

α and β replication origins of plasmid R6K show similar distortions of the DNA helix *in vivo*

(direction of replication/synchronization of initiation/DNA helix distortion)

YEHUDA FLASHNER AND AVIGDOR SHAFFERMAN*

Department of Biochemistry, Israel Institute for Biological Research, P.O. Box 19, Ness-Ziona 70450, Israel

Communicated by Donald R. Helinski, August 13, 1990 (received for review December 12, 1989)

ABSTRACT Plasmid R6K contains inverted repeats of an ≈ 100 -base-pair sequence separated by 5.5 kilobases. These long inverted repeats (LIRs) occur within the α and β origins of replication and are essential for origin function. In this study, primer-extension analysis of DNA modified *in vivo* by dimethyl sulfate or KMnO_4 revealed that both α and β LIRs acquire similar structural distortions of the DNA helix in a functional R6K replicon. These distortions were not seen in plasmids containing isolated LIR sequences. In the functional replicon, the dimethyl sulfate and KMnO_4 hyperreactive sites appear on complementary strands and are located to one side of an internal palindromic sequence within the LIRs. This asymmetry coincides with the primary direction of DNA replication from α and β origins *in vivo*. We also observed two intermediate structures when certain R6K cis- or trans-acting elements are missing. Sequences near the α origin are required for generation of the dimethyl sulfate hyperreactive sites, whereas sequences near the β origin are responsible for the appearance of KMnO_4 hyperreactive sites. We suggest that these structures represent a hierarchy that leads to a "locked" preinitiation complex, which functions to synchronize and determine the direction of replication from the α and β replication origins *in vivo*.

DNA replication in eukaryotes and some prokaryotes can be initiated from multiple DNA origins. Common to all these systems is the requirement for mechanisms of synchronization and coordination of initiation of DNA replication events. This fundamental biological problem has been the target of extensive research for over a decade, in which various approaches and model systems have been used (1). One such model system is provided by the prokaryotic plasmid R6K that carries three origins of replication designated α , β , and γ (2, 3).

The replication region of R6K, containing all three functional origins extends for ≈ 5.5 kilobases (kb) (4, 5) (Fig. 1). The α and β origins are used *in vivo* with similar frequencies, whereas the γ origin is generally silent (3). The basic R6K DNA replication elements required for function of all three R6K origins of replication are as follows: (i) the seven 22-base-pair (bp) direct repeats (7DR) near the *Hind*-9/4 junction (7, 8) and (ii) the R6K *pir* gene coding for the initiator π protein (9–12), which binds to the 22-bp sequences (13, 14). The α and β origins differ in their structural properties. Functional α origin can be delineated into two R6K elements separated by 3000 bp: the 277-bp fragment carrying the 7DR and a 580-bp fragment in *Hind*III-4 (Fig. 1) (5). Functionality of the α origin depends on the wild-type orientation of these two genetic elements in cis. The essential π initiator protein may be provided in trans for initiation of DNA replication at the α or γ origins (5, 11). A functional β origin depends on a

contiguous and uninterrupted R6K sequence of 1964 bp (4), which is composed of the 7DR region, the *pir* gene, the *bis* gene (15), and the β -origin sequence. *In vivo* DNA replication from α and β origins is sequentially bidirectional. Replication first proceeds in one direction only, and when the replication fork reaches the asymmetrically located DNA replication terminus, replication proceeds in the opposite direction from the origin (16–18). The primary directions of initiation from α and β are in opposite orientations.

Recently we have found that the α origin of replication contains a sequence of ≈ 100 bp, which appears as an inverted repeat in the β origin (5). These α and β long inverted repeats (LIRs) are 96% identical, and each is essential for functionality of the respective origin. According to our working hypothesis, the LIR sequences acquire topological structures *in vivo* that are involved in initiating and modulating R6K DNA replication. In this study we test our hypothesis by using primer extension to probe DNA modified *in vivo* by dimethyl sulfate (DMS) and by potassium permanganate (KMnO_4) (19–22). The hyperreactivity of purines (preferentially guanine residues) to DMS modification and of pyrimidines to KMnO_4 treatment is attributed to DNA distortions that make these residues more accessible to attack by the modifying agents (20, 22). We show here, that the DNA of both α and β LIRs (α LIR and β LIR, respectively) in a functional R6K replicon are helically distorted at specific and identical sites. Furthermore, analysis of R6K DNA elements that contribute to the *in vivo* formation of helix distortion in LIR reveals a possible hierarchy of discrete topological structures that may be associated with the preinitiation of DNA replication in plasmid R6K.

