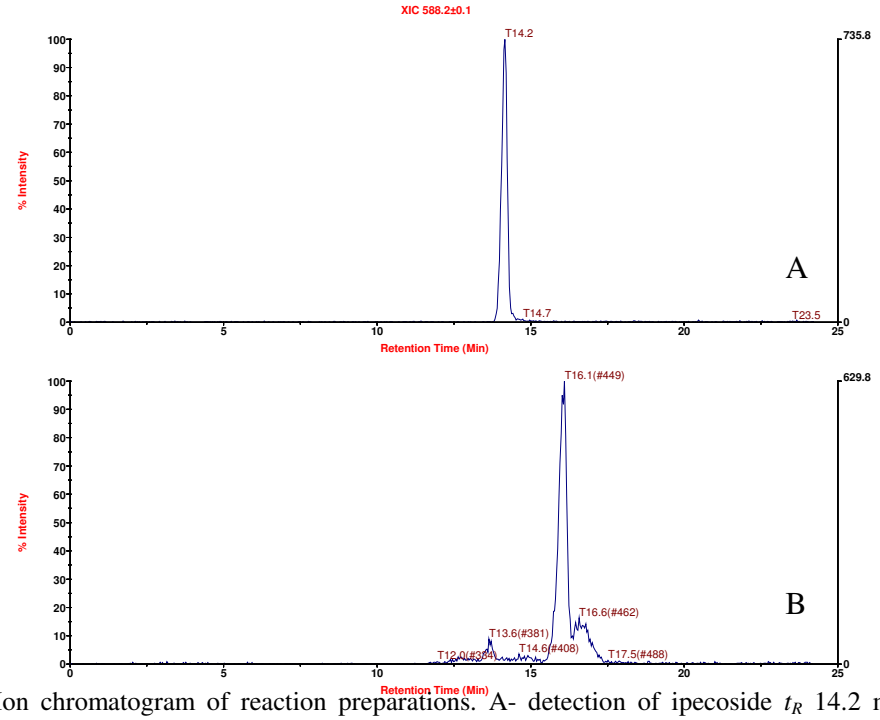

Figure 3.33. Ion chromatogram of reaction preparations. A- detection of ipecoside t_R 14.2 min in a control reaction and B- the detection only of ipecoside aglycone t_R 16.1 min in presence of active enzyme.

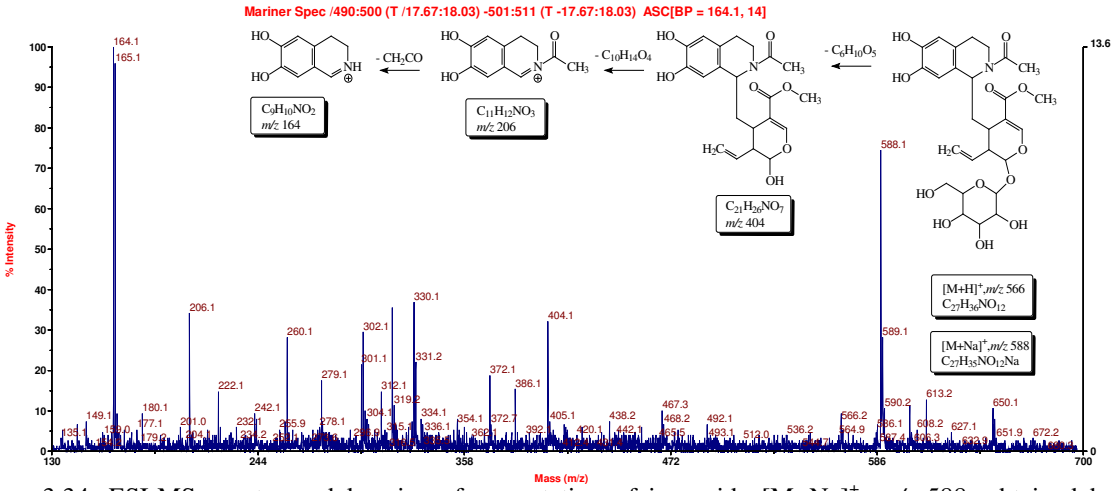

Figure 3.34. ESI-MS spectra and key ions fragmentation of ipecoside $[M+Na]^+$, *m/z* 588, obtained by LC-MS/TOF.

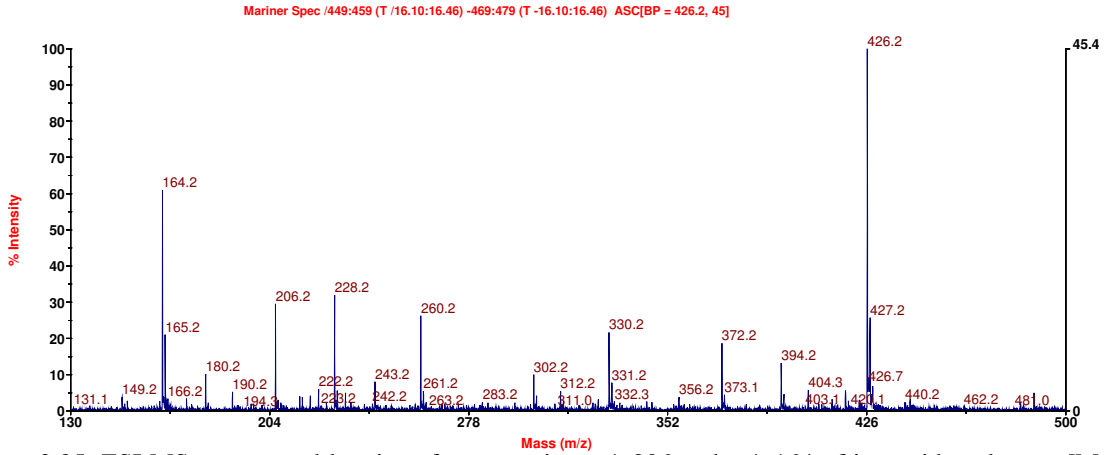

Figure 3.35. ESI-MS spectra and key ions fragmentation *m/z* 206 and *m/z* 164 of ipecoside aglycone $[M+Na]^+$, *m/z* 426, obtained by LC-MS/TOF.

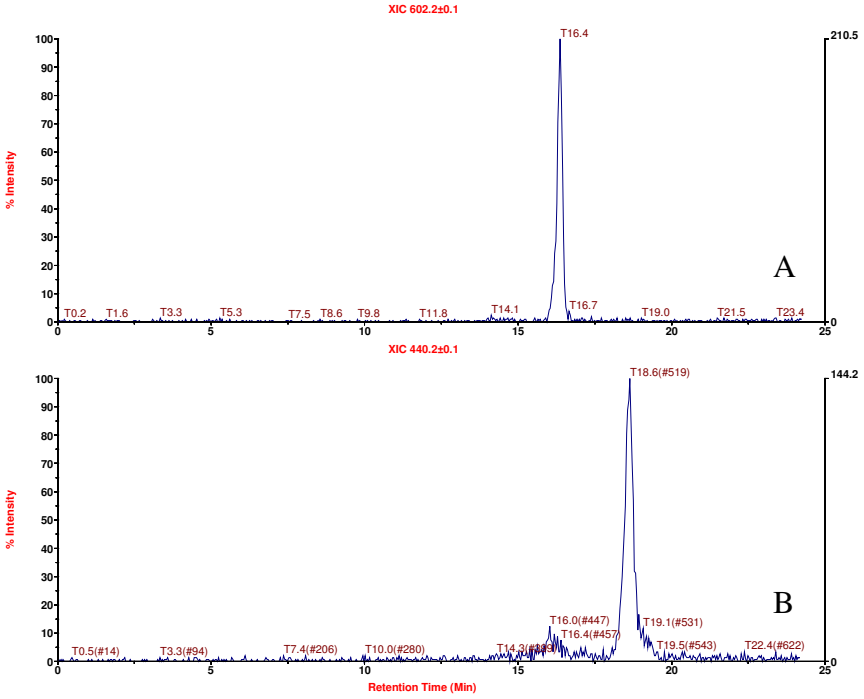

Figure 3.36. Ion chromatogram of reaction preparations. A- detection of *O*-methyl ipecoside t_R 16.4 min, in a control reaction and B- the detection only of *O*-methyl ipecoside aglycone t_R 18.6 min, in presence of active enzyme.

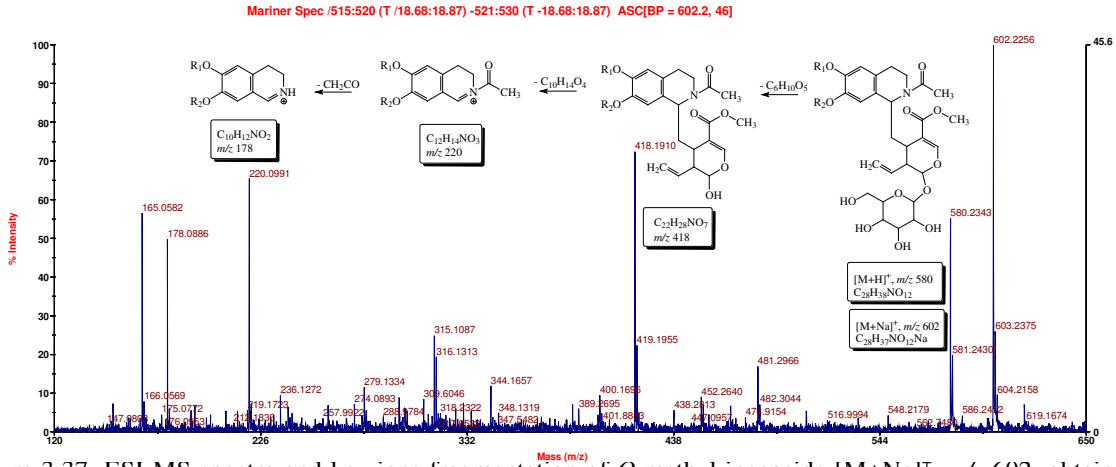

Figure 3.37. ESI-MS spectra and key ions fragmentation of *O*-methyl ipecoside [M+Na]$^{+}$, *m/z* 602, obtained by LC-MS/TOF.

