The EMBO Journal vol.14 no.16 pp.4083–4089, 1995

A protein that binds to the P1 origin core and the *oriC* 13mer region in a methylation-specific fashion is the product of the host *seqA* gene

Therese Brendler, Ann Abeles and Stuart Austin[1]

Laboratory of Chromosome Biology, ABL–Basic Research Program, NCI–Frederick Cancer Research and Development Center, Frederick, MD 21702-1201, USA

[1]Corresponding author

The P1 plasmid replication origin P1*oriR* is controlled by methylation of four GATC adenine methylation sites within heptamer repeats. A comparable (13mer) region is present in the host origin, *oriC*. The two origins show comparable responses to methylation; negative control by recognition of hemimethylated DNA (sequestration) and a positive requirement for methylation for efficient function. We have isolated a host protein that recognizes the P1 origin region only when it is isolated from a strain proficient for adenine methylation. The substantially purified 22 kDa protein also binds to the 13mer region of *oriC* in a methylation-specific fashion. It proved to be the product of the *seqA* gene that acts in the negative control of *oriC* by sequestration. We conclude that the role of the SeqA protein in sequestration is to recognize the methylation state of P1*oriR* and *oriC* by direct DNA binding. Using synthetic substrates we show that SeqA binds exclusively to the hemimethylated forms of these origins, forms that are the immediate products of replication in a methylation-proficient strain. We also show that the protein can recognize sequences with multiple GATC sites, irrespective of the surrounding sequence. The basis for origin specificity is primarily the persistence of hemimethylated forms that are over-represented in the natural DNA preparations relative to controls.

Keywords: methylation/*oriC*/P1*oriR*/replication/sequestration

Introduction

Replication of the plasmid prophage of bacteriophage P1 is stringently regulated from the P1*oriR* replication origin (Abeles *et al.*, 1984). The origin contains clustered GATC DNA adenine methylation sites, substrates for the DNA adenine (Dam) methylase. These are nested within short (heptamer) direct repeat elements (Figure 1). Although this feature is limited to a small group of low copy number plasmids that show homology to P1 (Kamio and Terawaki, 1983; Saul *et al.*, 1988; Gabant *et al.*, 1993), clustered Dam sites are common in the origins of bacterial chromosomes (Zyskind *et al.*, 1983), including the *Escherichia coli* origin, *oriC* (Meijer *et al.*, 1979). These origins also have GATC sites nested within repeat sequences, in this case 13mers (Figure 1). These 13mer repeats are involved in origin strand opening (Bramhill and Kornberg, 1988a). Despite their different length and sequence (Figure 1), it has been proposed that the P1 heptamers play a similar role (Bramhill and Kornberg, 1988b).

The *oriC* origin is subject to negative regulation, termed sequestration. This acts on the hemimethylated DNA created by replication and blocks secondary initiations (Messer *et al.*, 1985; Ogden *et al.*, 1988; Campbell and Kleckner, 1990). It is responsible for the synchrony of *oriC* initiation within the cell cycle (Bakker and Smith, 1989; Boye and Lobner-Olesen, 1990). Recently a gene (*seqA*) has been identified whose product is required for sequestration (Lu *et al.*, 1994). The *seqA* gene shows genetic interactions with that for the DnaA initiator protein and it was suggested that DnaA can recognize the methylation state of the origin and transmit this information via the *seqA* product to regulate origin sequestration (Lu *et al.*, 1994).

P1*oriR* and *oriC* show parallel responses to adenine methylation (Abeles *et al.*, 1993). Both are subject to sequestration. However, both also show a positive requirement for methylation in order to function efficiently. Plasmids replicated from them are unstably maintained at very low copy number in Dam$^-$ strains (Abeles *et al.*, 1993) and P1*oriR* templates also require origin methylation for efficient replication *in vitro* (Abeles and Austin, 1987; Abeles *et al.*, 1993).

What is the basis for the parallel responses of these origins to Dam methylation? The nesting of GATC sequences within direct repeats suggests that these regions may act in some more complex fashion than simply serving as targets for methylation. The integrity of at least the right-most of the 13mer repeats is essential for *oriC* function (Hwang and Kornberg, 1992) and the P1*oriR* repeats are highly sensitive to single point mutations, irrespective of whether or not the changes affect the nested GATC bases (Brendler *et al.*, 1991). The response to methylation of these repeat sequences could be explained if they recognize some component of the initiation mechanism in a methylation-dependent fashion. Using P1*oriR* as a model, we therefore screened cell extracts for factors that might bind to the repeats only when isolated from Dam$^+$ strains. We will refer to such DNA (which may or may not be fully methylated at its GATC sites) as Meth$^+$ DNA.

Results

Extracts of E.coli contain material that binds to the P1oriR core in a methylation-specific fashion

We used crude (FII) host extract in gel band retardation assays using labeled fragments of the P1 origin core (Figure 1) isolated from plasmids grown either in Dam$^+$

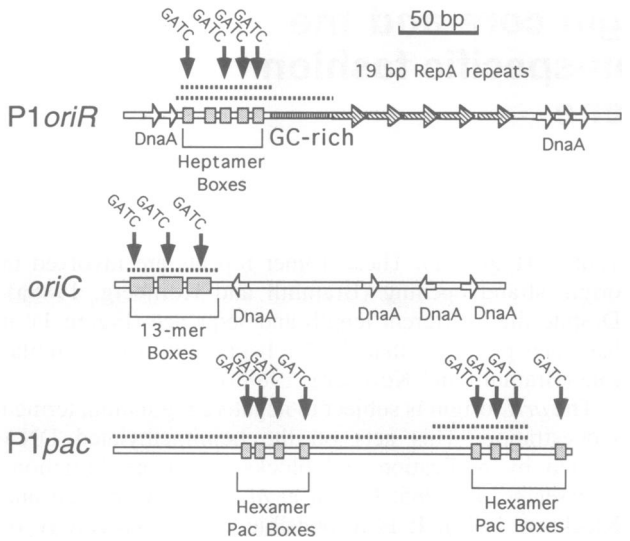

Fig. 1. DNA sites that have multiple GATC Dam methylation sites nested within direct repeats. P1*oriR* and P1*pac* are the replication origin and phage packaging site of bacteriophage P1. The *oriC* replication origin is that of the host chromosome. Only the GATC sites (vertical arrows) that are within direct repeat sequences (stippled boxes) are shown. The direction and orientation of DnaA boxes in the two origin sequences are shown. Dotted lines above the maps show the extent of the DNA fragments that were used in the binding experiments described below. Where two dotted lines are shown, the upper is the extent of the synthetic oligonucleotide used.

