Isolation and Characterization of a Histidine Biosynthetic Gene in Arabidopsis Encoding a Polypeptide with Two Separate Domains for Phosphoribosyl-ATP Pyrophosphohydrolase and Phosphoribosyl-AMP Cyclohydrase

Ko Fujimori and Daisaku Ohta*

Takarazuka Research Institute, Novartis Pharma K.K., 10-66 Miyuki-cho, Takarazuka 665-8666, Japan (K.F.); and Research Institute for Biological Sciences, Okayama, 754-9-1 Yoshikawa, Kayo-cho, Okayama 716-1241, Japan (D.O.)

Phosphoribosyl-ATP pyrophosphohydrolase (PRA-PH) and phosphoribosyl-AMP cyclohydrase (PRA-CH) are encoded by HIS4 in yeast and by hisIE in bacteria and catalyze the second and the third step, respectively, in the histidine biosynthetic pathway. By complementing a hisI mutation of Escherichia coli with an Arabidopsis cDNA library, we isolated an Arabidopsis cDNA (At-IE) that possesses these two enzyme activities. The At-IE cDNA encodes a bifunctional protein of 281 amino acids with a calculated molecular mass of 31,666 D. Genomic DNA-blot analysis with the At-IE a bifunctional protein of 281 amino acids with a calculated molecular mass of 31,666 D. Genomic DNA-blot analysis with the At-IE cDNA as a probe revealed a single-copy gene in Arabidopsis, and RNA-blot analysis showed that the At-IE gene was expressed ubiquitously throughout development. Sequence comparison suggested that the At-IE protein has an N-terminal extension of about 50 amino acids with the properties of a chloroplast transit peptide. We demonstrated through heterologous expression studies in E. coli that the functional domains for the PRA-CH (hisI) and PRA-PH (hisE) resided in the N-terminal and the C-terminal halves, respectively, of the At-IE protein.

Biochemistry and genetics of His biosynthesis (Fig. 1) have been extensively studied in a number of microorganisms (Winkler, 1987; Alifano et al., 1996). In eubacteria such as Escherichia coli and Salmonella typhimurium, the complete nucleotide sequences of the His operons have been determined (Carlomagno et al., 1988), and it was shown that eight structural genes are organized in a single operon encoding all of the enzymes catalyzing the 11 steps of the pathway (Carlomagno et al., 1988). In Lactococcus lactis the His biosynthetic genes appeared to be clustered in an operon containing several ORFs of unknown function (Delorme et al., 1992), whereas in archaeabacteria such as Methanococcus vannielii and Methanococcus jannaschii, these genes are scattered throughout the chromosome (Beckler and Reeve, 1986; Bult et al., 1996).

The complete genomic nucleotide sequence of Synecho- cystis sp. PCC6803 has recently been determined, and it was found that His biosynthetic genes do not organize an operon (Kaneko et al., 1996). In lower eukaryotes, including Saccharomyces cerevisiae, the His biosynthetic genes are found in different loci (Mortimer et al., 1994). It has also been established that several of these genes encode multifunctional enzymes (Alifano et al., 1996): hisIE encodes PRA-CH and PRA-PH, hisB codes for imidazoleglycerol-phosphate dehydratase and histidinolphosphate phosphatase, and hisD codes for histidinol dehydrogenase in E. coli and S. typhimurium (Carlomagno et al., 1988). However, the hisI and hisE reactions in Azospirillum brasilense (Fani et al., 1993) and also in some archaeabacteria (Beckler and Reeve, 1986, Bult et al., 1996) are catalyzed by separate protein molecules. On the other hand, multifunctional enzymes with the activities corresponding to the hisIE and hisD proteins are encoded by HIS4 in S. cerevisiae (Donahue et al., 1982) and Pichia pastoris (Crane and Gould, 1994), the his7" gene of Schizosaccharomyces pombe (Apolinaro et al., 1993), and the his-3 gene of Neurospora crassa (Legerton and Yanofsky, 1985). Genetic analysis of yeast his4 mutants suggested that the HIS4 protein can be divided into three subdomains, HIS4A, HIS4B, and HIS4C, which correspond to hisI, hisE, and hisD, respectively (Donahue et al., 1982).

