# MOLECULAR CLONING AND CHARACTERIZATION OF CINNAMYL ALCOHOL DEHYDROGENASE cDNA FROM INTERSPECIFIC HYBRID *ACACIA MANGIUM* X *ACACIA AURICULIFORMIS*

**Pang, S.L., Choong, C.Y.* and Wickneswari, R.**
School of Environmental and Natural Resource Sciences
Faculty of Science and Technology, Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor, Malaysia
*E-mail: cychoong@pkrisc.cc.ukm.my

## ABSTRACT

Cinnamyl alcohol dehydrogenase (CAD; EC 1.1.1.195) is an indicator of lignin biosynthesis because of its specific role in the reduction of hydroxycinnamaldehydes to hydroxycinnamyl alcohols. This protein catalyzes the final step in a branch of phenylpropanoid pathway specific for production of lignin monomers. In this study, a full length cDNA of CAD (AhCAD) from the inner bark tissue of interspecific hybrid *Acacia mangium* x *Acacia auriculiformis* was obtained by rapid amplification of cDNA ends (RACE). The full length AhCAD was 1346 bp; containing a 1089 bp open reading frame (ORF), which encodes a polypeptide of 363 amino acids, with a 5' untranslated region of 65 bp and a 3' untranslated region of 192 bp. The deduced protein had a calculated molecular weight of 39.99 kDa and an isoelectric point of 5.9. The AhCAD sequence was compared with deduced polypeptide sequences of other isolated plant CAD enzymes. A conserved zinc-containing alcohol dehydrogenase motif in all plant CADs was found in the encoded AhCAD amino acid sequence. The encoded polypeptide exhibited sequence similarity to CADs from different plants, the highest identities being to CADs from *Medicago sativa* (69 %) and *Nicotiana tabacum* (67 %).

*Keywords*: Cinnamyl alcohol dehydrogenase (CAD), molecular cloning, cDNA full length, interspecific hybrid *Acacia mangium* x *Acacia auriculiformis*

## INTRODUCTION

Lignin is a plant phenolic biopolymer made up of three main *p*-hydroxycinammyl alcohol precursors or monolignols, namely *p*-coumaryl, coniferyl and sinapyl alcohols. It is mainly found in the vascular tissues, where it waterproofs the cell wall and enables transport of water and solutes through the vascular system (Hawkins et al. 1997). Lignin is crucial for structural integrity of the cell wall, stiffness and strength of the stem and provides defense against pathogen attack (Boudet et al. 1995). Monolignols are synthesized by enzymes involved in general phenylpropanoid pathway followed by enzymes that are specifically involved in lignin biosynthesis, namely cinnamoyl-CoA reductase and cinnamyl alcohol

dehydrogenase (CAD). Monolignols are then polymerized to form lignin on secondary cell walls, and the polymerization is thought to be catalyzed by peroxidases and/or laccases (Lewis & Yamamoto 1990; Whetten & Sederoff 1995).

Among the enzymes involved in lignin biosynthesis, CAD has recently become the focus of a number of studies on molecular basis of lignification. Several CAD proteins and genes have been isolated and characterized in different plants but their roles have not been clearly understood (Baucher et al. 1998). Until recently, CAD proteins were generally believed to catalyze the final step of monolignol biosynthesis, which is the reduction of the hydroxycinnamyl aldehydes to the corresponding hydroxycinnamyl alcohols. CAD has been purified from *Nicotiana tabacum* (Halpin et al. 1992), *Eucalyptus gunnii* (Goffner et al. 1992; Hawkins & Boudet 1994), *Pinus taeda* (O'Malley et al. 1992) and *Ardia cordata* (Hibino et al. 1993b). cDNA clones have also been isolated for CAD from *Nicotiana tabacum* (Knight et al. 1992), *Pinus taeda* (O'Malley et al. 1992), *Eucalyptus gunni* (Grima-Pettenati et al. 1993) and *Aralia cordata* (Hibino et al. 1993a).

Isoforms of CAD have been detected in angiosperms, which can display an enhanced substract specificity preference for sinapylaldehyde and coniferaldehyde (Wyrambik & Grisebach 1975; Sarni et al. 1984; Goffner et al. 1992; Grima-Pettenati et al. 1993). This has lent support to a model in which the last step in the biosynthesis of guaiacyl and syringyl monolignols in angiosperms is mediated by a broad spectrum CAD capable of reducing both coniferaldehyde and sinapaldehyde (Boudet et al. 1995; Whetten & Sederoff 1995). In contrast, there is so far only one gene encoding CAD that has been detected and purified from gymnosperm lignifying tissues. This CAD is coniferaldehyde specific with insignificant catalytic activity toward sinapaldehyde, consistent with the biosynthesis of mainly guaicyl lignin in these species (Lüderitz & Grisebach 1981; O'Malley et al. 1992; MacKay et al. 1995; Zinser et al.1998). However, CAD activity has also been detected in apparently non-lignified tissues, and monolignols are involved in biosynthesis of non-lignin products such as lignan (Lewis & Yamato 1990) and surface polymers such as suberin and cutin (Kolattukudy 1981).

