Evolutionary Conservation of C-terminal Domains of Primary Sigma$^{70}$-type Transcription Factors between Plants and Bacteria*

Mohamed-Ali Hakimi, Isabelle Privat, Jean-Gabriel Valay, and Silva Lerbs-Mache‡

From the Laboratoire de Génétique Moléculaire des Plantes, "Plastes et Différenciation Cellulaire," Université Joseph Fourier and CNRS, B.P. 53X, F-38041 Grenoble, France

Three different cDNAs coding for putative plant plastid sigma$^{70}$-type transcription initiation factors have recently been cloned and sequenced from Arabidopsis thaliana. We have analyzed the evolutionary conservation of function(s) of the N-terminal and C-terminal halves of these three sigma factors by in vitro transcription studies using heterologous transcription systems and by complementation assays using Escherichia coli thermosensitive $rpoD$ mutants. Our results indicate differences and similarities of the three plant factors and their prokaryotic ancestors. The functions of the N-terminal parts of the plant sigma factors are considerably different from the function of the N-terminal part of the principal sigma$^{70}$ factor of $E. coli$. On the other hand, the C-terminal parts have kept at least two characteristics when compared with their prokaryotic ancestors: 1) they can distinguish between different promoter structures, and 2) one of them is capable of fully complementing $E. coli$ $rpoD$ mutants, i.e. recognizing all essential $E. coli$ promoters that are used by the $E. coli$ principal sigma$^{70}$ factor. This shows for the first time in vivo a strong evolutionary conservation of cis- and trans-acting elements between the prokaryotic and the plant plastid transcriptional machinery.

The higher plant plastid genome is transcribed by two different transcriptional systems (reviewed in Ref. 1). One of the two systems is phage-like and its evolutionary origin is still unclear (2, 3). The other system is of the prokaryotic type revealing the cyanobacterial origin of present-day chloroplasts (4–6). All polypeptides that are necessary to build up the core enzyme of the prokaryotic-type plastid RNA polymerase are still encoded in the plastid genome (7, 1). However, genes corresponding to sigma-like transcription initiation factors that are indispensable for the activity of this type of enzyme (reviewed in Ref. 8) are not found on the plastid genomes (7, 1). These genes have been transferred to the nucleus during evolution (9).

The existence of more than one sigma-like transcription factor in higher plant plastids had been suggested several years ago on the basis of biochemical approaches for spinach and mustard (10–12). However, the corresponding genes have not been cloned. Only recently, six different cDNAs showing strong sequence similarity with genes coding for prokaryotic-like sigma$^{70}$-type initiation factors have been cloned and sequenced from Arabidopsis thaliana (accession numbers: SIG1, dbj: AB004820 and emb: Y15362; SIG2, dbj: AB004821 and emb: Y14567; SIG3, dbj: AB004822; SIG4, gb: AF101075; SIG5, emb: Y18550; and SIG6, emb: AJ250812). For three of them, it has been shown that the corresponding proteins are transported into the chloroplasts (9, 14), but the function of these putative transcription factors in the regulation of plastid gene expression is not clear. In particular, it is not clear why several sigma$^{70}$-type factors have been conserved during evolution, despite an important reduction of (plastid) genome size.

Prokaryotic sigma factors can be classified into two families. Members of the first family are similar to the Escherichia coli sigma$^{70}$ factor, whereas members of the other family are similar to the $E. coli$ sigma$^{54}$ factor. The sigma$^{70}$ family of transcription factors can be further subdivided into two or three groups. One primary sigma factor is required to ensure that “housekeeping” functions are performed, and alternative sigma factors direct transcriptional responses to changing environmental conditions (8). Although both groups of sigma$^{70}$ factors are similar with respect to their amino acid sequence, they recognize different sequences at the two promoter elements localized −10 and −35 base pairs upstream of the transcription start sites (reviewed in 8 and 15).

In the present study, we attempted to obtain information about the evolutionary conservation of components of the transcriptional machinery between prokaryotes and higher plant plastids. With this aim, we analyzed differences of the three plant sigma factors with respect to promoter recognition by in vitro transcription, and we determined the plant analogue to the primary sigma factor of $E. coli$ by complementation of $E. coli$ $rpoD(ts)$ mutants. For the first time, we show in vivo complementation of a protein of the $E. coli$ transcriptional machinery by a protein of the plastid transcriptional machinery.

