Multifaceted Regulations of Gateway Enzyme Phenylalanine Ammonia-Lyase in the Biosynthesis of Phenylpropanoids

Xuebin Zhang and Chang-Jun Liu

Biological, Environmental and Climate Sciences Department
Brookhaven National Laboratory
Upton, NY 11973, USA

Submitted to Molecular Plant
October 2014
This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or any third party’s use or the results of such use of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof or its contractors or subcontractors. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
Multifaceted regulations of gateway enzyme phenylalanine ammonia-lyase in the biosynthesis of phenylpropanoids

Xuebin Zhang and Chang-Jun Liu*
Biological, Environmental & Climate Sciences Department, Brookhaven National Laboratory, Upton, NY 11973, USA.

*Corresponding Author: Chang-Jun Liu, email cliu@bnl.gov, phone: 631-344-2966
Fax: 631-344-3407,

Running title: Regulation of phenylalanine ammonia-lyase in plants

Summary:
Phenylalanine ammonia-lyase (PAL) is the first committed enzyme in the phenylpropanoid biosynthesis pathway. It is controlled under a multitude of regulatory mechanisms. In this review, a current overview on our understanding of the complicated regulatory mechanisms governing PAL's activity is presented.
Abstract

Phenylpropanoid biosynthesis in plants engenders a vast variety of aromatic metabolites critically important for their growth, development, and environmental adaptation. Some of these aromatic compounds have high economic value. Phenylalanine ammonia-lyase (PAL) is the first committed enzyme in the pathway; it diverts the central flux of carbon from primary metabolism to the synthesis of myriad phenolics. Over the decades, many studies have shown that exquisite regulatory mechanisms at multiple levels control the transcription and the enzymatic activity of PALs. In this review, we present a current overview on our understanding of the complicated regulatory mechanisms governing PAL's activity; we particularly highlight recent progresses in unraveling its post-translational modifications, its metabolite feedback regulation, and its enzyme organization.

Key words: Phenylpropanoids; phenylalanine ammonia-lyase; metabolic regulation.
Introduction

Phenylpropanoid biosynthesis converts L-Phe into diverse aromatic compounds, collectively termed (poly)phenolics. More than 8000 aromatic metabolites have been identified in plants, and categorized into different subclasses, including the benenoids, coumarins, flavonoids /anthocyanins, stilbenes, hydroxycinnamates, lignans, and the macromolecule lignin (Vogt, 2010; Fraser and Chapple, 2011). This large family of aromatic compounds fulfills numerous physiological functions essential for plant growth, development, and plant-environmental interactions. In particular, lignin, a structural component deposited in the vasculature, imparts mechanic support and engenders a hydrophobic environment that is important for conducting water and nutrients (Whetten and Sederoff, 1995; Vanholme et al., 2010); furthermore, many soluble phenolics, synthesized in a species-specific manner, act as anti-pathogenic phytoalexins, antioxidants, or UV-absorbing compounds, protecting plants from biotic- and abiotic-stresses, or act as pigments attracting pollinators, or as signaling molecules mediating plant-microbe interactions (Dixon and Paiva, 1995). The tissue-specific formation and accumulation of phenylpropanoids probably represents one of the most important prerequisites that ultimately facilitated the colonization of land by early plants (Weng and Chapple, 2010).

The biosynthesis of phenylpropanoids entails a sequence of central enzyme-regulated reactions (termed the general phenylpropanoid pathway) from which branch pathways emanate toward different aromatic end-products. In particular, this metabolic pathway assimilates about 30–40% of the biospheric organic carbon, ending up as the predominant lignin polymer in the cell walls. The general phenylpropanoid pathway in plants is linked to the shikimate pathway that produces aromatic amino acids from central carbon metabolisms. Phenylalanine ammonia-lyase (PAL, EC 4.3.1.5) is the entry-point enzyme of the general phenylpropanoid pathway that channels L-Phe from the primary metabolic pool to the synthesis of trans-cinnamic acid (t-CA). The produced t-CA is then further transformed into many phenolic compounds (Figure 1) (Bate et al., 1994; Cochrane et al., 2004). PAL activity has been found in all the higher plants analyzed so far, and in some fungi and a few bacteria, but not in animals (Hodgins, 1971; Fritz et al., 1976; Xiang and Moore, 2005). Since its discovery in 1961 by Koukol and Conn (1961), PAL has been extensively studied. Its activity is stimulated in developmental programming, and by a variety
of environmental cues, including pathogenic attacks, tissue wounding, UV irradiation, exposure to heavy metals, low temperatures, and low levels of nitrogen, phosphate, or ions (Dixon and Paiva, 1995; Weisshaar and Jenkins, 1998). The on and off, up and down activity of PAL in developmental processes and in stress-responses mirror its sophisticated regulatory control.

In the last few decades, studies in biochemistry, enzymology, structural biology, and molecular genetics have suggested /unveiled multifaceted regulatory mechanisms that plant cells variously employ to control PAL's activity. The recognized control of PAL can occur through several mechanisms, viz., the product inhibition, transcriptional- and translational-regulation, post-translational inactivation and proteolysis, enzyme organization/subcellular compartmentation, and metabolite-feedback regulation. In this review, we present an overview of the molecular bases and mechanisms controlling PAL's activity. In particular, we highlight herein recent progress in unraveling the molecular basis for proteolytic regulation of this gateway enzyme.

