ELSEVIER

Biochimica et Biophysica Acta 1229 (1995) 389–392

BBA Biochimica et Biophysica Acta

Short Sequence Paper

The genomic organization of the gene encoding a nitrate-inducible ferredoxin-NADP$^+$ oxidoreductase from rice roots ☆

Hideyuki Aoki [a], Kunisuke Tanaka [b], Shoji Ida [a,*]

[a] *Research Institute for Food Science, Kyoto University, Uji, Kyoto 611, Japan*
[b] *Department of Biochemistry, Kyoto Prefectural University, Kyoto 606, Japan*

Received 30 December 1994; revised 27 February 1995; accepted 27 February 1995

Abstract

A genomic clone of the gene encoding a nitrate-inducible ferredoxin-NADP$^+$ oxidoreductase (FNR) from rice (*Oryza sativa* L.) roots has been isolated and its nucleotide sequence determined. The clone contains 3897 nucleotides of the gene which consists of six exons interrupted by five introns. The transcription start site was determined by primer extension analysis which locates 64 bp upstream of the ATG translation initiation codon. The 5′-flanking region contains canonical TATA- and CAAT-boxes, and a potential Sp1-binding site. Four ATCAA(A/C) and two inverted TTTGAT sequences are localized in the promoter region and a TGTAA motif occurs three times in the 3′-untranslated region. No significant similarity was found when the 5′ flanking region was compared with that of the photosynthetic FNR gene.

Keywords: Ferredoxin-NADP$^+$ oxidoreductase; Gene structure; Nucleotide sequence; (Rice root); (*Oryza sativa* L.)

Ferredoxin-NADP$^+$ oxidoreductase (FNR, EC 1.18.1.2) provides NADPH from reduced ferredoxin (Fd) in the oxygenic photosynthetic electron transport, whereas the enzyme has been postulated to catalyze the reverse reaction to generate the reductant from NADPH for nitrite reduction and glutamate synthesis in the plastids of non-photosynthetic tissues [1,2]. Recently we reported two

☆ The nucleotide sequence data reported in this paper have been submitted to the DDBJ/EMBL/GenBank Data Libraries under the accession number D38445.

* Corresponding author. Fax: +81 774 333004.

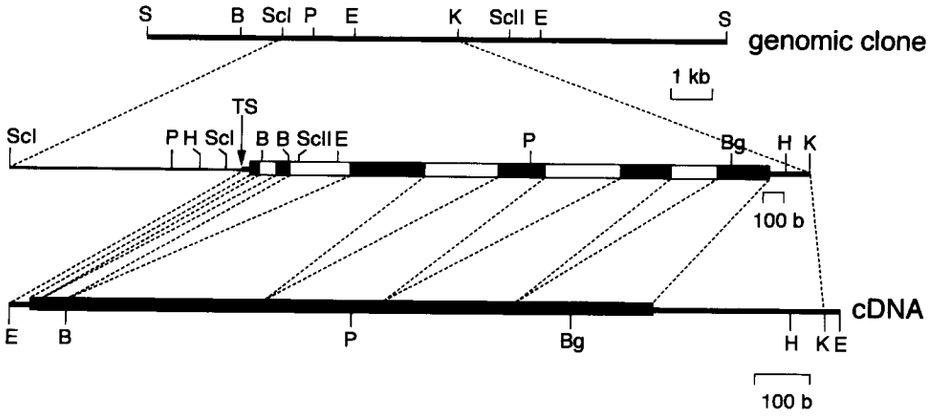

Fig. 1. Restriction map of the root ferredoxin-NADP$^+$ oxidoreductase gene of rice. The wide bar represents the protein coding regions, the open bar introns, the medium bar transcribed regions and the narrow bar untranscribed regions. The transcription start site (TS) is marked with a vertical arrow. B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; K, *Kpn*I; P, *Pst*I; S, *Sal*I; Sc I, *Sac*I and Sc II, *Sac*II.

SSDI 0005-2728(95)00032-1

structurally distinct FNR cDNAs in rice, one of which corresponds to the photosynthetic FNR [3] and the other occurring in the roots [4]. The expression of the former gene has been shown to be under phytochrome control in the photosynthetic tissues [5–7]. In contrast to the leaf gene, root FNR gene is nitrate-inducible [4,8] along with the genes involved in the nitrate assimilation pathway in the root tissue of higher plants [9–11]. A genomic clone encoding spinach leaf FNR gene has been described which comprises the promoter and partial coding regions [5]. We report here the genomic structure of the gene corresponding to an FNR cDNA from rice roots [4].

