Biochem. J. (1991) 277, 469–475 (Printed in Great Britain)

Isolation, characterization and sequence analysis of a full-length cDNA clone encoding acetohydroxy acid reductoisomerase from spinach chloroplasts

Renaud DUMAS,* Michel LEBRUN† and Roland DOUCE*‡

*Unité Mixte C.N.R.S./Rhône–Poulenc (Unité Associée au Centre National de la Recherche Scientifique, U.M. 41), and †Service de Biologie Moléculaire et Cellulaire Végétale, Rhône–Poulenc Agrochimie, 14–20 Rue Pierre Baizet, 69263 Lyon Cedex 09, France

Acetohydroxy acid reductoisomerase (AHRI), the second enzyme in the parallel isoleucine/valine-biosynthetic pathway, catalyses an unusual two-step reaction in which the substrate, either 2-acetolactate or 2-aceto-2-hydroxybutyrate, is converted via an alkyl migration and an NADPH-dependent reduction to give 2,3-dihydroxy-3-methylbutyrate or 2,3-dihydroxy-3-methylvalerate respectively. We have isolated and characterized a full-length cDNA from a λgt11 spinach library encoding the complete acetohydroxy acid reductoisomerase protein precursor. The 2050-nucleotide sequence contains a 1785-nucleotide open reading frame. The derived amino acid sequence indicates that the protein precursor consists of 595 amino acid residues including a presequence peptide of 72 amino acid residues. The *N*-terminal sequence of the first 16 amino acid residues of the purified AHRI confirms the identity of the cDNA. The derived amino acid sequence from this open reading frame shows 23 % identity with the deduced amino acid sequences of the *Escherichia coli* and *Saccharomyces cerevisiae* AHRI proteins. There are two blocks of conserved amino acid residues in these three proteins. One of these is a sequence similar to the 'fingerprint' region of the NAD(P)H-binding site found in a large number of NAD(P)H-dependent oxidoreductases. The other, a short sequence (Lys-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Ser-His-Gly-Phe) containing the amino acids lysine and histidine, could well be the catalytic site of the first step of the AHRI reaction. Southern-blot analysis indicated that AHRI is encoded by a single gene per haploid genome of about 7.5 kbp containing at least four introns.

INTRODUCTION

The enzyme acetohydroxy acid reductoisomerase (AHRI, EC 1.1.1.86) catalyses an unusual two-step reaction in the assembly of the carbon skeletons of valine and isoleucine. The substrate, either 2-acetolactate or 2-aceto-2-hydroxybutyrate, is converted via an alkyl migration and an NADPH-dependent reduction to give 2,3-dihydroxy-3-methylbutyrate or 2,3-dihydroxy-3-methylvalerate respectively. Interestingly, inhibition of AHRI is known to give rise to herbicidal effects (Schulz *et al.*, 1988; Aulabaugh & Schloss, 1990).

AHRI has been previously purified to homogeneity from the stroma of spinach leaf chloroplasts (Dumas *et al.*, 1989). The native enzyme is a tetramer with a subunit M_r of 59000. In addition, AHRI was eluted from Mono Q ion-exchange resins as three distinct peaks of activity. Each form exhibited an identical M_r of 59000 on SDS/PAGE, whereas they were easily distinguishable by PAGE under non-denaturing conditions (Dumas *et al.*, 1989). The origin of these forms can be attributed to combinations of similar but non-identical, polypeptide chains. In this case the multiplicity could be a result of the synthesis of different types of subunit coded by distinct genes at different loci or distinct alleles at a single locus, or the occurrence of post-transcriptional modifications.

To gain more information about the structure, function and possible active site of AHRI we have isolated and sequenced a spinach AHRI cDNA encoding the entire protein. Sequence analysis revealed that AHRI from spinach leaves contains the characteristics of an NADPH-binding protein and a continuous short stretch of four amino acid residues conserved in analogous sequences from *Saccharomyces cerevisiae* (Petersen & Holmberg, 1986) and *Escherichia coli* (Wek & Hatfield, 1986). Evidence is also presented to show that the AHRI in spinach is encoded by a single nuclear gene.

MATERIALS AND METHODS

Purification of AHRI

Chloroplasts were isolated from spinach (*Spinacia oleracea* L.) leaves and purified as described by Douce & Joyard (1982). AHRI was purified over 400-fold from the stroma of chloroplasts to a specific activity of 91 nkat/mg of protein with 2-aceto-2-hydroxybutyrate as substrate as described by Dumas *et al.* (1989).

Preparation of antibodies to AHRI

Purified AHRI was checked for purity by SDS/PAGE in gels containing a 7.5–15 % (w/v) acrylamide gradient. Bands corresponding to AHRI were excised, pooled and injected as an emulsion with Freund's adjuvant into two rabbits for raising antibodies. Injections (250 μg of protein each time) were given

Abbreviation used: AHRI, acetohydroxy acid reductoisomerase.

