Characterization of Plastid 5-Aminolevulinate Dehydratase (ALAD; EC 4.2.1.24) from Spinach (*Spinacia oleracea* L.) by Sequencing and Comparison with Non-Plant ALAD Enzymes*

Anke Schaumburg and Hansjörg A. W. Schneider-Poetsch

Botanisches Institut der Universität zu Köln, Gyrhofstraße 15, D-W-5000 Köln 41, Bundesrepublik Deutschland

in cooperation with

C. Eckerskorn

Max-Planck-Institut für Biochemie, Genzentrum, D-W-8033 Martinsried, Bundesrepublik Deutschland

Z. Naturforsch. **47c**, 77–84 (1992); received September 16/November 4, 1991

Plastid 5-Aminolevulinate Dehydratase (EC 4.2.1.24), Primary Structure Peculiarities, Transit Peptide, Mg^{2+}-Binding, Active Centre

We have sequenced 5-aminolevulinate dehydratase (ALAD; EC 2.4.1.24) of a plant. A full-length cDNA clone (1727 bp) encoding this enzyme has been identified by immunoscreening a lambda gt11 cDNA library of spinach. ALAD is not a plant-specific enzyme; however, the plant enzyme differs from the well known ALAD enzymes of bacteria, yeast and animals in structural and biochemical properties and in that it is located in the plastid. Differences and homologies can be traced back to the molecular level. The mature ALAD subunit, whose N-terminus was determined by automatic Edman degradation, is a protein of 367 amino acid residues and has a M_r of 40,132. This figure is in the range of molecular weights of non-plant ALADs. The active centre is highly conserved and the same is true for the ion-binding domain, except that 4 cysteines of the non-plant enzymes (binding Zn^{2+}) have disappeared and a total of 6 aspartic acids meets the demands of Mg^{2+}-binding. However, there are more distinct differences. Apart from a transit sequence of 56 amino acids targeting the plastid, the N-terminal part of the mature plant enzyme differs considerably from non-plant ALAD enzymes. It is rich in prolines and hydroxylated amino acids. The apparent M_r on SDS-PAGE is 45,000 or higher, but up to now posttranslational modifications have not been found.

Introduction

Porphyrin biosynthesis in plants differs from porphyrin biosynthesis in animals and bacteria by the spectrum of end products, the location, the mode of ALA (5-aminolevulinate) synthesis and the possibility of enzyme induction by light.

Although experimental data cannot wholly exclude that in plants some porphyrins may be synthesized in the cytoplasm and the mitochondria, there is much evidence that plant porphyrins are synthesized by enzymes which are located in the plastids and encoded in the nucleus [1–4]. It is mainly the capacity for the synthesis of ALA which is promoted by light. But experiments with tissues which do not contain etioplasts but proplastids show that also the synthesis of ensuing enzymes [5] is dependent on light. Illumination increases the amount of chlorophyll as well as the quantity of these enzymes in the plastids. The down-regulation by light of the NADPH-protochlorophyllide oxido-reductase (EC 1.6.99.–) is an interesting exception (see [6]).

Although porphyrin biosynthesis is a prerequisite for photosynthesis (chlorophylls and porphyrins), characterization of the involved enzymes lags behind that of the corresponding enzymes of non-plants or of proteins taking part in photosynthetic processes. In recent years, however, NADPH-protochlorophyllide oxido-reductase (*Hordeum* [7]), glutamate-1-semialdehyde aminotransferase (EC 5.4.3.8; *Hordeum* [8]) and porphobilinogen deaminase (EC 4.3.1.8; *Euglena* [9]) have been characterized on the molecular level. As another member of this pathway, we have formerly biochemically and immunologically characterized the enzyme in between the two latter enzymes, 5-aminolevulinate dehydratase (ALAD) from

* The sequence of spinach ALAD is available from the EMBL (European Molecular Biology Laboratory) data library (accession number X 57842).

Reprint requests to Hj. Schneider-Poetsch.

Verlag der Zeitschrift für Naturforschung, D-7400 Tübingen
0939–5075/92/0100–0077 $ 01.30/0

spinach [1, 10–12] synthesizing porphobilinogen from two molecules of ALA.

Here we report on a cDNA clone from spinach that codes for pre-ALAD. In some regions of the enzyme, the sequence data reveal strong homologies with animal, fungal and bacterial ALAD enzymes [13–17]. But differences exist particularly with respect to the site assumed to be involved in ion-binding, the N-terminal region of the mature enzyme and the existence of a transit sequence.

