JOURNAL OF BACTERIOLOGY, Dec. 1995, p. 7026–7032 Vol. 177, No. 24
0021-9193/95/$04.00+0

Characterization of Transposon Tn*5469* from the Cyanobacterium *Fremyella diplosiphon*

KATHERINE KAHN AND MICHAEL R. SCHAEFER*

Division of Molecular Biology and Biochemistry, School of Biological Sciences, University of Missouri-Kansas City, Kansas City, Missouri 64110

Received 7 August 1995/Accepted 29 September 1995

A transposon, designated Tn*5469*, was isolated from mutant strain FdR1 of the filamentous cyanobacterium *Fremyella diplosiphon* following its insertion into the *rcaC* gene. Tn*5469* is a 4,904-bp noncomposite transposon with 25-bp near-perfect terminal inverted repeats and has three tandemly arranged, slightly overlapping potential open reading frames (ORFs) encoding proteins of 104.6 kDa (909 residues), 42.5 kDa (375 residues), and 31.9 kDa (272 residues). Insertion of Tn*5469* into the *rcaC* gene in strain FdR1 generated a duplicate 5-bp target sequence. On the basis of amino acid sequence identities, the largest ORF, designated *tnpA*, is predicted to encode a composite transposase protein. A 230-residue domain near the amino terminus of the TnpA protein has 15.4% amino acid sequence identity with a corresponding domain for the putative transposase encoded by *Lactococcus lactis* insertion sequence S1 (IS*S1*). In addition, the sequence for the carboxyl-terminal 600 residues of the TnpA protein is 20.0% identical to that for the TniA transposase encoded by Tn*5090* on *Klebsiella aerogenes* plasmid R751. The TnpA and TniA proteins contain the D,D(35)E motif characteristic of a recently defined superfamily consisting of bacterial transposases and integrase proteins of eukaryotic retroelements and retrotransposons. The two remaining ORFs on Tn*5469* encode proteins of unknown function. Southern blot analysis showed that wild-type *F. diplosiphon* harbors five genomic copies of Tn*5469*. In comparison, mutant strain FdR1 harbors an extra genomic copy of Tn*5469* which was localized to the inactivated *rcaC* gene. Among five morphologically distinct cyanobacterial strains examined, none was found to contain genomic sequences homologous to Tn*5469*.

The cyanobacteria consist of a diverse and widely distributed group of prokaryotes that are characterized by their ability to perform oxygenic photosynthesis. Light harvesting in the cyanobacteria is accomplished by macromolecular antenna complexes, called phycobilisomes, which are composed of brightly pigmented phycobiliproteins and linker polypeptides (recently reviewed in reference 16). The filamentous cyanobacterium *Fremyella diplosiphon* belongs to a class of cyanobacteria that responds to changes in light quality by altering the phycobiliprotein composition of the phycobilisome (6, 15, 31). This acclimation response, termed complementary chromatic adaptation, is mediated primarily through differential expression of genes encoding two phycobiliproteins, the blue-pigmented inducible phycocyanin (PC_2) and the red-pigmented phycoerythrin (PE) (15). Red light promotes synthesis of PC_2 and suppresses synthesis of PE; cells grown in red light are pigmented blue-green. Alternatively, green light promotes synthesis of PE and suppresses that of PC_2; cells grown in green light are pigmented red.

Many *F. diplosiphon* mutants that exhibit aberrant pigmentation have been isolated. Because phycobiliproteins represent a significant fraction of total cellular protein, these mutants can be identified visually on plates. Spontaneous pigment mutants arise at a low frequency, probably in response to adverse environmental conditions such as nutrient limitation or desiccation which can occur during culture maintenance. Exposure of cells to mutagenic agents or electroporation increases the frequency at which pigment mutants appear. Genetic and biochemical characterization of several different classes of phycobiliprotein regulatory mutants (those distinguishable from mutants with lesions in genes encoding phycobilisome structural proteins) has been extremely important in defining the photosensory and signaling mechanisms involved in complementary chromatic adaptation (7, 16). In addition, recent complementation of several regulatory mutants has revealed that the complementary chromatic adaptation signaling pathway includes complex forms of bacterial two-component regulatory proteins (9, 19, 28).

In many prokaryotes, transposable genetic elements are agents of spontaneous mutations. The simplest forms are insertion sequence (IS) elements (reviewed in reference 14). IS elements range in size from 0.8 to 2.5 kbp and are found in the genomes of many different bacteria in multiplicities between a few to a few hundred per genome. Most IS elements share a common structure of one or a few genes required for transposition which are flanked by terminal inverted repeats (IRs). Almost all IS elements generate directly repeated duplications of a target DNA sequence upon insertion. The more complex forms of transposable elements include transposons which range in size from 2.5 to 60 kbp (5). Although they share many structural and functional features, transposons differ from IS elements in that, in addition to transposition functions, each encodes one or more functions that confer a phenotype on the bacterial host, such as resistance to a specific antibiotic. Often, transposition of IS elements or transposons causes mutations by directly inserting into a gene or operon. In other cases, transposition of an element mobilizes flanking genes or promotes a DNA rearrangement in its vicinity, each of which can produce a mutation.

To date, the reported distribution of transposable elements among the diverse cyanobacteria is limited to two filamentous genera. Fourteen cyanobacterial IS elements have been isolated, five of which have been sequenced, namely, IS*701* and

* Corresponding author. Mailing address: University of Missouri-Kansas City, School of Biological Sciences, 5100 Rockhill Rd., Kansas City, MO 64110. Phone: (816) 235-2573. Fax: (816) 235-5158. Electronic mail address: mschaefer@cctr.umkc.edu.