**Enzyme properties and product inhibition**

PAL catalyzes the non-oxidative elimination of ammonia from L-phenylalanine to yield trans-cinnamic acid (Cochrane et al., 2004). This enzyme belongs to a large amino acid ammonia-lyase/aminomutase superfamily, wherein it is phylogenetically clustered to the identified histidine ammonia-lyases (HALs) and tyrosine ammonia-lyases (TALs) found in bacteria (Poppe and Retey, 2005; Xiang and Moore, 2005). PAL from dicotyledonous plants predominantly and efficiently deaminates L-Phe, while PAL from yeast and some monocots, e.g., maize, converts L-Phe and L-Tyr to their corresponding cinnamic acid and p-coumaric acid (Fritz et al., 1976; Rösler et al., 1997).

In all plants studied so far, PAL is encoded by a multi-gene family, thus presenting as multiple isoforms (Cramer et al., 1989; Wanner et al., 1995). PAL occurs as a tetrameric enzyme in vivo with molecular weight of around 275-330KD (Havir and Hanson, 1973; Zimmermann and Hahlbrock, 1975; Appert et al., 1994). Multiple tetrameric forms with slightly different molecular weights, isoelectric point values, and substrate binding affinities have been purified from suspension cell cultures of bean (*Phaseolus vulgaris* L.), suggesting PAL can form heterotetramers (Bolwell et al., 1985). When different tobacco PAL isogenes are co-expressed in *E.coli*, the recombinant proteins indeed can form heterotetramers (Reichert et al., 2009).
Structurally, PAL is a helix-containing protein and its tertiary- and quaternary-sizes are much larger than those of HALs; this is mainly due to the existence in PAL of two additional structural segments: A mobile N-terminal extension that probably anchors the enzyme and interacts with other cellular components; and a specific shielding-domain sited over the active center that probably controls enzyme activity by restricting the access of the substrate to a narrow channel (Figure 2) (Calabrese et al., 2004; Ritter and Schulz, 2004). PAL and other members of the ammonia-lyase/aminomutase superfamily contain a common co-factor, 4-methylidene-imidazol-5-one (MIO) in their active site for catalysis (Figure 2); it is formed by the autocatalytic cyclization and dehydration of an internal tri-peptide segment, viz., Ala-Ser-Gly (Rother et al., 2002; Alunni et al., 2003). MIO resides on top of the positive pole of three polar helices in the active site, and serves as an prosthetic electrophile, arguably to initiate the direct elimination of the α-amino group, or for the stereo-specific abstraction of the β-proton to form aryl acid (Figure 2) (Langer et al., 2001; Retey, 2003; Calabrese et al., 2004; Ritter and Schulz, 2004; Louie et al., 2006).

PAL displays product-inhibition property; its activity is inhibited by trans-cinnamic acid, the product of its catalyzed reaction (O’Neal and Keller, 1970; Appert et al., 1994). Although there are no detailed in vitro studies on the inhibition mechanism, trans-cinnamic acid might act as a competitive inhibitor (Camm and Towers, 1973; Sato et al., 1982). Its derivatives, such as p-coumarate, O-chlorocinnamate, and the related compounds p-hydroxybenzoate and 2-naphthoate, also can inhibit PAL activity (Sato et al., 1982). In the crystal structure complex of tyrosine ammonia-lyase from Rhodobacter sphaeroides, trans-cinnamate or p-coumarate was observed to bind closely into the active site, implicating the structural basis of ammonia-lyase for its product inhibition (Appert et al., 1994). PAL purportedly follows an ordered Uni Bi type reaction sequence, in which trans-cinnamate is released before ammonia from the active site (Williams and Hiroms, 1967). Presumably, its release is a necessary step for inducing conformational change in PAL, preparing it for the next catalytic cycle. Therefore, a high concentration of trans-cinnamate may impair the release of the enzymatic product, and compete for the binding of Phe.

Earlier studies show that the purified PAL enzyme from parsley, French bean, or cultures of alfalfa cells exhibits negative cooperativity; it has a high binding affinity (low $K_m$) for phenylalanine at low concentration of substrate, and a relatively low binding affinity (high $K_m$) at a high substrate concentration (Zimmermann and Hahlbrock, 1975; Bolwell et al., 1985; Jorrin
These observations suggest that the substrate itself exerts an inhibitory effect on the rate of product formation through negative co-operative regulation of the enzyme subunits (Zimmermann and Hahlbrock, 1975). However, later experiments revealed that heterologously expressed recombinant PAL isoform in *E.coli* displays typical Michaelis-Menten kinetics and displays no co-operative behavior (Appert et al., 1994; Appert et al., 2003). This mechanistic discrepancy between the native- and recombinant-enzymes was ascribed to the presence of hetero-tetrameric isoforms in the purified native enzymes, or the occurrence of post-translational modification that may not occur in the bacterial expression system. However, the recombinant heterotetramer, composed of tobacco PAL1 and PAL2, or PAL1 and PAL4, also does not exhibit negative cooperative kinetics (Reichert et al., 2009). This suggests that the cooperative regulation of PAL activity observed in the earlier studies was either an artifact or caused by unknown factors rather than via the heteromerization of kinetically different isoforms.