In the past several years we have isolated cDNAs encoding enzymes involved in higher-plant His biosynthesis. They are the histidinol dehydrogenase from Brassica oleracea (Nagai et al., 1991) and the imidazoleglycerolphosphate dehydratase from Arabidopsis and wheat (Tada et al., 1994). Here we report the isolation of an Arabidopsis cDNA that encodes a bifunctional protein (At-IE) that has both PRA-PH and PRA-CH activities through genetic complementation of an E. coli hisI mutant defective in the-PRA-CH activity. Furthermore, we isolated and characterized a single-copy gene coding for the At-IE protein. Analysis of the At-IE gene and recombinant enzyme expression...

* Corresponding author; e-mail ohtad@orange.ocn.ne.jp; fax 81-866-56-9454.
studies revealed that the N-terminal and the C-terminal halves of the At-IE protein correspond to PRA-CH and PRA-PH, respectively.

MATERIALS AND METHODS

Plant Materials, Bacterial Strains, and Culture Media

Arabidopsis ecotype Columbia seedlings (Col-0, Lehle Seeds, Tucson, AZ) were germinated on germination medium 0.8% (w/v) agar plates (Valvekens et al., 1988) under sterile conditions, and the seedlings were cultivated in a growth chamber maintained at 23°C and 80% RH with a 16-h light/8-h dark cycle with the light intensity of 150 μE.

Bacterial strains used in this study are listed in Table I. Escherichia coli JM109 strain was used as the host for the propagation and manipulation of plasmid DNA. Luria-Bertani medium and M9 minimal medium for E. coli were prepared as described previously (Sambrook et al., 1989).

Isolation of an Arabidopsis cDNA Encoding PRA-CH (hisI) and PRA-PH (hisE)

A cDNA library of 7-d-old Arabidopsis seedlings (Mizutani et al., 1997) was converted to a phagemid stock by in vivo excision according to the manufacturer’s instructions (Stratagene). E. coli UTH903 cells were transformed with 5 μg of the CDNA library phagemid stock (1.9 × 10⁸ cells/μg plasmid). The transformation mixture was plated on M9 minimal plates containing 0.2% (w/v) Glc supplemented with 100 μg mL⁻¹ ampicillin, 25 μg mL⁻¹ streptomycin, and an amino acid mixture without L-His (Sambrook et al., 1989), and incubated at 37°C for 2 d. Plasmids were recovered from purified his+ colonies and tested for their ability to suppress the His auxotrophy of strain UTH903. The clone containing the longest insert (pKF323 = pAt-IE) was identified after restriction-enzyme analysis, and its DNA sequence was completely determined.

DNA Sequencing

Nucleotide sequences were determined from both strands by the dideoxy chain-termination method (Sanger et al., 1977) using a dye terminator cycle sequencing kit (PRISM, Applied Biosystems). Nucleotide and amino acid sequences were analyzed using DNASTAR version 3.4 software (Hitachi Software Engineering Co., Yokohama, Japan) and by performing the BLAST search (Altschul et al., 1990) of the National Center for Biotechnology Information.

Nucleic Acid Hybridization Analysis

Genomic DNA (10 μg) was prepared from 4-week-old Arabidopsis seedlings as described previously (Sambrook et al., 1989). After digestion with the restriction enzymes, DNA fragments were separated electrophoretically in a 0.7% (w/v) agarose gel and transferred to a Hybond N nylon membrane (Amersham) in 0.4 n NaOH (Sambrook et al., 1989). Hybridization was performed using the At-IE cDNA as a probe at 37°C overnight in a solution containing 40% (v/v) formamide, 5× Denhardt’s solution, 6× SSC, 0.5% (w/v) SDS, and 100 μg mL⁻¹ sheared salmon-sperm DNA (Sigma; Sambrook et al., 1989). The membrane was washed twice in 2× SSC/0.1% (w/v) SDS at room temperature for 10 min and then twice in 0.5× SSC/0.1% (w/v) SDS at 50°C for 15 min. Blots were exposed to a Hyperfilm-MP (Amersham) for 2 d at −80°C using an intensifying screen.

