*Nucleic Acids Research, 2003, Vol. 31, No. 8* ***2035–2044***
*DOI: 10.1093/nar/gkg320*

# Differential use of multiple replication origins in the ribosomal DNA episome of the protozoan parasite *Entamoeba histolytica*

**Soma Ghosh[1], S. Satish[1], Sonika Tyagi[2], Alok Bhattacharya[1,2] and Sudha Bhattacharya[3,*]**

[1]School of Life Sciences, [2]Bioinformatics Centre and [3]School of Environmental Sciences, Jawaharlal Nehru University, New Delhi-110067, India

Received February 3, 2003; Revised and Accepted February 21, 2003

## ABSTRACT

**The factors that control the initiation of eukaryotic DNA replication from defined origins (*ori*s) on the chromosome remain incompletely resolved. Here we show that the circular rDNA episome of the human pathogen *Entamoeba histolytica* contains multiple potential *ori*s, which are utilized in a differential manner. The primary *ori* in exponentially growing cells was mapped close to the promoter of rRNA genes in the upstream intergenic spacer (IGS) by two-dimensional gel electrophoresis. Replication initiated predominantly from the upstream IGS and terminated in the downstream IGS. However, when serum-starved cells were allowed to resume growth, the early *ori*s which became activated were located in other parts of the molecule. Later the *ori* in the upstream IGS became activated, with concomitant silencing of the early *ori*s. When the upstream IGS was located ectopically in an artificial plasmid, it again lost *ori* activity, while other parts of the rDNA episome could function as *ori*s in this system. Therefore, the activation or silencing of the *ori* in this episome is context dependent, as is also the case with many eukaryotic replicons. This is the first replication origin to be mapped in this primitive protozoan and will provide an opportunity to define the factors involved in differential *ori* activity, and their comparison with metazoans.**

## INTRODUCTION

Studies on initiation of eukaryotic DNA replication point to a high degree of conservation in the overall mechanism, across species (1–3). Homologues of initiator proteins required in this process have been reported in diverse organisms ranging from yeast to mammals, encouraging the belief that a common mechanism exists to ensure once per cell cycle initiation of DNA replication. The DNA sites to which the initiator proteins bind to form the replication complexes serve as replication origins (*ori*s). While the sites are very well defined in terms of DNA sequence in some protozoa, e.g. *Saccharomyces cerevisiae* (4), this is not true in other protozoa such as *Schizosaccharomyces pombe* (5) and in most metazoan chromosomes (6). Thus an element of complexity is introduced in defining the eukaryotic *ori*. It is believed that *ori*s, while generally not sequence specific, are rich in poly(dA/dT). Of all the potential *ori*s that exist on the chromosome, only a fraction are used actively under a given set of growth conditions.

One of the most important unresolved issues is to understand the factors that control differential *ori* usage in eukaryotes. This question has been addressed with developing embryos of *Xenopus* and *Drosophila* (7,8) where a transition from random to specific *ori*s takes place after the mid-blastula stage; and in the rDNA array of *S.cerevisiae* where clusters of active *ori*s are interspersed with large gaps of inactive *ori*s (9). Chromatin organisation and the transcription status of genes surrounding the *ori* are some of the factors implicated in differential activation of an *ori*.

Here we show that multiple *ori*s in the rDNA episome (10,11) of the human pathogen *Entamoeba histolytica* are utilised differentially. Although the primary replication *ori* in this molecule maps near the rRNA gene promoter, other potential *ori*s exist. The latter are transiently activated when serum-starved cells are allowed to resume growth, and are also active when cloned in an ectopic location. This system provides an opportunity to study the factors governing differential *ori* usage in this primitive protozoan, and to compare them with other eukaryotes.

## MATERIALS AND METHODS

### Strains and cell growth

*Entamoeba histolytica* strain HM-1:IMSS clone 6 and transfected lines were maintained in TYI-S-33 medium (12) at 36°C with appropriate antibiotics.

### DNA purification from *E.histolytica* trophozoites

Total DNA was purified from late-log phase-grown cells, as described (13). Briefly, cells were pooled from 40 tubes by chilling in ice water, and a cell pellet was obtained by centrifuging at 275 *g* for 7 min at 4°C. Cells ($5–10 \times 10^7$) were resuspended in 5 ml of buffer (100 mM NaCl, 10 mM EDTA, 10 mM Tris–HCl pH 8.0) and lysed by addition of

*To whom correspondence should be addressed. Tel: +91 11 26704308; Fax: +91 11 26172438; Email: sb@mail.jnu.ac.in

0.25% SDS. DNA was purified by phenol–chloroform extraction followed by ethanol precipitation (14). At every step of isolation, proper precautions (pipetting up and down of resuspended pellet or supernatant was avoided and cut tips were used for transferring supernatant or suspensions) were taken to avoid shearing of DNA and breakage of bubbles. The ethanol pellet was resuspended in 5 ml of buffer (10 mM NaCl, 1 mM EDTA, 10 mM Tris–HCl pH 8.0) and treated with RNase A (100 µg/ml) followed by proteinase K (100 µg/ml). The suspension was again extracted with phenol–chloroform and ethanol precipitated. The DNA pellet was washed with 70% ethanol, air dried and dissolved in TE (10 mM Tris–HCl, 1 mM EDTA, pH 8.0).

### Analysis of replication intermediates by neutral/neutral two-dimensional (2-D) gel electrophoresis

Neutral/neutral 2-D gel electrophoresis was carried out by the method of Brewer and Fangman (15,16). DNA samples (2 µg) were digested with appropriate restriction enzymes (30–40 U in a 40 µl reaction mixture). Electrophoresis was carried out in the first dimension in 0.4% agarose (FMC) at 1 V/cm for 14–16 h. Samples were loaded in duplicate and one of the lanes was stained with ethidium bromide to locate the DNA bands. The unstained lane was cut out and used for second dimension electrophoresis. This was done in 1–1.2% agarose (depending on fragment size) at 5 V/cm for 3–5 h. The time of electrophoresis in each dimension was optimised for best separation of the different fragment sizes. After electrophoresis, the DNA in the gel was depurinated using 0.25 M HCl for 15 min and then blotted onto Nylon membrane (Gene Screen Plus, NEN) following the manufacturer's protocol. Southern hybridisation with [$\alpha$-$^{32}$P]dATP-labelled DNA probes (17) was carried out as described (14).

### Plasmids and transfection

The rDNA fragments of EhR2 were cloned at the unique *Hin*dIII site of pTCV1 (18) (Fig. 7) and the constructs were transfected into *E.histolytica* cells by electroporation (19). The cell pellet was washed twice in phosphate-buffered saline (PBS) (0.11% $KH_2PO_4$, 0.37% $K_2HPO_4$ and 0.95% NaCl pH 7.2) and thereafter in cytomix buffer (10 mM $K_2HPO_4$/$KH_2PO_4$ pH 7.6, 120 mM KCl, 0.15 mM $CaCl_2$, 25 mM HEPES, 2 mM EGTA, 5 mM $MgCl_2$). For each transfection, $1 \times 10^7$ trophozoites and 200 µg of plasmid DNA were used along with 4 mM ATP and 10 mM glutathione. For stable transfection, two consecutive pulses were applied with an exponential discharge of 3000 V/cm (1.2 kV) and a capacitance of 25 µF (Gene Pulser II, Bio-Rad, USA), resulting in a time constant of 0.94 ms. Electroporated trophozoites were transferred into tissue culture tubes containing 12 ml of TYI-S-33, and were incubated in a slanted position at 36°C. Drug (neomycin) selection was started after 48 h of transfection at a lower concentration of 5 µg/ml for the first few subcultures until the transfected lines were stable. Subsequent drug selection was at 10 µg/ml. Stably transfected lines were used for 2-D gel analysis.