MATERIALS AND METHODS

Isolation of Clones Coding for A. thaliana Plastid Sigma Factors—A. thaliana cDNA and/or genomic sequences, similar to prokaryotic sigma factors, have been searched for in currently available data banks. Six such sequences have been found. Four of them have been detected as partial cDNA sequences in the data base expressed sequence tag (dbEST), and two of them have been identified in the genomic sequences resulting from the Arabidopsis Genomic Initiative. The corresponding full-length cDNA clones were isolated from an A. thaliana Matchmaker cDNA library (CLONTECH) either by screening with the corresponding A. thaliana EST clones or by direct PCR amplification (accession numbers: SIG1, Y15362; SIG2, Y14567; SIG5, Y18550; and SIG6, AJ250812).

Overproduction in E. coli and Purification of Recombinant Proteins—Sequences corresponding to full-length and truncated forms of SIG1, SIG2, and SIG3 polypeptides were PCR-amplified using the following primers at the 5′ ends: SIG1-KLR, 5′-GGCAGATCCAGCTTCTGTCGCCAGG-
Conservation of C-term-Domains of α70-Type Factors

TT-3; SIG1-RQR, 5'-CCGGATATCCAGCGCAACTGATACATAGGCG-3'; SIG2-STE, 5'-CCGGATATCCCTCTCATGAGAAGGC-3'; SIG2-VRG, 5'-CCGGATATCCTGATGCTTGTAGTAAAAGGTAAGT-3'; SIG3-MLV, 5'-CCGGATATCCATGTGTTCTTGTAGTACATCTC-3'; and SIG3-GLR (5'-CCGGATATCCTGATGCTTGTAGTACCAGAAGGC-3'). Primers corresponding to the 3' ends were as follows: SIG1, 5'-CCGGATATCCCTATGTTGGAACCAAGTA3'-3', SIG2, 5'-CCGGATATCCCTCAATCC-3', and SIG3, 5'-CCGGATATCCATGTTGGAACCAAGTA3'-3'.

PCR products were digested with BamHI, SalI, SacI, or EcoRI (primer sequences corresponding to the added restriction sites are in boldface letters) and cloned into the corresponding sites of the expression vectors pET28a (Novagen) and pQE30 (Qiagen) to generate pMA plasmids. Plasmids were amplified in *E. coli BL21* (DE3, pET28a) or M15 (pQE30). After induction of protein expression with isopropyl-β-D-thiogalactopyranoside, the His fusion proteins were recovered as inclusion bodies. After solubilization in lysis buffer (20 mM Tris/HCl, pH 8.0, 100 mM NaCl, 8 M urea) containing 1 mg/ml lysozyme, the fusion proteins were purified on Ni²⁺-NTA columns according to the supplier's protocol (Qiagen). After elution with 20 mM Tris/HCl, pH 8.0, 100 mM NaCl, 8 M urea and 300 mM imidazole, the proteins were renatured by dialysis against 50 mM Tris/HCl, pH 8.0, 1 mM EDTA, 1 mM dithiothreitol, 20% glycerol and stored at −20 °C until use. The purity of each protein preparation was checked by SDS-polyacrylamide gel electrophoresis followed by Coomassie Blue staining of the protein.

Construction of Hybrid Sigma Factors and Complementation of *E. coli* rpoD Mutants—cDNAs corresponding to full-length hybrid sigma factors were produced by PCR recombination (25) using four different oligonucleotides for each of the constructs. The 5' and 3' plant-specific primers were as indicated above (SIG1-KLR, SIG2-STE, and SIG3-MLV), and the 3' *E. coli* rpoD-specific primer was 5'-GGGGTACCTAGTGAGCAAAGACGACGCTGTCAG-3', with an added KpnI site marked in boldface letters. The complementary recombination primers were as follows: SIG1-KpDNA, 5'-gaatgggtaagggagcAAGACTGCTGTTTTG-3', 5'-AACGAGTCGAATGTTCAACATCAGC-3'; SIG2-KpDNA, 5'-ggatggtggagagcAATGTTCAACATCAGC-3', 5'-AACAAACGCAACTGTTCAACATCAGC-3'; SIG3-KpDNA, 5'-ggatggtggagagcAAGACTGCTGTTTTG-3', 5'-AACAAGACGACGACGCTGTCAGC-3'. The parts of the primers that correspond to the *E. coli* hybrid sigma factor of *E. coli* are shown in lowercase letters. After cleavage with the appropriate restriction enzymes, the PCR fragments were inserted into Bluescript KS vectors to give pMA plasmids.

For complementation studies, *E. coli* strains UQ285 and CAG1 were obtained from the *E. coli* Genetic Stock Center (Yale University, New Haven, CT) and transformed with the corresponding plasmids. Transformants were selected on t-broth-ampicillin plates at 32 °C and screened at 42 °C.