For RNA-blot analysis, total RNA was prepared as described previously (Lagrimini et al., 1987), and 10-μg aliquots of the sample were electrophoretically separated in a 2.2 m formaldehyde-1.2% (w/v) agarose gel in Mops buffer (Sambrook et al., 1989) and then transferred to a Hybond N nylon membrane in 6× SSC. The At-IE cDNA was labeled by the random-priming method (Feinberg and Vogelstein, 1983) using [α-32P]dCTP. Hybridization was carried out at 42°C overnight in a solution consisting of 50% (v/v) formamide, 5× Denhardt’s solution, 6× SSC, 0.5% (w/v) SDS, and 100 μg mL⁻¹ sheared salmon-sperm DNA (Sigma; Sambrook et al., 1989). The blots were washed twice in 2× SSC/0.1% (w/v) SDS at room temperature for 10 min and then twice in 0.2× SSC/0.1% (w/v) SDS at 55°C for 15 min. Blots were analyzed using a bioimaging analyzer (BAS2000, Fuji Photo Film Co., Tokyo, Japan).
Table 1. Summary of the E. coli strains, plasmids, and oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Material</th>
<th>Property</th>
<th>Reference or Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli strain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UTH903</td>
<td>λ^- hisI903 rpsL145(strE) malT1(f') malT1(f') yx1A5 mtl-1</td>
<td>CGSC*</td>
</tr>
<tr>
<td>JM109</td>
<td>recA1 endA1 gyrA96 thi hsdR17 supE44 relA1 Δ(lac-proAB)/F' (traD36 proAB lacI lacZ ΔM15)</td>
<td>Sambrook et al. (1989)</td>
</tr>
<tr>
<td>XLI-Blue</td>
<td>hsdR17 endA1 supE44 thi relA1 lacI Δ(lac-proAB) lacI lacZ ΔM15 .::Tn10</td>
<td>Stratagene</td>
</tr>
<tr>
<td>BL21</td>
<td>F' ompT[lacI] hsdS</td>
<td>Pharmacia</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasmid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pKF323</td>
<td>pBluescript bearing a 1.1-kb fragment for the At-IE cDNA (pAt-IE)</td>
<td>This study</td>
</tr>
<tr>
<td>pKF347</td>
<td>pMal-c2 bearing a full-length ORF</td>
<td>This study</td>
</tr>
<tr>
<td>pKF360</td>
<td>pBluescript bearing a 5.5-kb fragment for the At-IE gene</td>
<td>This study</td>
</tr>
<tr>
<td>pKF362</td>
<td>pMal-c2 bearing an N-terminal segment without its putative chloroplast transit peptide</td>
<td>This study</td>
</tr>
<tr>
<td>pKF363</td>
<td>pMal-c2 bearing a C-terminal segment</td>
<td>This study</td>
</tr>
<tr>
<td>pKF371</td>
<td>pMal-c2 bearing an N-terminal segment</td>
<td>This study</td>
</tr>
<tr>
<td>pKF372</td>
<td>pMal-c2 bearing an ORF without its putative chloroplast transit peptide region</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oligonucleotide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRI28b</td>
<td>5'-CGGAATTCGGGACTTCGTCGCGGATGCATGTA-3'</td>
<td>This study</td>
</tr>
<tr>
<td>PRI30c</td>
<td>5'-CGCTCCGACAACTTGCCCAAAATCAGCAAC-3'</td>
<td>This study</td>
</tr>
<tr>
<td>PRI44d</td>
<td>5'-GGCGATCCGTCGATGATCTGTCCGACG-3'</td>
<td>This study</td>
</tr>
<tr>
<td>PRI45w</td>
<td>5'-GGCCCTGCGTTCGACTCATCGTTAATGTTA-3'</td>
<td>This study</td>
</tr>
<tr>
<td>PRI46f</td>
<td>5'-GGGCGATCCGCTTCGGAAAACACGCTA-3'</td>
<td>This study</td>
</tr>
</tbody>
</table>

*Corresponding to the bases +1 to +16 of the At-IE gene on pKF360.  
Corresponding to the bases +236 to +2436 of the At-IE gene on pKF360.  
Corresponding to the bases +1343 to +1362 of the At-IE gene on pKF360.  
Corresponding to the bases +1760 to +1742 of the At-IE gene on pKF360.  
Corresponding to the bases +1855 to +1874 of the At-IE gene on pKF360.