TABLE 1. Strains and plasmids used

Strain or plasmid	Description	Reference(s) or source
Cyanobacterial strains		
F. diplosiphon		
UTEX 481	Wild-type; filamentous; also referred to as *Calothrix* sp. strain PCC 7601	Laboratory collection
Fd33	Short-filament mutant of UTEX 481; exhibits wild-type chromatic adaptation	10
FdR1	Red pigment mutant; isolated following electroporation of Fd33; *rcaC*::Tn*5469*	7, 9
Anabaena sp.		
PCC 7120	Wild-type; filamentous	Laboratory collection
ATCC 29413	Wild-type; filamentous	Laboratory collection
Nostoc sp. strain PCC 8009	Wild-type; filamentous	J. Meeks, University of California
Synechococcus sp. strain PCC 7942	Wild-type; unicellular	Laboratory collection
Synechocystis sp. strain PCC 6803	Wild-type; unicellular	W. Vermaas, University of Arizona
E. coli strains		
DH5α	F^- φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *deoR recA1 endA1 hsdR17*(r_K^- m_K^+) *supE44* λ^- *thi-1 gyrA96 relA1*	Bethesda Research Labs
DH5αMCR	F^- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *deoR recA1 endA1 supE44* λ^- *thi-1 gyrA96 relA1*	Bethesda Research Labs
Plasmids		
pGEM3zf(+)	Ap^r cloning and sequencing vector	Promega
pGCMS37	3.5-kbp *Cla*I fragment containing *rcaC* gene from *F. diplosiphon* Fd33; vector pUC118	9
pUMC227	5.7-kbp *Eco*RI fragment containing intact Tn*5469* and flanking *rcaC* sequences from *F. diplosiphon* FdR1; vector pGEM3zf(+)	This study
pJC1.6	1.6-kbp *Pvu*II-*Sal*I fragment containing *glnA* gene from *Synechococcus* sp. strain PCC 7942; vector pBluescript KS(+)	J. Curry and S. Robinson (unpublished data)

IS*702* from *Calothrix* sp. strain PCC 7601 (21), IS*891* from *Anabaena* sp. strain M-131 (4), and IS*892* (8) and IS*895* (1) from *Anabaena* sp. strain PCC 7120. IS*701*, IS*702*, and IS*892* are typical IS elements characterized by terminal IRs and duplicate target sequences, whereas IS*891* and IS*895* lack one or both of these features. Unlike the IS elements isolated from the *Anabaena* sp., which were first visualized as insertions on rescued recombinant plasmids, the IS elements isolated from *Calothrix* sp. strain PCC 7601 were first identified in spontaneous pigment mutants as an insertion at or near a known phycobiliprotein gene cluster on the chromosome (21). To our knowledge, no transposons have been reported for any cyanobacterial strain.

Recent complementation and genetic characterization of the *F. diplosiphon* pigment mutant strain FdR1 suggested that the activity of a transposon, as opposed to that of an IS element, was responsible for the FdR1 pigmentation phenotype. Strain FdR1 is a member of the red (FdR) class of phycobilisome regulatory mutants (7). FdR mutants are locked in the green-light regulatory mode and are characterized by constitutive synthesis of PE and no synthesis of PC_2 under any conditions of illumination (7). Complementation experiments demonstrated that the pigmentation phenotype of strain FdR1 was due to inactivation of the *rcaC* gene, which encodes a response regulator protein involved in complementary chromatic adaptation (9). Mapping experiments indicated that the *rcaC* gene in this mutant harbors a 4.9-kbp DNA insert, suggesting that the gene was insertionally inactivated following mobilization of an endogenous transposon. To examine this possibility, the interrupted region of the *rcaC* locus was cloned intact from the genome of strain FdR1 and the nucleotide sequence for the termini of the insert was determined. DNA sequence analysis revealed that the 4.9-kbp insert was a transposon, which was designated Tn*5469*. Here, we present a detailed characterization of Tn*5469* from *F. diplosiphon.*

MATERIALS AND METHODS

Strains and growth conditions. The strains and plasmids used in this study are listed in Table 1. Strain Fd33, which exhibits wild-type complementary chromatic adaptation, is a short-filament mutant of *F. diplosiphon* UTEX 481 (also referred to as *Calothrix* sp. strain PCC 7601) (10). Red mutant strain FdR1 was isolated following electroporation of Fd33 cells and has been phenotypically (7) and genotypically (9) characterized. All other cyanobacterial strains were from the Pasteur Culture Collection or the American Type Culture Collection and are referred to by their genus name followed by the collection number. Cyanobacteria were grown in liquid or on solid BG-11 medium (2) as described previously (7).

Escherichia coli DH5α was purchased from Bethesda Research Laboratories (Gaithersburg, Md.) and used as the host for most plasmids. *E. coli* DH5αMCR (Bethesda Research Laboratories) was employed for direct cloning of cyanobacterial genomic DNA. *E. coli* strains were propagated in liquid or on solid Luria-Bertani medium with antibiotics at standard concentrations (26).

DNA methods. DNA restriction endonucleases and modifying enzymes were purchased from Promega (Madison, Wis.). [α-^{32}P]dCTP was purchased from DuPont/NEN (Boston, Mass.), and deoxyadenosine 5′-[α-^{35}S]thio-triphosphate was purchased from Amersham (Arlington Heights, Ill.). DNA manipulations, including restriction digests, agarose gel electrophoresis, ligations, transformation of *E. coli*, and plasmid minipreparations, were performed essentially as described by Sambrook et al. (26). For Southern blot hybridization analysis, the digested DNA was transferred to a charged nylon membrane (Magnagraph; Micron Separations, Westboro, Mass.) by the method of Reed and Mann (23). Probes for *rcaC* (3.5-kbp *Cla*I fragment from pGCMS37) and Tn*5469* (3.8-kbp *Eco*RI-*Xba*I fragment from pUMC227) from *F. diplosiphon* and *glnA* (1.6-kbp *Xba*I-*Sal*I fragment from pJC1.6) from *Synechococcus* sp. strain PCC 7942 were generated from gel-purified DNA fragments by random-primer labeling with a kit from Promega. Hybridizations were performed at 62°C as described by Sambrook et al. (26). Double-stranded DNA sequencing templates were isolated and purified with a kit from Promega.

Isolation of cyanobacterial genomic DNA. For all strains, cells from a 50-ml culture were harvested by centrifugation at 16,000 × *g* for 10 min at 4°C, washed by resuspension in 30 ml of ice-cold lysis buffer (50 mM glucose, 25 mM Tris-HCl [pH 8.0], 10 mM EDTA), and reharvested. The cell pellet was resuspended in 2 ml of lysis buffer, and the suspension was frozen in liquid nitrogen. The frozen suspension was thawed at room temperature, adjusted to a final volume of 6 ml by adding lysis buffer supplemented with lysozyme at 7.5 mg · ml^{-1}, and incubated at 37°C for 15 min. The cell lysate was frozen in liquid nitrogen, thawed at room temperature, and brought to 1% (wt/vol) sodium dodecyl sulfate, and the DNA was extracted as described by Chiang et al. (9).

