J Mol Evol (1994) 38:610–618

Journal of Molecular Evolution

Evidence That Stilbene Synthases Have Developed from Chalcone Synthases Several Times in the Course of Evolution

S. Tropf, T. Lanz, S.A. Rensing, J. Schröder, G. Schröder

Institut für Biologie II, Universität Freiburg, Schänzlestr. 1, D-79104 Freiburg, Federal Republic of Germany

Received: 2 August 1993 / Accepted: 6 October 1993

Abstract. Chalcone (CHS) and stilbene (STS) synthases are related plant-specific polyketide synthases that are key enzymes in the biosynthesis of flavonoids and of stilbene phytoalexins, respectively. A phylogenetic tree constructed from 34 CHS and four STS sequences revealed that the STS formed no separate cluster but grouped with CHS from the same or related plants. This suggested that STS evolved from CHS several times independently. We attempted to simulate this by site-directed mutagenesis of an interfamily CHS/STS hybrid, which contained 107 amino acids of a CHS from *Sinapis alba* (N-terminal) and 287 amino acids of a STS from *Arachis hypogaea.* The hybrid had no enzyme activity. Three amino acid exchanges in the CHS part (Gln-100 to Glu, Val-103 to Met, Val-105 to Arg) were sufficient to obtain low STS activity, and one additional exchange (Gly-23 to Thr) resulted in 20–25% of the parent STS activity. A kinetic analysis indicated (1) that the hybrids had the same K_m for the substrate 4-coumaroyl-CoA but a lower V_{max} than the parent STS, and (2) that they had a different substrate preference than the parent STS and CHS. Most of the other mutations and their combinations led to enzymatically inactive protein aggregates, suggesting that the subunit folding and/or the dimerization was disturbed. We propose that STS evolved from CHS by a limited number of amino acid exchanges, and that the advantage gained by this enzyme function favored the selection of plants with improved STS activity.

Key words: Condensing enzyme — Phylogenetic tree — Polyketide synthase — Site-directed mutagenesis

Correspondence to: J. Schröder

Abbreviations: AA, amino acid; CHS, chalcone synthase; STS, stilbene synthase

The data are discussed on the level of the presently available CHS and STS sequences although many were published after beginning the experiments several years ago. The new information changed the CHS consensus in some details but otherwise confirmed the deductions on the potential significance of amino acid differences between CHS and STS

Introduction

Chalcone synthases (CHS) are plant-specific enzymes that synthesize naringenin chalcone, the precursor for a large number of flavonoids, including the anthocyanins responsible for red colors in many flowers. The reaction is complex. It involves the step-wise condensation of three two-carbon units to 4-coumaroyl-CoA and finally the formation of a new aromatic ring system from a postulated tetraketide intermediate (Fig. 1). Stilbene synthases (STS) occur in a limited number of widely unrelated higher plants. They synthesize the backbone of the stilbene phytoalexins which have antifungal properties and contribute to the plant defenses against pathogens. (See Schröder et al. 1993 for recent review.) The enzyme reaction appears to be identical with that of CHS up to the tetraketide intermediate, but the ring closure involves different atoms, and a different spatial arrangement of the intermediate is necessary to obtain a stilbene backbone as final product. The typical STS reaction also includes a decarboxylation which is not performed by CHS (Fig. 1).

The similarities of the reactions are reflected by similarities in structure and protein sequences. Both are homodimers of identical subunits with 388–400 amino

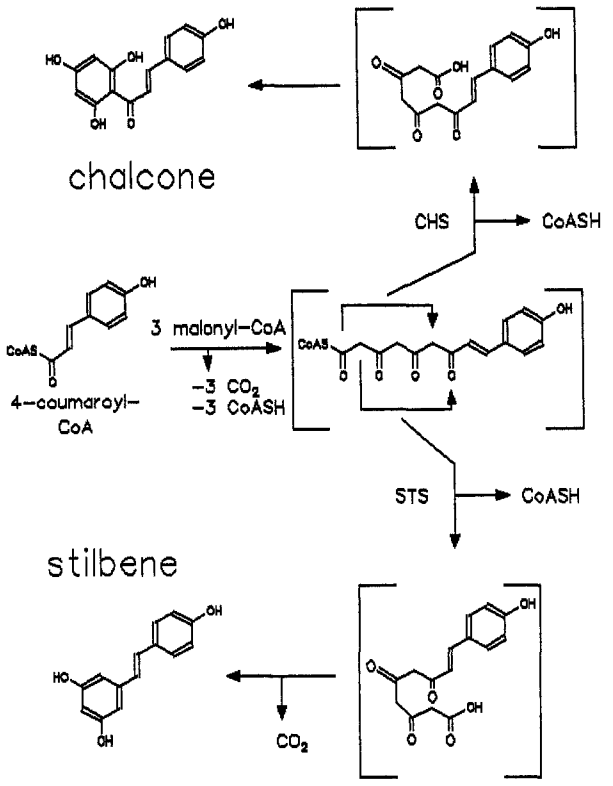

Fig. 1. Reactions of CHS- and STS-type enzymes. The postulated protein-bound intermediates are in brackets. The preferred starter substrate with the CHS from *S. alba* and the STS from *A. hypogaea* is 4-coumaroyl-CoA. Cinnamoyl-CoA (phenolic -OH replaced by -H) and dihydrocinnamoyl-CoA (aliphatic double bond in cinnamoyl-CoA reduced) are also accepted in vitro.

acids (AA), and both share in the same position a strictly conserved, essential cysteine, which represents the active site of the condensing reactions (Cys-169) (Lanz et al. 1991). The known CHS sequences share ≥60% identical positions, and the same results are obtained from comparisons of different STS to each other. The same values are also found for overall comparisons between CHS and STS, and only a detailed inspection reveals small differences. These are identified by deviations of STS from the CHS consensus or by consistent differences in charged/uncharged AA in the same position (Schröder and Schröder 1990). They are not confined to certain regions of the proteins, but are distributed throughout the sequence, and their significance with respect to the functional differences of CHS and STS is unknown. CHS and STS have no significant similarities to other eukaryotic condensing enzymes, and therefore represent an independent development.