Isolation of the At-IE Gene of Arabidopsis

A genomic clone containing a fragment of the At-IE gene was identified by screening approximately 5 × 10^6 plaques of an Arabidopsis genomic DNA library made with AZAPII (Stratagene). Plaques were transferred to Colony/Plaque Screen nylon membranes (NEN) and hybridized overnight at 42°C with the full-length At-IE cDNA as a probe. Membranes were washed once in 2× SSC/0.1% (w/v) SDS at room temperature for 30 min, once in 0.5× SSC/0.1% (w/v) SDS at 50°C for 30 min, and once in 0.2× SSC/0.1% (w/v) SDS at 50°C for 30 min. The blots were then exposed to Hyperfilm-MP films for 16 h. Positive plaques were rescreened until pure phages were obtained.

Expression of the At-IE cDNA in E. coli

For heterologous expression experiments, DNA fragments encoding different domains of the At-IE protein were amplified by PCR using specific sets of primers (Table I) and the At-IE cDNA as a template. The PCR products were double digested with BamHI and XhoI and cloned into a BamHI-SalI-digested pMAL-c2 vector (New England Biolabs, Inc., Beverly, MA) to obtain expression plasmids (Table I). A set of PRI28 and PRI30 was used for the amplification of a 931-bp fragment containing the entire coding region (pKF347); PRI44 and PRI30 for a 778-bp fragment encoding the At-IE protein without the region corresponding to the putative chloroplast transit peptide (pKF372); PRI28 and PRI45 for a 495-bp fragment for an N-terminal segment (pKF371); PRI44 and PRI45 for a 342-bp fragment for an N-terminal domain without the putative chloroplast transit peptide (pKF362); and PRI46 and PRI30 for a 436-bp fragment of a C-terminal segment (pKF363). The expressed recombinant fusion proteins were purified by amylose resin affinity-column chromatography (New England Biolabs.). These expression plasmids were also used to transform E. coli BL21 and UTH903 cells.

Complementation of the E. coli hisI Mutant

Strain UTH903 was transformed with either pAt-IE (= pKF323) or an empty pBluescript SK(-) (pBS SK(-)). pAt-IE (= pKF323) contains a full-length Arabidopsis hisI cDNA. After the transformation, strain UTH903 harboring either an empty pBluescript or pAt-IE was cultivated overnight in M9-Glc medium supplemented with 1 mM histidine at 37°C. After harvesting, E. coli cells were homogenized in 50 mM potassium phosphate (pH 7.5) buffer by mild sonication. After centrifugation at 10,000g for 10 min, the soluble fraction was passed through a Sephadex G-25 column (PD-10, Pharmacia). AICAR production was determined photometrically by measuring A550 by the Bratton-Marshall method (Ames et al., 1961). AICAR of 10 μM solution gave an A550 of 0.270 (Ames et al., 1961) in a 1-cm light path. Protein was assayed by the Bradford (1976) method.

