Plant Molecular Biology **32:** 475–484, 1996.

Molecular characterization of transketolase (EC 2.2.1.1) active in the Calvin cycle of spinach chloroplasts

Anke Flechner[1], Uta Dressen[2], Peter Westhoff[2], Katrin Henze[3], Claus Schnarrenberger[1] and William Martin[3,*]

[1] *Institut für Pflanzenphysiologie und Mikrobiologie Freie Universität Berlin, Königin-Luise-Strasse 12–16a, D-14195 Berlin, Germany;* [2] *Institut für Entwicklungs- und Molekularbiologie der Pflanzen Universitätsstrasse 1, Universität Düsseldorf, D-40225 Düsseldorf, Germany;* [3] *Institut für Genetik Technische Universität Braunschweig Spielmannstrasse 7, D-38106 Braunschweig, Germany (*author for correspondence)*

Received 5 March 1996; accepted in revised form 26 June 1996

Key words: carbon fixation, chloroplasts, evolution, isoenzymes

Abstract

A cDNA encoding the Calvin cycle enzyme transketolase (TKL; EC 2.2.1.1) was isolated from *Sorghum bicolor* via subtractive differential hybridization, and used to isolate several full-length cDNA clones for this enzyme from spinach. Functional identity of the encoded mature subunit was shown by an 8.6-fold increase of TKL activity upon induction of *Escherichia coli* cells that overexpress the spinach TKL subunit under the control of the bacteriophage T_7 promoter. Chloroplast localization of the cloned enzyme is shown by processing of the *in vitro* synthesized precursor upon uptake by isolated chloroplasts. Southern blot-analysis suggests that TKL is encoded by a single gene in the spinach genome. TKL proteins of both higher-plant chloroplasts and the cytosol of non-photosynthetic eukaryotes are found to be unexpectedly similar to eubacterial homologues, suggesting a possible eubacterial origin of these nuclear genes. Chloroplast TKL is the last of the demonstrably chloroplast-localized Calvin cycle enzymes to have been cloned and thus completes the isolation of gene probes for all enzymes of the pathway in higher plants.

Abbreviations: RPE, ribulose-5-phosphate 3-epimerase; RPI, ribose-5-phosphate isomerase; TKL, transketolase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; PGK, phosphoglycerate kinase; FBP, fructose-1,6-bisphosphatase; SBP, sedoheptulose-1,7-bisphosphatase; OPPP, oxidative pentose phosphate pathway; Rubisco, ribulose-1,5-bisphosphate carboxylase/oxygenase; FBA, fructose-1,6-bisphosphate aldolase; IPTG, isopropyl β-D-thiogalactoside; TPI, triosephosphate isomerase

Introduction

Transketolase (EC 2.2.1.1) is integral to both the Calvin cycle and the oxidative pentose phosphate pathway (OPPP) of higher-plant chloroplasts. In the Calvin cycle, it catalyses the transfer of a two-carbon ketol group from either D-fructose-6-phosphate or D-sedoheptulose-7-phosphate to D-glyceraldehyde-3-phosphate to yield D-xylulose-5-phosphate and either D-erythrose-4-phosphate or D-ribose-5-phosphate, respectively (Fig. 1). In the OPPP, these reversible reactions proceed in the opposite direction. From all sources studied to date, transketolase (TKL) is thiamine pyrophosphate-dependent and is a dimer of 74 kDa subunits [8, 29, 44]. The enzyme has been cloned from a number of eukaryotic sources [2] and the three-dimensional stucture of the yeast enzyme has been determined [30]. In photosynthetic purple bacteria, TKL is encoded in Calvin-Benson-Bassham (*cbb*) operons where it is expressed in a coordinated manner with other enzymes of the cycle [3, 13, 14]. Both *E. coli* [18] and yeast [37] have been shown to possess two active transketolase genes. In the methylo-

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers L76554 (TKL) and H55032 (pHHU58).

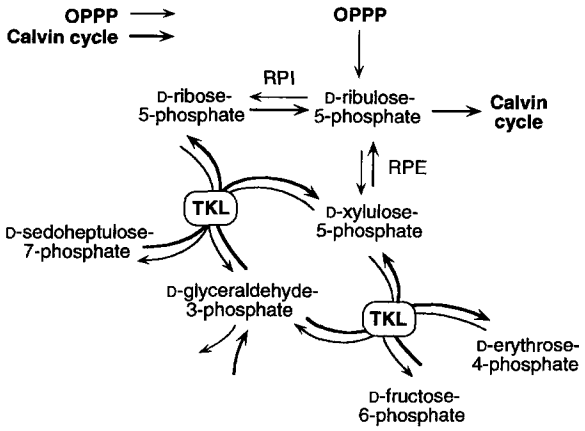

Figure 1. Function of transketolase in plant metabolism. Arrows indicate metabolite flux for the pathways indicated. RPI, ribose-5-phosphate isomerase; RPE, ribulose-5-phosphate 3-epimerase, TKL, transketolase; OPPP, oxidative pentose phosphate pathway.

trophic yeast *Hansenula polymorpha* TKL functions as a formaldehyde-accepting dihydroxyacetone synthase (DHAS) during methanol assimilation [19].

Comparatively little is known about the plant transketolase enzyme. Some previous studies indicated that higher plants in general might possess chloroplast and cytosolic TKL isoenzymes [21, 29], whereas others reported that most, if not all, TKL activity is localized to the chloroplast, at least in spinach leaves [11, 38]. At the molecular level, cDNA clones for three different transketolases were recently isolated from rehydrating tissues of the resurrection plant *Craterostigma plantagineum* [2] although it is not clear whether these cDNAs encode chloroplast (Calvin cycle) or cytosolic enzyme(s). They might play a role in the remarkable octulose metabolism of this totally dehydratable plant.

