

A novel composite locus of *Arabidopsis* encoding two polypeptides with metabolically related but distinct functions in lysine catabolism

Guiliang Tang, Xiaohong Zhu, Xiaohu Tang and Gad Galili*

Department of Plant Sciences, The Weizmann Institute of Science, Rehovot 76100, Israel

Received 8 February 2000; revised 24 March 2000; accepted 27 March 2000.

*For correspondence (fax +972 8 9344181; e-mail gad.galili@weizmann.ac.il).

Summary

Both plants and animals catabolise lysine via saccharopine by two consecutive enzymes, lysine-ketoglutarate reductase (LKR) and saccharopine dehydrogenase (SDH), which are linked on a single polypeptide. We recently demonstrated that *Arabidopsis* plants possess not only a bifunctional LKR/SDH but in addition a monofunctional SDH enzyme. We also speculated that these two enzymes may be controlled by a single gene (G. Tang *et al.*, *Plant Cell*, 1997, 9, 1305–1316). By expressing several epitope-tagged and GUS reporter constructs, we demonstrate in the present study that the *Arabidopsis* monofunctional SDH is encoded by a distinct gene, which is, however, nested entirely within the coding and 3' non-coding regions of the larger bifunctional LKR/SDH gene. The entire open reading frame of the monofunctional SDH gene, as well as some components of its promoter, are also parts of the translated coding sequence of the bifunctional LKR/SDH gene. These special structural characteristics, combined with the fact that the two genes encode simultaneously two metabolically related but distinct enzymes, render the LKR/SDH locus a novel type of a composite locus. Not all plant species possess an active monofunctional SDH gene and the production of this enzyme is correlated with an increased flux of lysine catabolism. Taken together, our results suggest that the composite LKR/SDH locus serves to control an efficient, highly regulated flux of lysine catabolism

Keywords: lysine-ketoglutarate reductase, saccharopine dehydrogenase, lysine catabolism, essential amino acids, transgenic plants, composite locus.

Introduction

In plant cells, similarly to animal cells, excess lysine is catabolised into glutamate and acetyl CoA via the α -amino adipic acid pathway (Figure 1) (Galili, 1995). The first enzyme in the lysine catabolic pathway, lysine-ketoglutarate reductase (LKR), condenses lysine and α -ketoglutarate into saccharopine, which is then converted by the second enzyme, saccharopine dehydrogenase (SDH), into α -amino adipic semi-aldehyde and glutamate (Goncalves-Butruille *et al.*, 1996; Markovitz *et al.*, 1984). Lysine catabolism plays an important physiological role in both animals and plants (Galili, 1995; Markovitz *et al.*, 1984). In plants, the LKR level was shown to be significantly up-regulated in inflorescence tissues and developing seeds, as well as in response to osmotic stress (Deleu *et al.*, 1999; Karchi *et al.*, 1994; Karchi *et al.*, 1995; Kemper *et al.*, 1999; Tang *et al.*, 1997). In addition,

LKR activity in tobacco seeds was shown to be stimulated by excess cellular lysine, via an intracellular signalling cascade involving Ca^{2+} and protein phosphorylation (Karchi *et al.*, 1994; Karchi *et al.*, 1995).

The control of metabolite flux via the LKR and SDH enzymes is still uncertain. In both plants and animals, these two enzymes are linked on a single bifunctional polypeptide (see for example Goncalves-Butruille *et al.*, 1996; Markovitz *et al.*, 1984; Miron *et al.*, 1997; Tang *et al.*, 1997). Yet, despite their physical linkage, these enzymes possess significantly different pH optima; physiological pH values are optimal for LKR but not for SDH activity, suggesting that metabolite flux via the SDH enzyme of LKR/SDH may be inefficient.

Metabolite flux via the α -amino adipic acid pathway may not be solely regulated by the bifunctional LKR/SDH

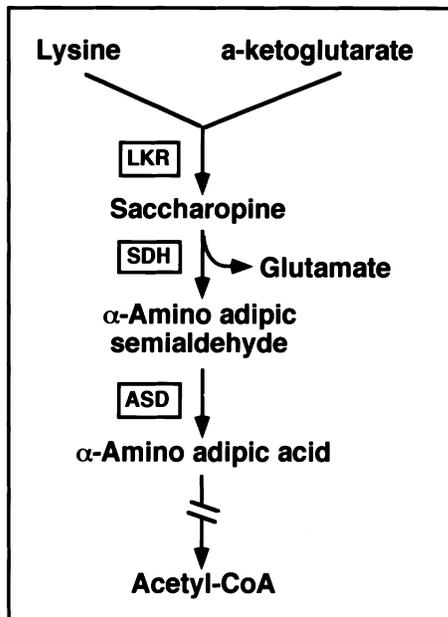


Figure 1. Schematic diagram of the lysine catabolism pathway. LKR, lysine ketoglutarate reductase; SDH, saccharopine dehydrogenase; ASD, amino adipic acid semi-aldehyde dehydrogenase. The broken arrow represents six enzymatic reactions leading to acetyl-CoA synthesis.

enzyme. We have recently demonstrated that *Arabidopsis thaliana* plants also possess a monofunctional SDH enzyme, in addition to the bifunctional LKR/SDH (Tang *et al.*, 1997). The genetic control of the bifunctional LKR/SDH and the monofunctional SDH was not elucidated. Yet, based on Southern blot and DNA sequence analyses, we speculated that these two enzymes may be encoded by the same gene (Tang *et al.*, 1997). In the present report, we show that the *Arabidopsis* bifunctional LKR/SDH and monofunctional SDH enzymes are encoded by distinct genes, but the monofunctional SDH gene is nested entirely within the coding and 3' non-coding region of the bifunctional LKR/SDH gene.

