J. Pestic. Sci. 29, 1–5 (2004)

Original Article

Secondary Structure Prediction of Acetolactate Synthase Protein in Sulfonylurea Herbicide Resistant *Limnophila sessiliflora*

Ying LIN,[#] Guang-Xi WANG,[*,†,††,#] Wei LI,[††] and Misako ITO[†]

Department of Biotechnology, South China University of Technology, Guangdong 510640, People's Republic of China
[†] *Graduate School of Agriculture, Kyoto University, Kyoto 606–8502, Japan*
[††] *Wuhan Institute of Botany, Chinese Academy of Sciences, Hubei 430074, People's Republic of China*

(Received July 18, 2003; Accepted October 14, 2003)

Here, we cloned and sequenced fragments encoding ALS genes from biotypes of *Limnophila sessiliflora* susceptible and resistant to SU-herbicide. Comparisons of deduced amino acid sequences and secondary protein structures were implemented. The results showed that Pro was substituted with Gln and a G-X-X-P motif was abolished in Domain A of resistant ALS, and the secondary protein structure was converted from extended strands to an α-helix around Domain A by the Pro mutation. We suggest that resistance to SU-herbicides is caused by the mutation of ALS and conformational change of its secondary structure.

Keywords: herbicide resistance, *Limnophila sessiliflora*, secondary structure, sulfonylurea herbicide.

INTRODUCTION

Acetolactate synthase (ALS, EC 4.1.3.18) is the first common enzyme in the biosynthesis of the branched chain amino acids valine, leucine and isoleucine. Several classes of herbicides are known to inhibit ALS, such as the sulfonylureas, imidazolinones, triazolopyrimidine sulfonanilides and pyrimidinyl oxybenzoates, by binding to a relic quinone-binding site.[1)] These highly selective ALS-inhibiting herbicides are indicative of their importance for weed management in a wide range of crops worldwide. However, biotypes resistant to the ALS-inhibiting herbicides have been reported in at least 20 monocotyledonous and 44 dicotyledonous plant species.[2)] Resistant (R) biotypes in many cases have modified ALS genes with one or more point mutations causing reduced sensitivity to the ALS-inhibiting herbicides.[3–6)] Sulfonylurea (SU) herbicides are commonly used for weed control in rice fields in Japan. Biotypes resistant to SU-herbicides have been found in several paddy weeds,[7–10)] including *Limnophila sessiliflora* Blume.[11)]

L. sessiliflora is a rooted, amphibious aquatic angiosperm with both submersed and emersed plant parts. It appears to be largely endemic to Asia and is documented as a major problem weed in paddy rice fields of India, China, Japan, and the Philippines.[11)] *L. sessiliflora* was confirmed to be resistant to SU-herbicides from herbicide dose-response relationships.[11)] In response to four SU-herbicides, bensulfuron-methyl (BSM) [methyl α-(4,6-dimethoxypyrimidin-2-ylcarbamoylsulfamoyl)-*o*-toluate], the first SU-herbicide used in Japan, pyrazosulfuron-ethyl (PSE) [ethyl 5-[[3-(4,6-dimethoxypyrimidin-2-yl)ureido]sulfonyl]-1*H*-1-methylpyrazole-4-carboxylate], imazosulfuron (IMS) [1-(2-chloroimidazo[1,2-α]pyridin-3-ylsulfonyl)-3-(4,6-dimethoxypyrimidin-2-yl)urea] and ethoxysulfuron (ETS) [3-(4,6-dimethoxypyrimidin-2-yl)-1-(2-ethoxyphenoxysulfonyl)urea], the GR_{50} (50% growth reduction) values for the resistant biotype were 300–900 times the GR_{50} for the susceptible (S) biotype.

The objective of the present study was to isolate the ALS genes from both R and S biotypes of *L. sessiliflora* and to compare the deduced amino acid sequences. A secondary objective was to explore the structural basis of SU-herbicide resistance by analyzing the secondary structure of ALS protein.

