

Molecules of Interest

Serine carboxypeptidase-like acyltransferases

Carsten Milkowski*, Dieter Strack

Leibniz-Institut für Pflanzenbiochemie, Weinberg 3, 06120 Halle (Saale), Germany

Received 19 December 2003; received in revised form 19 December 2003

Abstract

In plant secondary metabolism, an alternative pathway of ester formation is facilitated by acyltransferases accepting 1-*O*- β -acetal esters (1-*O*- β -glucose esters) as acyl donors instead of coenzyme A thioesters. Molecular data indicate homology of these transferases with hydrolases of the serine carboxypeptidase type defining them as serine carboxypeptidase-like (SCPL) acyltransferases. During evolution, they apparently have been recruited from serine carboxypeptidases and adapted to take over acyl transfer function. SCPL acyltransferases belong to the highly divergent class of α/β hydrolases. These enzymes make use of a catalytic triad formed by a nucleophile, an acid and histidine acting as a charge relay system for the nucleophilic attack on amide or ester bonds. In analogy to SCPL acyltransferases, bacterial thioesterase domains are known which favour transferase activity over hydrolysis. Structure elucidation reveals water exclusion and a distortion of the oxyanion hole responsible for the changed activity. In plants, SCPL proteins form a large family. By sequence comparison, a distinguished number of Arabidopsis SCPL proteins cluster with proven SCPL acyltransferases. This indicates the occurrence of a large number of SCPL proteins co-opted to catalyse acyltransfer reactions. SCPL acyltransferases are ideal systems to investigate principles of functional adaptation and molecular evolution of plant genes.

© 2004 Published by Elsevier Ltd.

Keywords: Acyltransferases; Serine carboxypeptidases; α/β Hydrolases; Sinapate esters; Molecular evolution

1. Introduction

One of the decisive steps in the evolution of plant secondary compounds is gene duplication and evolution of new catalytic properties of the duplicated gene product. Published examples are biosyntheses of pyrrolizidine alkaloids (Ober and Hartmann, 1999, 2000) and the benzoxazinone derivatives DIBOA/DIMBOA (Gierl and Frey, 2001); Deoxyhypusine synthase (EC 1.14.99.29), an enzyme modifying translation initiation factors, and a subunit of a tryptophan synthase (EC 4.2.1.20), a hydrolyase, are recruited to catalyse homospermidine and indole formation, respectively.

Recently, acyltransferases came also into focus of scientific interest on molecular evolution. The acyl

donors in acyltransferase reactions are mostly coenzyme A thioesters, however, β -acetal esters (1-*O*- β -glucose esters) have repeatedly been described as alternative substrates (Strack and Mock, 1993). UDP-glucose-dependent glucosyltransferases catalyse biosynthesis of the glucose esters, which exhibit high group transfer potentials (Mock and Strack, 1993; Leznicki and Bandurski, 1988).

Li and coworkers (1999) described a glucose ester-dependent acyltransferase from wild tomato (*Lycopersicon pennellii*) which catalyses both the disproportionation of 1-*O*-acyl- β -glucose to produce 1,2-di-*O*-acyl- β -glucose and anomeric acyl exchange between 1-*O*-acyl- β -glucose and glucose (Li et al., 1999). Sequence analysis of the corresponding cDNA revealed unexpected homology with peptide hydrolases of the serine-type carboxypeptidase (EC 3.4.16.1) (Li and Steffens, 2000). The same striking result was obtained for genes encoding enzymes involved in phenylpropanoid metabolism, i.e. 1-*O*- β -sinapoylglucose:L-malate sinapoyltransferase (SMT) from Arabidopsis (Lehfeldt

* Corresponding author. Tel.: +49-345-5582-1533; fax: +49-345-5582-1509.

E-mail addresses: cmilkows@ipb-halle.de (C. Milkowski), dstrack@ipb-halle.de (D. Strack).

et al., 2000) and 1-*O*- β -sinapoylglucose:choline sinapoyltransferase (SCT) from *Arabidopsis* (Shirley et al., 2001) and rape (*Brassica napus*) (Milkowski et al., 2004); sequence features classify these enzymes also as serine carboxypeptidase-like (SCPL) proteins.

SMT: \rightarrow 1-*O*- β -sinapoylglucose + L-malate
 \rightarrow sinapoylmalate + glucose

SCT: \rightarrow 1-*O*- β -sinapoylglucose + choline
 \rightarrow sinapoylcholine + glucose

The relationship with hydrolases makes this new class of SCPL acyltransferases a challenging target for research on fundamental questions on molecular evolution and mechanisms of enzyme-catalysed reactions. Steffens (2000) drew special attention to these enzymes in an article entitled “acyltransferases in protease’s clothing”. The most important question to be addressed is on the mechanism by which evolution recruits genes of primary metabolism into secondary metabolism to catalyse novel reactions.

2. SCPL character of glucose ester-dependent acyltransferases

Homology of glucose ester-dependent acyltransferases and hydrolases of the serine carboxypeptidase-type is illustrated by significant overall sequence identities (e.g. 37% of BnSCT with CPI from barley; 36% with SCPI from rice). This indicates that these acyltransferases and serine carboxypeptidases may have developed as a result of divergent evolution. The sequence alignment of SCPL acyltransferases and serine carboxypeptidases (SCP) defines highly conserved blocks of amino acid residues (Fig. 1). These blocks are preferentially located around the residues that make up the catalytic triad (Ser, His, Asp) in the active centre of SCPs suggesting an important functional role for these sequence elements in acyltransferases as well. Diisopropylfluorophosphate (DFP), a potent inhibitor of catalytic seryl residues, was shown to block the disproportionation reaction of the glucose acyltransferase from wild tomato (Li and Steffens, 2000). A similar inhibitory effect has been reported for SMT and SCT from *Arabidopsis* after treatment with phenylmethylsulfonyl fluoride (PMSF), another specific inhibitor of ‘serine enzymes’ (Lehfeldt et al., 2000; Shirley et al., 2001). Computational analyses identify two other characteristic sequence features: N-terminal signal peptides for targeting to the endoplasmic reticulum and potential N-glycosylation sites. This suggests that glucose ester-dependent acyltransferases are synthesized as pre-proteins. Via the endoplasmic reticulum and Golgi apparatus where the proteins are glycosylated, they should be transported

