**GOLD HULL AND INTERNODE2 Encodes a Primarily Multifunctional Cinnamyl-Alcohol Dehydrogenase in Rice**

Kawei Zhang, Qian Qian, Zejun Huang, Yiqin Wang, Ming Li, Lilan Hong, Dali Zeng, Minghong Gu, Chengcai Chu, and Zhukuan Cheng*

State Key Laboratory of Plant Genomics and National Center for Plant Gene Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, People’s Republic of China (K.Z., Z.H., Y.W., M.L., L.H., D.Z., C.C., Z.C.); State Key Laboratory of Rice Biology, China National Rice Research Institute, Hangzhou 310006, People’s Republic of China (Q.Q., D.Z.); and Graduate School of the Chinese Academy of Sciences, Beijing 100049, People’s Republic of China (K.Z., Z.H., L.H., D.Z.)

Lignin content and composition are two important agronomic traits for the utilization of agricultural residues. Rice (*Oryza sativa*) gold hull and internode phenotype is a classical morphological marker trait that has long been applied to breeding and genetics study. In this study, we have cloned the **GOLD HULL AND INTERNODE2** (*GH2*) gene in rice using a map-based cloning approach. The result shows that the *gh2* mutant is a lignin-deficient mutant, and *GH2* encodes a cinnamyl-alcohol dehydrogenase (*CAD*). Consistent with this finding, extracts from roots, internodes, hulls, and panicles of the *gh2* plants exhibited drastically reduced *CAD* activity and undetectable sinapyl alcohol dehydrogenase activity. When expressed in *Escherichia coli*, purified recombinant *GH2* was found to exhibit strong catalytic ability toward coniferaldehyde and sinapaldehyde, while the mutant protein *gh2* completely lost the corresponding *CAD* and sinapyl alcohol dehydrogenase activities. Further phenotypic analysis of the *gh2* mutant plants revealed that the *p*-hydroxyphenyl, guaiacyl, and sinapyl monomers were reduced in almost the same ratio compared to the wild type. Our results suggest *GH2* acts as a primarily multifunctional *CAD* to synthesize coniferyl and sinapyl alcohol precursors in rice lignin biosynthesis.

In addition to cellulose, lignin is the second major biopolymer component of the plant cell wall. Lignin cross-links with carbohydrates, such as cellulose and hemicellulose, to play an important role of the cell walls, enhancing the rigidity, conferring resistance to pathogens and mechanical stress, and enabling solute transport in the xylem (Brill et al., 1999; Chabannes et al., 2001; Jones et al., 2001). Lignin is a complex phenolic polymer that is mainly derived from the oxidative polymerization of *p*-coumaryl, coniferyl, and sinapyl alcohol precursors as the principal monomers, and the monomers are defined as *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units. The types and amount of lignin units vary depending on plant species, tissues, cell types, and cell wall layers and are influenced by developmental and environmental cues (Campbell and Sederoff, 1996; Donaldson, 2001).

Apart from its critical role in plant development, lignin has great economical and environmental value. Considerable monetary and environmental costs are incurred by industry in removing lignin from cellulose during production of pulp and paper (Hu et al., 1999). Similarly, lignin limits the digestibility and energy yield of forage crops (Jung and Vogel, 1986; Cherney et al., 1990). Studies on the lignin metabolic pathway help to understand plant development and modify plants for meeting the industrial and livestock needs (Baucher et al., 1996; Halpin et al., 1998). Many investigations were conducted on a series of maize (*Zea mays*) *brown midrib* (*bm*) mutants, which have been used as a model for digestibility and lignin studies (for review, see Barriere et al., 2004). *bm* mutants have a reddish-brown pigmentation in the leaf midrib and stalk pith, which is associated with lignified tissues. Researches show that these *bm* mutants are deficient in the enzymes that control the lignin content or composition in lignin biosynthesis (Vignols et al., 1995; Halpin et al., 1998; Piquemal et al., 2002). The mutants have aroused considerable agronomic interest in past decades because the changes in lignin content and composition may affect intake digestion and milk production by dairy cows (Ebling and Kung, 2004;...
Taylor and Allen, 2005). Rice (*Oryza sativa*) is one of the most important crops in the world, feeding more than half of the world’s population, so identification of the corresponding lignin mutants will have important implications for agriculture.

Cinnamyl-alcohol dehydrogenase (CAD) was one of the first enzymes studied in lignin biosynthetic pathway, and it catalyzes the final reduction of the hydroxyl-cinnamaldehydes to the corresponding alcohols (Mansell et al., 1974). For the important role in regulation of the lignin content and composition, many CAD homologs have been isolated from different plant species and used to modify plants by transgenic approach (Halpin et al., 1994; Baucher et al., 1996, 1999; Ralph et al., 1997). CAD was initially believed to be multifunctional to catalyze the reduction of three hydroxyl-cinnamaldehydes (p-coumaryl, coniferyl, and sinapyl aldehydes) to the corresponding alcohols (for review, see Boerjan et al., 2003). A sinapyl alcohol dehydrogenase (SAD) has previously been isolated from aspen (*Populus tremuloides*) and proven to be specifically involved in the reduction of sinapaldehyde, while CAD from aspen has been proven to be specifically involved in the reduction of coniferaldehyde (Li et al., 2001). The result supports a hypothesis that CAD is involved only in the reduction of coniferylaldehyde, while SAD is specifically involved in the reduction of sinapaldehyde in angiosperm plants. The genomes of Arabidopsis (*Arabidopsis thaliana*) and rice have been fully sequenced and annotated (Arabidopsis Genome Initiative, 2000; International Rice Genome Sequencing Project, 2005), which provides a good opportunity to conduct genome-wide analyses of CAD family members. So far, nine Arabidopsis CAD members have been annotated as the bona fide CAD homologs, and their biochemical properties have been well studied (Kim et al., 2004). *AthCADC* and *AthCADD* of the family have been studied in null mutants and double mutants (Sibout et al., 2003, 2005) and proven to be the primary CAD genes in Arabidopsis lignification. Tobias and Chow (2005) have recently shown that there are 12 CAD family members in rice based on bioinformatics criteria, but the corresponding biochemical properties and physiological functions remain unknown.