Cloning of Tn*5469*. Total DNA from mutant strain FdR1 was digested with

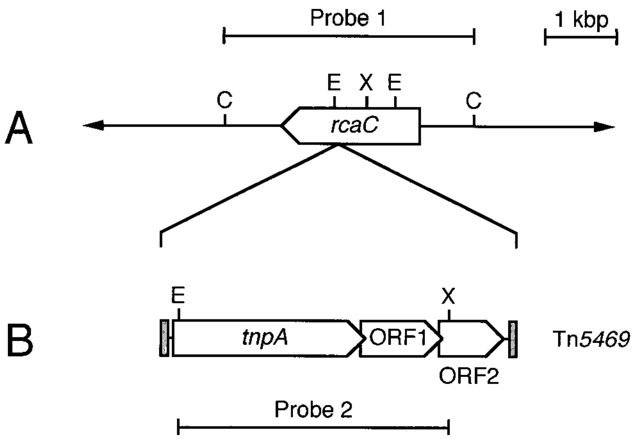

FIG. 1. Tn*5469* inactivation of the *rcaC* gene in *F. diplosiphon* FdR1. (A) Localization of Tn*5469* in the *rcaC* gene. The open box indicates the orientation of the *rcaC* ORF. The brackets show the site of Tn*5469* insertion. (B) Physical map of Tn*5469*. The open boxes indicate the orientation of three ORFs as determined by sequence analysis. The shaded vertical boxes indicate terminal IR sequences. Flanking and internal restriction sites are shown for enzymes used in cloning and mapping experiments. Labeled horizontal bars above or below maps identify regions that correspond to probes used in Southern hybridization analysis. C, *Cla*I; E, *Eco*RI; X, *Xba*I.

*Eco*RI, and the digestion products were separated by agarose gel electrophoresis. To isolate fragments in the 4- to 6-kbp size range, the corresponding region of the gel was excised and the DNA was recovered by electroelution (26). The gel-purified 4- to 6-kbp DNA fragments were ligated into *Eco*RI-digested pGEM3zf(+), and the ligation products were used to transform *E. coli* DH5α-MCR. Approximately 450 transformants were screened for plasmids containing *rcaC* sequences by Southern blot hybridization against a probe for *rcaC* (see Fig. 1A, probe 1) by a colony screening method (26). Two plasmids that hybridized to the *rcaC* probe contained the predicted 5.7-kbp insert, and one, designated pUMC227, was chosen for further analysis.

DNA sequence analysis. Double-stranded sequencing of the pUMC227 insert was performed by the dideoxynucleotide chain termination method (27) with Sequenase version 2.0 modified T7 DNA polymerase purchased from United States Biochemical (Cleveland, Ohio). Sequencing reactions were primed with M13 universal primers or with oligonucleotides synthesized on an Applied Biosystems (Foster City, Calif.) model 381A oligonucleotide synthesizer. DNA and protein sequences were analyzed and compared with sequences in the GenBank database by use of the BLAST (3) or MacVector (Eastman Kodak, Rochester, N.Y.) sequence analysis program.

Nucleotide sequence accession number. The complete DNA sequence of Tn*5469* has been deposited in the GenBank database under accession number U33002.

RESULTS

Cloning of Tn*5469* from *F. diplosiphon* red mutant strain FdR1. The phenotype of pigment mutant strain FdR1 is due to inactivation of the *rcaC* gene, which encodes a response regulator protein involved in chromatic adaptation (9). In strain FdR1, the *rcaC* gene harbors a 4.9-kbp insertion localized to the 432-bp region between the unique *Xba*I site and a downstream *Eco*RI site (Fig. 1A). To investigate whether the 4.9-kbp insertion represented an uncharacterized transposable element for *F. diplosiphon*, the interrupted region of the *rcaC* locus was cloned intact from the genome of strain FdR1 for analysis. Earlier mapping experiments suggested that the 4.9-kbp insert could be cloned intact with flanking *rcaC* sequences on a 5.7-kbp *Eco*RI fragment of FdR1 genomic DNA. Gel-purified 4- to 6-kbp *Eco*RI fragments of FdR1 genomic DNA were ligated into vector pGEM3zf(+), and the ligation products were used to transform *E. coli* DH5αMCR. Approximately 450 transformants were screened by colony hybridization for plasmids containing *rcaC* sequences. Two transformants were identified, both of which contained apparently identical plasmids carrying the predicted 5.7-kbp genomic insert. One of these plasmids, designated pUMC227, was the source of DNA for subsequent sequencing and genetic analysis. DNA sequence analysis confirmed that the pUMC227 clone contained flanking *rcaC* sequences and revealed that the 4.9-kbp insert exhibited structural features characteristic of bacterial transposons. On the basis of the initial data from DNA sequencing and Southern hybridization analysis, the presumed transposable element was designated Tn*5469* and further analyzed.

Transposition of Tn*5469*. Transposition of Tn*5469* into the *rcaC* gene of strain FdR1 was examined by Southern blot analysis. Total DNA from strains Fd33 and FdR1 was digested with *Xba*I and *Sal*I and hybridized to a probe for *rcaC* (Fig. 1A, probe 1) or Tn*5469* (Fig. 1B, probe 2). Insertion of Tn*5469* into the *rcaC* gene alters the restriction fragment pattern obtained with the *rcaC* probe. The 15-kbp fragment for strain Fd33 was detected as two fragments of 19 and 1.2 kbp in strain FdR1 (Fig. 2A, compare lanes 1 and 2) as a result of the introduction of the *Xba*I site in Tn*5469* (Fig. 1). Hybridization with the Tn*5469* probe revealed that strain Fd33 harbors multiple genomic copies of the transposable element (Fig. 2B, lane 1). Similar analysis of the same DNA digested with different restriction enzymes (data not shown) indicates a total of five genomic copies of Tn*5469* for strain Fd33 (the 8.5-kbp fragment identified in Fig. 2B represents a doublet). In comparison with strain Fd33, strain FdR1 harbors six genomic copies of Tn*5469* (Fig. 2B, compare lanes 1 and 2). Hybridization of the FdR1 genomic DNA with the Tn*5469* probe identified a 19-kbp fragment in addition to a restriction fragment pattern identical to that obtained for strain Fd33. This 19-kbp fragment was also detected with the *rcaC* probe (Fig. 2A, lane 2). These data indicate that Tn*5469* is mobile in *F. diplosiphon.*