to the vacuole as it was shown for SCPs (Zuber and Matile, 1968). First experimental proof for this assumption has been obtained by immunolabeling approaches. For *Arabidopsis* SMT it could be shown that the mature protein is located in the central vacuoles of the leaf tissue (Hause et al., 2002). Lectin binding and differences in molecular weight between SMT variants purified from *Arabidopsis* and overexpressed from *E. coli* indicate that the SMT is highly glycosylated. Beyond sequence features, similarities of serine carboxypeptidases and glucose ester-dependent acyltransferases seem to extend to their structures. Immunological studies on *Arabidopsis* SCT protein revealed post-translational processing that results in the endoproteolytic cleavage of the expected 50 kDa protein to form polypeptides of 30 kDa and 17 kDa characteristic for the mature protein (Shirley and Chapple, 2003). A similar kind of modification has been shown for carboxypeptidase I (CPI) from barley (Doan and Fincher, 1988) and carboxypeptidase II (CPDWII) from wheat (Breddam et al., 1987) forming both heterotetramers. For barley-CPI (CPI-Hv, Fig. 1) it has been shown the endoproteolytic excision of a 55-residue linker peptide ranging from Ile²⁶⁷ to Ser³²¹ of the precursor polypeptide. For CPDWII, it has been shown that the chain A of the mature protein ends with Ser²⁶³, whereas chain B starts with Thr²⁸⁵ of the pre-protein (Breddam et al., 1987). On the other hand, it is striking that SMT is not cleaved post-translationally as was confirmed by Western Blot analysis (Hause et al., 2002). This resembles yeast carboxypeptidase (CPDY) which does not undergo endoproteolysis during maturation.

3. Functional aspects: SCPL but not SCP

Despite of their SCPL character, the glucose ester-dependent acyltransferases isolated so far are not able to catalyse peptide hydrolysis. From studies on yeast carboxypeptidase Y (Hayashi et al., 1973, 1975; Bech and Breddam, 1989; Endrizzi et al., 1994) and serine carboxypeptidases from barley and wheat (Baulcombe et al., 1987; Doan and Fincher, 1988; Liao and Remington, 1990; Liao et al., 1992; Degan et al., 1994) it is known that carboxypeptidase activity is based on a charge relay system composed of seryl, histidyl and aspartyl residues in the catalytic centre (Fig. 2). By interaction with the imidazol-*N* of the histidyl residue, the seryl residue becomes a strong nucleophile that attacks the carbonyl carbon of a peptide bond resulting in the formation of an acyl-enzyme intermediate. This intermediate is produced via a negatively charged transition state with tetrahedral geometry. The transition state is stabilized by hydrogen bonds to groups located in a pocket of the enzyme called the oxyanion hole. The acyl-enzyme intermediate is then cleaved by water as a