### Serum starvation of *E.histolytica* cells

Mid-log phase cells growing in 15% adult bovine serum (Biological Industries, Israel) were collected by centrifugation and grown in medium containing 0.5% serum for 24 h. Normal serum was restored at the end of 24 h and DNA was isolated from cells at different time points for 2-D gel electrophoresis.

### Measurement of cell count and [methyl-$^3$H]thymidine incorporation in serum-starved cells

A total of 2–4 $\times$ $10^5$ cells/ml were transferred to 5 ml glass tubes in a total volume of 3.5 ml of complete medium, and were incubated for the required time at 36°C. Cell number was determined at different time points by counting in a haemocytometer. For radioactive incorporation, 15 µCi of [methyl-$^3$H]thymidine (specific activity 25.0 Ci/mmol; Amersham Life Sciences) was added to each tube. The cells were incubated at 36°C for 30 min, chilled in ice water, harvested by centrifugation and resuspended in 1 ml of ice-cold PBS. DNA was precipitated by the addition of 1 ml of 15% trichloroacetic acid (TCA) and kept on ice for 30 min. The precipitate was collected by filtration through GF/C filters. Free radioactivity was removed by repeated washing with 5% TCA and finally 95% ethanol. Filters were dried at 50°C for 30 min, and the radioactivity incorporated was measured in a Beckman scintillation counter.

### Dot blot analysis

The DNA samples were denatured with 0.25 M NaOH for 10 min and spotted on Nylon membrane, pre-soaked in 0.4 M Tris–HCl pH 7.5, using a manifold apparatus. After loading the samples, the apparatus was kept undisturbed for 30 min at room temperature, after which gentle suction was applied. The blots were air dried and used directly for hybridisation.

### GC skewness analysis

From the EhR2 nucleotide sequence the cumulative skew was calculated (20) as the sum of ([G] – [C])/([G] + [C]), where G and C are counts of the two complementary nucleotides in adjacent non-overlapping windows. Cumulative GC skew (100 bp windows with 3 bp displacement) was plotted.

## RESULTS

### Replication of EhR2 initiates predominantly from the upstream intergenic spacer

The rRNA genes in *E.histolytica* are located on circular episomes (11,21). EhR2, an rDNA circle of *E.histolytica* [Fig. 1; (10)] is 14.1 kb in size. It contains a single rDNA transcription unit (5.9 kb), which encodes the 18S, 5.8S and 28S rRNAs. Several families of short tandem repeats are located in the intergenic spacers (IGSs) upstream and downstream of the rDNA. The promoter of rRNA genes has been mapped 2.6 kb upstream of the mature 18S rRNA (22), between the *Ava*II and *Hin*fI repeats.

Neutral/neutral 2-D gel electrophoresis (Fig. 2; (15)] was used to analyse replication intermediates of EhR2. DNA was digested with the appropriate restriction enzymes (Fig. 3a) to generate fragments (A–L) in the size range of 3–7 kb, which is optimal for analysis by this method. After 2-D gel electrophoresis, Southern blots were hybridised with DNA probes specific for each fragment. As seen in Figure 3b, all fragments from EhR2 contained prominent Y arcs indicative of replication bubbles traversing the entire length of the fragment. However, signals corresponding to replication bubbles were

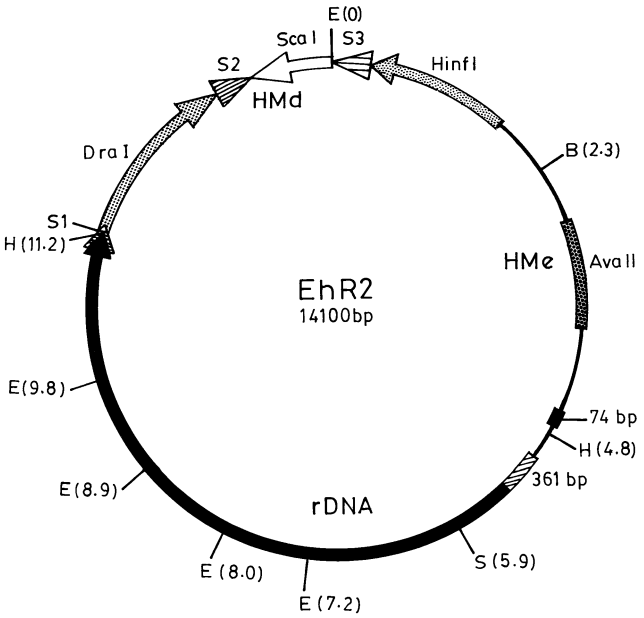


**Figure 1.** Sequence organisation of the *E.histolytica* HM-1:IMSS rDNA plasmid EhR2. The *Eco*RI site at the 12 o'clock position is used as reference for numbering. The bold arrow shows the rDNA unit and its direction of transcription. Nucleotide positions (in kb) are indicated at each restriction enzyme site. B, *Bsa*HI; E, *Eco*RI; H, *Hin*dIII; S, *Sac*II. Different families of short tandem repeats in the regions upstream and downstream of the rDNA are marked as *Hin*fI, *Ava*II, 74 bp, *Dra*I and *Sca*I. Arrows show their relative orientation, where relevant. S1, S2 and S3 are the spacer sequences that lie between the adjacent repeat families. *Eco*RI fragments named HMe (0–7.2 kb) and HMd (9.8–14.1 kb), and *Hin*dIII fragment (4.8–11.2 kb) have been cloned into the plasmid vector pBlueScript.

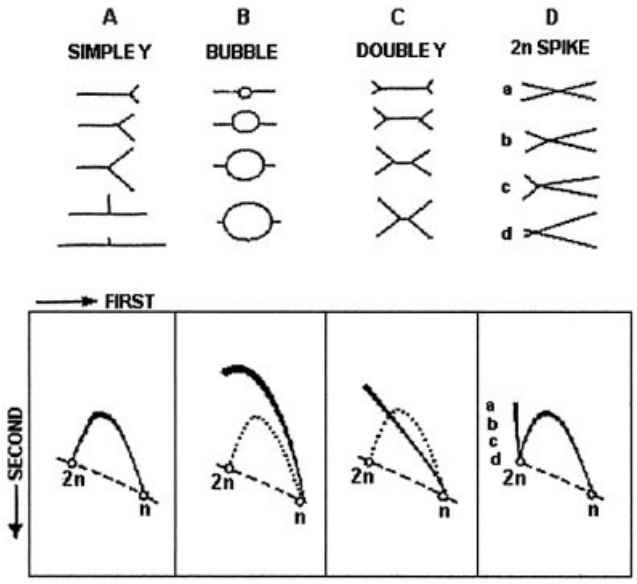


**Figure 2.** Replication intermediates as depicted by 2-D gel electrophoresis. The different types of replication intermediates are illustrated in the top panel. The expected migration of the resultant replication intermediates upon 2-D gel electrophoresis is shown below by bold lines. The dashed lines indicate the migration of linear double-stranded molecules in the 2-D gel. The positions of the unreplicated form (n) and fully replicated form (2n) are shown. The arc of simple Y is shown as dotted line for comparison. The figure is adapted from Friedman and Brewer (16).

seen only in fragments A–D, which encompass the upstream IGS. Fragment E (0–2.9 kb) and fragment F (2.3–5.9 kb) did not show any bubble signal even after prolonged exposure. Since the upstream IGS is a repeat-rich region, there is a paucity of restriction enzyme sites. In this region, the shortest restriction fragment that gave a complete bubble arc was fragment D (1.1–4.5 kb). From these results, it may be concluded that replication initiation takes place in a zone centering around 2.5 kb from the 12 o'clock *Eco*RI site. Replication bubbles arising in this zone would rapidly turn into Ys in fragments E and F and would, therefore, not be detectable. Fragments G–J spanning the rDNA transcription unit failed to show any bubble signal, even after prolonged exposure. Fragments K and L spanning the downstream spacer also gave no bubble signal. Thus, EhR2 is replicated predominantly from bubbles initiating upstream of the rDNA. The observation that fragments A–D contain prominent Y signals, in addition to bubbles, is explained below in the Discussion.