Enzyme Assay

Recombinant proteins corresponding to the putative functional domains were produced as described above and used in a coupling enzyme assay (Ames et al., 1961) to mimic the His biosynthetic pathway (Fig. 1). This coupling assay was used because neither the substrate nor the reac-
tion product of the At-IE was available. The reaction mixture (170 μL) contained 111 mM Tris-Cl (pH 8.5), 22.2 mM MgCl₂, 83.5 mM KCl, 5.6 mM ATP (Sigma), the recombinant proteins to be assayed, and a recombinant HIS1 protein of *Saccharomyces cerevisiae* and a recombinant hisA protein of *E. coli* as the coupling enzymes. These recombinant HIS1 and hisA proteins were produced in *E. coli* as fusion proteins with the maltose-binding protein using a pMAL-c2 vector (New England Biolabs) and affinity purified using amylose resin columns (New England Biolabs). The reaction was started by the addition of 10 μL of 10 mM PRPP (Sigma), and the reaction mixture was incubated at 30°C for 15 min. In this assay system, BBM III was produced from ATP and PRPP according to the His biosynthetic scheme (Fig. 1) through the activities of the HIS1 (the first step of His pathway), Arabidopsis At-IE (the second and third steps), and the hisA (the fourth step) proteins. The BBM III produced was hydrolyzed to AICAR in HCl at 95°C. AICAR was determined as described above.

**RESULTS**

**Cloning of an Arabidopsis cDNA That Suppresses an *E. coli* hisI Mutation**

An Arabidopsis cDNA encoding a bifunctional enzyme with PRA-PH and PRA-CH activities was isolated through genetic complementation of a bacterial His auxotrophic mutant. Thus, an *E. coli* hisI defective mutant (UTH903) was transformed with a phagemid library prepared from 7-d-old Arabidopsis seedlings, and 20 prototrophic colonies of 9.5 × 10⁷ transformants were identified after cultivating for 2 d on M9 minimal agar plates. Upon retransformation of strain UTH903, 16 of the isolated 20 plasmids were found to be able to suppress the His auxotrophy. The DNA inserts of these 16 plasmids exhibited identical restriction patterns, and DNA sequencing showed that they were derived from the same cDNA fragment (data not shown). The remaining four of the identified His prototrophic colonies might be revertants, since their plasmids contained DNA inserts of inconsistent nucleotide sequences and failed to recomplement the His auxotrophy (data not shown). One of the plasmids (pKF323 = pAt-IE) containing the longest insert was sequenced completely and used for further analyses.

Strain UTH903 was transformed with either pAt-IE or an empty pBluescript, and crude cell extracts were prepared for an AICAR production assay. The AICAR production observed with the cells transformed with pAt-IE (2.06 ± 0.19 nmol mg⁻¹·protein min⁻¹) was comparable to that with XL1-Blue as a control, whereas no AICAR production was detected with the UTH903 transformed with an empty pBluescript. The results were consistent with the complementation experiments (Fig. 2) in which the UTH903 transformed with pAt-IE was able to grow on M9-Glc minimum medium but no bacterial growth was observed when transformed with a pBluescript empty vector.

The At-IE cDNA contained an ORF of 843 bp encoding a polypeptide of 281 amino acids with a calculated molecular mass of 31,666 Da (Figs. 3 and 4). Nucleotide sequence analysis showed that the consensus motif surrounding a translation initiation codon (AACAAATGGC) in plants (Lütcke et al., 1987) was well conserved as TAAAATGGC in the At-IE cDNA. Several consensus sequences required for the correct 3' end formation of transcripts in plants were also found in the 3' untranslated region (Fig. 3). Thus, a putative polyadenylation signal sequence, AATAAA (Wahle and Keller, 1992), was found 23 bp upstream from the adenylation tail. The TTTGTA motif, which is considered to be involved in the stability of transcripts (Rothnie et al., 1994), was also identified at position +2363 (Fig. 3).

The predicted primary structure of the At-IE protein was compared with those of microorganisms available in the nucleotide database (Fig. 4). Sequence alignment indicated that an N-terminal segment (spanning residues Gly-65 to Phe-158) of the At-IE protein was highly homologous to the conserved region among the hisI proteins so far reported and that the C-terminal region encompassing residues Leu-179 to Arg-267 was homologous to the domain conserved among the microbial hisE proteins (Fig. 4). These results indicate that the At-IE cDNA encoded a bifunctional protein of PRA-CH and PRA-PH, of which domain organization has also been found in the PRA-CH (hisI) and PRA-PH
hisE) enzymes of eubacteria and lower eukaryotes but not in archaebacteria. This putative domain structure of the At-IE protein was investigated through both the gene structure analysis and the recombinant protein expression studies, as described later. It was also found that the N-terminal portion of approximately 50 amino acids showed no significant homology to those of the hisIE proteins from microorganisms but showed the properties characteristic of chloroplast transit peptides (von Heijne and Nishikawa, 1991).