In *Vito Transcription—In vitro* transcription reactions were performed at 37 °C in 25-µl assays. The reaction mixture contained 40 mM Tris/HCl (pH 8.0); 20 µM dithiothreitol; 14 mM MgCl₂; 40 µM EDTA; 50 mM NaCl; 300 µM each of GTP, ATP, and CTP; and 5 µM UTP, including 20 µCi of [α-³²P]UTP; 71 nt template DNA; and 0.05 units of *E. coli* holozyme (Roche Molecular Biochemicals) or 26 µM of *E. coli* core enzyme (Epicentre Technologies). Sigma proteins were added to give final concentrations of 100, 200, or 300 nM. Reaction mixtures were preincubated for 5 min without template at 37 °C. Transcription was started by addition of DNA template, and the reactions were stopped after 15 min by extraction with phenol/chloroform and precipitated with ethanol. Transcription products were analyzed on 8% acrylamide/urea gels. The template DNA was used in supercoiled form. It corresponded to the plasmid pTZ19 harboring the corresponding promoter regions and the *E. coli* throneine attenuator (2, 4).

Western Blot Analysis of *A. thaliana* Proteins—Antibodies were produced in rabbits, and horseradish peroxidase-coupled secondary antibodies were detected using the ECL™ Western blotting detection system (Amersham Pharmacia Biotech).

RESULTS

Overproduction of Entire and Truncated Plant Sigma Factors in *E. coli*—We have cloned and sequenced the cDNAs corresponding to six putative *A. thaliana* sigma factors as described under “Materials and Methods.” We have analyzed the function of three of them (named SIG1, SIG2, and SIG3, according to Ref. 9) in more detail. To this end, selected full-length clones have been used to amplify the coding regions and to clone them into two different expression vectors, pET28a and pQE30. Regions coding for supposed transit peptides have been eliminated in the cloning step, thus locating the N-terminal three amino acids to the sequences KLR, STE, and MLV, respectively, as indicated in Fig. 1A. Following transformation of *E. coli* BL21 or M15 and induction with isopropyl-1-thio-

![Fig. 1. Comparison of components of the plant plastid “E. coli”-like and the *E. coli* transcriptional machineries. A, schematic representation of the principal sigma-70 factor of *E. coli* and of the three putative *A. thaliana* plastid sigma-like factors, SIG1, SIG2 and SIG3. The localization of the conserved domains is indicated by filled, shaded, or patterned boxes as shown, and the N-terminal ends of the full-length or truncated plant proteins, which are used in the following studies, are marked by vertical arrows shown with the three N-terminal amino acids. B, growth curves of *E. coli* cultures transformed with expression vectors coding for plant sigma-like factors. The time of induction by isopropyl-1-thio-β-D-galactopyranoside is indicated by a vertical arrow. C, sequence comparison of the three plastid *E. coli*–like promoter structures that are used in these studies with the consensus sequence of the principal *E. coli* sigma-70 factor.
protein corresponding to the C-terminal part could be produced in large amounts without any harmful effects for *E. coli*.

Having thus shown that the N-terminal parts of two plant sigma-like factors might be incompatible with the *E. coli* transcriptional machinery, we recloned only the 3' portions of the three plant cDNAs that code for three essential functions of a sigma factor: promoter recognition (regions 2.4 and 4.2), DNA melting (region 2.3), and RNA polymerase-core interactions (regions 2.1 and 3.2, reviewed in Ref. 8). It has already been shown that truncated sigma factors retain DNA binding and core RNA polymerase binding activity (16–18). Therefore, we assumed that it should be possible to use such truncated proteins to characterize these two functions by *in vitro* transcription assays. Sequence alignment of the amino acid sequences of the three plant sigma factors with the one of *E. coli* shows that the C-terminal regions, but not the N-terminal parts, of the proteins are highly conserved. Each of these truncated proteins (detailed in Fig. 1A) could be produced in large amounts in *E. coli*, indicating that the N-terminal parts of SIG1 and SIG3 are responsible for the failure of overproduction of the entire proteins.

