

A motif conserved among the type I restriction-modification enzymes and antirestriction proteins: a possible basis for mechanism of action of plasmid-encoded antirestriction functions

Anatol A. Belogurov* and Eugene P. Delver

Department of Genetic Engineering, Cardiology Research Center, Moscow 121552, Russia

Received November 14, 1994; Revised and Accepted January 19, 1995

ABSTRACT

Antirestriction proteins Ard encoded by some self-transmissible plasmids specifically inhibit restriction by members of all three families of type I restriction-modification (R-M) systems in *E. coli*. Recently, we have identified the amino acid region, 'antirestriction' domain, that is conserved within different plasmid and phage T7-encoded antirestriction proteins and may be involved in interaction with the type I R-M systems. In this paper we demonstrate that this amino acid sequence shares considerable similarity with a well-known conserved sequence (the Argos repeat) found in the DNA sequence specificity (S) polypeptides of type I systems. We suggest that the presence of these similar motifs in restriction and antirestriction proteins may give a structural basis for their interaction and that the antirestriction action of Ard proteins may be a result of the competition between the 'antirestriction' domains of Ard proteins and the similar conserved domains of the S subunits that are believed to play a role in the subunit assembly of type I R-M systems.

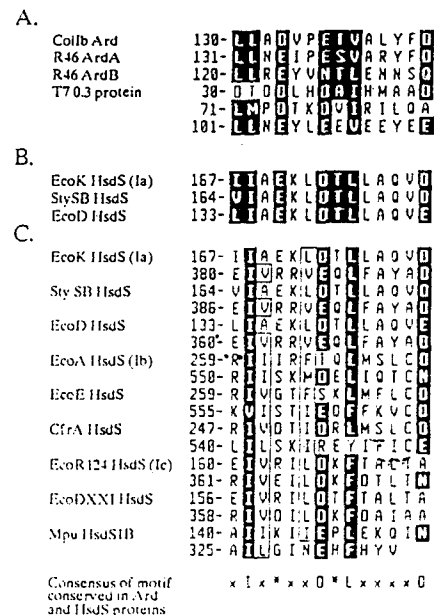


Figure 1. A motif conserved in the antirestriction proteins and DNA specificity subunit (HsdS) of type I R-M systems. Deduced amino acids from antirestriction proteins ('antirestriction' domain) (A) and HsdS polypeptides (the Argos repeat-like sequences) (B and C) were aligned using the CLUSTAL program. The sequences used for alignment were ColIb Ard (6), R46 ArdA (8), R46 ArdB (9), T7 0.3 protein (12), EcoK HsdS (18), EcoD HsdS (18), EcoA HsdS (30), EcoE HsdS (30), CfrA HsdS (17), EcoR124 HsdS (31), EcoDXXI HsdS (GenBank accession number X73984) and Mpu HsdS1B (32). The families to which the type I enzymes belong are indicated in parentheses. Note that the appropriate sequences of type Ia systems EcoB (18) and StySP (19) are identical to that of EcoK and therefore not shown. Black boxes represent similar (in one-letter notation, A.S.T; D.E.N.Q; H.K.R; F.L.L.M.V; Y.W) and identical amino acids in antirestriction proteins and their conservation in the HsdS polypeptides. Open boxes represent amino acid residues conserved only in the HsdS polypeptides. The numbering of residues of each polypeptide is shown on the left. Two Argos repeat-like sequences presented for each HsdS are located within its central (first line) and carboxyl (second line) conserved regions, respectively (see Fig. 2 for details). Symbols * and x in the consensus sequence represent polar amino acid residues and lack of consensus, respectively.

INTRODUCTION

Self-transmissible plasmids play an important role in the dissemination of variety of genes among different bacterial populations. However, DNA transfer mediated by these plasmids may be limited by host 'immigrant control', restriction endonucleases that recognize the 'molecular passport', the methylation pattern of guest DNA (see, for recent reviews 1-4). To overcome the host restriction barrier, some self-transmissible plasmids were found to encode antirestriction functions Ard (alleviation of restriction of DNA) (5-9). Recently, we have found that the broad-host-range IncN plasmid R46 encodes at least two types of antirestriction proteins, ArdA and ArdB. These proteins are non-homologous but functionally similar in that both efficiently inhibit restriction by members of all three families of type I systems in *E. coli* (EcoK, EcoB, EcoD, EcoA, and EcoR124). In addition, they slightly affect the type II (EcoRI) restriction and do