The rice gold hull and internode (*gh*) mutant was first described as early as 1917 (http://www.gramene.org/db/mutant/search_mutant?id=GR0060365). It exhibits a reddish-brown pigmentation in the hull and internode, instead of in the midrib. The hull and internode become golden yellow at maturation, so the mutant is named as *gh* rice. Based on the morphological and genetic descriptions, a series of *gh* mutants (*gh1, gh2, gh3, and gh4*) have been identified. Among them, *gh1* was located on chromosome 5 (Nagao and Takahashi, 1963), and *gh2* and *gh3* were both located on chromosome 2 (Iwata and Omura, 1971, 1977). Despite the fact that these mutants have long been used as marker genes in rice breeding and genetics study, the mechanism underlying the phenotype is still unknown. In this study, the *gh2* was identified to be a lignin-deficient mutant by a map-based cloning approach, and the *GH2* gene codes a CAD. Both native and recombinant *GH2* protein exhibit CAD activity and SAD activity, but native and recombinant gh2 protein lost the corresponding activities. Additionally, the H, G, and S monomers were similarly reduced in the *gh2* mutant plants compared with the wild type. Our data suggest that coniferyl alcohol precursors and sinapyl alcohol precursors are both derived from the GH2-catalyzing reduction of coniferaldehyde and sinapaldehyde in the rice hull and internode.

**RESULTS**

**Morphological Characterization of the *gh2* Mutant**

We evaluated the morphology of the *gh2* mutant and compared it with the wild-type variety, Zhefu802. *gh2* exhibits an obvious reddish-brown pigment in the panicle (hull), internode, and basal leaf sheath at the heading stage (Fig. 1, A and B). The reddish-brown pigments gradually become intensive starting from the basal internode to the apical internode. During plant maturing, seed coloration becomes darker and finally changes to golden yellow at maturation (Fig. 1C).

![Figure 1](https://www.plantphysiol.org/)
Except the reddish-brown coloration in the above specific tissues, the gh2 mutant plants show similar development with the wild type under normal cultivated condition.

**The GH2 Gene Encodes a CAD**

The gh2 gene has been genetically mapped to the region around 20 cM on chromosome 2 by the primary trisomics (Sanchez and Khush, 1994). To achieve a more precise map for the gh2 locus, we generated a large F2 mapping population between gh2 and Nipponbare, a japonica rice variety that has been fully sequenced. Of the approximately 13,000 F2 plants, 3,256 segregants showing the gh2 mutant phenotype were further used for fine mapping. Based on this, the gh2 gene was further located to a genomic region demarcated by sequence tag site (STS) markers P1 and S21668 (Fig. 2A). Between the two markers, we developed six additional STS markers and two cleaved amplified polymorphic sequence (CAPS) markers (Table I), leading to the mapping of this gene to a narrow region between two CAPS markers (P7 and P8; Fig. 2B). No recombinant near the STS marker P9 was found. The candidate genes near the P9 marker in the 30-kb region were chosen to be sequenced, and a gene encoding a CAD was found to be mutated by one nucleotide substitution G to A (554G > 554A) in the fourth exon, which substituted a zero-charge Gly to an electronegative Asp (Fig. 2C).

The identity of GH2 was confirmed by genetic complementation experiments. The plasmid pCGH, containing the entire open reading frame (ORF), 2,397-bp upstream region and 2,460-bp downstream region, and pCGHS, containing a partial coding region of the ORF (Fig. 2D), were introduced into the recessive mutants from the F3 generation of the mapping population, then 12 and eight independent transgenic lines were obtained from the two constructs, respectively. All 12 lines of pCGH showed complete complementation of the gh2 phenotype, whereas all eight lines of pCGHS failed to rescue the gh2 mutant. Additionally, the mutation in the gh2 produces an Aat II site in the genomic DNA, which can be used as a CAPS marker to determine the gh2 mutant background in the complementation test (Fig. 2E). Therefore, we conclude that the candidate gene controls the gh2 phenotype in rice.

**Phylogenetic Analysis of GH2**

To determine the evolutionary relationship between GH2 and CAD family members from the rice and Arabidopsis CAD families as well as other representative identified CAD homologs, an unrooted tree was built using the neighbor-joining method based on full-length protein sequences (Fig. 3). The result