Five genomic clones were isolated from $4 \cdot 10^5$ phage of an EMBL3 genomic library of rice (*Oryza sativa* L. cv. Nihonbare) using the rice root FNR cDNA as a probe. Restriction maps showed all the clones to be identical. A representative clone was chosen to digest with *Sal*I, which gave a 15 kb fragment. Further digestion of the DNA with *Sac*I, *Eco* RI and *Kpn*I yielded two fragments of 1.6 kb and 2.3 kb that comprised an entire region of the gene (Fig. 1). The resultant DNA was subcloned into pBluescript II SK + (Stratagene) and sequenced for both strands from progressive deletion with an Applied Biosystems 373A DNA sequencer according to the manufacturer's

```
-1122                                                                                                gagctcctcatggaggccttga
-1100 tcacatcatggggatgtgatcatctgaaacgaagctggctgtttgcaactcatcatatgcctgaagtcaggactcaggactagtccatctgtagttagcc
-1000 taccaaaaggattatacttttagttttaccttaaaaagaaatggcgtgtgccattgaaggataagaaatgtgttggttgtagttaaaaggtactgcttg
 -900 ctatcctgtgatgctgtaaaaatgcaacagaattgcgcagctctctgcacaagtagtttctcagcttttggtgaacagtccgaaatggagaaacaaaggta
 -800 aaaacatcagacttacatggttctcttctgtggtgttgcatgacccgattacatttgatgagcatagtaggcccataaaatttgctacctttggtcctt
 -700 gccttttttttaatttttttgggtaaatttgcaatgtgatcagtcacatgatggctcattggtgcaagggcaaacatctagccctttgatcacaaaagt
 -600 ttttttcccccctctatccttaaaaaagttttttttttcgccaagatttgacacattttggcagtaaaatgtccaacagatacacaccggcttgtgcact
 -500 acatgcacacgctgccaactcatgtactgttccaaaaaactaaaaagctaataatatacaagttgtatgtgcatctggctagctccatgggatacatgca
 -400 tctcttcaggccaaaggtgacccttgatgccaccacaaagccatttttatcagcagccatctggagcactcaagctgcagccaccatttctgctagtcaa
 -300 gcagccaggcaacccaaccctctgctgctacctccacctaacccaaacccccctaatctaatcaacatcaaaaacgaaaaaaacaaaatctttttttttaa
 -200 attatcaaaccccaaacccaaactctaaaacgaaattactaagcagctaaagcttaaatcaaccccaagagcgggccgggttaatcacgcaaatttaagcagc
 -100 taatagcttaagtcgccccttccagcacctcaagagctcacctcactcctccccctttctctgattctgtcctataaaaccccgcctcctccccctcccatctc
    1 atctcaactcatctcatctcccttctccagatcaaaagacccttgctttccctcaggatcggccATGGCGACCGCCGTTGCGTCCCAGgtacgcgccgtc
                                                                                    MetAlaThrAlaValAlaSerGln
  101 ctcgcggtttttggatcccctggttgtctgcgctgcgcgatggtgctgagttcatgggggttcttgtggctgtttctgacacagGTTGCTGTCTCTGCT
                                                                                                ValAlaValSerAla
  201 CCGGCTGGCTCGGATCGCGGCTTGAGGAGTTCTGGGATCCAGgtgcgtggtttggtgcccctgatggtagaacttgcttcctttgggtggaatcctgtg
      ProAlaGlySerAspArgGlyLeuArgSerSerGlyIleGln
  301 aggattccgcgggcttggattggtcaatgtgtagtaggtggatacaaatctatgggatatatctcagattggttcagtattgagcgctattgagagctag
  401 gcaacgccagatattgctgctgcttgttaattcgatgtgcagatcaattatcagggttcagggtttggtttagtaggctgtttatcctgaattcacagc
  501 gttcttgaatggtggtgctaaactatgtgctggatttcattctgtgtgattcctcagGGTAGCAACAATATTAGCTTTAGCAACAAATCATGGGTTGGCA
                                                                                    GlySerAsnAsnIleSerPheSerAsnLysSerTrpValGlyT
  601 CCACATTGGCGTGGGAGAGCAAGGCCACGCGACCGAGGCATGCGAACAAGGTGCTCTGCATGTCAGTTCAGCAAGCGAGCGAAAGCAAGGTTGCTGTCAA
      hrThrLeuAlaTrpGluSerLysAlaThrArgProArgHisAlaAsnLysValLeuCysMetSerValGlnGlnAlaSerGluSerLysValAlaValLy
  701 GCCTCTTGATTTGGAGAGTGCTAACGAGCCGCCGCTCAACACATACAAACCAAAGGAGCCTTACACCGCCACAATTGTCTCGGTTGAGAGGATCGTAGGC
