Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 6470–6474, October 1985
Biochemistry

Formaldehyde-mediated DNA–protein crosslinking: A probe for *in vivo* chromatin structures

(simian virus 40 chromosomes/nucleosome-free region/Pronase-resistant crosslinks/*lac* repressor/α-protein)

MARK J. SOLOMON AND ALEXANDER VARSHAVSKY

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Communicated by Gary Felsenfeld, June 6, 1985

ABSTRACT **Formaldehyde (HCHO) produces DNA–protein crosslinks both *in vitro* and *in vivo*. Simian virus 40 (SV40) chromosomes that have been fixed by prolonged incubation with HCHO either *in vitro* or *in vivo* (within SV40-infected cells) can be converted to nearly protein-free DNA by limit-digestion with Pronase in the presence of NaDodSO$_4$. The remaining Pronase-resistant DNA–peptide adducts retard the DNA upon gel electrophoresis, allowing resolution of free and crosslink-containing DNA. Though efficiently crosslinking histones to DNA within nucleosomes both *in vitro* and *in vivo*, HCHO does not crosslink either purified *lac* repressor to *lac* operator-containing DNA or an (A + T)-DNA-binding protein (α-protein) to its cognate DNA *in vitro*. Furthermore, a protein that does not bind to DNA, such as serum albumin, is not crosslinked to DNA by HCHO even at extremely high protein concentrations. These properties of HCHO as a DNA–protein crosslinker are used to probe the distribution of nucleosomes *in vivo*. We show that there are no HCHO-crosslinkable DNA–protein contacts in a subset of SV40 chromosomes *in vivo* within a 325-base-pair stretch that spans the "exposed" (nuclease-hypersensitive) region of the SV40 chromosome. This replication origin-proximal region has been found previously to lack nucleosomes in a subset of isolated SV40 chromosomes. We discuss other applications of the HCHO technique, including the possibility of obtaining base-resolution *in vivo* nucleosome "footprints."**

DNA–protein interactions can be probed *in vitro* by a number of techniques, including chemical modification (1) and nuclease- or drug-mediated "footprinting" (2–5). Most of these assays are kinetic in nature, since the probes used are not absolutely specific for the structure of interest. In a different set of approaches, DNA-bound proteins are first crosslinked to DNA by treatments with UV light (6), formaldehyde (HCHO) (7), dimethyl sulfate (8), or a variety of other agents (9). Some of the above methods, in particular DNA modification by dimethyl sulfate (10, 11) or UV light (12), have been used recently to footprint DNA–protein interactions *in vivo* (in intact cells). Since both treatments damage DNA extensively upon prolonged exposure, the use of the corresponding techniques is restricted to a kinetic assay, with attendant advantages and drawbacks of such a constraint.

On the other hand, HCHO produces DNA–protein crosslinks both *in vitro* and *in vivo* (7, 13–18) and at the same time displays virtually no reactivity toward free double-stranded DNA (19, 20). Since HCHO produces DNA–protein (7, 13–16), RNA–protein (17), and protein–protein (18) crosslinks, its addition to living cells results within minutes in formation of crosslinked networks of biopolymers, thus preventing any large-scale redistribution of cellular components upon prolonged ("limit") fixation. HCHO-induced crosslinks, in particular DNA–protein crosslinks, can be reversed under relatively mild conditions, as demonstrated previously (16, 18) and further refined in the present work.

Limit-digestion of HCHO-fixed eukaryotic chromosomes with relatively nonspecific proteinases, such as Pronase or proteinase K, does not yield a completely peptide-free DNA (7, 14, 21). Moreover, DNA fragments containing Pronase-resistant peptide–DNA adducts have reduced electrophoretic mobilities (7). We have exploited this latter observation, together with other properties of HCHO, to develop an approach to probing the distribution of nucleosomes *in vivo*.

MATERIALS AND METHODS

Cell Culture, Labeling, and Infection. African green monkey CV-1 cells and simian virus 40 (SV40) (strain 777) were propagated, and DNA was labeled with [*methyl*-^{3}H]thymidine as described (7, 22). Unfixed SV40 chromosomes were released from isolated CV-1 nuclei as described (22) except that the extraction buffer was 0.25% Triton X-100/0.12 M NaCl/10 mM Na_2EDTA/10 mM sodium Hepes, pH 7.5.

***In Vitro* HCHO Fixation.** Eleven percent HCHO (from a 37% HCHO/12% CH_3OH stock solution, Malinckrodt)/0.1 M NaCl/1 mM Na_2EDTA/50 mM sodium Hepes, pH 7.5, was added directly to the SV40 chromosome-containing nuclear extract to a final HCHO concentration of 1% (0.33 M). Fixation was carried out for 4 days (100 ± 10 hr) at 4°C followed by dialysis at 4°C for 2 days against multiple changes of 1 mM Na_2EDTA/0.5 mM Na_2EGTA/10 mM sodium Hepes, pH 7.5.