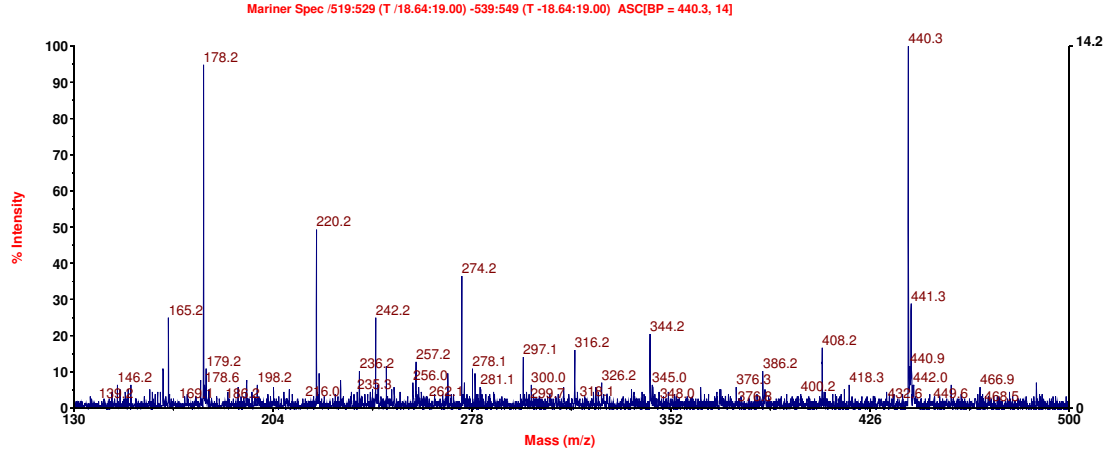

Figure 3.38. ESI-MS spectra and key ions fragmentation *m/z* 220 and *m/z* 178 of *O*-methyl ipecoside aglycone [M+Na]$^{+}$, *m/z* 440, obtained by LC-MS/TOF.

In preparations with the substrate strictosidine (t_R 16.4 min) in presence of active and inactive enzyme, it was detected cathenamine (t_R 22.0 min). The key ion fragments of strictosidine $[M+H]^+$, *m/z* 531 were clear identified (fig. 3.39) and also for cathenamine $[M+H]^+$, *m/z* 351 (fig. 3.40).

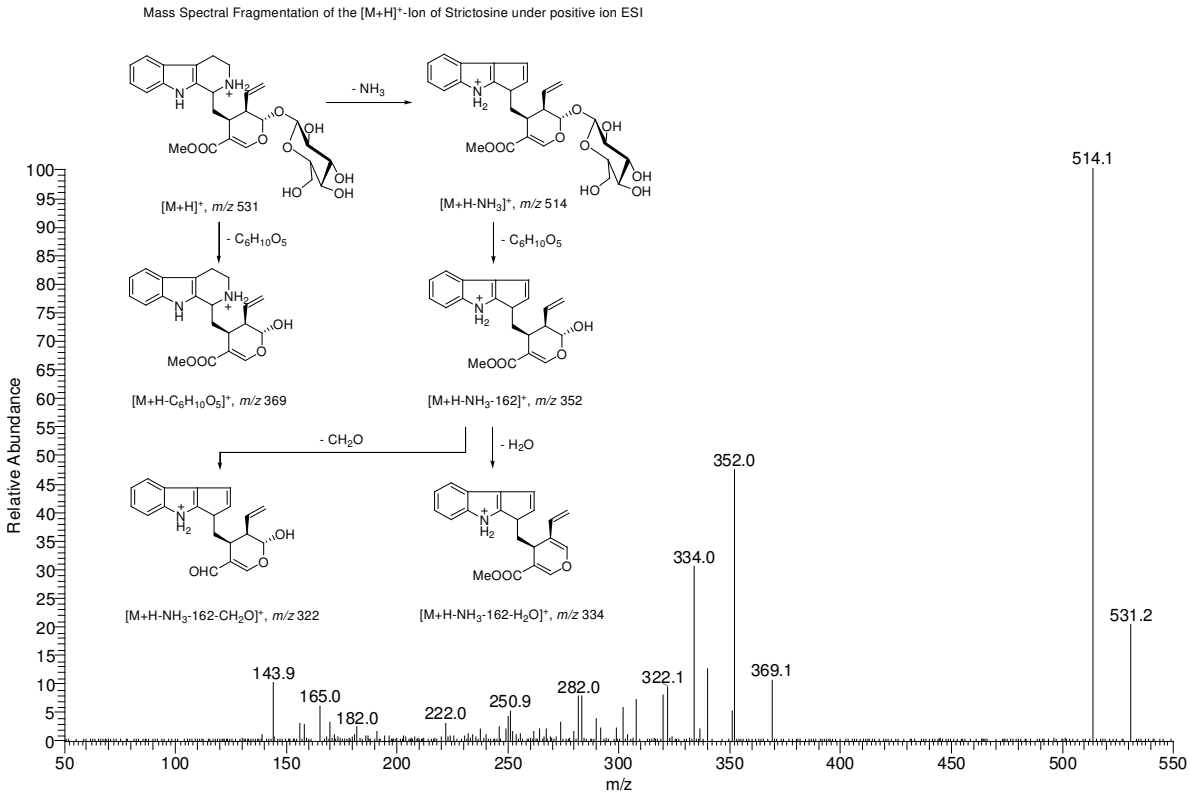

Figure 3.39. Positive ion ESI mass spectra and key ions fragmentation of strictosidine $[M+H]^+$, *m/z* 531, t_R 16.4 min.

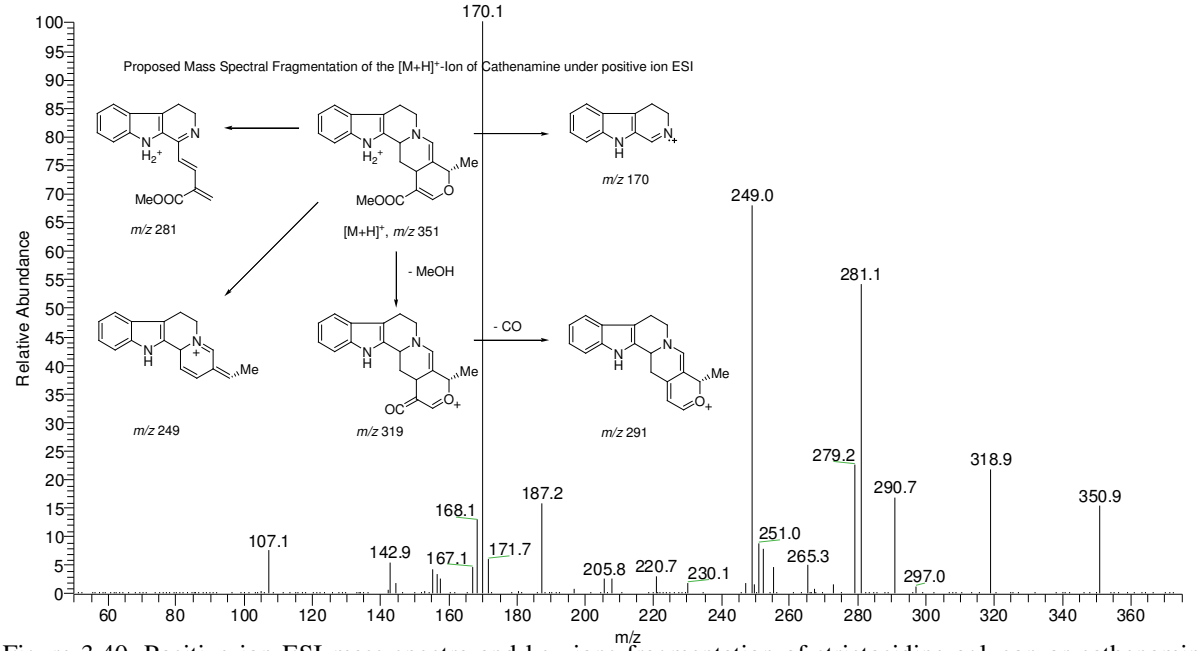

Figure 3.40. Positive ion ESI mass spectra and key ions fragmentation of strictosidine aglycon or cathenamine $[M+H]^+$, *m/z* 351, t_R 22.0 min.

4. Discussion

The increasing medicinal relevance of plant natural products had lead to a relatively comprehensive understanding of specific secondary biosynthetic pathways (De Luca and Laflamme, 2001; Kutchan, 1995). The use of modern techniques and methods in molecular biology and *in vitro* culture allows for a rapid and precise discovery of genes and enzymes involved in such processes (Kutchan, 1998). It is now realistic to attempt to discover each enzyme involve in a secondary biosynthetic pathway with the help of continuously improved biotechnological and biochemical tools. The final aim is often the deviation of the normal pathway routes to obtain tailored natural product profiles through the over expression or silencing of genes in selected points of the biosynthesis pathway (Frick *et al.*, 2006; Larkin *et al.*, 2006; Magnotta *et al.*, 2006; Frick *et al.*, 2004).

The present work was the first attempt to discover, isolate and characterize candidate cDNAs encoding enzymes involved in the biosynthetic pathway of the terpenoid isoquinoline (ipecac) alkaloids from tissue cultures of *Psychotria ipecacuanha* plant species.

From *P. ipecacuanha* plants, until now, no cDNA library or gene isolations have been reported in the literature. From *Alangium lamarckii* leaves, the enzyme deacetylipecoside synthase (DIS), which condenses dopamine and secologanin to the (*R*)-configured deacetylipecoside, was purified and partially characterized (De-Eknamkul *et al.*, 2000) and the presence of deacetylisoipecoside synthase (DIIS), which leads ultimately to the synthesis of emetine and cephaeline with the (*S*)-configuration, was also proposed (De-Eknamkul *et al.*, 1997). However, clones for these key enzymes as well as for any other specific step in the ipecac alkaloid biosynthetic pathways have not yet been isolated.

4.1. Tissue culture and alkaloid content

The tissue culture technique allowed, according to the objectives of this study, a successful maintenance and multiplication of plant material. Costa Rican *Psychotria ipecacuanha* plants were cultured under the *in vitro* conditions described in Hidalgo and Palma (1993). Without additional requirements, healthy *in vitro* plants were grown and could enable the multiplication of more than 4000 plants per year for various uses (fig. 3.1).

The generation and multiplication of root cultures was attempted considering the literature available (Yoshimatsu and Shimomura, 1991a; Yoshimatsu and Shimomura, 1993). Nevertheless, as it is well known, there are not universal culture conditions for all plant species. It could, therefore, be possible that for the Costa Rican plant material under the examined conditions in the section (3.1.2.), the root induction required the supplementation of 0.5 - 1.0 µg/l α-naphthalene acetic acid (NAA) and the root development was optimal without growth regulators in the culture media (fig. 3.4).

