京都大学
KYOTO UNIVERSITY

Title	Studies on the Histidine Biosynthesis in Arabidopsis thaliana(Dissertation_)
Author(s)	Fujimori, Ko
Citation	Kyoto University ( )
Issue Date	1999-01-25
URL	http://dx.doi.org/10.11501/3147528
Right	
Type	Thesis or Dissertation
Textversion	author

Studies on the Histidine Biosynthesis

in *Arabidopsis thaliana*

Ko Fujimori

1998

Contents

General Introduction

Many studies in the biochemistry and genetics of amino acid biosynthesis have begun in the 1950's. Majority of early investigations concerning biochemical mechanisms have been carried out in prokaryotic organisms. Recent contributions using genetic approaches have added novel informations.

Histidine (His) is one of the essential amino acids in mammal. The pathway of the His biosynthesis has been studied extensively in enterobacteriums such as *Salmonella typhimurium* and *Escherichia coli* since more than 30 years ago, biochemically and genetically (Winkler, 1987). In these species, His is synthesized from ATP and phosphoribosyl pyrophosphate (PRPP) through a complex pathway which does not involve any branch points, and does not form any other metabolically important products. In *E. coli* and *S. typhimurium*, the His biosynthetic genes were organized in an operon (Goldschmidt et al., 1970, Bruni et al., 1980), like lactose operon (Newton et al., 1965) and tryptophan operon (Imamoto et al., 1966). Carlomagno et al. was determined the complete nucleotide sequences of the whole His biosynthetic genes of these species (1988). In most prokaryote, the His operon consists of a cluster of nine structural genes (cistrons) (Carlomagno et al., 1988, Fleischmann et al., 1995). Early investigations in microorganisms were about the elucidation of regulation mechanisms of the His biosynthesis at the protein level (feedback inhibition, allosteric regulation) (Ames et al., 1961, Martin, 1963, Klungsoyr et al., 1968, Brenner and Ames, 1971) and at the DNA level (attenuation control) (Kasai, 1974, Artz and Broach, 1975, Johnson et al., 1980, Blasi and Bruni, 1981). To date, in microorganisms, many informations about the regulation mechanisms of the His biosynthetic pathway have been accumulated (Alifano et al., 1996). On the contrary, the His biosynthetic pathway in higher plants remains unclear.

The histidine biosynthesis in microorganisms

The studies on the His biosynthesis were started in enterobacteriums such as

1
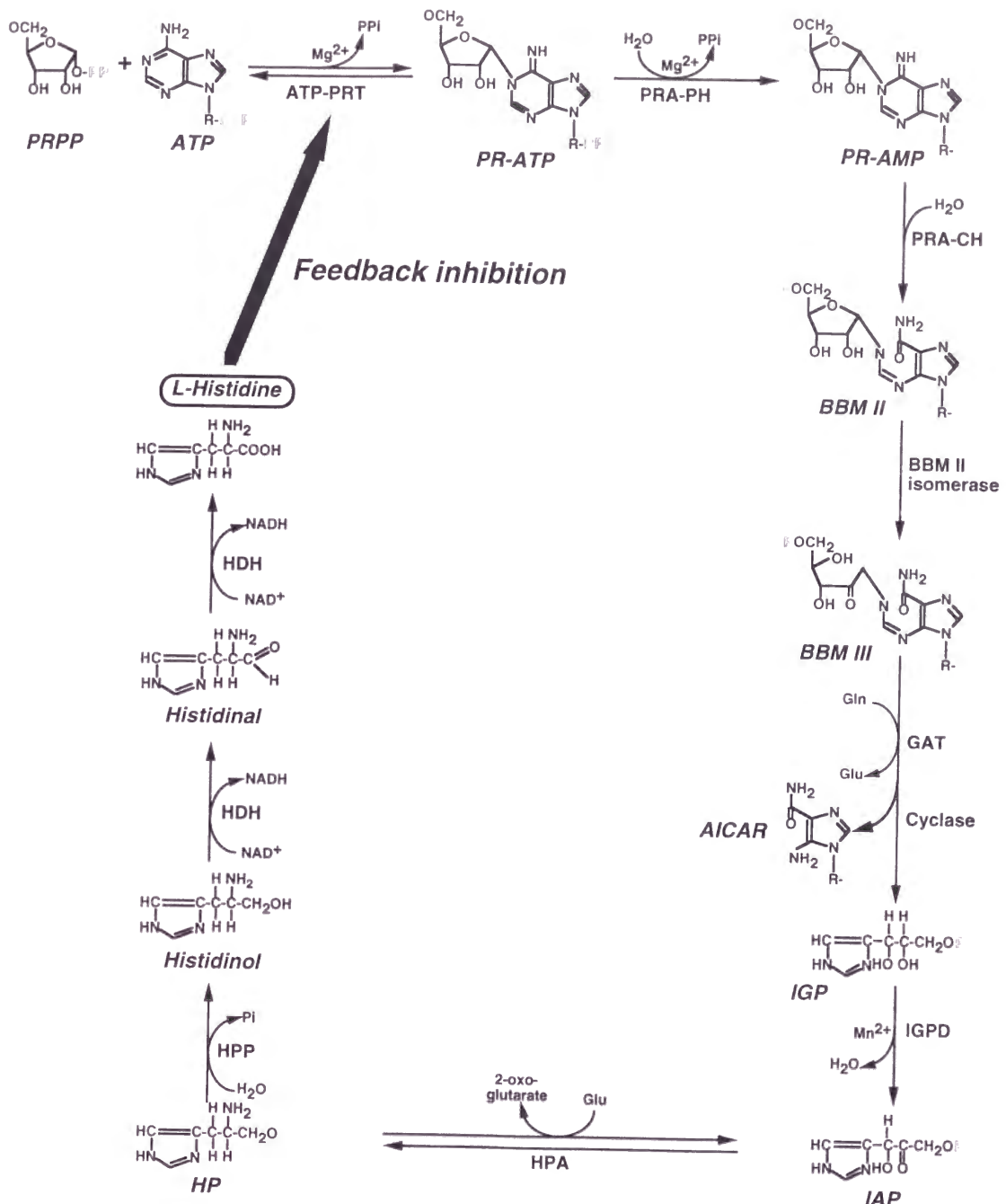
Fig. 1. Histidine biosynthetic pathway in microorganisms

PRPP; phosphoribosyl pyrophosphate, PR-ATP; phosphoribosyl-ATP, PR-AMP; phosphoribosyl-AMP, BBM II; N'-[(5'-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide ribonucleotide, BBM III; N'-[(5'-phosphoribulosyl)-formimino]-5-aminoimidazole-4-carboxamide ribonucleotide, AICAR; 5-aminoimidazole-4-carboxamide-1-β-D-ribofuranoside, IGP; imidazoleglycerolphosphate, IAP; imidazoleacetolphosphate, HP; histidinolphosphate, ATP-PRT; ATP phosphoribosyl transferase, PRA-PH; phosphoribosyl-ATP pyrophosphohydrolase, PRA-CH; phosphoribosyl-AMP cyclohydrolase, GAT; glutamine amidotransferase, IGPD; imidazoleglycerolphosphate dehydratase, HPA; histidinolphosphate aminotransferase, HPP; histidinolphosphate phosphatase, HDH; histidinol dehydrogenase.

E. coli and *S. typhimurium* (Winkler, 1987, Alifano et al., 1996). The His biosynthetic pathway is well-characterized, which is an unbranched pathway consisting of ten enzymatic steps catalyzed by eight enzymes (Fig. 1). The His biosynthetic pathway in these organisms appears to be basically same. The first step in the His biosynthetic pathway is a Mg^{2+} ion-dependent reversible reaction catalyzed by ATP phosphoribosyl transferase (ATP-PRT; EC 2.4.2.17) encoded by *hisG* in prokaryote. Next two reactions are also required for Mg^{2+} ion. Phosphoribosyl (PR)-ATP (PR-ATP) is hydrolyzed to PR-AMP by phosphoribosyl-ATP pyrophosphohydrolase (PRA-PH; EC 3.5.4.19) encoded by *hisE*, and PR-AMP cyclohydrolase (PRA-CH; EC 3.5.4.19) encoded by *hisI* opens the purine-ring. In enterobacteriums, these two enzymes are encoded by a single structural gene, *hisIE*. On the contrary, in archaebacteria, these two enzymes are encoded by independent two genes. The following step is an internal redox reaction known as Amadori rearrangement (Smith and Ames, 1964) catalyzed by *N'*-[(5'-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (BBM II) isomerase encoded by *hisA*. Next two steps is still unclear containing the intermediates of unknown structure (Rieder et al., 1994). *N'*-[(5'-phosphoribulosyl)-formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (BBM III) is cleaved to yield imidazoleglycerolphosphate (IGP), and 5-aminoimidazole-4-carboxamide-1-β-D-ribofuranoside (AICAR) which is the intermediate of the purine biosynthesis. Glutamine amidotransferase (GAT) encoded by *hisH* catalyzes the introduction of nitrogen atom from glutamine, followed by forming of imidazole-ring catalyzed by cyclase encoded by *hisF*. The following reaction is Mn^{2+} ion-dependent dehydration catalyzed by imidazoleglycerolphosphate dehydratase (IGPD; EC 4.2.1.19) encoded by *hisB*. The resultant imidazoleacetolphosphate (IAP) is done reversibly the transamination with a nitrogen atom from glutamate catalyzed by histidinolphosphate aminotransferase (HPA; EC 2.6.1.9) encoded by *hisC*. Histidinolphosphate (HP) is dephosphorylated by histidinolphosphate phosphatase (HPP; EC 3.1.3.15) which is one activity of bifunctional hisB protein. The last two steps are NAD^+-dependent oxidation to give a final product, L-His, catalyzed by histidinol dehydrogenase (HDH; EC 1.1.1.23) encoded by *hisD*.

Studies using many mutants of *S. typhimurium* demonstrated that the His

biosynthetic genes were clustered, differing from the fungi systems (Ames et al., 1960, Hartmann et al., 1960). Deletion analysis of 5'-regulatory region of the His operon becomes the complete loss of function of all His genes (Ames et al., 1963). The His operon was used for the model system, together with *lac* (Newton et al., 1965) and *trp* (Imamoto et al., 1966) operons. Recently, molecular biological techniques contributed to provide many informations about the His biosynthesis in microorganisms (Alifano et al., 1996). Many His genes have been identified by restoring the His auxotrophy of the microbial His-requiring mutants by cDNAs from various species. Nucleotide sequences of the whole His biosynthetic genes of *E. coli* and *S. typhimurium* were determined (Carlomagno et al., 1988). A lot of genes involved in the His biosynthetic pathway has been identified from many microorganisms; in prokaryotes including the eubacteria such as *Klebsiella pneumoniae* (Rodriguez and West, 1984), *Bacillus subtilis* (Henner et al., 1986), *E. coli* (Carlomagno et al., 1988), *S. typhimurium* (Carlomagno et al., 1988), *Azospirillum brasilense* (Fani et al., 1989), *Halobacterium volcanii* (Conover and Doolittle, 1990), *Streptomyces coelicolor* (Limauro et al., 1990), *Lactococcus lactis* (Delorme et al., 1992), *Acetobacter pasteurianus* (Takemura et al., 1993), *Haemophilus influenzae* (Fleischmann et al., 1995), *Zymomonas mobilis* (Gu et al., 1995), *Pseudomonas aeruginosa* (de Kievit et al., 1995), *Synechocystis* sp. PCC6803 (Kaneko et al., 1996), *Mycobacterium smegmatis* (Hinshelwood and Stocker, 1992), and archaebacteria such as *Methanococcus voltae* and *M. vannielii* (Cue et al., 1985), *M. thermolithotrophicus* (Weil et al., 1987), *M. jannaschii* (Bult et al., 1996), as well as lower eukaryote such as *Neurospora crassa* (Legerton and Yanofsky, 1985), *Candida maltosa* (Hikiji et al., 1989), *Saccharomyces kluyveri* (Weinstock and Strarthern, 1993), *Pichia pastoris* (Crane and Gould, 1994), *Candida albicans* (Pla et al., 1995), *Shizosaccharomyces pombe* (Erickson and Hannig, 1995), *Saccharomyces cerevisiae* (Alifano et al., 1996). In *E. coli, S. typhimurium*, and *H. influenzae*, genes encoding ten enzymes involved in the His biosynthetic pathway are organized in an operon (Carlomagno et al., 1988, Fleischmann et al., 1995). On the contrary, the His biosynthetic genes of other prokaryotes such as *L. lactis, Synechocystis* sp. PCC6803, and *M. jannaschii* are not organized in an operon (Delorme et al., 1992, Bult et al., 1996, Kaneko et al., 1996). In lower eukaryote such as *S. cerevisiae*, the His biosynthetic

4

genes are scattered throughout the chromosomes (Mortimer et al., 1994). Finally, it was found that the structures of the His biosynthetic gene are very variable, although their pathways, and molecular weight and secondly structures of each protein, in microorganisms are almost same.

Regulation of the histidine biosynthesis

The His biosynthesis is the energy consuming pathway requiring 41 ATP for synthesis of one His molecule (Brenner and Ames, 1971). Such energy consuming pathway should be strictly regulated. The His biosynthetic pathway in microorganisms is regulated at the protein level and at the DNA level. The rate of the His biosynthesis is regulated by varying the activity of first step enzyme of the His pathway, ATP-PRT (Ames et al., 1961, Martin, 1963). This enzymatic activity is inhibited by end-product of this pathway, L-His (feedback inhibition) or its derivatives (Ames et al., 1961, Martin, 1963, Klungsoyr et al., 1968, Brenner and Ames, 1971). Another way to transcriptionally control the regulation of the expression level by intracellular charged histidyl-tRNA^{His} level (Brenner and Ames, 1971), which was also found as the attenuation control (Kasai, 1974, Artz and Broach, 1975). In addition, transcriptional initiation at the His operon promoter is positively regulated by increasing guanosine 3', 5'-diphosphate (ppGpp) concentration in the absence of His (Morton and Parsons, 1977). In lower eukaryote *S. cerevisiae*, the general amino acid control response is mediated by a transcriptional activator, GCN4, that binds to GCRE (GCN4 responsive element) elements located upstream of many amino acid biosynthetic genes (Hinnebusch, 1988). Transcriptional activation by GCN4 leads to increase the expression level of those genes. The expression level of *GCN4* gene itself is up-regulated in response to amino acid starvation. On the contrary, amino acids in the medium cause to repress the expression of those genes. Such regulatory mechanism was also found in not only *S. cerevisiae* but *S. pombe* (Erickson and Hannig, 1995). In other species, *K. pneumoniae*, which is closely related to *E. coli* and *S. typhimurium*, has the well-conserved regulatory mechanism of the His biosynthesis (Rodriguez and West, 1984). In *S. coelicolor*, the His gene expression is seemed to be regulated by intracellular His levels

(Limauro et al., 1990). Finally, the regulation mechanisms of the His biosynthetic genes from another prokaryotes and eukaryotes have not elucidated yet.

The histidine biosynthesis in higher plants

The His biosynthesis in higher plants has been poorly understood (Table 1). In higher plants, the following enzymatic activities have been detected in shoot extracts from barley, oat and pea; ATP-PRT, IGPD, and HPP (Wiater et al., 1971). However, the difficulty of conventional enzyme study was attributed to the complex nature of the intermediates of the His biosynthetic pathway. The HDH from *Brassica oleracea* (Nagai and Scheidegger, 1991) and the IGPD from wheat germ (Mano et al., 1993) were

Table 1.　His genes and enzymes in different organisms

Gene			Enzyme
Procaryote[a]	Yeast[b]	Plant	
hisG	HIS1	n.i.[f]	ATP phosphoribosyl transferase (ATP-PRT)
hisE	HIS4B	n.i.[f]	Phosphoribosyl-ATP pyrophosphohydrolase (PRA-PH)
hisI	HIS4A	n.i.[f]	Phosphoribosyl-AMP cyclohydrolase (PRA-CH)
hisA	HIS6	n.i.[f]	BBM II isomerase
hisH	HIS7	n.i.[f]	Glutamine amidotransferase (GAT)
hisF	HIS7	n.i.[f]	Cyclase
hisB	HIS3	IGPD[c]	Imidazoleglycerolphosphate dehydratase (IGPD)
hisC	HIS5	HPA[d]	Histidinolphosphate aminotansferase (HPA)
hisB	HIS2	n.i.[f]	Histidinolphosphate phosphatase (HPP)
hisD	HIS4C	HDH[e]	Histidinol dehydrogenase (HDH)

[a] *E. coli* and *S. typhimurium* (Carlomagno et al., 1988), [b] *S. cerevisiae* (Alifano et al., 1996) [c] Tada et al. (1994), [d] El Malki et al. (1998), [e] Nagai et al. (1991), [f] n.i.; not identified

purified, and characterized. The corresponding cDNAs have been also isolated (Nagai et al., 1991, Tada et al., 1994). In prokaryotes and higher plants, HDH is a monofunctional enzyme, on the contrary, in fungi, it's a trifunctional peptide having the HDH, PRA-CH, and PRA-PH activities. In eukaryotes, IGPD is a monofunctional

protein, while, in prokaryotes, IGPD activity attaches with HPP activity. These protein structures are variable among organisms. However, until today, a few genes and proteins have been characterized from higher plants. Recently, El Malki et al. reported that the HPA cDNA was isolated from *Nicotiana tabacum* by genetic complementation using *E. coli hisC* defective mutant (1998). The other steps in this pathway have never been identified yet. Also, its regulation mechanisms of the plant His biosynthesis have not been yet elucidated.

In this study

The biochemical and genetical properties of the His biosynthetic pathway in microorganisms have been extensively studied. On the contrary, the plant His biosynthetic pathway has been little characterized, as described above. I identified the plant His biosynthetic genes for the purpose of the elucidation of the plant His biosynthetic pathway and of its regulation mechanisms. First of all, I isolated four cDNAs and genes involved in the His biosynthesis from *A. thaliana*. A cDNA (*At-IE*) encoding PRA-CH/PRA-PH, which is the second- and third-steps of the His biosynthetic pathway was isolated by functional complementation using *E. coli hisI* mutant. I also isolated *At-IE* gene and characterized the domain structure (Chapter I). A cDNA encoding BBM II isomerase which catalyzes 4th-step was isolated by genetic complementation using *E. coli hisA* defective mutant and characterized (Chapter II). A cDNA (*At-HF*) encoding bifunctional glutamine amidotransferase (GAT) and cyclase which catalyze the 5th- and 6th-steps isolated by library screening. Its function was confirmed by suppression of His auxotrophy of *S. cerevisiae his7* mutant (Chapter III). I also isolated two isoforms for ATP-PRT. Finally, In order to elucidate the regulation mechanisms of the plant His biosynthetic pathway, the expression manner of the plant His biosynthetic genes was investigated. Also, I characterized the plant ATP-PRT proteins using recombinant proteins to know the regulation mechanism at the protein level in the His biosynthesis of higher plants (Chapter IV).

Chapter I

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

Introduction

Biochemistry and genetics of histidine (His) biosynthesis 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, and it was shown that eight structural genes are organized in a single operon encoding all the enzymes catalyzing the eleven steps of the pathway (Carlomagno et al., 1988). In *Lactococcus lactis*, the His biosynthetic genes appeared to be clustered in an operon containing several open reading frames (ORFs) of unknown functions (Delorme et al., 1992), while in archaebacteria such as *Methanococcus vannielii* and *M. jannaschii*, these genes are scattered throughout the chromosome (Beckler and Reeve, 1986, Bult et al., 1996). The complete genomic nucleotide sequence of *Synechocystis* sp. PCC6803 has recently been determined, and it was found that the 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 been also established that several of these genes encode multifunctional enzymes (Alifano et al., 1996): *hisIE* encodes phosphoribosyl-AMP cyclohydrolase (PRA-CH) and phosphoribosyl-ATP pyrophosphohydrolase (PRA-PH); *hisB* codes for imidazoleglycerolphosphate dehydratase (IGPD) and histidinolphosphate phosphatase; *hisD* for histidinol dehydrogenase (HDH) in *E. coli* and *S. typhimurium* (Carlomagno et al., 1988). On the other hand, the hisI and hisE reactions in *Azospirillum brasilense* (Fani et al., 1993) and also in some archaebacteria (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), *his7+* gene of *Schizosaccharomyces pombe* (Apolinario et al., 1993) and *his-3* gene of *Neurospora crassa* (Legerton and Yanofsky, 1985), respectively. Genetic analysis of yeast *his4* mutants suggested that the HIS4 protein can be divided into three subdomains; HIS4A, HIS4B and HIS4C which are corresponding to hisI, hisE, and hisD, respectively (Donahue et al., 1982).

In the past several years, cDNAs encoding enzymes involved in the His biosynthesis of higher plants have been isolated. They are the HDH from *Brasicca oleracea* (Nagai et al., 1991), the IGPD from *Arabidopsis thaliana* and *Triticum aestivum* (Tada et al., 1994), and histidinolphosphate aminotransferase (HPA) from *Nicotiana tabacum* (El Malki et al., 1998). In Chapter I, I describe the isolation of an *Arabidopsis* cDNA encoding a bifunctional protein (At-IE) having both PRA-PH and PRA-CH activities through genetic complementation of an *E. coli hisI* mutant defective in the PRA-CH activity. Further, I isolated and characterized a single copy gene coding for the At-IE protein. Analysis of the *At-IE* gene and recombinant enzyme expression studies revealed that the N-terminal half and C-terminal half of the At-IE protein correspond to PRA-CH and PRA-PH, respectively.

Materials and methods

Plant materials, bacterial strains and culture media

Arabidopsis thaliana ecotype Columbia seedlings (Col-0, Lehle Seeds; Tucson, AZ, USA) were germinated on GM 0.8% (w/v) agar plates (Valvekens et al., 1988) under a sterile condition, and the seedlings were cultivated in a growth chamber maintained at 23 °C and relative humidity of 80% with a 16 hr light / 8 hr dark cycle with the light intensity of 150 $\mu E/m^2/s$. Bacterial strains used in this study are listed in Table 1. *E. coli* JM109 strain was used as the host for the propagation and manipulation of plasmid DNA. LB medium and M9 minimal medium for *E. coli* were prepared as described (Sambrook et al., 1989).

Table 1. Summary of the *E. coli* strains, plasmids, and oligonucleotides used in this study

Material	Property	Reference or source
E.coli strain		
UTH903	λ^- his1903 rpsL145(str^R) malT1(λ^R) xylA5 mtl-1	CGSC[a]
JM109	recA1 endA1 gyrA96 thi hsdR17 supE44 relA1) Δ(lac-proAB)/ F'(traD36 proAB lacI^q lacZΔM15)	Sambrook et al. (1989)
XLI-Blue	hsdR17 endA1 supE44 thi recA1 relA1 lac/F'(proAB lacI^q lacZΔM15::Tn10) gyrA96 relA1 lac[F'proAB lacI^q ZΔM15 Tn10]	Stratagene
BL21	F⁻ ompT[lon] hsdS	Pharmacia
Plasmid		
pKF323	pBluscript bearing a 1.1-kb fragment for the At-IE cDNA (pAt-IE)	This study
pKF347	pMAL-c2 bearing a full-length ORF	This study
pKF360	pBluscript bearing a 5.5-kb fragment for the At-IE gene	This study
pKF362	pMAL-c2 bearing an N-terminal segment without its putative chloroplast transit peptide region	This study
pKF363	pMAL-c2 bearing a C-terminal segment	This study
pKF371	pMAL-c2 bearing an N-terminal segment	This study
pKF372	pMAL-c2 bearing an ORF truncated its putative chloroplast transit peptide region	This study
Oligonucleotide		
PRI28[b]	5'-CGGAATTCGGATCCATGGCGGTATCGTACA-3'	This study
PRI30[c]	5'-CGCTCGAGCAAAGTTGCACAATGAACAAC-3'	This study
PRI44[d]	5'-GCGGATCCGTAGATAACTTGTTGGACCG-3'	This study
PRI45[e]	5'-CGCTCGAGCTACTCATCATTGTTTAATTGA-3'	This study
PRI46[f]	5'-GCGGATCCGCTTCAGGAAACAAGCTAGC-3'	This study

[a] E.coli Genetic Stock Center, Yale University, New Haven, CT, USA
[b] corresponding to the bases of +1 to +16 of the At-IE gene on pKF360
[c] corresponding to the bases of +2366 to +2346 of the At-IE gene on pKF360
[d] corresponding to the bases of +1343 to +1362 of the At-IE gene on pKF360
[e] corresponding to the bases of +1760 to +1742 of the At-IE gene on pKF360
[f] corresponding to the bases of +1855 to +1874 of the At-IE gene on pKF360

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

A cDNA library of 7-day-old *Arabidopsis* seedlings (Mizutani et al., 1997) was converted to a phagemid stock by *in vivo* excision according to the manufacture's instruction (Stratagene, La Jolla, CA, USA). *E. coli* UTH903 cells were transformed with 5 μg of the cDNA library phagemid stock (1.9x 10^5 cells/μg plasmid). The transformation mixture was plated on M9 minimal plates containing 0.2% (w/v) glucose supplemented with 100 μg/ml ampicillin, 25 μg/ml streptomycin and an amino acid mixture without L-His (Sambrook et al., 1989), and incubated at 37 °C for 2 days. 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 an ABI PRISM™ Dye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster city, CA, USA). Nucleotide and amino acid sequences were analyzed using DNASIS Ver. 3.4 software (Hitachi Software Engineering Co., Yokohama, Japan) and by performing the Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990) of the National Center for Biotechnology Information (NCBI).

Nucleic acid hybridization analysis

Genomic DNA (10 µg) was prepared from 4-week-old *Arabidopsis* seedlings as described (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, Buckinghamshire, UK) 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, 5x Denhart's solution, 6x SSC, 0.5% (w/v) SDS and 100 µg/ml shared salmon sperm DNA (Sigma Chemical Co., St. Louis, MO, USA) (Sambrook et al., 1989). Membrane was washed twice in 2x SSC/0.1% (w/v) SDS at room temperature for 10 min, and then twice in 0.5x SSC/0.1% (w/v) SDS at 50 °C for 15 min. Blots were exposed to a Hyperfilm-MP (Amersham) for 2 days at -80 °C using an intensifying screen.

For Northern blot analysis, total RNA was prepared as described (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 6x SSC. The At-IE cDNA was labeled by the random priming method (Feinberg and Vogelstein, 1983) using [α-^{32}P]dCTP. Hybridization was carried out at 42 °C overnight in a solution consisting of 50% (v/v) formamide, 5x Denhart's solution, 6x SSC, 0.5% (w/v) SDS and 100 µg/ml shared salmon sperm DNA (Sambrook et al., 1989). The blots were washed twice in 2x SSC/0.1% (w/v) SDS at room temperature for 10 min and twice in 0.2x SSC/0.1% (w/v) SDS at 55 °C for 15 min. Blots were analyzed using a

Bio Imaging Analyzer (BAS2000; Fuji Photo Film Co., Tokyo, Japan).

Isolation of the At-IE *gene of* Arabidopsis

A genomic clone containing a fragment of the *At-IE* gene was identified by screening approximately 5x 10^5 plaques of an *Arabidopsis* genomic DNA library made with λZAPII (Stratagene). Plaques were transferred to Colony/Plaque Screen™ nylon membranes (NEN™ Life Science Products, Boston, MA, USA) and hybridized overnight at 42 °C with the full-length At-IE cDNA as a probe. Membranes were washed once in 2x SSC/0.1% (w/v) SDS at room temperature for 30 min, once in 0.5x SSC/0.1% (w/v) SDS at 50 °C for 30 min, and once in 0.2x SSC/0.1% (w/v) SDS at 50 °C for 30 min. The blots were then exposed to Hyperfilm-MP films for 16 hr. Positive plaques were re-screened 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 1) and the At-IE cDNA as a template. The PCR products were double-digested with *Bam*HI and *Xho*I and cloned into a *Bam*HI-*Sal*I digested pMAL-c2 vector (New England Biolabs. Inc., Beverly, MA, USA) to obtain expression plasmids (Table 1). 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 for an N-terminal domain without the putative chloroplast transit peptide (pKF362); PRI46 and PRI30 for a 436-bp of 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*

E. coli strain UTH903 was transformed with either pAt-IE (= pKF323) or an empty

pBluescript SK(-) [pBS SK(-)]. pAt-IE (= pKF323) contains a full-length *Arabidopsis hisIE* cDNA. After the transformation, strain UTH903 harboring either an empty pBluscript or pAt-IE was cultivated overnight in M9-glucose medium supplemented with 1 mM histidinol 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 10,000x *g* for 10 min, the soluble fraction was passed through a Sephadex G-25 column (PD-10, Pharmacia). 5-aminoimidazole-4-carboxamide-1-β-D-ribofuranoside (AICAR) production was determined photometrically by measuring the absorbance at 550 nm by the Bratton-Marshall method (Ames et al., 1961). AICAR of 10 μM solution gave an absorbance of 0.270 at 550 nm (Ames et al., 1961) in 1 cm light path. Protein was assayed by the method of Bradford (1976).

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 in General Introduction). This coupling assay was employed because neither the substrate nor the reaction product of the At-IE was available. Briefly, the reaction mixture (170 μl) contained 111 mM Tris-Cl (pH 8.5), 22.2 mM $MgCl_2$, 83.3 mM KCl, 5.6 mM ATP (Sigma), the recombinant proteins to be assayed, and a recombinant HIS1 protein of *S. 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 in General Introduction) through the activities of the HIS1 (the 1st step of His pathway), *Arabidopsis* At-IE (the 2nd and 3rd steps) and the hisA (the 4th step) proteins. The produced BBM III was hydrolyzed to AICAR in HCl at 95 °C. AICAR was determined as descirbed above.

Results

Cloning of an Arabidopsis *cDNA that suppresses an* E. coli hisI *mutation*

An *Arabidopsis* cDNA encoding a bifunctional enzyme having the 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-day-old *Arabidopsis* seedlings, and 20 prototrophic colonies out of 9.5×10^5 transformants were identified after cultivating for 2 days 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 remaing 4 out 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 SK(-), 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 protein/min) was comparable to that with XL1-blue as a control, whereas no AICAR production was detected with the UTH903 transformed with an empty pBluescript SK(-). These results were consistent with the complementation experiments (Fig. 1) in which the UTH903 transformed with pAt-IE was able to grow on M9-glucose 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. 2 and 3). Nucleotide sequence analysis showed that the consensus motif surrounding a translation initiation codon (AACA<u>AT</u>GGC) in plants (Lütcke et al., 1987) was well-conserved as TAAA<u>AT</u>GGC 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. 2). Thus, a putative polyadenylation signal sequence, AATAAA

Fig. 1. Complementation of the His auxotrophy of *E. coli* strain UTH903 (*hisI*) by the *Arabidopsis* At-IE cDNA. *E. coli* UTH903 (*hisI*) transformed with either an empty pBluescript SK(-) plasmid [pBS SK(-)] or a pBluescript SK(-) carrying a 1.1-kb At-IE cDNA (pAt-IE) and were streaked onto a M9 minimal-glucose agar plate in the presence (+His) or absence (-His) of 1 mM L-His and incubated overnight at 37 °C.

(Wahle and Keller, 1992), was found 23-bp upstream 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. 2).