MATERIALS AND METHODS

1. Plant Materials

SU-resistant *L. sessiliflora* was collected from a population which had survived SU-herbicide treatments in rice fields of Sennan Village, Akita and SU-susceptible *L. sessiliflora* was collected from a population that had never been treated with SU-herbicide in marshes in Omagari, Akita, Japan. The former was confirmed to be resistant to SU-herbicides: BSM, PSE, IMS and ETS whereas the latter

* To whom correspondence should be addressed.
# These authors contributed equally to this work.

was susceptible to these herbicides.[11] Both populations were confirmed to be *L. sessiliflora* by Dr. T. Yamazaki, a plant taxonomist from the University of Tokyo.

2. DNA Extraction

Genomic DNA was extracted from leaf tissue of the resistant and susceptible biotypes of *L. sessiliflora* according to the method of Uchino and Watanabe.[7] Freshly collected leaves (approximately 0.05 g) were frozen in a 1.5 ml microtube and ground with a plastic pestle. After the addition of 300 μl of extraction buffer (100 mM tris-HCl, pH 8.0; 50 mM ethylenediaminetetraacetic acid (EDTA); 500 mM NaCl) and 50 μl of sodium dodecylsulfate, the extract was incubated at 65°C for 10 min. The solution was then mixed with 80 μl of 8 M potassium acetate, incubated on ice for 30 min and centrifuged at 10,000 $\times g$ for 20 min at 4° C. The supernatant was mixed with 0.6 volumes of 2-propanol, incubated for 5 min and centrifuged for 1 min at room temperature. The precipitate was rinsed with 70% ethanol, dried briefly and dissolved in 30 μl of TE (10 mM tris-HCl, pH 8.0; 1 mM EDTA).

3. Amplification and Analysis of DNA

Fragments of the *ALS* gene were amplified from the genomic DNA of the two biotypes using Taq polymerase (Stratagene, CA, USA) with 30 cycles of 30 sec at 94°C, 30 sec at 58°C, and 60 sec at 72°C. The primers used are common to both biotypes, having the sequences 5′-GCAACG TGCTGCCACGTCACGAGCAGGG and 5′-ACATGA GGCTGCTTGTTCTTTCCAATCTCAGC designed based on the *ALS* gene sequence of *Arabidopsis thaliana* (GenBank No. NM_114714), and available on request. Amplified fragments were subcloned using the TOPO TA vector (Invitrogen, CA, USA). Cloned fragments were sequenced with an ABI 310 Genetic Analyser (Perkin-Elmer, CT, USA).

4. Secondary Structure Prediction

Computer modeling for the structural prediction of the protein from both the R and S biotypes was implemented with the SSpro program[12] (http://www.igb.uci.edu/tools/scratch/), and consensus structures were identified with DPM,[13] GOR4,[14] HNNC,[15] PHD,[16] Predator,[17] SIMPA96,[18] SOPM[19] and NPS@ (Network Protein Sequence Analysis),[20] available online at http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_seccons.html