Ipecac alkaloids, the principal secondary metabolites present in *P. ipecacuanha* plants, were clearly identified in both leaves and roots, separately cultivated. Several intermediates were detected by their mass spectrometric fragmentation patterns (3.2.2.), but the absolute configuration was not determined. The fact that cephaeline, emetine and ipecoside were found leads to assumption of the presence of the, so far known, biosynthetic pathways for both (*R*) and (*S*)-configurations in the cultured tissues. Although ipecoside was not detected in leaves, the presence of *O*-methylipecoside (and/or its isoform) might suggest the presence also of the pathway with the (*R*)-configuration in this tissue (fig. 1.1). In *P. ipecacuanha*, nothing is known about the transport of the alkaloids, nevertheless the results indicated that the biosynthetic pathways seem to be active in leaves as well as in roots and therefore the alkaloids are found in both independently cultivated tissues.

A comparison of alkaloid contents among different cultured tissues and parental plants showed differences in amount and composition. In roots, cephaeline was 1.1-2.7 times more concentrated than emetine, whereas in leaves the cephaeline concentration was around 5 times higher than that of emetine. In roots of parental plants, similar levels of cephaeline and emetine were found (Yoshimatsu and Shimomura, 1993). The alkaloid concentrations in the tissues analyzed herein were measured as µg alkaloid per 100 mg fresh weight material. The ratio of cephaeline to emetine indicated that 20 times more cephaeline than emetine accumulated in roots, whereas only 1.3 times more cephaeline was present in leaves. Comparing alkaloid concentrations between the tissues, cephaeline was around 5 times lower in roots than leaves; in the literature almost no differences are reported. Emetine was approximately 50 times lower in roots than in leaves (fig. 3.10). On the contrary, the literature showed 3 times more emetine in roots than in leaves (Yoshimatsu and Shimomura, 1993).
It was not possible to obtain an absolute quantification of emetamine, ipecoside, *O*-methylipecoside, psychotrine and *O*-methylpsychotrine in this work due to a lack of

standard compounds (fig. 3.7). The relative concentration determined for each alkaloid did not permit to establish a significant comparison of both tissues (fig. 3.11).

The objective of the induction assay with methyl jasmonate was to increase alkaloid production (Gundlach et al., 1992; Memelink *et al.*, 2001) in order to facilitate alkaloid biosynthetic gene isolation based upon differential expression (Eiler *et al*., 1987; Roewer *et al*., 1992; Yamazaki *et al.*, 2003). A clear induction effect in either roots or leaves was not observed for any of the alkaloids. In an eventual further investigation, however, the potential increase of cephaeline and emetine concentration in leaves at 48 hours after induction might be analyzed more closely.

Although the concentration of the ipecac alkaloids could not be estimated, and induction in the alkaloid production was not detected, the identification of several ipecac alkaloids in tissue cultures of *P. ipecacuanha* indicate the presence an active biosynthetic pathway. The probability of isolating clones involved in the alkaloid biosynthesis in these tissues was considered probable.

4.2. cDNA library construction

The cDNA library construction was a critical step to obtain clones involved in ipecac alkaloid biosynthesis. In the preparation of the library, the source tissue for RNA isolation should have a maximal expression of the alkaloid biosynthetic genes compared to house keeping genes. This facilitates transcript (cDNA) identification. Different methods can be applied to increase the chance of detecting genes of interest, including the use of a specific developmental stage, selection of plant tissues in which biosynthesis is high, the use of gene induction by addition of compounds such as methyl jasmonate, as well as technical methods like the analysis of gene expression by array technologies in which transcriptome analysis and alkaloid profiling are compared (Ziegler *et al.,* 2006; Shukla *et al.,* 2006; Zulak *et al.,* 2006) or the correlative analysis of extensive metabolic profiling against the gene expression profiling (Fridman and Pichersky, 2005).

A cDNA library was generated from *P. ipecacuanha* root cultures. Advantages considered for this tissue include the avoidance of the highly expressed genes of the photosynthetic apparatus (clearly observed in the work of Shukla *et al.*, 2006), the higher alkaloid production found in roots than in unorganized tissues such as callus or cell cultures (Jha *et al.*, 1988; Teshima

et al., 1988) and the detection of at least five ipecac alkaloids in the Costa Rican *P. ipecacuanha* root cultures. Ultimately, the cDNA library was generated from mRNA isolated from cultured roots and estimated by plaque titration to contain 30,000 primary clones.

4.3. cDNA identification and analysis

The expressed sequence tag (EST) library was prepared from a randomly selected sample of the cDNA library in order to collect enough sequence information on cDNA clones. The similarity and identities with known sequences was determined by comparison using the software BLAST against the NCBI data base bank. From the *P. ipecacuanha* cDNA library 1050 single-pass sequences with an average of 324 nucleotides long were analyzed for this study. To judge the quality of the cDNA library preparation, the classification according to the matching homology (fig. 3.14) helped to evaluate the cDNA library contained representative clones. Compared to similar studies, the *P. ipecacuanha* EST library did not show obvious differences in the distribution of the EST classifications to *Papaver somniferum* (Ziegler *et al.*, 2005) or *Catharanthus roseus* (Murata *et al.*, 2006). Around 45 % of the ESTs were classified into no-homology and unknown sequences (fig. 3.14 -A) and within general cellular activity classification, 41 % of the sequences showed homology to those genes with metabolism function (fig. 3.14 -B). A classification of the homologous metabolic sequences differs from study to study, depending on the objectives of the work. Of the *P. ipecacuanha* ESTs that showed homology to those sequences assigned to metabolism in the public databases, 9 % correspond to secondary metabolism (fig. 3.14 -B).

Among the 1050 ESTs, sequences with homology to various types of proteins including ABC-transporters, cytochromes P450 and transcription factors with similarities between 72 and 98 % were identified, including a sequence with specific homology to a flavonoid *O*-methyltransferase from *C. roseus* (77 %). More directly related to alkaloid biosynthesis was a homolog to 10-hydroxygeraniol oxidoreductase from *C. roseus* (77 %) and a cytochrome P450 (secologanin synthase) (68 %) (table 3.1). Both are involved in the biosynthesis of terpenoid isoquinoline and terpenoid indole alkaloids through the synthesis of secologanin. The ESTs that were tentatively most related to terpenoid alkaloid biosynthesis were a strictosidine synthase-like and a raucaffricine β-D-glucosidase-like sequence.

Several ESTs with homology to important proteins involved with general and specific steps of several groups of alkaloid biosynthesis were in this work clearly identified. The correlation between the intensity of the expressed genes and a detailed quantification of the alkaloid intermediates under given developmental stages, plant tissues or gene inductions conditions could allow the targeting of clones that have not been previously identified. Macroarray and microarray gene expression analysis methods could make the detection of genes encoding enzymes that catalyze specific steps of ipecac alkaloid biosynthesis systematic and facile.

4.4. Strictosidine synthase-like cDNA (*ipstr*-like)
4.4.1. Sequence analysis

The partially sequenced cDNA had 78.2 % identity to a strictosidine synthase-like cDNA from *Arabidopsis thaliana*. The complete sequence of the *P. ipecacuanha str*-like cDNA clone (*ipstr*-like) was 1767 bp long with an open reading frame of 1056 bp. The identity of the clone was reduced to around 40% when compared to the characterized strictosidine synthase (STR1) from *O. pumila, R. serpentina* and *C. roseus*. The identification of a clone with similarity to STR1 could indicate the isolation of the gene involved in the terpenoid indole alkaloid biosynthesis, which is not expected for *P. ipecacuanha* plants, or a gene with a similar function from a common diverged origin, which could condense dopamine (instead of tryptamine) with secologanin, to synthesize the key intermediate deacetylipecoside (by DIS) or deacetylisoipecoside (by DIIS) along the ipecac alkaloid pathway. Both STR1 and DIS were similar with respect to temperature optimum, pH optimum, exhibiting high substrate specificity, sharing one common substrate secologanin, using the Pictet-Spengler reaction type and similar molecular size, 30 kDa for DIS and 38 kDa for STR1 (39.6 kDa for ipSTR-like). (Hemscheidt and Zenk, 1980; Kutchan, 1989; De-Eknamkul *et al.*, 2000).

The recombinant ipSTR-like protein was purified to near homogeneity (fig. 3.21), but did not show any activity with the substrate secologanin together with dopamine or tryptamine. Tryptophan, *L*-dopa, tyramine and tyrosine, also tested together with secologanin, did not serve as substrates for the protein.
Recently, the crystallization and structure elucidation of STR1 from *Rauvolfia serpentina* have been achieved. Several residues were identified in the structure that may be responsible for the binding or possibly function in the enzyme activity (Ma *et al.*, 2006). Two small helices held in proximity by a disulfide bridge between Cys-89 and Cys-101 appear to be conserved in the STR1 family and seem to be a distinct feature of the family. The covalently

bound Cys residues play an important role in the integrity of the substrate binding pocket and to the overall structure (Ma *et al.*, 2006). Based on the algorithms of the program MegAlign used for the sequence comparison, section (2.1.7), both Cys residues were also conserved in the ipSTR-like protein. In the same comparison, the residues that form the hydrophobic pocket and the two polar residues within the active site in the ipSTR-like sequence did not match to any single residue in the STR1 sequence. This includes Glu-309, proposed to be the essential residue for the catalysis by its amine deprotonation in STR1. However, with exception of the residues Val-167 and Phe-226 for STR1, the nature of the amino acids was conserved in *ipstr*-like cDNA (table 4.1). It is not aim of this work to elucidate the function of residues or the structure of ipSTR-like, but the presence of a glutamate residue next to the position, at the C-terminus, where the essential residue (Glu-309) occurs in STR1 was present in the ipSTR-like amino acid sequence (fig. 3.16).

Position	89	101	149	151	167	176	179	180	208	226	276	277	307	309	324
STR1	Cys	Cys	Trp	Tyr	Val	Val	Ile	Met	Val	Phe	Met	His	His	Glu	Phe
ipSTR-like	Cys	Cys	Phe	Asn	Ser	Phe	Leu	Val	Phe	Gly	Ile	His	Ala	Ser	Leu
STR1	di	di	hy	po	hy	hy	hy	hy	hy	hy	hy	ch	po	ca	hy
ipSTR-like	✓	✓	✓	✓	po	✓	✓	✓	✓	po	✓	✓	✓	✗	✓

Table. 4.1. Positions of the amino acid residues with possible binding and catalytic relevance in the STR1 and amino acid residues and the nature in the corresponding position of ipSTR-like are compared. The nature or function of the STR1 residues are abbreviated as follow: di, form a disulfide bridge; hy, hydrophobic; po, polar; ch, positively charged; ca, essential for catalysis.