### GC skewness of EhR2

Additional support that the *ori* in EhR2 is located at ~2.5 kb from the 12 o'clock *Eco*RI site in the upstream IGS came from GC skewness analysis of this molecule. Strand compositional asymmetry between the two DNA strands has been observed in various bacterial and viral genomes (23). There is a preference for G over C in the leading strand compared with the lagging strand. A switch in the sign of cumulative GC skew at the 'global minimum' (20) is correlated very well with the experimentally observed location of the replication origin and terminus in many bacterial and viral genomes. The plot of cumulative GC skewness in EhR2 shifts distinctly at ~2.5 kb (Fig. 4) where the most commonly used *ori* of this molecule has been mapped by 2-D gel analysis.

### Replication terminates in the downstream intergenic spacer

Replication forks, which converge in a termination zone, result in double Y molecules giving a typical migration pattern in 2-D gels (Fig. 2). Such a pattern was seen prominently in fragments K and L from the downstream spacer (Fig. 3b). Fragments G–J from the rDNA transcription unit showed no such pattern. Fragments A–D from the upstream spacer did not give a strong termination signal. However, a smear between the Y arc and the 2n signal was sometimes seen, which could correspond to asymmetric termination events. Therefore, replication of EhR2 terminates predominantly in the downstream spacer, although some termination may also occur in the upstream spacer.

### DNA fragments on both sides of the downstream spacer show 'spike' signals indicative of recombined molecules

Joint DNA molecules of size 2n may be generated either by Holliday junctions between molecules undergoing homologous recombination, or by branch migration of daughter strands annealing to each other at stalled replication forks. In our analysis of replication intermediates, we observed

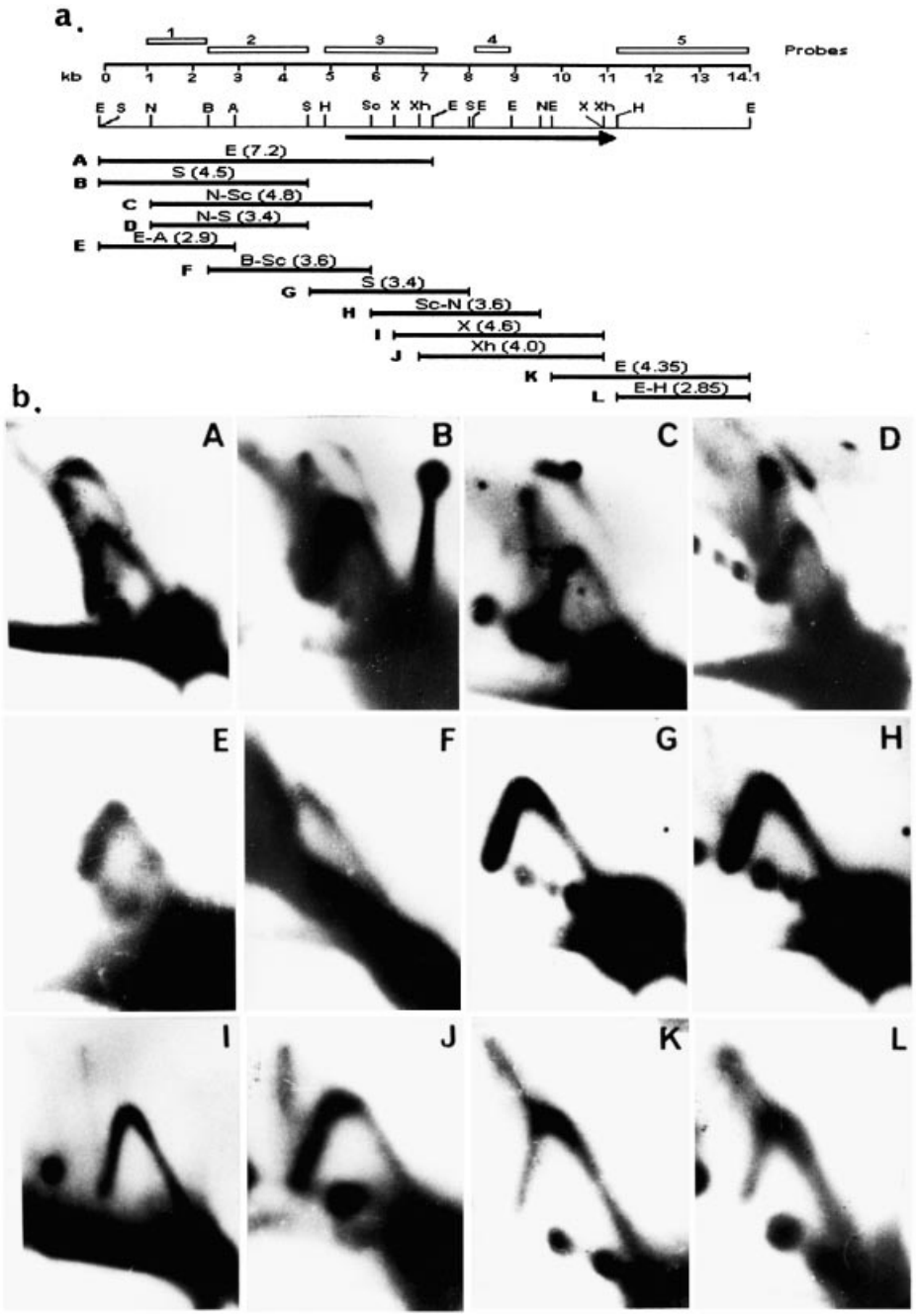


**Figure 3.** 2-D gel electrophoresis of replication intermediates in EhR2. (**a**) A linear map of the 14.1 kb plasmid cut at the 12 o'clock *Eco*RI site (as shown in Fig. 1) is shown. E, *Eco*RI; S, *Sca*I; N, *Nde*I; B, *Bsa*HI; A, *Ava*II; H, *Hin*dIII; Sc, *Sac*II; X, *Xba*I; Xh, *Xho*I. The DNA fragments analysed by 2-D gel electrophoresis are shown as filled bars below the EhR2 map. The restriction enzyme(s) used to generate each fragment and fragment sizes (in kb) are indicated above each bar. The various DNA probes used (probe 1 for fragments A, B, E; probe 2 for fragments C, D, F; probe 3 for fragment G; probe 4 for fragments H, I, J and probe 5 for fragments K, L) are shown as open bars above the linear map. (**b**) Autoradiograms of Southern blots of individual fragments. *E.histolytica* DNA was digested with indicated enzymes to generate fragments of interest. These were separated by neutral/neutral 2-D gel electrophoresis under conditions optimised for fragment size, as described in Materials and Methods.