Results

A single Arabidopsis LKR/SDH gene simultaneously encodes two mRNAs and two polypeptides

To dissect the genetic control of the *Arabidopsis* LKR/SDH and monofunctional SDH enzymes, we first tried to ascertain whether they were derived from a single gene or from two independent genes. We thus constructed a recombinant *Arabidopsis* LKR/SDH (*AtLKR/SDH*) gene to which a small DNA fragment encoding three copies of an HA epitope tag was fused in-frame at the C-terminus upstream of the TGA stop codon (Figure 2a). To analyse the expression pattern of the recombinant HA-tagged

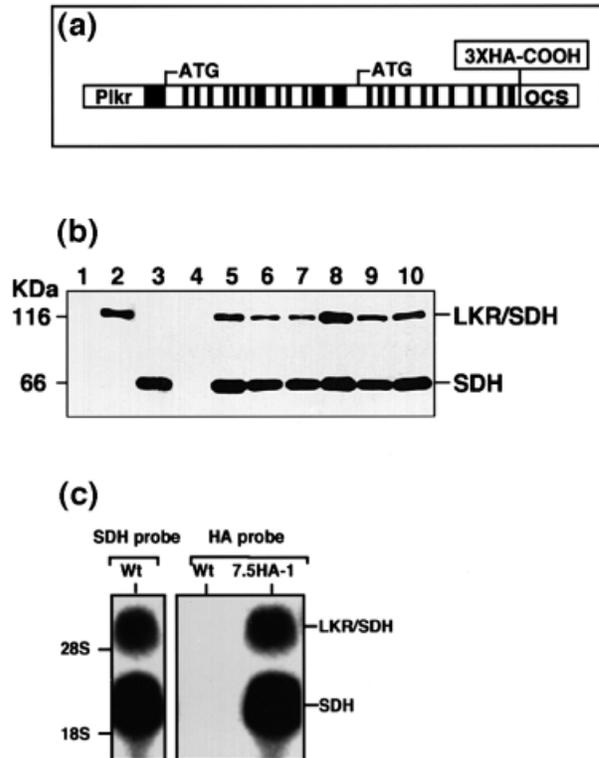


Figure 2. Expression of the chimeric construct encoding the HA-tagged *Arabidopsis* LKR/SDH in transgenic *Arabidopsis*.

(a) Schematic representation of the chimeric *gAtLKR/SDH::HA* gene encoding the *Arabidopsis* HA-tagged LKR/SDH. The triplet HA sequence at the C-terminus is shown in a box on the right. Black rectangles indicate introns; empty rectangles indicate exons. Plkr, upstream promoter of the *LKR/SDH* gene; ATG, LKR/SDH translation initiation codon; OCS, octopine synthase terminator. (b) Western blot analysis of protein extracts from yeast as well as from inflorescences plus developing siliques of *Arabidopsis* plants, using anti-HA antibodies. Lanes 1–3, extracts from yeast cells expressing, respectively, a vector alone, an HA-tagged *Arabidopsis* LKR/SDH, and HA-tagged *Arabidopsis* monofunctional SDH. Lane 4, extract from control non-transformed *Arabidopsis*. Lanes 5–10, extracts from independent transgenic *Arabidopsis* lines transformed with the chimeric *gAtLKR/SDH::HA* construct. The migration of the *Arabidopsis* bifunctional LKR/SDH and monofunctional SDH are shown on the right. The migration of molecular weight protein markers is indicated on the left. (c) Northern blot analysis of total RNA from reproductive organs of *Arabidopsis* plants. Wt, control non-transformed plants; 7.5HA-1, transgenic line 7.5HA-1. Blots were probed either with the SDH probe or the HA-specific probe (see Experimental procedures), as indicated at the top of the lanes. The migrations of the *Arabidopsis* bifunctional LKR/SDH (~3.5 kb) and the monofunctional SDH (~1.5 kb) mRNAs are indicated on the right. The migrations of 18S and 28S rRNA bands are indicated on the left.

AtLKR/SDH gene, we transformed this construct into *Arabidopsis* plants. Soluble proteins from stem sections containing inflorescences and developing siliques of several transgenic lines were reacted in a Western blot with anti-HA antibodies. As shown in Figure 2(b) (lanes 5–10), two polypeptides of approximately 116 and 66 kDa were revealed. These polypeptides migrated

identically with control HA-tagged *Arabidopsis* LKR/SDH (AtLKR/SDHp) and HA-tagged *Arabidopsis* monofunctional SDH (AtSDHp) proteins, produced in yeast cells upon expression of HA-tagged *Arabidopsis* cDNAs containing either the *AtLKR/SDH* or the monofunctional *AtSDH* open reading frames (Figure 2b, compare lanes 2 and 3 with lanes 5–10). Control non-transformed *Arabidopsis* had no detectable bands reacting with the anti-HA antibodies (Figure 2b, lane 4), confirming that both bands were produced by the transformed construct. The intensity of the smaller HA-tagged polypeptide was significantly higher than the one corresponding to the AtLKR/SDHp band in all transgenic lines.

We then performed a Northern blot analysis of RNA from control and one of the above mentioned transgenic lines to ascertain whether the two HA-tagged polypeptides were produced from the same or from different mRNAs. As shown in Figure 2(c), hybridization of RNA from the non-transformed plant with an *AtSDH*-specific probe revealed two bands representing the endogenous *AtLKR/SDH* and the monofunctional *AtSDH* mRNAs. Hybridization with an HA-specific probe yielded two similarly migrating bands in the transgenic line, but not in the control non-transformed plant.

The upstream promoter plus the first intron of the AtLKR/SDH gene are not essential for production of the smaller HA-tagged mRNA and polypeptide

Since the *AtLKR/SDH* gene contains a number of introns, the possibility existed that the smaller HA-tagged mRNA shown in Figure 2(c) originates from alternative splicing of a *AtLKR/SDH* precursor mRNA. To address this possibility, we deleted the entire upstream promoter and the first intron from the HA-tagged *AtLKR/SDH* gene (Figure 3a). Transgenic plants expressing this construct produced only the smaller HA-tagged mRNA and protein (Figure 3b,c), suggesting that these smaller mRNA and polypeptides did not originate by alternative splicing of an *AtLKR/SDH* precursor mRNA.

We next wished to confirm that the smaller mRNA shown in Figure 3(b) indeed represented an HA-tagged monofunctional *AtSDH* mRNA. To this end, total RNA from one of the transgenic lines expressing the *AtLKR/SDH* gene construct lacking the upstream promoter plus the first intron (Figure 3a) was subjected to RT-PCR using a forward primer starting from the internal in-frame ATG codon, and a reverse HA-specific primer. These primers gave rise to the expected ~1.5 kb RT-PCR band from RNA derived from the transgenic line, but not from a control non-transformed plant (data not shown). The identity of the ~1.5 kb RT-PCR product as a monofunctional *AtSDH* mRNA was further confirmed by DNA sequencing.