---

**Figure 2.** Cloning and confirmation of the GH2 gene. A, The GH2 locus was mapped to a region between markers P1 and S21668 in the short arm of rice chromosome 2 (Chr2). B, Fine mapping of the GH2 locus using primers developed from seven bacterial artificial chromosomes (BACs): BAC1, OJ1297_C09; BAC2, OJ1572_F02; BAC3, OJ1073_F05; BAC4, OJ1145_F01; BAC5, OSJNB0031B09; BAC6, OJ1136_D07; and BAC7, J1225_F07. P1 to P9 are the developed markers in the work (Table I). C, GH2 gene structure, showing the mutation site of the gh2. The start codon (ATG) and the stop codon (TGA) are indicated. Black and white boxes indicate the exon and intron sequences, respectively. The G to A substitution in the gh2 converted a zero-charge Gly to an electronegative Asp. D, Complementation constructs. The construct pCGH contains the entire GH2 gene, plus a 2,397-bp upstream region and 2,460-bp downstream region. The plasmid pCGHS contains a partial coding region of the ORF. E, Identification of transgenic plants. The G to A substitution in the gh2 creates an Aat II site that is used for identifying the mutant line, the complement line, and the wild-type line. Lane M, DL2000 marker; lane 1, gh2; lane 2, wild type; lane 3, the pCGH-transformed rice line 1.
indicates that all the CAD protein sequences are divided into two subfamilies: subA (GH2, sugarcane [Saccharum officinarum] CAD), maize CAD, aspen PtCAD, eucalyptus [Eucalyptus globules] CAD, tobacco [Nicotiana tabacum] CAD1, tobacco CAD2, Lucerne [Medicago sativa] MsaCad2, Arabidopsis CADC, Arabidopsis CADD, pine [Pinus taeda] CAD, and spruce [Picea abies] CAD, and subB (Lucerne MsaCad1, aspen PtSAD, OsCAD1, OsCAD3, OsCAD4, OsCAD5, OsCAD6, OsCAD7, OsCAD8A, OsCAD9, AtCAD1, AtCAD2, AtCAD3, AtCAD6, AtCAD7, AtCAD8, and AtCAD9). SubA members are evolutionally conserved CAD genes in gymnosperm and angiosperm, which has been regarded as the true CAD clade before angiosperm and gymnosperm split (Raes et al., 2003; Peter and Neale, 2004). In the complete CAD family of Arabidopsis, two CAD members (AtCADC and AtCADD) belong to subA subfamily. Only double mutation of the two subA CAD members (AtCADC and AtCADD) can cause the lignin-deficient phenotype. Although rice has 12 CAD members, only GH2 belongs to the subA CAD subfamily. The fact that the point mutation of GH2 in rice has led to the lignin-deficient phenotype demonstrates that GH2 is the only primary enzyme that catalyzes the last step of lignin precursor biosynthesis. SubB is an angiosperm-specific subfamily, which can be regarded as the duplication of subA, and the members are likely to gain new functions in evolution. So far, Lucerne MsaCAD1 and Arabidopsis AtCAD7 in SubB have been proven to participate in plant defense, aspen PtSAD has been proven to catalyze sinapyl alcohol biosynthesis (Kiedrowski et al., 1992; Brill et al., 1999; Li et al., 2001), but the functions of rice subB subfamily remain to be elucidated. Based on the role of subB members in other plants, we predict that the subB CAD members in rice may function in enhancing the cell wall for plant defense or act in assistant roles of lignification.

**GH2 Expression Pattern and CAD Activities of Total Proteins**

Since the mutation is a point mutation, GH2 gene expression was detected in the wild-type plants only at the heading stage (Fig. 4B). We detected GH2 gene transcription in all seven tissues, including panicles, hulls, blades, midribs, leaf sheaths, internodes, and young roots, with a 217-bp 3’ untranslated sequence as the specific probe. The results showed GH2 gene was expressed in the lignified tissues, including panicles, hulls, internodes, young roots, and leaf sheaths. In the leaves (blade and midrib), which are slightly lignified, the CAD transcripts cannot be obviously detected by northern blot. Furthermore, GH2 gene expressed stronger in the first internode than in the second internode, suggesting that the gene expressed gradually stronger starting from the basal internode to the apical internode (data not shown). Consequently, the GH2 expression pattern is closely related to the rice lignification pattern and the localization of reddish-brown pigments.

To study the catalytic ability of native GH2 in rice, CAD activity and SAD activity of the total proteins were assayed in all tissues of wild-type and gh2 plants at the heading stage. Strong CAD activity and weak SAD activity were detected in all the tissues of the wild type (Fig. 4A). In the gh2 sheaths, roots, panicles, hulls, and internodes in which GH2 gene was strongly expressed, the CAD activity was relatively strongly expressed, the SAD activity was not drastically reduced and the SAD activity was not detectable, whereas, in the leaves (blade and midrib) in which GH2 gene was weakly expressed, the CAD activity was drastically reduced and the SAD activity was not detectable, whereas, in the leaves (blade and midrib).
activity and SAD activity were slightly reduced. The reductions are consistent with the expression pattern of GH2 described above (Fig. 4B). Thus, the result shows that GH2 protein has CAD activity and SAD activity, and the CAD activity and SAD activity were drastically reduced in the gh2 plants. In the leaves in which the GH2 gene is slightly expressed, the total proteins of the gh2 mutant still exhibit high CAD activity and weak SAD activity. Additionally, the reduction ratios of CAD activity in the different tissues

![Figure 3. Phylogenetic analysis of GH2. A neighbor-joining tree was built on the full-length protein sequences of rice CAD family members, Arabidopsis CAD family members, and other representative identified CAD homologs. The scale bar is an indicator of genetic distance based on branch length. The subA subfamily includes rice GH2, sugarcane CAD, maize CAD, aspen PtCAD, eucalyptus CAD, tobacco CAD1, tobacco CAD2, Lucerne MsaCad2, Arabidopsis CADC, Arabidopsis CADD, pine CAD, and spruce CAD. The subB subfamily includes Lucerne MsaCad1, aspen PtSAD, rice OsCAD1, OsCAD3, OsCAD4, OsCAD5, OsCAD6, OsCAD7, OsCAD8A, OsCAD8B, OsCAD8C, OsCAD8D, OsCAD9, Arabidopsis AtCAD1, AtCAD2, AtCAD3, AtCAD6, AtCAD7, AtCAD8, and AtCAD9. Accession numbers of the sequences used to build the tree are as follows: AY288079 (AtCAD1), AY302077 (AtCAD2), AY302078 (AtCAD3), AY302081 (AtCAD4), AY302082 (AtCAD5), AY302075 (AtCAD6), AY302079 (AtCAD7), AY302080 (AtCAD8), AY302076 (AtCAD9), OsCAD1 (AAC09864), GH2 (DQ234272), OsCAD3 (AAP53892), OsCAD4 (BK003970), OsCAD5 (BK003971), OsCAD6 (CAD39907), OsCAD7 (CAE05207), OsCAD8A to D (BK003972), OsCAD9 (AAC05338), sugarcane CAD (AJ231135), maize CAD (AJ005702), spruce CAD (X72675), aspen CAD (AF217957), aspen SAD (AF273256), eucalyptus (AF038561), tobacco CAD1 (X62343), tobacco CAD2 (X62344), pine (Z37992), Lucerne MsaCad1 (AAC35846), and Lucerne MsaCad2 (AAC35845).
are unequal. So we speculate it is possible that additional CAD isoenzymes with CAD activity also unevenly exist in different tissues of the plants. Most importantly, the fact that the SAD activity was completely abolished due to the defective GH2 protein in panicles, hulls, internodes, and sheaths shows that no other SAD homolog with obvious SAD activity exists in these tissues.

