Structural Analysis of the **TNIT4** Genes Encoding Nitrilase-like Protein From Tobacco

Mitsuko DOHMOTO, Jun SANO, Hiroyuki TSUNODA,† and Kazuo YAMAGUCHI*

Institute for Gene Research, Kanazawa University, Takara-machi 13-1, Kanazawa 920-0934, Japan

(Received 16 June 1999; Revised 27 July 1999)

**Abstract**

Nitrilase (nitrile aminohydrolase, EC 3.5.5.1) catalyzes the hydrolysis of indole-3-acetonitrile (IAN) to indole-3-acetic acid (IAA). The *Arabidopsis thaliana* genome has four nitrilase genes (NIT1 to NIT4), while tobacco (*Nicotiana tabacum*) has only NIT4 homologs (TNIT4A and TNIT4B) and no NIT1 to NIT3 homologs. We have cloned the promoter region of TNIT4B and determined the transcriptional start sites of these genes. The TNIT4 genes are expressed in various organs at low levels. The positions of the exon-intron splicing junctions in *Arabidopsis* NIT1 to NIT3 are completely conserved in TNIT4A.

**Key words:** *Nicotiana tabacum*; nitrilase; TNIT4; transcription start point; exon-intron splicing junction

Despite the importance of auxins in plant growth and development, the pathways of auxin biosynthesis and their regulation are still unclear. Indole-3-acetaldehyde, indole-3-acetamide or indole-3-acetonitrile (IAN) is supposed to be the precursor of the most abundant auxin, indole-3-acetic acid (IAA).1 IAN appears to be the immediate precursor of IAA in the Brassicaceae family.2 A nitrilase (nitrile aminohydrolase, EC 3.5.5.1) that hydrolyzes IAN to IAA has been detected in some plant families, for example, Brassicaceae, Poaceae and Musaceae3 and has been partially purified from *Brassica campestis*.4

Four cDNAs (NIT1 to NIT4) homologous to microbial genes for nitrilase were cloned from *Arabidopsis thaliana*.5-7 Products of these genes were reported to be able to convert IAN to IAA in vitro when expressed in *Escherichia coli*.5,7 NIT1 to NIT3 are tandemly arranged on *Arabidopsis* chromosome III, while NIT4 is located on chromosome V.7 IAN has auxin activity for *A. thaliana* and inhibits its growth probably due to the hydrolysis of IAN to IAA. On the other hand, the tobacco (Nicotiana tabacum) genome has only NIT4 homologs (TNIT4A and TNIT4B)8 but not NIT1 to NIT3 homologs (unpublished observation). Tobacco seedlings are resistant to IAN. Introduction of not only *Arabidopsis* NIT2,9 but also NIT1 and NIT3 (unpublished observation) into tobacco conferred growth inhibition by IAN, while over-expression of NIT4 or ectopic expression of NIT4A had no effect on the sensitivity of transgenic tobacco to IAN (unpublished observation). Therefore, tobacco NIT4 proteins might have no or little nitrilase ability to convert IAN to IAA. Phylogenetic analyses of NIT genes from prokaryotes and higher plants show that *Arabidopsis* NIT1 to NIT3 and *Arabidopsis* NIT4 and tobacco TNIT4 are divided into different subgroups, although all NIT genes from higher plants are included in the same group. The genomic structure of NIT genes has been reported for only NIT1 to NIT3.10-13 It is important to clarify the genomic blueprints of the TNIT4 group to understand their functions and the difference from the NIT1 to NIT3 group. In this study, we analyzed the genomic structures of tobacco TNIT4 genes, and the characteristics of their gene expression including the transcriptional start sites.

1. Cloning of the **TNIT4B** Promoter Region

To clone the TNIT4B promoter region, we designed two primers, 5’ATCTCCCATATCCACTTCAGCAAAC3’ (nucleotide positions +66 to +42; nucleotide position 1 was designated as A of the first codon of TNIT4A) and 5’ACTGTAGTTCAGGCCTCTACCATC3’ (+94 to +117), toward outsides of TNIT4 cDNA for inverse PCR. DNA sequencing of an amplified fragment, 2 kb in length, revealed that it contained only a 669 bp-region up-
stream of the start codon of TNIT4B (data not shown). The λFixII- N. tabacum SR-1 genomic DNA library was searched for further upstream regions using the 669 bp-region as a probe. As a result, a 2.2 kb-upstream region of TNIT4B gene was cloned and sequenced (Fig. 1).

