IncI1 plasmid R64 encodes the ArsR protein that alleviates type I restriction

Sergei M. Rastorguev[a], Gennadii B. Zavilgelsky[a,*], Nickolai A. Tchurikov[b]

[a] *State Scientific Centre of Russian Federation GNIIGENETIKA, 1th Dorozhnii pr. 1, Box 825, Moscow 113545, Russia*
[b] *Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilov str. 32, Moscow 117984, Russia*

Received 20 January 1998; revised version received 5 February 1998

Abstract The host-controlled *Eco*K restriction of unmodified phage λ was five-fold alleviated in the wild-type *Escherichia coli* strain K12 carrying the R64 plasmid of the incompatibility group I1. The relevant gene was mapped between the origin of vegetative replication (*rep*, *ori*V) and the *tet*^{r} gene about 60 kbp downstream from the origin of transfer, *ori*T. We cloned this gene inside the 613 bp long *Eco*RI-*Pst*I fragment and sequenced it. Only one 351 bp long open reading frame (ORF) starting at 124 bp from the beginning of the insert was found in the sequence. Computer search in the current databases revealed that the putative protein is identical to the ArsR protein specified by the IncFI plasmid R773. ArsR is a repressor of the arsenical resistance (*ars*) operon, *ars*RDABC. There are no *ars*ABC genes in the R64 plasmid since plasmid R64- (or pSR8)-mediated resistance of *E. coli* K12 cells to the arsenicals arsenate and arsenite was not detected. The gene *ars*R and the antirestriction genes *ard* (*ard*A and *ard*B) are non-homologous. However, comparison of the deduced amino acid sequence of ArsR with the ArdA and ArdB sequences revealed only one small region of similarity, a 9 amino acid motif found in different antirestriction proteins that is hypothesized to be an interaction site for antirestriction proteins with restriction endonucleases.

Key words: Self-transmissible plasmid; Antirestriction; *ard* gene; *ars* operon; 'Antirestriction motif'

1. Introduction

Self-transmissible plasmids of the IncI1, B/O, K, FV and N incompatibility groups were shown to be able to weaken the action of type I restriction enzymes upon non-modified DNA (Ard phenotype) [1,2]. The antirestriction genes *ard*AB were mapped in the plasmids in the leading regions 4–9 kbp downstream from *ori*T and were shown to exhibit the specificity of inhibition of type I restriction enzymes [2–6].

In this paper we show that the IncI1 self-transmissible plasmid R64 [7,8] carries an *ard*-type gene that efficiently prevents *Eco*K restriction in *Escherichia coli* K12. The location of this gene, its nucleotide sequence, and some aspects of its function are reported. Comparison of the deduced amino acid sequence of the antirestriction protein of plasmid R64 and ArsR protein (the repressor of the *ars*RDABC operon) of IncFI plasmid R773 [9] revealed that these proteins have 100% identity. The ArsR and Ard (ArdA and ArdB) proteins are non-homologous, but share a small region of homology containing nine amino acid residues, the 'antirestriction motif'. Currently the *ars* genes in two self-transmissible plasmids, R773 (IncFI) and R46 (IncN), are known [9,10]. However, there were no data about the location of the *ars* genes in the R64 plasmid.

2. Materials and methods

2.1. Bacterial strains and media

AB1157 is an F$^-$ derivative of *E. coli* K12 having the genotype *thr*-1 *leu*-6 *pro*A2 *his*-4 *thi*-1 *arg*E3 *lac*Y1 *gal*K2 *ara*-14 *xyl*-5 *mtl*-1 *tsx*-33 *rps*L31 *sup*E44; Ab2463 is the *rec*A13 derivative of AB1157. JM101 *thi sup*E44 Δ(*lac-pro*AB) F′ (*tra*D36 *pro*AB$^+$ *lac*I^q *lac*ZΔM15); TG1 *hsd*R17 *hsd*M *thi rel*A1 *sup*E44 Δ(*lac-pro*AB) F′ (*tra*D36 *pro*AB$^+$ *lac*I^q *lac*ZΔM15). *E. coli* C $r_o m_o$.