GAC-Lp	-----EHFIVETLPGFHCKLPFTLETGYISVGEEEKVOLFVFVQSERDPRNDPLMTWLTGGPGCSSLSS--FVYEIGPLTDEYANS	80
SMT-At	-----ASIVKFLPGFECPLPFELETGYISGSDENVQFFYFIKSENNPKDDPLLIWNGGPGCSSLG--IIFENGPVGLKEFVF	79
SCT-At	-----SLLVKSLLPGFECPLPFELETGYISGSDVDFLYFVISENNPNDPLMTWLTGGPGCSSLIG--LLEFANGPLAFKGEY	79
SCT-Bn	-----ASLHVKYLPGFECPLPFELETGYISGSDVDFLYFVISENNPKDDPLMTWLTGGPGCSSLIG--LLEFANGPLAFKGEY	80
CPDY	-----KIKDKPKLID--PNVTQYTGIDVED--EDKHFFETFFSRNPAKDPVILWNGGPGCSSLTG--LFEEFGPSSIGDPLK	76
CPDWII	---VEPSGHAADRIARLPQGP--AVDFDMYSGYITVDEGAGRSILFYLLQEAPEAQAPAPLVILWNGGPGCSSLVAYGASEELG--AFRVKPR	84
CPI-Hv	-----APQGAEVTLPGFECALPSKHYAGYITVDEGHGRNLFYVVESEDPGKDPVILWNGGPGCSSLFDG--FVYEHGPFNFESGGS	82
CBP1-Os	GGGVCEAAPASAVVKSVPCEFCALPSKHYAGYITVEEQHGRNLFYLVSESDPAKDPVILWNGGPGCSSLFDG--FVYEHGPFNFESGGS	89
SCP1-Le	-----APQSALVTQLPGFECALPSKHYAGYITVDESHGKNLYFVSESDPAKDPVILWNGGPGCSSLFDG--FVYEHGPFNFDFGKP	82
GAC-Lp	SGNFKLELNSYSWAKVANIIFIDQAGGYSYANTSE--AYNCNLTSLVTLTMDFTPKMLMDPEYINN--PLYVGGDSYSGIFVALLTR	167
SMT-At	NGSAPSLFSTTYSWAKMANIIFLDQVGSYGSYSKTP--IDKTDGISEVKTTHFTLOKLSRPOYFSN--PLYVVGDSYSGMIVPALVQ	165
SCT-At	NGTVPLLELTSYSWAKVANILYLEAPAGSCYSYAKTRR--AFESDDTKQMHQIDQFTRSEVVKPEFISN--PFYVGGDSYSGKIVEGAVQ	166
SCT-Bn	NGTLPLLELTSYSWAKVANILYLEAPAGSCYSYAKTRR--AAETSDTKQIHQIDQFTRSEVVDPEFISN--SFYVGGDSYSGKIVEGVVQ	167
CPDY	PIG-----NPTYSNSNATVIFLDQVNVGFSYSGSSG--VSTNVAAEKQDVNFTLEEDQEPPEYVNGQDSEIAGESYAGHYIEVFS	157
CPDWII	GAG-----IVLNEFRMKNVANVLELDSBAGVGFYSYNTSSDIYTSNDRTAHDSEAFIAKWEEREPHYKYR--DFYIAGESYAGHYIVELISQ	169
CPI-Hv	VKSLKHLNLPYAWKSVSTMIYLDSPAGVGLSYSKNVS--DYETGDLKTATDSHTFLKWFQLYPEFLSN--PFYIAGESYAGHYIVETLSH	169
CBP1-Os	AKSLKHLNLPYWSKVSSVIYLDSPAGVGLSYSKNTS--DYNTGDLKTAADSHFTFLKWFQLYPEFLSN--PFYIAGESYAGHYIVETLSH	176
SCP1-Le	SGSLPSLHNNPYSWKVSNIIYLDSEVGVGLSYSGNKS--DYNTGDLKTAADSHTFLKWEETIPEFLKN--PFYISGESYAGHYIVETLAS	169
GAC-Lp	KIYDCEIEVGDRFRVNIKGYIQGNAITDRSIDFNGRUKYANHGLISDKIYQSAKANCNENYIDVDPNNILCLNDLQKVTRCLKNIRRAQI	257
SMT-At	EISQENYICCEPPINLOGYMGNPVTYMDFEONFRIPYAYGMGLISDEIYEPMKRICNENYINVDPSNTQCLKLTTEEYHKCTAKINIHII	255
SCT-At	QILLCNEKGLTFLINIGYVGNPTDKNIEITNYRVP--AHGMGLISDELFESELRSCGCKFENVDPSNARCSNNLQAYDHCMSEIYSEHI	256
SCT-Bn	QISLCKNEKGLTFLINIGYVGNPAVRTNLEPNHRVSP--AHRMGLISDELFESELRNCGCKFENVDPSNARCSNNLQAYDHCMSEIYSEHI	257
CPDY	EILSHKDRN-----ENLTSVLLGNGLTDPLOTQNYEPMACGEGGEPSVLPSEECSAMEDSLERCLGLIESCYDSQSWSVCVPATTYCNA	243
CPDWII	LVHRS--KN--EVLINLKGEMVGNGLIDDYHGVYGTFFEWNNHGTISDDTYRRLKEACLDHSTIHP--SPACDAATVATAEQNDIMYST	253
CPI-Hv	EVVKIQGGAKEITINFGYMGVNGVCDTIFDGNALVP--AHGMGLISDEIYQCASTSCHENYNA--TDGKCDTAISGLISGLNIYDI	257
CBP1-Os	EVVKELHDGVKETINFGYMGVNGVCDTVFDGNALVPAHGMALISDDIYQEAQTACHENYNT--TTDKCENALYKVDTSINDLNIYDI	264
SCP1-Le	EVIKIDAGVREAINFMGYMGVNGVADDIIDGNALVPEQHGMGLISDDIYEEAVVACHENFYEP--VDSNCSSEKLNKIDQVVVDINVYDI	257
GAC-Lp	LEP-----YCDLPYLMG-----ILQETPTNGQSVFPIAGPWCREKN-----YIYSYVWANDKAVQ	307
SMT-At	LTE-----DCDVTNVT-----PDCTYYP-----YHLIECWANDESVR	288
SCT-At	LLR-----NCKVDYVLADT-----PNIRTRRRVMKEFSVNDSSSL--PPPSCTFYR-----YFLSAFWANDENVR	315
SCT-Bn	LLE-----NCKVDYVLADISQTLNIRTSRRRELKEFSRNDSSSL--PPPSCTFYR-----YFLSAFWANDENVR	320
CPDY	QLA-----PYQR-TG-----RNVDIRKDEGG-----NLCYPTL-----QDIDDLINDQYV	285
CPDWII	YTE-----VCNITSSSS-----SSSSLSQQRSSRGYPWLTGSYDPT-----ERYSTAIYNRDVO	306
CPI-Hv	LEPCYHSRSIKEVNLQNSKLPQSFKDLGTTNKPFPVRTMLGRAWFLRAPVKAG--RVESWQEVA-----SGVPCMSDEVATANDNAVR	341
CBP1-Os	LEPCYHSKTIKKVTPANTKLPKSFQHLGTTTKPLAVTRMHRGAWFLRAPVRAG--RVESWQEFARGSRPSSGVPCMSDEVATANDNDVR	353
SCP1-Le	LEPCYHSKKPSVITGNSRLPMSFRKLGETERPLVRKRMFGRAWPYKAPVRAG--HVETWPEILN---SVEVPCTDDRVATLNLNADVR	343
GAC-Lp	KALNVREG--TTLEIVRQNESMHYRGKERTESVYDVPSEVDDHQHLLT--SKSCRALISGDHDMVPHLSTEEIETIKPLIADDEW-	391
SMT-At	EALHIEKG--SKGKKARON-----RTIPYNHDIIVSSIPYHMNS--ISGYRSLISGDHDIAPVLAQAIRSNISYPIHNWR-	363
SCT-At	RALGVKK--VGKKNRONS-----QNIPYTFEIFNAVYPYHNS--LKGRSLISGDHDSMVP--SSQAIRSNISYIVDDWR-	389
SCT-Bn	RALGVKKG--FGKNSRON-----QNIPYTYLIHNAIPYHNS--RKGRALISGDHDMMP--SSQAIRSNISYIVDDWR-	395
CPDY	EAVGAEDV--HYESON-----FDINRNLFAGDWKMYHTAVTLLNODLPILVYAGDKDFICNLGNKALTDVLPFWKYDEEAS	363
CPDWII	MALHANVYGGAMNYTATCS-----DTINTHWHDAPEMIPYRELI--AAGRIWVESGDTDAVVLTAIRYSIGALGLPTTTSWY-	385
CPI-Hv	SAIHAQSV--SAIGPWLCT-----DKLYVHDAGSMIAHKNLT--SCGYRALISGDHDMCVPTGSEAWTKSLGYGVVDSWR-	417
CBP1-Os	AATHAQPV--SSIGSNLCT-----NVLDIHDAGSMISYHKNLT--GCGYRALISGDHDMCVPTGTEATRSIGYGVVDSWR-	429
SCP1-Le	KATHAHPA--TVIGPWLCT-----DKIDLHDSGSMIPYHKNLT--AGCYRALISGDHDMCVPTGSAVATKSLGYPIVDWR-	419
GAC-Lp	----PWFVD--DOVAGYKVYLQNDYEMTATVKACHTAPEYKPEOCLPMVDRWFSGOPL-----	446
SMT-At	----PWMIN--NOTAGYTRAYSN--KITEATHKGGHTA--EYRPNETFIMFORWISGOPL-----	414
SCT-At	----PWWMS--NOVAGYTRTYAN--KITEATHKGGHTA--EYTEDQCSLMERRWIDGESL-----	441
SCT-Bn	----PWWMTS--NOVAGYTRTYAN--KITEATHKGGHTA--EYNPDOCSLMERRWIDGESL-----	447
CPDY	QKVRNWTASITDEVADEVKSYKH--FTMLRVFNGGHMVPFDVPENALSMVNEWTHGGFSL-----	421
CPDWII	----PWYDD--QEVGGNSQVYKG--LTLVSVRGAGHEVPLHRPRCALVLEQYELQCKPMPGQTKNAT	444
CPI-Hv	----PWITN--GOVSGYTEGYEH--GLTEATHKAGHTVPEYKPEALAFYSRWLAGSKL-----	469
CBP1-Os	----PWHLN--GOVSGYTEGYEH--GLTEATHKAGHTVPEYKPEALAFYSRWLAGSKL-----	481
SCP1-Le	----PWYVN--DOVAGYTGQYAN--NLIFMTAKAGHTVPEYKPEALAFYSRWLAGSKL-----	471

