Biochemical mechanisms of initiation and termination of plasmid DNA replication

Deepak Bastia

Department of Biochemistry and Molecular Biology, Medical University of South Carolina, Charleston, SC 29425, USA

My first exposure to research was a project on ultrastructure of interphase chromosomes with Swaminathan as the mentor[1]. Since that time I and my associates have focused on the biochemistry of replication initiation and termination and regulation of replication using plasmid chromosomes as model systems. Here, I will briefly review our work that was carried out first at the Yale Medical School, New Haven, CT (with Sherman Weissman as mentor), then in my own laboratory at the Duke University Medical Center, Durham, NC for many years and presently at the Medical University of South Carolina, Charleston. The dispensability of plasmids, the manageable circular DNA templates (chromosomes), and the characteristic copy number at which each plasmid is maintained in the cell makes it a convenient model for investigations of replication mechanisms. Here, I summarize some of the highlights of our work on the biochemistry and copy control mechanisms of the plasmid R6K.

Keywords: Plasmid DNA, replication, initiation, termination.

Replication origins and initiator protein of R6K

MORE than 40 years ago, Jacob *et al.*[2] had postulated that replication of chromosomes (at least prokaryotic chromosomes) is regulated autonomously and that each unit of replication is called a replicon. Each replicon contains a site of initiation (called origin of replication) and encodes an origin-binding initiator protein that positively controlled replication initiation. The drug resistance factor R6K has 3 origins of replication called *ori* α, *ori* β and *ori* γ (Figure 1 *a*). In a given chromosome only one of the 3 *ori* is active, α and β being used the majority of the times and γ used rather infrequently[3–6]; see Figure 1. The plasmid-encoded initiator protein π (refs 7–12) binds specifically to 7 tandemly repeated sequences (called iterons) at *ori* γ and bends the DNA[13]. Binding of π to the γ iterons, in collaboration with the DNA bending protein IHF (integration host factor) and the bacterial initiator DnaA causes melting of the AT-rich region near the iterons[14] and this open complex then recruits the replicative helicase DnaB and other replication proteins such as the 10-subunit DNA polymerase III holoenzyme and the primase DnaG to form the replisome[15–17]; see Figure 1 *b–d*. The assembled

