

Expression of the murine *RanBP1* and *Htf9-c* genes is regulated from a shared bidirectional promoter during cell cycle progression

Giulia GUARGUAGLINI, Alessandra BATTISTONI, Carmine PITTOGGI, Gigliola DI MATTEO, Barbara DI FIORE and Patrizia LAVIA*

CNR Centre of Evolutionary Genetics, c/o Department of Genetics and Molecular Biology, University 'La Sapienza', Via degli Apuli 4, Rome 00185, Italy

The murine *Htf9-a/RanBP1* and *Htf9-c* genes are divergently transcribed from a bidirectional promoter. The *Htf9-a* gene encodes the RanBP1 protein, a major partner of the Ran GTPase. The divergently transcribed *Htf9-c* gene encodes a protein sharing similarity with yeast and bacterial nucleic acid-modifying enzymes. We report here that both mRNA species produced by the *Htf9*-associated genes are regulated during the cell cycle progression, peak in S phase and decrease during mitosis. Transient expression experiments with reporter constructs showed that cell cycle expression is controlled at the transcriptional level, because the bidirectional *Htf9* promoter is down-

regulated in growth-arrested cells, is activated at the G₁/S transition and reaches maximal activity in S phase, though with a different efficiency for each orientation. We have delimited specific promoter regions controlling S phase activity in one or both orientations: identified elements contain recognition sites for members belonging to both the E2F and Sp1 families of transcription factors. Together, the results suggest that the sharing of the regulatory region supports co-regulation of the *Htf9-a/RanBP1* and *Htf9-c* genes in a common window of the cell cycle.

INTRODUCTION

The mouse *Htf9* locus was originally cloned for its association with a CpG-rich DNA sequence [1]; these sequences are now widely assumed to be diagnostic of protein-encoding genes [2], preferentially of the housekeeping type [3]. In the *Htf9* locus, two independent transcriptional units were identified on each DNA strand [4], which are both expressed in many tissues and cell types. The major transcription start sites of the divergently transcribed genes in the *Htf9* locus show coincident locations on complementary DNA strands (schematized in Figure 1).

The lower-strand gene, which we originally named *Htf9-a*, encodes Ran-binding protein 1 (RanBP1) [5,6], one of the molecular partners of Ran, a nuclear protein homologous with Ras in the GTP-binding domain. The Ran GTPase network is implicated in the control of several processes, including the initiation of DNA replication, cell cycle progression, monitoring completion of DNA replication before entry into mitosis, nuclear structure, nuclear protein import and RNA export (reviewed in [7,8]). Most structural genes of the Ran network identified so far, i.e. *Ran*, *RCC1* (regulator of chromosome condensation) and *RanGAP* (Ran GTPase-activating protein), are constitutively active. In contrast, we have found that the *RanBP1* gene is inactive in quiescent cells and becomes actively expressed on entry into the cell division cycle, being maximally expressed in S phase; cell cycle regulation of the *RanBP1* gene is essentially under the control of E2F activators and repressors of the retinoblastoma gene family [9]. This cyclic pattern of expression is consistent with the regulative role hypothesized for RanBP1 in control of Ran activity during S phase [10,11].

The upper-strand *Htf9-c* gene, divergently transcribed from *Htf9-a/RanBP1*, encodes a protein of unidentified function; results presented here show that the Htf9-c protein harbours a previously unidentified domain homologous with yeast and

bacterial nucleic acid-modifying enzymes. The unusual organization of the genes in the *Htf9* locus suggests that they might be transcribed in a co-ordinated fashion and therefore might provide a hint to identify a novel cell-cycle-related gene arranged head-to-head with *RanBP1*. To investigate this possibility, we have undertaken to analyse the expression of both *Htf9*-associated genes during cell cycle progression. Using different cell cycle synchronization techniques we have found that the divergently transcribed genes are indeed expressed in a cell cycle-dependent fashion; both show the highest level of mRNA expression during S phase and are down-regulated in mitotic cells. In addition, we show that S phase-dependent activity is controlled at the transcriptional level for both genes; the information directing S phase up-regulation in both orientations resides in a 273-bp region from the shared bidirectional promoter. Together these results suggest that the head-to-head arrangement of the *Htf9*-associated genes with juxtaposed 5' ends within the shared promoter might ensure their co-regulation in a common temporal window during cell cycle progression.

MATERIALS AND METHODS

Htf9-c cDNA subcloning and sequence analysis

The λ gt10 subclone L6 contains the *Htf9-c* complete cDNA. The cDNA insert had been previously digested with *EcoRI* and *PstI* restriction endonucleases and the resulting fragments had been subcloned in pUC vectors [4]. We have now generated novel overlapping subclones by single digestion of the original insert with *PvuII*, *XbaI* and *EcoRV*. We also sequenced fragments (available under accession numbers Y08060, Y07806 and Y08059) from the genomic clone pL9.5, flanking the *Htf9* CG-rich genomic region (see [1]), which contains intronic sequences

Abbreviations used: CAT, chloramphenicol acetyltransferase; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; RanBP1, Ran-binding protein 1.

* To whom correspondence should be addressed.

The nucleotide sequence for the the hypothetical Htf9-c protein reported here appears in DDBJ, EMBL and GenBank Nucleotide Sequence Databases under the accession number X56044.

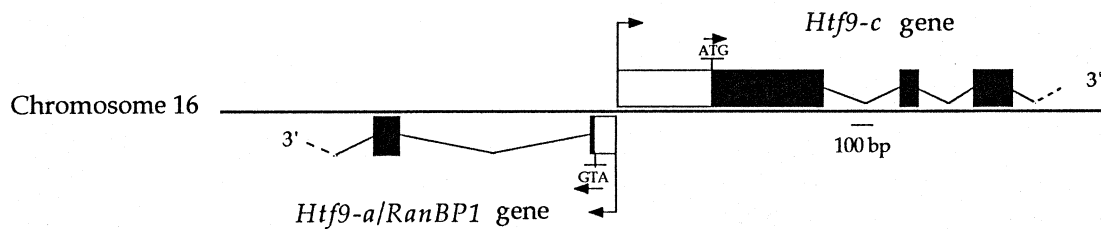


Figure 1 Map showing the head-to-head organization of the divergently transcribed *Htf9-a/RanBP1* and *Htf9-c* genes on murine chromosome 16

Only the 5' ends of both genes are shown. The major mRNA transcription start sites of both genes are indicated by vertical arrows; the empty boxes represent untranslated regions; the filled boxes represent the coding exonic sequences; the broken lines represent introns. The genomic location of the most 3' exons is not determined. The start codon in each open reading frame is marked; the orientation of transcription and translation is arrowed (data from [1,4]).

and part of exons 3 and 4 of the *Htf9-c* gene. All sequences were determined on both DNA strands by using the direct and reverse universal pUC primers as well as *Htf9-c*-specific primers (5'-ACTCTCTTCACCTTCGG-3' and 5'-AGGAGGCTGTGGAGGAT-3'). Sequencing data were analysed with the Fasta, Prosite and BLITZ search programs, the BLASTP and BEAUTY alignment program and the BIOSCAN multiple alignment program.

Cell cultures and synchronization

All experiments were performed with murine NIH/3T3 fibroblast cultures (ATCC CRL 1658) grown in Dulbecco's modified Eagle's medium supplemented with 10% (v/v) fetal calf serum in an air/CO₂ (19:1) atmosphere. Briefly, G₀/G₁ synchronization was obtained by inducing proliferation arrest after serum starvation for 48 h; subsequent re-entry into the cell division cycle was induced by adding fresh serum [15% (v/v) final concentration]. Synchronous progression through S phase was obtained by arresting the cells at the G₁/S transition in the presence of hydroxyurea (0.5 mM final concentration for 16 h) and subsequently transferring the cells to hydroxyurea-free medium. Finally, cells were synchronized in G₂ in the presence of 0.2 µg/ml nocodazole for 10 h; the block was released by transferring the cells to nocodazole-free medium. For all three types of synchronization, release of the block and progression into the next phase of the cycle were monitored in cells collected at regular intervals after the block release and subjected to FACS analysis after incubation with propidium iodide (50 µg/ml of cell sample). In certain experiments, cell cycle progression was analysed by determining the simultaneous incorporation of both propidium iodide and 5-bromodeoxyuridine into replicating DNA as previously described [9]. Cell samples were analysed in a FACStar Plus cytofluorimeter with either the Multicycle (DNA content) or the WinMDI (simultaneous determination of the DNA content and bromodeoxyuridine incorporation) software (10000 events per sample).

