

Construction and analysis of cells lacking the HMGA gene family

Brett Beitzel and Frederic Bushman*

Infectious Disease Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

Received May 21, 2003; Revised and Accepted June 26, 2003

DDBJ/EMBL/GenBank accession no. AY303673

ABSTRACT

The high mobility group A (HMGA) family of non-histone chromosomal proteins is encoded by two related genes, HMGA1 and HMGA2. HMGA proteins are architectural transcription factors that have been found to regulate the transcription of a large number of genes. They are also some of the most commonly dysregulated genes in human neoplasias, highlighting a role in growth control. HMGA1 and HMGA2 have also been found to stimulate retroviral integration *in vitro*. In this study, we have cloned chicken HMGA1, and used the chicken DT40 B-cell lymphoma line to generate cells lacking HMGA1, HMGA2 and both in combination. We tested these lines for effects on cellular growth, gene control and retroviral integration. Surprisingly, we found that the HMGA gene family is dispensable for growth in DT40 cells, and that there is no apparent defect in retroviral integration in the absence of HMGA1 or HMGA2. We also analyzed the activity of approximately 4000 chicken genes, but found no significant changes. We conclude that HMGA proteins are not strictly required for growth control or retroviral integration in DT40 cells and may well be redundant with other factors.

INTRODUCTION

High mobility group (HMG) proteins are a group of non-histone chromosomal proteins that are defined by their ability to be extracted from chromatin using 0.35 M NaCl and solubility in 5% perchloric acid (1). HMG proteins are divided into three families: HMGA (formerly known as HMG1/Y), HMGB (formerly known as HMG-1 and -2) and HMGN (formerly known as HMG-14 and -17). HMG proteins have been called 'architectural transcription factors' because they have no transcription activation or repression activities themselves, but they help assemble large, multiprotein complexes on DNA that can then activate or repress transcription (2).

The HMGA family consists of four proteins: HMGA1a, HMGA1b, HMGA1c and HMGA2. HMGA1a, 1b and 1c are all encoded by the same gene but vary in length due to

alternative splicing (3–5), and HMGA2 is encoded by a separate gene (6). The common structural motifs in this group include three DNA binding domains and an acidic C-terminus (7). The DNA binding domains are called A-T hooks because they bind short (4–6 bp) AT-rich sequences in the minor groove (8–10). Two or more AT hooks can bind separate AT-rich regions, causing bridging of two separate DNA molecules (8,9). If the AT-rich regions are located on the same stretch of DNA, binding by HMGA can induce bending, straightening, looping or unwinding depending on the context of the binding sites (10).

HMGA proteins have been shown to regulate the transcription of a large number of genes (11). One of the most well studied examples is the interaction of HMGA1 with the interferon (IFN)- β enhancer. At the IFN- β enhancer, HMGA1 binds and reverses an intrinsic DNA bend. Reversal of this bend allows other transcription factors to bind, and protein-protein interactions between HMGA1 and some of these transcription factors stabilize enhanceosome formation. The basal transcription machinery is then recruited to the enhanceosome and transcription is initiated (12–16).

HMGA proteins may also influence gene expression at the level of chromosomal organization by displacing the transcriptional repressor histone H1 at scaffold attachment regions (SARs), which are A-T-rich regions at the base of large loops of chromatin (17). Histone H1 is thought to be a general transcription repressor that acts by folding chromatin into higher order structures that are inaccessible to the transcription machinery (18–20). Removal of histone H1 mediated by HMGA1 de-represses transcription near SARs *in vitro*, and HMGA1 is enriched in histone H1 depleted, transcriptionally active chromatin fractions, indicating that it may play the same role *in vivo* (17).

In accordance with their many roles in transcription regulation, aberrant expression of HMGA proteins has been implicated in a large number of human cancers (reviewed in 21). Over-expression of full-length proteins has been detected in almost every tumor type analyzed, and there is a strong correlation with HMGA over-expression and tumor progression. Those tumors that have the highest HMGA expression level have been found to have increased malignancy and metastatic potential. Translocations that attach the DNA binding regions of HMGA proteins (most commonly HMGA2) to heterologous partners are frequently found in common benign mesenchymal tumors, and may be the most commonly rearranged genes in human neoplasias (21).

*To whom correspondence should be addressed. Tel: +1 858 453 4100; Fax: +1 858 554 0341; Email: bushman@salk.edu

HMGA proteins have also been shown to activate retroviral preintegration complexes (PICs) *in vitro*. PICs can be extracted from recently infected cells and if an exogenous target DNA is added to the extract, PICs can mediate the integration of the viral cDNA into the target DNA. If the PICs are first subjected to a high salt wash, however, the integration activity is lost. Extracts from uninfected cells could restore integration activity to high salt inactivated PICs and fractionation of these extracts identified HMGA1 as the active component in the extracts. Integration activity can also be restored by adding purified recombinant HMGA proteins to high salt inactivated PICs. HMGA1 can be detected in PIC extracts by western blotting. When these observations are taken together, it is reasonable to think that HMGA1 is a functional part of retroviral PICs (9,22).

Mouse embryonic stem cell lines in which HMGA1 was knocked out have recently been described and these lines were found to have increased expression of BRCA-1 (23). The phenotype of HMGA1 knockout mice has not been reported. Inactivation of HMGA2 in mice leads to a pygmy phenotype (24) and a reduction in adipose tissue (25).

In this study, we used the chicken DT40 B-cell lymphoma line to generate knockout cell lines lacking HMGA1, HMGA2 or both in combination. We have used these lines to investigate the role of the HMGA family in cell growth, the transcription of approximately 4000 chicken genes and retroviral integration.

MATERIALS AND METHODS

Cloning and mapping of chicken HMGA genes

To isolate the chicken ortholog of HMGA1, a stage 3 chicken embryo library (obtained from Martyn Goulding, The Salk Institute, La Jolla, CA) was screened using the human HMGA1 coding sequence as a probe. Low stringency hybridization identified several clones that hybridized to the human probe. Sequencing the insert of one of these, pBBcHMGA1b, identified a coding region that was highly similar to human HMGA1b (GenBank accession no. AY303673). The chicken HMGA1b coding sequence identified in this clone was used to screen a chicken genomic library (Stratagene, La Jolla, CA) and several clones were identified that contained the chicken HMGA1 locus. These clones were subcloned and subjected to restriction mapping to generate a genomic map of the chicken HMGA1 locus.