In spite of the similarities with the STR1 sequence, the recombinant ipSTR-like protein did not show STR1, DIS or DIIS enzyme activity. This could be related to the nature of the substrates, the assay conditions or folding problems with the recombinant protein. On the other hand, even if terpenoid indole alkaloids (TIAs) are not synthesized in *P. ipecacuanha*, it should be considered that strictosidine synthase-like cDNAs have been found in plants that do not produce complex alkaloids such as *A. thaliana* (De Luca and Laflamme, 2001, Facchini *et al.*, 2004); these are cDNAs to which the *ipstr*-like cDNA presented a high homology. The production of MIAs in other *Psychotria* species should also mean the activity of a STR1 in these species (Pasquali *et al.*, 2006; Both *et al.*, 2005; Henriques *et al.*, 2004; De santos *et al.*, 2001). Therefore the *ipstr*-like cDNA could represent a very ancestral STR1 that through evolution lost its functionality, a protein of no function in *P. ipecacuanha* from still useful machinery present in closely related species or even just an enzyme with a yet undetected novel activity in *P. ipecacuanha*.

4.4.2. cDNA isolation attempt by the degenerate primer method

A potential method for the isolation of a putative gene is the use of the degenerate PCR primers the sequences of which are based upon microsequences of internal peptides of the purified enzyme as well as the amino acid sequence of previously characterized homologous enzymes. Successful results have been obtained with enzyme alkaloid pathways in the isolation of cDNAs encoding a 2-oxoglutarate dependent-dioxygenase involved in the biosynthesis of vindoline (Vazquez-Flota *et al.*, 1997), methyltransferases in the biosynthesis of berberine and morphine alkaloids (Frick and Kutchan, 1999; Ounaroon *et al.*, 2003) and strictosidine glucosidase of the terpenoid indole alkaloid pathway (Warzecha *et al.*, 2000; Geerlings *et al.*, 2000; Gerasimenko *et al.*, 2002). With the aim to isolate cDNAs encoding the deacetylisoipecoside (DIIS) and deacetylipecoside (DIS) synthases, the degenerate primer method was attempted based on the STR1 sequence considering the similarities in size, activity conditions and reaction type found for DIS and STR1 (Hemscheidt and Zenk, 1980; Kutchan, 1989; De-Eknamkul *et al.*, 2000). Before the structure elucidation of STR1 was known from Ma *et al.* (2006), the analysis of the STR1 sequences from *O. pumila, R. serpentina* and *C. roseus* established four conserved regions (fig. 3.16) from which degenerate primers were designed (section 2.1.2.2). All combinations of forward and reverse primers were tested by RT-PCR and additional nested PCRs and re-amplifications were applied also with the same primers; however, the isolation of a clone with sequence similarity to STR1 could not been achieved. Once an *ipstr*-like cDNA was obtained from the *P. ipecacuanha* EST library, it became obvious from the amino acid sequence comparison that the conserved regions in functional STR1s showed a low homology. Only the conserved region LIKYDP had a 50 % homology including the last amino acid, proline, at the carboxyl terminus, which is a critical point in the success for the reverse transcription (Abd-Elsalam, 2003). Finally, from 15 residues with relevance to the binding and catalysis of STR1 indicated by Ma *et al.* (2006) (table 4.1) only Trp-149, Tyr-151 and Glu-309 were found in the selected conserved regions. Given these homology results, a successful isolation of an *ipstr*-like cDNA with degenerate primers was, in retrospect, unlikely.

4.5. Ipecoside glucosidase (Ipe-Gluc)
4.5.1. Isolated clones

The second clone potentially related to terpenoid alkaloid biosynthesis that was isolated from the cDNA library showed 67 % identity to raucaffricine-*O*-*β*-D-glucosidase (RG) from *R. serpentina* and 64 % strictosidine *β*-D-glucosidase (SG) from *C. roseus* and *R. serpentina*. The complete sequence of the *P. ipecacuanha glucosidase* cDNA (*ipe-gluc*) was 1775 bp long with an open reading frame of 1632 bp. The identity of Ipe-Gluc was 54% to RG, and 47 % to SGs. The full-length clone was obtained by the rapid amplification cDNA 5'-End (5'-RACE) (section 3.3.4.) using the partial cDNA sequence as reference and RNA isolated from leaf as template. A total of three glucosidase clones were isolated: two partial cDNAs from cultured root RNA and leaf RNA and a full-length cDNA from leaf RNA. The comparison of the three cDNA sequences showed 12 mismatching residues (table 3.2). In grey are shown matching residues between the first (ph0212_60) and the second (Glucosidase_2) clone and in yellow and red between the second and third (Ipe-Gluc) clone. The differences among the sequences may suggest the presence of isoforms. The ipecac alkaloids can be grouped according to the *R*- or *S*-conformation, for which it has been demonstrated that isomerization does not occur (Nagakura *et al.*, 1978; De-Eknamkul *et al.*, 1997; De-Eknamkul *et al.*, 2000). It could be expected that individual glucosidases act separately on each diastereomer.

4.5.2. Sequence analysis

Multiple sequence alignments of *β*-glucosidases in the family 1 show several conserved regions and motifs with possible binding or catalytic function (Henrissat, 1991). In figure 3.19, the principal conserved motifs of *β*-glucosidases in Ipe-Gluc, RG and two SG sequences are shown. In Ipe-Gluc, the essential residues in each motif are strictly conserved.
The motifs found in Ipe-Gluc that are considered signatures for the family 1 glycoside hydrolases are FIFGAGTSSYQIEGA (position 26-40) (Henrissat and Davies, 1997) and the motif IYITENGV (position 418-425), which contains glutamate as catalytic nucleophile (Withers *et al*., 1990; Trimbur *et al*., 1992; Keresztessy *et al.*, 2001). The highly conserved motif NEP, with the glutamate identified as the acid/base catalysis residue in the cyanogenic *β*-glucosidase (linamarase) (Keresztessy *et al.*, 1994), was found in position 185-187 in the Ipe-Gluc sequence. The carboxyl terminus of Ipe-Gluc contained the motif DxxRxxY at positions 441-447; aspartate is suggested as necessary for glucoside cleavage (Trimbur *et al*., 1992).

The comparison of the Ipe-Gluc with closely related glucosidase sequences identified fifty positions at which the residues were conserved in the SGs, but differed in the RG and in the Ipe-Gluc sequence (fig. 3.19). Three of these positions were next to the NEP motif, two flanked the motif and one was two-residues upstream from the motif. Keresztessy *et al.* (2001) found that the mutation Ala-210Val two residues from the C-terminus of the NEP motif in linamarase (*M. esculenta*) changed the hydrophobic environment of the glutamate in the motif and, therefore, the enzyme activity. A residue that is only conserved in SGs, but not in RG or in Ipec-Gluc, was identified adjacent to the nucleophile motif IYITENGV in the C-terminal direction and at the sixth residue in the DxxRxxY motif. In figure 3.19, some of the mentioned positions are identified with green arrowheads. An overview of the complete sequences shows an evidently higher concentration of this pattern in the C-terminal region, including four consecutive residues at the 3'-end of the sequences. This finding could be the result of higher conserved sequence between SGs genes than RG and Ipe-Gluc, but also the positions near to identified active motifs may take part in the substrate recognition or in the catalytic activity of the enzyme.

4.5.3. Recombinant enzyme

The heterologously expressed Ipe-Gluc enzyme was found partly soluble and partly insoluble in the *E. coli* protein extraction buffer. The formation of inclusion bodies with incorrectly folded protein might be the reason for the insolubility, since no membrane transport signal peptides could be detected. Incubation of soluble recombinant enzyme with the substrates ipecoside, strictosidine, raucaffricine, vincoside lactam, glucosylquercetin, apigenin, naringerin, galloyl, kaempferol, salicin and arbutin (fig. 3.29 and 3.30) with subsequent spectrophotometric detection clearly showed enzymatic deglycosylation of the substrates ipecoside and its methylated form to the respective aglycones (fig. 3.33-3.38). The methylated form of ipecoside was detected as an impurity in both substrate and aglycone product. The absolute configuration was not determined. The concentration of the substrate strictosidine was reduced in assays in the presence of active (fig. 3.30) and inactive enzyme (data not shown), however, only trace amounts of cathenamine (fig. 3.40), the reduced aglycone product of strictosidine, were detected in the assay containing active enzyme. Strictosidine glucosidase (SG) activity reported in Gerasimenko *et al.* (2002) showed activity trends similar to Ipe-Gluc. SG showed activity with strictosidine and no activity with vincoside lactam, raucaffricine and other related compounds. Nevertheless, 0.8 % relative activity was detected

with ipecoside. Raucaffricine glucosidase (RG) reported in Warzecha *et al*. (2000) showed maximal activity with raucaffricine and a reduced relative activity with strictosidine (6.5 %), while ipecoside was not accepted as substrate. Taking together the results from Gerasimenko *et al.* (2002) and Warzecha *et al*. (2000), the higher structural similarity between the substrates ipecoside and strictosidine compared to raucaffricine (fig. 3.29) and the amino acid sequence similarities between SG, RG and Ipe-Gluc, most evident in the motif regions (fig. 3.19), it is feasible that Ipe-Gluc shows reduced activity with the substrate strictosidine and no activity with raucaffricine. In the biosynthesis of ipecac alkaloids, Ipe-Gluc catalyzes a reaction that until now has not been detected. Nagakura *et al.* (1978) suggested that the *N*-acylated alkaloidal glycosides, ipecoside and alangiside, are metabolic dead-end products from intermediates of the wrong stereochemistry. Later on, however, benzopyridoquinolizine bases, biogenetically derived from alangiside, were isolated from *Alangium lamarckii* (Fujii and Ohba, 1983). Ipecoside aglycone was not detected in alkaloid extracts of leaves and roots in this work and no reports on related compounds were found in previous studies. The hydrolysis of ipecoside would produce a unique stable aglycone which could be an intermediate for up until now unknown alkaloids. On the other hand, in the biosynthetic pathway of ajmaline, a terpenoid indole alkaloid of *R. serpentina*, raucaffricine glucosidase hydrolyzes raucaffricine to vomilenine (Warzecha *et al.*, 1999; Warzecha *et al.*, 2000). Vomilenine is in turn converted to raucaffricine by action of vomilenine UDP-glucose transferase (Ruyter and Stockigt, 1991) (fig 1.3). A putative glucosyltransferase (*A. thaliana*) found in the *P. ipecacuanha* EST library might suggest the presence of a similar glucose transferase. Since in alkaloid extractions, ipecoside has been isolated, but neither ipecoside aglycone nor any biogenetically derived intermediate have not been reported, a similar cycle to raucaffricine-vomilenine may also be occurring. In Stöckigt (1995), raucaffricine was suggested to take the role of storage compound, but its biosynthetic function is not yet known. Ipecoside glucosidase could be then similar to raucaffricine glucosidase, and this deglycosidation is a step not completely understood in the metabolic pathway of ipecac alkaloids.