The bacteriophage referred to as λ was λ_{vir} kindly provided by Dr. R. Devoret, France. Unmodified phage, denoted λ.0, was grown on *E. coli* C $r_o m_o$ lacking restriction and modification functions. Modified phage, denoted as λ.K, was grown on the $r_k^+ m_k^+$ strain AB1157.

2.2. DNA manipulations and plasmid constructions

Standard procedures were used [11].

2.3. DNA sequence analysis

The nucleotide sequence of the 613 bp *Eco*RI-*Pst*I fragment from plasmid R64 was determined by the dideoxy chain termination method [12].

2.4. Measurement of antirestriction activity

The antirestriction activity of plasmids was defined as the ratio of the efficiency of plating (EOP) of unmodified phage λ.0 on the experimental (plasmid-bearing) strain to the EOP on the control (plasmidless) restriction strain [5].

3. Results and discussion

3.1. Cloning of the R64 fragments

Previously, we found that *Eco*K restriction of unmodified phage λ.0 was five-fold alleviated in the *E. coli* K12 strains harboring plasmid R64 [5] (Table 1). The locus specifying the function of alleviation of *Eco*K restriction was identified by subcloning various *Eco*RI restriction fragments of R64 in the pUC18 vector (Fig. 1). Recombinant plasmid pSM8 carrying the 9.3 kbp *Eco*RI-D fragment efficiently (20-fold) prevented *Eco*K restriction of phage λ.0 in *E. coli* K12 (Table 1). This increase in the antirestriction activity presumably reflects an increased copy number of the gene after subcloning in a multicopy vector. Subsequent subcloning allowed us to limit the antirestriction locus in a 0.6 kbp *Eco*RI-*Pst*I fragment in the recombinant plasmid pSR13 that also efficiently (20-fold) prevented *Eco*K restriction of phage λ.0 in *E. coli* K12 (Fig. 1 and Table 1). The IPTG inducer strengthens the antirestriction effect of plasmid pSR13 about three-fold (Table 2), suggesting the orientation of the antirestriction gene with respect to the *lac* promoter from the *Eco*RI to *Pst*I sites.