General features of Tn*5469*. The complete nucleotide se-

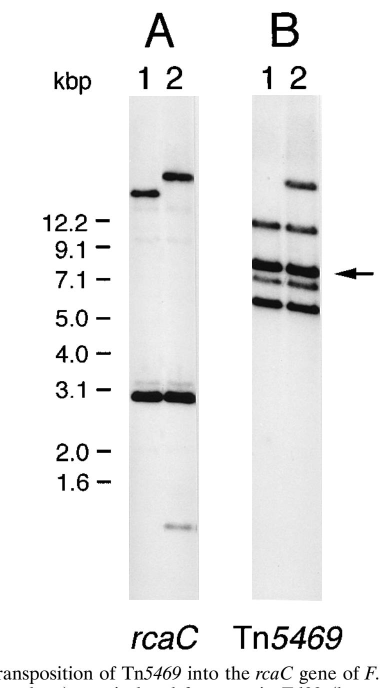

FIG. 2. Transposition of Tn*5469* into the *rcaC* gene of *F. diplosiphon*. Total DNA (5 μg per lane) was isolated from strain Fd33 (lane 1) and red mutant strain FdR1 (lane 2), digested with *Sal*I and *Xba*I, and subjected to Southern blot hybridization with a probe for *rcaC* (Fig. 1A, probe 1) or Tn*5469* (Fig. 1B, probe 2). (A) Southern blot hybridized with the *rcaC* probe; (B) Southern blot identical to that shown in panel A hybridized with the Tn*5469* probe. The arrow indicates the location of a doublet.

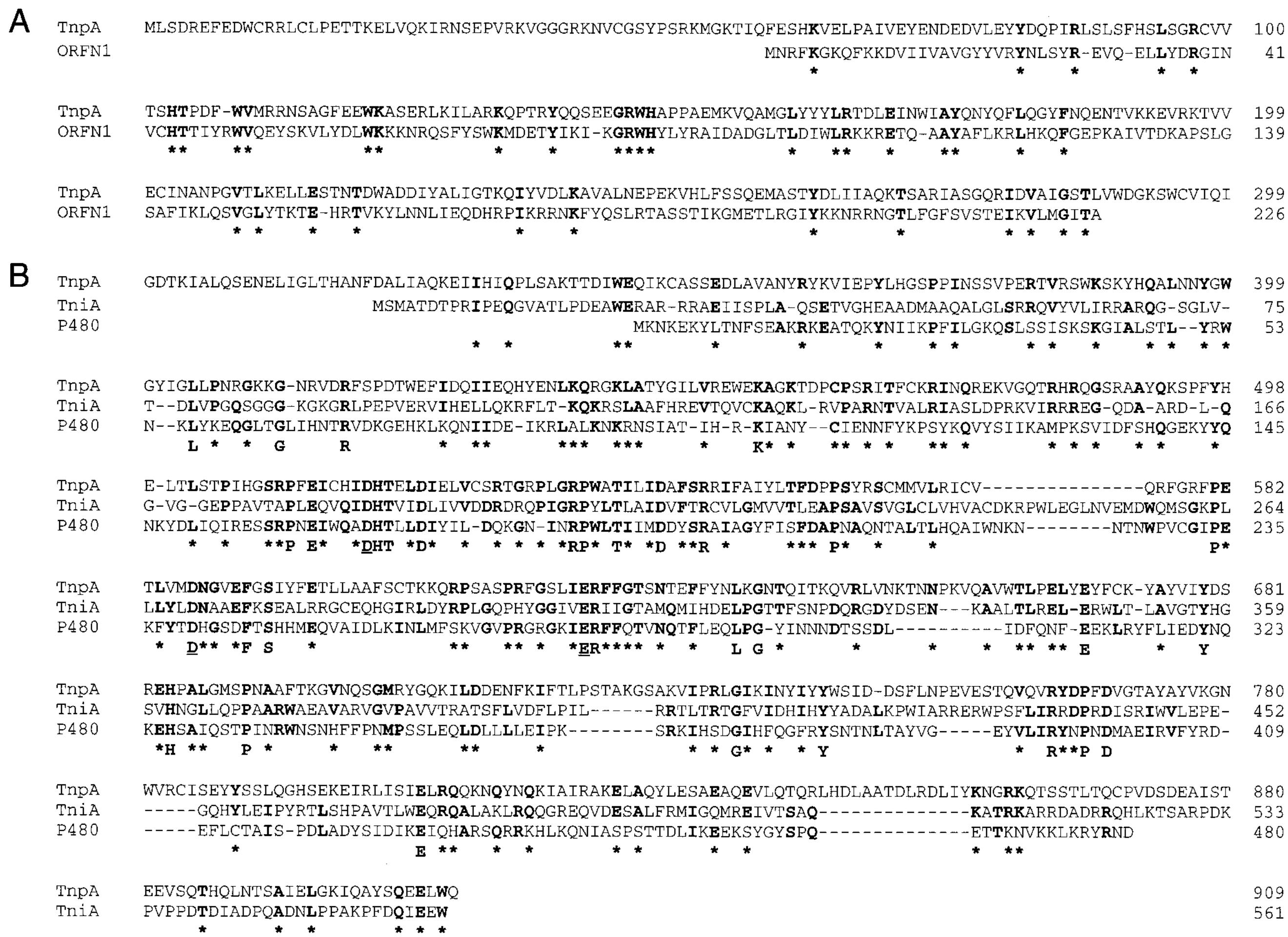

FIG. 3. Alignment of the TnpA amino acid sequence with amino acid sequences for several bacterial transposases. (A) Alignment of residues 1 to 299 with the sequence of ORFN1 transposase encoded by IS*S1* (17). (B) Alignment of residues 300 to 909 with the sequences of TniA from Tn*5090* (22) and P480 from Tn*552* (25). The numbering of the amino acids for the individual proteins is shown to the right of the sequences. Residues identical to at least two proteins are indicated by boldface letters. Amino acid identities between TnpA and an aligned sequence are marked with an asterisk below the sequences. Boldface letters below the sequences identify residues identical to TnpA, TniA, and P480. The conserved aspartate (D) and glutamate (E) residues constituting the D,D(35)E motif of the integrase and transposase superfamily (12, 20) are underlined.

