The EMBO Journal vol.13 no.1 pp.269-276, 1994

Assembly of phage φ29 genome with viral protein p6 into a compact complex

Crisanto Gutiérrez, Raimundo Freire, Margarita Salas[1] and José M.Hermoso

Centro de Biología Molecular 'Severo Ochoa' (CSIC-UAM), Universidad Autónoma, Cantoblanco, 28049 Madrid, Spain
[1]Correponding author

Communicated by M.Salas

The formation of a multimeric nucleoprotein complex by the phage φ29 dsDNA binding protein p6 at the φ29 DNA replication origins, leads to activation of viral DNA replication. In the present study, we have analysed protein p6-DNA complexes formed *in vitro* along the 19.3 kb φ29 genome by electron microscopy and micrococcal nuclease digestion, and estimated binding parameters. Under conditions that greatly favour protein-DNA interaction, the saturated φ29 DNA-protein p6 complex appears as a rigid, rod-like, homogeneous structure. Complex formation was analysed also by a psoralen crosslinking procedure that did not disrupt complexes. The whole φ29 genome appears, under saturating conditions, as an irregularly spaced array of complexes ~200-300 bp long; however, the size of these complexes varies from ~2 kb to 130 bp. The minimal size of the complexes, confirmed by micrococcal nuclease digestion, probably reflects a structural requirement for stability. The values obtained for the affinity constant ($K_{eff} \approx 10^5$ M^{-1}) and the cooperativity parameter ($\omega \approx 100$) indicate that the complex is highly dynamic. These results, together with the high abundance of protein p6 in infected cells, lead us to propose that protein p6-DNA complexes could have, at least at some stages during infection, a structural role in the organization of the φ29 genome into a nucleoid-type, compact nucleoprotein complex.

Key words: electron microscopy/histone-like proteins/nucleoprotein complex/phage φ29/psoralen crosslinking

Introduction

The genetic material of both prokaryotic and eukaryotic organisms does not exist, normally, as a naked nucleic acid molecule. Rather, it usually appears associated with very abundant, low molecular weight, basic proteins that allow packaging of the DNA molecule in a reduced space. In eukaryotes, wrapping of DNA around histone octamer cores determines the well-known structure of chromatin (van Holde, 1989; Wolffe, 1992). In some cases, nucleosomes act as general repressors of both transcription (Grunstein, 1990; Felsenfeld, 1992; Svare and Hörz, 1993; Workman and Buchman, 1993) and initiation of DNA replication (Simpson, 1990; DePamphilis, 1993). In prokaryotes, the so-called histone-like proteins, of which the most abundant and best characterized are HU and H-NS (also called H1), organize the DNA into a highly condensed structure and have functional implications through their effect on DNA topology (Drlica and Rouviere-Yaniv, 1987; Pettijohn, 1988; Schmid, 1990). However, the structural details of such complexes and the role of the proteins involved in their formation still remain unclear.

The 19 285 bp linear genome of the *Bacillus subtilis* bacteriophage φ29 encodes a small (103 amino acid) early protein, p6, that is absolutely required for φ29 DNA replication *in vivo* and greatly stimulates initiation of φ29 DNA replication *in vitro* (reviewed in Salas, 1991). Furthermore, it has been shown that protein p6 inhibits transcription both *in vivo* and *in vitro* from the early C2 promoter (Whiteley *et al.*, 1986; Barthelemy *et al.*, 1989). Protein p6 forms dimers in solution (Pastrana *et al.*, 1985), interacts with double-stranded DNA (dsDNA) (Prieto *et al.*, 1988) and shows a moderate specificity for the φ29 DNA replication origins, located at the genome ends, where it forms a multimeric nucleoprotein complex which activates the initiation of φ29 DNA replication (Serrano *et al.*, 1989). The structure of the complex has been inferred by footprinting studies that showed a repeated pattern of a 24 bp binding unit, with two contact sites flanked by two DNase I hypersensitive sites in each DNA strand, indicating that protein p6 binds to dsDNA as a dimer (Serrano *et al.*, 1990). In addition, protein p6 induces the formation of positive supercoils in covalently closed circular DNA (Prieto *et al.*, 1988). These results strongly suggest a model in which DNA forms a right-handed solenoid wrapping around a multimeric protein p6 core (Serrano *et al.*, 1990). Furthermore, the parameters defining the superhelical path described by the DNA in the complex have been determined (Serrano *et al.*, 1993a). Thus, 2.6 protein p6 dimers are wrapped by 63 bp of DNA, building up one superhelical turn in which DNA is compacted 4.2-fold, is bent 66° every 12 bp, and is undertwisted (11.5 bp/turn).

It has been estimated that 1.5×10^6 protein p6 dimers are present per cell at late times of infection (M.Serrano, C.Gutiérrez, R.Freire, A.Bravo, M.Salas and J.M. Hermoso, submitted). Assuming that the average length and diameter of a *B.subtilis* cell are 3.3 and 1 μm, respectively (Luria, 1960), the intracellular concentration of protein p6 dimers may be as high as 1 mM. This led us to analyse whether and how protein p6 is able to form multimeric complexes *in vitro* with the whole φ29 genome. Thus, in this paper we have studied φ29 DNA-protein p6 complex formation using nuclease digestion and electron microscopy after psoralen crosslinking. The results obtained indicated that, under conditions which highly favour binding of protein p6 to DNA, multiple complexes appear scattered, covering most of the whole φ29 genome. We propose that the effect of protein p6 on φ29 DNA replication and transcription may be due to a more general role in genome organization and, thus, protein p6 might behave as a phage counterpart of histone-like proteins.

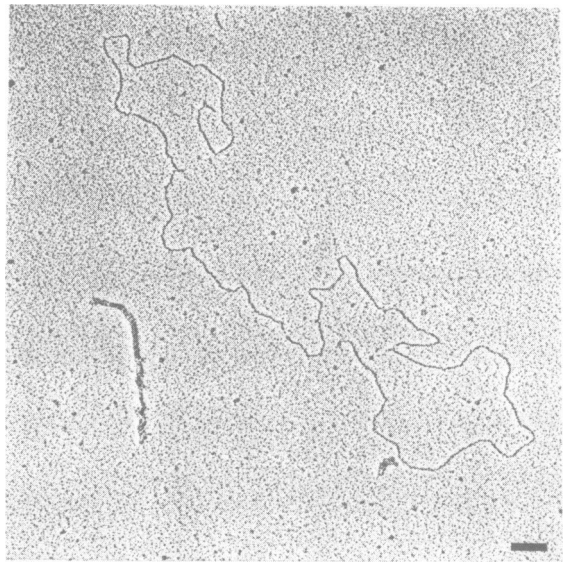

Fig. 1. ϕ29 DNA−protein p6 complexes. Protein p6 (12.6 μM, dimers) was incubated with ϕ29 DNA (0.5−2 ng) and samples fixed with glutaraldehyde and adsorbed directly to mica as described in Materials and methods. One complexed and one free DNA molecule from the same grid are shown. Bar corresponds to 500 bp.