RESULTS AND DISCUSSION

The amino acid sequence of ALS derived from the DNA codes was organized in the FASTA format, and the file was inputted into BioEdit to carry out the alignment manually. Figure 1 shows the alignment of sequence for the ALS of *L. sessiliflora.* It reveals high homology (88.5%) with the ALS of *A. thaliana* (GenBank No. NP_190425) between amino acid residues 141–410, which cover the conserved Domain A and D.[5] We compared the nucleotide sequence and deduced amino acid sequence in the ALS fragments from both biotypes of *L. sessiliflora.* A single base differs between *ALS-S* and *ALS-R* (C→A). The base change in the coding sequence of *ALS-R* results in the amino acid substitution of

```
                    141         151         161         171         181
A. thaliana         PRHEQGGVFA AEGYARSSGK PGICIATSGP GATNLVSGLA DALLDSVPLV
L. sessiliflora     PRHEQGGVFA AEGYARASGK PGVCIATSGP GATNLVSGLA DALLDSVPLV

                    191         201         211         221         231
                    Domain A    Domain D
A. thaliana         AITGQVPRRM IGTDAFQETP IVEVTRSITK HNYLVMDVED IPRIIEEAFF
L.sessiliflora      AITGQVPRRM IGTDAFQETP IVEVTRSITK HNYLVLDVED IPRIVKEAFF

                    241         251         261         271         281
A. thaliana         LATSGRPGPV LVDVPKDIQQ QLAIPNWEQA MRLPGYMSRM PKPPEDSHLE
L.sessiliflora      IARSGRPGPV LIDVPKDIQQ QMVVPNWNQP MMLAGYLSRL PKPPSELLLE

                    291         301         311         321         331
A. thaliana         QIVRLISESK KPVLYVGGGC LNSSDELGRF VELTGIPVAS TLMGLGSYPC
L.sessiliflora      QVVRLIAESK KPVLYVGGGC LNSSEELRRF VELTGIPVAS TLMGLGSYPG

                    341         351         361         371         381
A. thaliana         DDELSLHMLG MHGTVYANYA VEHSDLLLAF GVRFDDRVTG KLEAFASRAK
L.sessiliflora      DEEFALQMLG MHGTVYANYA VDKSDLLLAF GVRFDDRVTG KLEAFASRAK

                    391         401
A. thaliana         IVHIDIDSAE IGKNKTPHVS
L.sessiliflora-R    IVHIDIDSAE IGKNKQPHVS
```

Fig. 1. Alignment of deduced amino acid sequences for ALSs of *L. sessiliflora* with *A. thaliana.* The different amino acids are shown in boldface.

ALS-S ---GCAATAACTGGACAAGTGCCACGGCGGATGATTGGCACT---
ALS-S Aa ---A I T G Q V **P** R R M I G T ---

ALS-R ---GCAATAACTGGACAAGTGCAACGGCGGATGATTGGCACT---
ALS-R Aa ---A I T G Q V **Q** R R M I G T ---

Fig. 2. A single base mutation in Domain A results in an amino acid difference between ALS genes from the susceptible biotype and resistant biotype of *L. sessiliflora*. A segment of the nucleic acid sequence (*ALS-S* and *ALS-R* represent the susceptible and resistant biotype, respectively), and the corresponding deduced amino acid sequence (ALS-S and ALS-R for susceptible and resistant biotype, respectively) are shown. Boldface indicates differences between the biotypes.

Pro with Gln in Domain A (Fig. 2); this is a replacement of a polar residue with a non-polar one.

From the deduced amino acid sequence, the secondary structures of the ALS fragments were predicted using the SSpro program and compared between susceptible and resistant biotypes (Fig. 3). Substitution of Pro is likely to induce significant topological changes, as predicted by secondary structural analysis. The results show that the extended strands formed in Domain A of the wild-type protein are converted to α-helices by the mutation. A total of seven methods of secondary structure prediction were used on the fragments covering Domain A and D of ALS from susceptible and resistant biotypes of *L. sessiliflora*. A consensus prediction was generated (Fig. 4). The data obtained support that the Pro substitution tends to increase the number of predicted α-helices.

Biotypes resistant to ALS-inhibiting herbicides, including SU-herbicides, have been reported in over 70 plant species throughout the world.[21] Herbicide resistance is considered to

ALS-S:

PRHEQGGVFAAEGYARASGKPGVCIATSGPGATNLVSGLADALLDSVPLVAITGQV**P**RRMIGT
CCCHHHHHEHHHEEEEETSCTEEEEEECSTTHHHHHHHHHHHHHHTSCEEEEETCCCEEEEEC

DAFQETPIVEVTRSITKHNYLVLDVEDIPRIVKEAFFIARSGRPGPVLIDVPKDIQQQMVVPN
CHHCCCCEEEEEEEECTTCEEEEEHTCHHHHHHHHHHHHHHTSCCCCEEEECCHHHHHHEECCC

WNQPMMLAGYLSRLPKPPSELLLEQVVRLIAESKKPVLYVGGGCLNSSEELRRFVELTGIPVA
CTHHHHHHHHHCTCCCCCHHHHHHHHHHHHHHHTSSCEEEEECCCCCCHHHHHHHHHHHTCCHH

