Molecular characterization of the genes encoding acetohydroxy acid synthase in the cyanobacterium *Spirulina platensis*

A. MILANO, E. DE ROSSI, E. ZANARIA, L. BARBIERATO, O. CIFERRI and G. RICCARDI*

Dipartimento di Genetica e Microbiologia 'A. Buzzati Traverso', Via Abbiategrasso 207, 27100 Pavia, Italy

(Received 3 January 1992; revised 10 March 1992; accepted 16 March 1992)

---

The enzyme acetohydroxy acid synthase (AHS), which catalyses the first common step in the biosynthesis of isoleucine, leucine and valine, has been demonstrated to be present in *Spirulina platensis* in two isoenzymic forms. The complete nucleotide sequences of the genes *ilvX* and *ilvW* encoding these two enzymes have been determined. Sequence analysis revealed the presence of two open reading frames, of 1836 and 1737 nucleotides for *ilvX* and *ilvW*, respectively. The predicted amino acid sequences of the two isoenzymes, compared with the *Synechococcus PCC 7942* AHS enzyme and the large subunits of the *Escherichia coli* AHS I, II, III isoenzymes, revealed a notable degree of similarity. A small subunit has not been identified for either of the *S. platensis* AHS isoenzymes. Analysis by Northern blot hybridization demonstrated that the *ilvX* and *ilvW* genes are transcribed to give mRNA species of approximately 2.15 kb and 1.95 kb, respectively.

---

**Introduction**

The study of amino acid biosynthesis, which has substantially contributed to our present knowledge of microbial metabolism, has been very limited in cyanobacteria (Riccardi et al., 1989). We are studying the pathway of isoleucine, valine and leucine biosynthesis in *Spirulina platensis* as an approach to understanding the regulation of cyanobacterial metabolism (Riccardi et al., 1990). Acetohydroxy acid synthase (AHS, EC 4.1.3.18) is the first common enzyme in the biosynthetic pathways to valine, isoleucine and leucine. AHS has recently been demonstrated to be present in two new and structurally unrelated classes of herbicides, the sulphonylureas (Chaleff & Mauvais, 1984; Ray, 1984) and the imidazolinones (Shaner et al., 1984).

Three AHS isoenzymes have been found in *Escherichia coli* and *Salmonella typhimurium* (De Felice et al., 1982; Umbarger, 1983). The genes encoding these isoenzymes are organized in operons and include *ilvBN*, *ilvGM* and *ilvIH*, coding for isoenzymes I, II and III, respectively (Wek et al., 1985; Lawther et al., 1981; Squires et al., 1983). In these bacteria AHS is a tetramer composed of two large and two small subunits. The large subunits have a molecular mass of about 60 kDa, while the small subunits have molecular mass of 10 to 20 kDa (Eoyang & Silverman, 1988; Schloss & Van Dyk, 1988; Barak et al., 1988). Only one gene encoding functional AHS lacking a small subunit was isolated from yeast (Falco & Dumas, 1985; Falco et al., 1985). Genomic clones encoding AHS have been isolated and sequenced from the higher plants *Arabidopsis thaliana* and *Nicotiana tabacum* (Mazur et al., 1987). In *A. thaliana* the gene codes for an AHS of molecular mass 72593 Da. Friedberg & Seijffers (1990) isolated and sequenced a gene from *Synechococcus PCC 7942* that encodes a protein of 612 amino acids showing high sequence similarity with the AHSIII large subunit of *E. coli* and the monomeric tobacco AHS.

