Molecular analysis of the *recA* gene and SOS box of the purple non-sulfur bacterium *Rhodopseudomonas palustris* no. 7

Valérie Dumay,[†] Masayuki Inui and Hideaki Yukawa

Author for correspondence: Hideaki Yukawa. Tel: +81 774 75 2308. Fax: +81 774 75 2321.
e-mail: yukawa@rite.or.jp

Molecular Microbiology and Genetics Lab, Research Institute of Innovative Technology for the Earth, 9-2 Kizugawadai, Kizu-cho, Soraku-gun, Kyoto 619-0292, Japan

The *recA* gene of the purple non-sulfur bacterium *Rhodopseudomonas palustris* no. 7 was isolated by a PCR-based method and sequenced. The complete nucleotide sequence consists of 1089 bp encoding a polypeptide of 363 amino acids which is most closely related to the RecA proteins from species of *Rhizobiaceae* and *Rhodospirillaceae*. A *recA*-deficient strain of *R. palustris* no. 7 was obtained by gene replacement. As expected, this strain exhibited increased sensitivity to DNA-damaging agents. Transcriptional fusions of the *recA* promoter region to *lacZ* confirmed that the *R. palustris* no. 7 *recA* gene is inducible by DNA damage. Primer extension analysis of *recA* mRNA located the *recA* gene transcriptional start. A sequential deletion of the fusion plasmid was used to delimit the promoter region of the *recA* gene. A gel mobility shift assay demonstrated that a DNA–protein complex is formed at this promoter region. This DNA–protein complex was not formed when protein extracts from cells treated with DNA-damaging agents were used, indicating that the binding protein is a repressor. Comparison of the minimal *R. palustris* no. 7 *recA* promoter region with the *recA* promoter sequences from other α -Proteobacteria revealed the presence of the conserved sequence GAACA-N₆-G(A/T)AC. Site-directed mutations that changed this consensus sequence abolished the DNA-damage-mediated expression of the *R. palustris* *recA* gene, confirming that this sequence is the SOS box of *R. palustris* and probably plays the same role in other α -Proteobacteria.

Keywords: *recA*, photosynthetic bacteria, gene disruption, SOS box

INTRODUCTION

Rhodopseudomonas palustris no. 7 belongs to the group of purple non-sulfur bacteria which can grow either phototrophically under anaerobic conditions or chemotrophically under aerobic conditions. *R. palustris* no. 7 is of potential biotechnological interest because of its ability to produce hydrogen at a high rate using alcohol as carbon source (Fujii *et al.*, 1983). Several genes encoding enzymes involved in key metabolic pathways have already been characterized (Inui *et al.*, 1997). Amplification of such genes on plasmid vectors would enable genetic engineering of strains of biotechnological

relevance. However, the presence of multiple copies of a gene in a bacterium is often a source of genetic instability due to recombination between duplicated genes. Homologous recombination in bacteria is mainly governed by the *recA* gene product (Miller & Kokjohn, 1990). Consequently, various bacterial RecA[−] strains show greatly reduced frequencies of homologous recombination. A *recA* mutant of *R. palustris* no. 7 is thus expected to stably maintain modified host genes.

As a prerequisite to the construction of a recombination-deficient strain, we have characterized the *recA* gene of *R. palustris* no. 7. Isolation of the *recA* gene enabled us to study the SOS response in this bacterium. The RecA protein plays a major role in the induction of the SOS response (Walker, 1984). Following exposure to DNA-damaging agents, RecA promotes the autocleavage of LexA, the repressor of some 20 SOS genes including *recA* itself (Little, 1991). Although the RecA–LexA

[†] **Present address:** E. C. Slater Institute, Plantage Muidergracht 12, 1018 TV Amsterdam, The Netherlands.

Abbreviation: MMS, methyl methane sulfonate.

The GenBank/DBJ/EMBL accession number of the sequence reported in this paper is D84467.

pathway for the SOS response is well documented in *Escherichia coli*, little is known about the molecular mechanisms of SOS response in other Gram-negative bacteria. Purple non-sulfur bacteria together with *Rhizobiaceae* belong to the α -Proteobacteria. All α -proteobacterial *recA* genes so far isolated are intraspecifically inducible by DNA-damaging agents (Riera *et al.*, 1994). However, none of them are induced in *E. coli* (Riera *et al.*, 1994). Accordingly, no *E. coli*-like LexA binding site upstream of the coding regions of these genes has been identified. These data suggest that different regulatory sequences are involved in the regulation of the expression of the *recA* genes of α -Proteobacteria.

Here we describe the analysis of the *recA* gene of *R. palustris*, the construction of a *recA* mutant, and the regulation of the *recA* gene in this strain.

METHODS

Bacterial strains, plasmids and growth conditions. Strains and plasmids used in this study are listed in Table 1. Plasmid pGP704 is an R6K-based suicide delivery plasmid that is only maintained in bacteria producing the π protein encoded by the *pir* gene. It also carries the conjugal transfer origin of the RP4 plasmid and therefore can be transferred in a broad range of

Gram-negative bacteria when provided *in trans* with mobilization functions (Miller & Mekalanos, 1988). Plasmid pMG102 is an *E. coli*-*R. palustris* no. 7 shuttle vector (M. Inui and others, unpublished data). *E. coli* strains were routinely grown in LB medium (Miller, 1982). *R. palustris* no. 7 was cultivated aerobically at 30 °C in van Niel's medium (van Niel, 1944). Its sensitivity to commonly used antibiotics was evaluated by plating early-stationary-phase cells from aerobic culture on plates of van Niel's medium containing increasing levels of antibiotic. When appropriate, antibiotics were added to the media in the following amounts: for *R. palustris* no. 7, kanamycin was used at a final concentration of 200 $\mu\text{g ml}^{-1}$ for agar medium or 50 $\mu\text{g ml}^{-1}$ in liquid medium; for *E. coli*, ampicillin and kanamycin were used at a final concentration of 50 $\mu\text{g ml}^{-1}$.