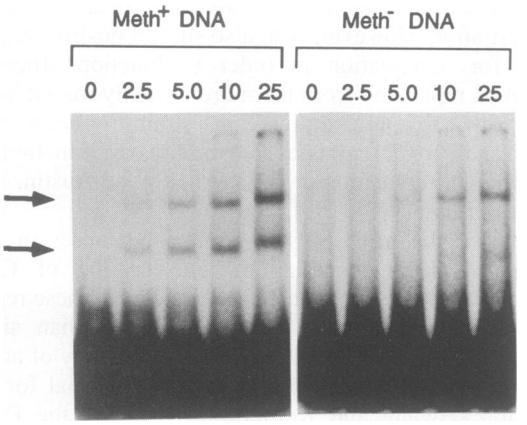

Fig. 2. Gel retardation assays. A radiolabeled fragment encompassing the heptamer repeat region of P1*oriR* (Figure 1) was incubated with crude extracts (FII) of *E.coli* and run on a polyacrylamide gel. Meth$^+$ DNA, adenine methylated fragment isolated from plasmids grown in a Dam$^+$ strain; Meth$^-$ DNA, unmethylated fragment from a Dam$^-$ strain. Material at the bottom of the gel is the unbound fragment. Arrows mark the position of two retarded complexes which form only on methylated DNA. The upper complex overlaps a band formed by non-methylation-specific binding material.

or Dam$^-$ cells. The DNA analog poly(dI–dC)·poly(dI–dC) was used to suppress non-specific DNA binding. Retarded bands with two distinct mobilities were observed using the Meth$^+$ fragment, whereas unmethylated DNA (Meth$^-$ DNA) gave only one of them (Figure 2).

The different patterns observed with Meth$^+$ and Meth$^-$ DNA are due to the fact that at least two different DNA binding proteins are acting, one of which binds only to Meth$^+$ DNA. Thus it was possible to separate the activities and purify the methylation-dependent binding protein.

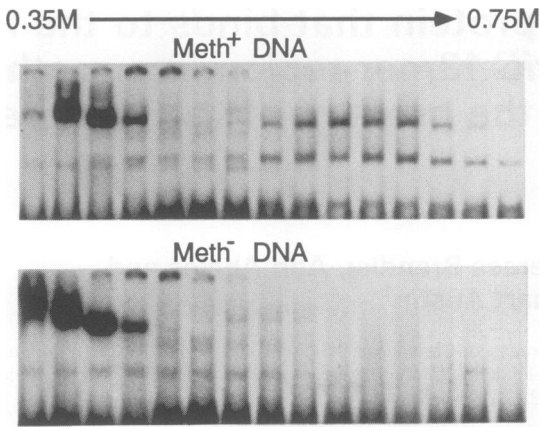

Fig. 3. Gel retardation assay of fractions from heparin–Sepharose chromatography. The labeled P1*oriR* fragment was used (Figure 1). (Top panel) Methylated DNA; (bottom panel) unmethylated DNA. The methylation-specific binding activity elutes as a peak at ~0.6 M NaCl (upper panel) and is clearly separated from a major non-methylation-specific binding activity that elutes at ~0.4 M NaCl.

Partial purification of a methylation-specific DNA binding protein

Figure 3 shows the result of fractionating the crude material by heparin–Sepharose chromatography and assaying binding to Meth$^+$ and Meth$^-$ origin core fragments. A clear separation of a methylation-dependent and a major methylation-independent activity is seen. The methylation-dependent binding activity was purified through six successive chromatography steps by pooling the active fractions (see Materials and methods). Further steps reveal additional methylation-independent activities, which give rise to retardation bands that were not evident when the crude extracts were examined. This is due to the presence of non-methylation-specific DNA binding activities whose affinity for the labeled DNA fragment was originally suppressed by interaction with other factors or by competition from the considerable amount of nucleic acid present in the crude extract. These activities are unmasked in early steps, when purification removes the competing materials (Figure 4). After the hydroxyapatite purification step (step 3) the methylation-dependent activity is separated from all non-methylation-specific activity, so that peak fractions from this and subsequent column steps bound only to the Meth$^+$ P1*ori* DNA (Figure 4). Inspection of Figures 3 and 4 shows that there are two retarded species due to methylation-dependent binding. One corresponds to that seen from the beginning of the purification and the second is a slower migrating species that was obscured by a non-methylation-specific band in the patterns from earlier steps. The two methylation-dependent bands appear to be caused by binding of the same protein; increasing concentrations of protein favor the appearance of the upper band at the expense of the lower one (data not shown).