The work described here analyzes the relationship between CHS and STS. A phylogenetic tree shows that the STSs do not form a separate cluster, but group with CHS from the same or related plants. This suggests that STS developed several times from CHS, and we describe experiments to test the hypothesis that a relatively small number of AA exchanges are sufficient to adapt CHS sequences to STS function.

Materials and Methods

Phylogenetic Tree. A total of 34 CHS and four STS AA sequences were analyzed. The alignment was performed with CLUSTAL V (Higgins et al. 1992) and was manually improved by visual inspection, leading to a total length of 436 positions. Phylogenetic analysis was done with the program package PHYLIP 3.5c (unpublished; see Felsenstein 1989 for version 3.2); 403 positions of the alignment were used for the calculation. The estimation was carried out by Felsenstein's protein parsimony method, which does not count synonymous changes. Bootstrap resampling (Felsenstein 1985; Wu 1986) of the original data was used as a pseudo-empirical test of the reliability of the tree topology. The consensus tree was constructed by use of the majority-rule and strict consensus algorithm implanted in PHYLIP.

Sources for CHS sequences: *Medicago sativa* (Junghans et al. 1993); *Pisum sativum* (Ichinose et al. 1992); *Pueraria lobata* (Nakajima et al. 1991); *Phaseolus vulgaris* (Ryder et al. 1987); *Glycine* max (Akada et al. 1990); *Pinus sylvestris* (Fliegmann et al. 1992); *Lycopersicon esculentum* (O'Neill et al. 1990); *Petunia hybrida* (Reif et al. 1985; Koes et al. 1989); *Antirrhinum majus* (Sommer and Saedler 1986); *Magnolia liliiflora, Ranunculus acer,* and *Zea mays* (Niesbach-Klösgen et al. 1987; Franken et al. 1991); *Hordeum vulgare* (Rohde et al. 1991); *Sinapis alba* (Ehmann and Schäfer 1988); *Arabidopsis thaliana* (Feinbaum and Ausubel 1988); *Matthiola incana* (Epping et al. 1990); *Petroselinum hortense* (Reimold et al. 1983). Sources for STS sequences: *Arachis hypogaea* (Schröder et al. 1988; Lanz et al. 1990); *Pinus sylvestris* (Fliegmann et al. 1992), and unpublished; *Vitis* var. optima. (Melchior and Kindl 1990). The outgroup was FABH from *E. coli* (Tsay et al. 1992).

CHS and STS Clones and Site-Directed Mutagenesis. The original constructs for expression of the CHS from *Sinapis alba* (CHS3) and the STS from *Arachis hypogaea* (RS3) in the vector pTZ19R have been described (Lanz et al. 1991). They express the proteins under control of a bacterial promoter.

The mutagenesis was performed with single-stranded DNA produced with helper phage M13K07 in *E. coli* strain RZ1032 (Kunkel 1985). Often several different oligonucleotides were used simultaneously to obtain several combinations of mutants from one round of mutagenesis. The mutations were verified by DNA sequence analysis with the dideoxynucleotide chain termination technique, and the expression of correct size subunits was confirmed by immunoblots (Lanz et al. 1991).

Recloning in Expression Vector pQE-6. The mutagenesis and the initial tests for STS activity were performed with the PTZ19R clones. The fragments were recloned into expression vector pQE-6 (Crowe and Henco 1992) for detailed quantifications and the kinetic analysis. The optimal in-frame insertion required the modification of the vector polylinker (Fig. 2). First, the *Bam*HI site was removed by digestion with *Bam*HI, treatment with S1 nuclease, and religation. Second, the vector was restricted with *Hin*dIII, and a *Bam*HI/*Not*I-adaptor was inserted after fill-in of the *Hin*dIII site with the Klenow-fragment of DNA polymerase. The changes were confirmed by sequence analysis. For insertion of the cDNAs, the modified vector was prepared by restriction with *Sal*I, fill-in of the protruding ends, and subsequent digestion with *Bam*HI. The cDNAs were then inserted as *Eco*RI (blunt)/*Bam*HI fragments isolated from the pTZ19R constructs. The final plasmids retained the optimal promotor/translation-start configuration of the vector. The polypeptides contained at the N-terminal five additional amino acids from the vector. Previous results have shown that such changes at the N-terminal have no significant effect on enzyme activity (Lanz et al. 1991). The STS from *A. hypogaea* was recloned into pQE-6 after introduction of a *Nco*I site into the start ATG by site-directed mutagenesis.

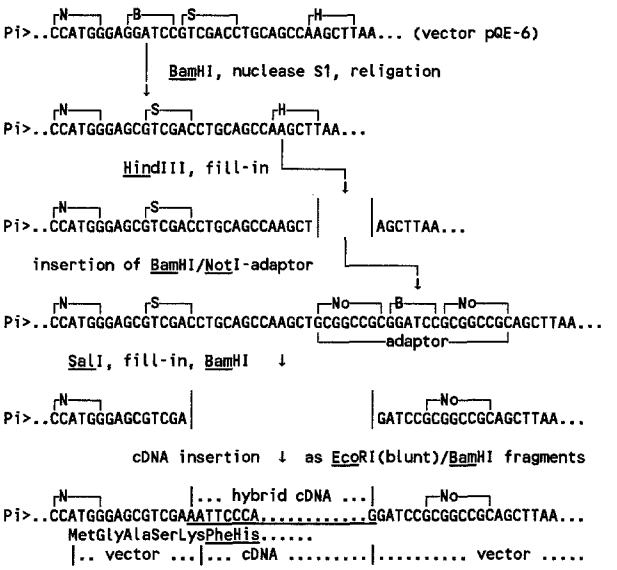

Fig. 2. Strategy for recloning of the hybrids for expression in *E. coli* vector pQE-6. The modifications in the vector polylinker were necessary for in-frame insertion of the protein-coding regions. The polypeptides were translational fusions with five vector-encoded amino acids at the N-terminal. Pi>, vector promotor. N, *Nco*I; B, *Bam*HI; S, *Sal*I; H, *Hin*dIII; No, *Not*I.