In a previous report, we exploited the differential expression of Calvin cycle genes in mesophyll and bundle sheath cells of the C4 plant *Sorghum bicolor* to isolate molecular probes with which to clone the Calvin cycle enzyme ribulose-5-phosphate 3-epimerase [31] from spinach. Here we have followed the same approach to clone the functional transketolase enzyme of spinach chloroplasts.

Materials and methods

cDNA cloning

Using the differential hybridization techniques previously described [31], bundle sheath-specific cDNA clones were isolated from a *Sorghum bicolor* bundle sheath cDNA library. One of these, designated as pHHU58, was shown by subcloning and terminal sequencing to encode a protein with very high similarity to published transketolase sequences. This indicated that the 1.2 kb *Xho*I-*Eco*RI insert of plasmid pHHU58 (accession number H55032) putatively encodes a partial cDNA for *Sorghum* TKL. The *Sorghum* cDNA insert was then used as a heterologous probe to screen our spinach cDNA library constructed from light-grown seedling mRNA. Details of construction of the spinach cDNA library are described elsewhere [16]. Plaque hybridization and washing during screening of the spinach cDNA library was performed as described by Henze *et al.* [16] but with 20 ng/ml of radioactive probe and at 60 °C. Other molecular methods were performed as described by Nowitzki *et al.* [31].

Construction of expressing strains

For expression under the control of the T_7 promoter [41], the region corresponding to the mature subunit was amplified in 25 μl for 30 cycles of 1 min at 93 °C, 1 min at 62 °C, and 2.5 min at 72 °C in 50 mM KCl, 10 mM Tris-HCl, 0.1% Triton, 1.0 mM $MgCl_2$, 200 μM of each dNTP, pH 9.0, 0.01 U/μl Tli polymerase (Promega) and 0.8 μM each of 5′-GCGTAGCAT**ATG**TTGGAGAGTACTGATATTGAC-3′ and 5′-GCGTAGGGATCC**CTA**ACAAACTGATT TTGCAGC-3′ (start and stop codons in bold face). Amplification products were purified on Microcon 30 devices (Amicon), cut with *Bam*HI and *Nde*I and ligated into *Bam*HI/*Nde*I-cut pET-3a [41]. *Ampf* transformants in *E. coli* nm522 (Stratagene) were checked by restriction and terminal sequencing. The plasmid pTK39$\Delta^{71}T_7$ from one of these was transformed into *E. coli* BL21lys [41] to yield the strain BL21lys/pTK39$\Delta^{71}T_7$ for further analyses.

Expression in E. coli

BL21lys/pTK39$\Delta^{71}T_7$ was grown overnight in 5 ml of LB medium at 37 °C containing 100 mg/l ampicillin. Aliquots of 1 ml were transferred to 25 ml of fresh LB medium containing 2% glucose, grown at 22 °C to a final OD_{600} of 0.6, and induced by addition of 0.4 mM IPTG. Samples (1 ml) were taken at 30 min intervals. Cells were harvested, washed with lysis buffer (50 mM Tris-HCl, 2 mM EDTA, pH 8.0), recentrifuged and suspended in 400 μl lysis buffer containing 100 μg/ml

lysozyme. After incubation for 15 min at 30 °C, cells were vortexed and centrifuged. Ten μl of supernatant was used for SDS-PAGE [36] and 15 μl for activity assay. Gels were stained for proteins with Coomassie Brilliant Blue G250. BL21lys containing pET-3a as a control was handled in the same manner. Protein was estimated by the method of Bradford as described [16].

Enzyme assay

Transketolase activity was assayed at 25 °C in 1 ml of 73 mM Tris-HCl pH 7.5, 3.7 mM $MgCl_2$, 0.25 mM thiamin pyrophosphate, 0.21 mM NADH, 1 U/ml glycerol-3-phosphate dehydrogenase, 1 U/ml triosephosphate isomerase, and 40 μl of substrate. Substrate was prepared by incubation of 50 mM ribose-5-phosphate with 5 U/ml each of ribulose-5-phosphate 3-epimerase and ribose-5-phosphate isomerase for 1 h. The commercially available yeast enzyme (Sigma) was used as a control. One unit is the amount of enzyme required to convert 1 μmol of substrate per minute.

In vitro *transcription, translation and import*

pTK39 was linearized with *Sac*I, purified, transcribed in the presence of 1 mM ^{7}mG(5′)ppp(5′)G (New England Biolabs) with T_7 RNA polymerase (Pharmacia) and translated as described [27]. Preparation of intact chloroplasts from 80 g of 14-day old pea leaves and import were performed according to Clausmeyer *et al.* [6] with a protocol kindly provided by R.-B. Klösgen (München). Proteins were electrophoresed on 12% denaturing acrylamide gels [36], impregnated with scintillator (Amplify, Amersham) subsequent to electrophoresis, dried, and subjected to autoradiography at −80 °C.

Genomic Southern blots

Twenty μg of spinach genomic DNA was digested with the appropriate enzyme, electrophoresed, blotted and hybridized as described by Henze *et al.* [16]. The hybridization probe was 10 ng/ml of the 2.5 kb *Not*I insert of pTK39 random-labelled to 10^8 cpm/μg with α-^{32}P dCTP. Filters were washed for 60 min at 68 °C in 2 × SSPE, 0.1% (w/v) SDS and autoradiographed for two days on XAR films.