e-mail: bastia@musc.edu

replisome either initiates replication locally at γ or at a distance at α or β. α and β are located at distances of 4000 and 1200 bp respectively, away from γ and the latter contacts α or β by π-mediated DNA looping[18–20] (see Figure 1 *a–c*). This was the first example of replication initiation by 'action at a distance' that is control by initiator protein-
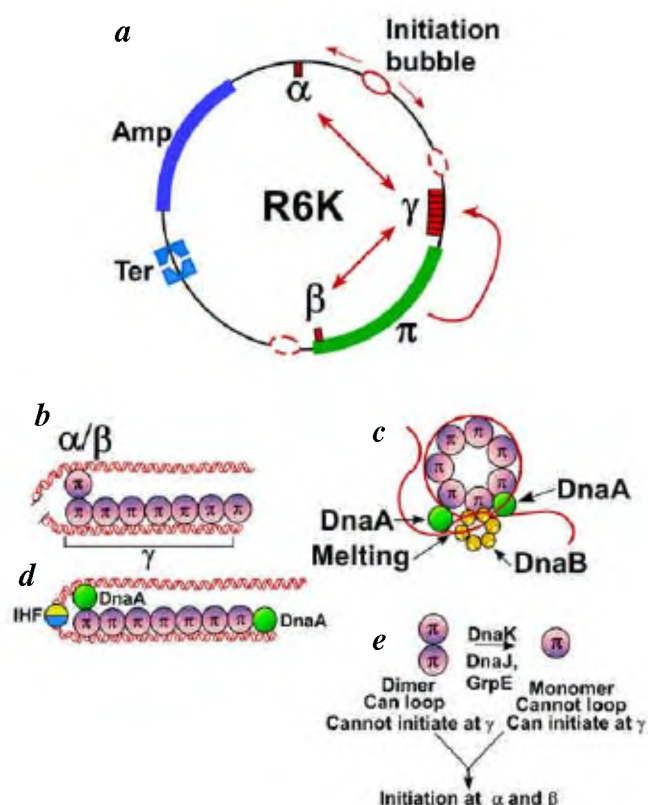
Figure 1. The R6K replicon. *a*, Physical map showing the locations of the three origins of replication called α, β and γ. *Ori* α and β are located ~4000 and ~1200 bp away from the 7 iterons of γ. The red arrows indicate looping between γ and the other two origins. Note that the replication bubble is initiated downstream of the α iteron. The pair of polar Ter sites is shown. *b*, Diagram showing looping between γ and α/β. At least one dimeric π is needed for looping; monomeric mutants cannot loop. *c*, The Iteron DNA is believed to wrap round the 7 π monomers and promotes contacts with DnaA molecule that is necessary to melt the AT rich region of *ori* γ. The melting of the AT rich region is a prerequisite for loading of the hexameric helicase DnaB. *d*, The DNA bending protein IHF bends the DNA and promotes contact among π, the melted region and DnaA. *e*, The WT dimeric π gets converted to monomeric π by the concerted action of DnaK, DnaJ and GrpE proteins of the chaperone system and requires hydrolysis of ATP. Whereas monomeric π is necessary and sufficient to initiate replication at *ori* γ, initiation at *ori* α and β is believed to require both dimeric and monomeric π.
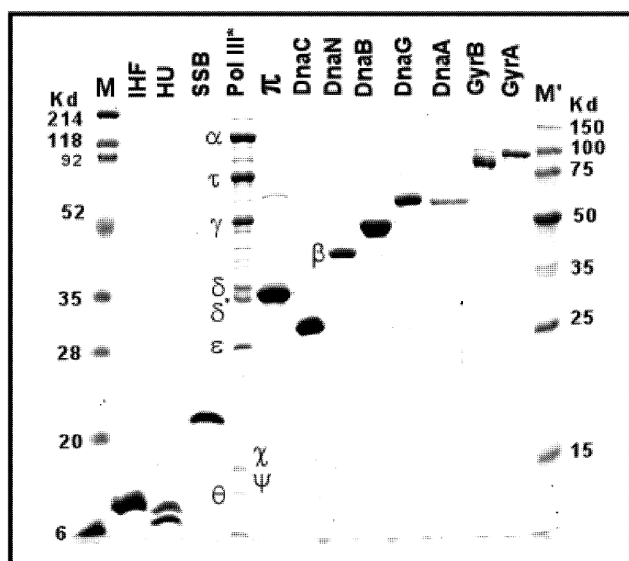
mediated long range DNA looping. The π initiator is the only plasmid-encoded component of the replisome and the remainder consisting of a minimum of 26 other proteins are encoded in the host *E. coli* chromosome.

Host and plasmid encoded proteins needed to replicate R6K

Which are the other host-encoded proteins needed for R6K replication? The most definitive analysis of the minimum set of proteins needed to initiate replication from plasmid replication origins, progression of the forks and termination at a specific terminus (or random termini) has been achieved by identification and purification of some 26 proteins encoded by the host and reconstitution of authentic plasmid replication *in vitro*[15,17]. Most of the replication proteins needed for *in vitro* replication are displayed in a SDS-polyacrylamide gel shown in Figure 2. The proteins functionally belong to several classes: initiators (π and DnaA), structural proteins (HU and IHF), elongation proteins (DnaB helicase, DnaG primase, DnaC (helicase loader), GyrA and GyrB (modulate superhelicity), the 10 subunits of DNA polymerase III, RNaseH (specificity factor), DnaK, DnaJ and GrpE (chaperone system that converts inactive dimeric π to active monomers; Figure 1 *e*) and Tus (replication termination).
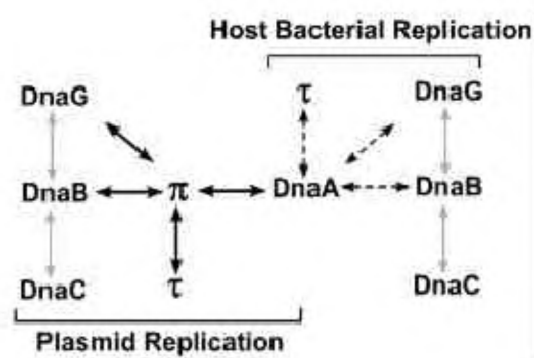
Protein–protein interaction