‡ To whom correspondence should be addressed.

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession number X57073.

four times at 3-week intervals. IgG was purified from rabbit antisera by the method of Saint-Blancard *et al.* (1981). For the expression screening of the λgt11 library, the antibodies were further immunopurified by affinity chromatography on pure AHRI (4 mg) bound to CNBR-activated Sepharose 4B according to the manufacturer's instructions (Pharmacia). Immunopurified antibodies were stored at −20 °C in 30 mM-Tris/HCl buffer, pH 7.5, containing 150 mM-NaCl, 0.02% (w/v) NaN_3 and 1% (w/v) BSA until use.

Nucleic acid extraction

RNA. A 25 g batch of frozen leaf material was ground to a fine powder in liquid N_2 with a mortar and pestle and homogenized for 4–5 min in 100 ml of lysis buffer [50 mM-Tris/HCl buffer, pH 8.0, containing 7 M-urea, 0.35 M-NaCl, 2 mM-EDTA, 1% (w/v) Sarkosyl and 5% (v/v) phenol] (Shure *et al.*, 1983). When all the aggregates had been solubilized, 1 vol. of phenol/chloroform (1:1, v/v) containing 0.5% (v/v) SDS was added, and the mixture was swirled for 10 min at room temperature in a conical flask and then centrifuged at 10000 *g* for 10 min. The aqueous phase was removed and extracted once more with phenol/chloroform (1:1, v/v). RNA was further purified by LiCl (2.5 M) precipitation. The pellet of total leaf RNA was dissolved in 5 ml of sterile distilled water and stored at −20 °C. The polyadenylated RNA fraction was prepared from the total RNA with the use of oligo(dT)–cellulose as described by Kingston (1987).

DNA. After propan-2-ol precipitation nucleic acids were resuspended in water. RNAase A (100 μg/ml) was added, and after incubation at 37 °C for 30 min the proteins were digested with of proteinase K (100 μg/ml) at 37 °C for 30 min. Genomic DNA was extracted with phenol/chloroform (1:1, v/v) and reprecipitated. The pellet was dissolved in 5 ml of TE buffer (10 mM-Tris/HCl buffer, pH 8.0, containing 1 mM-EDTA) and stored at −20 °C until use.

Construction of a spinach leaf cDNA library

cDNA (double-stranded, blunt-end cDNA) was synthesized from spinach leaf polyadenylated mRNA by using a cDNA Synthesis System Plus kit according to the manufacturer's instructions (Amersham International). Blunt-ended cDNA was ligated to *Eco*RI–*Sma*I synthetic linkers (Biolabs, Beverly, MA, U.S.A.) and size-fractionated by Sephadex G-100 chromatography. cDNAs greater than 500 bp in length were phosphorylated before ligation to λgt11 DNA that had been cut and dephosphorylated at the *Eco*RI restriction site (Stratagene). Ligation products were packaged with λ head and tail proteins to produce viable bacteriophage particles by using a Gigapack II Plus kit according to the manufacturer's instructions (Stratagene). The resulting library consisted of approximately 2×10^6 recombinant clones. Half of the library was amplified in *E. coli* Y 1088 to allow multiple screening.

Immunological screening

The cDNA library was screened (Huynh *et al.*, 1985) for AHRI-specific clones with immunopurified antibody previously prepared against denaturated AHRI from spinach chloroplasts. To screen with this antiserum 3×10^5 plaque-forming units were plated out on *E. coli* Y 1090. AHRI antigen was detected according to the Western-blot technique of Burnette (1981) and the colour reaction was performed with the Promega alkaline phosphatase kit. One positive plaque was replated and rescreened (four or five cycles) until all plaques yielded only positive signals. The size of the cDNA insert in this clone was determined by digestion of purified recombinant λDNA with *Eco*RI and subsequent agarose-gel electrophoresis.

The cDNA was inserted into pUC19 plasmids and used to transform *E. coli* strain DH5α for amplification. Plasmid DNA was prepared by the method of Birnboim & Doly (1979).

DNA sequencing and sequence analysis

DNA sequence analysis was carried out on both strands by the dideoxy chain-termination method (Sanger *et al.*, 1977) with the use of T7 DNA polymerase (Pharmacia). Several restriction fragments were subcloned into pUC19 and sequenced by using the M13 universal primer and the M13 reverse primer (Boehringer). For regions that were inaccessible through the subcloning process, several 18-mer primers were synthesized to complete the sequencing. Ambiguous band patterns were resolved by using 7-deaza-dGTP according to the manufacturer's instructions (Pharmacia). For data handling a Compaq (Deskpro 286 e) computer was used with the Genepro program (Riverside Scientific, Seattle, WA, U.S.A.).