A 22 kDa protein is responsible for methylation-specific binding to P1oriR

Protein gel analysis of fractions from the final column step (elution from S-Sepharose) shows two visible polypeptides that are partially separated (Figure 5). The peak activity of the fractions corresponds to the smaller of the two, a

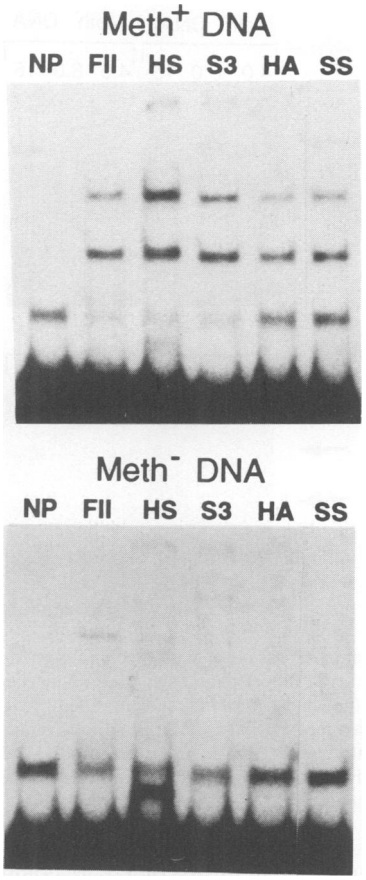

Fig. 4. Gel retardation assays of pooled fractions from selected steps in the protein purification. Labeled P1*oriR* fragment (Figure 1) was incubated with no protein (NP) or fraction II crude extract (FII) and pools of active fractions from heparin–Sepharose (HS), Sephacryl S-300 (S3), hydroxyapatite (HA) or S-Sepharose (SS) columns. Assays from pools from the phenyl-Sepharose and blue-Sepharose columns are not shown, but gave identical patterns to that of the hydroxyapatite pool. (Upper panel) Methylated DNA; (lower panel) unmethylated DNA.

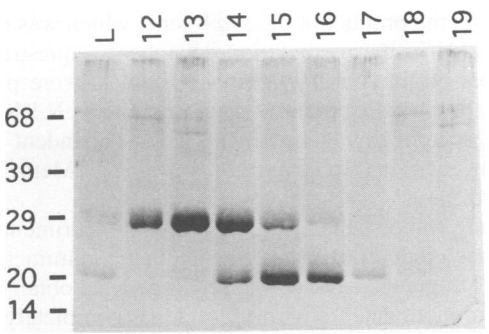

Fig. 5. Silver stained polyacrylamide gel of fractions from S-Sepharose chromatography. Lanes correspond to the column load (L) and fractions 12–19 from the S-Sepharose column. The peak of the methylation-dependent DNA binding activity corresponds to fractions 15 and 16 (data not shown). Thus the activity corresponds to the 22 kDa (lower) band material. The positions of molecular weight markers (molecular weights in kDa) are shown on the left side.

protein with an apparent molecular weight of 22 kDa. Fractions containing the 22 kDa protein, but very little of the larger protein (e.g. fraction 17), are active for methylation-specific binding. Fractions containing only the larger protein (e.g. fraction 12) are not active. Mixing

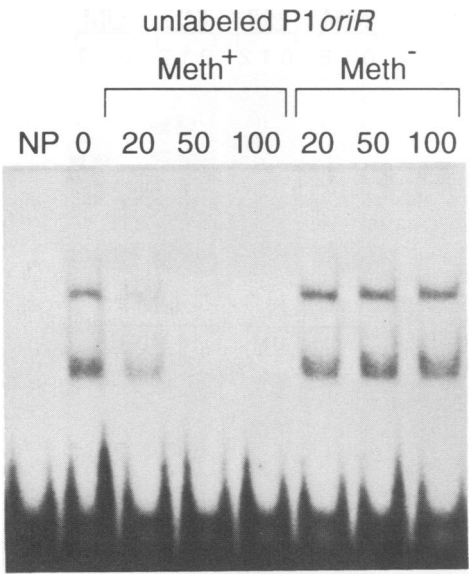

Fig. 6. P1*oriR* DNA competes with *oriC* DNA for protein binding. The labeled *oriC* fragment (10 ng) was incubated with purified methylation-specific binding protein from the S-Sepharose column. NP, no protein. Subsequent lanes, complete reaction plus 0, 20, 50 or 100 ng unlabeled competing DNA. Competing DNA was Meth$^+$ DNA from a methylation-proficient strain or Meth$^-$ unmethylated DNA.

of the two fractions had no additive effect (data not shown). Thus the binding activity is associated with the 22 kDa protein alone.

The 22 kDa protein binding is specific, but also binds to the host replication origin oriC

Unlabeled P1 origin fragments compete with the labeled DNA for binding to the 22 kDa protein, but only when the competing DNA is Meth$^+$. Control fragments from Dam$^+$ strains (the P1*parS* partition site and fragments of plasmid pBR322) failed to compete, irrespective of whether or not they contained GATC sites (data not shown). Thus the methylation-specific binding activity shows specificity for the P1 origin core as isolated from a methylation-proficient strain. We also assayed binding to two sequences that have similar motifs to the P1 origin core. The P1*pac* (phage packaging) site has multiple short repeat sequences, each of which contains a GATC site (Sternberg and Coulby, 1987; Figure 1). No evidence of binding of the protein to the *pac* fragment was detected in these experiments (see below). The second fragment was from *oriC* and spanned the 13mer GATC-containing repeats. Despite the fact that this fragment shares no obvious sequence similarity with the P1*ori* heptamers, except for the presence of the methylation sites, binding was observed (Figure 6). Moreover, a very similar pattern of retarded bands was obtained and the *oriC* binding event was also methylation-specific (Figure 6).