prominent 'spike' signals arising from the 2n position, which correspond to joint DNA molecules of the same size, in which the junction between the two molecules can occur over the entire length (Fig. 3b). Fragments showing the most prominent 'spike' signals were located in the upstream IGS (Fig. 3b, A–D). The signal was almost absent or very faint in fragments derived from the midpoint of the rDNA circle (G and H). It was again visible in fragments containing the middle and 3′ end of the rDNA unit (I and J), and was completely absent in fragments from the downstream spacer (K and L). The distribution of this signal in EhR2 indicates that joint DNA molecules are formed in regions on both sides of the downstream spacer, which contains the termination zone. By analogy with other systems (24), we conjecture that these

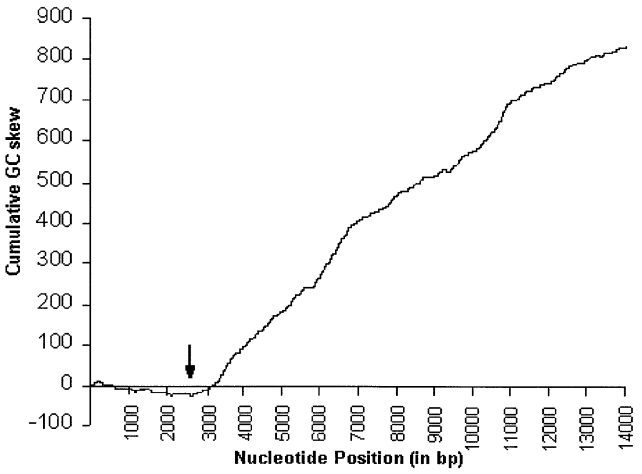


**Figure 4.** Cumulative GC skew (100 bp windows with 3 bp displacement) plotted for the EhR2 episome. The *y*-axis indicates the (G – C)/(G + C) values in sliding windows. The arrow indicates the 'global minimum'.

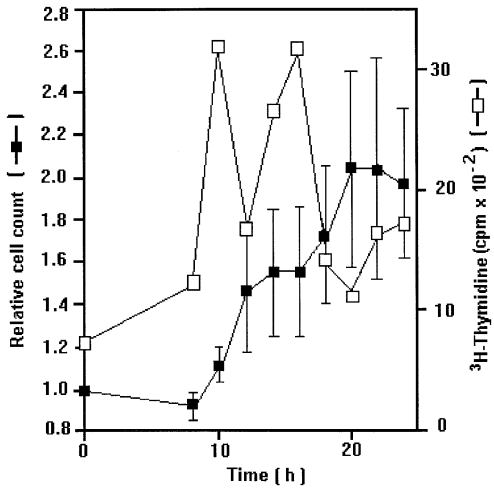


**Figure 5.** Synchronisation of cell growth by serum starvation. Mid-log phase cells growing in 15% serum were collected by centrifugation and grown in 0.5% serum for 24 h. Normal serum (15%) was restored at time zero. At different time points thereafter, samples were removed. Cell number was determined by counting in a haemocytometer, and DNA synthesis was measured by [methyl-$^3$H]thymidine incorporation, as described in Materials and Methods. Duplicate samples were taken for each time point. Cell count is normalised with respect to the value at zero time ($2.5 \times 10^5$ cells/ml). Cell count data are shown as the mean ± SD of four independent experiments.

structures may form as a consequence of recombination events that occur at replication forks stalling before they enter the termination zone.

### Different *ori*s are used transiently in cells recovering from serum starvation

The data presented so far are based on cells growing under normal conditions. We wanted to see whether the same initiation sites are utilised when DNA synthesis is reinitiated after perturbation of growth. Replication was arrested by subjecting the cells to serum starvation. Cells were restored to normal serum after 24 h of starvation. This treatment synchronised the cell cycle for at least one round of cell division (Fig. 5). DNA was purified from 24 h serum-starved cells and from cells restored to normal serum for different time periods, to map the *ori* in EhR2. The 24 h-starved cells did not show any replication intermediates [Fig. 6b(i)]. After restoration to normal serum, cells grown for 30 min showed a prominent bubble signal in the *Sca*I fragment from the rDNA transcription unit, while the *Nde*I–*Sac*II fragment from the upstream IGS did not show any bubble signal [Fig. 6b(ii)]. This fragment did not give a well defined Y signal either. Since at later time points, normal replication intermediates were seen in this fragment, the lack of these signals at the 30 min time point shows unproductive replication through this region in cells recovering from serum starvation. Cells grown for 2 h in normal serum continued to give a bubble signal in the *Sca*I fragment (and also in the *Xba*I fragment), and a faint bubble signal now also appeared in the *Nde*I–*Sac*II fragment [Fig. 6b(iii)]. Cells grown for 4 h in normal serum showed a bubble signal only in the *Nde*I–*Sac*II fragment, while the signal could not be detected in the *Sca*I and *Xba*I fragments even after prolonged exposure [Fig. 6b(iv)]. The same pattern was observed in cells grown for 12 h [Fig. 6b(v)]. These data suggest that while the upstream IGS is the preferred *ori* in exponentially growing cells, other potential *ori*s exist. The latter are probably recruited when the upstream *ori* is not in an active state, but become silenced when the upstream *ori* is activated.

### Effect of cloned rDNA fragments of EhR2 on replication and maintenance of the shuttle vector pTCV1

We were interested to see whether the upstream IGS of EhR2 could function as an *ori* when placed in an ectopic location. Restriction fragments of EhR2 were cloned in the plasmid vector pTCV1. pTCV1 is derived from an *Escherichia coli* plasmid in which the neomycin resistance gene is expressed from *E.histolytica* regulatory sequences (18,25). Since EhR2 shows preferential replication initiation from the upstream IGS (Fig. 3), we decided to check if the same preference is exhibited when the cells are transfected with pTCV1 containing cloned fragments of EhR2. The *E.histolytica* cells were transfected with pTCV1 alone, and with three different constructs: pTCV1-Up (containing the upstream IGS), pTCV1-rDNA (containing the rDNA transcription unit) and pTCV1-Down (containing the downstream IGS). In each case, G418-resistant cells were selected and subsequently grown in the absence of selection pressure for 8 weeks. The transfected DNA remained episomal in these cells, as demonstrated by Southern hybridisation of total DNA from the cells (data not shown). DNA was purified at 2-week intervals and the level of pTCV1 was determined by dot blot hybridisation (Fig. 7). Cell lines with pTCV1 alone showed a progressive decline in plasmid levels, and by 8 weeks retained only ~3% of the plasmid. Cells transfected with pTCV1-rDNA showed the maximum retention of plasmid in the absence of selection pressure (~40% after 8 weeks), followed by cells with pTCV1-Up (~8% after 8 weeks). Cells with pTCV1-Down lost the plasmid very rapidly, indicating a possible negative effect of this fragment. The relative stability of pTCV1-rDNA could mean that the rDNA fragment serves as a more efficient replication origin in pTCV1, or it assists in better segregation