Sequence analysis

Standard sequence processing was performed with the Wisconsin package [9]. Sequence alignments were created with CLUSTAL V [17] and refined by eye. Distance between sequences for phylogenetic inference with the neighbor-joining method [35] was measured as numbers of amino acid substitutions per site corrected for multiple substitutions with the Dayhoff matrix of PHYLIP [12].

Results

From the differential screening of the *Sorghum* cDNA library with mesophyll and bundle-sheath-specific radiolabelled cDNA probes we obtained 62 different cDNA clones which were characterized by partial sequencing and RNA expression analysis (data not shown). The encoded protein of one of these, pHHU58 (accession number H55032), revealed 40–45% identity over 315 amino acids to *Alcaligenes* and *Rhodobacter* transketolase proteins. The 930 bp pHHU58 insert was used to screen 100 000 recombinants of a non-amplified spinach seedling cDNA library, yielding 20 independent positively hybridizing cDNAs. Seven clones containing *Not*I inserts in the range of 2.0–2.3 kb in length were subcloned into pSK vectors (Stratagene) (pTK21, pTK24, pTK26, pTK29, pTK39 pTK40, pTK46) and shown by terminal sequencing to be identical to pTK39 (Fig. 2), suggesting that they represent transcripts from one and the same gene. A Southern blot of spinach DNA was probed with the complete 2.6 kb insert of pTK39 under low-stringency conditions (2 × SSPE, 0.1% SDS, 68 °C), revealing only a single band of 4.6 and 11 kb with *Eco*RV and *Hpa*I, respectively (Fig. 3). Taken together, these results indicate that pTK39 represents the major transketolase-related transcript in spinach leaves and that it is encoded by a single gene. pTK39 shares 77% DNA sequence identity over 930 bases with the *Sorghum* cDNA pHHU58 (data not shown).

The complete sequence of pTK39 is 2595 bp long with an open reading frame of 2223 bp, 52 bp of 5′-untranslated region (UTR) and a 235 bp 3′ UTR; four of the six cDNA clones isolated extended into the 5′ UTR of this 2.6 kb transcript (Fig. 2). pTK39 posesses a polyT tract at its 5′ end. It is neither clear whether this is cDNA cloning artefact or whether it is contained within the mRNA. We have observed similar homopolymer tracts at the 5′ end of clones for chloroplast triosephos-