3.2. Nucleotide sequence of the R64 antirestriction gene

The nucleotide sequence of the antirestriction locus was

*Corresponding author. Fax: (7) (095) 315-05-01.
E-mail: lab2@vnigen.msk.su

PII S0014-5793(98)00285-3

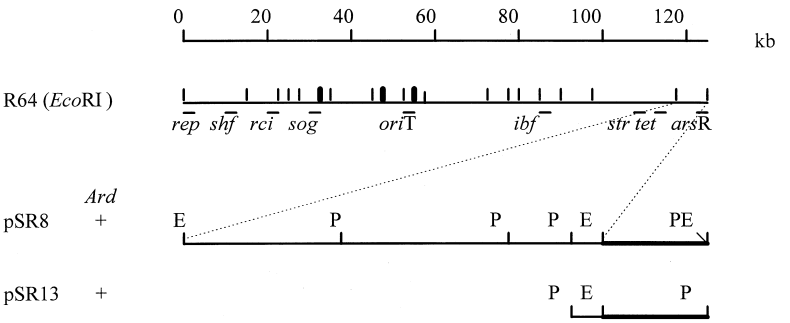

Fig. 1. Restriction map of recombinant plasmids carrying the R64 *ars*R locus. The top line indicates kilobase coordinates on the R64 map. The lower lines show restriction sites and the location of the origin of transfer (*ori*T), *rep* region, *sog, ibf, rci, tet, str,* and *ars*R genes. pUC18 vector is shown by the thick line. Transcription of *ars*R is from right to left. Restriction sites are: E, *Eco*RI; P, *Pst*I.

```
GAATTCCAAGTTATCTCACCTACCTTAAGGTAATAGTGTGATTAATCATA  50
          G      -35                  -10
TGCGTTTTTGGTTATATGTTGTTTGACTTAATATCAGAGCCGAGAGATAC 100
          SD
TTGTTTTCTACAAAGGAGGGAAAATGTTGCAACTAACACCACTTCAGTTA 150
                   M  L  Q  L  T  P  L  Q  L
TTTAAAAACCTGTCCGATGAAACCCGTTTGGGTATCGTGTTGTTGCTCAG 200
 F  K  N  L  S  D  E  T  R  L  G  I  V  L  L  L  R
GGAGATGGGAGAGTTGTGCGTGTGTGATCTTTGCATGGCACTGGATCAAT 250
  E  M  G  E  L  C  V  C  D  L  C  M  A  L  D  Q
CACAGCCCAAAATATCCCGTCATCTGGCGATGCTACGGGAAAGTGGAATC 300
S  Q  P  K  I  S  R  H  L  A  M  L  R  E  S  G  I
CTTCTGGATCGTAAACAGGGAAAATGGGTTCACTACCGCTTATCACCGCA 350
 L  L  D  R  K  Q  G  K  W  V  H  Y  R  L  S  P  H
TATTCCTTCATGGGCTGCCCAGATTATTGAGCAGGCCTGGTTAAGCCAAC 400
  I  P  S  W  A  A  Q  I  I  E  Q  A  W  L  S  Q
AGGACGACGTTCAGGTCATCGCACGCAAGCTGGCTTCAGTTAACTGCTCC 450
Q  D  D  V  Q  V  I  A  R  K  L  A  S  V  N  C  S
GGTAGCAGTAAGGCTGTCTGCATCTAAAAAATTTGCCTGAACATATATGT 500
 G  S  S  K  A  V  C  I
TTTATCAAATGCGAGGTATTTAAGATGAAAACGTTAATGGTATTTGACCC 550

GGCGATGTGTTGCAGCACCGGCGTCGCGGTACAGATGTTGATCAGGCTCT 600
          A
GGTCGATTTTTCTGCAG                                   617
```

Fig. 2. Nucleotide sequence of the *Eco*RI-*Pst*I fragment of R64. The sequence numbering begins at the *Eco*RI site located upstream of *ars*R and extends to the *Pst*I site. The deduced amino acid sequence (ORF) is shown. The potential −10 and −35 sequences and the putative Shine-Dalgarno site (SD) are underlined. The initiation ATG codon and the TAA stop codon are shown in bold. Several deviations are shown in bold, top symbols belong to R773. Other indications are defined in the text.

determined for both strands. The nucleotide sequence of the plus strand is shown in Fig. 2. Only one rather long ORF spacing from 124 to 475 bp was found in the sequence, which presumably encodes a polypeptide of 117 amino acid residues. The starting ATG codon is located at nucleotide 124 and follows 4 bp downstream of a strong Shine-Dalgarno GGAGG sequence. Upstream of the ORF, the ATACTT hexamer (at position −10) and the TTGACTT sequence (at position −35) are located at 97–102 and 73–80 bp, respectively.

We searched for the nucleotide sequence of the R64 *Eco*RI-*Pst*I fragment in the current nucleotide databases using the Blast 2.0 program. We have observed that our sequence is practically identical to the nucleotide sequence of the *ars*R gene from the R773 plasmid (incFI) [9] (see Fig. 2). It is known that the *ars* operon (*ars*RDABC genes) in *E. coli* R773 plasmid confers high-level resistance to arsenicals and antimonials [13,14]. The product of the first gene of the operon, the 117 residue ArsR protein, has been shown to be a *trans*-acting repressor that is a sensor of environmental As(III) and Sb(III) [14]. This 13 kDa protein belongs to the ArsR family of metalloregulatory proteins that respond to a variety of metals including As(III), Sb(III), Cd(II), and Zn(II) [15]. However, the basic *ars*ABC genes are not present in the R64 plasmid, because this plasmid and the recombinant pSR8 plasmid do not determine high resistance of *E. coli* cells to *meta*-arsenite (data not shown).