* To whom correspondence should be addressed

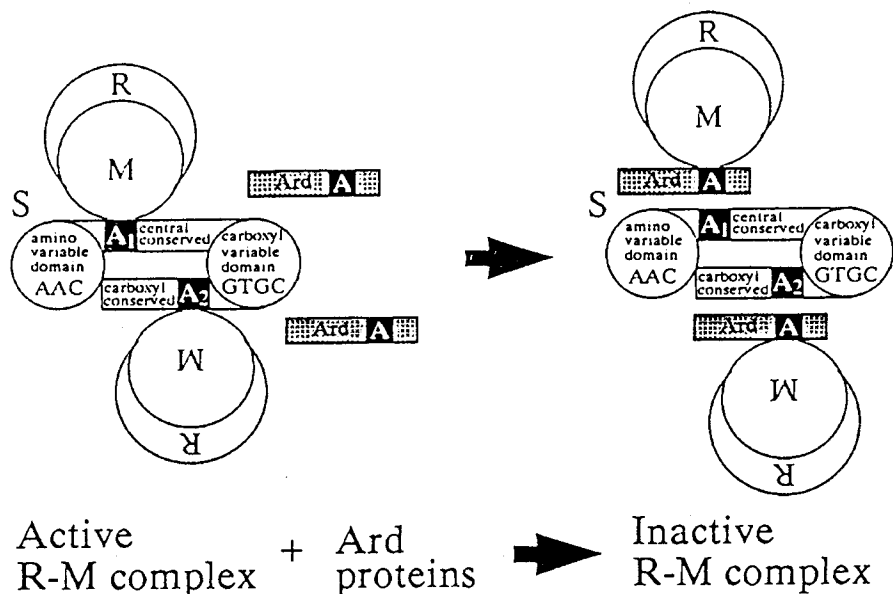


Figure 2. Schematic representation of proposed mode of interaction between antirestriction proteins Ard and type I R-M systems. Mechanism of this interaction is based on the sequence similarity of conserved elements (closed boxes) of Ard proteins ('antirestriction' domain, box A) and S subunits (the Argos repeat-like sequences, boxes A₁ and A₂) and their competition for the interaction with M subunits. The type I enzymes consist of S, M and R subunits that are responsible for DNA recognition, DNA methylation and DNA cutting, respectively. The domain structure of S subunit of *EcoK* common also for R-M enzymes of other families is presented (see 22,23 for details). This subunit contains two target recognition domains (open circles) and recognizes the sequence AAC(N)₆GTGC. These domains are separated by small regions (open bars) containing the short repeated sequences (the Argos repeats, closed boxes A₁ and A₂) that are conserved between S subunits of different type I systems and seem to be involved in interfaces with at least the M subunit (18–23).

not influence either the type III (*EcoP1*) restriction or the 5-methylcytosine specific restriction systems McrA and McrBC (8,9). It has been shown that unrelated *IncI* plasmid ColIb-P9 also encodes the antirestriction Ard-type protein homologous (about 60% identity) to R46 ArdA and that other *IncN* and *IncI* plasmids also express antirestriction functions (6–8).

A question arises about the mechanism of action of Ard-type antirestriction proteins. We suggest that the antirestriction activity of Ard proteins (i) may interact directly with the restriction-modification (R-M) complex or (ii) may interfere or compete with the restriction-modification system binding to its recognition sites.

Analysis of deduced amino acid sequence revealed that homologous R46 ArdA and ColIb Ard proteins are strongly acidic (6,8) and their binding to DNA seems, therefore, unlikely. We have suggested that the acidic Ard proteins might act as a well-studied acidic antirestriction *O.3* protein of T7 phage (Ocr protein) (2,10–15) and inhibit R-M systems by direct binding to the enzymes (6,8). The studies with purified proteins have shown that this phage protein appears to act as a polyanion that inactivates the R-M complex by binding near its DNA binding site (11,13). Comparison of the amino acid sequences of T7 *O.3* protein and three plasmid-encoded antirestriction proteins revealed only one small region of similarity (9, Fig. 1A). It is possible that this conserved amino acid region is responsible for binding to the R-M complex and play a role of universal 'antirestriction' domain. Note that this motif is located in the C-terminal half of Ard proteins, which is essential for their activities as shown for ColIb Ard protein (6). T7 *O.3* protein seems to contain three motifs similar to 'antirestriction' domain one of which (Fig. 1A, coordinates 101–114) is also located at the

end of C-terminal half. Interestingly, the mutant *O.3* protein that lacks this distal motif retains functional activity (12) and this observation is associated with the idea that the T7 *O.3* protein may have at least two functional domains.