Propensity measures for the formation of α-helices and β-sheets for the derived AHRI polypeptide sequence were determined by the Ralph & Smith program, which uses the algorithm of Chou & Fasman (1978). An alignment of the derived amino acid sequences for spinach, *Saccharomyces cerevisiae* and *E. coli* AHRIs was produced by using the Bestfit program (Smith & Waterman, 1981).

Northern-blot and Southern-blot analysis

Northern transfer of RNA, denatured with methylmercuric hydroxide, to Hybond N membranes (Amersham) was carried out by the procedure of Maniatis *et al.* (1982). Filters were prehybridized for 4 h at 55 °C in 6 × SSC/5 × Denhardt's solution containing 0.5% (w/v) SDS and 0.1 mg of denatured salmon sperm DNA/ml [1 × SSC is 0.15 M-NaCl/15 mM-sodium citrate buffer, pH 7.0; 1 × Denhardt's solution is 0.1% BSA/0.1% Ficoll (M_r 400000)/0.1% polyvinylpyrrolidone (M_r 40000)]. The DNA probe (full-length cDNA) was ^{32}P-labelled with the random primed labelling DNA kit (Boehringer) and hybridization was allowed to proceed overnight at 55 °C. Membranes were washed consecutively in (*a*) 2 × SSPE containing SDS, (*b*) 1 × SSPE containing 0.5% (w/v) SDS and 2% (w/v) $Na_4P_2O_7$ and (*c*) 0.7 × SSPE/0.5% (w/v) SDS, each time for 30 min at 65 °C (1 × SSPE is 0.18 M-NaCl/10 mM-sodium phosphate buffer, pH 7.7, containing 1 mM-Na_2EDTA). Filters were autoradiographed for 24 h at −80 °C. Methylene Blue staining of RNA on nylon membranes was carried out by the procedure of Maniatis *et al.* (1982).

Nuclear DNA was digested for 2 h with restriction enzymes and concentrated by ethanol precipitation. DNA was fractionated by agarose-(0.7%)-gel electrophoresis, and Southern transfer was carried out by the procedure of Maniatis *et al.* (1982). Hybridization with several labelled cDNA probes was conducted overnight at 65 °C in the nucleic acid hybridization solution described above. The blots were washed three times at 65 °C in SSPE for 30 min as described above. From the cDNA clone four probes were used. Probe PCR 1 was obtained by using the polymerase chain reaction (Saiki *et al.*, 1988). The three others were restriction fragments of the AHRI cDNA: *Eco*RI–*Pst*I, *Pst*I–*Hin*dIII and *Eco*RV–*Eco*RI.

Amino acid sequencing

Portions of protein were used for direct sequence analysis by automated serial Edman degradation with an Applied Biosystems 477 gas–liquid-phase protein Sequenator. Amino acid com-

position was determined after acid hydrolysis by using a Kontron analyser.

RESULTS

Characterization of AHRI

A three-step protocol was devised to purify AHRI, which can be released as soluble protein from spinach chloroplasts. After the final step of purification (anion-exchange f.p.l.c. on Mono Q) the native enzyme was eluted from the anion-exchange resin as three distinct peaks of activity. Each form exhibited an identical M_r of 59000 on SDS/PAGE whereas they were easily distinguishable by PAGE under non-denaturing conditions. The M_r values of each form were estimated to be 235000, 220000 and 205000 (see Dumas *et al.*, 1989). This would suggest that the native AHRI is a tetramer composed of four subunits with identical M_r 59000. The *N*-terminal sequence of each form was determined by using a gas–liquid-phase sequencer and gave a single sequence of 16 amino acid residues (Val-Ser-Ala-Pro-Ser-Ile-Asn-Thr-Pro-Ser-Ala-Thr-Thr-Phe-Asp-Phe). In addition, the determination of the overall amino acid composition does not show any difference between the forms (results not shown). Furthermore, antibodies raised against the purified AHRI reacted only with a polypeptide of M_r 59000 in a preparation of total chloroplast proteins separated by SDS/PAGE and transferred to nitrocellulose (results not shown). The origin of these forms can be attributed to combinations of similar, but non-identical, polypeptide chains. The multiplicity of subunit could arise from (*a*) the synthesis of different types of subunit coded by distinct genes at different loci, (*b*) the synthesis of different types of subunit coded by different alleles of one gene at a single locus or (*c*) the occurrence of post-transcriptional modifications of a single gene product.

Isolation and characterization of AHRI cDNA from spinach

A spinach leaf cDNA library in λgt11 was screened with the antibodies to AHRI, and one plaque out of a total of 3×10^5 gave a positive signal. This positive clone was further purified by five rounds of screening at increasingly large bacteriophage densities. Agarose-gel electrophoresis after digestion with *Eco*RI to cut out the insert revealed that this clone contained an insert of approx. 2.1 kbp. This clone was digested with *Eco*RI and subcloned into the *Eco*RI site of the pUC19 plasmid. Recombinant pUC19 plasmids were transformed into competent *E. coli* strain DH5α.

