Mitotic stability of an episomal vector containing a human scaffold/matrix-attached region is provided by association with nuclear matrix

A. Baiker*, C. Maercker*, C. Piechaczek*, S. B. A. Schmidt*, J. Bode†, C. Benham‡ and H. J. Lipps*§

*Institut für Zellbiologie, Universität Witten/Herdecke, Stockumer Strasse 10, D-58448 Witten, Germany
†Gesellschaft für Biotechnologische Forschung mbH, Mascheroder Weg 1, D-38124 Braunschweig, Germany
‡Department of Biomathematical Sciences, Mount Sinai School of Medicine, New York, New York 10029, USA
§e-mail: lipps@uni-wh.de

DNA replication occurs in tight association with the nuclear matrix, where binding of the replication origin to the nuclear matrix must precede the onset of S phase[1–4]. We have shown previously that the origin of replication of the simian virus 40 (SV40) genome linked to a human scaffold/matrix-attached region (S/MAR) allows sustained episomal replication (where an episome is autonomous, self-replicating DNA) that is independent of the expression of the virally encoded large T-antigen[5]. A vector with this combination of SV40 origin and potential for matrix association is maintained in cultured cells for at least 100 cell generations, in the absence of selection[5]. Here we show, by *in situ* hybridization and nuclear-fractionation procedures, that there is a specific interaction of this vector with the nuclear matrix and the chromosome scaffold, presumably through proteins that both structures have in common. This interaction correlates with replication of the vector as an episome. These observations allow a mechanistic explanation for the episomal replication and mitotic stability of this new type of vector.

Chinese hamster ovary (CHO) cells, transfected either with the S/MAR–origin-containing vector (pEPI-1, Fig. 1a) or the corresponding truncated vector lacking the S/MAR (pGFP-C1, Fig. 1d) were subjected to fluorescence *in situ* hybridization (FISH) on metaphase spreads. Confirming previous results[5], FISH analyses showed that the truncated vector was prone to integrate into the host-cell DNA. A single intense signal was seen at the same chromosomal locus of different cells derived from the same clone (Fig. 1e, f); however, the site of integration varied between different clones. In contrast to this predictable behaviour of integrated transgenes[6], clones established using the complete vector contained cells with a greater and variable number of fluorescent spots that were associ-