Two isoforms of AHS were detected in cell-free extracts of *Spirulina platensis* and separated both by ion-exchange chromatography and by hydrophobic interaction (Forlani et al., 1991). The isoforms differ in pH optimum, FAD requirement for both activity and stability, and heat lability (Forlani et al., 1991). The corresponding genes, *ilvX* and *ilvW* (named *ilvY* in our previous papers), were isolated from a λ library of *S. platensis* genomic DNA. The 4.2 kb *ClaI* (*ilvX*) and the 3.2 kb *SalI/ClaI* (*ilvW*) fragments were successively subcloned in the plasmid vector pAT153 (Maniatis et al., 1983).
The *ilv* gene was able to complement a suitable mutant of *E. coli* while the *ilv* gene supported poor growth of the same *E. coli* mutant (Riccardi et al., 1991).

We here present the complete nucleotide sequence and a preliminary transcriptional analysis of the two genes (*ilvX, ilvW*) encoding AHS isolated from *S. platensis*. We also compare the deduced amino acid sequences of the two AHS isoenzymes to that of the large subunit of the *E. coli* AHS isoenzymes I, II and III and to the *Synechococcus* PCC 7942 enzyme.

**Methods**

**Culture conditions, RNA isolation and Northern hybridization.** Culture conditions for *S. platensis* C1 have been previously described (Riccardi et al., 1981). RNA was extracted from 1 litre of *S. platensis* culture grown to mid-exponential phase and then harvested by filtration on paper discs. Cells were immediately frozen in liquid nitrogen, ground using a ceramic mortar and pestle, and resuspended in guanidine HCl buffer (8 M-guanidine HCl, 0.1 M-sodium acetate, 5 mM-dithiothreitol, 0.5 M-sodium lauryl sarcosinate, pH 5.2). The cell suspension was extracted with an equal volume of water-saturated phenol and the aqueous phase was re-extracted with chloroform/isomay alcohol. The RNA was precipitated by adding 2.5 volumes of sodium acetate (pH 5.2) followed by storage at −20 °C for at least 2 h. RNA was recovered by centrifugation at 10 000 × g for 10 min at 4 °C. The pellet was washed with ice-cold 70% (v/v) ethanol, recentrifuged briefly, and redissolved in a small volume of water. Three volumes of ethanol were added, and the preparation was stored at −70 °C until needed.

Aliquots containing 20 μg of the RNA were fractionated by formaldehyde-agarose gel electrophoresis (Maniatis et al., 1982) and transferred by an alkaline blotting procedure to Hybond-N filters (Amersham) according to the manufacturer's instructions. Filters were prehybridized and hybridized according to standard procedures (Maniatis et al., 1982). The hybridized blots were washed twice for 10 min each in 2× SSC (g l−1: 17.5 NaCl, 8.8 Na2HPO4, 2H2O, pH 7.2) and 0.1% (w/v) SDS at 50 °C and exposed to X-Omat AR X-ray films at −70 °C with an intensifying screen.

**DNA sequence determination.** Double-strand nucleotide sequencing of the two *ilv* genes from *S. platensis* was performed on both strands by the dideoxy-chain termination method (Sanger et al., 1977), after subcloning overlapping restriction fragments, ranging from 200 to 500 bp, in the plasmid pGEM4Z (Promega Corp.). A sequencing kit was used (Pharmacia, LKB); reactions were carried out according to the manufacturer's instructions. SP6 and T7 promoter sequences (Promega Corp.) were used as primers. To overcome secondary structures in *G* + *C*-rich stretches, 7-deaza-dGTP was used rather than dGTP in the sequencing reactions. DNA sequences were analysed with the DNA and Protein Sequence Analysis Programs of D. W. Mount and B. Conrad, University of Arizona. Protein sequence alignments were generated with the CLUSTAL program (Higgins & Sharp, 1989). Possible secondary structures downstream of *ilv* genes were identified using the RSEC program (Lang & Burger, 1986). Free energy of formation for these structures was also calculated using the RSEC program.