Northern blot hybridization experiments

Total RNA was extracted by the acid guanidinium/phenol/chloroform extraction protocol. Aliquots (40 µg) of total RNA dissolved in formamide/formaldehyde buffer were loaded on 1.5% (w/v) agarose/formaldehyde gels and run in 1 × Mops buffer. Gels were stained to reveal the 18 S and 28 S ribosomal bands, then blotted on GeneScreen membranes, UV cross-linked, hybridized and washed as described [9]. Northern blots included a control: a subclone of cDNA for glyceraldehyde-3-phosphate

dehydrogenase (GAPDH). The *RanBP1* probe was a gel-purified fragment corresponding to the *Htf9-a* cDNA (accession number X56045); the *Htf9-c* probe was gel-purified from the corresponding cDNA subclone. Gene expression was monitored throughout the cell cycle with cDNA probes corresponding to cyclin E and D1 (early G₁) as well as cyclin A (late G₁ and S phase expression). All probes were labelled by using the random-priming method. Results were analysed by densitometric scanning of autoradiographs and by direct counting of the radioactivity on hybridized filters with a ³²P instant imager (Canberra Packard).

Promoter expression assays

Reporter constructs were synthesized by inserting different portions of the *Htf9* genomic sequence (accession number X05830) in both orientations upstream of the coding sequence for the chloramphenicol acetyltransferase (CAT) enzyme in the pSV0-CAT expression vector. Most assayed constructs were described previously [9,12]. The pEA series of reporter constructs were obtained by inserting an *EarI*-*AluI* 74 bp fragment encompassing the major *Htf9* transcription start site in both orientations in the *HindIII* site of the pSV0-CAT vector. The pPS-C construct was obtained by ligating the *TaqI*-*AluI* fragment (50 bp from the 5' end of the pTS-C construct truncated at the *AluI* site; see sequence of entry X05830) to the promoter fragment in the pNS-C construct: thus the resulting promoter carries a 20 bp deletion immediately downstream of the *AluI* end that removes the sequence 5'-CCTTTCCTCCGCGTCTGGCG-3'; the E2Fa site is underlined. The pmES-A and pmES-C constructs carry mutated sequences replacing sites Sp1.3 and E2Fb (see maps in Figure 6) and were obtained by replacing the *Sau96I*-*XmaIII* 60 bp fragment from the native *Htf9* promoter with a corresponding oligonucleotide carrying the following substitutions: 5'-GAGATGCC-3' replacing GGGGCGGG (site Sp1.3), and 5'-TTACTCAGA-3' replacing TTTGGCGGG (site E2Fb). The mutated fragment was cloned in both orientations upstream of the CAT coding sequence. Control promoter constructs included pE1A-CAT, which carries the cell-cycle-independent promoter of the adenovirus 5 E1A gene upstream of the CAT coding sequence [13], and pA10-CAT₂, carrying a minimal promoter composed of a TATA box and two Sp1-binding sites. NIH/3T3 cells were lipofected with the DOTAP reagent (Boehringer), 5 µg of CAT reporter construct and 2 µg of pCMV-lac Z plasmid. For cell cycle analysis experiments, cells were passaged from one large culture flask on day 1. On day 2, constructs were lipofected; 6 h after lipofection, the medium was changed and replaced with low-serum medium (0.5%) to arrest proliferation. Fetal calf

serum (15%, v/v) was added back 48 h after starvation (day 4); cell samples were harvested on day 5 at regular intervals after re-stimulation (see Figure 6). Transfection experiments were repeated three to seven times for each time point on duplicate sets of cultures, which were subjected to determination of CAT activity and to FACS monitoring of the cycle progression. Promoter mapping analysis was performed by the same general procedure, except that re-stimulated cells were harvested 15–18 h after refeeding with serum. Promoter strengths were quantified by immunodetection of the CAT and β -galactosidase enzymes, with the CAT-ELISA and β -gal ELISA kits.

RESULTS

Htf9-c gene product harbours regions of similarity to nucleic acid-modifying enzymes

We have re-examined the coding sequence of the *Htf9-c* gene [4], and have identified sequencing errors in the 3' region that resulted in a prematurely terminated open reading frame. The correct sequence of the *Htf9-c* gene was newly determined from overlapping subclones generated from the original cDNA insert

by using single-cut restriction endonucleases, and was found to encode a hypothetical protein product of 676 amino acid residues (75 kDa). The predicted sequence of the Htf9-c protein remains structurally unrelated to that of the RanBP1 protein encoded by the divergently transcribed *Htf9-a* gene, as previously reported [4]. Sequence inspection of the Htf9-c protein product (Figure 2A) revealed significant similarities in the C-terminal region to several characterized proteins, shown in Figure 2(B). These proteins include the *Escherichia coli* and *Haemophilus influenzae* tRNA (uracil-5-)-methyltransferase enzymes; two related bacterial sequences that are also thought to encode RNA-binding proteins, i.e. the hypothetical HI0333 and Y958 *S*-adenosyl-methionine-dependent tRNA methyltransferases; the *Chlamidia trachomatis* pCTHom1 gene, which encodes a late developmental protein with some homeoprotein similarity; and finally the hypothetical *Saccharomyces pombe* YAD7 nuclease and the homologous *S. cerevisiae* NucR endo-exonuclease, which has hydrolase, exonuclease and endonuclease activity on single-stranded and double-stranded DNA. The murine *Htf9-c* coding sequence specifies a larger N-terminal region, rich in potential modification sites, within which we noticed a putative nuclear