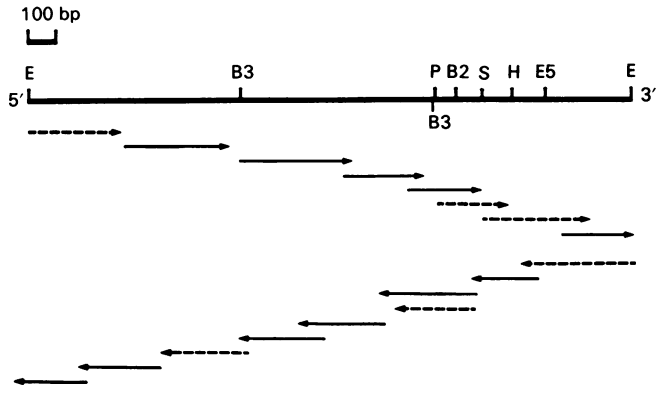


Fig. 1. Restriction-endonuclease map and sequencing strategy of the spinach AHRI cDNA

Arrows under the solid bar indicate the direction and extent of the sequencing reactions. Regions indicated by broken arrows were sequenced with the use of either the M13 universal primer or the M13 reverse primer. Unbroken arrows indicate regions that were sequenced with the use of several 18-mer primers complementary to the AHRI cDNA. Restriction sites: E, *Eco*RI; B3, *Bsp*MI; P, *Pst*I; B2, *Bg*lII; S, *Sph*I; H, *Hin*dIII; E5, *Eco*RV.

The restriction map of the full-length cDNA, together with the sequencing strategies employed, are shown in Fig. 1. The first ATG triplet (nucleotides 29–31) (Fig. 2) encountered downstream from the 5′-end is designated as the translation start site because the surrounding sequence (ATCAATGGC) agrees well with the favoured sequence (AACAATGGC) flanking the functional plant initiator codon (Lütcke *et al.*, 1987; Joshi, 1987). The synthesis of the AHRI polypeptide will begin therefore with the dipeptide methionylalanine, as do the majority of plant proteins. The insert contained an open reading frame of 1785 bp flanked at the 5′-end by 28 bases of non-coding sequence and at the 3′-end by 186 bases of non-coding sequence and a 51-base poly(A) tail. No typical eukaryotic poly(A) addition signal (AATAAA) (Proudfoot & Brownlee, 1976) was found after the stop codon. A Northern-blot analysis of mRNA for AHRI was carried out with the α-^{32}P-labelled cDNA as the probe. An mRNA of approx. 2100 bases in length was detected (results not shown). These results demonstrate, therefore that the cDNA obtained was full length.

Starting with the first in-frame methionine residue the open reading frame codes for a protein of 595 amino acid residues. The identity of the encoded protein was confirmed by comparison of the deduced amino acid sequence with the *N*-terminal sequence of the mature protein. The *N*-terminus of the mature protein corresponds to Val-73 in the deduced sequence, indicating that the protein is synthesized with a 72-amino acid-residue (M_r 6918) *N*-terminal presequence. The mature AHRI is predicted to consist of 523 amino acid residues, giving a protein of M_r 56853, which compares favourably with the M_r of 59000 for AHRI on SDS/PAGE. Finally, the derived amino acid sequence corresponds closely to amino acid composition data from the purified protein (results not shown).

The presequence peptide exhibits a typical enrichment in alanine (21%) and serine (24%) and contains no negatively charged residues (aspartic acid or glutamic acid) (Heijne *et al.*, 1989). In addition, the *N*-terminal region of the polypeptide is basic, a feature that has been noted previously for proteins that are transported into the chloroplasts, including acetolactate synthase from *Arabidopsis thaliana* and *Nicotiana tabacum* (Mazur *et al.*, 1987).

A comparison of the spinach, *E. coli* (see Wek & Hatfield, 1986) and *Saccharomyces cerevisiae* (see Petersen & Holmberg, 1986) AHRIs reveals a low overall degree of identity of 23% at the amino acid level. In addition, a region of divergence between the yeast AHRI and spinach AHRI proteins occurs at amino acid residues 330–470 of the spinach AHRI protein (results not shown). This stretch of amino acid residues is not found in the yeast AHRI protein and may be the result of gene modification (deletion or insertion) during evolution of higher-plant versus yeast AHRI. If one compares the most conserved regions of the AHRIs two major blocks of sequence similarity are shared by these three proteins. First, a region in the mature AHRIs is very similar to the 'fingerprint' structural motif in the NAD(P)H-binding domain of many enzymes: a $\beta\alpha\beta$-fold centred around a highly conserved sequence Gly-Xaa-Gly-Xaa-Xaa-Ala-(or Gly)-Xaa-Xaa-Xaa-Ala-(or Gly)- (where Xaa is any amino acid) that constitutes a tight turn at the end of the first strand of a β-sheet and marks the beginning of the succeeding α-helix (Scrutton *et al.*, 1990; Wierenga *et al.*, 1985) (Fig. 3). Secondly, a sequence