Materials and Methods

Cultivation of spinach

Spinacia oleracea L. (cultivar. Aestivato) was grown in the greenhouse of the Botany Institute. RNA was extracted from leaves of plants grown for about 4 weeks.

Determination of the N-terminal amino acid sequence

ALAD was immunopurified (see [10]) and electrophoresed (12% SDS-PAGE) according to standard procedures [18]. Thereafter, the protein(s) were electro-blotted onto siliconized glass fiber sheets (Glassybond, Biometra, Göttingen), detected by Coomassie Blue staining, excised and applied to a gas-phase sequencer (Type 470 A, Applied Biosystems, [19]).

Isolation of poly(A)$^+$-mRNA

Total RNA was isolated by extraction with "hot phenol". After shock-freezing in liquid nitrogen, 10 g of leaves were powdered in a precooled mortar and resuspended in 20 ml of 70 °C hot isolation buffer (50 mM Tris/HCl, 10 mM EDTA, 2% SDS). Proteinase K (0.2 ml of a 20 mg/ml solution) was immediately added. The mixture was stirred at 50 °C for 10 min and then transferred to a 100 ml screwcap flask containing 20 ml of 60 °C hot phenol/chloroform/isoamylalcohol (25:24:1; see [20]). The mixture was shaken vigorously for 10 min at 60 °C; the two phases were separated by centrifugation and the procedure was repeated with the aqueous phase. A third extraction was performed with chloroform/isoamylalcohol (24:1) alone at room temperature. The aqueous phase was adjusted to a NaCl concentration of 0.5 M. The absorption on oligo-dT cellulose and elution of poly(A)$^+$-mRNA followed standard procedures (see [20]). The yield of poly(A)$^+$-mRNA ($OD_{260/280}$ >1.9) was *ca.* 60 µg. An aliquot of 5 µg was used for the synthesis of cDNA.

Preparation of the cDNA-library

cDNA ligated to Eco RI adaptors was synthesized according to the protocol provided with the cDNA synthesis kit of Pharmacia (Freiburg, Germany). The cDNA was cloned into the expression vector lambda gt11 (GIBCO/BRL, Eggenstein, Germany, No. 5272 SA) and packaged (Gigapack II Gold kit, No. 200216; Stratagene, Heidelberg, Germany). The resulting library consisted of about 5×10^{-5} independent recombinant clones.

Screening for clones expressing 5-ALAD fusion proteins

A mixture of five monoclonal antibodies (7–1, 34–1, 210–5, 255–8, 288–2) formerly raised against spinach ALAD [11] was used to screen the expression library. The detection of positive clones followed the protocol provided by Amersham-Buchler (Braunschweig, Germany; No. RPN 1281 Y). Several antibody-positive clones were determined. The clone with the longest insert was further characterized. Because of Eco RI sites, obviously spoilt during cloning, the insert could not be cut out by Eco RI; instead, Kpn I and Sac I were chosen. The insert was isolated with parts of lambda gt11 on either side.

Subcloning and sequencing

Restriction analysis of the isolated cDNA insert revealed an internal restriction site for Sal I, which was used to reclone the resulting two fragments into pUC 19 (Kpn I–Sal I and Sac I–Sal I). Transformation with *E. coli* JM 101 as host was performed as described [21]. The preparation of pUC 19 for sequencing followed standard protocols [20]. Using primer, reverse primer, device and protocol of Pharmacia, the first sequence data of the inserts became available. Then, ALAD-specific primers were synthesized (see below) and the residual sequences determined. S-35-ThioATP was obtained from Amersham-Buchler (SJ 1304).

Synthesis of primers

Synthetic oligonucleotides 12 to 14 bases long were synthesized on a Pharmacia Gene Assembler

Plus using reagents and protocol of the manufacturer. The concentration of the primer was adjusted to 0.8 μM before annealing. In contrast to the given protocol, three times more primer was used.

Sequence comparisons

Amino acid homology searches were performed by FASTA (program of [22]) provided by EMBL, Heidelberg.

For some specific comparisons, Hydrophobic Cluster Analysis [23] was brought in. According to this method, a sequence is arranged (notwithstanding the natural existing conformation) α-helical. The helix is then cut parallel to the axis of the helix cylinder and unrolled. This representation is duplicated in order to make the sequence easier to follow and to give a better impression of the environment of each amino acid. Isoleucine, L, F, W, M, Y, V are considered as hydrophobic and are highlighted. Although the regions analyzed by HCA can scarcely be assumed to be naturally arranged α-helical [24], we used this method because it advantageously brings sequence peculiarities into prominence.