Lignin is an undesired plant product during the processing of pulp and paper. In order to better understand the role of CAD in the lignin biosynthesis in interspecific hybrid *A. mangium* x *A. auriculiformis*, cloning and characterization of this gene has been carried out in this study. We report here the isolation and sequence analyses of AhCAD from the inner bark tissues of the hybrid *Acacia*.

## MATERIALS AND METHODS

### Plant materials

Inner bark tissues of an interspecific hybrid *A. mangium* x *A. auriculiformis* were obtained from Terrace W, Plant Biotechnology Laboratories, Universiti Kebangsaan Malaysia, Bangi, Malaysia. The tree was pollarded at 0.5 m height. The bark was

peeled and inner bark tissues were scraped into a clean plastic bag, dropped directly into liquid nitrogen and stored at 80°C until needed.

### RNA isolation

Total RNA was extracted using RNeasy Midi Kit (Qiagen, Germany), and mRNA was isolated from total RNA using Dynabeads oligo-$(dT)_{15}$ (Dynal, Norway) according to the manufacturer's instructions. RNA concentration was determined by UV spectrophotometry.

### 5'-RACE, PCR cloning and DNA sequencing

5'-RACE was performed by using *SMART™ RACE cDNA Amplification Kit* (Clontech, USA). cDNA was prepared according to the manufacturer's instructions. Primer used for 5'RACE was designed based on 3' ESTs for CAD from the cDNA library of *Acacia* hybrid (Wickneswari et al. 2004) by using *Primer3* (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3). The sequence of the primer was 5' GAC CAG CCT ATC ACC TGG CAC ACT G 3'. The PCR was performed according to the suggested touch-down PCR cycles in the manufacturer's manual. The fragment was purified and cloned into pGEMT-Easy cloning vector (Promega, USA). The cloned cDNA fragment was sequenced in both directions by the dideoxy chain termination method using BigDye terminator cycle sequencing kit with Ampli-Taq DNA polymerase FS (PE Applied Biosystems). The sequences were checked and edited by using Chromas Lite 1.0 (http://www.technelysium.com.au/chromas14x.html).

### Sequence analysis

BLASTP (Altschul et al. 1997) was used to identify homologues of the AhCAD protein by comparing the nucleotide sequences translated into open reading frames against protein databases (http://www.ncbi.nih.gov). Analyses of the molecular weight and theoretical isoelectric point were performed by using Compute pI/Mw tool at http://us.expasy.org/tools/pi_tool.html. Patterns and motifs of the amino acid sequence were analyzed by using ScanProsite available at http://us.expasy.org/tools/scanprosite/. Multiple alignment of the deduced amino acid sequence of AhCAD and other higher plant species was done using BOXSHADE 3.21 program (www.ch.embnet.org/software/BOX_form.html).

## RESULTS AND DISCUSSION

For the first time in *Acacia* species, a cDNA coding for the CAD enzyme has been studied. The primer, designed for 5'-RACE yielded a fragment of 750 bp which was cloned and sequenced in both directions. The partial 5'-sequence was then analyzed using BLASTN (Altschul et al. 1997) to confirm the gene before being digitally combined to the partial 3'-sequence generated from the cDNA library to get the full length cDNA sequence for AhCAD.

The full length AhCAD sequence (GenBank accession number AY769938) was 1346 bp long with one major open reading frame of 363 amino acids extending from 66-1155 bp, with a 5' untranslated region of 65 bp and a 3' untranslated region of 192 bp (Figure 1). The nucleotide sequences around the ATG start codon are in agreement with both the eukaryotic consensus sequence of Kozak with a purine in position −3, and a guanine in position +4 (Kozak 1984). A potential polyadenylation site recognized by the sequence ATTAAA (position 1267), and a stop codon sequence TTAAA that agrees with the stop codon sequence preferentially utilized by higher plants (Gallie 1993) were also found.

