A Novel Pair of Terminal Protein and Telomere-Associated Protein for Replication of the Linear Chromosome of *Streptomyces griseus* IFO13350

Hirokazu Suzuki, Kazuya Marushima, Yasuo Ohnishi, and Sueharu Horinouchi

Department of Biotechnology, Graduate School of Agriculture and Life Sciences, The University of Tokyo, Bunkyo-ku, Tokyo 113-8657, Japan

Received July 2, 2008; Accepted July 25, 2008; Online Publication, November 7, 2008

The linear chromosome of *Streptomyces griseus* IFO13350 contains not only atypical telomere sequences but also probable pseudogenes for two typical telomeric proteins. Two identical operons (SGR98t- SGR97t near one telomere and SGR7041t- SGR7042t near the other telomere) in the terminal inverted repeat sequence were predicted to encode a novel pair of telomeric proteins. SGR97t, a 185-amino-acid protein showing only 18% amino acid sequence identity to typical terminal proteins of *Streptomyces*, was found to be attached to the chromosomal ends, as determined by immunological analysis. On the other hand, electrophoretic mobility shift assays showed that SGR98t, an 837-amino-acid protein having a DnaB-like helicase C-terminal domain, was capable of binding specifically to the single-stranded terminal DNA corresponding to the 3′ overhang of the replication intermediate. These results indicate that SGR97t (and SGR7042t) and SGR98t (and SGR7041t) were the functional telomeric proteins in the replication of the linear chromosome of *S. griseus* IFO13350.

**Key words:** *Streptomyces griseus*; linear chromosome; telomere; terminal protein; DNA replication

*Streptomyces* species are high G+C, Gram-positive, soil-inhabiting, filamentous eubacteria. Unlike most chromosomes in eubacteria, *Streptomyces* contains a linear chromosome. In addition, many *Streptomyces* members have linear plasmids with sizes of 12 to 1,700 kb. Linear chromosomes and plasmids of *Streptomyces* have terminal inverted repeats (TIRs) that carry terminal proteins (TPs) covalently bound to the 5′ ends.1,2) Replication of *Streptomyces* linear chromosomes and plasmids proceeds bidirectionally from the centrally located replication origin toward the telomeres.3–5) Replication of this type leaves a single-stranded overhang of about 280 nucleotides at the 3′ telomeric ends as a replication intermediate. The single-stranded overhang, whose nucleotide sequence is usually conserved, can form a foldback structure with a Y-shaped duplex and three additional hairpins6) (Fig. 1). To resolve the “end-of-replication problem” or to convert the 3′ overhang to a double strand, a telomere-associated protein (TAP) binds the 3′ overhang and probably recruits TP to the telomere, resulting in TP-primed DNA synthesis.7–9) In accordance with the conserved telomere sequences, the two telomere-binding proteins, TP and TAP, are highly conserved among many *Streptomyces* species. The conserved TPs (referred to as Tpg) are similar in size, 184–185 amino acids (aa), and in amino acid sequence. The conserved TAPs (referred to as Tap hereafter) are approximately 750 aa proteins with a putative helix-turn-helix DNA-binding domain at their N-terminal ends. Both Tpg and Tap proteins are essential for the propagation of *Streptomyces* linear replicons in a linear, but not a circular, form.