**Transcriptional regulation of the expression of PAL genes**

PAL long has been known to be regulated at the transcriptional level in response to environmental stimuli, and to developmental cues, e.g., the demand for lignin in specific tissues (Edwards et al., 1985; Liang et al., 1989; Dixon and Paiva, 1995; Anterola et al., 2002; Pawlak-Sprada et al., 2011). In plants, PAL is encoded by a multi-gene family, ranging from a few members in many species, to more than a dozen copies in potato and tomato (Chang et al., 2008). Specific PAL isogenes often exhibit a unique but overlapping spatial- and temporal-expression pattern, suggesting that PAL isoforms might have distinct but redundant functions in plant growth, development, and in plant-environmental interactions. For instance, in *Populus* spp., two PAL cDNAs are expressed differentially in developing shoot- and root-tissues. *PtPAL1* is mainly expressed in the cells of non-lignifying stems, leaves, and roots wherein cells accumulate a high level of condensed tannins, whereas *PtPAL2* is expressed both in the heavily lignified structural cells of shoots, and in the non-lignifying cells of root tips, highlighting the potential spatial association of PAL isoforms with specific metabolic activities (Kao et al., 2002). In Arabidopsis, four PAL isogenes are expressed in inflorescence stems, a tissue rich in lignifying cells; *AtPAL-1*, -2 and -4 have much higher expression levels than *AtPAL-3* (Wanner et al., 1995; Raes et al., 2003; Rohde et al., 2004). Genetic evidence from studies of mutant lines deficient in *AtPAL* isogenes has established the close association of *AtPAL1*, 2, and 4 with lignin biosynthesis (Oh
et al., 2003; Rohde et al., 2004; Huang et al., 2010). Meanwhile, it is found that AtPAL1 and PAL2 also dominate flavonoid synthesis; the pal1 pal2 double knock-out mutant is almost devoid of flavonoids, although PAL activity reaches 50% of the level of the wild type due to the activity of other isoforms (Rohde et al., 2004). Nitrogen depletion particularly induces the expression of PAL1 and PAL2; together with the presumable induction of the downstream phenylpropanoid-flavonoid biosynthetic genes, it increases the accumulation of flavonoids (kaempferol and quercetin), anthocyanins, and sinapic acid (Olsen et al., 2008). These data indicate that AtPAL isogenes differentially respond to environmental stresses, and that AtPAL1 and AtPAL2 have functional specialization in environmentally-triggered phenolic synthesis.

The developmental- and inducible-expression of PALs is regulated by the functioning of their promoter activity. Both AtPAL1 and AtPAL2 promoter-GUS activities are readily detected throughout plant’s lifetime, from the development of the early seedling to fully grown adult plants where they are strongly expressed in the vascular tissues of roots, leaves, and inflorescence stems but are not active in young and fast-growing tissues, like the root tip or the shoot apical meristem (Ohl et al., 1990; Wong et al., 2012).

Mapping the promoter region of AtPAL1 via various deletions reveals that the 1.8 kb promoter possesses a possible negative element between -1832 to -816, where the deletion increases the promoter activity by 30%; two potential positive elements are located thereafter between -816 to -290, where the deletion severely lowers the promoter’s activity to about 2% of that of the full-length one (Ohl et al., 1990). Interestingly, this study revealed that the proximal region of the promoter to location-290 sustains the tissue- and organ-specific expression pattern as well as a full length promoter does; furthermore, the proximal region to -540 is responsive to environmental stimuli (Ohl et al., 1990). These discoveries can help to further identify the particular transcription factors that interact with these promoter regions for controlling PAL’s tissue-specific and stimuli-inducible expression.

The dissected promoter regions typically contain multiple cis-regulatory elements (Yang et al., 2001; Rohde et al., 2004). In particular, the AC-rich motifs (AC-I, ACCTACC; AC-II, ACCAACC; and AC-III, ACCTAAC), the target binding sites of the MYB transcription factors, commonly occur in the promoters of PAL isogenes from different species. These AC cis-elements also exist in the promoters of many other phenylpropanoid biosynthetic genes, including the chalcone synthase (CHS) gene, and most of the monolignol biosynthetic genes,
such as 4CL (Hatton et al., 1995; Hatton et al., 1996; Lois et al., 1989; Zhong and Ye, 2009; Zhao and Dixon, 2011), highlighting a key role of such MYB binding-elements in coordinating the expression of a set of phenylpropanoid biosynthetic genes for the tissue-specific synthesis of phenolics (Leyva et al., 1992).