4.5.4. Preliminary enzyme characterization

Once ipecoside was established as a substrate for Ipe-Gluc, the pH and temperature optimum for enzyme activity were determined. The pH optima for most β-glucosidases are in the range of pH 4.0 - 6.2 (Esen, 1993). The pH optimum for Ipe-Gluc was determined to be between

3.5 - 4.0. Figure 3.28-A shows, under the conditions of this work, a minimum ipecoside concentration between pH 3.5 and 5.5; the ipecoside aglycone concentration was increasing from pH 5.5 until its maximum at pH 3.5. Warzecha *et al.* (2000) assayed RG at pH 5.0 and for assays with SG from *C. roseus* at pH 6.3. An SG was isolated from cell cultures of *Tabernamontana divaricata* that showed a relatively constant activity in the pH interval 4.5 - 8.0 with two optima at 4.5 - 5.0 and 7.0 (Luijendijk *et al.*, 1996). Gerasimenko *et al.* (2002) showed that an SG from *R. serpentina* had a pH optima between 5.0 - 5.2 and 50 % maximal activity at pH 4.2 that slowly decreased until pH 8.0. For Ipe-Gluc no specific temperature between 20 and 50 °C could be determined as optimal with the substrate ipecoside at pH 3.5 (fig. 3.28-B). The temperature used for assays with RG and SG was around 30 °C (Warzecha *et al.*, 2000; Geerlings *et al.*, 2000). However, Gerasimenko *et al.* (2002) determined 50 °C to be the optimum temperature for SG from *R. serpentina*. Finally, test assays with enzyme extracts stored under various conditions showed a stable Ipe-Gluc enzyme; the enzyme maintained around 90 % activity after three weeks (fig. 3.31).

Thorough functional characterization and a set of structural data of ipecoside glucosidase would clarify the catalytic mechanism when this enzyme is compared with similar glucosidases.

5. Summary

Psychotria ipecacuanha (Brot.) Stokes (Rubiaceae) is a plant species from the tropical hot, humid forests of Central and South America. The best-known active principles in the plant are the terpenoid isoquinoline (ipecac) alkaloids emetine and cephaeline, localized principally in the root and rhizome. These ipecac alkaloids are used mainly in medicine as emetics, expectorants, and amoebicides. The protein-inhibiting effects of emetine and other ipecac alkaloids may account for the ability of the plant to inhibit growth of or kill several types of microorganisms. *P. ipecacuanha* was used for a biosynthetic study of ipecac alkaloids. By condensation of dopamine and secologanin in a Pictet-Spengler manner, the intermediates deacetylipecoside and deacetylisoipecoside are formed. The resulting deacetylipecoside diastereomer is then inactivated by acetylation or by lactam formation, depending of the species plant. Deacetylisoipecoside, on the other hand, is the intermediate for the major ipecac alkaloids including cephaeline and emetine in *P. ipecacuanha* plants in yet to be defined steps.

The present work was the first attempt to discover, isolate, and characterize candidate cDNAs involved in the biosynthetic pathways of the terpenoid isoquinoline (ipecac) alkaloids from tissue cultures of *P. ipecacuanha* plant species.

1- The tissue culture technique allowed, according to the objectives of this study, a successful maintenance and multiplication of plant material. Healthy *in vitro* plants were grown and could enable the multiplication of more than 4000 plants per year for various uses. For the Costa Rican plant material, the root induction required the supplementation of 0.5 - 1.0 µg/l α-naphthalene acetic acid and the root development was optimal without growth regulators in the culture media.

2- Several intermediates, of both (*S*)- and (*R*)-configuration such as cephaeline, emetine, ipecoside and psychotrine, were identified by their mass spectrometric fragmentation patterns, but the absolute configuration was not determined. The results indicated that the biosynthetic pathways in *P. ipecacuanha* seem to be active in leaves as well as in roots and therefore, the alkaloids were found in both independently cultivated tissues.

3- A cDNA library was generated from mRNA isolated from cultured roots of *P. ipecacuanha* and estimated by plaque titration to contain 30,000 primary clones. An EST library of 1050 single-pass sequences with an average length of 324 nucleotides was generated and used in this study. The EST sequences matching to secondary metabolism were 4.8 % of the total cDNAs sequenced. Several ESTs with homology directly related to alkaloid biosynthesis, such as 10-hydroxygeraniol oxidoreductase and cytochrome P450 (secologanin synthase) were found. The ESTs that were tentatively most related to terpenoid alkaloid biosynthesis were identified as a strictosidine synthase-like and a raucaffricine β-D-glucosidase-like sequence.

4- The complete sequence of the *P. ipecacuanha* strictosidine synthase-like cDNA clone showed an identity of approximately 40% when compared to the characterized strictosidine synthase (STR1) from *O. pumila, R. serpentina* and *C. roseus*. The residues with possible relevance to the active site in the known structure of STR1 did not match to the strictosidine synthase-like sequence. Therefore, the strictosidine synthase-like cDNA might represent a very ancestral STR1 that through evolution changed functionality, might be a protein that lost function in *P. ipecacuanha* during the evolution from otherwise still useful machinery present in closely related species, or even just an enzyme with as-of-yet undetected activity in *P. ipecacuanha*.

5- With the aim of isolating cDNAs encoding deacetylisoipecoside (DIIS) and / or deacetylipecoside (DIS) synthase, the degenerate primer method was applied based on the STR1 sequence considering the functional similarities found for DIS and STR1. The amino acid sequence comparison of the strictosidine synthase-like cDNA with the conserved regions in functional STR1s showed a low homology. Of the 15 amino acids residues with relevance in the binding and catalysis of STR1, only three were found in the selected conserved regions. Given these homology results, a successful isolation of the strictosidine synthase-like cDNA with degenerate primers was, in retrospect, unlikely.

6- The complete sequence of the ipecoside glucosidase (Ipe-Gluc) cDNA, the second clone tentatively related to terpenoid alkaloid biosynthesis, showed an identity of 54 % to raucaffricine glucosidase (RG), and 47 % to strictosidine glucosidase (SG). The motifs that are considered signatures for the family 1 glucosyl hydrolases and contain glutamate as

catalytic nucleophile were found in ipecoside glucosidase. The highly conserved motifs NEP and DxxRxxY with described acid/base catalysis and cleavage functions, respectively, were identified.

7- The comparison of the ipecoside glucosidase (Ipe-Gluc) with SG and RG sequences identified fifty positions at which the residues were conserved in the SGs, but differed in the RG and in the Ipe-Gluc sequence. Five of these positions were identified close to the established glucosidase conserved motifs NEP, DxxRxxY, and YITENGV. This pattern could be the result of higher conserved sequence between SGs genes than RG and Ipe-Gluc, but also these amino acids close to identified active motifs may take part in the substrate recognition or in the catalytic activity of the enzymes.

8- Incubation of soluble recombinant ipecoside glucosidase enzyme with the different *β*-glucoside substrates clearly showed enzymatic deglycosylation of the substrates ipecoside and of its methylated form to the respective aglycones. From the results of previous studies with SG and RG, the higher structural similarity between the substrates ipecoside and strictosidine compared to raucaffricine and from the amino acid sequence similarities between SG, RG and ipecoside glucosidase, it is feasible that ipecoside glucosidase showed reduced activity with the substrate strictosidine and no activity with raucaffricine. Thorough functional characterization and a set of structural data of ipecoside glucosidase would clarify the catalytic mechanism of this enzyme compared with similar glucosidases.

6. References

Abd-Elsalam, K.A. 2003. Bioinformatic tools and guideline for PCR primer design. *African Journal of Biotechnology* **2**(5): 91 – 95.

Adams M.D., Kelley J.M., Gocayne J.D., Dubnick M., Polymeropoulos M.H., Xiao H., Merril C.R., Wu A., Olde B., Moreno R.F., Kerlavage A.R., McCombie W.R., Craig Venter J. 1991. Complementary DNA sequencing: Expressed sequence tags and Human Genome Project. *Science* **252**: 1651 – 1656.

Andersson L. 2002. Re-establishment of *Carapichea* (Rubiaceae, Psychotrieae). *Kew. Bull.* **57**: 363 - 374

Ausubel, F.M., Brent R., Kingston R.E., Moore D.D., Seidman J.G., Smith J.A., Struhl K. 2004. *Current Protocols in Molecular Biology*. Massachusetts, USA. Units: 4.9.1 - 4.9.19.

Baldwin I.T. 1999. Inducible nicotine production in native *Nicotiana* as an example of adaptative phenotypic plasticity. *J. Chem. Ecol*. **25**: 3 – 30.

Bird D.A., Franceschi V.R., Facchini P.J. 2003. A tale of three cell types: Alkaloid biosynthesis is localized to sieve elements in opium poppy. *The Plant Cell*. **15**(11): 2626 – 2635.

Barleben L., Ma X., Koepke J., Peng G., Michel H., Stöckigt J. 2005. Expression, purification, crystallization and preliminary X-ray analysis of strictosidine glucosidase, an enzyme initiating biosynthetic pathways to a unique diversity of indole alkaloid skeletons. *Biochimica Et Biophysica Acta* **1747**: 89 - 92

Barrett T., Suresh C.G., Tolley S.P., Dodson E.J., Hughes M.A. 1995. The crystal structure of a cyanogenic β-glucosidase from white clover, a family 1 glycosyl hydrolase. *Structure* **3**: 951 - 960.

Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, Benfey PN. 2003. A Gene Expression Map of the *Arabidopsis* Root. *Science* **302**: 1956 - 1960.

de Boer H.J., Thulin M. 2005. Lectotypification of *Callicoca ipecacuanha* Brot. and neotypification of *Cephaelis acuminata* H.Karst., with reference to the drug ipecac. *Taxon* **54**(4): 1080 - 1082.

Bonaldo M., Lennon G., Soares M.B. 1996. Normalization and subtraction: two approaches to facilitate gene discovery. Genome Res. 6: 791 – 806.

Bourgaud F., Bouque V., Guckert A. 1999. Production of flavonoids by *Psoralea* hairy root cultures. *Plant Cell, Tissue and Organ Culture* **56** (2): 96 - 103.