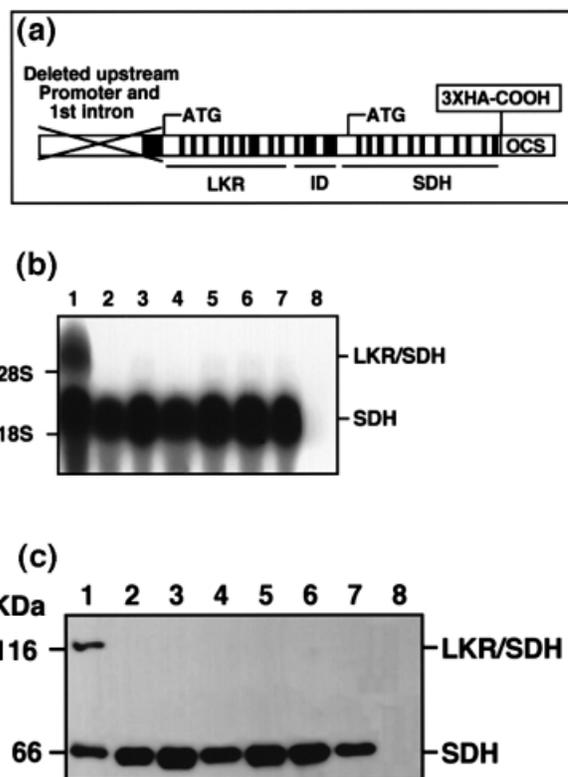


Figure 3. An *Arabidopsis* LKR/SDH gene lacking its entire upstream promoter and first intron still produces a monofunctional SDH. (a) Schematic representation of a chimeric *gAtLKR/SDH::HAΔPro* gene from which the entire upstream promoter and first intron were deleted. Black rectangles indicate introns; empty rectangles indicate exons. The triplet HA sequence at the C-terminus is indicated on the right. ATG, *LKR/SDH* translation initiation codon; OCS, octopine synthase terminator. The regions encoding the LKR domain, inter-domain region ID and SDH domain are indicated below. (b) Northern blot analysis of total RNA from inflorescences plus developing siliques of *Arabidopsis* plants. Blots were hybridized with HA-specific probe. Lane 1, line 7.5HA-2 expressing the *gAtLKR/SDH::HA* gene. Lanes 2–7, independently transformed lines expressing *gAtLKR/SDH::HAΔPro*. Lane 8, control non-transformed *Arabidopsis*. The migrations of the *Arabidopsis* bifunctional *LKR/SDH* (~3.5 kb) and the monofunctional *SDH* (~1.5 kb) mRNAs are indicated on the right. The migrations of 18S and 28S rRNA bands are indicated on the left. (c) Western blot analysis of protein extracts from reproductive organs plus developing siliques of *Arabidopsis* plants, using anti-HA antibodies. The lanes in (c) contain proteins from the same plants shown in lanes with identical numbers in (b). The migration of the *Arabidopsis* bifunctional LKR/SDH and monofunctional SDH are shown on the right. The migration of molecular weight protein markers is indicated on the left.

The coding DNA region of the AtLKR/SDH gene contains promoter-related sequences

Based on the results presented in Figure 3, we hypothesized that the smaller monofunctional *SDH* mRNA was transcribed from an internal promoter, located within the coding region of the *AtLKR/SDH* gene. Hence, we searched for potential promoter-related elements within the coding DNA region of the *AtLKR/SDH* gene, upstream of the *SDH* domain. As shown in Figure 4(a), putative CAAT and TATA

(a)

```

1653 TGGGGCAGAGCGCTGATGCTGAATCGTAACTCAGAACTTGAAGtaagttt
      G R A L M L N R N S E L E V
1703 aacctattcccacttgttaagaacccatcttgcatattcttggtagGTT
      CAAT Box
1753 GGTGGCGATGATAAGAGAGATTTGGATCAAAATCATTGATTCATTAACTCG
      A D D K R V L D Q I I D S L T R
1803 GTTAGCTAATCCAAATGAAGATATAATATCCCCACATAGAGAAGCAATA
      L A N P N E D Y I S P H R E A N
1853 AGATCTCACTGAAGATTTGGTAAAGTCCAGCAAGAAAATGAGATAAAAGAG
      K I S L K I G K V Q Q E N E I K E
SDH translation initiation codon
1903 AAGCCTGAAATGACGAAAAAATCAGGTGTTTGTGATTCCTGGTGCTGGACG
      K P E M T K K S G V L I L G A G R

```

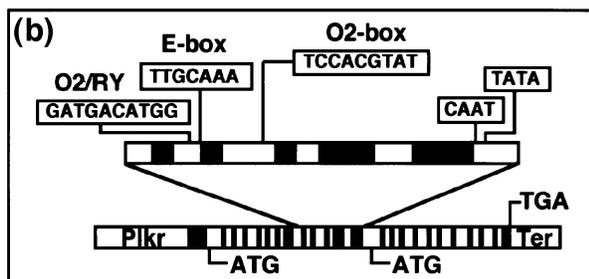


Figure 4. Identification of putative promoter elements within the coding region of the *Arabidopsis* LKR/SDH gene.

(a) Computer-predicted minimal promoter located in exon 14, upstream of the monofunctional SDH open reading frame. The predicted 'CAAT' and 'TATA' sequences and the initiator 'ATG' codon of the SDH open reading frame are in bold. (b) Schematic representation of a DNA fragment within the *Arabidopsis* LKR/SDH coding region, which contains the following putative promoter elements: CAAT and TATA boxes, Opaque2 box (O2-box), endosperm box (E-box), and a combined Opaque2/legumin RY box (O2/R1). The region containing these putative promoter elements is blown up at the upper part of the scheme of the entire LKR/SDH gene. Black rectangles indicate introns; empty rectangles indicate exons (bottom scheme). Plkr, upstream promoter of the LKR/SDH gene; ATG, LKR/SDH and monofunctional SDH translation initiation codons; TGA, the translation termination codon; Ter, the terminator of the LKR/SDH gene.