Fig. 1. Alignment of deduced amino acid sequences of 1-*O*- β -glucose ester-dependent acyltransferases and serine carboxypeptidases. N-terminal leader peptides have been removed. Fully conserved residues are shaded in black. Grey shading indicates conservation of at least 70%. Amino acid residues forming the catalytic triad of serine carboxypeptidases (S, D, H) are marked in red. GAC-Lp (SwissProt accession Q9LKY6)=glucose acyltransferase from *Lycopersicon penellii*; SMT-At (O64809)=sinapoylglucose:L-malate sinapoyltransferase from *Arabidopsis*; SCT-At (Q941P1)=sinapoylglucose:choline sinapoyltransferase from *Arabidopsis*; SCT-Bn (AAQ91191)=SCT from *Brassica napus*; CPDY (P00729)=carboxypeptidase Y from *Saccharomyces cerevisiae*; CPDWII (P008819)=carboxypeptidase II from *Triticum aestivum*; CPI-Hv (P07519)=serine carboxypeptidase I from *Hordeum vulgare*; CBP1-Os (P37890)=carboxypeptidase I from *Oryza sativa*; SCP1-Le (Q9M513)=serine carboxypeptidase I from *Lycopersicon esculentum*. The alignment was performed with the CLUSTALW algorithm using the BLOSUM62 matrix.

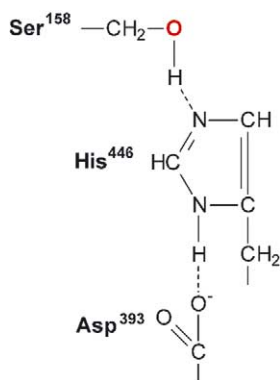


Fig. 2. A schematic view of the catalytic triad from CP1-Hv (sequence in Fig. 1) catalysing cleavage of the C-terminal peptide bond in polypeptides. Interaction with the imidazol-*N* of the histidyl residue 446 enables the nucleophilic atom (shown in red) of the seryl residue 158 to attack the carbonyl carbon of the peptide bond, resulting in liberation of the C-terminal amino acid and enzyme acylation by the peptide fragment that is finally hydrolytically cleaved from the seryl residue. The aspartyl residue 393 in the catalytic triads of SCPs makes the histidyl residue a better proton acceptor to transform the seryl hydroxyl group into a substrate-attacking alkoxide ion.

nucleophile liberating the bound peptide and regenerating the active seryl residue.

As there is no striking difference in the sequence context of the amino acid residues forming the catalytic triad in serine carboxypeptidases (Fig. 1), it is tempting to suggest an analogous mechanism for transacylation: The strongly nucleophilic serine at the catalytic centre of glucose ester-dependent acyltransferases could attack the carbonyl carbon of the ester substrate (e.g. sinapoylglucose) producing an acyl-enzyme intermediate (e.g. sinapoylated SMT protein). The ester-bond of this protein-substrate intermediate should then be cleaved by the nucleophilic attack of the second substrate (e.g. L-malate).

The most challenging question in the case of SCPL acyltransferases is to explain their specificity at the molecular level. Why do these enzymes not act as SCPs given the sequence features classifying them as SCPL proteins? Li and Steffens (2000) report that there is no sequence conservation of glucose acyltransferase from wild tomato and yeast carboxypeptidase Y in a protein region forming the *H*-bond network necessary for peptide recognition by the latter (Endrizzi et al., 1994; Mortensen et al., 1994). This may prevent peptide substrates from being bound by the SCPL acyltransferases. Even more difficult it seems to explain why these acyltransferases do not display hydrolytic activity on the glucose ester substrates. In general, there are two scenarios discussed: (i) prevention of water from the catalytic centre by hydrophobic shielding of the active site and (ii) mechanisms giving the second substrate a higher nucleophilic potential than water. The latter mechanism could be based on a general base mechanism as proposed recently for a C–C hydrolase of *E. coli* (Fleming

et al., 2000). This would require the catalytic serine acting as a general base to deprotonate the hydroxyl group of the second substrate (e.g. L-malate in SMT reaction or choline in SCT reaction) resulting in a direct nucleophilic attack on the glucose ester substrate. Whereas the mechanism of water exclusion would be consistent with a double-displacement mechanism of reaction (ping-pong mechanism), the latter model would require a random or sequential bi-bi mechanism. Kinetic analysis published so far give a controversial picture. A study on the sinapoyltransfer of the native SMT from radish suggested a random bi-bi mechanism (Gräwe et al., 1992). On the other hand, kinetic studies on SCT purified from rape (Vogt et al., 1993) and Arabidopsis (Shirley and Chapple, 2003) give rise to a double displacement mechanism of catalysis. As it turns out from the preliminary data, neither biochemical and kinetic studies nor sequence analysis will be sufficient to explain the differences between SCPs and SCPL acyltransferases at the molecular level. This would finally require structure elucidation of SCPL acyltransferases and comparison of these structures with those of related peptidases.

4. Evolution

With a closer look at related enzymes, the homology of glucose ester-dependent acyltransferases and SCPs loses a bit of its initially stated peculiarity. Both enzyme groups belong to a highly divergent family whose members catalyse a wide array of different, mostly hydrolytic, reactions. At the sequence level there is no pronounced similarity detectable, but all enzymes of this family contain a central catalytic domain of unique topology and three dimensional structure designated as “ α/β hydrolase fold” (Ollis et al., 1992). It seems that the α/β hydrolase fold structure has been preserved during evolution because it constitutes a simple, stable and effective way of building a variety of different catalytic triads composed of a nucleophile, an acid and histidine with the potential to set up a charge relay system as described for proteinases. Moreover, this structure element allows to arrange oxyanion hole motifs and different binding pockets (Brändén, 1986). Thus, α/β hydrolase fold enzymes catalyse a wider variety of hydrolysis reactions than any other class of catalytic triad enzymes (Ollis et al., 1992). This diversity is illustrated by the considerable number of different enzyme groups sharing the α/β hydrolase fold character: carboxypeptidases (Liao and Remington, 1990), prolyl-oligopeptidases (Fulop et al., 1998), lipases (Winkler et al., 1990; Lang et al., 1998; Nardini et al., 2000), esterases (Bourne et al., 2000), thioesterases (Lawson et al., 1994; Li et al., 2000; Bellizzi et al., 2000; Devedjiev et al., 2000), or hydroxynitrile lyases (Wajant et al., 1994; Lauble et al., 2001).