DNA techniques. Standard methods were used for Southern, lambda plaque and colony hybridizations (Sambrook *et al.*, 1989). Restriction endonucleases and other commercially available enzymes were employed as recommended by the manufacturers. Sequencing was performed using the dideoxynucleotide chain-termination method and a PRISMReady Reaction Dye Cycle Sequencing kit (Applied Biosystems) with fluorescently labelled M13 and RP1 primers on an Applied Biosystems 373A automated DNA sequencer. Plasmid template DNA was linearized prior to reaction and 10% DMSO was added to the sequencing reaction mixture in order to improve resolution in read-out of the highly GC-rich *Rhodospseudomonas* DNA. DNA sequence data were analysed

Table 1. Strains and plasmids

Strain/plasmid	Relevant characteristics	Reference/source
<i>R. palustris</i>		
No. 7	Wild-type	Fujii <i>et al.</i> (1983)
<i>recA</i> derivative	<i>recA</i> :: ΩKm^r	This study
<i>E. coli</i>		
JM109	F' $\Delta\text{lac-pro hsdR recA1}$	Yanisch-Perron <i>et al.</i> (1985)
JM83	F' $\Delta\text{lac-pro}$ ($\phi 80 \text{ lacZ}\Delta\text{M15}$)	Vieira & Messing (1982)
SM10(λpir)	Km ^r $\lambda\text{pir recA}$::RP4-2-Tc::Mu	Miller & Mekalanos (1988)
Plasmids		
pUC119	Ap ^r , <i>lacZ</i> α	Vieira & Messing (1987)
pUC4-K	Source of Km ^r cartridge	Vieira & Messing (1982)
pGP704	Ap ^r , <i>oriR6K mobRP4</i>	Miller & Mekalanos (1988)
pMC1871	Tet ^r , source of promoterless <i>lacZ</i>	Casadaban <i>et al.</i> (1983)
pMG102	Km ^r , shuttle vector	Authors' laboratory
pMG302	<i>recA</i> in pUC119	This study
pMG307	pGP704 <i>recA</i> :: ΩKm	This study
pMG308	pMG102 <i>recA</i> :: <i>lacZ</i>	This study
pMG310	pMG102 <i>recA</i> :: <i>lacZ</i> ΔEcoRI	This study
pMG318	pMG308 (5' $\Delta 436$ bp of <i>recA</i> :: <i>lacZ</i>)	This study
pMG319	pMG308 (5' $\Delta 528$ bp of <i>recA</i> :: <i>lacZ</i>)	This study
pMG320	pMG308 (5' $\Delta 611$ bp of <i>recA</i> :: <i>lacZ</i>)	This study
pMG315	pMG318 (–30AGAA→CTAG)	This study
pMG314	pMG318 (–28AACA→TACT)	This study
pMG316	pMG318 (–40TTG→AAC)	This study
pMG321	pMG315 + (–38G→A)	This study
pMG322	pMG314 + (–38G→A)	This study
pMG325	pMG318 (–16A→C)	This study
pMG326	pMG318 (–26C→T)	This study

Table 2. Oligonucleotides

Primer	Sequence (5'–3')*	Strand	Position/description
Primers used to generate 5' progressive deletions of the <i>recA</i> promoter			
PX18	GCACGTCGTC <u>TAG</u> ATTGGG	+	–145 to –127; <i>Xba</i> I site italic
PX19	GACAAGTCTAGATTGCGAACC	+	–52 to –32; <i>Xba</i> I site italic
PX20	GCGAATCCGTC <u>TA</u> GAAGGAGC	+	+29 to +49; <i>Xba</i> I site italic
PN	GGATTTGTCCATGGAAGAACC	–	+78 to +99; <i>Nco</i> I site italic
Primers used to generate probes for mobility shift assays			
GR1	GCACGTCGTCAGGATTGGG	+	–145 to –127
GR2	TCCGGATTTTCGACAAGAA	+	–62 to –45
ST	CCGGCAAAATCAACGTACC	–	–19 to –1
Primers used to introduce base changes in the <i>recA</i> promoter and operator			
PM14	CGAACGGAGTACTAATAGGG	+ / –	–37 to –18; <i>Sca</i> I site italic
PM15	CGAACGGCTAGCAAATAGGG	+ / –	–37 to –18; <i>Nhe</i> I site italic
PM16	CAAGAACATAA <u>ACC</u> GAACGG	+ / –	–50 to –31
PM25	GGGTCCGTTGATTTTGCCGG	+ / –	–19 to –1
PM26	CGAACGGAGAATAAATAGGG	+ / –	–37 to –18

* Italic denotes restriction sites (see column 4); underlining indicates mutated nucleotides.

using the INHERIT (Applied Biosystems/Perkin-Elmer) and the GENETYX (Software Development, Tokyo, Japan) programs.

Primer design and DNA amplification. The degenerate primers used in the amplification of *R. palustris* no. 7 *recA* internal fragment were: P1, 5'-TT(TC)ATTGA(TC)GCIGA(AG)CA-(TC)GC-3'; and P2, 5'-CCICCI GTIGTIGTTT(CT)IGG-3' (see Fig. 1). Synthetic oligonucleotides used to generate deletions, mutations or probes for gel mobility shift experiments are listed in Table 2. Amplification reactions were performed in a thermocycler for 30 cycles (1 min at 94 °C for denaturation, 1 min at 37 °C for primer annealing and 2 min at 72 °C for strand synthesis). DMSO (5%) was added to the PCR reaction mixture to improve denaturation of the GC-rich DNA template.