quence of Tn*5469* has been deposited in the GenBank database (accession no. U33002). The 4,904-bp transposon contains 25-bp near-perfect (22 out of 25 bp) terminal IRs that do not share significant identity with corresponding sequences of known IS elements or transposons. On the basis of the orientation of several putative genes (see below), the IRs were designated IR_L for the left end and IR_R for the right end of the element. In the pUMC227 clone, Tn*5469* is flanked by the 5-bp duplicate target sequence 5′-GACAA-3′. The DNA sequence analysis also revealed an internal *Eco*RI site at nucleotide position 258 in Tn*5469* which repeatedly escapes cleavage in *Eco*RI digestion of FdR1 genomic DNA but is cleaved in *Eco*RI digestion of pUMC227 isolated from *E. coli*. The mechanism by which the internal *Eco*RI site escapes cleavage in digestion of FdR1 genomic DNA is unknown; however, it facilitated isolation of the intact Tn*5469* element from the strain.

The nucleotide sequence of Tn*5469* predicts three open reading frames (ORFs) arranged in tandem on the element (Fig. 1B), assuming that each initiates at the first methionine codon in the corresponding reading frame. Like many characterized cyanobacterial genes, two of the three ORFs on Tn*5469* are not preceded by an obvious ribosome binding sequence. The leftmost and largest ORF (nucleotide positions 175 to 2902), which is preceded by an *E. coli*-like promoter, predicts a 909-residue protein with a molecular mass of 104.7 kDa and a pI of 8.8. A BLAST (tblastn) search of the GenBank database (release 15.0) indicated that the 909-residue protein shares domains of significant amino acid sequence identity with several different bacterial transposases. On the basis of these similarities, the ORF was designated *tnpA*.

The protein encoded by the *tnpA* gene appears to be a composite of two different transposase forms. An alignment of the TnpA amino acid sequence with the sequence for several transposases most similar to TnpA is shown in Fig. 3. A 230-residue domain near the amino terminus of TnpA has 15.4% amino acid sequence identity with the putative transposase, designated ORFN1, encoded by IS*S1* from *Lactococcus lactis* (17). In addition, the carboxyl-terminal two-thirds of the TnpA protein (approximately 600 residues) has 20.0% amino acid sequence identity with the TniA transposase encoded by Tn*5090* on the *Klebsiella aerogenes* plasmid R751 (22) and 18.1% amino acid sequence identity with the P480 transposase encoded by Tn*552* from *Staphylococcus aureus* (25). The TnpA, TniA, and P480 transposases possess the D,D(35)E motif characteristic of a recently defined protein superfamily composed of eukaryotic retroviral and retrotransposon integrase proteins and bacterial transposases (12, 20). Another distinguishing feature of this class of mobile genetic elements is the dinucleotide 5′-TG which always occurs at the termini; this feature is shared by Tn*5469*.

The central ORF (designated ORF1; nucleotide positions 2895 to 4020) on Tn*5469* predicts a protein of 375 residues with a molecular mass of 42.5 kDa and a pI of 9.2. The DNA sequence indicates a seven-nucleotide overlap between the 3′ end of the *tnpA* coding region and the 5′ end of ORF1. The first methionine start codon for ORF1 is preceded by a characteristic ribosome binding site, and a potential −35 promoter sequence is located upstream of the ORF1 coding region. A BLAST comparison of the ORF1 nucleotide and polypeptide sequences with sequences in the GenBank database revealed no significant matches, with the exception that a 50-residue domain of the ORF1 protein is 38% identical in amino acid sequence to a region of the ATP-binding erythromycin resistance protein encoded by the *Staphylococcus epidermidis msrA* gene (24). Outside of this domain, no other similarities between the two proteins were observed, including the ATP-binding domain of MsrA. To examine whether Tn*5469* confers erythromycin resistance to *F. diplosiphon*, wild-type cells were cultured in liquid or on solid medium supplemented with different amounts of the antibiotic (data not shown). This experiment revealed that *F. diplosiphon* was sensitive to even very low levels of the antibiotic, arguing against a role for the ORF1 protein in erythromycin resistance. A related analysis was performed with *E. coli* DH5α; in liquid culture supplemented with up to 200 μg of the antibiotic ml^{-1}, the rate of growth of strain DH5α/pUMC227 was indistinguishable from that of control strain DH5α/pGEM3zf(+) (data not shown). Whether Tn*5469* confers resistance to one or more known antibiotics remains to be determined.

The rightmost and smallest ORF (ORF2; nucleotide positions 4012 to 4828) on Tn*5469* predicts a 272-residue protein with a molecular mass of 31.92 kDa and a pI of 5.6. In a manner similar to that of the *tnpA*-ORF1 junction, the ORF1-ORF2 junction on Tn*5469* exhibits an eight-nucleotide overlap, assuming that ORF2 initiates with the methionine codon at nucleotide position 4017. A potential −35 promoter sequence is located upstream of the ORF2 coding region. A BLAST comparison of the ORF2 nucleotide and polypeptide sequences with sequences in the GenBank database revealed no significant matches.