Bradford M.M. 1976. A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem*. **7** (72): 248 - 254.

Bruneton J. 1999. Pharmacognosy: phytochemistry medicinal plants (Isoquinoline alkaloids: Monoterpenoid isoquinolines). 2nd edition, Lavoisier publishing, Paris, France p. 961 - 966.

Brzobohatý B., Moore I., Kristoffersen P., Bako L., Campos N., Schell J., Palme K. 1993. Release of active cytokinin by a β-glucosidase localized to the maize root meristem. *Science* **262**: 1051 - 1054.

Campos N., Bako L., Feldwisch J., Schell J., Palme K. 1992. A protein from maize labeled with azido-IAA has novel β-glucosidase activity. *Plant J.* **2**:675 - 684.

Capell T, Christou P. 2004. Progress in plant metabolic engineering. *Current Opinion in Biotechnology* **15**: 148 - 54.

Cenci A, Chantret N, Kong X, Gu Y, Anderson OD, Fahima T, Distelfeld A, Dubcovsky. 2003. Construction and characterization of a half million clone BAC library of durum wheat (*Triticum turgidum* ssp. durum). *Theor Appl Genet* **107**(5): 931 - 939.

Chatel G., Montiel G., Pre M., Memelink J., Thiersault M., Saint-Pierre B., Doireau P., Gantet P. 2003. CrMYC1, a *Catharanthus roseus* elicitor- and jasmonate-responsive bHLH transcription factor that binds the G-box element of the strictosidine synthase gene promoter. *J Exp Bot.* **54**(392): 2587 - 2588.

Chomczynski P., Sacchi N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem.* **162**(1): 156 - 159.

Coutinho P.M., Henrissat B. 1999. Carbohydrate-active enzymes: an integrated database approach. In "Recent Advances in Carbohydrate Bioengineering", H.J. Gilbert, G. Davies, B. Henrissat and B. Svensson (eds.), *The Royal Society of Chemistry*, Cambridge, pp. 3 - 12. (URL: http://afmb.cnrs-mrs.fr/CAZY, accessed september-2006).

Creelman R.A., Mullet J.E. 1997. Biosynthesis and action of jasmonates in Plants. *Annu Rev Plant Physiol Plant Mol Biol* **48**: 355-381.

Croteau, R., Kutchan, T.M. und Lewis, N.G. 2000. Natural products (secondary metabolites). *In* Buchanan, B. Gruissem, W. Jones, R., eds, Biochemistry and Molecular Biology of Plants. American Society of Plant Biologists, Rockville, MD, pp. 1250 - 1268.

Czjzek M., Cicek M., Zamboni V., Bevan D.R., Henrissat B., Esen A. 2000. The mechanism of substrate (aglycone) specificity in β-glucosidases is revealed by crystal structures of mutant maize β-glucosidase-DIMBOA, -DIMBOAGlc and –dhurrin complexes. *Proc Natl Acad Sci USA* **97**: 13555 - 13560.

Czjzek M., Cicek M., Zamboni V., Burmeister W.P., Bevan D.R., Henrissat B., Esen A. 2001. Crystal structure of a monocotyledon (maize ZMGlu1) β-glucosidase and a model of its complex with p-nitrophenyl β-D-thioglucoside. *Biochem J* **354**: 37 - 46.

De Luca V., Marineau C., Brisson N. 1989. Molecular cloning and analysis of cDNA encoding a plant tryptophan decarboxylase: comparison with animal dopa decarboxylases. *Proc Natl Acad Sci*. **86**(8): 2582 - 2586.

De Luca V., St Pierre B. 2000. The cell and developmental biology of alkaloid biosynthesis. *Trends in plant science* **5**(4): 168 - 173

De Luca V., Laflamme P. 2001. The expanding universe of alkaloid biosynthesis. *Current Opinion in Plant Biology* **4**: 225 - 233.

De-Eknamkul W., Ounaroon A., Tanahashi T., Kutchan T.M., Zenk M.H. 1997. Enzymatic condensation of dopamine and secologanin by cell-free extracts of alangium lamarckii *Phytochemestry* **45**: 477 - 484

De-Eknamkul W., Suttipanta N., Kutchan T.M. 2000. Purification and characterization of deacetylipecoside synthase from *Alangium lamarckii* Thw. *Phytochemestry* **55**: 177 - 181.

Devies G., Henrissat B. 1995. Structures and mechanisms of glycosyl hydrolases. *Structure* **3**(9): 853 – 859.

Deutscher M.P. (ed) 1990. Methods in enzymology vol. 182: Guide to protein purification. Academic press, inc. California, USA. p. 893.

Dharmawardhana D.P., Ellis B.E., Carlson J.E. 1999. cDNA cloning and heterologous expression of coniferin β-glucosidase. *Plant Mol. Biol.* **40**:365 - 372.

Dietz K-J., Sauter A., Wichert K., Messdaghi, Hartung W. 2000. Extracellular β-glucosidase activity in barley involved in the hydrolysis of ABA glucose conjugate in leaves. *J Exp Bot.* **51:**937 - 944.

Doran, P.M. (ed) 1997. Hair roots: Culture and application. Hardwood academic publishers. Netherlands. p 239.

Eliert U., De Luca V., Constabel F., Kurz W.G.W. 1987. Elicitor-mediated induction of tryptophan decarboxylase and strictosidine synthase activities in cell suspension cultures of *Catharanthus roseus*. *Arch. Biochem. Biophys*. **254**: 491 – 497.

El-Sayed M., Verpoorte R. 2002. Effect of phytohormones on growth and alkaloid accumulation by a *Catharanthus roceus* cell suspension cultures fed with alkaloid precursors tryptamine and loganin. *Plant Cell, Tissue and Organ Culture* **68**(3): 265 - 270.

Esen A. 1993. β-Glucosidases: biochemistry and molecular biology. ACS symposium series 533:1-14, ISBN 0-8412-2697-0.

Evans, W.C. 1989. *Trease and Evans' Pharmacognosy,* 13th ed. London: Baillière Tindall pp. 595 - 599.

Facchini PJ, Bird DA, St-Pierre B. 2004. Can *Arabidopsis* make complex alkaloids? *TRENDS in Plant Science* **9**(3): 116 - 122.

Falk A., Rask L. 1995. Expression of a zeatin-*O*-glucoside-degrading *β*-glucosidase in *Brassica napus*. *Plant Physiol.* **108**:1369 – 1377.

Ferry N, Edwards MG, GatehouseJA, Gatehouse AMR. 2004. Plant–insect interactions: molecular approaches to insect. Resistance. *Current Opinion in Biotechnology* **15**: 155 - 161.

Fisher H.H. 1973. Origin and Uses of Ipecac. *Economic Botany* **27**: 231 - 234.

Frick, S., Chitty, J.A., Kramell, R., Schmidt, J., Allen, R.S., Larkin, P.J. and Kutchan, T.M. 2004. Transformation of opium poppy (*Papaver somniferum* L.) with antisense berberine bridge enzyme 1 (anti-bbe1) via somatic embryogenesis results in an altered ratio of alkaloids in latex but not in roots. *Transgenic Res.* **13**: 607 - 613.

Frick S, Kutchan TM. 1999. Molecular cloning and functional expression of *O*-methyltransferases common to isoquinoline alkaloid and phenylpropanoid biosynthesis. *Plant J* **17**(4): 329 - 39.

Frick S., Kramell R., Kutchan T.M. 2006. Metabolic engineering of a morphine biosynthetic P450 in opium poppy surpasses breeding. *Metabolic Engineering*. Article in press.

Fridman E., Pichersky E. 2005. Metabolomics, genomics, proteomics, and the identification of enzymes and their substrates and products *Current Opinion in Plant Biology* **8**: 242 - 248.

Fujii T., Ohba M. 1983. Ipecac alkaloids and β-carboline congeners. *The Alkaloids* **22**: 1 - 50.

Fujii T., Ohba M. 1998. The ipecac alkaloids and related bases. *The Alkaloids* **51**: 271 - 321.

Gebler J., Gilkes N.R., Claeyssens M., Wilson D.B., Béguin P., Wakarchuk W.W., Kilburn D.G., Miller R.C. Jr., Warren R.A., Withers S.G. 1992. Stereoselective hydrolysis catalyzed by related beta-1,4-glucanases and beta-1,4-xylanases *J. Biol. Chem.* **267**: 12559 - 12561.

Geerlings A., Ibañez M.M., Memelink J., van der Heijden R., Verpoorte R. 2000. Molecular cloning and analysis of strictosidine ß-D-glucosidase, an enzyme in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *J. Biol. Chem.* **275**: 3051 - 3056.

Gerasimenko I., Sheludko Y., Ma X., Stöckigt J. 2002. Heterologous expression of a *Rauvolfia* cDNA encoding strictosidine glucosidase, a biosynthetic key to over 2000 monoterpenoid indole alkaloids. *Eur. J.Biochem.* **269**: 2204 - 2213.

Grollman A.P. 1968. Inhibitors of protein biosynthesis: effects of emetine on protein and nucleic acid biosynthesis in HeLa cells. *The journal of biological chemistry* **243**(15): 4089 - 4094.

Gundlach H., Müller M.J., Kutchan T.M., Zenk M.H. 1992. Jasmonic acid is a signal transducer in elicitor-induced plant cell cultures *Proc. Natl. Acad. Sci*. **89**: 2389 - 2393.

Gustafsson C, Govindarajan S, Minshull J. 2004. Codon bias and heterologous protein expression. *TRENDS in Biotechnology* **22**(7): 346 - 353.

Hemscheidt T., Zenk MH. 1980. Glucosidases involved in indole alkaloid biosynthesis of *Catharanthus* cell cultures. *FEBS Lett*. **110**(2): 187 - 191.

Henriques A.T., Lopes S.O., Paranhos J.T., Gregianini T.S., Von Poser G.L., Fett-Neto A.G., Schripsema J. 2004. Links N, beta-D-Glucopyranosyl vincosamide, a light regulated indole alkaloid from the shoots of *Psychotria leiocarpa*. *Phytochemistry*. **65**(4): 449 - 454.

Henrissat B. 1991. A classification of glycosyl hydrolases based on amino-acid sequence similarities. *Biochem. J.* **280**: 309 - 316 .

Henrissat B., Bairoch A. 1993. New families in the classification of glycosyl hydrolases based on amino- acid sequence similarities. *Biochem. J.* **293**: 781 - 788.

Henrissat B., Bairoch A. 1996. Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* **316**: 695 - 696.