MATERIALS AND METHODS

Bacterial Strains and Plasmids. The following strains of *Escherichia coli* were used: LE392 (F^- *hsdR514* (r^- , m^+) *supE44* *supF58* *lacY1* or Δ (*lacIZY*)6 *galK2* *galT22* *metB1* *trpR55* λ^-) and C2110 (*polA1* *his* *rha*). Plasmids pAS910 and pAS895 have been described (4, 23), and their R6K sequences indicated (see Fig. 1). pAS807A is a pACYC184 derivative containing R6K DNA sequences depicted in Fig. 1. Plasmids p2AC and p4AC were constructed by insertion of R6K *Hind*III-2 or *Hind*III-4 DNA fragments (Fig. 1), respectively, into the specific pACYC184 *Hind*III site. pAS909 was generated from pAS910 (Fig. 1) by *Eco*RI (partial digestion) and *Sal* I, the protruding ends were filled in by DNA polymerase I (Klenow enzyme) and then ligated with T4 DNA ligase (Fig. 4). A 4200-bp *Cla* I–*Bgl* II DNA fragment from pRK43 (18) was inserted in pACYC184 between the *Cla* I and *Bam*HI sites to generate pAS913 (Fig. 4). The R6K sequence from position 243 to 2548 was isolated as a *Sna*BI–

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: LIR, long inverted repeat; DMS, dimethyl sulfate; α LIR and β LIR, α and β LIRs, respectively; 7DR, seven 22-base-pair direct repeats.

*To whom reprint requests should be addressed.