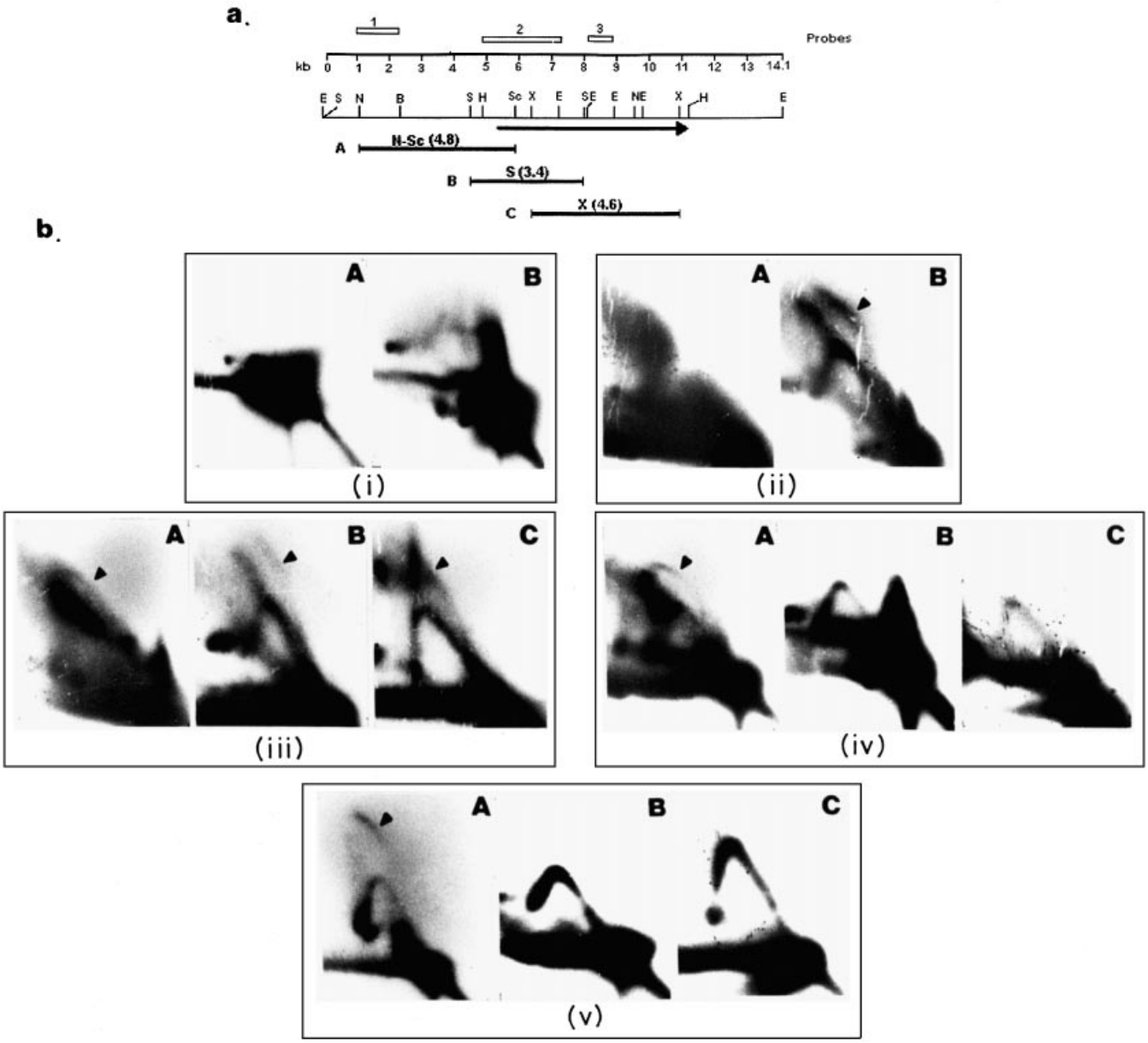


**Figure 6.** 2-D gel electrophoretic analysis of EhR2 after serum starvation and post-serum replenishment. (**a**) A linear map of EhR2 cut at the 12 o'clock *Eco*RI site is shown. E, *Eco*RI; S, *Sca*I; N, *Nde*I; B, *Bsa*HI; H, *Hin*dIII; Sc, *Sac*II; X, *Xba*I. The DNA fragments analysed by 2-D gel electrophoresis are shown as filled bars below the EhR2 map. The restriction enzyme(s) used to generate each fragment and the fragment sizes (in kb) are indicated above each bar. The various DNA probes used (probes 1, 2, 3 used for fragments A, B, C, respectively) are shown as open bars above the linear map. (**b**) Autoradiograms of Southern blots of the indicated fragments separated by 2-D gel electrophoresis as described before. Cells were grown in serum-deprived conditions for 24 h (i) as described in Materials and Methods. Serum was then restored and cells harvested after (ii) 30 min, (iii) 2 h, (iv) 4 h and (v) 12 h of growth. The bubble signals are indicated by an arrowhead.

of the plasmid to progeny cells, or both. The ability of this fragment to serve as a replication origin was tested by 2-D gel electrophoresis.

### 2-D gel electrophoresis of transfectants with pTCV1 constructs

DNA was purified from stable *E.histolytica* transfectants (pTCV1, pTCV1-Up and pTCV1-rDNA) grown with selective pressure, and analysed by 2-D gel electrophoresis. pTCV1 DNA was digested with the appropriate restriction enzymes to yield a fragment containing the bacterial plasmid sequence (Ib, Fig. 8A) and another containing the neomycin reporter gene and flanking amoebic sequence from the 3′ region of the *hgl1* gene (Ia, Fig. 8A). Two fragments each were analysed from pTCV1-Up (Fig. 8B) and pTCV1-rDNA (Fig. 8C). One contained the bacterial plasmid sequence and the other

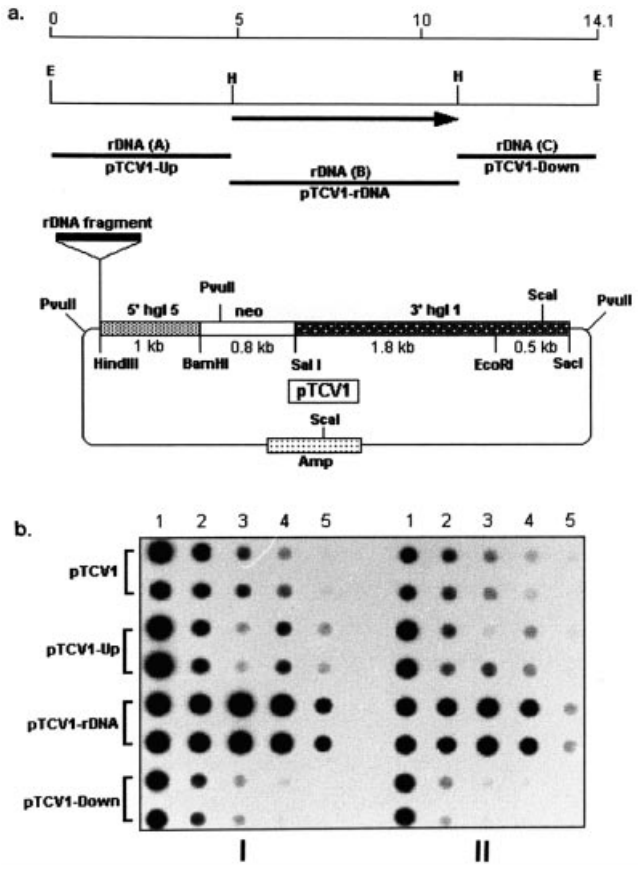


**Figure 7.** (**a**) Restriction map of pTCV1 indicating the unique *Hin*dIII site in which three EhR2 fragments were cloned. The positions of the restriction enzymes *Eco*RI (E) and *Hin*dIII (H) are shown in the linear map of EhR2 above the pTCV1 map. The three EhR2 fragments that were cloned (filled bars) in pTCV1 are indicated. (**b**) Dot blot analysis of DNA samples isolated from *E.histolytica* cells transfected with the pTCV1 constructs shown in (a). Cells were grown in the absence of G418 selection. DNA samples for each construct were loaded in duplicate and hybridised with the neo probe. The intensity of each spot was quantified. DNA was loaded in the following order: column 1, normal transfected cells in the presence of neomycin; columns 2–5, cells grown in the absence of neomycin for: 2, 2 weeks; 3, 4 weeks; 4, 6 weeks; and 5, 8 weeks. DNA loaded on each spot was 100 ng (panel I) and 50 ng (panel II).

