GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Cloning and sequence analysis of the safflower betaine aldehyde dehydrogenase gene

Y.B. Wang[1]*, L.L. Guan[2]*, Y.W. Xu[1], H. Shen[1] and W. Wu[1]

[1]Agronomy College, Chengdu Campus, Sichuan Agricultural University, Wenjiang, China
[2]Tropical Crops Genetic Resources Institute, Chinese Academy of Tropical Agricultural Sciences, Haikou, China

*These authors contributed equally to this study.
Corresponding author: W. Wu
E-mail: ewuwei@sicau.edu.cn / ewuwei@gmail.com

Genet. Mol. Res. 13 (1): 344-353 (2014)
Received August 9, 2013
Accepted October 24, 2013
Published January 21, 2014
DOI http://dx.doi.org/10.4238/2014.January.21.2

ABSTRACT. In response to salinity or drought stress, many plants accumulate glycine betaine, which is a regulator of osmosis. In plants, the last step in betaine synthesis is catalyzed by betaine aldehyde dehydrogenase (BADH), a nuclear-encoded chloroplastic enzyme. Based on the conserved oligo amino acid residues of the published *BADH* genes from other higher plant species, a cDNA sequence, designated *CtBADH*, was isolated from safflower (*Carthamus tinctorius* L.) using a polymerase chain reaction approach. The clones were 1.7 kb on average, and contained an open reading frame predicting a polypeptide of 503 amino acids with 84.5% identity to that of *Helianthus annuus*. The deduced amino acid sequence showed a decapeptide, Val-Thr-Leu-Geu-Leu-Gly-Gly-Lys-Ser-Pro and Cys, which is essential for proper functioning of BADH. Phylogenetic analysis indicated that *CtBADH* grouped with other dicotyledonous plant *BADH* genes, and subgrouped in the composite family. Prediction of secondary structure and subcellular localization suggested that the

protein encoded by *CtBADH* contains 33 coils, 15 alpha helixes, and 21 beta strands, and most likely targets the chloroplast or mitochondria.

Key words: Betaine aldehyde dehydrogenase; Phylogenetic analysis; Safflower; Secondary structure; Subcellular localization

INTRODUCTION

Environmental stress, such as low temperature, high temperature, high salinity, and drought, restrict the distribution and productivity of plants. When subject to salt stress or drought, some vascular plants typically respond with an increased accumulation of glycine betaine, an important osmoprotectant that is produced in response to salt and other osmotic stresses. In plants, the last step in betaine synthesis is catalyzed by betaine aldehyde dehydrogenase (BADH), a nuclear-encoded chloroplastic enzyme. Overexpression of BADH results in increased tolerance of salt and osmotic stress in many organisms (Kishitani et al., 1994). The biosynthesis of betaine consists of two steps. Oxidation of choline is catalyzed by choline monooxygenase, and then betaine aldehyde transforms to glycine betaine via deoxygenation. Both steps were shown to be localized in chloroplasts in spinach, where the plastid BADH was identified as a dimeric protein of approximately 60 kDa.

To date, BADHs have been purified from several species. *BADH* has already been cloned from spinach (*Spinacia oleracea* L.), sugar beet (*Beta vulgaris* L.), barley (*Hordeum vulgare* L.), sorghum (*Sorghum bicolor*), rice (*Oryza sativa* L.), and mangrove (*Avicennia marina*) among others. Furthermore, some studies have reported that the salt tolerance of *BADH* transgenic plants is much higher than that of controls, with significantly increased BADH activity (Wu et al., 2008; Zhou et al., 2008). The promoter of the *BADH* gene was isolated from *Suaeda liaotungensis*, and analysis of the promoter sequence revealed the existence of several putative *cis*-elements, such as a TATA-box, a CAAT-box, a GC-motif, EIRE, MRE, WUN-motif, a heat shock element, and ABRE (Zhang et al., 2008). The clones averaged 1.7 kb in length, and contained an open reading frame (ORF) predicting a polypeptide of 503 amino acids, containing the decapeptide Val-Thr-Leu-Geu-Leu-Gly-Gly-Lys-Ser-Pro and Cys, which are related to the function of BADH.