Distribution of Tn*5469*. The distribution of Tn*5469* among several morphologically distinct strains of cyanobacteria was examined by Southern blot analysis. Total DNA from *F. diplosiphon* UTEX 481 (the Fd33 parental strain), *F. diplosiphon* Fd33, *Synechocystis* sp. strain PCC 6803, *Anabaena* sp. strain PCC 7120, *Synechococcus* sp. strain PCC 7942, *Nostoc* sp. strain PCC 8009, and *Anabaena* sp. strain ATCC 29413 was digested with *Xba*I and *Sal*I and hybridized to the Tn*5469* probe (Fig. 4A). As a control, a similar blot was hybridized to a probe for the *glnA* gene from *Synechococcus* sp. strain PCC 7942 (Fig. 4B). Both Southern blots presented in Fig. 4 were overexposed intentionally for this analysis. Among the strains examined, only DNA from *F. diplosiphon* hybridized to the Tn*5469* probe. The parental strain UTEX 481 and short-filament mutant strain Fd33 exhibited identical restriction fragment patterns (compare lanes 1 and 2). In a similar analysis of wild-type *F. diplosiphon* strains from different stock collections, the Southern blot profile for Tn*5469* was invariant (data not shown). On the control blot, the *glnA* probe hybridized with different signal intensities to DNA from all of the cyanobacterial strains. To decrease the hybridization signal intensity from the *Synechococcus* sp. strain PCC 7492 *glnA* probe, only 2.5 μg of total DNA from that strain was used for the Southern blot shown in Fig. 4B (lane 5).

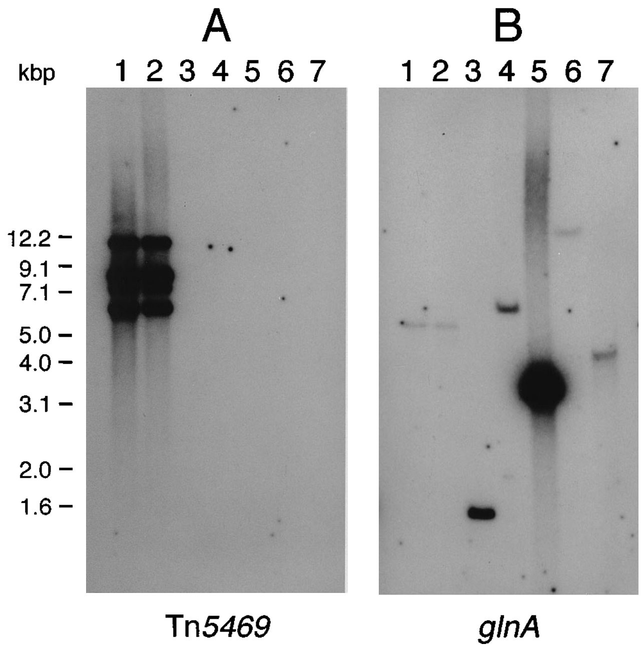

FIG. 4. Distribution of Tn*5469* among morphologically distinct cyanobacterial genera. Total DNA (5 μg per lane) was digested with *Sal*I and *Xba*I and subjected to Southern blot hybridization with a probe for Tn*5469* or *glnA* from *Synechococcus* sp. strain PCC 7942. (A) Southern blot hybridized with the Tn*5469* probe; (B) Southern blot identical to that shown in panel A hybridized with the *glnA* probe. Both of the Southern blot autoradiograms were intentionally overexposed for this analysis. Lane 5 in panel B contained 2.5 μg of DNA to decrease signal intensity from the homologous *glnA* probe. DNA from the following strains was analyzed: *F. diplosiphon* UTEX 481 (lane 1), *F. diplosiphon* Fd33 (lane 2), *Synechocystis* sp. strain PCC 6803 (lane 3), *Anabaena* sp. strain PCC 7120 (lane 4), *Synechococcus* sp. strain PCC 7942 (lane 5), *Nostoc* sp. strain PCC 8009 (lane 6), and *Anabaena* sp. strain ATCC 29413 (lane 7).

DISCUSSION

We have isolated a transposon, designated Tn*5469*, endogenous to the filamentous cyanobacterium *F. diplosiphon*. The transposon was first identified in pigment mutant strain FdR1 as a 4.9-kbp DNA insert responsible for inactivation of the *rcaC* gene, which encodes a response regulator protein involved in complementary chromatic adaptation (9). In the absence of RcaC regulator activity, cells of strain FdR1 are locked in the green-light regulatory mode and remain pigmented red as a result of constitutive synthesis of PE. In comparison with the wild-type strain, FdR1 harbors an extra genomic copy of Tn*5469* in an otherwise indistinguishable genetic background. For this study, this extra copy of Tn*5469* was cloned intact from the inactivated *rcaC* gene of strain FdR1.

The general structure of Tn*5469* places the element in the noncomposite transposon class. Tn*5469* is 4,904 bp in length, with 25-bp terminal IR sequences that differ in three positions. Upon transposition of Tn*5469* into the *rcaC* gene in strain FdR1, the 5-bp target sequence 5′-GACAA-3′ was duplicated as a direct repeat which flanks the element. Tn*5469* contains three tandemly arranged, slightly overlapping ORFs. The leftmost and largest ORF, designated *tnpA*, is preceded by an *E.*

coli-like promoter and encodes a putative transposase. The central and rightmost ORFs, designated ORF1 and ORF2, respectively, encode proteins of unknown function. Although both ORF1 and ORF2 are preceded by a −35 promoter sequence, their overlapping structures suggest that *tnpA*, ORF1, and ORF2 may be cotranscribed, perhaps from the *E. coli*-like promoter upstream of the *tnpA* gene.

The putative transposase encoded by the *tnpA* gene is an unusual protein composite of two complete transposase forms. The amino-terminal one-third of the TnpA protein resembles the putative ORFN1 transposase encoded by the widely distributed lactococcal IS element IS*S1* (17, 29). The IS*S1* transposase has significant sequence identity with other transposases encoded by IS elements found in both gram-negative and gram-positive bacteria (29), suggesting a common ancestral gene for this transposase form. In addition to the amino-terminal transposase equivalent, the carboxyl-terminal two-thirds of the TnpA protein is similar to transposases that belong to a recently defined superfamily of transposition proteins. Members of this protein superfamily include bacterial transposases and eukaryotic retroviral and retrotransposon integrase proteins, all of which are characterized by a so-called constellation of highly conserved amino acids which includes the invariant D,D(35)E motif (20). Amino acid substitution experiments demonstrated that for the integrase proteins of the Rous sarcoma virus and human immunodeficiency virus, the invariant aspartate and glutamate residues are critical for processing (cleavage) of the replicated viral genome and insertion of the viral DNA into the host genome (20). A similar function for a bacterial transposase with the D,D(35)E motif remains to be demonstrated.