Figure 6 shows the results of a competition experiment in which unlabeled P1 origin fragment was used to compete for protein binding to labeled *oriC* fragment. The Meth$^+$ P1 DNA was an efficient competitor, but the unmethylated form showed no significant competitive binding. We conclude that the same 22 kDa protein is involved in binding both the P1 origin core and the *oriC* 13mer region in a methylation-dependent fashion. It is

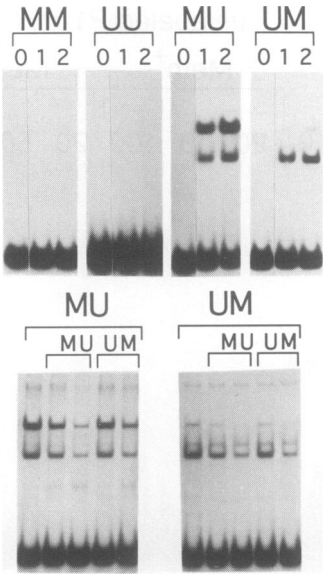

Fig. 7. The 22 kDa protein binding is specific for the hemimethylated DNA forms. MM, UU, MU and UM are the fully methylated, unmethylated and the two hemimethylated P1*oriR* oligonucleotides respectively (MU, top strand methylated; UM, bottom strand methylated; with respect to the map in Figure 1). (Upper panel) Each labeled oligonucleotide was assayed with 1 or 2 μl partially purified SeqA protein. (Lower panel) Competition experiments using labeled MU or UM hemimethylated fragments (large letters) in competition with unlabeled MU or UM fragments (small letters). SeqA–DNA complexes formed with 2 μl partially purified SeqA protein were competed with an equimolar (left lane of bracketed pairs) or 5-fold excess of unlabeled hemimethylated P1 origin DNA. DNA binding assays contained 25 fmol (1.15 ng) appropriate hemimethylated DNA. The final DNA concentration per assay was 10^{-9} M.

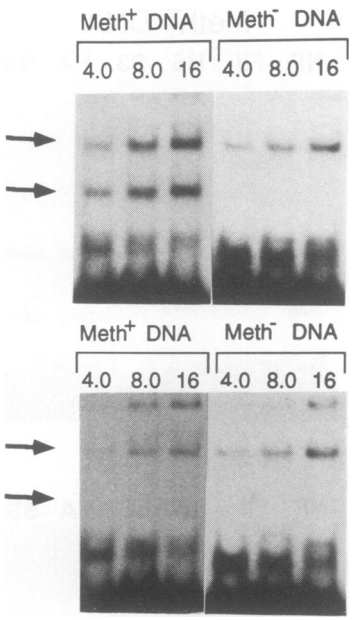

Fig. 8. Gel retardation assays of wild-type and ΔseqA extracts. FII extracts of wild-type (upper panel) and ΔseqA (lower panel) cells were assayed using the labeled P1*oriR* fragment. The arrows mark positions to which the two methylation-specific complexes migrate. These are not formed with Δ*seqA* extracts. Additional bands are due to non-specific DNA binding activities and are *seqA*-independent. Numbers refer to the amount of FII extract used per reaction ($OD_{260} \times 10^3$).

interesting that although the relatively pure 22 kDa protein readily recognizes *oriC*, no methylation-dependent binding to *oriC* is seen with crude (FII) extracts (data not shown). This is presumably due to the presence of an inhibitor or competitor of *oriC* binding that is not present in more purified material.

The 22 kDa binding protein is specific for hemimethylated DNA

Synthetic oligonucleotides were prepared corresponding to the single-stranded components of the P1 origin fragment. One pair of these oligomers had methyladenine at each of the normal methylation sites, whereas the second pair lacked methylation. By annealing the appropriate strands, we constructed unmethylated, fully methylated and each of two hemimethylated origin species (see Materials and methods). Figure 7 shows that the 22 kDa protein binds to both of the hemimethylated species, but fails to bind to the fully methylated or unmethylated forms. We conclude that the 22 kDa protein is a hemimethylation-specific binding protein. Although both hemimethylated P1*oriR* fragments are efficient substrates for binding, the band patterns produced by the two species differ (Figure 7). The origin of the two bands and the significance of their relative abundance is not presently understood. Figure 7 (lower panel) shows the ability of the two P1*oriR* hemimethylated fragments to compete with each other for binding to the protein. We conclude that the two hemimethylated species compete for binding to a common binding region on the protein. As expected, unmethylated and fully methylated forms failed to compete with the hemimethylated forms (data not shown). We also found that none of the single-stranded oligonucleotides was able to compete for binding, even if a large excess was present (data not shown).

The 22 kDa protein is the product of the host seqA gene

The involvement of the protein in *oriC* binding and its apparent molecular weight of 22 kDa suggested a connection with the product of the *seqA* gene, which was recently implicated in the methylation-dependent sequestration of the host origin (Lu *et al.*, 1994). Extracts were prepared from a *seqA* null mutant (a generous gift of N.Kleckner). These specifically lacked methylation-dependent P1*oriR* binding activity to naturally occurring Meth$^+$ DNA (Figure 8).

Figure 9 shows the results of binding experiments using the same crude extracts and the synthetic hemimethylated P1*oriR* fragments. The strong binding signal obtained with the hemimethylated oligonucleotides is completely absent in the Δ*seqA* extract. We conclude that the hemimethylated DNA binding activity depends on the product of the *seqA* gene. The 22 kDa protein was then identified as the *seqA* product by use of an antibody that recognizes the SeqA protein (kindly supplied by N.Kleckner; data not shown).

The hemimethylated state and the specificity of SeqA binding

Binding is much more efficient using the hemimethylated oligonucleotides than when using the original Meth$^+$ DNA preparations (c.f. Figures 8 and 9). Unlike Meth$^+$ DNA, virtually all of the synthetic hemimethylated DNA can be driven into the retarded band with sufficient protein. As

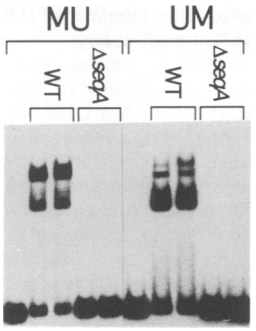

Fig. 9. Binding of the two hemimethylated P1*oriR* oligonucleotides using crude extracts. Gel retardation assays of crude FII extracts of wild-type (WT) and Δ*seqA* cells using the two hemimethylated forms (MU and UM) of the P1*oriR* oligonucleotide were done. The *seqA* mutation eliminates all detectable binding activity. No detectable binding activity was seen with either extract using fully methylated or unmethylated DNA under these conditions (not shown). The first and sixth lanes are MU and UM DNA respectively, without added protein. Bracketed pairs of lanes have 5 and 10 U protein respectively ($OD_{260} \times 10^3$).