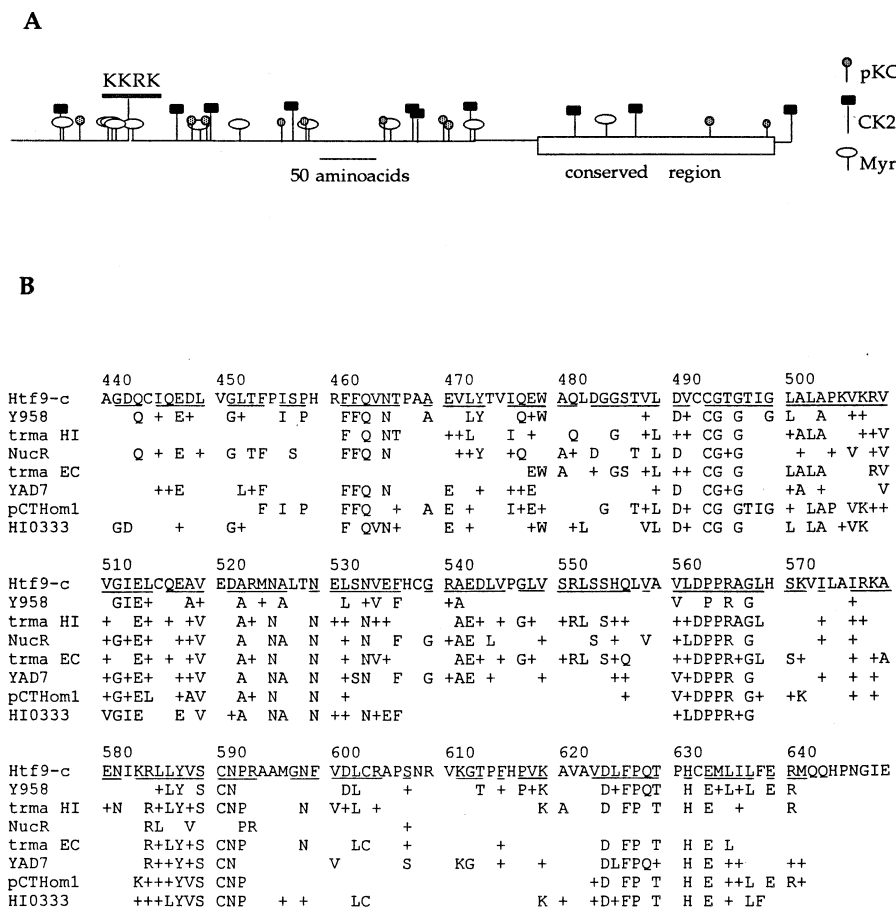


Figure 2 *Htf9-c* protein

(A) Schematic representation of the hypothetical Htf9-c protein. Potential phosphorylation sites for protein kinase C (pKC) and casein kinase II (CK2) and N-myristylation (Myr) sites are indicated. The KKRK core indicates a putative nuclear targeting signal. The box represents a conserved domain whose sequence is shown in (B). (B) Alignment of the C-terminal domain of the hypothetical Htf9-c protein and GenBank protein sequences. The diagram summarizes the results of single and multiple alignment searches with the BLITZ, BLASTP, BEAUTY and BIOSCAN programs. Identical amino acids are indicated by the single-letter code; conserved amino acids are indicated by +. The aligned sequences and accession numbers from top to bottom are: Y958, *H. influenzae* hypothetical protein HI0958 (P440837); trma HI, *H. influenzae* tRNA (uracil-5-)-methyltransferase (P31812); NucR, *S. cerevisiae* NucR endo-exonuclease (Z25734); trma EC, *E. coli* tRNA (uracil-5-)-methyltransferase (P23003); YAD7, *S. pombe* hypothetical 59.6 kDa protein (Q09833); pCTHom1, *Chlamidia trachomatis* pCTHom1 gene product (M94254); HI0333, *H. influenzae* hypothetical protein HI0333 (C64148). Htf9-c amino acids that are identical or conserved in at least two aligned sequences are underlined.

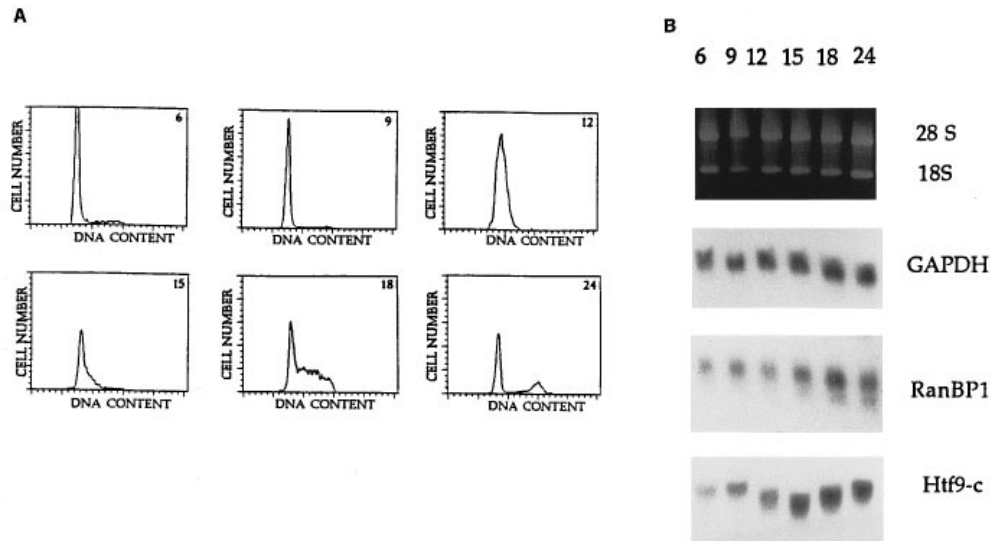


Figure 3 Expression of *RanBP1* and *Htf9-c* transcripts in serum-stimulated cells

(A) FACS analysis of synchronously cycling murine NIH/3T3 cell cultures subjected to serum starvation and re-stimulation. The DNA content, as determined by measuring the propidium iodide fluorescence at the indicated times (h) after re-stimulation, is indicated on the x-axis; the cell number is indicated on the y-axis. (B) Ethidium bromide staining of total RNA extracted from NIH/3T3 synchronized cells at the times indicated above each lane: upper panel, the 18 S and 28 S rRNA bands are stained as a loading control; lower panels, Northern blots were sequentially hybridized with probes corresponding to the coding sequences for GAPDH, *RanBP1* and *Htf9-c*.

targeting signal similar to that carried by the SV40 large T antigen (Figure 2A).

Expression of the *Htf9-a/RanBP1* and *Htf9-c* endogenous genes is activated on entry into the cell division cycle

In a previous study we found that the *Htf9-a/RanBP1* gene is transcribed in a proliferation-dependent manner [9]. To establish whether expression of the divergently transcribed *Htf9-c* gene was also cell-cycle-dependent, we analysed transcription of both endogenous genes in cell cultures that were synchronously progressing through the cell division cycle.

Murine NIH/3T3 cells were first synchronized in G_0/G_1 by serum starvation for 48 h and subsequently re-stimulated to enter the cell division cycle by again raising the serum concentration; cell cycle re-entry was monitored by flow cytometric analysis of the DNA content in the cell population (Figure 3A). Northern blot experiments with the *Htf9-a* probe confirmed our previous observations that *RanBP1* transcription was gradually up-regulated on cell cycle re-entry; the highest levels of *RanBP1* mRNA were observed 15–18 h after stimulation, i.e. when cells reached S phase (Figure 3B). The *Htf9-c* transcript also reached maximal abundance 18 h after re-stimulation, i.e. when most of the cells were in S phase. Up-regulation of both *Htf9* divergent transcripts in S phase was specific, as indicated by comparison with the expression of the unrelated mRNA for GAPDH.

To examine the expression of both *Htf9*-associated genes during S phase progression in more detail, cells were synchronized by using the hydroxyurea block/release protocol; hydroxyurea is a powerful inhibitor of DNA replication, and its addition to the culture medium yields cell populations that are arrested at the G_1/S transition (Figure 4A, time 0). After removal of the drug, S phase progressed more slowly than during an ordinary cell cycle, as shown by the FACS profiles in Figure 4(A); thus this method of synchronization enabled us to resolve accurately the progression through S phase. Synchronization was controlled by

monitoring the expression of cyclin genes: Figure 4(B) shows that both cyclin E and D1 mRNA species peaked at time 0, i.e. in cells arrested at the G_1/S transition, whereas the expression of cyclin A mRNA was first observed 5 h after the block release. The results obtained with probes derived from each *Htf9*-associated gene are shown in Figure 4(C). The *RanBP1* transcript was expressed as early as the G_1/S transition (time 0), and continued to increase throughout S phase (5–9 h after the removal of hydroxyurea). Transcription of the *Htf9-c* mRNA showed a lower level until 7 h after the block release, then increased sharply 7–9 h after the block release. Thus the peak of expression of both genes coincided in time, but the *Htf9-c* mRNA accumulated in a more restricted window of the cell cycle, corresponding to late S phase, than *RanBP1*. Interestingly, the expression of both *Htf9* transcripts was drastically decreased 13 h after the removal of hydroxyurea, i.e. when cells were either completing the cell division cycle or were in early G_1 phase of the next cycle.

Expression of the *Htf9-a/RanBP1* and *Htf9-c* endogenous genes is down-regulated during completion of the cell division cycle

Results obtained from the hydroxyurea synchronization experiments, as well as confirming the indications obtained from serum starved and re-stimulated cells that expression of the *Htf9* genes was up-regulated in S phase, also indicated that the expression of both mRNA transcripts was down-regulated in the late stages of the cell cycle. On the basis of available biochemical and functional data, the disappearance of the *RanBP1* mRNA in mitotic cells might determine the inactivity of the Ran GTPase cycle (reviewed in [7,8]), and might thus indicate a regulative loop in the mechanisms connecting the Ran cycle and cell cycle control. It was therefore important to analyse the expression in mitotic cells in more detail, to ascertain whether transcription of the *RanBP1* mRNA did actually cease, and, if so, whether the *Htf9-c* gene was also transcriptionally inactive.