Safflower (*Carthamus tinctorius* L.) is a shade-tolerant plant in the composite family, and is cultivated mainly for its seed, which is used as edible oil and as birdseed. It shows a high tolerance to drought, heat, and other stressors. It is a highly branched, herbaceous, thistle-like annual plant, usually containing long sharp spines on its leaves. Safflower has great stress tolerance, is adapted to a wide range of moisture conditions, and retains green color even in drought and cold seasons. It is widely distributed and cultivated around the world. Safflower production is highest in the USA, Mexico, Ethiopia, Argentina, and Australia. Although China also has significant safflower crops, the florets are harvested for use in traditional medicines and the crop is not reported internationally. Environmental stress tolerance is one of the most important factors limiting safflower growth and survival. Therefore, increasing environmental stress tolerance has been one of the major objectives in safflower breeding programs. When subjected to salt stress or drought, safflower responds with an increased accumulation of glycine betaine. Because there are no reports characterizing its *BADH* gene, we cloned and characterized the safflower *BADH*.

In this study, based on similarity of its primary structure, homologous primers were designed according to the highly conserved region of BADH genes of other species, and then the full-length cDNA sequence of *BADH* was isolated from safflower using reverse transcription-polymerase chain reaction (RT-PCR) and rapid amplification of cDNA ends (RACE). Sequence analyses were conducted with DNAMAN, DNAStar, Mega, and BLAST (http://www.ncbi.nlm.nih.gov/) softwares.

MATERIAL AND METHODS

Plant materials

Safflower Chuan Hong 1 (*C. tinctorius*) was grown in a natural environment at the experimental fields of Sichuan Agricultural University. Young leaves were harvested and snap frozen in liquid nitrogen and stored at -80°C until RNA extraction.

RNA extraction and RT reactions

Total RNA was isolated from young leaves using the Plant RNA Extraction Kit (Tiangen, Beijing). RNA concentration was determined spectrophotometrically, and verified by ethidium bromide staining on agarose gel. Total RNA was then treated with RNase-free DNaseI (TaKaRa) to remove any contaminating genomic DNA, and approximately 3 μg DNA was used as template for the first cDNA synthesis using TaKaRa reverse transcription reagents, following manufacturer instructions, and was stored at -20°C.

Isolation of partial cDNA clones

Clones were amplified using the cDNA described above as a template and PF1/PR1 primers designed from the *BADH* sequences conserved among other plant species. The primer sequences were: PF1 (5'-TCGACGGAGAATGGAGAGA-3') and PR1 (5'-CCATTAGTCCAGAAGCA-ACCA-3'). PCR was performed as follows: 95°C for 5 min followed by 35 cycles at 94°C for 30 s, 42°C for 30 s, and 72°C for 1 min, then 72°C for 7 min. One DNA fragment with the same expected size was generated, subcloned into the vector pMD19 -T (Promega), and sequenced.

RACE

The 5'-ends of BADH cDNAs were amplified using the SMART and RACE methods on the 5'-RACE System (BD-Clontech, Palo Alto, CA, USA) according to manufacturer instructions with two different antisense primers: PR2 (5'-TGGTTGCTTCTGGACTAATGGG-3'), PR3 (5'-CCATTAGTCCAGAAGCAACCAA-3'), PR4 (5'-GGTGGCGGGATTGATAACAG G-3'), and PR5 (5'-CTGTTATCAATCCCGCCACC-3'), which were deduced from the 5'-region specific to each of the above mentioned clones (Alam et al., 2010). Amplification of the 3'-end was performed by the 3'-RACE method as described by Schmidt et al. (2010). The primers used for 3'-RACE were the oligo(dT)17 adaptor primer (for first-strand cDNA synthesis) and the forward primer PF2, according to manufacturer instructions.

Based on the sequence information of 5'- and 3'-ends, the sense/antisense primer pairs, PF4 (5'-TCGACGGAGAATGGAGAGAA-3') and PR6 (5'-CTCCGACGACGGTTTCAGT-3'), were designed to amplify the ORF of BADH. PCR fragments were subcloned into the vector pMD19-T (Promega) and transformed into *Escherichia coli* JM109. Subsequently, the nucleotide sequence was determined by Invitrogen (Shanghai).