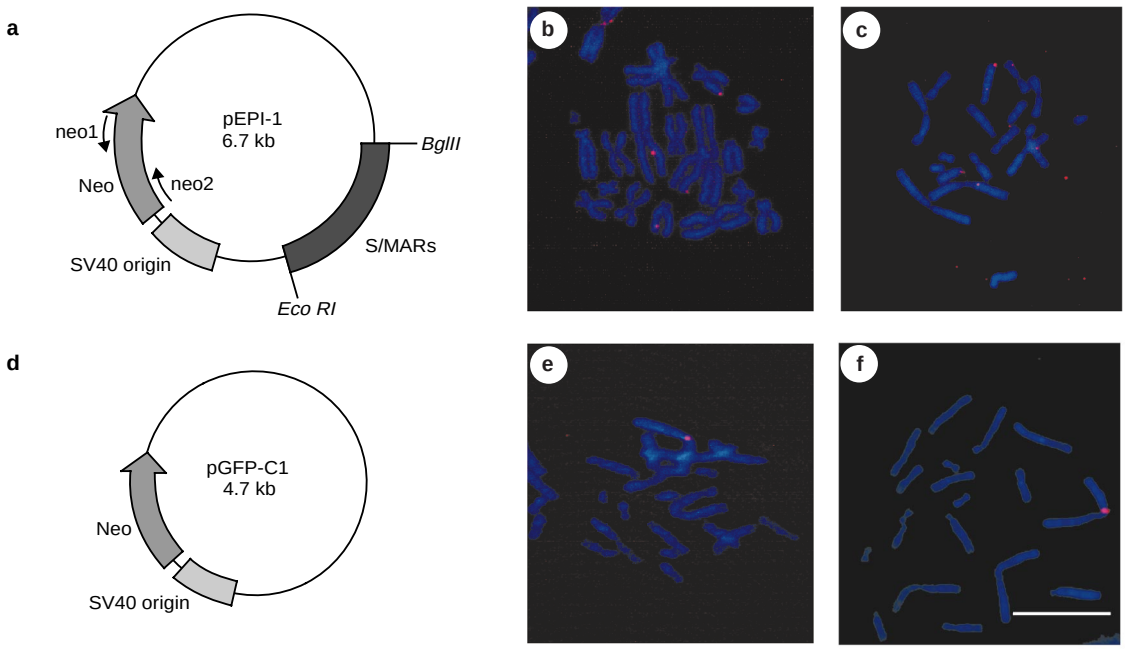

Figure 1 *In situ* hybridization of S/MAR–origin-containing vector (pEPI-1) and its S/MAR-depleted version (pGFP-C1) to CHO metaphase chromosomes. **a**, Restriction map of pEPI-1 (ref. 5). Only the functional elements of the vector are indicated. Arrows indicate the positions of primers used for PCR analysis. **b**, **c**, *In situ* hybridization to a CHO clone transfected with pEPI-1. **b**, Spreading of metaphase chromosomes was done as gently as possible. **c**, During spreading, increased shear forces were applied as described in Methods. **d**, Map of pGFP-C1 (ref. 5). **e**, **f**, *In situ* hybridization to a CHO clone transfected with pGFP-C1. Spreading of metaphase chromosomes was done in **e** and **f** as in **b** and **c** respectively. *In situ* hybridization was done according to ref. 17 using biotin-labelled pGFP-C1 as a probe. Scale bar represents 5 μm. kb, kilobase.

ated, but not strictly coincident, with the chromosomes (Fig. 1b, c). By analysing metaphase spreads of 10 different clones, we estimated the number of vector molecules per cell to be between 4 and 13 (average copy number between 5 and 9; standard deviations between 0.83 and 3.1) (see Supplementary information). These figures did not change significantly between 20 and 50 generations after transfection in the absence of selection. The position of these signals varied from cell to cell and the non-covalent association of the vector and host chromosomes was evident, as hybridization was seen always on only one of the sister chromatids in metaphase spreads. The non-covalent nature of this association was further shown by applying increased shear forces during spreading. Under gentle lysis conditions more than 95% of the signals were associated with the chromosome (Fig. 1b); however, when cells were dropped onto the slides from a distance exceeding 10 cm, a significant number of vector molecules appeared to be separated from the chromosome structures (Fig. 1c).

To confirm the association of the S/MAR-containing vector with the nuclear matrix, we used nuclear-fractionation procedures[7]. We separated the various protein fractions by SDS–polyacrylamide gel electrophoresis (SDS–PAGE) and visualized the DNA-binding constituents by southwestern analysis. In the presence of a 1,000-fold excess of nonspecific DNA, both vectors showed comparable binding to a protein in fraction 2 (0.25 M ammonium sulphate), which contains the soluble nuclear components including histone H1. Subsequent digestion with DNaseI removed DNA and the remaining histones (fraction 3), leaving the complete matrix in the pellet fraction. After extraction with 2 M NaCl, matrix proteins not part of the core filament network will be in the supernatant (fraction 4). The core filament network was then solubilized with SDS-sample buffer to yield fraction 5. In the matrix-containing fractions, a major protein band bound strongly to the S/MAR-containing vector but not to its truncated form (Fig. 2a, b). However, although migration of the S/MAR-binding protein and the protein in fraction 2 indicates that they might belong to the histone H1 group, this correspondence could not be confirmed by using several available antibodies against common H1 subspecies that, in a control, showed crossreaction with endogenous CHO-cell histone H1.