Extracts, Enzyme Assays, and Quantification. These were performed as described recently (Schanz et al. 1992), with the modification that the lysozyme concentration during lysis of the *E. coli* cells was raised to 2 mg/ml. The assay conditions were the same with all proteins and starter CoA-esters. The products were analyzed by thin-layer chromatography (TLC) with 20% acetic acid as solvent. Under these conditions the STS and CHS products were clearly distinguished by their R_f values. The radioactive products were routinely quantified with a TLC analyzer. The identity of the products had been previously established by HPLC chromatography and gas chromatography–mass spectrometry (Lanz et al. 1991).

The enzyme activities were not based on total protein in the extracts because these contained an additional protein (lysozyme) and because the rate of protein expression varied among the mutants. Each enzyme extract (soluble proteins after a centrifugation of 15 min at 15,000g) was analyzed by immunoblots, and the enzyme activity was based on the quantity of the immuno-decorated protein (Lanz et al. 1991). The proteins were visualized with an antiserum recognizing only the C-terminal half of the STS from A. hypogaea (Lanz et al. 1991). This ensured that the immuno-staining reflected the amount of hybrid protein regardless of the position and number of mutations in the CHS part of the hybrids.

Miscellaneous Techniques. Other molecular biology techniques were carried out according to published procedures (Sambrook et al. 1989).

Results

The phylogenetic tree (Fig. 3) showed the separation of the CHS proteins into some subfamilies which were largely in agreement with the systematic families, and the stable grouping within the subfamilies was reflected by good bootstrap values. The inter-subfamily relationships were not quite clear, nor was the branching order of some sequences near the outgroup. Nevertheless, it was obvious that the STS proteins did not form a cluster of their own and that they grouped with the CHS from the same or related plants rather than with each other. This suggested that there was no ancestral STS gene and that the STS (present in only a rather small number of higher plants) developed from CHS several times independently. If this hypothesis is correct, one would expect that a limited number of AA changes would be sufficient to convert a CHS into a protein with STS activity, and the experiments described below tested that possibility.

In view of the complexity of the CHS and STS reactions (Fig. 1), and in the absence of three-dimensional protein structures, it was not advisable to attempt the conversion of a complete CHS into a protein with STS function. We therefore simplified the problem by using a CHS/STS hybrid that contained 107 AAs from the N-terminal of a CHS and 287 AAs from a STS (Fig. 4). This choice was based on two considerations: (1) the CHS part encompassed the entire first exon (AAs 1–65), and this was of interest, because that exon is strictly conserved in CHS and STS genes, and (2) the STS contained a convenient restriction site facilitating the cloning procedures. The apparent simplification of the problem was compensated for by using CHS and STS from two different families (*S. alba:* Brassicaceae; *A. hypogaea:* Fabaceae) because this appeared a rigorous test of the hypothesis. The two enzymes share an overall identity of 72.2%, and another 11.3% are considered similar according to standard criteria. The hybrid C1 (Fig. 4) had no detectable STS or CHS activity. It was the starting point for site-directed mutagenesis experiments in the CHS part to obtain STS activity.

Figure 5A summarizes a comparison of the CHS (*S. alba*) and the STS (*A. hypogaea*) in the region pertinent to the experiments. The sequences shared 62.6% identity and 13.1% similarity in these parts of the proteins, and 24.3% (26 AAs) were different. The STS contained eight positions deviating from the CHS consensus from 34 CHS, and these were considered targets for the mutagenesis because the chances appeared high that they were of functional significance. Six of the exchanges could be interpreted as conservative replacements, but the other two were not likely to belong to that category (CHS Pro-42 vs STS Ala; CHS Gln-100 vs STS Glu). The latter was of particular interest because the positions 104 deviated from the CHS consensus as well, and the STS Arg-105 appeared of potential significance because all CHS contained either Val or a non-charged AA in that position. The CHS Val-103 (Met in STS) was included in the experiments because it was part of the CHS consensus until very recently (Junghans et al. 1993). The regions around AA 42 and AAs 100–105 also revealed differences in the secondary structure predictions (Fig. 5A).

The mutagenesis experiments are summarized in Fig.