The AhCAD gene encodes a protein with calculated molecular weight of about 40 kDa, and an isoelectric point of 5.9. The first CAD from woody angiosperm isolated from *Eucalyptus gunni* was 1.39 kb and 38.8 kDa (Feuillet et al. 1993). In *Arabidopsis thaliana*, 9 putative CAD clones showed different lengths ranging from 1.07 to 1.13 kb and molecular mass from 38.7 to 40.9 kDa (Kim et al. 2004). The *Fxacad1* cDNA sequence isolated from *Fragaria* X *ananassa* cv. Chandler was 1.4 kb long with the predicted molecular mass of 39.3 kDa (Blanco-Portales et al. 2002). Thus, the predicted molecular mass of AhCAD agrees well with the molecular weights of other CADs.

Multiple alignment of AhCAD and other plant CAD sequences (Figure 2) revealed several regions of high homology. A motif for zinc-containing alcohol dehydrogenase: (GHEXXGXXXXXGXXV) found between residues 68-82 in this polypeptide sequence is in agreement with other plant CADs (Grima-Pettenati et al. 1993; Hibino et al. 1993b) This indicates that the AhCAD protein can be classified as a zinc-containing alcohol dehydrogenase. Besides, a GXGXXG motif was found between the amino acid residues 188-193, which is identified as an NADP-binding domain (Jornvall et al. 1987).

The AhCAD has several features present in the zinc-containing alcohol dehydrogenase. The putative ligands of the catalytic zinc, Cys-47, His-69 and Cys-163 are found in the AhCAD amino acid sequence. Cysteines 100, 103, 106 and 114 are four putative ligands that have been identified as structural zinc (Vallee & Aulds 1990; Grima-Pettenati et al. 1993). An NADP-specific residue, Ser-216, which is suggested to determine the cofactor specificity of CAD enzymes, was also found highly conserved with other plant CADs (Vallee & Aulds 1990; Knight et al. 1992; Grima-Pettenati et al. 1993). The AhCAD also exhibits a SKL motif, which putatively involved in peroxisomal targeting (Gould et al. 1988) and is specific in CAD protein (Blanco-Portales et al. 2002).

The comparison of *Acacia* hybrid AhCAD nucleotide and protein sequences in the GenBank revealed a significant identity with other CAD proteins ranging from 78% to 83 % and from 42 % to 69 % at the nucleotide and amino acid levels respectively. This high degree of consensus between the AhCAD amino acid sequence and other sequences corresponding to CAD enzymes from higher plants