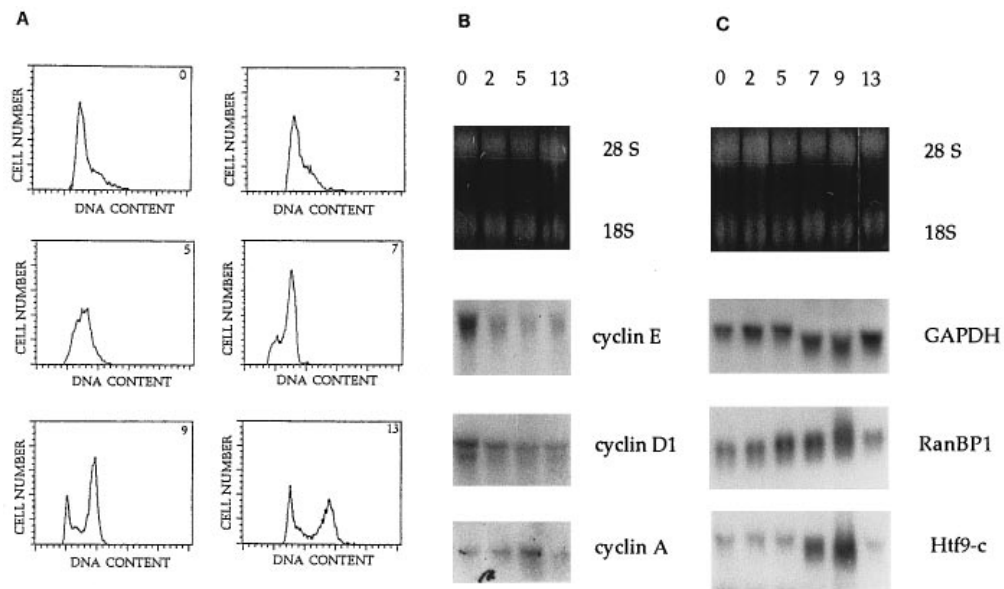


Figure 4 Expression of *RanBP1* and *Htf9-c* transcripts in hydroxyurea-arrested/released cells

(A) FACS analysis of murine NIH/3T3 cell cultures synchronized at the G_1/S boundary by exposure to hydroxyurea (time 0) and released from the hydroxyurea block. The DNA content was determined at the indicated times, as in Figure 3. (B) Ethidium bromide staining of total RNA extracted at the times indicated above each lane: upper panel, the 18 S and 28 S rRNA bands are stained as a loading control; lower panels, the corresponding Northern blot was hybridized with probes corresponding to cyclin E, D1 and A coding sequences. (C) Ethidium bromide staining of total RNA extracted at the times indicated above each lane (top), sequentially probed with GAPDH, *RanBP1* and *Htf9-c* coding sequences (lower panels).

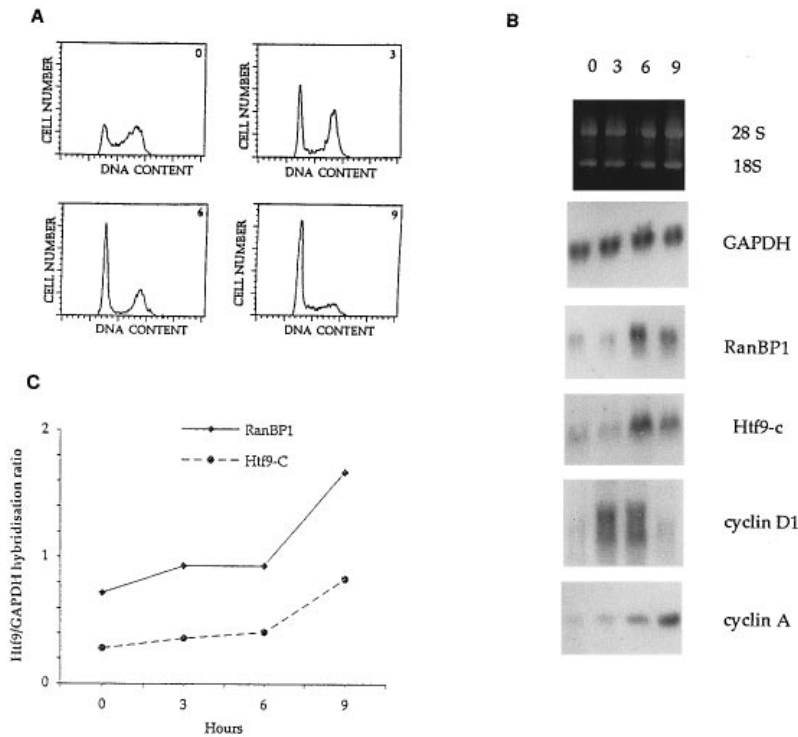


Figure 5 Expression of *RanBP1* and *Htf9-c* transcripts in nocodazole-arrested/released cells

(A) FACS analysis of murine NIH/3T3 fibroblasts arrested in mitosis in the presence of nocodazole (time 0) and G_1 re-entry after nocodazole removal. The DNA content was determined as in Figure 3. (B) Ethidium bromide staining of total RNA extracted at the times indicated above each lane: upper panel, the 18 S and 28 S rRNA bands are stained as a loading control; lower panels, Northern blot hybridization assays with probes corresponding to the GAPDH, *RanBP1*, *Htf9-c*, cyclin D1 and cyclin A coding sequences. (C) Quantitative estimates obtained by processing the filters through an instant imager: radioactivity obtained with each *Htf9* probe was counted; the counts obtained for each time point relative to those of the GAPDH probe are shown in the diagram.

Cells were synchronized in G_2/M by exposure to nocodazole, whose microtubule-depolymerizing activity is well established, for 10 h before harvesting. Although nocodazole synchronization proved to be somewhat less stringent than that obtained by either serum starvation or hydroxyurea addition, the nocodazole method enabled us to follow the progression through and the exit from the mitotic division. Because the addition of nocodazole to growing cell cultures can introduce some distortion in the ordinary cell cycle controls [14], release from the nocodazole block was again monitored by following the transcription of cyclin genes (Figure 5B) coupled to FACS analysis (Figure 5A). The results indicated that the highest proportion of cells exposed to nocodazole were indeed arrested before completion of the mitotic division, as indicated by their G_2/M DNA content (Figure 5A, time 0) and by the absence of cyclin D1 or A expression (Figure 5B); 3 h after block release, the cell population was composed of cells still completing mitosis, as judged by their G_2/M DNA content, and of cells that had re-entered the cycle, which showed a G_1 DNA content and had resumed cyclin D1 mRNA expression. At 6 h later, a proportion of lagging cells still showed a G_2/M DNA content, whereas most had entered a new cell cycle, as shown by the continuing expression of cyclin D1 and the gradual appearance of cyclin A mRNA. After 9 h the cycle was progressing towards S phase, as indicated by the higher abundance of cyclin A than cyclin D1 mRNA. Thus exposure to nocodazole had significantly delayed the exit from mitosis in a proportion of cells but had not irreversibly altered the internal cycle controls in the cell population. Northern blot hybridization experiments (Figure 5B) showed a very low level of expression of both *Htf9*-associated transcripts 0–3 h after release of the mitotic arrest; at these times, the cells were either arrested in mitosis (0 h) or completing mitotic exit and re-entering G_1 (3 h). The low level of both *Htf9*-associated mRNA species in these cultures did not reflect a generally toxic effect of the nocodazole treatment, because expression of the mRNA for GAPDH was not significantly decreased at 0 h or 3 h compared with the level observed 9 h after block release, and thus indicated a specific down-regulation of both *Htf9*-associated transcripts in mitotic and early G_1 cells (Figure 5B). Because autoradiography might not linearly reflect low hybridization signal intensities, the actual representation of target RNA sequences hybridizing with each probe was directly quantified by scanning the filters through an instant imager; the results plotted in Figure 5(C) show the levels of both *Htf9*-associated transcripts relative to that of the *GAPDH* gene, and confirm that both transcripts were barely detectable during mitotic arrest; the amount of hybridization relative to that of *GAPDH* significantly increased 9 h after release of the nocodazole block, when most cells were progressing towards the S phase of a novel division cycle.