Chicken HMGA2 had been cloned previously by Martin Groenen (26), who kindly sent us two cDNA clones, HCC3 and HCC4, and one genomic clone, H34, containing the 3' end of the HMGA2 gene. We PCR amplified the HMGA2 coding region from HCC3 and used the PCR product to screen a chicken genomic library (Stratagene) to obtain clones containing the 5' end of the HMGA2 gene. Several of these clones were subjected to restriction digest mapping to generate a map of the HMGA2 locus.

Generation of knockout constructs

pBluescript SK+ was digested with BamHI, filled in with T4 polymerase and religated to generate pBluescriptnoBam. Oligos BB37 (5'-ATAACTTCGTATAATGTATGCTATACGAAGTTATGGATCCATAACTTCGTATAATGTATGCT-

ATACGAAGTTATTGCA-3') and BB38 (5'-ATAACTTCGTATAGCATACATTATACGAAGTTATGGATCCATAACTTCGTATAGCATACATTATACGAAGTTATTGCA-3') were annealed and ligated into PstI-cut pBluescriptnoBam to generate pBluescript-lox. Cassettes containing histidinol dehydrogenase (*hisD*), puromycin-*N*-acetyl-transferase (*puro*), blasticidin *S*-acetyltransferase (*bsr*) and xanthine-guanosine phosphoribosyltransferase (*gpt*), all under the control of the chicken beta-actin promoter, were kindly provided by J. M. Buerstedde (Heinrich-Pette-Institut for Experimental Virology and Immunology, Hamburg, Germany) and cloned into the BamHI site of pBluescript-lox to generate pBB102, pBB103, pBB104 and pBB105, respectively.

A 7.5 kb ApaI-SspI fragment downstream of the HMGA1 coding region was cloned into pBB104 and pBB105 to generate pBB104+ds and pBB105+ds. An 8 kb NotI fragment (one NotI site provided by the vector) upstream of the HMGA1 coding region was cloned into pBB104+ds and pBB105+ds to generate pBB104KO and pBB105KO.

Oligos BB296 (5'-AATTGGCGCGCCGGCCGGCCGTAC-3') and BB297 (5'-GGCCGGCCGGCGCGCC-3') were annealed and ligated into EcoRI/KpnI-cut pBB102 and pBB103 to add AscI and FseI sites, generating pBB102mod and pBB103mod. An ~8 kb region downstream of HMGA2 exon III was amplified using primers BB300 (5'-GGCCGCCCTTCCAGGAGAGGATCAGC-3') and BB301 (5'-GGCGCGCCGGTACCGAAAAGTTAACCACAGCTCCC-3'), generating a fragment flanked by AscI-KpnI proximal to exon III and FseI distal to exon III. This fragment was cloned into AscI/FseI-digested pBB102mod and pBB103mod to generate pBB102+ds and pBB103+ds. Oligos BB294 (5'-GCGGCCGCGCGCTCCGCATGTTGTG-3') and BB295 (5'-ACTAGTGGCGAGCCGAGAGCTCAG-3') were used to amplify a 1 kb region upstream of HMGA2 exon I. This fragment was digested with NotI and SpeI and cloned into NotI-SpeI-digested pBB102+ds and pBB103+ds to generate pBB102KO and pBB103KO.

Culturing DT40 cells

DT40 cells were maintained in RPMI (Invitrogen, Carlsbad, CA) plus 10% FBS plus 1% chicken serum plus 50 μ M beta-mercaptoethanol. During selection for drug-resistant clones, cells were grown at 42°C. DT40 cells have a doubling time of 8–10 h at 42°C and growth at this temperature speeded the isolation of drug-resistant clones. Cells were generally maintained at 37°C at all other times.

Transfection and selection

A total of 10⁷ DT40 cells were washed once with PBS and resuspended in 500 μ l of PBS in a 0.4 cm electroporation cuvette. Fifty micrograms of linearized knockout construct was added to the cells and the mix was incubated on ice for 10 min. Cells were electroporated at room temperature in a BioRad Gene Pulser at 300 V and 975 μ F. After electroporation, cells were incubated on ice for 10 min, resuspended in 20 ml of normal growth medium and incubated at 42°C for 18 h. After recovery, the volume of medium was increased to 80 ml, the appropriate drug selection was added and the mix was plated in four 96-well plates. Resistant clones generally appeared after 5–7 days.

Screening for homologous recombinants

Genomic DNA from drug-resistant clones was analyzed by Southern blotting to determine which clones had undergone homologous recombination to knockout one allele. Genomic DNA from putative HMGA1 knockout clones was digested with SphI and probed with a 300 bp SspI–SphI DNA fragment located directly downstream of the genomic region included in the knockout construct. Wild-type alleles will generate an ~10 kb band and knockout alleles will generate an ~12 kb band in this Southern blot analysis. Genomic DNA from putative HMGA2 clones was digested with KpnI and probed with a DNA fragment generated by PCR amplification of a region immediately downstream of the genomic DNA region included in the knockout construct. Oligos used for this amplification were: BB298, 5'-ACCTAACACAGAATAA-ATTGAAG-3'; BB299, 5'-CTAGGGCAGGCTAATATG-AGG-3'. Wild-type alleles will generate an ~10.5 kb band and knockout alleles will generate an ~8 kb band in this Southern blot analysis. A subset of clones that had lost one endogenous allele were transfected with the second knockout construct targeting the same locus and double drug-resistant clones were screened to isolate knockout lines. HMGA1/HMGA2 double knockout lines went through four transfection and screening rounds.

Analysis of RNA expression in putative knockout clones

To confirm that clones identified as knockouts by Southern blotting were indeed knockouts, total RNA from clones was analyzed for the presence of HMGA1 and HMGA2 mRNA. Northern blotting using a chicken HMGA1b cDNA as a probe was used to confirm loss of expression of HMGA1. RT-PCR was performed on total RNA using primers BB185 (5'-CCGGGCGGAGGGCTGAG-3') and BB186 (5'-GATT-CTTGTGACGATGTTTCTT-3') to confirm loss of expression of HMGA2. RT-PCR was performed using a Superscript One-Step RT-PCR kit (Invitrogen) and the following cycle temperatures and times: 92°C/2 min; (92°C/15 s, 65–55°C/15 s, 72°C/45 s) × 20 cycles (annealing temperature started at 65°C and decreased 0.5°C every cycle to 55°C); (92°C/15 s, 55°C/15 s, 72°C/45 s) × 35 cycles.