Results

Electron microscopy of saturated ϕ29 DNA−protein p6 complexes after glutaraldehyde fixation

Phage ϕ29 DNA molecules fully covered with protein p6 were visualized after adsorption to mica and compared with protein-free molecules (Figure 1). In the absence of the fixation treatment, only uncomplexed DNA molecules were found. The nucleoprotein filaments appeared as rod-like, rigid structures with a rather homogeneous morphology throughout their length, suggesting a lack of a higher-order structure of the nucleoprotein complex. The DNA molecules analysed appeared either naked or fully saturated with protein p6, the latter being ~10% of the total molecules. This is likely as a consequence of the instability of the complex under the crosslinking conditions with glutaraldehyde. The complexes were ~6.5-fold shorter than the naked DNA molecules. A 4.2-fold reduction in length was obtained previously for pUC19 DNA (Serrano *et al.*, 1993a) using a spreading technique in which the complexes appeared more stretched than in the direct adsorption method used here.

Analysis of ϕ29 DNA−protein p6 complexes by electron microscopy after psoralen crosslinking

To analyse at a higher resolution DNA regions involved in ϕ29 DNA−protein p6 complex formation, we used a psoralen crosslinking procedure. It is known that treatment of chromatin with psoralen, followed by short-wave (366 nm) UV light irradiation, produces specific interstrand crosslinks in nucleosome-free DNA regions. Psoralen-crosslinked DNA, in contrast to uncrosslinked DNA, appears in electron micrographs as dsDNA under denaturing conditions. This technique has allowed the unambiguous distinction between nucleosome-containing and nucleosome-free DNA regions and has been applied to the analysis of chromatin from different sources (Cech *et al.*, 1977; Sogo *et al.*, 1986; Gruss *et al.*, 1990).

Recently, we have shown that binding of protein p6 to dsDNA prevents it from extensive psoralen crosslinking (Serrano *et al.*, 1993a). Therefore, we incubated ϕ29 DNA with different amounts of protein p6, irradiated the samples in the presence of psoralen, and spread the purified DNA for electron microscopy under denaturing conditions. Three examples are shown in Figure 2. As expected, in the absence of protein p6, ϕ29 DNA molecules appeared as dsDNA since the multiple crosslinks introduced by psoralen prevented them from denaturation (Figure 2A). Occasionally, very small single-stranded DNA (ssDNA) bubbles were also present as a result of psoralen crosslinking failure (see quantitative study below). When protein p6 was bound to ϕ29 DNA at concentrations 20-fold lower than saturation, DNA molecules appeared largely as dsDNA with small ssDNA bubbles, scattered over the entire molecule length (Figure 2B and C). At saturating protein p6 concentrations, the DNA molecule has interacted with protein p6 over most of its length, and both the size and, to a lesser extent, the number of ssDNA bubbles (Figure 2D and E, long arrows) increased and the proportion of dsDNA regions (Figure 2D and E, short arrows) decreased. It is worth noting that, frequently, ssDNA bubbles appeared separated by a crosslinked region (Figure 2D and E, arrowheads) whose length was below the resolution limit of our measurements. The boundary between two consecutive ssDNA bubbles was of this type in ~55% of the cases and they may well represent crosslinking events within the protein p6 complex, as will be discussed later (see also DNase I footprints in Figure 4). Therefore, in the quantitative analysis described below we have considered, instead of individual ssDNA bubbles, ssDNA tracts flanked by well-defined dsDNA regions, irrespective of the fact that they were formed by single ssDNA bubbles or by an array of contiguous ssDNA bubbles.

A quantitative analysis of both dsDNA and ssDNA regions was carried out at different protein p6 concentrations. The case corresponding to saturating protein p6 concentration is shown, as an example, in Figure 3A and B. The size distribution of the dsDNA regions was relatively narrow with a modal size of 60 bp (Figure 3A). The few ssDNA bubbles found in the absence of protein p6 had a modal size of 26 bp (Figure 3B, unfilled bars). In the presence of protein p6, as shown in Figure 3B (shaded bars), the size distribution of ssDNA tracts, indicative of protein p6−DNA complexes, significantly changed. Three aspects should be noted in such a size distribution. First, although there is a predominant ssDNA tract size of 200−300 bp (modal size of 260 bp), protein p6−DNA complexes have a wide range of sizes from ~130 bp up to 2 kb, which results in a large standard deviation of the size distribution. Second, there is no apparent indication of a bi- or trimodal nature of the ssDNA size distribution that could suggest the occurrence of repeated complexes of similar size, as is the case in eukaryotic chromatin. Third, the minimal ssDNA tract size found was 130 bp, significantly greater than the ssDNA bubbles found in the absence of protein p6. At different protein p6 concentrations, the ssDNA size distributions (not shown) were clearly as asymmetrical as that shown in Figure 3B. The variation of the modal ssDNA size with the concentration of protein p6 is shown in Figure 3C. At low protein p6 concentration, the modal size was 65 bp while at the highest concentration used, the mode was 260 bp.