Characterization of Truncated Sigma Factors by *in Vitro* Transcription—SIG2 was the only sigma factor that could be obtained as an entire protein by overproduction in *E. coli*. We used this protein to analyze whether our assumption was correct, i.e. that the truncated proteins could be used to characterize the two essential functions of a sigma factor: promoter recognition and RNA polymerase-core interaction. To this aim we compared the full-length SIG2 factor (SIG2-STE) with its truncated form (SIG2-VRG) in an *in vitro* transcription assay using three different plant plastid promoters with the *E. coli* core enzyme (Fig. 2). The sequences of the three different plastid promoters and their comparison to the *E. coli* consensus promoter sequence, which is recognized by the *E. coli* primary sigma factor, RpoD, are shown in Fig. 1C. These particular plastid promoters have been chosen for the following reason: The *rbcL* promoter represents one highly transcribed plastid promoter, regulating the transcription of the most abundant chloroplast protein and thus implying that it should be recognized by a primary sigma factor. The two other plastid promoters, *rrn*P1 and *rrn*P2, have plastid-specific regulatory functions in the expression of the *rrn* operon (4). Therefore, they represent good candidates to be recognized by an alternative sigma factor that fulfills specific functions in the down-regulation of *rrn* expression during plastid development.

The *E. coli* holoenzyme recognizes all three plastids promoters in *in vitro* transcription assays (Fig. 2, lanes 15 and 18), albeit with different efficiencies (asterisk notes *rrn*P2; note that *rrn* transcription is exposed four times longer than *rbcL* transcription). The *E. coli* core enzyme produces a longer RNA than the one that is initiated at the –180 *rbcL* promoter (Fig. 2, lanes 1 and 8, asterisk). We did not map the exact location of the 5’ end of this RNA, but the production of this RNA can serve to verify the saturation by sigma factor of commercially available *E. coli* RNA polymerase preparations (see Fig. 6, lanes 1–3 and 18). Reconstitution of the *E. coli* core enzyme with either the full-length (Fig. 2, lanes 1–7; exposure times were 12 h for lanes 1–4 and 2 days for lanes 3’, 4’, and 5) or the truncated form of SIG2 (Fig. 2, lanes 8–14; exposure time was 2 days) gave similar results. The nonspecific transcription product that was obtained when using the cloned *rbcL* promoter as template diminished gradually with increasing SIG2 concentrations, and very small amounts of the specific transcript (–180) were produced. The appearance of the specific transcript does not quantitatively correlate with the disappearance of the nonspecific transcript. We interpret this result to indicate that the factor has a high affinity for the *E. coli* core enzyme but a very low affinity for the –180 *rbcL* promoter. A slightly different result was obtained for the *rrn* promoter region. Neither of the two SIG2 proteins allows initiation, even at low levels, at either of the two *E. coli*-like *rrn* promoters (Fig. 2, lanes 5–7 and 12–14). If SIG2-VRG is added to transcription assays that are performed with the *E. coli* holoenzyme, the transcription that initiates at the *rrn*-P1 promoter is specifically diminished (Fig. 2, lanes 15–20). This suggests that of the three promoters tested, SIG2 recognizes only the *rrn*-P1 promoter efficiently. The full-length form of SIG2, SIG2-STE, gave the same result as the truncated protein (not shown). Altogether, this indicates that SIG2 interacts with the *E. coli* core enzyme. This sigma factor can specifically recognize the *rrn*-P1 promoter, but it cannot activate transcription efficiently, not even when present as a full-length native protein.

In the case of SIG1, we analyzed two different truncated forms of the protein (RLV and RQR; see Fig. 1). SIG1-RLV corresponds in length to the truncated form of SIG2 and harbors a supposed 1.2 region at its N terminus. We found that only SIG1-RLV fulfills a sigma-like function in the *in vitro* transcription assay (Fig. 3, lanes 1–16; exposure time was 12 h). SIG1-RLV does not diminish the production of the nonspecific RNA (Fig. 3, asterisk) by the *E. coli* core enzyme, i.e. SIG1 has much less affinity to the *E. coli* core enzyme than...
SIG2. However, activation of transcription at the –180 rbcL promoter by SIG1 is much stronger than that obtained by SIG2 (compare Fig. 2, lanes 1–4, with Fig. 3, lanes 13–16; note that all lanes were exposed for 12 h). SIG1 recognizes only the –180 rbcL promoter (Fig. 3, lanes 13–16, 12 h of exposure; lanes 17–19, 4 h of exposure) but does not recognize either of the two E. coli-like rrr promoters (Fig. 3, lanes 1–4 and 9–12, 12 h of exposure; lanes 20–22, 4 h of exposure).

Addition of SIG3 (GLR) to the in vitro transcription assays did not change the transcription pattern obtained with the E. coli core enzyme (not shown), suggesting that the truncated form of the factor does not interact with any of the three tested promoter regions or with the core RNA polymerase.