***In Vivo* HCHO Fixation.** SV40-infected, [^{3}H]thymidine-labeled CV-1 cell monolayers in 15-cm plates (Falcon) at 40 hr after infection (see above) were rinsed once with 0.14 M NaCl/0.8 mM $MgSO_4$/1.8 mM $CaCl_2$/3 mM KCl/8 mM Na_2HPO_4/1.5 mM KH_2PO_4, pH 7.2, at 37°C followed by immediate addition of 20 ml of 1% HCHO in the same buffer. Alternatively, 11% HCHO/3% CH_3OH/0.1 M NaCl/1 mM Na_2EDTA/50 mM sodium Hepes, pH 7.5, was added directly to the cell's growth medium to a final HCHO concentration of 1%. The results obtained by using these protocols were indistinguishable. Plates remained at 37°C for 10 min followed by incubation at 4°C for 4 days (100 ± 10 hr).

HCHO Treatment of *lac* Repressor–DNA and α-Protein–DNA Complexes. Purified *lac* repressor (23, 24) was a gift from K. Matthews (Rice Univ., Houston, TX). Purified α-protein from green monkey CV-1 cells (25) and plasmid pFS522 containing the single 172-base-pair (bp) *Hin*dIII repeat of green monkey α-satellite DNA (25) were gifts from F. Strauss (J. Monod Institute, Paris, France). pJW270, a plasmid containing the *lac* operator (12), was a gift from J. Wang (Harvard Univ., Cambridge, MA). The 92-bp *Bam*HI–*Eco*RI DNA fragment containing the *lac* operator

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: HCHO, formaldehyde; SV40, simian virus 40; bp, base pair(s).

(12, 24) was prepared from pJW270 (12). DNA fragments were end-labeled with ^{32}P as described (25).

Restriction Endonuclease Digestions. When up to 0.1% $NaDodSO_4$ was present in the sample to be digested, 10% Triton X-100 was added to a final concentration of 1%. Most of the endonucleases tested were fully active in buffers containing 1% Triton X-100/0.1% $NaDodSO_4$.

Gel Electrophoresis. Horizontal agarose gels containing 0.1% $NaDodSO_4$ and TAE buffer (1 mM Na_2EDTA/5 mM sodium acetate/40 mM Tris·HCl, pH 8.0) were used.

RESULTS

Pronase-Resistant Crosslinks in HCHO-Fixed Chromosomes Retard DNA During Electrophoresis. Extensive HCHO fixation of both cellular (13–15) and SV40 (7, 16) chromosomes prevents subsequent dissociation of proteins in the presence of either ionic detergents or high-salt buffers. Isolated SV40 chromosomes that have been limit-fixed *in vitro* with HCHO migrate as a broad band upon electrophoresis in an agarose gel containing $NaDodSO_4$ (Fig. 1, lane 1). Limit-digestion of these HCHO-fixed SV40 chromosomes with Pronase in the presence of $NaDodSO_4$ yields a much sharper band migrating reproducibly slower than free form I SV40 DNA (Fig. 1, lanes 3 and 4; see also ref. 7). Fixed SV40 chromosomes that have been limit-digested with trypsin migrate much slower than Pronase-digested chromosomes (Fig. 1, lanes 2 and 3), apparently reflecting the amount of undigested DNA-bound peptides resistant to further trypsin treatment. Furthermore, the greater the time of HCHO fixation of chromosomes at 4°C, the greater the electrophoretic retardation observed following limit-digestion with Pronase, with the characteristic limit-mobility (Fig. 1, lane 3) approached after ≈100 hr of fixation (data not shown). Finally, HCHO fixation of purified SV40 DNA under conditions identical to those used for chromosome fixation does not affect the electrophoretic mobility of DNA relative to that of untreated naked DNA (data not shown).

Pronase-Resistant Peptide–DNA Crosslinks Can Be Thermally Reversed. Both protein–protein and protein–DNA crosslinks produced by HCHO are reversible by high-temperature treatments (16, 18). We have explored a variety of milder protocols for thermal reversal of the electrophoretic retardation observed with HCHO-fixed, Pronase-digested SV40 chromosomes (Fig. 2). For example, 2 days at 23°C has no effect on retardation (Fig. 2*A*), whereas the same treatment at 37°C partially reverses the retardation (Fig. 2*B*). On the other hand, incubation for 6 hr at 60°C virtually completely reverses the retardation (Fig. 2*C*). In Fig. 2*D* the high range of temperatures is explored in more detail; our standard protocol for thermal reversal of Pronase-resistant crosslinks adopted on the basis of these data is incubation at 65°C for 6 hr.