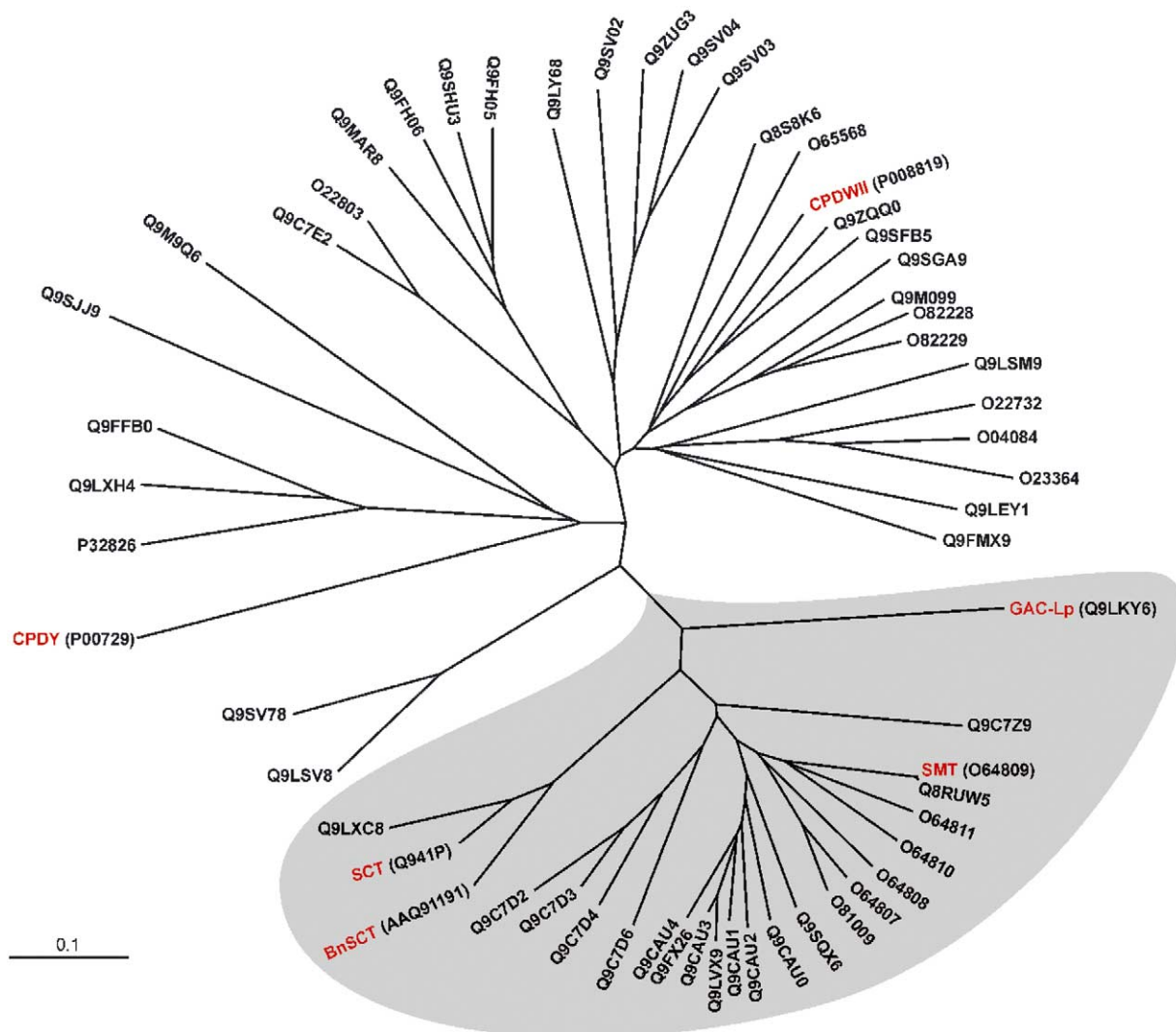


Fig. 3. Un-rooted similarity tree of Arabidopsis-SCPL proteins. The tree was constructed from an multiple alignment (CLUSTALW) of 53 SCPL proteins from Arabidopsis found in SwissProt database using the neighbour joining method of TREEVIEW program. Proven glucose ester-dependent acyltransferases from *B. napus* (BnSCT) and *L. penellii* (GAC-Lp) and carboxypeptidases from *S. cerevisiae* (CPDY) and *T. aestivum* (CPDWII) were included and indicated in red colour. A group of proteins clustered by similarity to acyltransferases is marked with grey shading. Proteins are denoted by SwissProt accession numbers.

Hydroxynitrile lyases (HNLs) catalyse the cleavage of cyanohydrins into aldehydes or ketones and hydrogen cyanide. Crystal structures reveal that the enzyme from *Manihot esculenta* (MeHNL) uses a catalytic triad of Ser, His and Asp to facilitate substrate decomposition (Lauble et al., 2001) whereas the enzyme from *Sorghum bicolor* (SbHNL) should apply another mechanism. Although sequence analysis detected a potential catalytic triad (Ser, His, Asp) in SbHNL, crystal structure showed that a two amino acids deletion immediately downstream from the seryl residue causes a geometry that prevents interaction of these amino acid residues as charge relay system (Lauble et al., 2002). The authors suggest an alternative reaction mechanism involving general base catalysis by the carboxy-terminal Trp

carboxyl group and proton transfer toward the leaving nitrile group by an active site water molecule.

Experimental data reveal that the α/β hydrolase fold structure harbours the potential for catalysing transfer reactions as well. It has been shown that yeast carboxypeptidase Y, a typical SCP, catalyses transesterification and transesterification reactions under nonphysiological alkaline conditions (Widmer and Johansen, 1979; Widmer et al., 1980). From nonribosomal peptide synthetases (NRPS), thioesterase (TE) domains are known which do not act as hydrolases but catalyse regio- and stereospecific macrolactonisation or macrolactamisation (Keating et al., 2001). SrfTE, the TE domain of surfactin synthetase from *Bacillus subtilis* catalyses macrolactonisation and belongs to the α/β hydrolase fold

enzymes. Structure elucidation gave first evidence for elements designed to favour lactonisation over hydrolysis (Bruner et al., 2002). SrfTE contains a helical lid for shielding the active centre which was found to be adjacent to a particularly hydrophobic region. An unusual distortion of the oxyanion hole is thought to play a specific role in promoting ring closure over hydrolytic deacylation (M.T. Stubbs, personal communication).