Kinashi and his colleagues10) cloned the telomere of *Streptomyces griseus* 2247, showing that the sequence was different from the typical ones conserved among other *Streptomyces* species, such as *Streptomyces lividans*. The unique telomere sequence contains extensive palindromes, but cannot form a foldback structure (Fig. 1). The telomeric proteins of the linear chromosome of *S. griseus* 2247 are unknown. Recently, the complete nucleotide sequence of the linear chromosome of *S. griseus* IFO13350 was determined.11) It consists of 8,545,929 bp with extremely long TIRs of 132,910 bp each. Its telomere sequence is different from the unique telomere sequence of *S. griseus* 2247 and the typical *Streptomyces* ones. Like the telomere of *S. griseus* 2247, the telomere of *S. griseus* IFO13350 has several palindromes, but cannot form a foldback structure (Fig. 1). Reflecting the difference in telomere sequence and structure, *S. griseus* IFO13350 appears not to produce
any conserved Tpg or Tap protein, because the chromosome has probable pseudogenes for them (SGR6986 and SGR6987 respectively). As shown in Fig. 2A and B, SGR6986 and its upstream region encoded two polypeptides showing very high sequence similarity with conserved Tpg, but a 73-bp deletion should disrupt the conserved tpg. On the other hand, SGR6987 with its upstream region is probably a ruin of the conserved tap (Fig. 2A and C). A large part of the 5’ portion of the conserved tap is missing. Several mutations, including deletions and frame-shifts, are also present in the 3’ portion. The unique telomere sequence and the absence of functional tpg and tap genes suggest occurrence of a novel pair of telomeric proteins in S. griseus IFO13350. Hence, we searched the S. griseus IFO13350 genome, especially near the chromosomal ends, for a novel pair of genes encoding telomeric proteins. We found two copies of candidates for such an operon (SGR98t-SGR97t near one telomere and SGR7041t-SGR7042t near the other telomere) in both TIRs. In this paper, we propose that SGR97t (= SGR7042t, named GtpA) and SGR98t (= SGR7041t, named GtpB) represent the functional TP and TAP, respectively, in S. griseus IFO13350.

**Materials and Methods**

**Bacterial strains and media.** S. griseus IFO13350 (= NBRC102592) was obtained from the Institute of Fermentation (Osaka, Japan). It was grown at 30 °C in YPD medium (0.2% yeast extract, 0.4% Bacto peptone, 0.5% NaCl, 0.5% MgSO₄·7H₂O, 1% glucose, and 0.5% glycine, pH 7.2). S. lividans TK21 and Streptomyces coelicolor A3(2) M130 were obtained from D.A. Hopwood. These strains were grown at 30 °C in YEME medium.

**Preparation of chromosomal DNA.** Total mycelium of 100 ml culture was suspended in 20 ml of a buffer consisting of 50 mM Tris–HCl, pH 8.0, 10 mM EDTA, 30% sucrose, 2 mg/ml lysozyme, and 10 μg/ml RNase A. In the preparation of protease-treated DNA, protease K at a final concentration of 0.1 mg/ml was also added. Otherwise, (p-amidinophenyl)methanesulfonfluoride (final concentration, 0.1 mM) was added to the mixture to avoid possible protein degradation by endogenous proteases. The mixture was incubated at 37 °C for 30 min, followed by the addition of SDS (final, 0.1%), and further incubated at 55 °C for 30 min. NaCl (final, 1 M), cetyltrimethylammonium bromide (final, 0.5%), and 5 ml of chloroform were successively mixed into the mixture. The mixture was centrifuged at 10,000 x g for 5 min, total DNA in the supernatant was precipitated with isopropyl alcohol and subjected to ultracentrifugation at 543,000 x g at 18 °C for 16 h in TE buffer (10 mM Tris–HCl and 1 mM EDTA, pH 7.5) containing 0.53 g/ml CsCl, 0.25 g/ml guanidine hydro-
amplified by PCR using primers 5'-TCTCGAGCTAGGACGCAGCAGGACGAGCCACCAGGAAGA-3' (the start codon of SGR97t in boldface and an NdeI site underlined) and 5'-ACAATGAAACCAGGAAGATCCCTGACC-3' (the stop codon of SGR98t in boldface and an XhoI site underlined), cloned in pCR4Blunt-TOPO (Invitrogen, Carlsbad, CA), and sequenced to confirm the absence of PCR errors. From this plasmid, the recombinant SGR97t sequence was excised and placed between the NdeI and XhoI sites of pET-17b (Novagen, Darmstadt, Germany), resulting in pET-SGR97t. The SGR98t sequence was amplified by PCR using primers 5'-ACATAGAAACCAGGAAGATCCCTGACC-3' (the start codon of SGR98t in boldface and an NdeI site underlined) and 5'-TCTCGAGCTAGGACGCAGCAGGACGAGCCACCAGGAAGA-3' (the stop codon of SGR98t in boldface and an XhoI site underlined), cloned in pCR4Blunt-TOPO, and sequenced to confirm the absence of PCR errors. The SGR98t sequence was placed between the NdeI and XhoI sites of pIJ4123, resulting in pIJ4123- SGR98t.