boxes were identified within exon 14, which encodes the linker region between the LKR and SDH domains, 50 bases upstream of the internal ATG codon. Computer analysis further suggested that these CAAT and TATA boxes have a high probability of functioning together as a minimal promoter. In addition, we also identified a number of putative promoter elements in the DNA region upstream of the putative CAAT and TATA boxes, between exons 10 and 14 (Figure 4b). These promoter elements included an Opaque2 box, an endosperm box, and a combined element containing both an Opaque2 box and a legumin box (see for example Dickinson *et al.*, 1988; Lohmer *et al.*, 1991; Muller and Knudsen, 1993; Pysh *et al.*, 1993). Elements of these types are common in the promoters of

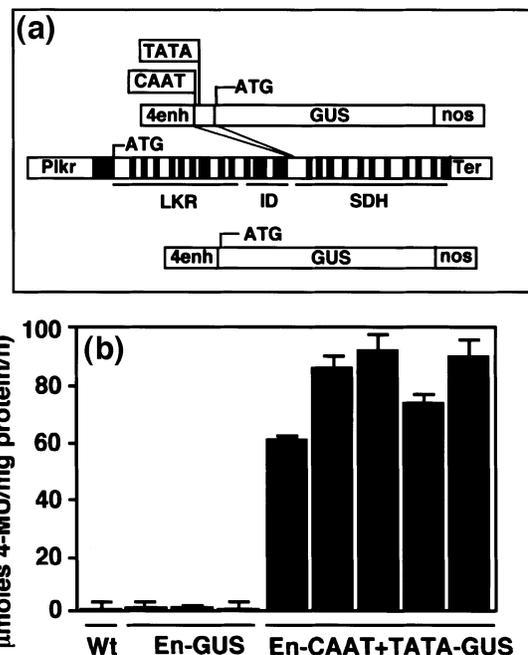


Figure 5. The internal CAAT and TATA box-containing region of the *Arabidopsis* LKR/SDH gene functions as a minimal promoter.

(a) Schematic representation of the chimeric *35S.enh-Psdh.minimal-GUS* construct (top) and its control *35S.enh-GUS* construct (bottom). The *Arabidopsis* LKR/SDH gene from which the internal CAAT plus TATA sequence was taken is shown in the centre. Black rectangles in the *Arabidopsis* LKR/SDH gene indicate introns; empty rectangles indicate exons. Plkr, upstream promoter of the LKR/SDH gene; ATG, translation initiation codons; Ter, the terminator of the LKR/SDH gene; nos, the nopaline synthase gene terminator; 4enh, quadruplicated enhancer of the 35S promoter. Regions encoding the LKR domain, the inter-domain region (ID) and the SDH domain are indicated below the central scheme. (b) GUS activity levels in extracts from inflorescences plus developing siliques of control non-transformed *Arabidopsis* (Wt), control independently transformed *Arabidopsis* plants expressing the *35S.enh-GUS* construct (En-GUS), and independently transformed *Arabidopsis* plants expressing the *35S.enh-Psdh.minimal-GUS* construct (En-CAAT+TATA-GUS). Bars represent the standard error of the mean of triplicate repeats.

a number of genes encoding seed storage proteins and metabolic enzymes and are known to regulate the expression of these genes in a complex developmental and physiological manner. Similar elements were also identified in the upstream promoter of the *AtLKR/SDH* gene (data not shown).

The coding DNA region upstream from the SDH domain functions as an autonomous promoter

To ascertain the functionality of the predicted CAAT and TATA boxes shown in Figure 4(a), a 156bp long DNA fragment containing these boxes was fused between a 5' quadruplicated 35S enhancer and a 3' GUS reporter sequence (Figure 5a, top scheme). As a control, the quadruplicated 35S enhancer was fused directly to the

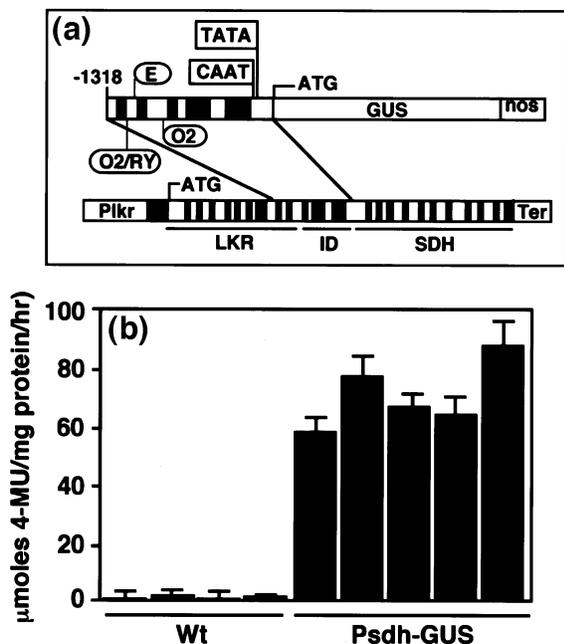


Figure 6. GUS activity levels in transgenic *Arabidopsis* plants expressing the chimeric *Psdh*-GUS gene.

(a) Schematic representation of the chimeric *Psdh*-GUS gene (top) and the *Arabidopsis* *LKR/SDH* gene from which the putative internal SDH promoter was excised (bottom). The positions of the putative TATA and CAAT sequences, as well as the Opaque2 (O2), legumin (RY) and endosperm (E) boxes are also indicated. Black rectangles indicate introns, while empty rectangles indicate exons of the *Arabidopsis* *LKR/SDH* gene (bottom scheme). Plkr, the upstream promoter of the *LKR/SDH* gene; ATG, translation initiation codons; nos, the nopaline synthase gene terminator. The regions encoding the LKR domain, inter-domain region (ID) and SDH domain are indicated below. (b) GUS activity levels in extracts from inflorescences plus developing siliques of control non-transformed *Arabidopsis* (Wt) and independently transformed *Arabidopsis* lines expressing the *Psdh*-GUS construct (Psdh-GUS). Bars represent the standard error of the mean of triplicate repeats.

GUS reporter sequence (Figure 5a, bottom scheme). These constructs were transformed into *Arabidopsis* plants and GUS activity was monitored in protein extracts from inflorescences plus developing siliques of several independently transformed lines. As shown in Figure 5(b), transgenic *Arabidopsis* plants expressing the construct that includes the CAAT and TATA box-containing region had significantly higher GUS activity than control transgenic plants expressing the construct containing only the quadruplicated 35S enhancer.