The catalytic triad composed of nucleophile, acid and histidine is not restricted to the family of α/β hydrolase fold enzymes. It has been found in three other groups of enzymes: the serine proteases, subtilisin and cysteine proteases (Ollis et al., 1992; Dodson and Wlodawer, 1998). There is no global similarity between members of these enzyme groups revealing that the catalytic triad has been developed independently in each of these groups as a result of convergent evolution. Thus, the catalytic triad seems to represent a very efficient and robust configuration for accomplishing hydrolytic and related functions. This would qualify it to be recruited for related enzymatic activities like acyltransfer.

In higher plants, SCPL proteins form a large and diverse family. In contrast, the genomes of bacteria, fungi and lower plants known so far contain only very few genes encoding this type of proteins. Thus, the remarkable spread out of SCPL proteins should be a typical feature of higher plants and one could hypothesise that this spreading has paralleled the evolution of plant secondary metabolism providing a source of robust enzymes to be recruited for functions in diverse pathways. The ongoing efforts in sequencing of whole genomes will produce the necessary data for proving this hypothesis in the near future. The proteom of *Arabidopsis* harbours 53 SCPL proteins according to SwissProt database. An unrooted similarity tree (Fig. 3) shows an interesting clustering. Twenty one proteins including the glucose ester-dependent acyltransferases SMT and SCT form a distinct group. It is remarkable that SCPL acyltransferases functionally proven from other plants so far (BnSCT from rape and GAC-Lp from wild tomato) cluster to this group, whereas SCPs (e.g. yeast and wheat) fall into other groups. This indicates at sequence level that the large family of plant SCPL proteins harbours a substantial fraction of related enzymes with different activities. Members of the Brassicaceae, such as *Arabidopsis*, rape or radish, accumulate complex patterns of sinapate esters and it is tempting to assume that these plants express corresponding patterns of SMT- and SCT-related SCPL acyltransferases.

5. Conclusions

Diversity, illustrated by a tremendous number of plant secondary metabolites (estimated over 200,000) such as phenolics, alkaloids, terpenoids or polyketides,

is a characteristic feature of plant metabolism. Accordingly, an appropriate number of specialised enzymes must have been developed during plant evolution to facilitate the formation of secondary metabolites. Recent findings have shed light on an interesting class of plant secondary enzymes catalysing β -acetal ester-dependent acyltransferase reactions. Molecular and biochemical features give rise to the assumption that these acyltransferases have been recruited from hydrolases of the serine carboxypeptidase type. Future studies will focus on questions on the molecular basis of changes in enzyme activities from hydrolysis to acyltransfer and on the time when SCPL acyltransferases appeared in evolution.

In order to define consensus sequence motifs which make up the difference between SCPs and SCPL acyltransferases, it will be helpful to acquire more sequence information about the latter ones. A useful approach could include enzyme purification or homology-based cloning from plants which have been shown to produce metabolites via the glucose ester-dependent acyltransferase reactions. Possible targets for such a strategy would be enzymes catalysing the synthesis of chlorogenic acid in sweet potato (*I. batatas*) (Villegas and Kojima, 1986), gallotannins in oak (*Quercus robur*) (Gross, 1983), IAA-inositol in maize (*Zea mays*) (Michalczyk and Bandurski, 1980, 1982), 1,2-disinapoylglucose in radish (*Raphanus sativus*) (Dahlbender and Strack, 1986), 2-*O*-acetyl-3-*O*-(*p*-coumaroyl)-*meso*-tartrate in spinach (*Spinacia oleracea*) (Strack et al., 1987), caffeoylglucarate in red cestrum (*Cestrum elegans*) (Strack et al., 1988), betanidin hydroxycinnamoylglycosides in members of the plant order Caryophyllales (Bokern et al., 1992), and a cyanidin sinapoylglycoside in wild carrot (*Daucus carota*) (Gläbgen and Seitz, 1992). Recently, the purification of 1-*O*-(indole-3-acetyl)- β -D-glucose:*myo*-inositol indoleacetyl transferase (IA-*myo*-inositol synthase) from *Zea mays* has been reported (Kowalczyk et al., 2003). Based on sequence information of a short peptide fragment, the authors claim that IA-*myo*-inositol synthase could belong to the family of SCPL acyltransferases. A laborious but rewarding strategy would use the *Arabidopsis* model. A systematic screen for mutants defective in genes encoding SCPL enzymes combined with a thorough biochemical phenotypic characterisation may provide insight into unpredictable functions of at least some of these proteins. Thus, SCPL acyltransferases provide targets of studying questions on the molecular evolution of plant secondary metabolism enzymes in focus of prime scientific interest.

Acknowledgements

Research of SCPL proteins in Halle is supported by the DFG priority programme 1152, "Evolution of

Metabolic Diversity". We thank T. Hartmann (Braunschweig) for comments on the manuscript.