Taken together, the results so far indicate that the expression of both *Htf9* divergently transcribed genes is modulated in a cell-cycle-dependent manner; expression of both transcripts is suppressed in non-proliferating cells and is activated on entry in the cell division cycle, although with different kinetics in opposite orientations, and indicates a more gradual activation in the direction of *RanBP1* transcription. Both transcripts are maximally expressed in S phase and are down-regulated in mitosis.

Timing of activation of the *Htf9* promoter during the cell cycle reproduces the regulation of the endogenous mRNA species

Given the close proximity between the 5' ends of the *Htf9*-associated genes (Figure 1), we wondered whether their expression during the cell cycle was controlled at the transcriptional

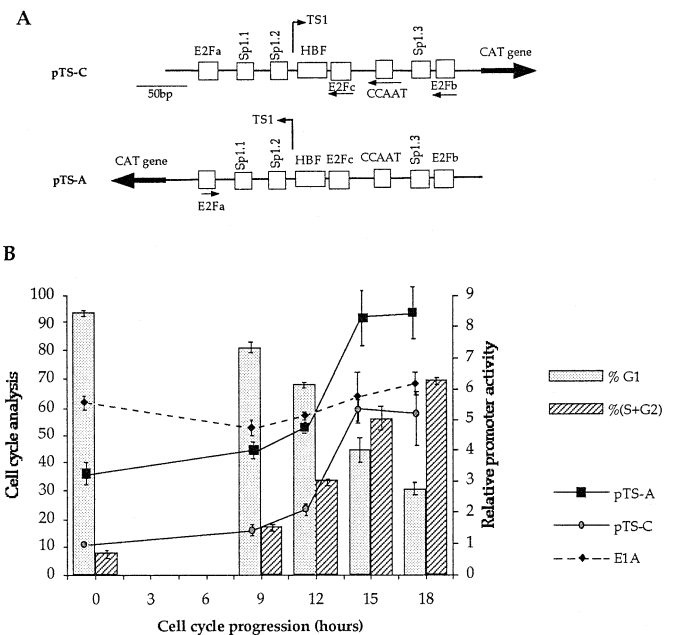


Figure 6 Activity of the *Htf9* bidirectional promoter during the cell cycle

(A) The *Htf9* bidirectional promoter: transcription factor-binding sites identified in previous studies [9,12,15] are shown. Arrows indicate factor-binding sites that are arranged in the reverse orientation relative to the direction of transcription. The HBF box indicates a footprinting activity surrounding the major transcription start site (TS1) of both genes. The orientation of the CAT coding sequence is indicated by heavy arrows. (B) Transient expression assays of the pTS-A and pTS-C promoter constructs in synchronously cycling cells. The shaded columns show the proportion of cells with a G_1 DNA content, whereas the hatched columns show the proportion of cells with an (S+ G_2) DNA content, as measured by FACS analysis; cell percentages are indicated on the left-hand y-axis. Points represent the relative activity of promoter constructs, determined by immunoenzymic assay of synthesized CAT enzyme. Data were normalized relative to the amount of β -galactosidase synthesized from a co-transfected construct in each experiment. Relative promoter activities are indicated on the right-hand y-axis. The lowest promoter activity in each experiment was associated with pTS-C in arrested cells and was taken as 1; mean values were calculated from six (pTS-A and pTS-C) or three (pE1A-CAT) experiments. Bars represent S.E.M..

level. Previous promoter-mapping experiments had identified a 273 bp regulatory region required for the expression of both divergently transcribed genes, thereby defining the *Htf9* bidirectional promoter [12]. The arrangement of transcription-factor-binding sites that were identified in previous DNA-protein interaction studies [9,12,15] is shown schematically in Figure 6(A). The common promoter region, cloned in opposite orientations upstream of a promoterless CAT reporter gene (pTS-A and pTS-C constructs), was assayed in transient expression experiments in serum-starved, i.e. proliferation-arrested, cells, and in cells collected at various intervals after release of the proliferation block. Control experiments were performed with a reporter construct in which CAT expression was driven from the adenovirus *E1A* oncogene promoter (pE1A-CAT construct), whose expression is not cell cycle dependent.

In these experiments the activity of the *Htf9* promoter was found to be down-regulated in both orientations in arrested cells (Figure 6B, time 0). Transcriptional activation occurred substantially between 12 and 15 h of the cycle re-entry, corresponding to the G_1/S boundary as determined by FACS analysis. Up-regulation of the bidirectional *Htf9* promoter was specific and was not simply a consequence of resumed transcriptional activity on exit from quiescence, because the pE1A-CAT construct did not show any significant variation during cell cycle progression.

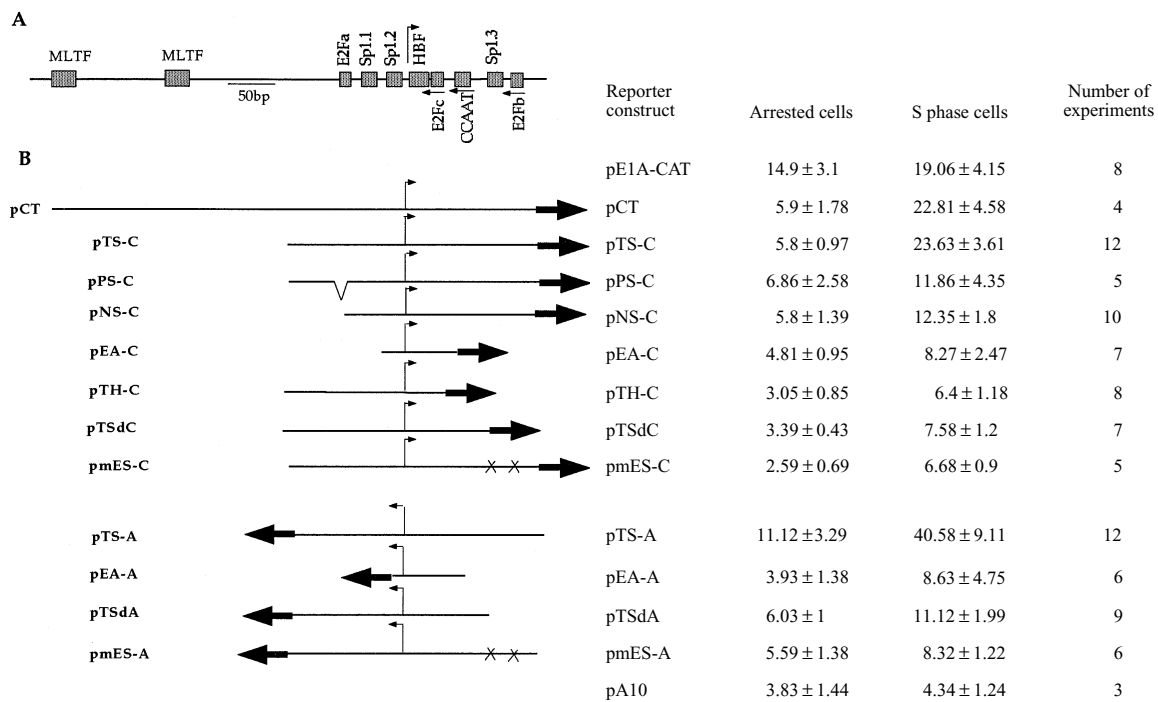


Figure 7 Functional analysis of the *Htf9* promoter in arrested and in S phase NIH/3T3 cells

(A) The region assayed in promoter reporter constructs is shown in the orientation of the *Htf9-c* gene. Boxes indicate identified transcription-factor-binding elements; arrows indicate factor-binding sites that are arranged in a reverse orientation relative to the direction of *Htf9-c* transcription; the major transcription start site is marked by the vertical arrow. MLTF, major late promoter transcription factor; HBF, *Htf-9*-binding factor. (B) The maps show the extent of the assayed region in each reporter construct; heavy arrows indicate the orientation of the CAT coding sequence, and crosses indicate the locations of mutated sites. The activity of each construct was calculated as the mean \pm S.E.M. of synthesized CAT enzyme in pg/ μ g of protein extract after normalization relative to the synthesis of β -galactosidase, as in Figure 6. pA10 and E1A are control constructs in which CAT transcription is controlled by a minimal and a cell-cycle-independent promoter respectively.