Monitoring cell growth rate

Cells were examined microscopically to make sure that cell viability was >90%, and counted with a Coulter Counter. Initial cultures were seeded at 1×10^5 cells/ml. Approximately 18 h after seeding, cells were counted again, and this count was used as the starting time point. Cells were counted in a Coulter Counter approximately every 12 h thereafter for 36 h.

Single round HIV-based vector infection

HIV-based vector particles transducing luciferase were generated as described previously (27; J. M. Kilzer, T. H. Stracker, B. F. Beitzel, K. Meek, M. D. Weitzman and F. Bushman, submitted for publication). Cells were spin infected (28) at 1200 g for 2 h and then incubated at 37°C for 2–3 days prior to quantitation of marker expression. Luciferase expression was quantitated using LucLite (Perkin Elmer, Boston, MA).

Spreading infection of avian retrovirus

RCASBP(C) and RCASBP(M2C)GFP were kindly provided by Stephen Hughes (National Cancer Institute, Frederick, MD). The GFP cassette was excised from RCASBP-(M2C)GFP by ClaI digest and cloned into ClaI-digested RCASBP(C) to generate RCASBP(C)GFP. RCASBP(C)GFP was transfected into DF-1 cells to generate replication-competent virus. Supernatant was removed from the transfected DF-1 cells and filtered with a 0.22 µm syringe filter to remove any suspended cells. This filtered supernatant was added to the DT40 cell lines to initiate infection and the spread of the infection was monitored by FACS analysis to detect GFP fluorescence.

Global analysis of cellular transcription using spotted cDNA arrays

RNA labeling and cDNA array hybridizations were performed by the Fred Hutchinson Cancer Research Center Genomics Core (Seattle, WA) as described previously (29).

RESULTS

Disruption of HMGA1

A cDNA encoding chicken HMGA1b was isolated by screening a stage 3 chicken embryo cDNA library with a human HMGA1 probe. Low stringency washes after hybridization identified several clones and sequencing of one of these clones (pBBcHMGA1b) identified a 1089 bp cDNA with a 288 bp coding region. *In silico* translation of the coding region gave a 96 amino acid protein with a predicted molecular weight of 10.5 kDa that was 71% identical to human HMGA1b and had completely conserved A-T hook DNA binding domains (Fig. 1A). This chicken HMGA1b cDNA was subsequently used as a probe to screen a chicken genomic library to identify clones containing the chicken HMGA1 locus. These clones were restriction mapped and a partial map is shown in Figure 1B.

Mapping of the HMGA1 locus showed that all of the coding exons were contained in a 1.9 kb region. Using this information, we generated knockout constructs pBB104KO and pBB105KO. Homologous recombination between these constructs and the endogenous locus replaces all of the coding exons with a blasticidin-S resistance cassette (pBB104KO) or a mycophenolic acid resistance cassette (pBB105KO). Following transfection of the knockout constructs into DT40 cells and selection with the appropriate antibiotic, genomic DNA from the resistant clones was isolated and screened by Southern hybridization. Clones were identified that had undergone homologous recombination, resulting in the disruption of one HMGA1 allele. These clones were then transfected with the second knockout construct (bearing the other antibiotic selection marker) and selected for resistance to both antibiotics. Doubly resistant clones were screened by Southern hybridization and those clones that had undergone a second round of homologous recombination, resulting in disruption of the second HMGA1 allele, were identified (Fig. 2A). Several knockouts were generated in this manner and three were chosen for further study. They were designated A1-1, A1-2 and A1-3. The order of transfections used in generating each of these clones is indicated in Table 1.

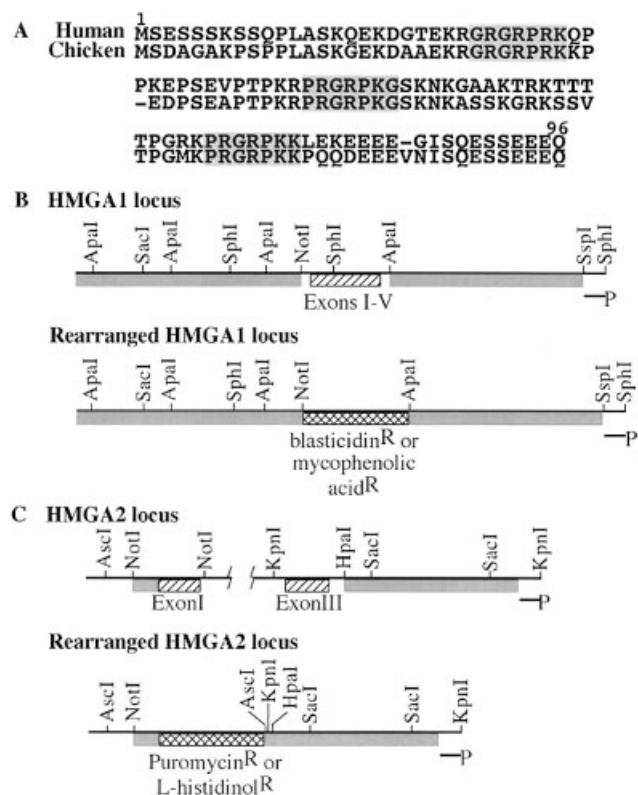


Figure 1. (A) Alignment of human and chicken HMGA1b. Gray boxes indicate position of A-T hook DNA binding domains. (B) Partial map of chicken HMGA1 locus and the locus after recombination with a knockout construct. Gray boxes indicate regions of genomic DNA included in the knockout construct. Hatched box represents HMGA1 coding exons (note: intron-exon boundaries were not mapped). Cross-hatched box represents antibiotic selection cassette. P indicates region of probe used on Southern blot to screen for homologous recombinants. (C) Partial map of 5' end of chicken HMGA2. Representations are as in (B). (B) and (C) are not drawn to scale.

Disruption of HMGA2

It has been shown previously that HMGA2 has activity similar to that of HMGA1 in *in vitro* integration assays (9). Using the chicken HMGA2 cDNA sequence (26), we designed primers for RT-PCR and analyzed DT40 RNA for the presence of HMGA2 message. We found that DT40 cells do express

HMGA2 (data not shown), so we decided to attempt to knockout HMGA2 as well as HMGA1. We obtained two cDNA clones and one genomic clone from Martin Groenen (26). The genomic clone, H34, only contained the 3' end of the coding region of HMGA2, so we used one of the cDNA clones, HCC3, to screen a chicken genomic DNA library to obtain clones that contained the 5' end of the HMGA2 coding region. After restriction mapping the 5' end of the HMGA2 locus, we decided on a knockout strategy that would remove exons 1–3, containing the A-T hook DNA binding domains of HMGA2 (Fig. 1C). We generated knockout constructs pBB102KO and pBB103KO and used those to target the endogenous HMGA2 alleles. Transfection and screening as described above resulted in the generation of HMGA2 knockout lines A2-1 and A2-2 (Fig. 2B and Table 1).