To prevent chromosomal rearrangements from occurring during fractionation, we added a crosslinking step to the protocol. *Cis*-DDP is a reagent that links matrix proteins to endogenous S/MAR with high specifity[8]. *Cis*-DDP can be applied to the living cell and activated by reducing the concentration of chloride ions. In these experiments, we replaced digestion with DNaseI during the fractionation procedure by digestion with six restriction enzymes that do not cut pEPI-1, together with either one restriction enzyme that linearizes the vector (*Eco*RI or *Bgl*II, Fig. 1a) or two enzymes that excise the S/MAR element (*Eco*RI and *Bgl*II, Fig. 1a). After nuclear fractionation, the crosslinked complexes were treated with proteinase K and the purified DNA was amplified by the polymerase chain reaction (PCR) to trace the complete vector and its S/MAR-depleted variant. Vectors without the S/MAR were found in fraction 3 and, to a lesser extent, in fraction 4 (Fig. 2c, PCRII) whereas S/MAR-containing vectors were found in the nuclear-matrix fraction 5 and to a lesser extent in fraction 4 (Fig. 2c, PCRI).

Further control experiments support our view that the combination of an SV40 replication origin and a scaffold/matrix-attached region enables replication in the episomal state and that this functional state correlates with a specific interaction with components of the nuclear matrix. Plasmids in which either the S/MAR (Fig. 1d–f) or the replication origin (data not shown) was deleted lost this property and became integrated. In addition, other control sequences cloned in place of the S/MAR element did not support the episomal state. The AT-rich NTS-1 and NTS-2 sequences are associated with an endogenous origin of replication in mouse cells[9]. We cloned these sequences in place of the S/MAR element either as monomers or as multimers (up to a tetramer). In no case was episomal replication of these vectors observed. While the 69% AT-rich

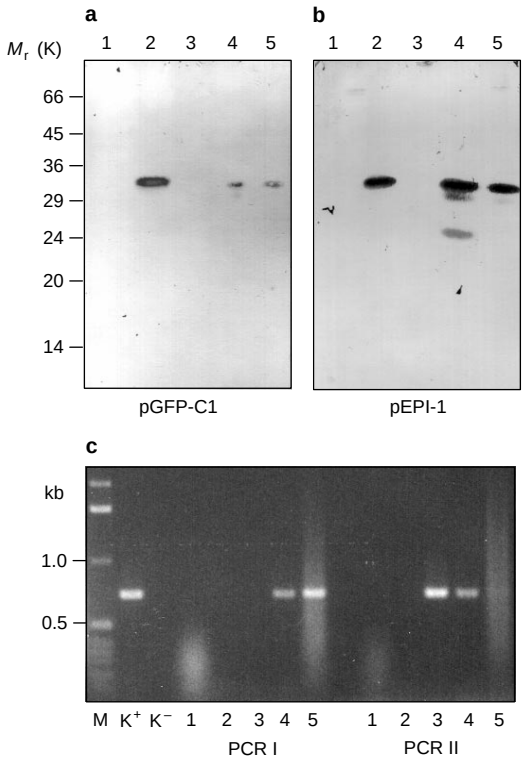

Figure 2 Nuclear fractionation and identification of the S/MAR-containing vector by southwestern and PCR analysis. **a**, **b**, Southwestern analysis of nuclear proteins fractionated according to ref. 7. The proteins were separated by SDS-gel electrophoresis, blotted and probed either with DIG-labelled pGFP-C1 (**a**) or with pEPI-1 (**b**). Fraction 1 was extracted with cytoskeletal buffer including 100 mM NaCl; fraction 2 was extracted with extraction buffer containing 250 mM ammonium sulphate; fraction 3 was extracted with digestion buffer after digestion with DNaseI; fraction 4 is the supernatant obtained after adjusting to 2 M NaCl; fraction 5 is the pellet fraction resulting after extraction with 2 M NaCl. M_r (K), relative molecular mass (in thousands). **c**, Distribution of vector DNA in nuclear fractions 1–5, after crosslinking with *cis*-DDP[8], of cells transfected with pEPI-1. For this experiment, cellular DNA was not digested by DNaseI but by restriction digestion using six non-cutting restriction enzymes together with either a restriction enzyme that linearizes the vector (the S/MAR is retained with the portion of the vector amplified by PCR (PCRI)) or two restriction enzymes (the S/MAR is released from the portion of the vector amplified by PCR (PCRII)). The same volumes of each fraction were used for PCR analysis; primers were derived from the neomycin-resistance genes indicated in Fig. 1a; the DNA concentrations in the corresponding fractions were identical in PCRI and PCRII. K^+ and K^-, control PCRs using either pEPI-1 as template (K^+) or no template (K^-); M, 1-kb ladder (Gibco). PCR conditions were as described[20].