Transformation of *R. palustris* no. 7. When using derivatives of shuttle vector pMG102, *R. palustris* cells were transformed by electroporation according to Donohue & Kaplan (1992). Plasmid pMG307 was mobilized in conjugation experiments with *E. coli* SM10(λ pir) as the donor and *R. palustris* no. 7 as the recipient, as previously reported (Inui *et al.*, 1997). Transconjugants were selected on van Niel's plates supplemented with kanamycin and streptomycin. Streptomycin was used for counterselection of *E. coli* donor cells as *R. palustris* no. 7 was resistant to this antibiotic.

Determination of sensitivity to DNA-damaging agents. *R. palustris* no. 7 was cultured on van Niel's medium and diluted to a final OD₆₀₀ of 0.5 in 0.9% NaCl. The cell suspension was diluted 10⁵-fold and 200 µl of the diluted suspension plated on van Niel's medium. Plates were exposed to a UV lamp (TDM-20, 302 nm UV lamp; Ultraviolet Products) at a distance of 5 cm for various periods of time as indicated in Fig. 2. After irradiation, plates were incubated for 6 d in the dark at 30 °C. For testing sensitivity to methyl methane sulfonate (MMS), cells were plated on plates containing various amounts of the DNA-damaging agent as indicated in Fig. 2.

Primer extension analysis. Total cellular RNA was isolated from *R. palustris* no. 7 grown aerobically to mid-exponential

phase by the hot phenol method as described by Wilkinson (1991). The resulting RNA pellet was dissolved in diethylpyrocarbonate-treated water and stored at –80 °C. The concentration of RNA was determined spectrophotometrically. The sequence of the oligonucleotide used as primer for primer extension experiments is 5'-AGAGCGCCTTG-GATTTGTCC-3'. The primer hybridizes to nucleotides +679 to +688 relative to the translational starting site. This primer was end-labelled with [γ -³²P]ATP and T4 polynucleotide kinase. Primer extension analysis was carried out as previously described (Inui *et al.*, 1997). A sequencing reaction with the same primer and pMG302 as template was run in parallel on a denaturing 6% polyacrylamide gel as a reference for determining the end-point of the extension product. Autoradiography was performed using a Fujix BAS2000 Image Analyzer System (Fuji).

Construction of *recA::lacZ* fusion plasmids. A fusion of *recA* with *lacZ* was constructed by inserting the 950 bp *Xmn*I–*Hinc*II fragment at the *Sma*I site of pMC1871 (Casadaban *et al.*, 1983). The 4 kb *Sall* fragment containing the gene fusion was then cloned into the *Sall* site on shuttle vector pMG102, yielding plasmid pMG308.

5' deletions of the *recA* promoter region were generated by PCR to produce plasmids pMG318, pMG319 and pMG320. Mutations were introduced into the sequences of the 5' primers PX18, PX19 and PX20 (Table 2) to create *Xba*I restriction sites, used for construction of pMG318, pMG319 and pMG320, respectively. The 3' primer was PN (Table 2), which includes a *Nco*I site present within the *R. palustris* no. 7 *recA* gene. The *Xba*I/*Nco*I-digested PCR fragments were cloned into pMG308 in place of the original 950 bp *Xba*I–*Nco*I fragment.

Site-directed mutations in the *recA* promoter region were generated by a multi-step PCR mutagenesis as described by Higuchi *et al.* (1988). Two overlapping fragments were synthesized in separate amplifications using plasmid pMG302 as template. The 3' fragment was amplified with the flanking primer PN and mutagenic primers PM14, PM15, PM16, PM25 and PM26 (Table 2) carrying the mutations introduced in pMG314, pMG315, pMG316, pMG325 and pMG326, re-

spectively. The complementary mutagenic oligonucleotides were used with flanking primer PX18 in a separate PCR to generate the 5' fragment. Both fragments were combined in a subsequent PCR step using the external primers PX18 and PN. The fusion product was digested with *Xba*I/*Nco*I and cloned into pMG308. Four independent clones for each mutation were verified by sequencing of the double-stranded DNA. An unexpected second-site mutation was found for only one fusion product between fragments 5' PX18-PM14 and 3' PM14-PN. The second-site mutation consisted of a G to A change at position -38, within the -35 element of the putative *E. coli*-type σ^{70} promoter. This double mutant, referred to as pMG322, was kept for further expression analysis. Double mutant pMG321 was constructed by replacing the *Xba*I-*Nhe*I fragment from pMG315 by a *Xba*I-*Nhe*I PCR product generated in a PCR reaction using primers PX18 and PM15 with pMG322 as template.

Assay of β -galactosidase activity. β -Galactosidase assays were performed as described by Miller (1982). The cells were harvested at different time intervals and permeabilized with toluene. The *recA* promoter was induced by addition of mitomycin C ($10 \mu\text{g ml}^{-1}$).

Gel mobility shift assay. *R. palustris* cells grown aerobically in van Niel's medium were harvested by centrifugation in the middle of the exponential growth phase. Cells were washed once in PEP buffer (50 mM sodium phosphate buffer pH 7, 0.5 mM EDTA, 0.5 mM PMSF), resuspended in the same buffer and disrupted by sonication. The supernatant obtained after centrifugation for 10 min at 14000 r.p.m. was centrifuged in a Hitachi type 50 rotor for 30 min at 40000 r.p.m. in a Hitachi type α ultracentrifuge.

The 145 bp and 62 bp probes used for mobility shift assays were synthesized by PCR amplification with [γ - ^{32}P]ATP end-labelled primers GR1 and ST, and GR2 and ST (Table 2), respectively. The PCR fragments were separated by electrophoresis in a native polyacrylamide gel (5%). Labelled bands were cut out and the DNA was recovered by electroelution.