RNA-Blot Analysis

Upon northern-blot analysis, the transcript size of the At-IE gene appeared to be approximately 1.2 kb (Fig. 5), which was in good agreement with the predicted size from the At-IE cDNA (Fig. 3). The At-IE gene was expressed ubiquitously in plants throughout development. The highest expression level for the At-IE mRNA was observed in roots of 3-week-old plants and in inflorescence stems of 4-week-old plants.

Cloning and Sequencing of an At-IE Genomic Clone

To examine the number of the At-IE genes in Arabidopsis, Southern-blot analysis was performed using the At-IE cDNA as a probe, and a phagemid harboring a 5.5-kb EcoRI-EcoRI fragment was identified to contain the At-IE gene (Figs. 3 and 7). The At-IE gene consists of five exons divided by four introns (Figs. 3 and 7). Intron-splice sites of all of the introns follow the “GU-AG” rule, which is observed at intron-splice sites of all eukaryotic organisms, including higher plants (Breathnach and Chambon, 1981; Simpson and Filipowicz, 1996). Amino acid sequence alignment (Figs. 3 and 4) showed that intron 1 was located at the junction of the putative transit peptide portion encoded by exon 1 and the PRA-CH (hisI) domain encoded by exons 2 and 3. It was also found that intron 3 was located at the putative binding site for the "GU-AG" rule, which is observed at intron-splice sites of all eukaryotic organisms, including higher plants (Breathnach and Chambon, 1981; Simpson and Filipowicz, 1996). Amino acid sequence alignment (Figs. 3 and 4) showed that intron 1 was located at the junction of the putative transit peptide portion encoded by exon 1 and the PRA-CH (hisI) domain encoded by exons 2 and 3. It was also found that intron 3 was located at the putative binding site for the "GU-AG" rule, which is observed at intron-splice sites of all eukaryotic organisms, including higher plants (Breathnach and Chambon, 1981; Simpson and Filipowicz, 1996).
ary between the PRA-CH (hisI) domain and the PRA-PH (hisE) domain, which was encoded by exons 4 and 5. Two homologous regions among the bacterial hisIE proteins so far reported were also found to be conserved in the At-IE protein (Fig. 4).

In the 5' untranslated region several putative regulatory elements were found (Fig. 3). Possible TATA and CAAT elements were identified at positions -2166 and -2195, respectively. Furthermore, a sequence motif, TAACTC, similar to the S. cerevisiae GCN4-responsive element, GCRE (Arndt and Fink, 1986), was located at position -2288 (Fig. 3).

Characterization of the Domain Structure of the At-IE Protein

Sequence comparison (Fig. 4) suggested that the At-IE protein molecule consisted of a putative chloroplast transit peptide and two separate catalytic domains corresponding to the PRA-CH (hisI) and PRA-PH (hisE) proteins, respectively. This overall putative domain structure of the At-IE protein was confirmed through heterologous expression studies.

Both pKF347 and pKF372 were able to suppress the His auxotrophy of E. coli strain UTH903 (Fig. 7). The insert of pKF372 encoded an At-IE protein of which the N-terminal extension had been truncated (Fig. 7). Therefore, the successful suppression of the His auxotrophy of UTH903 with pKF372 indicated that the N-terminal extension was not essential for the catalytic activity, supporting the idea that

Figure 4. Alignment of the amino acid sequence predicted from the Arabidopsis At-IE cDNA and the corresponding proteins of microbial origins. Ec, E. coli (accession no. X13462; Carlomagno et al., 1988); Sy, Synechocystis sp. PCC6803 (accession no. D90917; Kaneko et al., 1996); Rs, Rhodobacter sphaeroides (accession nos. X87256 and X82010; Oriol et al., 1996); Mj, Methanococcus jannaschii (accession nos. U67484 and U67585; Bult et al., 1996); Sc, S. cerevisiae (accession no. J01331; Donahue et al., 1982). Asterisks show the stop codon and dashes inserted to maximize the alignment. Residues conserved among all of the compared sequences are shaded.