Further support for the hypothesis about 'antirestriction' domain comes from the considerable sequence similarity between the conserved 'antirestriction' domain of Ard proteins and the Argos repeats of S polypeptides of type Ia R-M systems (Fig. 1A and B). The S subunits are known to confer sequence specificity to the type I R-M systems that include also R and M subunits essential for restriction and modification functions (see for recent reviews, 1–4). Detailed studies of S subunit structure revealed that they consist of two variable DNA recognition domains (150–180 amino acids) (Fig. 2, open circles) separated by small conserved regions (open bars) each of which contains the sequence that is conserved even between S subunits of different type I systems (closed boxes A₁ and A₂). This conserved sequence has been identified by Argos as a repeat within S polypeptides of the type Ia family (16) and then Kannan *et al.* (17) have shown its conservation within S polypeptides of all three families of type I systems. Comparison of the Argos repeat-like sequences encoded by all three families of type I systems with the amino acid sequences of 'antirestriction' domains of Ard proteins also revealed the sequence similarity (Fig. 1A and C).

In recent years, considerable evidence has accumulated which indicates that the conserved regions of S subunits containing the Argos repeat-like sequences are involved in the interaction with M subunits (18–23), suggesting that this amino acid repeat may play an essential role in the functional integrity of type I R-M systems. It seems likely, therefore, that the conserved 'antirestriction' domain of Ard proteins may compete with the similar Argos

repeat-like domains of the S subunits for the interaction with M subunits and prevent the functional integrity of type I systems (Fig. 2). This proposal is consistent with the observation that the binding site of *O.3* protein of phage T7 appears to be in one of the small subunits of *EcoK* enzyme, either S or M, and that after reaction with T7 *O.3* protein, the holoenzyme is unable to bind to DNA and is totally inactive as a nuclease, methylase or ATPase (13). It is interesting to note that, like the acidic T7 *O.3* protein, plasmid-encoded acidic *ArdA*-type proteins also efficiently inhibit the restriction and modification activities of type I R-M systems in *E. coli* cells (5-8). On the other hand, the slightly acidic antirestriction protein, *ArdB*, has been shown to affect only the restriction activity and does not influence the modification function of R-M systems in *E. coli* (9). These findings are consistent with other observations that the electrostatic interactions between polypeptides may play the important role in the functional organization of R-M complexes (24). However, further work with purified *Ard* proteins will be needed to clarify in detail the mechanism of interaction of antirestriction proteins and R-M enzymes.

It is interesting to note that in addition to the phage T7 and plasmid-encoded functions, the antirestriction functions that are induced in *E. coli* when cell's DNA is damaged (by UV irradiation) (25-27), modified (by 2-aminopurine) (28) or unmethylated (in *dam* mutants deficient in adenine methylase activity) (27,29) also affect only the type I R-M systems of *E. coli* and influence types II and III restriction in *E. coli* only slightly or not at all. It seems likely, therefore, that the type I R-M systems are programmed to interact with these (phage, plasmid and cellular) antirestriction functions to control the permeability of host restriction barrier. It is possible that such an interaction between restriction and antirestriction functions provides the effective mechanism for controlling the transfer of foreign genes into cells and may make bacterial cells more adaptable and evolutionary flexible. We speculate that the conserved motif found in the S polypeptides of type I systems and antirestriction proteins may be a basis for such interactions.

ACKNOWLEDGEMENTS

This work was supported by a grant No M20000 from the International Science Foundation and a grant No 98-94 from Russian Government Program 'Priority Studies in Genetics'.