Radiolabelled DNA fragments (4 fmol) were incubated with the indicated amount of protein extract (Fig. 5) from *R. palustris* no. 7 in 20 μl binding buffer (10 mM HEPES/NaOH pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM DTT, $1 \mu\text{g BSA } \mu\text{l}^{-1}$, 10 ng poly(dI-dC) μl^{-1} , 10%, v/v, glycerol) for 20 min at room temperature. Samples were electrophoresed for 2 h on a 5% acrylamide/0.13% bisacrylamide (w/v) gel at a constant voltage of 130 V. Gel and running buffers were 22.5 mM Tris base, 22.5 mM boric acid and 0.5 mM EDTA. The gel was dried and radioactivity detected by a radioisotope imaging system (Fujix BAS2000 Image Analyzer System, Fuji).

RESULTS AND DISCUSSION

Cloning strategy for the *recA* gene from *R. palustris* no. 7

A PCR-based approach was used (Duwat *et al.*, 1992). Alignment of protein sequences of previously characterized *recA* genes reveals highly conserved stretches which provide a basis for the design of degenerate oligonucleotide primers. The forward primer P1 corresponded to the conserved RecA sequence FIDAEHA located at positions 92–98 of the *E. coli* RecA. The reverse primer P2 corresponded to the reverse sequence PETTTGG between residues 206 and 212. When used in a PCR reaction with *R. palustris* no. 7 chromosome as

template, these primers generated a 350 bp PCR product whose sequence showed high similarities to *recA* sequences of other bacteria. Using the cloned PCR *recA* fragment as a probe in a Southern hybridization of *R. palustris* no. 7 chromosomal DNA located the *recA* gene within a 1.9 kb *Bam*HI DNA fragment. In order to obtain the full-length *recA* gene, the same probe was used to screen a λ FixII library of *R. palustris* no. 7 chromosomal DNA. The hybridizing 1.9 kb *Bam*HI fragment was subcloned from a recombinant phage into pUC119 and its sequence was determined on both strands.

Nucleotide sequence and its analysis

The complete nucleotide sequence and the deduced amino acid sequence are shown in Fig. 1. The high GC content (64.7 mol%) of the region sequenced was in agreement with the DNA composition described for this genus of Gram-negative bacteria. Sequence analysis revealed the existence of a large ORF with two possible start codons, at positions 625 and 666 (Fig. 1). Based upon comparison with the most closely related RecA proteins [those from *Rhizobium phaseoli* (Michiels *et al.*, 1991), *Rhizobium* (*Sinorhizobium*) *meliloti* (Selbitschka *et al.*, 1991) and *Agrobacterium tumefaciens* (Wardhan *et al.*, 1992)] the ATG codon at position 625 (7 nt downstream of a putative Shine-Dalgarno sequence) is the most probable translational start codon. We assume that the *recA* gene of *R. palustris* no. 7 therefore consists of 1089 nt, encoding 363 amino acids. A sequence which shows strong homology to the consensus of an *E. coli* σ^{70} promoter (Hawley & McClure, 1983) is located 95 nt upstream of the ORF. At position 323–349 upstream of the *R. palustris* no. 7 *recA* coding region, the sequence CTG-N₁₀-CAG occurs, which resembles a canonical *E. coli* SOS box, known to bind the LexA regulatory protein (Walker, 1984).

The deduced amino acid sequence displayed the closest similarity with the RecA proteins of *R. (S.) meliloti* (82% identity) and *A. tumefaciens* (79% identity). It showed 75% and 74% identity with *Rhodobacter capsulatus* (Fernandez de Henestrosa *et al.*, 1995) and *Rhodobacter sphaeroides* (Calero *et al.*, 1994) RecA proteins, respectively. These comparisons reflect the taxonomic position of *R. palustris* no. 7 within the α -Proteobacteria. Alignment of the deduced amino acid sequences revealed that amino acid residues associated with functional activities in *E. coli* RecA protein (Kawashima *et al.*, 1984) are conserved in *R. palustris* no. 7 RecA. The ATP binding domain (R----G---SGKT) is present in *R. palustris* no. 7 RecA between residues 72 and 85. Domains involved in filament formation and single-stranded DNA binding (Lys21 to Gly41), ATP hydrolysis (Ala103 to Gly120) or λ repressor recognition (Gly216 to Arg239) also exhibit strong homologies. Residues involved in LexA interactions and proteolysis so far identified for *E. coli* protein are also conserved in the *R. palustris* no. 7 RecA polypeptide sequence. A higher degree of divergence between RecA proteins occurred at the carboxy-terminal region of the proteins.

Fig. 1. Nucleotide sequence of the *recA* gene from *R. palustris* no. 7. The deduced amino acid sequence is shown in one-letter code. A putative Shine-Dalgarno sequence is indicated by asterisks. The two putative starting methionines are circled. *E. coli* σ^{70} -like promoter elements are underlined and the transcriptional start nucleotide is indicated by an arrow. An *E. coli* LexA-like binding site is underlined with a dashed line. *R. palustris* no. 7 putative SOS boxes are boxed. Arrows show the sequences used to synthesize the oligonucleotide primers for PCR reactions.

However, the carboxy-terminal regions of the RecA proteins of *R. palustris* no. 7, *R. (S.) meliloti* and *A. tumefaciens* share a high degree of homology.

Construction and characterization of a *recA* derivative of *R. palustris* no. 7

The 1.9 kb *Bam*HI fragment of *R. palustris* no. 7 DNA carrying the *recA* gene was cloned, after blunting, into the *Sal*I site of the mobilizable plasmid pGP704 (Miller & Mekalanos, 1988). Subsequently, the centrally located *Sal*I fragment present within the 1.9 kb *Bam*HI insert was replaced by a kanamycin-resistance cassette to create plasmid pMG307. The reintroduction of the *in vitro*-inactivated *recA* gene was performed by con-

jugation between *E. coli* SM10(*λpir*) containing pMG307 and *R. palustris* no. 7, using a plate mating technique (Inui *et al.*, 1997). Kanamycin-resistant colonies arose at a mean frequency of 1.5×10^{-5} , estimated as the ratio of recipient cells that had received the kanamycin marker to the total number of recipients. No plasmid could be detected as stable minichromosome element in 10 transconjugants tested, indicating that pMG307 was unable to replicate in *R. palustris* no. 7. The antibiotic-resistance marker gene was integrated into the chromosome as a result of recombination between vector and chromosomal homologous DNA following either a single crossover or a double crossover event. In order to discriminate between the two events, colony hybridization analysis was carried out using a pGP704

Fig. 2. Resistance of wild-type *R. palustris* no. 7 and the *recA* derivative to DNA-damaging agents: (a) UV and (b) MMS. ○, Wild-type; ●, *recA* strain.