Fig. 2. Nucleotide and predicted amino acid sequences of the cDNA of spinach chloroplast AHRI

Nucleotide and amino acid residues are numbered on the right of the sequence. The translation start site for the reading frame coding for AHRI is underlined with a dashed line, and the corresponding stop codon is marked with an asterisk. The vertical arrow indicates the end of the presumptive transit peptide, and the *N*-terminal sequence of the mature polypeptide is underlined. Amino acids corresponding to the 'fingerprint' region of the NAD(P)H-binding site of oxydoreductases (see Fig. 3) are indicated by triangles.

Ser-His-Gly-Phe, which is highly conserved in AHRIs from *Saccharomyces cerevisiae and E. coli*, is present in AHRI from spinach at amino acid residues 225–228 (Fig. 4). Interestingly, in all the AHRI proteins there is a lysine residue approximately seven residues upstream of the conserved four-amino-acid-sequence containing histidine.

Southern-blot analysis

Southern-blot analysis was used to examine the number of genes encoding AHRI in *Spinacia oleracea*. Spinach genomic DNA digested with six restriction enzymes was fractionated on a 0.7% agarose gel and Southern blots were probed with an α-^{32}P-labelled cDNA fragment (*Eco*RV–*Eco*RI restriction fragment). The results shown in Fig. 5 reveal that in each digestion only a single genomic DNA restriction fragment hybridizes to this probe, even after a low-stringency wash. These results strongly suggest that spinach contains a single AHRI gene. The utilization of other probes (PCR1, *Eco*RI–*Pst*I and *Pst*I–*Hin*dIII) derived from the cDNA also indicates the presence of only one nuclear gene (results not shown).

The restriction map of this gene was determined by analysis of the Southern blots of digested spinach genomic DNA hybridized with various ^{32}P-labelled cDNA fragments (PCR1, *Eco*RI–*Pst*I, *Pst*1–*Hin*dIII and *Eco*RV–*Eco*RI). In spinach, AHRI is encoded by an interrupted nuclear gene of about 7.5 kbp containing at least four introns (Fig. 6).

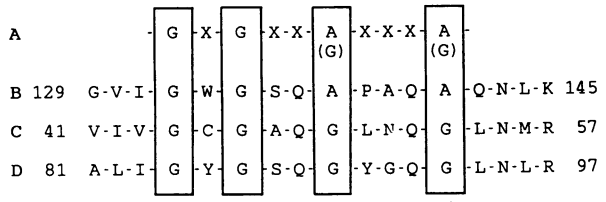


Fig. 3. Protein sequence similarities in AHRIs around the putative NAD(P)H-binding site

A, 'fingerprint' region of the NAD(P)H-binding site (Wierenga *et al.*, 1985); B, deduced sequence of spinach protein; C, deduced sequence of *E. coli* protein (Wek & Hatfield, 1986); D, deduced sequence of *Saccharomyces cerevisiae* protein (Petersen & Holmberg, 1986). The boxes indicate conserved amino acid residues.

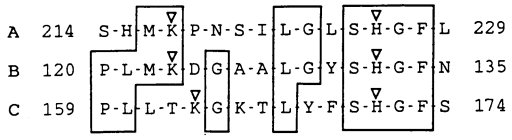


Fig. 4. Comparison of a highly conserved region in the deduced amino acid sequences of AHRIs from spinach, *E. coli* and *Saccharomyces cerevisiae*

The boxes indicate the conserved amino acids in the proteins from (A) spinach, (B) *E. coli* (Wek & Hatfield, 1986) and (C) *Saccharomyces cerevisiae* (Petersen & Holmberg, 1986). Triangles show amino acid residues that could be involved in the isomerization of 2-acetolactate and 2-aceto-2-hydroxybutyrate.