Preparation of anti-SGR97t antibody. E. coli BL21(DE3) [pET- SGR97t] ([] denotes the plasmid-carrier state) produced the recombinant SGR97t protein in the insoluble fraction. The recombinant SGR97t protein was purified under denaturing conditions with Ni2+-nitrilotriacetic acid agarose (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. Finally, the protein was precipitated with trichloroacetic acid, and the precipitate was dissolved in a buffer (20 mM sodium phosphate, pH 7.2, 0.15 M NaCl, and 0.1% SDS). This sample was used in the production of rabbit anti-SGR97t polyclonal antibody. The anti-SGR97t antibody was purified from the antisera by ammonium sulfate precipitation (50% saturation) and column chromatography using protein A agarose (GE Healthcare, Chalfont St. Giles, UK), according to the manufacturer’s instructions. The purified antibody was dialyzed against 20 mM sodium phosphate (pH 7.0) and concentrated to 30 mg/ml.

Western blotting. Proteins were separated by SDS-12.5% polyacrylamide gel electrophoresis (PAGE) and transferred to a Millipore Immobilon-P membrane (Millipore, Billerica, MA). After the membrane was treated with the recommended dilution of a blocking reagent (GE Healthcare) at room temperature for 1 h, the blots were probed with the anti-SGR97t antibody (1:200 dilution), followed by the donkey anti-rabbit IgG-horseradish peroxidase conjugate (1:25,000 dilution) (GE Healthcare). The immune complexes were visualized by means of a peroxidase-catalyzed chemiluminescence reaction using ECL Plus Western Blotting Detection Reagents (GE Healthcare).

Immunoprecipitation-PCR. The BglII-digested S. griseus IFO13350 chromosome (0.5 μg), prepared without proteinase K treatment, was incubated with the anti-SGR97t antibody at 4 °C for 6 h in buffer A (20 mM sodium phosphate, pH 7.0, and 0.5 mM NaCl) containing 0.1 mg/ml bovine serum albumin in a total volume of 100 μl. Then 10 μl of protein A agarose was added to the mixture, followed by incubation at 4 °C for 14 h with
Preparation of SGR98t. *S. lividans* TK21 [pIJ4123- SGR98t] was cultured at 30 °C for 2 d in YEME medium supplemented with 5 μg/ml kanamycin. The tipA promoter was then induced by the addition of thiostrepton to a final concentration of 5 μg/ml, and cultivation was continued for additional 2 d. All operations described below were carried out at 4 °C. The harvested cells (5 ml wet volume) were suspended in a buffer (50 mM sodium phosphate, pH 8.0, 0.5 M NaCl, 20% glycerol, and 7 mM 2-mercaptoethanol) containing 2 mg/ml lysozyme, and incubated on ice for 30 min. The cells were disrupted by sonication, followed by centrifugation at 10,000 × g for 20 min to remove cell debris. From this soluble fraction, the recombinant SGR98t protein was purified with a HisTrap column (1 ml, GE Healthcare). The protein solution was concentrated and applied to a gel filtration column (HiLoad Superdex 200 16/60 prep grade, GE Healthcare) on fast protein liquid chromatography with isocratic elution in buffer B (20 mM Tris–HCl, pH 8.0, 0.1 M NaCl, 10% glycerol, 1 mM dithiothreitol, and 1 mM EDTA) at a flow rate of 1 ml/min.

Electrophoretic mobility shift assay. The terminal 356-bp sequence (nt positions 1 to 356) of the *S. griseus* IFO13350 chromosome was amplified by PCR using primers 1F (5′-GCAGGATATAGAATACACC-3′) and 5′-ACGAGCACAAACGCTGATC-3′. A 589-bp region (nt positions 95,677 to 96,265) in the *SGR97t*–*SGR98t* operon was detected as an internal region of the chromosome using primers 5′-GAGCGTGCACGTCTGGT-3′ and 5′-CCA-TGGCGGGATATGAGAATCCATC-3′. A terminal 522-bp region of the *S. coelicolor* A3(2) chromosome (nt positions 491 to 1,012) was detected using primers 5′-ACG-GATTCCGACGTCCGGGTGAC-3′ and 5′-TTGGTCTGTCGCGGCGTAGTG-3′.

