Higher-plant chloroplast and cytosolic fructose-1,6-bisphophosphatase isoenzymes: origins via duplication rather than prokaryote-eukaryote divergence

William Martin^{1,*}, Abdel-Zaher Mustafa¹, Katrin Henze¹ and Claus Schnarrenberger² ¹ Institut für Genetik, Technische Universität Braunschweig, Spielmannstrasse 7, D-38106 Braunschweig, Germany (*author for correspondence); ² Institut für Pflanzenphysiologie und Mikrobiologie, Freie Universität Berlin, Königin-Luise-Strasse 12–16a, D-14195 Berlin, Germany

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

Key words: Calvin cycle, sedoheptulose-1,7-bisphosphatase, isoenzymes, endosymbiosis, evolution

Abstract

Full-size cDNAs encoding the precursors of chloroplast fructose-1,6-bisphosphatase (FBP), sedoheptulose-1,7bisphosphatase (SBP), and the small subunit of Rubisco (RbcS) from spinach were cloned. These cDNAs complete the set of homologous probes for all nuclear-encoded enzymes of the Calvin cycle from spinach (*Spinacia oleracea* L.). FBP enzymes not only of higher plants but also of non-photosynthetic eukaryotes are found to be unexpectedly similar to eubacterial homologues, suggesting a eubacterial origin of these eukaryotic nuclear genes. Chloroplast and cytosolic FBP isoenzymes of higher plants arose through a gene duplication event which occurred early in eukaryotic evolution. Both FBP and SBP of higher plant chloroplasts have acquired substrate specificity, i.e. have undergone functional specialization since their divergence from bifunctional FBP/SBP enzymes of free-living eubacteria.

Abbreviations: FBP, fructose-1,6-bisphosphatase; SBP, sedoheptulose-1,7-bisphosphatase; FBA, fructose-1,6-bisphosphate aldolase

Introduction

Carbon enters higher-plant metabolism through the Calvin cycle in chloroplasts. Several enzymes of the pathway also occur as cytosolic isoenzymes of the glycolytic and gluconeogenetic pathways in a number of higher plants: phosphoglycerate kinase (PGK), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), triosephosphate isomerase (TPI), fructose-1,6-bisphosphate aldolase (FBA) and fructose-1,6bisphosphatase (FBP). The importance of these chloroplast-cytosol isoenzmyes for plant metabolism is obvious, since they are directly involved in the fixation and distribution of photosynthate. But they are also interesting from an evolutionary standpoint because they directly test a prediction of endosymbiotic theory known as the gene transfer corollary. Namely, one would expect the cytosolic isoenzymes to reflect the evolutionary history of the host whereas the chloroplast isoenzymes should reflect the evolutionary history of chloroplasts, having been transferred to the nucleus during the course of chloroplast genome reduction.

Yet as more sequences become known for chloroplast-cytosol isoenzymes of the Calvin cycle and glycolysis/gluconeogenesis, it is becoming apparent, surprisingly, that they do not bear out this prediction at all. The genes for choroplast and cytosolic TPI, for example, arose through gene duplication early in plant evolution, whereby the chloroplast's copy was functionally replaced [22, 41]. The genes for choroplast and cytosolic PGK also arose through gene duplication during plant evolution, whereby chloroplast's copy

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers L76557 (RbcS), L76555 (FBP) and L76556 (SBP).

was transferred to the nucleus and the host's copy was functionally replaced [6]. The gene duplication which gave rise to choroplast and cytosolic FBA occurred earlier, probably during the initial phases of eukaryotic evolution, whereby it appears that the chloroplasts's copy was functionally replaced [42]. The genes for choroplast and cytosolic GAPDH arose through a very ancient eubacterial gene duplication, both genes were transferred to the nucleus, the chloroplast enzyme apparently from the antecedents of cyanobacteria and the cytosolic enzyme apparently from the entecedents of mitochondria [21, 27]. These findings indicate that the evolutionary process which shaped the distribution of chloroplast-cytosol isoenzymes in contemporary plants was much more complex than the commonly assumed organelle-to-nucleus gene transfer product reimport scenario.

Chloroplast and cytosolic fructose-1,6-bisphosphatases (FBP) are the only remaining pair of chloroplastcytosol isoenzymes of the Calvin cycle and glycolysis/gluconeogenesis for which the evolutionary origin of the genes for the respective isoenzymes in higher plants has not been investigated. We have isolated and characterized cDNA clones for fructose-1,6-bisphosphatase (FBP), sedoheptulose-1,7-bisphosphatase (SBP), and the small subunit of Rubisco (RbcS) from spinach chloroplasts. These clones provide a complete set of homologous full-size cDNAs for all nuclear-encoded enzymes of the Calvin cycle from a single species. Here we examine the origins of higher-plant chloroplast and cytosolic FBP isoenzymes and the evolutionary relationship of chloroplast SBP to FBP in order to better understand the evolutionary history of Calvin cycle enzymes in the context of their glycolytic/gluconeogenetic homologues.