The predicted primary structure of the At-IE protein was compared with those of microorganisms available in the nucleotide database (Fig. 3). Sequence alignment indicated that an N-terminal segment (spanning residues 65-Gly to 158-Phe) of the At-IE protein was highly homologous to the conserved region among the hisI proteins so far reported, and the C-terminal region encompassing residues 179-Leu to 267-Arg was homologous to the domain conserved among the microbial hisE proteins (Fig. 3). These results indicated that the At-IE cDNA encoded a bifunctional protein of PRA-CH (hisI) and PRA-PH (hisE), of which domain organization has also been found in the PRA-CH (hisI) and PRA-PH (hisE) enzymes of eubacteria and lower eukaryote, 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

15

```
-300  AGTTGAAAATGATAACTCGGAGAGCTAGCTCTTCTAGGGTTTTTGGATCTCCTGGAGACGAAGAATAGAGAAAACTTAAGTTATGGCGGAATTTGAAATT
                 GCRE-like
-200  TCAAGCAATTAATAGGGATTGCTTTTGTAATTCCTAATATTAAGACATAAATAGGGGCATTTTAGTATTTGCACACATTGCTGAGTCGGTCTTTAAAAAG
           CAAT                              TATA
-100  CCCTTCACTATCACATTGACTTCAGAACAATTAGGGTTAATCTCTTCCGTTGCTTTTTTTTTTTCAATCCGCCGCCACTTTTCGTTTCTCGTAGCGTTAAA

   1  ATGGCGGTATCGTACAATGCATTAGCTCAGTCTTTAGCGAGAAGTAGCTGCTTCATCCCCAAACCTTATTCCTTTAGAGATACTAAGCTGAGAAGCAGAT
   1   M  A  V  S  Y  N  A  L  A  Q  S  L  A  R  S  S  C  F  I  P  K  P  Y  S  F  R  D  T  K  L  R  S  R

 101  CCAATGTCGTATTCGCGTGCAATGATAATAAGAACATTGCTCTTCAAGCTAAGGTAAACATCTTCTGTGTAATCTTAACTTATAGATAGAATGTTCTGCA
  34   S  N  V  V  F  A  C  N  D  N  K  N  I  A  L  Q  A  K

 201  ATGGAATCGTGTTAAGTTTAATTGTTATCTAAGAACGATTAGAATTTAGATTGAATATTGAACCAGACATTGTCAGGTGCTTACTCTTTGACATATGAGT

 301  AATTACTGTTACATGTCGGGTAGAGAAAAAAGAGAGAAGGAAAAGCTTTTTAACCTGAGCTTCATTTTTAGAGTTTGGATTTCTTTTTTTAAGATGCTCAT

 401  CGTTTGCTTGGTCATCTGCATTTGGTGATGCCTGAGCTACATCTGTTCATCACATTTTATTAAAGCATTTAAGTATATTGTTATGCCAAAAGTAATTTAG

 501  TTCTATGAACATTGTCTGGAGTAAAATTGTATATGTAAAGTTTGGATTTTTGCCAGAGATGGTCAGGTTCCTACCCTTTGACATGTGGGTTTGGACTGTTA

 601  CATCTCTGGTATAGAAAAGAGAGAGAAGGAAAGGCTTTTTGATGGTTTCCATCTTTAGGTTTGGATGTCTCTTTCAAGGTGCTCATCGTTTGCTTGGTCA

 701  TCTTTGTTACGTGATGACTTGGCTCCATTTGTTCATCACAGTTGATATCTTTTCAAGGAAATATGTATTGTCTTGCGCTCTCAGAGTTGGATTAATTCAT

 801  CACTGAATTACTGTGATGCCTTCCGAAGTAACTTATGTTTTAATTTAGAGAAAAGGAGGTTCATGTAAGATGTGTCATTTGTTAATCTATTTTAGTTACG

 901  TCATGTACAATTGAGATTTGGCCAGAGATGGTCAGGTTCCTACTCTTTGACAAATGAGTTTGAACTGTTACATCTCTGGTAAAGAAAAAAGAGAGAAGGA

1000  AAGGCTTTTTGATGGTTTCCATCTTTTGGTCTGGAGTTTTCTCTAAAGGTGTACATCGTTTGCTTGGTCATCTGAATCTGTGATGGCATGGGCTACATCT

1100  GTCCTCCACAACCTGAATAATCTTCATTGAGACATTTCTCTTTTGGCTTTAATATCTAGTCCGATAATGCATTTACTGTTTATACTATTAGAGAGGGGCT

1200  TGTTTGATTGAAATGACAATGTTTTTTTCTGTGTTGGTAGTGTTGCCATTGTGCCTTTTATTACTGTCTCAATAAGCAATGTATATGCGGCTGTTGTGTT

1300  CCCTGGTAGCTAATCATGGTTTCTGTTGTTTTCTTGTGTTAGGTAGATAACTTGTTGGACCGCATTAAATGGGATGACAAAGGATTAGCTGTGGCAATAG
  52                                            V  D  N  L  L  D  R  I  K  W  D  D  K  G  L  A  V  A  I

1400  CACAAAACGTTGATACGGGAGCAGTATTGATGCAAGGCTTTGTTAATAGGGAGGCCCTCTCCACAACCATCAGTTCTCGGAAAGCTACATTCTTTAGTCG
  71   A  Q  N  V  D  T  G  A  V  L  M  Q  G  F  V  N  R  E  A  L  S  T  T  I  S  S  R  K  A  T  F  F  S  R

1500  ATCAAGATCTACCTTATGGACTAAGGGAGAGACATCCAATAACTTCATCAATATTCTTGATGTGTATGTTGATTGTGATCGTGATTCGGTATGTTTGCGT
 105   S  R  S  T  L  W  T  K  G  E  T  S  N  N  F  I  N  I  L  D  V  Y  V  D  C  D  R  D  S

1600  GGACCACAGATATGAATCAATGCAGTGTTATTCTTCATAAACTCATTGGTCTCTATTTTCTCAGATTATTTACCTTGGAACACCTGATGGACCTACCTGT
 134                                                                 I  I  Y  L  G  T  P  D  G  P  T  C

1700  CACACAGGGGAAGAGACTTGTTACTACACATCGGTTTTTGATCAATTAAACAATGATGAGGTTTGCCTTTGTTGAATTTGGTTTTTACCTTTTCCTGTAT
 146   H  T  G  E  E  T  C  Y  Y  T  S  V  F  D  Q  L  N  N  D  E

1800  CAAGAGACGAATTGAAATCTGAAGAATATCAACTACTTTCTTACTGAATCACAGGCTTCAGGAAACAAGCTAGCATTAACAACATTGTACTCGCTAGAAT
 166                                                      A  S  G  N  K  L  A  L  T  T  L  Y  S  L  E

1900  CAATCATTTCCAAGCGGAAAGAAGAATCAACAGTTCCTCAAGAAGGTAAACCATCATGGACTCGACGGTTGTTGACGGATGACGCTCTGCTTTGCTCAAA
 181   S  I  I  S  K  R  K  E  E  S  T  V  P  Q  E  G  K  P  S  W  T  R  R  L  L  T  D  D  A  L  L  C  S  K

2000  GATCAGGTACTTTGTCGGTTATCTGTATAATTGAATTTGAGCTATAGATTCTTATGAAGCTTATTGGGATATTCGGCTTCAGGGAAGAAGCTGACGAGTT
 215   I  R                                                                              E  E  A  D  E  L

2100  ATGCAGAACACTGGAGGATAATGAGGAAGTTTCAAGAACACCATCAGAGATGGCTGATGTTTTATACCACGCAATGGTGCTTCTATCTAAAAGGGGTGTG
 223   C  R  T  L  E  D  N  E  E  V  S  R  T  P  S  E  M  A  D  V  L  Y  H  A  M  V  L  L  S  K  R  G  V

2200  AAGATGGAAGATGTTCTTGAAGTTCTTAGGAAACGCTTCTCTCAATCTGGAATCGAGGAGAAGCAAAACCGTACAAAGTAACCATTTTCTTCGATTGTCT
 256   K  M  E  D  V  L  E  V  L  R  K  R  F  S  Q  S  G  I  E  E  K  Q  N  R  T  K  *

2300  GTTTAGTATTTGCGAGAATTTTTTTAGTTACTCGGAACTTATACTGTTGTTCATTGTGCAACTTTGTATTGTATAAATAGATATCCTTTGTGATAAAAGA
                                                                   ------

2400  AATAAATGAGTCCTTTTTCCATCTTTCTACTTTCATTAACTTTTTCAGTTAACTTAAAAGATATAAATGAGTCGCTTCTCTCCAAAATGTGTCAACCATC
      --------                          ^

2500  TCTGATCTAGTTTATGCTCTATTTGACATCATGGTTAAATATCCTTCTAAAAGAGCACCAAACTTGAGCTTCCAATGCCGATCGTAGTTTGGATCGAGTAA

2600  CCCATAAATATGATCTTAAGTCCTTAGTGGCAGTCCAGCGAAAAGAAGTCATTTGATCAAACTTGCATGCATACTTGCTCTTTATCCACCCTGAATCTGG

2700  TCATATCCATTCATCTTCCAGATAGCATTATTCTGCCACCAACATATAGGTGATTTTCGGGTTTCCTTGAAGACATGATACAACTTGCTGCGTCAGTCAA

2800  TTTATTAATCAAGAGGTCCTTACTCCAACATTTTTTAACTCTTGCAAAGCAAATCTTTAAGACCTTGAGGAATTC
```

Fig. 2.   Nucleotide sequence of the *Arabidopsis At-IE* gene and the amino acid sequence predicted from the At-IE cDNA.

Nucleotide number refers to the A (+1) of the first ATG in the open reading frame. The putative polyadenylation signal and TTTGTA motif are underlined and double-underlined, respectively. Vertical arrow indicates the polyadenylation site. Possible TATA and CAAT elements and a putative GCN4 recognition element (GCRE) in the *At-IE* promoter region are underlined.

16

```
At-IE     MAVSYNALAQSLARSSCFIPKPYSFRDTKLRSRSNVVFACNDNKNIALQA    50
Ec hisIE                                                  ML     2
Sy hisIE                           MSSRSAGRNSPILPLMSHSDLPL    23
Sc HIS4   TAEHLVEQLNVPKERVVVEENGVFSNQFMVKQKFSQDKIVSIKKLSKDML   135
Mj hisI                                             MDVEDT     6
Rs hisI

                  ▼
At-IE     KVDNLLDRIKWDDKGLAVAIAQNVDTGAVLMQGFVNREALSTTISSRKAT   100
Ec hisIE  TEQQRRELDWEKTDGLMPVIVQHAVSGEVLMLGYMNPEALDKTLESGKVT    52
Sy hisIE  ANAVPLDKIRYNDQGLVPAIAQDYLDGTVLMLAWMNEAALAKTLATGQVW    73
Sc HIS4   TKEVLGEVRTDRPDGLYTTLVVDQYER-CLGLVYSSKKSIAKAIDLGRGV   184
Mj hisI   VKKLNLKFRNIEGERLILAITCDE-NKNVLMVAFMNEEALKKTLETGYMH    55
Rs hisI    MAFDPASLTFDANGLIPAVAQDHATGEVLMMAWMNAEAVARTVETGCVT    49

                                      ▼
At-IE     FFSRSRSTLWTKGETSNNFINILDVYVDCDRDSIIYLGTPDGP-TCHTGE   149
Ec hisIE  FFSRTKQRLWTKGETPGNFLNVVSIAPDCDNDTLLVLANPIGP-TCHKGT   101
Sy hisIE  YWSRSRQBLWHKGATSGHFQKLLGIRYDCDSDALLLTIEQKGDIACHTGE   123
Sc HIS4   YYSRSRNEIWIKGETSGNGQKLLQISTDCDSDALKFIVEQENVGFCHLET   234
Mj hisI   YYSTSRKKLWRKGBESGNVQKLIKFYRDCDGDALLFIVEQKGV-ACHEGY   104
Rs hisI   YWSRSRQAFWVKGETSGHVQRLIELRIDCDRDCLLLLIEQEGP-ACHTNR    98

                ▼
At-IE     ETCYYTSVFDQLNNDEASGNKLALTTLYSLESIISKRKEESTVPQEGKPS   199
Ec hisIE  SSC-----FGDTAHQWLF--------LYQLEQLLAERKSA-----DPETS   133
Sy hisIE  RSC-----FHQLDGHKSPPPADM---LTELARVIGDRRDHP----TPE-S   160
Sc HIS4   MSC-----FGEFKHG----------LVGLESLLKQRLQD-----APEES   263
Mj hisI   YSC-----FHYKIEDGELKITGEYYSKR*                        127
Rs hisI   RSC-----FYT-ALREGEERIILDPMV*                         119
Mj hisE                         MILEEVYEIIKQRIKEKPEGS    21
Rs hisE                         MTVLERLAATVEARKGADPDSS    22

                              ▼
At-IE     WTRRLLTDDA-LLCSKI----REEADELCRTLEDNEEVSRTPSEMADVLY   244
Ec hisIE  YTAKLYASGTKRIAQKV----GEEGVETALAATVHDRFELT-NEASDLMY   178
Sy hisIE  YTCKLLAGGDNKILKKI----GEESAEVVMACKDDDPEAIA-GEVADLFY   205
Sc HIS4   YTRRLFNDSA-LLDAKI----KEEAEELTEA-KGKKELS---WEAADLFY   304
Mj hisE   YVAKLTTDDKKTAINKICEKIGEESTELILAAKDDKKDEII-YEAADLIF    70
Rs hisE   WTAKLFAKGPEKCAEKF----GEEAVBAIIEAVRGDRAKLA-SEAADVLY    67

At-IE     HAMVLLSKRGVKMEDVLEVLRKRFSQSGIEEKQNRTK*               281
Ec hisIE  HLLVLLQDQGLDLTPVIENLRKRHQ*                           203
Sy hisIE  HTLVALAHHNVDLRAVYRKLGDRRR*                           230
Sc HIS4   FALAKLVANDVSLKDVENNLNMKHLKVTRRKGDAKPKFVGQPKAEEEKLT   354
Mj hisE   HTMVLLAYKNIEFEELLKEFERRKK*                            95
Rs hisE   HLLVMLAARDVTLAEVMAVLEAREGTSGIAEKAGPG*                103
```

Fig. 3. Alignment of the amino acid sequence predicted from the *Arabidopsis* At-IE cDNA and the corresponding proteins of microbial origins. Ec: *E. coli* (GenBank accession No. X13462; Carlomagno et al., 1988), Sy: *Synechocystis* sp. PCC6803 (DDBJ accession No. D90917; Kaneko et al., 1996), Sc: *S. cerevisiae* (GenBank accession No. J01331; Donahue et al., 1982), Mj: *M. jannaschii* (GenBank accession No. U67484, U67585; Bult et al., 1996), and Rs: *Rhodobacter sphaeroides* (GenBank accession Nos. X87256, X82010; Oriol et al., 1996). Asterisks show the stop codon and dashes were inserted to maximize the alignment. Conserved residues are shaded. ▼ indicates the intron insertion site in the genomic DNA.

acids showed no significant homology to those of the hisIE proteins from microorganisms, but showed the properties characteristic to chloroplast transit peptides (von Heijne and Nishikawa, 1991).

Northern blot analysis

Upon Northern blot analysis, the transcript size of the *At-IE* gene appeared to be approximately 1.2-kb (Fig. 4), which was in good agreement with the predicted size
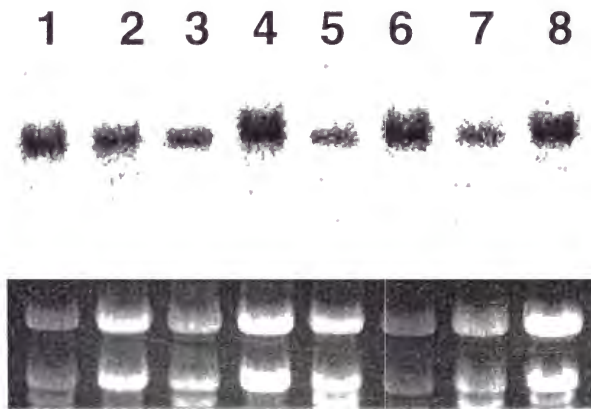
Fig. 4.   RNA blot analysis of the *At-IE* mRNA levels.
Lane 1; 1-week-old plants, lane 2; roots from 2-week-old plants, lane 3; leaves from 2-week-old plants, lane 4; roots from 3-week-old plants, lane 5; leaves from 3-week-old plants, lane 6; roots from 4-week-old plants, lane 7; leaves from 4-week-old plants, lane 8; siliques from 4-week-old plants. Membrane was hybridized with a ^{32}P-labeled *Pst*I-*Eco*RV fragment of the At-IE cDNA. Total RNA (10 µg) prepared from *Arabidopsis* seedlings was electrophoresed in each lane. The photograph of the ethydium bromide-stained gel for the blotting is also shown to the bottom of the RNA blot analysis.

from the At-IE cDNA (Fig. 2). 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 full-length At-IE cDNA as a probe. The digestion with *Hinc*II,

which cuts once within the cDNA, gave rise to two bands (Fig. 5). *Bgl*II also has a single restriction site, but the digestion yielded a single hybridization signal which was thought to be derived from overlapping two bands. The digestion with either *Bam*HI, *Eco*RI or *Xba*I, of which no restriction sites were found in the cDNA sequence, yielded only a single hybridization signal even after long exposure (data not shown). These
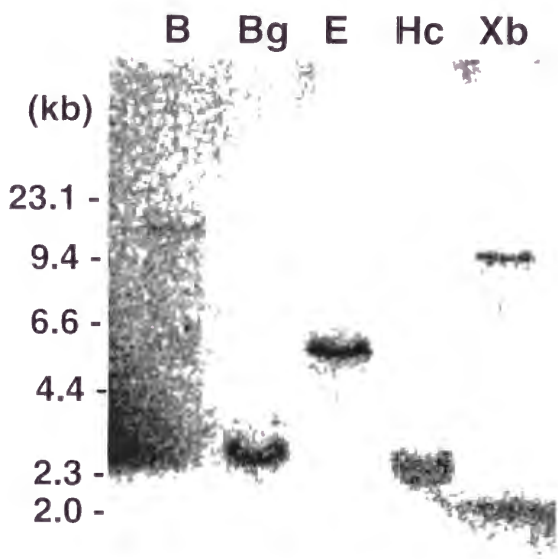
Fig. 5.   Genomic Southern blot analysis.
Genomic DNA (10 μg) was prepared from *Arabidopsis* leaves, and was digested with restriction enzymes (B: *Bam*HI, Bg: *Bgl*II, E: *Eco*RI, Hc: *Hinc*II, Xb: *Xba*I). Hybridization was performed using a ^{32}P-labeled *Pst*I-*Eco*RV fragment of the At-IE cDNA. The lambda DNA digested with *Hind*III is shown as a molecular size marker.

hybridization patterns indicated that the *At-IE* gene exists as a single copy in the genome of *Arabidopsis*.

The *At-IE* gene was cloned by screening an *Arabidopsis* genomic library using the At-IE cDNA as a probe, and a phagemid harboring a 5.5-kb *Eco*RI fragment was identified to contain the *At-IE* gene (Figs. 2 and 6). The *At-IE* gene consists of five exons divided by four introns (Figs. 2 and 6). Intron-splice sites of all the introns follow the

19

"GU-AG" rule, which is observed at intron-splice sites of eukaryotic organisms including higher plants (Breathnach and Chambon, 1981, Simpson and Filipowicz, 1996). Amino acid sequence alignment (Figs. 2 and 3) showed that the intron 1 was located at the

Fig. 6. Complementation of the *hisI* mutation of strain UTH903 with putative functional domains of PRA-PH and PRA-CH. Exon-intron relationship between the *At-IE* gene structure and the At-IE cDNA for the bifunctional PRA-PH/PRA-CH protein is schematically shown. Expression plasmids were designed to contain the putative catalytic domains and used for the complementation assay for the His auxotrophy of *E. coli hisI* mutant. The symbols + and - indicate the ability and inability of the plasmids to suppress the *E. coli* UTH903 *hisI* mutation, respectively. *E. coli* UTH903 cells were transformed with either pKF347 representing the full-length of the At-IE cDNA, pKF372 carrying the full-length insert truncated 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 on M9-glucose minimal agar plates. The portion of the putative chloroplast transit sequence is shaded.
[a]The expression vectors (Table 1) used for the production of the recombinant functional domains.
[b]Schematic diagram of the recombinantly expressed proteins.

junction of the putative transit peptide portion encoded by the exon 1 and the PRA-CH (hisI) domain encoded by the exons 2 and 3. It was also found that the intron 3 located at the putative boundary between the PRA-CH (hisI) domain and the PRA-PH (hisE)

domain, which was encoded by the exons 4 and 5. Two homologous regions among the bacterial hisIE proteins so far reported were found to be also conserved in the At-IE protein (Fig. 3).

In the 5'-untranslated region, several putative regulatory elements were found (Fig. 2). A possible TATA and CAAT elements were identified at position -166 and at -195, respectively. Furthermore, a sequence motif, TAACTC, similar to the *S. cerevisiae* GCN4 responsive element, GCRE (Arndt and Fink, 1986) was located at position -288 (Fig. 2).

Characterization of the domain structure of the At-IE protein

Sequence comparison (Fig. 3) 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. 6). The insert of pKF372 encoded an At-IE protein of which the N-terminal extension had been truncated (Fig. 6). 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 this N-terminal extension corresponded to a chloroplast transit peptide. The recombinantly expressed protein using pKF372 (Fig. 6) was enough to support the AICAR production in the assay system containing the recombinant HIS1 protein of *S. cerevisiae* and the hisA protein of *E. coli* (Fig. 7). Thus, pKF372 contained the cDNA encoding a protein catalyzing the hisIE reactions (Fig. 1 in General Introduction). On the other hand, the recombinant protein prepared with pKF362 (the N-terminal domain), which was able to complement the *hisI* mutation of UTH903 (Fig. 6), did not work for the AICAR production (Fig. 7). Thus, the N-terminal domain catalyzed the hisI (PRA-CH) reaction but did not have the hisE (PRA-PH) activity. The cDNA insert of pKF363 was derived from the exons 4 and 5 coding for the C-terminal domain (Fig. 6). This plasmid failed to suppress the His auxotrophy of UTH903 (Fig. 6) and could not support for the AICAR production

(Fig. 7), 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. 7), indicating that the C-terminal domain corresponded to the hisE (PRA-CH) domain. No histidinol dehydrogenase
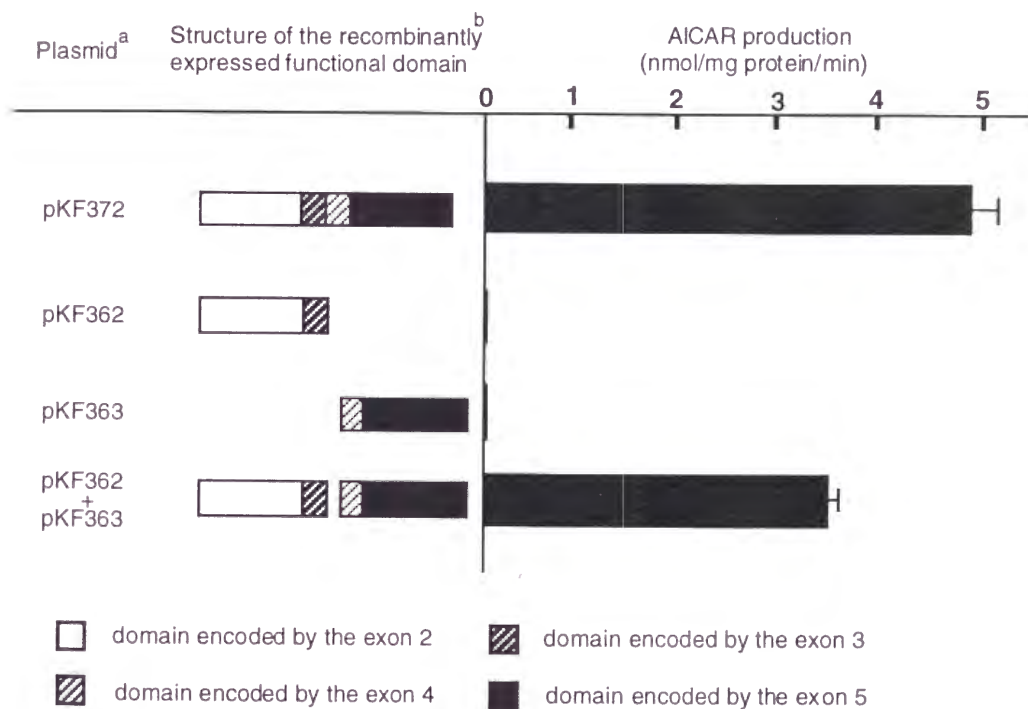
Fig. 7. The At-IE dependent AICAR production.
AICAR production was determined in the assay mixture containing *S. cerevisiae* HIS1, the hisA protein of *E. coli* and one of the recombinantly expressed proteins. The At-IE protein without the putative chloroplast transit peptide, the putative PRA-PH (hisE) 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.
[a]The expression vectors used were the same as those presented in Fig. 6.
[b]Schematic diagram of the recombinantly expressed proteins.

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, Chapter II; Fujimori et al., 1998b). 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 the hisI (PRA-CH) and hisE (PRA-PH) activities. The amino acid sequence predicted from the At-IE cDNA is significantly homologous to both the hisIE proteins of eubacteria and the corresponding counterparts of yeasts and fungi. It has been reported that the hisI and hisE domains are located at the N-terminal and the C-terminal halves, respectively of the microbial hisIE proteins (Donahue et al., 1982, Fig. 2). In the At-IE protein, the region encompassing residues 65-Gly to 158-Phe is homologous to the bacterial hisI and *S. cerevisiae* HIS4A proteins having the PRA-CH activity, and the second region from 179-Leu to 267-Arg is homologous to the bacterial hisE and the HIS4B of *S. cerevisiae* with the PRA-PH activity.

These two functional domains for hisI (PRA-CH) and hisE (PRA-PH) were able to be separately expressed with the corresponding cDNA fragments retaining their individual catalytic activities (Figs. 6 and 7). Thus, the At-IE protein molecule is shown to be composed of two independent, functional domains connected by a gap region, which shows no significant homology with other corresponding proteins (Fig. 6). It was also demonstrated that the At-IE protein does not contain the histidinol dehydrogenase domain encoded by *HIS4* of *S. cerevisiae*. It has been already shown that histidinol dehydrogenase is encoded by a separate gene in higher plants (Nagai et al., 1993).

Genomic Southern blot analysis showed that the *At-IE* exists as a single copy gene consisting of the five exons divided by the four introns, encoding the three putative

23

functional units; the chloroplast transit peptide (exon 1), the PRA-CH domain (exons 2 and 3), and the PRA-PH domain (exons 4 and 5). This gene organization suggests that the gene for the chloroplast transit peptide and the genes for the protein corresponding to the PRA-CH and PRA-PH have fused during evolution. Examples for similar gene organization containing an intron between a putative chloroplast transit peptide and a mature protein have also been reported for a few nuclear genes encoding chloroplastic proteins (Wolter et al., 1988, Gantt et al., 1991). It is thought that the sequences for the chloroplast transit peptide have been attached to the genes for chloroplastic proteins during the process of gene-transfer from chloroplast to nucleus (Gantt et al., 1991). The yeast HIS4 and fungal his3 proteins also have N-terminal leader sequences of unknown functions, whereas no significant N-terminal homology was seen between the At-IE protein and these microbial proteins.

It has been demonstrated that other higher plant His biosynthetic enzymes, imidazoleglycerolphosphate dehydratase (Tada et al., 1995) and histidinol dehydrogenase (Nagai et al., 1993), are localized in chloroplasts. Furthermore, the cDNAs encoding N'-[(5'-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (BBM II) isomerase, glutamine amidotransferase/cyclase and ATP phosphoribosyl transferase (ATP-PRT; EC 2.4.2.17) of A. thaliana (Chapter II; Fujimori et al., 1998b, Chapter III; Fujimori and Ohta, 1998c, Chapter IV) and histidinolphosphate aminotransferase from N. tabacum (El Malki et al., 1998) have also been shown to contain the regions corresponding to putative N-terminal transit peptides. It is therefore possible that the entire His biosynthesis may be completed in chloroplasts. According to a calculation, the His biosynthesis is an extremely energy consuming process, which requires 41 ATP molecules for each His molecule produced (Brenner and Ames, 1971). In other words, the compartmentalization of the His pathway in chloroplasts is favorable to ensure the efficient energy supply. At least six genes encoding 8 steps of the His biosyntheis in plants exhibited constitutive expression patterns throughout development, and there were no clear tissue specific expression patterns (Nagai et al., 1993, Tada et al., 1994, El Malki et al., 1998, Chapter II; Fujimori et al., 1998b, Chapter III; Fujimori and Ohta, 1998c, Chapter IV). In microorganisms, the His biosynthesis is regulated through the controlled gene expression and the feedback regulation of ATP-PRT activity by L-His

(Alifano et al., 1996). Isolation and characterization of ATP-PRT are essential for understanding the mechanism that regulates the His biosynthesis of higher plants.

Summary

Phosphoribosyl-ATP pyrophosphohydrolase (PRA-PH) and phosphoribosyl-AMP cyclohydrolase (PRA-CH) are encoded by *HIS4* in yeast and by *hisIE* in bacteria and catalyze the second- and the third-step in histidine (His) biosynthetic pathway, respectively. By complementing a *hisI* mutation of *Escherichia coli* with an *Arabidopsis* cDNA library, I isolated an *Arabidopsis* cDNA (At-IE) having these two enzyme activities. The At-IE cDNA encodes a bifunctional protein of 281 amino acids with a calculated molecular mass of 31,666 Da. Genomic DNA blot analysis with the At-IE cDNA as a probe revealed a single copy gene in *Arabidopsis*, and Northern 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 chloroplast transit peptide. I 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 half and C-terminal half of the At-IE protein, respectively.