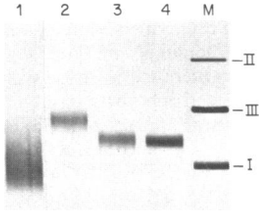

FIG. 1. Effect of HCHO-induced, Pronase-resistant DNA–protein crosslinks on electrophoretic mobility of DNA. Lane 1, isolated, HCHO-fixed SV40 chromosomes before proteolysis. Lane 2, same as lane 1 but limit-digested with trypsin. Lane 3, same as lane 1 but limit-digested with Pronase. Lane 4, same as lane 3 but heated at 65°C for 10 min after Pronase digestion. Lane M, free SV40 DNA markers [supercoiled (I), nicked (II), and linear (III)]. Electrophoresis was in a 1.3% agarose gel containing 0.1% $NaDodSO_4$. [^3H]thymidine-labeled SV40 DNA was detected by fluorography. Proteolytic limit-digestion of fixed and dialyzed chromosomes with Pronase (type VI, Sigma) at 0.1 mg/ml was carried out in the presence of 0.1% $NaDodSO_4$ at 37°C for 6 hr with readdition of an equal amount of enzyme at 3 hr. Alternatively, digestions were carried out with trypsin (lane 2) at 0.1 mg/ml but in the absence of $NaDodSO_4$. Proteolysis was terminated by addition of phenylmethylsulfonyl fluoride to 1 mM from a 0.1 M stock solution in absolute ethanol.

HCHO-Produced, Pronase-Resistant Crosslinks Leave Most Restriction Endonuclease Cleavage Sites in SV40 DNA Accessible. Free SV40 DNA (control) and HCHO-fixed, Pronase-digested chromosomes were limit-digested with various restriction endonucleases followed by complete thermal reversal of the Pronase-resistant crosslinks and gel electrophoresis (Fig. 3*A*). The following conclusions can be drawn from the data in Fig. 3. (*i*) To a first approximation, each DNA restriction site for a variety of endonucleases tested has the same probability of being protected from cleavage in HCHO-fixed, Pronase-digested chromosomes (Fig. 3*A*). (*ii*) Most blocked DNA restriction sites in HCHO-fixed, Pronase-digested chromosomes become accessible after thermal reversal of the crosslinks (Fig. 3*C*). (*iii*) Fragments produced by *Hae* III digestion of HCHO-fixed, Pronase-digested SV40 chromosomes retain the relative electrophoretic retardation observed with whole SV40 chromosomes (Fig. 3*B*; compare to Fig. 1; see also Figs. 5 and 6). Thus, to a first approximation, the linear density of crosslinks is uniform along the DNA. (*iv*) A small SV40 DNA fragment in Fig. 3*B* (lane 1, indicated by arrow) migrates as a more discrete band, with mobility identical to the mobility of the 325-bp-long, replication origin-proximal *Hae* III DNA fragment. In contrast to all of the other *Hae* III DNA fragments derived from HCHO-fixed, Pronase-digested SV40 chromosomes, this fragment is devoid of Pronase-resistant crosslinks (see below).

Serum Albumin Is Not Crosslinked to DNA by HCHO. As shown in Fig. 4*A*, incubation of SV40 DNA with 1% HCHO and bovine serum albumin at up to 50 mg/ml does not alter the electrophoretic mobility of the DNA. The concentration of albumin in the crosslinking assay (Fig. 4*A*) is comparable to the total intracellular protein concentration (estimated to be $\approx 2 \times 10^2$ mg/ml), which is lower by a factor of only 2–4 than the protein concentration in typical protein crystals (28). We conclude that HCHO does not crosslink free proteins to DNA.

***In Vitro* DNA–Protein Crosslinking by HCHO Is Highly Target-Dependent.** As shown in Fig. 4*B* (lanes 4 and 5), HCHO does not crosslink purified *lac* repressor (23) to a DNA fragment containing the *lac* operator even when *lac* repressor is present in up to a 1000-fold excess over the amount needed to bind one-half of the target operator (see *Materials and Methods* and the legend to Fig. 4*B* for details). Furthermore, as shown in Fig. 4*B* (lanes 1–3), HCHO also does not crosslink the green monkey α-satellite DNA-binding protein (α-protein; ref. 25) to α-satellite DNA even when α-protein is present in up to an 80-fold excess over the amount needed to bind one-half of the target DNA fragment.

α-Protein, a specific high-mobility group (HMG) nuclear protein of ≈10 kilodaltons, binds tightly and specifically to any stretch of pure (A + T)-DNA that is at least 6 bp long (ref. 25; unpublished data). These and other properties of α-protein indicate that the recognition of double-stranded (A + T)-DNA by α-protein is largely through contacts in the minor groove of the DNA and therefore is very different from the mechanism of DNA recognition by *lac* repressor and structurally related proteins such as *cI* and *cro* of bacteriophage λ (29).