Next we wished to test whether the coding region upstream from the SDH domain can also function as an autonomous promoter. Thus, we produced a chimeric construct in which a 1318 bp DNA fragment containing part of the coding region upstream from the SDH domain was fused upstream to a GUS reporter gene (Figure 6a). This construct was transformed into *Arabidopsis* plants and GUS levels were measured in inflorescences plus developing siliques from several independently transformed lines. As shown in Figure 6(b), transgenic plants expres-

canola TATTGGATCAAAATAATTGATTCATTAACCTGGTTAGCTAATCCA .GATGA
Caulif. TATTGGATCAAAATAATTGATTCATTAACCTGGTTAGCTAATCCA .GATGA
Arab. TATTGGATCAAAATAATTGATTCATTAACCTGGTTAGCTAATCCA .AATGA
Soybean TTCTGGATCAAAATAATTGATTCATTAACCTGGTTAGCTAATCCA .ACTGA
Maize CATTGGATAAAATATTGATTCCTTGACTCTTTAGCTAATGAA .CATGG

Canola TGAAGAAGACTATATAATATCCACGCAAGAAACAAATAAAATTTTCAC
Caulif. TGAAGAAGAC .TATA.TATCCACGTAGGAATCAAATAGATTTTCAC
Arab. AG ATATA.TATCCCCACATAGGAAGCAAATAGATTTTCAC
Soybean A CATGATA.GATTTCAAATCAAGTTCAAGTAAAATTTTCAC
Maize TGA GATCAGGAGCCGGCAAGAAATGAATTAGCTC

Canola TGAAGATTGAAAAGTCCAGCAAGAAAACCA . . . AGAGAAGCCTGAAGAA
Caulif. TGAAGATTGAAAAGTCCAGCAAGAAAACCA . . . AGAGAAGCCTGAAGAA
Arab. TGAAGATTGAAAAGTCCAGCAAGAAAATCGATATAAAGGAAGCCTGAA
Soybean TTAAGCTTGAAAAGT . . . TGAAGAAATGCGATAGAGAGGAATCTGCAC
Maize TGAAGATAGAAAAGT . . . CATGAATATGAATGACTCCAATGTGATAAA

Canola ATGACAAAGAGATCAGCGGTTTTGATTCCTGGCGCTGGAGTGTGTTGTCG
Caulif. ATGACAAAGAGATCCTCGGTTTTGATTCCTGGCGCTGGAGTGTGTTGTCG
Arab. ATGACGAAAATAATCAGCGGTTTTGATTCCTGGTGTGGAGTGTGTTGTCG
Soybean CCCAGAAAGAGGCTGCGGTTTTAATTCCTGGAGCTGGTGGGCTCTGTCA
Maize GGAGGCCAAGATTTAATTTTGGAGCTGAAGAGTCTGTGGCAGCAGCCG

Figure 7. The internal CAAT and TATA are conserved in canola and cauliflower, but not in maize and soybean *LKR/SDH* genes.

The DNA sequences within the *Arabidopsis*, canola, cauliflower, maize and soybean *LKR/SDH* inter-domain region are aligned using the 'PILEUP' program of the GCG software. The conserved TATA and CAAT boxes as well as the SDH initiator ATG in the *Arabidopsis*, canola and cauliflower *LKR/SDH* genes are in bold and underlined.

sing this GUS construct had significantly higher levels of GUS activity than non-transformed control plants.

Do *LKR/SDH* genes of other plant species possess a promoter upstream of the SDH coding sequence?

To study whether an internal promoter also exists in other plant species, we cloned DNA sequences spanning the inter-domain region of *LKR/SDH* genes from two additional cruciferae plants, *Brassica napus* (canola) and *Brassica oleracea* (cauliflower) (GenBank accession numbers AF191667 and AF191666). The DNA spanning the inter-domain region of the *AtLKR/SDH* gene was then compared with the analogous sequences in canola and cauliflower, as well as with the published maize and soybean *LKR/SDH* cDNAs (GenBank accession number AF003551; Falco *et al.*, 1998). As shown in Figure 7, the CAAT and TATA boxes, as well as the in-frame ATG codon at the beginning of the SDH domain, were conserved among the *LKR/SDH* genes of all three cruciferae plants, but not in the maize and soybean *LKR/SDH* genes.

Discussion

The monofunctional AtSDH is encoded by a distinct gene that is located within the coding and 3' non-coding region of the bifunctional AtLKR/SDH gene

In the present report we analysed the genetic control of the bifunctional *AtLKR/SDH* and the monofunctional *AtSDH*

enzymes in *Arabidopsis*. We showed that a single recombinant *AtLKR/SDH* gene containing an HA epitope tag was able to produce these two enzymes in transgenic *Arabidopsis*. Yet our results also showed that the monofunctional *AtSDH* mRNA was not produced by alternative splicing of a precursor *AtLKR/SDH* mRNA. Rather, three lines of evidence demonstrate that the monofunctional SDH is encoded by a distinct gene that resides within the coding and the 3' non-coding sequence of the *AtLKR/SDH* gene. First, the monofunctional *AtSDH* was produced from a construct lacking the promoter plus the first intron of the *AtLKR/SDH* gene. Second, a 1318 bp DNA fragment located upstream of the SDH domain was proven by its fusion to the GUS reporter to have promoter activity. Third, fusion of the predicted CAAT and TATA boxes of this internal promoter to the GUS reporter showed that these boxes were functional and essential for the production of the monofunctional SDH enzyme. Both the bifunctional *AtLKR/SDH* and the monofunctional *AtSDH* mRNAs are transcribed from the same DNA strand, and the open reading frame of the monofunctional *AtSDH* is a part of the larger *LKR/SDH* open reading frame. In addition, some of the exons of the *AtLKR/SDH* gene, particularly exon 14, serve a dual function, being parts of the translated *AtLKR/SDH* open reading frame as well as parts of the monofunctional *AtSDH* gene promoter. These characteristics render the *AtLKR/SDH* and monofunctional *AtSDH* genes a novel type of a composite locus encoding two metabolically related but distinct enzymes, i.e. a bifunctional *AtLKR/SDH* and a monofunctional *AtSDH*. Although many metabolic enzymes in plants exist as bifunctional polypeptides (for example the bifunctional aspartate kinase/homoserine dehydrogenase enzymes of lysine biosynthesis; reviewed in Galili, 1995), none of the genes encoding them were shown to contain a nested gene encoding only one of the two enzymes. We are also unaware of any previous report describing other types of composite loci encoding functionally distinct proteins in plants. Furthermore, aside from being a novel plant composite locus, *AtLKR/SDH* is also clearly distinct from previously described composite loci in animal cells. The animal composite loci generally fall into two classes: (i) partially overlapping genes, which encode non-related proteins from different DNA strands (Normark *et al.*, 1983; Shintani *et al.*, 1999); and (ii) large genes containing small non-related genes within one of their introns (Henikoff *et al.*, 1986; Pohar *et al.*, 1999).