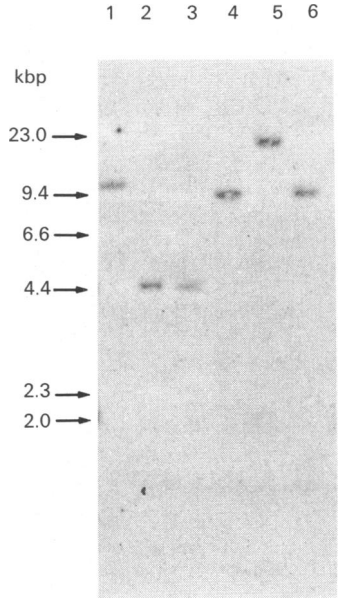


Fig. 5. Southern-blot hybridization of spinach genomic DNA

Portions (10 μg) of genomic DNA were digested with *Eco*RI (lane 1), *Bam*HI (lane 2), *Eco*RV (lane 3), *Hin*dIII (lane 4), *Bgl*II (lane 5) and *Pst*I (lane 6). Digested genomic DNA was fractionated on a 0.7% agarose gel, transferred to nitrocellulose and probed with an α-^{32}P-labelled *Eco*RV–*Eco*RI restriction fragment of the cDNA.

DISCUSSION

This is the first report of the isolation and characterization of a cDNA corresponding to the mRNA for plant AHRI. Confidence in its identity comes from the following considerations. Firstly, the predicted M_r of spinach mature AHRI is 56853, almost identical with that expected based on its mobility in SDS/7.5–15% PAGE. Secondly, the predicted *N*-terminal sequence matches the sequence derived from Edman degradation of the intact protein. Thirdly, the deduced sequence of the mature AHRI corresponds closely to the amino acid composition of the purified protein. The open reading frame encodes a precursor of 595 amino acid residues, the first 72 residues of which have been found to bear some of the hallmarks of a chloroplast transit peptide, notably an unusually high content of hydroxylated and basic residues and a deficiency in acidic residues. However, aside from the 'positive–hydrophobic–polar design' in primary structure no precise sequence conservation could be found in different transit peptides studied so far (Douce & Joyard, 1990). Such an observation strongly supports the idea that the tertiary structure of chloroplast protein precursors is of major importance in their interactions with envelope membrane receptors or with the lipid bilayer itself (Heijne, 1990). The AATAAA element signalling polyadenylation in higher eukaryotes was not found at the 3′-end of spinach AHRI cDNA. Interestingly, the 3′-terminal untranslated regions of a number of plant genes do not contain this sequence (Hunt *et al.*, 1987). It is therefore clear that the specific sequences required for 3′-terminal processing and polyadenylation of some plant RNAs have not yet been identified (Shirley *et al.*, 1990).

Southern-blot analysis indicated that the AHRI gene in spinach is most likely to be present in a single copy per haploid genome. Likewise, in *Saccharomyces cerevisiae* and *E. coli* AHRI is coded by a single gene. These findings therefore demonstrate that the previously reported presence of multiple forms of AHRI in spinach chloroplasts (Dumas *et al.*, 1989) is attributable to the expression of distinct allelic forms of one gene containing identical restriction sites, and therefore indistinguishable by Southern-blot experiments, or to post-transcriptional modifications of a single gene product. In fact it appeared that the forms do not differ either in their substrate and Mg^{2+} saturation curves or in the pattern of pH-dependence. If these forms were encoded by distinct structural genes with their own pattern of regulation, it might be expected that they would differ in their biochemical properties, including feedback inhibition by various branched-chain amino acids. For example, in *E. coli* and *Salmonella typhimurium* three acetolactate synthase isoenzymes that differ in their biochemical properties have been characterized (De Felice *et al.*, 1982; Schloss *et al.*, 1985; Gollop *et al.*, 1989), and the genes encoding each isoenzyme from *E. coli* have been cloned and sequenced (Wek *et al.*, 1985).

The results described in the present paper indicate that the spinach, yeast and bacterial AHRIs share little sequence similarity along their entire polypeptide chain, indicating their distant evolutionary relationship. Such a result is not surprising because in yeast AHRI is very likely to be located inside the mitochondria (Ryan & Kohlhaw, 1974) whereas in spinach it is present in chloroplasts (Dumas *et al.*, 1989). Regions that are highly conserved between spinach, *Saccharomyces cerevisiae* and *E. coli* may include important sites for enzyme activity. Of particular interest here is the finding of a the short sequence Ser-His-Gly-Phe that is highly conserved in all the AHRI proteins studied so far. In addition, in all eukaryotic and prokaryotic AHRIs examined there is a lysine residue approximately seven residues upstream of the conserved four-amino acid-residue sequence containing histidine. This region Lys-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Xaa-Ser-His-Gly-Phe is a candidate for the catalytic site of the reductoisomerase. Indeed, the first step of the AHRI reaction (alkyl migration) can be effected by both acidic and basic reagents (Armstrong & Lipscomb, 1983). In this case, the tertiary ketol rearrangement could be catalysed by both histidine (pK 6.5) and lysine (pK 10.2), as shown in Scheme 1. In support of this