**Results and Discussion**

**Sequencing of the *S. platensis* genes encoding AHS**

In a previous study, a *S. platensis* genomic library was shown to contain a 4.2 kb *ClaI* fragment and a 3.2 kb *ClaI/SalI* fragment carrying the presumptive genes encoding AHS (Riccardi et al., 1991). The sequences of these two genes and of their deduced polypeptides are shown in Fig. 1(a, b). A total of 3373 and 3170 nucleotides including respectively the entire *ilvX* and *ilvW* genes and their 5' and 3' flanking regions were sequenced. The coding region of *S. platensis* *ilvX* includes 1836 nucleotides, corresponding to 612 amino acids, while that of the *ilvW* gene has 1737 nucleotides, corresponding to 579 amino acids. The molecular masses predicted by the sequences are about 67 kDa for *ilvX* and 63 kDa for *ilvW*. In *ilvW*, GTG appears to be the initiation codon. The same codon was found in the corresponding gene of *Synechococcus* PCC 7942 (Friedberg & Seijffers, 1990), as well as in other cyanobacterial genes (Reddy et al., 1988; Yasui et al., 1988). GTG is also a translation start codon in other organisms (Kozak, 1983). Each gene is preceded by a polypurine tract which resembles the sequences proposed by Shine & Dalgarno (1974) for a ribosome-binding site (RBS). The *ilvX* gene is preceded by the sequence 5' GAGGAG 3', 6 to 11 nucleotides upstream from the apparent translation start codon. The *ilvW* gene is preceded by the sequence 5' GAGGAA 3', 5 to 10 nucleotides upstream from the apparent translation start signal. The sequence of the 16S rRNA from the cyanobacterium *Anacystis nidulans* (Tomioka & Sugiiura, 1983) is very similar to that of the 16S rRNA from *E. coli*, suggesting that analogous RBS sequences are likely to occur in both organisms. Indeed, sequences such as AAGG, GGAG have been found upstream from most of the cyanobacterial genes sequenced so far (Tandeau de Marsac & Houmard, 1987).

AHS in other bacteria studied is composed of two large and two small subunits whose genes are organized in the DNA and Protein Sequence Analysis Programs of D. W. Mount and B. Conrad, University of Arizona. Protein sequence alignments were generated with the CLUSTAL program (Higgins & Sharp, 1989). Possible secondary structures downstream of *ilv* genes were identified using the RSEC program (Lang & Burger, 1986). Free energy of formation for these structures was also calculated using the RSEC program.
(a) Sequence of ilrX