Figure 5. RNA-blot analysis of the At-IE mRNA levels. Lane 1, One-week-old plants; lane 2, roots from 2-week-old plants; lane 3, leaves from 2-week-old plants; lane 4, roots from 3-week-old plants; lane 5, leaves from 3-week-old plants; lane 6, roots from 4-week-old plants; lane 7, leaves from 4-week-old plants; lane 8, siliques from 4-week-old plants. Membrane was hybridized with a 32P-labeled Psrl-EcoRV fragment of the At-IE cDNA. Total RNA (10 μg) prepared from Arabidopsis seedlings was electrophoresed in each lane. The photograph of the ethidium bromide-stained gel for the blotting is also shown at the bottom of the RNA-blot analysis.

Figure 6. Genomic Southern-blot analysis. Genomic DNA (10 μg) was prepared from Arabidopsis leaves and was digested with restriction enzymes (B, BamHII; Bg, BglII; E, EcoRI; Hc, HindIII; and Xb, XbaI). Hybridization was performed using a 32P-labeled Psrl-EcoRV fragment of the At-IE cDNA. The λ-DNA digested with HindIII is shown as a molecular size marker.
this N-terminal extension corresponded to a chloroplast transit peptide. The expressed recombinant protein using pKF372 (Fig. 7) was enough to support the AICAR production in the assay system containing the recombinant HIS1 protein of \textit{S. cerevisiae} and the hisA protein of \textit{E. coli} (Fig. 8). Thus, pKF372 contained the cDNA that encodes a protein catalyzing the hisIE reactions (Fig. 1). However, the recombinant protein prepared with pKF362 (the N-terminal domain), which was able to complement the hisI mutation of UTH903 (Fig. 7), did not work for the AICAR production (Fig. 8). Thus, the N-terminal domain catalyzed the hisI (PRA-CH) reaction but did not have hisE (PRA-PH) activ-

---

**Figure 7.** Complementation of the hisI mutation of strain UTH903 with putative functional domains of PRA-PH and PRA-CH. A, Exon-intron relationship between the At-IE gene structure and the At-IE cDNA for the bifunctional PRA-PH: PRA-CH protein is shown schematically. B, Expression plasmids were designed to contain the putative catalytic domains and used for the complementation assay for the His auxotrophy of the \textit{E. coli} hisI mutant. The symbols + and \(-\) indicate the ability and inability, respectively, of the plasmids to suppress the \textit{E. coli} UTH903 hisI mutation. \textit{E. coli} UTH903 cells were transformed with pKF347, representing the full-length of the At-IE cDNA; pKF372 carrying the full-length insert truncated in its putative chloroplast transit sequence; pKF371, corresponding to the N-terminal segment; pKF362 for the N-terminal segment without the putative chloroplast transit sequence; or pKF363 for the C-terminal half of the At-IE protein. After transformation, cells were plated onto M9-Glc minimal agar plates. The portion of the putative chloroplast transit sequence is shaded.

**Figure 8.** At-IE-dependent AICAR production determined in the assay mixture containing \textit{S. cerevisiae} HIS1, the hisA protein of \textit{E. coli}, and one of the expressed recombinant proteins. The At-IE protein without the putative chloroplast transit peptide, the putative PRA-PH (hisE), and the putative PRA-CH (hisI) domains were expressed as the fusion proteins with a maltose-binding protein using a pMAL-c2 bacterial expression vector. The expression vectors used were the same as those presented in Figure 7.
ity. The cDNA insert of pKF363 was derived from exons 4 and 5 coding for the C-terminal domain (Fig. 7). This plasmid failed to suppress the His auxotrophy of UTH903 (Fig. 7) and could not support the AICAR production (Fig. 8), indicating that the C-terminal domain did not catalyze the hisI (PRA-CH) reaction but was involved in the hisE (PRA-PH) reaction. On the other hand, the AICAR production was reconstituted when the recombinant proteins encoded by pKF362 and pKF363 were mixed in the reaction mixture (Fig. 8), indicating that the C-terminal domain corresponded to the hisE (PRA-CH) domain. No histidinol dehydrogenase activity was observed with the protein produced with pKF347 (data not shown).