STLMGLGSYPGSDEEFALQMLGMHGTVYANYAVDKSDLLLAFGVRFDDRVTGKLEAFASRAKI
HHHTTCCCCTTCHHHHHHHHHTCTHEEEEEHEECHHHHHHHHTCEECCHHHSHHHHHHHHHEE

VHIDIDSAEIGKNKQPHVS
EEEEECEHEECTTCCCCCC

ALS-R:

PRHEQGGVFAAEGYARASGKPGVCIATSGPGATNLVSGLADALLDSVPLVAITGQV**Q**RRMIGT
CCCHHHHHEHHHEEEEETSCTEEEEEECSTTHHHHHHHHHHHHHHHTSCEEEEHHHHHHHHHHC

DAFQETPIVEVTRSITKHNYLVLDVEDIPRIVKEAFFIARSGRPGPVLIDVPKDIQQQMVVPN
HHHCCCCEEEEEEEECTTCEEEEEHTCHHHHHHHHHHHHHHTSCCCCEEEECCHHHHHHEECCC

WNQPMMLAGYLSRLPKPPSELLLEQVVRLIAESKKPVLYVGGGCLNSSEELRRFVELTGIPVA
CTHHHHHHHHHCTCCCCCHHHHHHHHHHHHHHHTSSCEEEEECCCCCCHHHHHHHHHHHTCCHH

STLMGLGSYPGSDEEFALQMLGMHGTVYANYAVDKSDLLLAFGVRFDDRVTGKLEAFASRAKI
HHHTTCCCCTTCHHHHHHHHHTCTHEEEEEHEECHHHHHHHHTCEECCHHHSHHHHHHHHHEE

VHIDIDSAEIGKNKQPHVS
EEEEECEHEECTTCCCCCC

Fig. 3. Deduced amino acid sequence (upper) and the secondary structures (lower) for herbicide-susceptible and -resistant ALS of *L. sessiliflora* predicted by the SSpro program. The secondary structure is denoted by H for α-helix; E for extended strand; T for turn; S for bend and C for rest. Domain A is underlined.

```
                                  Domain A      Domain D
ALS-S      ADALLDSVPLVAITGQVPRRMIGTDAFQETP
DPM        cchhhceeceeeeeeceechheeccchhhhcc
GOR4       ccccccccceeeeccccccccccccceeeeec
HNNC       ccccccccceeeeeccccceeeccccccccc
PHD        cceeccccceeeeeeccccceeeeccccccc
Predator   ccccccccccceeeccccccccccccccccc
SIMPA96    chhhhccceeeeeccccccceeccccccccc
SOPM       hhhhhttcceeeeetccceeeeeccccccc
Sec.Cons.  cc??ccccceeeeccccccceeccccccc

ALS-R      ADALLDSVPLVAITGQVQRRMIGTDAFQETP
DPM        cchhhheeeeeeeeeeehheecchhhhhcc
GOR4       ccccccccchhhhchhhhhhhhccceeeeec
HNNC       ccccccccceeeeeccchheeeccccccccc
PHD        ccccccccceeeeeeccccceccccccccc
Predator   ccccccccccceeeehhhhhhhccccccccc
SIMPA96    chhhhccceeeeeccchhhhccccccccccc
SOPM       hhhhhttcceeeeetcchheeccccccccc
Sec.Cons.  ccccccccceeeeeecchhhhe?cccccccc
```

Fig. 4. Consensus predictions (Sec. Cons.) of herbicide-susceptible and -resistant ALS from *L. sessiliflora* using the DPM, GOR4, HNNC, PHD, Predator, SIMPA96 and SOPM methods. The secondary structure is denoted by h for α-helix; e for extended strand; t for beta turn; c for random coil and ? for ambiguous state.