Enzyme Kinetics of the Recombinant GH2 and gh2

To evaluate the way in which the G to A nucleotide substitution impaired enzymatic function in the gh2 mutant plant, we expressed the wild-type protein GH2 and the point-mutated protein gh2 (G185D) in *Escherichia coli*. The recombinant proteins were purified by glutathione-Sepharose 4B affinity chromatography and cleaved by PreScission Protease on column (Fig. 5). The enzyme activities of the cleaved recombinant proteins were assayed toward the substrates coniferaldehyde and sinapaldehyde. GH2 exhibited strong CAD and SAD activities, while gh2 completely lost the CAD and SAD activities (Fig. 6; Table II). The result shows the point mutation G185D completely abolished the enzyme activities of GH2 for substrates coniferaldehyde and sinapaldehyde. Furthermore, the result confirms that native GH2 also completely lost the enzyme activities, and additional CAD isoenzymes with CAD activity exist in different tissues of rice.

To further know the substrate specificity of GH2 in vitro, we study the kinetics of the purified recombinant GH2 toward the substrates coniferaldehyde and sinapaldehyde by HPLC analysis. The overall catalytic properties were revealed by $K_m$, $V_{max}$, $k_{cat}$ and $k_{cat}/K_m$ (calculated by dividing $k_{cat}$ by $K_m$) values (Table II) calculated on the basis of the Lineweaver-Burk plots (Fig. 7). The $K_m$ value of GH2 for sinapaldehyde is 4.7-fold higher than that for coniferaldehyde, showing GH2 has higher affinity for coniferaldehyde than sinapaldehyde. The $V_{max}$ value of GH2 for sinapaldehyde is 2.6-fold higher than that for coniferaldehyde, indicating that the GH2 exhibits higher catalytic ability for coniferaldehyde than sinapaldehyde. Based on the above, the specific constant $K_{cat}/K_m$ for coniferaldehyde is 1.8-fold higher than that for sinapaldehyde. The results suggest that coniferaldehyde and sinapaldehyde both are the favored substrates of GH2, but coniferaldehyde is the main substrate and sinapaldehyde is the secondary substrate.

Phloroglucinol and Mäule Staining of the Wild Type and gh2 Mutant

Phloroglucinol, a reagent traditionally used to detect cinnamyl aldehyde and lignin (Speer, 1987), appears to be more sensitive for cinnamyl aldehyde detection than other methods (Halpin et al., 1998). So the accumulated cinnamyl aldehydes in the internode sections, young roots, and young branch panicles were detected with phloroglucinol reagent. Obvious differences of staining patterns between wild type and gh2 were observed. In the wild-type plants, little phloroglucinol staining was evident, whereas in the mutant plants strong staining occurred in the internode sections, young branch panicles, and roots (Fig. 8, B, D, G, and H). When the fresh hand-cut internode sections

Figure 4. CAD activities of total proteins and GH2 expression pattern. A, CAD activity and SAD activity of the total proteins from different tissues of the wild-type and gh2 plants at the heading stage. The SE bars were obtained from triplicate independent measurements. *, Not detectable. B, Northern-blot analysis of GH2 expression in different tissues of the wild type at the heading stage. Ethidium bromide staining of the 28S rRNA band is shown as the loading control. For both assays, the tissues include panicles, hulls, leaf blades, midribs, sheaths, the first two internodes, and young roots. The leaves eliminated of midrib (blades), midribs, and sheaths were all derived from flag leaves.

Figure 5. Expression and purification of recombinant GH2 and gh2 proteins in *E. coli*. Lane 1, The molecular mass ladder; lanes 2, 3, and 4, the crude extracts from induced recombinant *E. coli* of pGEX6P-1, pGEXGH2, and pGEgh2; lanes 5 and 6, the purified glutathione S-transferase (GST)-GH2 and GST-gh2 (about 66 kD); lanes 7 and 8, the cleaved fusion proteins GH2 and gh2 (about 40 kD). The arrowheads point to the GST fusion proteins and the cleaved fusion proteins. Proteins were visualized by Coomassie Blue R-250 staining.
were observed under light microscopy, the reddish-brown coloration and phloroglucinol staining are found in the same cortex (Co) region and vascular bundle (VB) region (Fig. 8, D and E). The results suggest that accumulated hydroxyl-cinnamaldehydes deposited in the high-lignified tissues or regions of the gh2 plants due to the deficient GH2 protein. Conifer-aldehyde residues had been proven to be responsible for the reddish-brown color in woody internodes of caffeic acid O-methyltransferase (COMT) down-regulated transgenic plants (Tsai et al., 1998). Considering the same localization of the reddish-brown pigments and the hydroxyl-cinnamaldehydes, we speculate the phenotype of gh2 is caused by the polymerization of accumulated aldehydes into lignin.