The binary interactions between π and other host proteins needed for replication initiation and propagation are shown in Figure 3. The interactions indicated by solid black arrows are necessary for R6K replication[14,21,22] and those indicated by grey arrows are needed for both plasmid and host replication. The interaction between π-DnaA (latter is the bacterial initiator protein) in the presence of the DNA bending protein IHF (Figure 1 *d*), is needed for initial DNA melting at the *ori* γ that is a prerequisite for loading the replicative helicase DnaB that requires the molecular matchmaker DnaC. The significances of DnaG-π and π–τ [τ, encoded by the gene DnaN, is a component of DNA polymerase III holoenzyme; (Figure 2)] interactions are not completely clear at this time. The proteins shown in Figure 2 were individually purified and the replication reaction reconstituted *in vitro*[15,17]. The template used was a circular *ori* γ plasmid called pMA-DT and the reaction was monitored by counting acid precipitable radioactivity using ^{3}H-labelled dTTP and unlabelled other dNTPs as precursors (Figure 4). Note that whereas the wild type dimeric π is inert in the *in vitro* reaction, the monomeric mutant form called π^* promotes vigorous replication. What converts the wild type inactive π to an active form? We have

Figure 3. Protein–protein interactions. The interactions in (heavy black arrows) are needed for R6K replication; interactions indicated by dotted arrows are needed for both the replication of the plasmid and the host chromosomes. The interactions shown by grey arrows are needed for only host replication.
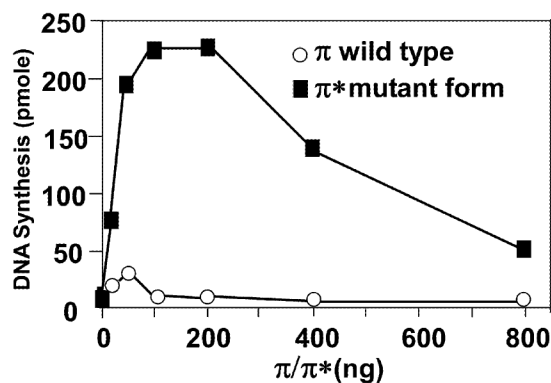

Figure 2. Coomassie Blue stained SDS-polyacrylamide gel showing the degrees of purity of the replication proteins. Tus protein and RNaseH are not shown. The 10 subunits of DNA polymerase III holoenzyme are shown.
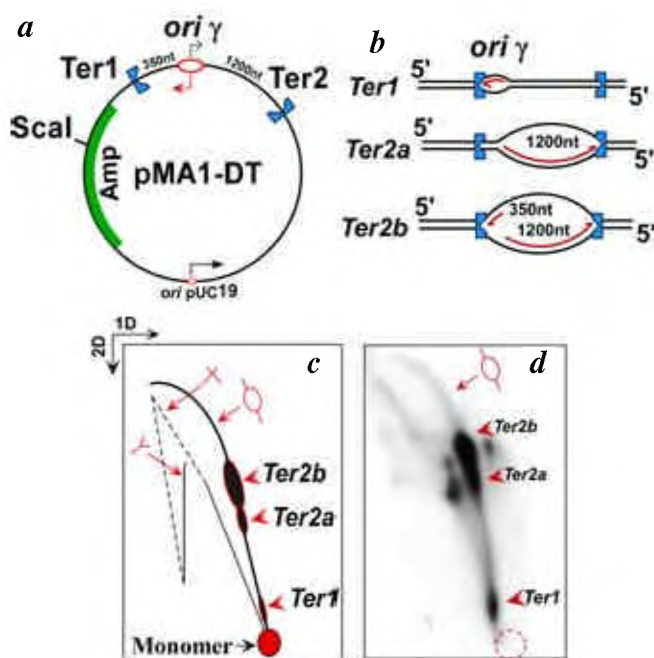