SeqA binds exclusively and efficiently to hemimethylated DNA, the inefficient binding seen with the naturally isolated Meth$^+$ DNA must be due to binding of a minority of hemimethylated DNA species that persists in the DNA isolated from Dam$^+$ cells. This impacts on the question of the specificity of the protein, as the specific binding to origin sequences relative to controls seen using Meth$^+$ DNA could reflect the abundance of hemimethylated forms of these sequences, rather than the ability of the protein to discriminate between extended sequences.

Figure 10 shows the results of binding experiments using synthetic forms of the *oriC* fragment and a portion of the P1*pac* fragment (see Materials and methods). Fully methylated and unmethylated forms of these fragments gave no significant binding (data not shown). However, the hemimethylated forms of both fragments were efficiently bound, even though P1*pac* does not bind when isolated from Dam$^+$ cells. This implies that the naturally isolated P1*pac* DNA contains few hemimethylated molecules in comparison with P1*oriR* and *oriC* DNA.

Nesting of GATC sequences within extended repeats is not essential for SeqA binding

To further probe the specificity of binding, we constructed oligonucleotides with the same distribution and positioning of the three GATC sites as the synthetic P1*pac* fragment but with all other bases shuffled (see Materials and methods). The hemimethylated forms of this sequence also bind to SeqA (Figure 10). We conclude that the nesting of methylation sites within direct repeat sequences as seen in both P1*pac* and the two origin sequences is not essential for recognition by SeqA. It should be noted that while the hemimethylated forms of the naturally occurring sequences tend to form two retarded bands (Figures 4, 6 and 10, upper panel), the shuffled sequence only forms one, irrespective of the hemimethylated form used (Figure 10, lower panel).

Discussion

We have isolated a protein that binds to the core sequences of two replication origins, P1*oriR* and *oriC*, and only

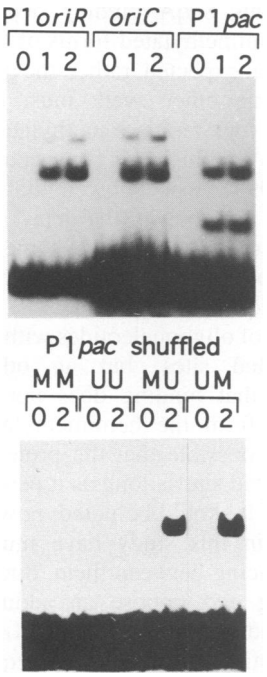

Fig. 10. Binding of purified SeqA protein to hemimethylated *oriC*, P1*pac* and P1*pac* shuffled sequences. (Upper panel) Labeled oligonucleotides (see Figure 1) with the nested GATC sequences hemimethylated (the lower strands with respect to the maps in Figure 1 were methylated). (Lower panel) Fully methylated, unmethylated (MM and UU) and the two hemimethylated forms of the labeled P1*pac* shuffled sequence. Numbers refer to μl purified SeqA protein used per 25 μl binding reaction.

when these are isolated from a Dam$^+$ strain. The protein is the product of the *seqA* gene, which is required for negative regulation of *oriC* initiation by the mechanism referred to as sequestration. As plasmids driven by P1*oriR* are also subject to sequestration, it is probable that the *seqA* protein regulates P1 replication in a similar way. Sequestration involves the recognition of newly synthesized DNA by detecting its hemimethylated state. Our results indicate that this recognition involves direct DNA binding of the SeqA protein to the hemimethylated origin sequences formed as a result of replication. In the simplest hypothesis, SeqA would block initiation directly by preventing binding of essential initiation factors or by interfering with some key event, such as strand opening. This would prevent re-initiation until the unmethylated strands are methylated, thus releasing the SeqA protein. At least in the case of *oriC*, this takes a considerable time (Lu *et al.*, 1994), so that premature re-initiation is prevented. This model provides an explanation for why P1*ori*, like *oriC*, is subject to negative control by sequestration, but it fails to explain why the origin functions inefficiently in Dam$^-$ strains. We will present evidence that there is a positive requirement for P1 origin methylation that is independent of the sequestration pathway (Austin and Brendler, manuscript in preparation).

We show that, in large part, the specificity of the SeqA protein for origin sequences, rather than sequences such as P1*pac*, is due to the persistence of hemimethylated forms of these origins in Dam$^+$ DNA. Hemimethylated P1*pac* DNA binds SeqA efficiently, yet the hemimethylated forms of P1*pac* are not apparently detectable in

Dam$^+$ DNA. Thus SeqA binding does not cause the persistence of hemimethylated forms of origin sequences, as might have been expected, rather, the persistence causes the binding. Some other event must recognize origin sequences to prevent rapid re-methylation of the origin GATC sites. This conclusion is consistent with the results of Lu *et al.* (1994), who showed that *seqA* mutations diminish, but do not abolish, the delay in re-methylation of *oriC* replication products. Thus some mechanism that recognizes the post-replication state of the origin appears to act prior to SeqA binding.