Results

We immunoscreened an expression library constructed in lambda gt11 from poly(A)$^+$-mRNA of young green spinach leaves. Several antibody-positive clones were identified and isolated. In a first attempt, the clone apparently containing the largest insert was subcloned into pUC 19 and sequenced.

Nucleotide sequence and primary structure of spinach ALAD

The isolated insert consists of 1727 base pairs (Fig. 1). According to general translation start conditions (see [25]), translation of the ALAD gene in the plant starts with the clone's nucleotide 123. Translation is terminated at position 1391. Polyadenylation signals (AATAAA and ATTAAA) are contained in the 3′-untranslated region, but the insert ends without reaching the polyadenylated region.

According to automated Edman degradation of immunopurified spinach ALAD, the N-terminal part of the enzyme starts with the amino acids xERRDNLTQQKTGL. This sequence corresponds to the amino acids 57 to 70 of the sequence deduced from the nucleotide sequence (Fig. 1). Hence, the mature enzyme consists of 367 amino acids (1101 bp) and the preceding 56 amino acids constitute the transit sequence split after the enzyme's transport into the plastid. A minor fraction of ALAD running on SDS-PAGE sporadically ahead of the main fraction [10] starts with the sequence AAVVAGNAPSAPP. In case of slow purification procedures, this fraction becomes prominent (Fig. 4).

The transit sequence of spinach ALAD

The essential structural requirements constituting transit sequences are not known. However, some details common to peptides allowing the transport of proteins into the chloroplast [26] are also found in the transit peptide of spinach ALAD: *i.e.* a net positive charge, absence of tryptophan and tyrosine, occurrence of proline, relative abundance in hydroxylated amino acids. The first charged amino acid in the ALAD transit peptide is lysine at position 15.

General similarities and differences between spinach ALAD and non-plant ALAD enzymes

The predicted primary structure of mature spinach ALAD shows a high degree of homology with non-plant ALAD enzymes (Fig. 2). However, a N-terminal overhang, two extensions within the N-terminal part and two gaps later distinguish the plant enzyme from ALADs from *E. coli* [17], yeast [16], man [13, 14] and rat [15]. In the arrangement of Fig. 2, homologies become visible downstream from position 37 of spinach ALAD. With a slight lead over other ALADs, rat ALAD appears to be most similar to the plant enzyme. As referred to the whole spinach sequence, 36.2 per cent of the positions of spinach and rat ALAD are occupied by identical amino acids. The figure is 41.8 per cent if the overhang and the two extensions at the beginning of the spinach ALAD sequence are omitted.