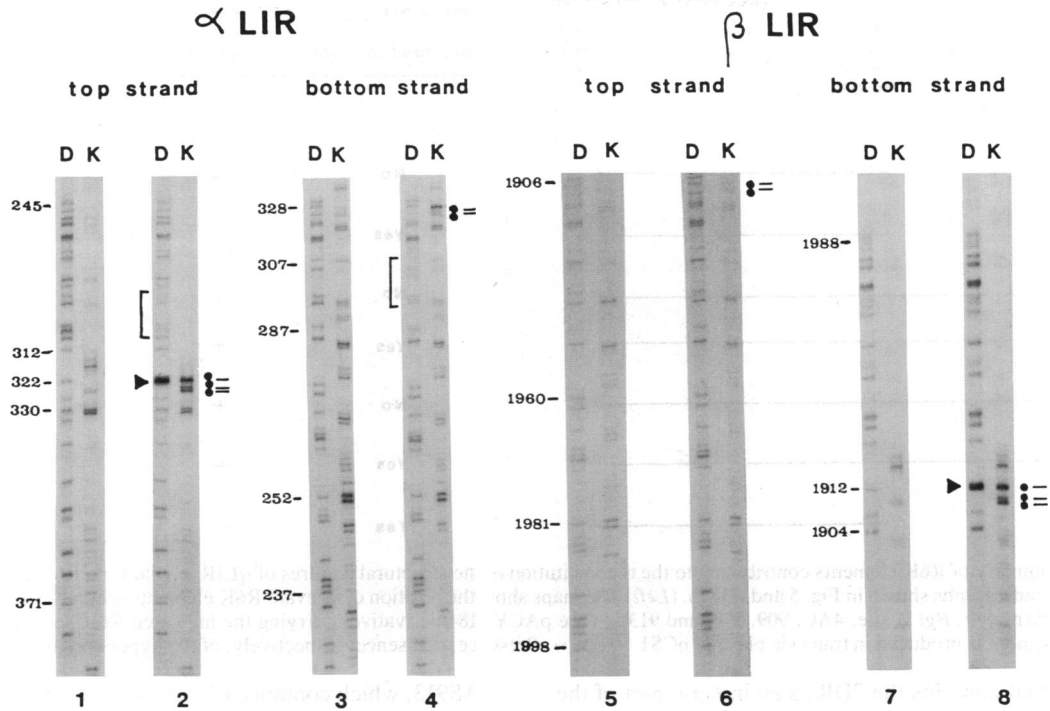


FIG. 2. Primer-extension autoradiographs of α - and β LIR sequences after *in vivo* DMS or KMnO_4 modifications. Cells harboring p4AC (sets 1 and 3), p2AC (sets 5 and 7), or pAS807A (sets 2, 4, 6, and 8) were treated with either DMS (D) or KMnO_4 (K). Nucleotides along the α - or β LIR are numbered according to ref. 5. Brackets refer to protected region. \blacktriangleright , DMS hypersite; \bullet , KMnO_4 hypersites.

Various derivatives of p4AC were constructed, but we were unable to detect the DMS or the permanganate hypersites observed for functional R6K replicons (pAS807A and R6K) unless an R6K element left of α LIR was also present, as demonstrated by plasmid pAS909 (Fig. 4). Plasmid pAS909 contains, in addition to the R6K *Hind*III-4 fragment, an R6K 2-kb fragment located between the *Hind*III-4 and the *Bgl* II site in *Hind*III-7. When cells harboring pAS909 were treated with DMS, a hypersite in α LIR on the top strand at position 322 was detected (Fig. 5 lane 2D). As shown above (Fig. 2 lanes 1 and 3), the p4AC did not reveal DMS hyperreactivity at this position. In contrast, the primer-extension patterns of pAS909 after KMnO_4 treatment were identical to those of p4AC (Fig. 5 lanes 2K and 1K). This result shows that the topological structure found in α LIR in pAS807A (or in R6K) can be dissected into intermediate discrete substructures in which the DMS hypersite at position 322 exists (in pAS909) in the absence of KMnO_4 hypersites. To attempt to reconstitute the KMnO_4 hypersites in pAS909, we provided in

trans (via plasmid pCS1) the *pir-bis-beta*LIR sequences. Under these conditions the KMnO_4 hypersites on both strands of α LIR could be generated (top strand displayed in Fig. 5 lane 4K). KMnO_4 hypersites could not be observed in p4AC, even when the *pir-bis-beta*LIR element was provided in trans by pCS1 (Fig. 5 lane 1). This result suggests that to generate the KMnO_4 hypersites via *pir-bis-beta*LIR transcomplementation, it is essential first to generate the adjacent DMS hypersite in nucleotide 322 (Fig. 3).

To generate a DMS protection of nucleotides in α LIR an additional R6K element was required—the 7DR region. When plasmid pAS910 was present in a cell together with plasmid pCS1 (to provide *pir-bis-beta*LIR) we observed clearly the protection of nucleotides between positions 292 and 306 (Fig. 5 lanes 5D and 6D). In cells harboring only pAS910 we detected neither the DMS protection nor the KMnO_4 hypersites, (Fig. 5 lanes 5D and K). To determine whether DMS protection in α LIR required the presence of the 7DR region in cis or in trans, cells harboring pAS909 were transformed