**Results**

*Protein capping of S. griseus IFO13350 telomere*

The presence of a covalently linked TP at the telomeres can be confirmed by retardation of the telomere DNA during electrophoresis. We extracted the chromosomal DNA from *S. griseus* IFO13350 cells with and without proteinase K treatment. The chromosome DNAs were digested with *PstI* and analyzed by Southern hybridization using probe I for a 350-bp region near the chromosomal ends and probe II for an internal 406-bp region (Fig. 3A). A signal for the internal 406-bp *PstI* fragment (nt positions 718 to 1,123) was similarly detected for the chromosomal DNAs prepared with and without proteinase K treatment (Fig. 3B) by probe II. However, a signal for the 717-bp terminal *PstI* fragment (positions 1 to 717) detected by probe I was clearly retarded only by the DNA prepared without the proteinase K treatment (Fig. 3B), indicating that a protein(s) were attached covalently to the telomere DNA, probably to its 5′ end, as observed for conserved telomere systems of Streptomyces.
we tentatively designated the former SGR98t and the latter SGR97t. SGR97t encodes a 185-aa protein showing very low sequence similarity (18% identity) to the putative ORFs. The positions of SGR6986 and SGR6987 (probable pseudogenes for conserved Tpg and Tap respectively) on the chromosome are also shown. The nucleotide and amino acid sequences of the intervening region between SGR7041t and SGR7042t are shown. B, Amino acid alignment of Tpg and GtpA. GtpA is the SGR97t (=SGR9721) gene product. The helix-turn-helix (HTH) DNA-binding motif of TpgL, suggested by Yang et al., is indicated.

**Fig. 4.** A Novel Pair of TP and TAP in *S. griseus* IFO13350.

A, Schematic representation of two identical operons for a novel pair of TP and TAP in the TIRs. Black arrows indicate the extent and direction of the putative ORFs. The positions of SGR6986 and SGR6987 (probable pseudogenes for conserved Tpg and Tap respectively) on the chromosome are also shown. The nucleotide and amino acid sequences of the intervening region between SGR7041t and SGR7042t are shown. B, Amino acid alignment of Tpg and GtpA. GtpA is the SGR97t (=SGR9721) gene product. The helix-turn-helix (HTH) DNA-binding motif of TpgL, suggested by Yang et al., is indicated.

**Fig. 5.** Covalent Binding of SGR97t to the Termini of the *S. griseus* IFO13350 Linear Chromosome.

A, Western blotting using anti-SGR97t antibody. Crude proteins (1 μg) of *S. griseus* IFO13350 mycelium with (lane 1) and without (lane 2) recombinant SGR97t protein (2 ng) were examined. Proteins that were co-purified with chromosomal DNAs (10 μg) were examined: *S. griseus* IFO13350 chromosomal DNA digested with *MboI* (lane 3), HaeIII plus benzonase (lane 6), and HaeIII (lane 7), and *S. coelicolor* A3(2) chromosomal DNA digested with HaeIII plus benzonase (lane 5, a negative control). As a positive control, the recombinant SGR97t protein (0.5 ng) was examined (lane 4). The signals arising from SGR97t are indicated by arrowheads *a* to *d*. A non-specific signal arising from a large amount of benzonase is indicated by arrowhead Bz. The smeared signals in high molecular weight regions shown in lanes 3, 5, 6, and 7 are also non-specific. B, Pull-down assays of the terminal DNA of the *S. griseus* IFO13350 (5g) chromosome with anti-SGR97t antibody. The anti-SGR97t antibody used for immunoprecipitation was 0 (lane 1), 60 (lane 2), and 600 μg/ml (lanes 3 to 5). Using the immunoprecipitated DNAs (precipitate) and the remaining DNAs (supernatant) as templates, PCR was performed to amplify part of the terminal DNA fragment (TM) and an internal DNA fragment (IN). As a negative control, the XhoI-digested *S. coelicolor* A3(2) (Sc) chromosome was used (lane 5).