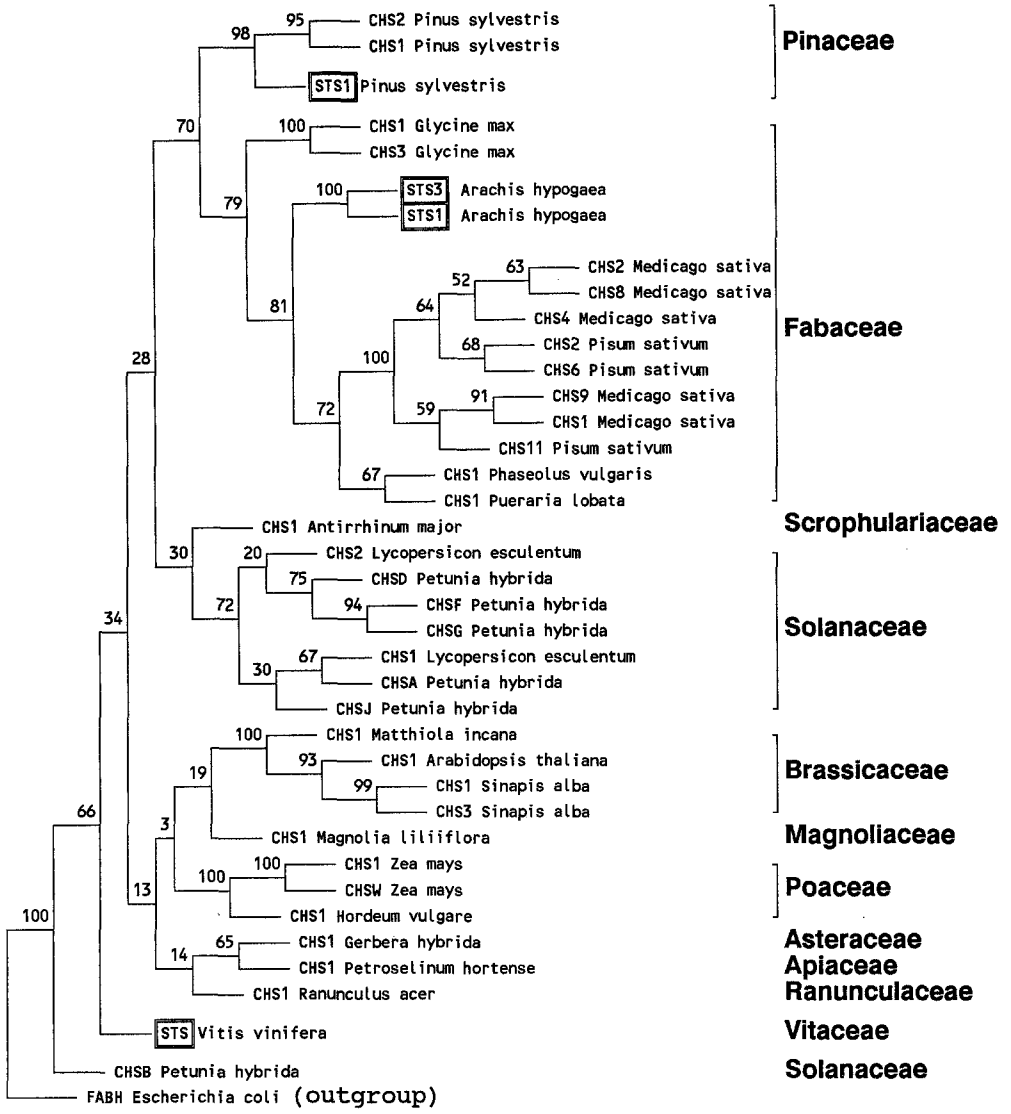

Fig. 3. Consensus phylogenetic tree of chalcone and stilbene synthase proteins. The FABH sequence of *E. coli* served as outgroup. *Numbers at the forks* indicate how often the group to the right appeared among 100 bootstrap replicates. An average of 5.4 equal parsimonious trees was found within the bootstrapped data sets. The number of required changes ranged from 1,530 to 1,714 with an average of 1,600. The systematic families are shown to the right. CHS = chalcone synthase, STS = stilbene synthase (framed); *letters and numbers* identify the individual proteins.

5B. The exchange of CHS Pro-42 to the Ala of STS had no effect at all—i.e., the protein was completely inactive (C1a). The same result was obtained with the single exchange of the strictly conserved CHS Gln-100 into the Glu of STS (C1as). Mutations of the two amino acids 103 and 105 toward STS (C1b), also in combination with the Pro-42–to–Ala exchange (C1c), were unsuccessful as well. The simultaneous mutation of the AA 100, 103, and 105 (C1r), however, led to an STS activity corresponding to a few percent of that observed with the parent STS. An additional Pro-42–to–Ala exchange (C1g) did not result in a drastic improvement of the activity. The results showed that changes of three AAs were sufficient to obtain STS activity. The Gln-100–to–Glu conversion appeared to be essential, and this conclusion was supported by four other mutants with several additional exchanges but lacking this particular mutation; all were enzymatically inactive (data not shown).

The next experiments attempted to improve the STS activity by additional mutations of AA that might have effects on either function or structure of the hybrid proteins, and this series included the complete exchange of the 11 N-terminal AAs against a sequence which leads to a fully functional enzyme when present in STS (Lanz et al. 1991). In view of the large number of possibilities it was not possible to investigate all combinations. The 14 different mutants tested, however, revealed no large increases of activity, and many had lower activities than C1r or none at all (Fig. 5B, Claw to rC1ao).

The inactive hybrid rC1ao carried all mutations performed so far. One additional exchange (CHS Gly-23

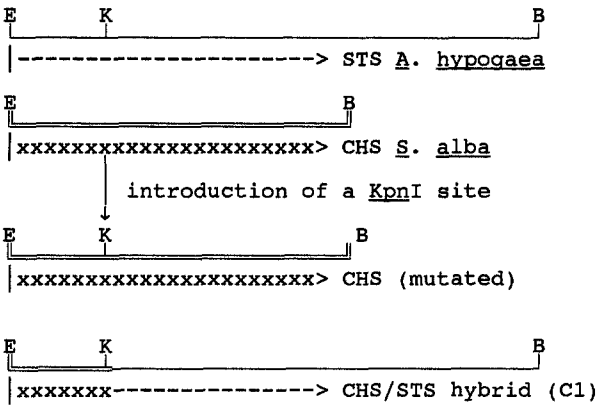

Fig. 4. Construction of a protein hybrid between CHS from *S. alba* and STS from *A. hypogaea*. The STS sequence contains a *Kpn*I site at the AA position 107/108. This restriction site was introduced into the CHS by site-directed mutagenesis at the corresponding position, and the *Eco*RI/*Kpn*I fragment in STS was then replaced by the fragment from CHS (CHS/STS hybrid C1). E, *Eco*RI; K, *Kpn*I, B, *Bam*HI. ═══, CHS; ———, STS.

to STS Thr) restored at least some STS activity (rC1ap), suggesting that this particular AA might be important. The subsequent experiments investigated whether this exchange would lead to higher enzyme activities in combination with previously tested mutations. In the series rC1aq to C1au (Fig. 5B), the highest activity was obtained with C1au (about 25% of the unchanged STS). This hybrid contained only four exchanges in the CHS part: the Gly-23–to–Thr modification and the three AAs identified before as important for STS activity (positions 100, 103, 105; C1r). Control experiments showed that mutants containing only the Gly-23–to–Thr exchange (C1av), or three exchanges, but not the Gln-100–to–Glu modification (C1an), were completely inactive. The results demonstrated that four AA changes in a CHS sequence of 107 AAs were sufficient to obtain quite high STS activities. None of the mutants had CHS activity (formation of chalcone instead of stilbene). To investigate the effect of the four mutations on CHS activity, the mutated CHS part was retransferred from C1au to the CHS from *S. alba*. The replacement led to an almost complete loss of CHS activity (≤0.6% of the parent CHS), and STS activity was not detectable. The results indicated that these four AAs are important for CHS activity but that they were not sufficient to convert the complete CHS protein into a functional STS.