Sequence analyses

The nucleotide sequence from our cDNA clone and the deduced amino acid (aa) sequence were identified using the National Center for Biotechnology Information (NCBI) BLAST program (http://www.ncbi.nlm.nih.gov/BLAST/). Sequence alignment was conducted with the DNAMAN software (version 4.0, Lynnon Biosoft, Quebec, Canada). The nucleotide and deduced amino acid analyses were conducted using programs in the NCBI network. The cDNA sequence and deduced aa sequence, designated as *CtBADH*, has been submitted to the GenBank database with the assigned accession No. HQ439601.

RESULTS AND DISCUSSION

Nucleotide sequence analysis

Using the degenerate primers PF1/PR1, a distinct cDNA fragment was amplified with the expected size of approximately 600 bp, containing an ORF encoding a 200-aa protein. Alignment of the deduced aa sequences showed a high degree of identity between our cloned fragment and the central coding regions of known plant BADH sequences. The fragment was then used to design gene-specific primers for cloning the full-length *BADH* sequence. A 200-bp fragment of 5'-RACE and an 800-bp fragment of 3'-RACE were amplified, and both sequences were determined. BLAST searches of the deduced aa sequences revealed correspondence with known *BADH* genes. Sequence comparisons of the 5'- and 3'-ends with the central part of the gene indicated that the overlapping regions matched perfectly, suggesting that these sequences represent the missing parts of *BADH* genes. Based on this sequence data, specific primers were designed, and the full-length cDNAs were amplified, cloned, and sequenced, revealing 100% identities to the expected *BADH* sequence.

Our sequence analysis indicated that the cloned full-length cDNA (1746 bp) contains a 1509-bp ORF encoding 380-aa residues with an estimated molecular mass of 50.12 kDa. The coding region was flanked by a 144 bp 5'-untranslated region (UTR) of the mRNA along with a full 93-bp 3'-UTR (Figure 1).

Amino acid sequence analysis

The 1509-bp ORF encodes a 503-aa sequence. Using ProtParam of the ExPASy program (http://www.expasy.ch/), the safflower BADH protein was analyzed. The results showed that the safflower BADH protein contains 503 amino acids, comprising 64 negatively charged residues and 41 positively charged residues. The molecular weight was determined to be 54.489 kDa. When the safflower BADH aa sequence was compared with that of other plants (*Arabidopsis thaliana*, *Glycine max*, *Chrysanthemum lavandulifolium*, and *Helianthus annuus*), no significant differences were found. Alignment of the deduced CtBADH aa sequence

with other plant BADH polypeptides showed an identity of approximately 87% between the predicted amino acid sequences derived from the safflower *BADH* and other plants (*G. max*, *Gossypium hirsutum*, *H. annuus*, *Leymus chinensis*, *O. sativa*, *Populus trichocarpa*, *Zea mays*, and *C. tinctorius*) (Figure 2).

Figure 1. Nucleotide sequence of the *CtBADH* gene from safflower. The ORF of the sequence were shaded in gray, and the initiation codon and termination codon were indicated in blank.