The properties of oligonucleotides with various numbers of hemimethylated sites, but an otherwise random sequence, show that binding does not require specific sequences apart from the hemimethylated GATC sites themselves. It is possible that the protein can recognize any hemimethylated site as long as it persists long enough to allow binding. It should be noted, however, that all the substrates used in this study have multiple sites with rather regular spacing between them. Such features might facilitate binding and impose an additional degree of specificity on binding to loci such as P1*oriR* and the host *oriC*. The properties of the shuffled sequence show that the nesting of GATC sites within extended direct repeats seen in the naturally occurring sequences is not needed for binding. However, these repeat features could still play a role in recognition. One or both hemimethylated forms of the naturally occurring sequences always give two retarded bands, whereas the hemimethylated shuffled sequences give only one. Thus it is possible that the extended repeats promote some novel topology or stoichiometry of the bound complex that is reflected in the two band pattern.

Materials and methods

Buffers

Buffer A was 25 mM HEPES–KOH, pH 7.5, 0.1 mM EDTA, 2 mM dithiothreitol (DTT), 15% (v/v) glycerol; buffer B was 20 mM potassium phosphate, pH 6.8, 2 mM DTT, 15% (v/v) glycerol; buffer C was 50 mM potassium phosphate, pH 7.0, 0.1 mM EDTA, 2 mM DTT, 15% (v/v) ethylene glycol; buffer D was 40 mM HEPES–KOH, pH 7.5, 0.1 mM EDTA, 2 mM DTT, 15% (v/v) glycerol.

Strains and plasmids used

The FII extracts were prepared from C600, MM294 and the derivative MM294*seqA*Δ::*tet* (Lu *et al.*, 1994). The DNA fragments were from plasmids pALA1202 and pALA1208 (P1*ori*), pL70 (P1*pac*; Sternberg and Colby, 1987) and pALA1290 (*oriC*) grown on strain DH5, for methylated DNA, or RS5033 (*dam4*; Lu, 1987), for the unmethylated DNA. The regions contained in the inserts of these plasmids are indicated by the dotted lines in Figure 1. Plasmid pALA1208 contains a synthetic oligomer spanning the P1*oriR* region of the five heptamer repeats plus an adjacent 39 base GC-rich region (bases 405–502; Brendler *et al.*, 1991). To facilitate cloning, a *Bgl*II site was added on the left end and a *Bam*HI site on the right end. This was inserted into the *Bam*HI site of pUC19 (Sambrook *et al.*, 1989). Plasmid pALA1202 is similar, but the fragment encompasses only bases 405–465, the five heptamer repeats and six bases to their right (Figure 1). Plasmid pALA1290 contains the three 13mer repeats of *oriC* plus the adjacent GATC site (bases 17–80; Woelker and Messer, 1993). The oligomers have an *Xba*I site on the left end and adjacent *Xba*I and *Kpn*I sites on the right end. The resulting fragments were inserted between the *Xba*I and *Kpn*I sites of plasmid pUC19. Plasmid pL70 contains seven of the eight hexamer repeats of the P1*pac* site inserted between the *Eco*RI and *Bam*HI sites of pUC19 (Sternberg and Colby, 1987).

Purification and radioactive labeling of DNA fragments for electrophoretic mobility shift assay

Plasmid DNA was prepared by alkaline lysis and cesium chloride/ethidium bromide gradient centrifugation (Sambrook *et al.*, 1989). This DNA was treated with ATP-dependent DNase (US Biochemicals, used as recommended by the supplier) to remove contaminating chromosomal DNA. The DNA was digested with *Eco*RI and *Hin*dIII and the desired fragments were separated by electrophoresis and purified as described (Sambrook *et al.*, 1989). The DNA fragments were radiolabeled using Klenow enzyme and [α-^{32}P]dATP (Sambrook *et al.*, 1989). Following the labeling reaction, the DNA was diluted to 100 µl, extracted with phenol and chloroform and purified by Sephacryl S-200 HR chromatography.

Preparation of synthetic DNA fragments

Deoxyoligonucleotides were synthesized on an Applied Biosystems DNA synthesizer (Model 380B) using the phosphite triester method. Single-stranded deoxyoligonucleotides were prepared corresponding to both strands of P1*oriR*, *oriC*, P1*pac* and a shuffled P1*pac* sequence. For each oligonucleotide a second version was made in which the GATC sequences were 2′-*N*6-adenomethylated using 5′-dimethoxytrityl-2′-*N*6-methyldeoxyadenosine-3′-[(2-cyanoethyl)-(*N*,*N*-diisopropyl)]-phosphoramidite (Glen Research Corporation, Sterling, VA). By annealing the appropriate pairs, double-stranded oligonucleotides were obtained corresponding to the unmethylated, fully adenomethylated and the two hemimethylated forms of each sequence. After purification of the oligonucleotides by Sephacryl S-100 HR chromatography, 1 µg of each complementary oligonucleotide was annealed in 70 µl 50 mM NaCl, 10 mM $MgCl_2$, 10 mM Tris–HCl, 1 mM DTT, pH 7.9. The DNA was heated to 65°C for 15 min and then cooled slowly to room temperature for 1 h. The DNA was further cooled to 4°C overnight. The DNA solution was brought to 0.1 mg/ml bovine serum albumin and the DNA digested with *Xba*I for 3 h at 30°C. The resulting DNA fragment was purified on an 8% polyacrylamide gel. The 69 or 70 base pair fragment was excised, extracted for 3 h at 37°C, purified by alcohol precipitation and resuspended in 10 mM Tris–HCl, 0.1 mM EDTA, pH 8.0 (Sambrook *et al.*, 1989). The P1*oriR* and *oriC* sequences were bracketed by *Xba*I sites. After *Xba*I digestion the products were identical to the 69 base pair *Xba*I fragment from pALA1202 (P1*oriR*) and the 70 base pair *Xba*I fragment from pALA1290 (*oriC*). The 91 base pair P1*pac* oligonucleotides contained only the right three hexamer repeats as shown in Figure 1 (bases 70–124; Sternberg and Colby, 1987). The left end contained an *Xba*I site and the right end contained sequences identical to the *Bam*HI–*Xba*I fragment of pUC19. *Xba*I digestion of annealed duplexes yielded a 69 base pair fragment. The P1*pac* shuffled oligonucleotide retains the P1*pac* GATC sequences and *Xba*I sites, but the rest of the bases were shuffled at random. The resulting sequence was: 5′-GGTCGACTCTAGAAGGTGGACATTCCTAAAGCTCTAAG**GATC**C-AAGCACC**GATC**GACTCTG**GATC**AGAACAACATCTAGAGTCGA-CCTGCA-3′.