strongly suggests *Acacia* hybrid cDNA encoding a CAD enzyme has been successfully isolated.

```
   1 GACATTCTTTCTTCTTCTTCTTCTTCTCTCTTTTTTCCTCCCTGTCTCTCTCTGTGTGAGAC  62
                 *   ****
  63 AACATGGGAAGCATTGAAGGAGAAAGAACAACAGTTGGATGGGCTGCAAGAGACCCTACT   122
       M  G  S  I  E  G  E  R  T  T  V  G  W  A  A  R  D  P  T
 123 GGGATTCTCTCTCCATACACCTTCAATCTCAGGAACACGGGGCCTGATGATGTGTATATT   182
       G  I  L  S  P  Y  T  F  N  L  R  N  T  G  P  D  D  V  Y  I
 183 AAGGTTCACTACTGTGGAATCTGCCACTCCGATCTGCATCAGATTAAGAATGATCTTGGC   242
       K  V  H  Y  C  G  I  C  H  S  D  L  H  Q  I  K  N  D  L  G
 243 ATGTCCAATTATCCCATGGTTCCTGGGCATGAGGTGGTTGGAGAAGTGATCGAGGTGGGT   302
       M  S  N  Y  P  M  V  P  G  H  E  V  V  G  E  V  I  E  V  G
 303 TCCAATGTGACCAAGTTTAAGGTCGGAGAAGTGGTTGGAGCAGGACTCATCGTTGGCAGC   362
       S  N  V  T  K  F  K  V  G  E  V  V  G  A  G  L  I  V  G  S
 363 TGCCGGAATTGCAGGGCATGCAAATCTGACATTGAGCAATATTGTGGCAAGAAGATCTGG   422
       C  R  N  C  R  A  C  K  S  D  I  E  Q  Y  C  G  K  K  I  W
 423 AATTACAACGATGTGTACACCGATGGAAAACCCCCCCAAGGTGGCTTTGCTGAAACCATG   482
       N  Y  N  D  V  Y  T  D  G  K  P  P  Q  G  G  F  A  E  T  M
 483 ATCGTCGATCAAAACTTCGTTGTGAAGATACCGGAGGGGATGTCACCGGAGCAAGTGGCG   542
       I  V  D  Q  N  F  V  V  K  I  P  E  G  M  S  P  E  Q  V  A
 543 CCACTGTTATGCGCCGGCGTGACGGTGTACAGTCCACTGTCACACTTTGGGCTGAGAGAT   602
       P  L  L  C  A  G  V  T  V  Y  S  P  L  S  H  F  G  L  R  D
 603 AGTGGGCTAAGAGGAGGAATATTGGGACTTGGAGGAGTGGGACACATGGGTGTGAAAGAT   662
       S  G  L  R  G  G  I  L  G  L  G  G  V  G  H  M  G  V  K  D
 663 AGGGGAAGGGCGGATTGGAAACACCCAGTGTGCCAGGTGATAGGCTCGTCGGAGAAGAGA   722
       R  G  R  A  D  W  K  H  P  V  C  Q  V  I  G  S  S  E  K  R
 723 AAGAAGGAGGCTAGGGAGACCTTGGAGCAAACGATTATGTGGTTAAGTTCAGATGAAACT   782
       K  K  E  A  R  E  T  L  E  Q  T  I  M  W  L  S  S  D  E  T
 783 CAGATGCAGAAGATTGCTGATTCTCTGGATTATATTATTGATACCGTACCGGTGGGTCAC   842
       Q  M  Q  K  I  A  D  S  L  D  Y  I  I  D  T  V  P  V  G  H
 843 CCTCTTGAGCCTAATCTTTCTTTGCTAAAAATTGATGGCAAGTTGATCTTAATGGGTGTT   902
       P  L  E  P  N  L  S  L  L  K  I  D  G  K  L  I  L  M  G  V
 903 ATCAACACTCCTTTGCAATTCGTTAGCCCCATGGTCATGCTTGGGAGGAAGACGATAACA   962
       I  N  T  P  L  Q  F  V  S  P  M  V  M  L  G  R  K  T  I  T
 963 GGAAGCTTCATAGGGAGCATAAAGGAGACAGAAGAGATGTTAGAATTCTGGAAAGAGAAG  1022
       G  S  F  I  G  S  I  K  E  T  E  E  M  L  E  F  W  K  E  K
1023 GGTCTGAGTTCAATGATAGAGATGGTGAAGATGGATTACATAAACAAGGCACTTGAGAGG  1082
       G  L  S  S  M  I  E  M  V  K  M  D  Y  I  N  K  A  L  E  R
1083 TTGGAGAAGAACGATGTGAGGTATAGGTTCGTGGTTGACGTTGCTGGCAGCAAACTTGAT  1142
       L  E  K  N  D  V  R  Y  R  F  V  V  D  V  A  G  S  K  L  D
1143 GATCATCAGTGAAACAATATTTACAGCAGCCATTCCTATACTTTCTTACATAACCTTTTG  1202
       D  H  Q
1203 CTAAATGTATGATTATGAATATGATGCCTCACAGATGTTTGTCGTGTTACGTTGTGTGGA  1262
1263 ATATATTAAATGTGCTCTGTTTTCCCTTTTCAAAGAATTCCCAAATTTGATGTCGGAAAA  1322
1323 AAAAAAAAAAAAAAAAAAAAAAA  1346
```

Figure 1. Nucleotide and amino acid sequences for AhCAD. Putative start codon, stop codon and the polyadenilation site are underlined. Asterisks indicate nucleotides matching the consensus start codon found in plants (Gallie, 1993). Motifs for zinc-containing alcohol dehydrogenase and $NAD^+/NADP^+$ are boxed. Conserved amino acid pattern for plant CAD is in bold. The SKL motif, only present in CAD protein motif, is highlighted in black.

The information obtained from the AhCAD sequence and the characterization of the gene is vital for understanding the lignin biosynthesis pathway in the *Acacia* hybrid. The AhCAD sequence is being employed to study the