The sequence in this region is very A+T rich (71%), as is the Arabidopsis NIT1 promoter region. The sequence, AAACATGGC (−4 to +5), around the initiation codon resembles AACATGGC conserved in the plant initiation codon. DNA elements for transcription factors were detected by PLACE (http://www.dna.afric.go.jp/htdocs/PLACE/). Two TATA box sequences (TATATAAA) are located at positions −395 to −389 and −183 to −177. A CAAT box (−430 to −427) is present upstream of the former, but this complete minimal promoter sequence might be too far from the transcriptional start site. No CAAT box is found upstream of the latter TATA box. These characteristics of the TNIT4B sequence may be related to low level expression of this gene. Interestingly, this region contains sequences similar to an abscisic acid-responsive element (ABRE, positions −2061 to −2055) and an ethylene-responsive element (ERE, positions −971 to −964).

2. Determination of TNIT4 Transcriptional Start Sites

The oligo-cap method can be used to amplify specifically the 5' region of mRNA with the methylated guanine cap at the 5' end. We used this method to determine transcriptional start sites of TNIT4 genes. Sequences common to both TNIT4 genes were selected as downstream primers for nested PCR. Cloning and sequencing of the PCR products revealed two kinds of sequences; one was derived from TNIT4A, and another from TNIT4B judging from the sequences of their coding regions. The 5' ends of both sequences started at position −56, suggesting strongly that these are the transcriptional start sites of the TNIT4 genes.

3. Expression of TNIT4 Genes

Four nitrilase genes of A. thaliana are differentially expressed in an organ-specific manner. Accumulation of the transcripts from NIT4 was predominantly observed in leaves, stems and siliques. We performed semi-quantitative PCR to detect mRNA of tobacco TNIT4 genes (Fig. 2A), since it was hardly detected in any tissues tested by Northern blotting analysis, probably due to its low expression. A primer set was designed to amplify the DNA region spanning intron 1, in order to discriminate the mRNA-derived PCR signal and the genomic DNA-derived PCR signal. As a control, an actin gene cDNA was amplified. Figure 2B shows PCR cycle numbers in which TNIT4 and the actin gene cDNA were exponentially amplified, and the difference in these cycle numbers from the start of exponential phase were indicative of the putative initiation codon of translation. A double underline shows the sequence conserved around the plant initiation codon (ACACATGGC). The TATA box sequences are presented by boxed letters. ABRE (Abscisic Acid Responsive Element) motif (positions −2061 to −2055) and ERE (Ethylene Responsive Element) motif (position −971 to −964) are underlined. The arrowhead indicates the transcription start site.
numbers (Δcycle). In the case of cotyledons, PCR signals of TNIT4 increased exponentially at 2 cycles later than those of actin. In another experiment, amounts of actin cDNA amplified from 8- to 10-fold diluted total RNA from cotyledons was roughly equal to those of TNIT4 cDNA amplified from non-diluted sample. Although TNIT4 genes were expressed in all tissues tested, their levels were relatively lower in roots and flowers. These expression patterns might resemble those of A. thaliana NIT4.

The TNIT4B promoter region contains two phytohormone-responsive elements, ABRE and ERE, as described above. We analyzed the expression of TNIT4 by semi-quantitative RT-PCR after treatments with 100 μM abscisic acid or 100 μM 1-aminocyclopropane-1-carboxylic acid, a precursor of ethylene, for 1 hr or 24 hr. Expression of the osmotin gene was induced by both abscisic acid and ethylene (Δcycle of osmotin against actin increased by 4) as described previously, while expression levels of TNIT4 were unchanged (data not shown).

4. Analysis of Exon-Intron Junctions

A. thaliana NIT1 to NIT9 are tandemly arranged on chromosome III, and the positions of their exon-intron junctions are completely conserved. The NIT4 gene is located on chromosome V, and its exon-intron junctions are currently unknown. We searched for the exon-intron junctions of TNIT4 by comparison of PCR products derived from mRNA with those from genomic DNA. The following primers were designed according to the sequence of TNIT4.