      sProLeuAspLeuGluSerAlaAsnGluProProLeuAsnThrTyrLysProLysGluProTyrThrAlaThrIleValSerValGluArgIleValGly
  801 CCCAAGGCTCCAGGAGAGACATGCCACATTGTTATTGATCATGGTGGCAATGTGCCTTACTGGGAGGGGCAAAGCTATGGCATTATTCCTCCAgtaagta
      ProLysAlaProGlyGluThrCysHisIleValIleAspHisGlyGlyAsnValProTyrTrpGluGlyGlnSerTyrGlyIleIleProPro
  901 attcctgctttcaatcattttccttgtttaaaatagaacattatgcttatgctttgtcggtttcgttcatagtctatttatctgattggatattgacgat
 1001 tgtctcgttgacactattaattaagaacttacaaagatatagtcggtagtatattaggattatttctgtagacactgtttccaatgttctattcaatggc
 1101 ttaaaattgtaaacctaggtgagaaaacctgctcttgataggagacttgagtatgtgcacacactattaatacagttccagtggtatctccatgggcatc
 1201 accctcttcactgttctctaaaccctgtgcatttttgcttgttagGGGGAGAACCCGAAGAAGCCTGGTGCACCACATAATGTCCGTCTTTATTCAATTG
                                                                                GlyGluAsnProLysLysProGlyAlaProHisAsnValArgLeuTyrSerIleA
 1301 CATCTACAAGGTATGGAGATTCATTCGATGGAAGGACCACTAGTTTATGTGTGCGCCGTGCCGTTTATTATGATCCTGAAACTGGCAAGGAGGACCCCTC
      laSerThrArgTyrGlyAspSerPheAspGlyArgThrThrSerLeucysValArgArgAlaValTyrTyrAspProGluThrGlyLysGluAspProSe
 1401 AAAAAATGGTGTCTGCAGTAACTTCCTATGTAATTCAAAACCAGGGGACAAGGTTAAAGTGACAGgttagttgatttctttcagcttcagatgaatgcaa
      rLysAsnGlyValCysSerAsnPheLeuCysAsnSerLysProGlyAspLysValLysValThrG
 1501 aggctttgtgagtttcctatgaggttgtttaacttctaattccaagcggagcaaaattagttgttttagttgcgtttttcctgaaacccacctttggtag
 1601 acaacctatttctgagttcatacataccaacaacaaataaacggttctctgcaacaaaggttccaatagcctctcctctaatatgattatgccccttcta
 1701 caatatgattatcttatcttgctctttccttcttctactaaagtttgcaatcgcatcttagtttaccctttttcttttgttagtcatcatgtcctttgtggc
 1801 acatgatggatgccaagtatctcacagtgatttccttctgaacaccagGTCCGTCAGGCAAAATAATGCTCCTGCCTGAGGAAGATCCAAATGCAACTCA
                                                                                lyProSerGlyLysIleMetLeuLeuProGluGluAspProAsnAlaThrHi
 1901 CATCATGATAGCTACTGGCACTGGTGTTGCTCCATTCCGTGGCTACCTACGCCGTATGTTCATGGAAGATGTCCCAAAGTACAGATTTGGTGGCTTGGCC
      sIleMetIleAlaThrGlyThrGlyValAlaPriPheArgGlyTyrLeuArgArgMetPheMetGluAspValProLysTyrArgPheGlyGlyLeuAla
 2001 TGGCTCTTCCTTGGTGTGGCTAACACTGACAGCCTTCTCTATGATGAAGAGTTCACAAGCTACCTTAAGCAGTATCCAGACAATTTCAGgttgctagaca
      TrpLeuPheLeuGlyValAlaAsnThrAspSerLeuLeuTyrAspGluGluPheThrSerTyrLeuLysGlnTyrProAspAsnPheAr
 2101 tcatctgtcttgactcttcgtattaattattagtactacatgcatacttggctgaaacagcgccacataatgaacatagatgcctctgtcaactgatgcc
 2201 atcatagaattattgagctttgacttgagagtaccacataaatgagctgtttttgttaagtatattcgtgcatattgctataattgaattgtgttgaat
 2301 ttgtagGTATGACAAAGCGCTAAGCAGGGAGCAGAAAAACAAGAACGCTGGCAAGATGTATGTCCAGGACAAGATCGAGGAGTACAGCGACGAGATCTTC
          gTyrAspLysAlaLeuSerArgGluGlnLysAsnLysAsnAlaGlyLysMetTyrValGlnAspLysIleGluGluTyrSerAspGluIlePhe
 2401 AAGCTCTTGGATGGCGGCGCGCACATCTACTTCTGTGGTTTGAAGGGGATGATGCCTGGGATTCAAGACACCCTCAAGAAAGTGGCGGAGCAGAGAGGGG
      LysLeuLeuAspGlyGlyAlaHisIleTyrPheCysGlyLeuLysGlyMetMetProGlyIleGlnAspThrLeuLysLysValAlaGluGlnArgGlyG
 2501 AGAGCTGGGAGCAGAAGCTATCCCAGCTCAAGAAGAACAAGCAATGGCACGTTGAGGTCTACTAGgatctaagtgtccaaggattatgattgttgcgcag
      luSerTrpGluGlnLysLeuSerGlnLeuLysLysAsnLysGlnTrpHisValGluValTyr***
 2601 tgaaaaagagaaaacaaaacgcatgatctgatgattcttgtagggtggtgtaaaatcatcatttttttttctgaatatgaatcataaaatcacccatgtaa
 2701 ttcataagcttctgcatcacatgatgaacgaaaggaagcatgtaacttttgcctgtcactattgcagctggtacc
```