A few R2R3 MYB transcription factors are known to be able to trans-activate PAL promoters to control PALs' tissue-specific expression (Martin and Paz-Ares, 1997). AmMYB305, which is expressed in the carpels of young snapdragon flowers (Antirrhinum majus) activates the PAL promoter when they are co-expressed in tobacco protoplasts (Sablowski et al., 1994). In the anthers of tobacco, a specific transcription-factor, NtMYBAS1/2, can bind to the MYB binding-site and activate NtPAL1 expression, thus positively regulating the expression of the NtPAL1 and phenylpropanoid synthesis in sporophytic tissues, such as the tapetum, stomium, and the vascular tissues of the young tobacco anthers (Yang et al., 2001). In Arabidopsis, a hierarchical regulatory network controls the formation of secondary cell walls, in which three MYB transcription factors, MYB58, MYB63, and MYB85, are specifically involved in regulating lignin biosynthesis (Zhou et al., 2009). Mutation of the AC elements abolishes the binding of MYB58. It is unknown whether, and how, these transcription factors can directly target on the AtPAL promoter and control the temporal- and spatial-expression of AtPALS in coordinating the distribution of carbon flux to different cell-wall polymers. In addition to the MYB transcription factors, the LIM domain-containing proteins that are composed of a double, semi-conservative zinc finger motifs have also been shown to bind on the AC rich motif, the Pal-box [CCA(C/A)(A/T)A(C/A)(C/T)CC], and positively regulate the transcription levels of NtPAL and other phenylpropanoid genes (Kawaoka et al., 2000; Kawaoka and Ebinuma, 2001).

Furthermore, the KNOX family of transcription factors that function in maintaining cells in an indeterminate state (Tsiantis, 2001) also affect PAL expression. Mutations in the KNOX gene BREVIPEDICELLUS (BP) up-regulate the expressions of Arabidopsis PAL1, 4CL1, C4H, and CAD1, and cause premature lignification of the vasculature with phenotype of shorter internodes, downward pointing siliques, and an epidermal stripe of disorganized cells along the stem, all of which suggest that the KNOX transcription factor AtBP functions as a negative regulator for the expression of AtPAL and other lignin biosynthetic genes (Figure 1) (Mele et al., 2003). Again, it will be interesting to determine whether this family of transcription factors directly or indirectly
targets PAL and other phenylpropanoid biosynthetic genes, and how they coordinately control cell-wall lignification with other regulators in the network.

Post-translational modifications of PALs
Ubiquitination and proteolysis of PAL
Under biotic and abiotic stresses, PAL activity is induced as a defense-response mechanism, concomitant with the onset of the biosynthesis of different phenolics (Dixon and Paiva, 1995). As demonstrated for the PAL of French bean, its induction essentially reflects de novo enzyme synthesis after the up-regulation of mRNA levels (Edwards et al., 1985). However, after the peak of the activity, PAL rapidly declines, engendering a transient induction pattern. This behavior suggests the imposition of a post-translational regulation (Lamb et al., 1979; Jones, 1984). Early studies proposed an inactivation mechanism that reversibly converts PAL into its inactive form: The elicitation triggers the synthesis of a putative proteinaceous inhibitor that can react specifically with the PAL enzyme, thus engendering an inactive enzyme-inhibitor complex (Attridge and Smith, 1974; French and Smith, 1975). Purportedly, a high molecular-weight protein fraction from sunflower leaves possesses a PAL-inactivation system (Creasy, 1976). However, this has not been confirmed nor characterized so far. In examining the fluctuation of PAL activity in the injured root tissue of sweet potato, Tanaka and Uritani (1977) reported that following the de novo synthesis of PAL, its subsequent decline in activity is likely reflects the proteolytic degradation of the existing enzyme; however, its molecular basis has not been resolved until recently.

In eukaryotic systems, the ubiquitin-26S proteasome system dominates the selective protein degradation (Vierstra, 2003). To initiate its breakdown, the protein first is modified by the covalent attachment of the ubiquitin, a 76 amino-acid protein, to its lysine residues. Then, the protein is recognized and degraded by the 26S proteasome (Smalle and Vierstra, 2004). Using ubiquitin-affinity and nickel-chelate affinity chromatography coupled with mass spectrometry, an array of ubiquitinated proteins from Arabidopsis was identified (Saracco et al., 2009; Kim et al., 2013a). Among them, several phenylpropanoid-monolignol biosynthetic enzymes, including AtPAL-1, -2, and -3, were found to be potentially ubiquitinated (Kim et al., 2013). The (poly)ubiquitination of four Arabidopsis PAL isoforms was confirmed in our recently
immunoblot analysis of transiently expressed AtPAL proteins against the anti-ubiquitin antibody (Zhang et al., 2013).

Ubiquitination of a target protein requires the sequential action of three enzymes: The ubiquitin activating enzyme (E1), ubiquitin conjugating enzymes (E2), and ubiquitin protein ligase (E3) (del Pozo and Estelle, 2000; Lechner et al., 2006). When Arabidopsis PAL1 and PAL2 are used as the baits to screen an Arabidopsis cDNA expression library in a yeast two-hybrid assay, three interacting partners are captured, denoted kelch motif-containing F-box proteins, KFB0-1, -20, and -50. Subsequent pair-wise verification via Y2H, BiFC, and co-IP assays corroborates the physical interactions of all four AtPAL isozymes with these KFB proteins (Zhang et al., 2013). Furthermore, a close homolog, KFB39, also interacts with AtPALs (X.Z. and C.-J. L. unpublished data). This finding agrees well with the discovery in a genome-wide study of binary protein-protein interactions, that AtPAL1, 2, and/or 4 are found to potentially interact with KFB01 and KFB50 (Consortium, 2011).