Co-disruption of HMGA1 and HMGA2

We were also able to generate lines in which both genes were disrupted. Using the reagents and methods described above, we generated double knockout lines A1/A2-1 and A1/A2-2 (Table 1).

HMGA expression in knockout lines

To confirm that cell lines designated as knockouts by Southern hybridization were indeed knockouts, total RNA was isolated from all cell lines and subjected to northern blotting to detect HMGA1 mRNA and RT-PCR to detect HMGA2 mRNA. As shown in Figure 3, clones A1-1, A1-2 and A1-3 have no detectable HMGA1 RNA, as expected. Clones A2-1 and A2-2 have no HMGA2 RNA detectable by RT-PCR, and clones A1/A2-1 and A1/A2-2 have no detectable HMGA1 RNA or HMGA2 RNA. Loss of HMGA2 does not perturb transcription of HMGA1, indicating that HMGA2 does not regulate HMGA1 transcription. Loss of HMGA1 does not appear to perturb the transcription level of HMGA2, but modest effects could have been missed in the RT-PCR used to detect the HMGA2 message.

Growth rate of cell lines is not dependent on HMGA expression status

As shown in Figure 4, the growth rate of the various HMGA wild-type and knockout cell lines was not dependent on the expression of HMGA1, HMGA2 or both together. There was some variability in the growth rate of the cell lines, but that variability does not correlate with HMGA expression. For

Table 1. Genotype and order of knockouts of DT40 cell clones used in this study

Cell line	Genotype	Knockout construct transfections				Gene disruption order
		1st	2nd	3rd	4th	
A1-1	HMGA1 $-/-$	pBB104KO	PBB105KO	NA	NA	NA
A1-2	HMGA1 $-/-$	pBB104KO	PBB105KO	NA	NA	NA
A1-3	HMGA1 $-/-$	pBB105KO	PBB104KO	NA	NA	NA
A2-1	HMGA2 $-/-$	pBB102KO	PBB103KO	NA	NA	NA
A2-2	HMGA2 $-/-$	pBB103KO	PBB102KO	NA	NA	NA
A1/A2-1	HMGA1 $-/-$ HMGA2 $-/-$	pBB102KO	PBB103KO	pBB104KO	pBB105KO	HMGA2 then HMGA1
A1/A2-2	HMGA1 $-/-$ HMGA2 $-/-$	pBB104KO	PBB105KO	pBB102KO	pBB103KO	HMGA1 then HMGA2

NA, not applicable to single gene knockout cell lines.

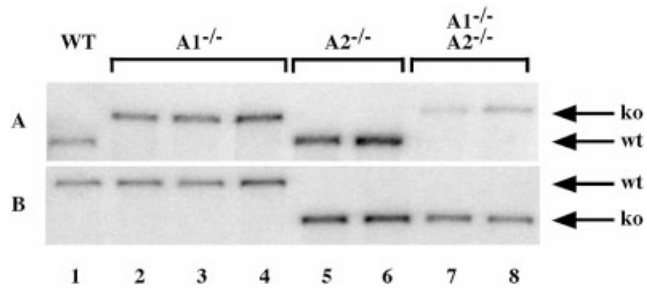


Figure 2. Southern blots showing genotype of various HMGA cell lines. Arrows indicate size of wild-type and knockout alleles. Cell lines are: (1) wild-type DT40, (2) A1-1, (3) A1-2, (4) A1-3, (5) A2-1, (6) A2-2, (7) A1/A2-1 and (8) A1/A2-2. (A) Southern blot of SphI-digested genomic DNA probed with DNA from region indicated in Figure 1B. (B) Southern blot of KpnI-digested genomic DNA probed with DNA from region indicated in Figure 1C.

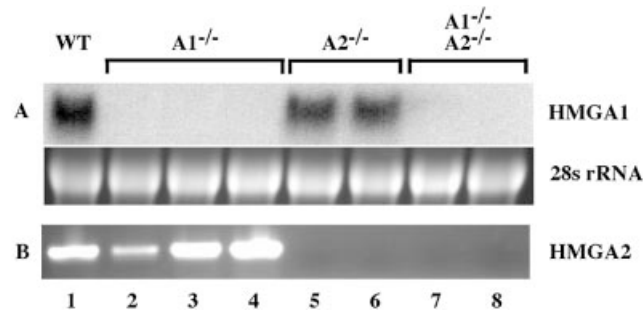


Figure 3. Expression of HMGA1 and HMGA2 in knockout cell lines. RNA samples are as follows: (1) wild type DT40, (2) A1-1, (3) A1-2, (4) A1-3, (5) A2-1, (6) A2-2, (7) A1/A2-1 and (8) A1/A2-2. (A) Northern showing HMGA1 expression. The ethidium bromide stained 28S rRNA image is included to show equal loading. (B) RT-PCR analysis of HMGA2 expression.

example, one double knockout line (A1/A2-1) was the fastest growing cell line in this experiment and the other double knockout line (A1/A2-2) was the slowest growing. This does indicate that quantitative changes in phenotype can be observed during subculture of particular cell lines (also see below). We did not observe any growth rate differences whether the cells were grown at 37 or 42°C with either 1 or 10% chicken serum in the growth medium.