A N-terminal region of spinach ALAD (positions 21 to 66) which shows little coincidences with the non-plant ALADs, is remarkable for its high proline content. In a HCA (Hydrophobic Cluster Analysis, [23]) presentation all prolines appeare

```
        10         20         30         40         50         60
CTCCATTTTCCACTCAATTTGTTCTTCACTTCTGCAAATCATAGCAAAAAATCCCATCTT
  p  f  s  t  q  f  v  l  h  f  c  k  s  *y q  k  i  p  s  s
        70         80         90        100        110        120
CAAAACCCCAATTTAAATTAATCTCCGCTTTTAGGGGTTTTGGGACGATAAGATTTGGGA
  k  p  q  f  k  l  i  s  a  f  r  g  f  g  t  i  r  f  g  m
       130        140        150        160        170        180
TGATGGCATCAACATTTAACATCCCCTGTAATGCTGGTACGATAAAGAATTTCAACAATT
  M  A  S  T  F  N  I  P  C  N  A  G  T  I  K  N  F  N  N  S
       190        200        210        220        230        240
CTCAGAGGAATTTAGGGTTCAGCTCAAATTTGGGGATTAATTTCGCGAAAACTAGGTTTT
  Q  R  N  L  G  F  S  S  N  L  G  I  N  F  A  K  T  R  F  S
       250        260        270        280        290        300
CGAATTGCGGAGATTCTGGCCGGATTCCATCGCAATTGGTGGTGAGGGCAAGTGAGAGAC
  N  C  G  D  S  G  R  I  P  S  Q  L  V  V  R  A  S  E  R  R
       310        320        330        340        350        360
GAGATAATTTGACCCAACAGAAGACAGGGTTGAGCATTGAAGAATGCGAAGCTGCTGTTG
  D  N  L  T  Q  Q  K  T  G  L  S  I  E  E  C  E  A  A  V  V
       370        380        390        400        410        420
TGGCTGGGAATGCACCCTCTGCTCCTCCTGTTCCTCCCACACCTAAAGCACCTTCTGGAA
  A  G  N  A  P  S  A  P  P  V  P  P  T  P  K  A  P  S  G  T
       430        440        450        460        470        480
CACCGTCTGTCTCTCCCCTTTCATTGGGTCGACGTCCACGTCGTAACCGGACATCACCTG
  P  S  V  S  P  L  S  L  G  R  R  P  R  R  N  R  T  S  P  V
       490        500        510        520        530        540
TATTTAGAGCAGCATTTCAAGAGACAACCTTGTCCCCTGCAAATGTGGTTTACCCATTGT
  F  R  A  A  F  Q  E  T  T  L  S  P  A  N  V  V  Y  P  L  F
       550        560        570        580        590        600
TCATTCATGAGGGAGAAGAGGACACGCCTATTGGCGCAATGCCTGGATGTTATAGACTTG
  I  H  E  G  E  E  D  T  P  I  G  A  M  P  G  C  Y  R  L  G
       610        620        630        640        650        660
GGTGGAGGCATGGCCTTGTAGAAGAGGTAGCAAAAGCACGTGATGTTGTTGTTAATAGCA
  W  R  H  G  L  V  E  E  V  A  K  A  R  D  V  V  V  N  S  I
       670        680        690        700        710        720
TTGTTGTTTTTCCAAAACCTGATGCTTTGAAGTCACCAACCGGTGATGAAGCATACAATG
  V  V  F  P  K  P  D  A  L  K  S  P  T  G  D  E  A  Y  N  E
       730        740        750        760        770        780
AAAATGGTTTAGTTCCCCGGACTATACGAATGCTGAAGGACAAATTTCCTGATCTTATAA
  N  G  L  V  P  R  T  I  R  M  L  K  D  K  F  P  D  L  I  I
       790        800        810        820        830        840
TTTACACGGATGTTGCCTTAGATCCATATTATTATGATGGGCATGATGGTATCGTTACAC
  Y  T  D  V  A  L  D  P  Y  Y  Y  D  G  H  D  G  I  V  T  Q
       850        860        870        880        890        900
AACATGGTGTAATTATGAATGATGAGACAGTGCACCAATTATGCAAACAAGCTGTAGCCC
  H  G  V  I  M  N  D  E  T  V  H  Q  L  C  K  Q  A  V  A  Q
       910        920        930        940        950        960
AGGCCCGTGCTGGAGCGGATGTTGTCAGTCCTAGTGATATGATGGATGGTCGTGTAGGTG
  A  R  A  G  A  D  V  V  S  P  S  D  M  M  D  G  R  V  G  A
       970        980        990       1000       1010       1020
CAATTCGTGCAGCGCTTGATGCAGAAGGATATTCAAATGTATCGATCATGTCCTATACGG
  I  R  A  A  L  D  A  E  G  Y  S  N  V  S  I  M  S  Y  T  A
      1030       1040       1050       1060       1070       1080
CAAAGTATGCAAGTTCATTCTATCCACGCTTTGGGGACAAGAAAACTTATCAAATGAACC
  K  Y  A  S  S  F  Y  P  R  F  G  D  K  K  T  Y  Q  M  N  P
      1090       1100       1110       1120       1130       1140
CAGCAAATTATAGAGAGGCTTTGATTGAAACTCAAGAAGATGAGTCGGAAGGAGCTGATA
  A  N  Y  R  E  A  L  I  E  T  Q  E  D  E  S  E  G  A  D  I
      1150       1160       1170       1180       1190       1200
TCCTATTGGTAAAACCCGGACTTCCATACTTGGACATTATCAGGCTTCTTCGGGACAACT
  L  L  V  K  P  G  L  P  Y  L  D  I  I  R  L  L  R  D  N  S
      1210       1220       1230       1240       1250       1260
CAGATTTGCCTATTGCTGCATATCAGGTTTCTGGTGAGTACTCCATGATAAAGGCGGGCG
  D  L  P  I  A  A  Y  Q  V  S  G  E  Y  S  M  I  K  A  G  G
      1270       1280       1290       1300       1310       1320
GGGTTCTGAAAATGATTGATGAGGAGAAGGTCATGTTAGAGTCGCTTTTGTGTCTGCGTC
  V  L  K  M  I  D  E  E  K  V  M  L  E  S  L  L  C  L  R  R
      1330       1340       1350       1360       1370       1380
GAGCTGGTGCCGACATCATTCTTACATATTTTGCTTTACAAGCTGCAAGATGTTTGTGTG
  A  G  A  D  I  I  L  T  Y  F  A  L  Q  A  A  R  C  L  C  G
      1390       1400       1410       1420       1430       1440
GAGAGAAGAGATGATTCCCTTTTGGGGTCACCATGGATTACATGATTGGAGATGAGCAGT
  E  K  R  *
      1450       1460       1470       1480       1490       1500
GAAATTTGCAGTTGTTCTCTGTTGTTCTTAGACGAGTGTTTAGCTAGTTTTGTCTAGTTT

      1510       1520       1530       1540       1550       1560
TTTACCGCCTTAATTAGTTCTCGAGAAGAACATTAGGGGAATTAGGTTAAAAATAGGAAG

      1570       1580       1590       1600       1610       1620
AGCTTTTACGCAATTATATTCCTAATATTGTACCAAGAGGTCTTTGCACGTACCATGTGT

      1630       1640       1650       1660       1670       1680
ACAATAAATTTACTGTAGAGGGGCTTGTAAGTTGCATCACCCCCTTGCCGCTGTATTTGT

      1690       1700       1710       1720
AAATGTCTAATGGTAATATATGATTAAAAGTTCTGCTAGGTTGTTTT
```