Figure 4. *In vitro* R6K replication with the purified proteins. The extent of replication as a function of both the dimeric wt and the monomeric mutant form π^* is shown. Note that while the monomeric protein catalyses vigorous replication, the dimer is inert.

330

recently reported that the inactive dimeric π protein gets converted to active monomer by the chaperone system DnaJ–DnaK–GrpE and the process requires ATP hydrolysis[15] (see Figure 1 *e*).

Reaction product analysis

In order to make sure that replication *in vitro* had initiated from the correct origin and had the correct topology (θ-shaped or Cairn's form) we constructed the template pMA-DT that had the *ori* γ flanked by two replication termini (*Ter*) placed at different distances from the *ori* (Figure 5 *a*). The rationale was to trap the replication bubble between the terminator pairs and to identify the replication intermediates, origin location, and replication topology. The key technique employed was 2D gel electrophoresis in which a population of circular DNA linearized by a single restriction cut is resolved strictly by molecular mass in the first dimension and then the DNA intercalating dye ethidium bromide is added to the gel and fractionation is carried out in the second dimension to discriminate between the linear DNA from bubbles, Y-shaped DNA generated by the single cut within a bubble and X-shaped DNA representing the last stages of replication[23].

Replication was initiated *in vitro* in the presence or absence of purified replication terminator protein called Tus[17].

Trapping the replication bubble at the two Ter sites should generate 3 types of terminated structures labeled as Ter1, Ter2a and Ter2b (Figure 5 *b*). It should be noted that the *in vitro* reaction lacks ligase, thereby preventing the joining of Okazaki fragments to form the lagging strand. The intermediates were analysed by 2D gel electrophoresis and the expected pattern is shown in Figure 6 *c*. Note that the gel resolves X-shaped intermediates from those containing a replication loop and Y-shaped intermediates generated by a restriction cut within a loop. The actual data are shown in the autoradiogram in Figure 6 *d*. The three spots corresponding to Ter1, Ter2a and Ter2b are all resolved and the pattern is therefore consistent with the replication mode shown in Figure 6 *b* (ref. 17).

Copy control by DNA looping and 'handcuffing'

R6K is maintained at a copy number of about 10–20 per cell. What controls the copy number? The experimental approach to address this question was to first isolate high copy number mutations in the gene encoding π that were manifested in increased resistance to ampicillin when ampicillin resistance was used as the selection marker in the plasmid (e.g., see Figure 1 *a*). Several high copy number mutants are displayed in Figure 6. The plasmid with a known copy number (pACYC174) was used as an internal
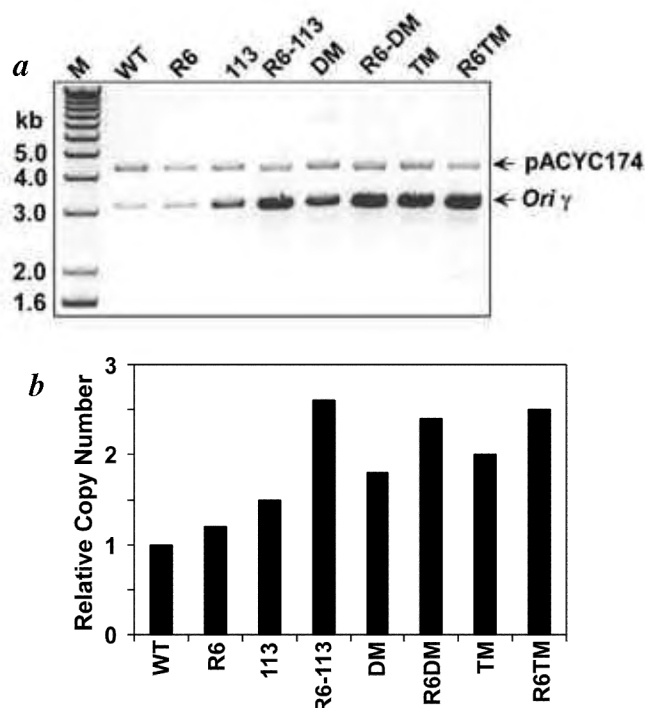