expression pattern of this gene through real-time RT-PCR. The available AhCAD sequence could also be used to study AhCAD polymorphisms in the development of molecular markers for high quality wood pulp in marker-assisted selection breeding program for *Acacia* hybrid.

```
                                        0
P.tomentosa   1 -
MGSLETERKIVGWAATDSTGHLAPYTYSLRDTGPEDVFIKVISCGICHTDIHQIKNDLG
P.deltoides   1 -
MGSLETERKIVGWAATDSTGHLAPYTYSLRDTGPEDVFIKVISCGVCHTDIHQIKNDLG
A.cordata     1 -
MGSLEAERKTTGWAARDPSGVLSPYTYTLRETGPEDVFIKIIYCGICHTDIHQIKNDLG
E.gunnii      1 -
MGSLEKERTTTGWAARDPSGVLSPYTYSLRNTGPEDLYIKVLSCGVCHSDIHQIKNDLG
N.tabacum     1 -
MGGLEVEKTTIGWAARDPSGVLSPYTYTLRNTGPEDVEVKVLYCGLCHTDLHQVKNDLG
AhCAD          1 -
MGSIEGERTTVGWAARDPTGILSPYTFNLRNTGPDDVYIKVHYCGICHSDLHQIKNDLG
P.taeda       1 -
MGSLESEKTVTGYAARDSSGHLSPYTYNLRKKGPEDVIVKVIYCGICHSDLVQMRNEMG
P.abies       1 -
MGSLESERTVTGYAARDSSGHLSPYTYTLRNKGPEDVIVRVIYCGICHSDLVQMHNEMG
A.thaliana    1 --
MGKVLQKEAFGLAAKDNSGVLSPFSFTRRETGEKDVRFKVLFCGICHSDLHMVKNEWG
Fragaria      1
MSIEQEHPNKASGWAARDSSGVLSPFNFSRRETGEKDVMFKVLYCGICHSDHHMVKNEWG

                  0                    * * *    *
P.tomentosa   60
MSHYPMVPGHEVVGEVVEVGSDVTRFKVGDVVGVGVIVGSCKNCHPCKSELEQYCNKKIW
P.deltoides   60
MSHYPMVPGHEVVGEVVEVGSDVTRFKVGDVVGVGVIVGSCKNCHPCKSEIEQYCNKKIW
A.cordata     60
ASNYPMVPGHEVVGEVVEVGSDVTKFKVGDCVGDGTIVGCCKTCRPCKADVEQYCNKKIW
E.gunnii      60
MSHYPMVPGHEVVGEVLEVGSEVTKYRVGDRVGTGIVVGCCRSCSPCNSDQEQYCNKKIW
N.tabacum     60
MSNYPLVPGHEVVGEVVEVGPDVSKFKVGDTVGVGLLVGSCRNCGPCKRDIEQYCNKKIW
AhCAD          60
MSNYPMVPGHEVVGEVIEVGSNVTKFKVGEVVGAGLIVGSCRNCRACKSDIEQYCGKKIW
P.taeda       60
MSHYPMVPGHEVVGIVTEIGSEVKKFKVGEHVGVGCIVGSCRSCGNCNQSMEQYCSKRIW
P.abies       60
MSNYPMVPGHEVVGVVTEIGSEVKKFKVGEHVGVGCIVGSCRSCSNCNGSMEQYCSKRIW
A.thaliana    59
MSTYPLVPGHEIVGVVTEVGAKVTKFKTGEKVGVGCLVSSCGSCDSCTEGMENYCPKSIQ
Fragaria      61
FSTYPLVPGHEIVGEVTEVGSKVQKFKVGDRVGVGCIVGSCRSCENCTDHLENYCPKQIL

                                        0

P.tomentosa   120
SYNDVYTDGKPTQGGFAESMVVDQKFVVRIPDGMSPEQAAPLLCAGLTVYSPLKHFGLKQ
P.deltoides   120