Analysis of the Promoter Specificity of Plant Plastid Sigma-like Factors by Complementation of E. coli σ70 Mutants—To study in vivo the function of the three plant sigma-like factors, we used two different E. coli strains harboring thermosensitive rpoD mutations, UQ285 (19) and CAG1 (20). In both cases, we obtained identical results. Fig. 4 shows the results obtained with the strain CAG1. To make the plant factors functionally comparable for its usage in the E. coli system, we fused the 3’ ends of the plant cDNAs, coding for the regions 2-1 up to 4-2, to the 5’ end of the primary E. coli σ70 factor (Fig. 4A). These constructs result in the production of three hybrid sigma factors that are identical to the primary E. coli σ70 factor in their N-terminal part (region 1) but different in the C-terminal parts (regions 2–4). We found that of the three different hybrid factors (Fig. 4A), only one (pMA 55, N′-E. coli σ70-C′-plant SIG1; Fig. 4B) complements the E. coli rpoD mutant strain with the same efficiency as does the homologous E. coli σ70 factor (pMRG1, Ref. 21), i.e. SIG1 represents a primary sigma factor when assayed in E. coli. As a control, we also attempted to complement the two E. coli mutations with the three full-length plant sigma factors. All of these assays were negative, indicating that the plant-specific N-terminal part of SIG1 prevents the complementation. In Fig. 4B, only the result obtained with full-length SIG1 is shown (pMA 40).

SIG3 Might Contain an Inhibition Domain—The low, but obvious, complementation of the two E. coli rpoD mutants by SIG3 was surprising when compared with the complete absence of specific initiation that was observed by in vitro transcription. Therefore, we compared the amino acid sequences of two C-terminal parts that have been used for either in vivo complementation or for in vitro transcription in more detail. Sequence alignment with the prokaryotic sigma factors known thus far shows sequence similarity to the sigma K factor of Bacillus subtilis in a region of about 100 amino acids (Fig. 5). Sigma K is synthesized as a precursor protein with a 20-amino acid pro sequence. In the presence of this sequence, the factor is inactive, and its activation is brought about by proteolytic removal of the pro sequence (22). The two SIG3 constructs that have been used for in vitro transcription and in vivo complementation differ in the presence and absence, respectively, of the region that is similar to the inhibiting pro sequence (see Fig. 5). Therefore, low level of complementation could be related to the absence of this domain.

To test this hypothesis, we produced two additional N-terminal truncated proteins. One of them lacks only the supposed pro peptide (SIG3-ASL), and the other one lacks the entire N-terminal part that has amino acid sequence similarity to Sigma K (SIG3-HTR). We performed reconstitution assays of the E. coli core enzyme with the two different truncated forms of SIG3 using the same templates as already described. SIG3-ASL does not affect in vitro transcription (not shown). However, SIG3-HTR stimulates specific initiation at all three promoters, rrr-P1, rrr-P2, and rbcL (Fig. 6, lanes 4–9; 12 h of exposure). Transcription of the cloned rDNA promoter produces many different transcripts. Therefore, we confirmed the P1- and P2-initiated RNAs by transcribing three different mutated promoter regions. These mutations have been previously constructed in our laboratory to analyze the regulation of plastid rDNA transcription by the transcription factor CDF2 (Ref. 4; Fig. 6, lanes 10–17; 2 days of exposure). As already shown by transcription of these three templates using E. coli holoenzyme (4), the mutation in the –35 region of the P2 promoter diminishes initiation at P2 (Fig. 6, compare lanes 11 and 13). The mutation in the –10 region abolishes initiation at P2 (Fig. 6, compare lanes 11 and 15), and the mutation in the CDF2 binding site abolishes initiation at the P1 promoter (Fig. 6, compare lanes 11 and 17). Competition assays of SIG3-HTR with the E. coli holoenzyme confirm the recognition of all three promoters by SIG3-HTR (Fig. 6, lanes 18–23; 2 h of exposure).

Activation of SIG3 Is Not Regulated by Proteolytic Cleavage of the Hypothetical Inhibition Domain—To analyze whether the activity of SIG3 might be regulated in vivo by proteolytic cleavage of the N-terminal region, we have prepared antibodies against the SIG3-HTR protein. These antibodies specifically recognize SIG3. They do not cross-react with SIG1 or with SIG2 (not shown). Western blot analysis of proteins isolated from Arabidopsis cotyledons revealed only one polypeptide of about 60 kDa, corresponding to the size of the full-length protein (Fig. 7).