The F-box proteins are the structural components of the canonical SCF type E3 protein-ubiquitin ligase, responsible for specifying the target proteins for ubiquitination and degradation (del Pozo and Estelle, 2000; Lechner et al., 2006). The KFB proteins are one subclass of F-box proteins. The physical interaction of these KFBs with PAL isozymes suggests that they might act as the key mediators specifying PALs' ubiquitination and subsequent degradation. Indeed, co-expressing any one of them with PAL isozyme in tobacco leaves impairs the latter’s stability and activity. Moreover, with treatment of MG132, a specific inhibitor of the 26S proteasome, the degradation of PAL protein is retarded or mitigated; meanwhile, the ubiquitination of PAL is enhanced. Furthermore, manipulating (either up- or down-regulating) KFB expression in Arabidopsis reciprocally affects PAL's cellular concentration and activity, and the consequent accumulation of phenolic esters, flavonoids, anthocyanins, and the lignin deposited in the cell walls of the stem (Zhang et al., 2013). These data denote that the identified KFBs are negative regulators governing PAL's stability via the ubiquitination-26S proteasome pathway.

E3 ligase-mediated ubiquitination and the subsequent protein degradation are critical components in the complex regulatory network underlying cellular processes (Vierstra, 2003). In most cases, the E3 ligase components themselves are regulated by the plant’s developmental cues and by environmental stresses, thus transducing developmental or environmental signals to the corresponding physiological processes (del Pozo and Estelle, 2000; Lechner et al., 2006).
Examining the expressions of the identified KFBs reveals that they differ in tissues and at different growth stages; for example, KFB20 is highly expressed in leaf and stem tissues, while KFB01 mainly is expressed in flowers, thus denoting their potential distinct roles in mediating PAL degradation and controlling a particular set of phenylpropanoid synthesis. Moreover, the expressions of four KFBs are affected by stimulators. Exposing Arabidopsis seedlings to a high source of carbon (4% sucrose in the medium) or to UVB irradiation substantially suppresses KFBs’ expression, and is accompanied by an increased accumulation of flavonoids/anthocyanin pigments (Zhang et al., 2013).

The tissue-specific-, temporal-, and inducible-expression of the identified KFB genes, and the reciprocal correlation of their expression with the cellular concentration of PAL and the accumulation of phenolics strongly suggest that KFB-mediated PAL ubiquitination and degradation function as a basic regulatory mechanism, by which plant cells finely gear carbon flux towards phenylpropanoid metabolites.

The close homologs of the identified Arabidopsis KFB genes can be recognized across a range of plants, from moss (Physcomitrella patens) to angiosperm- and gymnosperm-trees (e.g., Populus and Picea). All those species synthesize and accumulate different types of phenylpropanoids (Richardson et al., 2000; Morreel et al., 2004; Richter et al., 2012). Although the precise functions of those KFB homologs remain unknown, the evolution of those E3 ligase subunits in land plants, along with the emergence of phenolic metabolites, implies that the KFB-mediated selective turnover of PAL may be an universal regulatory mechanism adopted by a variety of terrestrial plants to control their phenylpropanoid synthesis.

Besides mediating PAL turnover, the same set of KFBs were found to physically interact with several type-B Arabidopsis response regulators (ARRs); accordingly, they are proposed to negatively regulate cytokinin responses (Kim et al., 2013b). Cytokinins are a group of plant hormones with vital roles in regulating cellular division, and modulating the activities of both the shoot- and root-meristems in building plant architectures (Sakakibara, 2006). Several studies suggested that disturbing the translocation, distribution, or signal perception of cytokinin in plants severely degrades their vascular formation and lignification (Werner et al., 2003; Zhang et al., 2014), inferring the existence of common links orchestrating the cytokinin signaling responses, xylogenesis, and lignin deposition. Further research will clarify whether the KFB proteins identified have potential roles in coordinating cytokinin signaling and phenylpropanoid-
lignin biosynthesis and/or deposition. It also would be interesting to systematically identify any other potential targets of those KFBs, and to explore whether the altered plant lignification in turn affects other KFB-mediated biological processes. Since the expression patterns of four KFB genes are different (Zhang et al., 2013), defining and delineating their potential different biological functions *in planta* would be valuable.