Fig. 2. Nucleotide sequence of the rice root ferredoxin-NADP$^+$ oxidoreductase gene. Sequence in bold capital letters indicate the exons and lower-case letters represent the 5′ and 3′ flanking and intron sequences. Numbering of nucleotides begins at the transcription start site (designated +1). The TATA, CAAT, and GC motifs are boxed. The ATCAA(A/C) and AAAGTTTTTTTT sequences are underlined and the TGTAA is double-underlined.

specifications based on the dideoxy chain-termination method [12]. Nucleotide and the deduced amino acid sequences of the gene are shown in Fig. 2. The exon/intron boundary follows the AG/GT rule of splice junction [13]. The coding region consists of 6 exons interrupted by 5 introns. The protein coding sequence is completely identical with that from the rice root FNR [4]. The first exon contains the ATG translation start codon. The binding site for Fd is in the 4th exon and NADP-PP_i binding region stretches over the 4th and 5th exons. The $NADP^+$ binding site is also located in the 5th exon. The short stretch for the active site domain is found in the last coding segment (see [4] for an assignment of the functional domains of FNR).

The transcription start site (TS) was determined by primer extension analysis with a synthetic primer complementary to the root FNR gene. The reverse transcription yielded one major extension product of 66 bases, indicating that the TS is the nucleotide A that locates 64 bp upstream from the methionine initiation (ATG) codon (Fig. 3). A TATA box (TATAA) and CAAT box (CAAT) are found at −30 and −112, respectively, as found in many eukaryotic promoters. A Sp 1-binding GC box-like sequence (GGCCGG) is also noticed at −128. Besides those core promoter elements, there are four ATCAA(A/C) sequences between −138 and −241, two of which is contiguous and its inverted sequence TTTGAT occurs twice at −609 and −742.

It is interesting to find that the TATCAA motif, closely related sequence of a core sequence TATCTA (GATA on the other strand) of the binding site for the NIT2 of *Neurospora crassa* [15], is the same as located at −193 to −198 of the root FNR gene (Fig. 2). The NIT2 is the product of *nit*-2, the major positive-acting regulatory gene, turns on the expression of various nitrogen catabolic structural genes of *N. crassa* under nitrogen-limitation conditions [16]. The existence of a NIT2-like protein named NTL1 has been reported in *Nicotiana plumbiginifolia* and its full-length cDNA clone was shown to encode a single zinc-finger domain [17] as in the case of the NIT2 of *N. crassa* [16]. Another characteristic sequence, AAAGTTTTTTTT, is present at two positions −575 and −604. Although the hexameric and T-rich motifs are suggestive of a function in the transcriptional control, their significance is as yet unknown. In the 3′-untranslated region, the TGTAA motif is localized three times as found in the maize root FNR cDNA [8].