Chapter II

Molecular Cloning and Characterization of the Gene Encoding *N'*-[(5'-Phosphoribosyl)-Formimino]-5-Aminoimidazole-4-Carboxamide Ribonucleotide (BBM II) Isomerase from *Arabidopsis thaliana*

Introduction

Histidine (His) biosynthesis has been extensively studied in a variety of
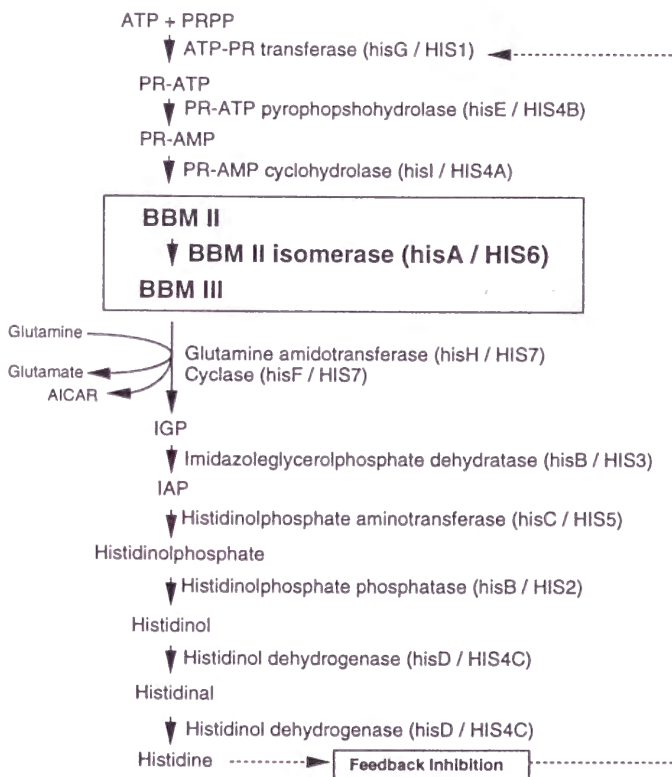
Fig. 1. Histidine biosynthetic pathway *in E. coli and S. cerevisiae.*
BBM II isomerase catalyzes the fourth step of the pathway. PRPP; phosphoribosyl pyrophosphate; ATP, adenosine triphosphate; PR-ATP, phosphoribosyl-ATP; AICAR, 5-aminoimidazole-4-carboxamide-1-β-D-ribofuranoside; IGP, imidazoleglycerolphosphate; IAP, imidazoleacetolphosphate. Corresponding counterparts for *E. coli* and S. *cerevisiae* are indicated (*E. coli* / S. *cerevisiae*), respectively.

microorganisms. In bacteria, N'-[(5'-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide ribonucleotide isomerase (BBM II isomerase; EC 5.3.1.16) is encoded by *hisA* cistron in the His operon and catalyzes the conversion of N'-[(5'-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide-ribonucleotide (BBM II) to N'- [(5'-phosphoribulosyl)-formimino]-5-aminoimidazole-4-carboxamide-ribonucleotide (BBM III) at the fourth step of the His biosynthetic pathway (Fig. 1). In yeast, this enzyme is encoded by *HIS6*. *hisA* homologues have been isolated from a variety of organisms; *Methanococcus voltae*, *M. vannielii* (Cue et al., 1985), *M. thermolithotrophicus* (Weil et al., 1987), *M. jannaschii* (Bult et al., 1996), *E. coli* and *S. typhimurium* (Carlomagno et al., 1988), *Streptomyces coelicolor* (Limauro et al., 1990), *Lactococcus lactis* (Delorme et al.,1992), *Azospirillum brasilense* (Fani et al., 1993), *Haemophilus influenzae* (Fleischmann et al., 1995); *Saccharomyces cerevisiae* (Fani et al., 1997) and *Schizosaccharomyces pombe* (Barrel et al., unpublished). Of these, the *hisA* genes of *M. voltae*, *M. vannielii*, *M. thermolithotrophicus*, and *A. brasilense* have been cloned by genetic complementation of *E. coli hisA* mutants. It was also reported that the *S. cerevisiae HIS6* gene was able to complement an *E. coli hisA* mutant (Fani et al., 1997). BBM II isomerase has been purified to apparent homogeneity from *S. typhimurium* (Margolies and Goldberger, 1966), and its biochemical properties were also characterized (Margolies and Goldberger, 1967). However, in plants, even the existence of BBM II isomerase has not yet been demonstrated.

In Chapter II, I described about the isolation of a BBM II isomerase cDNA from an *Arabidopsis* cDNA library through genetic complementation of an *E. coli hisA* mutant, HfrG6. The primary structure predicted from the cDNA showed a high degree of identity to the HIS6 protein of *S. pombe*. In a coupling assay containing a series of enzymes involved in the earlier steps of the His pathway, I demonstrated that the cDNA encoded a functional BBM II isomerase that catalyzed the conversion of BBM II to BBM III, in a reaction known as an Amadori rearrangement (Smith and Ames, 1964).

Materials and methods

Plant materials and bacterial strains

Arabidopsis thaliana ecotype Columbia (Col-0, Lehle Seeds, Tucson, AZ, USA) seedlings were grown described in Chapter I (Fujimori and Ohta, 1998a).

Bacterial strains used in this study are listed in Table 1. An *E. coli hisA* mutant, HfrG6 (CGSC No.5388) (Matney et al., 1964) was obtained from the E. coli Genetic Stock Center (CGSC; Yale University, New Haven, CT, USA). *E. coli* JM109 was used as a host for the propagation and manipulation of plasmid DNAs. *E. coli* XL1-Blue MRF', ExAssist™ helper phage, and pBluescript were purchased from Stratagene (La Jolla, CA, USA).

Table 1.   Summary of the *E. coli* strain and plasmids used in this study

Material	Property	Reference or source
E. coli strain		
HfrG6	λ^- *hisA323*	CGSC[a]
JM109	*recA1 endA1 gyrA96 thi hsdR17 supE44 relA1 Δ(lac-proAB)/ F'(traD36 proAB lacI^q lacZΔM15)*	Yanisch-Perron et al. (1985)
XLI-Blue MRF'	*Δ(mcrA)183 Δ(mcrCB-hsdSMR17-mrr)173 endA1supE44 thi-1 recA1 gyrA96 relA1 lac[F' proAB lacI^q ZΔM15 Tn10]*	Stratagene
Plasmids		
pKI-1	pBluescript bearing an *Arabidopsis* BBM II isomerase cDNA	This study
pKI-2	pBluescript bearing an *Arabidopsis* BBM II isomerase cDNA	This study
pKF410	pBluescript bearing a 4.7-kb fragment for the *Arabidopsis* BBM II isomerase gensis	This study
pT11D6	pBeloBACII bearing an *Arabidopsis* BBM II isomerase gene	ABRC[b]

[a] E. coli Genetic Stock Center, Yale University, New Haven, CT, USA
[b] Arabidopsis Biological Resource Center, Ohio State University, Columbus, OH, USA

Genetic complementation of an E. coli hisA mutant with Arabidopsis cDNAs

A cDNA library was constructed in the λZAPII vector system (Stratagene) from poly (A)⁺ RNA isolated from 7-day-old *Arabidopsis* seedlings (Mizutani et al., 1997). Recombinant pBluescript cDNA phagemids were excised *in vivo* from the λZAPII vector by the infection of XL1-Blue MRF' cells with the M13 helper phage ExAssist™, according to the manufacturer's instruction (Stratagene). A 100 μl aliquot of the cDNA library phagemid stock (6.1×10^5 transducing units/ml) was mixed with 2 ml of a mid-log phase culture of *E. coli* strain HfrG6. The mixture was incubated at 37 °C for 15 min. After centrifugation, the pellet was washed in minimal E medium (Vogel and Bonner,

1956) and plated on minimal E agar plates supplemented with ampicillin (100 μg/ml). Plates were incubated at 37 °C overnight. Plasmids were extracted from His prototrophic colonies and again tested for their ability to recomplement the *hisA* mutation in strain HfrG6. The plasmids thus obtained were designated pKI-1 and pKI-2 (Table 1).

cRACE

Circular first-strand cDNA-mediated rapid amplification (cRACE; Maruyama et al., 1995) was used to complete cDNA sequences. Design of gene-specific primers was based on the sequence of the isolated cDNA clone, pKI-1 (Table 1). First-strand cDNA was synthesized from 5 μg of total RNA extracted from the leaves of 2-week-old *Arabidopsis* seedlings, using a cDNA Cycle Kit (Invitrogen, San Diego, CA, USA) with a gene-specific reverse primer HY1. After hydrolysis of the template RNA with 0.5 N NaOH at 37 °C for 15 min, the first-strand cDNA was synthesized and circularized by T4 RNA ligase (New England Biolabs. Inc., Beverly, MA, USA). The ligation mixture was subjected to the first-step PCR amplification using a forward primer HY2 and a reverse primer HY3. A second-step PCR amplification was performed with another set of a forward primer HY4 and a reverse primer HY5. The amplified PCR products were fractionated electrophoretically in 1% (w/v) agarose gels and eluted using a Sephaglas BandPrep Kit (Pharmacia, Uppsala, Sweden). The resulting major DNA fragments were cloned into a pCR2.1 vector (TA Cloning Kit, Invitrogen).

Isolation of the BBM II isomerase gene from Arabidopsis

Approximately 1x 10^5 recombinant phages from an *Arabidopsis* λZAPII genomic library (Stratagene) were screened with the cDNA insert of pKI-1 after ^{32}P-labeled using a Random Primer DNA Labeling Kit (Takara Shuzo, Kyoto, Japan) and [α-^{32}P]dCTP (Amersham, Buckinghamshire, UK) (Feinberg and Vogelstein, 1983). Prehybridization, hybridization and washing were performed as described (Sambrook et al., 1989). Filters were exposed to Hyperfilm-MP films (Amersham) with an intensifying screen at -80 °C. Positive signals were purified to apparent homogeneity.

DNA sequencing and analysis

DNA sequencing was performed using a PRISM™ Dye Terminator Cycle Sequencing Core Kit (Applied Biosystems, Foster City, CA, USA) and an automated DNA sequencer (Model 373A; Applied Biosystems). Sequence analysis was performed using the software, DNASIS version 3.4 (Hitachi Software Engineering Co., Yokohama, Japan). Nucleotide and predicted amino acid sequences were analyzed by the Basic Local Alignment Search Tool (BLAST) and Gapped BLAST programs of the National Center for Biotechnology Information (NCBI) (Altschul et al., 1990, Altschul et al., 1997).

Preparation of crude extract from bacteria

Soluble proteins were extracted from HfrG6 cells transformed with either pKI-1 or an empty pBluescript vector. The bacteria were grown overnight in LB medium and collected by centrifugation. After washing with 50 mM potassium phosphate buffer (pH 7.5), approximately 0.5 g of cells were resuspended in 3 ml of buffer A containing 20 mM Tris-Cl (pH 7.5), 200 mM NaCl and 1 mM EDTA, and cells were then broken open by sonication with short pulses for 2 min. After cell debris was removed by centrifugation at 10,000x *g* for 15 min, proteins were precipitated by the addition of $(NH_4)_2SO_4$ to a concentration of 80% (w/v). The pellet was redissolved in buffer A and desalted by passing through a Sephadex G-25 column (NAP-5; Pharmacia). Protein was assayed by the method of Bradford (1976) using BSA as standard.

Enzyme assay

The crude bacterial extracts were used for the assay of BBM II isomerase in the presence of the His biosynthetic enzymes; ATP-phosphoribosyl transferase (ATP-PRT) and the bifunctional PR-AMP cyclohydrolase (PRA-CH)/PR-ATP pyrophosphohydrolase (PRA-PH). For the preparation of these coupling enzymes, the ATP-PRT (HIS1) of *S. cerevisiae* (Hinnebusch and Fink, 1983) and the *hisIE* gene product of *E. coli* (the bifunctional protien with PRA-PH and PRA-CH activity) were overexpressed in *E. coli* as fusion proteins with the maltose binding proteins, using a pMAL-c2 vector (New England Biolabs.), and were purified by column chromatography on an amylose resin. In this assay system, BBM II (the substrate for BBM II isomerase) is produced

from ATP and phosphoribosyl pyrophosphate (PRPP) by the coupled enzyme activities. The amount of BBM III finally produced by the activity of BBM II isomerase was estimated by measuring its ability to reduce a dye mixture (Smith and Ames, 1964). This dye mixture was prepared freshly and contained 2-*p*-iodophenyl-3-*p*-nitrophenyl-5-phenyltetrazolium chloride (3.2 mg/ml) (Sigma Chemical Co., St Louis, MO, USA), phenazine methosulfate (0.4 mg/ml) (Sigma), and 0.2% (w/v) gelatin at a ratio of 5:1:1(v/v/v). The assay mixture consisted of 100 mM Tris-Cl (pH 8.5), 20 mM $MgSO_4$, 100 mM KCl, 4 mM ATP, 2 mM PRPP, 80 µl of the dye mixture and enzymes in a total volume of 200 µl. The enzyme mixture contained recombinant *S. cerevisiae* ATP-PRT (HIS1), the *E. coli hisIE* gene product (the bifunctional PRA-PH and PRA-CH enzymes) and crude extracts of strain HfrG6 transformed with either pKI-1 or an empty pBluescript. The reaction was started by adding the enzyme mixture. After incubation at 30 °C for 1.5 hr, the reaction was stopped by the addition of a 1/5 vol. of 0.67 N HCl, and then the absorbance was read at 520 nm and compared with that of a control incubated in the absence of the enzyme mixture. Purified *E. coli* HisA was used for a positive control to check that the coupling reaction proceeded appropriately. The rate of increase in absorbance at 520 nm reflects the amount of BBM III formed. BBM III has a molar extinction coefficient of 11,800 at 520 nm (Smith and Ames, 1964).

Southern and Northern hybridization analyses

Total genomic DNA (10 µg) was prepared from 2-week-old *Arabidopsis* seedlings according to the method described by Sambrook et al. (1989). After digestion with either *Bam*HI, *Bgl*II, *Eco*RV, *Hind*III, or *Xho*I, DNAs were separated electrophoretically in a 0.7% (w/v) agarose gel, and blotted onto a Hybond N⁺ nylon membrane (Amersham) in 0.4 N NaOH (Sambrook et al., 1989). Hybridization with a full-length cDNA random-labeled with [α-^{32}P]dCTP (Feinberg and Vogelstein, 1983) was carried out for 16 hr at 37 °C in 40% (v/v) formamide, 5x Denhart's solution, 6x SSC, 0.5% (w/v) SDS and 100 mg/ml shared salmon sperm DNA (Sigma) (Sambrook et al., 1989). The membrane was washed twice in 2x SSC/0.1% (w/v) SDS at room temperature for 10 min and then twice in 0.5x SSC/0.1% (w/v) SDS at 50 °C for 15 min.

Tissues were harvested from *Arabidopsis* seedlings at different developmental

stages, and total RNA was prepared by phenol/chloroform extraction followed by lithium chloride precipitation as described (Lagrimini et al., 1987). Total RNA (10 µg) from each sample was analyzed by Northern hybridization using a full-length cDNA as a hybridization probe labeled with [α-^{32}P]dCTP by the random priming method (Feinberg and Vogelstein, 1983). Prehybridization, hybridization, and washing were performed as described by Sambrook et al. (1989).

Results and Discussion

Isolation of Arabidopsis *cDNAs that complement an* E. coli hisA *mutant*

Genetic complementation has been extensively applied as a means of isolating *hisA* gene homologues from various organisms such as *M. voltae, M. vannielii* (Cue et al.,
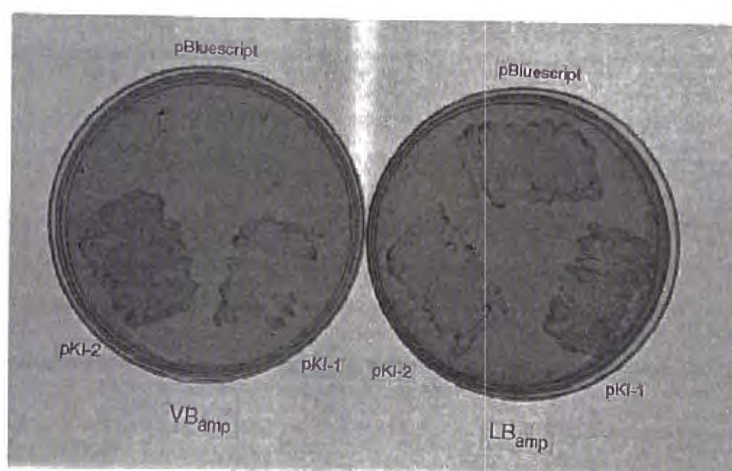
Fig. 2. Complementation of an *E. coli hisA* mutant, HfrG6.
Cells were transformed with either pBluescript, pKI-1 or pKI-2 and were plated on minimal VB agar plates containing ampicillin (100 µg/ml) or LB agar plates containing ampicillin (100 µg/ml).

1985), *M. thermolithotrophicus* (Weil et al., 1987), and *A. brasilense* (Fani et al., 1993). Recently, the *HIS6* gene of *S. cerevisiae* has also been shown to complement an *E. coli hisA* mutant (Fani et al., 1997). Assuming that the bacterial BBM II isomerase encoded by *hisA* gene and the its homologue in higher plants might be functionally conserved, I employed a genetic complementation approach in an attempt to isolate a *hisA* counterpart

```
-268 gctttggtggaggattaggtattcggtaaatctcaggtaacggtggcgtaatcggcgccggaagatgatgagactttgag

-190 acgagactgataacgttcttcttgttccatcgcatcactctctgtttgtcaaacgaatcgataaatgttgggcattttta

-110 tcgatttccaattttaaaccaaaccggaataaaccgatatctaaaccctagaatatctagaaccagatacgtcactggtg

 -30 cttactgattacagtttcagttccaaagATGAGAACTTTGAGCTCCCAATTATACTCCAATGGTGGTCTCACTTGGTTCC
                                 M  R  T  L  S  S  Q  L  Y  S  N  G  G  L  T  W  F    17

  50 AGAAGAAGAATCAGAGCTCTCTGTTCATCAAGCATCTCAGAGTCTCTAAACCATCGAGGGTTCAGTCATATCTGCTGTT
      Q  K  K  N  Q  S  S  L  F  I  K  H  L  R  V  S  K  P  S  R  V  Q  L  I  S  A  V   44

 130 CAATTCCGTCCGTGCATCGATATTCATAAGgtatgaaatcaatttcaccgattagaaaatcctcactggctgttgagtgc
      Q  F  R  P  C  I  D  I  H  K                                                      54

 210 agtgaattcatttatggatctatctctgtgtgatgtgattttcgatagGGAAAAGTGAAACAGATAGTTGGATCCACACT
                                           HY5 ←
                                              G  K  V  K  Q  I  V  G  S  T  L           65

 290 TCGTGACTTGAAAGAGGATGGTTCAGTTCTCGTTACGAATTTTGAATCGGATAAGTCTGCGGAAGAGTATGCAAAGATGT
          HY3 ←            HY2 ←
      R  D  L  K  E  D  G  S  V  L  V  T  N  F  E  S  D  K  S  A  E  E  Y  A  K  M      91

 370 ACAAAGAAGATGGGCTTACGGGTGGTCATGTGATAATGCTTGGAGCAGACCCTTTAAGCCAAGCTGCAGCTATTGGAGCA
      Y  K  E  D  G  L  T  G  G  H  V  I  M  L  G  A  D  P  L  S  Q  A  A  A  I  G  A   118

 450 TTACATGCTTATCCTGgtatgtgtttagttgaatgtcacatggtgtttgttttatattgtttgaatttgaaatgtaaagt
      HY4
      L  H  A  Y  P                                                                     123

 530 tttgatgtgttatatatgatacaagatgatgaattctagGCGGTTTGCAAGTTGGAGGTGGAATCAATTCTGAGAATTGC
                                          →
                                        EcoRI  G  G  L  Q  V  G  G  G  I  N  S  E  N  C  137

 610 ATGAGTTATATTGAGGAGGGAGCAAGCCATGTCATTGTCACTTCGgtatagctgtcatctagcttgattgaatatgttat
      M  S  Y  I  E  E  G  A  S  H  V  I  V  T  S                                       152

 690 catagtaagtatagctgatgataattatgtttttcgttttgtatttcagTATGTATTTAACAATGGAAAGATTGATCTTGA
                                                       Y  V  F  N  N  G  K  I  D  L  E   163

 770 AAGACTCAAAGATATTGTGAGCATTGTTGGGAAACAAAGACTGATATTGGACCCTTAGCTGCAGAAAGAAGgtgattctct
                                 HY1 ←
      R  L  K  D  I  V  S  I  V  G  K  Q  R  L  I  L  D  L  S  C  R  K  K              186

 850 gactagtttcatcttcagtttaatgttattggtcatgaaaagaagatcggcatttgaggtatttatatggttattttgtt

 930 gcgcagGATGGAAGATACGCAATTGTTACTGATAGATGGCAGAAGTTCAGCGACGTCATCCTTGACGAGAAATCATTGGA
            D  G  R  Y  A  I  V  T  D  R  W  Q  K  F  S  D  V  I  L  D  E  K  S  L  E   211

1010 GTTTCTTGGAGGATTTTCAGATGAGTTTCTGGTGCATGGCGTAGACGTTGAAGGAAAGAAgtaagctttaaatatctcaa
      F  L  G  G  F  S  D  E  F  L  V  H  G  V  D  V  E  G  K  K                        231

1090 ccctgtgattttagtgtcttagagtaaattcgttttgcattgaaaatacagtatttagtcagtatgtactgatcagtggt

1170 gttaacacttcttctctcaaagGCTAGGGATCGATGAAGAGCTTGTAGCATTGCTTGGCAACTACTCTCCGgtgagtaat
                             L  G  I  D  E  E  L  V  A  L  L  G  N  Y  S  P             247

1250 aacatgcagtttcttatgattcggaactaatcttatggacttaacgataattaacagaacttcactctattcaaaatgct

1330 tgtgatgttagtgaaaccggaattggattagatgtagtcgttatacttcactttagacatagagataagctcgattgcaa

1410 acttgttcatttgagcttgatcttaaatgttttttgaccttgacagATTCCAGTAACTTATGCGGGAGGTGTGACGGTAA
                                                     I  P  V  T  Y  A  G  G  V  T  V    258

1490 TGGATGATGTAGAGAGGATCAAGGACGCAGGAAAGGGACGTGTAGATGTAACAGTGGGAAGTGCCTTAGACATTTTCGGG
      M  D  D  V  E  R  I  K  D  A  G  K  G  R  V  D  V  T  V  G  S  A  L  D  I  F  G   285

1570 GGAAATCTTCCATACAAAGACGTCGTCGCTTGGCATCACAAGCAGCACTCACTCCATTGAtatcgccatttttcacctttg
      G  N  L  P  Y  K  D  V  V  A  W  H  H  K  Q  H  S  L  H  *                        304

1650 ctcatctttttttatttgcattaccccaaaccttaagcttcaatatttgcaacattaatcttcagctttaaacttttgtgc
                                    HindIII

1730 ttctttctaaactcttcaagaatttgtcaattttgaatatttttcgattaatcaagtttatggtttggtttgactacta

1810 ctcatcaaggaaaatcgcattccaacaacacgtaaacagaaacaaatacattttgtttgacaatagcatatacagatata

1890 ataacatttaaactgaactatttacaacggaaactagaagaatacaaaaacagtgcggtgggtgtgaaatcaccgctttc

1970 gcgttgccgtgtttctttctgaataaaaacgttgatcagctg
                                        PvuII
```

Fig. 3. *Arabidopsis* BBM II isomerase gene structure.
The DNA sequence determined from the cDNA is shown with capital letters, and the predicted amino acid sequence is shown in the single letter expression. The 5' end of the cDNA determined through the cRACE experiment was shown in bold letters. Nucleotide and amino acids are numbered on the right and left of the sequence, respectively. The stop codon is marked with an asterisk. The DDBJ accession numbers are AB006139 (*Arabidopsis* BBM II isomerase cDNA) and AB008929 (*Arabidopsis* BBM II isomerase gene).

from *Arabidopsis* using the *E. coli hisA* mutant, HfrG6 as a recipient.

A phagemid library prepared from 7-day-old *Arabidopsis* seedlings was used to infect strain HfrG6, and the infected culture was plated on minimal agar plates. Two colonies out of 6×10^4 transformants grew on minimal medium lacking exogenous L-His, whereas no suppression of the *E. coli hisA* mutation was observed when cells were transformed with the empty pBluescript vector. The plasmids designated pKI-1 and pKI-2 were isolated from these two His prototrophic colonies, and their ability to suppress the *hisA* mutation was confirmed by re-transforming strain HfrG6 to His prototrophy (Fig. 2).

The inserts from the two plasmids, pKI-1 and pKI-2, were approximately 1.1-kb in length, and the DNA sequences were determined to be identical, except for the variation in the lengths of their 5'-termini (data not shown). However, the translational initiation codon was not included in the putative N-terminal region of the longest open reading frame (ORF) found in the pKI-1 insert (Fig. 3), indicating that both cDNAs were incomplete at the 5'-end.

To obtain the missing portion of the cDNA, I employed circular first-strand cDNA-mediated rapid amplification (cRACE; Maruyama et al., 1995) on total RNA from the leaves of 2-week-old *Arabidopsis* as a template. After initial reverse transcription and additional two rounds of PCR amplification, I obtained a major DNA fragment of 410-bp and cloned it into the pCR2.1 vector. DNA sequencing revealed that the isolated cDNA clones contained the 5' region of the pKI-1 insert spanning the HY1 primer sequence and an additional 44 nucleotides at the 5'-end (Fig. 3, bold letters). By combining the sequences of the insert from pKI-1 and the cRACE product, the full-length sequence of a 1115-bp cDNA was determined (Fig. 3). It contains a 912-bp ORF encoding a protein of 304 amino acids with calculated molecular weight of 33,363; the initial codon was found to be located five codons upstream of the 5' terminus of the cDNA insert in pKI-1 (Fig. 3).

A genomic clone (TAMU T11D6: pT11D6) was identified in the non-redundant GenBank database that contained the promoter region and the ORF of *Arabidopsis* BBM II isomerase gene. To isolate the 3' non-coding region, an *Arabidopsis* genomic library was screened with the BBM II isomerase cDNA as a probe, and a clone containing

an approximately 4.7 kb insert corresponding to the 3' region of BBM II isomerase was isolated. The BBMII isomerase gene structure was thus determined by combining the DNA sequences of the inserts of pT11D6 and this newly isolated genomic clone (Fig. 3). The intron/exon junction sequences are highly homologous to the plant consensus sequences (Simpson and Filipowicz, 1996). Also, the genomic structure provided further evidence that the N-terminus determined by the cRACE experiment was indeed the actual translational initiation site.

Enzymatic activity of Arabidopsis *BBM II isomerase*

To confirm that the inserts from pKI-1 and pKI-2 actually encoded a functional BBM II isomerase, strain HfrG6 was transformed with either pKI-1 or the empty
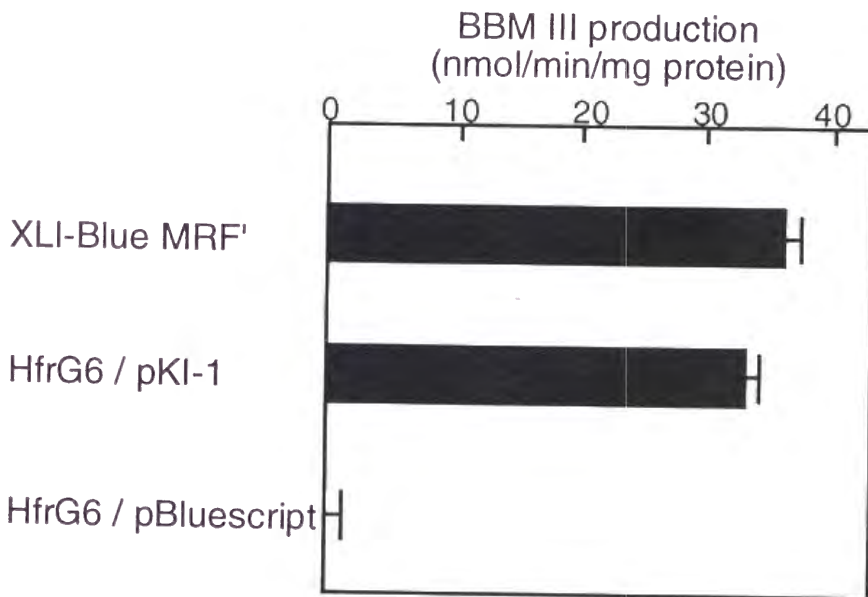
Fig 4.   BBM II isomerase activity in wild-type (XLI-Blue MRF'), *hisA* mutant (HfrG6) carring each of pKI-1 (*Arabidopsis* cDNA) or pBluescript SK(-) (empty).
The bars presented are the means ± standard deviation (SD) for three independent experiments.

pBluescript SK(-) vector, and crude bacterial extracts were prepared for enzyme assay. In this assay system, BBM II is produced from ATP and PRPP through the coupled enzymatic activities of the HIS1 of *S. cerevisiae* and the HisIE of *E. coli*, according to the

His biosynthetic scheme (Fig. 1). In the presence of the *hisA* gene product (BBM II isomerase), BBM II is further converted to BBM III, which can be colorimetrically determined (Smith and Ames, 1964). No BBM III production was detected when strain HfrG6 was transformed with a pBluescript SK(-) empty vector (Fig. 4). In contrast, the level of BBM III production in the extracts from HfrG6 transformed with pKI-1 was comparable (31 nmol/mg protein/min) to that observed with the XLI-Blue MRF' extract (37 nmol/mg protein/min) (Fig. 4). pKI-2 gave the same results as pKI-1 (data not shown). These results demonstrated that the isolated cDNA inserts of pKI-1 and pKI-2 actually encoded the BBM II isomerase of *Arabidopsis*.