Henrissat B., Davies G.J. 1997. Structural and sequence-based classification of glycoside hydrolases. *Curr. Op. Struct. Biol.* **7**: 637 - 644.

Hofmann, K., Bucher, P., Falquet, L. & Bairoch, A. 1999. The PROSITE database, its status in 1999. *Nucleic Acid Res* **27**: 215 - 219.

Hösel W., Fiedler-Preiss A., Borgmann E. 1982. Relationship of coniferin *β*-glucosidase to lignification in various plant cell suspension cultures. *Plant Cell Tiss. Org.* **1**:137 - 148.

Hu W, Wang Y, Bowers C, Ma H. 2003. Isolation, sequence analysis, and expression studies of florally expressed cDNAs in *Arabidopsis*. *Plant Mol Biol* **53**(4): 545 - 563.

Ikan R. (*ed*) 1999. Naturally occurring glycosides. John Wiley & Sons Ltd. Endland. p. 444.

Ikeda K., Teshima D., Ayoama T., Satake M., Shimomura K. 1988. Clonal propagation of *Cephaelis ipecacuanha*. *Plant cell reports* **7**: 288 - 291.

Irmler S., Schröder G., St-Pierre B., Crouch N.P., Hotze M., Schmidt J., Strack D., Matern U., Schröder J. 2000. Indole alkaloid biosynthesis in *Catharanthus roseus*: new enzyme activities and identification of cytochrome P450 CYP72A1 as secologanin synthase. *Plant J.* **24**: 797 - 804.

Ishihara T, Sakurai N, Sekine KT, Hase S, Ikegami M, Shibata D, Takahashi H. 2004. Comparative analysis of expressed sequence tags in resistant and susceptible ecotypes of *Arabidopsis thaliana* infected with cucumber mosaic virus. *Plant Cell Physiol* **45**(4): 470 - 480.

Itoh A., Ikuta Y., Baba Y., Tanahashi T., Nagakura N. 1999. Ipecac alkaloids from *Cephaelis acuminata*. *Phytochemistry* **52**: 1169 - 1176.

Jha S., Sahu N.P., Mahato S.B. 1988. Production of the alkaloids emetine and cephaeline in callus cultures of *Cephaelis ipecacuanha*. *Planta Médica* January: 504 - 506.

Jha S., Sahu N.P., Sen J., Jha T.B., Mahato S.B. 1991. Production of emetine and cephaeline from cell suspension and excised root cultures of *Cephaelis ipecacuanha*. *Phytochemistry* **30**(12): 3999 - 4003.

Keresztessy Z., Brown K., Dunn M.A., Hughes M.A. 2001. Identification of essential active-site residues in the cyanogenic *β*-glucosidase (linamarase) from cassava (*Manihot esculenta* Crantz) by site-directed mutagenesis. *Biochem J.* **15**; 353(Pt 2): 199 - 205.

Koepke, J., Ma, X., Fritzsch, G., Michel, H., Stöckigt, J. 2005. Crystallization and preliminary X-ray analysis of strictosidine synthase and its complex with the substrate tryptamine. *Acta Cryst.* **D61:** 690 - 693.

Kutchan T.M., Hampp N., Lottspeich F., Beyreuther K., Zenk M.H. 1988. The cDNA clone for strictosidine synthase from *Rauvolfia serpentine*. *FEBS letters* **237**(1,2): 40 - 44.

Kutchan T.M. 1989. Expression of enzymatically active clone strictosidine synthase from the higher plant *Rauvolfia serpentia* in *Escherichia coli*. *FEBS letters* **257**(1): 127 - 130.

Kutchan T.M. 1993. Strictosidine: from alkaloid to enzyme to gene. *Phytochemistry* **32**(3): 493 - 506.

Kutchan T.M. 1995. Alkaloid biosynthesis - The basis for metabolic engineering of medicinal plants. *The plant cell* **7**: 1059 - 1070.

Kutchan T.M. 1998. Molecular genetics of plant alkaloid biosynthesis. *The Alkaloids*, edited by Cordell G.A. by Academic Press. **50**: 258 - 309.

Kutchan T.M. 2005. A role for intra- and intercellular translocation in natural product biosynthesis. *Current Opinion in Plant Biology* **8**: 292 – 300.

Larkin P.J., Miller J., Allen R.S., Millgate A.G., Chitty J.A., Gerlach W.L., Frick S., Kutchan T.M., Fist A.J. 2003. Increasing morphinan alkaloid production by overexpressing codeinone reductase in transgenic *Papaver somniferum*. *Plant Biotech. J.* doi: 10.1111/j.1467-7652.2006.00212.x.

Lloyd J.U. 1897. *Cephaelis ipecacuanha*. Reprint from: The Western Druggist, Chicago, August: 1 - 9.

Luijendijk T.J.C., Nowak A., Verpoorte R. 1996. Strictosidine glucosidase from suspension cultured cells of *Tabernaemontana divaricata*. *Phytochemistry* **41**: 1451 - 1456.

Ma X.Y., Koepke J., Fritzsch G., Diem R., Kutchan T.M., Michel H., Stöckigt J. 2004. Crystallization and preliminary x-ray crystallographic analysis of strictosidine synthase from *Rauvolfia* – The first member of a novel enzyme family. *Biochim. Biophys. Acta* **1702**: 121 - 124.

Ma X.Y., Panjikar S., Koepke J., Loris E., Stöckigt J. 2006. The structure of *Rauvolfia serpentina* strictosidine synthase is a novel six-bladed *β*-propeller fold in plant proteins. *The Plant Cell* **18**: 907 - 920.

Martins E., Oliveira L., Garcia R., Barros W., Pereira M., Vieira I., Pereira T. 2000. Conservation genetics of native populations of *Psychotria ipecacuanha* from south-eastern Brazil. International Conference on Science and Technology for Managing Plant. Genetic Diversity in the 21st Century. K.L. Hilton International Kuala Lumpur, Malaysia. p. 69.

McKnight T.D., Roessner C.A., Devagupta R., Scott A.I., Nessler C.L. 1990. Nucleotide sequence of a cDNA encoding the vacuolar protein strictosidine synthase from *Catharanthus roseus*. *Nucleic Acids Res*. **18**(16): 4939.

McPherson M.J., Jones K.M., Gurr S.J. 1991. PCR with highly degenerate primers. *In*: McPherson M J, Quirke P, Taylor G R. (*ed*). PCR: a practical approach. Oxford, United Kingdom: Oxford University Press; pp. 171 - 186.

Memelink J., Verpoorte R., Kijne J.W. 2001. ORCAnization of jasmonate-responsive gene expression in alkaloid metabolism. *Trends in Plant Science* **6**(5): 212 - 219.

Menke F.L., Champion A., Kijne J.W., Memelink J. 1999. A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene Str interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2. *EMBO J.* **18**(16): 4455 - 4463.

Muhammad I., Dunbar D.C., Khan S.I., Tekwani B.L., Bedir E., Takamatsu S., Ferreira D., Walker L.A. 2003. Antiparasitic alkaloids from *Psychotria klugii*. *J. Nat. Prod.* **66**(7): 962 - 967.

Murashige T., Skoog F. 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol Plant* **15**: 473 - 497.

Murata J, Bienzle D, Brandle JE, Sensen CW, De Luca V. 2006. Expressed sequence tags from Madagascar periwinkle (Catharanthus roseus). *FEBS Lett.* **580**(18): 4501 - 4507.

Nagakura N., Höfle G., Coggiola D., Zenk M.H. 1978 Biosynthesis of the ipecac alkaloids and of ipecoside and alangiside. *Planta Medica* **34**: 381 - 389.

Nakai K., Kanehisa M. 1992. A knowledge base for predicting protein localization sites in eukaryotic cells. *Genomics* **14**: 897 - 911.

Oliveira L.O., Martins E.R. 2002. A quantitative assessment of genetic erosion in ipecac (*Psychotria ipecacuanha*). *Genet. Resour. Crop Evol*. **49**: 607 - 617.

Ounaroon A., Decker G., Schmidt J., Lottspeich F., Kutchan T.M. 2003. (R,S)-Reticuline 7-*O*-methyltransferase and (R,S)-norcoclaurine 6-O-methyltransferase of *Papaver somniferum* - cDNA cloning and characterization of methyl transfer enzymes of alkaloid biosynthesis in opium poppy. *Plant J.* **36**(6): 808 – 819.

Palma T., Hidalgo N. 1993. Micropropagación *in vitro* de raicilla *Psychotria ipecacuanha*. Instituto Tecnológico de Costa Rica. Costa Rica, p 58.

Pascuali G., Goddijn O.J.M., De Waal A., Verpoorte R. 1992. Coordinated regulation of two indole alkaloid biosynthetic genes from *Catharanthus roseus* by auxin and elicitors. *Plant molecular biology* **18**: 1121 - 1131.

Pauli H.H., Kutchan T.M. 1998. Molecular cloning and functional heterologous expression of two alleles encoding (S)-N-methylcoclaurine 3'-hydroxylase (CYP80B1), a new methyl jasmonate-inducible cytochrome P-450-dependent mono-oxygenase of benzyliso-quinoline alkaloid biosynthesis. *Plant J* **13**(6): 793 - 801.

Pitta-Alvarez S.I., Giulietti A.M. 1995. Advantages and limitations in the use of hairy root cultures for the production of tropane alkaloids: use of anti-auxins in the maintenance of normal root morphology. *In vitro Cell. Dev. Biol- Plant* **31**: 215 - 220.

Poulton J.E. 1990. Cyanogenesis in Plants. *Plant Physiol.* **94**: 401 - 405.

Rodriguez S., Compagnon V., Crouch N.P., St-Pierre B., De Luca V. 2005. Jasmonate-induced epoxidation of tabersonine by a cytochrome P-450 in hairy root cultures of *Catharanthus roseus*. *Phytochemistry* **64**: 401 - 409.

Roewer I.A., Cloutier N., Nessler C.L., De Luca V. 1992. Transiente induction of tryptophan decarboxylase (TDC) and strictosidine synthase (ss) genes in cell suspension cultures of *Catharanthus roseus*. *Plant cell reports* **11**: 86 - 89.

Rout G.R., Samantaray S., Das P. 2000. *In vitro* somatic embryogenesis from callus cultures of *Cephaelis ipecacuanha* A. Richard. *Scientia Horticulturae* **86**: 71 – 79.

Ruyter C.M., Stockigt J. 1991. Enzymatic formation of raucaffricine, the major indole alkaloid of *Rauwolfia serpentina* cell-suspension cultures. Helv. Chim. Acta **74**: 1707 - 1712.