The relationship between Tn*5469* and other prokaryotic and eukaryotic mobile genetic elements encoding transposases characterized by the D,D(35)E motif extends to the DNA level. A distinguishing feature of these elements is that they all terminate with the dinucleotide 5′-TG (11, 12). For the retroviruses and retrotransposons, the terminal 5′-TG is generated by the removal of one or two terminal nucleotides from the proviral DNA prior to its insertion into the host genome (11, 32). This processing reaction as well as host DNA cleavage and joining of the virus and host DNA is carried out by the multifunctional integrase protein (20). On the basis of the structural and functional similarities to the retroviral integration mechanism, it is likely that in the related bacterial forms, the terminal 5′-TG dinucleotide serves a specific role in transposase-mediated site-specific cleavage and strand transfer of the element to the target DNA during transposition.

On the basis of significant amino acid sequence identities, the transposase encoded by Tn*5469* is most similar to the bacterial transposases encoded by Tn*5090* (22), Tn*552* (25), and Tn*7* (13), all of which contain the D,D(35)E motif described above. However, despite the unifying transposase, the genetic structure of Tn*5469* is markedly different from the other transposons. In terms of genetic organization and complexity, Tn*5090* is most similar to Tn*7* (22). Both of these transposons encode multiple proteins with defined roles in transposition, and both are characterized by an integrase system that functions in acquisition of antibiotic resistance cassettes. Tn*552* also encodes multiple and defined transposition proteins; however, this element is more structurally related to members of the Tn*3* family of transposons (25). In comparison with Tn*5090*, Tn*552*, and Tn*7*, the genomic structure of Tn*5469* is remarkably simple; Tn*5469* contains three genes, only one of which encodes a defined transposition protein. Collectively, these structural differences suggest that the individual transposons arose independently but evolved from a common ancestor.

It is difficult to categorize Tn*5469* in the absence of identifiable roles for the proteins encoded by ORF1 and ORF2. Presumably, one of these ORFs encodes a protein that confers a specific phenotype on the host, a feature that distinguishes transposons from IS elements. On the basis of the small domain of sequence identity with the ATP-binding erythromycin resistance protein encoded by the *S. epidermidis msrA* gene (24), the protein encoded by ORF1 would seem a likely candidate. However, preliminary experiments with *F. diplosiphon* and *E. coli* indicated that Tn*5469* does not confer resistance to erythromycin. It is possible that ORF1 (or ORF2) encodes a form of multidrug resistance protein or transport protein. In the absence of corroborating data, this possibility is purely speculative.

ORF1 or ORF2 might encode a protein required for transposition of Tn*5469*. In addition to a transposase gene, many noncomposite transposons encode one or more proteins required for their mobilization. The function of a transposition protein can often be defined by identifying a homologous form in a database search. In this study, no significant homologies were obtained for either ORF1 or ORF2 in searches of the entire GenBank database. Another way to identify a transposition protein is to demonstrate that mutagenesis of the corresponding gene inhibits transposition. Towards this end, we have initiated development of an assay for Tn*5469* transposition in *E. coli* that may allow us to determine whether ORF1 or ORF2 plays a role in mobilization of the element.

We know very little about the mechanism of Tn*5469* transposition in *F. diplosiphon*. Southern hybridization analysis showed that strain Fd33 contains five genomic copies of the element, whereas red mutant strain FdR1 harbors an extra genomic copy of Tn*5469*. The Southern blot profile for Tn*5469* in strain FdR1 was identical to that for strain Fd33, with the exception of a sixth genomic copy of the element which was localized to the *rcaC* gene (Fig. 2). We have recently extended this analysis to a collection of second-site pigment mutants generated by electroporation of FdR1 cells; these secondary mutants exhibit pigmentation phenotypes epistatic to that for strain FdR1 (18). Of 15 secondary mutants examined, 5 (33%) were found to harbor an extra (seventh) genomic copy of Tn*5469*. With the exception of the seventh copy, the Southern blot Tn*5469* profile for each of these five secondary mutants was identical to that for strain FdR1. In the generation of strain FdR1 and these five secondary pigment mutants, no preexisting copy of Tn*5469* changed location or disappeared from the parental genome. These data suggest a replicative mechanism for transposition of Tn*5469*; however, it must be emphasized that they do not rule out the possibility of a conservative transposition mechanism. Because this cyanobacterium harbors multiple cellular copies of its genome (30), determining whether the transposition mechanism is replicative or nonreplicative will require a more sophisticated molecular analysis.

This study shows that Tn*5469* is not widely distributed among morphologically distinct genera of cyanobacteria. Sequences homologous to Tn*5469* were not detected in genomic DNA from *Synechocystis* sp. strain PCC 6803, *Anabaena* sp. strain PCC 7120, *Synechococcus* sp. strain PCC 7942, *Nostoc* sp. strain PCC 8009, or *Anabaena* sp. strain ATCC 29413 (Fig. 4). This result was not unexpected, since most of the characterized cyanobacterial IS elements are similarly limited in their distribution. In a broad survey, neither IS*701* nor IS*702* from *Calothrix* sp. strain PCC 7601 was found to be widely distributed among a large number of diverse strains (21). Of the two, IS*702* appeared to be the most dispersed; sequences homologous to IS*702* were detected in heterocystous, nonheterocystous, and even unicellular forms. The distribution of IS*892* (8) and IS*895* (1) from *Anabaena* sp. strain PCC 7120 is also limited; sequences homologous to IS*892* and IS*895* were found in closely related filamentous species but not in several unicellular genera. The Tn*5469* data suggest that the element was acquired by *F. diplosiphon* after divergence of the examined strains.

ACKNOWLEDGMENTS

We thank P. Shubert and J. Wright for excellent technical help, R. Hirschberg, J. Meeks, and W. Vermaas for providing cyanobacterial strains, and W. K. Thomas for helpful discussions and comments. We are very grateful to L. Hutt-Fletcher and J. Golden for critically reviewing the manuscript.

This research was partially supported by grant 3493050 from the University of Missouri Research Board to M.R.S.