DNA fragment as probe. Colonies that did not give a hybridization signal were the result of a gene replacement event. Three clones out of 140 were isolated using this procedure. Disruption of the *recA* gene was confirmed by Southern hybridization experiments with a *recA* gene probe (data not shown).

The sensitivity of the *recA*-deficient strain to DNA-damaging agents was determined. The *recA* mutant exhibited increased UV sensitivity compared to the wild-type strain (Fig. 2a). The *recA* strain was also sensitive to MMS: at 0.02% MMS, there were fewer than 1% *recA* survivors whereas there was 100% survival of the wild-type cells (Fig. 2b). The increased sensitivity to DNA-damaging agents of the *recA*-deficient strain compared to the wild-type *R. palustris* no. 7 clearly suggests a role for RecA in DNA repair of *R. palustris* no. 7.

Induction of *recA* transcription by DNA-damaging agents

To analyse the expression of *R. palustris* no. 7 *recA*, a fusion that placed the *lacZ* reporter gene under control of both transcriptional and translational signals of the *recA* gene was constructed in shuttle vector pMG102 (see Methods and Fig. 3a). The resulting plasmid, pMG308, enabled us to study *recA* gene expression in both *R. palustris* no. 7 and *E. coli* host strains.

Plasmid pMG308 was introduced into *R. palustris* no. 7 by electroporation. When this strain was grown on agar medium containing X-Gal, blue colonies developed. By contrast, wild-type *R. palustris* no. 7, as well as *R. palustris* no. 7 containing pMG102, grew as red colonies on this medium. Quantitative assays showed that no β -galactosidase activity was detectable in the latter two strains. *R. palustris* no. 7 harbouring pMG308 had a constant β -galactosidase activity of about 800 Miller units, throughout the exponential growth phase (Fig. 3b). Moreover, the β -galactosidase activity sharply increased upon addition of mitomycin C to the medium

(Fig. 3b). These results demonstrate that the *recA* gene of *R. palustris* no. 7 is inducible by DNA damage.

When transformed into *E. coli* JM109 (*recA*) or JM83 (*recA*⁺), pMG308 gave blue clones, indicating that the *R. palustris* no. 7 *recA* promoter may be functional in *E. coli*. This was further confirmed by a β -galactosidase assay. While no β -galactosidase activity could be measured in strains harbouring pMG102 as a control, the basal level of β -galactosidase activity in JM109 and JM83 harbouring pMG308 was about 2000 Miller units (Fig. 3b). The activity of the promoter in *E. coli* might be explained by the presence of σ^{70} -like promoter elements in the vicinity of the *R. palustris* no. 7 *recA* promoter. Addition of mitomycin C to the medium did not induce the *R. palustris* no. 7 *recA* promoter in any of the *E. coli* strains. The putative *E. coli*-like SOS box is therefore not functional in *E. coli*. The inability of the *E. coli* LexA repressor to bind the *recA* promoter results in constitutive expression that is unaffected by treatment with DNA-damaging agents. Similar data have been reported by Riera *et al.* (1994) for the *R. (S.) meliloti recA* gene. Although an *E. coli*-like SOS box is found 193 nt upstream of the coding region (Selbitschka *et al.*, 1991), the *R. meliloti recA* gene is not inducible in *E. coli*.

Localization of *recA* promoter/operator sequences

The 5' end-point of the *recA* mRNA was mapped by primer extension using RNA templates isolated from uninduced cells or cells treated with mitomycin C. In both cases, the 5' end-point of the *recA* mRNA occurred at a cytosine residue located 55 nt upstream of the putative translation start codon of the *recA* gene (Fig. 4). The amount of transcript initiated at this start-point was drastically increased upon treatment by mitomycin C. The *E. coli*-type σ^{70} -like promoter sequences are located at -15 and -40 with respect to the transcriptional start site.

To further define the extent of the sequences necessary for expression of *R. palustris* no. 7 *recA*, progressive deletions at the 5' end of the *recA-lacZ* fusion were generated by PCR (see Methods and Fig. 3a). β -Galactosidase activity was measured in the absence and in the presence of mitomycin C (Fig. 3c). Deletions made from the 5' side to -145 in pMG318 maintained β -galactosidase levels comparable to those obtained with pMG318, under both uninduced and induced conditions. The 145 bp of DNA upstream of the transcription start site contained in pMG318 was sufficient to confer the same regulation of β -galactosidase activity as observed for the 537 bp DNA fragment contained in pMG308. The LexA-like binding site located at position -247 with respect to the transcriptional start site is therefore dispensable for induction of the *R. palustris* no. 7 *recA* gene by mitomycin C. Deletion up to nucleotide -41 in pMG319 maintained the fivefold induction factor in the presence of mitomycin C. However, the basal level of β -galactosidase was considerably less, perhaps because of the partial deletion of transcription signals at the 5' side

Fig. 3. Expression and deletion analysis of the *recA* promoter. (a) Schematic representation of the 5' upstream region of the *recA* gene with simplified maps of *recA-lacZ* fusions. The appropriate restriction sites, the potential transcriptional initiation site (+1), the -10 and -35 boxes of the *E. coli* σ^{70} -like promoter (at positions -40 and -17) and the ribosome-binding site (RBS) are indicated. DNA fragments fused to *lacZ* in the indicated plasmids are represented by horizontal lines. (b) Expression of the *recA-lacZ* fusion from pMG308 introduced into *R. palustris* no. 7 (○, ●) or into *E. coli* JM109 (▲, △) during growth in rich medium without (open symbols) or with (filled symbols) the addition of mitomycin C at time zero. (c) β -Galactosidase activities generated by the indicated plasmids in *R. palustris* no. 7 following 5 h growth in the absence (grey bars) and presence (black bars) of mitomycin C ($10 \mu\text{g ml}^{-1}$).

of the σ^{70} -like promoter sequence. Promoter activity was eliminated by deletion up to +29 (pMG320). From this set of data, we conclude that the information necessary for *R. palustris* no. 7 *recA* gene expression and regulation by DNA-damaging agents is contained between positions -145 and +29.