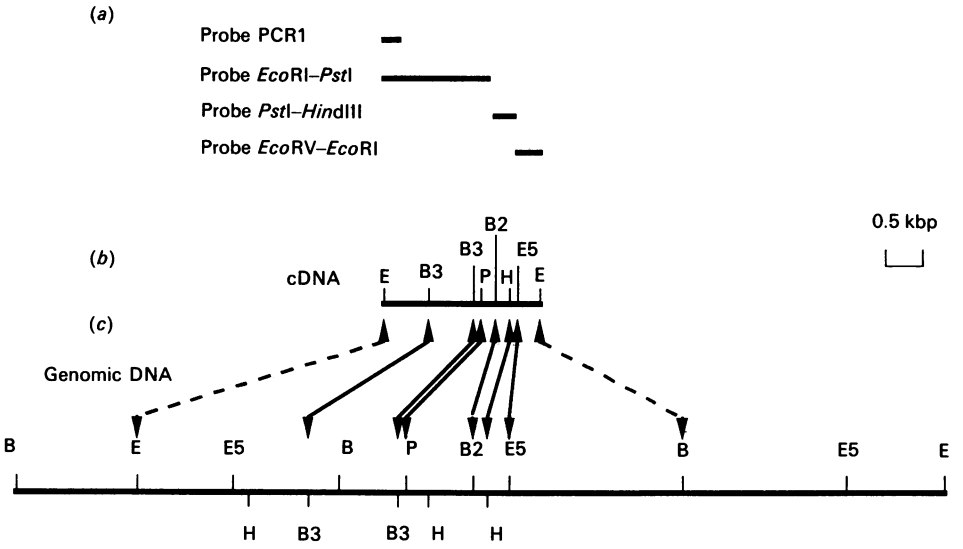
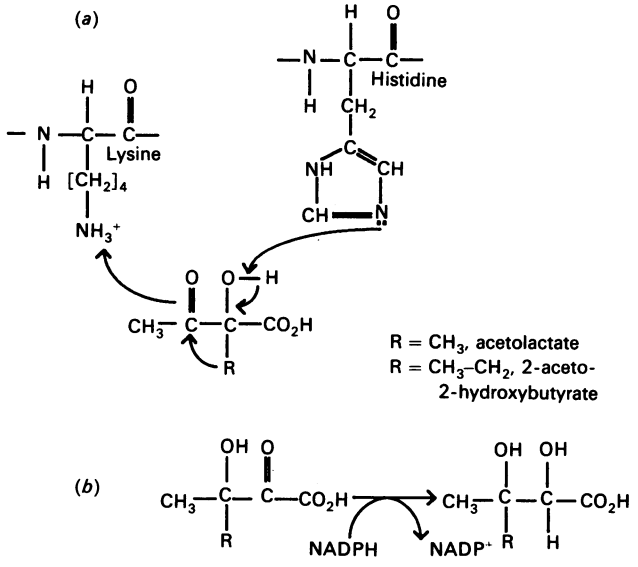


Fig. 6. Restriction map of the AHRI genomic DNA

(*a*) Probes used in the Southern-blot experiment. (*b*) cDNA. (*c*) Genomic DNA. Unbroken arrows indicate the corresponding restriction sites in the cDNA and the genomic DNA. Broken arrows show the ends of the cDNA and the corresponding hybridization sites in the genomic DNA. Restriction sites: E, *Eco*RI; B3, *Bsp*MI; P, *Pst*I; B2, *Bgl*II; H, *Hin*dIII; E5, *Eco*RV; B, *Bam*HI.



Scheme 1. Mechanism of the AHRI reaction

(*a*) First step of the reaction: the tertiary ketol rearrangement could be catalysed by both lysine (217) and histidine (226) of the spinach AHRI. (*b*) Second step of the reaction: NADPH-dependent reduction of the carbonyl residue.

suggestion we have previously shown that the AHRI reaction is effective in a broad pH range between the pK values of histidine and lysine (Dumas *et al.*, 1989). Although these theoretical considerations are consistent with the tertiary ketol re-arrangement model shown, other interpretations are possible. Crystallographic analyses of NAD(P)H-dependent oxidoreductases have shown that these proteins are functionally divisible into two domains, the NAD(P)H-binding domain and the catalytic domain (Tobey & Grant, 1986). The NAD(P)H-binding site found in other oxidoreductases (Scrutton *et al.*, 1990) is also present in spinach AHRI.

Furthermore, the predicted secondary structure for this sequence in spinach AHRI also matches the known $\beta\alpha\beta$-fold structure (Wieranga *et al.*, 1985) for NAD(P)H-binding domains (results not shown). These sequence characteristics of AHRI provide additional evidence that the identified open reading frame in the cDNA codes for AHRI and should aid in the structure–function analysis of AHRI and NAD(P)H-dependent oxidoreductases in general.