The rapid and high-yield production of DNA–histone crosslinks by HCHO both *in vivo* and *in vitro* and the

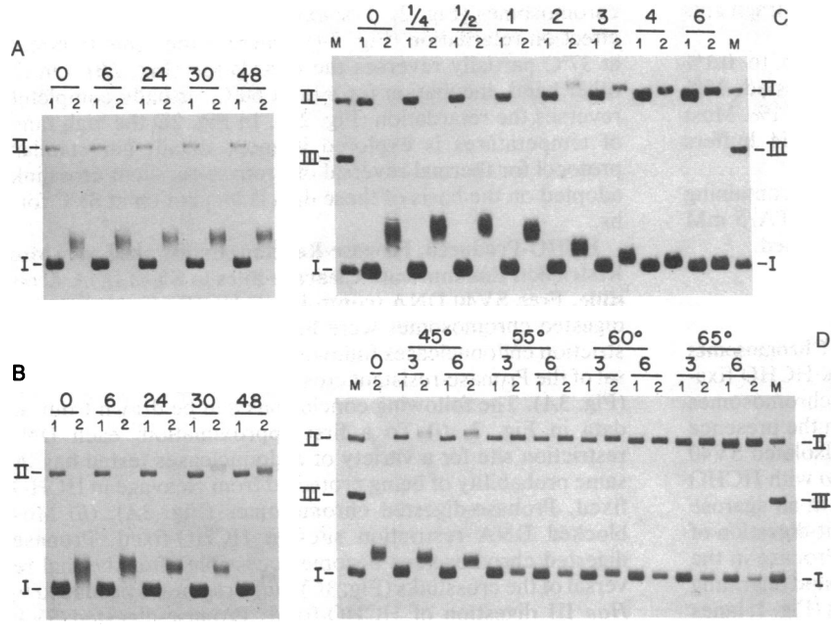

FIG. 2. Thermal reversal of HCHO-induced, Pronase-resistant DNA–protein crosslinks. Control SV40 DNA (lanes 1) and *in vitro* HCHO-fixed, Pronase-digested SV40 chromosomes (lanes 2) were incubated at various temperatures in 1 mM Na_2EDTA/10 mM Tris·HCl, pH 8.0. Incubations were carried out at 23°C (*A*), 37°C (*B*), or 60°C (*C*) for the indicated number of hours (0, 6, 24, 30, and 48 hr in *A* and *B*; 0, 0.25, 0.5, 1, 2, 3, 4, and 6 hr in *C*). In *D*, incubations were at 45°C, 55°C, 60°C, or 65°C for either 3 or 6 hr, as indicated. Electrophoresis, DNA detection, and markers (M) were as in Fig. 1.

apparently complete inability of HCHO to crosslink two strongly different nonhistone DNA-binding proteins to their cognate DNAs *in vitro* reveal the striking target selectivity of HCHO. This selectivity can be exploited to probe the distribution of nucleosomes *in vivo* as shown below.

HCHO Fixation of SV40 Chromosomes Within SV40-Infected CV-1 Cells. Although ≈50% of the total SV40 DNA is released from *in vivo* HCHO-fixed, SV40-infected CV-1 cells at intermediate extents of Pronase digestion, most of this

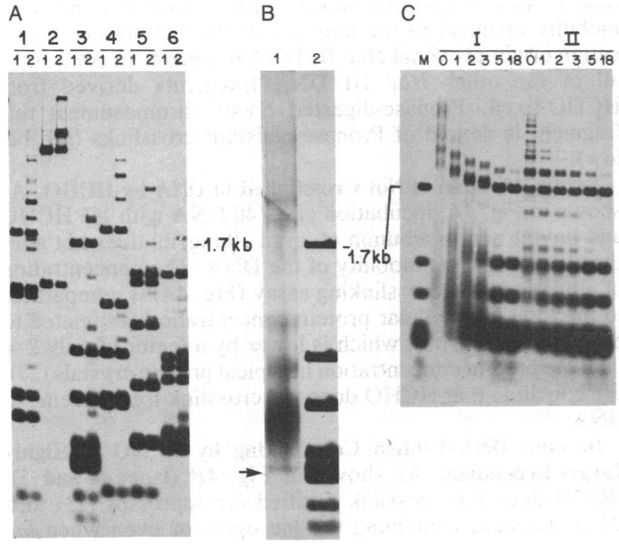

FIG. 3. Restriction site accessibilities of HCHO-fixed, Pronase-digested SV40 chromosomes. (*A*) Control SV40 DNA (lanes 1) and *in vitro* HCHO-fixed, Pronase-digested SV40 chromosomes (lanes 2) were limit-digested with *Hin*dIII (group 1), *Pst* I (group 2), *Hae* III (group 3), *Hin*fI (group 4), *Mbo* I (group 5), or *Mbo* II (group 6). Following endonuclease digestion, Pronase-resistant crosslinks were thermally reversed (see *Results*) and the samples were electrophoresed on a 2% agarose gel containing 0.1% $NaDodSO_4$. kb, Kilobases. (*B*) *In vitro* HCHO-fixed, Pronase-digested SV40 chromosomes (lane 1) and control SV40 DNA (lane 2) were limit-digested with *Hae* III and electrophoresed on a long (33 cm) 2% agarose gel containing 0.1% $NaDodSO_4$. The arrow indicates the replication origin-proximal, 325-bp *Hae* III DNA fragment whose mobility is that expected of free DNA (see text and Fig. 6). (*C*) *In vitro* HCHO-fixed, Pronase-digested SV40 chromosomes were incubated in 1 mM Na_2EDTA/10 mM Tris·HCl, pH 8.0, at 65°C for 0, 1, 2, 3, 5, or 18 hr, as indicated, followed by limit-digestion with *Hae* III. The samples were electrophoresed on a 2% agarose/$NaDodSO_4$ gel either immediately (group I) or after an additional 6-hr incubation at 65°C to remove Pronase-resistant crosslinks in all samples (group II). Lane M, *Hae* III fragments of free SV40 DNA.