REFERENCES

- 1 Wilson, G.G. and Murray, N.E. (1991) *Annu. Rev. Genet.* **25**, 585-627.
- 2 Bickle, T.A. and Kruger, D.H. (1993) *Microbiol. Rev.* **57**, 434-450.
- 3 Heitman, J. (1993) In Setlow, J.K. (ed.), *Genetic Engineering*. Plenum Press, New York, pp. 57-108.
- 4 Roberts, R.J. and Halford, S.E. (1993) In Linn, S.M., Lloyd, R.S. and Roberts, R.J. (eds), *Nucleases*. 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 35-88.
- 5 Belogurov, A.A., Yussifov, T.N., Kotova, V.U. and Zavilgelsky, G.B. (1985) *Mol. Gen. Genet.* **198**, 509-513.
- 6 Delver, E.P., Kotova, V.U., Zavilgelsky, G.B. and Belogurov, A.A. (1991) *J. Bacteriol.* **173**, 5887-5892.
- 7 Read, T.D., Thomas, A.T. and Wilkins, B.M. (1992) *Mol. Microbiol.* **6**, 1933-1941.
- 8 Belogurov, A.A., Delver, E.P. and Rodzevich, O.V. (1992) *J. Bacteriol.* **174**, 5079-5085.
- 9 Belogurov, A.A., Delver, E.P. and Rodzevich, O.V. (1993) *J. Bacteriol.* **175**, 4843-4850.
- 10 Spoerl, N.P., Herrlich, P. and Bickle, T.A. (1979) *Nature (London)* **278**, 30-34.
- 11 Mark, K.-K. and Studier, F.W. (1981) *J. Biol. Chem.* **256**, 2573-2578.
- 12 Dunn, J.J., Elzinga, M., Mark, K.-K. and Studier, F.W. (1981) *J. Biol. Chem.* **256**, 2579-2585.
- 13 Bandyopadhyay, P.K., Studier, F.W., Hamilton, D.L. and Yuan, R. (1985) *J. Mol. Biol.* **182**, 567-578.
- 14 Kruger, D.H., Schroeder, C., Reuter, M., Bogdarina, I.G., Buryanov, Y.I. and Bickle, T.A. (1985) *Eur. J. Biochem.* **150**, 323-330.
- 15 Moffatt, B.A. and Studier, F.W. (1988) *J. Bacteriol.* **170**, 2095-2105.
- 16 Argos, P. (1985) *EMBO J.* **4**, 1351-1354.
- 17 Kannan, P., Cowan, G.M., Daniel, A.S., Gann, A.A.F. and Murray, N.E. (1989) *J. Mol. Biol.* **209**, 335-344.
- 18 Gough, J.A. and Murray, N.E. (1983) *J. Mol. Biol.* **166**, 1-19.
- 19 Gann, A.A.F., Campbell, A.J.B., Collins, J.F., Coulson, A.F.W. and Murray, N.E. (1987) *Mol. Microbiol.* **1**, 13-22.
- 20 Meister, J., MacWilliams, M., Hubner, P., Jutte, H., Skrzypek, E., Piekarczyk, A. and Bickle, T.A. (1993) *EMBO J.* **12**, 4585-4591.
- 21 Abadjieva, A., Patel, J., Webb, M., Zinkevich, V. and Firman, K. (1993) *Nucleic Acids Res.* **21**, 4435-4443.
- 22 Cooper, L.P. and Dryden, D.T.F. (1994) *J. Mol. Biol.* **236**, 1011-1021.
- 23 Willcock, D.F., Dryden, D.T.F. and Murray, N.E. (1992) *EMBO J.* **13**, 3902-3908.
- 24 Dryden, D.T.F., Cooper, L.P. and Murray, N.E. (1993) *J. Biol. Chem.* **268**, 13228-13236.
- 25 Day, S.R. (1977) *J. Virol.* **21**, 1249-1251.
- 26 Thoms, B. and Wackernagel, W. (1982) *Mol. Gen. Genet.* **186**, 111-117.
- 27 Kelleher, J.E. and Raleigh, E.A. (1994) *J. Bacteriol.* **176**, 5888-5896.
- 28 Efimova, E.P., Delver, E.P. and Belogurov, A.A. (1988) *Mol. Gen. Genet.* **214**, 317-320.
- 29 Efimova, E.P., Delver, E.P. and Belogurov, A.A. (1988) *Mol. Gen. Genet.* **214**, 313-316.
- 30 Cowan, G.M., Gann, A.A.F. and Murray, N.E. (1989) *Cell* **56**, 103-109.
- 31 Price, C., Lingner, J., Bickle, T.A., Firman, K. and Glover, S.W. (1989) *J. Mol. Biol.* **205**, 115-125.
- 32 Dybvig, K. and Yu, H. (1994) *Mol. Microbiol.* **12**, 547-560.