We thank Dr. Didier Faucher and Dr. François Clerc (Rhône–Poulenc Santé, Vitry, France) for amino sequence analysis, Dr. Gerard Leclerc (Institut Pasteur, Lyon, France) for the determination of the amino acid composition, Dr. Alfred Greiner and Jean-Yves Ortholand for providing us with 2-acetolactate and 2-aceto-2-hydroxybutyrate, and Dr. Mike May, Dr. John Lunn and Dr. Vincent Wingate for helpful discussions.

REFERENCES

Armstrong, F. B. & Lipscomb, E. L. (1983) J. Chem. Soc. Perkin Trans. 1 1197–1201
Aulabaugh, A. & Schloss, J. V. (1990) Biochemistry **29**, 2824–2830
Birnboim, H. C. & Doly, J. (1979) Nucleic Acids Res. **7**, 1513–1523
Burnette, W. R. (1981) Anal. Biochem. **112**, 195–203
Chou, P. Y. & Fasman, G. D. (1978) Annu. Rev. Biochem. **47**, 251–276
De Felice, M., Lago, C. T., Squires, C. H. & Calvo, J. M. (1982) Ann. Microbiol. **133**, 251–256
Douce, R. & Joyard, J. (1982) in Methods in Chloroplast Molecular Biology (Edelman, M., Hallick, R. B. & Chua, N. H., eds.), pp. 239–256, Elsevier Biomedical, Amsterdam
Douce, R. & Joyard, J. (1990). Annu. Rev. Cell Biol. **6**,173–216
Dumas, R., Joyard, J. & Douce, R. (1989) Biochem. J. **262**, 971–976
Gollop, N., Damri, B., Barak, Z. & Chipman, D. M. (1989) Biochemistry **28**, 6310–6317
Heijne, G. (1990) J. Membr. Biol. **115**, 195–201
Heijne, G., Steppuhn, J. & Herrmann, R. G. (1989) Eur. J. Biochem. **180**, 535–545
Hunt, A. G., Chu, N. M., Odell, J. T., Naggy, F. & Chua, N. H. (1987) Plant Mol. Biol. **8**, 23–35
Huynh, T. V., Young, R. A. & Davis, R. W. (1985) in DNA Cloning: A Practical Approach (Glover, D. M., ed.), vol. 1, pp. 49–78, IRL Press, Oxford
Joshi, C. P. (1987) Nucleic Acids Res. **15**, 9627–9640

Kingston, R. E. (1987) in Current Protocols in Molecular Biology (Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Smith, J. E., Seidman, J. G. & Struhl, K., eds.), pp. 4.5.1–4.5.3, Wiley–Interscience, New York
Lütcke, H. A., Chow, K. C., Mickel, F. S., Moss, K. A., Kern, H. F. & Scheele, G. A. (1987) EMBO J. **6**, 43–48
Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor
Mazur, B. J., Chui, C. F. & Smith, J. K. (1987) Plant Physiol. **85**, 1110–1117
Petersen, J. G. L. & Holmberg, S. (1986) Nucleic Acids Res. **14**, 9631–9651
Proudfoot, N. J. & Brownlee, G. G. (1976) Nature (London) **263**, 211–214
Ryan, E. D. & Kohlhaw, G. B. (1974) J. Bacteriol. **120**, 631–637
Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuki, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) Science **239**, 487–491
Saint-Blancard, J., Foucart, J., Limonne, F., Girot, P. & Boschetti, E. (1981) Ann. Pharm. Fr. **39**, 403–409
Sanger, F., Nicklen, S. & Coulson, A. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463–5467
Schloss, J. V., Van Dyk, D. E., Vasta, J. F. & Kutny, R. M. (1985) Biochemistry **24**, 4952–4959
Schulz, A., Spönemann, P., Köcher, H. & Wengenmayer, F. (1988) FEBS Lett. **238**, 375–378
Scrutton, N. S., Berry, A. & Perham, R. N. (1990) Nature (London) **343**, 38–43
Shirley, B. W., Ham, D. P., Senecoff, J. F., Berry-Lowe, S. L., Zurflub, L. L., Shah, D. M. & Meagher, R. B. (1990) Plant Mol. Biol. **14**, 909–925
Shure, M., Wessler, S. & Fedoroff, N. (1983) Cell **35**, 225–233
Smith, T. F. & Waterman, M. S. (1981) Adv. Appl. Math. **2**, 482–489
Tobey, K. L. & Grant, G. A. (1986) J. Biol. Chem. **261**, 12179–12183
Wek, R. C. & Hatfield, G. W. (1986) J. Biol. Chem. **261**, 2441–2450
Wek, R. C., Hauser, C. A. & Hatfield, (1985) Nucleic Acids Res. **13**, 3995–4010
Wierenga, R. K., De Maeyer, M. C. H. & Hol, W. G. (1985) Biochemistry **24**, 1346–1357

Received 23 October 1990/18 December 1990; accepted 5 March 1991