The deduced aa sequence comprises a decapeptide, Val-Thr-Leu-Geu-Leu-Gly-Gly-Lys-Ser-Pro and Cys, which are essential for the function of BADH. Similar to some other plant *BADH* genes, the safflower BADH peptide lacks typical transit peptide sequences that are needed for entry into chloroplasts. Instead, it contains a C-terminal Ser-Lys-Leu (SKL) signal (Figure 2), which indicates targeting to peroxisomes (Reumann, 2004). Nakamura et al. (2001) reported that the SKL sequence was present in all monocotyledonous BADH sequences examined, whereas most of the dicotyledonous BADHs were of the chloroplast type. Our results appear to conflict with those of Nakamura et al. (2001), as safflowers are dicot plants. In 2009, Fitzgerald et al. investigated most of the plant BADH enzymes. The results showed that some monocot plant BADHs did not contain an SKL C-terminal sequence, such as *Triticum aestivum* BADH2 and *H. vulgare* BADH2. On the other hand, some dicot plant BADHs did contain the SKL signal, such as BADH proteins of *A. thaliana*, *S. oleracea*, and *B. vulgaris*.

```
Glycine max          ..MSIPI H Q  ID D KVPVLKN I II  S QHII D  AA K  VDL  A  KA LS NK      63
Gossypium hirsutum   ..MAVQV S Q  ID E REPILKE L TI  A EEII N  AA A  VEL  A  RR LS NK      63
Helianthus annuus    ..MAISI F Q  ID E REPVRKN I VI  A EEIV D  AA A  IDI  K  RR LK DG      63
Leymus chinensis     MASPAI. Q Q  ID E RAPALGR L VI  T EASI E  SG S  VDA  A  RA LK NR      64
Oryza sativa         ..MATAI Q Q  VA E RAPALGR L VV  A ESPI E  AG A  VDA  A  RE LK NR      63
Populus trichocarpa  ..MAIHL N Q  ID E RETVLKK I VI  A EQII D  AA A  VEI  E  KK FS NK      63
Zea mays             MASQAMV L Q  VD E RPPAQGR L VV  T EAHI E  AG A  VDA  A  RA LK NR      65
Carthmus tinctorius  ..MAMAI S L  ID D KEPVKKN I VI  A ELII D  AA A  VDI  E  RG LK NG      63

Glycine max          AD AS S SV  R     A IT K PELAK  AI C   LD  A  ID      FY DL  K     128
Gossypium hirsutum   KD AT P AV  K     A VT R TELAK  AI C   LD  V  IE      YY DL  G     128
Helianthus annuus    KE AS S AH  K     A VT K DMFAK  AI C   LD  A  MD      YN DL  A     128
Leymus chinensis     RD SR P AV  K     A MI R ADLAR  AL C   LD  A  MD      FF GH  A     129
Oryza sativa         RD AR P AV  K     A II R SELAR  TL C   LD  A  MD      YF DL  S     128
Populus trichocarpa  KD SS S AY  R     A IT R SELGK  AI S   LD  L  MD      YY DL  G     128
Zea mays             RD AR P AV  K     A VI R QELAK  AL C   YD  A  MD      YF DQ  A     130
Carthmus tinctorius  KE AS S AH  K     S IM K SELSK  AI C   LE  A  MD      YN DL  E     128

Glycine max           AQ KAH S  MDTF SYVLK  I V A         M      PA    A I    L  V     193
Gossypium hirsutum    AR KAP S  METF SYVLK  I V G         M      PS    A I    L  I     193
Helianthus annuus     AK NAP N  MDTF CHIIR  I V G         M      SA    A V    L  V     193
Leymus chinensis      KR NAA S  EN.F CHLKK  I V S         M      PA    T V    L  V     193
Oryza sativa          KR NAP S  MENL CYLRK  I V G         M      PA    T V    S  V     193
Populus trichocarpa   TK KAP S  METF SFVLK  L V A         L      PA    T I    L  V     193
Zea mays              KR NSP S  METF CHLRR  I V G         M      PA    A V    L  V     195
Carthmus tinctorius   RK NAS A  METF CHIIR  I . G         M      PA    A I    L  V     192

Glycine max              AEICKEVG  P V  ILT   PE    AA  D   I     SA  SKI TA  QLI P S   258
Gossypium hirsutum       AKVVKGSR  P V  ILA   PE    AS  D   I     SA  SKI AA  QMV P S   258