- 878 AACCCTGACCTGATGACATAGGGATCAGTCTCAAGGCGGATGATAGAGAAGACTACG

- 815 TTCCCCAGTTGGGAGAGGCCTCAGGCAATGATATACGGGAAATGCTTTTAATGGCTCTTA

- 752 GGTCTCTTTCTCCTGTCGGATACGAATACTTTTCTGCTAACTGAGGAGAGAAATATCTTACG

- 689 CTCATATCAGCAGATCAGGATGATATATTTTCTGCTAAGACTGACATACGGCAGGAT

- 626 CTTGACCACAAATTTCCATTTGAGTTTTTTTGTAGAAGCTGGAACACTGATATTTCGGAACCCGAT

- 563 AAGGCTCAGTCTGACCTTTTGGGCTGCGGAAAAACCCGCGAACCTTGCGGAAACCTGCTATTCCGACT

- 500 TGCCAGCATTGACAGAGGGAAATCTTGACTCTCAAGATTTAATGCGGCAAATATGGGCTCTC

- 437 ATCTTTTCTGTTTTTCTGTTTCATACGCTATCTTGGAGAAATCTTCATACCTCTGCGATAGTGGG

- 374 TTTACTGCTGCTGACAGAGGTGGTGCAGAACTTTGGAGTGTGAGTACAGCTGACCC

- 311 AGCTATCTAGATATAGTACAGGTGCTGATGCAAATCTCTCAGAGGAGTACAGCTGACCCGAT

- 248 TTATCCGACATCATACTGGGCAATAGGGGAAACCCTTTTAGGATATCTAAAGGGCACCATG

- 185 CTGTGATACAAATTCGAGAAGAGCCTGCGACCCCGAGCAATACCCGGGCTCTAGTTGAAAAT

- 121 AGTGGGTGCTGAAACAATGTAAATCTTGTGAATAACCGCTACGTCGCCTTTAAGAAA

- 58 TATTATCTAGATATAGGAGTTTTAATCTCAAGAAATATAGAATAGAGGAGGCGACGC

    M Q

    L Q T K I A A A K R AT G A F B L I D S L K

    R H G V Q H I F G Y P G G A I L P I Y D E

    L Y R A E A E G D I Q H I L V P H E Q G A

    S H A A D R A T G R V G V C F G T S

    G P G A T N L V T G I A T A H M D S I P M

    V I I T Q V R A P A I G T D A F Q D S D

    I F G I T L P I V K H S Y V V R E P G D M

    A R I V A H I A T S T G R P G P V L I

    D V P K D V G L E E F D Y I P V N P G E V

    S L P G Y R P T V K G N V R Q I N Q A I K

    L I E E A E R P L M Y V G G G A I S A T A
700  CATGCAGAAATTGGGAAATTTTCCAGATTCCGTTAATCTACCCATACCAATAGGGG
HAEIAELAEBLFEQIFVTVTTLMLMG
762
730  AAAGGATCGTTTCTGATGAAAGACATCCCCCTACTGTTGGGAAGTGGGGATGCACGCC
KGFEDHEKLMGLMGHGTAK
825
790  TATGCTAACTTGGCTGGTTGAATGCTGATCTTCTATGCTGGCTGCTGAGCACTATG
YANFAVCDFLIAVGAARFFDD
988
890  CGGCTACCGCCGAAAAACCTTTGGATGCTAGCTGGAGATCTATAGGAAAATGAG
RVTGKDLBFSRBAVKHIDID
951
1010  CGGCGGCAAGTGGTAAACAAAACCCCGAACCAGGAGTGCTCAAATAGGGGTACG
PALPEVKGKNTPEVPIVGDVRQV
1014
1015  TTGATACACTCTGCTTGGCCGCTCCGAGAAATGGAGATGCTGGGAAGCTAGAATCAACA
LHTLLRERCEIDVGNQNTFQ
1077
1078  TCTGTTAGAGAGGTATATGCTGGCAGAAGATATTATCTCTCCTGGTCTCTATTCT
SWLERINPEDEPLVVFPSYS
1140
1111  GATAGGCTCTCCTCCCCAAAGGGCATGCAATAATAGGAAAATTGCAACCCGATGGCAGTACATTAT
DSLSPFQEVIALGKMAPDGYY
1203
1204  ACCACCGGATGGTGTCAGCAACATAATGGGCGGCTCAATTTTGGAAAATGCGCCCGTCAAG
TTDGQHMWAAQFLKNGFRQ
1266
1267  TGATATATCTACGTCAGTCTGGGAACTATGTTGCTGTATCCGTCTGGCTATGGGAACCTAAG
WISSAGLGMGYGISMPAMGAK
1329
1330  GTGTGCTTGGGCAGTACGCAGGTGTATTGTTTATGCTGGGTAGCTCTGTGAAATGACATTC
VALERSVQVICIAGDASVQMN
1391
1392  CAAGAAGCTCGGAAACCATTAGCACAAATATGCTAATAATATGAAAGCCGGAACGGTAAT
QELGTVQYGINVKTVINNG
1455
1456  TGCCAAAGGTATATGGTCCGCGCAGTGGCAACAAGGCTTTATGGTGGAGCGTTATGCGGCCTCAAC
WGGMVRQWQQAEYGERYSAHSN
1513
1519  ATGGAAAAACGGAAGATTTTGGAAATGTGACTAGCTTCTGCTTAGGAAAGCTGGTACCG
MEIGMFDPEMLARSGYVKGVM
1581
1582  GATGAAAGCTGCTGACAAATTACACAACAGCAGTTAGAAGATCTGCTATTGATGCTCTTG
VKSRDELHQAALAEMLAYDGPV
1644
1645  CTTATGGGTATATGTATCTTTCAGGTAAGAACCTTGTATCCTATGGTATCCAGGAAAGGAGT
LMDDVHVTVDENCPYMVAFGEB
1707
1708  AATGCCTAGATGTAGATTACCTGAAAGGCCAGCCAGATGAAGAAACGGTGGAGTATACTAT
NAQMIQLPERRQLKEAVELIY
1770
1771  TGGCGTAAATGGGCGCCGAAAATGGCTGGCAAGGTTACAAACTTTCTGTCTGCTGAGTGGCAGT
CSNCGAKNVASNNNFCPECGBK
1833
1834  CTATTAAAACACTCCTCTTATTACCTGTGACACATGAGAATTGTGCTGGCTATTGATTATGC
L
1896
1897  GATAGTGGGAGACAGGCTTGGCTCTCTCTTACTCTGTCATTGTAGAGGGTATGCTG
1959
1960  ACACCTAGGCTCAGGAAATAGTGGAGAGCCTTACCCCTGCCGGAAATAGAAATCAAT
2022
2023  GGTAGCGAGAACATCTGGATTTTTAAGATGTAGTGGATATTGCTGGGCTGCAAACATTATCA
2085
ilv genes from *Spirulina platensis*