The sequence of the Calvin cycle from spinach

To isolate clones for FBP, SBP and RbcS, 40000 recombinants from the unamplified spinach cDNA library previously described [22] were screened. Phages were plated on *Escherichia coli* POP13 and screened by plaque hybridization on 82 mm nitrocellulose filters (Gelman Sciences) at 60 °C in $3 \times$ SSPE, 0.1% (w/v) SDS, 0.02% (w/v) PVP, 0.02% (w/v) Ficoll-400 and 50 µg/ml calf thymus DNA containing 20 ng/ml of hybridization probe. As hybridization probes, we used the inserts of cDNA clones for the chloroplast FBP and SBP from wheat [38, 39] (kindly provided by C. Raines) and for RbcS from potato [45]. Each

hybridization was carried out separately. Probes were radioactively labelled by random priming as described [22] to a specific activity of 5×10^7 cpm/µg. Filters were washed for 60 min at 60 °C in $2 \times$ SSPE, 0.1% (w/v) SDS and autoradiographed over night on XAR films. *Not*I inserts of positively hybridizing cDNAs were subcloned into pBluescriptSK plasmids (Stratagene) for further analysis. Other molecular methods were performed as described by Nowitzki *et al.* [34].

The spinach chloroplast FBP cDNA clone (L76555) encodes a 57 amino acid transit peptide followed by a 358 amino acid mature subunit which differs from the published protein sequence of purified FBPase from spinach chloroplasts [24] in only three positions: E9 instead of Q9, E31 instead of P31, and D246 instead of P₂₄₆. These minor differences may be lineor allele-specific. The cDNA clone for SBP (L76556) encodes a 387 amino acid open reading frame, the Cterminal 330 amino acids of which share roughly 90% amino acid identity to the previously published SBPase sequence of wheat [38]. The mature subunit of the RbcS cDNA clone (L76557) encodes a mature subunit which possesses an average of roughly 74% identity to numerous other higher-plant RbcS proteins, but, it is also only 72% identical to the published sequence of the purified spinach RbcS protein [24]. This indicates that in spinach, as in other higher plants [13, 45], RbcS is encoded by a multigene family. Southern blots with the spinach RbcS cDNA confirmed this, revealing 5-6 major hybridizing bands (data not shown) using the same low-stringency wash conditions described previously [16]. Southern blots of spinach DNA digested with enzymes which do not cut in the respective inserts revealed only two hybridizing bands each for chloroplast FBPase and SBPase probes under the same conditions (data not shown), indicating that these enzymes are encoded as single genes or very small gene families. All three clones encode N-terminal regions with typical properties of transit peptides [44].

With the characterization of these three clones, the complete sequences of full-size cDNAs for all enzymes active in the Calvin cycle of chloroplasts in spinach leaves are now known. The sources of the sequences and accession numbers are: ribulose-1,5-bisphosphate carboxylase/oxygenase, LSU (RbcL, V00168 [48]); phosphoribulokinase (PRK, X07654 [30]); glyceraldehyde-3-phosphate dehydrogenase, subunit B (GapB, X15189 [5]); 3-phosphoglycerate kinase (PGK, X68430 [4]); fructose-1,6-bisphosphate aldolase (FBA, X66814 [36]); triosephosphate isomerase (TPI, L36387 [22]); ribulose-5-phosphate 3-epimerase (RPE, L42328 [34]), ribose-5-phosphate isomerase (RPI, L43068 [28]); transketolase (TKL, L76554 [16]); glyceraldehyde-3-phosphate dehydrogenase, subunit A (GapA, L76552 [3]); sedoheptulose-1,7bisphosphatase (SBP, L76556, this paper); fructose-1,6-bisphosphatase (FBP, L76555, this paper); ribulose-1,5-bisphosphate carboxylase/oxygenase, small subunit (RbcS, L76557, this paper). We are currently investigating the coordinated expression and regulation of these genes.

Higher-plant chloroplast FBP and SBP: functional specialization

The fructose-1,6-bisphosphate aldolase (FBA) and sedoheptulose-1,7-bisphosphate aldolase reactions of the Calvin cycle are catalyzed by a single chloroplast FBA enzyme in spinach which has dual substrate specificity [7, Flechner and Schnarrenberger, unpublished]. However, the bisphosphatase reactions subsequent to these two aldolase reactions in higherplant chloroplasts are catalysed by distinct, separate and specific FBP [47] and SBP enzymes [10]. In eubacteria by contrast, single 'FBP' enzymes are known which possess dual substrate specificity for both sedoheptulose-1,7-bisphosphate and fructose-1,6-bisphosphate [1, 2, 18, 43]. This dual specificity has also been demonstrated for the cloned enzymes from Rhodobacter [19] and Alcaligenes [46], which possess sufficient activity with either substrate to catalyze both reactions in the Calvin cycle. Accordingly, these are designated as fructose-1,6/sedoheptulose-1,7-bisphosphatases (FBP/SBP). Some proteo- and cyanobacteria have been suggested to possess two [18] or even three [46] enzymes with FBP activity. Some non-photosynthetic eubacteria, such as E. coli, possess FBP which lacks SBP activity [17].