DISCUSSION

Three different cDNAs coding for potential plastid-localized sigma70-type transcription initiation factors have recently been cloned and sequenced from A. thaliana (9, 23), and the existence of a multigene family of sigma factors has been suggested for Zea mays (24). Meanwhile, the list of Arabidopsis sigma-
Conservation of C-terminal Domains of $\sigma^{70}$-Type Factors

**A**

| Hybrid sigma factors |
|----------------------|
| **E. coli** N-terminal domain | **plant C-terminal domain** |
| 1-1 | 2-1 |
| 1-2 | 2-2 |
| 1-3 | 2-3 |
| 1-4 | 2-4 |
| 1-5 | 2-5 |

- SIG1: NIR (pMA 55)
- SIG2: NVR (pMA 58)
- SIG3: TRS (pMA 59)

**B**

![Diagram showing hybrid sigma factors](Image)

**Fig. 4. Complementation of *E. coli* rpoD mutants with hybrid sigma factors.** A, schematic representation of the three hybrid sigma factors being composed in the N-terminal half of the *E. coli* principal sigma70 factor and in the C-terminal half of the three *A. thaliana* sigma factors. The N-terminal borders of the *A. thaliana* sigma factors are indicated by a vertical arrow and the three most N-terminal amino acids (NIR, NVR, and TRS). pMA 55, 58, and 59 are the plasmids used for complementation. B, complementation of *E. coli* strain CAG1 harboring the rpoD800(ts) mutation with the plasmids pMA 55, 58, and 55. The plasmid pMRG1 corresponds to the control plasmid harboring the cDNA coding for the *E. coli* principal $\sigma^{70}$ factor, and the plasmid pMA 40 corresponds to the SIG1-KLR control.

**Table 1:**

| Sigma Factor | Sequence | Note |
|--------------|----------|------|
| SIG3 | MTGVKVADQEMRTLYQETEVWIFYCVAAGAGNNNLNLGNNYNGW | N-terminus used for *in vitro* transcription |
| SIGK | HYPFV YPAGLYKEVLFYKDRFDQGE YSKEEYKILAIK | N-terminus used for *in vivo* complementation |

**Fig. 5. Sequence alignment of SIG3 and sigma K.** The pro sequence of sigma K is underlined, and the cleavage site is indicated by a vertical arrow. The 5' ends of the truncated SIG3 proteins are labeled by stars. The 5' ends of the two supplementary SIG3 truncated proteins that have been used to analyze the function of the hypothetical inhibition domain are boxed.

**Experiments and Results:**

1. **Introduction:** The sigma70-like sigma factors have been extended to six (SIG1–SIG6; see accession numbers listed above). Considering the small size of the plastid genome compared with that of bacteria or cyanobacteria, the existence of several sigma-like factors in plastids is quite surprising. It suggests that these factors might play different roles in plastid gene expression.

2. **The sigma70 transcription factor family is well characterized in prokaryotes such as *E. coli* and *B. subtilis*. It is subdivided into two (8) or three (25) groups including primary and alternative sigma factors. In the present study, we analyzed whether the two-group system has been conserved in higher plant plastids and to characterize the function(s) of the three of the six different plant plastid sigma-like factors. Expression and complementation studies (Figs. 1 and 4) of full-length and truncated plant plastid sigma polypeptides in *E. coli*

3. **Fig. 4:**

   - Complementation of *E. coli* rpoD mutants with hybrid sigma factors. A, schematic representation of the three hybrid sigma factors being composed in the N-terminal half of the *E. coli* principal sigma70 factor and in the C-terminal half of the three *A. thaliana* sigma factors. The N-terminal borders of the *A. thaliana* sigma factors are indicated by a vertical arrow and the three most N-terminal amino acids (NIR, NVR, and TRS). pMA 55, 58, and 59 are the plasmids used for complementation. B, complementation of *E. coli* strain CAG1 harboring the rpoD800(ts) mutation with the plasmids pMA 55, 58, and 55. The plasmid pMRG1 corresponds to the control plasmid harboring the cDNA coding for the *E. coli* principal $\sigma^{70}$ factor, and the plasmid pMA 40 corresponds to the SIG1-KLR control.

4. **Fig. 5:**

   - Sequence alignment of SIG3 and sigma K. The pro sequence of sigma K is underlined, and the cleavage site is indicated by a vertical arrow. The 5' ends of the truncated SIG3 proteins are labeled by stars. The 5' ends of the two supplementary SIG3 truncated proteins that have been used to analyze the function of the hypothetical inhibition domain are boxed.