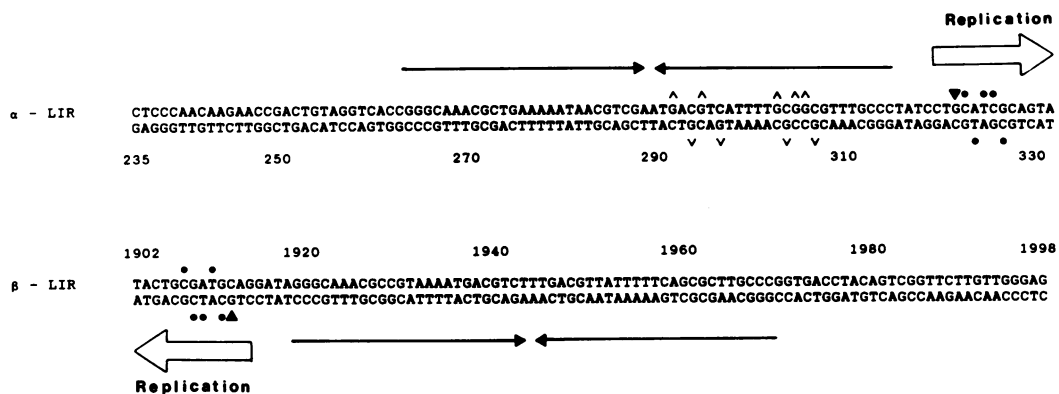


FIG. 3. DMS and KMnO_4 hypersites and DMS-protected nucleotides in the sequence of α - and β LIRs. \blacktriangledown , DMS hypersite; \bullet , KMnO_4 hypersites. \blacktriangle , Guanine residues protected from DMS modification. Thin arrows mark the palindromic region in α - and β LIRs. Thick arrow points to the initial primary direction of replication from each origin according to Crosa *et al.* (2, 24).

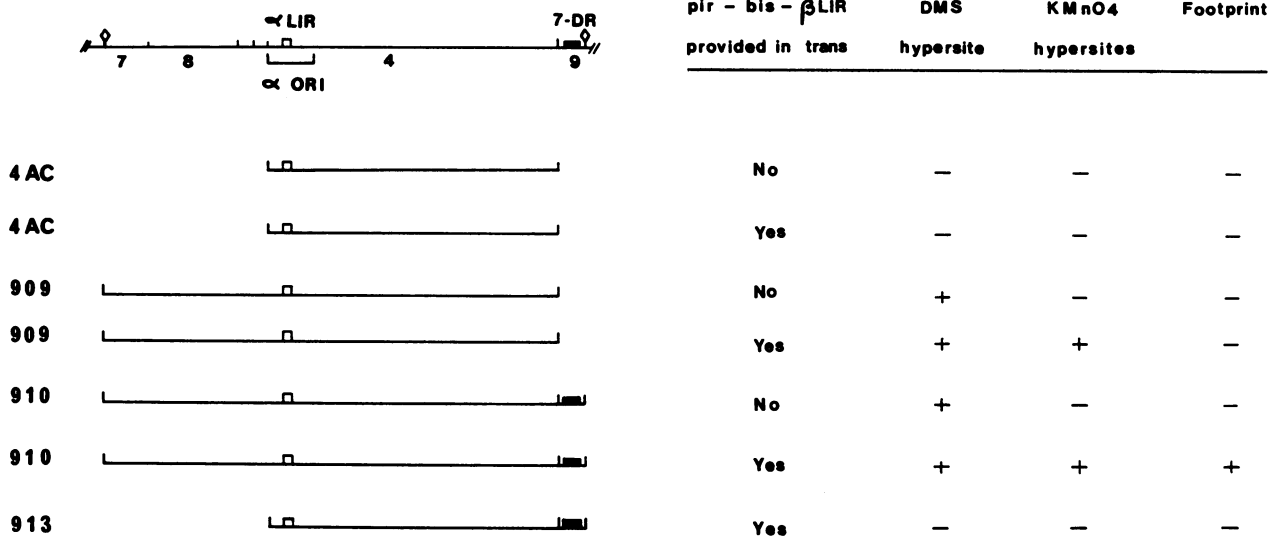


FIG. 4. A summary of R6K elements contributing to the reconstitution of the structural features of α LIR in wild-type R6K replicon. The table is based on autoradiographs shown in Fig. 5 and others. (Left) The maps show the position of relevant R6K elements around α LIR. Numbers refer to *Hind*III fragments; \diamond , *Bgl* II site. 4AC, 909, 910, and 913 are the pACYC184 derivatives carrying the indicated R6K sequences. (Right) The *pir-bis- β LIR* element is produced in trans via plasmid pCS1. + and -, Presence or absence, respectively, of the hypersites and footprint in α LIR.

with pAS895 that contains the 7DR as an integral part of the functional β replicon (Fig. 1). In such cells (harboring both pAS909 and pAS895) the DMS pattern in α LIR was identical to those found in cells harboring pAS909 and pCS1 (Fig. 5 lane 4D), and no clear footprint in α LIR was detected. These results suggest that for DMS protection of α LIR sequences the 7DR is required to be in cis to the LIR sequences.