contained the cloned rDNA sequence. The enzymes used to release the latter fragments from pTCV1 were *Bam*HI + *Pvu*II (pTCV1-Up) and *Sac*II + *Pvu*II (pTCV1-rDNA). *Bam*HI and *Pvu*II do not have any sites in EhR2, and *Sac*II has only one site in it. Therefore, the indigenous EhR2 would not interfere with the analysis of the cloned rDNA fragments. The results of 2-D gel analysis show that in pTCV1, the fragment containing the bacterial plasmid (Ib) served as an *ori* while the fragment with the neomycin gene (Ia) showed no bubble signal. On the other hand, in addition to a prominent Y, it showed a signal corresponding to doubleY molecules typical of a termination zone (Fig. 8A). In the pTCV1-Up construct, a bubble signal was clearly visible in the fragment containing the bacterial plasmid (IIb), but no such signal was seen in the upstream IGS fragment (IIa, Fig. 8B). Overexposure of the blot also did not show any bubble signal. In the pTCV1-rDNA construct, a prominent bubble signal was seen in the rDNA fragment as well as the fragment with the bacterial plasmid (Fig. 8C). Therefore, the better retention of pTCV1-rDNA by *E.histolytica* cells may be the result of enhanced replication of this molecule.

## DISCUSSION

In the protozoan parasite *E.histolytica*, multiple potential *ori*s exist in the rDNA episome, and a specific sequence may not define the replication origin. This has also been demonstrated for the rDNA episome, EhR1, of *E.histolytica* (26). Here we show that under normal growth conditions, the *ori* in the upstream IGS of the rDNA episome, EhR2, is utilised predominantly. Replication initiates from the upstream IGS and terminates in the downstream IGS. With the limited availability of restriction enzyme sites in this molecule, the origin in the upstream IGS could be mapped to a zone centering at ~2.5 kb from the 12 o'clock *Eco*RI site on the rDNA map (Fig. 3). Since all bubble-containing fragments (Fig. 3b, A–D) had complete bubble arcs, and none showed a partial bubble arc that progresses into a Y arc, we infer that replication initiates from a zone and not from a narrow location. The presence of complete Y arcs in these fragments is expected if initiation events take place from a relatively broad zone. Other factors could also contribute to the Y signals, e.g. breakage of bubbles during DNA purification, and replication of multimeric forms of the rDNA episomes. Replication originating from relatively broad zones, which may, in turn, be composed of closely spaced *ori*s, is a common occurrence in metazoan chromosomes (2,6) but is rare in episomal replicons. That the 'primary' *ori* in EhR2 is located in the upstream IGS appears to be borne out by the GC skewness data (Fig. 4). GC skewness reflects a property of the genome resulting from a replication mechanism operating over evolutionary time (20). For EhR2, it probably means that the *ori* in the upstream IGS is the most commonly used *ori* in this episome under natural conditions in which the *E.histolytica* genome has evolved. Other potential *ori*s are used much less commonly, or are recruited under special conditions.

Replication from *ori*s other than those in the upstream IGS of EhR2 was particularly in evidence when normal serum was restored to serum-starved cells (Fig. 6). At early time points, replication initiations occurred in the rDNA transcription unit, and only later did the *ori* in the upstream IGS become activated. Progressively, this *ori* predominated and, at later time points, few, if any, initiations were seen in other regions. Several factors may contribute to the shift in *ori* usage. Interference of closely spaced *ori*s has been observed in both *S.cerevisiae* (27) and *S.pombe* (28,29). Factors such as chromatin organisation (8), nucleosome positioning (30), DNA methylation (1) and transcriptional status (3) are all known to influence replication initiation. One or more of these may determine *ori* usage in our system. The proximity of the 'primary' *ori* in EhR2 with the transcription start point makes it likely that transcription factors may be involved. This will be the subject of future analysis. Physical association of promoters and replicators is a common occurrence (31–33) and, in a number of studies, transcription factors have been shown to influence the initiation of DNA replication (34–36). A direct link between transcription and *ori* usage has been reported in *Xenopus* oocytes, where random initiations are