The overall efficiency of the pTS-A promoter was higher than that of the pTS-C (opposite orientation) construct at all tested times; the timings of promoter induction at the G₁/S transition coincided for both orientations. Thus the shared promoter region contains differential information on each strand that results in higher activity in the direction of *Htf9-a/RanBP1* gene transcription, yet directs S-phase-dependent activation in both divergent orientations.

Identification of S phase control regions in the *Htf9* bidirectional promoter

Cell cycle activity of the pTS-A promoter construct had been examined in a previous study and had been found to be essentially controlled by E2F activators, whose effect could be antagonized in the presence of repressors of the retinoblastoma gene family [9]. To further identify promoter elements conferring growth control in either orientation, mutant promoter constructs (maps in Figure 7) were transfected in cell cultures that were either growth-arrested by serum withdrawal or collected 15–18 h after serum re-stimulation and were predominantly in S phase. The results are shown in Figure 7(B). Testing of different promoter regions in the direction of *RanBP1* transcription confirmed our previous findings [9] that S phase activity was remarkably efficient in the pTS-A construct, whereas it was essentially abolished in the pTSdA deletion, lacking a 60 bp region containing several potential factor-binding sites. Among those, sites Sp1.3 and E2Fb (see map in Figure 7A) had previously been found to be consistently protected *in vivo* [16], and both might have represented functional targets for the observed repression by

factors of the retinoblastoma group [9]. We therefore decided to inactivate both binding sites by site-directed mutagenesis. The activity of the resulting construct, pmES-A, was similar to that of pTSdA, and both mutant promoters were markedly affected in S-phase-dependent activation compared with the full-length pTS-A construct. Thus virtually all the information required to up-regulate *RanBP1* transcription in S phase is conferred by the neighbouring Sp1.3 and E2Fb sites. Finally, a 74 bp fragment surrounding the transcriptional initiation region (pEA-A construct) was still capable of promoter activity, in comparison with the minimal pA10 promoter, composed of two Sp1-binding sites and a TATA box, indicating that elements downstream of the *Htf9-a/RanBP1* transcription start site did not significantly contribute to basal transcription.

Similar experiments were performed in the orientation of the *Htf9-c* gene transcription. The two longest assayed regions (530 bp in the pCT and 273 bp in the pTS-C constructs) showed similar profiles of induction in S phase, indicating that the upstream pCT promoter region, which contains two footprinted elements harbouring binding sites for MLTF/USF factors [12], did not significantly contribute to cell cycle control. Deleted constructs derived from pTS-C were then assayed. Growth response was partly impaired in the pNS-C construct, lacking a 75 bp upstream region: therefore the region upstream of the deletion end-point contains one or more responsive elements to cell cycle activator(s). A previously identified footprint in the region removed in the pNS-C construct encompassed the sequence 5'-TCTGGCGC-3' [12], which resembles characterized E2F-binding sites (indicated as E2Fa in Figure 7A). To establish whether that footprint identified a regulatory element, we intro-

duced a targeted deletion that removed a 20 bp sequence immediately upstream of the pNS-C end-point and abolished the E2Fa site. The activity of the resulting pPS-C construct was found to be extremely close to that of pNS-C, because both the pPS-C and pNS-C promoters lost nearly 50% of the activity in S phase cells compared with the pTS-C promoter. Thus cell-cycle-dependent activation mediated by the region removed in pNS-C can be ascribed to the E2Fa element deleted in pPS-C.

We then examined the effect of deletions downstream of the major transcription start site. Two progressive deletions in the pTH-C and pTSdC constructs, lacking 100 and 60 bp respectively, yielded a very low level of activity, both in G₀ (50% decrease) and in S phase cells (70% decrease), compared with the pTS-C promoter. Therefore the downstream promoter region contains elements controlling both basal transcription and cell cycle induction of the *Htf9-c* gene. A comparable level of CAT expression was also measured in G₀ and in S phase cells after simultaneous mutagenesis of both the Sp1.3 and E2Fb sites (compare pTH-C, pTS-dC and pmES-C with pTS-C). This finding indicates that the defective activation observed with both deletions could be attributed to the removal of the elements inactivated in pmES-C. From these experiments it can therefore be concluded that the upstream E2Fa site is necessary for transcriptional activation of the *Htf9-c* promoter in S phase, whereas the Sp1.3 and E2Fb elements contribute to efficient expression in either culture condition, although most significantly in S phase. Finally, the pEA-C construct, lacking both the upstream and the downstream regions and retaining 74 bp surrounding the transcription start site, shows a basal level of expression and therefore represents the minimal *Htf9-c* promoter.

DISCUSSION

In the present study we have examined the expression of the divergently transcribed *Htf9-c* and *RanBP1* genes during the cell cycle. We have employed combined synchronization protocols to dissect the expression pattern of both *Htf9* endogenous genes. First, we have confirmed that expression of the *Htf9-a/RanBP1* gene was down-regulated in cells that cease proliferation; expression was up-regulated at the G₁/S transition and peaked in S phase. Maximal expression of the *Htf9-c* mRNA also occurred in S phase, and, unlike the unrelated *GAPDH* transcript, both *Htf9*-associated transcripts were specifically down-regulated in mitotic cells. However, certain differences were apparent in the expression pattern of the divergently transcribed genes, because the window of expression of the *Htf9-c* mRNA was more restricted, and its peak level in S phase sharper, than those of the *Htf9-a/RanBP1* mRNA; this was particularly evident after synchronization with hydroxyurea (Figure 4), which enabled us to resolve differences in the timing of expression during S phase. Thus it would seem that the *Htf9-c* transcript is not necessarily required in large amounts but rather in a defined window of time during continuing replication.

Transient expression assays indicated that 273 bp from the *Htf9* promoter contain sufficient information to direct activation at the G₁/S transition in both orientations in reporter plasmids carrying the same post-transcriptional control signals: therefore cell cycle expression of the entire *Htf9* locus is essentially controlled at the transcriptional level. The efficiency of transcription was different in opposite promoter orientations: expression was more efficient in the orientation corresponding to the *RanBP1*, compared with the *Htf9-c*, promoter at all examined phases of the cell cycle, consistent with results obtained from previous promoter assays, which had estimated the *Htf9-c* promoter strength to be approx. 30% of that of *RanBP1* in

asynchronously proliferating cells [12]. The present results indicate that the pTS-A promoter was not completely silent in arrested cells, and thus induction was essentially achieved by up-regulation of transcription at the G₁/S boundary, whereas the pTS-C promoter was subjected to true transcriptional activation starting from a non-expressed state (Figure 6). Thus although the extent of induction was comparable in both orientations, repression in arrested cells was more effective in the direction of the *Htf9-c* gene transcription. These observations are compatible with the patterns of expression detected by Northern blot experiments, which had indicated a sharp increase in the *Htf9-c* mRNA level during S phase, whereas the *RanBP1* mRNA was already expressed at low levels at the G₁/S transition.