Preparation of FII extract

FII extracts of *E.coli* were prepared as described by Fuller *et al.* (1981), with the following modifications. The cells were grown to an OD_{600} of 1.5–1.6 before pelleting. They were washed and resuspended in 25 mM HEPES, pH 7.5, 1 mM EDTA, 10% sucrose and 2 mM DTT before freezing. After thawing the cell pellet, a 0.11 volume of 10× lysis salts (2.5 M KCl, 0.2 M EDTA, 0.2 M spermidine–HCl, 0.02 M DTT) was added and the cells were treated with lysozyme and another freeze/thaw cycle as described previously (Fuller *et al.*, 1981). Centrifugation (183 000 *g*, Beckman 50Ti rotor) was for 45 min. The proteins were precipitated with 0.35 g/ml $(NH_4)_2SO_4$ for 1 h. The pellet was dialyzed extensively against buffer D plus 200 mM ammonium sulfate.

Electrophoretic mobility shift assay

Each 25 µl assay contained 20 mM Tris–acetate, pH 7.5, 1 mM DTT, 10 mM EDTA, 0.1% Nonidet NP-40, 5% glycerol, 0.5–2.0 µg poly(dI–dC)·poly (dI–dC), 10 000 c.p.m. (2 fmol) labeled DNA and potassium glutamate to give 80–100 mM cation. An aliquot of 0.5–5.0 µl protein was added and the reaction incubated at 30°C for 20 min. Loading dye [3 µl, 20 mM Tris–acetate, pH 7.5, 1 mM DTT, 0.1 mM EDTA, 50% glycerol and 0.25% (w/v) bromophenol blue] was added and the mixture was electrophoresed at 25 mA in an 8% polyacrylamide gel in 1× glycerol tolerant gel buffer (US Biochemical). When required, the protein fractions were diluted with 200 mM potassium glutamate, 20 mM Tris–acetate, pH 7.5, 0.1 mM EDTA, 1 mM DTT, 10% glycerol, or they were dialyzed against buffer D plus 200 mM ammonium sulfate. Following

electrophoresis, the gels were washed in 10% acetic acid and 15% methanol for 15 min, dried under vacuum and exposed to X-Omat film (Kodak).

Purification of the methylation-specific DNA binding protein (SeqA)

All chromatography media were from Pharmacia Biotech Inc., except Macro-Prep Ceramic hydroxyapatite, which was obtained from BioRad Laboratories. FII extracts were thawed and diluted with buffer A to a conductivity equivalent to buffer A containing 150 mM NaCl. Extracts were loaded onto a heparin–Sepharose CL-6B column (2.6×20 cm, 100 ml) equilibrated with buffer A containing 150 mM NaCl and eluted with a linear gradient of 0.15–1.5 M NaCl in buffer A. The active fractions (0.46–0.78 M NaCl) were pooled and concentrated by precipitation with 0.35 g/ml ammonium sulfate. The precipitate was resuspended and dialyzed extensively against buffer D plus 200 mM ammonium sulfate. This pool was applied to a Sephacryl S-300 HR column (1.6×34 cm, 68 ml) equilibrated with buffer B plus 0.1 M NaCl. An activity peak corresponding to 65 kDa was pooled and diluted to a conductivity equivalent to 0.05 M NaCl with buffer B. The pool was applied to a column of Macro-Prep Ceramic hydroxyapatite, 40 μm mesh (1.6×5 cm, 10 ml), equilibrated with buffer B containing 0.05 M NaCl and eluted with a linear gradient of 0–0.6 M ammonium sulfate in buffer B plus 0.05 M NaCl. The active pool (0.16–0.32 M ammonium sulfate) was adjusted to 1.0 M ammonium sulfate by the dropwise addition of buffer C containing 2.0 M ammonium sulfate with 15% (v/v) glycerol, instead of ethylene glycol, and applied to a 1 ml phenyl-Sepharose HP HiTrap column equilibrated with buffer C containing 1.0 M ammonium sulfate. The column was developed with a linear descending gradient of 1.0–0 M ammonium sulfate in buffer C. The activity peak eluted between 0.80 and 0.55 M ammonium sulfate. This pool was dialyzed extensively against buffer D plus 200 mM ammonium sulfate, diluted with buffer D to a conductivity equivalent to buffer D containing 0.1 M KCl, made 0.01 M in magnesium acetate and applied to a 1 ml blue-Sepharose HP HiTrap column equilibrated with buffer D containing 0.1 M KCl and 0.01 M magnesium acetate. Using buffer D containing 0.01 M magnesium acetate, stepwise elutions were performed with 0.5, 1.5 and 2.0 M KCl. The 1.5 M KCl eluate was dialyzed extensively against buffer D plus 0.1 M KCl, applied to a 1 ml SP-Sepharose HP HiTrap column equilibrated with buffer D containing 0.1 M KCl and eluted with a linear gradient of 0.1–1.5 M KCl in buffer D. The methyladenine-dependent DNA binding activity eluted between 0.58 and 0.78 M KCl. Individual column fractions were frozen and stored at −70°C for further use.

Western blotting analysis

Proteins were separated on a 12.5% (w/v) polyacrylamide mini gel (Laemmli, 1970) and transferred to an Immobilon-P PVDF membrane as described (Pluskal *et al.*, 1986), using a pH 9.9 carbonate buffer (Dunn, 1986). A 1:1000 dilution of anti-SeqA antibody (a generous gift of Dr N.Kleckner) and the PHOTOBLOT Chemiluminescent System (Life Technologies) were used to visualize the bands.