*Phosphorylation of PAL*

Phosphorylation is another major type of post-translational modification of proteins in eukaryotic cellular processes. Bolwell et al. (1992) reported their detection of PAL phosphorylation in suspension cultures of French bean cells. Their subsequent purification led to the identification of a 55 KD PAL kinase activity that phosphorylates the threonine/serine residues of the conserved polypeptides of the PALs of both bean and poplar; such PAL kinase displays the property of a calmodulin-like domain protein kinase (CDPK) (Bolwell, 1992; Allwood et al., 1999). In assaying eight protein kinases from Arabidopsis, a calcium-dependent protein kinase, AtCDPK, indeed could phosphorylate the short polypeptide of PAL and a poplar recombinant PAL enzyme (Cheng et al., 2001). The sequence of polypeptide of PAL used in the assay is largely conserved in homologous enzymes from other species; therefore, this finding implies that this threonine/serine type of phosphorylation of PAL might be a ubiquitous modification mechanism occurring in higher plants (Allwood et al., 1999). Structural determination of PAL from parsley suggests that the potential phosphorylation site of PAL might be located at the end of a flexible helix connecting the shielding domain to the core domain. This position displays the highest mobility, suggesting that phosphorylation might affect the shift of the shielding domain and change access to the active center of the enzyme (Ritter and Schulz, 2004). Additionally, PAL phosphorylation also was predicted to mark particular subunits for turnover, or to target them to the membrane for subcellular compartmentalization (Cheng et al., 2001). However, recent phospho-proteomic studies failed to confirm the phosphorylation of PAL and any other monolignol proteins from poplar (*P. trichocarpa*). In one study, 147 phosphoprotein groups were identified from differentiated xylem wherein monolignol biosynthetic enzymes are abundant. None of them are proteins known to participate in the monolignol biosynthetic pathway (Chen et al., 2013). In another work, Liu et al. (2011) detected 151 phosphoproteins in the dormant terminal buds of *P. simonii x P. nigra* using the same strategy of
phosphopeptide-enrichment; similarly, none are monolignol proteins (Liu et al., 2011). Further study may clarify whether the apparent disparity between the early and current studies reflects the limitations of the methodology used, and whether phosphorylation of PAL and/or other monolignol enzymes is a species-specific or a condition-specific event.

**Metabolite feedback regulation**

Apart from transcriptional regulation in response to environmental stimulus and nutrient balance, PAL and phenylpropanoid biosynthetic activity appear to be metabolically regulated by particular biosynthetic intermediates or chemical signals (Figure 1). *Trans*-cinnamic acid long has been proposed as a signal molecule, regulating flux into the pathway (Lamb, 1979; Bolwell et al., 1986). Both PAL activity and PAL gene transcription are negatively regulated by *trans*-cinnamic acid, either via the exogenous application of the compound or by blocking the downstream C4H activity for metabolizing cinnamic acid (Lamb, 1979; Bolwell et al., 1986; Mavandad et al., 1990; Blount et al., 2000). When PAL from French bean is overexpressed in tobacco wherein the C4H is co-suppressed, the transgenic plants present significantly lower PAL activity than those plants harboring the bean PAL alone, suggesting PAL activity might be feedback-regulated by the accumulated *trans*-cinnamate due to the reduction of C4H *in planta* (Blount et al., 2000). The mechanism by which PAL is feedback down-regulated may be complex. Excessive *trans*-cinnamate can directly inhibit the enzyme activity itself, since PAL exhibits a product-inhibition property, or the accumulated *trans*-cinnamate impairs PAL transcription (Lamb, 1979; Bolwell et al., 1986; Mavandad et al., 1990). In addition, our current study also suggests that feeding *trans*-cinnamate to Arabidopsis seedlings induces the expression of the identified KFB proteins that modulate PAL turnover and the pathway activity (X.Z and C.-J.L., unpublished data).

The feedback regulation is not only triggered by PAL’s immediate product, but also by the intermediates from branch pathways or exogenous chemical signals (Figure 1). Feeding caffeic acid to soybean (*Glycine max*) seedlings substantially inhibits PAL activity (Bubna et al., 2011). In the Arabidopsis double-mutant deficient in *AtUGT78D1* and *UGT78D2*, which compromises the 3-*O*-glycosylation of flavonols, AtPAL activity and the transcripts of both *AtPAL1* and *PAL2*, the two isogenes specified for flavonoid synthesis (Olsen et al., 2008), and *chalcone synthase* (*CHS*) are inhibited, so repressing flavonol synthesis itself. Moreover, upon blocking flavonol
synthesis either in the tt4chs or flavonol synthase 1 (fls1) mutant lines, the repression of PAL can be released completely, suggesting that flavonol aglycones might serve as signal molecule triggering PAL feedback regulation (Figure 1) (Yin et al., 2012). The ugt78d1 ugt78d2-dependent suppression of the flavonol branch and AtPAL1 and PAL2 activity does not affect the synthesis of other phenylpropanoids; in particular, lignin biosynthesis appears normal, corroborating the functional specification of AtPAL1 and PAL2 in flavonoid synthesis.