To begin to identify control regions conferring cell-cycle regulation yet specific patterns of expression to the divergently transcribed genes, deletion mapping and mutagenesis experiments were performed in both orientations of the bidirectional promoter. The analysis reported here (Figure 7) identifies at least three regions with different functional properties: (1) a 74 bp region, harbouring site Sp1.2 and the major transcription start site, was sufficient for basal activity in both orientations; the level of CAT transcription was comparable with that promoted by the pA10 minimal promoter and was not significantly activated on cell cycle induction; (2) an upstream region, containing site E2Fa, was important for cell cycle control of *Htf9-c* transcription; and (3) a common region, containing sites Sp1.3 and E2Fb, contributed to transcriptional activation in both promoter orientations, and influenced both basal and S-phase-activated transcription, although the effects were more pronounced on cell cycle induction.

Most S-phase-dependent genes are under the control of E2F activators (reviewed in [17–19]). The gene encoding RanBP1 falls within this group: we have previously shown that expression of the E2F-1 factor counteracted transcriptional repression of the *RanBP1* gene in growth-arrested cells; in contrast, the pTS-A promoter was negatively regulated by expressing the retinoblastoma gene product pRb, or its relative p107, in co-transfection experiments. That control had been found to require the 60 bp region removed in the pTSdA deletion [9]; this formally indicated the existence of an E2F-responsive element in the distal region of the *RanBP1* promoter. We have now found that the pmES-A construct, harbouring mutations for sites E2Fb and Sp1.3, completely reproduces the failure to undergo cell cycle activation observed with the pTSdA deletion. These data, together with the observation that E2Fb is an effective E2F-binding site *in vitro* [9] and *in vivo* [16], implicate site E2Fb, alone or in conjunction with site Sp1.3, as the responsive element to activation by E2F factors. Furthermore the impaired activation of the pmES-C, compared with the pTSC, promoter indicates that the same region also contributes to S phase activity in the opposite orientation.

A footprinted element containing the TCTGGCGC sequence was also identified in the orientation of *Htf9-c* transcription in the region removed in the pNS-C construct [12]. The low level of induction of pPS-C compared with the full-length pTS-C construct indicates that the footprinted element is necessary for efficient activation of *Htf9-c* transcription in S phase cells. Previous binding assays *in vitro* did not enable us to establish the identity of the footprinting activity; however, the observation that site E2Fa is homologous with a functional E2F-responsive site in the human thymidylate synthetase gene promoter [17] strongly suggests that the *Htf9-c* promoter is a novel cellular target of E2F regulation. Five related genes encode distinct E2F activators capable of differential interactions with repressors of the retinoblastoma group and with cyclin-dependent kinases

[17–19]. Site E2Fa upstream of *Htf9-c* and site E2Fb upstream of *RanBP1* are not identical in sequence, and it is possible that the differences observed in the expression pattern of the divergently transcribed genes reflect a different response of the regulatory elements of each gene to specific E2F activators.

The *Htf9* bidirectional promoter also harbours several Sp1 sequences capable of interacting with protein factor(s) *in vitro* [12,15]. At least two Sp1 elements, i.e. Sp1.2, adjacent to the major transcription start site, and Sp1.3, removed in the pTSdA and pTSdC deletions, are clearly protected *in vivo* in proliferating cell cultures [16], which suggests that both sites are functional promoter elements. Sp1 elements are well known to be bidirectionally active [20] and represent versatile functional elements in different cellular backgrounds and promoter contexts. Sp1-binding sites can contribute to basal transcription from minimal promoters by accommodating the interaction of Sp1 with general factors of the preinitiation complex [21,22]. In addition it is becoming increasingly clear that Sp1 elements can confer either negative or positive cell cycle control of gene expression [23–27]. Comparison of pmES-C with pTS-C, and pmES-A with pTS-A, indicates that the integrity of site Sp1.2 is not in itself sufficient to confer high levels of expression in S phase in either orientation. In contrast, site Sp1.3 is included in a region important for efficient transcription especially in S phase. Site Sp1.3 might contribute to cell cycle activation in the *RanBP1* orientation and potentiate the transcriptional induction mediated by site E2Fb, as suggested by reports that have shown biochemical and functional co-operation in cell-cycle-regulated promoters harbouring neighbouring Sp1- and E2F-binding sites [28,29]. That possibility is currently being examined in our laboratory by using promoter constructs carrying individual mutations to assess the role played by each single element and their responses to specific regulators; a full account of this analysis will be reported elsewhere (G. Guarguaglini, C. Pittoggi and P. Lavia, unpublished work). In the opposite orientation, the experiments with the mutated pmES-C construct strongly implicate site Sp1.3 in cell cycle activation of *Htf9-c* transcription, because E2F elements, although present in certain bidirectional promoters such as that shared by the *dhfr/rep-3* genes [26,30], are not bidirectionally active; it is therefore unlikely that site E2Fb, which is implicated in *RanBP1* control, also contributes to *Htf9-c* control. Thus the impaired activation of pmES-C compared with pTS-C most probably reflects a specific role exerted by Sp1.3 in determining high levels of transcriptional induction during the cell cycle. Together these observations suggest that the Sp1.2 and Sp1.3 sites might serve different functions in the bidirectional *Htf9* promoter.

In summary, the bidirectional *Htf9* promoter contains several target elements for factors contributing to both basal and S-phase-activated transcription, some of which are shared by the divergently transcribed genes. The possibility that control elements are bound either by related factors with subtle functional differences, as might be expected of E2F-binding sites, or by versatile factors such as Sp1, capable of a wide variety of interactions with adjacent promoter elements and transcription factors, might contribute to the observed differences in expression of the divergently transcribed genes within the common frame of cell-cycle-regulated transcription.

The arrangement of the regulatory regions of both genes deserves a final comment. Divergent transcription has been reported at a few loci in mammalian genomes; in most instances examined, bidirectional promoters are shared by genes encoding either related products, such as the $\alpha 1$ and $\alpha 2$ chains of collagen type IV [31,32] and the H2A and H2B histones [33,34], or proteins required in a common metabolic or regulatory pathway.

For example, two related genes whose products catalyse distinct steps in purine nucleotide synthesis are divergently transcribed from a bidirectional promoter [35,36]; similarly, a bidirectional promoter regulates the co-ordinated expression of two genes involved in transport of the major histocompatibility class I antigens [37]. In other bidirectional loci the function is not clarified for both transcriptional units identified on each side of the common promoter, yet the evolutionary conservation of the bidirectional organization among species suggests a functional requirement for the tight association between the genes; this is so, for example, in the *surf-1* and *surf-2* genes from the *surfeit* locus, whose products are unknown, which show a highly conserved organization from *Drosophila* to humans [38]. Similarly, the bidirectional promoter of the dihydrofolate reductase (*dhfr*) gene and the *rep-3* gene, whose function is also unknown, is conserved in rodents and humans (reviewed in [39]). These observations have led to the suggestion that the bidirectional organization might provide one mechanism for the co-regulation and co-expression of genes; as such, this arrangement might represent the eukaryotic equivalent of a prokaryotic operon and has been named a 'dioskourion' [35]. The finding that the common promoter shared by the *RanBP1* and *Htf9-c* genes directs S phase expression in both directions suggests that the gene products might be related in the timing of their requirement, notwithstanding the higher level of expression of the *RanBP1* gene discussed above. The Htf9-c protein shares significant similarities in the C-terminal domain with the yeast NucR nuclease [40], which has been implicated in repair and recombination [41], and with prokaryotic proteins with tRNA modification activity. The *E. coli trmA* gene, encoding the best characterized of these proteins, a tRNA methyltransferase, specifies a vital function independent of the tRNA methylation activity [42], and the essential part of the gene falls within the region of homology with the *Htf9-c* gene product. Interestingly, the *E. coli* gene is also expressed in a growth-dependent manner [43]. In yeast, cell cycle control of proteins related to the Htf9-c product was not directly addressed; however, the finding that NucR activity is under the control of the *rad 52* gene [41] also suggests a genetic link between the yeast protein and the cell cycle. On the basis of the sequence features and similarities, the C-terminal region of the Htf9-c gene product might be expected to be active in nucleic acid modification and/or processing, whereas the N-terminal region is rich in potential regulatory signals (Figure 2). Some of the functions controlled by RanBP1, such as the control of S phase and of RNA export, splicing and stability [44], might well overlap with those involving Htf9-c. The common window depicted for maximal expression of both *Htf9* genes in S phase, and the sharing of regulatory elements, suggest that their products might be required in a co-ordinated fashion during the cell cycle.