REFERENCES

1. **Alam, J., J. M. Vrba, Y. Cai, J. A. Martin, L. J. Weislo, and S. E. Curtis.** 1991. Characterization of the IS*895* family of insertion sequences from the cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **173:**5778–5783.
2. **Allen, M. M.** 1968. Simple conditions for growth of unicellular blue-green algae on plates. J. Phycol. **4:**1–4.
3. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
4. **Bancroft, I., and C. P. Wolk.** 1989. Characterization of an insertion sequence (IS*891*) of novel structure from the cyanobacterium *Anabaena* sp. strain M-131. J. Bacteriol. **171:**5949–5954.
5. **Bennett, P. M.** 1991. Transposable elements and transposition in bacteria, p. 323–364. *In* U. N. Streips and R. E. Yasbin (ed.), Modern microbial genetics. Wiley Publishers, New York.
6. **Bogorad, L.** 1975. Phycobiliproteins and complementary chromatic adaptation. Annu. Rev. Plant Physiol. **26:**369–401.
7. **Bruns, B. U., W. R. Briggs, and A. R. Grossman.** 1989. Molecular characterization of phycobilisome regulatory mutants of *Fremyella diplosiphon*. J. Bacteriol. **171:**901–908.
8. **Cai, Y.** 1991. Characterization of insertion sequence IS*892* and related elements from the cyanobacterium *Anabaena* sp. strain PCC 7120. J. Bacteriol. **173:**5771–5777.
9. **Chiang, G. G., M. R. Schaefer, and A. R. Grossman.** 1992. Complementation of a red-light-indifferent cyanobacterial mutant. Proc. Natl. Acad. Sci. USA **89:**9415–9419.
10. **Cobley, J. G., E. Zerweck, R. Reyes, A. Mody, J. R. Seludo-Unson, H. Jaeger, S. Weerasuriya, and S. Navankasattusas.** 1993. Construction of shuttle plasmids which can be efficiently mobilized from *Escherichia coli* into the chromatically adapting cyanobacterium, *Fremyella diplosiphon*. Plasmid **30:**90–105.
11. **Coffin, J. M.** 1991. Retroviridae and their replication, p. 645–708. *In* B. N. Fields and D. M. Knipe (ed.), Fundamental virology, 2nd ed. Raven Press, New York.
12. **Fayet, O., P. Ramond, P. Polard, M. F. Prere, and M. Chandler.** 1990. Functional similarities between retroviruses and the IS*3* family of bacterial insertions sequences? Mol. Microbiol. **4:**1771–1777.
13. **Flores, C. C., M. I. Qadri, and C. P. Lichtenstein.** 1990. DNA sequence analysis of five genes, *tnsA*, *tnsB*, *tnsC*, *tnsD*, and *tnsE*, required for Tn*7* transposition. Nucleic Acids Res. **18:**901–911.
14. **Galas, D. J., and M. Chandler.** 1989. Bacterial insertion sequences, p. 109–162. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.
15. **Grossman, A. R.** 1990. Chromatic adaptation and the events involved in phycobilisome biosynthesis. Plant Cell Environ. **13:**651–666.
16. **Grossman, A. R., M. R. Schaefer, G. G. Chiang, and J. L. Collier.** 1993. The phycobilisome, a light-harvesting complex responsive to environmental conditions. Microbiol. Rev. **57:**725–749.
17. **Haandrikman, A. J., C. Van Leeuwen, J. Kok, P. Vos, W. M. de Vos, and G. Venema.** 1990. Insertion elements on lactococcal proteinase plasmids. Appl. Environ. Microbiol. **56:**1890–1896.
18. **Kahn, K., and M. R. Schaefer.** Unpublished data.
19. **Kehoe, D., and A. R. Grossman.** Personal communication.
20. **Kulkosky, J., K. S. Jones, R. A. Katz, J. P. G. Mack, and A. M. Skalka.** 1992. Residues critical for retroviral integrative recombination in a region that is highly conserved among retroviral/retrotransposon integrases and bacterial insertion sequence transposases. Mol. Cell Biol. **12:**2331–2338.
21. **Mazel, D., C. Bernard, R. Schwarz, A. M. Castets, J. Houmard, and N. Tandeau de Marsac.** 1991. Characterization of two insertion sequences, IS*701* and IS*702*, from the cyanobacterium *Calothrix* species PCC 7601. Mol. Microbiol. **5:**2165–2170.
22. **Radstrom, P., O. Skold, G. Swedberg, J. Flensburg, P. H. Roy, and L. Sundstrom.** 1994. Transposon Tn*5090* of plasmid R751, which carries an integron, is related to Tn*7*, Mu, and the retroelements. J. Bacteriol. **176:**3257–3268.
23. **Reed, K. C., and D. A. Mann.** 1985. Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acids Res. **13:**7207–7221.
24. **Ross, J. I., E. A. Eady, J. H. Cove, W. J. Cunliffe, S. Baumberg, and J. C. Wootton.** 1990. Inducible erythromycin resistance in staphylococci is encoded by a member of the ATP-binding transport super-gene family. Mol. Microbiol. **4:**1207–1214.
25. **Rowland, S.-J., and K. G. H. Dyke.** 1990. Tn*552*, a novel transposable element from *Staphylococcus aureus*. Mol. Microbiol. **4:**961–975.
26. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
27. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
28. **Schaefer, M. R., G. G. Chiang, and A. R. Grossman.** 1993. RcaC, a novel bacterial regulator protein involved in complementary chromatic adaptation, p. 157–160. *In* N. Murata (ed.), Research in photosynthesis, vol. I. Kluwer Academic Publishers, Dordrecht, The Netherlands.
29. **Schafer, A., A. Jahns, A. Geis, and M. Teuber.** 1991. Distribution of the IS elements IS*S1* and IS*904* in lactococci. FEMS Microbiol. Lett. **80:**311–318.
30. **Tandeau de Marsac, N.** Personal communication.
31. **Tandeau de Marsac, N.** 1983. Phycobilisomes and complementary chromatic adaptation. Bull. Inst. Pasteur **81:**201–254.
32. **Varmus, H., and P. Brown.** 1989. Retroviruses, p. 53–108. *In* D. E. Berg and M. M. Howe (ed.), Mobile DNA. American Society for Microbiology, Washington, D.C.