The lignin composition of the wild-type and gh2 plants was also estimated by Mäule regent, which can stain G residues yellow and S residues red (Meyer et al., 1998). In the Mäule staining, the fresh hand-cut internode sections of wild type were stained reddish-yellow in the Co region and purely yellow in the VB region, suggesting that the rice Co region contains abundant G monomers and a small amount of S monomer while the VB region mainly contains G monomers (Fig. 8C). In the corresponding tissues of the gh2 plants, the Co region and VB region were stained black-yellow by Mäule reagent (Fig. 8F). The phenomenon showed that the lignin monomers of the Co and VB regions were both altered in the gh2 mutant for the deficient GH2 protein, but the reason the tissues were stained

Table II. Kinetics properties of the recombinant GH2 protein

<table>
<thead>
<tr>
<th>Substrate</th>
<th>(K_m) (\mu M)</th>
<th>(V_{max}) (\mu g \cdot m^{-1} \cdot s^{-1})</th>
<th>(K_{cat}) (s^{-1})</th>
<th>(K_{max}) (\mu g \cdot m^{-1} \cdot s^{-1})</th>
<th>Specific Activity (\mu g \cdot m^{-1} \cdot s^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coniferaldehyde</td>
<td>4.4 \pm 0.2</td>
<td>32.9 \pm 1.7</td>
<td>1.3 \pm 0.1</td>
<td>288,640</td>
<td>23.3 \pm 1.0</td>
</tr>
<tr>
<td>Sinapaldehyde</td>
<td>20.8 \pm 2.3</td>
<td>87.0 \pm 8.1</td>
<td>3.4 \pm 0.3</td>
<td>161,540</td>
<td>59.5 \pm 0.7</td>
</tr>
</tbody>
</table>

\(^{a}\)Values are means \(\pm\) se for three independent assays.
black-yellow is unknown. The lignin properties of the gh2 mutants will be studied to explain the phenomenon.

**gh2 Plants Have Altered Lignin Content and Composition**

To further detect the lignin deficiency of the gh2 mutant plants, lignin content and composition were estimated in the wild-type and gh2 plants. Cell wall residue (CWR) was prepared from the dried mature hulls and internodes. Klason lignin analysis of the CWR showed that the total lignin content of the gh2 plants was reduced by 5% to 6% compared with the wild-type plants (Table III). The result demonstrates that GH2 is a critical enzyme that regulates lignin content in rice. Lignin composition of the wild-type and gh2 plants was also detected in the hulls and internodes by the derivatization followed by reductive cleavage (DFRC) method. The results show rice lignin consists of the three kinds of basal monomers, the H, G, and S monomers (Table III). In wild-type rice, G monomer is the major monomer and the G:S ratio is 60.66 in the hulls and 6.05 in the internodes, which is consistent with the Mäule staining (Fig. 8C). In the gh2, the three monomers are all reduced by approximately 35% in the hulls and approximately 25% in the internodes compared with the wild type. Although the reduction ratios are different in different tissues (hull and internode), the reductions of H, G, and S monomers in the same tissue are almost similar in degree. The results show that the GH2 enzyme takes the role of synthesizing H, G, and S monomer precursors in the hulls and internodes. The unequal reductions in hulls and internodes may be caused by the different lignification pattern between the hulls and internodes. Such analogous differences have been observed among Arabidopsis floral stems, roots, and hypocotyls (Sibout et al., 2003, 2005). Compared with the dramatic reductions in lignin monomers (25%–35%), the total lignin content reductions are quite low (only 5%–6%; Table III). The difference shows there are other substrates besides the three basal monomers participating in lignin biosynthesis.

**DISCUSSION**

**The gh2 Plant Is Defective in Lignin Biosynthesis**

ghan mutants have long been used as morphological markers and classified into coloration mutant categories (Zeng et al., 2003; Kurata et al., 2005). We found the GH2 gene encodes a CAD by map-based cloning, which shows the gh2 mutant is a lignin-deficient mutant. Maize bm lignin mutants have been used as...
a model to study crop lignin digestibility. Following the first bm mutant documented in maize in 1924 (Barrière et al., 2004), bm phenotype has also been identified in sorghum (Sorghum bicolor) and millet (Panicum miliaceum), which are all diploid C4 plants (Cherney et al., 1991). However, no corresponding bm mutant has been found in rice, to our knowledge. In this study, the gh2 mutant has been proven to be the so-called rice bm mutant.

Comparing the maize CAD mutant bm1 with the rice gh2 mutant in our study, the phenotype difference between bm1 and gh2 is that maize bm1 mutant plants have reddish-brown pigments in the leaf midrib and stalk pith, whereas rice gh2 mutant plants have the pigments in the hull, internode, and basal leaf sheath. In truth, the reddish-brown pigments in the two mutants are all associated with the high-lignified tissues. When the expression pattern of CAD genes was investigated, maize CAD gene was expressed strongly in the two mutants are all associated with the high-lignified tissues. When the expression pattern of CAD genes was investigated, maize CAD gene was expressed strongly in the two mutants but GH2 gene was slightly expressed in the same tissue (Fig. 4B). So it is likely that the phenotype difference may be caused by the expression pattern or loss-of-function extent of the CAD genes between bm1 and gh2. The identification of the gh phenotype provides an important clue for finding lignin mutants in grass crops such as wheat (Triticum aestivum) and barley (Hordeum vulgare).