The measurements described so far were performed with the proteins expressed from vector pTZ19R. This was adequate for the screening of the mutants, but a more detailed characterization required recloning in an expression vector. This was performed with the STS itself and representative mutants, and the results are summarized in Table 1. They confirmed the data obtained with the pTZ19R vector system, except for mutant C1ao, which under the improved expression conditions revealed higher activities than before. This hybrid contained the Pro-42–to–Ala exchange in addition to the four important mutations identified with C1au. Its activity was still slightly lower than that of C1au, confirming the previous conclusion that the Pro-42 absolutely conserved in CHS was not a critical determinant in the conversion of CHS to STS activity.

Table 1 also summarizes data which are of interest for the interpretation of the mutants. The enzyme measurements were performed with the supernatants of a 15,000*g* centrifugation, but the presence and percentage of immunoreactive proteins in the pellet was routinely tested as well. With unmodified STS, about 60% were in the soluble fraction, and the remaining 40% in the pellet were enzymatically inactive. This was expected, because a certain percentage of foreign proteins expressed in *E. coli* is usually found in inclusion bodies (Marston 1986), and insolubility has been used as a diagnostic tool for misfolded protein aggregates (Mitraki and King 1992). Interestingly, the hybrids with enzyme activity had a distribution similar to that of STS, while the inactive proteins were mostly found in the insoluble fraction (Table 1; the exception C1an will be discussed below). The correlation between protein solubility and activity was also observed with other hybrids, and sizing chromatography (gel filtration) confirmed that the insoluble proteins were aggregates and not the expected dimers (data not shown). The hybrid C1an was an interesting exception, because the percentage of the protein in the soluble fraction was quite high in the absence of detectable enzyme activity. C1an lacked the Gln-100–to–Glu mutation of the very active C1au hybrid, and this confirmed the important role of that AA for enzyme activity.

The CHS/STS hybrids C1ao and C1au had STS activity, but it was lower than that of the original STS. It was therefore of interest whether this reflected differences in substrate affinity (K_m), rate of reaction (V_{max}), or both. The kinetic analysis with 4-coumaroyl-CoA as starter CoA-ester showed that the K_ms were very similar and that a lower V_{max} was responsible for the lower activities of the mutants (Fig. 6).

The *A. hypogaea* STS and the *S. alba* CHS prefer 4-coumaroyl-CoA as starter substrate, corresponding to the stilbene or naringenin chalcone derivatives found in these plants. Like all tested STS and CHS, these proteins also accept related CoA-esters in vitro—for example, cinnamoyl-CoA and dihydrocinnamoyl-CoA (Fliegmann et al. 1992; Schröder and Schröder 1992; Schanz et al. 1992). *A. hypogaea* does not contain the products of these STS reactions (pinosylvin and dihydropinosylvin, or their derivatives), but they do occur in other plants—e.g., in pines (Gorham 1989; Kindl 1985). The hybrids represented new enzyme creations, and we therefore tested whether they possessed the same substrate preferences as the parent enzymes. Table 2 shows