Figure 5. Product analysis of the replication intermediates. *a*, The template pMA-DT has double Ter sites flanking the origin and traps the replication bubble in the presence of the Ter binding protein Tus; *b*, Diagram of the expected termination products; *c*, The expected 2D gel patterns of the replication intermediates including the 3 termination products, bubble-shaped DNA, Y-shaped DNA and X-shaped intermediates. *d*, Typical autoradiogram of a 2D gel showing the actual profile of replication carried out *in vitro* with Tus protein present.

Figure 6. Copy number mutants of R6K. *a*, Photograph of agarose gel showing the relative copy number of linearized *ori* γ plasmid in comparison with the linearized internal standard pACYC 174 DNA. M = molecular size markers, WT, R6, etc. are the wild type and the various mutant forms of π. *b*, Quantification of the *ori* γ DNA normalized with the internal standard.

standard. The mutation R6 (P42L) in π caused loss of DNA looping and a modest increase in copy number. However, a combination of P42L with 3 other mutations namely P106L, F107S and P113S [called R6TM; see ref. 15] causes a significant increase in plasmid copy number (Figure 6). What is the molecular basis of the relaxation of copy control? Biochemical analyses have shown that high copy number directly correlates with increased iteron-promoted monomerization (i.e. the mutant form is dimeric in solution but monomerizes upon binding to the iteron DNA) of π without the need for the chaperones DnaK, DnaJ and GrpE. One should recall that in the *in vitro* reaction, where WT dimeric π was inert, the mutant form π^* (the same as R6TM) supported vigorous replication (Figure 4). Taken together, the data show that the mechanism of copy control is principally through monomerization of the dimeric protein that is active in replication initiation. As noted before, the WT π, needs interaction with monomerization by the DnaK chaperone system (Figure 1 *e*) to activate replication and the equilibrium between inactive dimeric form and the active monomeric form that is maintained by chaperone action controls copy number by regulating the frequency of initiation. The high copy mutant forms make the initiator into constitutive monomers.
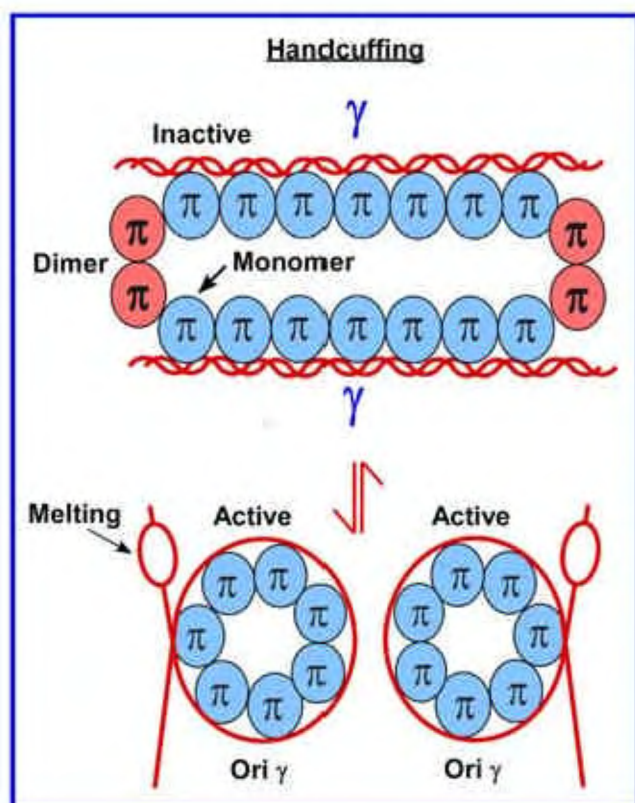