Rice is the first crop plant whose genome has been sequenced (International Rice Genome Sequencing Project, 2005). Many characterized genes in maize have homologs in rice (Gaut and Doebley, 1997; Devos and Gale, 2000). With the advantages of a fully sequenced genome and an effective transformation system, rice seems to be the best model in the grass family to study lignin biosynthesis. Therefore, cloning and characterization of the other rice GH genes will be critical for us to get a deep understanding of lignin biosynthesis in grasses.

**GH2 Is a Multifunctional Enzyme**

We conclude GH2 is the primary CAD that acts as a multifunctional CAD enzyme to synthesize coniferyl alcohol and sinapyl alcohol precursors in lignified tissues of rice for the following reasons. First, in the rice CAD family, GH2 is the only one CAD belonging to the subA subfamily, which is widely proven to participate in constitutive lignin biosynthesis (Fig. 3). Second, when the GH2 gene was mutated, the CAD and SAD activities of both the native and purified recombinant GH2 protein were drastically reduced relative to the wild type (Figs. 4A and 6). Third, the kinetics properties of recombinant GH2 on the substrates of coniferaldehyde and sinapaldehyde show GH2 exhibits strong catalytic ability toward coniferaldehyde and sinapaldehyde (Table II). Finally, the G monomers and S monomers were reduced to a similar degree in the gh2 mutant plants compared with the wild type (Table III).

The difference between gymnosperm and angiosperm lignin is that gymnosperm plants contain mainly G monomers and monogenic CAD gene, whereas angiosperm plants consist of G monomers, S monomers, and one CAD family (Galliano et al., 1993; MacKay et al., 1997; Baucher et al., 1999). Consistent with this, CAD enzymes extracted from gymnosperm plants are quite specific for the reduction of coniferaldehyde (Kutsuki et al., 1982; Galliano et al., 1993), whereas CAD enzymes extracted from the gymnosperm plants have a significant affinity for both coniferaldehyde and sinapaldehyde (Mansell et al., 1974; Hawkins and Boudet, 1994; Brill et al., 1999). The relations between lignin composition and CAD enzyme activities in different plants have led to a hypothesis that a specialized CAD protein is involved in either sinapyl alcohol or coniferyl alcohol biosynthesis (Hawkins and Boudet, 1994). This hypothesis is supported by the identification of aspen PtCAD and PtSAD, which can specifically synthesize coniferyl alcohol and sinapyl alcohol, respectively (Li et al., 2001). On the contrary, Sibout et al. (2005) have recently proven that AtCAD2 and AtCAD3 are the primary CAD genes participating in lignin biosynthesis of the floral stems in Arabidopsis and speculated that no specific SAD is needed to synthesize sinapyl alcohol. Furthermore, aspen PtCAD cDNA has been found to be able to complement the cad-c cad-d double mutant (Sibout et al., 2005), completely restoring the level of conventional G and S units, while PtSAD cDNA only partially complemented the S monomer reduction. In the study, GH2 protein has been proven to have strong CAD activity and SAD activity (Fig. 6; Table II), and no additional SAD with obvious SAD activity was detectable in the highly lignified tissues, such as hull, sheath, and internode (Fig. 4A).

### Table III. Klasson and DFRC analysis of the wild type and gh2 mutanta,b

<table>
<thead>
<tr>
<th>Lignin Content</th>
<th>Lignin Monomer Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>% CWR</td>
<td>Total H Monomers</td>
</tr>
<tr>
<td>Wild-type hull</td>
<td>34.6 ± 0.1</td>
</tr>
<tr>
<td>gh2 hull</td>
<td>32.5 ± 0.1</td>
</tr>
<tr>
<td>Wild-type internode</td>
<td>15.0 ± 0.1</td>
</tr>
<tr>
<td>gh2 internode</td>
<td>14.2 ± 0.1</td>
</tr>
</tbody>
</table>

aThe data correspond to the means and ses from triplicate analyses.
bKlasson lignin content (percentage of weight of the extract-free dried cell wall from the hull and internode).

---

Zhang et al.
Our results suggest GH2 exhibits multifunctional activities and no additional specific SAD exists in rice highly lignified tissues. The available data lead to a paradox that questions whether there is a specific SAD in angiosperm plants. We think the conflicting data may be interpreted by elucidating the source of the S monomers. The proportion of S monomers in the total lignin monomers changes greatly among various kinds of species (Sederoff et al., 1999). In detail, S monomers exist in the wood angiosperm plants with high proportion, in the herbaceous angiosperm plants with relatively low proportion, and cannot be detected in the gymnosperm plants (Baucher et al., 1996; Campbell and Sederoff, 1996; Sibout et al., 2005). At present, F5H, COMT, and SAD (including the SAD activity of CAD) are regarded as the three factors that can affect the biosynthesis of S monomers (Boerjan et al., 2003). The alterations of the expression pattern of the three kinds of genes in plants have been found to be able to influence the proportion of S lignin monomers in lignin (Vignols et al., 1995; Meyer et al., 1998; Piquemal et al., 2002). Among different species, especially between wood angiosperm and herbaceous angiosperm, the different proportions of S monomers in total monomers may be caused by the diversities of the three kinds of genes, including gene copy members, expression patterns, or protein biochemical properties. So in some species such as aspen, it is also possible for some subfamily members to gain SAD function to act in an assistant role in maintaining enough sinapyl alcohol concentration for lignin polymerization in specific development phase or tissue. Although GH2 has been proven to be the primary CAD gene in rice lignin biosynthesis, the reason rice lignin consists of a low proportion of S monomers remains unknown. Maize is a grass plant that has a relatively high proportion of S monomers (Halpin et al., 1998). Further study on the three factors of F5H, COMT, and CAD genes in rice and maize, which is under way, will be very useful for understanding the mechanism of low S monomer composition in rice.