Protein binding to the promoter region of *R. palustris* no. 7 *recA*

Gel retardation experiments were carried out to test whether the promoter region of the *R. palustris* no. 7 *recA* gene binds a regulatory protein. Two DNA fragments corresponding to nucleotides -145 to -1 and -62 to -1, respectively, were synthesized by PCR and used in a gel retardation assay with protein extracts from *R. palustris* no. 7. As shown in Fig. 5(a), lanes 2-4, and 5(b), lanes 2 and 3, incubation with cell extract caused a shift in the mobility of both DNA fragments.

To determine the chemical nature of the molecule binding to the DNA fragment, the crude extract was treated with heat, RNase or proteinase K prior to incubation with the larger DNA probe. RNase had no effect on the gel shift (Fig. 5a, lane 9) whereas heat treatment and proteinase K completely eliminated the mobility shift (Fig. 5a, lanes 8 and 10, respectively). We therefore assume that the shift results from interaction of the *recA* promoter DNA sequence with one (or more) polypeptide(s). The specificity of the binding was demonstrated by competition experiments. The unlabelled DNA probe was added in a 50-fold molar excess to the reaction mixture. As seen in Fig. 5(a), lane 11, addition of non-labelled homologous DNA reduced the amount of radioactive DNA that was shifted. Finally, the mobility shift of the probe was not observed when using a protein extract obtained from mitomycin C-induced cells (Fig. 5a, lanes 5 to 7). This result strongly argues in favour of the binding of a repressor to the *recA* promoter region.

Fig. 4. Primer extension experiment to map the 5' end of *recA* mRNA. Equal amounts of RNA isolated from cells harvested after 4 h induction by 10 µg mitomycin C ml⁻¹ (lane 1) and from uninduced cells (lane 2) were used as template for reverse transcriptase. The complement of the DNA sequence illustrated on the sequencing gel is shown on the right. The potential start site of transcription is indicated by an asterisk.

Comparison of *recA* promoter sequences from various α -Proteobacteria

Data from gel retardation experiments clearly indicated that the promoter region between nucleotides -61 and

+1 was the target for the binding of a putative repressor. With the aim of identifying the targeted sequence, we carefully compared the sequence of the *recA* promoter region from *R. palustris* no. 7 with the sequences of *recA* promoter regions from other α -Proteobacteria. As shown in Fig. 6(a), the consensus sequence GAACA-N₆-G(A/T)AC is present within all of the α -proteobacterial *recA* promoter regions examined. Two consensus sequences are present in direct repeat in the promoter regions of both *Rhizobiaceae* and *R. palustris* no. 7 *recA*, whereas they occur inverted with respect to each other in the *recA* promoter region from *Rhodobacter* species (Fig. 6b). Within the *R. palustris* no. 7 *recA* promoter, the two consensus sequences are located at nucleotides -47 to -33 and -29 to -15, therefore overlapping the *E. coli*-type σ^{70} promoter elements. The binding of a protein at these positions would interfere with the initiation of transcription by RNA polymerase, as expected for a repressor.

Mutational analysis of the *R. palustris* no. 7 *recA* promoter and operator

Deletion experiments reported above showed that deletion up to nucleotide -41 in pMG319 maintained the fivefold induction factor in the presence of mitomycin C. Therefore, the 5' half (GAACA) of the consensus located between positions -47 and -33 is unlikely to be involved in mitomycin-C-mediated induction of the *recA* gene. We then performed site-specific mutagenesis of the second consensus sequence lying between nucleotides -29 and -15. Three mutations, GAACA→GTACT in pMG314, AGAACA→CTAGCA in pMG315 and GAACA→GAATA in pMG326, were generated in the 5' half of the consensus and analysed for

Fig. 5. Gel retardation of the *recA* promoter/operator region by protein extracts from *R. palustris* no. 7. (a) 145 bp labelled probe incubated without protein extract (lane 1); with 20, 40 and 60 µg protein extract from uninduced cells (lanes 2, 3 and 4, respectively); with 20, 40 and 60 µg protein extract from mitomycin-C-induced cells (lanes 5, 6 and 7, respectively); with 40 µg protein extract from uninduced cells previously treated by heat (5 min at 95 °C, lane 8), RNase (5 µg, 30 min at 37 °C, lane 9) or proteinase K (10 µg, 30 min at 37 °C, lane 10); or with 40 µg protein extract from uninduced cells and a 50-fold excess of unlabelled probe (lane 11). (b) 62 bp labelled probe incubated without protein extract (lane 1) or with 20 and 40 µg protein extract from uninduced cells (lanes 2 and 3, respectively).