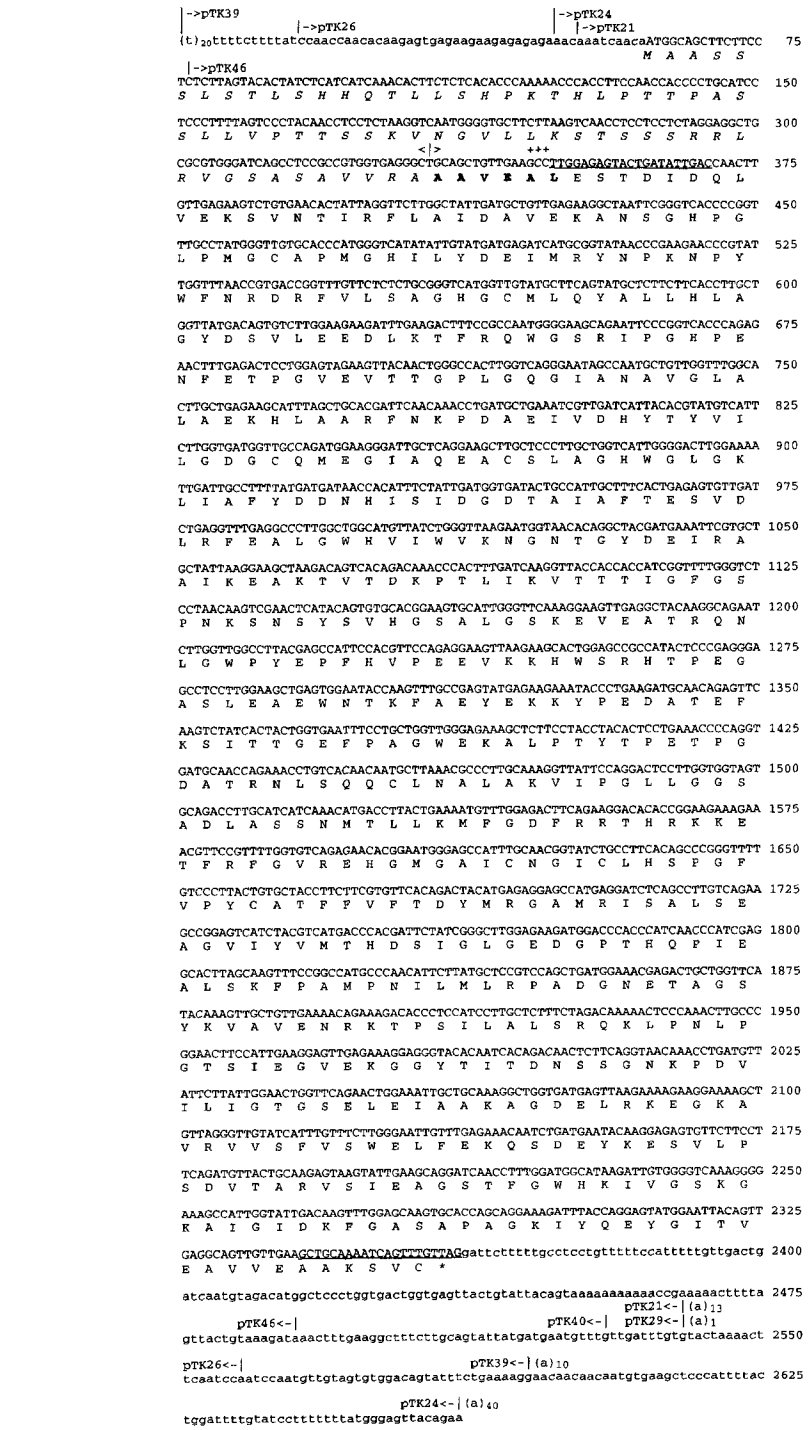

Figure 2. Sequence of cDNA clones for transketolase from spinach. Terminal nucleotides of clones pTK21, pTK46, pTK40, pTK29, pTK26, pTK39, and pTK24 are indicated by vertical bars. The 66 amino acid transit peptide is shown in italics. Primers used for construction of expression plasmids in the T_7 system [41] are underlined, the start codon used in T_7 expression is marked (+ + +). The first residues of the N-terminus of purified protein [43] is indicated in bold type. Non-coding regions are shown in lower case.

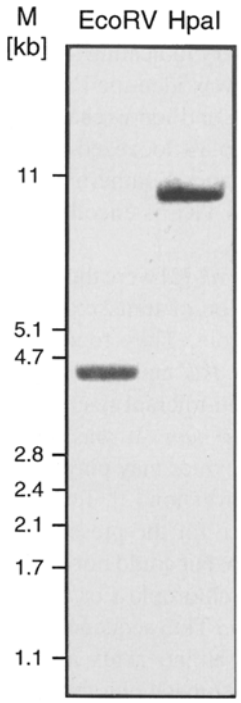

Figure 3. Southern blot of genomic DNA from spinach probed with 2.5 kb *Not*I insert of pTK39. Twenty μg of spinach DNA was loaded per lane. The film was exposed for 48 h at −80 °C.

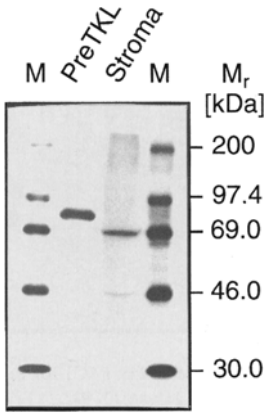

Figure 4. Import of *in vitro* synthesized spinach transketolase precursor into pea chloroplasts. The molecular masses of methylated ^{14}C-labelled standards (Amersham) are indicated. M, molecular mass standards. PreTKL, precursor synthesized *in vitro* with *in vitro* transcribed mRNA from pTK39 prior to transport; Stroma, stroma proteins after transport. The film was exposed for seven days at −80 °C.

phate isomerase and chloroplast ribulose-5-phosphate 3-epimerase which were isolated from this same spinach cDNA preparation, but only in the longest cDNAs for each of these genes [16, 31].

pTK39 encodes the cytosolic precursor of chloroplast TKL

The predicted protein encoded by pTK39 is 741 amino acids long and has an expected size of 80.2 kDa. *In vitro* transcription and translation of the pTK39 encoded protein yields a product of 80 kDa (Fig. 4), in good agreement with the predicted size. The coding region of pTK39 which shares sequence similarity to transketolases from other sources (see below) is preceded by a 60–70 amino acid serine-rich amino-terminal extension with typical properties of chloroplast transit peptides [45]. The 80 kDa translation product of pTK39 is processed to ca. 70 kDa upon import *in vitro* into pea chloroplasts (Fig. 4), indicating that the transit peptide-like region is indeed functional, and that the cDNA encodes a chloroplast protein. The N-terminal sequence of purified spinach chloroplast transketolase was recently determined [43] and is found in the protein encoded by pTK39 (in bold face in Fig. 2), thereby defining the processing site of the 66 amino acid transit peptide. The predicted size of the 675 amino acid spinach chloroplast TKL mature subunit is 73.5 kDa, very close to the approximate value of 70 kDa observed in Fig. 4 and in very good agreement with previous estimates of 74 kDa obtained from studies of the native protein [11].

Transketolase encoded by pTK39 is enzymatically active

In order to determine whether pTK39 encodes a functional enzyme, the region corresponding to the mature chloroplast protein was amplified and cloned into pET-3a and transformed into *E. coli* BL21lys as described in Materials and methods to yield the strain BL21lys/pTK39$\Delta^{71}T_7$. When BL21lys/pTK39$\Delta^{71}T_7$ cultures were induced with IPTG at 37 °C, large amounts of a roughly 70 kDa protein were accumulated in an induction-dependent manner, but no increase of TKL activity was observed relative to the control lacking the 2 kb cDNA insert (data not shown). When cultures were induced and cultivated at 22 °C, a less prominent accumulation of the 70 kDa protein was observed (Fig. 5), but transketolase activity increased in BL21lys/pTK39$\Delta^{71}T_7$ cultures to levels 8.6-fold greater than that observed in the pET-3a control (Table 1). The specific TKL activity in cleared BL21lys/pTK39$\Delta^{71}T_7$ crude extracts 90 min after

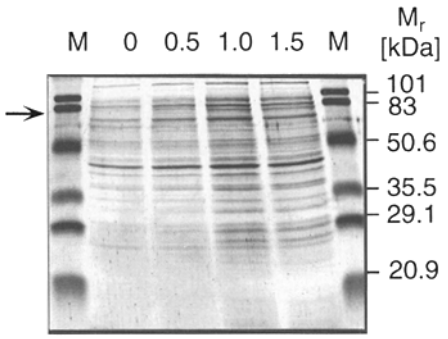

Figure 5. Coomassie-stained SDS-PAGE of induction kinetics of TKL expressed in *E. coli* BL21lys/pTK39Δ^{71}T$_7$. M, molecular mass standards. Aliquots were taken from the culture at various times [h] after induction. The arrow indicates the induced 74 kDa TKL protein. Corresponding specific TKL activities are given in Table 1.

induction was about 4 U/mg, which compares favourably with the specific activity of 15–25 U/mg for the purified yeast enzyme commercially available from Sigma.