Fig. 1. Nucleotide sequence of the isolated cDNA clone and the deduced primary structure of spinach ALAD. The sequence is numbered starting with the first 5′-nucleotide of the cDNA. Polyadenylation signals (AATAAA, ATTAAA) are underlined. The sequence of mature ALAD whose N-terminus was determined by peptide sequencing is drawn in bold letters. The transit sequence of ALAD is underlined. A double line marks the peptide at the beginning of mature ALAD which was determined by Edman degradation and a second peptide starting a proteolytic degraded mature ALAD. Amino acids only present in the fusion protein of lambda gt11 are shown in small letters (*y: TAG, in the sup F host translated as thyrosine).

Fig. 2. Sequence comparison of spinach ALAD and ALAD enzymes of man [13, 14], rat [15], yeast [16] and *E. coli* [17]. The numbering starts with the first amino acid of mature spinach ALAD and includes gaps that are demanded by the alignment. Asterisks (*) indicate amino acids of spinach ALAD finding identical counterparts in any of the other sequences.

close together and lined up (Fig. 3A). In addition, this domain is characterized by an unusual high content of alanines, glycines, hydroxylated and basic amino acids. It partly reminds of a transit sequence.

Homologies in regions of known biological function

The active centre of the enzyme, localized by homology, lies within the most conserved region. Compared with the rat enzyme, spinach ALAD is identical at 15 out of 16 positions. This region includes the invariant lysine residue (position 334) involved in Schiff base formation with ALA [12, 27]. An arginine identified by biochemical modification studies of spinach ALAD [12] and assumed to be essential for catalysis is near by (position 344). It is conserved in other ALADs except the rat enzyme (see also Fig. 3C). Invariant arginines, however, are found at the positions 106, 313 and 392. Exchanges of amino acids in the immediate environs of the active centre are often conservative ones.

The latter applies also to the ion-binding region (vicinity of residue 205). However, the different ion-requirements of plant (Mg^{2+}) and non-plant ALADs (Zn^{2+}) are well reflected by the primary structures. The four cysteine residues which are thought to be involved in Zn^{2+} binding have all disappeared. Instead, five more aspartic acids than in non-plant ALADs are present in spinach ALAD (Fig. 3B).