Potential Application of the gh2 Mutant in Agriculture

Rice is an important crop that feeds more than half of the world’s population. Effective utilization of rice by-products will definitely increase the income of rice producers. Rice residue has the potential to be used for bio-oil extraction by pyrolysis, for mushroom growth, and even for dietary fiber preparation (Kong et al., 2005). However, the most important applications are for livestock feeding, pulp, and paper making (Reddy and Yang, 2005). Lignin content and composition are the most important traits that limit the above applications. Point mutation of gh2 has been proven to affect lignin content and composition without influencing plant development under normal cultivar conditions, and further study on the mutant may improve the qualities of the rice residues. The molecular cloning of other gh mutants will help to discover new genes controlling lignin content and composition. Just like the wide application of maize bnm mutants in agricultural production, cloning of GH genes and their application in rice will arouse many agricultural interests in the future.

MATERIALS AND METHODS

Materials

The original rice (Oryza sativa) gh2 mutant (RGS no. 6) was kindly provided by Dr. Khush at the International Rice Research Institute, The Philippines. The gh2 gene was transferred to Zhefu802 (opp. indica) by 10 rounds of backcrosses with Zhefu802 (Zeng et al., 2003). We obtained the isogenic line with the gh2 gene and also named it as gh2, while Zhefu802 represents the wild type. The gh2 was crossed with a japonica rice variety, Nipponbare, to construct the F2 mapping population.

Genetic Analysis and Marker Development

The primary locus of GH2 locus was determined by newly developed marker P1 and reported STS marker S21668. To fine map the gh2, seven STS markers and two CAPS markers (Table I) were developed on comparisons of original or CAPS length between indica var. 9311 and japonica var. Nipponbare according to the data published in http://www.ncbi.nlm.nih.gov.

Complementation Test

A 8.44-kb genomic DNA fragment containing the entire GH2 coding region, the 2,397-bp upstream sequence, and the 2,460-bp downstream sequence was inserted into the binary vector pCAMBIA1300 to generate the transformation plasmid pCGH for complementation test. A control plasmid, pCGHS, containing truncated GH2 gene also was constructed according to the strategy described previously (Li et al., 2005a). The two binary plasmids were introduced into Agrobacterium tumefaciens EHA105 by electroporation, and transformed to rice for complementation testing according to a published method (Hei et al., 1994).

Phylogenetic Analyses

The full-length CAD protein sequences were retrieved from GenBank and used for phylogenetic analyses according to the methods described by Li et al. (2003b). Multiple sequence alignments were conducted using ClustalX version 8.0 with the PAM matrix. A neighbor-joining tree was built using MEGA version 2.1 adopting Poisson correction distance and was presented using radiation TreeView. Support for the tree obtained was assessed using the bootstrap method with 1,000 replicates. The Arabidopsis (Arabidopsis thaliana) CAD family was named as described by Kim et al. (2004) and Sibout et al. (2005). The rice CAD family was named as described by Tobias and Chow (2005) except that OsCAD2 was renamed as GH2.

RNA Extraction and Gel Blot

Total RNA was extracted from tissues of wild-type plants at the heading stage with TRIZOL reagent (Invitrogen). Twenty micrograms of total RNA per lane was electrophoresed on 1% agarose-formaldehyde gels and capillary blotted onto Hybond-N* membranes (Amersham). The 217-bp 3′-terminal untranscriptional fragment of the GH2 gene was labeled with α-32P-dCTP using the primer-a-gene labeling system (Promega) as the specific probe, and ethidium bromide staining of the 28S ribosomal RNA band was used as the loading control. The membrane was incubated in standard prehybridization solution at 65°C for 6 h and then hybridized with the α-32P-labeled probe at 65°C for 16 h. Following hybridization and sequential washing, the radioactive membrane was then exposed to x-ray film.

Plant Protein Extraction and Enzyme Assay

Plant total proteins were extracted as the method described by Halpin et al. (1998) with a little modification. Rice fresh tissues (about 300 mg) at the
opening step was separately milled to fine powder in liquid N2, then extracted with 1,000 μL extraction buffer (100 mM Tris-HCl, pH 7.5; 2% polyethylene glycol 6000, 5 mM dithiothreitol, 2% polyvinylpyrrolidone) for 2.5 h at 4°C. The supernatant was decanted and about 100 μg total proteins were used for CAD activity and SAD activity assay. Assays of crude enzyme were carried out using coniferyl alcohol and sinapyl alcohol as the substrates, and the formation of hydroxy-cinnamaldehydes was monitored spectrophotometrically (TU-1201, GENERAL) at 400 nm using the following molar extinction coefficients: coniferaldehyde 2.10 × 10^4 M^−1 cm^−1 and sinapaldehyde 1.68 × 10^4 M^−1 cm^−1, as described by Sibout et al. (2003). Assays were carried out at 30°C for 10 min in 500 μL of 100 mM Tris-HCl, pH 8.8, 100 μM NADP, and 250 μL of coniferyl alcohol or sinapyl alcohol. The enzyme reactions were initiated by enzyme addition and stopped by holding at 85°C for 10 min. An assay without NADP was used as the control. Protein concentration was determined using the Bradford reagent. Resulting units are defined as the amount of activity that converts 1 nmol of hydroxy-cinnamyl alcohol into the corresponding aldehyde per second (1 nKatal) per milligram of crude protein extract.