Figure 7. A diagram showing the handcuffing model. *Bottom*, DNA wrapped around monomeric π causes melting of the AT-rich segment at the ori needed for initiation. *Top*, Isomerization by pairing (handcuffing) of the two origins in *trans* mediated by a dimeric protein bridge causes abolition of the melted region, thereby inhibiting replication initiation.

How is replication initiation shut down by π dimers? Various lines of evidence[24,25] suggest that DNA wrapped around monomeric π can form the open complex needed for active replication. However, pairing *in vivo* between origin-π complex causes isomerization to the unraveled structure that are held together by a dimeric π bridge contacting the monomeric arrays (Figure 7). This structure is believed to block origin melting (Zzaman and Bastia in preparation), thereby blocking initiation. This process of negative regulation of replication has been termed as 'handcuffing'.

In addition to negative regulation by handcuffing, π-mediated DNA looping between γ and α or γ and β activates the two distant origins and thus constitutes a mechanism of positive regulation of replication (Figure 1 *b*).

Mechanism of termination of DNA replication

Replication of many prokaryotic circular chromosomes and special regions of rDNA of yeast to human contain specific, short DNA sequences that arrest and terminate replication. These sites are polar, i.e. they arrest forks coming from one direction with respect to the *ori* but allow the forks coming from the opposite direction to pass through (see Figure 1 *a*; Figure 5; and ref. 26). These sequences called Ter sites that bind to a protein called Tus that arrests the progression of the hexameric replicative helicase DnaB (the ATP-dependent motor that drives the fork) in a polar mode[27,28]. We have determined the atomic structure of the first terminator protein namely RTP (**R**eplication **T**erminator **P**rotein) of *B. subtilis*[29]. This work was followed by
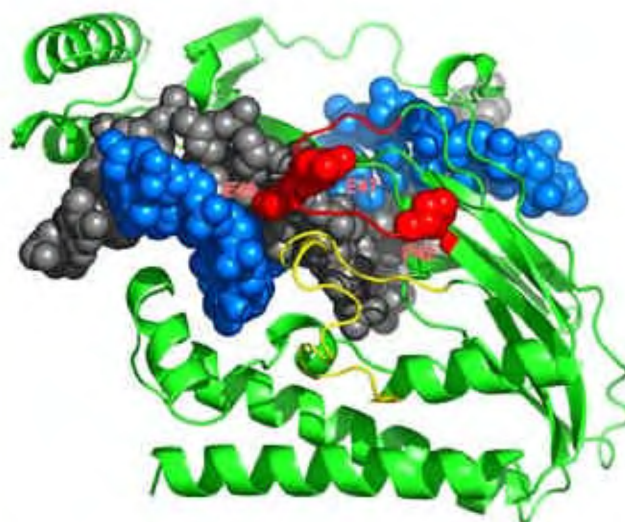
Figure 8. A ribbon diagram of the atomic structure of Tus protein bound to Ter DNA (space filling model). The face that arrests DnaB helicase and forks has the L1 loop shown in red. Single mutations at the positions 42, 47 and 49 cause failure to arrest the helicase without eliminating DNA binding. Thus DNA protein interaction is necessary but not sufficient to arrest helicase/replication forks.

the determination of the analogous protein of *E. coli*[30]. The atomic structure of Tus bound to a Ter site DNA is shown in Figure 8. Note that the protein is an asymmetric monomer that binds to Ter DNA by invading the major groove with the β strands of the protein. What is the molecular basis of asymmetry? Why does it stop the helicase approaching the blocking face that has the projecting L1 loop (shown in red; Figure 8) but not from the opposite direction (that apparently denies access to the L1 loop)? We have reported that the mutations at the amino acid residues 42, 47 and 49, that are located in the L1 loop, cause abolish fork arrest and physical interaction with DnaB by the mutant forms of Tus while still retaining DNA binding activity[31]. These crucial experiments show that mere binding of Tus to Ter DNA is necessary but not sufficient to cause polar fork arrest. It is the contact of the helicase with probably the L1 loop present on the blocking surface that arrests the helicase in a polar mode (Figure 8).