To determine what, if any, was the effect of psoralen

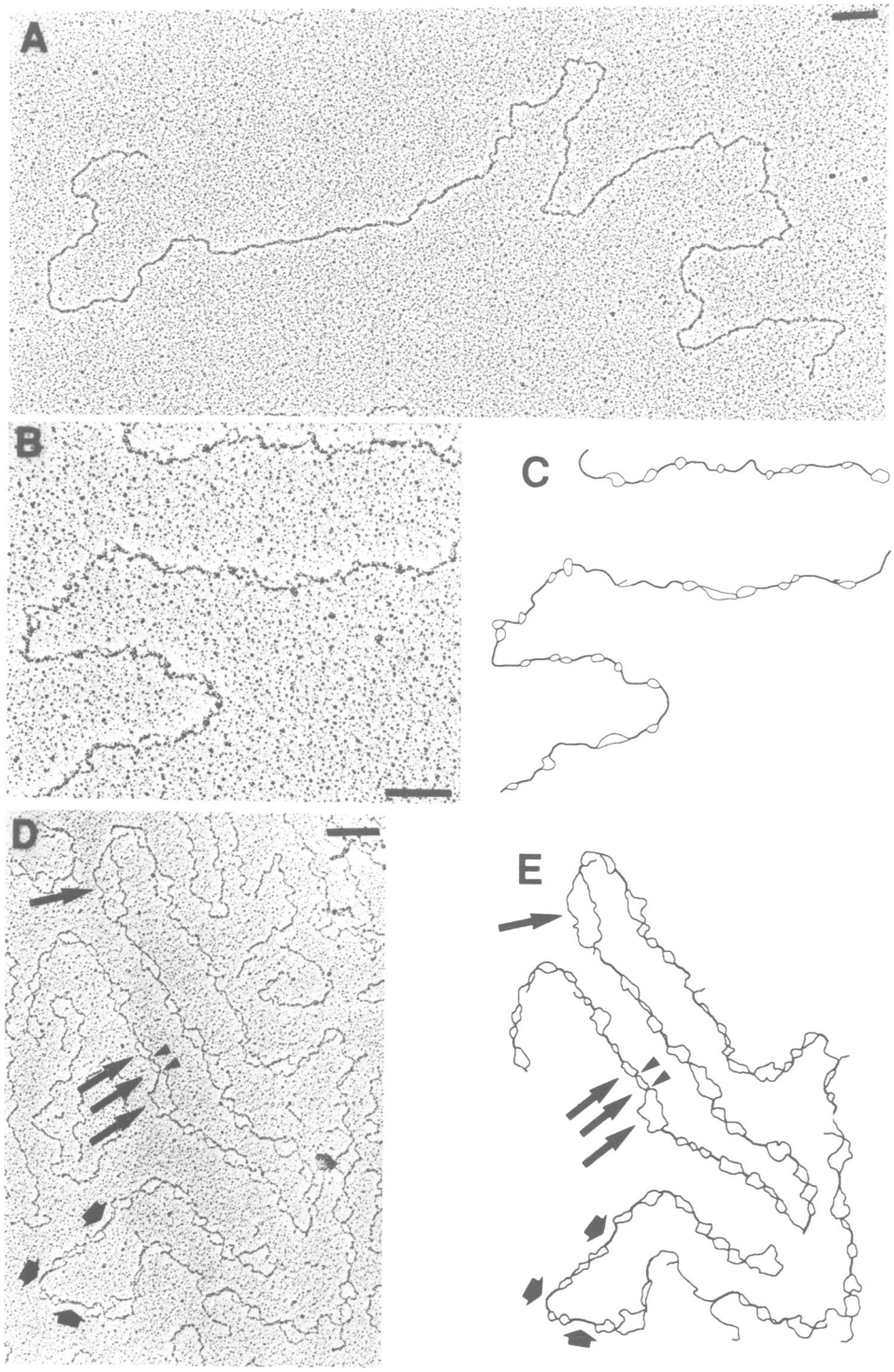

Fig. 2. Electron microscopy of psoralen crosslinked DNA–protein p6 complexes. ϕ29 DNA (1 μg) was incubated in the absence (**A**) or in the presence of protein p6 [2.1 μM (**B**) or 42 μM (**D**) dimers], and treated with psoralen, and the DNA was purified and prepared for electron microscopy as described in Materials and methods. Drawings of the molecules of interest shown in B and D appear in **C** and **E**, respectively. Short and long arrows point to dsDNA regions and ssDNA bubbles, respectively. Note that, very frequently, contiguous ssDNA bubbles (arrowheads) are present. Bars in A, B and D correspond to 500 bp.

crosslinking on the protein p6–DNA complex, we monitored complex stability by DNase I footprinting. Protein p6 bound to a DNA fragment containing the right ϕ29 DNA replication origin gives a characteristic pattern of hypersensitive bands every 24 bp (Prieto *et al.*, 1988; see also Figure 4, lane a). When protein p6–DNA complexes were subjected to psoralen crosslinking, the characteristic protein p6 footprint pattern was observed together with some discrete bands of retarded mobility (Figure 4, lane b). The presence of discrete bands instead of the smear obtained with protein-free DNA (Figure 4, lane f), indicates the existence of DNA fragments crosslinked at a few, preferred sites. To

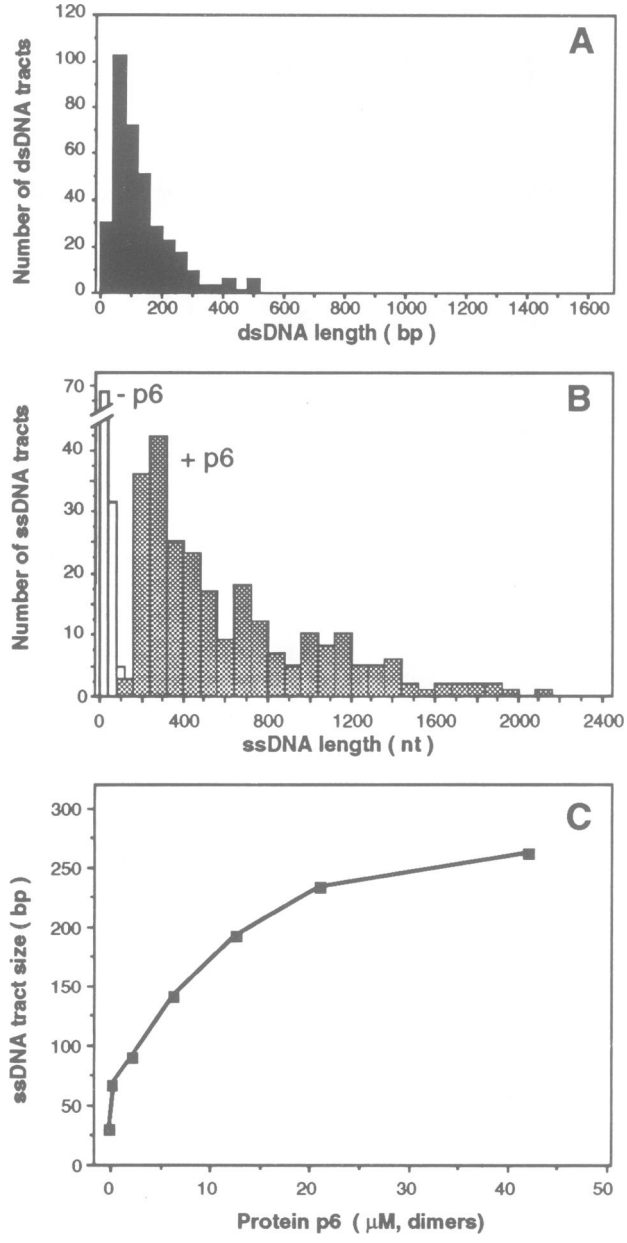

Fig. 3. DNA size distribution of dsDNA regions (**A**) and ssDNA tracts (**B**) in ϕ29 DNA molecules. Experimental conditions were as described for Figure 2D. Panel B shows the ssDNA tract size distributions found in ϕ29 DNA molecules incubated in the absence (unfilled bars) or presence (shaded bars) of protein p6 (42 μM, dimers). (**C**) Effect of protein p6 concentration on the size of ssDNA regions. Modal ssDNA values were calculated as $L + \{[\Delta_1/(\Delta_1 + \Delta_2)] \cdot c\}$ where L is the lower limit of the modal class, Δ_1 and Δ_2 are the differences of the modal class frequency with the anterior and posterior classes, respectively, and c is the size interval of the modal class.

determine whether psoralen treatment disrupts protein p6–DNA complexes, crosslinks were reversed by irradiation with 254 nm light (Sogo *et al.*, 1984). As shown in Figure 4, lanes c and d, the amount of retarded material was greatly reduced by this treatment, as was the case with protein-free DNA (Figure 4, lanes g and h). Moreover, the intensity of the characteristic hypersensitive bands increased and no extra bands appeared (Figure 4, lanes c and d) as compared with the control protein p6 footprint (Figure 4, lane a). This implies that DNase I digested both crosslinked and uncrosslinked complexes at the same sites.