S/MAR element[10] used in this study has a normalized binding affinity of 94 ± 5% to the nuclear matrix[11] and an extended base-unpairing region over its entire length[12], the amplification-promoting sequences[9] NTS-1 (370 base pairs, 56% A+T, binding affinity <20%) and NTS-2 (424 base pairs, 65% A+T, binding affinity 50 ± 10%) contain only one (NTS-1) or two (NTS-2) restricted unpairing elements, which are not sufficient to confer S/MAR character[12].

The association of the vector molecules with the host chromosomes, as shown by *in situ* hybridization, is reminiscent of the behaviour of episomally replicating viruses such as Epstein–Barr virus, Kapos's-sarcoma-associated herpesvirus or bovine papilloma virus, in which maintenance of the viral genome is mediated by virally encoded proteins (EBNA-1, LANA and E2, respectively) that ensure efficient segregation of the replicated genome during mitosis[13–15]. In the case of our vector it may be that association with elements of the

nuclear matrix is also essential for segregation during mitosis; furthermore, such an association must be vital in allowing the plasmid to replicate as an episome. In SV40 the expression of the large T-antigen is essential for SV40 replication: it unwinds the SV40 origin and allows the replication complex to bind[16]. In the S/MAR-containing vector the S/MAR might complement some of the functions of the large T-antigen and, as it is bound to the nuclear matrix, the SV40 origin could be co-replicated during cellular DNA replication using the replication machinery that forms on the chromosome. To our knowledge, this is the first report that the persistence of an episomal vector that is capable of sustained, selection-independent replication in the absence of virally encoded *trans*-acting factors correlates with its association with the nuclear matrix. □

Methods

Culturing of CHO cells, transfection of CHO cells and selection of transfected clones were done as described[5].

FISH analysis was done according to ref. 17 using biotin-labelled pGFP-C1 (Fig. 1d) as a probe. Spreading of metaphase chromosomes was done as gently as possible (Fig. 1b), or increased shear forces were applied by dropping the cells onto the slides from at least a 10-cm distance (Fig. 1c). The labelled probe was detected by an avidin/tetramethyl rhodamine isothiocyanate (TRITC)-labelled anti-avidin antibody/biotin (Sigma) sandwich procedure[17]. After immunostaining, metaphase chromosomes were counterstained with 4,6-diamidinophenylindole (DAPI; 0.1 μg ml^{-1}; Sigma). Fluorescence microscopy was done using a Leitz DM RB microscope and photographed using Kodak Elite Chrome 100 ASA films.

Nuclear proteins were fractionated according to ref. 7, separated on a 12.5% SDS gel[18] and blotted onto a nylon membrane (Amersham); a southwestern analysis[19] was performed using either digoxygenin (DIG)-labelled pGFP-C1 or pEPI-1 as a probe. Before the incubation with the labelled probes, filters were saturated with a 500–1,000-fold excess of sonicated *Escherichia coli* DNA.

To trace the vector DNA in the different nuclear fractions by PCR analysis, we included a *cis*-DDP crosslinking step[8]. The digestion with DNaseI during the fractionation process was replaced in these experiments by digestion with restriction enzymes, using six restriction enzymes that do not cut pEPI-1 and, in addition, either a restriction enzyme (*Eco*RI or *Bgl*II) that linearizes the vector (Fig. 2c, PCRI) or two restriction enzymes (*Eco*RI and *Bgl*II) that delete the S/MAR from the vector (Fig. 2c, PCRII)[5]. PCR analysis was done according to ref. 20. The same volumes of each fraction were used for PCR analysis and the DNA concentrations in the corresponding fractions were identical in PCRI and PCRII. Amplification was done using primers derived from the neomycin-resistance gene (neo 1: 5′-GGAGAGGCTATTCGGCTATGAC; neo 2: 5′-CGTCAAGAAGGCGATAGAAGGC). Primer positions are indicated in Fig. 1a.