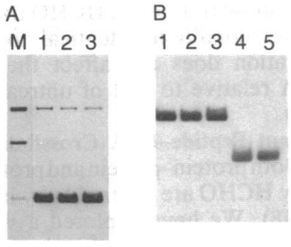

FIG. 4. Target selectivity of HCHO in protein–DNA crosslinking. (*A*) Samples in 0.15 M NaCl/1 mM Na_2EDTA/10 mM triethanolamine HCl, pH 7.5, containing SV40 DNA alone (lane 1) or in the presence of 10 mg of bovine serum albumin per ml (lane 2) or 50 mg of bovine serum albumin per ml (lane 3) were treated with 1% HCHO under conditions identical to those used to fix SV40 chromosomes followed by electrophoresis on a 1.3% agarose gel containing 0.1% $NaDodSO_4$. Fluorographic detection of DNA and markers (lane M) was as in Fig. 1. (*B*) Lanes 1–3, samples containing the ^{32}P-labeled, 172-bp α-satellite DNA fragment in 4% glycerol/0.1% Triton X-100/70 mM NaCl/5 mM 2-mercaptoethanol/20 mM triethanolamine HCl, pH 7.5, either alone (lane 1) or in the presence of 20 times (lane 2) or 80 times (lane 3) the amount of the α-satellite DNA-binding protein [α-protein (25)] necessary to bind one-half of the DNA fragments were treated with HCHO as in *A*. Lanes 4 and 5, samples containing the ^{32}P-labeled, 92-bp fragment bearing *lac* operator in 3% glycerol/1% glucose/0.1% Triton X-100/1 mM Na_2EDTA/10 mM 2-mercaptoethanol/60 mM potassium phosphate, pH 7.6, either alone (lane 4) or in the presence of 1000 times (lane 5) the amount of *lac* repressor necessary to bind one-half of the *lac* operator-bearing DNA fragments were treated with HCHO as in *A*. The samples then were electrophoresed on a 0.1% $NaDodSO_4$/7% polyacrylamide gel overlaid with a 2-cm spacer of 1% agarose to prevent crosslinking of DNA fragments to the polyacrylamide matrix, which has been observed in the presence of HCHO (data not shown). [^{32}P]DNA was detected by autoradiography of dried gels. Both α-protein/α-satellite DNA and *lac* repressor/*lac* operator binding assays were carried out in parallel on low-ionic-strength polyacrylamide gels (25–27) to monitor DNA–protein binding (data not shown).