**GH2 and gh2 Expression in Escherichia coli**

Both GH2 and gh2 full-length cDNAs were isolated by reverse transcriptase-PCR from the total RNA from gh2 and wild-type plants roots with RT-PCR using the respective primers 1 (5’-CTCTCTGGTTGCTTTGT-3’) and primer 2 (5’-GCCGGATATTATCCGGGAAG-3’). The PCR products were inserted into pGEM-T vectors and sequenced to obtain the correct clones pGEMGH2 and pGEMgh2. The ORF sequences of GH2 and gh2 were amplified from cDNA clones pGEMGH2 and pGEMgh2 using primer 3 (5’-CGCGG-GATCCATGGGGAGCTCGCCGGCGCA-3’) and primer 4 (5’-CGGGAATTC-TAGGGCCATCTGCTGCGGCT-3’). The primers incorporated a BamHI site at the N-terminal end and an EcoRI site at the C-terminal end of the ORF. The PCR products were digested and cloned into the corresponding site of pGEX-6P-1 to generate pGEXGH2 and pGEXgh2, sequenced to assure the ORF sequence, and then introduced into E. coli BL21 for protein expression. The bacteria containing pGEXGH2 or pGEXgh2 were cultured overnight at 37°C, diluted to 400 times, and grown at 37°C for about 3 h to an OD600 of 0.6, then transferred to 16°C for 16 h after adding isopropyl-β-D-thiogalactopyranoside (IPTG). The cells were collected and lysed by moderate ultrasonic. The recombinant GH2 and gh2 protein were purified by glutathione-Sepharose 4B affinity chromatography and cleaved by PreScission Protease on a column according to the manufacturer’s instructions (Pharmacia Biotech). The purified proteins were quantified by the bicinchoninic acid method and used for the enzyme kinetics analysis.

**HPLC Analysis of Enzyme Functions and Reaction Kinetics**

For the substrate-specific activity test, the enzyme reaction mixture contained 100 mM Tris-HCl, pH 8.8, buffer, 200 μM NADPH, 200 μM aldehyde substrate in a final volume of 250 μL, and 1.2 μg of purified recombinant GH2 or gh2 protein were used. The reactions were for 10 min at 30°C and stopped by adding glacial acetic acid (10 μL). For the kinetic analyses, the enzyme reaction mixture contained 100 mM Tris-HCl, pH 8.8, buffer, 200 μM NADPH, 0.5 μM purified recombinant protein, and varying concentrations (10-200 μM) of coniferaldehyde (Aldrich, High Purity) or sinapaldehyde (Apin Chemicals, High Purity) in a final volume of 250 μL. Enzymatic reactions were initiated by cinnamyl alcohol addition and, after 4 min incubation at 30°C, were stopped by adding isopropyl-β-D-thiogalactopyranoside (IPTG). The reactions were carried out with NADPH-free mixture as the control. An aliquot (50 μL) of each assay mixture was subjected to reverse-phase HPLC analysis with UV detection for the product identification. Authentic coniferyl alcohol (Aldrich, High Purity) mixture was subjected to reverse-phase HPLC analysis with UV detection for authentic coniferyl alcohol (Aldrich, High Purity) in a final volume of 250 μL. Molar extinction coefficients: coniferaldehyde 2.10 × 10^4 M^−1 cm^−1 and sinapaldehyde 1.68 × 10^4 M^−1 cm^−1, as described by Sibout et al. (2003). Assays were carried out at 30°C for 10 min in 500 μL of 100 mM Tris-HCl, pH 8.8, 100 μM NADP, and 250 μL of coniferyl alcohol or sinapyl alcohol. The enzyme reactions were initiated by enzyme addition and stopped by holding at 85°C for 10 min. An assay without NADP was used as the control. Protein concentration was determined using the Bradford reagent. Resulting units are defined as the amount of activity that converts 1 nmol of hydroxy-cinnamyl alcohol into the corresponding aldehyde per second (1 nKatal) per milligram of crude protein extract.

**Lignin Analysis**

**Histochemical Staining**

For histochemical localization of the accumulated cinnamyl aldehydes, phloroglucinol staining was performed according to a standard protocol (Halpin et al., 1998) with a little modification. Fresh hand-cut sections from rice second internode, young roots, and branch panicles at the heading stage were incubated for 10 min in phloroglucinol solution (1% in ethanol:water [70:30]), the phloroglucinol was removed and treated with 18% HCl for 5 min, then photographed using a color digital camera directly or under light microscope (Olympus BX51). Mâle staining was according to Chapple’s method (Chapple et al., 1992) with a few modifications. Fresh hand-cut sections of the second internode were treated with 0.5% KMnO4 for 10 min, rinsed with water, treated with 10% HCl for 2 min, rinsed with water again, then mounted in concentrated H3PO4 and examined immediately under light microscopy and photographed as above.

**Lignin Content and Composition Analysis**

Dried mature internodes and hulls were collected after the removal of leaves, leaf sheaths, and the seeds. The milled hull or internode power was sieved through a 0.5-mm screen, washed in phosphate buffer (50 mM, pH 7.2) three times, and extracted with 70% ethanol twice at 70°C for 1 h. The CWR was dried in a vacuum and used to measure lignin content and composition. Lignin content was quantified according to Li et al. (2003b). Lignin compositions were assayed by the DFRC protocol (Lu and Ralph, 1997).

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession number DQ234227 (GH2).

**ACKNOWLEDGMENTS**

We thank Dr. John Ralph (University of Wisconsin, Madison, WI) for supplying authentic standards for DFRC analysis, Dr. Khush (International Rice Research Institute, The Philippines) for providing the original gh2 mutant, Dr. Mu Zhang (Institute of Genetics and Developmetal Biology, Chinese Academy of Sciences) for assistance in gas chromatography-mass spectrometry analysis. We also thank the two anonymous reviewers and the editor for critical comments on the manuscript.

Received October 17, 2005; revised January 17, 2006; accepted January 17, 2006; published January 27, 2006.

**LITERATURE CITED**


Rice GOLD HULL AND INTERNODE2 Cloning


Downloaded from on June 23, 2017 - Published by www.plantphysiol.org
Copyright © 2006 American Society of Plant Biologists. All rights reserved.