The molecular weight of spinach ALAD

According to our sequence data, the molecular weight of spinach ALAD is 40,132. This M_r is largely in accordance with the molecular weights

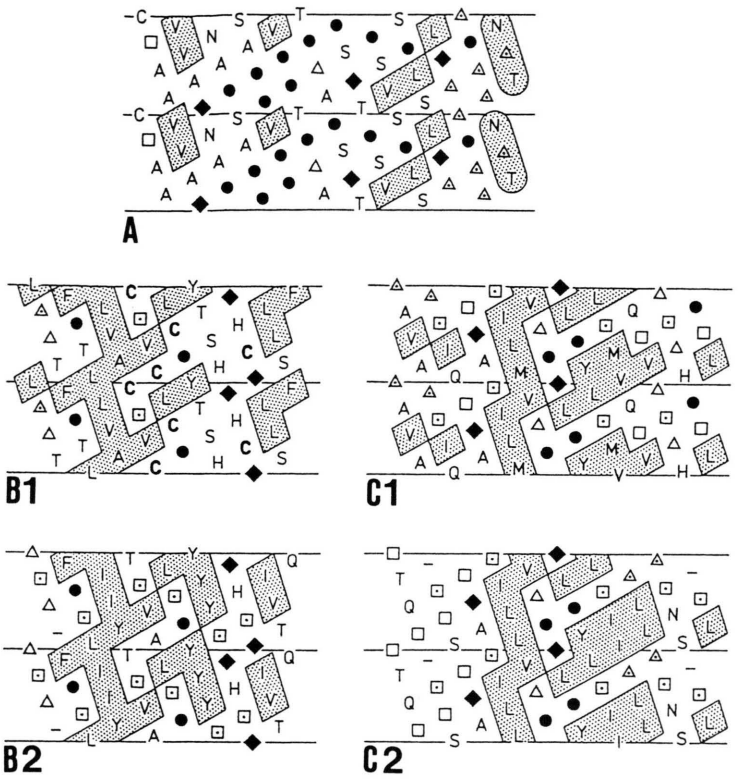

Fig. 3. Alpha-helical arrangement (doubled, see Methods) of prominent sequences of ALAD (Hydrophobic Clusters Analysis according to [23]). The amino acids proline (●), glycine (◆), glutamic acid (□), aspartic acid (⊡), lysine (△), and arginine (◬) are represented by symbols. Hydrophobic amino acids are enclosed in boxes. The numbering of positions refers to Fig. 2.
A: Cluster representation of the proline-rich region at the N-terminus of spinach ALAD (position 19–64). A possible glycosylation site [N T] is highlighted.
B: Comparison of the Zn^{2+} binding site of 1) rat ALAD (position 186–215) with the corresponding region of 2) spinach ALAD (position 165–215). The functional cysteines are in bold letters.
C: Comparison of the active centre of 1) rat ALAD (position 318–353) with the homologous region of 2) spinach ALAD (position 318–353).

of non-plant ALADs, but conflicting with apparent molecular weights of 45,000 or even higher which have been determined by SDS-PAGE ([10], Fig. 4). Up to now, standard chemical and enzymatic procedures failed to demonstrate any post-translational modification (glycosylation or ubiquitination). Although glycosylation was indicated by the DIG Glycan Detection Kit (Boehringer, Mannheim, No. 1142372), a particular saccharide could not be determined.

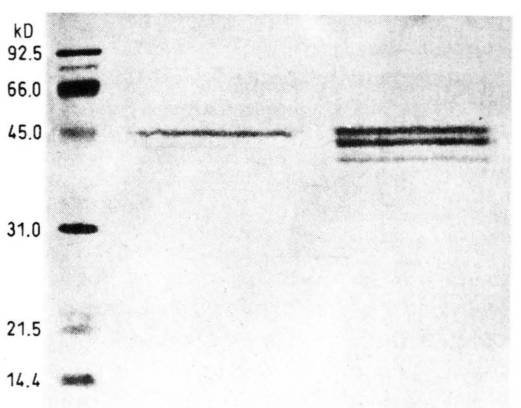

Fig. 4. Determination of the subunit molecular weight of ALAD by SDS-PAGE. Left panel: mature ALAD; right panel: mature ALAD and proteolytic degradation products. The two heavy bands were subjected to micro-sequencing and resulted in the sequences xERRDNLTQQKTGL... and AAVVAGNAPSAPP... (see Fig. 1).

Discussion

ALAD differs from most of the plastid enzymes in that it is not plant-specific. Corresponding enzymes are found in bacteria, fungi and animals. However, the maintenance of enzyme function within different locations and cellular environ-

ments during evolution has required altered structures. So, ALAD from spinach shows details which are specific to this enzyme and sometimes unexpected and other details which are common, *i.e.* conserved.

