Journal of Applied Phycology **8**: 563–564, 1997.

Gene bulletin

Recombination protein A gene, *recA*, from *Spirulina platensis* IAM-M135

Ajay K. Vachhani & Avigad Vonshak*
Microalgal Biotechnology Laboratory, The Jacob Blaustein Institute for Desert Research, Ben-Gurion University of the Negev, Sede-Boker Campus, 84990 Israel. (Author for correspondence: fax 972-6570198; email: avigad@bgumail.bgu.ac.il)

Received 3 September 1996; revised and accepted 5 September 1996

Database accession number: U33924

Key words: Cyanobacteria, gene-transfer, recombination, *recA*, *Spirulina platensis*

Abstract

We report the cloning and sequencing of the *recA* gene from *Spirulina platensis*. A genomic library of *Spirulina* was constructed in pUC19 and screened by PCR using oligonucleotides corresponding to the conserved amino acid sequences of *Anabaena variabilis* and *Synechococcus* RecA proteins. The *Spirulina recA* gene consists of an open reading frame (ORF) of 1095 nucleotides encoding a protein (365 residues) which shares an identity of 79%, 70% and 57% with the RecA proteins of *Anabaena variabilis, Synechococcus* and *Escherichia coli* respectively. The *recA* gene is located close to one end of the cloned *Bgl*II fragment and has only 53 bp of 5′ nucleotides. The isolation of this gene has implications for the development of gene transfer system(s) for *Spirulina*.

Results and discussion

A single hybridising band of approximately 9.0 kbp was obtained when *Bgl*II digested *Spirulina platensis* genomic DNA was probed with a radioactively labeled *recA* probe from *Synechococcus*. Genomic DNA digested with *Bgl*II was ligated to *Bam*HI digested pUC19. The clones were screened by carrying out PCR with two consensus *recA* primers (A: 5′ CTCCATGCGATCGCCGAAGT 3′ and B: 5′ GGTTTGGATGCGGCGGATATCTA 3′) which were based on the conserved amino acid sequences LHAIAEV and LDIRRIQT in the *Anabaena variabilis* and *Synechococcus* RecA proteins. A putative clone containing an insert of 9 kbp was obtained and automated DNA sequencing initiated with the primers described above resulted in the identification of an open reading frame (ORF) of 1095 nucleotides followed by a stop codon.

The *Spirulina recA* gene is located towards one end of the cloned *Bgl*II fragment (close to the *Sma*I site in the pUC19 polylinker) and has only 53 bp of DNA sequence upstream of the gene. A ribosome binding site having the typical bacterial ribosome binding sequence of 5′ AGGA 3′ is located 14 bp upstream of the start codon. A survey of *E. coli* σ70-dependent consensus promoters has confirmed the prevalence of two consensus hexamer sequences upstream of the transcription start sites at position +1: the sequence TTGACA centred at 35 bp upstream (−35) and the sequence TATAAT centred at 10 bp upstream (−10). Although no base within either hexamer motif is invariant among all promoters of the database, the first three positions of the −35 element (TTGNNN) and three positions in the −10 element (TANNNT) are conserved in 80–90% of the promoters. Comparisons of *E. coli* and cyanobacterial promoters have shown that the most highly conserved nucleotides of the *E. coli* −10 hexamer (TANNNT) are conserved in >70% of the cyanobacterial sequences (Curtis & Martin, 1995). The *Spirulina recA* gene has a sequence similar to the −10 conserved sequence at 50 bp upstream of the start codon. This sequence TATTTT matches the consensus −10 element in three of the most important positions. Thus the cloned *Spirulina recA* gene lacks a −35 sequence and this may be the reason why it is unable to functionally complement *E. coli recA* mutants.

The *S. platensis recA* gene has an identity of 72%, 68% and 60% when compared to the *recA* genes of *Anabaena variabilis* (Owttrim & Coleman, 1989) *Synechococcus* PCC 7002 (Murphy et al., 1990) and *Escherichia coli* (Sancar et al., 1980). Comparison of the amino acid sequences reveals an identity of 79%, 70% and 57% with the RecA proteins of *Anabaena variabilis*, *Synechococcus* and *Escherichia coli* respectively.