SYNDVYTDGKPTQGGFAESMVVHQKFVVRIPDGMSPEQAAPLLCAGLTVYSPLKHFGLKQ
```

```
A.cordata 120
SYNDVYTDGKPTQGGFSGHMVVDQKFVVKIPDGMAPEQAAPLLCAGVTVYSPLTHFGLKE
E.gunnii 120
NYNDVYTDGKPTQGGFAGEIVVGERFVVKIPDGLESEQAAPLMCAGVTVYSPLVRFGLKQ
N.tabacum 120
NCNDVYTDGKPTQGGFAKSMVVDQKFVVKIPEGMAPEQAAPLLCAGITVYSPLNHFGFKQ
AhCAD 120
NYNDVYTDGKPPQGGFAETMIVDQNFVVKIPEGMSPEQVAPLLCAGVTVYSPLSHFGLRD
P.taeda 120
TYNDVNHDGTPTQGGFASSMVVDQMFVVRIPENLPLEQAAPLLCAGVTVESPMKHFAMTE
P.abies 120
TYNDVNHDGTPTQGGFASSMVVDQMFVVRIPENLPLEQAAPLLCAGVTVYSPMKHFGMTE
A.thaliana 119
TYGFPYYDNTITYGGYSDHMVCEEGFVIRIPDNLPLDAAAPLLCAGITVYSPMKYHGLDK
Fragaria 121
TYGANYYDGTTTYGGCSDIMVAHEHFVVRIPDNLPLDGAAPLLCAGITTYSPLRYFGLDK

                                    ↓

P.tomentosa 180 -SGLRGGILGLGGVGHMGVK-IAKAMGHH--
VTVISSSDKKREEAMEHLGADEYLVSSDV
P.deltoides 180 -SGLRGGILGLGGVGHMGVK-IAKAMGHH--
VTVISSSDKKREEAMEHLGADEYLVSSDV
A.cordata 180 ISGLRGGILGLGGVGHMGVK-LAKAMGHH--
VTVISSSDKKKEEAIDHLGADAYLVSSDA
E.gunnii 180 -SGLRGGILGLGGVGHMGVK-IAKAMGHH--
VTVISSSDKKRTEALEHLGADAYLVSSDE
N.tabacum 180 -SGLRGGILGLGGVGHMGVK-IAKAMGHH--
VTVISSSNKKRQEALEHLGADDYLVSSDT
AhCAD 180 -
SGLRGGILGLGGVGHMGVKDRGRADWKHPVCQVIGSSEKRKKEARETLEQTIMWLSSDE
P.taeda 180 -PGKKCGILGLGGVGHMGVK-IAKAFGLH--
VTVISSSDKKKEEAMEVLGADAYLVSKDT
P.abies 180 -PGKKCGILGLGGVGHMGVK-IAKAFGLH--
VTVISSSDKKKEEALEVLGADAYLVSKDA
A.thaliana 179 -PGMHIGVVGLGGLGHVGVK-FAKAMGTK--
VTVISTSEKKRDEAINRLGADAFLVSRDP
Fragaria 181 -PGMHVGVVGLGGLGHVAVK-FAKAMGVK--
VTVISISPKKEEEALKHLGADSFLVSRDQ


P.tomentosa 236
ESMQKAADQLDYIIDTVPVVHPLEPYLSLLKLDGKLILMGVINTPLQFGTPMVMLGRKSI
P.deltoides 236
ESMQKAADQLDYIIDTVPVVHPLEPYLSLLKLDGKLILMGVINAPLQFVTPMVMLGRKSI
A.cordata 237
TQMQEAADSLDYIIDTVPVFHPLEPYLSLLKLDGKLILMGVINTPLQFISPMVMLGRKAI
E.gunnii 236
NGMKEATDSLDYIFDTIPVVHPLEPYLALLKLDGKLILTGVINAPLQFISPMVMLGRKSI
N.tabacum 236
DKMQEASDSLDYIIDTVPVGHPLEPYLSLLKIDGKLILMGVINTPLQFISPMVMLGRKSI
AhCAD 239
TQMQKIADSLDYIIDTVPVGHPLEPNLSLLKIDGKLILMGVINTPLQFVSPMVMLGRKTI
P.taeda 236
EKMMEAAESLDYIMDTIPVAHPLEPYLALLKTNGKLVMLGVVPEPLHFVTPLLILGRRSI
P.abies 236
EKMQEAAESLDYIMDTIPVAHPLEPYLALLKTNGKLVMLGVVPEPLHFVTPLLILGRRSI
```