Future directions

Where does the work go from here? An obvious major gap in our knowledge is the lack of information on the atomic structure of the π protein bound to an iteron DNA. We have made major progress in this direction and have crystallized π-DNA complex that diffracts to 3 Å and by introducing bromine atoms into the DNA, we have begun to get the phases and expect that the atomic structure will soon be solved (Swan, Bastia and Davies, in preparation). Working out the molecular details of negative and positive regulation using all purified proteins and unraveling the molecular details of Ter protein-helicase contact are also significant unfinished tasks that are being addressed in our laboratory.

1. Bastia, D. and Swaminathan, M. S., Ultrastructure of interphase chromosomes. *Exp. Cell Res.*, 1967, **48**, 18–26.
2. Jacob, F., Brenner, S. and Cuzin, F., On the regulation of DNA replication in bacteria. *Cold Spring Harbor Symp. Quant. Biol.*, 1964, **28**, 329–348.
3. Crosa, J. H., Luttrop, L. K. and Falkow, S., Mode of replication of the conjugative R-plasmid R6K in *Escherichia coli. J. Bacteriol.*, 1976, **126**, 454–456.
4. Lovett, M., Sparks, R. A. and Helinski, D. R., Bidirectional replication of plasmid R6K DNA in *E. coli*: Correspondence between origin of replication and a single strand break in relaxed complexes. *Proc. Natl. Acad. Sci. USA*, 1975, **72**, 2905–2909.
5. Inuzuka, N., Inuzuka, M. and Helinski, D. R., Activity *in vitro* of three replication origins of the antibiotic resistance plasmid RSF1040. *J. Biol. Chem.*, 1980, **255**, 11071–11074.
6. Shon, M., Germino, J. and Bastia, D., The nucleotide sequence of the replication origin beta of the plasmid R6K. *J. Biol. Chem.*, 1982, **257**, 13823–13827.
7. Germino, J. and Bastia, D., Primary structure of replication initator protein of plasmid R6K. *Proc. Natl. Acad. Sci. USA*, 1982, **79**, 5475–5479.
8. Germino, J. and Bastia, D., The replication initiator protein of plasmid R6K tagged with b galactosidase shows sequence specific DNA binding. *Cell*, 1983, **32**, 131–140.
9. Germino, J. and Bastia, D., Interaction of plasmid R6K encoded replication initiator protein with its binding site on DNA. *Cell*, 1983, **34**, 125–134.
10. Germino, J. and Bastia, D., Rapid purification of a cloned gene product by genetic fusion and site-specific proteolysis. *Proc. Natl. Acad. Sci. USA*, 1984, **81**, 4692–4696.
11. Stalker, D. M., Kolter, R., Helinski, D. R., Tolun, A. and Helinski, D. R., Plasmid R6K DNA replication. I. Complete nucleotide sequence of an autonomously replicating segment separation of the minimal replication region of the F plasmid into a replication origin segment and a trans-acting segment. *J. Mol. Biol.*, 1982, **161**, 33–43.
12. Kolter, R., Inuzuka, M. and Helinski, D. R., Trans-complementation-dependent of a low molecular weight ori fragment from plasmid R6K. *Cell*, 1978, **15**, 1199–1258.
13. Mukherjee, S., Patel, I. and Bastia, D., Conformational changes in a replication origin induced by an initiator protein. *Cell*, 1985, **43**, 189–197.
14. Lu, Y. B., Datta, H. J. and Bastia, D., Mechanistic studies of initiator–initiator interaction and replication initiation. *EMBO J.*, 1998, **17**, 5192–5200.