Comparison of the primary structures of BBM II isomerase between Arabidopsis *and microorganisms*

Significant sequence homology was found among the protein sequences predicted from the *hisA* genesfrom various prokaryotic microorganisms (Fig. 5). The primary structure predicted from the ORF of the *Arabidopsis* BBM II isomerase cDNA (Fig. 3) was 50% and 42% identical, respectively, to the HIS6 proteins of *S. pombe* and *S. cerevisiae* (Fig. 5). In contrast, the degree of sequence similarity between *Arabidopsis* BBM II isomerase and the bacterial *hisA* gene products was very small, although there were several short homologous stretches were found (Fig. 5). Sequence comparison also revealed that *Arabidopsis* BBM II isomerase contained an N-terminal extension. This N-terminal extension of approximately 40 amino acids showed the general features of chloroplast transit peptides (von Heijne and Nishikawa, 1991); specifically, it was rich in Ser and contained no acidic residues, implying that *Arabidopsis* BBM II isomerase is localized in chloroplasts. It should be noted that the cDNAs encoding the HDH of cabbage (Nagai et al., 1991) and the IGPD of *Arabidopsis* (Tada et al., 1994) also encoded proteins having putative chloroplast transit peptides at their N-termini. Protein blot analyses demonstrated the chloroplast localization of HDH (Nagai et al., 1993) and IGPD (Tada et al., 1995) proteins. Taken together, these data make it likely that His biosynthesis in higher plants occurs entirely in the chloroplasts.

```
At   MRTLSSQLYSNGGLTWFQKKNQSSLFIKHLRVSKPSRVQLISAVQFRPCI         50
Sp                                        MEKSQHTLFRPCI         13
Sc                                            MTKFIGCI          8
Ec                                       MIIPALDLIDGTVV        14
Ll                                      MKIIPAIDLQNGEAV        15
Ab                                      MIIYPAIDLKDGACV        15
Sy                                      MKILPAIDLLGGKCV        15
Mv                                      MYIIPAVDMKEGKCV        15

At   DIHKGKVKQIVGSTLRDLKEDGSVLVTNFESDKSAEEYAKMYKEDGLTGG        100
Sp   DIHQGQVKQIVGGTLGNDAD----VKTNYVSLKPSSYYAELYKLNHLEGA         59
Sc   DLHNGEVKQIVGGTLTSKKED--VPKTNFVSQHPSSYYAKLYKDRDVQGC         56
Ec   RLHQGDYGK---------------QRDYGNDPLPR--LQDYAAQGAEVL          46
Ll   RLYKGDYDK---------------KTVYSKNPLEI--AQKFERMGATDL          47
Ab   RLLRGEMSQ---------------ATVFNTEPADQ--ARLFESQGFEWL          47
Sy   RLYQGDYDQ---------------SQVYHEDPVEV--ARQWQAQGASRL          47
Mv   QLIQGDPTK---------------RHVEYDNPDEI--AKMWIENGAEML          47

At   HVIMLGA----DPLSQAAAIGALHAYPGGLQVGGGI-NSENCMSYIEE-G        144
Sp   HVIMLG------PNCEEAAKTALHTWPGALQIKGGGI-NISNAQNWIQE-G       101
Sc   HVIKLG-----PNNDDAAREALQESPQFLQVGGGI-NDTNCLEWLKW--         97
Ec   HLVDLTGAKDPAKRQIPLIKTLVAGVNVPVQVGGGVRTEEDVAALLBA-G        95
Ll   HLVDLDGAKIGQTRNLELVRKIKDETRLKIEIGGGIRDFDTVRMYLEQIG        97
Ab   HLVDLNGAFEGKPVNGKAVESILGAVTVPVQLGGGIRDLKTIALWLEK-G        96
Sy   HLVDLDGAKEGQPVNLTAIANIVEALTIPVQVGGGLRDRNRVKQLLDL-G        96
Mv   HLVDLDGAIDGERVNLPCVKKIIQESKVPVQMGGGIRTIQDVEELVDL-G        96

At   ASHVIVTSYVF-NNGKIDLERLKDIVSIVGKQRLILDLSCRK-KDGRYAI        192
Sp   ASKVIVTSWLF-PDAQFDLDRLKAISSKIGKDRLVVDVSCRR-KDNRWFA        149
Sc   ASKVIVTSWLFTKEGHFQLKRLERLTELCCGKDRIVVDLSCRKTQDGRWIV       147
Ec   VARVVVGSTAVKSQ-----DMVKGWFERFGADALVVALDVRIDEDGNKQV        140
Ll   VERVILGTAAVEKP-----DFLKELLIKYGPSRIVVGVDIR----EGFV        137
Ab   VSRVILGTVALREP-----ELVREARREF-PGKVAVGIDAR-----EGYV        135
Sy   VGRVILGTIAVENP-----DLVGELCAEF-PGQIVVGIDAR-----NGKV        135
Mv   INKVIIGTVAVQNP-----DFVEQLAKKVGSDKIMVALDAK-----DGKV        136

At   VTDRWQKFSDVILDEKSLEFLGGFSDEFLVHGVDVEGKKLGIDEELVALL        242
Sp   AINKWQVMTAFELNEENLDLLSQYCSEFLIHAADVEGLCKGIDEDLVIKL        199
Sc   AMNKWQTLTDLELNADTFRELRKYTNEFLIHAADVEGLCGGIDELLVSKL        197
Ec   AVSGWQENSGVSLEQLVETYLPVGLKHVLCTDISRDGTLAGSNVSLYEEV        190
Ll   STSGWLEKTSLPYLSFLKKLERIGVKTTIITDISKDGTLTGPNFKLYDEI        187
Ab   AVAGWAETSTIKALDLALKFEDSGVAAIIYTDINRDGAMGGVNVEATSDL        185
Sy   ATRGWLETSTVEAGELAQRMEKLGAAAIIYTDIHRDGTMQGPNLDALRQL        185
Mv   VIKGWKEKTEYTPVQMGKILEEKGAGSILFTNVDSEGLLNGINITPTKEL        186

At   G----NYSPIPVTYAGGVTVMDDVERIKDAGK-GRVDVTVGSALDIFGGN        287
Sp   GE----WVKIPTTYAGGGRSIEDLDLVDKLSK-GKVDLTIGSALDIFGGV        244
Sc   FEWTKDYDDLKIVYAGGAKSVDDLKLVDELSH-GKVDLTFGSSLDIFGGN        246
Ec   CA---RYPQVAFQSSGGIGDIDDVAALR---GTGVRGVIVGRAL---LEGK        232
Ll   S----KENSLNVIISGGVKDNSDIQRA---TRSDFYGIIVGKAY--YEGK        228
Ab   A----FHLTTPVIASGAVSSIDDLIALKKEEDTGIQGVICGRAL--YDGR        229
Sy   A----SQLTIPVIASGGVSQVEDLLNLLSLESLGVNGVIIGKAL--YTGD        229
Mv   V----DNLNIPIIASGGVTTIEDLIEFKK---IGVAGVVVGSAL--YKNN        227

At   LPYKDVVA-MHHKQHSLH*                                        304
Sp   LEFTRVVA-WNRRFTPLPLQD*                                     264
Sc   LVKFEDCCRWNEKQG*                                           261
Ec   FTVKEAIACKONA*                                             245
Ll   INLEKEFRNAN*                                               239
Ab   IDPKTALDLLSSVSVSGTPVTGKAG*                                 254
Sy   IQLAEAIRAVGNGRWQDVPPPLDFPRLG*                              256
Mv   FKLQDAINTVNNFKI*                                           242
```

Fig. 5. Alignment of protein sequences predicted from *Arabidopsis* BBM II isomerase cDNA and microbial *hisA* and *HIS6* genes.
At; *A. thaliana*, Sp; *S. pombe* (Barrel et al., 1996; accession number Z69369), Sc: *S. cerevisiae* (Fani et al., 1997; accession number Z46881), Ec: *E. coli* (Carlomagno et al., 1988; accession number X13462), Ll; *L. lactis* (Delorme et al., 1992; accession number M90760), Ab; *A. brasiliense* (Fani et al., 1993; accession number X61207), Sy; *Synechocystis* sp. PCC6803 (Kaneko et al., 1996 accession number D90916), Mv; *M. voltae* (Cue et al., 1985; accession number M11218). Identical amino acid residues are shaded. Asterisks represent the stop codon, and dashes were inserted to maximize the sequence homology.

38

Hybridization analyses

Northern blot analysis showed that the size of the BBM II isomerase transcript in *Arabidopsis* is approximately 1.1 kb in length (Fig. 6), which is consistent with the insert size of pKI-1 and the 5' extension detected by cRACE. The BBM II isomerase mRNA was detected in all the samples examined, and the levels of mRNA did not differ notably different between roots and shoots throughout development (Fig. 6).
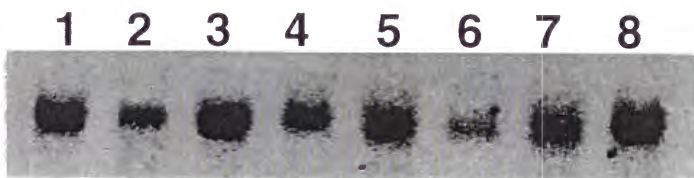
Fig. 6. Expression of BBM II isomerase mRNA in *Arabidopsis*.
Total RNA was isolated from different organs at various times; 1-week-old plants (lane 1), roots from 2-week-old plants (lane 2), leaves from 2-week-old plants (lane 3), roots from 3-week-old plants (lane 4), leaves from 3-week-old plants (lane 5), roots from 4-week-old plants (lane 6), leaves from 4-week-old plants (lane 7), siliques from 4-week-old plants (lane 8). RNA gel blots of total RNA (10 μg) were probed with a cDNA insert of pKI-1.

The number of BBM II isomerase genes in *Arabidopsis* was examined by genomic Southern analysis (Fig. 7). Digestion with enzymes lacking the recognition sites within the cDNA (*Bgl*II and *Xho*I) gave a single hybridization signal. Digestion with enzyme either of *Bam*HI, *Eco*RV or *Hind*III, each of which has a single restriction site, gave rise to two bands. At low stringency, the hybridization profile was in agreement with that obtained at high stringency (data not shown). These results indicated that the BBM II isomerase gene existed as a single copy in *Arabidopsis* genome.

Three lines of evidence indicate that the *Arabidopsis* cDNA clone described here encodes BBM II isomerase. First, the cDNA complemented an *E. coli hisA* mutant. Second, the predicted protein sequence showed a high degree of identity to *S. pombe* HIS6, including several highly conserved regions (Fig. 5). Third, the cDNA encodes a protein with an enzymatic activity catalyzing the conversion of BBM II to BBM III. While complete elucidation of the genes involved in His biosynthesis in plants must await
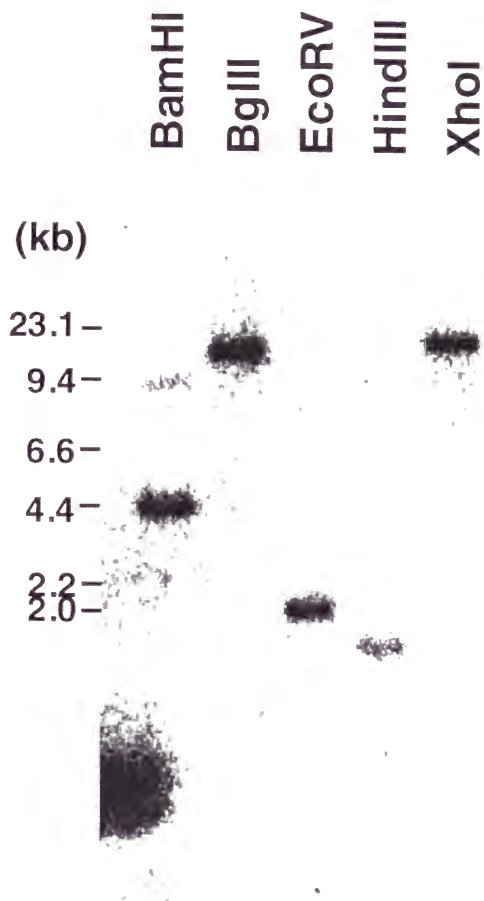
Fig. 7. DNA gel-blot analysis of *Arabidopsis* genomic DNA.
Total genomic DNA (10 μg) was digested with the restriction endonucleases indicated, and a DNA gel-blot was probed with a cDNA insert of pKI-1. The positions of molecular weight standards are indicated to the left.

further work, the functional coupling of the *Arabidopsis* BBMII isomerase with the His biosynthetic enzymes of microbial origins indicates that the enzymes involved in the His biosynthetic pathway are conserved across organisms.

Summary

I have isolated an *Arabidopsis* BBM II isomerase cDNA with the aid of functional complementation of an *E. coli hisA* mutant, strain HfrG6, using an *Arabidopsis* cDNA library. The isolated cDNA encoded a polypeptide consisting of 304 amino acids with calculated molecular weight of 33,363. Sequence comparison with the HIS6 proteins of yeasts revealed that *Arabidopsis* BBM II isomerase protein contained an N-terminal extension of approximately 40 amino acids having general properties of chloroplast transit peptides. This finding was consistent with the chloroplastic localization of other His biosynthetic enzymes such as imidazoleglycerolphosphate dehydratase and histidinol dehydrogenase in higher plants. The rest of the primary structure was 50% and 42% identical to those of the HIS6 proteins of *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae*, respectively, while prominent sequence similarity to the bacterial BBM II isomerase was not found. That the isolated *Arabidopsis* cDNA actually encoded a functionally active BBM II isomerase activity was confirmed in an *in vitro* enzyme assay using the crude extract prepared from strain HfrG6 transformed with the *Arabidopsis* BBM II isomerase cDNA.

Chapter III

An *Arabidopsis* cDNA Encoding a Bifunctional Glutamine Amidotransferase/Cyclase Suppresses the Histidine Auxotrophy of a *Saccharomyces cerevisiae his7* Mutant

Introduction

In most eubacteria including enterobacteriums such as *Escherichia coli* and *Salmonella typhimurium*, the His biosynthetic genes were organized in an operon (Carlomagno et al., 1988), while those in the lower eucaryotes are scattered throughout the chromosomes (Mortimer et al., 1994). In the archaebacteria such as *Methanococcus vanielii and M. jannaschii*, the His biosynthetic genes are not organized as an operon (Beckler et al., 1986, Bult et al., 1996). Not only the overall gene organization but the structures of each gene are variable among different organisms. For example, in *S. cerevisiae, HIS4* and *HIS7* code for multi-functional enzymes (Donahue et al., 1982, Kuenzler et al., 1993). The *HIS7* gene of *S. cerevisiae* encodes the bifunctional glutamine amidotransferase (HisH) and the cyclase (HisF), although both proteins in eubacteria are encoded by independent cistrons, *hisH* and *hisF*, which are interrupted by the *hisA* cistron (Carlomagno et al., 1988, Fleischmann et al., 1995). In higher plants, the histidinol dehydrogenase (HDH) protein catalyzes the tenth- and eleventh-steps of the His biosynthetic pathway (Nagai et al., 1991), as does the HisD protein in eubacteria (Carlomagno et al., 1988), while the imidazoleglycerolphosphate dehydratase (IGPD) in higher plants is not accompanied by a histidinolphosphate phosphatase domain (Tada et al., 1994), both of which are encoded by *hisB* in most eubacteria (Carlomagno et al., 1988, Fleischmann et al., 1995). On the other hand, phosphoribosyl (PR)-AMP cyclohydrolase (PRA-CH) and PR-ATP pyrophosphohydrolase (PRA-PH) encoded by a single gene, *hisIE* in eubacteria (Carlomagno et al., 1988) and in *Arabidopsis thaliana* described in Chapter I (Fujimori and Ohta, 1998a), were encoded by independent genes in *M. vanielii and M. jannaschii* (Beckler et al., 1986, Bult et al., 1996) and *Azospirillum*

brasilense (Fani et al., 1993).

In Chapter III, I describe the isolation and characterization of an *Arabidopsis* cDNA encoding a bifunctional glutamine amidotransferase/cyclase. Functional expression of the *Arabidopsis At-HF* cDNA in a *S. cerevisiae his7* defective mutant demonstrated that the *At-HF* cDNA encodes a single polypeptide having the bifunctional activities for glutamine amidotransferase (HisH) / cyclase (HisF).

Materials and Methods

Plant material, growth and microbial strains

Arabidopsis thaliana ecotype Columbia (Col-0) (Lehle Seeds, Tucson, AZ, USA) were grown described in Chapter I (Fujimori and Ohta, 1998a). *S. cerevisiae* strain SH782 (*MATa ura3-52 leu2-3,112*) was a generous gift from S. Harashima (Osaka University, Suita, Japan) for the construction of the *HIS7* disruption mutant. *E. coli* strain DH5α was used as a host for the propagation and manipulation of plasmid DNAs. The media for yeast were as described (Rose et al., 1990).

Isolation of an Arabidopsis *cDNA encoding a bifunctional glutamine amidotransferase/ cyclase*

An *Arabidopsis* EST clone (186B18T7; GenBank accession No. H37732) was identified as a putative cyclase (HisF) of *Arabidopsis* through the Basic Local Alignment Search Tool (BLAST) search against the *Arabidopsis* ESTs (Altschul et al., 1990, Newman et al., 1994) with the aid of the primary structure deduced from the *S. cerevisiae HIS7* gene (Kuenzler et al., 1993). In order to isolate full-length cDNAs corresponding to the EST clone, total RNA was isolated from 7-day-old *Arabidopsis* seedlings by the acid guanidium thiocyanate-phenol-chloroform extraction method (Chomczynski and Sacchi, 1987), and a DNA fragment corresponding to the EST clone was amplified employing a reverse transcription (RT)-polymerase chain reaction (PCR) strategy. Briefly, first strand cDNA corresponding to the EST clone was synthesized using a specific antisense primer PRI60 (5'-CCAACTGCTTGCGATCACCGG-3') and Superscript II™ RNaseH⁻ Reverse Transcriptase (Gibco-BRL, Rockville, MD, USA). Prior to the RT reaction, the total

RNA (5 µg) was heat-denatured at 72 °C for 3 min, and the first strand cDNA synthesis was performed at 42 °C for 60 min followed by heat-inactivation at 95 °C for 3 min. PCR was performed using a set of specific primers of PRI59 (5'-GAAATCAGGCAGTGGTTGTAAG-3') and PRI60, ExTaq DNA Polymerase (Takara Shuzo, Kyoto, Japan) and the RT reaction products (Saiki et al., 1988). After the initial denaturation at 95 °C for 5 min, PCR was carried out with 1 min at 94 °C, 1 min at 55 °C and 1 min at 72 °C for 30 cycles. A PCR-amplified 320-bp DNA fragment was directly cloned into a pCR2.1 vector (Invitrogen, San Diego, CA, USA), yielding a plasmid, pKF405. The nucleotide sequence of the insert DNA of pKF405 was confirmed to contain the expected DNA fragment (data not shown).

An *Arabidopsis* cDNA library (7-day-old) (Mizutani et al., 1997) constructed with a λZAPII vector system (Stratagene, La Jolla, CA, USA) was screened with the ^{32}P-labeled cDNA insert of pKF405. Prehybridization, hybridization and washing were carried out as described (Sambrook et al., 1989).

The nucleotide sequences were determined on both strands by the dideoxynucleotide chain-termination method (Sanger et al., 1977) using an ABI PRISM™ Dye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA).

Construction of the S. cerevisiae his7 *mutant*

A *HIS7* disruption mutant of *S. cerevisiae* BY1006 was constructed as follows. A region encoding the HIS7 ORF of *S. cerevisiae* was amplified by PCR using PRI75 (5'-GCAAGCTTATGCCGGTCGTTCACGTGATTGAC-3') and PRI76 (5'-GCCTCGAGCCACATTACTCTTCATCCATTC-3') as primers and chromosomal DNA of *S. cerevisiae* strain S288C (*MAT*a *mal mel gal2*) as the template. PCR products double-digested with *Hind*III and *Xho*I were cloned into a pYES2 vector (Invitrogen) to yield a plasmid, pKF416. The 1.0-kb *Sal*I-*Sal*I fragment from pKF416 was replaced with the 2.2-kb *Sal*I-*Xho*I fragment of *LEU2* gene (Andreadis et al., 1982) to obtain a plasmid, pKF426. *S. cerevisiae* SH782 was transformed with a 2.9-kb *Hind*III-*Xho*I fragment from pKF426 (Ito et al., 1983) and incubated on synthetic complete (SC) plates supplemented with 2% (w/v) glucose (Glu) and an amino acids mixture without Leucine (Leu) (SC+Glu-Leu) at 30 °C for 3 days. The selection for Leu$^+$ phenotype in the

44

presence of L-His resulted in strain BY1006 (*MATa ura3-52 leu2-3,112 his7::LEU2*).

Expression in S. cerevisiae his7 *defective mutant*

      The following regions were amplified by PCR using sets of gene specific primers; a 1900-bp of a coding region (nucleotide positions +1 to +1900, taking the A of the first ATG codon as +1) using PRI102 (5'-CG<u>GAATTC</u>A**TG**GAGGCTACGGCGGCGCC-3') and PRI103 (5'-CG<u>CTCGAG</u>GATTCACTCAGACCTATTGG-3'), a 1717-bp of a coding region truncated its putative chloroplast transit peptide (nucleotide positions +184 to +1900) using PRI122 (5'-CG<u>GAATTC</u>A**TG**GTTGTGACTTTGCTTGACTAC-3') and PRI103. Each of those PCR products digested with *Eco*RI and *X ho*I was cloned into a pYES2 vector to obtain plasmids, pKF433 and pKF434, respectively. Each of those plasmids was used for the transformation of *S. cerevisiae* BY1006 (*his7*). The transformants were selected by uracil auxotrophy and incubated on SC plates containing 2% (w/v) galactose (Gal) and appropriate amino acids (SC+Gal-Ura-Leu) for 3 days at 30 °C. The pKF416 carrying the DNA insert containing the open reading frame (ORF) of *S. cerevisiae* HIS7 protein was used for the transformation of strain BY1006 as a positive control. The *S. cerevisiae* transformants were analyzed for their ability to grow on minimal-galactose plates without L-His (SC+Gal-Leu-His).

Southern and Northern blot hybridization analyses

      Genomic southern and Northern blot analyses were carried out as described by Sambrook et al. (Sambrook et al., 1989).

Results and Discussion

Isolation of an Arabidopsis *cDNA encoding bifunctional glutamine amidotransferase and cyclase*

      Recent genetic studies revealed the substantial conservation of the protein primary structures of the His biosynthetic enzymes of prokaryotic and lower eukaryotic origins (Alifano et al., 1996). This sequence conservation has also been verified with respect to the His biosynthetic pathway enzymes of higher plants (Nagai et al., 1991, Tada et al.,

PRI102

```
 -5   TATCAATGGAGGCTACGGCGGCGCCATTCTCTTCAATTGTCTCTTCCAGACAAAACTTCTCTTCATCTTCTTCGATTCGC
       M  E  A  T  A  A  P  F  S  S  I  V  S  S  R  Q  N  F  S  S  S  S  I  R         25

 75   GCTTCTTCTCCGGCTTCTTTATTCCTCTCCCAGAAGAGTATTGGCAATGTTAATCGCAAATTCAAATCTCCCAGAAGCCT
       A  S  S  P  A  S  L  F  L  S  Q  K  S  I  G  N  V  N  R  K  F  K  S  P  R  S  L  52
```
PRI122
```
155   CTCCGTCCGCGCATCTTCTACCTCAGATTCTGTTGTGACTTTGCTTGACTACGGAGCTGGAAATGTTCGGAGCATCCGCA
       S  V  R  A  S  S  T  S  D  S  V  V  T  L  L  D  Y  G  A  G  N  V  R  S  I  R    78

235   ATGCTCTTCGTCATCTCGGCTTCAGCATCAAAGACGTTCAAACGCCGGGAGACATTCTGAATGCTGATCGACTCATATTT
       N  A  L  R  H  L  G  F  S  I  K  D  V  Q  T  P  G  D  I  L  N  A  D  R  L  I  F 105

315   CCAGGCGTTGGGCCTTTTGCACCCGCCATGGATGTACTTAACAGAACTGGGATGGCTGAAGCTTTGTGCAAATATATTGA
       P  G  V  G  P  F  A  P  A  M  D  V  L  N  R  T  G  M  A  E  A  L  C  K  Y  I  E 132

395   GAATGACCGTCCATTTCTAGGCATATGTCTTGGTCTACAACTACTTTTCGATTCTAGTGAACAGAATGGACCAGTCAAAG
       N  D  R  P  F  L  G  I  C  L  G  L  Q  L  L  F  D  S  S  E  Q  N  G  P  V  K    158

475   GTCTTGGTGTGATACCGGGAATAGTTGGACGCTTTGATGCTTCAGCTGGTATAAGAGTACCCCACATTGGCTGGAATGCT
       G  L  G  V  I  P  G  I  V  G  R  F  D  A  S  A  G  I  R  V  P  H  I  G  W  N  A 185

555   TTGCAAGTTGGGAAGGATTCTGAAATTTTGGATGATGTTGGAAACCGTCATGTCTATTTTGTTCATTCGTACAGGGCCAT
       L  Q  V  G  K  D  S  E  I  L  D  D  V  G  N  R  H  V  Y  F  V  H  S  Y  R  A  I 212

635   TCCATCAGATGAAAATAAGGACTGGATTTCGTCTACCTGTAATTATGGTGAATCATTTATATCTTCCATAAGAAGGGGAA
       P  S  D  E  N  K  D  W  I  S  S  T  C  N  Y  G  E  S  F  I  S  S  I  R  R  G    238

715   ATGTGCATGCAGTTCAATTCCATCCTGAAAAGAGCGGGGAGGTGGGGCTTTCTGTTTTAAGAAGGTTCTTGCATCCAAAA
       N  V  H  A  V  Q  F  H  P  E  K  S  G  E  V  G  L  S  V  L  R  R  F  L  H  P  K 265

795   TTACCTGCAACACAGAAGCCAATGGAAGGAAAGGCCTCAAAACTTGCAAAGAGGGTGATTGCTTGTCTTGATGTGAGGAC
       L  P  A  T  Q  K  P  M  E  G  K  A  S  K  L  A  K  R  V  I  A  C  L  D  V  R  T 292

875   GAATGATAAAGGAGATCTCGTAGTTACTAAAGGGGATCAGTATGATGTGAGAGAGCAATCTAATGAAAACGAGGTTCGAA
       N  D  K  G  D  L  V  V  T  K  G  D  Q  Y  D  V  R  E  Q  S  N  E  N  E  V  R    318

955   ACCTTGGCAAACCTGTTGATTTGGCTGGGCAGTATTACAAAGATGGTGCAGATGAGATTAGCTTTTTAAACATAACTGGA
       N  L  G  K  P  V  D  L  A  G  Q  Y  Y  K  D  G  A  D  E  I  S  F  L  N  I  T  G 345

1035  TTCCGCGATTTTCCTCTAGGGGATTTGCCGATGATTCAGGTGTTGAGGCAGACATCAAAGAATGTCTTTGTACCACTAAC
       F  R  D  F  P  L  G  D  L  P  M  I  Q  V  L  R  Q  T  S  K  N  V  F  V  P  L  T 372

1115  TGTTGGAGGTGGTATTAGAGACTTTACAGATGCTAGTGGCAGGTACTATTCTAGCTTGGAAGTTGCTGCTGAGTATTTCA
       V  G  G  G  I  R  D  F  T  D  A  S  G  R  Y  Y  S  S  L  E  V  A  A  E  Y  F    398

1195  GATCCGGTGCTGATAAGATGTCCATAGGAAGTGACGCTGTTTTTGCTGCAGAGGAGTTCATAAAATCAGGGGTGAAGACA
       R  S  G  A  D  K  M  S  I  G  S  D  A  V  F  A  A  E  E  F  I  K  S  G  V  K  T 425
```
PRI59
```
1275  GGAAAGAGTAGTTTAGAACAGATATCCAGAGTTTATGGAAATCAGGCAGTGGTTGTAAGTATTGATCCTCGTAGAGTTTA
       G  K  S  S  L  E  Q  I  S  R  V  Y  G  N  Q  A  V  V  V  S  I  D  P  R  R  V  Y 452

1355  TGTGAACCATCCGGATGATGTGCCATACAAAGTCATCAGAGTAACTAATCCAGGCCCAAATGGAGAAGAATATGCCTGGT
       V  N  H  P  D  D  V  P  Y  K  V  I  R  V  T  N  P  G  P  N  G  E  E  Y  A  W    478

1435  ATCAGTGCACGGTCAGTGGAGGACAAGAAGGTCGACCTATTGGAGCATTTGAGCTTGCGAAAGCGGTTGAAGAATTAGGT
       Y  Q  C  T  V  S  G  G  Q  E  G  R  P  I  G  A  F  E  L  A  K  A  V  E  E  L  G 505

1515  GCCGGTGAAATACTATTGAACTGCATAAACTGTGATGGTCAAGGGAAAGGATTCGACATAGACTTAGTAAAGCTCATCTC
       A  G  E  I  L  L  N  C  I  N  C  D  G  Q  G  K  G  F  D  I  D  L  V  K  L  I  S 532
```
PRI60
```
1595  AGATTCAGTAGGCATACCGGTGATCGCAAGCAGTGGAGCAGGTACTCCCGACCACTTTTCCGAGGTGTTTGAAGAAGACA
       D  S  V  G  I  P  V  I  A  S  S  G  A  G  T  P  D  H  F  S  E  V  F  E  E  D    558

1675  AACGCATCTGCCGCGCTTGCTGCCGGCATTTTCCACCGGAAAGAGGGTACCAATCCCAATCTGTGAAAGAGCACTTACAA
       K  R  I  C  R  A  C  C  R  H  F  P  P  E  R  G  Y  Q  S  Q  S  V  K  E  H  L  Q 585

1755  GAGGAGCGCATAGAAGTCAGGATCTGAGAATTTTCTGGTCTGCGTGTTACCATAATTCATGACACATTCAACAAGATAGA
       E  E  R  I  E  V  R  I  *                                                      593
```
PRI103
```
1835  ATCTTTACCTTCAATAAAATTAAATTAAGTCGGAACCATCGGAGCTCCAATAGGTCTGAGTGAATCCTCTTATAGTTGATT

1915  TCCTTTTGTTGACAGTAAGACTGATACAATTTATGGATTTCCCTTGAAAGTAAGAAAGCAACGTATGTTCACTTAAAAAAA

1995  AAAAAAAAAAA
```

Fig. 1.   Nucleotide and predicted amino acid sequences of the *Arabidopsis At-HF* cDNA. Numbers shown left and right to the columns refer nucleotide and amino acid sequences, respectively. Asterisk indicates the stop codon. Primers used were indicated by arrows. Possible polyadenylation signal was underlined. The DDBJ accession numbers are AB006210 (*Arabidopsis At-HF* cDNA) and AB016783 (*Arabidopsis At-HF* gene).

1994). In this Chapter, I describe a novel cDNA encoding a glutamine amidotransferase/cyclase, the At-HF from *A. thaliana*. A database search against the expression sequence tags (ESTs) using the HIS7 protein sequence hit an EST clone (186B18T7) of *A. thaliana*. The amino acid sequence encoded by this clone showed a high degree of similarity to the C-terminal portion of the bifunctional glutamine amidotransferase/cyclase (HIS7) from *S. cerevisiae*. An *Arabidopsis* cDNA library was then screened for a full length cyclase cDNA using a PCR-amplified 186B18T7 sequence as a probe, and 4 positive clones out of 50,000 plaques were finally isolated. From the partial DNA sequencing, these four clones were identical except for different lengths at their 5'-ends (data not shown). The plasmid harboring the longest insert was designated pKF412. The cDNA insert of pKF412 is 2011-bp long and contains an ORF encoding a polypeptide of 593 amino acids with a molecular mass of 64,720 Da (Fig. 1). Analysis of the 5'-untranslated region of the *Arabidopsis At-HF* gene revealed an in-frame stop codon (TAG) at 84 nucleotides upstream the putative translation initiation codon (ATG) of the pKF412 insert, indicating that the ORF found in the pKF412encodes a full length transcript.

Amino acid sequence comparison

The amino acid sequence deduced from the *At-HF* cDNA was compared to those of microorganisms currently available on the GenBank/EMBL/DDBJ/SwissProt databases (Fig. 2). Sequence alignment indicated that the N-terminal domain encompassing 64-Val to 262-Leu and the C-terminal domain from 280-Leu to 555-Phe of the At-HF protein showed significant homology to the glutamine amidotransferase (HisH) and the cyclase (HisF) from microorganisms, respectively (Fig. 2).

In yeast *S. cerevisiae*, the glutamine amidotransferase (HisH) and the cyclase (HisF) domains are located at the N-terminal half and C-terminal half of a single polypeptide encoded by *HIS7*, respectively (Kuenzler et al., 1993).