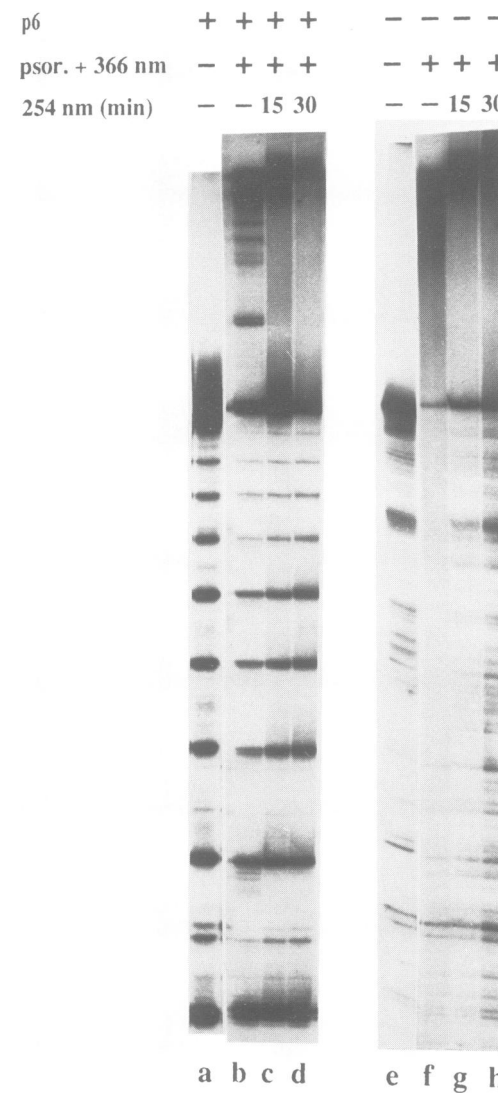

Fig. 4. DNase I footprint of ϕ29 DNA right replication origin. The 273 bp *Hin*dIII L fragment (~2 ng) from the right ϕ29 DNA end, labelled at the 3′-ends with [α-^{32}P]dATP and Klenow DNA polymerase I, was incubated without (lanes e–h) or with (lanes a–d) protein p6 (2.1 μM, dimers). Psoralen crosslinking was performed by irradiation at 366 nm as indicated. Psoralen crosslinks were reversed by irradiation at 254 nm for the indicated times (lanes c, d, g and h). After psoralen crosslinking and reversion, DNase I footprinting was carried out as described in Materials and methods.

Analysis of ϕ29 DNA–protein p6 complexes by micrococcal nuclease digestion

Digestion with nucleases, e.g. micrococcal nuclease, has been widely used to study nucleoprotein complexes as it gives useful information about structural details, such as periodicities or higher-order structures (Noll and Kornberg, 1977; Chrysogelos and Griffith, 1982; Broyles and Pettijohn, 1986). Thus, protein p6 complexes similar to those shown previously (Figure 2D) were obtained with uniformly labelled ϕ29 DNA, and digested with micrococcal nuclease. As expected, naked DNA was rapidly degraded by the nuclease treatment while in the presence of protein p6 DNA was greatly protected from degradation, even at incubation times longer than those used for naked DNA (Figure 5). In the presence of protein p6, the maximum size of digested products was reduced as the time increased, remaining fairly constant (~300 bp) between 10 and 22 min of digestion.

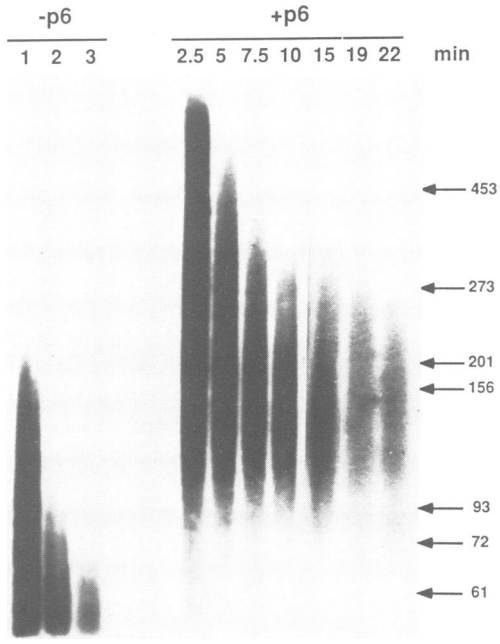

Fig. 5. Micrococcal nuclease digestion pattern of protein p6–DNA complexes. Uniformly labelled ϕ29 DNA (~16 ng) was incubated in the absence or presence of protein p6 (17.6 μM, dimers), digested with micrococcal nuclease for the indicated times, and analysed as described in Materials and methods. DNA size markers in bp are shown by arrows.

Interestingly, the minimum size remained essentially unchanged (~80–90 bp). The absence of protected fragments shorter than this size suggests that complexes below this size are unstable, in good agreement with the psoralen crosslinking results. These experiments showed neither a discrete size of protected DNA nor the presence of oligomers of a predominant DNA fragment size, strongly suggesting the formation of complexes of heterogeneous sizes, in agreement with electron microscopic observations after psoralen crosslinking.

Estimation of protein p6 binding parameters: intrinsic binding constant (K) and cooperativity parameter (ω)

A more detailed picture of the protein p6–DNA complex would include the knowledge of thermodynamic parameters directing complex formation. The intrinsic binding constant (K) and the cooperativity parameter (ω) can be determined provided that we know (i) the DNA saturation fraction (θ), that is the fraction of the DNA molecule involved in complex formation, calculated as the fraction of ssDNA with respect to the total DNA in the psoralen crosslinking experiments shown before, and (ii) the binding site size (n), that is, the number of base pairs occupied by a protein p6 dimer in the complex, which is 24 for the protein p6 dimer (Serrano *et al.*, 1990). As described in Materials and methods, both the free and bound protein p6 concentrations can be calculated, and then K and ω can be estimated graphically.