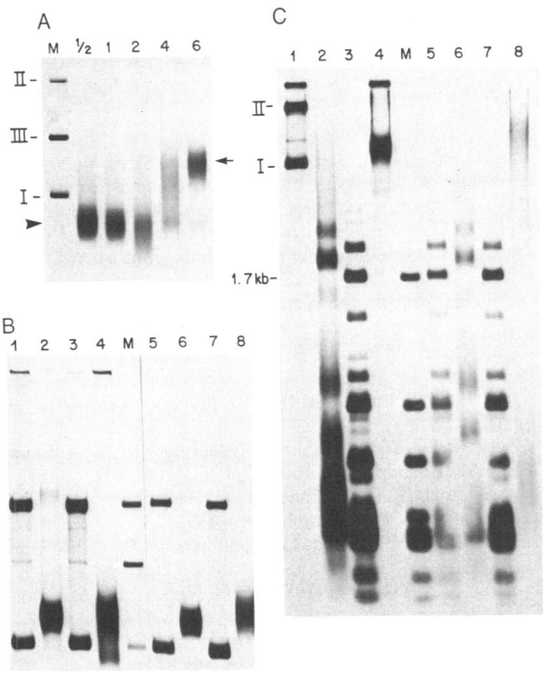

FIG. 5. Isolation of SV40 DNA from SV40-infected, HCHO-fixed CV-1 cells. (*A*) SV40-infected cells were fixed with HCHO, washed, and digested with Pronase in the presence of $NaDodSO_4$ for the indicated number of hours. At the end of digestion the samples were centrifuged at 12,000 × *g* for 10 min and the supernatants were retained. Equal [^{3}H]DNA cpm were electrophoresed on a 1.3% agarose gel containing 0.1% $NaDodSO_4$. Pronase-released [^{3}H]DNA cpm, expressed as a percent of the 6-hr value, were 36% at 0.5 hr, 51% at 1 hr, 60% at 2 hr, and 102% at 4 hr. (*B*) Samples of either *in vitro*- or *in vivo*-fixed, Pronase-digested SV40 chromosomes were electrophoresed as in Fig. 1 either before (lanes 1–4) or after (lanes 5–8) purification of DNA by sedimentation in a sucrose gradient containing 0.1% $NaDodSO_4$ (see below). Lanes 2, 3, 6, and 7 correspond to *in vitro*-fixed samples electrophoresed either before (lanes 2 and 6) or after (lanes 3 and 7) thermal reversal of crosslinks. Lanes 1, 4, 5, and 8 correspond to *in vivo*-fixed samples electrophoresed either before (lanes 4 and 8) or after (lanes 1 and 5) thermal reversal of crosslinks. (*C*) Same treatments as in *B* except that, in addition, samples were limit-digested with *Hae* III either immediately after limit digestion with Pronase (lanes 1–4) or after initial purification of DNA by sedimentation in a $NaDodSO_4$-containing sucrose gradient (lanes 5–8). Processing of HCHO-fixed cells (for *B* and *C*): fixed monolayers were rinsed with 0.1% $NaDodSO_4$/1 mM Na_2EDTA/0.5 mM Na_2EGTA/10 mM sodium Hepes, pH 7.5, incubated with an excess of the same buffer for 1 hr, and then limit-digested with Pronase (0.1 mg/ml) in 1 ml of the same buffer per plate at 37°C for 6 hr with readdition of Pronase at 3 hr. The samples were centrifuged at 12,000 × *g* for 10 min. The supernatants (containing >99% of the total SV40 DNA) were centrifuged at 20,000 rpm for 16 hr at 20°C in an SW41 rotor (Beckman) on a 5–30% linear sucrose gradient containing 0.1% $NaDodSO_4$, 1 mM Na_2EDTA, 0.5 mM Na_2EGTA, 10 mM sodium Hepes (pH 7.5), and freshly added 0.1 mM phenylmethylsulfonyl fluoride.

material migrates as a distinct band *ahead* of free form I SV40 DNA (Fig. 5*A*, arrowhead). The electrophoretic mobility of this rapidly migrating species is close to the mobility of *in vitro* HCHO-fixed SV40 virions (data not shown), suggesting the identity of the corresponding particles. In contrast, *limit*-digestion with Pronase solubilizes >99% of the total SV40 DNA from infected, HCHO-fixed CV-1 cells and converts all of it into a distinct species with the same electrophoretic mobility as *in vitro* HCHO-fixed, Pronase-digested SV40 chromosomes (Fig. 5*A*, lane 6, arrow). Digestions of SV40-infected, HCHO-fixed cells with either trypsin, V8 proteinase, or proteinase K or with Pronase in the presence of either Triton X-100 or Sarkosyl NL97 (instead of $NaDodSO_4$) release virtually none of the SV40 DNA (data not shown).

Restriction endonuclease digests of the *in vivo*-fixed, Pronase-digested SV40 chromosomes yield little of the characteristic DNA bands seen with the *in vitro*-fixed, Pronase-digested chromosomes (Fig. 5*C*, compare lane 6 with lane 8). However, thermal reversal after *Hae* III digestion of the *in vivo* HCHO-fixed, Pronase-digested chromosomes (Fig. 5*C*, lane 5) results in DNA restriction patterns virtually indistinguishable from those of *in vitro*-fixed chromosomes (Fig. 5*C*, lane 7). One interpretation of these results is that, in contrast to *in vitro*-fixed chromosomes, the *in vivo*-fixed, Pronase-digested SV40 chromosomes contain crosslinks not only *within* individual restriction DNA fragments but also *between* such fragments. If so, it should be possible to use HCHO as a probe not only for local DNA–protein contacts but also for higher-order aspects of *in vivo* chromosome structure.

Replication Origin-Proximal Region in a Subset of SV40 Chromosomes Lacks HCHO-Crosslinkable Proteins *in Vivo*. As mentioned above, all but one restriction fragment derived from *in vitro* HCHO-fixed, Pronase-digested, *Hae* III-digested SV40 chromosomes migrate in electrophoresis slower than their free DNA counterparts (Fig. 3*B*). To address this result in greater detail, *in vivo* HCHO-fixed, Pronase-digested SV40 chromosomes were limit-digested with *Hae* III and electrophoresed in a first-dimension agarose/$NaDodSO_4$ gel either alone (Fig. 6*B*) or with naked SV40 DNA markers produced by *Hae* III (Fig. 6*A*). DNA–peptide crosslinks were thermally reversed before carrying out second-dimension electrophoresis in a gel of the same composition. The rightward major diagonal (R) in Fig. 6*A* represents DNA