Alignment of the deduced amino acid sequence revealed that many of the domains and residues assigned functional roles in *E. coli* RecA are conserved in the *Spirulina* RecA. Numbering of the *E. coli* RecA residues begins with the first alanine (Dunkin & Wood, 1994). The ATP fold (R.....G...SGKT) (Walker et al., 1982) is present in *Spirulina* RecA between amino acids 62 and 77. This region coincides with the ATP binding site which has been proposed by Kawashima et al. (1984). However Tyr-264 identified by Knight and McEntee (1985) as the residue responsible for ATP binding is not present in the *Spirulina* sequence. Other *E. coli* residues associated with protease activity (Ser-25, Glu-158 and Gly-157) are conserved in nearly the same positions in the *Spirulina* RecA protein (Ser-24, Glu-159 and Gly-160). Some of the *E. coli* residues associated with homologous recombination (Gly-160, Leu-51, Arg-60 and Gly-301) are also similarly conserved (Gly-160, Leu-51, Arg-61 and Gly-302). The *Spirulina* RecA protein has retained the highly conserved proline residues observed by Tolmasky et al. (1992), the amino acids which constitute the RecA signature peptide, ALKF(F,Y)(S,T,A)$_2$VR, a sequence which has been defined as RecA specific by the program Motifs in the Genetics Computer Group package of programs (Devereux et al., 1984), and the conserved Arg-324 (Arg-325 in *Spirulina* RecA) in the highly variable C-terminal region (Dunkin & Wood, 1994).

Although mass production of *Spirulina* has reached a commercial stage and the annual production of *Spirulina* biomass is estimated to be some 1000 tons per year very little is known about the genetics of these cyanobacteria. Despite many attempts to develop a gene transfer system for *Spirulina* there has not been any significant progress. We are presently working on the development of a gene transfer system and the isolation and characterisation of *Spirulina recA* gene may significantly assist in a better understanding of the genetics of these cyanobacteria.

Acknowledgements

We wish to thank Donald Bryant for the gift of plasmids and gratefully acknowledge helpful discussions with Ali Nejidat, Simon Barak and Srinivas Garlapati. This work was supported by US AID CDR program grant TA-MOU-95-C15-003. Publication No. 82 of the Microalgal biotechnology laboratory.

References

Curtis SE, Martin JA (1995) The transcription apparatus and regulation of transcription initiation. In Donald B (ed.) The molecular biology of cyanobacteria. Kluwer Academic Publishers, Dordrecht: 613–639.

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12: 387–395.

Dunkin SM, Wood DO (1994) Isolation and characterization of the *Rickettsia prowazekii recA* gene. J. Bact. 176: 1777–1781.

Kawashima H, Horii T, Ogawa T, Ogawa H (1984) Functional domains of *Escherichia coli* RecA protein deduced from the mutational sites in the gene. Mol. Gen. Genet. 193: 288–292.

Knight KL, McEntee K (1985) Tyrosine 264 in the RecA protein from *Escherichia coli* is the site of modification by the photoaffinity label 8-azidoadenosine 5′ triphosphate. J. Biol. Chem. 260: 10185–10191.

Murphy RC, Gasparich GE, Bryant DA, Porter RD (1990) Nucleotide sequence and further characterisation of the *Synechococcus* sp. strain PCC 7002 *recA* gene: Complementation of a cyanobacterial *recA* mutation by the *Escherichia coli recA* gene. J. Bact. 172: 967–976.

Owttrim GW, Coleman JR (1989) Regulation of expression and nucleotide sequence of the *Anabaena variabilis recA* gene. J. Bact. 171: 5713–5719.

Sancar A, Stachelek C, Konigsberg W, Rupp WD (1980) Sequences of the *recA* gene and protein. Proc. Natl. Acad. Sci. USA 77: 2611–2615.

Tolmasky M, Gammie AE, Crosa JH (1992) Characterisation of the *recA* of *Vibrio anguillarum*. Gene 110: 41–48.

Walker JE, Saraste M, Runswick MJ, Gay NJ (1982) Distantly related sequences in the α- and β-subunits of ATP synthase, myosin kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. 1: 945–951.