These results have demonstrated that the single-copy gene, the At-IE gene of Arabidopsis, encodes the bifunctional protein of which the N-terminal and C-terminal domains separately catalyze the two successive reactions of PRA-CH (hisI) and PRA-PH (hisE), respectively, in the His biosynthetic pathway. Also, it was found that the At-IE protein, like most bacterial enzymes, was not accompanied by a histidinol dehydrogenase domain, which is encoded by HIS4 in S. cerevisiae (Donahue et al., 1982).

DISCUSSION

A number of eukaryotic genes, including those of plant origin, have been isolated by performing heterologous genetic complementation of E. coli or S. cerevisiae mutants (Minet et al., 1992; Senecoff and Meagher, 1993; Tada et al., 1994). This method was also successful in isolating the At-IE cDNA from Arabidopsis with the use of strain UTH903 defective in hisI (PRA-CH) activity, which encodes a bifunctional protein with hisI (PRA-CH) and hisE (PRA-PH) activities. The amino acid sequence predicted from the At-IE cDNA is significantly homologous to both the His biosynthetic enzymes, imidazolglycerolphosphate dehydratase (Tada et al., 1995) and histidinol dehydrogenase (Nagai et al., 1993), are localized in chloroplasts. Furthermore, the cDNAs encoding N-[5’-PR]-formimidino]-5-aminoimidazole-4-carboxamide ribonucleotide isomerase and Gln amidotransferase/cyclase of Arabidopsis have also been shown to contain the regions corresponding to putative N-terminal transit peptides (K. Fujimori and D. Ohta, unpublished results). It is therefore possible that the entire His biosynthesis may be completed in chloroplasts, whereas the first enzyme of the His pathway, ATP-PR transferase, has not been characterized yet. His biosynthesis is an extremely energy-consuming process, which requires 41 ATP molecules for each His molecule produced (Alifano et al., 1996). In other words, the compartmentalization of the His pathway in chloroplasts is favorable to ensure efficient energy supply. At least six genes encoding eight steps of His biosynthesis in plants exhibited constitutive expression patterns throughout development, and there were no clear tissue-specific expression patterns (Nagai et al., 1993; Tada et al., 1995; K. Fujimori and D. Ohta, unpublished results). In microorganisms, His biosynthesis is regulated through controlled gene expression and the feedback regulation of ATP-PR transferase activity by L-His (Alifano et al., 1996). Isolation and characterization of ATP-PR transferase are essential for understanding the mechanism that regulates higher-plant His biosynthesis.

ACKNOWLEDGMENT

We gratefully acknowledge Nobuko Udome for skillful experimental assistance.

Received March 10, 1998; accepted June 16, 1998.
Copyright Clearance Center: 0032-0889/98/118/0275/09.
The nucleotide sequence data reported in this paper will appear in the nucleotide sequence databases with the accession nos. AB006082 (At-IE cDNA) and AB006083 (At-IE gene).

LITERATURE CITED

Isolation and Characterization of hisIE/HIS4 in Arabidopsis


Arndt K, Fink GR (1986) GGN4 protein, a positive transcription factor in yeast, binds general control promoters at all 5′-TGACTC 3′ sequences. Proc Natl Acad Sci USA 83: 8516–8520


Gantt JS, Baldauf SL, Calie PJ, Weeden NF, Palmer JD (1991) Transfer of rpl22 to the nucleus greatly preceded its loss from the chloroplast and involved the gain of an intron. EMBO J 10: 3073–3078


Rothnie HM, Reid J, Hohn T (1994) The contribution of AUAAAA and the upstream element UUUUGA to the efficiency of mRNA 3′-end formation in plants. EMBO J 13: 2200–2210


Copyright © 1998 American Society of Plant Biologists. All rights reserved.