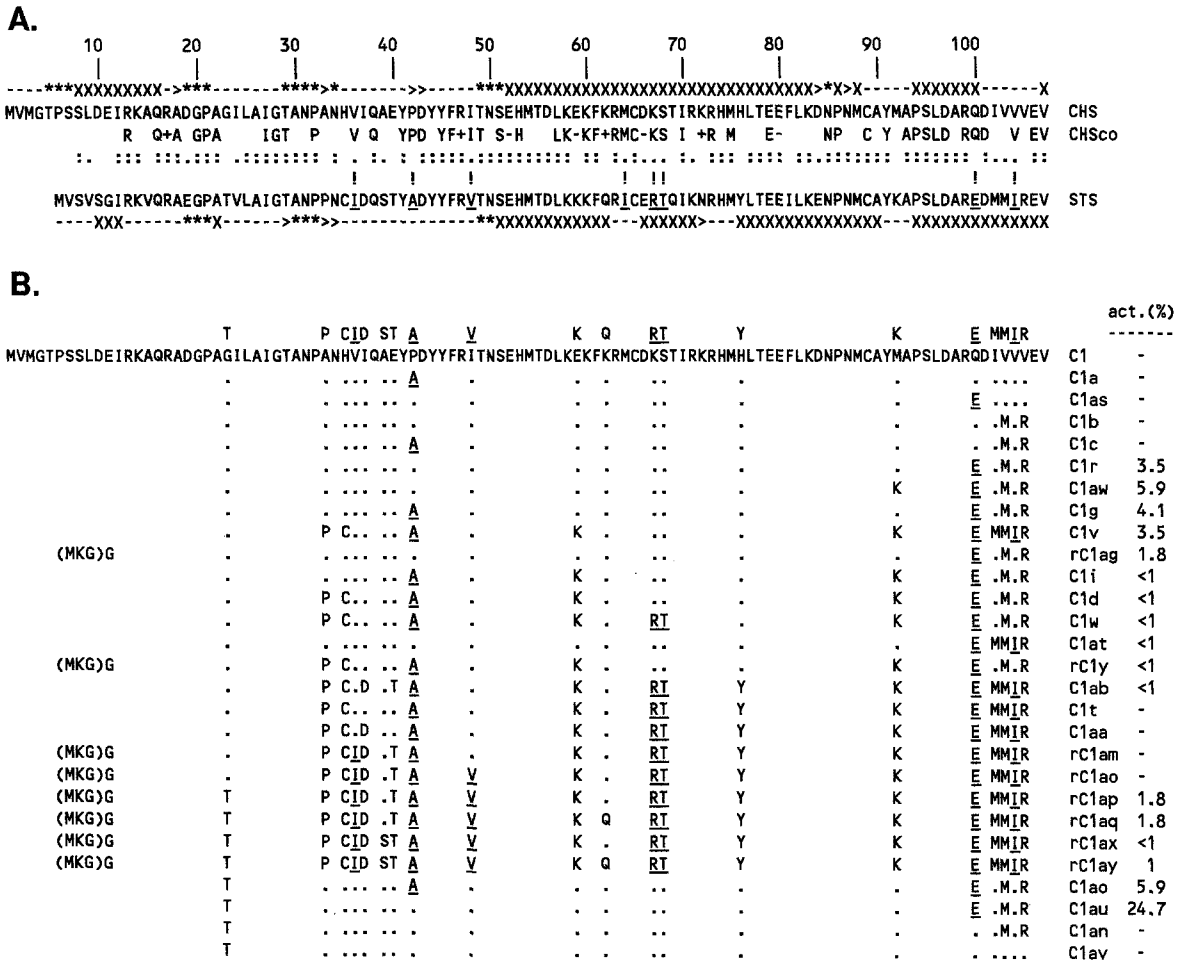

Fig. 5. **A** The N-terminal 107 AAs of the CHS from *S. alba*, CHS consensus (CHSco, from 34 CHS), alignment with the STS from *A. hypogaea*, and secondary structure predictions. *Double dots*, identity; *dots*, similarity. + and − in CHSco: AAs which contain a positive or negative charge in all CHS. *Exclamation marks:* deviations in STS from the CHS consensus. Symbols for predicted conformations: *X*, helical; −, extended; >, turn; *, coil (program GARNIER, PCGENE software package). **B** CHS sequence of the CHS/STS hybrid C1 (printed fully) and of the mutants (only the exchanged AA; *dots* indicate no AA change). The activities are given in % of the STS from *A. hypogaea;* −, not detectable. On top of C1, summary of all exchanged AA. *Underlined* in the mutants: changes from the CHS consensus. (MKG) indicates that the first three AAs in these hybrid proteins originated from translational fusions with vector-encoded AA.

that this was not the case, because the CHS/STS hybrids preferred dihydrocinnamoyl-CoA against 4-coumaroyl-CoA.

Discussion

The phylogenetic tree shows that the CHS proteins grouped into subfamilies which largely corresponded to the systematic families. Many species were represented by only one sequence, and this was most likely the reason for the uncertainty in the branching order of some proteins near the outgroup. Interestingly, often the members of CHS families did not form a species-specific cluster but instead formed subgroups with CHS from other plants (*Petunia hybrida, Medicago sativa, Pisum sativum, Lycopersicon esculentum*). This could indicate a specialization of enzyme properties within the CHS families, but heterologous expression and investigations of the enzyme properties have not been performed. CHSB from *Petunia hybrida* represented an extreme example, because it was positioned next to the outgroup and quite apart from the other *Petunia* CHS (Fig. 3). It had been noted before that the sequence of CHSB was unusual, and specialized enzyme properties or a gene in the process of being inactivated were offered as possible explanations (Koes et al. 1989).

Most importantly, the tree showed that the STS did not form a cluster of their own. They grouped with the CHS from the same or related plants rather than with each other. This could be explained by highly convergent evolution from originally different functions, driven by functional and structural requirements imposed on proteins of CHS or STS activity. In this case one would expect similar proteins, but not necessarily the same gene structures. These, however, are similar to the extent that all known CHS and STS share a conserved intron at the same position, splitting a cysteine strictly

Table 1. Properties of selected CHS/STS hybrids (experiments performed with the hybrids recloned in expression vector pQE-6)

	Position						
	23	42	100	103	105		
	Exchange						
	G → T	P → A	Q → E	V → M	V → R	STS activity (%)[a]	Protein in supernatant (%)[b]
STS						100	60
C1						—[c]	≤5
C1r			+	+	+	3–5	40
C1au	+		+	+	+	20–25	50
C1ao	+	+	+	+	+	15–20	40
C1a		+				—	≤5
C1b				+	+	—	≤15
C1an	+			+	+	—	50

[a] With 4-coumaroyl-CoA; assays under conditions of linear product formation; specific activities normalized to protein amounts in immunoblots
[b] Ratio of soluble (supernatant of a 15-min centrifugation at 15,000*g*) to total immunoreactive protein (both determined by immunoblots)
[c] No detectable activity

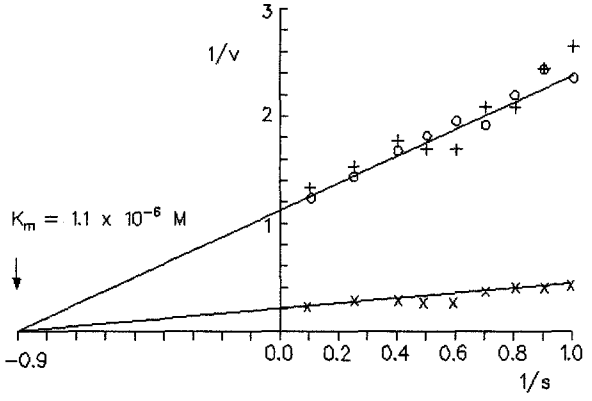

Fig. 6. Kinetic analysis of the STS from *A. hypogaea* (*x*), and the CHS/STS hybrids c1au (+) and C1ao (*o*) with 4-coumaroyl-CoA as starter CoA-ester. See Table 1 for description of the mutants. The experiments were performed with the hybrids recloned in expression vector pQE-6.

Table 2. Substrate preferences of STS (*A. hypogaea*), CHS/STS hybrids with STS activity, and CHS (*S. alba*) with three different starter CoA-esters[a]

	Relative activity (%) with		
protein	4-coumaroyl-CoA	cinnamoyl-CoA	dihydro-cinnamoyl-CoA
STS	100	50 ± 10	85 ± 10
C1au	40 ± 10	55 ± 5	100
C1ao	35 ± 10	70 ± 10	100
C1r	45 ± 10	55 ± 10	100
CHS	100	65 ± 5	20 ± 5

[a] The activity with the best substrate was set as 100%. See Table 1 for the relative activities

conserved in both enzymes. This point clearly argues for a common evolutionary origin of CHS and STS. The fact that the STS grouped with CHS and not as a separate cluster suggested that there was no ancestral STS gene in the strict sense, that STS evolved from CHS, and that this occurred several times independently in the course of evolution.