Cells harboring both pAS909 and pAS895 were used to test the ability of β LIR to acquire DNA distortions. In these cells the β LIR DMS and KMnO₄ hypersites were indistinguishable from those found in the functional R6K replicon (Fig. 5 lane 8). However, in cells harboring pAS895 alone neither type of hypersites was seen, and the DMS and KMnO₄ patterns did not differ from those found in p2AC (Fig. 2 lanes 5 and 7). Moreover, in cells harboring both p4AC and pAS895 no hypersite could be revealed in β LIR (Fig. 5 lane 7). These results suggest that (i) like α LIR, β LIR also depends on the R6K 2-kb element left of α LIR for formation of the basic structure related to the DMS hypersite at position 1912. (ii) The factor(s) encoded by the 2-kb fragment can be supplied in trans for generating the DMS hypersite in β LIR. It is interesting that, although pAS895 carries a functional β replicon, it does not exhibit the topological β LIR structures found in the intact R6K replicon. Likewise, in plasmid

pAS913, which contains a functional α origin but is devoid of the 2-kb element left to α LIR, the DMS and KMnO₄ hypersites in α LIR were not observed. This was also true even when the *pir* gene product (which can initiate replication from the α origin in pAS913) is provided in trans to pAS913 via plasmid pCS1 (Fig. 4). Thus, unless the 2-kb element is present in the cell, the isolated functional α or β origins do not display the topological structures in LIR.

DISCUSSION

We have reported (5) that the inverted repeats, α LIR and β LIR, are required for functionality of the α and β replicons, respectively. We now show that similar topological constraint structures are found within the α LIR and β LIR sequences *in vivo* (Fig. 3). The fact that the hypersites are found on both strands within a short stretch of DNA in LIR would be consistent with an untwisted or melting of the DNA helix in this region. The *in vivo* distorted DNA structures in α - or β LIR are asymmetrically located relative to the palindromic sequence in LIR (Fig. 3). This asymmetry coincides with the *in vivo* primary direction of initiation of DNA replication (Fig. 3) from the α origin (clockwise) or the β origin (counterclockwise) (24). These topological structures in α - and β LIRs are generated in the intact functional R6K replicon (pAS807A or

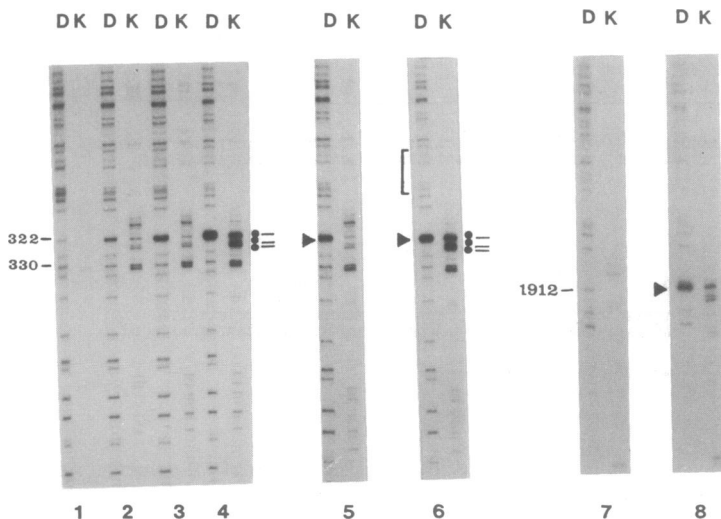


FIG. 5. *In vivo* reconstitution of DMS (D) and KMnO₄ (K) patterns in α LIR and β LIR in various R6K derivatives. Analysis of α LIR sequences by lanes: 1, p4AC in the presence of pCS1 (providing *pir-bis- β LIR* in trans); 2, pAS909; 3, pAS910; 4, pAS909 in the presence of pCS1; 5, pAS910; 6, pAS910 in the presence of pCS1. Analysis of β LIR sequences in lanes: 7, pAS895 in the presence of p4AC; 8, pAS895 in the presence of pAS909. The R6K sequences of the various constructs are shown in Figs. 1 and 4. LE392 cells harboring the different plasmids were treated as described. \blacktriangle , DMS hypersite; \bullet , KMnO₄ hypersites. Brackets refer to the DMS-protected region.