Infection with HIV-based lentiviral vectors

In order to assess whether retroviral integration can occur in the absence of HMGA1 or HMGA2, we infected the HMGA knockout lines with vesicular stomatitis virus G protein (VSV-G) pseudotyped HIV-based lentiviral vectors (27) which transduce a luciferase expressing construct (J. M. Kilzer, T. H. Stracker, B. F. Beitzel, K. Meek, M. D. Weitzman and F. Bushman, submitted for publication). These vectors were produced by transfecting three plasmids into 293T cells. The first encodes all of the viral structural proteins and enzymes, the second encodes the VSV-G protein, and the third encodes a genomic RNA that is packaged into virions and contains a luciferase expression cassette. Two days after transfection, the vector-containing supernatant is removed

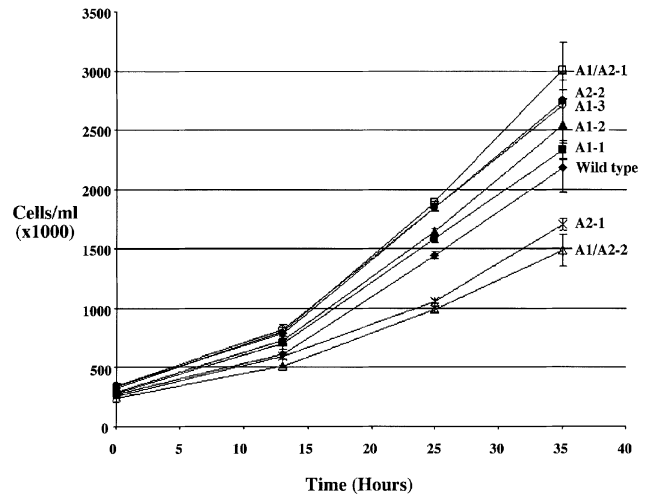


Figure 4. Growth rate of HMGA knockout lines. Wild-type DT40, closed diamonds; A1-1, closed squares; A1-2, closed triangles; A1-3, open circles; A2-1, crossed lines; A2-2, closed circles; A1/A2-1, open squares; A1/A2-2, open triangles.

from the 293T cells, filtered to remove any detached cells, and added to the various DT40 cell lines.

Approximately 3 days after infection, luciferase activity in the infected cells was assayed. As shown in Figure 5C, luciferase activity in all eight cell lines was relatively similar. There was some variation in the amount of luciferase activity from infected cell lines, but this variation did not correspond with HMGA genotype. In particular, clone A1/A2-2, a double knockout line, showed an ~2-fold increase in luciferase activity, but clone A1/A2-1, another double knockout line, did not show this increase. Luciferase activity was integrase dependent, as infection with an equal amount of vector carrying an integrase with a mutation in the catalytic site (E152A) had greatly reduced luciferase activity. From this we conclude that HMGA proteins are not strictly required for HIV-based vector integration in DT40 cells.

Spreading infection with RCAS vectors

Although HMGA is not strictly required for integration, as shown above, it is possible that the integration process could be slowed in the absence of HMGA, and this would not necessarily be detected in a single round infection assay. The activity of avian retroviral integrase is stimulated *in vitro* by the addition of HMGA1 (30), so we decided to see whether a spreading infection by a replication competent avian retrovirus was slowed in HMGA knockout lines. RCASBP(C)GFP was added to the cells and the spread of the infection was monitored by FACS analysis. We did detect an overall increase in the percentage of GFP-positive cells over the course of the 11 day infection (Fig. 6), consistent with viral spread. However, the spread of the GFP marker appeared inefficient, potentially due to sluggish growth of the virus or loss of the GFP marker during multiple rounds of replication. Clone A1/A2-2 showed more rapid spread than other clones at early time points, a result which may be related to the increased infection of this line by HIV-based vectors. Considering all of the cell lines together, the rate of increase in the percentage of GFP-positive cells did not correlate with

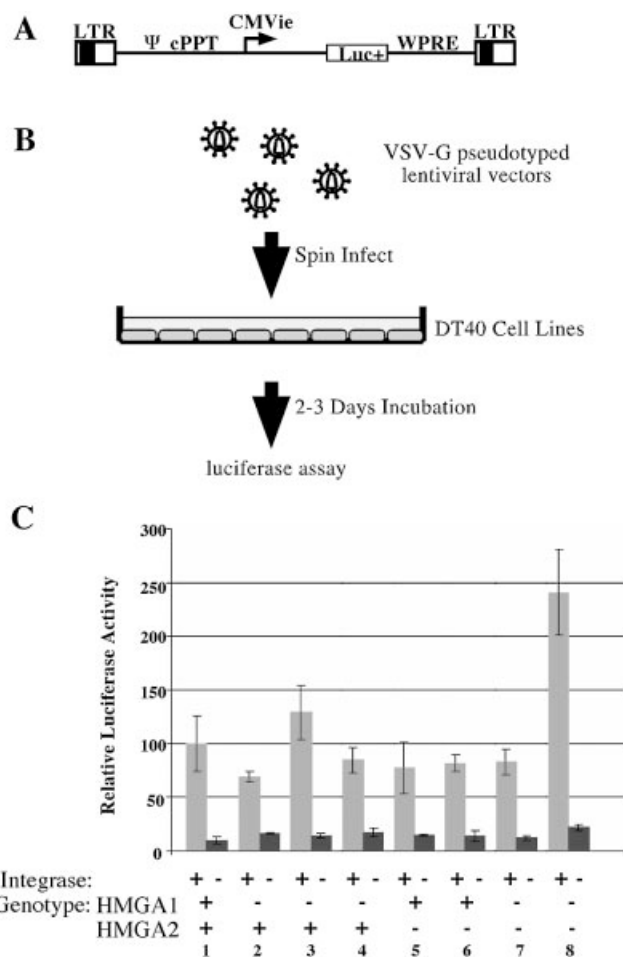


Figure 5. Single round infection of HMGA knockout lines with HIV-based lentiviral vectors. (A) Diagram of the HIV-based lentiviral vector construct. LTR, long terminal repeat; Ψ , genome packaging signal; cPPT, central polypurine tract; CMVie, CMV immediate early promoter; Luc+, luciferase coding region; WPRE, woodchuck hepatitis virus post transcriptional regulatory element. (B) Strategy for infecting DT40 cell lines. Lentiviral vectors were generated by transfection of 293T cells with vector-producing constructs. Supernatant was collected from cells and filtered to remove cellular debris. DT40 cells were spin-infected and incubated for 2–3 days at 37°C. After the incubation period, cells were lysed and luciferase activity was quantified. (C) Relative luciferase activity after infection with lentiviral vectors. All values are normalized to wild-type DT40 luciferase activity. Integrase (+) values represent the average of three experiments. Integrase (-) values represent the average of two experiments. Cell lines are: (1) wild-type DT40, (2) A1-1, (3) A1-2, (4) A1-3, (5) A2-1, (6) A2-2, (7) A1/A2-1 and (8) A1/A2-2.

HMGA expression status, indicating that the spreading infection was not slowed by the absence of HMGA. From this we concluded that the rate of integration is not significantly slower in HMGA knockout lines.