RECEIVED 4 NOVEMBER 1999; REVISED 10 JANUARY 2000; ACCEPTED 27 JANUARY 2000; PUBLISHED 10 FEBRUARY 2000.

1. Dijkwell, P. A. & Hamlin, J. L. *Int. Rev. Cyt.* **162A,** 455–484 (1995).
2. Cook, P. *Science* **284,** 1790–1795 (1999).
3. Cossons, N. *et al. J. Cell. Biochem.* **67,** 439–450 (1997).
4. Bode, J., Benham, C., Knopp, A. & Mielke, C. *Crit. Rev. Eukar. Gene Expr.* (in the press).
5. Piechaczek, C., Fetzer, C., Baiker, A., Bode, J. & Lipps, H. J. *Nucleic Acids Res.* **27,** 426–428 (1999).
6. Robins, D. M., Ripley, S., Henderson, A. S. & Axel, R. *Cell* **23,** 29–39 (1981).
7. Nickerson, J. A., Krockmalnic, G. & Penman, S. in *Cell Biology, a Laboratory Handbook* (ed. Celis, J. E.) 184–192 (Academic, London, 1998).
8. Ferraro, A., Cervoni, L., Eufemi, M., Altieri, F. & Turano, C. *J. Cell. Biochem.* **62,** 495–595 (1996).
9. Wegner, M. *et al. Nucleic Acids Res.* **17,** 9909–9931 (1989).
10. Mielke, C., Kohwi, Y., Kohwi-Shigematsu, T. & Bode, J. *Biochemistry* **29,** 7475–7485 (1990).
11. Kay, V. & Bode, J. *Methods Mol. Cell Biol.* **5,** 186–194 (1995).
12. Benham, C., Kohwi-Shigematsu, T. & Bode, J. *J. Mol. Biol.* **274,** 181–196 (1997).
13. Calos, M. P. *Proc. Natl Acad. Sci. USA* **95,** 4084–4085 (1998).
14. Lehman, C. W. & Botchan, M. R. *Proc. Natl Acad. Sci. USA* **95,** 4338–4343 (1998).
15. Ballestas, M. E., Chatis, P. A. & Kaye, K. M. *Science* **284,** 641–644 (1999).
16. Fanning, E. & Knippers, R. *Annu. Rev. Biochem.* **61,** 55–85 (1992).
17. Liu, C. G., Maercker, C., Castanon, M. J., Hauptmann, R. & Wiche, G. *Proc. Natl Acad. Sci. USA* **93,** 4278–4283 (1996).
18. Laemmli, U. K. *Nature* **227,** 680–685 (1970).
19. Miskimis, W. K., Roberts, M. P., McClelland, A. & Ruddle, F. H. *Proc. Natl Acad. Sci. USA* **82,** 6791–6794 (1985).
20. Saiki, R. K. *et al. Science* **239,** 487–491 (1988).

ACKNOWLEDGEMENTS

This work was supported by the Alfried Krupp von Bohlen und Halbach foundation (H.J.L.), by the EC (grant BIO4-CT98-0203) and DFG (grant Bo 419/5-3) (J.B.) and by the NSF (grant DBI99-04549) and NIH (grant ROI-GM47012) (C.B.). We thank F. Grummt for providing the NTS-1 and NTS-2 sequences and C. Fetzer for his help with the graphs.

Correspondence and requests for material should be addressed to H.J.L.

Supplementary information is available on *Nature Cell Biology*'s World-Wide Web site (http://cellbio.nature.com) or as paper copy from the London editorial office of *Nature Cell Biology*.