R6K) but not in naked LIR sequences (p4AC or p2AC, Fig. 2). From the above observations, it would be tempting to speculate that the helically distorted DNA sequences are carried on molecules "caught" in the early stages of the DNA replication process. However, DMS and KMnO_4 hypersites can be generated in nonfunctional α origin pAS909 (Fig. 4), as well as in nonfunctional β origin (data not shown). We, therefore, conclude that the topological structures in LIR do not reflect active replication in R6K molecules.

Although the topological structures are not a consequence of active early DNA initiation events, the following facts suggest that these structures are still intimately linked to R6K DNA replication. (i) The topological structures reside in sequences essential for functionality of α and β origins. (ii) The asymmetric location of DMS and KMnO_4 hypersites in α - or β LIR coincide with the *in vivo* primary direction of initiation from α or β origin, respectively. (iii) The formation of some of the structures and the footprint in α LIR (Fig. 5) depend in trans (*pir-bis- β LIR*) and in cis (the 7DR region) on elements known to be tightly linked to R6K DNA replication events.

We therefore suggest that the topological structures in α LIR or β LIR are related to "locked" preinitiation configurations on the R6K replicon. Together with R6K and possibly cellular factors as well, the locked preinitiation structures could be signals for intramolecular synchronization of DNA initiation events from the multiorigin R6K plasmid and for determining the direction of initiation in this composite replicon. Formation of such structures would not be a prerequisite for productive initiation from an isolated functional origin (compare the β replicon in pAS895 or α replicon pAS913 plus *pir*). It is possible to envisage a situation whereby the α - and β LIR, which are separated by 5.5 kb on the R6K genome, can be brought into spatial proximity by R6K, and host factors mediated DNA looping. As a consequence of this proximity, activation of one origin might preclude the simultaneous initiation from the other origin, either by simple steric hindrance and/or in combination with a factor(s) interacting simultaneously with both the similar α - and β LIR structures. It was demonstrated *in vitro* (25) that the initiation π protein can induce long-range intramolecular DNA looping in R6K via its interaction with the 22-bp repeat elements. Within the R6K genome, the sequences belonging to the 22-bp repeat family are highly abundant, and some of these are located close to β LIR and α LIR (5, 13). The *E. coli* integration host factor has also been reported to bend the DNA by its interaction with the 7DR region *in vitro* (26).

One of the predictions emerging from our model is that under conditions where topological structures in the LIR sequence are not allowed, the mode of initiation of DNA replication would not be expected to be confined to the specific initial directions found *in vivo* in the intact R6K. Indeed, DNA replication studies done *in vitro* (where these topological structures are not expected to exist due to the method of preparation of the substrate DNA) show that replication from γ , as well as from the α or the β origin, proceed unidirectionally (3). Moreover, in these studies the directionality of initiations from the β origin was found to occur at equal frequencies in clockwise and counterclockwise directions, whereas the major initial direction of replication for this origin *in vivo* is counterclockwise, in accordance with the asymmetric *in vivo* location of DMS and KMnO_4 hypersites in β LIR (see Fig. 3).

The analysis of R6K factors that promote formation of the presumably locked preinitiation topological structures in LIR revealed possible hierarchical order of intermediate structures that culminate with the final structure found in the intact R6K replicon. We suggest that the primary basic structure is that which renders the guanine residue in position 322 of α LIR hypersensitive to DMS (see pAS909 in Figs. 4 and 5). This helix distortion depends on R6K sequences within a 2-kb

fragment left of α LIR (Fig. 4). This R6K 2-kb element encodes a factor(s) that, when provided in trans, generate also the homologous DMS hypersite at position 1912 in β LIR. The structure related to DMS hypersensitive residue at position 322 in α LIR and position 1912 in β LIR appears to be the "nucleation" of subsequent structures because only on top of this structure could we detect the KMnO_4 hypersites in the adjacent nucleotides in LIR (Figs. 3 and 4). R6K factor(s) supplied in trans from the *pir-bis- β LIR* locus (Fig. 5) were required for generation of KMnO_4 hypersites but were insufficient unless the 2-kb element (near α LIR) was also present in the cell. *In vitro* DNA replication studies on *E. coli* oriC, λ phage, and simian virus 40 origins (27–32) showed that formation of the prepriming complex structure is preceded by a sequence of topological stages that result in a localized unwinding of an asymmetric nucleoprotein structure. This structure then serves to target subsequent factors that lead to production of a complete initiation complex and opening of the DNA helix.