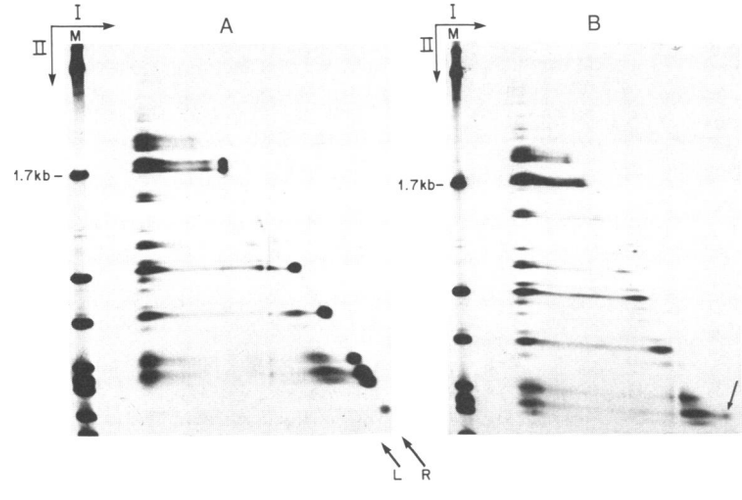

FIG. 6. Absence of Pronase-resistant DNA–protein crosslinks within the replication origin-proximal region in a subset of *in vivo*-fixed SV40 chromosomes. SV40 [^{3}H]DNA was purified from SV40-infected, HCHO-fixed, Pronase-digested CV-1 cells as described in the legend to Fig. 5. This material was limit-digested with *Hae* III either in the presence (*A*) or in the absence (*B*) of added free SV40 [^{3}H]DNA and thereafter was electrophoresed on a 2% agarose gel containing 0.1% $NaDodSO_4$. Following thermal reversal of Pronase-resistant DNA–peptide crosslinks within the first-dimension gel strip, second-dimension electrophoresis was carried out in a gel of the same composition. The rightward major diagonal (R) in *A* represents DNA that migrated with the mobility of naked DNA in *both* dimensions. The leftward major diagonal (L) represents DNA that migrated *slower* in the first dimension then in the second one due to Pronase-resistant, thermally reversible crosslinks. The arrow in *B* indicates a DNA spot corresponding to the replication origin-proximal region in SV40 chromosomes within which no Pronase-resistant crosslinks were formed during *in vivo* HCHO fixation.

fragments that lacked electrophoretically detectable crosslinks during first-dimension electrophoresis, whereas the leftward diagonal (L) represents peptide-bearing DNA fragments that were retarded in the first dimension.

As can be seen from Fig. 6*B*, the major spot on the rightward ("naked DNA") diagonal (R) corresponds to the 325-bp *Hae* III DNA fragment that is devoid of recognizable nucleosomes in a subset of SV40 chromosomes *in vitro* (7, 22, 30–32). From the intensity of this spot relative to the total intensity of DNA spots of the same size in Fig. 6*B* we estimate that the nucleosome-free region occurs *in vivo* in 10–15% of the SV40 chromosomes. This figure is comparable to the 20–25% estimate obtained by using *in vitro* approaches (31, 32). Thus, the HCHO technique reveals a clear distinction between nucleosome-free and nucleosome-containing regions of *in vivo* HCHO-fixed SV40 chromosomes, suggesting that this approach could also detect and map nucleosome-free regions within cellular chromosomes *in vivo*.

DISCUSSION

The HCHO Technique and Detection of the SV40 Nucleosome-Free Region *in Vivo*. Previous *in vitro* data, obtained with SV40 chromosomes by using both biochemical (7, 22, 30) and electron microscopic approaches (31, 32), have shown that the ≈400-bp SV40 control region is free of nucleosomes in a subpopulation of isolated SV40 chromosomes. Analogous "exposed" or nuclease-hypersensitive sites have since been detected in particular at the 5′ ends of active genes (ref. 33; reviewed in ref. 34). Using the HCHO technique, we find that the SV40 control region is also exposed *in vivo* in that it appears to be devoid of HCHO-crosslinkable DNA–protein contacts in a subpopulation of SV40 chromosomes. The striking target selectivity of HCHO that efficiently produces histone–DNA crosslinks both *in vitro* and *in vivo* but fails to crosslink either *lac* repressor or α-protein to their cognate DNA sequences *in vitro* (see *Results*) strongly suggests that nucleosomes (and possibly histones themselves) are absent from the exposed SV40 control region in a subpopulation of SV40 chromosomes *in vivo*. At the same time, these data do not exclude the possibility of specific nonhistone proteins interacting with the nucleosome-free region *in vivo*, since such proteins may not be crosslinkable to DNA by HCHO.

The fact that free proteins, such as serum albumin, do not crosslink to DNA by HCHO *in vitro* even at extremely high free protein concentrations and the absence of detectable crosslinks within the SV40 control region in the *in vivo* HCHO-fixed SV40 chromosomes (see *Results*) strongly suggest that most of the HCHO-mediated DNA–protein crosslinks that occur *in vivo* are due to physiological DNA–protein interactions. The use of HCHO as a probe for mapping *in vivo* DNA–protein interactions is complementary to *in vivo* uses of other probes such as dimethyl sulfate (10, 11), since the latter allows detection of DNA-bound nonhistone proteins but not of nucleosomes (5, 10, 11).