15. Zzaman, S., Reddy, J. M. and Bastia, D., The DnaK–DnaJ–GrpE chaperone system activates inert wild type pi initiator protein of R6K into a form active in replication initiation. *J. Biol. Chem.*, 2004, **279**, 50886–50894.
16. Abhyankar, M. M., Reddy, J. M., Sharma, R., Bullesbach, E. and Bastia, D., Biochemical investigations of control of replication initiation of plasmid R6K. *J. Biol. Chem.*, 2004, **279**, 6711–6719.
17. Abhyankar, M. M., Zzaman, S. and Bastia, D., Reconstitution of R6K DNA replication *in vitro* using 22 purified proteins. *J. Biol. Chem.*, 2003, **278**, 45476–45484.
18. Mukherjee, S., Erickson, H. and Bastia, D., Detection of DNA looping due to simultaneous interaction of a DNA-binding protein with two spatially separated binding sites on DNA. *Proc. Natl. Acad. Sci. USA*, 1988, **85**, 6287–6291.
19. Mukherjee, S., Erickson, H. and Bastia, D., Enhancer–origin interaction in plasmid R6K involves a DNA loop mediated by initiator protein. *Cell*, 1988, **52**, 375–383.
20. Miron, A., Mukherjee, S. and Bastia, D., Activation of distant replication origins *in vivo* by DNA looping as revealed by a novel mutant form of an initiator protein defective in cooperativity at a distance. *EMBO J.*, 1992, **11**, 1205–1216.
21. Ratnakar, P. V. A. L., Mohanty, B. K., Lobert, M. and Bastia, D., Replication initiator protein pi of the plasmid R6K specifically interacts with the host encoded helicase DnaB. *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 5522–5526.
22. Abhyankar, M. M., Reddy, J. M., Sharma, R., Bullesbach, E. and Bastia, D., Biochemical investigations of control of replication initiation. *J. Biol. Chem.*, 2003, **276**, 45467–45484.
23. Brewer, B. J. and Fangman, W. L., The localization of replication origins on ARS plasmids in *S. cerevisiae*. *Cell*, 1987, **51**, 463–471.
24. McEachern, M., Bott, M. A., Tooker, P. A. and Helinski, D. R., Negative control of plasmid R6K replication : possible role of intermolecular coupling of replication origins. *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 7942–7946.
25. Pal, S. K. and Chattoraj, D. K., P1 plasmid replication: Initiator sequestration is inadequate to explain control by initiator-binding sites. *J. Bacteriol.*, 1988, **170**, 3554–3560.
26. Bastia, D. and Mohanty, B. K., Mechanisms for completing DNA replication. In *DNA Replication in Eukaryotic Cells* (ed. DePamphilis, M.), Cold Spring Harbor Laboratory Press, Plainview, NY, 1996, pp. 177–215.
27. Khatri, G. S., MacAllister, T., Sista, P. R. and Bastia, D., The replication terminator protein of *E. coli* is a DNA sequence-specific contra-helicase. *Cell*, 1989, **59**, 667–674.

28. Lee, E. H., Kornberg, A., Hidaka, M., Kobayashi, T. and Horiuchi, T., *Escherichia coli* replication termination protein impedes the action of helicases. *Proc. Natl. Acad. Sci. USA*, 1989, **86**, 9104–9108.
29. Bussiere, D. E., Bastia, D. and White, S. W., Crystal structure of replication terminator protein of *B. subtilis* at 2.6 Å. *Cell*, 1995, **80**, 651–660.
30. Kamada, K., Horiuchi, T., Ohsumi, K., Shimamoto, M. and Morikawa, K., Structure of a replication terminator protein complexed with DNA. *Nature*, 1996, **383**, 598–603.
31. Mulugu, S. *et al.*, Mechanism of termination of DNA replication of *Escherichia coli* involves helicase–contrahelicase interaction. *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 9569–9574.