Transcriptional profiling

In order to carry out a global analysis of transcriptional activity in HMGA knockout cell lines, we used a chicken cDNA array containing approximately 4000 cDNAs (29) Each RNA sample was analyzed on two chips, one with Cy3-labeled wild-type RNA and Cy5-labeled knockout RNA, and one with the dyes reversed to prevent dye bias in the

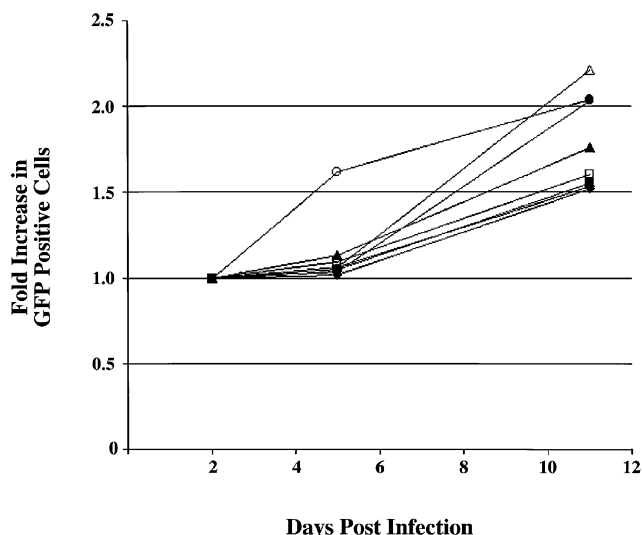


Figure 6. Spreading infection of RCASBP(C)GFP. Wild-type DT40, closed diamonds; A1-1, closed squares; A1-2, closed triangles; A1-3, closed circles; A2-1, open diamonds; A2-2, open squares; A1/A2-1, open triangles; A1/A2-2, open circles.

analysis. A total of 10 chips were used to analyze RNA from HMGA1 knockout lines (lines A1-1 and A1-3 were analyzed twice each) and four chips each were used to analyze HMGA2 knockout lines and double knockout lines. We did not see any changes in gene expression >3-fold increased or decreased, and some of the more modest changes we saw were not observed in northern blot analysis of those same genes (data not shown). Evidently, HMGA proteins do not play a large role in the regulation of the genes included in the array in DT40 cells.

DISCUSSION

In this work we have cloned chicken HMGA1, generated cell lines lacking HMGA1, HMGA2 or both in combination, and analyzed these cell lines for effects on cell growth, cellular transcription and retroviral integration.

DT40 cells lacking HMGA1 and HMGA2 are viable, have no apparent defects in cell growth, and show no change in the expression levels of approximately 4000 genes. This is a surprising result for several reasons. First, HMGA1 and HMGA2 have been shown to play a direct role in the regulation of a large number of genes (11) and HMGA proteins are thought to play a more general role in the regulation of transcription over large areas of chromatin by removing histone H1 from SARs, thereby de-repressing transcription in those regions (17). Secondly, HMGA dysregulation is seen in a large number of human neoplasias and there is a strong correlation between high HMGA expression levels and increased malignancy (21). Lastly, manipulation of HMGA levels in mice causes dramatic phenotypes. Although mouse ES cells in which HMGA1 has been knocked out have been described (23), there have been no reports of HMGA1 knockout mouse lines, consistent with HMGA1 being essential for mouse development or fertility. Knockout mice deficient for HMGA2 have a pygmy phenotype (24) and have less adipose tissue (31).

Some previously identified HMGA-regulated genes, such as CD44 (32), were present on the cDNA array used in this study, but most of the well characterized HMGA-regulated genes, such as IFN- β (12,14,33,34), IL-2 (35,36) and IL2R α (37–39), were not represented. Since only a small subset of chicken genes are represented on the cDNA array used in this study, many possible HMGA-regulated genes may have been missed.

One possibility for the lack of a phenotype in our DT40 cell knockouts could be that other HMGA family members are encoded in the chicken genome. To check for this, we searched the BBSRC Chicken EST Database (www.chick.umist.ac.uk/cgi-bin/chicken_database.cgi) for ESTs related to already identified chicken HMGA genes (40). This database contains 340 000 ESTs from 21 different chicken tissues and is the largest chicken EST database publicly available. We searched this database using the tblastx program using permissive criteria (Blosum45 matrix). We identified four EST clones that contained coding sequences identical, or nearly so, to pBBcHMGA1b (one clone contained two single nucleotide insertions). One of the clones may represent a splice variant, as its 5' untranslated region was unrelated to the other clones. Similar searches with the chicken HMGA2 sequence failed to locate clones with sequences other than those already identified. Although these searches are not conclusive, they do support the idea that there are no other HMGA family members in chickens.

Since the HMGA knockout clones all originated as single drug-resistant cells, it is possible that during selection they adapted to growth without HMGA proteins so that the phenotypes of the lines do not faithfully represent the role of HMGA proteins. Although this is not strictly ruled out, we have partially addressed this question by using RNA interference (RNAi) to knock down HMGA1 expression in human cell lines (data not shown). Using RNAi, we have been able to suppress HMGA1 expression by up to 95% without an apparent phenotype in most cell lines that we have tested. A few cell lines appear to have a slower growth rate when HMGA1 is suppressed, indicating that HMGA1 requirements are cell-type specific.

Another class of HMG proteins, HMGN, has also been knocked out in DT40 cells, and was similarly found to lack an obvious phenotype (41,42). It is possible that HMG proteins may have some functional redundancy and loss of only one class could be complemented by another. Alternatively, HMGA and HMGN expression may only be required in certain cell types or at certain times during development or may not be required at all during growth in cell culture. It is also possible that viability in the absence of these proteins is a peculiarity of DT40 cells.

Cell lines lacking HMGA did not show reduced efficiency of retroviral replication. Single round infection assays showed no difference in integration efficiency in the absence of HMGA proteins and viral spread was not slowed in a spreading infection assay with an avian retrovirus. DT40 cells may express other factors that could be involved in integration, including BAF (43) or LEDGF (44). Alternatively, retroviral integration may not require host-encoded cofactors, relying instead on virus-encoded cofactors such as HIV nucleocapsid (22,45–50).

The cell lines described in this work add to a growing number of DT40 cell knockout lines. A large pool of DT40 knockouts, combined with the reagents generated during the production of those lines, creates a powerful tool to probe the interactions of a large number of proteins. As long as a gene has been knocked out previously, it is a relatively easy task to combine knockouts and analyze the effects of multiple genes on processes in DT40 cells.