References

- Baulcombe, D.C., Barker, R.F., Jarvis, M.G., 1987. A gibberellin responsive wheat gene has homology to yeast carboxypeptidase Y. *J. Biol. Chem.* 262, 13726–13735.
- Bech, L.M., Breddam, K., 1989. Inactivation of carboxypeptidase Y by mutational removal of the putative essential histidyl residue. *J. Biol. Chem.* 262, 13726–13735.
- Bellizzi III, J.J., Widom, J., Kemp, C., Lu, J.Y., Das, A.K., Hofmann, S.L., Clardy, J., 2000. The crystal structure of palmitoyl protein thioesterase 1 and the molecular basis of infantile neuronal ceroid lipofuscinosis. *PNAS* 97, 4573–4578.
- Bokern, M., Heuer, S., Strack, D., 1992. Hydroxycinnamic acid transferase in the biosynthesis of acylated betacyanins—purification and characterization from cell cultures of *Chenopodium rubrum* and occurrence in some other members of the Caryophyllales. *Bot. Acta* 105, 146–151.
- Bourne, P.C., Isupov, M.N., Littlechild, J.A., 2000. The atomic-resolution structure of a novel bacterial esterase. *Structure Fold. Des.* 8, 143–151.
- Bränden, C.-I., 1986. In Fletterick, R., Zoller, M. (Eds). *Computer Graphics and Molecular Modelling*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 45–51.
- Breddam, K., Sørensen, S.B., Svendsen, I., 1987. Primary structure and enzymatic properties of carboxypeptidase II from wheat bran. *Carlsberg Res. Commun.* 52, 297–311.
- Bruner, S.D., Weber, T., Kohli, R.M., Schwarzer, D., Marahiel, M.A., Walsh, C.T., Stubbs, M.T., 2002. Structural basis for the cyclization of the lipopeptide antibiotic surfactin by the thioesterase domain SrfTE. *Structure* 10, 301–310.
- Dahlbender, B., Strack, D., 1986. Purification and properties of 1-(hydroxycinnamoyl)-glucose:1-(hydroxycinnamoyl)glucose hydroxycinnamoyltransferase. *Phytochemistry* 25, 1043–1046.
- Degan, F.D., Rocher, A., Cameron-Mills, V., Von Wettstein, D., 1994. The expression of serine carboxypeptidases during maturation and germination of the barley grain. *Proc. Natl. Acad. Sci. U.S.A.* 91, 8209–8213.
- Devedjiev, Y., Dauter, Z., Kuznetsov, S.R., Jones, T.L., Derewenda, Z.S., 2000. Crystal structure of the human acyl protein thioesterase I from a single X-ray data set to 1.5 Å. *Structure Fold. Des.* 8, 1137–1146.
- Doan, N.P., Fincher, G.B., 1988. The A- and B-chains of carboxypeptidase I from germinated barley originate from a single precursor polypeptide. *J. Biol. Chem.* 263, 11106–11110.
- Dodson, G., Wlodawer, A., 1998. Catalytic triads and their relatives. *Trends Biochem. Sci.* 23, 347–352.
- Endrizzi, J.A., Breddam, K., Remington, S.J., 1994. 2.8-Å structure of yeast serine carboxypeptidase. *Biochemistry* 33, 11106–11120.
- Fleming, S.M., Robertson, T.A., Langley, G.J., Bugg, T.D.H., 2000. Catalytic mechanism of a C-C hydrolase enzyme: evidence for a gem-diol intermediate, not an acyl enzyme. *Biochemistry* 39, 1522–1531.
- Fulop, V., Bocskei, Z., Polgar, L., 1998. Prolyl oligopeptidase: an unusual beta-propeller domain regulates proteolysis. *Cell* 94, 161–170.
- Gierl, A., Frey, M., 2001. Evolution of benzoxazinone biosynthesis and indole production in maize. *Planta* 213, 493–498.
- Gläbgen, W.E., Seitz, H.U., 1992. Acylation of anthocyanins with hydroxycinnamic acids by protein preparations from cell cultures of *Daucus carota* L. *Planta* 186, 582–585.
- Gräwe, W., Bachhuber, P., Mock, H.-P., Strack, D., 1992. Purification and characterization of sinapoylglucose:malate sinapoyltransferase from *Raphanus sativus* L. *Planta* 187, 236–241.
- Gross, G.G., 1983. Synthesis of mono-, di- and trigalloyl-β-D-glucose by β-glucogallin-dependent galloyltransferases from oak leaves. *Z. Naturforsch.* 38c, 519–523.
- Hause, B., Meyer, K., Viitanen, P.V., Chapple, C., Strack, D., 2002. Immunolocalization of 1-O-sinapoylglucose:malate sinapoyltransferase in *Arabidopsis thaliana*. *Planta* 215, 26–32.
- Hayashi, R., Moore, S., Stein, W.H., 1973. Serine as the active center of yeast carboxypeptidase. *J. Biol. Chem.* 248, 8366–8369.
- Hayashi, R., Bai, Y., Hata, T., 1975. Evidence for an essential histidine in carboxypeptidase Y. Reaction with the chloromethyl ketone derivative of benzyloxycarbonyl-L-phenylalanine. *J. Biol. Chem.* 250, 5221–5226.
- Keating, T.A., Ehmann, D.E., Kohli, R.M., Marshall, C.G., Trauger, J.W., Walsh, C.T., 2001. Chain termination Steps in nonribosomal peptide synthetase assembly lines: directed acyl-S-enzyme breakdown in antibiotic and siderophore biosynthesis. *ChemBioChem* 2, 99–107.
- Kowalczyk, S., Jakubowska, A., Zielinska, E., Bandurski, R.S., 2003. Bifunctional indole-3-acetyl transferase catalyses synthesis and hydrolysis of indole-3-acetyl-myoinositol in immature endosperm of *Zea mays*. *Physiol. Plantarum* 119, 165–174.
- Lang, D.A., Mannesse, M.L., de Haas, G.H., Verheij, H.M., Dijkstra, B.W., 1998. Structural basis of the chiral selectivity of *Pseudomonas cepacia* lipase. *Eur. J. Biochem.* 254, 333–340.
- Lauble, H., Miehlisch, B., Förster, S., Wajant, H., Effenberger, F., 2001. Mechanistic aspects of cyanogenesis from active-site mutant Ser80Ala of hydroxynitrile lyase from *Manihot esculenta* in complex with acetonecyanohydrin. *Protein Sci.* 10, 1015–1022.
- Lauble, H., Miehlisch, B., Förster, S., Wajant, H., Effenberger, F., 2002. Crystal structure of hydroxynitrile lyase from *Sorghum bicolor* in complex with the inhibitor benzoic acid: a novel cyanogenic enzyme. *Biochemistry* 41, 12043–12050.
- Lawson, D.M., Derewenda, U., Serre, L., Ferri, S., Szittner, R., Wie, Y., Meighen, E.A., Derewenda, Z.S., 1994. Structure of a myristoyl-ACP-specific thioesterase from *Vibrio harveyi*. *Biochemistry* 33, 9382–9388.
- Lehfeldt, C., Shirley, A.M., Meyer, K., Ruegger, M.O., Cusumano, J.C., Viitanen, P.V., Strack, D., Chapple, C., 2000. Cloning of the SNG1 gene of *Arabidopsis* reveals a role for a serine carboxypeptidase-like protein as an acyltransferase in secondary metabolism. *Plant Cell* 12, 1295–1306.
- Leznicki, A.J., Bandurski, R.S., 1988. Enzymic synthesis of indole-3-acetyl-1-O-β-D-glucose. I. Partial purification and characterization of the enzyme from *Zea mays*. *Plant Physiol.* 88, 1481–1485.
- Li, A.X., Steffens, J.C., 2000. An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. *Proc. Natl. Acad. Sci. U.S.A.* 97, 6902–6907.
- Li, A.X., Eannetta, N., Ghangas, G.S., Steffens, J.C., 1999. Glucose polyester biosynthesis. Purification and characterization of a glucose acyltransferase. *Plant Physiol.* 121, 453–460.
- Li, J., Derewenda, U., Dauter, Z., Smith, S., Derewenda, Z.S., 2000. Crystal structure of the *Escherichia coli* thioesterase II, a homologue of the human Nef binding enzyme. *Nat. Struct. Biol.* 7, 555–559.
- Liao, D.-I., Remington, S.J., 1990. Structure of wheat serine carboxypeptidase II at a 3.5-Å resolution. *J. Biol. Chem.* 265, 6528–6531.
- Liao, D.-I., Breddam, K., Sweet, R.M., Bullock, T., Remington, S.J., 1992. Refined atomic model of wheat serine carboxypeptidase II at 2.2-Å resolution. *Biochemistry* 31, 9796–9812.
- Michalczyk, L., Bandurski, R.S., 1980. UDP-glucose:indoleacetic acid glucosyl transferase and indoleacetyl-glucose:myoinositol indoleacetyl transferase. *Biochem. Biophys. Res. Commun.* 93, 588–592.
- Michalczyk, L., Bandurski, R.S., 1982. Enzymic synthesis of 1-O-indol-3-ylacetyl-β-D-glucose and indol-3-ylacetyl-myoinositol. *Biochem. J.* 207, 273–281.