(b) Sequence of *ilvW*

- 522  TGCACCTGGCAGATCGACCGAAGCTGCAATTTTAGAAGTAGCGCAGGGTAAAGGGACCCTTCCTCCTCTCTCCT - 460
- 459  CGATCCTAGCCAGCAGCCTTGGAGAAGGCCCTGCAAATTTCTACCCAGAATAGCCGAATT - 397
- 396  CACTGCGCCATTATAAAGTGGCAGATTTTTAGATGGTAAAGCCGGAACCTCTGCTCCAAAATCTCAGT - 334
- 333  GTATGGTAGAGTCTAGTATAATGCTACCCCTGGAGTACCGTAGAGCTCTCTACTACATCTC - 271
- 270  TAGAGCCTATCTATCAGCCGTGCAAGATGGTTAGAGATTAGTGATAATGTCTGGAAACAGCA - 208
- 207  ACCCATCCATCCGTATATGATACGGGAATCTCTACTGTAATATAGCTGATATACGCTTA - 145
- 144  GGATGGCTATGGGATCGCATATAAACCAAAATGGCAGTAATGTATCTGAAATTTG - 32
- 81  CACAAATATAAATATGCTGACATGTGCTGATAAGCTCCAAAAAGTAATACAGGAGCCTT - 19
- 13  CACAGATGGGAAATGGGTGTGACGAGGACAAACAGCGGATAAGAGAGTAACAGGAGGCTTTT - 46