Helianthus annuus        GEVCREVG  P I  IVT   PE    AA  D   I     SA  SKI TA  QNV P T   258
Leymus chinensis         ADVCKEVG  P V  VVT   PE    SS  D   V     YA  QKI VA  PTV A T   258
Oryza sativa             ADVCKEVG  S V  IVT   SE    SS  G   V     YE  KKI AS  PMV P S   258
Populus trichocarpa      GEVCREVG  P V  ILT   TE    AS  H   V     TA  SRI AS  QMV P S   258
Zea mays                 ADICKEVG  P V  IVT   PD    SA  D   V     PE  KKI AA  PMV P T   260
Carthmus tinctorius      GEICREVG  P V  ILT   PE    AS  N   I     SV  SKV TA  QNV P T   257

Glycine max          L       I  ED . LD  A  TI                IV ES  TE LNRI K V N     322
Gossypium hirsutum   L       I  ED . LD  A  TA                IV EN  RE LDRL K T N     322
Helianthus annuus    L       V  DD . ID  V  AL                IL ES  KE LDKL K A N     322
Leymus chinensis     L       V  DD . ID  V  TL                LI KN  KE VDRM A S N     322
Oryza sativa         L       V  DD . VE  V  TL                IL KK  KE QERM A A N     322
Populus trichocarpa  M       I  ED . LD  A  TL                LV ES  SE LDKL K T K     322
Zea mays             L       V  DD . ID  V  TL                LV TK  KK NERM A A N     324
Carthmus tinctorius  L       V  DN . ID  A  AL                LV ES  EE LDKL M A N     321

Glycine max          I   L    R DPI  EG  E ILK ISN  SE   I T  S  E LK   FVE TV T VT      387
Gossypium hirsutum   I   F    R GPV  GG  E VLK IST  SE   I S  V  E LK   FVE TI T VT      387
Helianthus annuus    I   L    R GPV  AG  E VLK VET  RE   V F  K  Q LT   YME AI T VT      387
Leymus chinensis     V   L    R GPV  EG  E IKK ISN  SE   I T  V  K ME   FIE TI T IN      387
Oryza sativa         V   L    M GPV  EG  E IKQ VST  SQ   I T  V  K LE   YIE TI T VD      387
Populus trichocarpa  I   F    R GPL  GE  D ILK IAT  SE   I S  D  K LN   FVE TI I VT      387
Zea mays             V   L    R GPV  EG  E IKK ILN  SE   I T  V  A LE   FIE TI T IT      389
Carthmus tinctorius  I   L    R GAV  GA  E VLK IST  SE   I F  E  E LE   YIN TI T VT      386

Glycine max          Q   E          T ST E  ID     V   GS VI N L   E ITKAFK  IV I        452
Gossypium hirsutum   Q   E          T RT E  LE     H   GA VI N L   D VSKNLQ  IV V        452
Helianthus annuus    Q   D          T ST Q  IE     H   GS II N L   D VAKAFE  IV V        452
Leymus chinensis     E   E          E ST E  IE     H   AG VI G R   Q LAEEID  CI V        452
Oryza sativa         Q   E          E ST E  IE     H   AG VL G R   Q LTEEID  II V        452
Populus trichocarpa  Q   E          T ST D  ID     H   GA VI N P   D VAKAFR  IV I        452
Zea mays             E   E          E ST D  IE     Q   AG VI G R   Q LSEEID  II V        454
Carthmus tinctorius  Q   D          T SA Q  IE     H   GS II N L   D VAKAFD  IV V        451

Glycine max            T      I           W LD   V    Q I DEP    QS SRL.                503
Gossypium hirsutum     C      N           W LD   V    Q V DEP    RS SKL.                503
Helianthus annuus      S      K           W LD   V    R I NEP    TP SKL.                503
Leymus chinensis       C      N           G ID   I    E T DAP    KA AN..                502
Oryza sativa           C      N           G ID   V    E A DEP    KS SKL.                503
Populus trichocarpa    C      I           W LE   V    R I EEP    QA SKL.                503
Zea mays               C      N           G ID   V    E I DEP    RS SKL.                505
Carthmus tinctorius    C      K           W LD   V    R I NDA    TP SSKL                503
```