Discussion

Cloning of chloroplast TKL of spinach

Six cDNA clones from spinach which encode chloroplast transketolase were isolated of which four represented full-size transcripts for a ca. 2.6 kb mRNA (Fig. 2). The serine-rich N-terminal extension preceding the mature subunit was shown *via in vitro* transcription, translation, and import into pea chloroplasts to encode a functional transit peptide (Fig. 4), indicating that the mRNA codes for a cytosolic precursor of the mature chloroplast enzyme. We expressed the pTK39-encoded protein without its transit peptide in *E. coli* under the control of a phage T$_7$ promoter (Fig. 5). The 74 kDa subunit was accumulated at high levels though inactive when expressed at 37 °C. When *E. coli* was incubated at 22 °C subsequent to IPTG induction, lower levels of protein accumulation were observed, but transketolase activity in the cleared supernatant increased to 4.2 U/mg, almost 9-fold over the transketolase activity in control cells (Table 1). These data demonstrate the cloning of active, nuclear-encoded transketolase present in chloroplasts of spinach leaves. As summarized in Table 2, transketolase is the last of the Calvin cycle enzymes for which the demonstrably chloroplast-localized protein has been cloned from higher plants.

We detected no other forms of TKL-related cDNA clones in the library indicating that the cloned enzyme is the TKL that was identified in previous studies by cell fractionation and ion-exchange techniques as the prevalent, chloroplast-localized TKL [11, 38]. Furthermore, low stringency Southern hybridization (Fig. 3) suggests that this TKL is encoded by a single nuclear gene in spinach.

Bernacchia *et al.* [2] were the first to report molecular characterization of transketolase from any photosynthetic eukaryote. They found three transketolase transcripts – *tkt3*, *tkt7* and *tkt10* – which are expressed in the desiccation-tolerant resurrection plant *Craterostigma plantagineum*. It was suggested that these transketolase enzymes may play a role in desiccation-related sugar conversion [2]. Teige *et al.* [42] recently reported evidence for the presence of only one TKL enzyme in potato, but could not discern whether it was localized to the chloroplast or the cytosol. Although the *Craterostigma* Tkt3 sequence is incomplete (about 160 N-terminal amino acids are lacking [2]), it is more similar to spinach chloroplast TKL than Tkt10 and Tkt7 are (Fig. 6). Based on sequence similarity to spinach chloroplast TKL, it is possible that the cDNA clones Tkt3 from *Craterostigma* and that from *Solanum* also encode chloroplast enzymes. The other two clones from *Craterostigma* (Tkt10 and Tkt7) may encode enzymes related to the particular desiccation physiology of this plant and conceivably could be cytosolic isoenzymes, since their sequences appear to lack recognizable transit peptides [2].

Transketolase gene evolution

An alignment of spinach chloroplast transketolase with homologous sequences from eubacterial and eukaryotic sources (available upon request) reveals a rather uniform distribution of highly conserved regions across 760 total positions. As evident in previous TKL alignments [2], the region of greatest variability is found in the central third of the protein, where the sequences from mammals (man, mouse, and rat) possess a ca. 40 amino acid deletion. Transketolase sequences from eubacteria and non-mammalian eukaryotes share (roughly) 50% amino acid identity, whereas plant TKL shares only (roughly) 30% identity with mammalian homologues. The neighbor-joining tree shown in Fig. 6 reveals that the formaldehyde-accepting transketolase (dihydroxyacetone synthase) of *Hansenula polymorpha*, which is involved in methanol assimilation [19] is clearly related to conven-

Table 1. Transketolase activity in cleared crude extracts of *E. coli* BL21lys/pTK39$\Delta^{71}T_7$ strain and the same strain harboring pET-3a without the spinach transketolase cDNA insert. Both strains were grown at 22 °C after IPTG induction.

Time after induction (h)	Transketolase activity (U/mg)	
	BL21lys/pET-3a	BL21lys/pTK39$\Delta^{71}T_7$
0	0.09	0.29
0.5	0.33	1.09
1.0	0.25	1.77
1.5	0.48	4.16

Table 2. Calvin cycle enzymes of higher plant chloroplasts. Except for RbcL, all of the proteins are nuclear-encoded. References give the initial report of molecular characterization of the chloroplast enzyme. * this paper. ‡ Fructose-1,6-bisphosphate aldolase catalyzes the aldol condensation of dihdroxyacetone phosphate with either glyceraldehyde-3-phosphate or erythrose-4-phosphate, and thus also perfoms the sedoheptulose-1,7-bisphosphate aldolase reaction in the cycle [5].