Expression of AtLKR/SDH and AtSDH is coordinately regulated during Arabidopsis plant development

Using Northern blot analysis, we have previously shown that the levels of the *AtLKR/SDH* and monofunctional *AtSDH* mRNAs are very low in various vegetative tissues, but are coordinately increased in inflorescences and

developing siliques (Tang *et al.*, 1997). This coordinated expression is supported by our present finding that the two promoters possess similar putative binding sites for the transcription factors Opaque2, endosperm box and legumin box. Notably, expression of the *LKR/SDH* gene in maize was recently shown to be regulated by the Opaque2 transcription factor (Kemper *et al.*, 1999). Despite their similar up-regulation in inflorescences and developing siliques, the relative levels of the monofunctional *AtSDH* mRNAs and proteins were several fold higher than their *AtLKR/SDH* counterparts in all of the transgenic lines (Figure 2). Whether these differences reflect differential transcription or post-transcription control remains to be elucidated.

The promoter of the internal monofunctional AtSDH gene is conserved in some but not all plant species

DNA sequence comparisons have shown that the CAAT, TATA and the internal ATG codon of the *AtLKR/SDH* gene are fully conserved in the cruciferae species *Arabidopsis*, canola and cauliflower. In contrast, these elements are not fully conserved in maize and soybean *LKR/SDH* genes, suggesting that the latter two species may not possess a composite *LKR/SDH* locus. This supposition is supported by independent functional observations. Whereas *Arabidopsis* and canola were shown to possess both bifunctional *LKR/SDH* and monofunctional SDH gene products (Deleu *et al.*, 1999; Tang *et al.*, 1997), maize and soybean lack a monofunctional SDH gene product (Kemper *et al.*, 1999; Miron *et al.* 2000). Experiments are currently in progress in our laboratory to ascertain whether the composite *LKR/SDH* locus is unique to cruciferae plants or also exists in other plant species.

Does the composite AtLKR/SDH locus play a regulatory role in lysine catabolism?

An intriguing characteristic of bifunctional *LKR/SDH* polypeptides is the distinct pH optima of their two linked enzymes: around pH 7 for LKR and above pH 9 for SDH (see for example Goncalves-Butruille *et al.*, 1996; Markovitz *et al.*, 1984; Miron *et al.*, 2000). Since plant *LKR/SDH* enzymes are localized in the cytosol (Kemper *et al.*, 1999), the activity of SDH at the neutral pH of this compartment is expected to be relatively inefficient, resulting in transient accumulation of saccharopine (the product of LKR and the substrate of SDH; see Figure 1). On the other hand, plants possessing an additional monofunctional SDH are expected to have a more efficient flux of lysine catabolism. These have indeed been demonstrated empirically (Falco *et al.*, 1995; Mazur *et al.*, 1999). Falco and co-workers found that lysine-overproducing seeds of soybean (lacking a monofunctional SDH) accumulate saccaropine, while

lysine-overproducing seeds of canola (possessing both the bifunctional LKR/SDH and the monofunctional SDH) accumulate α -amino adipic acid as intermediate products of lysine catabolism (see Figure 1). The situation in maize is different to that soybean and canola. Although there is no evidence for the existence of a monofunctional SDH in this plant (Kemper *et al.*, 1999), lysine overproducing maize seeds accumulate relatively low levels of both saccharopine and α -amino adipic acid (Mazur *et al.*, 1999). The molecular basis for this observation has still to be elucidated. The significance of the different pH optima of LKR and SDH and their linkage is not understood. Possibly, these characteristics have evolved to enable cross-regulation of LKR activity by the SDH domain. Indeed, Kemper *et al.* (1998) recently showed that proteolytic fragments from the SDH domain inhibit the activity of LKR. Taken together, we hypothesize that the composite *AtLKR/SDH* locus provides a novel mechanism of metabolic regulation. On the one hand, this locus allows a regulated flux of lysine catabolism via LKR by its linked SDH domain; on the other hand, it also allows an efficient flux of lysine catabolism via SDH by producing excess amounts of this monofunctional enzyme (Figure 2).

The presence of a composite *LKR/SDH* locus in some plant species only, such as the cruciferae, is also puzzling. Since the end-product of lysine catabolism is acetyl-CoA (Figure 1), a metabolite required for various biochemical processes, particularly for lipid biosynthesis (Lehninger, 1975), it is possible that the composite *LKR/SDH* locus has evolved in plant species that require high levels of acetyl-CoA. Cruciferae plants do indeed require high levels of acetyl-CoA for efficient synthesis of storage oils.

Experimental procedures

Materials

Arabidopsis thaliana ecotype C24, as well as canola and cauliflower plants, were grown either in culture or in the greenhouse, as previously described (Shaul *et al.*, 1999). The *Arabidopsis* λ FIX genomic library was kindly provided by Dr F. M. Ausubel via the *Arabidopsis* Resource Center in Columbus, OH, USA. The Ti plasmid, pPZP111, was kindly provided by Dr P. Maliga.

Gene isolation, plasmid construction and plant transformation

Cloning of the *Arabidopsis cAtLKR/SDH* and *cAtSDH* cDNAs has been described previously (Tang *et al.*, 1997). The *Arabidopsis* LKR/SDH gene (GenBank accession no. U95758) was cloned from an *Arabidopsis* genomic library by screening with ³²P-labelled *cAtLKR/SDH* cDNA as a probe, as previously described (Tang *et al.*, 1997). A 9.4 kb *Xba*I fragment containing the LKR/SDH gene, including ~2 kb promoter and ~0.5 kb terminator sequences, was subcloned into the *Xba*I site of Bluescript SK- plasmid to generate *gAtLKR/SDH*.