We calculated the saturation fraction when a constant amount of ϕ29 DNA was challenged with different amounts of protein p6, under conditions in which, as shown in previous sections, protein p6 binding is highly favoured. As expected, in the absence of protein p6, psoralen crosslinking was virtually complete, rendering a negligible background

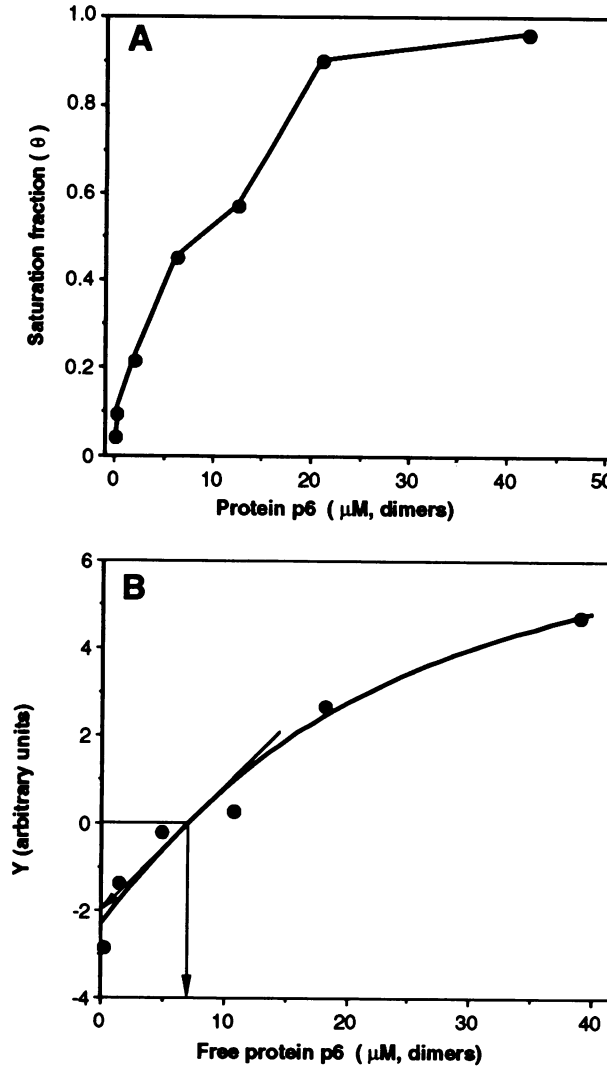

Fig. 6. Estimation of DNA saturation fraction, the effective (K_{eff}) and the intrinsic (K) binding constants of protein p6 to DNA, and its cooperativity parameter (ω). (**A**) DNA saturation values (θ) were calculated for ϕ29 DNA molecules incubated at different concentrations of protein p6 by determining the fraction of ssDNA with respect to the total ϕ29 DNA length. (**B**) Graphical estimation of binding parameters was carried out as described in Materials and methods.

(θ = 0.03 ± 0.01). Figure 6A shows that the saturation values at increasing concentrations of protein p6 followed a profile similar to that observed for the ssDNA tract sizes (Figure 3C). At low protein concentrations, when bubble size was slightly larger than in the protein-free controls, small saturation values were obtained (θ = 0.09). This value approached full saturation (θ = 0.96) at the highest protein p6 concentration used.

To estimate binding parameters, we used the procedure developed by Schwartz and Watanabe (1983), based on the model of McGhee and von Hippel (1974) for the binding of large ligands, e.g. protein p6, to a linear polymer, e.g. ϕ29 DNA. Figure 6B shows the results of such an analysis which allowed us to estimate the value of the effective binding constant of protein p6 to dsDNA ($K_{eff} \approx 10^5\ M^{-1}$) under binding conditions which highly favoured binding to dsDNA. This value represents the product of the intrinsic binding constant ($K \approx 10^3\ M^{-1}$) and the unitless cooperativity parameter ($\omega \approx 100$). Compared with other

DNA binding proteins, this is not a specially high value of K, while the value obtained for ω indicates a moderate binding cooperativity. However, it should be pointed out that these binding parameters might not necessarily be those corresponding to the DNA sequences present at the ϕ29 DNA replication origins (Serrano *et al.*, 1993a).

Discussion

Protein p6–DNA complexes along the ϕ29 genome

Multimeric nucleoprotein complexes adopt a variety of structures according to the way that proteins associate with the dsDNA molecule (Serrano *et al.*, 1993b). In addition, continuous nucleoprotein complexes, in which protein subunits associate with ssDNA to form long, multimeric complexes, are typically exemplified by proteins such as T4 gp32 (Delius *et al.*, 1972) and Rec A (Radding, 1991). In other cases, homo- or hetero-oligomers of proteins associate with DNA in clusters. The occurrence of such discrete entities separated by regions of protein-free DNA characterizes discontinuous nucleoprotein complexes. This is the case for eukaryotic chromatin, formed by an ordered array of nucleosomes (van Holde, 1989) and *Escherichia coli* single-strand binding protein (SSB) bound to ssDNA (Chrysogelos and Griffith, 1982), where a repeating unit of a constant size appears separated by a protein-free DNA region.

Phage ϕ29 protein p6 activates the initiation of ϕ29 DNA replication through the formation of a nucleoprotein complex at the replication origins (Serrano *et al.*, 1989, 1990). Complex formation was studied mainly by footprinting, where only DNA fragments of a limited size could be used. To address the question of whether complexes were formed along the ϕ29 genome at locations other than the replication origins, we have visualized the entire ϕ29 DNA molecule by electron microscopy. Thus, when protein p6–ϕ29 DNA complexes fixed with glutaraldehyde under saturating protein p6 concentrations were analysed, the entire DNA molecule appeared as a continuous complex.

To improve the resolution, we have used a psoralen crosslinking technique (Sogo and Thoma, 1989). Since protein p6 binding protects DNA very efficiently from psoralen crosslinking, complexes are visualized as ssDNA bubbles under denaturing conditions. A qualitative analysis of ϕ29 DNA molecules incubated with protein p6 showed multiple ssDNA bubbles, scattered along the ϕ29 genome even at low protein p6 concentrations, suggesting that binding is nucleated from many DNA sites. Although this pattern may seem reminiscent of the ssDNA bubbles observed in eukaryotic chromatin (Cech *et al.*, 1977; Sogo *et al.*, 1984) or in SV40 minichromosomes (Sogo *et al.*, 1986; Gruss *et al.*, 1990), protein p6–DNA complexes clearly differ from eukaryotic nucleosomes. In addition to the already known structural differences between nucleosomes (Richmond *et al.*, 1984; Hayes *et al.*, 1990; Arents *et al.*, 1991) and protein p6–DNA complexes (Serrano *et al.*, 1990, 1993a), the present study has shown further differences. The quantitative analysis of electron micrographs indicated that the length distribution of ssDNA regions is very wide, in agreement with micrococcal nuclease experiments. Furthermore, no ladder of modal bubble sizes was observed in the electron microscopic analysis, and nor was a ladder of protected DNA fragments seen in the micrococcal nuclease digestion experiments, indicating a lack of a repeating pattern of homogeneously sized complexes.