**Fig. 4.** A Novel Pair of TP and TAP in *S. griseus* IFO13350.

A, Schematic representation of two identical operons for a novel pair of TP and TAP in the TIRs. Black arrows indicate the extent and direction of the putative ORFs. The positions of SGR6986 and SGR6987 (probable pseudogenes for conserved Tpg and Tap respectively) on the chromosome are also shown. The nucleotide and amino acid sequences of the intervening region between SGR7041t and SGR7042t are shown. B, Amino acid alignment of Tpg and GtpA. GtpA is the SGR97t (=SGR9721) gene product. The helix-turn-helix (HTH) DNA-binding motif of TpgL, suggested by Yang et al., is indicated.

**Fig. 5.** Covalent Binding of SGR97t to the Termini of the *S. griseus* IFO13350 Linear Chromosome.

A, Western blotting using anti-SGR97t antibody. Crude proteins (1 μg) of *S. griseus* IFO13350 mycelium with (lane 1) and without (lane 2) recombinant SGR97t protein (2 ng) were examined. Proteins that were co-purified with chromosomal DNAs (10 μg) were examined: *S. griseus* IFO13350 chromosomal DNA digested with *MboI* (lane 3), HaeIII plus benzonase (lane 6), and HaeIII (lane 7), and *S. coelicolor* A3(2) chromosomal DNA digested with HaeIII plus benzonase (lane 5, a negative control). As a positive control, the recombinant SGR97t protein (0.5 ng) was examined (lane 4). The signals arising from SGR97t are indicated by arrowheads *a* to *d*. A non-specific signal arising from a large amount of benzonase is indicated by arrowhead Bz. The smeared signals in high molecular weight regions shown in lanes 3, 5, 6, and 7 are also non-specific. B, Pull-down assays of the terminal DNA of the *S. griseus* IFO13350 (5g) chromosome with anti-SGR97t antibody. The anti-SGR97t antibody used for immunoprecipitation was 0 (lane 1), 60 (lane 2), and 600 μg/ml (lanes 3 to 5). Using the immunoprecipitated DNAs (precipitate) and the remaining DNAs (supernatant) as templates, PCR was performed to amplify part of the terminal DNA fragment (TM) and an internal DNA fragment (IN). As a negative control, the XhoI-digested *S. coelicolor* A3(2) (Sc) chromosome was used (lane 5).
Specific binding of SGR98t to the single-stranded terminal probe of the telomeric DNA

Tap is able to bind specifically to the single-stranded telomeric DNA of linear plasmids and chromosomes, suggesting that it is involved in the recruitment of Tpg to the ends of telomeres.8) Hence we examined the ability of SGR97t to bind specifically to the 3′ end of the telomeric DNA. Although the molecular function of SGR97t remains to be elucidated, we propose that SGR97t bound covalently to the ends of the chromosome.

Specific binding of SGR98t to the single-stranded 3′ overhang of the telomeric DNA

Tap is able to bind specifically to the single-stranded telomeric DNA of linear plasmids and chromosomes, suggesting that it is involved in the recruitment of Tpg to the ends of telomeres.8) Hence we examined the ability of SGR98t to bind single-stranded DNA by gel mobility shift assay. Recombinant SGR98t having an N-terminal His6-tag was produced by S. lividans TK21 [pIJ4123-SGR98] and purified by immobilized metal affinity chromatography and gel filtration column chromatography (Fig. 6A). On gel filtration column chromatography, the recombinant SGR98t was eluted predominantly (about 80%) at a retention time representing 180 kDa. Therefore, SGR98t formed a dimer in solution, while a small population (about 20%) was eluted broadly at earlier retention times.

Gel mobility shift assays were carried out using the protein eluted at a retention time representing 180 kDa and the single-stranded DNA, including the 3′ end of the chromosome, as the 32P-labeled probe. As shown in Fig. 6B, the single-stranded terminal probe was bound by the recombinant SGR98t and shifted in gel mobility (lanes 1 to 5). The shifted signal almost completely disappeared in the presence of an excess of cold single-stranded DNA (lane 6). However, non-specific single-stranded DNA fragments did not affect the binding of SGR98t to the single-stranded probe (lane 7). On the other hand, the recombinant SGR98t did not bind to the double-strand probe, when similarly determined by gel mobility shift assay (Fig. 6B, lanes 8 to 12). These results indicate that SGR98t possessed the ability to bind specifically to the 3′ single-strand overhangs of the telomeric DNA. Although the molecular function of SGR98t remains to be elucidated, we propose that SGR98t is the functional TAP of S. griseus IFO13350.