Future Development and Applications of the HCHO Technique. We envision three versions of the HCHO technique. The first one, described in the present work, is designed to map *in vivo* protein–DNA contacts (largely, if not exclusively, histone–DNA contacts) at the resolution provided by restriction endonuclease mapping and the electrophoretic "retardation" effect of Pronase-resistant peptide–DNA adducts. The second version, to be developed, aims to map the sites of peptide–DNA adducts on the DNA to the nucleotide level, thus permitting *in vivo* footprint analysis of DNA–protein interactions. In this approach, the electrophoretic retardation effect is no longer necessary to detect the presence of a crosslink. Ultimately, one should be able to determine *both* the position of a peptide–DNA crosslink on the DNA and the identity of the corresponding peptide component.

We are greatly indebted to Kathleen Matthews, François Strauss, and James Wang for their gifts of *lac* repressor, α-protein, and plasmid pJW270, respectively. We also thank Daniel Finley for helpful comments on the manuscript and Barbara Doran for secretarial assistance. This work was supported by a grant to A.V. from the National Cancer Institute (CA30367). M.S. was supported by a predoctoral fellowship from the National Science Foundation.

1. Siebenlist, U., Simpson, R. B. & Gilbert, W. (1980) *Cell* **20,** 269–281.
2. Galas, D. J. & Schmitz, A. (1978) *Nucleic Acids Res.* **5,** 3157–3170.
3. Cartwright, I. L., Hertzberg, R. P., Dervan, P. B. & Elgin, S. C. R. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3213–3217.
4. Kohwi-Shigematsu, T., Gelinas, R. & Weintraub, H. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 4389–4393.
5. McGhee, J. D. & Felsenfeld, G. (1980) *Annu. Rev. Biochem.* **49,** 1115–1156.
6. Gilmour, D. S. & Lis, J. T. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 4275–4279.
7. Varshavsky, A., Sundin, O. & Bohn, M. (1979) *Cell* **16,** 453–466.
8. Karpov, V. L., Preobrazhenskaya, O. V. & Mirzabekov, A. D. (1984) *Cell* **36,** 423–431.
9. Welsh, J. & Cantor, C. R. (1984) *Trends Biochem. Sci.* **9,** 505–508.
10. Church, G. M. & Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1991–1995.
11. Ephrussi, A., Church, G. M., Tonegawa, S. & Gilbert, W. (1985) *Science* **227,** 134–140.
12. Becker, M. M. & Wang, J. C. (1984) *Nature (London)* **309,** 682–687.
13. Ilyin, Y. V. & Georgiev, G. P. (1969) *J. Mol. Biol.* **41,** 299–303.
14. Brutlag, D., Schlehuber, C. & Bonner, J. (1969) *Biochemistry* **8,** 3214–3218.
15. Varshavsky, A. & Ilyin, Y. V. (1974) *Biochim. Biophys. Acta* **340,** 207–217.
16. Jackson, V. & Chalkley, R. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 6081–6085.
17. Moller, K., Rinke, J., Ross, A., Buddle, G. & Brimacombe, R. (1977) *Eur. J. Biochem.* **76,** 175–187.
18. Jackson, V. (1978) *Cell* **15,** 945–954.
19. McGhee, J. D. & von Hippel, P. H (1975) *Biochemistry* **14,** 1281–1296.
20. Trifonov, E. N., Lazurkin, Y. S. & Frank-Kamenetsky, M. D. (1967) *Mol. Biol. (Engl. Transl.)* **1,** 142–151.
21. Doenecke, D. (1978) *Hoppe-Seyler's Z. Physiol. Chem.* **359,** 1343–1352.
22. Varshavsky, A., Sundin, O. H. & Bohn, M. J. (1978) *Nucleic Acids Res.* **5,** 3469–3478.
23. Matthews, K. S. (1979) *J. Biol. Chem.* **254,** 3348–3353.
24. Barkley, M. D. & Bourgeois, S. (1978) in *The Operon*, eds. Miller, J. H. & Reznickoff, W. S. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 177–220.
25. Strauss, F. & Varshavsky, A. (1984) *Cell* **37,** 889–901.
26. Fried, M. & Crothers, D. M. (1981) *Nucleic Acids Res.* **9,** 6505–6525.
27. Garner, M. M. & Revzin, A. (1981) *Nucleic Acids Res.* **9,** 3047–3060.
28. Fulton, A. B. (1982) *Cell* **30,** 345–347.
29. Pabo, C. O. & Sauer, R. T. (1984) *Annu. Rev. Biochem.* **53,** 293–321.
30. Scott, W. A. & Wigmore, D. J. (1978) *Cell* **15,** 1511–1518.
31. Jacobovitz, E. G., Bratosin, S. & Aloni, Y. (1980) *Nature (London)* **285,** 263–265.
32. Saragosti, S., Moyne, G. & Yaniv, M. (1980) *Cell* **20,** 65–73.
33. Wu, C. (1980) *Nature (London)* **286,** 854–860.
34. Elgin, S. C. R. (1984) *Nature (London)* **309,** 213–214.