Enzyme	EC number	Protein symbol	Ref.
Ribulose-1,5-bisphosphate carboxylase/oxygenase, SSU	4.1.1.39	RbcS	1
Ribulose-1,5-bisphosphate carboxylase/oxygenase, LSU	4.1.1.39	RbcL	26
Glyceraldehyde-3-phosphate dehydrogenase, GAPDH subunit A	1.2.1.13	GapA	23
Glyceraldehyde-3-phosphate dehydrogenase, GAPDH subunit B	1.2.1.13	GapB	40
Fructose-1,6-bisphosphatase	3.1.3.11	FBP	33
Phosphoribulokinase	2.7.1.19	PRK	28
3-Phosphoglycerate kinase	2.7.2.3	PGK	22
Sedoheptulose-1,7-bisphosphatase	3.1.3.37	SBP	34
Fructose-1,6-bisphosphate aldolase‡	4.1.2.13	FBA	32
Triosephosphate isomerase	5.3.1.1	TPI	16
Ribulose-5-phosphate 3-epimerase	5.1.3.1	RPE	31
Ribose-5-phosphate isomerase	5.3.1.6	RPI	24
Transketolase	2.2.1.1	TKL	*

tional transketolases from the pentose phosphate pathways and was probably derived from these (Fig. 6). However its long branch, similar to that for mammalian homologues, suggests an elevated substitution rate for the enzyme which probably relates to its unusual formaldehyde-accepting function. Gene duplications have resulted in *tkl* gene families of three members in *Craterostigma plantagineum*, and two members each in *E. coli* and yeast. The *E. coli* and *Craterostigma* duplications appear to have occurred prior to plant speciation processes reflected in the tree. The complete sequence of the *Haemophilus influenzae* genome contains only one *tkl* gene, suggesting that differential gene loss has occurred subsequent to the duplication event which gave rise to *E. coli tklA* and *tklB*, as in the case of eubacterial *rbc*L [25] and GAPDH genes [15].

With the exception of DHAS from *Hansenula* and the deletion-containing mammalian sequences, divergence within eubacterial and eukaryotic TKL genes surveyed, respectively, is not appreciably greater than across the eukaryote-eubacterial boundary; many branches separating proteobacteria are more robust than those separating eubacterial and eukaryotic sequences (Fig. 6). This is a noteworthy find-

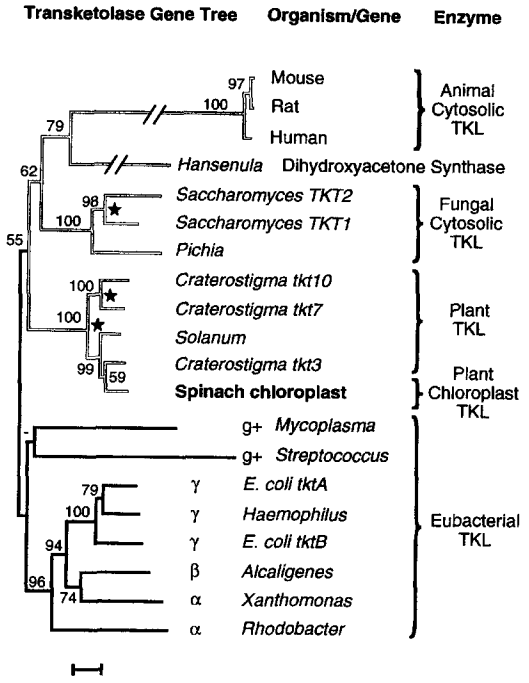

Figure 6. Transketolase gene phylogeny. The tree was constructed by the neighbor-joining method from a matrix of estimated numbers of amino acid substitutions per site calculated with the Dayhoff option of PHYLIP. Numbers near branches indicate the bootstrap proportion for 100 replicas using the same method; – indicates a bootstrap proportion of less than 50. The scale bar indicates 0.1 substitutions per site. Genes encoded in nuclear DNA are borne on open branches, those encoded in eubacterial DNA are borne on solid branches. α, β, and γ, proteobacterial groups for eubacteria; g+, gram-positive. Sequences were extracted from GenBank. The position of the gram positive bacteria may be a branch attraction artefact. For convenience, the very long branches leading to mammalian TKL and DHAS of *Hansenula* are drawn to only half of their true length (indicated by //). Recognizeable gene duplications are indicated by stars.

ing, because eubacterial and eukaryotic lineages are believed to have separated very early in evolution [10, 46, 47], and one would expect to find greater divergence across that split in Fig. 6, were *tkl* genes related through eubacterial-eukaryote separation. The general pattern of similarity observed between eubacterial and plant *tkl* genes is strikingly similar to that found for plant GAPDH [15] and plant PGK [4], where the respective genes for both chloroplast and cytosolic plant enzymes have been shown to derive from (endosymbiotic) eubacteria. Although TKL sequences from early branching eukaryotes and archaebacteria are not available for comparison, the most straightforward interpretation of the current findings is that TKL genes of plants, and perhaps of other eukaryotes, are also simply of eubacterial origin. For TKL, the phylogeny of proteobacterial sequences differs from that observed for rRNA genes [7]. This may be due to paralogy ensuing from differential loss of members of a eubacterial gene family, as for *rbc*L [25] and GAPDH [15] or may simply relate to the sequence characteristics and small bacterial sample available for analysis. It also cannot be excluded that interspecies transfer of genes for sugar phosphate metabolism between eubacteria may also have occurred in evolution. Interspecies gene transfer is surprisingly common between contemporary eubacteria [20], and it may have spread earlier in evolution as well. Although such events would complicate the eubacterial phylogeny, they would not be able to account for the finding that divergence within eubacterial *tkl* genes is as great as that found across the eubacterial/eukaryotic split, whereas a eubacterial origin of eukaryotic *tkl* genes easily could.