The presence of multiple, heterogeneous complexes scattered along the ϕ29 genome must be related both to the conditions used, which highly favour protein p6 binding, and to the nature of the signals recognized by protein p6. Protein p6 binding depends on the structural ability of a set of sequences to follow the right-handed superhelical path imposed by the bound protein p6 core. Therefore, the signals sensed by protein p6 must be present in DNA sequences with a tendency to be bent (or kinked) every 12 bp. The ability of a particular sequence to meet this requirement, which could be fulfilled by multiple sequences throughout the ϕ29 genome, would determine both the formation of a nucleation site and the propagation of the complex.

These observations lead us to conclude that, in general terms, protein p6–DNA complex positioning may share some of the basic principles that govern eukaryotic chromatin formation. Indeed, the same algorithm developed to predict histone octamer positioning on a given DNA sequence (Satchwell *et al.*, 1986) allowed us to predict protein p6 high-affinity binding sequences (Serrano *et al.*, 1989, 1993a). Thus, it is known that in cases such as the sea urchin 5S rRNA gene, histone octamers have some preference for interaction with certain DNA sequences, thus leading to the formation of well positioned, highly stable nucleosomes; in other cases, histone octamers may also interact more randomly with DNA, leading to the formation of nucleosomes which are not in the same positions in all DNA molecules (reviewed in Thoma, 1992).

Functional implications of ϕ29 DNA–protein p6 complexes

It is generally accepted that histone-like proteins, the major components of bacterial nucleoid, are involved in genome organization and packaging. In addition, histone-like protein–DNA complexes affect processes depending on which DNA sequences are involved in complex formation. For example, HU protein has a stimulatory effect on the initiation of DNA replication *in vitro* at the *E.coli* replication origin (Dixon and Kornberg, 1984), can either stimulate or inhibit transcription depending on the template or the conditions used (Drlica and Rouviere-Yaniv, 1987), and stimulates site-specific recombination (Haykinson and Johnson, 1993) and transposition (Craigie *et al.*, 1985). H-NS (H1), the other main prokaryotic histone-like protein, also influences recombination, transposition and gene expression (Higgins *et al.*, 1990). Although no sequence homology has been found between phage ϕ29 protein p6 and histone-like proteins, they share some properties such as being small and abundant proteins that compact DNA and show little sequence specificity for DNA binding.

It has been estimated that the intracellular concentration of protein p6 dimers is 1 mM under conditions in which 1.8×10^3 phages per infected cell are produced (Serrano *et al.*, submitted). This amount of protein would be enough to cover the intracellular viral DNA. Thus, it is conceivable that such a type of nucleoprotein complex could also be formed *in vivo*. Therefore, we propose that protein p6 has the potential to serve a structural role in ϕ29 genome organization, holding the DNA in an appropriate conformation and providing the adequate structural framework for multiple processes (see Figure 7), as happens

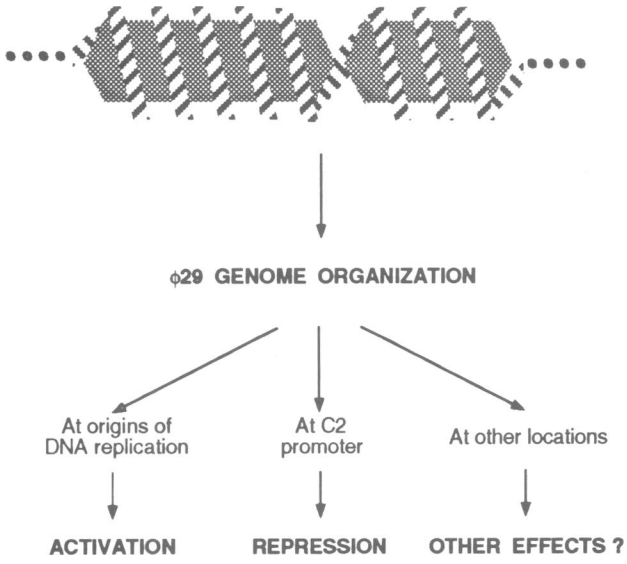

Fig. 7. Functional implications of φ29 DNA–protein p6 complexes.

with other multimeric nucleoprotein complexes (Serrano *et al.*, 1993b). Activation of φ29 DNA initiation of replication depends on complex formation at the replication origins located at the genome ends (Serrano *et al.*, 1989). Complex formation at the right DNA end may also explain the repression of φ29 C2 early promoter, observed both *in vivo* and *in vitro* (Whiteley *et al.*, 1986; Barthelemy *et al.*, 1989), as the location of protein p6 in the complex would prevent C2 promoter recognition by RNA polymerase. Complexes present at other regions throughout the genome could have further implications; thus, it is conceivable that protein p6, acting as a viral counterpart of other prokaryotic proteins which participate in genome organization, may displace host proteins, e.g. HBsu (Micka *et al.*, 1991), from the φ29 genome. A rapid increase in the free protein p6 concentration at early times after infection could ensure the replacement of cellular DNA binding proteins by the viral one. A similar situation has been proposed during adenovirus infection in which adenovirus DNA binding protein may displace nucleosomes from the viral DNA (Stuiver *et al.*, 1992). Mechanisms for replacing *E.coli* SSB protein bound to ssDNA by the phage M13 gene 5 protein (Alma *et al.*, 1983) or for sequestering the *E.coli* histone-like H-NS protein by the phage T7 gene 5.5 protein (Liu and Richardson, 1993) have been described. Once the φ29 genome is covered by protein p6, the highly dynamic nature of protein p6–DNA complexes would facilitate the continuous assembly and disassembly which must occur during DNA replication and transcription of a nucleoprotein template.

In summary, we have studied the *in vitro* assembly of the φ29 genome with viral protein p6 into a compact nucleoprotein complex and determined its main binding parameters. We propose that the effects of protein p6 in DNA replication and transcription could be the consequence of a structural role of protein p6 in organizing the φ29 genome into a nucleoid-type, compact nucleoprotein complex.

Materials and methods

DNA, proteins and chemicals

Proteinase K-treated φ29 DNA was prepared as described by Inciarte *et al.* (1976). Protein p6 was purified by J.M.Lázaro essentially as described by Pastrana *et al.* (1985). Proteinase K, micrococcal nuclease, *E.coli* DNA polymerase I and unlabelled deoxynucleotides were from Boehringer Mannheim, restriction endonucleases from New England Biolabs, 4,5′,8-trimethylpsoralen from Sigma and [α-^{32}P]dATP (400 Ci/mmol) from Amersham International. Reagents and materials for electron microscopy were from Balzers except benzyldimethylalkylammonium chloride (BAC) which was from Bayer.