5. **Experiments and Results:** To reveal and analyze specificity in promoter recognition of the three plant sigma-like factors, we used truncated polypeptides that harbor only the conserved regions 2–4. In this way, we avoid differences in promoter-RNA polymerase interactions that may arise from different functions of the different N-terminal sequences of these factors. In addition, it has been shown previously that the intact sigma70 factor of *E. coli* does not bind to DNA but that the cleavage of the N-terminal part of the factor transforms it into a DNA-binding protein (16). Therefore, we supposed that it should be possible to reveal promoter specificity by using truncated sigma factors. We analyzed transcription factor-RNA polymerase-promoter interactions either by reconstitution of the truncated polypeptides with the *E. coli* core enzyme or by competition of the truncated polypeptides with the *E. coli* holoenzyme followed by *in vitro* transcription. These experiments clearly demonstrate differences in the recognition of three selected plastid promoters, *rrn*P1, *rrn*P2, and *rbcL*. SIG3 recognizes all three promoters, i.e., it is the least specific of the three regulatory proteins. SIG2 recognizes specifically the *rrn*P1 promoter, and SIG1 recognizes only the *rbcL* promoter (Figs. 2, 3, and 6). The analysis of several truncated SIG3 proteins shows that SIG3 might contain an inhibition domain that is similar to the pro sequence of *B. subtilis* sigma K. However, we could not detect proteolytic cleavage of SIG3 in *Arabidopsis*. Nevertheless, our experiments suggest a function of this domain in the regulation of SIG3 activity. Instead of proteolytic cleavage, the activity of SIG3 might be regulated by posttranslational conformational changes that modify the accessibility of this putative inhibition domain.

6. **If we compare the three plastid promoter structures with the consensus sequence recognized by the principal sigma70 factor of *E. coli*, we find that the *rbcL* promoter has the highest similarity with the *E. coli* consensus sequence (Fig. 1C). SIG1 recognizes specifically only the *rbcL* promoter, as analyzed by *in vitro* transcription, suggesting that SIG1 is the plant analogue to the primary sigma factor of *E. coli*. This hypothesis is further supported by the *in vivo* analysis of the hybrid sigma factors (Fig. 4). For this analysis, the N-terminal part of the *E. coli* primary sigma70 factor (region 1) was fused to the different C-terminal parts of the three plant sigma factors (regions 2–4).
Region 1 had been shown to be important for open and ternary complex formation (26) and to induce conformational changes into the holoenzyme that are important for correct promoter recognition (27). Parts of regions 2 and 4 are important for the recognition of consensus –10 and –35 DNA sequences (for review, see Ref. 8). Thus, differences in promoter recognition and initiation of these hybrid sigma factors ought to be due to differences in the 3'9 plant-specific part of the constructs. Results show that of the three tested hybrid sigma factors, only SIG1 fully complements the E. coli thermosensitive rpoD mutants.

Our results indicate that the three higher plant plastid sigma 70-like proteins have at least three characteristics in common with their prokaryotic ancestors: 1) they are composed of specific functional domains; 2) they distinguish between different promoter structures; and 3) only one of them (SIG1) is capable of recognizing all essential E. coli promoters that are recognized by the E. coli principal sigma 70 factor in vivo in E. coli. Therefore, we consider SIG1 to be the plant analogue to the primary sigma factor of E. coli. Interestingly, sequence alignment of all six Arabidopsis sigma factors to all five sigma factors localized on the Synechocystis genome (13) shows that SIG1 has the strongest sequence similarity and/or identity to all of the cyanobacterial sigma factors (Table I). Thus, SIG1 is the most prokaryotic-like plant sigma factor. The other plant sigma factors might have evolved in coordination with the transformation of a unicellular organism, the cyanobacterial ancestor, into an integrated part of a multicellular organism, the present-day plastid. The in vivo function of the SIG2-specific rrn-P1 promoter is still unclear (4), and the activity of SIG3 might be regulated by posttranslational modification(s). Therefore, our results suggest specific functions for SIG2 and SIG3 that are related either to plant development and/or changes of environmental conditions. Experiments are in progress in our laboratory to analyze the function of the three plant sigma factors during A. thaliana development using an antisense approach.

**Acknowledgments**—We are grateful to H. Geiselmann for helpful discussions during the preparation of the manuscript, and we thank H. Pesey for photographic work.