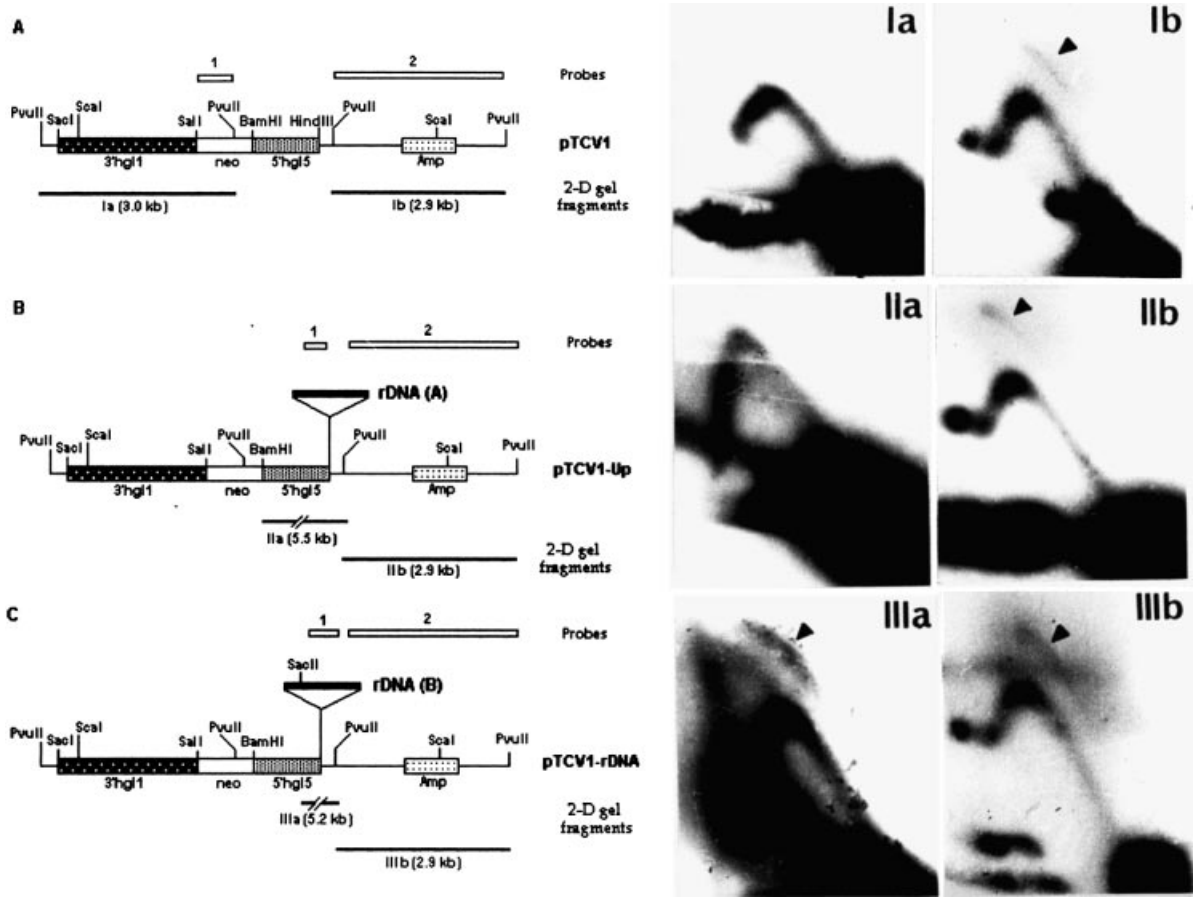


**Figure 8.** 2-D gel analysis of DNA from cells transfected with plasmids pTCV1, pTCV1-Up and pTCV1-rDNA. Linear maps of (**A**) pTCV1, (**B**) pTCV1-Up and (**C**) pTCV1-rDNA. The fragments analysed are shown as thick lines below the map, and the sizes of the fragments are also indicated. The probes used are shown as open bars above the map. DNA was purified from stably transfected cells grown in the presence of G418. Autoradiograms of Southern blots of each fragment, separated by 2-D gel electrophoresis as described before, are shown in adjacent panels (Ia, Ib for pTCV1; IIa, IIb for pTCV1-Up; and IIIa, IIIb for pTCV1-rDNA). The bubble signals are indicated by an arrowhead.

seen in the rDNA locus in early embryogenesis when cells divide rapidly without transcription. At mid-blastula, when the rDNA is transcriptionally active, the *ori* shifts to the IGS (7). Future studies in our laboratory will address the question of a link, if any, between the transcriptional status of the rDNA and choice of *ori* in EhR2.

The upstream IGS failed to function as an *ori* when it was placed in an ectopic location by cloning the fragment into a plasmid vector. In this system, the rDNA transcription unit functioned efficiently as an *ori* (Fig. 8). This was also observed with the rDNA episome, EhR1, of *E.histolytica* (26). Such context dependence of *ori* usage has also been demonstrated in other systems. The *ori* of the β-globin locus is controlled by an upstream region called the locus control region (LCR). Deletion of the LCR represses transcription of the locus and also the activity of the *ori* (37). However, when the β-globin *ori* is placed in an ectopic location, it can function without the LCR (38). Amongst other factors, differences in chromatin organisation of the same sequences in two different contexts may influence their ability to initiate replication.

An interesting observation in this study is the existence of joint DNA molecules where the cross-over point is seen predominantly in two parts of EhR2 (Fig. 3), namely the upstream IGS and the 3′ end of the rRNA-coding region, which flank the termination zone. It appears unlikely that the joint DNA molecules are due to homologous recombination between fully replicated EhR2 circles, since these structures are absent in the termination zone in which frequent recombination occurs between the repeats (21). It is likely that these joint DNA molecules arise due to a replication-mediated recombination event. It is known that replication forks arrested near the termination site, or for other reasons, including single-strand interruptions in the template DNA, or the presence of DNA–protein complexes (24), are processed by homologous recombination enzymes, leading to replication restart (39,40). In EhR2, the prevalence of joint DNA molecules in regions flanking the termination zone could indicate that a recombination-mediated process is required to restore replication from forks that may be arrested near the termination zone.

The data presented herein constitute one of the more detailed analyses of replication initiation in a protozoan parasite. Apart from the maxicircle and minicircle replicons of trypanosomatids (41) and the plastid DNA of *Plasmodium*

*falciparum* (42), the structure of an *ori* in these organisms is unknown. The EhR2 episome of *E.histolytica* provides an opportunity to dissect out the *cis*- and *trans*-factors that are necessary for *ori* function in this primitive eukaryote. The assemblage of proteins (ORCs, Cdc6, MCMs and Cdt1) that form the pre-replicative complex has been characterised extensively in *S.cerevisiae* and *Xenopus* (43,44). Although the six-subunit origin replication complex (45,46) is a conserved feature of chromosomal replication in all eukaryotes studied, the sequence of individual subunits is highly divergent across species (47). We have searched the *E.histolytica* genome database for homologues of these proteins. Homologues of Orc1p, Orc4p, Cdc6p, Mcm2-7p and Cdc45p (43) could be found by BLAST analysis of the database. With the availability of more *E.histolytica* sequences, the genes not found so far may also be detected. Thus, comparative genomics will aid in understanding the conserved and divergent features of the *E.histolytica* replication initiation machinery. The latter could provide useful targets for drug design.

## ACKNOWLEDGEMENTS

We thank Dr Anindya Dutta and Dr Suman K. Dhar for useful suggestions and critical reading of the manuscript. This work was supported by the Department of Science and Technology, Government of India. Financial support is acknowledged from the University Grants Commission for research fellowship to S.G., S.S. and S.T.