3.3. Comparison of ArsR and Ard proteins

The nucleotide sequences of *ars*R and *ard*A or *ard*B genes are non-homologous. However, ArsR, ArdA and ArdB proteins, which are also non-homologous, reveal a functional similarity in inhibition of *Eco*K (type I) restriction. Analysis of the deduced amino acid sequence of ArsR revealed that in contrast to the strongly acidic ArdA (an excess of 25–29 negatively charged amino acids) and the slightly acidic ArdB (an excess of only seven acidic residues) [2,6,16], ArsR is charged slightly positively (+2). The antirestriction proteins (ArdA, ArdB and T7 phage 0.3) share only one small region of homology containing nine amino acid residues ('antirestrictase motif') [6]. Interestingly, this region is also conserved in the R64 ArsR protein (Table 3). In contrast to Ard proteins, in ArsR the 'antirestrictase motif' is located at the N-terminus of the polypeptide. However, the N-terminal part of the ArdA protein (90–95 amino acid residues from 166) is non-essential for Ard activity [16]. Similarly, the C-terminal 27 amino acid

Table 1
Effects of antirestriction locus of R64 on *Eco*K restriction in *E. coli* K12 JM101 r^+m^+

Plasmid	EOP of test phage λ.0, K_+[a]	Relief of restriction, R[b]	Antirestriction phenotype
pUC18	2.0×10^{-4}	1	−
R64	1.0×10^{-3}	5	+
pSR8	4.0×10^{-3}	20[c]	+
pSR13	4.0×10^{-3}	20	+

[a]The EOP (K) of unmodified test phage λ.0 was determined as the ratio of phage titer on the restricting strain *E. coli* JM101 to the titer on the non-restricting strain *E. coli* TG1. K_- was determined on the JM101 strain without plasmid, K_+ on the same strain carrying plasmid.
[b]$R = K_+/K_-$.
[c]On the AB1157 strain with pSR8 plasmid R=70–90. In all cases the plating efficiency of modified phage λ.K was equal to 1.

Table 2
Effects of the inducer (IPTG) on the pSR13-induced alleviation of *Eco*K restriction of phage λ.0 in *E. coli* K12 JM101[a]

Plasmid	EOP of test phage λ.0, K_+		Relief of restriction, R	
	−IPTG	+IPTG	−IPTG	+IPTG
pUC18	2.0×10^{-4}	2.0×10^{-4}	1	1
pSR13	4.0×10^{-3}	1.2×10^{-2}	20	60

[a]See footnote to Table 1.

Table 3
'Antirestriction motif' in amino acid sequences of ArsR and antirestrictase proteins

Gene	Plasmid	Sequence	Reference
*ard*A	ColIb-P9 (IncI1)	130-**LL**A**D**VP**E**TV	[16]
*ard*A	R16 (IncB)	131-**LL**A**D**VP**E**TV	[2]
*ard*A	pKM101 (IncN)	131-**LL**N**E**IP**E**SV	[18]
*ard*A	F_{olac} (IncFV)	132-**LL**N**E**VP**E**PL	[2]
*ard*B	pKM101 (IncN)	120-**LL**R**E**YV**E**TL	[6]
*ars*R	R64 (IncI1)	24-**LL**R**E**MG**E**LC	This study
ocr (0.3)	phage T7	101-**LL**N**E**YL**E**EV	[19]

Identical amino acids and negatively charged Asp and Glu (D and E) are shown in bold.

Table 4
Effects of arsenite on the pSR13-induced alleviation of *Eco*K restriction in *E. coli* K12 JM101[a]

Plasmid	EOP of test phage λ.0		Relief of restriction, R	
	−arsenite	+arsenite	−arsenite	+arsenite
pUC18	2×10^{-4}	2×10^{-4}	1	1
pSR13	1×10^{-2}	2×10^{-3}	50	10

[a]See footnote to Table 1.

residues in ArsR protein are not essential for repressor activity [17].

The ArsR sequence 30-ELCVCDL has been proposed to form a portion of the metal As(III) binding domain [20]. This domain partly overlaps with the antirestriction domain (Fig. 2). When binding with As(III), ArsR loses activity as a repressor of the *ars* operon [20]. Therefore we propose that the binding of As(III) with ArsR could lead to the loss of both repressor and antirestriction activities. As seen in Table 4, the antirestriction activity of ArsR on a medium containing *meta*-arsenite is decreased about five-fold. We cannot increase the concentration of arsenite in the medium because it would be lethal for the *E. coli* cells.