In order to examine in more detail the origin of higher plant chloroplast and cytosolic FBP isoenzymes and the nature of their relationship to chloroplast SBP and the bifunctional eubacterial enzymes, we extracted all available eubacterial and several eukaryotic homologues from the data base, produced an alignment and subjected the sequences to phylogenetic analysis. The complete alignment (available upon request) contains 411 positions for comparison. From the partial alignment shown in Fig. 1 it is evident that several of the most highly conserved regions of FBP sequences are also conserved in SBP, confirming that SBP and FBP enzymes share a common ancestor, chloroplast SBP shares about 30% identity with FBP sequences from various sources. Higher-plant chloroplast and cytosolic FBP share an average of 50% amino acid identity whereas eukaryotic FBPs share about 40% identical residues with their eubacterial counterparts. A phylogenetic tree constructed for higher plant FBP and SBP sequences in the context of eubacterial and eukaryotic homologues is shown in Fig. 2.

At face value Fig. 2 suggests a eubacterial origin of higher-plant nuclear genes for chloroplast SBP, since they branch within the eubacterial segment of the tree, although the SBP branch reveals no specific affinity with any particular group of eubacterial sequences and the degree of divergence between SBP and FBP enzymes makes placement of the SBP root difficult. The long branch separating chloroplast SBP enzymes from other bisphosphatases (Fig. 2) suggests elevated substitution rate during the phase of SBP evolution from an FBP or FBP/SBP ancestor towards sedoheptulose-1,7-bisphosphate substrate specificity, which could reflect either a relaxation of functional constraint, positive selection or both. The divergence between spinach and wheat chloroplast FBP (81% amino acid identity over 359 positions) is the same as that between chloroplast SBP for the same species (81% over 338 positions), indicating that subsequent to acquiring substrate specificity, both enzymes evolved at the same conservative pace, probably due to similarly strong functional constraints. Spinach chloroplast SBP shows considerable substrate specificity for sedoheptulose-1,7-bisphosphate [10], indicating that functional specialization of SBP has occurred in the evolution of this enzyme, although it should be noted that the substrate specificity of other plant SBPs has not been extensively characterized.

The phylogeny of proteobacterial and cyanobacterial FBP and FBP/SBP sequences in Fig. 2 differs, albeit with low bootstrap support, from that observed for rRNA genes [12]. This may be due to paralogy ensuing from differential loss of members of a eubacterial gene family, as for *rbcL* [29] and GAPDH [21] or may simply relate to the sequence characteristics and small bacterial sample available for analysis. Interspecies transfer of FBP/SBP genes may also be commonplace, as has been suggested for *rbcL* [35], and in this context it is noteworthy that two of the proteobacterial FBP/SBP sequences are plasmid-borne (designated with asterisks in Fig. 2), so that lateral transfer of genes surveyed between eubacteria cannot be excluded.