Upstream primers;
U1 (5'-CTAAGAAAAACATGGCTTTGGTCCC-3', -10 to +14),
U2 (5'-CAGCTGTTGTTGTITCTGAAGC-3', +187 to +209),
U3 (5'-GTTGTTGATTATTTATGAGATGAGC-3', +905 to +879),
U4 (5'-CTGTTGGAATGTTATTATCACC-3', +1113 to +1089), and
U5 (5'-AAAGAACAGGGTTTTTCAATC-3', +1303 to +1279).

Downstream primers;
D1 (5'-CTGTTGCAATCTCCATGCAGCACC-3', +268 to +244),
D2 (5'-CCATCTCTCTCAATACAC-3', +400 to +382),
D3 (5'-CAGATCGTGTTGTATGAGAGC-3', +905 to +879),
D4 (5'-GAGATTTCTGACACTTTGCAGG-3', +1113 to +1089), and
D5 (5'-AAAGAACAGGGTTTTTCAATC-3', +1303 to +1279).

Figures 3B and 3C show that PCR products were about 2300 bp (primers U1 and D1), 630 bp (primers U2...
Figure 3. Exon-intron splicing junctions of TNIT4A. (A), positions of primers (arrows) for PCR. Open boxes: exons, closed boxes: introns, and shaded boxes: untranslated regions. (B) and (C), RT-PCR of total RNA (B) and PCR of genomic DNA (C). PCR from genomic DNA was carried out using LA-PCR Kit Ver. 2 (Takara Shuzo). The reaction mixture was heated at 94°C for 1 min, followed by 30 cycles of PCR (98°C for 20 sec and 68°C for 20 min) and terminated at 72°C for 10 min. RT-PCR was carried out as described in the legend of Fig. 2. PCR was performed under the conditions of 95°C for 2 min, 35 cycles of 95°C for 1 min and 50°C for 1 min, then 60°C for 7 min. The PCR products were cloned into pUC19 vector and sequenced. Lane M, molecular weight marker; lanes 1, primers U1 and D1; lanes 2, primers U2 and D2; lanes 3, primers U3 and D3; lanes 4, primers U4 and D4; lanes 5, primers U1 and D5. (D), Nucleotide sequence of exon-intron junctions. Exons are given in capital letters, and introns in small letters. The figures above the sequences indicate nucleotide positions of TNIT4A cDNA. The junction motif of introns, gt-ag, is underlined.

and D2), 1260 bp (primers U3 and D3) and 1100 bp (primers U4 and D4) longer than RT-PCR products. The sum of these additional DNA lengths almost equals the difference (5150 bp) between the products of PCR (Fig. 3C, lane 5) and RT-PCR (Fig. 3B, lane 5). The four PCR products (Fig. 3C, lanes 1-4) were cloned and sequenced. Four exon-intron junctions were found between nucleotide positions +142 and +143, +322 and +323, +616 and 617, and +898 and 899 (Fig. 3D). These correspond to the positions between the first and the second codons of Val-47, Gly-108, Gly-206 and Asp-300. When the amino acid sequences of TNIT4A, NIT1, NIT2 and NIT3 were aligned, the locations of four exon-intron junctions in TNIT4A were completely conserved in NIT1, NIT2 and NIT3; for example, four junctions in NIT1 are located between the first and the second codons of Asp-43, Gly-105, Gly-203 and Asp-297.11 No other exon-intron junctions were found in the TNIT4A coding region.

The 5'-splice junctions of TNIT4A are divergent in nucleotide sequence from the consensus sequence, (C/T)AG-gtaag, of plant introns18 except for the core sequence AG:gt. In the 3'-splice junctions of TNIT4A, t-clusters in addition to the core sequence, cag:G, are well conserved.

In conclusion, the NIT1 to NIT3 group and the NIT4 group might be derived from the same ancestor. However, IAN has no “auxine activity” on tobacco seedlings,9 that is, tobacco is IAN-resistant, and the tobacco genome has only NIT4 homologs. TNIT4 proteins might have another enzymatic activity than IAA-producing activity. These NIT4 homologs might be distributed widely in higher plants, since a nitrilase-like gene found in the rice genome (Accession no. D15299) is a member of the NIT4 subfamily. The function of the NIT4 group remains to be clarified.

Acknowledgements: We thank colleagues in our laboratory for stimulating discussions, especially G. Isaji and R. Chiba for constructing plasmids and other collaborations, and K. Mukai for her technical assistance.

References
1. Gibson, R. A., Scheider, E. A., and Wightman, F. 1972, Biosynthesis and metabolism of indole-3-acetic acid, J. Exp. Bot., 23, 381–399.
2. Ludwig-Müller, J. and Hilgenberg, W. 1988, A plasma membrane-bound enzyme oxidizes L-tryptophan to indole-3-acetaldoxime, Physiol Plant, 74, 240–250.
3. Thimann, K. V. and Mahadevan, S. 1964, Nitrilase I. Occurrence, preparation, and general properties of the enzyme, Arch. Biochem. Biophys., 105, 133–141.
4. Rausch, T. and Hilgenberg, W. 1980, Partial purification of nitrilase from Chinese cabbage, *Phytochemistry*, 19, 747–750.

5. Bartling, D., Seedorf, M., Mithofer, A., and Weiler, E. W. 1992, Cloning and expression of an *Arabidopsis* nitrilase which can convert indol-3-α-cetonitrile to the plant hormone, indol-3-acetic acid, *Eur. J. Biochem.*, 205, 417–424.

6. Bartling, D., Seedorf, M., Schmidt, R. C, and Weiler, E. W. 1994, Molecular characterization of two cloned nitrilases from *Arabidopsis thaliana*: key enzymes in biosynthesis of the plant hormone indole-3-acetic acid, *Proc. Natl. Acad. Sci. USA*, 91, 6021–6025.

7. Bartel, B. and Fink, G. R. 1994, Differential regulation of an auxin-producing nitrilase gene family in *Arabidopsis thaliana*, *Proc. Natl. Acad. Sci. USA*, 91, 6649–6653.

8. Tsunoda, H. and Yamaguchi, K. 1995, The cDNA sequence of an auxin-producing nitrilase homologue in tobacco (Accession No. D63331), *Plant Physiol.*, 109, 339.

9. Schmidt, R. C, Muller, A., Hain, R., Bartling, D., and Weiler, E. W. 1996, Transgenic tobacco plants expressing the *Arabidopsis thaliana* nitrilase II enzyme, *Plant J.*, 9, 683–691.

10. Zhou, L., Bartel, B., and Thornburg, R. 1996, Nucleotide sequence of the *Arabidopsis thaliana* nitrilase 1 gene (Accession No. U38845), *Plant Physiol.*, 110, 337.

11. Hillebrand, H., Tiemann, B., Bartling, D., and Weiler, E. W. 1996, Structure of the gene encoding nitrilase 1 from *Arabidopsis thaliana*, *Gene*, 170, 197–200.

12. Zhou, L., Bartel, B., and Thornburg, R. 1996, Nucleotide sequence of a pathogen induced nitrilase gene from *Arabidopsis thaliana*: Nit2 (Accession No. U47114), *Plant Physiol.*, 110, 1048.

13. Hillebrand, H., Bartling, D., and Weiler, E. W. 1998, Structural analysis of the nit2/nit1/nit3 gene cluster encoding nitrilases, enzymes catalyzing the terminal activation step in indole-acetic acid biosynthesis in *Arabidopsis thaliana*, *Plant Mol. Bio.*, 36, 89–99.

14. Lutcke, H. A., Chow, K. C, Mickel, F. S., Moss, K. A., Kern, H. F., and Scheele, G. A. 1987, Selection of AUG initiation codons differs in plants and animals, *EMBO J.*, 6, 43–48.

15. Maruyama, K. and Sugano, S. 1994, Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides, *Gene*, 138, 171–174.

16. Thangavelu, M., Belostotsky, D., Bevan, M. W., Flavell, R. B., Rogers, H. J., and Lonsdale, D. M. 1993, Partial characterization of the *Nicotiana tabacum actin* gene family: Evidence for pollen-specific expression of one of the gene family members, *Mol. Gen. Genet.*, 240, 290–295.

17. Nelson, D. E., Raghothama, K. G., Singh, N. K., Hasegawa, P. M., and Bressan, R. A. 1992, Analysis of structure and transcriptional activation of an osmotin gene, *Plant Mol. Biol.*, 19, 577–588.

18. Brown, J. W. S. 1986, A catalogue of splice junction and putative branch point sequences from plant introns, *Nucl. Acids Res.*, 14, 9549–9559.