Saito K., Sudo H., Yamazaki M. 2001. Feasible production of camptothecin by hairy root culture of *Ophiorrhiza pumila*. *Plant cell reports* **20**: 267 - 271.

Sato F., Hashimoto T., Hachiya A., Tamura K., Choi K.B., Morishige T., Fujimoto H., Yamada Y. 2001. Metabolic engineering of plant alkaloid biosynthesis. *Proc Natl Acad Sci.* **98**(1): 367 - 72.

Sakurai N., Nakagawa-Goto K., Ito J., Sakurai Y., Nakanishi Y., Bastow K.F., Cragg G., Lee K.H. 2006. Cytotoxic Alangium alkaloids from *Alangium longiflorum*. *Phytochemistry*. **67**: 894 - 897.

Sambrook J., Fritsch E.F., Maniatis T. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schmeller T., Wink M. 1998. Utilization of alkaloids in modern medicine. In: Roberts M, Wink M (eds). *Alkaloids—Biochemistry, Ecology and Medicinal Applications*. New York: Plenum Press, pp 435 - 459

Skorupa L.A., Assis M.C. 1998. Collecting and conserving Ipecac (*Psychotria ipecacuanha,* Rubiaceae) germplasm in Brazil. *Economic Botany* **52**(2): 209 - 210.

Sonnhammer E.L.L., von Heijne G., Krogh A. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. In Proceedings of the of Sixth International Conference on Intelligent Systems for Molecular Biology (Glasgow, J., Littlejohn, T., Major, F., Lathrop, R., Sankoff, D. & Sensen, C., eds), Menlo Park, CA: AAAI Press. pp 175 - 182.

Stöckigt, J. 1995. Biosynthesis in *Rauwolfia serpentina* – modern aspects of an old medicinal plant. *In*: Cordell G.A. (ed). The Alkaloids. Academic Press, New York. Vol **47**, pp. 115 - 172.

Stöckigt J., Zenk M.H. 1977. Isovincoside (strictosidine), the key intermediate in the enzymatic formation of indole alkaloids. *FEBS Lett*ers **79**: 233 - 237.

Souza M., Martins E.R., Pereira T.N.S., Oliveira L.O. 2006. Reproductive studies on ipecac (*Cephaelis ipecacuanha* (Brot.) A. Rich; Rubiaceae): meiotic behavior and pollen viability. *Braz. J. Biol.* **66**(1A): 151 - 159.

Tan G.T., Kinghorn A.D., Hughes S.H., Pezzuto J.M. 1991. Psychotrine and its *O*-metil ether are selective inhibitors of human immunodeficiency Virus-1 Reverse transcriptase. *The journal of biological chemistry* **226**(35): 23529 - 23536.

Teshima D., Ikeda K., Satake M., Aoyama T., Shimomura K. 1988. Production of emetic alkaloid by *in vitro* culture of *Cephaelis ipecacuanha* A. Richard. *Plant cell reports* **7**: 278 - 280

Teshima D., Suzuki A., Otsubo K., Higuchi S., Aoyama T., Shimozono Y., Saita M., Noda K. 1990. Efficacy of emetic and united state pharmacopoeia Ipecac Syrup in prevention of drug absortion. *Cham. Pharm. Bull*. **38**(8): 2242 - 2245.

Trimburs D.E., Warren A.R.J., Withers S.G. 1992. Region-directed mutagenesis of residues surrounding the active site nucleophile in *β*-glucosidase from *Agrobacterium faecalis*. *J. Bio. Chem.* **267**(15): 10248 - 10251.

Van der Fits L, Zhang H, Menke FL, Deneka M, Memelink J. 2000. A *Catharanthus roseus* BPF-1 homologue interacts with an elicitor-responsive region of the secondary metabolite biosynthetic gene Str and is induced by elicitor via a JA-independent signal transduction pathway. *Plant Mol Biol.* **44**(5): 675 - 685.

Vazquez-Flota F., De Carolis E., Alarco A.M., De Luca V. 1997. Molecular cloning and characterization of desacetoxyvindoline-4-hydroxylase, a 2-oxoglutarate dependent-dioxygenase involved in the biosynthesis of vindoline in *Catharanthus roseus* (L.) G. Don. *Plant Mol Biol.* **34**(6): 935 - 948.

Veeresham C., Kokate C., Venkateshwarlu V. 1994. Influence of precursors on production of isoquinoline alkaloids in tissue cultures of *Cephaelis ipecacuanha*. *Phytochemistry* **35**(4): 947 - 949.

Verdoucq L., Morinière J., Bevan D.R., Esen A., Vasella A., Henrissat B., Czjze M. 2004. structural determinants of substrate specificity in family 1 *β*-glucosidases. *J. Biol. Chem.* **279**(30): 31796 - 31803.

Warzecha H., Obitz P., Stöckigt J. 1999. Purification, partial amino acid sequence and structure of the product of raucaffricine-*O*-*β*-D-glucosidase from plant cell cultures of *Rauvolfia serpentina*. *Phytochemistry* **50**: 1099 - 1109.

Warzecha H., Gerasimenko I., Kutchan T.M., Stöckigt J. 2000**.** Molecular cloning and functional bacterial expression of a plant glucosidase specifically involved in alkaloid biosynthesis. *Phytochemistry* **54**: 657 - 666.

Wink M. 1999. Plant secondary metabolites: biochemistry, function and biotechnology. *In* Biochemistry of Plant Secondary Metabolism *Annual Plant Reviews* **2**: 1 - 16.

Withers S.G., Warren R.A.J., Street I.P., Rupitz K., Kempton J.B., Aebersold R. 1990. Unequivocal demonstration of the involvement of a glutamate residue as a nucleophile in the mechanism of a retaining glycosidase *J. Am. Chem. Soc.* **112**: 5887 - 5889.

Wöhler F., Liebig J. 1837. Über die bildung des bittermandelöls. *Annalen der Pharmacie* **22**:11 – 14.

Wright C.W., Phillipson J.D. 1990. Natural products and the development of selective antiprotozoal drugs. *Phytother Res*. **4**: 127 - 139.

Yoshimatsu K., Shimomura K. 1991a. Efficient shoot formation on internodal segments and alkaloid formation in the regenerates of *Cephaelis ipecacuanha* (A. Richard). *Plant Cell Reports* **9**: 567 - 570.

Yoshimatsu K., Shimomura K. 1991b. Emetic alkaloid formation in root culture of *Cephaelis ipecacuanha*. *Phytochemistry* **30**: 505 - 507.

Yoshimatsu K., Shimomura K. 1993. *Cephaelis ipecacuanha* A. Richard (Brazilian Ipecac): Micropropagation and the Production of emetine and Cephaeline. *Biotecnology in agriculture and forestry* **21**: 87 - 103.

Yoshimatsu K., Shimomura K., Yamazaki M., Saito K., kiuchi F. 2003. Transformation of ipecac (*Cephaelis ipecacuanha*) with *Agrobacterium rhizogenes*. *Plant Med.* **69**: 1018 - 1023.

Yun D-J., Hashimoto T., Yamada Y. 1992. Metabolic engineering of medicinal plants: Trangenic *Atropa belladonna* with an improved alkaloid composition. *Proc. Natl. Acad. Sci.* **89**: 11799 - 11803.

Ziegler J., Diaz-Chavez M.L., Kramell R., Ammer C., Kutchan T.M. 2005. Comparative macroarray analysis of morphine containing *Papaver somniferum* and eight morphine free *Papaver* species identifies an *O*-methyltransferase involved in benzylisoquinoline biosynthesis. *Planta*. **222**(3): 458 - 471.

Ziegler J., Voigtlander S., Schmidt J., Kramell R., Miersch O., Ammer C., Gesell A., Kutchan T.M. 2006. Comparative transcript and alkaloid profiling in Papaver species identifies a short chain dehydrogenase/reductase involved in morphine biosynthesis. *Plant J.* **48**(2): 177 - 192.

Zulak K.G., Cornish A., Daskalchuk T.E., Deyholos M.K., Goodenowe D.B., Gordon P.M., Klassen D., Pelcher L.E., Sensen C.W., Facchini P.J. 2006. Gene transcript and metabolite profiling of elicitor-induced opium poppy cell cultures reveals the coordinate regulation of primary and secondary metabolism. *Planta* Nov. [Epub ahead of print].

Acknowledgments

I would like to express my sincere gratitude to Prof. Dr. Toni M. Kutchan for introducing me in the topic of this work, the supervision of the project and valuable discussions on the work. I am also thankful for all motivating advice and the excellent work conditions facilitated to achieve of this work.

I want to express my gratitude to Dr. Robert Kramell for valuable discussions and advise concerning theoretical and practical aspects of the work especially on HPLC and LC-MS methods.

I am especially grateful to Dr. Jörg Ziegler for the help in producing of cDNA library from root cultures, stimulating discussions and advice.

I am particularly grateful to Dr. Thomas Vogt for the substrates kindly provided, stimulating discussions and critical reading of the manuscript.

I would like to thank to Dr. Jürgen Schmidt for the very appreciated help with spectrometric analysis, LC-ESI-MS/MSSI-MS/MS.

I want to express my thanks to Prof. Dr. Wasternack for valuable advice and help in several ways.

I would like to thank Dr. G. Herrmann for advice in culture tissues. Thanks also to Frau Domenika Arndt for the help with *Psychotria ipecacuanha* plants.

I am also thankful to Prof. Dr. Stöckigt for the invaluable substrates ipecoside, arbutin, raucaffricine, strictosidine and vincoside lactam kindly provided.

I want to express also my thanks to Dr. Miersch, Dr. Page and Dr. Frick.

A big personal thanks to my dear colleges and friends Andreas G., Tobias K., María Luisa D. Ch., Khaled S., Stefanie H., Nadia G., Katja K., Ursula S., Verona D., Jana N., Frau Wegener, Carolin D., Marco S., Birgit O., Christian B., Silke P., Kathleen B., Christine D., Frau Hillert, Anja S. and all the co-workers that helped me in different ways. Special thanks to Nicole F. and Prof. J. Jacobsen.

Thanks a lot to my family for the support in all aspects of my life, their love and confidence.

This work was done at the Institute of Plant Biochemistry (IPB) in Halle (Saale). I would like to thank all the co-workers of the institute, especially from the "Naturstoff Biotechnologie Abteilung" for the scientifically stimulating environment that has promoted this work.

This work has been supported by the "Deutscher Akademischer Austauschdienst (DAAD) "
I would like to thank the DAAD for the opportunity and facilities to develop this work.