The *gAtLKR/SDH::HA* clone (illustrated schematically in Figure 2a) was constructed by a modification of *gAtLKR/SDH*. A PCR-derived DNA fragment containing triplet codons for 3 \times haemagglutinin (HA) epitope (3 \times 'YPYDVPDYA'), followed by a TGA stop codon and a 3' terminator derived from the octopine synthase gene of *Agrobacterium tumefaciens*, was fused in-frame to *gAtLKR/SDH*, replacing its natural TGA stop codon plus the 3' terminator. The chimeric *gAtLKR/SDH::HA* gene was subcloned into the polycloning site of the binary Ti plasmid pPZP111 (Hajdukiewicz *et al.*, 1994) to generate pPZP-*gAtLKR/SDH::HA*.

The *gAtLKR/SDH Δ Pro* clone, lacking the upstream *LKR/SDH* gene promoter and the first intron in the 5' non-coding region (illustrated schematically in Figure 3a), was constructed from *gAtLKR/SDH* by digestion at the unique *Xba*I and *Esp*II sites following by blunt end formation and self-ligation of the plasmid. This construct was subcloned into pPZP111 to generate *pPZP-gAtLKR/SDH Δ Pro*.

Psdh-GUS and *Psdh.minimal*-GUS (illustrated schematically in Figures 5a and 6a, respectively) were constructed by combining the respective forward oligonucleotides 5'-GCTCTAGAGTTAA-CCGAGCTACTTTAAT-3' and 5'-GCGGATGATAAGAGAGATTG-3' with the same reverse oligonucleotide 5'-TCCCCTGGTTCA-GGCTTCTCTTTTATCT-3' to amplify the specific genomic fragments of *gAtLKR/SDH*. These sequences were subcloned into an expression cassette containing GUS coding sequences followed by a nopaline synthase terminator of *A. tumefaciens*. The chimeric genes were subcloned into the binary Ti plasmid vector pPZP111 to form *pPZP-Psdh*-GUS and *pPZP-Psdh.minimal*-GUS.

Cloning of the partial *LKR/SDH* genomic DNA sequences from canola and cauliflower was performed by PCR of genomic DNAs, using the two oligonucleotide 5'-GAGCATCGTTGATCATTTTC-3' and 5'-TGATTATGAGCTAATTGTTGG-3'. Genomic DNAs were extracted using the Tri-Reagent (MRC Inc., Cincinnati, OH, USA), according to the protocol provided by the manufacturer. Construction of chimeric genes encoding HA-tagged *Arabidopsis* LKR-SDH and monofunctional SDH in yeast expression vectors will be described elsewhere (Zhu *et al.*, submitted).

All cloned DNAs were sequenced by an automatic sequencer (Model 373A, Version 1.2.0, Applied Biosystems, Foster City, CA, USA).

Transformation of whole *Arabidopsis* plants was performed as previously described (Clough and Bent, 1998).

Northern blot analysis

Extraction of total RNA, fractionation of RNA samples (30 μ g) on 1% agarose gels, transfer to a Hybond-N nylon membrane, hybridization and autoradiography detection were essentially as previously described (Tang *et al.*, 1997). The ³²P-labelled probes included either the 1454 bp *Sal*I to *Nde*I fragment of *cAtSDH* (SDH probe), or a 100 bp PCR fragment containing the 3 \times HA sequence (HA probe). Hybridization was performed in 5 \times SSC (1 \times SSC is 0.15 M NaCl, 0.015 M sodium citrate), 5 \times Denhardt's solution (1 \times Denhardt's solution is 0.02 Ficoll, 0.02% PVP and 0.02% BSA) and 1% SDS. Blots were washed twice for 10 min at 65°C in 1 \times SSPE (1 \times SSPE is 0.15 M NaCl, 10 mM sodium phosphate, 1 mM EDTA, pH 7.5), 0.1% SDS, followed by another wash in 0.1 \times SSPE, 0.1% SDS. Radioactive bands were detected by autoradiography.

SDS-PAGE and Western blot analysis

Plant tissues were homogenized in 25 mM potassium phosphate buffer, pH 7.5, containing 1 mM EDTA, 1 mM DTT, 10 mM PMSF,

10 µg ml⁻¹ leupeptin, 10 µg ml⁻¹ aprotinin and 10 µg ml⁻¹ pepstatin, at 4°C. The homogenate was centrifuged at 18 000 g for 10 min and the supernatant was collected. Yeast cells expressing the *Arabidopsis* LKR/SDH and the monofunctional SDH were homogenized in the same buffer and broken by vortexing with glass beads for 30 min at 4°C. Protein samples (30 µg) were fractionated on 7.5% polyacrylamide–SDS gels (Laemmli, 1970) and transferred to a PVDF membrane, using a Bio-Rad Protein Trans-Blot apparatus. The membranes were completely dried, stained with Coomassie blue R-250, blocked overnight at 4°C in a solution of 5% non-fat dry milk, and then reacted with the anti-HA antibodies for 2 h at room temperature followed by incubation with horseradish-peroxidase-conjugated anti-mouse IgG under the same conditions. Immunodetection was performed with an enhanced chemiluminescence immunodetection kit (Pierce) in accordance with the manufacturer's instructions.

Protein determination and fluorometric analysis of GUS activity

Protein levels were determined according to Bradford (1976). Fluorometric analysis of GUS activity was performed as previously described (Jefferson *et al.*, 1987). GUS activity was expressed as nmol 4-methylumbelliferone (4-MU) mg⁻¹ protein h⁻¹.

Computerized promoter prediction

Prediction of a minimal promoter was obtained by a special Web program (<http://www-hgc.lbl.gov/projects/promoter.html>). DNA and protein sequence comparisons were performed by the 'PILEUP' program of the Genetic Computer Group (GCG) software package (version 8) from the University of Wisconsin, Madison, WI, USA.

Acknowledgements

We thank Professors Jonathan Gressel and Yosef Shaul as well as Mr Yigal Avivi for critical reading of the manuscript and helpful comments. This work was supported by grants from the Framework Program of the Commission of the European Communities, and the Israel Academy of Sciences and Humanities, National Council for Research and Development, Israel, and by grant no. BIO4-CT97-2182. G.T. is supported in part by a Leon and Kathe Fallek scholarship. G.G. is an incumbent of the Bronfman Chair of Plant Sciences.