	+++++++++++ 100
Triticum SBP Spinacia SBP Chlamydomonas SBP Haemophilus Anabaena Synechocystis Xanthobacter Rhodobacter F/SBP Nitrobacter Spinacia cp FBP Triticum cp FBP Spinacia cyt FBP Solanum cyt FBP Kluveromyces Saccharomyces	PAGRQSKAASRAALTTRCA IGDSLEEFLTKAT PDKNL IRLLICMGEAMRT IAFKVRTASCGGTACVNSFGDEQLAVDMLADKLLF TTKSSVRVNKAKNSSLUTKCELGDSLEEFLAKATTDKGL IRLMYCMGEALRTIGFKVRTASCGGTACVNSFGDEQLAVDMLADKLLF RARVQSRRTAVLTQAKIGDSLAEFLVEAT PDFKLRVVMMSMEATRTIAHKVRTASCGGTACVNSFGDEQLAVDMVADKLLF KTLSEF IVERQAEYP.NAKGELGILSSIRLLAKI IRRDINKAGLTN.ILGQSGIENVQGESQMKLDLFAHNTMK AKASESLDLSVNESTDKALDRDCTTLSRHVLQQLQSFS.ADAQDLSALMNRIALAGKLVARRLSRAGLMEGVLGFTGENVQGESQMKLDLFAHNTMK AKASESLDLSVNESTDKALDRDCTTLSRHVLQQLQSFS.PEAQDLAALMQRIGLAAKLIARRLSRAGLMEGVLGFTGENVQGESQMKLDLFAHNTMK AKASESLDLSVNESTDKALDRDCTTLSRHVLQQLQSFS.PEAQDLAALMQRIGLAAKLIARRLSRAGLMEGVLGFTGENVQGESQMKLDUFAHNTMK AKASESLDLSVNESTDKALDRDCTTLSRHVLDQLQSFS.PEAQDLAALMQRIGLAAKLIARRLSRAGLMEGVLGFTGENVQGESQMKLDVANDVFI LEPNADHRAAVAQAAGVAASRITLTVMLDEWAGADAR.RSDVAALMARAVADAGAGAARISISRNGEBEDL.ARTLSSGEAGEGQKALDVIADDAFR AIELEDLGL.SPDVADVMQRLARVGAGIARISSNGLERDL.GQAQGTNAGGDQKDLDVRADQIIR SQELTLQQRMDGEAGSDPL.RQAVKBAVAALARAAVDISDLTCRGALAGISN.LTGIQQAVNIQGEDQKKLDVXSNEVFS AVVDTASAPAPAAARKRSSYDMITLITWLL.KQEQEG.VIDAELTIVLSSISLACKQIASLVQRAGISN.LTGIQQANNIQGEDQKKLDVLSNEVFS DHAADRHRTDLMTITRFVLNEQSKRP.ESRGDFSILLSHIVLGCKFVCTAVNKAGLAK.LIGLAGETNGGEDQKKLDVLSNEVFY DHAADRHRTDLMTITRFVLNEQSKRP.ESRGDFSILLSHIVLGCKFVCTAVNKAGLAK.LIGLAGENTGGEQQKKLDVLSNEVFY DHAADRHRTDLMTITRFVLNEQKKPP.ESRGDFSILLSHIVLGCKFVCTAVNKAGLAK.LIGLAGSNSTGGDQKKLDVLSNEVFY DHAADRHRTDLMTITSFILDQQHLSA.KNATGEFSMLLNSLQFAFKFISQTIRRAELVN.LIGLAGASNSTGGDQKKLDVLSNEVFY AGIKHRRDSAESINTDIIILSRFILDQQHLSA.KNATGEFSMLLNSLQFAFKFISQTIRRAELVN.LUGLAGASNFTGDQQKKLDVLSNEVFY
Triticum SBP Spinacia SBP Chlamydomonas SBP Haemophilus Anabaena Synechocystis Xanthobacter Rhodobacter F/SBP Nitrobacter Spinacia cy FBP Triticum cp FBP Triticum cy FBP Solanum cyt FBP Kluveromyces Saccharomyces	***** ***** ************************************
Triticum SBP Spinacia SBP Chlamydomonas SBP Haemophilus Anabaena Synechocystis Xanthobacter Rhodobacter F/SBP Nitrobacter Spinacia cp FBP Triticum cp FBP Spinacia cyt FBP Solanum cyt FBP Kluveromyces Saccharomyces	300 MGIYGPRTTFVVALKDC PGTHEFLL · LDEGKWQHVKDTT·····SIGEGKMF SPGNLRATFDNPDYDKLVNYVKEKYTLRYTGGMVPDVNQIIVK MGIYGPRTTVCIALKDYPGTHEFLL · LDEGKWQHVKDTT····· SIGEGKMF SPGNLRATFDNPDYBLVRUIQYYIKEKYTLRYTGGMVPDVNQIIVK MGIYGPRTVCIALKDAPGCHEFLL · MDGKWMHVKETT······ HIGEGKMFAPGNLRATFDNPAYERLINFYLGEKYTLRYTGGNVPDVNQIIVK YIVYGSSTMLVYTT·· GNCVNGFTVDPSIGTFCLSHEMMQMPKEGKIYSINEGNFLPQVKKYI · KYCQEEDK · ATHRPYVSRYIGSLVADFHRNLLK YILYGPSTMLVYTM·· GTGVHSFTLDPSIGTFCLSHEMMQMPKEGKIYSINEGNFWQWEESMREYI · KYCQEEDK · ATHRPYVSRYIGSLVADFHRNLLK YILYGPSTMLVYTM·· GGCVHVFVLDPSIGEFILSEENIRIPDHGAVYSVNEGNFWQWEESMREYI · KYCHREF····· AYSGRYSGALVADFHRLLQ YVLYGPTLVYTM·· GQCVHVFVLDPSIGEFILSEENIRIPDHGAVYSVNEGNFWQWEESMREYI · KYCHREF····· AYSGRYSGALVADFHRLLQ YVLYGPTLVYTM·· GQCVHVFVLDPSIGFFILSEENIRIPDHGAVYSVNEGNFWQWEESMREYI · KYCHREF····· AYSGRYSGALVADFHRLLQ YVLYGPTLVTI··· GACVHFVLDPSIGFVLAGSIJQIPNGQIYSVNEGNFWQWEPGYRQYI · REMHRRE····· AYSGRYSGALVADFHRLLQ YVLYGPTSLVLTI··· GACVHFVLDPSIGFVLASSIQIPTDANEFAINASNRHWDVPVRDYVDCLAGKKG · PRENDFNMRWASMVADAHRIFQR YIIYGPQTSLVLTI··· GRCVDIFTLDRADFVFKLTGSSVQIPTDANEFAINASNRHWDDPVRDYVDCLAGKG · PRENDFNMRWIASLVAETHRILMR YCMYSSSVIFVLTI··· GRCVDIFTLDRADFVFKLTGSSVQIPTDANEFAINASNRHWDDPVRAYIDECLAGADG · PCGKNFNMRWIGSLVAEAFRILIR YCMYSSSCILVLST··· GSGVNGFTLDPMGEFVLTQEKVQIPKSGKIYSFNEGNYALWDDKLKKYMDDLKEPGES ··· GKPYSRYIGSLVGDFHRTLLY YCMYGSSCTLVLST··· GSGVNGFTLDPSLGEYILTHPDIKIPNKKIYSNEGNAKNWDDFTKIVVEKCKFPTDG ··· SSPKSLRYIGSMVADVHRTLLY YCMYGSSTHLVLTI··· GAGVNGFTLDPNLGEFILTPPIKIPNKKIYSNEGNAKNWDSFTKIVVQSCKYPADG···· SSPKSLRYIGSMVADVHRTLLY YMYGSSTHLVLTI··· GGCVNGFTLDNLGEFILTPPIKIPLKIPKKKIYSNEGNAKNWDSFTKIVQSCKYPADG···· SSPKSLRYIGSMVADVHRTLLY YMYGSSTHLVLTI··· GGCVNGFTLDNNLGEFILTPPEKLPPEKSIYSINEGNTLYWETIRTFIEKVKQQAD·NNNKFYSARYGSMVADVHRTLLY YMYGSSTHLVLTI··· GGCVNGFTLDNNLGEFILTPPEKLPPEKSIYSINEGNTLYWETIRTFIEKVKQQAD·NNNKFYSARYGSMVADVHRTLLY YMYGSSTHLVLTI··· GDCVNGFTLDNNLGEFILTHPNLRIPPEKABPKAKIYSNEGNALWMETIRTFIEKVKQAAD·NNKFFSARYGSMVADVHRTLLY
Triticum SBP Spinacia SBP Chlamydomonas SBP Haemophilus Anabaena Synechocystis Xanthobacter Rhodobacter F/SBP Nitrobacter Spinacia cp FBP Triticum cp FBP Triticum cp FBP Spinacia cyt FBP Solanum cyt FBP Kluveromyces	400 EKGIF ··· TRVTSPTAKAKLRLLFEVAPLGFLIEKAGGHSSDGKQSV ·· LDKVISVLDERTQVAYGSKNEIIRFEETLYG · SSRLAASATVGATA EKGIF ··· TRVTSPTAKAKLRLLFEVAPLGFLIEKAGGHSSEGTKSV ·· LDIEVKNLDDRTQVAYGSLNEIIRFEETLYG · SSRLAEASATVGATA EKGVF ··· TRVTSPTAKAKLRLLFEVAPLGFLIEKAGGHSSEGTKSV ·· LDIEVKNLDDRTQVAYGSLNEIIRFEETLYG · SSRLAEASATVGATA EKGVF ··· TRVTSPTAKAKLRLLFEVAPLGFLIEKAGGASSCDGKAVSALDIPILVCOQRTQCTQGSLOEVRFEETMYGTSPFSEKVVA G·GIYIYPSATNYP ·· NGKLRLLYEGNPIAFLAEQAGGVATGC ·· RRILDIEPTALHERVPLFVGSEDMVKKAQEMMEEFKE G·GVFLYPGTIQNP ·· EGKLRLLYESAPLAFLIQQAGGRATTG ·· VRILDIEPTALHERVPLFVGSEDMVKKAQEMMEEFKE G·GVFLYPETTVKNP ·· TGKLRLLYEAPPAFLMEQAGGGATGG ·· VRILDIEPTALHERVPLIGSKEDVAKVESFIQNGH G·GIYLVPEOGRKGGYTHGRLRLLYEAPPAFLMEQASGSATDG ·R · GAILDLSATGLHQRVPFIIGSSDEVARVSSRYHLEPNGHGERSPLFARRGLFI G·GVFLYPGDERKGYTEGRLRHVYECAPIAFLINVGGGATDG ·C · ADILTALDRLHARTPFVFCGASKVARVAXYH ·· GLACEETSALFGSRGLFRS G·GIFLYPGDARKGYEGGRL· VVYEAHPMAFIVEQAGGGASTG ·R · KRVLDIVPDSLHQRVPLIMGSIKNVQRLEHMHTVPDVALEANAPLFGNRGLFRV G·GIYGYPRDAKSK ·· NGKLRLLYECAPMSFINEQAGGGASGG ·H · QRILDIQPTEHRQVPLYGSVEEVEKVEKFLSSE G·GIFLYPGDKKSP ·· NGKLRULYEVPPMSFINEQAGGGAFTG ·K · QRALDLIPTKIHERSPVFLGSVDDVEEIKALYAAQEKTA G·GIFLYPGDKKSP ·· NGKLRULYEVPPMSFIMEQAGGGAFTG ·K · QRALDLIPTKIHERSPVFLGSVDDVEEIKALYAAQEKTA G·GIFLYPGDKKSP ·· NGKLRULYEVPPMSFLMEQAGGGAAVND · RGERILDLVPQHIHDKSSIWLGSSGDVDKYLKHIGKL G·GLFSYPGDKKNP ·· NGKLRLLYEAPPMAFLWEQAGGKAVND · RGERILDLVPSHIHDKSSIWLGSSGDVDKYLKHIGKL G·GLFAYPCDKKSP ·· NGKLRLLYEAPPMAFLMEQAGGKAVND · RGERILDLVPSHIHDKSSIWLGSSGDVDKYLKHIGKL