Fig. 6. Consensus sequence found common to all *recA* promoter regions so far examined in species of α -Proteobacteria. (a) DNA sequences of the α -proteobacterial *recA* promoter region. Whenever possible^a, positions are indicated relative to the transcriptional start nucleotide, otherwise^b relative to putative translational initiation. *E. coli* σ^{70} -like promoter elements identified in the *R. palustris* no. 7 *recA* promoter region are underlined. (b) Schematic comparison of the position of the SOS boxes within the α -proteobacterial *recA* promoter region (from -50 to +20). The transcriptional start nucleotide is indicated by a black box. The -10 and -35 sequences are represented by shaded boxes. Position and orientation of the SOS boxes is indicated by white arrows.

their effects on *recA-lacZ* fusion expression (Fig. 7a, b). All three mutations significantly enhanced the basal level of β -galactosidase activity as compared to the wild-type fusion in pMG318. Consequently, the induction factor obtained with these mutants was lower than that obtained with the wild-type fusion (1.3 for pMG315 and pMG326 versus 5 for pMG308). Mutagenesis of this GAACA motif leads to an almost constitutive expression of *R. palustris* no. 7 *recA*, confirming the binding of a repressor at the level of the *R. palustris* no. 7 *recA* promoter, as previously shown in gel mobility shift assays.

During construction of mutant pMG314, we obtained a spontaneous second-site mutation within the -35 promoter element. As shown in Fig. 7(b), this second-site mutation (in pMG322) drastically reduced basal β -galactosidase levels and completely abolished induction by mitomycin C. The same mutation was introduced into pMG315. The double mutant thus obtained (pMG321) also showed a drop in β -galactosidase production.

Attempts to obtain any single mutation within the -35 element of the *recA* promoter were unsuccessful due to

Fig. 7. Effects of mutations in the promoter/operator region of *R. palustris* no. 7 *recA*. (a) Arrows show the locations of specific mutations made within the σ^{70} -like sequence (asterisks) and the consensus sequence (bold letters). (b) β -Galactosidase activities generated by the various mutants in *R. palustris* no. 7 following 5 h growth in the absence (grey bars) and presence (black bars) of mitomycin C ($10 \mu\text{g ml}^{-1}$).

the instability of the clones *in vivo* for unknown reasons. However, a 3 bp substitution in the -35 -like region (TTG $\rightarrow$ AAC in pMG316) reduced β -galactosidase activity by 60%. Similarly, a 1 bp substitution in the -10 -like region (TAC $\rightarrow$ TTC in pMG325) led to a decrease in β -galactosidase levels. Taken together, these results show that the σ^{70} -like sequence is used for the initiation of *R. palustris* no. 7 *recA* transcription. It will require additional experiments to determine whether some nucleotides in the -35 and -10 regions are also involved in DNA-damage-mediated induction of *recA*, as suggested from consensus positions (see Fig. 6). Mutations in these areas are likely to affect both basal expression and induction processes. The use of gel retardation assays with mutated fragments could be an alternative to test this hypothesis.

Conclusions

A *recA*-deficient strain was constructed by using an efficient system for gene replacement in *R. palustris* no. 7 (Inui *et al.*, 1997). The *recA*-deficient strain provides a new tool for the development of a stable host-vector

system in this bacterium. Moreover, this *recA* strain has allowed us to study the function of the *recA* gene of *R. palustris* no. 7. The role of *R. palustris* no. 7 RecA in DNA damage repair was established by showing the increased sensitivity to DNA-damaging agents of the *recA*-deficient strain compared to the wild-type *R. palustris* no. 7. Both primer-extension and gene-fusion experiments further showed that *R. palustris* no. 7 *recA* is inducible upon treatment by DNA-damaging agents. Gel retardation experiments demonstrated a specific interaction between *R. palustris* protein extracts and the *recA* promoter region, an interaction which was not detected using extracts obtained from mitomycin-C-treated cells. This result suggests that a repressor is bound to the *recA* promoter under uninduced conditions. Such a regulatory pattern is similar to the repression of *recA* by LexA regulatory protein in *E. coli* (Walker, 1984). However, deletion experiments clearly showed that the *E. coli*-like SOS box found in the upstream region of the *recA* gene was not involved in repression of *R. palustris recA*. We therefore searched further for the operator sequence responsible for *recA* gene regulation. Sequence comparisons of *recA* promoters from α -Proteobacteria led us to identify the consensus sequence GAACA-N₆-G(A/T)AC. Two such consensus sequences are found within the *R. palustris* no. 7 *recA* promoter region. Interestingly, they overlap with the σ^{70} -like promoter elements such that the binding of a regulatory molecule at these sites could interfere with the initiation of transcription by RNA polymerase. Mutagenesis analysis of the promoter region demonstrated (i) that transcription at *recA* indeed utilizes the σ^{70} -like promoter and (ii) that at least one of the two consensus sequences is the binding site for the repressor regulatory protein. However, we cannot exclude a role of the second sequence in DNA-damage-mediated induction of the *recA* gene. We believe this to be the first characterization of an SOS box for a *Rhodopseudomonas* species.

The GAACA-N₆-GAAC sequence has recently been reported to play a role in regulation of *Rhodopseudomonas sphaeroides recA* by DNA-damaging agents (Fernandez de Henestrosa *et al.*, 1998). A second direct repeat, GTTC-N₇-GTTC, was also found to be involved in regulation of both *recA* and *uvrA* in *R. sphaeroides*, leading the authors to favour this sequence as the genuine SOS box. These same authors had previously proposed that the palindromic sequences TTG-N₁₁-CAA and TTGT-N₁₅-ACAA were the SOS regulatory sequences in *Rhizobium etli* (Tapias *et al.*, 1997) and *R. capsulatus* (Fernandez de Henestrosa *et al.*, 1997), respectively. It is interesting that in both species, the 3' half of the palindrome overlaps the 5' half of the GAACA-N₆-GAAC consensus. Moreover, the authors mentioned that the TTG motif was dispensable for DNA-damage-mediated induction of the *R. etli recA* gene while mutagenesis analysis indicated that *R. capsulatus recA* would be controlled by a positive regulator. These observations do not completely rule out the involvement of the GAACA-N₆-G(A/T)AC

sequence in the regulation of these *recA* genes. It would be of interest to investigate in more detail the role of this sequence in these micro-organisms.