For example, the subunit M_r of spinach ALAD has proved to be very unexpected. It is plainly less than that determined by SDS-PAGE and in the M_r range of non-plant ALADs. According to its amino acid composition, the M_r of spinach ALAD amounts to 40,132 kg/mol, but on SDS gels, the enzyme behaves like a protein with a M_r around 45,000 or even higher. However, up to now efforts to demonstrate posttranslational modifications have failed.

Likewise, the high degree of homology of spinach ALAD with non-plant ALAD enzymes has been unexpected; it had not been indicated by available immunological data or data on ion-requirements and subunit composition [11, 12]. Rat and spinach ALAD have about 42 per cent identical amino acids at identical positions. In addition, conservative exchanges of amino acids are frequent in the parts interrupting homologous regions.

It is mainly the N-terminal parts of spinach ALAD which are different. There is an overhang and two extensions. Two greater gaps occur in the more homologous regions. The high content of prolines and hydroxylated amino acids of the N-terminal region is remarkable and may be characteristic for plant ALAD enzymes and their final location.

A sequence of 56 amino acids preceding the mature enzyme corroborates the finding [4] that ALAD is one of the chlorophyll and porphyrin biosynthetic enzymes which are all located in the plastid and are encoded in the nucleus. At its beginning, the transit sequence exhibits a motive which seems to be common to plastid enzymes of nucleus-cytoplasm provenance (see *e.g.* [26, 28–30]): a stretch of uncharged amino acids followed by positively charged ones. Other motives spread in transit sequences are also present, but it is not known as to whether they describe essential requirements. The transit sequence of spinach ALAD, however, shares in addition the amino acid sequence SxxxxSxxVVRA C-terminally contained in the transit peptide of barley glutamate-1-semialdehyd aminotransferase [8].

As far as general features can be extracted from transit sequences, which determine different final locations, transit sequences of proteins transported into mitochondria [30] start earlier with positively charged amino acids. The only known transit sequence of an enzyme located in glyoxysomes [31] has a negative charge close to its N-terminus.

A fraction of spinach ALAD preparations with a M_r slightly below that of the mature enzyme may originate from the action of a protease different from the transit peptidase. The homologies shared by all known ALAD enzymes (see also our ALAD sequence of the fern *Selaginella,* EMBL accession # 61652) make it highly improbable that a hypothesized, non-plastid located plant ALAD enzyme is more dissimilar than the known ALAD enzymes. Low stringency hybridization of ALAD cDNA with cut genomic DNA should reveal such an enzyme's gene.

One of the highly conserved regions mentioned above is the region ascribed to the active centre. The position of a lysine assumed to form the Schiff base with the substrate ALA [12, 27] as well as the positions of many other amino acids are conserved. In the vicinity of the lysine the sequence also contains an arginine. In previous enzyme modification experiments [12] the inhibition of ALAD activity by butanedione that blocks arginine had indicated an arginine taking part in substrate conversion.

In contrast, highly interesting alterations have occurred in the ion-binding centre. None of the 4 cysteines assumed to be involved in Zn^{2+}-binding is present in the Mg^{2+}-requiring plant enzyme. The disappearance of cysteine in the ion-binding centre matches the finding that spinach ALAD is not blocked by SH-reagents [12]. Instead of cysteines, 5 more aspartic acids (a total of 6) are present.It seems very likely that in this centre Mg^{2+} has been substituted for Zn^{2+}, the more so as Mg^{2+} and Zn^{2+} have almost the same ion radii. The coordination number of Zn^{2+} is 4 or 6, that of Mg^{2+} is regularly 6 rather than 4. Models should show as to whether the steric arrangement can satisfy coordination requirements. In order to enable comparisons with the K^+-dependent ALAD from *Rhodopseudomonas sphaeroides* [32], sequence data of this enzyme would be welcome.

Little information as to light regulation of ALAD can be drawn from the sequence at pres-

ence. Motives described for light regulated proteins [33] are mostly located hundreds of base pairs more upstream than present in cDNAs. This region should be addressed after probing genomic DNA with the cDNA clone. Since the primary structures of three light-inducible enzymes at the beginning of the plant porphyrin biosynthesis chain (Glutamate-1-semialdehyd aminotransferase [8], ALAD, porphobilinogen deaminase [9]) are known, a comparative study of their promoters might add to the understanding of light-regulated porphyrin biosynthesis in plants.

Acknowledgements

This study was made possible by a generous grant of the Volkswagen Stiftung to Hj. S.-P.