- Milkowski, C., Baumert, A., Schmidt, D., Nehlin, L., Strack, D., 2004. Molecular regulation of sinapate ester metabolism in *Brassica napus*: expression of genes, properties of the encoded proteins and correlation of enzyme activities with metabolite accumulation. *Plant J.*, in press.
- Mock, H.-P., Strack, D., 1993. Energetics of the uridine 5'-diphosphoglucose: hydroxycinnamic acid acyl-glucosyltransferase reaction. *Phytochemistry* 32, 575–579.
- Mortensen, U.H., Remington, S.J., Breddam, K., 1994. Site-directed mutagenesis on (serine) carboxypeptidase Y. A hydrogen bond network stabilizes the transition state by interaction with the C-terminal carboxylate group of the substrate. *Biochemistry* 33, 508–517.
- Nardini, M., Lang, D.A., Liebeton, K., Jaeger, K.E., Dijkstra, B.W., 2000. Crystal structure of *Pseudomonas aeruginosa* lipase in the open conformation. The prototype for family I.1 of bacterial lipases. *J. Biol. Chem.* 275, 31219–31225.
- Ober, D., Hartmann, T., 1999. Homospermidine synthase, the first pathway-specific enzyme of pyrrolizidine alkaloid biosynthesis, evolved from deoxyhypusine synthase. *Proc. Natl. Acad. Sci. U.S.A.* 96, 14777–14782.
- Ober, D., Hartmann, T., 2000. Phylogenetic origin of a secondary pathway: the case of pyrrolizidine alkaloids. *Plant Mol. Biol.* 44, 445–450.
- Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., Sussman, J.L., Verschueren, K.H.G., Goldman, A., 1992. The α/β hydrolase fold. *Protein Eng.* 5, 197–211.
- Shirley, A.M., Chapple, C., 2003. Biochemical characterization of sinapoylglucose:choline sinapoyltransferase, a serine carboxypeptidase-like protein that functions as an acyltransferase in plant secondary metabolism. *J. Biol. Chem.* 278, 19870–19877.
- Shirley, A.M., McMichael, C.M., Chapple, C., 2001. The *sng2* mutant of *Arabidopsis* is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose:choline sinapoyltransferase. *Plant J.* 28, 83–94.
- Steffens, J.C., 2000. Acyltransferases in protease's clothing. *Plant Cell* 12, 1253–1255.
- Strack, D., Mock, H.P., 1993. Hydroxycinnamic acids and lignins. In: Dey, P.M., Harborne, J.B. (Eds.), *Methods in Plant Biochemistry* Vol. 9, Lea, P.J., (Ed.), *Enzymes in Secondary Metabolism*. London: Academic Press, pp. 45–97.
- Strack, D., Heilemann, J., Boehnert, B., Grotjahn, L., Wray, V., 1987. Accumulation of 2-O-acetyl-3-O-(*p*-coumaroyl)-*meso*-tartaric acid. *Phytochemistry* 26, 107–111.
- Strack, D., Gross, W., Heilemann, J., Keller, H., Ohm, S., 1988. Enzymic synthesis of hydroxycinnamic acid esters of glucaric acid and hydroaromatic acids from the respective 1-O-hydroxycinnamoylglucoside and hydroxycinnamoyl-coenzyme A thioester as acyldonors with protein preparations from *Cestrum elegans* leaves. *Z. Naturforsch.* 43c, 32–36.
- Villegas, R.J.A., Kojima, M., 1986. Purification and characterization of hydroxycinnamoyl-D-glucose. Quinate hydroxycinnamoyl transferase in the root of sweet potato, *Ipomoea batatas* Lam. *J. Biol. Chem.* 261, 8729–8733.
- Vogt, T., Aebersold, R., Ellis, B., 1993. Purification and characterization of sinapine synthase from seeds of *Brassica napus*. *Arch. Biochem. Biophys.* 300, 622–628.
- Wajant, H., Mundry, K.-W., Pfizenmaier, K., 1994. Molecular cloning of hydroxynitrile lyase from *Sorghum bicolor* (L.). Homologies to serine carboxypeptidases. *Plant Mol. Biol.* 26, 735–746.
- Widmer, F., Breddam, K., Johansen, J.T., 1980. Carboxypeptidase Y catalysed peptide synthesis using amino acid alkyl esters as amine components. *Carlsberg Res. Commun.* 45, 453–463.
- Widmer, F., Johansen, J.T., 1979. Enzymatic peptide synthesis: carboxypeptidase catalyzed formation of peptide bonds. *Carlsberg Res. Commun.* 44, 37–46.
- Winkler, F.K., D'Arcy, A., Hunziker, W., 1990. Structure of human pancreatic lipase. *Nature* 343, 771–774.
- Zuber, H., Matile, P.H., 1968. Acid carboxypeptidases: Their occurrence in plants, intracellular distribution and possible function. *Z. Naturforsch.* 23b, 663–665.