ACKNOWLEDGEMENTS

We are especially grateful to Dr Takaaki Fujii (Chiba University) for the generous gift of *R. palustris* no. 7 strain. We thank Dr Kenneth Zahn and Dr Alain Vertès for fruitful discussions during the course of this work and critical reading of the manuscript.

This work was supported by a grant from the New Energy and Industrial Technology Development Organization.

REFERENCES

- Calero, S., Fernandez de Henestrosa, A. R. & Barbé, J. (1994). Molecular cloning, sequence and regulation of expression of the *recA* gene of the phototrophic bacterium *Rhodobacter sphaeroides*. *Mol Gen Genet* **242**, 116–120.
- Casadaban, M. J., Martinez-Arias, A., Shapira, S. K. & Chou, J. (1983). β -Galactosidase gene fusions for analyzing gene expression in *Escherichia coli* and yeast. *Methods Enzymol* **100**, 293–308.
- Donohue, T. J. & Kaplan, S. (1992). Genetic techniques in Rhodospirillaceae. *Methods Enzymol* **204**, 459–485.
- Duwat, P., Ehrlich, S. D. & Gruss, A. (1992). A general method for cloning *recA* genes of gram-positive bacteria by polymerase chain reaction. *J Bacteriol* **174**, 5171–5175.
- Fernandez de Henestrosa, A. R., Rivera, E. & Barbé, J. (1995). Non-reciprocal regulation of *Rhodobacter capsulatus* and *Rhodobacter sphaeroides* *recA* genes expression. *FEMS Microbiol Lett* **129**, 175–182.
- Fernandez de Henestrosa, A. R., Labazi, M., Mar Lopez, M. & Barbé, J. (1997). Functional analysis of the *recA* promoter of *Rhodobacter capsulatus*. *Mol Gen Genet* **255**, 487–494.
- Fernandez de Henestrosa, A. R., Rivera, E., Tapias, A. & Barbé, J. (1998). Identification of the *Rhodobacter sphaeroides* SOS box. *Mol Microbiol* **28**, 991–1003.
- Fujii, T., Nakazawa, A., Sumi, N., Tani, H., Ando, A. & Yabuki, M. (1983). Utilization of alcohols by *Rhodopseudomonas* sp. No. 7 isolated from *n*-propanol-enrichment culture. *Agric Biol Chem* **47**, 2747–2753.
- Hawley, D. K. & McClure, W. R. (1983). Compilation and analysis of *Escherichia coli* promoter DNA sequences. *Nucleic Acids Res* **11**, 2237–2255.
- Higuchi, R., Krummel, B. & Saiki, R. (1988). A general method of *in vitro* preparation and specific mutagenesis of DNA fragments: study of protein and DNA interactions. *Nucleic Acids Res* **16**, 7351–7367.
- Inui, M., Dumay, V., Zahn, K., Yamagata, H. & Yukawa, H. (1997). Structural and functional analysis of the phosphoenolpyruvate carboxylase gene from the purple nonsulfur bacterium *Rhodopseudomonas palustris* No.7. *J Bacteriol* **179**, 4942–4945.
- Kawashima, H., Horii, T., Ogawa, T. & Ogawa, H. (1984). Functional domains of *Escherichia coli* RecA protein deduced from mutational sites in the gene. *Mol Gen Genet* **193**, 288–292.
- Little, J. W. (1991). Mechanism of specific LexA cleavage: autodigestion and the role of RecA coprotease. *Biochimie* **73**, 411–422.
- Michiels, J., Broek, A. V. & Vanderleyden, J. (1991). Molecular cloning and nucleotide sequence of the *Rhizobium phaseoli* *recA* gene. *Mol Gen Genet* **228**, 486–490.
- Miller, J. H. (1982). *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Miller, R. V. & Kokjohn, T. A. (1990). General microbiology of *recA*: environmental and evolutionary significance. *Annu Rev Microbiol* **44**, 365–394.
- Miller, V. L. & Mekalanos, J. J. (1988). A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. *J Bacteriol* **170**, 2575–2583.
- van Niel, C. B. (1944). The culture, general physiology, morphology and classification of the non-sulfur purple and brown bacteria. *Bacteriol Rev* **8**, 1–118.
- Riera, J., Fernandez de Henestrosa, A. R., Garriga, X., Tapias, A. & Barbé, J. (1994). Interspecies regulation of the *recA* gene of gram-negative bacteria lacking an *E. coli*-like SOS operator. *Mol Gen Genet* **245**, 523–527.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Selbitschka, W., Arnold, W., Priefer, U. B., Rottschäfer, T., Schmidt, M. & Pühler, A. (1991). Characterization of the *recA* genes and *recA* mutants of *Rhizobium meliloti* and *Rhizobium leguminosarum* biovar *viciae*. *Mol Gen Genet* **229**, 86–95.
- Tapias, A., Fernandez de Henestrosa, A. R. & Barbé, J. (1997). Characterization of the promoter of the *Rhizobium etli* *recA* gene. *J Bacteriol* **179**, 1573–1579.
- Vieira, J. & Messing, J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene* **19**, 259–268.
- Vieira, J. & Messing, J. (1987). Production of single-stranded plasmid DNA. *Methods Enzymol* **153**, 3–11.
- Walker, G. C. (1984). Mutagenesis and inducible responses to deoxyribonucleic acid damage in *Escherichia coli*. *Microbiol Rev* **48**, 60–83.
- Wardhan, H., McPherson, M. J., Harris, C. A., Sharma, E. & Sastry, G. R. K. (1992). Molecular analysis of the *recA* gene of *Agrobacterium tumefaciens* C58. *Gene* **121**, 133–136.
- Wilkinson, M. (1991). Purification of RNA. In *Essential Molecular Biology: a Practical Approach*, vol. 1, p. 71. Edited by T. A. Brown. New York: Oxford University Press.
- Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains; nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.

Received 5 November 1998; accepted 19 January 1999.