**Macromolecular organization of PAL with other phenylpropanoid enzymes**

Srere in 1985 proposed the term “metabolons”, referring to the super-complex formed by the interactions of sequential metabolic enzymes and their interaction with supporting cellular structural elements (Srere, 1985). Such organization of metabolic pathways is believed to increase the local concentrations of the enzymes and of their substrates, and to improve intermediate handing-over between consecutive enzymes. In addition, it also defines some specific domains in the spaces of cell to ensure the swift conversion of the labile- or toxic-intermediates, and avoids unnecessary competition with other metabolites present in the cell (Srere, 1987; 2000; Jorgensen et al., 2005). In eukaryotes, most of the cytochromes P450 are anchored via their N-terminus membrane- targeting peptides on the cytoplasmic surface of the endoplasmic reticulum (ER), with their main catalytic domains protruding on the membrane’s surface (Bayburt and Sligar, 2002). This unique character of the P450 enzymes provides perfect nucleation sites for metabolons to form, viz., a gathering of some of the soluble enzymes, so to increase metabolic efficiency. The phenylpropanoid pathway has long been proposed to have such macromolecular organization (Winkel-Shirley, 1999; Winkel, 2004; Jorgensen et al., 2005). By differential centrifugation, sedimentation of microsomes, or gradient separation of different membrane fractions, PAL activity from different plant species is found to associate with the ER membrane (Czichi and Kindl, 1975; 1977; Hrazdina and Wagner, 1985). The membrane fraction from potato tubers converts Phe more efficiently into p-coumaric acid than does exogenously supplied cinnamic acid. This finding engenders an interesting assumption, i.e., the potential association of PAL with the consecutive P450 enzyme, C4H, for metabolite channeling. Using isotope-labeled Phe and cinnamic acid, Rasmussen and Dixon (1999) showed that the exogenous application of cinnamate into suspension cultures of tobacco cells does not dilute/equilibrate with the endogenous pool of cinnamic acid produced from PAL, supporting the concept of metabolic
channeling between PAL and C4H. This proposed concept further was corroborated in
subsequent studies of the co-localization of tobacco PAL and C4H (Achnine et al., 2004).
Immunoblot analysis on the prepared microsomal fractions and immunofluorescence confocal-
microscopic observation reveal that the operationally soluble tobacco PAL1 itself partially
bounds to the ER membrane and co-localizes with NtC4H, whereas its isozyme, NtPAL2,
appears mostly as a cytosolic protein. However, when co-overexpressed with C4H, both
isoforms shift to the membrane fraction. These data suggest that membrane-bound C4H might
serve as the nucleation site for binding the PAL isoforms, and that NtPAL1 binds more strongly
to C4H than does NtPAL2 (Achnine et al., 2004).

The scope of metabolon sitting on the general phenylpropanoid pathway to the monolignol
biosynthetic branch was expanded recently. Using different biochemical approaches, Chen et al
(2011) demonstrated the interaction between two poplar C4H isoforms and their interactions
with C3H. Those isoforms form heterodimeric- (PtrC4H1/C4H2, PtrC4H1/C3H3, and
PtrC4H2/C3H3) and heterotrimeric- (PtrC4H1/C4H2/C3H3) membrane-protein complexes.
Moreover, the authors found that the assembly of those hydroxylase isoforms significantly
improves the enzyme kinetics with any of the complexes by some 70-6,500 times higher than
that of the individual proteins. Furthermore, the association of PtC4H and PtC3H likely creates
two potential hydroxylation pathways for monolignol synthesis: One through p-coumaric acid to
caffeic acid; and the other, through p-coumaroyl shikimic acid to caffeoyl shikimic acid.

By protein co-purification and fluorescence resonance-energy transfer study, Bassard et al
(2013) reported that Arabidopsis C4H and C3H, like their homologs in poplar, also interact with
each other to form homo- or heteromers and both are mobile within the ER. Further, AtHCT, the
first branch-point enzyme for monolignol synthesis and an operationally soluble protein, was
revealed to partially associate with the ER, and the expression of AtC3H in the cells promotes
HCT's ER association. Under AtC3H overexpression, another soluble enzyme, 4-coumarate:
CoA ligase (4CL), is also redistributed closer to the ER. Their data suggest that AtC4H, AtC3H,
and two functionally connecting soluble enzymes, At4CL and AtHCT, likely form a loosely
associated supra-molecular complex in the cells. Moreover, when plant cells are under wounding
stress that induces cell-wall lignification and the defensive production of phenolics, such protein
association is enhanced, suggesting the formation of metabolon appropriates for particular
metabolic needs (Bassard et al., 2013). Taken together, these experimental data substantiate the
concept that phenylpropanoid biosynthetic enzymes, including different PAL isozymes, are
nucleated by membrane-bound P450 proteins; in this manner they form enzyme complexes to
regulate pathway activity and metabolic complexity.

Conclusions and perspectives

The gateway enzyme, PAL, plays a key role in mediating carbon flux from primary
metabolism into the phenylpropanoid pathway. Over the decades, studies have shown that
exquisite regulatory mechanisms at multiple levels control the transcription and the enzymatic
activity of PALs, thereby coordinating the distribution of metabolic flux into a set of aromatic
natural products, and the abundant biopolymer, lignin.

At the transcriptional level, the PAL promoter contains characteristic cis-regulatory elements
that recruit particular transcription factors to regulate PAL developmental- and inducible-gene
expression. Recognizing PAL’s ubiquitination and the identifying a group of ketch-repeat F-box
proteins that mediate PAL proteolytic degradation has shed light on our understanding of the
post-translational regulation of phenylpropanoid biosynthesis and of the cross-talk of phenolic
metabolisms with other biological processes. Moreover, there is growing evidence pointing to
the organization of phenylpropanoid-pathway enzymes. More detailed studies are needed to gain
a comprehensive insight into the complexity of the molecular regulation of PAL activity, and the
related phenylpropanoid biosynthesis. For instance,

1) Although a few transcription factors are implicated in the control of the PAL gene
expression, a detailed characterization remains to be conducted to define their interactions with
cis-regulatory elements of PAL, and their regulatory roles in modulating specific
phenylpropanoid synthesis in response to development and environmental clues.