Current results indicated that the glutamine amidotransferase/cyclase activities resided in a single polypeptide encoded by a single gene in *A. thaliana* like that found in *S. cerevisiae*. On the contrary, in enterobacteria such as *E. coli* and *S. typhimurium*, the proteins are encoded by two independent cistrons in the same operon, *hisH* and *hisF*

```
At-HF     1   MEATAAPFSSIVSSRQNFSSSSSIRASSPASLFLSQKSIGNVNRKFKSPRSLSVRASSTSDSVVTLLDYG
Sc HIS7   1                                                                 MPVVHVIDVE
Ec hisH   1                                                                  MNVVILDTG
Sy hisH   1                                                                 MGYIAVVDYD
Mj hisH   1                                                                 MKMIGIIDYN

At-HF    51   AGNVRSIRNALRHLGFSIKDVQTPGD..ILNADRLIFPGVGPFAPAMDVL..NRTCMAEALCKYIENDRE
Sc HIS7  11   SGNLQSLTNAIEHLGYEVQLVKSPKDFNISGTSRLILPGVGNYAHFVDNI..FNRGFEKPIREYIESGKE
Ec hisH  10   CANLNSVKSAIARHGYEPKVSRDPDV..VLLADKLFLPGVGTAQAAMDQV..REFELFDLIK...ACTQP
Sy hisH  11   MGNLHSVCKGLEKVGGNPLVTDQAHV..IDGATAIVLPGVGSFDPAVQHL..RARGLEEVLKTAIAKGIP
Mj hisH  11   AGNLRSIQKA.VELYDKVIITNNSEE..LLACDKIIILPGVGNFGSAMENLAPLKETIYKIVDDR....VP

At-HF    97   FLGICLGLQLLFDSSEQNGPVKGLGVIPGIVGRFDASAGIRVPHIGWNALQVGKDSEIL..DDVGNRHVY
Sc HIS7  59   IMGICVGLQALFAGSVESPKSTGLNYIDFKLSRFDDSEK.PVPEIGWNS.CIPSENLFF..GLDPYKRYY
Ec hisH  53   VLGICLGMQLLGRRSEESNGVDLLGIIDEDVPKMT.DFGLPLPHMGWNRVYPQAGNR.LFQGIEDGAYFY
Sy hisH  57   FLGICLGSQLLFDSSEEGQET.GLGIIPGQVKHFRSEPGLTIPHMGWNGLQFNQPDLCLWQDLPPEPQVY
Mj hisH  41   FLGICLGMQILFEESEKRGIKGLGIIKGNYIKFK.DVE.KLPHMGWNSVKIVKDCP.LFEGIKNNSYFY

At-HF   145   FVHSYRAIESDENK......DWISSTCNYG.ESFISSIRRGNVHAVQFHPEKSGEVGLSVIRRFLH....
Sc HIS7 105   FVHSFAAILNSEKKKNLENDGWKIAKAKYGSEEFIAAVNKNNIFATQFHPEKSGKAGLNVIENFLKQQSP
Ec hisH 101   FVHSYAM.FVNPWT........IAQ.CNYGEP.FTAAVQKDNFYGVQFHPERSGAAGAKLLKNFLEM*
Sy hisH 106   FVHSYYMAPLDPQV........IAASTTHGSQTIAAAIAKDNVMAVQFHPEKSSTLGIKILANFVKKVAL
Mj hisH 141   FVHSYHVNPDEDCI........VGK.TEYGR.EFPSVINKDNVFATQFHPEKSGKIGLKIIENFVELL*

At-HF   184   ..PKLPATQKPM...EGKASKLAKRVIACLDVRTNDKGDLVVIKGDQYDVREQSNENEVRNLGKPVDLAG
Sc HIS7 155   PIPNYSAEEKELLMNDYSNYGLTRRIIACLDVRTNDQGDLVVIKGDQYDVREKSDGKGVRNLGKPVQLAQ
Ec hisH
Sy hisH 148   SDI*
Mj hisH
Ec hisF   1              MLAKRIIPCLDVRDGQ.....VVKGVQF........RNHEIIGDIVELAK
Sy hisF   1             MTLAKRILPCLDVNACR.....VVKGINF........VDLQDAGDPVELAR
Mj hisF   1              MLTKRIIPCLDIKDGR.....VVKGTKF........LNLRDAGDPVELAQ

At-HF   229   QYYKDCADEISFLNITGFRDFPLGDLPMIQVLRQISKNVFVPLTVGGGIRDFTDASGRYYSSLEVAAEYF
Sc HIS7 205   KYYQQGADEVIFLNITSFRDCPLKDTPMLEVLKQAAKTVFVPLTVGGGIKDIVDVDGTKIPALEVASLYF
Ec hisF  38   RYAEEGADELVFYDITASSD...GRVVDKSWVSRVAEVIDIPFCVAGGIKSLEDAA..........KIL
Sy hisF  39   AYNEAGADELVFLDITATHE...QRDTIIDVVYRTAEEVFIPLTVGGGISTLEHIK..........NLL
Mj hisF  38   YYDDEGADELVFLDITASAE...KRDIIIDVVERTAEKVFIPLTVGGGIKSIEDFR..........RIL

At-HF   279   RSGADKMSIGSDAVFAAEEFIKSGVKT.GKSSLEQISRVYGCNQAVVVSIDPRRVYVNHPDDVPYKVIRVT
Sc HIS7 255   RSGADKVSIGTDAVYAAEKYYELGNRGDGTSPIETISKAYGCAQAVVISVDPKRVYVNSQADTKNKVFETE
Ec hisF  74   SFGADKISINSEAL.............ADPTLITRLADRFGVQCIVVGIDTWYDAETGKYHVNQYT....
Sy hisF  75   RAGADKVSVNSSAV.............RDPDFISRASDRFGRQCIVVAIDARRRLDADNPG.........
Mj hisF  74   RAGADKVSINTAAV.............KNPNLIKEASEIFGSQCVVVAIDAKRHYVNEDEIDKINKNVVK

At-HF   328   NPGPNGEEMAMYQCTVSGGQEGRPIGAFELAKAVEELGAGEILLNCINCDGQGKGFDIDLVKLISDSVGI
Sc HIS7 305   YPGPNGEKYQYYQCTIKGGRESRDLGVWELTRACEALGAGEILLNCIDKDGSNSGYDLELIEHVKDAVKI
Ec hisF 107   .............GDESRTRVIQWETLDWVQEVQKRGAGEIVLNMMNQDGVRNGYDLEQLKKVREVCHV
Sy hisF 103   .........W.DVYVRGGRENIGLDAIAMAEEVAKRGAGEILVTSMDGDGTQAGYDLALTAAIAERVBI
Mj hisF 111   ....VEDGVQVFEVYIYGGRKEIGIDAINQAKKVEELGAGEILLTSIDKDGTKSGYDLILTKEISKSVKL

At-HF   378   PVIASSGAGTPDHFSEVFEEDK.RICRACCRHFPPERGYQSQSVREHLQEERIEVRI*
Sc HIS7 355   PVIASSGAGVPEHFDEAFLKTRADACLGAGM.FHRG.EFTVNDVREYLLEHGLKVRMDEE*
Ec hisF 143   PLIASGGAGTMEHFLEAFRDADVDGALAASV.FHKQIINIGELKAYLATQGVEIRIC*
Sy hisF 143   PVIASGGAGNCQHVYEAFTEGKABAALLASL.LHYGQLTIGELKTFLEARQIPVRHTAVCG*
Mj hisF 157   PVIASGGAGKPEHVYEAFVYGKADAALMAGI.LHYREYTIFEIKKYCADRGIPMRLL*
```

Fig. 2.   Alignment of the predicted amino acid sequences of *Arabidopsis At-HF*. *S. cerevisiae* HIS7 (accession No. X69815) (Kuenzler et al., 1993) and *hisH*, *hisF* gene products from *E. coli* (accession No. X13462) (Carlomagno et al., 1988), *Synechocystis* sp. PCC6803 (accession No. D64004, D90912) (Kaneko et al., 1996) and *M. jannaschii* (accession No. U67493, U67500) (Bult et al., 1996). Conserved amino acid residues are shaded.

(Carlomagno et al., 1988). On the other hand, in archaebacteria such as *M. vanielii* and *M. jannaschii* (Beckler et al., 1986, Bult et al., 1996), and *Synechocystis* sp. PCC6803 (Kaneko et al., 1996), the two genes scattered throughout the chromosome. These observations imply different evolutional processes or unique chromosomal gene organization processes of the glutamine amidotransferase/cyclase genes in these organisms, but the high sequence similarity and catalytic domains have been conserved (Fig. 2).

Sequence comparison (Fig. 2) revealed an N-terminal extension of approximately 60 amino acids of the At-HF protein exhibiting properties typical of chloroplast transit peptides, being rich in hydroxylated residues and a few negatively charged residues (von Heijne and Nishikawa, 1991). This is consistent with the previous findings that the proteins of the IGPD of *Triticum aestivum* (Tada et al., 1995) and the HDH of *Brassica oleracea* (Nagai et al., 1993) were immunochemically detected in intact chloroplast fractions.

Functional complementation of a S. cerevisiae his7 *mutant with the* At-HF *cDNA of* Arabidopsis

In order to investigate whether or not the *Arabidopsis* cDNA cloned in pKF412 actually encodes a bifunctional glutamine amidotransferase (HisH) and cyclase (HisF), we employed complementation analysis using a *S. cerevisiae his7* mutant, of which the DNA region of HIS7 ORF has been replaced with *S. cerevisiae LEU2* gene (Fig. 3). This His auxotrophic mutant of *S. cerevisiae* (BY1006) is defective in both enzymatic activities of the glutamine amidotransferase and cyclase, and is this no longer able to grow on minimal plate without L-His (SC+Glu-Leu-His) (Fig. 4). Mutant strain BY1006 (*his7*) was transformed with either pKF416 (containing the *S. cerevisiae HIS7* ORF), pKF433 (bearing the *At-HF* ORF), pKF434 (harboring the *At-HF* ORF truncated its putative chloroplast transit peptide portion) or pYES2, and was cultivated on SC+Glu-Ura-Leu plates for 3 days at 30 °C. Strain BY1006 carrying either pKF433 or pKF434 was able to grow on the SC+Gal-Leu-His plates at the same level as that transformed with pKF416 (Fig. 4). These results indicated that the *At-HF* cDNA cloned in pKF412 encodes a functional glutamine amidotransferase/cyclase of *Arabidopsis*.
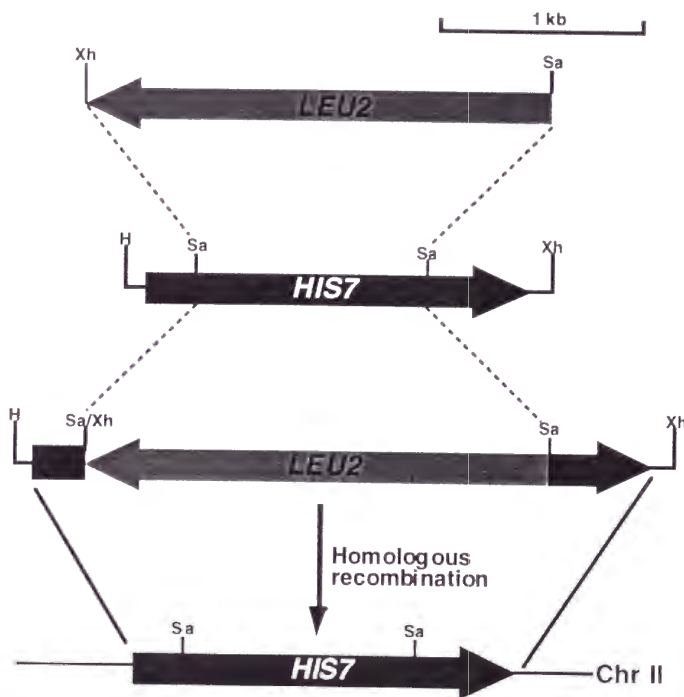
Fig. 3. Construction of the *S. cerevisiae his7::LEU2* null allele, BY1006.
Restriction enzyme sites: H; *Hind*III, Sa; *Sal*I, Xh; *Xho*I. Chr II indicates chromosome 2.

Since there was not significant difference in the growth between BY1006/pKF433 and BY1006/pKF434, it can be concluded that the N-terminal extension of the At-HF protein was not required for the enzymatic activities for glutamine amidotransferase/cyclase (Fig. 4). It is therefore possible that this N-terminal portion was in fact a chloroplast
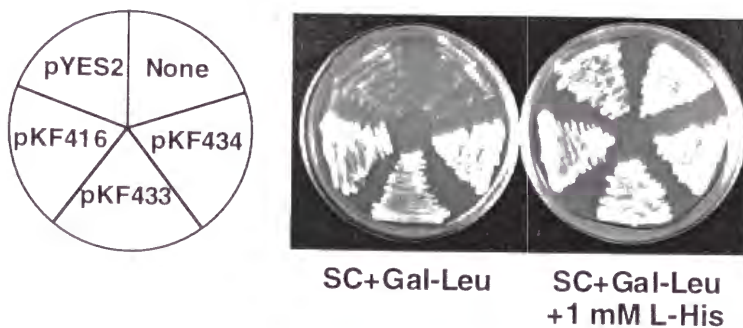
Fig. 4. Suppression of the His auxotrophy in *S. cerevisiae his7* mutant, BY1006. Strain BY1006 was transformed with either pKF416 (containing the *S. cerevisiae HIS7* ORF), pKF433 (bearing the *A. thaliana At-HF* ORF), pKF434 (harboring the *A. thaliana At-HF* ORF truncated its putative chloroplast transit peptide region) or pYES2 and cultivated on SC+Gal-Leu plate in the presence or absence of 1 mM L-His.

50

transit peptide, as is found with other His biosynthetic enzymes of higher plants (Nagai et al., 1991, Tada et al., 1994, El Malki et al, 1998, Chapter I; Fujimori and Ohta, 1998a, Chapter II; Fujimori et al., 1998b, Chapter IV). Further biochemical studies are needed to confirm the chloroplastic localization of the At-HF protein.

Southern blot analysis

Fig. 5 shows a genomic Southern blot analysis for the *At-HF* gene of *Arabidopsis*. *Arabidopsis* genomic DNA was digested with the enzymes described below and hybridized with the $[\alpha\text{-}^{32}P]$-labeled full length *At-HF* cDNA. Digestion with the enzymes *Sma*I or *Xho*I (no restriction sites in the *At-HF* cDNA) gave a single hybridization signal, whereas digestion with *Bgl*II, *Eco*RV or *Hind*III (each of which contains a single restriction site in the cDNA) gave rise to two or more hybridization bands (Fig. 5). No
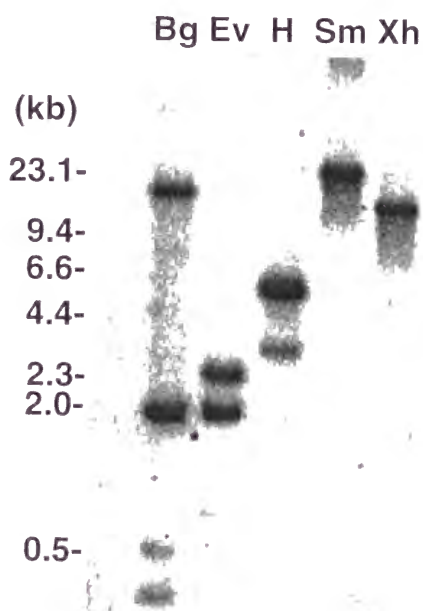
Fig. 5. Genomic Southern blot analysis.
Total genomic DNA prepared from 2-week-old grown *Arabidopsis* seedlings was digested with several restriction enzymes as indicated (Bg; *Bgl*II, Ev; *Eco*RV, H; *Hind*III, Sm; *Sma*I and Xh; *Xho*I). Ten µg of the digested DNA was separated on a 0.7% (w/v) agarose gel and hybridized with the ^{32}P-labeled full-length *At-HF* cDNA under low stringency condition. Molecular size markers are shown on the left.

51

more additional signals were observed even after long exposure (data not shown). These results indicated that there was a single copy of the *At-HF* gene in the *Arabidopsis* genome.

Expression of the At-HF *gene in* Arabidopsis

To study the expression patterns of the *At-HF* gene, Northern blot analysis was performed with total RNA samples from various tissues using a $[\alpha\text{-}^{32}P]dCTP$ labeled full
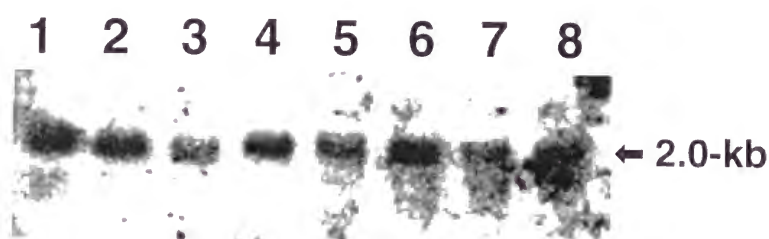
Fig. 6. Northern blot analysis of the *At-HF* gene expression.
Total RNA (10 µg) prepared from 1-week-old *Arabidopsis* (lane 1), roots from 2-week-old plants (lane 2), leaves from 2-week-old plants (lane 3), roots from 3-week-old plants (lane 4), leaves from 3-week-old plants (lane 5), roots from 4-week-old plants (lane 6), leaves from 4-week-old plants (lane 7), inflorescence stems from 4-week-old plants (lane 8) was electrophoresed in a 2.2 M formaldehyde-1.2% (w/v) agarose gel, and hybridized with the ^{32}P-labeled full-length *At-HF* cDNA. Equal loading of RNA was confirmed by staining the gel with ethidium bromide before tranfer onto the membrane (data not shown).

length *At-HF* cDNA as a probe. The size of the hybridization signal of approximately 2.0-kb corresponds well to the transcript size predicted from the cDNA size (Fig. 6). The *At-HF* gene was expressed similarly in all tissues throughout development. A ubiquitous expression pattern was also noted for other His biosynthetic genes (Tada et al., 1994, El Malki et al, 1998, Chapter I; Fujimori and Ohta, 1998a, Chapter II; Fujimori et al., 1998b, Chapter IV). These results indicated that there are no specific organs/tissues that produce His, but rather this amino acid is synthesized and supplied throughout the plant.

At least two plant His biosynthetic enzymes, IGPD (Tada et al., 1995) and HDH (Nagai et al., 1993), have been shown to be localized at chloroplasts. The His biosynthesis is an energy consuming process, where 41 ATP molecules are utilized for

the synthesis of each His molecule (Brenner and Ames, 1971). It is therefore advantageous for plants to compartmentalize the entire His biosynthetic pathway in chloroplasts. Intracellular localization of the At-HF together with other His biosynthetic pathway enzymes remains to be clarified through biochemical means.

Summary

A cDNA encoding a glutamine amidotransferase and cyclase catalyzing the fifth- and sixth-steps of the histidine (His) biosynthetic pathway has been isolated from *Arabidopsis thaliana*. The N- and C-terminal domains of the primary structure deduced from a full-length *Arabidopsis hisHF* (*At-HF*) cDNA showed a significant homology to the glutamine amidotransferase and cyclase of microorganisms, respectively. Effective suppression of the His auxotroph of a *Saccharomyces cerevisiae his7* mutant with the *At-HF* cDNA confirmed that the At-HF protein has the bifunctional glutamine amidotransferase (HisH) and cyclase (HisF) activities.

Chapter IV

Characterization of Two *Arabidopsis* isoforms for ATP Phosphoribosyl Transferase Regulating the Histidine Biosynthesis

Introduction

Histidine (His) is one of the essential amino acids for animal life, while it is *de novo* synthesized in microorganisms and plants (Alifano et al., 1996). Biochemistry and genetics of the 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 genes involved in the His biosynthetic pathway have been determined, and it was shown that the ten enzymatic activities are encoded by eight genes organized in a single operon (Carlomagno et al., 1988). Moreover, the His biosynthetic genes have also been isolated from a variety of organisms including lower eukaryotes such as *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and *Neurospora crassa* (Alifano et al., 1996). The rate of the His biosynthesis in microorganisms is regulated by the attenuation control of the expression of the His operon (Kasai, 1974, Artz and Broach, 1975, Johnson et al., 1980, Blasi and Bruni, 1981). The regulation at the protein level is also involved in the fine-tuning of the pathway activity to reflect cellular metabolic status. The biochemical regulation of the His biosynthesis is solely achieved at the first step of the His biosynthetic pathway, which is catalyzed by ATP phosphoribosyl transferase (ATP-PRT; EC 2.4.2.17) through the feedback inhibition by L-His (Ames et al., 1961, Martin, 1963, Klungsoyr et al., 1968, Brenner and Ames, 1971). It has been well-established that cellular conditions including energy charge parameter regulate the activity of ATP-PRT, thereby controlling the rate of the His biosynthesis in the cells of *E. coli* (Klungsoyr et al., 1968) and *S. typhimurium* (Morton and Parsons, 1977). In addition, it has also been

```
                          ATP + PRPP
S. cerevisiae / E. coli
    HIS1 / hisG            ATP-PRT  ◄ ·······················

    HIS4B / hisE           PRA-PH

    HIS4A / hisI           PRA-CH

    HIS6 / hisA            BBM II isomerase

    HIS7 / hisH            GAT  ─────────────►  de novo purine
                                                biosynthesis
    HIS7 / hisF            Cyclase

    HIS3 / hisB            IGPD

    HIS5 / hisC            HPA

    HIS2 / hisB            HPP

    HIS4C / hisD           HDH

    HIS4C / hisD           HDH

                          L-His  ·······  feedback inhibition  ···
```
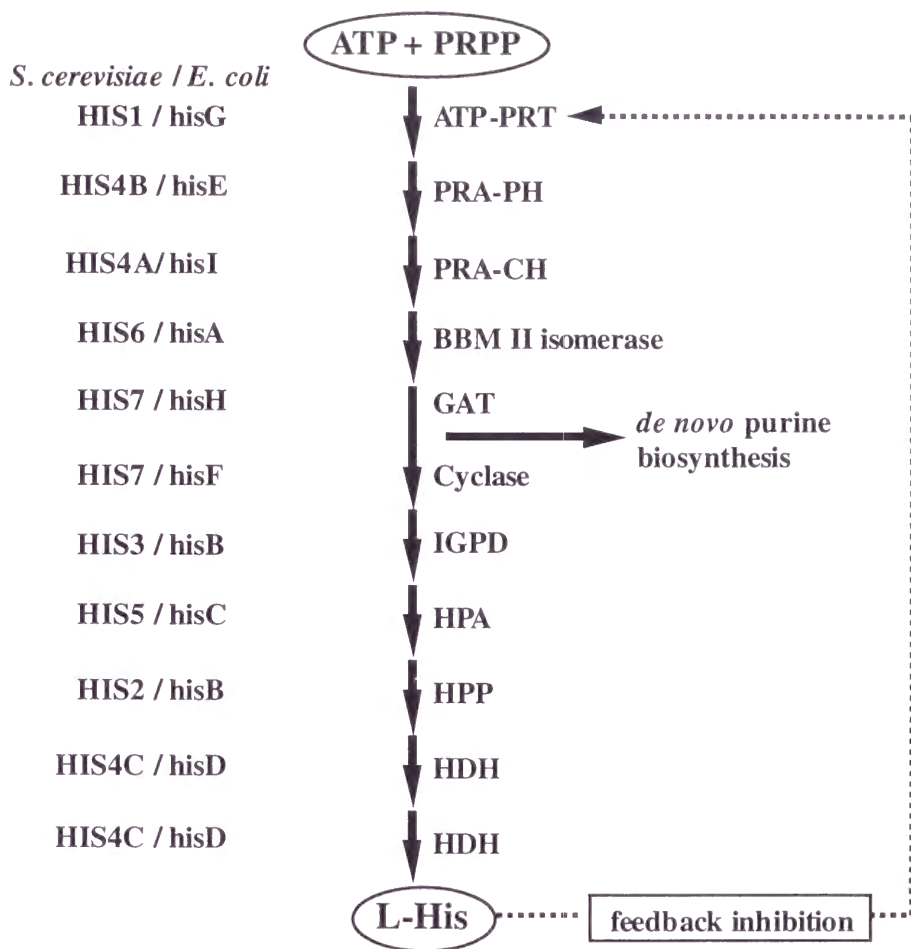
Fig. 1.   Metabolic scheme of the histidine biosynthesis in microorganisms.
Abbreviations: ATP-PRT; ATP phosphoribosyl transferase, PRA-PH; phosphoribosyl-ATP pyrophosphohydrolase, PRA-CH; phosphoribosyl-AMP cyclohydrolase, BBM II; *N'*-[(5'-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide ribonucleotide, GAT; glutamine amidotransferase, IGPD; imidazoleglycerolphosphate dehydratase, HPA; histidinolphosphate aminotransferase, HPP; histidinolphosphate phosphatase, HDH; histidinol dehydrogenase.

reported that excess accumulation of His causes the repression of the His operon expression (Kasai, 1974, Artz and Broach, 1975, Johnson et al., 1980, Blasi and Bruni, 1981).

Studies on the His biosynthesis in higher plants can be traced back to 1971 when Wiater et al. detected the enzyme activities for ATP-PRT, imidazoleglycerolphosphate dehydratase (IGPD; EC 4.2.1.19), and histidinolphosphate phosphatase (HPP; EC 3.1.3.15) in crude extracts from the shoots of barley, oat, and pea (Wiater et al.,

1971). Nonetheless, purification and characterization of these plant enzymes remains difficult, and it was quite recently that the histidinol dehydrogenase (HDH; EC 1.1.1.23) from cabbage (Nagai and Scheidegger, 1991) and the IGPD from wheat germ (Mano et al., 1993) were purified and the corresponding cDNAs have also been cloned (Nagai et al., 1991, Tada et al., 1994). On the other hand, recent progress of molecular biology has considerably accelerated the elucidation of the plant His biosynthetic pathway. Thus, I have isolated and characterized other His biosynthetic genes from *Arabidopsis* including those for the bifunctional phosphoribosyl (PR)-ATP pyrophosphohydrolase (PRA-PH)/PR-AMP cyclohydrolase (PRA-CH) (in Chapter I; Fujimori and Ohta, 1998a), *N*'-[(5'-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (BBM II) isomerase (in Chapter II; Fujimori et al., 1998b), and the bifunctional glutamine amidotransferase/cyclase (in Chapter III; Fujimori and Ohta, 1998c). In addition, a cDNA for histidinolphosphate aminotransferase was cloned from *Nicotiana tabacum* by functional complementation of an *E. coli hisC* auxotroph (El Malki et al., 1998). These results suggested that the His biosynthetic route in plants is essentially the same as those operating in microorganisms, leaving ATP-PRT and histidinolphosphate phosphatase for further investigation.

In this Chapter, I describe biochemical and molecular biological characterization of two isoforms for ATP-PRT (AtATP-PRT1 and AtATP-PRT2) from *Arabidopsis*. These two *Arabidopsis* ATP-PRT cDNAs were able to suppress the His auxotrophy of *S. cerevisiae his1* defective mutant, indicating that both are encoding active enzymes, respectively. Gene specific probes revealed that both transcripts in all the tissues examined, and protein-blot analyses with polyclonal antibodies against recombinant AtATP-PRT1 protein revealed ATP-PRT proteins were detected throughout development. The biochemical properties of the recombinantly expressed ATP-PRT proteins indicated that the His biosynthesis in plants is also metabolically regulated by L-His.

Materials and methods

Plant materials and microbial strains

Seeds of *A. thaliana* (ecotype Columbia; Lehle Seeds, Tucson, AZ, USA) were

surface-sterilized and cultivated as described in Chapter I (Fujimori and Ohta, 1998a).

The *E. coli* and *S. cerevisiae* strains used in this work were listed in Table 1. *E. coli* JM109 was used as a host for the propagation and manipulation of plasmid DNAs.

Table 1. Microbacterial strains and plasmids used in this study

Material	Property	Reference or source
E.coli		
JM109	*recA1 endA1 gyrA96 thi hsdR17 supE44 relA1 Δ (lac-proAB)/ F'(traD36 proAB lacI^q lacZ ΔM15)*	Sambrook et al. (1989)
XLI-Blue	*hsdR17 endA1 supE44 thi recA1 relA1 lac/F'(proAB lacI^q lacZΔM15::Tn10) gyrA96 relA1 lac[F proAB lacI^q ZΔM15 Tn10]*	Stratagene
S.cerevisiae		
S288C	*MATa mal mel gal2*	Rose et al. (1990)
SH782	*MATa ura3-52 leu2-3,112*	S. Harashima[a]
BY1001	*MATa ura3-52 leu2-3,112 his1::LEU2*	This work
Plasmids		
pAR5-13	pCRII bearing *A. thaliana AtATP-PRT1* fragment	This work
pAtATP-PRT1	pBluescript bearing *A. thaliana AtATP-PRT1* cDNA	This work
pAtATP-PRT2	pBluescript bearing *A. thaliana AtATP-PRT2* cDNA	This work
pMA1	pMAL-c2 bearing *A. thaliana AtATP-PRT1* coding region	This work
pMA2	pMAL-c2 bearing *A. thaliana AtATP-PRT2* coding region	This work
pKF110	pYES2 bearing *S. cerevisiae HIS1* coding region	This work
pKF157	plasmid for disruption of *S. cerevisiae HIS1* gene	This work
pKF251	pYES2 bearing *A. thaliana AtATP-PRT1* coding region	This work
pKF252	pYES2 bearing *A. thaliana AtATP-PRT2* coding region	This work
pKF411	pBluescript bearing *A. thaliana AtATP-PRT1* gene	This work

[a] Department of Biotechnology, Osaka University, Suita-shi, Osaka, Japan

The media for *E. coli* and *S. cerevisiae* were as described previously (Sambrook et al., 1989, Rose et al., 1990).

Isolation of Arabidopsis ATP-PRT *cDNAs*

In order to isolate the *Arabidopsis ATP-PRT* cDNAs, first of all, the DNA fragments were amplified by PCR using the primers, AR5 corresponding to the amino acid residues, Y-I-F-D-E-D-T which was determined from purified *T. aestivum* ATP-PRT (Münzer et al., 1992) and SK derived from vector sequence (Table 2), and the bacteriophage λZAPII DNA bearing cDNAs prepared from *A. thaliana* cDNA library (7-day-old seedlings) (Mizutani et al., 1997) as the template. The resultant PCR products were cloned into pCRII vector (TA Cloning Kit; Invitrogen, San Diego, CA, USA) to obtain pAR5-13.

To isolate the full-length *Arabidopsis ATP-PRT* cDNAs, an *Arabidopsis* cDNA

58

library (Mizutani et al., 1997) was screened by using insert DNA from pAR5-13 labeling with [α-^{32}P]dCTP (Amersham, Buckinghamshire, UK) by the random priming method (Feinberg and Vogelstein, 1983). Prehybridization, hybridization and wash were performed as described in Chapter I (Fujimori and Ohta, 1998a). Twenty six out of 6x 10^5 recombinant phages obtained through the twice plaque-purifications were converted to phagemids by *in vivo* excision method in accordance with the manufacturer's

Table 2.   Oligonucleotides used in this work

AR5	5'-GTCTCCTCGTCAAATATGTA-3'
SK	5'-TCTAGAACTAGTGGATC-3'
EF3	5'-CGGGATCCATGAAGCGTGACCAGATTCGTCTTG-3'
ERXH	5'-GCTCTAGAAGCTTCAGCATATGCATCTTCC-3'
AP1R	5'-CAAGACGAATCTGGTCACGCTC-3'
AP2F	5'-CGGGATCCCGGGAGCAGATTCGTCTT-3'
AP2R	5'-CGAAGCTTGAGAAGCAGCATCAAAGGCCG-3'
AP21R	5'-CAAGACGAATCTGCTCCCGCCC-3'
H1F	5'-GGAATTCGGATCCAGAAAAATGGATTTGGTGAACCATC-3'
H1R	5'-GATCTAGACGTTCTATCTTATACACGACAATTAG3'

instruction (Stratagene, La Jolla, CA, USA). The restriction enzyme analysis and the determination of partial DNA sequence indicated that these inserts were able to be classified to two groups. These two isoforms were designated as *AtATP-PRT1* and *AtATP-PRT2*, respectively.

Isolation of Arabidopsis ATP-PRT1 *gene*

Approximately 5x 10^5 recombinant phages of *Arabidopsis* λZAPII genomic library (Stratagene) were screened using an insert DNA from pAtATP-PRT1 as a probe. The probe was made by the random priming method with [α-^{32}P]dCTP (Feinberg and Vogelstein, 1983). Hybridization of replica filters, and wash were carried out as described in Chapter I (Fujimori and Ohta, 1998a). The screenings were continued until pure phages were obtained.