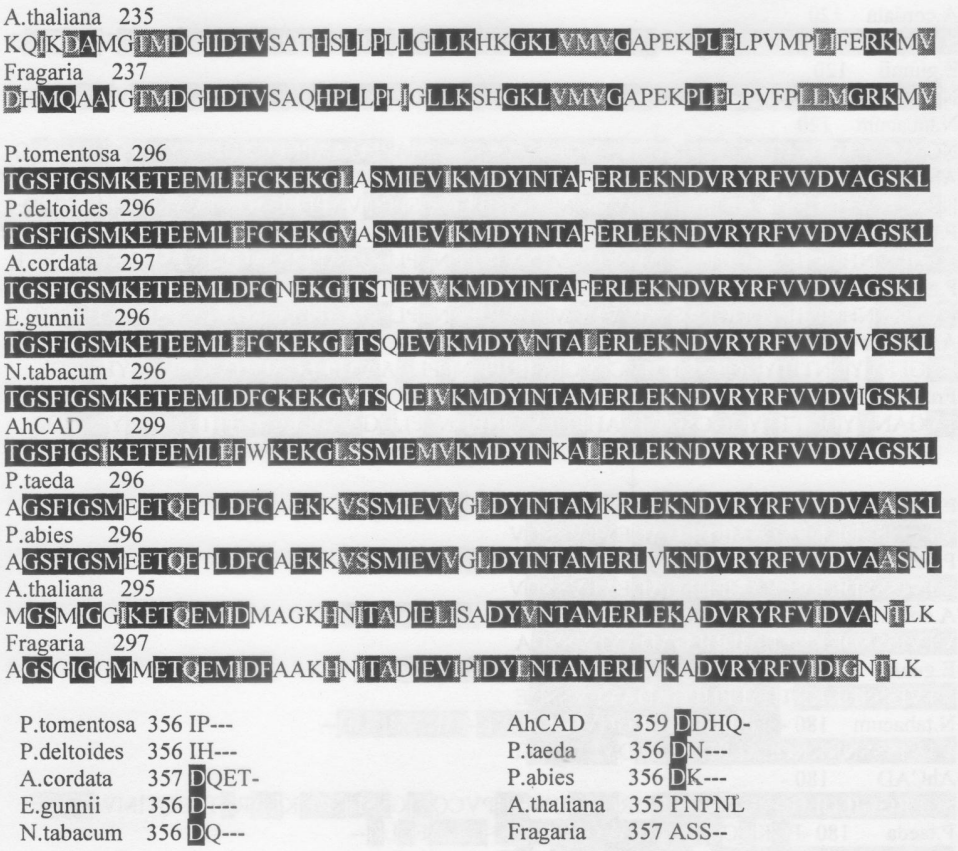


Figure 2. Multiple alignment of the deduced amino acid sequence of AhCAD and other higher plant species. Consensus amino acids are in black boxes and conserved amino acids are shaded. Open dot (0) indicates the Cys-47, His-69 and Cys-163 residues conserved in all plant CADs. Asterisks indicates the cysteine residues identified as structural ligands. Arrow indicates the residue Ser-216 which distinguishes NADP from NAD in the enzyme family. The GenBank accession numbers are AAR83343 for *Populus tomentosa*, T09141 for *Populus deltoids*, BAA03099 for *Aralia cordata*, P31655 for *Eucalyptus gunnii*, S23525 for *Nicotiana tabacum*, AY769938 for *Acacia mangium x Acacia auriculiformis* (AhCAD), S49444 for *Pinus taeda*, S39509 for *Picea abies*, S28043 for *Arabidopsis thaliana* and AAK28509 for *Fragaria* X *ananassa*.

## ACKNOWLEDGEMENTS

This work was supported by IRPA grant 01-02-02-0010-PR0003/03-01 of the Ministry of Science, Technology and Innovation, Malaysia.

## REFERENCES

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped-Blast and Psi-Blast: a new generation of protein database search programs. *Nucleic Acids Research* **25**: 3389–3402.

Baucher M, Monties B, van Montagu M, Boerjan W. 1998. Biosynthesis and genetic engineering of lignin. *Critical Review of Plant Science* **17**: 25–197.

Blanco-Portales R, Medina-Escobar N, Lopez-Raez JA, Gonzalez-Reyes JA, Villalba JM, Moyano E, Caballero JL, Munoz-Blanco J. 2002. Cloning, expression and immunolocalization pattern of a cinnamyl alcohol dehydogenase gene from strawberry (*Fragaria x annassa* cv. Chandler). *Journal of Experimental Botany* **53**(375): 1723-1734.

Boudet AM, Lapierre C, Grima-Pettenati J. 1995. Biochemistry and molecular biology of lignification. *New Phytology* **129**: 203–236.

Feuillet C, Boudet AM, Grima-Pettenati J. 1993. Nucleotide sequence of a cDNA encoding cinnamyl alcohol dehydrogenase from *Eucalyptus*. *Plant Physiology* **103**: 1447.

Gallie D. 1993. Post-transcriptional regulation of gene expression in plants. *Annual Review of Plant Physiology Plant Molecular Biology* **44**: 77-105.

Goffner D, Joffroy I, Grima-Pettenati J, Halpin C, Knight ME, Schuch W, Boudet AM. 1992. Purification and characterization of isoforms of cinnamyl alcohol dehydrogenase from *Eucalyptus* xylem. *Planta* **188**: 48–53.

Gould SJ, Keller GA, Subramani S. 1988. Identification of peroxisomal targeting signals located at the carboxy terminus of four peroxisomal proteins. *Journal of Cell Biology* **107**: 897-905.

Grima-Pettenati J, Feuillet C, Goffner D, Borderies G, Boudet A.M. 1993. Molecular cloning and expression of a *Eucalyptus gunnii* cDNA clone encoding cinnamyl alcohol dehydrogenase. *Plant Molecular Biology* **21**: 1085-1095.