There is no significant sequence homology in the promoter region between the rice root and spinach leaf FNR genes, although the protein sequences are about 50% identical as a whole and much higher homologies are conserved in the cofactor binding regions. The lack of substantial similarity in the 5′-flanking regions suggests that the transcription regulation might be different between the two types of FNR gene.

This research was supported in part by a Grant-in-Aid for General Scientific Research (No. 0560093) from the Ministry of Education, Science and Culture of Japan. H.A. is a J.S.P.S. Fellow.

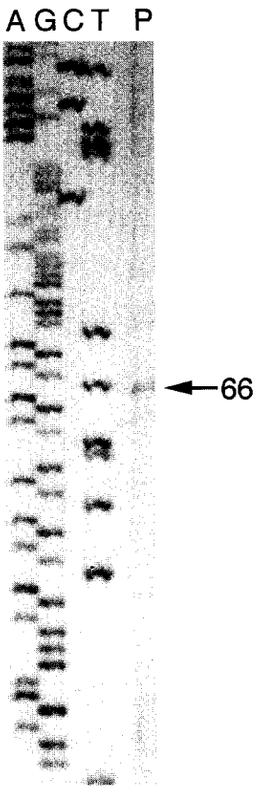

Fig. 3. Determination of the transcription start site by primer extension analysis. Primer extension was carried out according to Triezenberg [14] on a total RNA (50 μg) that was prepared from the rice roots induced with nitrate for 1 h as described [4]. The primer used was 5′-biotinylated oligomer (5′-ATGGCCGATCCTGAGGGAAA-3′) complementary to nucleotides 47–66 of the gene. The products were analyzed on a 6% denaturing (8 M urea) acrylamide sequencing gel and visualized with a chemiluminescent detection kit (Toyobo). The size of the nucleotide product is shown in the right of the figure. The transcription start site (TS) is the nucleotide T which corresponds to an A in the coding strand.

References

[1] Bowsher, C.G., Hucklesby, D.P. and Emes, M.J. (1989) Planta 177, 359–366.
[2] Bowsher, C.G., Boulton, E.L., Rose, J., Nayagam, S. and Emes, M.J. (1992) Plant J. 2, 893–898.
[3] Aoki, H., Doyama, N. and Ida, S. (1994) Plant Physiol. 104, 1473–1474.
[4] Aoki, H. and Ida, S. (1994) Biochim. Biophys. Acta 1183, 553–556.
[5] Oelmüller, R., Bolle, C.E.H., Tyagi, A.K., Niekrawietz, N., Breit, S. and Hermann, R.G. (1993) Mol. Gen. Genet. 237, 261–272.
[6] Lübberstedt, T., Bolle, C.E.H., Sopory, S., Flieger, K., Hermann, R.G., and Oelmüller, R. (1994) Plant Physiol. 104, 997–1006.
[7] Bowler, C. and Chua, N.H. (1994) Plant Cell 6, 1529–1541.
[8] Ritchie, S.W., Redinbaugh, M.G., Shiraishi, N., Vrba, J.M. and Campbell, W.H. (1994) Plant Mol. Biol. 26, 679–690.
[9] Melzer, J.M., Kleinhofs, A. and Warner, R.L. (1989) Mol. Gen. Genet. 217, 341–346.

[10] Kramer, V., Lahners, K., Back, E., Privalle, L.S. and Rothstein, S. (1989) Plant Physiol. 90, 1214–1220.
[11] Redinbaugh, M.G. and Campbell, W.H. (1993) Plant Physiol. 101, 1249–1255.
[12] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[13] Breathnach, R. and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349–383.
[14] Triezenberg, S.J. (1992) Current Protocols in Molecular Biology, Supp. 20 (Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K., eds.), pp. 4.8.1–4.8.5, John-Wiley & Sons, New York.
[15] Fu, Y.H. and Marzluf, G.A. (1990) Proc. Natl. Acad. Sci. USA 87, 5331–5335.
[16] Marzluf, G.A. (1993) Annu. Rev. Microbiol. 47, 31–55.
[17] Daniel-Vedele, F. and Caboche, M. (1993) Mol. Gen. Genet. 240, 365–373.