Determination of DNA sequences

DNA sequences were determined from both strands by the dideoxy chain termination method (Sanger et al., 1977) by using an ABI PRISM™ Dye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster city, CA, USA). DNA and amino acid sequences were analyzed using DNASIS Ver. 3.4 software (Hitachi Software Engineering Co., Yokohama, Japan).

Isolation and analyses of nucleic acids

Total genomic DNA (10 μg) prepared from 2-weeks-old grown *Arabidopsis* seedlings as described (Sambrook et al., 1989) was digested with indicated restriction enzymes, separated electrophoretically on 0.7% (w/v) agarose gel, and blotted onto a nylon membrane (Hybond N$^+$; Amersham). Hybridization with indicated ^{32}P-labeled probes was for 18 hr either at 42 °C (high stringency) or at 37 °C (low stringency) in the solution containing 50% (v/v) (high stringency) or 40% (v/v) (low stringency) formamide, 6x SSC, 5x Denhardt's solution, 0.5% (w/v) SDS and 100 μg/ml denatured salmon sperm DNA (Sigma Chemical Co, St Louise, MO, USA). Blots were washed twice in 2x SSC/0.1% (w/v) SDS at room temperature for 15 min and twice either in 0.2x SSC/0.1% (w/v) SDS at 55 °C for 10 min (high stringency) or in 0.5x SSC/ 0.1% (w/v) SDS at 50 °C for 10 min (low stringency).

Total RNA was extracted as described (Lagrimini et al., 1987). Gene-specific probes for both cDNAs of *AtATP-PRT1* (+1 to +240) and *AtATP-PRT2* (+1 to +247) were prepared by PCR using SK primer and gene-specific primers, AP1R for *AtATP-PRT1* and AP21R for *AtATP-PRT2* (Table 2), incorporating [α-^{32}P]dCTP. Hybridization and wash were carried out described in Chapter I (Fujimori and Ohta, 1998a).

Construction of S. cerevisiae his1 *defective strain, and suppression by* Arabidopsis ATP-PRT *cDNAs*

S. cerevisiae strain disrupted *HIS1* gene was constructed as described in Chapter III (Fujimori and Ohta, 1998c). The coding region of *S. cerevisiae HIS1* gene was amplified by PCR using H1F and H1R as primers (Table 1) and *S. cerevisiae* strain

S288C genomic DNA as the template. PCR products digested with *Eco*RI were cloned into galactose-inducible expression vector, pYES2 (Invitrogen), yielding plasmid pKF110. The *Bam*HI-*Xho*I fragment of *S. cerevisiae LEU2* gene (Andreadis et al., 1982) was replaced with *Bgl*II-*Sal*I fragment of pKF110 to obtain plasmid pKF157. *S. cerevisiae* strain SH782 was transformed with 3-kb *Bam*HI-*Xho*I fragment from pKF157 by the method of Ito et al. (1983). The resultant *his1* disruption strain was designated BY1001.

To construct the expression vectors for the yeast, *S. cerevisiae*, the coding regions truncated the putative chloroplast transit peptide of *Arabidopsis AtATP-PRT1* or *AtATP-PRT2* cDNAs were amplified by PCR using the sets of gene-specific primer (Table 2). The resultant PCR products double-digested with *Bam*HI and *Hind*III were cloned into the corresponding sites of pYES2 to obtain the pKF251 and pKF252 shown in Table 1. The *S. cerevisiae his1* defective mutant BY1001 was transformed with each of pKF251 or pKF252 (Ito et al., 1983). After 4 days of growth at 30 °C on the synthetic medium containing 2% (w/v) glucose (Glu) and amino acids mixture without leucine (Leu) and uracil (Ura) (SC/Glu-Leu-Ura) plate, the selected colonies were streaked on the SC/Gal-His-Leu -Ura to test the ability to suppress the His auxotrophy of strain BY1001. Strain BY1001 transformed with pKF110 was used for a positive control.

Evolutional analysis

For phylogenetic analysis, the multiple sequence alignment of ATP-PRTs using CLUSTAL W program (Thompson et al., 1994) was performed. A matrix of the estimated number of amino acid differences between sequence pairs was calculated by using the program, PROTDIST of the PHYLIP 3.5c software (Felsenstein, 1993). To estimate the confidence limits of branch points in the phylogenetic trees, the neighbor-joining (N-J) analysis was performed with the programs, SEQBOOT, NEIGHBOR, and CONSENSUS from the PHYLIP 3.5c package.

Expression of recombinant Arabidopsis *ATP-PRT proteins*

The coding region truncated putative chloroplast transit peptide of AtATP-PRTs were amplified by PCR using gene-specific primer sets, EF3 and ERXH for

61

AtATP-PRT1 and AP2F and AP2R for AtATP-PRT2 shown in Table 2. The resultant PCR products double-digested with *Bam*HI and *Hind*III were cloned into pMAL-c2 vector (New England Biolabs, Beverly, MA, USA), resulting in pMA1 and pMA2, respectively. The each expression plasmid was introduced into *E. coli* strain XL1-Blue for production of recombinant protein fused with maltose binding protein (MBP). Soluble crude extracts prepared in column buffer containing 20 mM Tris-Cl (pH 7.5), 1 mM EDTA, and 200 mM NaCl were applied onto the amylose resin column chromatography (New England Biolabs.). The fusion proteins were eluted with column buffer containing 10 mM maltose. The ATP-PRT proteins were separated from MBP after digestion with Factor Xa (New England Biolabs), and then maltose was removed by the hydroxylapatite column chromatography. ATP-PRT proteins separated from MBP were obtained by the second amylose resin column chromatography. The protein concentration was determined by the method of Bradford (1976), and the proteins were analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) (Laemmli, 1970), followed by staining with coomassie brilliant blue (CBB).

Antibody production and immunoblot analysis

Purified recombinant protein of *Arabidopsis* AtATP-PRT1 was injected intradermally with complete adjuvant into two rabbits. Boosts were carried out in incomplete adjuvant.

Crude extracts prepared from *Arabidopsis* tissues were fractionated by SDS-PAGE, and electroblotted onto the polyvinylidene difluoride (PVDF) membrane (Immobilon P; Millipore Corp., Bedford, MA, USA). The membrane was incubated with 1:5000 dilution of *Arabidopsis* ATP-PRT1 polyclonal antibody for 1 hr. The immunoreactive proteins were detected using peroxisome-conjugated goat anti-rabbit IgG (BioRad, Hercules, CA, USA) and the ECL Western analysis system (Amersham).

Assay for ATP-PRT activity

Enzymatic activity for ATP-PRT was assayed by the method of Martin (1963). The reaction mixture contained 111.1 mM Tris-Cl (pH 8.5), 22.2 mM $MgCl_2$, 83.3 mM KCl, 5.6 mM ATP, 1 mM dithiothreitol (DTT), 0.56 mM phosphoribosyl pyrophosphate

(PRPP) and enzyme in a final volume of 360 μl. After preincubation at 30 °C for 5 min, the reaction was started by adding of PRPP, and then monitored at 30 °C for 2 min at 290 nm. One unit of activity is defined as a change in absorbance of 0.02 per min which corresponds to the transformation of 1.67 nmol of substrate per min and is equivalent to 10.4 units (Voll et al., 1967).

Results

Isolation of Arabidopsis *ATP-PRT cDNAs*

I have isolated the full-length cDNAs encoding ATP-PRT from *A. thaliana*. Firstly, cDNA fragments encoding ATP-PRT were amplified using AR5 and SK as primers (Table 2) and an *Arabidopsis* cDNA library constructed in λZAPII vector as the template. The resultant PCR products were directly cloned into the TA cloning vector, pCRII to obtain pAR5-13. Using the insert from pAR5-13 as a screening probe, an *Arabidopsis* cDNA library ($6x\ 10^5$ plaques) was screened, and 26 positive plaques were identified. From the partial DNA sequencing and the restriction enzyme analysis, these were classified to two groups, namely *AtATP-PRT1* (4 clones) and *AtATP-PRT2* (22 clones). The longest inserts from pAtATP-PRT1 or pAtATP-PRT2 had a length of 1485-bp or 1535-bp, respectively (Figs. 2a, b). Both inserts from pAtATP-PRT1 and pAtATP-PRT2 contained the long ORFs consisting of 411 and 413 amino acids, calculated molecular mass of 44.6-kDa and 44.8-kDa, respectively. In the 3'-noncoding region, one typical polyadenylation signal, AATAAA was found 24-bp upstream the polyadenylation tail in *AtATP-PRT1* cDNA (Wahle and Keller, 1992). On the other hand, no putative polyadenylation signal could not be found in *AtATP-PRT2* cDNA. Previous work was demonstrated that the polyadenylation signals in higher plants are not as highly conserved or precise as those found in vertebrate (Rothnie et al., 1994).

Characterization of Arabidopsis ATP-PRT *genes*

The *AtATP-PRT1* gene was isolated by the plaque hybridization. While the database search provided an *Arabidopsis* genomic clone (F21M12, GenBank Accession No. AC000132, Vysotskaia et al., unpublished) containing *AtATP-PRT2* gene. Sequence

a

```
-318  ccacacttgtaccattcgatccaaccgtccaattcattcttcaaatgttgcataaatcctacaaccacttccttttttctcaatttttaattctttctt
                                    ‾‾‾‾‾
                                    CCAAT

-218  tcacgtttttctcataaagattccactaaaataagaaaagccaccacaaaaaaaaaaaaaaaaaaagcatggatcaggcgttaagtgttaaaatggcatca

-118  agaatctaatagtcagaatctgcctctcccagagtgaaagtgtgttgtgtcttattcgttatataattccttctccggtctaaaagaaaccctttcactc
                                                          ‾‾‾‾‾‾‾‾
                                                          TATA

-18   tcgaaaagccacacaacaATGTCTCTCCTTCTCCCTACGAATTTACAACAATACCCTTCTTCTTCCTTCCTTCCCATCTTCAACACCTATCCTATCTCCGC
                     M  S  L  L  L  P  T  N  L  Q  Q  Y  P  S  S  S  S  F  P  S  S  T  P  I  L  S  P      27

 83   CTCCTTCCACCGCTTTCTCCGTCATCGTACCTCGTCGGAGATGTCTCAGATTGGTTACTTCTTGTGTCTCCACCGTTCAAAGCTCCGTCGCAACAAACGG
       P  P  S  T  A  F  S  V  I  V  P  R  R  R  C  L  R  L  V  T  S  C  V  S  T  V  Q  S  S  V  A  T  N  G  60

183   TTCCTCTCCAGCTCCTGCTCCGGCCGCTGTTGTCGTTGACCGTGACCAGATTCGTCTTGGTCTTCCTAGTAAAGGACGTATGGCTGCTGATGCAATCGAT
       S  S  P  A  P  A  P  A  A  V  V  V  E  R  D  Q  I  R  L  G  L  P  S  K  G  R  M  A  A  D  A  I  D   93

283   CTTCTCAAGgtctgagatttaacattgcgattgttaatctcagatttaattgttttagcttcactcagcttgtatttagttcttagagactgagaaatgtt
       L  L  K

383   tgtcattggagattttgatagagactgatttatgattttttatattagctctctctgtttctgtgttcgtcattctctttcactatataatttgattatttg

483   tgattttgaaaacttcaaatgggttctacagactgcttatttttactatactttgtttttgtgtgctttttagtgttaagttttggttaattagttggagca

583   agttgggatctttaactgatataaattggtttgatttgaatcctaggcatgtgaagtgttgttttgcgtttgatgatcgctttcttgaatagttgccaaag

683   cttcatttgaactctttactcactgacatgttttttttccttacttttgagtgaattgttatctcatggcttgatttgacttttctcattgtgtggtttc

783   ttttatgaagGACTGTCAACTGTTTGTTAAACAAGTCAATCCTAGGCAATATGTTGCACAGATTCCCCAGgtatcgtttccaagcgaattcaaaaccata
                D  C  Q  L  F  V  K  Q  V  N  P  R  Q  Y  V  A  Q  I  P  Q                                 116

883   gtgtgtttgtgggtttttgtgtaccatgacttatgttggtattttttgtatgtccttcagTTACCAAACACTGAAGTCTGGTTTCAACGGCCAAAAGATATT
                                                                   L  P  N  T  E  V  W  F  Q  R  P  K  D  I  130

983   GTCAGAAAGTTACTCTCAGGAGATTTGGATCTAGGTATCGTTGGTCTTGCACACACTTAGTGAATATGGTCAGgtaagtgatgatcataaaaatattctt
       V  R  K  L  L  S  G  D  L  D  L  G  I  V  G  L  D  T  L  S  E  Y  G  Q                              154

1083  tgaagagctttgatatcatgttcaggatgctgaattagcagtggtgattttttatcattatcattatgctggaaccagtttacttttattctctcatcat

1183  tagtatgccatattgttgcttagcgtatagtgtttaaaattaaagttttccctttgattgtctctgataagcccccttgcatcataatttagatgtttg

1283  aattttctttctagGAAAATGAAGATCTTATCATTGTCCATGAAGCTCTCAACTTTGGAGACTGTCACCTGTCTATTGCGgtatgctttaatctattgtg
                    E  N  E  D  L  I  I  V  H  E  A  L  N  F  G  D  C  H  L  S  I  A                        176

1383  gttaaggtcctcaggtcattaccgtggatggttttctgattgtatctcaaatggtcacagATTCCAAACTATCGGATATTTGAGAATATAAATTCTCTGA
                                                                    I  P  N  Y  G  I  F  E  N  I  N  S  L    189

1483  AGGAGCTAGCGCAAATGCCCCAATGGAGTGAAGAGAGACCCTTACGCTTAGCTACTGGCTTCACTTATgtatgatgagttacagattcttccttttggtgcg
       K  E  L  A  Q  M  P  Q  W  S  E  E  R  P  L  R  L  A  T  G  F  T  Y                                 212

1583  ctgttgtcatctaatttctctcagagtccatttcgcttcaacataatattttgggaatgattcaactggttttttgcagCTCCGGCCCCAAATTTATGAAAGA
                                                                                     L  G  P  K  F  M  K  E  220

1683  AAAATGGCATAAAGCATGTGGTGTTTTCAACTGCAGACGGACCCACTGGACGGCAGCTCCAGCGgtgagattttcaacgaaaagcttaacatttggcattaag
       N  G  I  K  H  V  V  F  S  T  A  D  G  A  L  E  A  A  P  A                                          240

1783  ttgcttgtgtatctttttctaaaaacaagtgttaaactgatatgctctgtggaggataatgttgtcagcttgtagtgtccgatttcatatcttactgtgcc

1883  ttaccagcataatctgttgactcatatgtgctgttaattagATGGGGATAGCTGATCCCATTTTGGATCTTGTGAGTAGTCGGTATAAACACTCAAAGAGA
                                               M  G  I  A  D  A  I  L  D  L  V  S  S  G  I  T  L  K  E     259

1983  ACAACTTGAAAGAAATTGAAGGAGGTGTTGTGGCTGGAAAAGCCAGgtgatgtgaagttctggatccatgaactgtctgtttagaatcgttggcggccaT
       N  N  L  K  E  I  E  G  G  V  V  L  E  S  Q                                                         274

2083  cattcctgaacattttttttttttctttgtaatttcagGCGGCACTTGTGCCAAGTAGAAGAGCATTAAACGAGAGAAAAGGGCCACTAAACACAGTACACG
                                             A  A  L  V  A  S  R  R  A  L  N  E  R  K  G  A  L  N  T  V  H  295

2183  AGATTCTTCAGAGATTCGGAGGCCCATCTAAAGCCGGATGGCCAATTCACTgtacgcacttggttccatgatgtgttttttccttattcaccccttttcgcttt
       E  I  L  E  R  L  E  A  H  L  K  A  D  G  F  T                                                      312

2283  cttttctctaataatttcattcagGTTGTTGCAAACATGAGAGGGAAATAGTGCTCACGAAGTGGCTGAGCGTGTGCTGAGCCAACCATCATTGTCACGA
                                V  V  A  N  M  R  G  N  S  A  Q  E  V  A  E  R  V  L  S  Q  P  S  L  S  G    337

2383  TTGCAGgtaatgtttcagaatgtgggagttcttaagccttagcttttttagttcttgctgattgtttatgggtctataacaagcagGGACCGACAATAAG
       L  Q                                                                               G  P  T  I  S    344

2483  CCCAGTGTACTGTACACAAAATGGAAAAGTATCGGTTGACTACTATGCCATCGTGATTTGTGTACCAAAAAAAGCCCCTATACGACTCTGTGAAGCAACTT
       P  V  Y  C  T  Q  N  G  K  V  S  V  D  Y  Y  A  I  V  I  C  V  P  K  K  A  L  Y  D  S  V  K  Q  L    377

2583  AGAGCGgttagtaatcccctgccatttctctcttagtaatcccctgccatttttttctcaattatcttactcccaaccgatctttaaaacttattgatcctc
       R  A                                                                                                379

2683  tttctactttttgcaccaaccacgttccacataaatctgtgtttatcgttccctctttggtttggtcatttttcataaatgttattgcagCCCGGAGGCAG
                                                                                                 A  G  G  S  393

2783  TGGGGTATTAGTTTCACCTTTGACCTACATTTTTGATGAGGATACTCCAAGATGGGGTCAGCTCCTGAGAAACCTCGGGATTTAAtgcttt ggtaagga
       G  V  L  V  S  P  L  T  Y  I  F  D  E  D  T  P  R  W  G  Q  L  L  R  N  L  G  I  *                  411

2883  agatgcatatgctgaacaagtttt aagaagactctttggtgtaaatctctgagttttttt tggtaccttgcgtgtctctttttttttttttttctttcccttt

2983  tctgctgcagtcgtcgtcctccgtgcctgttgtaattttttacatctcttttctgtatggatcctattcgttgttgttaatccccgttggaacaataaaaaca
                                                                                                   ‾‾‾‾‾‾‾‾‾

3083  agccttttttttaataatacatctttgcagaactcatttgttaataagttatcacttgttgttgttctgtgtttttgctaacctggcgggggtttggttccaga
             ↑

3183  aaccaattccttttgctcacacgggctcggtgttgcgagttacctgtctctaga
                                               ‾‾‾‾‾‾‾‾
                                               XbaI
```

Fig. 2. Nucleotide and amino acid sequences of the *Arabidopsis AtATP-PRT* genes and their proteins.
a. AtATP-PRT1 gene, *b. AtATP-PRT2* gene. The deduced amino acid sequence is presented in single-letter. A putative polyadenylation signal is shown *double-underlined*. Vertical arrow indicates the polyadenylation site. Asterisk indicates the stop codon. Introns are shown in small letters. Putative TATA box and CCAAT box are shown *underlined* and indicated. Numbers shown left and right to the columns refer nucleotide and amino acid sequences, respectively. Nucleotide number refers to the A (+1) of the first ATG in the open reading frame.
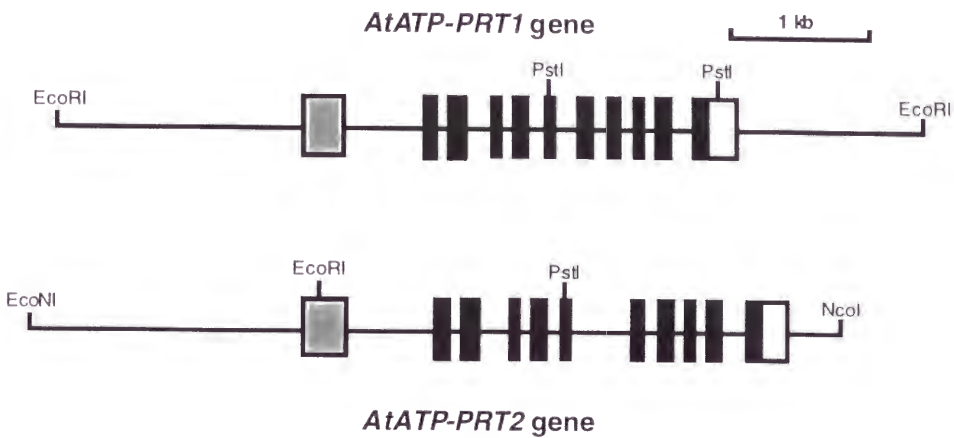
b

```
 -318   gtcaacacattgtggatctgtagctgctctcttgaagcaaatcaagaagtaacaacaataagcaaaggactaatccctctaaacctgctagaggcatcag

 -218   aagagactacaaagcaaatttctctttctgtctgtataaatattaaatagcaaataaaaaacatctttgagtgtaataaataattccggctattattttt

 -118   taattaaatttgaaatttgtgagtatccacgtgtacactgggatttTATAAcggcctagctttgtgcaattATATAtccgccgaccatctctctcctccgt
                                                      TATA                      TATA

  -18   caccgaggGagagcaaccATGCCAATCTCAATTCCCCTTAACGCCACTCTACAATACTCGTCTCCTTCTTCTTCTTCTTCTTCTTCTCTCGTACCTT
                       M  P  I  S  I  P  L  N  A  T  L  Q  Y  S  S  P  S  S  S  S  S  S  S  L  V  P    27
   83   CTTCCCCTCTTTTCTCTCCGATTCCTTCAACCACTGTTTCCCTTACCGGAATTCGCCAGCGATGCCTCAGGATGGTCACTTCTTGCGTCTCCAACGCTCA
        S  S  P  L  F  S  P  I  P  S  T  T  V  S  L  T  G  I  R  Q  R  C  L  R  M  V  T  S  C  V  S  N  A  Q    60
  183   GAAATCCGTCCTGAATGGTGCCACCGATTCCGTCTCTGTCGTCGGGCGGGAGCAGATTCGTCTTGGTCTTCCTAGCAAAGGACGCATGGCCGCCGATTCG
        K  S  V  L  N  G  A  T  D  S  V  S  V  V  V  G  R  E  Q  I  R  L  G  L  P  S  K  G  R  M  A  A  D  S    93
  283   CTGGATCTTCTCAAGgtgagtggtttttgggggagtggttgtgtttcccttacaaaattgaggattgtgctgacaaatttccatggagatttggagtttttt
        L  D  L  L  K                                                                                           98
  383   tgtatctgagttgccgaaacttgaatttgttctgggaattgctcctcacaatgctcgccattttttggcactggagtctcctttaattctcagtttatcac

  483   tggagataætcgtagattactgggatataaaatatcactggagaagatatagtaggggagctatacttactcaaggactagtctttctatatttgagcta

  583   ttgactctgtttttgaagcctaaatattgattgctgagcttgcacatcgctcccttggggaagttaggtatttgagtgttctactgttgtggttttgtctc

  683   tgtttgatggtctcttgcttatagatcacattctccttgtttgctttcttgggttttttgcaaggtttcattttttgatcctttgatttcttaaatgattt

  783   ttcactctcagtatctcttagcaacgcatgattttcaatctcatagatccttacaactcttttacttttcccctttctcatgaagGATTGCCAATTGTTTG
                                                                                             D  C  Q  L  F   103
  883   TCAAACAAGTCAATCCACGACAATATGTTGCTCAAATTCCTCAGgtatgcatttagagagaagggcaaacgcgtattgacttactaatctatcattgtttg
        V  K  Q  V  N  P  R  Q  Y  V  A  Q  I  P  Q                                                             118
  983   tttggtgtaacgcaatttgtaatggtaatctcaatcttcttgtcctccagTTACCAAATACTGAAGTTTGGTTTCAACGACCAAAAGACATTGTCAGAAA
                                                          L  P  N  T  E  V  W  F  Q  R  P  K  D  I  V  R  K   135
 1083   GTTGCTATCAGGAGATCTGGATCTAGGCATCGTTGGTCTTGACATAGTTGGTGAATTTGGTCAGgtaaatgacatcataatcatatatcttgattagcag
        L  L  S  G  D  L  D  L  G  I  V  G  L  D  I  V  G  E  F  G  Q                                            156
 1183   acacctgttgagattctagggtatcttagtgctggtgcctggtgcttaatgattatactgattttcatggtagagacaaagaaacgagtctaatcttgtt

 1283   atttgttcttctgcggttttccttgacattgctaaaccatactgtttgcgttagcttatatgacttagacttgtcttccgaaggccatttggcatcaaaa

 1383   tttagagattttcgagttctattttcagGGAAATGAAGATCTTATCATTGTCCACGAAGCTCTCAACTTTGGAGATTGTCATCTCTCCCTTGCCgtacgtt
                                    G  N  E  D  L  I  I  V  H  E  A  L  N  F  G  D  C  H  L  S  L  A            178
 1483   caaatctattgtagttgtggcttatcatcaatacagatgacttcttgattgctctttcatatgattttcagATACCCAATTATGGAATATTTGAGAATAT
                                                                                I  P  N  Y  G  I  F  E  N  I   188
 1583   AAAATCTCTGAAGGAGCTAGCACAAATGCCTCAATGGACCGAGGAAAGACCTTTACGAGTTGCTACAGGATTCACTTATgtatgatgagttacaagtgtT
        K  S  L  K  E  L  A  Q  M  P  Q  W  T  E  E  R  P  L  R  V  A  T  G  F  T  Y                            214
 1683   cctttggctcattattgtctttaatctttttccggaatgcaatttaattgtttgaagtttatccaacttgtgtttgcagCTAGGCCCCAAATTTATGA
                                                                                       L  G  P  K  F  M       220
 1783   AAGACAACGGTATAAAGCATGTGACTTTCTCAACTGCAGATGGAGCCCTAGAGGCAGCGCCTGCGgtgagatttacaatgtttcaaatcagccctagat
        K  D  N  G  I  K  H  V  T  F  S  T  A  D  G  A  L  E  A  A  P  A                                         242
 1883   ttcgaactatgcacccttattagactaccaagtttccattaggattaaaaaagcttggaatttgtatgaaatgttgagtacatttagatattttttacaag

 1983   agagttatcggattcttctcttatctttcagattgtattacttgctgatatgagattaacgttccaaattgtgaaagtattttgctattatgtcccatcag

 2083   cgtgagatgatggattatagttatagagaaagaaactactgttgtgtgcagtgtttgttgacacatatgctcttgttaattagATGGGGATAGCTGATGC
                                                                                              M  G  I  A  D  A   248
 2183   CATTTTAGACCTTGTGAGCAGTGGGACAACCCTTAAAGAGAACAACTTAAAAGAAATTGAAGGAGGGGTTGTGCTGGAAAGTCAGgtatgacatgaagat
        I  L  D  L  V  S  S  G  T  T  L  K  E  N  N  L  K  E  I  E  G  G  V  V  L  E  S  Q                      276
 2283   ctttggaattgattattttttcgcacttggcaaccggtgtaccaatttttttttcttttacgtactatcagGCTGCACTAGTGGCGAGTAGAAGGGCAT
                                                                                A  A  L  V  A  S  R  R  A      285
 2383   TGACTGAGAGAAAAGGAGCACTAGAAACAGTGCATGAGATTCTCGAGAGATTAGAGGCTCATTTGAAGGCAAATGGTCAATTCACTgtacgtgcatatta
        L  T  E  R  K  G  A  L  E  T  V  H  E  I  L  E  R  L  E  A  H  L  K  A  N  G  Q  F  T                    314
 2483   attgtgatgtttcaccttatttctacttttgcttctcttcatccgatgtttcactccagGTGGTTGCAAACATGAGAGGAACGGATGCTGAAGAAGTGG
                                                                   V  V  A  N  M  R  G  T  D  A  E  E  V       327
 2583   CTGAACGTGTGAAAACCCAACCATCACTATCAGGATTGCAGgtaaagtttgataacatttggatatcttcgatgtagttccactgagtttgtaggtgat
        A  E  R  V  K  T  Q  P  S  L  S  G  L  Q                                                                341
 2683   cgttcagacgattgtgcttattttgtgatgtgcagGGACCAACAATAAGTCCAGTTTATTGTAAACGAGATGGAAAAGTAACAATTGAGTACTATGCCAT
                                           G  P  T  I  S  P  V  Y  C  K  R  D  G  K  V  T  I  E  Y  Y  A  I   363
 2783   TGTGATATGTGTACCAAAAAAAGCGCTTTATGAGTCTGTGCAACTGAGAGCCGgttagtaatgatatccacctttgcttccacctccccctttttactc
        V  I  C  V  P  K  K  A  L  Y  E  S  V  Q  Q  L  R  A                                                    381
 2883   taagattgtcattatttaaggggatttacaaatacatgcaaaccacattgtttttggtactgagatcagtctttgtctcaaatctcggatgtgttctta

 2993   agtaatgtgtcacatgtgattgcagGTGGGGAGGAAGTGGGGTACTAGTTTCTCCTGTAACATACATTTTTCATGAGGAGACTCCCGATGGAGTCAGCTT
                                 V  G  G  S  G  V  L  V  S  P  V  T  Y  I  F  H  E  E  T  P  R  W  S  Q  L     406
 3093   CTGAGTAACCTCGGCCTTTGATgctgcttctcagacatactttggtttagtcagatgcattcgttgtggaggæctcttttggtcaaacactacaatgtctt
        L  S  N  L  G  L  *                                                                                     412
 3193   tctttttTgggttttatctttgtggatcttatagtagaaaaaacagaggttgtctctctttttaagccttttaagtccatctatgccatttcagtttttcagg

 3293   agtcaacæcttttaaaaaccaattttatcaaagcattcgtagaaggaggtcaaaacattccaaatgaaatætcgaaattgatccgcccgtacatcacta
                                                        ↑
 3393   tttccagttcattggattcagtaactcaggagtgatggtgtaacacataagaactcagtatctctatctattcagctaatcagcttctgaacaaaaacca

 3493   atattggtggactcctcgcccattgttgcggaaatgaagtaaattcacgtattcatggaaactcatatctgæcacgggcaagatacagaccagatgcagga

 3593   gccatgg
        NcoI
```

Fig. 3. Structure for *Arabidopsis AtATP-PRT* genes.
Shaded, solid and open boxes indicate putative chloroplast transit peptide, open reading frame (mature region), and 3'-noncoding region, respectively.
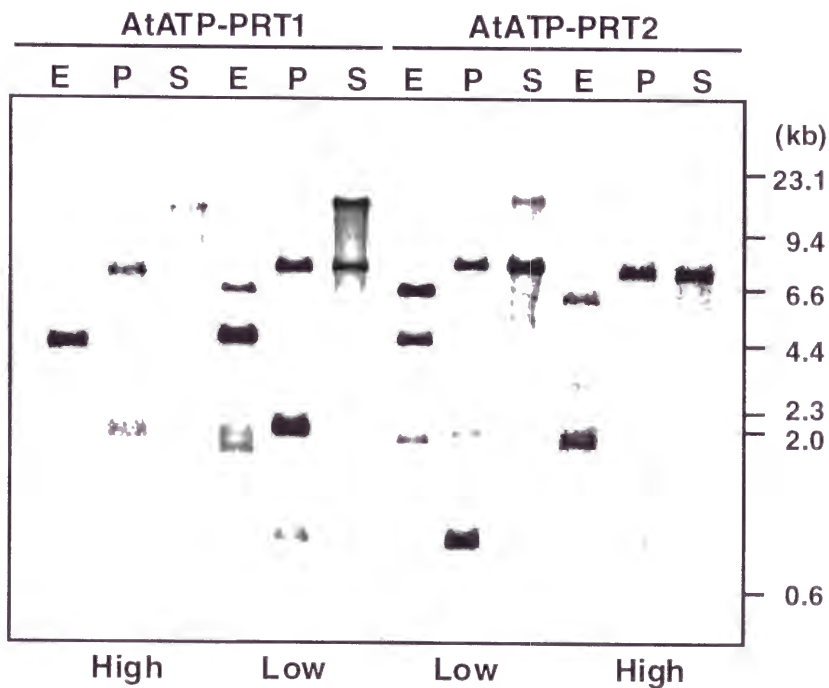

Fig. 4. Genomic Southern blot hybridization analysis.
Each of both *A. thaliana AtATP-PRT1* and *AtATP-PRT2* was hybridized to restricted *A. thaliana* genomic DNA, under low- or high-stringency conditions (see *Materials and methods*). The restriction enzymes used are indicated at the top of the figure, E; *Eco*RI, P; *Pst*I and S; *Sal*I. The molecular size marker is shown at the right side. High and Low indicate high and low stringency conditions, respectively.