Halpin C, Knight ME, Grima-Pettenati J, Goffner D, Boudet A, Schuch W. 1992. Purification and characterisation of cinnamyl alcohol dehydrogenase from tobacco stems. *Plant Physiology* **98**: 12-16.

Hawkins S, Samaj J, Lauvergeat V, Boudet A, Grima-Pettenati J. 1997. Cinnamyl alcohol dehydrogenase: identification of new sites of promoter activity in transgenic poplar. *Plant Physiology* **113**: 321-325.

Hawkins SW, Boudet AM. 1994. Purification and characterization of cinnamyl alcohol dehydrogenase isoforms from the periderm of *Eucalyptus gunnii* hook. *Plant Physiology* **104**: 75-84.

Hibino T, Shibata D, Chen JQ, Higuchi T. 1993a. Cinnamyl alcohol dehydrogenase from *Aralia cordata*-cloning of the cDNA and expression of the gene in lignified tissues. *Plant Cell Physiology* **l34**: 659-665.

Hibino T, Shibata D, Umezawa T, Higuchi T. 1993b. Purification and partial sequences of *Avalia cordata* cinnamyl alcohol dehydrogenase. *Phytochemistry* **32:** 565-567.

Jornvall H, Person B, Jeffery J. 1987. Characterization of alcohol/polyol dehydrogenases: The zinc-containing long chain alcohol dehydrogenases. *European Journal of Biochemistry* **167**: 195–201.

Kim SJ, Kim MR, Bedgar DL, Moinuddin SGA, Cardenas CL, Davin LB, Kang C, Lewis NG. 2004. Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in *Arabidopsis*. *Proceedings of National Academic Science U.S.A.* **101**(6): 1455-1460.

Knight ME, Halpin C, Schuch W. 1992. Identification and characterization of cDNA clones encoding cinnamyl alcohol dehydrogenase from tobacco. *Plant Molecular Biology* **19**: 793–801.

Kolattukudy PE. 1981. Structure, biosynthesis and biodegradation of cutin and suberin. *Annual Review of Plant Physiology* **32**: 539-567.

Kozak M. 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Research* **12**: 857–872.

Lewis NG, Yamamoto E. 1990. Lignin: occurrence, biogenesis and biodegradation. *Annual Review of Plant Physiology and Plant Molecular Biology* **41**: 455-496.

Lüderitz T, Grisebach H. 1981. Enzymic synthesis of lignin precursors: Comparison of cinnamoyl-CoA reductase and cinnamyl alcohol:NADP_ dehydrogenase from spruce (*Picea abies* L.) and soybean (*Glycine max* L.). *European Journal of Biochemistry* **119**: 115–124.

MacKay JJ, Liu W, Whetten R, Sederoff RR, O'Malley D. 1995. Genetic analysis of cinnamyl alcohol dehydrogenase in loblolly pine: Single gene inheritance, molecular characterization and evolution. *Molecular Gene Genetics* **247**: 537–545.

O'Malley DM, Porter S, Sederoff RR. 1992. Purification, characterization, and cloning of cinnamyl alcohol dehydrogenase in loblolly pine *(Pinus taeda* L.). *Plant Physiology* **98**: 1364-1371.

Sarni F, Grand G, Boudet AM. 1984. Purification and properties of cinnamyl-CoA reductase and cinnamyl alcohol dehydrogenase from poplar stems. *European Journal of Biochemistry* **139:** 259–265.

Vallee BL, Aulds DS. 1990. Zinc coordination, function and structure of zinc enzymes and other protein. *Biochemistry* **29**: 5647-5659.

Whetten R, Sederoff R. 1995. Lignin biosynthesis. *Plant Cell* **7**: 1001–1013.

Wickneswari R, Cheong PL, Pang SL, Choong CY, Harikrishna JA, Elliott RC, Vaillancourt RE. 2004. Analysis of a batch of expressed sequence tags from inner bark cDNA library of *A.mangium* x *A.auriculiformis* interspecific hybrid (In Malay). In Proceedings of the *7th National Biology Symposium*, Awana Genting Highlands, Malaysia, 18-20 May 2004. Pp. 477-480.

Wyrambik D, Grisebach H. 1975. Purification and properties of isoenzymes of cinnamyl-alcohol dehydrogenase from soybean cell-suspension cultures. *European Journal of Biochemistry* **59**: 9–15.

Zinser C, Ernst D, Sandermann H Jr. 1998. Induction of stilbene synthase and cinnamyl alcohol dehydrogenase mRNAs in Scots pine (*Pinus sylvestris* L.) seedlings. *Planta* **204**: 169–176.