Conclusion

As for two other enzymes common to the OPPP and Calvin cycle in spinach leaves, ribulose-5-phosphate 3-epimerase (RPE) and ribose-5-phosphate isomerase (RPI) [38], transketolase occurs predominantly, if not exclusively, as a single chloroplast enzyme and apparently lacks a cytosolic counterpart in this tissue [11, 38]. It thus appears that chloroplast transketolase functions in both the reductive and oxidative pentose phosphate pathways in spinach leaves. The evolutionary analysis suggests TKL may be the next link in an increasingly long chain of unexpectedly eubacterial genes in eukaryotic genomes.

Acknowledgements

W.M. and C.S. gratefully acknowledge financial support from the Deutsche Forschungsgemeinschaft (Ma 1426/3-2 and Schn 28/21-2). The work of P.W. was funded in part by the European Communities' Biotech Programme, as part of the Project of Technological Priority 1993–1996, and by a grant from the German-Israeli Foundation (I 0259-153.03/92).

References

1. Bedbrook JR, Smith SM, Ellis RJ: Molecular cloning and sequencing of cDNA encoding the precursor to the small

subunit of chloroplast ribulose-1,5-bisphosphate carboxylase. Nature 287: 692–697 (1980).
2. Bernacchia G, Schwall G, Lottspeich F, Salamini F, Bartels D: The transketolase gene family of the resurrection plant *Craterostigma plantagineum*: differential expression during the rehydration phase. EMBO J 14: 610–618 (1995).
3. Bowien B, Bednarski R, Kusian B, Windhövel U, Freter A, Schäferjohann J, Yoo J-G: Genetic regulation of CO_2 assimilation in chemoautotrophs. In: Murrel JC, Kelley DP (eds) Microbial Growth on C1 Compounds, pp. 481–491. Intercept Scientific, Andover, UK (1993).
4. Brinkmann H, Martin W: Higher plant chloroplast and cytosolic 3-phosphoglycerate kinases: a case of endosymbiotic gene replacement. Plant Mol Biol 30: 65–75 (1996).
5. Brooks K, Criddle RS: Enzymes of the carbon cycle of photosynthesis. I. Isolation and properties of spinach chloroplast aldolase. Arch Biochem Biophys 117: 650–659 (1966).
6. Clausmeyer S, Klösgen R-B, Herrmann RG: Protein import into chloroplasts: the hydrophilic lumenal proteins exhibit unexpected import and sorting specificities in spite of structurally conserved transit peptides. J Biol Chem 268: 13869–13876 (1993).
7. De Rijk P, Van de Peer Y, Van den Brock I, De Wachter R: Evolution according to large ribosomal subunit RNA. J Mol Evol 41: 366–375 (1995).
8. de la Haba G, Leder IG, Racker E: Crystalline transketolase from bakers' yeast: isolation and properties. J Biol Chem 214: 409–426 (1955).
9. Devereux J, Haeberli P, Smithies O: A comprehensive set of sequence analysis programms for the VAX. Nucl Acids Res 12: 387–395 (1984).
10. Doolittle WF, Brown JR: Tempo, mode, the progenote, and the universal root. Proc Natl Acad Sci USA 91: 6721–6728 (1994).
11. Feierabend J, Gringel G: Plant transketolases: subcellular distribution, search for multiple forms, site of synthesis. Z Pflanzenphysiol 110: 247–258 (1983).
12. Felsenstein J: PHYLIP: Phylogeny Inference Package (version 3.2). Cladistics 5: 164–166 (1989).
13. Gibson JL, Tabita FR: Localization and mapping of CO_2 fixation genes within two gene clusters in *Rhodobacter sphaeroides*. J Bact 170: 2153–2158 (1988).
14. Gibson JL, Tabita FR: Nucleotide sequence and functional analysis of CbbR, a positive regulator of the Calvin cycle operons of *Rhodobacter sphaeriodes*. J Bact 175: 5778–5784 (1993).
15. Henze K, Badr A, Wettern M, Cerff R, Martin W: A nuclear gene of eubacterial origin in *Euglena gracilis* reflects cryptic endosymbioses during protist evolution. Proc Natl Acad Sci USA 92: 9122–9126 (1995).
16. Henze K, Schnarrenberger C, Kellermann J, Martin W: Chloroplast and cytosolic triosephosphate isomerase from spinach: purification, microsequencing and cDNA sequence of the chloroplast enzyme. Plant Mol Biol 26: 1961–1973 (1994).
17. Higgins DH, Sharp PM: CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. Gene 73: 237–244 (1988).
18. Iida A, Teshiba S, Mizobushi K: Identification and characterization of the *tktB* gene encoding a second transketolase in *Escherichia coli* K-12. J Bact 175: 5375–5383 (1993).
19. Janowicz ZA, Eckart MR, Drewke C, Roggenkamp RO, Hollenberg CP, Maat J, Ledeboer AM, Visser C, Verrips CT: Cloning and characterization of the DAS gene encoding the major methanol assimilatory enzyme from the methylotrophic yeast *Hansenula polymorpha*. Nucl Acids Res 13: 3043–3062 (1985).
20. Lan R, Reeves PR: Gene transfer is a major factor in bacterial evolution. Mol Biol Evol 13: 47–55 (1996).
21. Latzko E, Gibbs M: Distribution and activity of enzymes of the reductive pentose phosphate cycle in spinach leaves and in chloroplast isolated by different methods. Z Pflanzenphysiol 59: 184–194 (1968).
22. Longstaff M, Raines CA, McMorrow E, Bradbeer JW, Dyer TA: Wheat phosphoglycerate kinase: evidence for recombination between the chloroplastic and cytosolic enzymes. Nucl Acids Res 17: 6569–6580 (1989).
23. Martin W, Cerff R: Prokaryotic features of a nucleus encoded enzyme: cDNA sequences for chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenases from mustard (*Sinapis alba*). Eur J Biochem 159: 323–331 (1986).
24. Martin W, Henze K, Kellermann J, Flechner A, Schnarrenberger C: Microsequencing and cDNA cloning of the Calvin cycle/OPPP enzyme ribose-5-phosphate isomerase (EC 5.3.1.6) from spinach chloroplasts. Plant Mol Biol 30: 795–805 (1996).
25. Martin W, Somerville CC, Loiseaux-de Goër S: Molecular phylogenies of plastid origins and algal evolution. J Mol Evol 35: 385–403 (1992).
26. McIntosh L, Poulsen C, Bogorad L: Chloroplast gene sequence for the large subunit of ribulose bisphosphatecarboxylase of maize. Nature 288: 556–560 (1980).
27. Meyer-Gauen G, Cerff R, Schnarrenberger C, Martin W: Molecular characterization of a novel, nuclear-encoded, NAD^+-dependent glyceraldehyde-3-phosphate dehydrogenase in plastids of the gymnosperm *Pinus sylvestris* L. Plant Mol Biol 26: 1155–1166 (1994).
28. Milanez S, Mural RJ: Cloning and sequencing of cDNA encoding the mature form of phosphoribulokinase from spinach. Gene 66: 55–63 (1988).
29. Murphy DJ, Walker DA: The properties of transketolase from photosynthetic tissue. Planta 155: 316–320 (1982).
30. Nikkola M, Lindqvist Y, Schneider G: Refined structure of transketolase from *Saccharomyces cerevisiae* at 2.0 Å resolution. J Mol Biol 238: 387–404 (1994).
31. Nowitzki U, Wyrich R, Westhoff P, Henze K, Schnarrenberger C, Martin W: Cloning of the amphibolic Calvin cycle/OPPP enzyme D-ribulose-5-phosphate 3-epimerase (E.C. 5.1.3.1) from spinach chloroplasts: functional and evolutionary aspects. Plant Mol Biol 29: 1279–1291 (1995).
32. Pelzer-Reith B, Penger A, Schnarrenberger C: Plant aldolase: cDNA and deduced amino-acid sequence of the chloroplast and cytosol enzyme from spinach. Plant Mol Biol 21: 331–340 (1993).
33. Raines CA, Lloyd JC, Longstaff M, Bradley D, Dyer T: Chloroplast fructose-1,6-bisphosphatase: the product of a mosaic gene. Nucl Acids Res 16: 7931–7942 (1988).
34. Raines CA, Lloyd JC, Willingham NM, Potts S, Dyer TA: cDNA and gene sequences of wheat chloroplast sedoheptulose-1,7-bisphosphatase reveal homology with fructose-1,6-bisphosphatases. Eur J Biochem 205: 1053–1059 (1992).
35. Saitou N, Nei M: The neighbor-joining method: A new method for the reconstruction of phylogenetic trees. Mol Biol Evol 4: 406–425 (1987).
36. Sambrook J, Fritsch E, Maniatis T: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, Laboratory, Cold Spring Harbor, NY (1989).
37. Schaaff-Gerstenschläger I, Mannhaupt G, Vetter I, Zimmermann FK, Feldmann H: *TKL2*, a second transketolase gene