Electron microscopy of protein–DNA complexes and psoralen-crosslinked DNA

To analyse φ29 DNA–protein p6 complexes, φ29 DNA was incubated with protein p6 at room temperature for 15 min in 20 μl of a solution containing 10 mM triethanolamine-HCl, pH 7.5 and 2 mM magnesium acetate. Then, glutaraldehyde was added to a final concentration of 0.6–0.8%. After 5 min at room temperature, samples were directly adsorbed to freshly cleaved mica and replicas obtained as previously described (Sogo *et al.*, 1987). Electron micrographs were taken at 80 kV, routinely at a magnification of 20 000×.

The psoralen crosslinking treatment to analyse the interaction between φ29 DNA molecules and protein p6 by electron microscopy was as described by Sogo and Thoma (1989) with some modifications (Serrano *et al.*, 1993a). Briefly, φ29 DNA was incubated, in a final volume of 20 μl, with different amounts of protein p6, as indicated, in a buffer containing 10 mM Tris–HCl, pH 7.5 and 50 mM NaCl at 0°C for 15 min. After binding, 0.05 vols of 4,5′,8-trimethylpsoralen (200 μg/ml in 100% ethanol) were added and the samples maintained on ice for 5 min in the dark. Then, the samples were irradiated with 366 nm UV light for 30 min, as described by Sogo and Thoma (1989). This irradiation time was sufficient to produce essentially complete crosslinking of the DNA molecules in the absence of protein p6. After psoralen crosslinking, the samples were digested with proteinase K (500 μg/ml) for 2 h at 56°C and extracted with phenol, and the DNA was precipitated with ethanol. Denaturation and spreading of the psoralen-crosslinked DNA for electron microscopy were carried out according to the BAC technique as described by Sogo *et al.* (1987). Electron micrographs were taken at 80 kV, routinely at a magnification of 30 000×. Contour length measurements were carried out on photographic prints using a Summagraphic-type digitizer tablet.

DNase I footprinting of protein p6–DNA complexes

DNase I footprinting (Galas and Schmitz, 1978) was carried out using the 273 bp *Hin*dIII L terminal fragment of φ29 DNA, which contains the φ29 DNA right replication origin, labelled at the 3′-end with [α-^{32}P]dATP and the Klenow fragment of *E.coli* DNA polymerase I. The labelled DNA fragment was incubated without or with protein p6 in the presence of a large excess (500 ng) of unlabelled φ29 DNA in a reaction mixture (20 μl) containing 50 mM Tris–HCl, pH 7.5, 5 mM $MgCl_2$ and 2 mM $(NH_4)_2SO_4$. After incubation at 0°C for 10 min, samples were psoralen crosslinked, when indicated, by irradiation at 366 nm UV light, during 30 min at 0°C, as described above. Then, digestion with DNase I (200 ng) was carried out for 2 min at 0°C, and the reactions stopped by adding EDTA up to 50 mM. When indicated, psoralen crosslinks were reverted by irradiation at 254 nm UV light for the indicated times. Finally, samples were extracted with phenol, the DNA precipitated with ethanol in the presence of 0.3 M sodium acetate, and subjected to denaturing electrophoresis in 6% polyacrylamide gels. After electrophoresis, gels were dried and autoradiographed.

Micrococcal nuclease digestion

φ29 DNA was labelled by nick-translation with DNA polymerase I and [α-^{32}P]dATP. DNA was incubated for 5 min on ice with protein p6 in a solution containing 50 mM Tris–HCl, pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$ and 20 mM $(NH_4)_2SO_4$. After addition of $CaCl_2$ to 2 mM and 25 U of micrococcal nuclease to a final volume of 10 μl, the samples were incubated at 25°C. At the indicated times, aliquots were taken and the reaction was stopped by addition of EDTA to a final concentration of 10 mM and electrophoresed through 8% polyacrylamide gels in Tris–borate buffer. *Hin*dIII and *Hha*I digests of φ29 DNA were run in parallel as molecular weight markers. Densitometric scans of autoradiographies were carried out using a Molecular Dynamics 300A instrument.

Estimation of binding parameters: effective binding constant (K_{eff}) and cooperativity parameter (ω)

If we assume that protein p6–DNA complex formation satisfies the requisites to apply the model of McGhee and von Hippel (1974) for protein–DNA interaction, the graphical procedure applied to the interaction of protamines with DNA (Watanabe and Schwartz, 1983) can be used for the estimation of K_{eff} and ω. Briefly, the saturation fraction data (θ) obtained in the

electron microscopic analysis as the fraction of ssDNA with respect to the total DNA, were expressed as a binding function in terms of the effective binding constant (K_{eff}) and the concentration of free protein p6 dimers ($p6_F$), as indicated below. The $p6_F$ values for each experimental situation were calculated from the mass conservation equation, $p6_F = p6_T\ p6_B$, where $p6_T$ is the total dimer concentration in the assay and $p6_B$ is the concentration of p6 dimers bound to DNA. To calculate the term $p6_B$, we used the expression $p6_B = \theta(BP/n)$, where θ is the experimentally determined saturation value, BP is the molar concentration of DNA in the assay, expressed in bp, and n is the protein p6 dimer binding site size, that is, the number of bp covered by a protein p6 dimer (n = 24, Serrano *et al.*, 1990). Thus, as described previously (Schwartz and Watanabe, 1983), the value of θ is related to K_{eff} and $p6_F$ by the following expression:

$$Y = (K_{eff} \cdot p6_F - 1)\sqrt{(\omega/n)}$$

where Y is a function of θ [$Y = (2\theta - 1)/\sqrt{\theta(1-\theta)}$]. Although the plot of Y versus $p6_F$ is a straight line, it becomes more curved for values of $\sqrt{(\omega/n)}$ smaller than ~7. However, they still provide a convenient evaluation of K_{eff} and ω. Thus, the intercept with the $p6_F$ axis (at Y = 0) gives K_{eff} as $1/p6_F$ while the intercept with the Y axis (at $p6_F = 0$) gives $\omega = n \cdot Y^2$. The intrinsic binding constant (K) of protein p6 to DNA may, then, be estimated since $K_{eff} = K\omega$.

Acknowledgements

We are indebted to J.M.Lázaro for the purification of protein p6 and to L.Villar for the preparation of proteinase K-treated ϕ29 DNA. We are also grateful to M.Serrano and J.M.Sogo for their helpful discussions and suggestions on the manuscript. This work has been supported by grants 5R01GM27242-14 from the National Institutes of Health, BIOT-CT 91-0268 from the European Economic Community and PB90/0091 from the Dirección General de Investigación Científica Técnica. The institutional help of Fundación Ramón Areces is also acknowledged. R.Freire was recipient of a predoctoral fellowship from Comunidad Autónoma de Madrid.