References

- 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.
- Clough, S.J. and Bent, A.F.** (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735–743.
- Deleu, C., Coustaut, M., Niogert, M.-F. and Larpher, F.** (1999) Three new osmotic stress-regulated cDNAs identified by differential display polymerase chain reaction in rapeseed leaf discs. *Plant Cell Environ.* **22**, 979–988.
- Dickinson, C.D., Evans, R.P. and Nielsen, N.C.** (1988) RY repeats are conserved in the 5'-flanking regions of legume seed-protein genes. *Nucl. Acids Res.* **16**, 371.
- Falco, S.C., Guida, T., Locke, M., Mauvais, J., Sandres, C., Ward, R.T. and Webber, P.** (1995) Transgenic canola and soybean seeds with increased lysine. *Bio/Technology*, **13**, 577–582.
- Falco, S.C., Mcdevitt, R.E. and Epelbaum, S.** (1998) *Chimeric Genes and Methods for Increasing the Lysine Content of the Seeds of Plants*. PCT Patent Application No. WO 98/42831.
- Galili, G.** (1995) Regulation of lysine and threonine synthesis. *Plant Cell*, **7**, 899–906.
- Goncalves-Butruille, M., Szajner, P., Torigoi, E., Leite, A. and Arruda, P.** (1996) Purification of characterization of the bifunctional enzyme lysine-ketoglutarate reductase–saccharopine dehydrogenase from maize. *Plant Physiol.* **110**, 765–771.
- Hajdukiewicz, P., Svab, Z. and Maliga, P.** (1994) The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. *Plant Mol. Biol.* **25**, 989–994.
- Henikoff, S., Keene, M.A., Fichtel, K. and Fristrom, J.W.** (1986) Gene within a gene: nested *Drosophila* genes encode unrelated proteins on opposite DNA strands. *Cell*, **44**, 33–42.
- Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W.** (1987) GUS fusions: β-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6**, 3901–3907.
- Karchi, H., Miron, D., Ben-Yaacov, S. and Galili, G.** (1995) The lysine-dependent stimulation of lysine catabolism in tobacco seeds requires calcium and protein phosphorylation. *Plant Cell*, **7**, 1963–1970.
- Karchi, H., Shaul, O. and Galili, G.** (1994) Lysine synthesis and catabolism are coordinately regulated during tobacco seed development. *Proc. Natl Acad. Sci. USA*, **91**, 2577–2581.
- Kemper, E.L., Cord-Neto, G., Capella, A.N., Goncalves-Butruille, M., Azevedo, R.A. and Arruda, P.** (1998) Structure and regulation of the bifunctional enzyme lysine-oxoglutarate reductase–saccharopine dehydrogenase in maize. *Eur. J. Biochem.* **253**, 720–729.
- Kemper, E.L., Neto, G.C., Papes, F., Moraes, K.C., Leite, A. and Arruda, P.** (1999) The role of Opaque2 in the control of lysine-degrading activities in developing maize endosperm. *Plant Cell*, **11**, 1981–1994.
- Laemmli, U.K.** (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
- Lehninger, A.L.** (1975) *Biochemistry*. New York: Worth Publishers Inc., pp. 572.
- Lohmer, S., Maddaloni, M., Motto, M., Di Fonzo, N., Hartings, H., Salamini, F. and Thompson, R.D.** (1991) The maize regulatory locus Opaque-2 encodes a DNA-binding protein which activates the transcription of the b-32 gene. *EMBO J.* **10**, 617–624.
- Markovitz, P.J., Chuang, D.T. and Cox, R.P.** (1984) Familial hyperlysinemias: purification and characterization of the bifunctional aminoadipic semialdehyde synthase with lysine-ketoglutarate reductase and saccharopine dehydrogenase activities. *J. Biol. Chem.* **259**, 11643–11646.
- Mazur, B., Krebbers, E. and Tingey, S.** (1999) Gene discovery and product development for grain quality traits. *Science*, **285**, 372–375.
- Miron, D., Ben-Yaacov, S., Karchi, H. and Galili, G.** (1997) *In vitro* dephosphorylation inhibits the activity of soybean lysine-ketoglutarate reductase in a lysine-regulated manner. *Plant J.* **12**, 1453–1458.
- Miron, D., Ben-Yaacov, S., Reches, C., Schupper, A. and Galili, G.** (2000) Purification and characterization of bifunctional lysine-ketoglutarate reductase/saccharopine dehydrogenase from developing soybean seeds. *Plant Physiol.*, in press.
- Muller, M. and Knudsen, S.** (1993) The nitrogen response of a

- barley C-hordein promoter is controlled by positive and negative regulation of the GCN4 and endosperm box. *Plant J.* **4**, 343–355.
- Normark, S., Bergström, S., Edlund, T., Grundström, T., Jaurin, B., Lindberg, F.P. and Olsson, O.** (1983) Overlapping genes. *Annu. Rev. Genet.* **17**, 499–525.
- Pohar, N., Godenschwege, T.A. and Buchner, E.** (1999) Invertebrate tissue inhibitor of metalloproteinase: structure and nested gene organization within the synapsin locus is conserved from *Drosophila* to human. *Genomics*, **57**, 293–296.
- Pysh, L.D., Aukerman, M.J. and Schmidt, R.J.** (1993) OHP1: a maize basic domain/leucine zipper protein that interacts with opaque2. *Plant Cell*, **5**, 227–236.
- Shaul, O., Hilgemann, D., Engler, J., Van Montagu, M., Inze, D. and Galili, G.** (1999) Cloning and characterization of a novel Mg²⁺/H⁺ exchanger. *EMBO J.* **18**, 3973–3980.
- Shintani, S., O’Huigin, C., Toyosawa, S., Michalová, V. and Klein, J.** (1999) Origin of gene overlap: the case of TCP1 and ACAT2. *Genetics*, **152**, 743–754.
- Tang, G., Miron, D., Zhu-Shimoni, J.X. and Galili, G.** (1997) Regulation of lysine catabolism through lysine-ketoglutarate reductase and saccharopine dehydrogenase in Arabidopsis. *Plant Cell*, **9**, 1305–1316.