MQDQKQAKAVKRVGTAG

47  CCCTGATTAGCCCTGGCCGCGCGCCACCTTGGAGATATTTTGGCTTTACCCAGAGGGCTCAA
    ALIDSRLRHGHVHIGYPAGGS

110  ACTTACCATTTACAGGAAATATATATAGAGCCGAAACAGCGGGGAAATCAACATATTATCTGG
    NLPYIDEIYRAEQAEGIEKHYL

173  TGGGGCAAGGACAGAGAGGGGCCACAGCTGGCGGATACTGGGATGGGATACCAGGAAGGGG
    VRHHEQCAAHADDYAGYSTGVK

236  GGATGTGTCTACCCACCTCTGGAGCCGGGGGACCAAAATTTAGTGACAGGCTTAGCCACCTGTT
    GVCLATSGFGATNLTGTALATA

299  ATCTCGATCTCTACCCCGGTTAGCCCATTTACCCGAGCAAGTCCCCCGCGACGGCTTGAGAAGAG
    YLDHSVPLAIQGVPRSAALT

362  ATGGGTTTTCAGGAAATGATTTTGGGATTACCCATTTTGGTTAAAACATTTTATTCTAG
    DAFQIDEIFGILTLPIKHSYL

425  TCCCGGAACCTTTCGGAATTTGCAACTACCTGATTGCTTAAACATTTATCTAG
    VREPSVRELIXBEAFHLAMSG

488  GACCAGGACCTGGTAAAGACATTCCCAGAAGATGGGATATCGCCAATATTGCTATTATTCT
    RGPGVPYLIIDIPKDVGNANQIDYI

551  CGTGAAGACGGGTCACTACGAGGGTGGGATTACGTCCTACCCACGGACCGGGGAAATCTCGAC
    PVEPGSVVRVGYRFTERGNPR
614 AAATATCAGGGTTTACAGTTGATTTCCGAACTACTAAACCCCTCTTATATGTTGGGGGAG 676
616 QINQALQLISEATKPLLLYVGG
677 GTGCAATTATGGGCGGCTCAGTGAAATGCGGAAATTGAGGGAACCTTACCAATACCAG 739
680 GAIMAHABIAELSERQIF
740 TTAACTAGCACTTTAGGGAAATGGCGCTTTTTGATGAAAATGCACCCCCATATCCTGGGAATCG
782 VTSMLMGKGRFDENNHPFLSGLI
803 TGGAGAATGCTTGGAATGCACGGCGGACTATCTAATTTTGCGTGATGGAAATTGATTTGCG
365 VMGLMHTGAYANFAVMELDE
866 TGATGCGGCTTGGGCCTTTGATGACGGCGGTCGCTGGGCTAGATCAATTGCCCAATA
928 VIAVGVBFDDVRAGTGDQFAH
929 GCCCTAAGGTGATTACTATGATATATGACCCGGCGGAATGGGGAAAATCGCACTGAGTATG
991 SAKVHIDPEAVGKNRSTD
992 TTCTCATATTGGGAGATGTGGGACAGGTCTTGGGAGATGCTACAAATGCTAGCTATATTGGG
1052 VPIVGDVRQVQLGMQLRTYHW
1055 AACGGAATTAACGCGGAAACAAAACCGGATGGAACCTTAAATCGATTGGTAACCCA
1117 ERLKRSRNPRGTLDNLQFEP
1188 TACCCCCGACGGTTTCAACCCGGGAAAGATGTTATCCTTCATCAAGATGGTATGGGAAACTTA
1150 IPLTVPHPEDGISPQDGDWEL
1181 GCCATCAATGCTCTGATGTTCTATCTACTATGTTGGCTACCCGATAATGTTGGGGGGCC
1243 SHCQCPFDAFYTDVQGVHQMWMAG
1244 AATTCTGCTAAAAATGGCCCCCGACGTTTGGATGACTACGCGGCCGTTTGGAGTTAGTTACG
1306 QFVQNGPRRWMGSTGGLGTMGY
1307 GHTTACCTGCTGCTGGGCTAATGGTGTGTCTACCTCTCATGACTACGGTTACCTGATATTAGCG
1369 GLPAAVGVKYVHPHDTVTCSIS
1370 GTGATGGAATTTCCAAAATGAATATGCAAAAGATTTGGGAAATATTGCTAGTAGTTTGGG
1432 GDGSFQMNMQELGTIAQYGIGC
1431 TTTATTCTTAATATAATGGCTTGTTTGGGAATGTTGGCGCAGTTGGGCAACATATTGTIXTTT
1495 VKVILILNNGWLMVRWQWHMF
1496 ATAATGACCGCTATGAGGGACTAATACCCGGAATGGAACCCCAAGAATTTGGCAGCTTAGCG
1553 YNDRYEBATHNLEDGTFEFARL
1559 ATGTTTATGGGACTGCGGATGGCTGCGCAACACCAGAATATTACCCAGGCTGCA
1621 DVYVGLEAMNVQSKRIYQRRLF
1622 AAGCTCTATCTCAAAGGTTGCGATGTACCTATGTTGGGCTAACCCGTGATAGGAGACTGT
1684 KALSHKGMIPLDVVRTRDDEC
1685 ACCCGATGGAAGGCGGTACACGATAAACACCGACATGATTGGGTTTATCCAGTTAGTGGCTCT
1747 YPMVPDHNSDDMMLSS
1748 AACCCTCAGCCCCCCTTTAGTCTTATGTTGGGGGTTATTCCTGTGAGAACAATCTTAATT
1810 CAACCTGAAAAACCGCCATATTGATTTTAAATCGAAAAATATAATGTTTTTTGGAT
1873 AAATCGATTGAAAAATTTGGTACAAATAATCTTCAAAATAATGTTTACCTTCAATTCC
1936 TAACCGGAAATGATATCTATCTATATGAGCGGGAACACTAAGGAAATCGGAGAGCAGCAAAGTCC
1999 TAACCGGAAATGATATCTATCTATATGAGCGGGAACACTAAGGAAATCGGAGAGCAGCAAAGTCC
2000 CACGGTATCCCCGGTGTTGAATGTTTGGAAAATCTGGATTTTCCAAAGATGATTTTCCTGTC
operons (Umbarger, 1983). Attempts to find additional open reading frames in the 3' flanking regions of both ilvX and ilvW were unsuccessful. We hypothesize that the AHS small subunit either is not present in S. platensis, similar to the situation in Arabidopsis thaliana and Nicotiana tabacum (Mazur et al., 1987), Brassica napus (Wiersma et al., 1989), yeast (Falco et al., 1985) and Synechococcus PCC 7942 (Friedberg & Seijffers, 1990), or is not organized in an operon with ilvX or ilvW. However, the existence of a small subunit in the S. platensis AHS isoenzymes cannot be excluded. We intend to purify the S. platensis ilvX and ilvW gene products from recombinant E. coli and to prepare antibodies against them in order to purify the isoenzymes from S. platensis by affinity chromatography. This approach could further elucidate the subunit structure of the cyanobacterial AHS.