We speculated that a limited number of AA exchanges could convert a CHS into a STS. The successful hybrids showed that three changes in a CHS sequence of 107 AAs were sufficient to obtain STS activity with the hybrid protein and that one additional mutation led to a remarkable further increase (Table 1). Our approach was mostly based on consensus observations and secondary structure predictions. One of the deductions was correct (region 100–105), in particular with respect to the Gln-100–to–Glu exchange which was essential for catalytic activity (C1an vs C1au, Table 1). The second was not correct (Pro-42), and the effect of the CHS Gly-23 change to the Thr in STS was remarkable, but hardly predictable, because most other CHSs contain Thr in this position, and Gly appears to be confined to the Brassicaceae family. Interestingly, this AA is in the first exon (AAs 1–65), while the other successful exchanges are in the second exon. It should be noted, however, that the modifications performed in this work were not necessarily the only possible solution, and there may be even simpler ways to solve the problem even with complete CHS proteins.

The kinetic analysis indicated that the differences between the STS and the active hybrids were in V_{max}, not in K_m (Fig. 6), and this appears plausible, because the binding site for the starter CoA-ester (active site for the condensing reaction) is located in the STS part of the hybrid (Cys-169) (Lanz et al. 1991). Interestingly, the hybrids had a different substrate preference than the parent enzymes. The side chain in the dihydrocinnamoyl residue is more flexible than the double-bond configuration in the other two substrates, and the resulting higher flexibility of the tetraketide intermediate may be easier to handle with imperfect proteins like the hybrids.

The condensing reactions of CHS and STS are very similar or identical (Schröder and Schröder 1990; Lanz et al. 1991), but the configuration of the enzyme-bound

tetraketide intermediate must be different to obtain either a chalcone or a stilbene as final product (Fig. 1). One would therefore predict that the conversion of CHS to STS involves changes in the three-dimensional structure of the protein without affecting the protein-folding pathway or the capacity of the subunits for dimerization. This is a prerequisite for testing the effects of AA exchanges on enzyme activity. Most of the hybrids formed insoluble aggregates, indicating that these and the original hybrid C1 failed in the first place because they could not fold or dimerize correctly. The interplay of structure and function is not understood, and the use of an interfamily hybrid certainly represented an additional complication. To the best of our knowledge, interspecies chimeras have been used rarely so far to create new enzyme activities (e.g., Pompon and Nicolas 1989), although site-directed mutagenesis is a standard technique for that purpose. (See Aoyama et al. 1989; Clarke et al. 1989; Iwasaki et al. 1993; Juvonen et al. 1991 for some examples.)

In conclusion, the experiments showed that three AA exchanges were sufficient to adapt the N-terminal 107 AAs of a CHS to STS function in an artificial protein with enzyme parts from different plant families. This suggests that comparably simple changes (but not necessarily the same) were sufficient to change a CHS into a protein with STS activity in the course of evolution. Such an enzyme would not have to be very efficient initially (like C1r), because the advantage of producing a new phytoalexin as defense against pathogen attack would favor the selection of mutational changes toward more effective enzymes. The hypothesis is consistent with the phylogenetic tree which indicates that STS evolved from CHS several times independently.

Acknowledgments. We thank Y. Helariutta for the unpublished sequence of the *Gerbera hybrida* CHS, and U.-G. Maier and M. Scheufens for placing computer facilities at our disposal. This work was supported by Deutsche Forschungsgemeinschaft (SFB206) and Fonds der Chemischen Industrie.

References

Akada S, Kung SD, Dube SK (1990) The nucleotide sequence of gene 3 of the soybean chalcone synthase multigene family. Nucleic Acids Res 18:5899

Aoyama T, Korzekwa K, Nagata K, Adesnik M, Reiss A, Lapenson DP, Gillette J, Gelboin HV, Waxman DJ, Gonzalez FJ (1989) Sequence requirements for cytochrome P-450IIB1 catalytic activity. Alteration of the stereospecificity and regioselectivity of steroid hydroxylation by a simultaneous change of two hydrophobic amino acid residues to phenylalanine. J Biol Chem 264:21327–21333

Clarke AR, Atkinson T, Holbrook JJ (1989a) From analysis to synthesis: new ligand binding sites on the lactate dehydrogenase framework. Part I. Trends Biochem Sci 14:101–105

Clarke AR, Atkinson T, Holbrook JJ (1989b) From analysis to synthesis: new ligand binding sites on the lactate dehydrogenase framework. Part II. Trends Biochem Sci 14:145–148

Crowe J, Henco K (1992) The QIAexpressionist. Qiagen Inc, Chatsworth, CA

Ehmann B, Schäfer E (1988) Nucleotide sequences encoding two different chalcone synthases expressed in cotyledons of SAN9789 treated mustard (*Sinapis alba* L.). Plant Mol Biol 11:869–870

Epping B, Kittel M, Ruhnau B, Hemleben V (1990) Isolation and sequence analysis of a chalcone synthase cDNA of *Matthiola incana* R. Br. (Brassicaceae). Plant Mol Biol 14:1061–1063

Feinbaum RL, Ausubel FM (1988) Transcriptional regulation of the *Arabidopsis thaliana* chalcone synthase gene. Mol Cell Biol 8:1985–1992

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Felsenstein J (1989) PHYLIP—phylogeny inference package (version 3.2). Cladistics 5:164–166

Fliegmann J, Schröder G, Schanz S, Britsch L, Schröder J (1992) Molecular analysis of chalcone and dihydropinosylvin synthase from Scots pine (*Pinus sylvestris*), and differential regulation of these and related enzyme activities in stressed plants. Plant Mol Biol 18:489–503

Franken P, Niesbach-Klösgen U, Weydemann U, Maréchal-Drouard L, Saedler H, Wienand U (1991) The duplicated chalcone synthase genes *C2* and *Whp* (white pollen) of *Zea mays* are independently regulated; evidence for translational control of *Whp* expression by the anthocyanin intensifying gene *in*. EMBO J 10:2605–2612