of *Saccharomyces cerevisiae*: cloning sequence and deletion analysis of the gene. Eur J Biochem 217: 487–492 (1993).
38. Schnarrenberger C, Flechner A, W Martin: Enzymatic evidence indicating a complete oxidative pentose phosphate pathway in the chloroplasts and an incomplete pathway in the cytosol of spinach leaves. Plant Physiol 108: 609–614 (1995).
39. Schnarrenberger C, Pelzer-Reith B, Yatsuki H, Freund S, Jacobshagen S, Hori K: Expression and sequence of the only detectable aldolase in *Chlamydomonas reinhardtii*. Arch Biochem Biophys 313: 173–178 (1994).
40. Shih M-C, Lazar G, Goodman HM: Evidence in favor of the symbiotic origin of chloroplasts: primary structure and evolution of tobacco glyceraldehyde-3-phosphate dehydrogenases. Cell 47: 73–80 (1986).
41. Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW: Use of T_7 RNA polymerase to direct expression of cloned genes. Meth Enzymol 185: 60–89 (1990).
42. Teige M, Kopriva S, Bauwe H, Süß K-H: Chloroplast ribulose-5-phosphate 3-epimerase from potato: cloning, cDNA sequence, and tissue-specific enzyme accumulation. FEBS Lett 377: 349–352 (1995).
43. Teige M: Ph.D. Thesis. Universität Hannover (1994).
44. Villafranca JJ, Axelrod B: Heptulose synthesis from nonphosphorylated aldolases and ketoses by spinach transketolase. J Biol Chem 246: 3126–3131 (1971).
45. von Heijne G, Stepphuhn J, Herrmann RG: Domain structure of mitochondrial and chloroplast targeting peptides. Eur J Biochem 180: 535–545 (1989).
46. Woese CR, Kandler O, Whellis ML: Towards a natural system of organisms: Proposal for the domains Archae, Bacteria and Eucarya. Proc Natl Acad Sci USA 87: 4576–4579 (1990).
47. Zillig W: Comparative biochemistry of Archaea and Bacteria. Curr Opin Genet Devel 1: 544–551 (1991).