Figure 2. Alignment of predicted amino acid sequences of safflower *CtBADH* and other plant *BADH* polypeptides. Sequences used for the analysis were obtained from GenBank. Latin names and accession numbers were as follows: *Glycine max* adn03184, *Gossypium hirsutum* ay461804, *Helianthus annuus* ACU65243, *Leymus chinensis* bad86758, *Oryza sativa* abi84118, *Populus trichocarpa* xm 002318594, *Zea mays* np001105781. C-terminal signal of *CtBADH* aa sequence was indicated in blank.

Evolutionary relationship of *BADH* genes

To elucidate the phylogenetic relationships of the safflower BADH protein, the deduced aa sequence was aligned with those of other plant species and a neighbor-joining tree was constructed. Sequences were selected to cover most of the plant species available and to avoid redundancies. As shown in Figure 3, our results revealed that except for *Selaginella moellendorffii*, these plant BADH proteins were classified into two major branches: monocotyledonous plants and dicotyledonous plants. Safflower BADH was located in the branch of dicotyledonous plants, and was subgrouped in the composite families (*C. lavandulifolium* and *H. annuus*). This indicated that our sequence isolated from safflower might have a similar enzyme function to sunflower *BADH* genes.

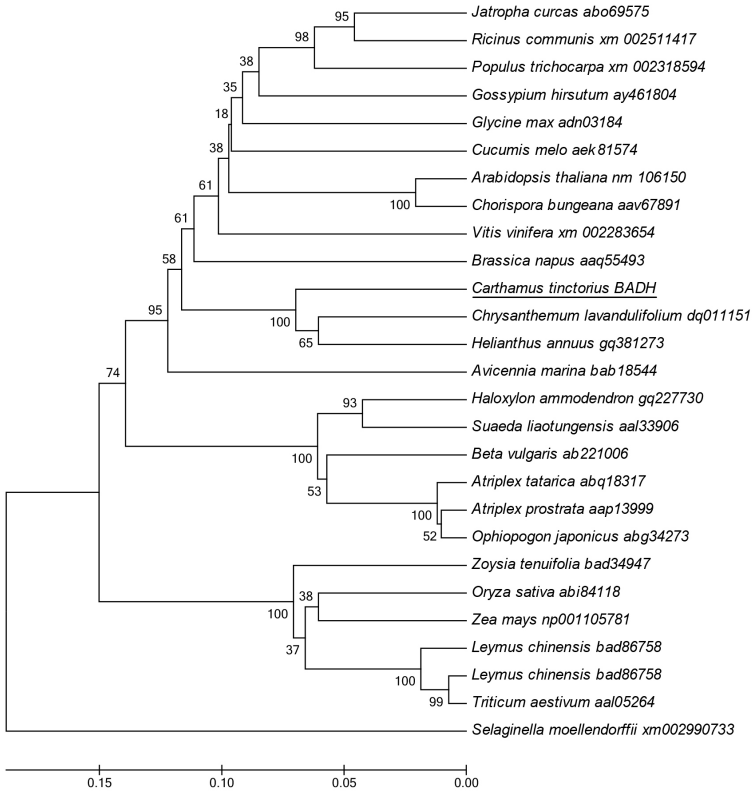

Figure 3. Phylogenetic tree of the *BADH* protein from safflower and other plant species. Sequences used for the analysis were obtained from GenBank. Latin names and accession numbers were as follows: *Arabidopsis thaliana*, nm 106150; *Atriplex prostrate*, aap13999; *Atriplex tatarica*, abq18317; *Avicennia marina*, bab18544; *Beta vulgaris*, ab221006; *Brassica napus*, aaq55493; *Chorispora bungeana*, aav67891; *Chrysanthemum lavandulifolium*, dq011151; *Cucumis melo*, aek81574; *Glycine max*, adn03184; *Gossypium hirsutum*, ay461804; *Haloxylon ammodendron*, gq227730; *Helianthus annuus*, gq381273; *Hordeum vulgare*, bab62846; *Jatropha curcas*, abo69575; *Leymus chinensis*, bad86758; *Ophiopogon japonicas*, abg34273; *Oryza sativa*, abi84118; *Populus trichocarpa*, xm 002318594; *Ricinus communis*, xm 002511417; *Selaginella moellendorffii*, xm_002990733; *Suaeda liaotungensis*, aal33906; *Triticum aestivum*, aal05264; *Vitis vinifera*, xm 002283654; *Zea mays*, np 001105781; *Zoysia tenuifolia*, bad34947.

Prediction of protein secondary and senior structures

Prediction of secondary structure of the deduced amino acids of *BADH* was conducted with the DNAstar software (Figure 4). Results showed that the safflower BADH protein mainly contains alpha helixes and beta strands, which made up 80% of the amino acids of the sequence. Analysis of the hydrophilic plot of our aa sequence showed that the hydrophobic areas were mainly in the two ends (N-terminal and C-terminal) of the sequence, and the hydrophilic regions were mainly in the middle of safflower BADH.