Discussion

SGR97t was attached to the chromosomal ends and SGR98t was capable of binding specifically to the
single-stranded DNA of the chromosomal ends. Hence we propose that SGR97t and SGR98t are the functional TP and TAP respectively, both of which are responsible for replication of the telomeres in S. griseus IFO13350. Hence we designated SGR97t and SGR98t GtpA (S. griseus IFO13350 telomeric protein) and GtpB respectively.

GtpA is a 185-aa protein, similar in size to the conserved Tggs of Streptomyces, but GtpA shows very low amino acid sequence similarity (18% identity) with them. Furthermore, GtpA has a very low predicted pI value of 5.77, whereas not only the conserved Tggs of Streptomyces (pI 11–12) but also two different TPs of Streptomyces, pRL2.3c-encoded protein (pI 10.32) and Tpc (pI 11.52),15 have much higher predicted pI values. GtpA is the first example of TP with a low predicted pI value in Streptomyces. While the conserved Tggs perhaps have a helix-turn-helix DNA-binding motif near the N-terminus (aa positions 36 to 57, Fig. 4B),9 no significant helix-turn-helix DNA-binding motif is predicted on the GtpA sequence by MOTIF search (http://motif.genome.jp/). Because a function of Tap recruiting Tpg to the telomere has been suggested,8 the DNA-binding activity of Tpg might not be required in the telomere system of Streptomyces. Indeed, the in vitro binding of TpgL of S. lividans to both single-stranded and double-stranded DNAs is non-specific.8

GtpB is an 837-aa protein having a domain (aa positions 174 to 341) that resembles the DnaB-like helicase C-terminal domain. The pRL2.4c-encoded protein, which is a TAP of a different family, also has a domain (aa positions 175 to 520) that is homologous to the superfamily II helicase of Thiothrix denitrificans.14 Therefore, a domain related to the DNA helicase presumably has an important role in the molecular function of these novel TAPS. It is noteworthy that the approximately 200-aa C-terminal regions of these novel TAPS are somewhat homologous; a C-terminal region (from aa 617 to 831) of GtpB (837 aa) showed low sequence similarity (26% identity and 39% similarity) with a C-terminal region (from aa 860 to 1084) of the pRL2.4c-encoded protein (1100 aa). GtpB can bind specifically to the S. griseus IFO13350 terminal single-stranded DNA corresponding to the 3’ overhang of the replication intermediate. We assume that GtpB recognizes some secondary structures of the unique 3’ overhang, although the manner of DNA-binding of GtpB remains to be elucidated. A possible interaction between GtpB and GtpA also remains to be elucidated.

The diversity of the telomere sequences of Streptomyces linear plasmids and their telomere proteins has recently become apparent.14,15 In this study, we discovered the possibility that linear chromosomes of Streptomyces also have diversity in their telomere sequences and telomeric proteins. The telomere sequence of S. griseus IFO13350 is different from the unique telomere sequence of S. griseus 2247.10 S. griseus 2247 presumably contains a different pair of TP and TAP, which recognizes and replicates the atypical telomere sequence. S. griseus IFO13350 has probable pseudogenes, SGR6986 and SGR6987, that once encoded conserved Tpg and Tap, respectively, for the typical Streptomyces telomere, but underwent rearrangement and mutations. In addition, the functional telomeric proteins, GtpA and GtpB, are encoded as an operon (SGR98t–SGR97t near one telomere and SGR7041t–SGR7042t near the other telomere) in the TIRs. Therefore, we assume that S. griseus IFO13350 acquired the unique telomere together with a novel pair of TP and TAP from a certain linear plasmid during its evolution, and that the original tpg and tap became unnecessary and decayed. If this is the case, an unidentified linear plasmid with the same telomere sequence as that of the S. griseus IFO13350 chromosome will probably be isolated from some Streptomyces strain. In screening for such a plasmid, colony PCR for amplification of the coding sequence of GtpA or GtpB is a promising strategy.

**Acknowledgments**

We are grateful to C.-H. Huang and C. W. Chen for discussion on the unique telomere sequence and probable pseudogenes for Tpg and Tap of S. griseus IFO13350. This work was supported by a Grant-in-Aid for Scientific Research on Priority Area “Applied Genomics” from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

**References**