Acknowledgements

We would like to gratefully acknowledge Marilyn Powers of the Oligonucleotide Synthesis Laboratory, SAIC, for synthesizing both the unmethylated and adenomethylated oligonucleotides used in this study. This research was sponsored by the National Cancer Institute, DHHS, under contract with ABL. The contents of this publication do not necessarily reflect the views or policies of the Department of Health and Human Services nor does mention of trade names, commercial products or organizations imply endorsement by the US Government.

References

Abeles,A.L. and Austin,S.J. (1987) P1 plasmid replication requires methylated DNA. *EMBO J.*, **6**, 3185–3189.

Abeles,A., Snyder,K. and Chattoraj,D.J. (1984) P1 plasmid replication: replicon structure. *J. Mol. Biol.*, **173**, 307–324.

Abeles,A.L., Brendler,T. and Austin,S. (1993) Evidence for two levels of control of P1 *oriR* and host *oriC* replication origins by DNA adenine methylation. *J. Bacteriol.*, **175**, 7801–7807.

Bakker,A. and Smith,D. (1989) Methylation of GATC sites is required for precise timing between rounds of replication in *Escherichia coli*. *J. Bacteriol.*, **171**, 5738–5742.

Boye,E. and Lobner-Olesen,A. (1990) The role of *dam* methyl transferase in the control of DNA replication in *E. coli*. *Cell*, **62**, 981–989.

Bramhill,D. and Kornberg,A. (1988a) Duplex opening by dnaA protein at novel sequences in initiation of replication at the origin of the *E. coli* chromosome. *Cell*, **52**, 743–755.

Bramhill,D. and Kornberg,A. (1988b) A model for initiation at origins of DNA replication. *Cell*, **54**, 915–918.

Brendler,T. Abeles,A. and Austin,S. (1991) Critical sequences in the core of the P1 plasmid replication origin. *J. Bacteriol.*, **173**, 3935–3942.

Campbell,J.L. and Kleckner,N. (1990) *E. coli oriC* and the *dnaA* gene promoter are sequestered from *dam* methyltransferase following the passage of the chromosomal replication fork. *Cell*, **62**, 967–979.

Dunn,S.D. (1986) Removal of the epsilon subunit from *Escherichia coli* F1-ATPase using monoclonal anti-epsilon antibody affinity chromatography. *Anal. Biochem.*, **157**, 144–153.

Fuller,R.S., Kaguni,J.M. and Kornberg,A. (1981) Enzymatic replication of the origin of the *Escherichia coli* chromosome. *Proc. Natl Acad. Sci. USA*, **78**, 7370–7374.

Gabant,P., Newnham,P., Taylor,D. and Couturier,M. (1993) Isolation and location on the R27 map of two replicons and an incompatibility determinant specific for IncHI1 plasmids. *J. Bacteriol.*, **175**, 7697–7701.

Hwang,D.S. and Kornberg,A. (1992) Opposed actions of regulatory proteins, DnaA and IciA. in opening the replication origin of *Escherichia coli*. *J. Biol. Chem.*, **267**, 23087–23091.

Kamio,Y. and Terawaki,Y. (1983) Nucleotide sequence of an incompatibility region of mini-Rts1 that contains five direct repeats. *J. Bacteriol.*, **155**, 1185–1191.

Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (Lond.)*, **227**, 680–685.

Lu,A.-L. (1987) Influence of GATC sequences on *Escherichia coli*: DNA mismatch repair *in vitro*. *J. Bacteriol.*, **169**, 1254–1259.

Lu,M., Campbell,J.L., Boye,E. and Kleckner,N. (1994) SeqA: a negative modulator of replication initiation in *E. coli*. *Cell*, **77**, 413–426.

Meijer,M., Beck,E., Hansen,F.G., Bergmans,H.E.N., Messer,W., von Meyenburg,K. and Schaller,H. (1979) Nucleotide sequence of the origin of replication of the *Escherichia coli* K-12 chromosome. *Proc. Natl Acad. Sci. USA*, **76**, 580–584.

Messer,W., Bellekes,V. and Lother,H. (1985) Effect of *dam* methylation on the activity of the *E. coli* replication origin, *oriC*. *EMBO J.*, **4**, 1327–1332.

Ogden,G.B., Pratt,M.J. and Schaechter,M. (1988) The replicative origin of the *E. coli* chromosome binds to cell membranes only when hemimethylated. *Cell*, **54**, 127–135.

Pluskal,M.G., Prezekop,M.B., Kavonian,M.R., Vecoli,C. and Hicks,D.A. (1986) Immobilon PVDF transfer membrane: a new membrane substrate for Western blotting of proteins. *BioTechniques*, **4**, 272–283.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Saul,D., Lane,D. and Bergquist,P.L. (1988) A replication region of the IncHI plasmid, R27, is highly homologous with the RepFIA replicon of F. *Mol. Microbiol.*, **2**, 219–225.

Sternberg,N. and Coulby,J. (1987) Recognition and cleavage of the bacteriophage P1 packaging site (*pac*). II. Functional limits of *pac* and location of *pac* cleavage termini. *J. Mol. Biol.*, **194**, 469–479.

Woelker,B. and Messer,W. (1993) The structure of the initiation complex at the replication origin, *oriC*, of *Escherichia coli*. *Nucleic Acids Res.*, **21**, 5025–5033.

Zyskind,J.W., Cleary,J.M., Brusilow,W.S.A., Harding,N.E. and Smith,D.W. (1983) Chromosomal replication origin from the marine bacterium *Vibrio harveyi* functions in *Escherichia coli*: *oriC* consensus sequence. *Proc. Natl Acad. Sci. USA*, **80**, 1164–1168.

Received on March 16, 1995; revised on May 24, 1995