be due to the expression of an altered ALS, which has been selected for as a direct result of repeated exposure of field populations of weeds to ALS-inhibiting herbicides. The mutation of the ALS gene is usually caused by an amino acid substitution in conserved sequences.[7,22] The mutation results in ALS activity that is less sensitive to herbicide action. In our research, analysis of partial sequences of ALS from SU-herbicide resistant *L. sessiliflora* showed that a point mutation occurs in the highly conserved Domain A, where Pro (CCA) was substituted with Gln (CAA). The mutation resulted in insensitivity to the herbicide, but the retaining of the catalytic activity of ALS (data not shown). The Pro residue in Domain A of ALS has been considered to play a key role in herbicide action. When the functional Pro is substituted with another amino acid such as His, Ser, Gln, Thr, Arg, Leu, Ala, Ile or Lys[3,5] during the evolution of resistance to SU-herbicides, the activity of ALS is not affected by SU-herbicides.

Pro is unique among the natural amino acids, because the α-nitrogen is part of a pyrrolidine ring that imparts unique constraints on the peptide backbone and prevents it from serving as a hydrogen bond donor. The most conserved Pro residue makes the structure more flexible.[23] Furthermore, there is a G-X-X-P motif in Domain A, which was suggested to be responsible for the backbone flexibility in the protein.[24] The flexibility of the herbicide-binding site in the ALS molecule is important to the herbicide action. The conserved site concerned with Pro is essentially associated with the function as a herbicide. It has been suggested that the SU-herbicides induce the formation of a docking pocket and bind to ALS in S biotypes, and then either change the conformation of the active site or block the active site.[25] ALS is inactivated irreversibly after the binding. The Pro in Domain A of ALS is a mutation hot spot in S-biotype weeds. Gln and amino acids described above could replace Pro and restore hydrogen bonding. Consequently, the flexibility is disrupted by the Pro mutation; and then ALS becomes less sensitive to herbicides.

SU-herbicides do not act as analogues of substrates or cofactors, and thus the inhibitory mechanism is complex. To fully understand the interplay among SU's binding and conformational changes, high-resolution structures of SU-bound ALS are required. Unfortunately, experimental difficulties abound for obtaining such structures.[26] However, secondary structural content is a very important feature both for experimental and for theoretical studies in protein science. Various efforts at predicting protein secondary structures have been made.[27] The secondary structures of the ALS fragments were predicted and compared between S- and R-biotypes. Substitution of Pro induces significant topological changes, as predicted by secondary structural analysis. The conversion of the peptide segment from an extended strand to an α-helix by a single-site mutation observed in the secondary structure of the ALS mutant (Pro to Gln) occurred at the point involved in the herbicide binding. This peptide shows a "chameleon-like" character since it can adopt either an α-helix or a β-strand structure in evolution. The structure of the Pro mutation provides an additional explanation as to why the extended strand structure in Domain A is important to the flexibility and herbicidal action. This result confirms the structural importance of the proline residue located at the binding site of herbicides. This conformational change in the mutated ALS might inhibit formation of a herbicide docking pocket or binding to that portion of the protein. Inhibition of ALS by SU-herbicides is time dependent and is not competitive with substrate, and the herbicide-binding site is located near the catalytic center. The change of the Domain A fragment into an α-helical structure would likely destabilize the existing protein scaffolding,[27] and reduce the sensitivity of ALS to herbicides. In essence, our hypothesis suggests that SU-herbicides do not simply bind to ALS, rather, they change the ALS structure so that ALS-substrate affinity is weakened. The resistance to SU-herbicides is generated when the binding site is destroyed or no longer exists.

The Pro mutation is suggested to cause a significant structural perturbation of ALS and confers herbicide resistance. In the development of new herbicides, it is desirable to have novel ALS-inhibiting herbicides binding a different site, as imidazoliones do,[28] or differing structurally from SU-herbicides, such as pyrimidinyl carboxy herbicides.[29]

ACKNOWLEDGMENTS

We express our appreciation to Dr. Bruce A. Auld of the University

of Sydney, and Dr. K. Itoh and Dr. A. Uchino of the National Agricultural Research Center for the Tohoku Region, Japan for discussion and comments on the manuscript. We also thank the technical editor and two reviewers for their comments.