ACKNOWLEDGEMENTS

We would like to thank Jean-Marie Buerstedde, Steve Hughes, Martyn Goulding and Martin Groenen for reagents and advice, and members of the Infectious Disease Laboratory at the Salk Institute for their help and useful discussions. This work was supported by grants AI34786 and AI53820 from the National Institutes of Health, the James B. Pendleton Charitable Trust, the Berger Foundation, the Margaret Morris Foundation and Frederic and Robin Withington.

REFERENCES

- Goodwin,G.H., Sanders,C. and Johns,E.W. (1973) A new group of chromatin-associated proteins with a high content of acidic and basic amino acids. *Eur. J. Biochem.*, **38**, 14–19.
- Grosschedl,R., Giese,K. and Pagel,J. (1994) HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. *Trends Genet.*, **10**, 94–100.
- Johnson,K.R., Disney,J.E., Wyatt,C.R. and Reeves,R. (1990) Expression of mRNAs encoding mammalian chromosomal proteins HMG-I and HMG-Y during cellular proliferation. *Exp. Cell Res.*, **187**, 69–76.
- Johnson,K.R., Lehn,D.A. and Reeves,R. (1989) Alternative processing of mRNAs encoding mammalian chromosomal high-mobility-group proteins HMG-I and HMG-Y. *Mol. Cell. Biol.*, **9**, 2114–2123.
- Naggal,S., Ghosn,C., DiSepio,D., Molina,Y., Sutter,M., Klein,E.S. and Chandraratna,R.A. (1999) Retinoid-dependent recruitment of a histone H1 displacement activity by retinoic acid receptor. *J. Biol. Chem.*, **274**, 22563–22568.
- Ashar,H.R., Cherath,L., Przybysz,K.M. and Chada,K. (1996) Genomic characterization of human HMGIC, a member of the accessory transcription factor family found at translocation breakpoints in lipomas. *Genomics*, **31**, 207–214.
- Reeves,R. (2001) Molecular biology of HMGA proteins: hubs of nuclear function. *Gene*, **277**, 63–81.
- Huth,J.R., Bewley,C.A., Nissen,M.S., Evans,J.N., Reeves,R., Gronenborn,A.M. and Clore,G.M. (1997) The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nature Struct. Biol.*, **4**, 657–665.
- Li,L., Yoder,K., Hansen,M.S., Olvera,J., Miller,M.D. and Bushman,F.D. (2000) Retroviral cDNA integration: stimulation by HMG I family proteins. *J. Virol.*, **74**, 10965–10974.
- Maher,J.F. and Nathans,D. (1996) Multivalent DNA-binding properties of the HMG-1 proteins. *Proc. Natl Acad. Sci. USA*, **93**, 6716–6720.
- Reeves,R. and Beckerbauer,L. (2001) HMGI/Y proteins: flexible regulators of transcription and chromatin structure. *Biochim. Biophys Acta*, **1519**, 13–29.
- Yie,J., Merika,M., Munshi,N., Chen,G. and Thanos,D. (1999) The role of HMG I(Y) in the assembly and function of the IFN-beta enhanceosome. *EMBO J.*, **18**, 3074–3089.
- Munshi,N., Merika,M., Yie,J., Senger,K., Chen,G. and Thanos,D. (1998) Acetylation of HMG I(Y) by CBP turns off IFN beta expression by disrupting the enhanceosome. *Mol. Cell*, **2**, 457–467.
- Yie,J., Liang,S., Merika,M. and Thanos,D. (1997) Intra- and intermolecular cooperative binding of high-mobility-group protein I(Y) to the beta-interferon promoter. *Mol. Cell. Biol.*, **17**, 3649–3662.
- Falvo,J.V., Thanos,D. and Maniatis,T. (1995) Reversal of intrinsic DNA bends in the IFN beta gene enhancer by transcription factors and the architectural protein HMG I(Y). *Cell*, **83**, 1101–1111.