ArsR protein has a moderate Ard activity. In the pUC18 vector in the normal orientation with respect to the *lac* promoter and in the presence of IPTG, the *ars*R-induced alleviation of *Eco*K restriction of unmodified phage λ.0 is up to 50–100-fold (Table 2). In the same conditions of cloning, the *ard*A genes exhibit alleviation of the *Eco*K restriction of λ.0 up to 6000–8000-fold [2,6,16].

A number of arsenic/antimony resistance (*ars*) operons have been identified from both Gram-negative and Gram-positive bacterial plasmids [21,22]. All of them share a common organization and have the gene specifying the regulatory ArsR protein. We suppose that the metalloregulatory proteins could possess antirestriction activity indispensable for the presence of the 'antirestriction motif' in the amino acid sequence.

Acknowledgements: This work was supported by the Russian Foundation for Basic Research (Grants 97-04-49897 and 96-15-97779).

References

[1] Zavilgelsky, G.B., Bakalova, T.L., Duzhii, D.E. and Kotova, V.Yu. (1994) Russ. J. Genet. 30, 1582–1586.
[2] Chilley, P.M. and Wilkins, B.M. (1995) Microbiology 141, 2157–2164.
[3] Zavilgelsky, G.B., Mershavka, V.Yu., Yussiffov, T.N. and Belogurov, A.A. (1984) Mol. Biol. 18, 1590–1596.
[4] Belogurov, A.A., Yussiffov, T.N., Kotova, V.Yu. and Zavilgelsky, G.B. (1985) Mol. Gen. Genet. 198, 509–513.
[5] Kotova, V.Yu., Zavilgelsky, G.B. and Belogurov, A.A. (1988) Mol. Biol. 22, 270–276.
[6] Belogurov, A.A., Delver, E.P. and Rodzevitch, O.V. (1993) J. Bacteriol. 175, 4843–4850.
[7] Furuichi, T., Komano, T. and Nisioka, T. (1984) J. Bacteriol. 158, 997–1004.
[8] Komano, T., Funayama, N., Kim, S.-R. and Nisioka, T. (1990) J. Bacteriol. 172, 2230–2235.
[9] San-Francisco, M.J.D., Hope, C.L., Owolabi, J.B., Tisa, L.S. and Rosen, B.P. (1990) Nucleic Acids Res. 18, 619–624.
[10] Bruhn, D.F., Li, J., Silver, S., Roberts, F. and Rosen, B.P. (1996) FEMS Microbiol. Lett. 139, 149–153.
[11] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[12] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5468.
[13] Wu, J. and Rosen, B.P. (1991) Mol. Microbiol. 5, 1331–1336.
[14] Chen, C.-M., Mistra, T.K., Silver, S. and Rosen, B.P. (1986) J. Biol. Chem. 261, 15030–15038.
[15] Xu, C., Shi, W. and Rosen, B.P. (1996) J. Biol. Chem. 271, 2427–2432.
[16] Delver, E.P., Kotova, V.Yu., Zavilgelsky, G.B. and Belogurov, A.A. (1991) J. Bacteriol. 173, 5887–5892.
[17] Xu, C. and Rosen, B.P. (1997) J. Biol. Chem. 272, 15734–15738.
[18] Belogurov, A.A., Delver, E.P. and Rodzevitch, O.V. (1992) J. Bacteriol. 174, 5079–5085.
[19] Dunn, J.J., Elzings, M., Mark, K.K. and Studier, F.W. (1981) J. Biol. Chem. 256, 2579–2585.
[20] Shi, W., Wu, J. and Rosen, B.P. (1994) J. Biol. Chem. 269, 19826–19829.
[21] Wu, J. and Rosen, B.P. (1993) J. Biol. Chem. 268, 52–58.
[22] Neyt, C., Iriarte, M., Thi, V.H. and Cornelis, G.R. (1997) J. Bacteriol. 179, 612–619.