REFERENCES

1) M. K. Tan and R. W. Medd: *Plant Sci.* **163**, 195–205 (2002).
2) L. L. Saari, J. C. Cotterman and D. C. Thill: "Herbicide Resistance in Plants: Biology and Biochemistry," ed. by S. B. Powles and J. A. M. Holtum, Lewis Publishers, Boca Raton, pp. 83–139, 1994.
3) M. J. Guttieri, C. V. Eberlein and D. C. Thill: *Weed Sci.* **43**, 175–178 (1995).
4) T. R. Wright, N. F. Bascomb, S. F. Sturner and D. Penner: *Weed Sci.* **46**, 13–23 (1998).
5) P. Boutsalis, J. Karotam and S. B. Powles: *Pestic. Sci.* **55**, 507–516 (1999).
6) C. V. Eberlein, M. J. Guttieri, P. H. Berger, J. K. Fellman, C. A. Mallory-Smith, D. C. Thill, R. J. Baerg and W. R. Belknap: *Weed Sci.* **47**, 383–392 (1999).
7) A. Uchino and H. Watanabe: *Weed Biol. Manag.* **2**, 104–109 (2002).
8) G.-X. Wang, H. Kohara and K. Itoh: Brighton Crop Protection Conference-Weeds, Vol. 1, Brighton, pp. 311–318, 1997.
9) K. Itoh, G.-X. Wang and S. Ohba: *Weed Res.* **39**, 413–423 (1999).
10) A. Koarai and H. Morita: *J. Weed Sci. Technol.* **47**, 20–28 (2002).
11) G.-X. Wang, H. Watanabe, A. Uchino and K. Itoh: *Pestic. Biochem. Physiol.* **68**, 59–66 (2000).
12) G. Pollastri, D. Przybylski, B. Rost and P. Baldi: *Proteins: Structure, Function, and Genetics* **47**, 228–235 (2002).
13) G. Deleage and B. Roux: *Protein Eng.* **1**, 289–294 (1987).
14) J. Garnier, J. F. Gibrat and B. Robson: *Method. Enzymol.* **266**, 540–553 (1996).
15) Y. Guermeur: Dissertation of Université de Paris VI (1997).
16) B. Rost and C. Sander: *J. Mol. Biol.* **232**, 584–599 (1993).
17) D. Frishman and P. Argos: *Protein Eng.* **9**, 133–142 (1996).
18) J. M. Levin, B. Robson and J. Garnier: *FEBS Lett.* **205**, 303–308 (1986).
19) C. Geourjon and G. Deleage: *Protein Eng.* **7**, 157–164 (1994).
20) C. Combet, C. Blanchet, C. Geourjon and G. Deléage: *Trends Biochem. Sci.* **25**, 147–150 (2000).
21) I. Heap: http://www.weedscience.com (2002).
22) F. S. Frank, N. B. Joanne and S. P. Mark: *J. Mol. Biol.* **323**, 951–960 (2002).
23) J. Jacob, H. Duclohier and D. S. Cafiso: *Biophys. J.* **76**, 1367–1376 (1999).
24) Y. An and R. A. Friesner: *Proteins: Structure, Function, and Genetics* **48**, 352–366 (2002).
25) S. S. Pang, L. W. Guddat and R. G. Duggleby: *J. Biol. Chem.* **278**, 7639–7644 (2003).
26) C. Geourjon and G. Deleage: *Protein Eng.* **7**, 157–164 (1994).
27) W. Z. Yang, T. P. Ko, L. Corselli, R. C. Johnson and H. S. Yuan: *Protein Sci.* **7**, 1875–1883 (1998).
28) D. L. Shaner, P. C. Anderson and M. A. Stidham: *Plant Physiol.* **76**, 545–546 (1984).
29) M. Tamaru, J. Inoue, R. Hanai and S. Tachikawa: *J. Agric. Food Chem.* **45**, 2777–2783 (1997).