2) PAL isozymes form homo- or hetero- tetramer. Both Y2H- and BiFC-assays indicate that
the identified AtKFB proteins likely interact differentially with AtPAL isozymes. It will be
interesting to determine whether the homo- or hetero-tertramerization of PALs and their
differential interactions with KFB proteins have any significance in regulating the pathway
activity.

3) The identified AtKFB genes exhibit tissue-specific- and developmental expression-
patterns. They also respond differentially to environmental stimuli. The differential regulatory
roles of KFBs in balancing carbon- and nutrient-sources for building plant architecture, cell-wall lignification, and defense responses remain to be evaluated.

4) Besides ubiquitination, it remains unclear whether PALs undergo any other types of post-translational modifications; particularly, we need to re-examine whether phosphorylation modification occurs on PAL and/or other monolignol enzymes, and whether this type of modification is a species-specific-, or a condition-specific-event.

5) PAL is feedback-regulated by its own product (at least in the C4H- disruption plants), or by the intermediate of the downstream branch pathway. More studies of genetics, biochemistry, and structure biology are needed to detail the molecular mechanisms underlying such feedback regulation, and to elucidate whether the metabolites have a direct or an indirect effect on PAL activity.

Phenylpropanoid biosynthesis controls up to 40% of organic compounds circulating in the biosphere. They contribute to plants’ hardiness, color, taste, and odor. Many are of high economic value. Knowledge about the complicated regulatory mechanisms of the phenylpropanoid pathway, in particular the entry point enzyme PAL, will facilitate our sophisticated manipulation of this economically important metabolic pathway towards the desired agro-industrial and horticultural traits.

FUNDING
This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division under contract number DEAC0298CH10886 and the National Science Foundation through grant MCB-1051675 to CJL. No conflict of interest declared.

Figure legends:

Figure 1. Schema of phenylpropanoid biosynthesis pathway, highlighting molecular regulation on the gateway enzyme PAL, and the physical interaction of the early phenylpropanoid pathway enzymes. The data are summarized primarily from studies in Arabidopsis, poplar, and tobacco. The identified transcriptional regulation (MYB, LIM, KNOX transcription factors), kelch repeat containing F-box protein-mediated ubiquitylation and degradation, and metabolic feedback regulation on PAL are indicated, respectively, in green, red,
and blue. The dashed line indicates the negative regulation; the double arrow indicates multiple steps of the pathway. PAL: Phenylalanine ammonia-lyase, C4H: Cinnamate 4-hydroxylase, 4CL: 4-Coumarate-CoA ligase, HCT: Hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase, CHS: Chalcone synthase, FLS: Flavonol synthase, UGT78D1: Flavonol 3-O-rhamnosyltransferase, UGT78D1: flavonoid 3-O-glucosyltransferase. ARR: Arabidopsis response regulator (to cytokinins).

Figure 2. Comparison of the tertiary structures of PAL monomer from parsley (Petroselinum crispum) (A) and HAL from Pseudomonas putida (B). The ribbon representations of PcPAL and PpHAL are re-built based on the coordinates of PAL (1W27) and HAL (1B8F) (Ritter and Schulz, 2004) via the PyMol program. The polypeptide chains are colored green for the MIO domain, magenta for the helical bundle core- domain, and cyan for the inserted shielding domain (in PcPAL). The first 24 residues (the extension domain) in the crystal structure of PcPAL were missing (Ritter and Schulz, 2004). The MIO is shown in the ball-and-stick model.

References


Figure 1. Schema of phenylpropanoid biosynthesis pathway, highlighting the molecular regulation on the gateway enzyme PAL and the physical interaction of the early phenylpropanoid pathway enzymes. The data are summarized primarily from studies in Arabidopsis, poplar, and tobacco. The identified transcriptional regulation (MYB, LIM, KNOX transcription factors), kelch repeat containing F-box protein- mediated ubiquitylation and degradation, and metabolic feedback regulation on PAL are indicated in green, red and blue, respectively. The dashed line indicates the negative regulation; double arrows indicate multiple steps of the pathway. PAL: Phenylalanine ammonia-lyase, C4H: Cinnamate 4-hydroxylase, 4CL: 4-Coumarate-CoA ligase, HCT: Hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase, CHS: Chalcone synthase, FLS: Flavonol synthase, UGT78D1: Flavonol 3-O-rhamnosyltransferase, UGT78D1: flavonoid 3-O-glucosyltransferase, ARR: Arabidopsis response regulator (to cytokinins).
Figure 2. Comparison of the tertiary structures of PAL monomer from parsley (*Petroselinum crispum*) (A) and HAL from *Pseudomonas putida* (B). The ribbon representations of PcPAL and PpHAL are re-built based on the coordinates of PAL (1W27) and HAL (1B8F) (Ritter and Schulz, 2004) via PyMol program. The polypeptide chains are colored green for MIO domain, magenta for helical bundle core domain and cyan for the inserted shielding domain (in PcPAL). The first 24 residues (the extension domain) in crystal structure of PcPAL were missing (Ritter and Schulz, 2004). The MIO is shown in the ball-and-stick model.