Amino acid sequences encoded by ilvX and ilvW

The deduced amino acid sequences of these genes can be aligned with Synechococcus PCC 7942 and E. coli AHS proteins as shown in Fig. 2. The six AHS isoenzymes show close amino acid sequence similarity. Previously, a comparison of the ilvG, ilvB and ilvI gene products revealed three regions of similarity (Wek et al., 1985). These three regions of similarity can also be identified in the ilvX and ilvW gene products. The overall degree of similarity is not equally distributed over the entire lengths of the six polypeptides: several of these blocks extend for 20 to 30 amino acids with few mismatches and other regions show much less similarity. However, the degree of overall identity indicates that these genes encode a family of related polypeptides. In addition to these regions there is conserved similarity between the carboxyl termini of the ilvX and ilvW gene products and the ilvB and ilvG gene products that is not shared with the ilvI product.

Comparison of the translated ORFs revealed that the two S. platensis AHS isoenzymes are 66% identical. The deduced amino acid sequences of ilvX and ilvW are respectively 70% and 61% identical with that of Synechococcus PCC 7942; 48% and 50% with that of Synechococcus PCC 7942; 42% and 44% with that of ilvG; 42% and 44% with that of ilvB; and 46% and 44% with that of ilvI.

Haughn et al. (1988) have shown that in A. thaliana, a single C to T transition within codon 197, resulting in proline being replaced by serine, is responsible for the mutant chlorosulphuron-resistant phenotype. Lee et al. (1988) have also shown that in the N. tabacum gene encoding AHS, the equivalent proline may be replaced by glutamine or alanine to give resistant forms of the enzyme. As shown in Fig. 2, motifs homologous to those of Brassica napus, A. thaliana and N. tabacum (AITGQVPRRMIGT; Wiersma et al., 1989) are also present in ilvX and ilvW. It is noteworthy that proline is replaced by alanine in the ilvX protein. Since S. platensis is resistant to 0.4 mM-chlorsulphuron (data not shown), this may be due to the presence of a resistant form of AHS. This result was confirmed in the E. coli mutant complemented with S. platensis ilvX (data not shown).