Figure 1. Alignment of several sedoheptulose-1,7-bisphosphatase (SBP), fructose-1,6-bisphosphatase (FBP) and bifunctional S/FBP amino acid sequences. Regions of strongest similarity between SBP and FBP enzymes are indicated with '+++'. Asterisks indicate strictly conserved residues in the alignment, gaps are indicated by dots. cp, chloroplast; cyt, cytosolic. The activity of FBP from chloroplasts and from several eubacterial sources is known to be regulated by light *via* the thioredoxin system [9].

Origin of higher-plant chloroplast and cytosolic FBP isoenzymes

The topology in Fig. 2 clearly suggests that plant chloroplast and cytosolic FBP are related through gene duplication, rather than through prokaryote-eukaryote

divergence. The duplication which gave rise to chloroplast/cytosol FBP isoenzymes apparently took place in eukaryotic genomes prior to the divergence of plants, animals and fungi and thus strongly parallels that for the isoenzymes of fructose-1,6-bisphosphate aldolase [42]. The lack of FBP sequences from archaebacteria

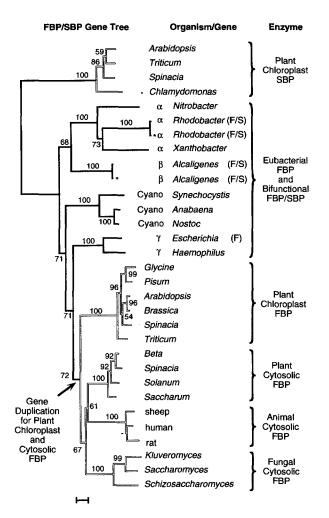


Figure 2. Gene phylogeny for fructose-1,6-bisphosphatase and sedoheptulose-1,7-bisphosphatase proteins. Sequences were extracted from GenBank. The tree was constructed by the neighbor-joining method [40] from a matrix of estimated numbers of amino acid substitutions per site calculated with the Dayhoff option of phylip [15]. 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; Cyano, cyanobacteria. Enzymes for which dual substrate specificity for fructose-1,6-bisphosphate and sedoheptulose-1,7-bisphosphate has been directly demonstrated are indicated by (F/S). (F) indicates fructose-1,6-bisphosphate specificity. Asterisks denote sequences which are encoded on plasmids.

for comparison makes it difficult to tell whether the ancestral gene that duplicated to give rise to the plant FBP isoenzymes was ancestral to the eukaryotic lineage or was acquired from eubacteria (for a discussion of this problem, see [26]). But, as in the case of GAPDH [21], PGK [6] and TKL [16], the FBP gene

tree is characterized by (1) greater diversity within eubacterial genes than across the eubacterial-eukaryote boundary and (2) a lack of distinct separation between eubacterial and eukaryotic genes. The most straightforward interpretation for this finding is that all FBP genes of eukaryotes surveyed here are of eubacterial origin. It is noteworthy that a surprisingly large number of eubacterial genes are currently being discovered in eukaryotic genomes [26], even in the genomes of amitochondriate protists [8, 11, 21].

Whereas gene transfers between eubacteria could account for the incongruency of the FBP gene tree with rRNA-based bacterial phylogeny [12], they cannot account for the unexpected similarity between eubacterial and eukaryotic FBP/SBP genes studied here. By contrast, a eubacterial origin of the eukaryotic nuclear genes can [6, 21, 27]. Some nuclear-encoded genes for Calvin cycle enzymes in higher-plant chloroplasts are certainly of eubacterial origin (GAPDH [21], PGK [6], rbcS [31]). Others clearly appear to be of eubacterial origin on the basis of available data, but a large eubacterial sample and archaebacterial homologues necessary to make a watertight case for this are still missing (RPE [34]; TKL [16]; FBP and SBP, this paper). Still others descend from duplication of 'pre-existing' genes of the eukaryotic nucleus (TPI [22, 41], FBA [42]), whereby a eubacterial origin of such 'pre-existing' nuclear homologues early in eukaryotic evolution seems increasingly likely [26]. Strong evidence is now accumulating to suggest that no eukaryotes are primitively amitochondriate [14, 33] and that endosymbiotic gene transfer from eubacteria has contributed very extensively eukaryotic nuclei [26]. Thus, for those enzymes of the Calvin cycle such as chloroplast FBP that (1) ultimately appear to be of eubacterial origin, but (2) lack clear cyanobacterial affinity and (3) share a recent common ancestor with homologues from the cytosol of non-photosynthetic eukaryotes, a new and explicit working hypothesis for the origin of their corresponding genes is necessary. For such enzymes, we suggest that the ancestors of mitochondria relinquished the gene to the nucleus, giving rise to the cytosolic enzyme. During the subsequent course of plastid endosymbiosis, gene(s) for the preexisting cytosolic enzyme acquired a transit peptide via duplication events, allowing them to be rerouted into the plastid, permitting loss of the chloroplast copy. This scenario, outlined in Fig. 3, could easily account both for the similarity between eukaryotic and eubacterial FBP enzymes, and for the similarity between

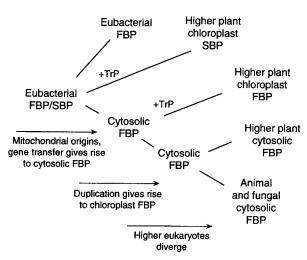


Figure 3. Schematic diagram of the evolution of genes chloroplast and cytosolic FBP which can account for available data. +TrP indicates the gain of a transit peptide for chloroplast import. Events indicated with arrows refer explicitly to processes of FBP (not SBP) evolution. See text for details.

choloroplast FBP of plants and cytosolic FBP of nonphotosynthetic eukaryotes.