16. Thanos,D. and Maniatis,T. (1995) Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. *Cell*, **83**, 1091–1100.
17. Zhao,K., Kas,E., Gonzalez,E. and Laemmli,U.K. (1993) SAR-dependent mobilization of histone H1 by HMG-I/Y *in vitro*: HMG-I/Y is enriched in H1-depleted chromatin. *EMBO J.*, **12**, 3237–3247.
18. Weintraub,H. (1985) Assembly and propagation of repressed and depressed chromosomal states. *Cell*, **42**, 705–711.
19. Laybourn,P.J. and Kadonaga,J.T. (1991) Role of nucleosomal cores and histone H1 in regulation of transcription by RNA polymerase II. *Science*, **254**, 238–245.
20. Kamakaka,R.T. and Thomas,J.O. (1990) Chromatin structure of transcriptionally competent and repressed genes. *EMBO J.*, **9**, 3997–4006.
21. Tallini,G. and Dal Cin,P. (1999) HMGI(Y) and HMGI-C dysregulation: a common occurrence in human tumors. *Adv. Anat. Pathol.*, **6**, 237–246.
22. Farnet,C.M. and Bushman,F.D. (1997) HIV-1 cDNA integration: requirement of HMG I(Y) protein for function of preintegration complexes *in vitro*. *Cell*, **88**, 483–492.
23. Baldassarre,G., Battista,S., Belletti,B., Thakur,S., Pentimalli,F., Trapasso,F., Fedele,M., Pierantoni,G., Croce,C.M. and Fusco,A. (2003) Negative regulation of BRCA1 gene expression by HMGA1 proteins accounts for the reduced BRCA1 protein levels in sporadic breast carcinoma. *Mol. Cell. Biol.*, **23**, 2225–2238.
24. Zhou,X., Benson,K.F., Ashar,H.R. and Chada,K. (1995) Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. *Nature*, **376**, 771–774.
25. Anand,A. and Chada,K. (2000) *In vivo* modulation of Hmgic reduces obesity. *Nature Genet.*, **24**, 377–380.
26. Ruyter-Spira,C.P., Herbergs,J., Limpens,E., Marsh,J.A., van der Poel,J.J., Ayoubi,T.A. and Groenen,M.A. (1998) Nucleotide sequence of the chicken HMGI-C cDNA and expression of the HMGI-C and IGF1 genes in autosomal dwarf chicken embryos. *Biochim. Biophys Acta*, **1399**, 83–87.
27. Naldini,L., Blomer,U., Gallyas,P., Ory,D., Mulligan,R., Gage,F.H., Verma,I.M. and Trono,D. (1996) *In vivo* gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science*, **272**, 263–267.
28. O'Doherty,U., Swiggard,W.J. and Malim,M.H. (2000) Human immunodeficiency virus type 1 spinoculation enhances infection through virus binding. *J. Virol.*, **74**, 10074–10080.
29. Neiman,P.E., Grbic,J.J., Polony,T.S., Kimmel,R., Bowers,S.J., Delrow,J. and Beemon,K.L. (2003) Functional genomic analysis reveals distinct neoplastic phenotypes associated with c-myc mutation in the bursa of Fabricius. *Oncogene*, **22**, 1073–1086.
30. Hindmarsh,P., Ridky,T., Reeves,R., andrake,M., Skalka,A.M. and Leis,J. (1999) HMG protein family members stimulate human immunodeficiency virus type 1 and avian sarcoma virus concerted DNA integration *in vitro*. *J. Virol.*, **73**, 2994–3003.
31. Benson,K.F. and Chada,K. (1994) Mini-mouse: phenotypic characterization of a transgenic insertional mutant allelic to pygmy. *Genet. Res.*, **64**, 27–33.
32. Foster,L.C., Wiesel,P., Huggins,G.S., Panares,R., Chin,M.T., Pellacani,A. and Perrella,M.A. (2000) Role of activating protein-1 and high mobility group-I(Y) protein in the induction of CD44 gene expression by interleukin-1beta in vascular smooth muscle cells. *FASEB J.*, **14**, 368–378.
33. Thanos,D. and Maniatis,T. (1992) The high mobility group protein HMG I(Y) is required for NF-kappa B-dependent virus induction of the human IFN-beta gene. *Cell*, **71**, 777–789.
34. Bonnefoy,E., Bandu,M.T. and Doly,J. (1999) Specific binding of high-mobility-group I (HMGI) protein and histone H1 to the upstream AT-rich region of the murine beta interferon promoter: HMGI protein acts as a potential antirepressor of the promoter. *Mol. Cell. Biol.*, **19**, 2803–2816.
35. Himes,S.R., Reeves,R., Attema,J., Nissen,M., Li,Y. and Shannon,M.F. (2000) The role of high-mobility group I(Y) proteins in expression of IL-2 and T cell proliferation. *J. Immunol.*, **164**, 3157–3168.
36. Himes,S.R., Coles,L.S., Reeves,R. and Shannon,M.F. (1996) High mobility group protein I(Y) is required for function and for c-Rel binding to CD28 response elements within the GM-CSF and IL-2 promoters. *Immunity*, **5**, 479–489.
37. Reeves,R., Leonard,W.J. and Nissen,M.S. (2000) Binding of HMG-I(Y) imparts architectural specificity to a positioned nucleosome on the promoter of the human interleukin-2 receptor alpha gene. *Mol. Cell. Biol.*, **20**, 4666–4679.
38. John,S., Reeves,R.B., Lin,J.-X., Child,R., Leiden,J.M., Thompson,C.B. and Leonard,W.J. (1995) Regulation of cell-type-specific interleukin-2 receptor alpha-chain gene expression: potential role of physical interactions between Elf-1, HMG I(Y) and NF KB family proteins. *Mol. Cell. Biol.*, **15**, 1786–1796.
39. John,S., Robbins,C.M. and Leonard,W.J. (1996) An IL-2 response element in the human IL-2 receptor alpha chain promoter is a composite element that binds Stat5, Elf-1, HMG-I(Y) and a GATA family protein. *EMBO J.*, **15**, 5627–5635.
40. Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E., Fong,W.T., Tickle,C., Brown,W.R., Wilson,S.A. and Hubbard,S.J. (2002) A comprehensive collection of chicken cDNAs. *Curr. Biol.*, **12**, 1965–1969.
41. Li,Y. and Dodgson,J.B. (1995) The chicken HMG-17 gene is dispensable for cell growth *in vitro*. *Mol. Cell. Biol.*, **15**, 5516–5523.
42. Li,Y., Strahler,J.R. and Dodgson,J.B. (1997) Neither HMG-14a nor HMG-17 gene function is required for growth of chicken DT40 cells or maintenance of DNaseI-hypersensitive sites. *Nucleic Acids Res.*, **25**, 283–288.
43. Lee,M.S. and Craigie,R. (1998) A previously unidentified host protein protects retroviral DNA from autointegration. *Proc. Natl Acad. Sci. USA*, **95**, 1528–1533.
44. Cherepanov,P., Maertens,G., Proost,P., Devreese,B., Van Beeumen,J., Engelborghs,Y., De Clercq,E. and Debysse,Z. (2003) HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *J. Biol. Chem.*, **278**, 372–381.
45. Buckman,J.S., Bosche,W.J. and Gorelick,R.J. (2003) Human immunodeficiency virus type 1 nucleocapsid zn(2+) fingers are required for efficient reverse transcription, initial integration processes and protection of newly synthesized viral DNA. *J. Virol.*, **77**, 1469–1480.
46. Cimarelli,A., Sandin,S., Høglund,S. and Luban,J. (2000) Rescue of multiple viral functions by a second-site suppressor of a human immunodeficiency virus type 1 nucleocapsid mutation. *J. Virol.*, **74**, 4273–4283.
47. Carreau,S., Gorelick,R.J. and Bushman,F.D. (1999) Coupled integration of human immunodeficiency virus type 1 cDNA ends by purified integrase *in vitro*: stimulation by the viral nucleocapsid protein. *J. Virol.*, **73**, 6670–6679.
48. Lapadat-Tapolsky,M., De Rocquigny,H., Van Gent,D., Roques,B., Plasterk,R. and Darlix,J.L. (1993) Interactions between HIV-1 nucleocapsid protein and viral DNA may have important functions in the viral life cycle. *Nucleic Acids Res.*, **21**, 831–839.
49. Carreau,S., Batson,S.C., Poljak,L., Mouscadet,J.F., de Rocquigny,H., Darlix,J.L., Roques,B.P., Kas,E. and Auclair,C. (1997) Human immunodeficiency virus type 1 nucleocapsid protein specifically stimulates Mg²⁺-dependent DNA integration *in vitro*. *J. Virol.*, **71**, 6225–6229.
50. Gao,K., Gorelick,R.J., Johnson,D.G. and Bushman,F. (2003) Cofactors for human immunodeficiency virus type 1 cDNA integration *in vitro*. *J. Virol.*, **77**, 1598–1603.