References

Alma,N.C.M., Harmsen,B.J.M., de Jong,E.A.M., Ven,J.v.d. and Hilbers,C.W. (1983) *J. Mol. Biol.*, **163**, 47–62.
Arents,G., Burlingame,R.W., Wang,B.C., Love,W.E. and Moudrianakis,E.N. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 10148–10152.
Barthelemy,I., Mellado,R.P. and Salas,M. (1989) *J. Virol.*, **63**, 460–462.
Broyles,S.S. and Pettijohn,D.E. (1986) *J. Mol. Biol.*, **187**, 47–60.
Cech,T., Potter,D. and Pardue,M.L. (1977) *Biochemistry*, **16**, 5313–5321.
Chrysogelos,S. and Griffith,J. (1982) *Proc. Natl Acad. Sci. USA*, **79**, 5803–5807.
Craigie,R., Arndt-Jovin,D. and Mizuuchi,K. (1985) *Proc. Natl Acad. Sci USA*, **82**, 7570–7574.
Delius,H., Mantell,N.J. and Alberts,B. (1972) *J. Mol. Biol.*, **67**, 341–350.
DePamphilis,M.L. (1993) *Trends Cell Biol.*, **3**, 161–167.
Dixon,N. and Kornberg,A. (1984) *Proc. Natl Acad. Sci USA*, **81**, 424–428.
Drlica,K. and Rouviere-Yaniv,J. (1987) *Microbiol. Rev.*, **51**, 301–319.
Felsenfeld,G. (1992) *Nature*, **355**, 219–224.
Galas,D.J. and Schmitz,A. (1978) *Nucleic Acids Res.*, **9**, 3157–3170.
Grunstein,M. (1990) *Trends Genet.*, **6**, 395–400.
Gruss,C., Gutierrez,C., Burhans,W.C., DePamphilis,M.L., Koller,T. and Sogo,J.M. (1990) *EMBO J.*, **9**, 2911–2922.
Hayes,J.J., Tullins,T.D. and Wolffe,A. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 7405–7409.
Haykinson,M.J. and Johnson,R.C. (1993) *EMBO J.*, **12**, 2503–2512.
Higgins,C.F., Hinton,J.C.D., Hulton,C.S.J., Owen-Hughes,T., Pavitt,G.D. and Seirafi,A. (1990) *Mol. Microbiol.*, **4**, 2007–2012.
Inciarte,M.R., Lázaro,J.M., Salas,M. and Viñuela,E. (1976) *Virology*, **74**, 314–323.
Liu,Q. and Richardson,C.C. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 1761–1765.
Luria,S.E. (1960) In I.C.Gunsalus and R.Y.Stanier (eds), *The Bacteria*. Academic Press, New York and London, Vol. I, p. 9.
McGhee,J.D. and von Hippel,P.H. (1974) *J. Mol. Biol.*, **86**, 469–489.
Micka,B., Groch,N., Heinemann,U. and Marahiel,M.A. (1991) *J. Bacteriol.*, **173**, 3191–3198.
Noll,M. and Kornberg,R.D. (1977) *J. Mol. Biol.*, **109**, 393–404.
Pastrana,R., Lázaro,J.M., Blanco,L., García,J. A., Méndez,E. and Salas,M. (1985) *Nucleic Acids Res.*, **13**, 3083–3100.
Pettijohn,D.E. (1988) *J. Biol. Chem.*, **263**, 12793–12796.
Prieto,I., Serrano,M., Lázaro,J.M., Salas,M. and Hermoso,J.M. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 314–318.
Radding,C.M. (1991) *J. Biol. Chem.*, **266**, 5355–5358.
Richmond,T.J., Finch,J.T., Rushton,B., Rhodes,D. and Klug,A. (1984) *Nature*, **311**, 532–537.
Salas,M. (1991) *Annu. Rev. Biochem.*, **60**, 39–71.
Satchwell,S.C., Drew,H.R. and Travers,A.A. (1986) *J. Mol. Biol.*, **191**, 659–675.
Schmid,M.B. (1990) *Cell*, **63**, 451–453.
Schwartz,G. and Watanabe,F. (1983) *J. Mol. Biol.*, **163**, 467–484.
Serrano,M., Gutiérrez,J., Prieto,I., Hermoso,J.M. and Salas,M. (1989) *EMBO J.*, **8**, 1879–1885.
Serrano,M., Salas,M and Hermoso,J.M. (1990) *Science*, **248**, 1012–1016.
Serrano,M., Gutiérrez,C., Salas,M. and Hermoso,J.M. (1993a) *J. Mol. Biol.*, **230**, 248–259.
Serrano,M., Salas,M. and Hermoso,J.M. (1993b) *Trends Biochem. Sci.*, **18**, 202–206.
Simpson,R.T. (1990) *Nature*, **343**, 387–389.
Sogo,J.M. and Thoma,F. (1989) *Methods Enzymol.*, **170**, 142–165.
Sogo,J.M., Ness,P.J., Widmer,R.M., Parish,R.W. and Koller,T. (1984) *J. Mol. Biol.*, **178**, 897–928.
Sogo,J.M., Stahl,H., Koller,T. and Knippers,R. (1986) *J. Mol. Biol.*, **189**, 189–204.
Sogo,J.M., Stasiak,A., de Bernardin,W., Losa,R. and Koller,T. (1987) In J.Sommerville and U.Scheer (eds), *Electron Microscopy in Molecular Biology. A Practical Approach*. IRL Press, Oxford, pp. 61–79.
Stuiver,M.H., Bergsma,W.G., Arnberg,A.C., van Amerongen,H., van Grondelle,R. and van der Vliet,P.C. (1992) *J. Mol. Biol.*, **225**, 999–1011.
Svare,J. and Hörz,W. (1993) *Curr. Opin. Genet. Dev.*, **3**, 219–225.
Thoma,F. (1992) *Biochim. Biophys. Acta*, **1130**, 1–19.
van Holde,K.E. (1989) *Chromatin*. Springer Verlag, Berlin.
Watanabe,F. and Schwartz,G. (1983) *J. Mol. Biol.*, **163**, 485–498.
Whiteley,H.R., Ramey,W.D., Spiegelman,G.B. and Holder,R.D. (1986) *Virology*, **155**, 392–401.
Wolffe,A. (1992) *Chromatin. Structure and Function*. Academic Press. London.
Workman,J.L. and Buchman,A.R. (1993) *Trends Biochem. Sci.*, **18**, 90–95.

Received on July 27, 1993; revised on October 11, 1993