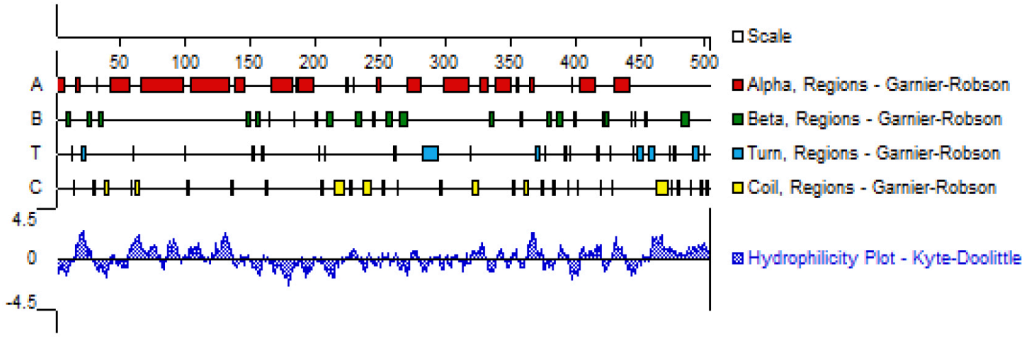

Figure 4. Prediction of secondary structure of *CtBADH* deduced amino acid sequence. Barrel structures are noted above the amino acid sequence.

Using the Swiss-model online service (Benkert et al., 2011) (http://swissmodel.expasy.org/workspace/index.php), the three-dimensional structure of the safflower BADH protein was predicted (Figure 5). The suggested structure was composed of 33 coils, 15 alpha helixes, and 21 beta strands. Alpha helixes and beta strands appeared alternately; beta strands were concentrated in the central region, and were surrounded by the external alpha helixes (Arnold et al., 2006).

Subcellular localization of the obtained protein

BADH isozymes have been found to target different subcellular compartments: chloroplasts, peroxisomes, or cytosol. Spinach BADH isozymes were found to consist of a major isozyme targeted to the chloroplast and a minor cytosolic isozyme (Weigel et al., 1986). In rice, both BADH1 and BADH2 have been shown to target to peroxisomes (Nakamura et al., 1997; Shirasawa et al., 2006). In barley, BADH1 appears to reside in the cytosol (Nakamura et al., 2001). The mangrove, *A. marina*, possesses a BADH isozyme that is targeted to the chloroplasts, and another that is targeted to peroxisomes (Hamilton III and Heckathorn, 2001). To determine the subcellular localization of our CtBADH protein, the online targetP tool (http://www.cbs.dtu.dk/services/) was used. Results showed that CtBADH most likely targets chloroplasts or mitochondria (Figure 6).