In summary, we now have a relatively complete picture of the origin of chloroplast-cytosol isoenzymes involved in the Calvin cycle and glycolysis in higher plants. The cloning and evolutionary analysis of all of these enzymes now allows us to make the simple statement that none of the isoenzyme pairs known to exist for these pathways (GAPDH, FBA, TPI, PGK, FBP) in higher plants reflect prokaryote-eukaryote divergence. Rather, they are related by a complex and continuous series of gene duplications that accompanied the endosymbiotic origins of organelles. Since many of these duplications occurred early in eukaryotic evolution, it is not surprising that in early branching photosynthetic protists both Calvin cycle enzymes [20, 32] and their cytosolic homologues [37] in some cases have an evolutionary origin which is distinct from that found in higher plants.

Acknowledgements

W.M. and C.S. gratefully acknowledge financial support from the Deutsche Forschungsgemeinschaft (Ma 1426/3-2, Schn 28/21-2). We thank Ch. Köhler for excellent technical assistance and A. Flechner for critical reading of the manuscript.

References

- Abdelal AT, Schlegel HG: Purification and regulatory properties of fructose-1,6-bisphosphatase from *Hydrogenomonas eutropha*. J Bact 120: 304–310 (1974).
- Amachi T, Bowien B: Characterization of two fructose bisphosphatase isoenzymes from the hydrogen bacterium *Nocardia* opaca 1b. J Gen Microbiol 113: 347–356 (1979).
- Baalmann E, Scheibe R, Cerff R, Martin W: Functional studies of chloroplast glyceraldehyde-3-phosphate dehydrogenase subunits A and B expressed in *Escherichia coli*: Formation of highly active A₄ and B₄ homotetramers and evidence that aggregation of the B₄ complex is mediated by the B subunit carboxyterminus. Plant Mol Biol 32: 505–513 (1996).
- Bertsch U, Schlicher TB, Schröder I, Soll J: Sequence of mature phosphoglycerate kinase from spinach chloroplasts. Plant Physiol 103: 1449–1450 (1993).
- Brinkmann H, Cerff R, Salomon M, Soll J: Cloning and sequence analysis of cDNAs encoding the cytosolic precursors of subunits GapA and GapB of chloroplast glyceraldehyde-3-phosphate dehydrogenase from pea and spinach. Plant Mol Biol 13: 81–94 (1989).
- 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).
- 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).
- Brown JR, Doolittle WF: Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. Proc Natl Acad Sci USA 92: 2441–2445 (1995).
- 9. Buchanan BB: Role of light in the regulation of chloroplast enzymes. *Annu Rev. Plant Physiol* 31: 341–374 (1980).
- Cadet F, Meunier J-C: pH and kinetic studies of chloroplast sedoheptulose-1,7-bisphosphatase from spinach (*Spinacia* oleracea). Biochem J 253: 249–254 (1988).
- Clark CG, Roger AJ: Direct evidence for secondary loss of mitochondria in *Entamoeba histolytica*. Proc Natl Acad Sci USA 92: 6518–6521 (1995).
- 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).
- DeRocher EJ, Quigley F, Mache R, Bohnert HJ: The six genes of the Rubisco small subunit multigene family from *Mesem*bryanthemum crystallinum, a facultative CAM plant. Mol Gen Genet 239: 450–462 (1993).
- Doolittle WF: Some aspects of the biology of cells and their possible evolutionary significance. Symp Soc Gen Microbiol, in press (1996).
- Felsenstein J: PHYLIP: Phylogeny Inference Package (version 3.2). Cladistics 5: 164–166 (1989).
- Flechner A, Dreßen U, Westhoff P, Henze K, Schnarrenberger C, Martin W: Molecular characterization of transketolase (EC 2.2.1.1) from spinach chloroplasts completes cloning of the Calvin cycle. Plant Mol Biol 132: 475–484 (1996).
- Fraenkel DG, Pontremoli S, Horecker BL: The specific fructose diphosphatase of *Escherichia coli*: properties and partial purification. Arch Biochem Biophys 114: 4–12 (1966).
- Gerbling KP, Steup M, Latzko E: Fructose 1,6-bisphosphatase form B from *Synechococcus leopoliensis* hydrolyzes both fructose and sedoheptulose bisphosphate. Plant Physiol 80: 716–620 (1986).