We thank Pidder Jansen-Durr for the gift of cyclin cDNA subclones; Simona Moretti and Enrico Cundari for flow cytometry; Graziano Bonelli for photography; and Franco Palitti and Enrico Cundari for comments and suggestions on the manuscript. This work was supported in part by the European Community (contract BMH4-CT96-1529), by CNR Progetto Strategico 'Cell Cycle and Apoptosis' and by the Italian Association for Cancer Research (AIRC). G.G. was supported by a grant from Fondazione A. Buzzati-Traverso, A. B. was supported in part by a grant from the MURST Ministry of Public Education, C. P. was supported by grants from Fondazione Cenci-Bolognietti and Fondazione A. Buzzati-Traverso, and G. D. M. was supported by a fellowship from AIRC.

REFERENCES

- 1 Lavia, P., Macleod, D. and Bird, A. (1987) *EMBO J.* **6**, 2773–2779
- 2 Antequera, F. and Bird, A. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1195–1199
- 3 Larsen, F., Gundersen, G., Lopez, R. and Prydz, H. (1992) *Genomics* **13**, 1095–1107

- 4 Bressan, A., Somma, M. P., Lewis, J., Santolamazza, C., Copeland, N., Gilbert, D., Jenkins, N. and Lavia, P. (1991) *Gene* **103**, 201–209
- 5 Coutavas, E., Ren, M., Oppenheim, J., D'Eustachio, P. and Rush, M. (1993) *Nature (London)* **366**, 585–587
- 6 Bischoff, R. F., Krebber, H., Smirnova, E., Dong, W. and Ponstingl, H. (1995) *EMBO J.* **14**, 705–715
- 7 Rush, M. G., Drivas, G. and D'Eustachio, P. (1996) *BioEssays* **18**, 103–112
- 8 Sazer, S. (1996) *Trends Cell Biol.* **6**, 81–85
- 9 Di Matteo, G., Fuschi, P., Zerfass, K., Moretti, S., Ricordy, R., Cenciarelli, C., Tripodi, M., Jansen-Duerr, P. and Lavia, P. (1995) *Cell Growth Diff.* **6**, 1213–1224
- 10 Ren, M., Villamarin, A., Shih, A., Coutavas, E., Shannon Moore, M., Lo Curcio, M., Clarke, V., Oppenheim, J., D'Eustachio, P. and Rush, M. G. (1995) *Mol. Cell. Biol.* **15**, 2117–2124
- 11 Richards, S. A., Lounsbury, K. M. and Macara, I. G. (1995) *J. Biol. Chem.* **270**, 14405–14411
- 12 Somma, M. P., Pisano, C. and Lavia, P. (1991) *Nucleic Acids Res.* **19**, 2817–2824
- 13 Caruso, M., Martelli, F., Giordano, A. and Felsani, A. (1993) *Oncogene* **8**, 267–278
- 14 Johnson, R. T., Downes, C. S. and Meyn, R. E. (1993) in *The Cell Cycle: A Practical Approach* (Fantes, P. and Brooks, R., eds.), pp. 1–24, IRL Press/Oxford University Press, Oxford
- 15 Somma, M. P., Gambino, I. and Lavia, P. (1991) *Nucleic Acids Res.* **19**, 4451–4458
- 16 Stapleton, G., Somma, M. P. and Lavia, P. (1993) *Nucleic Acids Res.* **21**, 2465–2471
- 17 La Thangue, N. (1994) *Trends Biochem. Sci.* **19**, 108–114
- 18 Mueller, R. (1995) *Trends Genet.* **11**, 173–178
- 19 Nevins, J. (1992) *Science* **258**, 424–429
- 20 Kadonaga, T., Jones, K. A. and Tijan, R. (1986) *Trends Biochem. Sci.* **11**, 20–23
- 21 Pugh, B. F. and Tijan, R. (1990) *Cell* **61**, 1187–1197
- 22 Kaufmann, J. and Smale, S. T. (1994) *Genes Dev.* **8**, 821–829
- 23 Robbins, P. D., Horowitz, J. M. and Mulligan, R. C. (1990) *Nature (London)* **346**, 668–671
- 24 Kim, S.-J., Onwuta, U. S., Lee, Y. I., Li, R., Botchan, M. R. and Robbins, P. D. (1992) *Mol. Cell. Biol.* **12**, 2455–2463
- 25 Chen, L. I., Nishinaka, T., Kwan, K., Kitabayashi, I., Yokohama, K., Fu, Y. H., Grunwald, S. and Chiu, R. (1994) *Mol. Cell. Biol.* **14**, 4380–4389
- 26 Shilling, L. J. and Farnham, P. J. (1995) *Cell Growth Diff.* **6**, 541–548
- 27 Datta, P. K., Raychaudhuri, P. and Bagchi, S. (1995) *Mol. Cell. Biol.* **15**, 5444–5452
- 28 Karlseder, J., Rothender, H. and Wintersberger, E. (1996) *Mol. Cell. Biol.* **16**, 1659–1667
- 29 Lin, S.-Y., Black, A. R., Kostic, D., Pajovic, S., Hoover, C. N. and Azizkhan, J. C. (1996) *Mol. Cell. Biol.* **16**, 1668–1675
- 30 Wells, J., Held, P., Illenye, S. and Heintz, N. (1996) *Mol. Cell. Biol.* **16**, 634–647
- 31 Burbelo, P. D., Martin, G. R. and Yamada, Y. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 9679–9682
- 32 Heikkila, P., Soininen, R. and Tryggvason, K. (1993) *J. Biol. Chem.* **268**, 24677–24682
- 33 Sturm, R. A., Dalton, S. and Wells, J. R. (1988) *Nucleic Acids Res.* **16**, 8571–8586
- 34 Dobner, T., Wolf, I., Mai, B. and Lipp, M. (1991) *DNA Seq.* **1**, 409–413
- 35 Gavalas, A., Dixon, J. E., Brayton, K. A. and Zalkin, H. (1993) *Mol. Cell. Biol.* **13**, 4784–4792
- 36 Brayton, K., Chen, Z., Zhou, G., Nagy, P., Gavalas, A., Trent, J., Deaven, L., Dixon, J. and Zalkin, H. (1994) *J. Biol. Chem.* **269**, 5313–5321
- 37 Wright, K. L., White, L. C., Kelly, A., Beck, S., Trowsdale, J. and Ting, J. P. (1995) *J. Exp. Med.* **181**, 1459–1471
- 38 Colombo, P., You, J., Garson, K. and Fried, M. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 6358–6362
- 39 Shilling, L. J. and Farnham, P. J. (1994) *Crit. Rev. Euk. Gene Exp.* **4**, 19–53
- 40 Van Vliet-Reedijk, J. C. and Planta, R. J. (1993) *Yeast* **9**, 1139–1147
- 41 Moore, P. D., Simon, J. R., Wallace, L. J. and Chow, T. Y.-K. (1993) *Curr. Genet.* **23**, 1–8
- 42 Persson, B. C., Gustafsson, C., Berg, D. E. and Bjork, G. R. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 3995–3998
- 43 Gustafsson, C., Lindstrom, P. H., Hagervall, T. G., Esberg, K. B. and Bjork, G. R. (1991) *J. Bacteriol.* **173**, 1757–1764
- 44 Schlenstedt, G., Wong, D. H., Koepp, D. and Silver, P. A. (1995) *EMBO J.* **14**, 5367–5378