We thank Dr. M. Leitner for preparation of DNA primers and Drs. D. Figurski, S. Cohen, and M. Filutowicz for fruitful discussions and helpful suggestions during the preparation of this manuscript. This work was performed in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Y.F.)

- Campbell, J. L. (1986) *Annu. Rev. Biochem.* **55**, 733–771.
- Crosa, J. H. (1990) *J. Biol. Chem.* **265**, 11075–11077.
- Inuzuka, N., Inuzuka, M. & Helinski, D. R. (1980) *J. Biol. Chem.* **255**, 11071–11074.
- Shafferman, A. & Helinski, D. R. (1983) *J. Biol. Chem.* **258**, 4083–4090.
- Shafferman, A., Flashner, Y., Hertman, I. & Lion, M. (1987) *Mol. Gen. Genet.* **208**, 263–270.
- Kolter, R. & Helinski, D. R. (1978) *Plasmid* **1**, 571–580.
- Stalker, D., Kolter, R. & Helinski, D. R. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1150–1154.
- Kolter, R. & Helinski, D. R. (1982) *J. Mol. Biol.* **161**, 45–56.
- Stalker, D., Kolter, R. & Helinski, D. R. (1982) *J. Mol. Biol.* **161**, 33–43.
- Inuzuka, M. & Helinski, D. R. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5381–5385.
- Kolter, R., Inuzuka, M. & Helinski, D. R. (1978) *Cell* **15**, 1199–1208.
- Shafferman, A., Kolter, R., Stalker, D. & Helinski, D. R. (1982) *J. Mol. Biol.* **161**, 57–76.
- Germino, J. & Bastia, D. (1983) *Cell* **34**, 125–134.
- Filutowicz, M., Uhlenhopp, E. & Helinski, D. R. (1986) *J. Mol. Biol.* **187**, 225–239.
- Mukhopadhyay, P., Filutowicz, M. & Helinski, D. R. (1986) *J. Biol. Chem.* **261**, 9534–9539.
- Lovett, M. A., Sparks, R. B. & Helinski, D. R. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 2905–2909.
- Crosa, J. H., Luttrupp, L. K. & Falkow, S. (1978) *J. Mol. Biol.* **124**, 443–468.
- Kolter, R. & Helinski, D. R. (1978) *J. Mol. Biol.* **124**, 425–441.
- Gralla, J. D. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3078–3081.
- Borowiec, J. A. & Gralla, J. D. (1986) *Biochemistry* **25**, 5051–5057.
- Flashner, Y. & Gralla, J. D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8968–8972.
- Sasse-Dwight, S. & Gralla, J. D. (1989) *J. Biol. Chem.* **264**, 8074–8081.
- Shafferman, A., Flashner, Y. & Hertman, I. (1983) *J. Biomol. Struct. Dyn.* **1**, 729–742.
- Crosa, J. H., Luttrupp, L. K. & Falkow, S. (1976) *J. Bacteriol.* **126**, 454–466.
- Mukherjee, S., Erickson, H. & Bastia, D. (1988) *Cell* **52**, 375–383.
- Filutowicz, M. & Appelt, K. (1988) *Nucleic Acids Res.* **16**, 3829–3843.
- Bramhill, D. & Kornberg, A. (1988) *Cell* **52**, 743–755.
- Alfano, C. & McMacken, R. (1989) *J. Biol. Chem.* **264**, 10699–10708.
- Dodson, M., McMacken, R. & Echols, H. (1989) *J. Biol. Chem.* **264**, 10719–10725.
- Roberts, J. M. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 3939–3943.
- Borowiec, J. A. & Hurwitz, J. (1988) *EMBO J.* **7**, 3149–3158.
- Schnos, M., Zahn, K., Inman, R. B. & Blattner, F. R. (1988) *Cell* **52**, 385–395.