- Gibson JL, Tabita FR: Localization and mapping of CO₂ fixation genes within two gene clusters in *Rhodobacter* sphaeroides, J Bact 170: 2153–2158 (1988).
- Gross W, Bayer MG, Schnarrenberger C, Gebhart UB, Maier TL, Schenk HEA: Two distinct aldolases of the class II type in the cyanoplasts and in the cytosol of the alga *Cyanophora paradoxa*. Plant Physiol 105: 1393–1398 (1994).
- 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).
- 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).
- Ladror US, Latshaw SP, Marcus F: Spinach cytosolic fructose-1,6-bisphosphatase. Purification, enzyme properties and structural comparisons. Eur J Biochem 189: 89–94 (1990).
- Marcus F, Harrisch PB: Amino acid sequence of spinach chloroplast fructose-1,6-bisphosphatase. Arch Biochem Biophys 279: 151-7 (1990).
- Martin PG: Amino acid sequence of the small subunit of ribulose-1,5-bisphosphate carboxylase from spinach. Aust J Plant Physiol 6: 401–408 (1979).
- 26. Martin W: Is something wrong with the tree of life? BioEssays 18: 523–527 (1996).
- Martin W, Brinkmann H, Savonna C, Cerff R: Evidence for a chimeric nature of nuclear genomes: eubacterial origin of glyceraldehyde-3-phosphate dehydrogenase genes. Proc Natl Acad Sci USA 90: 8692–8696 (1993).
- 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).
- Martin W, Somerville CC, Loiseaux-de Goër S: Molecular phylogenies of plastid origins and algal evolution. J Mol Evol 35: 385–403 (1992).
- Milanez S, Mural RJ: Cloning and sequencing of cDNA encoding the mature form of phosphoribulokinase from spinach. Gene 66: 55–63 (1988).
- Morden CW, Delwiche CF, Kuhsel M, Palmer JD: Gene phylogenies and the endosymbiotic origin of plastids. BioSystems 28: 75–90 (1992).
- Morse D, Salois P, Markovic P, Hastings JW: A nuclearencoded form II RuBisCO in dinoflagellates. Science 268: 1622–1624 (1995).
- 33. Müller M: Energy metabolism of amitochondriate protists, an evolutionary puzzle. In: Schlegel M, Hausmann K (eds) Christian Gottfried Ehrenberg-Festschrift anlässlich der 14. Wissenschaftlichen Jahrestagung der Deutschen Gesellschaft fur Protozoologie, 9.–11. März 1995, Delitzsch (Sachsen), pp. 63– 76. Leipziger Universitätsverlag, Leipzig (1996).

- Nowitzki U, Westhoff P, Henze K, Schnarrenberger C, Martin W: Cloning of the amphibolic Calvin cycle/OPPP enzyme D-ribulose-5-phosphate 3-epimerase (EC 5.1.3.1) from spinach chloroplasts: functional and evolutionary aspects. Plant Mol Biol 29: 1279–1291 (1995).
- Palmer JD: Rubisco rules fall; gene transfer triumphs. BioEssays 17: 1005–1008 (1995).
- 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).
- Pelzer-Reith B, Wiegand S, Schnarrenberger C: Plastid class I and cytosol class II aldolase of *Euglena gracilis*. Plant Physiol 106: 1137–1144 (1994).
- 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).
- 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-1,6bisphosphatases. Eur J Biochem 205: 1053–1059 (1992).
- Saitou N, Nei M: The neighbor-joining method: a new method for the reconstruction of phylogenetic trees. Mol Biol Evol 4: 406–425 (1987).
- 41. Schmidt M, Svendsen I, Feierabend J: Analysis of the primary structure of the chloroplast isozyme of triosephosphate isomerase from rye leaves by protein and cDNA sequencing indicates a eukaryotic origin of its gene. Biochim Biophys Acta 1261: 257–264 (1995).
- 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).
- Springgate CF, Stachow CS: Fructose 1,6-diphosphatase from *Rhodopseudomonas palustris*. I. Purification and properties. Arch Biochem Biophys 152: 1–12 (1972).
- von Heijne G, Stepphuhn J, Herrmann RG: Domain structure of mitochondrial and chloroplast targeting peptides. Eur J Biochem 180: 535–545 (1989).
- Wolter F, Fritz C, Willmitzer L, Schell J, Schreier P: *rbcS* genes in *Solanum tuberosum*: conservation of transit peptide and exon shuffling during evolution. Proc Natl Acad Sci USA 85: 846–850 (1988).
- Yoo JG, Bowien B: Analysis of the *cbbF* genes from *Alcali-genes eutrophus* that encode fructose-1,6-/sedoheptulose-1,7-bisphosphatase. Curr Microbiol 31: 55–61 (1995).
- Zimmermann G, Kelley GJ, Latzko E: Efficient purification and molecular properties of spinach chloroplast fructose 1,6bisphosphatase. Eur J Biochem 70: 361–367 (1976).
- Zurawski G, Perrot B, Bottomley W, Whitfeld PR: The structure of the gene for the large subunit of ribulose 1,5bisphosphate carboxylase from spinach chloroplast DNA. Nucl Acids Res 9: 3251–3270 (1981).