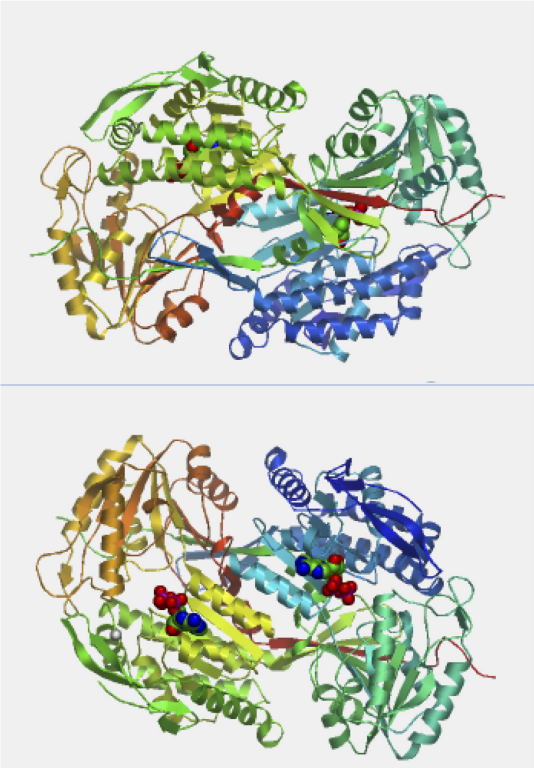

Figure 5. Prediction of third structure of the *BADH* deduced amino acid sequence.

```
### targetp v1.1 prediction results ###################################
Number of query sequences:  1
Cleavage site predictions not included.
Using PLANT networks.

Name                  Len     cTP    mTP     SP  other  Loc  RC
-----------------------------------------------------------------------
Sequence              503   0.338  0.048  0.166  0.738    _    4
-----------------------------------------------------------------------
cutoff                      0.000  0.000  0.000  0.000
```

Figure 6. Subcellular localization of *BADH* deduced amino acid sequence.

ACKNOWLEDGMENTS

Research supported by the Youth Science and Technology Foundation of Sichuan Province in China. The experiments were conducted in the Laboratory of Gene Resource and Molecular Breeding of Sichuan Agricultural University.

REFERENCES

Alam MM, Sharmin S, Nabi Z, Mondal SI, et al. (2010). A putative leucine-rich repeat receptor-like kinase of jute involved in stress response. *Plant Mol. Biol. Rep*. 28: 394-402.

Arnold K, Bordoli L, Kopp J and Schwede T (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22: 195-201.

Benkert P, Biasini M and Schwede T (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 27: 343-350.

Fitzgerald TL, Waters DLE and Henry RJ (2009). Betaine aldehyde dehydrogenase in plants. *Plant Biol*. 11: 119-130.

Hamilton EW III and Heckathorn SA (2001). Mitochondrial adaptations to NaCl. Complex I is protected by anti-oxidants and small heat shock proteins, whereas complex II is protected by proline and betaine. *Plant Physiol.* 126: 1266-1274.

Kishitani S, Watanabe K, Yasuda S, Arakawa K, et al. (1994). Accumulation of glycinebetaine during cold acclimation and freezing tolerance in leaves of winter and spring barley plants. *Plant Cell Environ.* 17: 89-95.

Nakamura T, Yokota S, Muramoto Y, Tsutsui K, et al. (1997). Expression of a betaine aldehyde dehydrogenase gene in rice, a glycinebetaine nonaccumulator, and possible localization of its protein in peroxisomes. *Plant J*. 11: 1115-1120.

Nakamura T, Nomura M, Mori H, Jagendorf AT, et al. (2001). An isozyme of betaine aldehyde dehydrogenase in barley. *Plant Cell Physiol*. 42: 1088-1092.

Reumann S (2004). Specification of the peroxisome targeting signals type 1 and type 2 of plant peroxisomes by bioinformatics analyses. *Plant Physiol*. 135: 783-800.

Schmidt T, Hillebrand A, Wurbs D, Wahler D, et al. (2010). Molecular cloning and characterization of rubber biosynthetic genes from *Taraxacum koksaghyz*. *Plant Mol. Biol. Rep*. 28: 227-284.

Shirasawa K, Takabe T, Takabe T and Kishitani S (2006). Accumulation of glycinebetaine in rice plants that overexpress choline monooxygenase from spinach and evaluation of their tolerance to abiotic stress. *Ann. Bot*. 98: 565-571.

Weigel P, Weretilnyk EA and Hanson AD (1986). Betaine aldehyde oxidation by spinach chloroplasts. *Plant Physiol.* 82: 753-759.

Wu W, Su Q, Xia XY, Wang Y, et al. (2008). The *Suaeda liaotungensis* kitag betaine aldehyde dehydrogenase gene improves salt tolerance of transgenic maize mediated with minimum linear length of DNA fragment. *Euphytica* 159: 17-25.

Zhang Y, Yin H, Li D, Zhu W, et al. (2008). Functional analysis of *BADH* gene promoter from *Suaeda liaotungensis* K. *Plant Cell Rep.* 27: 585-592.

Zhou S, Chen X, Zhang X and Li Y (2008). Improved salt tolerance in tobacco plants by co-transformation of a betaine synthesis gene *BADH* and a vacuolar Na^+/H^+ antiporter gene *SeNHX1*. *Biotechnol. Lett.* 30: 369-376.