Gorham J (1989) Stilbenes and phenanthrenes. In: Harborne JB (ed) Methods in plant biochemistry, vol 1, plant phenolics. Academic Press, London, pp 159–196

Higgins DG, Bleasby AJ, Fuchs R (1992) CLUSTAL V: improved software for multiple sequence alignment. Cabios 8:189–191

Ichinose Y, Kawamata S, Yamada T, An C, Kajiwara T, Shiraishi T, Oku H (1992) Molecular cloning of chalcone synthase cDNAs from *Pisum sativum*. Plant Mol Biol 18:1009–1012

Iwasaki M, Darden TA, Pedersen LG, Davis DG, Juvonen RO, Sueyoshi T, Negishi M (1993) Engineering mouse $P450_{coh}$ to a novel corticosterone 15α-hydroxylase and modeling steroid-binding orientation in the substrate pocket. J Biol Chem 268:759–762

Junghans H, Dalkin K, Dixon RA (1993) Stress responses in alfalfa (*Medicago sativa* L.). 15. Characterization and expression patterns of members of a subset of the chalcone synthase multigene family. Plant Mol Biol 22:239–253

Juvonen RO, Iwasaki M, Negishi M (1991) Structural function of residue-209 in coumarin 7-hydroxylase (P450coh). Enzyme-kinetic studies and site-directed mutagenesis. J Biol Chem 266: 16431–16435

Kindl H (1985) Biosynthesis of stilbenes. In: Higuchi T (ed) Biosynthesis and biodegradation of wood components. Academic Press, New York, pp 349–377

Koes RE, Spelt CE, van den Elzen PJM, Mol JNM (1989) Cloning and molecular characterization of the chalcone synthase multigene family of *Petunia hybrida*. Gene 81:245–257

Kunkel TA (1985) Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc Natl Acad Sci USA 82:488–492

Lanz T, Schröder G, Schröder J (1990) Differential regulation of genes for resveratrol synthase in cell cultures of *Arachis hypogaea*. Planta 181:169–175

Lanz T, Tropf S, Marner F-J, Schröder J, Schröder G (1991) The role of cysteines in polyketide synthases: site-directed mutagenesis of resveratrol and chalcone synthases, two key enzymes in different plant-specific pathways. J Biol Chem 266:9971–9976

Marston FAO (1986) The purification of eukaryotic polypeptides synthesized in *Escherichia coli*. Biochem J 240:1–12

Melchior F, Kindl H (1990) Grapevine stilbene synthase cDNA only slightly differing from chalcone synthase cDNA is expressed

in *Escherichia coli* into a catalytically active enzyme. FEBS Lett 268:17–20

Mitraki A, King J (1992) Amino acid substitutions influencing intracellular protein folding pathways. FEBS Lett 307:20–25

Nakajima O, Akiyama T, Hakamatsuka T, Shibuya M, Noguchi H, Ebizuka Y, Sankawa U (1991) Isolation, sequence and bacterial expression of a cDNA for chalcone synthase from the cultured cells of *Pueraria lobata.* Chem Pharm Bull (Tokyo) 39: 1911–1913

Niesbach-Klösgen U, Barzen E, Bernhardt J, Rohde W, Schwarz-Sommer Z, Reif HJ, Wienand U, Saedler H (1987) Chalcone synthase genes in plants: a tool to study evolutionary relationships. J Mol Evol 26:213–225

O'Neill SD, Tong Y, Spörlein B, Forkmann G, Yoder JI (1990) Molecular genetic analysis of chalcone synthase in *Lycopersicum esculentum* and an anthocyanin-deficient mutant. Mol Gen Genet 224:279–288

Pompon D, Nicolas A (1989) Protein engineering by cDNA recombination in yeasts: shuffling of mammalian cytochrome P-450 functions. Gene 83:15–24

Reif HJ, Niesbach U, Deumling B, Saedler H (1985) Cloning and analysis of two genes for chalcone synthase from *Petunia hybrida.* Mol Gen Genet 199:208–215

Reimold U, Kroeger M, Kreuzaler F, Hahlbrock K (1983) Coding and 3′ non-coding nucleotide sequence of chalcone synthase mRNA and assignment of amino acid sequence of the enzyme. EMBO J 2:1801–1805

Rohde W, Dörr S, Salamini F, Becker D (1991) Structure of a chalcone synthase gene from *Hordeum vulgare.* Plant Mol Biol 16: 1103–1106

Ryder TB, Hedrick SA, Bell JN, Liang X, Clouse SD, Lamb CJ (1987) Organization and differential activation of a gene family encoding the plant defense enzyme chalcone synthase in *Phaseolus vulgaris.* Mol Gen Genet 210:219–233

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor

Schanz S, Schröder G, Schröder J (1992) Stilbene synthase from Scots pine (*Pinus sylvestris*). FEBS Lett 313:71–74

Schröder G, Brown JWS, Schröder J (1988) Molecular analysis of resveratrol synthase: cDNA, genomic clones and relationship with chalcone synthase. Eur J Biochem 172:161–169

Schröder G, Schröder J (1992) A single change of histidine to glutamine alters the substrate preferences of a stilbene synthase. J Biol Chem 267:20558–20560

Schröder J, Schanz S, Tropf S, Kärcher B, Schröder G (1993) Phytoalexin biosynthesis: stilbene synthase and co-action of a reductase with chalcone synthase. In: Fritig B, Legrand M (eds) Mechanisms of plant defense responses. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 257–267

Schröder J, Schröder G (1990) Stilbene and chalcone synthases: related enzymes with key functions in plant-specific pathways. Z Naturforsch 45c:1–8

Sommer H, Saedler H (1986) Structure of the chalcone synthase gene of *Antirrhinum majus.* Mol Gen Genet 202:429–434

Tsay J-T, Oh W, Larson TJ, Jackowski S, Rock CO (1992) Isolation and characterization of the β-ketoacyl carrier protein synthase III gene (*fab*H) from *Escherichia coli* K-12. J Biol Chem 267:6807–6814

Wu CFJ (1986) Jackknife, bootstrap and other resampling plans in regression analysis. Ann Statist 14:1261–1295