---

**Table I**

Amino acid sequence similarities and identities between cyanobacterial and plant σ factors

|   | SIG1 | SIG2 | SIG3 | SIG4 | SIG5 | SIG6 |
|---|------|------|------|------|------|------|
| Syn1 | 59.9 | 58.2 | 52.2 | 57.2 | 52.4 | 57.3 |
| Syn2 | 38.2 | 36.7 | 30.9 | 35.8 | 31.4 | 36.2 |
| Syn3 | 58.8 | 58.5 | 53.4 | 55.0 | 55.2 |
| Syn4 | 32.7 | 33.1 | 27.1 | 32.5 | 29.3 | 33.2 |
| Syn5 | 59 | 58.1 | 49.1 | 54.6 | 51.9 | 56.1 |
| Syn6 | 36.8 | 32.4 | 28.5 | 34 | 30.2 | 36.6 |
| Syn7 | 55.7 | 54.2 | 49.5 | 54.7 | 51.6 | 53.2 |
| Syn8 | 34.5 | 31.6 | 26.9 | 31.7 | 30.1 | 33.7 |
| Syn9 | 65.5 | 55.5 | 49.6 | 60.1 | 53.1 | 53.6 |
| Syn10 | 41.9 | 31.6 | 25.7 | 34.4 | 29.7 | 30.5 |

---

*S. Lerbs-Mache, unpublished results.*
REFERENCES

1. Hess, W. R., and Börner, T. (1999) Int. Rev. Cytol. 190, 1–59
2. Lerbs-Mache, S. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 5509–5513
3. Hedtke, B., Börner, T., and Weihe, A. (1997) Science 277, 809–811
4. Iratni, R., Baeza, L., Andreева, A., Mache, R., and Lerbs-Mache, S. (1994) Genes Dev. 8, 2928–2938
5. Hu, J., and Bogorad, L. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 5509–5513
6. Hedtke, B., Börner, T., and Weihe, A. (1997) Science 277, 809–811
7. Shinozaki, K., Ohme, M., Tanaka, M., Wakasugi, T., Hayashida, N., Matusabayashi, T., Zaita, N., Chunwongse, J., Obokata, J., Yamaguchi-Shinozaki, K., Ohno, C., Torazawa, K., Meng, B. Y., Sugita, M., Deno, H., Kamogashira, T., Yamada, K., Kusuda, J., Takaiwa, F., Kato, A., Tohdoh, N., Shimada, H., and Suguri, M. (1986) EMBO J. 5, 2043–2049
8. Helmann, J. D., and Chamberlin, M. J. (1988) Annu. Rev. Biochem. 57, 839–872
9. Iseno, K., Shimizu, M., Yoshimoto, K., Niwa, Y., Satoh, K., Yokota, A., and Kebayashi, H. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 14948–14953
10. Tiller, K., Eisermann, A., and Link, G. (1991) Eur. J. Biochem. 198, 93–99
11. Tiller, K., and Link, G. (1993) Plant Mol. Biol. 21, 503–513
12. Lerbs, S., Braunigam, E., and Mache, R. (1988) Mol. Gen. Genet. 211, 459–464
13. Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirose, M., Sugura, M., Sasamoto, S., Kimura, T., Hosoe, T., Matsuura, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., Tanaka, K., Itoh, S., Seki, M., Katagiri, T., Nakamura, M., Mochizuki, N., Nagatani, A., Shinozaki, K., Tanaka, K., Takahashi, H. (1999) Plant Physiol. 121, 503–513
14. Isaksson, L. A., Skold, S. E., Skjoldebrand, J., and Takata, R. (1977) Mol. Gen. Genet. 156, 223–227
15. Liebke, H., Gross, C., Walter, W., and Burgess, R. (1980) Mol. Gen. Genet. 177, 277–282
16. Gruber, T. M., and Bryant, D. A. (1997) J. Bacteriol. 179, 1734–1747
17. Wilson, C., and Dombroski, A. J. (1997) J. Mol. Biol. 267, 60–74
18. Wilson Bowers, C., and Dombroski, A. J. (1999) EMBO J. 18, 709–716
Evolutionary Conservation of C-terminal Domains of Primary Sigma$^{70}$-type Transcription Factors between Plants and Bacteria
Mohamed-Ali Hakimi, Isabelle Privat, Jean-Gabriel Valay and Silva Lerbs-Mache

*J. Biol. Chem.* 2000, 275:9215-9221.
doi: 10.1074/jbc.275.13.9215

Access the most updated version of this article at [http://www.jbc.org/content/275/13/9215](http://www.jbc.org/content/275/13/9215)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 27 references, 10 of which can be accessed free at [http://www.jbc.org/content/275/13/9215.full.html#ref-list-1](http://www.jbc.org/content/275/13/9215.full.html#ref-list-1)