66

analysis has demonstrated that both genes consist of 11 exons and 10 introns, and all of the introns follows the "GU-AG rule" for the exon-intron junctions (Breathnach and Chambon, 1981) (Figs. 2a, b). Although the nucleotide sequences of 5'- and 3'-untranslated regions and introns were divergent, the intron positions of both genes were completely conserved with different intron lengths (Figs. 2a, b, and 3). Southern hybridization analyses were performed using each of full-length cDNAs as probes under both low- or high-stringency conditions (Fig. 4). *Arabidopsis* genomic DNA digested with each of *Sal*I (no restriction site in both cDNAs), *Eco*RI (one restriction site in *AtATP-PRT2* cDNA) or *Pst*I (two restriction sites in *AtATP-PRT1* cDNA and one restriction site in *AtATP-PRT2* cDNA), and subjected to Southern blot hybridization analysis. Under high-stringency condition, simple patterns of bands using each of both probes were detected which accounted for most of the bands identified under low-stringency conditions. In addition, the results from low-stringency conditions were suspected that no more isoforms exist in *Arabidopsis* genome.

Functional expression of Arabidopsis ATP-PRT *cDNAs in* S. cerevisiae his1 *defective mutant*

To gain the evidence that two *Arabidopsis* isoforms exactly have the ATP-PRT activity, I employed the functional expression study using the *S. cerevisiae his1* defective mutant. At first, I constructed the *S. cerevisiae* strain defective in the activity for ATP-PRT encoded by *HIS1* gene (Fig. 5a). *S. cerevisiae HIS1* gene was replaced with *S. cerevisiae LEU2* gene. The resultant strain, BY1001 was unable to grow on the SC/Glu-His plate (Fig. 5b). Fig. 5b showed that the plasmids, pKF251 (harboring *AtATP-PRT1* coding region truncated the putative chloroplast transit sequence) and pKF252 (containing *AtATP-PRT2* coding region truncated the putative chloroplast transit sequence) were able to suppress the His auxotrophy in a *S. cerevisiae his1* defective mutant BY1001 same as pKF110 (carrying *S. cerevisiae HIS1* coding region) as a positive control. It revealed that both *Arabidopsis* cDNAs encode the active ATP-PRT enzymes. In addition, amino-terminal extension of both proteins was not required for the enzyme activities.
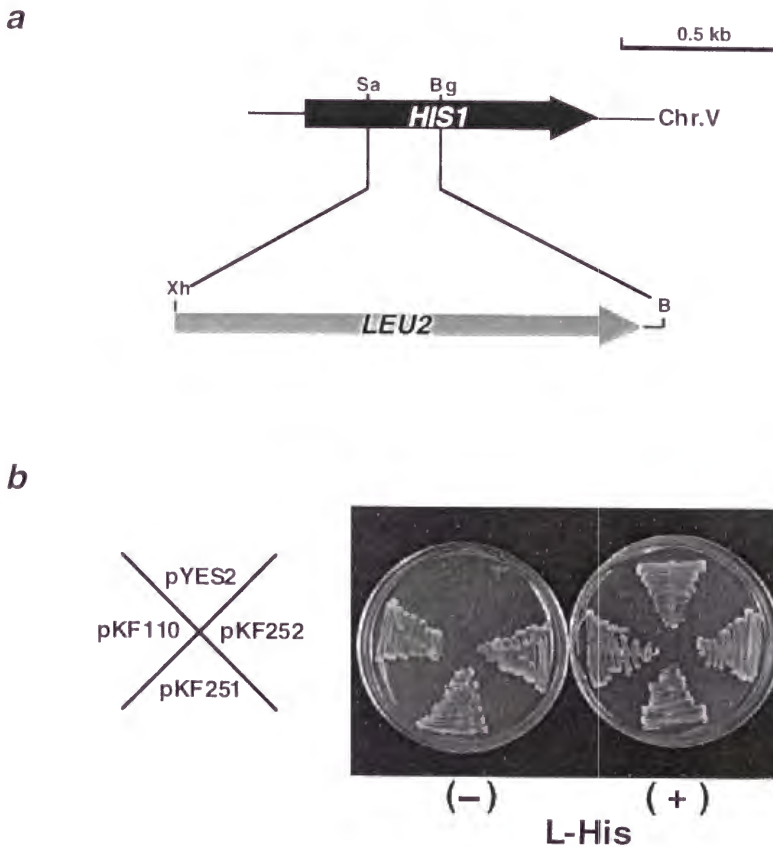
Fig. 5. Complementation of a *S. cerevisiae his1* null mutant, BY1001 with the *A. thaliana AtATP-PRT* cDNAs.

a. Construction of *his1::LEU2* null allele. Restriction enzyme sites are designated B; *Bam*HI, Bg; *Bgl*II, Sa; *Sal*I, X; *Xho*I. Chr.V means chromosome 5.   *b.* Growth of *S. cerevisiae his1* defective mutant, BY1001. Strain BY1001 transformed with each of expression plasmids, pYES2 (empty plasmid), pKF110 (carrying *S. cerevisiae HIS1* coding region), pKF251 (harboring *AtATP-PRT1* coding region truncated the putative chloroplast transit sequence) or pKF252 (containing *AtATP-PRT2* coding region truncated the putative chloroplast transit sequence) were cultivated on minimal-galactose plate supplemented with amino acids mixture without L-His, Leu and Ura (SC/Gal-His-Leu-Ura) at 30 °C for 3 days.

Comparison of amino acid sequences and phylogenetic analysis

The primary structures of putative mature AtATP-PRT1 and AtATP-PRT2 compared with those of *S. cerevisiae* HIS1 (Hinnebusch and Fink, 1983), *C. albicans* HIS1 (Pla et al., 1995), *E. coli* hisG (Carlomagno et al., 1988), *Synechocystis* sp. PCC6803 hisG (Kaneko et al., 1996) and *M. jannaschii* hisG (Bult

68

A.t ATP-PRT1   1   MSLLLPTNLQQYPSSSSFPSSTPILSPPPSTAFSVIVPRRRCLRLVTSCVSTVQSSVATNGSSPAPAPAAVV
A.t ATP-PRT2   1   MPISIPLNATLQYSSPSSSSSSSSLVPSSPLFSPIPSTTVSLTGIRQRCLRMVTSCVSNAQKSVLNGATDSVS-----V
T.g ATP-PRT   1   MSIANTILQNSPSCLVPSFSLLSPTPSAAVSSKGIRPRCLRPVTSCVSYARTSVQNGATDAVP-----V
S.c HIS1   1   MDL
C.a HIS1   1   MDL
E.c hisG
M.j hisG
S.y hisG

A.t ATP-PRT1   74   ERDQIRLGL--ESKGFMAAEAIDLLKDEQL--FVKQVNPFQYVAQIPQLENTDEWFQEPKDIVRKLLSEDLDLGIVGLDT
A.t ATP-PRT2   76   GREQIRLGL--ESKGFMAAESLDLLKDEQL--FVKQVNPFQYVAQIPQLENTDEWFQEPKDIVRKLLSEDLDLGIVGLDI
T.g ATP-PRT   66   GRDQIRLGL--ESKGFMATETLDLLKDEQL--FVKQVNPFQYVAQIPQLENTDEWFQEPKDIVRKLLSEDLDLGIVGLDI
S.c HIS1   5   NHLTDRILFAIEPKKGRLYSKSVSILNGADITFHR---SQFLDIALSTSLFEVAIAI-FLPAADIPTFVGEGKCDLGITGMDQ
C.a HIS1   5   NHLPDRILFAVEPKKGRLYEKCCNELSGADIQFRR---SNFLDIALSTNLFIALI-BLEAADIPVFVGEGNCDLGITGLDQ
E.c hisG   1   MTDNTRLRIAMQKSGRLSDESRELLARGGIKINL---HTQRLIAMAENMEIGVL-RVRDDDIPGLVMDCVVDLGIIGENV
M.j hisG   1   MIMFAIPLNKGRISEPVMIVLEKAGLKITV---KGPSLFANTVDDNIKVM-EARARDIPEFVADGVADLIGATGYDL
S.y hisG   1   MLTIAIPKKGSLLRCSIERFQQIGLDFSAFLEPKNRQLQILDPTHRAKALLVRAWGVEVYLEYEQAQLGIVGYEI

A.t ATP-PRT1   150 LSE-----YGQENEDLIIVHEALNFGDGHLSIAIPNYGIFENINSLKELAQMPQWSEERPLRVATGTYLGPKFMKENGI
A.t ATP-PRT2   152 VGE-----FGQGNEDLIIVHEALNFGDGHLSIAIPNYGIFENIKSLKELAQMPQWTEERPLRVATGTYLGPKFMKDNGI
T.g ATP-PRT   142 ISE-----FGAGNEDIIIVHEALNFGDGYLSIAIPNYGIFQNINSLKELAQMPQWTEKRPLRVATGTYLGPKFMKENEL
S.c HIS1   81 V------RESNV--RVDLAILQCFGNGKLQVQVFVNCEYKKPEQLI-------------GKTIVTSVKLAEBYFADLEG
C.a HIS1   81 I------KEAFQFDNIEDLLQLMKFGSGKLQIQVFADLEYEKPEQLV-------------GKKIVSSLTKLSTDYFKQLSD
E.c hisG   77 LEEELLNRRAAGEDPRYFTLRRLDFGGGRLSLATPVDEAWDGPLSLN-------------GKLATSYPHLLKRYLDQKE-
M.j hisG   72 V------LERNVEDKVDFLLD-PGFGFAKLVLAAPESSNINSLDDIKE-------------EMRVATEBPNLTKNYFEKLN-
S.y hisG   75 LRE--------KKPAVAQLAD-LNFGCGRLSVAVEQASPYRRSVDLPP-------------HGRVASKEVHCALEHFRRLD-

A.t ATP-PRT1   225 KHVVFSTAD-----GALEAPAMGIADAILDLVSSGITLKENNLKEIEGGVVLESQAAIVASRRAINERKEALNTVHEII
A.t ATP-PRT2   227 KHVTFSTAD-----GALEAASAMGIADAILDLVSSGITLKENNLKEIEGGVVLESQAAIVASRRAITERKEALETVHEII
T.g ATP-PRT   217 KHVTFSTAD-----GALEAAPAMGIADAILDLVSSGITLKENNLKEIEGGIVLQSQCAIVASRRAITERTEALETVHEII
S.c HIS1   140 TTVEKMTTRIKFVSGSVEASCALGIGDAIVDLVESGETMBRAAGLVDIA--TVILSTSAYLIESKNP-----KSDKSLIATIK
C.a HIS1   142 -----KPTNIRYVGGSVEASCAIGVADAIVDLVESGETMKAAGLKAIE--TILETSAILISSKKS-----KFPEMVNIIV
E.c hisG   144 -----ISFKSCLLNGSVEVAPRACLADAICDLVSTGATLEANGLFEVE--MIYRSKACLIQRDGE--MEESKQQLIDKLI
M.j hisG   133 -----KKVEIIELSGADEIEPFIGIADLISDLSTGCTLRINELKMID--EIVSSTTRLIANKNSLK-DKEKREKINQIV
S.y hisG   134 -----LPVEIIPLYGSVELGPITEMSEAIVDLVSTERTLQENGLMEIE--VLFDSSARLIAHPLSYR-VNGDRLLDWVEK

A.t ATP-PRT1   300 ERLFAHIKADGQFTVVANMRGNSAQEVAERVLSQPSLSGLQGPTISPVVYCTQNGKVSVDLYAIVICVPKKALYDSVKQLR
A.t ATP-PRT2   302 ERLFAHIKPNGQFTVVPYMRGTDABEVAERVKTQPSLSGLQGPTISPVVYCKRDGKVTIEYYAIVICVPKKALYESLQQLR
T.g ATP-PRT   292 ERLFAHIKADGQFTVRANMRGTDABEVAARVMAQPSLSGLQGPTMSPIYSLRDGKVSTEYYAIVICVPKKALYESLQQLR
S.c HIS1   214 SFIEGVMTAQRFVSCIYNAPEDKLPELL------KVTPGRRAPTISKI-------DDEGWVAVSSMIERKTKGVVLDEIL
C.a HIS1   210 QRLQGVLAAQEYVLCNYNAPKSIQAKCL------TITPGRRAAIVSTL--DKHSDDEEDWVAISSMVNRKEIGNVMDEUK
E.c hisG   215 TPIQGVIQARESKYIMMHAPTERLDEVI------ALLPGAERPTILPL------AGDQQRAAMHMVSSETLFWETMEKUK
M.j hisG   205 IAIKSVLFAETKRLIMMNARKDKVEEIR------KLIPGMAGPIVSKM------LSDDNMVATHAVNEDEIFTLMPKUH
S.y hisG   206 LESVI*

A.t ATP-PRT1   380 AAFCSGVLVSELTYHFDEDTPRWGQLLRNLGI*
A.t ATP-PRT2   382 MVPGSGVLVTEVTYHFHEETPRWSQLLSNLGL*
T.g ATP-PRT   372 AIICCSGVLVSELTYHFDEETPRWSQLLSNLGL*
S.c HIS1   281 RLGASDIMVFEISNCRV*
C.a HIS1   282 KAGAITDILVLEISNCRV*
E.c hisG   283 ALGASSILVLPIEKMME*
M.j hisG   273 ALGARDILNVPIERUL*

Fig. 6.   Alignment of ATP-PRT protein sequences.
The amino acid sequences of *Arabidopsis* AtATP-PRT1 (At ATP-PRT1) and AtATP-PRT2 (At ATP-PRT2) deduced from the corresponding cDNA sequences are aligned with those of *T. goesingense* ATP-PRT; T.g ATP-PRT (Yan et al., GenBank Accession No. AF003347), *S. cerevisiae* HIS1; S.c HIS1 (Hinnebusch and Fink, 1983, Acc. No. J01329), *Candida albicans* HIS1; C.a HIS1 (Pla et al., 1995, Acc. No. X83871), *E.coli* hisG; E.c hisG (Carlomagno et al., Acc. No. V00284), *M. jannaschii* hisG; M.j hisG (Bult et al., 1996, Acc. No. U67562), and *Synechocystis* sp. PCC6803; S.y hisG (Kaneko et al., 1996, Acc. No. D64006). *Dashes* indicate gaps introduced to allow for optimal alignment of the sequences. Conserved amino acid residues show *shaded*. Single-letter codes for amino acid residues are used and *asterisks* indicate termination codons for translation.

69

et al., 1996). Recently, cDNA encoding ATP-PRT from *Thlaspi goesingense* was isolated (Yan et al., unpublished). Amino acid identities among higher plants showed extremely high. Interestingly, the identity between *Arabidopsis* AtATP-PRT2 and *T. goesingense* ATP-PRT (81.6 %) was higher than that between *Arabidopsis* AtATP-PRT1 and AtATP-PRT2 (74.6 %). On the other hands, the identities between higher plants and microorganisms were not so high, although several conserved regions were observed in all organisms (Fig. 6). Additionally, the amino-terminal approximately
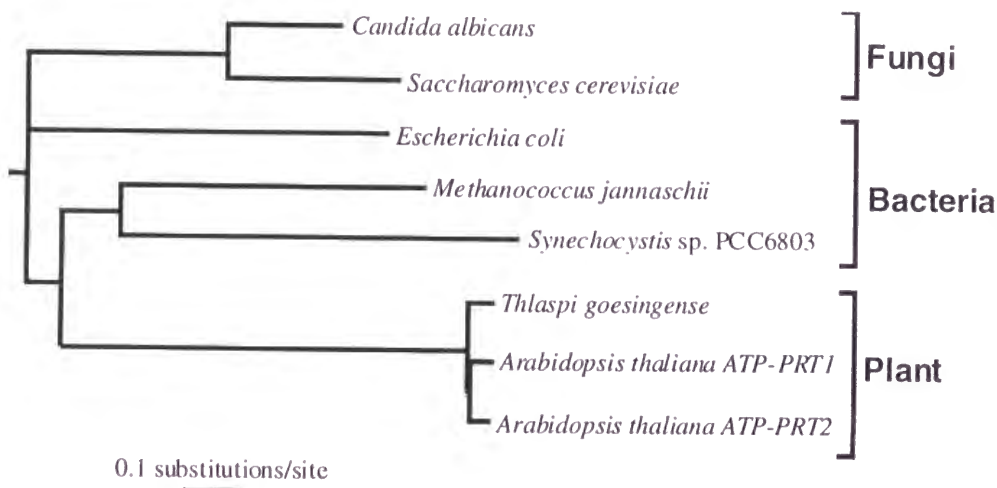
Fig. 7.   N-J phylogenetic analysis of ATP-PRT proteins.
The species are indicated the right and refer to GenBank/EMBL/DDBJ accession numbers as indicated in Fig. 6.

80 amino acid residues of the ORFs deduced from both *Arabidopsis* cDNAs showed features for signal sequences required for transport of proteins encoded by nuclear genes into chloroplasts (von Heijne and Nishikawa, 1991), which show no homology to amino acid sequences from microbial ATP-PRTs (Fig. 6). In order to determine the evolutional relationship between ATP-PRTs of higher plants and those of microorganisms, I performed the comparative phylogenetic analysis. Fig. 7 showed that plant ATP-PRTs are more closely related to bacterial ATP-PRTs than fungi ATP-PRTs. While fungi ATP-PRTs seemed to have branched off very early in evolution and thus forms a separate branch, the functional complementation of *S. cerevisiae his1*

70

mutant points to a very high conservation of functional motifs in plant ATP-PRT protein sequences (Fig. 7).

Expression of Arabidopsis ATP-PRT *genes*

Total RNA prepared from different tissues of *Arabidopsis* was subjected to Northern blot hybridization analysis with gene-specific probes prepared by PCR incorporating $[\alpha\text{-}^{32}P]$dCTP. Each probe detected transcripts of approximately 1.6-kb in length, respectively (Fig. 8). The hybridization data suspected that both cDNA clones
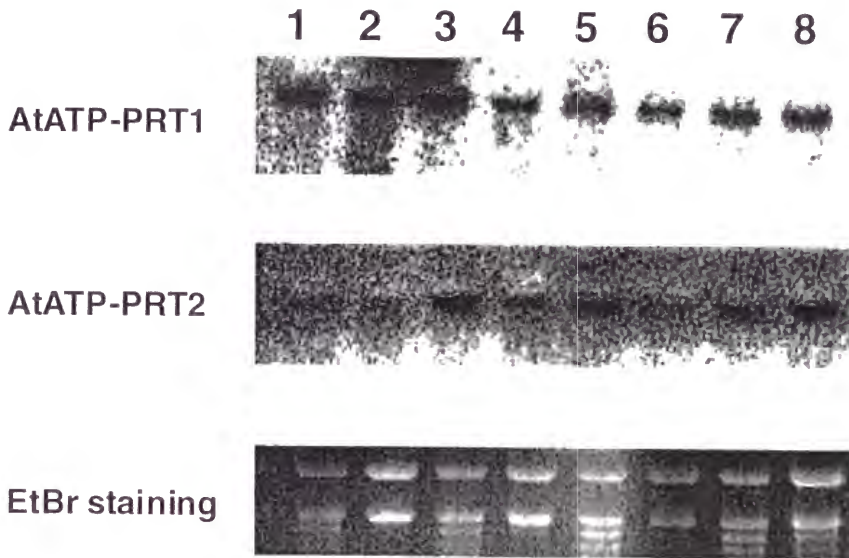
Fig. 8.   Expression profile of *Arabidopsis AtATP-PRT* genes.
Total RNA (10 µg) extracted from 1-week-old germinating seeds (lane 1), roots and leaves from 2-week-old plants (lanes 2 and 3, respectively), roots and leaves (lanes 4 and 5, respectively) and roots, leaves and siliques from 4-week-old plants (lanes 6-8, respectively) was subjected to the Northern blot hybridization analysis. The blot was probed with *Arabidopsis AtATP-PRT1* (top) or *AtATP-PRT2* cDNAs (middle). The bottom photograph shows the ethidium bromide (EtBr) staining of the ribosomal RNA bands for each RNA preparation.

were nearly full-length. The *AtATP-PRT1* mRNA was expressed in a constitutive manner, whereas the *AtATP-PRT2* mRNA was also detected in all tissues investigated here, but clearly lower in roots (Fig. 8).

Analysis of Arabidopsis *ATP-PRT protein*

Two *Arabidopsis* AtATP-PRTs were recombinantly expressed in *E. coli* and purified. AtATP-PRT1 protein was produced and able to be purified (Fig. 9a). In contrast, AtATP-PRT2 was unstable after purification although the protein was recombinantly expressed in *E. coli* (data not shown). Recombinantly produced AtATP-PRT1 protein fused with maltose binding protein (MBP) was purified by the amylose resin column chromatography. Purified *Arabidopsis* AtATP-PRT1 protein was obtained by separation from MBP after digestion with specific endoprotease, Factor Xa, followed by additional two chromatographic steps (Fig. 9a). Polyclonal antibodies against the purified recombinant AtATP-PRT1 protein were made to investigate the intracellular localization of ATP-PRT proteins. Polyclonal antibodies against AtATP-PRT1 protein could be also detected AtATP-PRT2 protein (data not shown). Anti-AtATP-PRT1 polyclonal antibodies strongly reacted with approximately 42-kDa protein prepared from leaves (Fig. 9b,
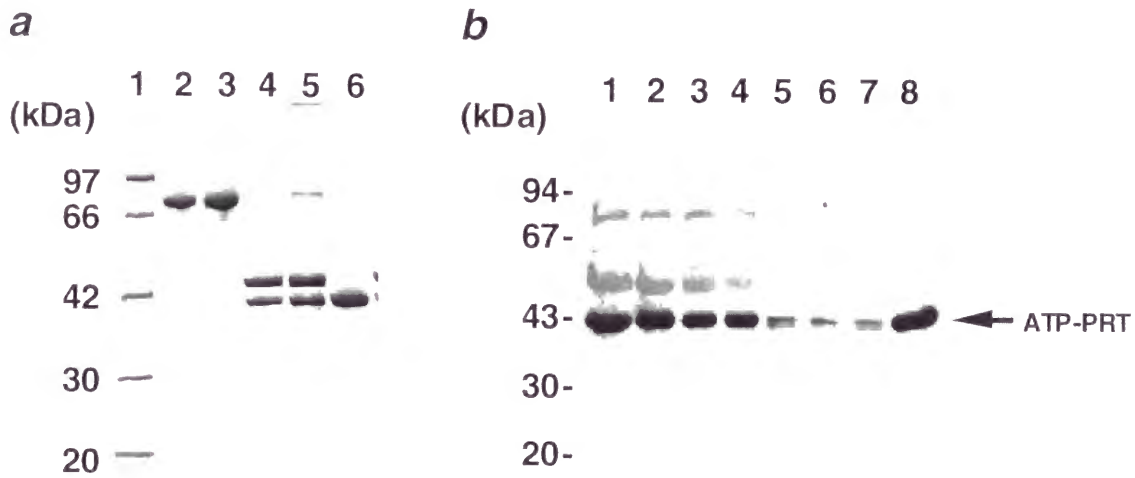
Fig. 9. Analysis of *Arabidopsis* ATP-PRT proteins.
a. Purification of recombinant AtATP-PRT1 protein. Protein size marker (lane 1), IPTG-induced *E. coli* lysate (lane 2), elution from first amylose resin chromatography (lane 3), after digestion with Factor Xa (lane 4), elution from hydroxylapatite chromatography (lane 5), and elution from second amylose resin chromatography (lane 6). *b*. Crude plant extracts (10 μg) prepared from leaves and roots were separated by SDS-PAGE, and were electrophoretically transferred to PVDF membrane for immunodetection with anti-*Arabidopsis* ATP-PRT1 polyclonal antibody. One week-old leaves (lane 1), 2-week-old leaves (lane 2), 3-week-old leaves (lane 3), 4-week-old leaves (lane 4), 2-week-old roots (lane 5), 3-week-old roots (lane 6), 4-week-old roots (lane 7) and purified *Arabidopsis* ATP-PRT1. Molecular size (in kDa) of protein standards is indicated on the left. Each sample was subjected to 10-20 % SDS-PAGE.

lanes 1-4). While, in crude extract prepared from roots, weaker bands were detected that are almost same size as those from leaves (Fig. 9b, lanes 5-7).

Kinetics for ATP-PRT proteins

In order to investigate the kinetic properties of *Arabidopsis* AtATP-PRT isozymes, recombinant *Arabidopsis* ATP-PRT proteins were produced in *E. coli*, and purified. Recombinant AtATP-PRT2 protein was not purified as described above. So, I performed the investigation of *Arabidopsis* AtATP-PRT isozymes using crude extracts from *E. coli* expressing each AtATP-PRT isozymes. No ATP-PRT activity was detected in crude

Table 3.　Properties of ATP-PRTs from higher plants and microorganism

Apparent K_m values

| | crude extract | | purified | | |
	AtATP-PRT1	AtATP-PRT2	AtATP-PRT1	*T. aestivum*[*]	*S. typhimurium*[**]
PRPP (mM)	0.37	0.57	0.13	0.13	0.067, 0.056
ATP (mM)	0.89	0.51	0.60	0.78	0.20, 0.43

IC_{50} values

| | crude extract | | purified | | |
	AtATP-PRT1	AtATP-PRT2	AtATP-PRT1	*T. aestivum*[*]	*S. typhimurium*[**]
L-His (μM)	40	320	45	75	60-80, 70-80
1,2,4-TA(mM)[***]	0.47	2.4	0.65	N.D.[****]	N.D.[****]

[*]Münzer et al. (1992), [**]Martin (1963), Whitfield (1971), [***] 1,2,4-TA; 1,2,4-triazolealanine
[****]N.D.; not determined
Each crude extract was prepared from *E. coli* expressing recombinant protein, and "purified" means the purified recombinant protein.

extracts from *E. coli* harboring an empty pMAL-c2 vector after IPTG induction (data not shown). Both isozymes in crude extracts from *E. coli* were significantly inhibited by L-His (Table 3). On the contrary, D-His did not inhibit the enzyme activities of each isozyme (data not shown). IC_{50} values for L-His, and 1,2,4-triazolealanine (TA), which is a His analogous were determined (Table 3). AtATP-PRT1 was more sensitive to those

compounds than AtATP-PRT2, although these K_m values for ATP and PRPP showed resemble. While, in purified AtATP-PRT1 protein, apparent K_m values were 0.13 mM for PRPP and 0.60 mM for ATP, being similar to those of crude extracts, native *T. aestivum* and *S. typhimurium* (Table 3). IC_{50} values for L-His and 1,2,4-triazolealanine of purified AtATP-PRT1 protein were indistinguishable from those of crude extracts from *A. thaliana*, native *T. aestivum* (Münzer et al., 1992) and *S. typhimurium* (Martin, 1963, Whitfield, 1971) (Table 3).

Discussion

The biosynthesis of His has been well-studied in microorganisms (Alifano et al., 1996). Previous studies provided us that synthesis of one His molecule is needed 41 ATP molecules (Brenner and Ames, 1971). I suspect such a pathway requiring much energy should be strictly regulated. Indeed, this pathway in microorganisms is regulated genetically, known as attenuation control in *S. typhimurium* and *E. coli* (Johnson et al., 1980, Blasi and Bruni, 1981) or general control in *S. cerevisiae* (Hinnebusch, 1988), and biochemically, known as feedback inhibition by L-His (Martin, 1963, Bell et al., 1971). On the contrary, although several steps of this pathway have been identified and characterized in higher plants (Nagai et al., 1991, Tada et al., 1994, El Malki, et al., 1998, Chapter I; Fujimori and Ohta, 1998a, Chapter II; Fujimori et al., 1998b, Chapter III; Fujimori and Ohta, 1998c), no one has found out the regulation mechanism of the plant His biosynthetic pathway yet.

Genomic Southern blot analysis demonstrated that *AtATP-PRT* genes exist two isoforms in *Arabidopsis* genome. Each of bands detected under low stringency condition was completely identical to the results under high stringency condition (Fig. 4). It means that only two genes encoding isozymes for ATP-PRT existed in *Arabidopsis* genome. The gene for AtATP-PRT1 has been isolated by the library screening, while, *AtATP-PRT2* gene have been identified in the *Arabidopsis* genome project. The sequence analysis of both genes showed that each of both genome structures is well-resemble including the intron positions. It indicates that two isoforms were evolutionally developed as the intron-dependent manner.

74

Messenger RNAs of both *AtATP-PRT* genes were detected in all tissues investigated here throughout the development. Other His biosynthetic genes were expressed in all tissues throughout the development (Nagai et al., 1991, Tada et al., 1994, Chapter I; Fujimori and Ohta, 1998a, Chapter II; Fujimori et al., 1998b, Chapter III; Fujimori and Ohta, 1998c). However, *AtATP-PRT2* expression in roots is significantly lower than those in other tissues (Fig. 8). This data is consistent that protein levels in roots are also lower than other tissues (Fig. 9b). The expression of the plant His biosynthetic genes is not seemed to be regulated by L-His, differing from microbial mechanisms. Guyer et al. reported that a specific inhibitor for IGPD of the His biosynthesis (Mori et al., 1995) caused to increase the expression level of genes involved in the biosynthesis of His, aromatic amino acids, lysine, and purines in *A. thaliana* (Guyer et al., 1995), like a general control of *S. cerevisiae* (Hinnebusch, 1988). It might mean that the existence of the genetic regulatory mechanism for the His biosynthesis in higher plants. Further analysis of the promoter region of the His biosynthetic genes will provide the understanding for the regulation mechanism genetically.

In microorganisms, ATP-PRT plays a key role in the biochemical regulation of the His biosynthesis. However, in higher plants, the regulatory mechanim of ATP-PRT has not been precisely elucidated yet. In this Chapter, I isolated two full-length cDNA clones encoding the first step enzyme of the His biosynthesis, ATP-PRT from *A. thaliana*, designated as *AtATP-PRT1* and *AtATP-PRT2*, and investigated the regulation mechanisms of the His biosynthesis. The present data clearly reveals the important regulatory role of ATP-PRT in higher plants. *Arabidopsis* AtATP-PRT1 protein produced in *E. coli* was purified and characterized. The assay using crude extracts prepared from *E. coli* expressing each isozyme show similar properties about apparent K_m values for ATP and PRPP, and IC_{50} values for L-His and 1,2,4-triazolealanine (TA) (Table 3). In purified AtATP-PRT1 protein, ATP-PRT activity is significantly inhibited by L-His (IC_{50} = 45 µM), which is similar to bacterias. The K_m values of recombinant ATP-PRT1 (0.13 mM for PRPP and 0.60 mM for ATP) were extremely resemble to those of native *T. aestivum* and *S. typhimurium* (Martin, 1963, Whitfield, 1971, Münzer et al., 1992). In higher plants, the His biosynthesis is seemed to be regulated only biochemically. Therefore, ATP-PRT, which is a first enzyme of His biosynthetic pathway, plays a

75

central role to regulate the intracellular His mass in higher plants.