## REFERENCES

1. Bogan,J.A., Natale,D.A. and Depamphilis,M.L. (2000) Initiation of eukaryotic DNA replication: conservative or liberal? *J. Cell Physiol.*, **184**, 139–150.
2. Bielinsky,A.K. and Gerbi,S.A. (2001) Where it all starts: eukaryotic origins of DNA replication. *J. Cell Sci.*, **114**, 643–651.
3. Gilbert,D.M. (2001) Making sense of eukaryotic DNA replication origins. *Science*, **294**, 96–100.
4. Marahrens,Y. and Stillman,B. (1994) Replicator dominance in a eukaryotic chromosome. *EMBO J.*, **13**, 3395–3400.
5. Kim,S.M. and Huberman,J.A. (1998) Multiple orientation-dependent, synergistically interacting, similar domains in the ribosomal DNA replication origin of the fission yeast, *Schizosaccharomyces pombe*. *Mol. Cell. Biol.*, **18**, 7294–7303.
6. Depamphilis,M.L. (1999) Replication origins in metazoan chromosomes: fact or fiction? *Bioessays*, **21**, 5–16.
7. Hyrien,O., Maric,C. and Mechali,M. (1995) Transition in specification of embryonic metazoan DNA replication origins. *Science*, **270**, 994–997.
8. Sasaki,T., Sawado,T., Yamaguchi,M. and Shinomiya,T. (1999) Specification of regions of DNA replication initiation during embryogenesis in the 65-kilobase DNApolalpha-dE2F locus of *Drosophila melanogaster*. *Mol. Cell. Biol.*, **19**, 547–555.
9. Pasero,P., Bensimon,A. and Schwob,E. (2002) Single-molecule analysis reveals clustering and epigenetic regulation of replication origins at the yeast rDNA locus. *Genes Dev.*, **16**, 2479–2484.
10. Ghosh,S., Zaki,M., Clark,C.G. and Bhattacharya,S. (2001) Recombinational loss of a ribosomal DNA unit from the circular episome of *Entamoeba histolytica* HM-1:IMSS. *Mol. Biochem. Parasitol.*, **116**, 105–108.
11. Bhattacharya,S., Som,I. and Bhattacharya,A. (1998) The ribosomal DNA plasmids of *Entamoeba*. *Parasitol. Today*., **14**, 181–185.
12. Diamond,L.S., Harlow,D.R. and Cunnick,C. (1978) A new medium for axenic cultivation of *Entamoeba histolytica* and other *Entamoeba*. *Trans. R. Soc. Trop. Med. Hyg.*, **72**, 431–432.
13. Bhattacharya,S., Bhattacharya,A. and Diamond,L.S. (1988) Comparison of repeated DNA from strains of *Entamoeba histolytica* and other *Entamoeba*. *Mol. Biochem. Parasitol.*, **27**, 257–262.
14. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
15. Brewer,B.J. and Fangman,W.L. (1987) The localization of replication origins on ARS plasmids in *S.cerevisiae*. *Cell*, **51**, 463–471.
16. Friedman,K.L. and Brewer,B.J. (1995) Analysis of replication intermediates by two-dimensional agarose gel electrophoresis. *Methods Enzymol.*, **262**, 613–627.
17. Feinberg,A.P. and Vogelstein,B. (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.*, **132**, 6–13.
18. Vines,R.R., Purdy,J.E., Raglans,B.D., Samuelson,J., Mann,B.J. and Petri,W.A.,Jr (1995) Stable episomal transfection of *Entamoeba histolytica*. *Mol. Biochem. Parasitol.*, **71**, 265–267.
19. Hanmann,L., Nickel,R. and Tannich,E. (1995) Transfection and continuous expression of heterologous genes in the protozoan parasite *Entamoeba histolytica*. *Proc. Natl Acad. Sci. USA*, **92**, 8975–8979.
20. Grigoriev,A. (1999) Strand-specific compositional asymmetries in double-stranded DNA viruses. *Virus Res.*, **60**, 1–19.
21. Sehgal,D., Mittal,V., Ramachandran,S., Dhar,S.K., Bhattacharya,A. and Bhattacharya,S. (1994) Nucleotide sequence organization and analysis of the nuclear ribosomal DNA circle of the protozoan parasite *Entamoeba histolytica*. *Mol. Biochem. Parasitol.*, **67**, 205–214.
22. Michel,B., Lizardi,P.M., Alagon,A. and Zurita,M. (1995) Identification and analysis of the start site of ribosomal RNA transcription of *Entamoeba histolytica*. *Mol. Biochem. Parasitol.*, **73**, 19–30.
23. Mrazek,J. and Karlin,S. (1998) Strand compositional asymmetry in bacterial and large viral genomes. *Proc. Natl Acad. Sci. USA*, **95**, 3720–3725.
24. Michel,B. (2000) Replication fork arrest and DNA recombination. *Trends Biochem. Sci.*, **25**, 173–178.
25. Dhar,S.K., Vines,R.R., Bhattacharya,S. and Petri,W.A.,Jr (1998) Ribosomal DNA fragments enhance the stability of transfected DNA in *Entamoeba histolytica*. *J. Eukaryot. Microbiol.*, **45**, 656–660.
26. Dhar,S.K., RoyChoudhury,N., Mittal,V., Bhattacharya,A. and Bhattacharya,S. (1996) Replication initiates at multiple dispersed sites in the ribosomal DNA plasmid of the protozoan parasite *Entamoeba histolytica*. *Mol. Cell. Biol.*, **16**, 2314–2324.
27. Brewer,B.J. and Fangman,W.L. (1993) Initiation at closely spaced replication origins in a yeast chromosome. *Science*, **262**, 1728–1731.
28. Dubey,D.D., Zhu,J., Carlson,D.L., Sharma,K. and Huberman,J.A. (1995) Three ARS elements contribute to the *ura4* replication origin region in the fission yeast *Schizosaccharomyces pombe*. *EMBO J.*, **13**, 3638–3647.
29. Gomez,M. and Antequera,F. (1999) Organization of DNA replication origins in the fission yeast genome. *EMBO J.*, **18**, 5683–5690.
30. Lipford,J.R. and Bell,S.P. (2001) Nucleosomes positioned by ORC facilitate the initiation of DNA replication. *Mol. Cell*, **7**, 21–30.
31. Gallagher,R.C. and Blackburn,E.H. (1998) A promoter region mutation affecting replication of the *Tetrahymena* ribosomal DNA minichromosome. *Mol. Cell. Biol.*, **18**, 3021–3033.
32. Patnaik,P.K., Fang,X. and Cross,G.A.M. (1994) The region encompassing the procyclic acidic repetitive protein (PARP) gene promoter plays a role in plasmid DNA replication in *Trypanosoma brucei*. *Nucleic Acids Res.*, **22**, 4111–4118.
33. Pierron,G., Pallota,D. and Benard,M. (1999) One-kilobase DNA fragment upstream of the *ardC* gene of *Physarum polycephalum* is both a replicator and a promoter. *Mol. Cell. Biol.*, **19**, 3506–3514.
34. Li,R., Yu,D.S., Tanaka,M., Zheng,L., Berger,S.L. and Stillman,B. (1998) Activation of chromosomal DNA replication in *Saccharomyces cerevisiae* by acidic transcriptional activation domains. *Mol. Cell. Biol.*, **18**, 1296–1302.
35. Stucki,M., Stagljar,I., Jonsson,Z.O. and Hubscher,U. (2000) A coordinated interplay: proteins with multiple functions in DNA replication, DNA repair, cell cycle/checkpoint control and transcription. *Prog. Nucleic Acid Res. Mol. Biol.*, **65**, 261–298.
36. Van der Vliet,P.C. (1996) Roles of transcription factors in DNA replication. In DePamphilis,M.L. (ed.), *DNA Replication in Eukaryotic Cells*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 87–118.

37. Aladjem,M.I., Rodewald,L.W., Kolman,J.L. and Wahl,G.M. (1998) Genetic dissection of a mammalian replicator in the human beta-globin locus. *Science*, **281**, 1005–1009.
38. Aladjem,M.I., Groudine,M., Brody,L.L., Dieken,E.S., Fournier,R.E., Wahl,G.M. and Epner,E.M. (1995) Participation of the human beta-globin locus control region in initiation of DNA replication. *Science*, **270**, 815–819.
39. Benard,M., Maric,C. and Pierron,G. (2001) DNA replication-dependent formation of joint DNA molecules in *Physarum polycephalum. Cell*, **7**, 971–980.
40. Zou,H. and Rothstein,R. (1997) Holliday junctions accumulate in replication mutants via a Rec A homolog-independent mechanism. *Cell*, **90**, 87–96.
41. Carpenter,L.R. and Englund,P.T. (1995) Kinetoplast maxicircle DNA replication in *Crithidia fasciculata* and *Trypanosoma brucei. Mol. Cell. Biol.*, **15**, 6794–6803.
42. Williamson,D.H., Preiser,P.R., Moore,P.W., McCready,S., Strath,M. and Wilson,R.J. (2002) The plastid DNA of the malaria parasite *Plasmodium falciparum* is replicated by two mechanisms. *Mol. Microbiol.*, **45**, 533–542.
43. Bell,S.P. and Dutta,A. (2002) DNA replication in eukaryotic cells. *Annu. Rev. Biochem.*, **71**, 333–374.
44. Dutta,A. and Bell,S.P. (1997) Initiation of DNA replication in eukaryotic cells. *Annu. Rev. Cell. Dev. Biol.*, **13**, 293–332.
45. Bell,S.P. and Stillman,B. (1992) ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature*, **357**, 128–134.
46. Romanowski,P., Madine,M.A., Rowles,A., Blow,J.J. and Laskey,R.A. (1996) The *Xenopus* origin recognition complex is essential for DNA replication and MCM binding to chromatin. *Curr. Biol.*, **6**, 1416–1425.
47. Dhar,S.K. and Dutta,A. (2000) Identification and characterization of the human ORC6 homolog. *J. Biol. Chem.*, **275**, 34983–34988.