[1] H. A. W. Schneider, Z. Pflanzenphysiol. **62,** 328–342 (1970).
[2] H. A. W. Schneider, in: Pigments in Plants (F.-C. Czygan, ed.), pp. 227–308, Stuttgart 1980.
[3] A. G. Smith, Biochem. J. **249,** 423–428 (1988).
[4] W. W. Beisenherz and H. A. W. Schneider, Ber. Deutsch. Bot. Ges. **87,** 161–166 (1974).
[5] H. A. W. Schneider, Enzyme **16,** 108–114 (1973).
[6] I. Häuser, K. Dehesh, and K. Apel, Planta **170,** 453–460 (1987).
[7] R. Schulz, K. Steinmüller, M. Klaas, C. Forreiter, S. Rasmussen, C. Hiller, and K. Apel, Mol. Gen. Genet. **217,** 355–361 (1989).
[8] B. Grimm, Proc. Natl. Acad. Sci. U.S.A. **87,** 4169–4173 (1990).
[9] A. L. Sharif, A. G. Smith, and C. Abell, Eur. J. Biochem. **184,** 353–359 (1989).
[10] W. Liedgens, R. Grützmann, and H. A. W. Schneider, Z. Naturforsch. **35c,** 958–962 (1980).
[11] H. A. W. Schneider and W. Liedgens, Z. Naturforsch. **36c,** 44–50 (1981).
[12] W. Liedgens, C. Lütz, and H. A. W. Schneider, Eur. J. Biochem. **135,** 75–79 (1983).
[13] J. G. Wetmur, D. F. Bishop, L. Ostasiewicz, and R. J. Desnick, Gene **43,** 123–130 (1986).
[14] J. G. Wetmur, D. F. Bishop, C. Cantelmo, and R. J. Desnick, Proc. Natl. Acad. Sci. U.S.A. **83,** 7703–7707 (1986).
[15] T. R. Bishop, P. J. Cohen, S. H. Boyer, A. N. Noyes, and L. P. Frelin, Proc. Natl. Acad. Sci. U.S.A. **83,** 5568–5572 (1986).
[16] A. M. Myers, M. D. Crivellone, T. J. Koerner, and A. Tzagoloff, J. Biol. Chem. **262,** 16822–16829 (1987).
[17] J. M. Li, C. S. Russell, and S. D. Cosloy, Gene **75,** 177–184 (1989).
[18] U. K. Laemmli, Nature **227,** 680–685 (1979).
[19] C. Eckerskorn, W. Mewes, H. Goretzki, and F. Lottspeich, Eur. J. Biochem. **176,** 509–519 (1988).
[20] J. Sambrook, E. F. Fritsch, and T. Maniatis, Molecular Cloning, second edition, Cold Spring Harbor Laboratory Press, New York 1989.
[21] D. Hanahan, in: DNA cloning (D. M. Glover, ed.), Vol. **I,** IRL Press, Oxford, Washington 1985.
[22] W. R. Pearson and D. L. Lipman, Proc. Natl. Acad. Sci. U.S.A. **85,** 2444–2448 (1988).
[23] C. Gaboriaud, V. Bissery, T. Benchetrit, and J. P. Mornon, FEBS Lett. **224,** 149–155 (1987).
[24] J. Garnier, D. J. Osguthorpe, and B. Robson, J. Mol. Biol. **120,** 97–120 (1978).
[25] G. Heidecker and J. Messing, An. Rev. Plant Physiol. **37,** 439–466 (1986).
[26] S. H. Phua, B. R. Srinivasa, and A. R. Sabramanian, J. Biol. Chem. **264,** 1968–1971 (1989).
[27] D. Shemin, Naturwissenschaften **57,** 185–190 (1970).
[28] G. A. Karlin-Neumann and E. M. Tobin, EMBO J. **5,** 9–13 (1986).
[29] A. Seidler and H. Michel, EMBO J. **9,** 1743–1748 (1990).
[30] C. E. Hurt and A. P. G. M. van Loon, Trends Biochem. Sci. **11,** 204–207 (1986).
[31] C. Gietl, Proc. Natl. Acad. Sci. U.S.A. **87,** 5773–5777 (1990).
[32] S. van Heyningen and D. Shemin, Biochemistry **10,** 4676–4682 (1971).
[33] P. M. Gilmartin, L. Sarokin, J. Memelink, and N.-H. Chua, The Plant Cell **2,** 369–378 (1990).