In the present point, the existence of other regulatory mechanisms for the plant His biosynthesis is undeniable. As plant ATP-PRT isozymes might be localized in plastid, the expression of these genes may be affected by energy dependent regulation, for instance, light and darkness. Furthermore, plant ATP-PRT enzyme may form hexameric structure, like bacterial one (Voll et al., 1967). The combinations of subunit of two isozymes might show different kinetic properties.

Finally, all genes involved in the His biosynthetic pathway in higher plants except for histidinolphosphate phosphatase which catalyzes 9th-step of the His pathway have been identified. The plant His biosynthetic pathway basically is same as the microbial pathways. In addition, IGPD and HDH involved in the His biosynthesis are localized in chloroplast (Nagai et al., 1993, Tada et al., 1995). Protein sequences deduced from 7 genes, ATP-PRT (1st-step), PRA-PH/PRA-CH (2nd-/3rd-steps), BBM II isomerase (4th-step), glutamine amidotransferase/cyclase (5th-/6th-steps), IGPD (7th-step), HPA (8th-step) and HDH (10th-/11th-steps) have the putative chloroplast transit peptides at their N-terminus (Nagai et al., 1991, Tada et al., 1994, El Malki et al., 1998, Chapter I; Fujimori and Ohta, 1998a, Chapter II; Fujimori et al., 1998b, Chapter III; Fujimori and Ohta, 1998c). These findings indicate that the whole reactions involved in the His biosynthetic pathway are performed in chloroplasts. Compartmentalization of the His biosynthesis in chloroplast, which is energy consuming pathway is great advantage to the plant.

Summary

I have characterized biochemically and molecular biologically two isoforms (*AtATP-PRT1* and *AtATP-PRT2*) of ATP phosphoribosyl transferase (ATP-PRT) from *Arabidopsis thaliana*, catalyzing the first step of the pathway of histidine (His) biosynthesis. DNA gel blot analysis indicated that ATP-PRTs in *Arabidopsis* are encoded by two separate genes having closely similar gene organizations. The primary structures deduced from the cDNA sequences of *AtATP-PRT1* and *AtATP-PRT2* share an overall amino acid identity of 74.6% and contain N-terminal chloroplast transit peptide sequences. RNA gel blot analyses demonstrated ubiquitous expression of both isoforms throughout development. Both *AtATP-PRT1* and *AtATP-PRT2* cDNAs were able to suppress the His auxotrophy of a mutant yeast strain defective in the ATP-PRT activity, indicating that both genes encode functional enzymes. Protein blot analysis displayed that ATP-PRT proteins were predominantly accumulated in leaves. The K_m values for ATP (0.6 mM) and phosphoribosyl pyrophosphate (PRPP) (0.13 mM) of a recombinantly expressed AtATP-PRT1 were comparable to those of native ATP-PRTs from higher plants and bacteria. The activity of recombinant AtATP-PRT1 is inhibited by L-His (IC$_{50}$ = 45 μM), indicating that the His biosynthesis is regulated in plants by the feedback inhibition by L-His, while no apparent regulation of the His biosynthetic genes have been observed.

Conclusion

In this thesis, His biosynthetic pathway of higher plants was investigated by isolating cDNAs and genes from *A. thaliana* encoding the enzymes involved in the His biosynthesis. The regulation mechanisms of the His biosynthesis in higher plants were also discussed.

The cDNA (At-IE) and the gene encoding bifunctional PRA-PH (hisE)/PRA-CH (hisI) were isolated from *A. thaliana* by complementation of the *hisI* defective *E. coli* mutant, as described in Chapter I. The N- and C terminal portion of the recombinant proteins produced in *E. coli* had PRA-CH and PRA-PH activity, respectively. The *At-IE* gene was constitutively expressed in all tissues throughout the development.

The cDNA and the gene encoding BBM II isomerase were isolated from *A. thaliana* by complementation of *hisA* defective *E. coli* mutant, as described in Chapter II. The BBM II isomerase activity of the recombinant protein produced in *E. coli* was confirmed by the *in vitro* enzyme assay. This gene was constitutively expressed in all tissues throughout the development.

The cDNA (At-HF) encoding bifunctional GAT/cyclase was isolated from *A. thaliana*, as described in Chapter III. His auxotrophy of *his7* defective *S. cerevisiae* mutant was suppressed by harboring *At-HF* cDNA to the *S. cerevisiae* mutant. The *At-HF* gene was expressed constitutively in all tissues throughout all developmental stages.

Complementary DNAs and genes for two isoforms of ATP-PRT (*AtATP-PRT1* and *AtATP-PRT2*) were isolated from *A. thaliana*, as described in Chapter IV. The AtATP-PRT proteins expressed in *E. coli* showed the biochemical properties similar to those of the enzymes isolated from microorganisms and *T. aestivum*. The *AtATP-PRT2* gene was constitutively expressed in all tissues throughout the development except for roots in which the expression was lower.

As described above, 4 cDNAs and 3 genes involved in the His biosynthetic pathway were isolated in the present investigation. These cDNAs and genes encoded six

enzymes involved in the His biosynthesis. Among them, At-IE and At-HF encoded bifunctional enzyme proteins, PRA-PH/PRA-CH and GAT/cyclase, respectively. In addition to IGPD, HPA and HDH reported previously, 9 steps in the His biosynthesis have been revealed. These results suggest that the His biosynthetic pathway in higher plants is the same as (or almost similar to) that in microorganisms such as *E. coli* and *S. cerevisiae*.

Arabidopsis 4 genes related to the His biosynthesis were expressed ubiquitously, and it does not appear that the transcriptional regulation does not exist for the expression of His biosynthetic genes in higher plants. However, the enzyme activity of the recombinant AtATP-PRT1 was significantly inhibited by L-His. This result suggests that the feedback inhibition of ATP-PRT by L-His is a major regulatory mechanism for the His biosynthesis in higher plants.

Amino terminal extensions of all genes investigated here had chloroplast transit peptide-like sequences. It was reported that two enzymes, IGPD and HDH, were localized in chloroplasts. These results, therefore, suggest that the whole His biosynthesis is performed in plastids including chloroplasts in higher plants. The compartmentalization of the whole pathway in chloroplasts is advantageous to plants, because the His biosynthesis requires much energy.

References

Alifano, P., Fani, R., Liò, P., Lazcano, A., Bazzicalupo, M., Carlomagno, M.S., and Bruni, C.B. (1996) Histidine biosynthetic pathway and genes: structure, regulation, and evolution. *Microbiol. Rev.* **60**: 44-69.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**: 403-410.

Altschul, S.F., Madden, T.L, Schäffer, A.A., Zhang, J., Zhang Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* **25**: 3389-3402.

Ames, B.N., Garry, B., and Herzenberg, L.A. (1960) The genetic control of the enzymes of histidine biosynthesis in *Salmonella typhimurium. J. Gen. Microbiol.* **22**: 369-378.

Ames, B.N., Martin, R.G., and Garry, B.J. (1961) The first step of histidine biosynthesis. *J. Biol. Chem.* **236**: 2019-2026.

Ames, B.N., Hartman, P. E., and Jacob, F. (1963) Chromosomal alterations affecting the regulation of histidine biosynthetic enzymes in *Salmonella. J. Mol. Biol.* **7**: 23-42.

Andreadis, A., Hsu, Y.P., Kohlhaw, G.B., and Schimmel, P. (1982) Nucleotide sequence of yeast *LEU2* shows 5'-noncoding region has sequences cognate to leucine. *Cell* **31**: 319-325.

Apolinario, E., Nocero, M., Jin, M., and Hoffman, C.S. (1993) Cloning and manipulation of the *Schizosaccharomyces pombe his7⁺* gene as a new selectable marker for molecular genetic studies. *Curr. Genet.* **24**: 491-495.

Arndt, K., and Fink, G.R. (1986) GCN4 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.

Artz, S.W., and Broach, J.R. (1975) Histidine regulation in *Salmonella typhimurium*: an activator-attenuator model of gene regulation. *Proc. Natl. Acad. Sci. USA.* **72**: 3453-3457.

Beckler, G.S., and Reeve, J.N. (1986) Conservation of primary structure in the *hisI* gene of the archaebacterium, *Methanococcus vannielli*, the eubacterium *Escherichia coli*, and the eucaryote *Saccharomyces cerevisiae. Mol. Gen. Genet.* **204**: 133-140.

Bell, R.M., and Koshland.Jr., D.E. (1971) Allosteric properties of the first enzyme of the histidine operon. *Bioorganic Chem.* **1**: 409-423.

Blasi, F., and Bruni, C.B. (1981) Regulation of the histidine operon: Translation controlled transcription termination (A mechanism common to several biosynthetic operons). *Curr. Top. Cell. Regul.* **19**: 1-45.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**: 248-254.

Breathnach, R., and Chambon, P. (1981) Organization of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.* **50**: 349-383.

Brenner, M., and Ames, B.N. (1971) In Metabolic Pathways, Vol. 5. Vogel, H.S., ed: Metabolic Regulation. Academic Press. New York. pp. 349-387.

Bruni, C.B., Musti, A.M., Frunzio, R., and Blasi, F. (1980) Structural and physiological studies of the Escherichia coli histidine operon inserted into vectors. *J. Bacteriol.* **60**: 17-28.

Bult, C.J., White, O., Olsen, G.J., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, J.A., FitzGerald, L.M., Clayton, R.A., Gocayne, J.D., Kerlavage, A.R., Dougherty, B.A., Tomb, J-F., Adams, M.D., Reich, C.I., Overbeek, R., Kirkness, E.F., Weinstock, K.G., Merrick, J.M., Glodek, A., Scott, J.L., Geoghagen, N.S.M., Weidman, J.F., Fuhrmann, J.L., Nguyen, D., Utterback, T.R., Kelley, J.M., Peterson, J.D., Sadow, P.W., Hanna, M.C., Cotton, M.D., Roberts, K.M., Hurst, M.A., Kaine, B.P., Borodovsky, M., Klenk, H-P., Fraser, C.M., Smith, H.O., Woese, C.R., and Venter, J.C. (1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii. Science* **273**: 1058-1073.

Carlomagno, M.S., Chiariotti, L., Alifano, P., Nappo, A.G., and Bruni, C.B. (1988) Structure and function of the *Salmonella typhimurium* and *Escherichia coli* K-12 histidine operons. *J. Mol. Biol.* **203**: 585-606.

Chomczynski, P., and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**: 156-159.

Conover, R.K., and Doolittle, W.F. (1990) Characterization of a gene involved in histidine biosynthesis in *Halobacterium (Haloferax) volcanii*: isolation and rapid mapping by transformation of an auxotroph with cosmid DNA. *J. Bacteriol.* **172**: 3244-3249.

Crane, D.I., and Gould, S.J. (1994) The *Pichia pastoris HIS4* gene: nucleotide sequence, creation of a non-reverting *his4* deletion mutant, and development of *HIS4*-based replicating and integrating plasmids. *Curr. Genet.* **26**: 443-450.

Cue, D., Beckler, G.S., Reeve, J.N., and Konisky, J. (1985) Structure and sequence divergence of two archaebacterial genes. *Proc. Natl. Acad. Sci. USA.* **82**: 4207-4211.

de Kievit, T.R., Dasgupta, T., Schweizer, H., and Lam, J.S. (1995) Molecular cloning and characterization of the *rfc* gene of *Pseudomonas aeruginosa* (serotype O5). *Mol. Microbiol.* **16**: 565-574.

Delorme, C., Ehrlich, S.D., and Renault, P. (1992) Histidine biosynthesis genes in *Lactococcus lactis* subsp. *lactis. J. Bactriol.* **174**: 6571-6579.

Donahue, T.F., Farabaugh, P.J., and Fink, G.R. (1982) The nucleotide sequence of the *HIS4* region of yeast. *Gene* **18**: 47-59.

El Malki, F., Frankard, V., and Jacobs, M. (1998) Molecular cloning and expression of a cDNA sequence encoding histidinol phosphate aminotransferase from *Nicotiana tabacum*. *Plant Mol. Biol.* **37**: 1013-1022.

Erickson, F.Les., and Hannig, E.M. (1995) Isolation of *S. pombe* cDNAs encoding homologs of the *S. cerevisiae HIS1* (ATP phosphoribosyl transferase) and *HIS3* (imidazoleglycerol-phosphate dehydratase) genes. *Yeast* **11**: 157-167.

Fani, R., Bazzicalupo, M., Damiani, G., Bianchi, A., Schipani, C., Sgaramella, V., and Polsinelli, M. (1989) Cloning of histidine genes of *Azospirillum brasilense*: organization of the *ABFH* gene cluster and nucleotide sequence of the *hisB* gene. *Mol. Gen. Genet.* **216**: 224-229.

Fani, R., Alifano, P., Allotta, G., Bazzicalupo, M., Carlomagno, M.S., Gallori, E., Rivellini, F., and Polsinelli, M. (1993) The histidine operon of *Azospirillum brasilense*: organization, nucleotide sequence and functional analysis. *Res. Microbiol.* **144**: 187-200.

Fani, R., Liò, P., and Lazcano, A. (1995) Molecular evolution of the histidine biosynthetic pathway. *J. Mol. Evol.* **41**: 760-774.

Fani, R., Tamburini, E., Mori, E., Lazcano, A., Lio, P., Barberio, C., Casalone, E., Cavalieri, D., Perito, B., and Polsinelli, M. (1997) Paralogous histidine biosynthetic genes: evolutionary analysis of the *Saccharomyces cerevisiae HIS6* and *HIS7* genes. *Gene* **197**: 9-17.

Feinberg, A.P., and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **132**: 6-13.

Felsenstein, J. (1993) Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**: 368-376.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J-F., Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J.D., Scott, J., Shirley, R., Liu, L-I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T.,

Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Geoghagen, N.S.M., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O., and Venter, J.C. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**: 496-512.

Fujimori, K., and Ohta, D. (1998a) Isolation and characterization of a histidine biosynthetic gene in *Arabidopsis* encoding a polypeptide with two separate domains for phosphoribosyl-ATP pyrophosphohydrolase and phosphoribosyl-AMP cyclohydrolase. *Plant Physiol.* **118**: 275-283.

Fujimori, K., Tada, S., Kanai, S., and Ohta, D. (1998b) Molecular cloning and characterization of a gene encoding *N'*-[(5'-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (BBM II) isomerase from *Arabidopsis thaliana. Mol. Gen. Genet.* **259**: 216-223.

Fujimori, K., and Ohta, D. (1998c) An *Arabidopsis* cDNA encoding a bifunctional glutamine amidotransferase/cyclase suppresses the histidine auxotrophy of a *Saccharomyces cerevisiae his7* mutant. *FEBS Lett.* **428**: 229-234.

Gantt, J.S., Baldauf, S.L., Calie, P.J., Weeden, N.F., and Palmer, J.D. (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.

Goldschmidt, E.P., Cater, M.S., Matney, T.S., Bulter, M.A., and Greene, A. (1970) Gnetic analysis of the histidine operon in Escherichia coli K12. *Genetics* **66**: 219-229.

Gu, W., Zhao, G., Eddy, C., and Jensen, R.A. (1995) Imidazoleacetolphosphate aminotransferase in *Zymomonas mobilis*: molecular genetic, biochemical, and evolutionary analyses. *J. Bacteriol.* **177**:1576-1584.

Guyer, D., Patton, D., and Ward, E. (1995) Evidence for cross-pathway regulation of metabolic gene expression in plants. *Proc. Natl. Acad. Sci. USA.* **92**: 4997-5000.

Hartmann, P.E., Hartman, Z., and Serman, D. (1960) Complementation-mapping by abortive transduction of histidine-requiring *Salmonella* mutants. *J. Gen. Microbiol.* **22**: 354-368.

Henner, D.J., Band, L., Flaggs, G., and Chen, E. (1986) The organization and nucleotide sequence of the *Bacillus subtilis hisH*, *tyrA* and *aroE* genes. *Gene* **49**:147-152.

Hikiji, T., Ohkuma, M., Takagi, M., and Yano, K. (1989) An improved host-vector system for *Candida maltosa* using a gene isolated from its genome that complements the *his5* mutation of *Saccharomyces cerevisiae. Curr. Genet.* **16**: 261-266.

Hinnebusch, A.G., and Fink, G.R. (1983) Repeated DNA sequences upstream from *HIS1* also occur at several other co-regulated genes in *Saccharomyces cerevisiae. J. Biol. Chem.* **258**: 5238-5247.

Hinnebusch, A.G. (1988) Mechanism of gene regulation in the general control of amino acid biosynthesis in *Saccharomyces cerevisiae. Microbiol. Rev.* **5 2**: 248-273.

Hinshelwood, S., and Stoker, N.G. (1992) Cloning of mycobacterial histidine synthesis genes by complementation of a *Mycobacterium smegmatis* auxotroph. *Mol. Microbiol.* **6**: 2887-2895.

Imamoto, F., Ito, J., and Yanofsky, C. (1966) Polarity studies with the tryptophan operon. *Cold Spring Harbor Symp. Quant. Biol.* **3 1**: 235-249.

Ito, H., Fukuda, Y., Murata, K., and Kimura, A. (1983) Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol.* **1 5 3**: 163-168.

Johnson, H.W., Barnes, W.M., Chumley, F.G., Bossi, L, and Roth, J.R. (1980) Model for regulation for histidine operon of *Salmonella. Proc. Natl. Acad. Sci. USA.* **7 7**: 508-512.

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., and Tabata, S . (1996) Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* **3**: 109-136.

Kasai, T. (1974) Regulation of the expression of histidine operon in *Salmonella typhimurium. Nature* **2 4 9**: 523-527.

Klungsöyr, L., Hageman, J.H., Fall, L., and Atkinson, D.E. (1968) Interaction between energy charge and product feedback in the regulation biosynthetic enzymes: aspartokinase, phosphoribosyladenosine triphosphate synthetase, and phosphoribosyl pyrophosphate synthetase. *Biochemistry* **7**: 4035-4040.

Kuenzler, M., Balmelli, T., Egli, C.M., Paravicini, G., and Braus, G.H. (1993) Cloning, primary structure, and regulation of the *HIS7* gene encoding a bifunctional glutamine amidotransferase: cyclase from *Saccharomyces cerevisiae. J. Bacteriol.* **1 7 5**: 5548-5558.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **2 2 7**: 680-685.

Lagrimini, L.M., Burkhart,W., Moyer, M., and Rothstein, S. (1987) Molecular cloning of complementary DNA encoding the lignin-forming peroxidase from tobacco: molecular analysis and tissue-specific expression. *Proc. Natl. Acad. Sci. USA.* **8 4**: 7542-7546.

Legerton, T.L., and Yanofsky, C . (1985) Cloning and characterization of the multifunctional *his-3* gene of *Neurospora crassa. Gene* **3 9**: 129-140.

Limauro, D., Avitabile, A., Cappellano, C., Puglia, A.M., and Bruni, C.B. (1990) Cloning and characterization of the histidine biosynthetic gene cluster of *Streptomyces coelicolor* A3(2). *Gene* **90**: 31-41.

Lütcke, 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.

Mano, J., Hatano, M., Koizumi, S., Tada, S., Hashimoto, M., and Scheidegger, A. (1993) Purification and properties of a monofunctional imidazoleglycerol-phosphate dehydratase from wheat. *Plant Physiol.* **103**: 733-739.

Margolies, M.N., and Goldberger, R.F. (1966) Isolation of the fourth enzyme (isomerase) of histidine biosynthesis from *Salmonella typhimurium*. *J. Biol. Chem.* **241**: 3262-3269.

Margolies, M.N., and Goldberger, R.F. (1967) Physical and chemical characterization of the isomerase of histidine biosynthesis in *Salmonella typhimurium*. *J. Biol. Chem.* **242**: 256-264.

Martin, R.G. (1963) The first enzyme in histidine biosynthesis: The nature of feedback inhibition by histidine. *J. Biol. Chem.* **238**: 257-268.

Maruyama, I.N., Rakow, T.L., and Maruyama, H.I. (1995) cRACE: a simple method for identification of the 5' end of mRNAs. *Nucl. Acids Res.* **23**: 3796-3797.

Matney, T.S., Goldschmidt, E.P., Erwin, N.S, and Scroggs, R.A. (1964) A preliminary map of genomic sites for F-attachment in *Escherichia coli* K12. *Biochem. Biophys. Res. Commun.* **17**: 278-281.

Minet, M., Dufour, M-E., and Lacroute, F. (1992) Complementation of *Saccharomyces cerevisiae* auxotrophic mutants by *Arabidopsis thaliana* cDNAs. *Plant J.* **2**: 417-422.

Mizutani, M., Ohta, D., and Sato, R. (1997) Isolation of a cDNA and a genomic clone encoding cinnamate 4-hydroxylase from Arabidopsis and its expression manner in planta. *Plant Physiol.* **113**: 755-763.

Mori, I., Fonne-Pfister, R., Matsunaga, S., Tada, S., Kimura, Y., Iwasaki, G., Mano, J., Hatano, M., Nakano, T., Koizumi, S., Scheidegger, A., Hayakawa, K., and Ohta, D. (1995) A novel class of herbicides: specific inhibitors of imidazoleglycerolphosphate dehydratase. *Plant Physiol.* **107**: 719-723.

Mortimer, R.K., Romano, P., Suzzi, G., and Polsinelli, M. (1994) Genome renewal: a new phenomenon revealed from a genetic study of 43 strains of *S . cerevisiae* derived from natural fermentation of grape musts. *Yeast* **10**: 1543-1552.

Morton, D.P., and Parsons, S.M. (1977) Synergistic inhibition of ATP phosphoribosyl transferase by guanosine tetraphosphate and histidine. *Biochem. Biophys. Res. Commun.* **74**: 172-177.

Münzer, S., Hashimoto-Kumpaisal, R., Scheidegger, A., and Ohta, D. (1992) Purification and properties of ATP-phosphoribosyl transferase from wheat germ. Research in Photosynthesis. Vol. IV, Norio, M. ed. Kluwer Academic Publishers, Netherlands, pp. 91-94.

Nagai, A., and Scheidegger, A. (1991) Purification and characterization of histidinol dehydrogenase from cabbage. *Arch. Biochem. Biophys.* **284**: 127-132.

Nagai, A., Ward, E., Beck, J., Tada, S., Chang, J-Y., Scheidegger, A., and Ryals, J. (1991) Structural and functional conservation of histidinol dehydrogenase between plants and microbes. *Proc. Natl. Acad. Sci. USA* **88**: 4133-4137.

Nagai, A., Suzuki, K., Ward, E., Moyer, M., Mano, J., Beck, J., Tada, S., Hashimoto, M., Chang, J-Y., Ryals, J., Scheidegger, A., and Ohta, D. (1993) Histidinol dehydrogenase in higher plants. Purification, cloning and expression. Research in Photosynthesis. Vol. IV. Dordrecht, ed. Kluwer Academic Publishers, Netherlands, pp.95-98.

Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E. and Somerville, C. (1994) Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones. *Plant Physiol.* **106**: 1241-1255.

Newton, W.A., Beckwith, J.R., Zipser, D., and Brennert, S. (1965) Nonsense mutants and polarity in the *lac* operon of *Escherichia coli. J. Mol. Biol.* **14**: 290-296.

Oriol, E., Méndez-Alvarez, S., Barbé, J., and Gilbert, I. (1996) Cloning of the *Rhodobacter sphaeroides hisI* gene: unifunctionality of the encoded protein and lack of linkage to other *his* genes. *Microbiology* **142**: 2071-2078.

Pla, J., Pérez-Díaz, M., Navarro-García, F., Sánchez, M., and Nombela, C. (1995) Cloning of the *Candida albicans HIS1* gene by direct complementation of a *C. albicans* histidine auxotroph using an improved double-ARS shuttle vector. *Gene* **165**: 115-120.

Rieder, G., Merrick, M.J., Castorph, H., and Kleiner, D. (1994) Function of *hisF* and *hisH* gene products in histidine biosynthesis. *J. Biol. Chem.* **269**: 14386-14390.

Rodriguez, R.L., and West, R.W.Jr. (1984) Histidine operon control region of *Klebsiella pneumoniae*: analysis withan *Escherichia coli* promoter-probe plasmid vector. *J. Bacteriol.* **157**: 764-771.

Rose, M.D., Winston, F., and Hieter, P. (1990) Methods in Yeast Genetics: A laboratory course manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Rothnie, H.M., Reid, J., and Hohn, T. (1994) The contribution of AAUAAA and the upstream element UUUGUA to the efficiency of mRNA 3'-end formation in plants. *EMBO J.* **13**: 2200-2210.

Saiki, R.K., Gelfand, D.H., Stofeel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **239**: 487-491.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular cloning: A Laboratory Manual. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* **74**: 5463-5467.

Senecoff, J.F., and Meagher, R.B. (1993) Isolating the *Arabidopsis thaliana* genes for *de novo* purine synthesis by suppression of *Escherichia coli* mutants. *Plant Physiol.* **102**: 387-399.

Sipmson, G.G., and Filipowicz, W. (1996) Splicing of precursors to mRNA in higher plants: mechanism, regulation and sub-nuclear organization of the spliceosomal machinery. *Plant Mol. Biol.* **32**: 1-41.

Smith, D.W.E., and Ames, B.N. (1964) Intermediates in the early steps of histidine biosynthesis. *J. Biol. Chem.* **239**: 1848-1855.

Tada, S., Volrath, S., Guyer, D., Scheidegger, A., Ryals, J., Ohta, D., and Ward, E. (1994) Isolation and characterization of cDNAs encoding imidazoleglycerolphosphate dehydratase from *Arabidopsis thaliana*. *Plant Physiol.* **105**: 579-583.

Tada, S., Hatano, M., Nakayama, Y., Volrath, S., Guyer, D., Ward, E., and Ohta, D. (1995) Insect cell expression of recombinant imidazoleglycerolphosphate dehydratase of Arabidopsis and wheat and inhibition by triazole herbicides. *Plant Physiol.* **109**: 153-159.

Takemura, H., Horinouchi, S., and Beppu, T. (1993) Suppression of an ethanol-sensitive mutation of *Acetobacter pasteurianus* by overexpression of the *his1* gene encoding histidinolphosphate aminotransferase. *J. Ferment. Bioeng.* **76**: 224-228.

Thompson, J.D., Higgins D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.* **22**: 4673-4680.

Valvekens, D., Montagu, M.V., and Lusebettens, M.V. (1988) *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. *Proc. Natl. Acad. Sci. USA.* **85**: 5536-5540.

Vogel, H.J., and Bonner, D.M. (1956) Acetylornithase of *Escherichia coli*: partial purification and some properties. *J. Biol. Chem.* **218**: 97-106.

Voll, M.J., Appella, E., and Martin, R. (1967) Purification and composition studies of phosphoribosyl adenosine triphosphate: pyrophosphate phosphoribosyltransferase, the first enzyme of histidine biosynthesis. *J. Biol. Chem.* **242**: 1760-1767.

von Heijne, G., and Nishikawa, K. (1991) Chloroplast transit peptides: the perfect random coil? *FEBS Lett* **278**: 1-3.

Wahle, E., and Keller, W. (1992) The biochemistry of 3'-end cleavage and polyadenylation of messenger RNA precursors. *Annu. Rev. Biochem.* **61**: 419-440.

Weil, C.F., Beckler, G.S., and Reeve, J.N. (1987) Structure and organization of the *hisA* gene of the thermophilic archaebacterium *Methanococcus thermolithotrophicus*. *J. Bacteriol.* **169**: 4857-4860.

Weinstock, K.G., and Strathern, J.N. (1993) Molecular genetics in *Saccharomyces kluyveri:* the *HIS3* homolog and its use as a selectable marker gene in *S. kluyveri* and *Saccharomyces cerevisiae. Yeast* **9**: 351-361.

Whitfield, H.J.Jr (1971) Purification and properties of the wild type and a feedback-resistant phosphoribosyl adenosine triphosphate. *J. Biol. Chem.* **246**: 899-908.

Wiater, A., Krajewska-Grynkiewicz, K., and Klopotowski, T. (1971) Histidine biosynthesis and its regulation in plants. *Acta Biochem. Pol.* **18**: 299-307.

Winkler, M.E. (1987) Biosynthesis of histidine. *Escherichia coli* and *Salmonella typhimurium*. Neidhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M., Umbarger, H.E., eds: Cellular and Molecular Biology, Vol. I. American Society of Microbiology, Washington, D.C., pp. 395-411.

Wolter, F.P., Fritz, C.C., Willmitzer, L., Schell, J., and Schreier, P.H. (1988) *rbcS* genes in *Solanum tuberosum*: conservation of transit peptide and exon shuffling during evolution. *Proc. Natl. Acad. Sci. USA.* **85**: 846-850.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119.

List of Publications

1. Matsumoto, M., **Fujimori, K.**, Okuno, T., Arai, K., Shindo, M., Takeda, T., Okanoue, T., Kashima, K., and Sokawa, Y. (1991) Change of the levels of hepatic and serum HCV-RNA by interferon therapy. *Acta Hepatologica Jpn.* **3 2**: 983-989.

2. Anamnart, S., Tomita, T., Fukui, F., **Fujimori, K.**, Harashima, S., Yamada, Y., and Oshima, Y. (1997) The *P-OLE1* gene of *Pichia angusta* encodes a Δ9-fatty acid desaturase and complements the *ole1* mutation of *Saccharomyces cerevisiae. Gene* **184**: 299-306.

3. **Fujimori, K.**, Anamnart, S., Nakagawa, Y., Sugioka, S., Ohta, D., Oshima, Y., Yamada, Y., and Harashima, S. (1997) Isolation and characterization of mutations affecting expression of the Δ9-fatty acid desaturase gene, *OLE1*, in *Saccharomyces cerevisiae. FEBS Lett.* **413**: 226-230.

4. **Fujimori, K.**, and Ohta, D. (1998) An *Arabidopsis* cDNA encoding a bifunctional glutamine amidotransferase/cyclase suppresses the histidine auxotrophy of a *Saccharomyces cerevisiae his7* mutant. *FEBS Lett.* **428**: 229-234.

5. **Fujimori, K.**, Tada, S., Kanai, S., and Ohta, D. (1998) Molecular cloning and characterization of the gene encoding *N*'-[(5'-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide ribonucleotide (BBM II) isomerase from *Arabidopsis thaliana. Mol. Gen. Genet.* **259**: 216-223.

6. **Fujimori, K.**, and Ohta, D. (1998) Isolation and characterization of a histidine biosynthetic gene in *Arabidopsis* encoding a polypeptide with two separate domains for phosphoribosyl-ATP pyrophosphohydrolase and phosphoribosyl-AMP cyclohydrolase. *Plant Physiol.* **118**: 275-283.

7. Ohta, D., **Fujimori, K.**, Mizutani, M., Münzer, S., Nakayama, Y., and Kozaki, A. Characterization of two *Arabidopsis* isoforms for ATP phosphoribosyl transferase regulating the histidine biosynthesis. (*In preparation*)

Acknowledgments