Transcriptional analysis

RNA isolated from S. platensis, grown in minimal medium, was subjected to Northern blot analyses using, respectively, the 1.5 kb EcoRV fragment from ilvX (nucleotides -215 to 1270, Fig. 1a) and the 0.66 kb
Fig. 2. Comparison of some known AHS amino acid sequences. The deduced amino acid sequences of genes encoding AHS are shown for: A, E. coli isoenzyme I; B, E. coli isoenzyme II; C, E. coli isoenzyme III; D, Synechococcus PCC 7942 enzyme; E, Spirulina platensis isoenzyme I; F, S. platensis isoenzyme II. The overlined sequences indicate blocks of similarity associated with sulphonylurea-resistant phenotypes. Identical residues are indicated by asterisks, and conservative amino acid substitutions by dots. Dashes represent gaps inserted to optimize the protein alignment.

Fig. 3. Northern blot analysis of ilvX and ilvW transcripts in S. platensis. Total RNA extracted from S. platensis cells grown in minimal medium was fractioned on formaldehyde-agarose gel, transferred to a Hybond-N+ filter and hybridized with the 1.5 kb EcoRV fragment from ilvX (A) and the 0.66 kb EcoRVIHaeIII fragment from ilvW (B) probes.

EcoRV/HaeIII fragment from ilvW (nucleotides 580 to 1240, Fig. 1b) as internal probes (Fig. 3). The preponderant message species was about 2.15 kb for ilvX and 1.95 kb for ilvW, corresponding roughly to the coding length required for the genes (1836 kb for ilvX and 1737 kb ilvW, Fig. 1).

The ilvBN and ilvGM operons are both regulated by transcription attenuation (Hauser & Hatfield, 1983; Lawther & Hatfield, 1980). The presence of a leader region (encoding a polypeptide rich in branched-chain amino acids) was not apparent in the 5' flanking sequences of either ilvX or ilvW. Analysis of the 5' non-coding region of ilvX and ilvW disclosed putative promoter sequences homologous to known cyanobacterial promoters (Tandeau de Marsac & Houmard, 1987). Some hypothetical −10 and −35 regions of ilvX have resemblances to the E. coli consensus promoter sequences (TTGACA, TATAAT). This could be an explanation for the difference between the two S. platensis genes in the complementation of the E. coli mutant (Riccardi et al., 1991). To overcome this problem we are trying to place the ilvW coding region under the control of an E. coli promoter. The ability to express this cyanobacterial enzyme in E. coli offers the opportunity to take advantage of the fast growth of E. coli cells to study a protein whose analysis is difficult in the organism of origin, essentially due to its high lability and to the presence of large amounts of pigments. It is noteworthy that the 5' flanking regions of ilvX and ilvW contain, near a hypothetical −10 region, an inverted-repeat sequence that could play a role in the expression of these genes (operator?). Downstream from the S. platensis ilvX and ilvW genes are, respectively, 30 bp and 28 bp sequences (underlined converging arrows in Fig. 1) which can code for an RNA capable of forming a stable stem and loop structure. The free energy of formation of such a conformation from these sequences is approximately −25.8 kcal mol−1 (1 cal = 4.184 J) in ilvX and −22.7 kcal mol−1 in ilvW. Whether these structures act as a transcription terminator and/or as a barrier against
The nucleotide sequences of 1408

The authors would like to thank Professor M. De Felice for helpful advice.

The present work was supported by MURST 60% and CNR Progetto Finalizzato PB no. 90.01304.14.

References


