The EMBO Journal Vol.17 No.8 pp.2426–2435, 1998

Initiation of DNA replication at CpG islands in mammalian chromosomes

Sonia Delgado, María Gómez, Adrian Bird[1] and Francisco Antequera[2]

Instituto de Microbiología Bioquímica, CSIC/Universidad de Salamanca, Edificio Departamental, Campus Miguel de Unamuno, 37007-Salamanca, Spain and [1]Institute of Cell and Molecular Biology, Darwin Building, King's Buildings, Edinburgh EH9 3JR, UK

[2]Corresponding author
e-mail: CpG@gugu.usal.es

S.Delgado and M.Gómez contributed equally to this work

CpG islands are G+C-rich regions ~1 kb long that are free of methylation and contain the promoters of many mammalian genes. Analysis of *in vivo* replication intermediates at three hamster genes and one human gene showed that the CpG island regions, but not their flanks, were present in very short nascent strands, suggesting that they are replication origins (ORIs). CpG island-like fragments were enriched in a population of short nascent strands from human erythroleukaemic cells, suggesting that islands constitute a significant fraction of endogenous ORIs. Correspondingly, bulk CpG islands were found to replicate coordinately early in S phase. Our results imply that CpG islands are initiation sites for both transcription and DNA replication, and may represent genomic footprints of replication initiation.

Keywords: cell cycle/CpG islands/DNA methylation/DNA replication/replication origins

Introduction

The biochemistry of DNA replication is relatively well understood in several viruses, organelles and prokaryotes whose genomes have unique and well-defined origins of replication (Kornberg and Baker, 1992; Kelman and O'Donell, 1994). The larger genomes of eukaryotes require large numbers of ORIs that must be coordinately activated once per cell cycle at the beginning of S phase. The best characterised eukaryotic ORIs are the Autonomously Replicating Sequences (ARS) of the yeast *Saccharomyces cerevisiae* which were identified by their ability to confer autonomous replication on plasmids. A comparable experimental strategy has been only partially successful in mammalian systems, because of the difficulty in isolating short DNA fragments capable of maintaining stable replication of plasmids (Heinzel *et al.*, 1991). A different approach based upon the isolation of DNA fragments that replicate early during S phase, led to the identification of the first mammalian ORI downstream of the dihydrofolate reductase gene (DHFR) in Chinese hamster ovary (CHO) cells (Heintz and Hamlin, 1982). This ORI has been studied in detail over the years using a wide variety of methods (Vassilev *et al.*, 1990; Hamlin and Dijkwel, 1995; Pelizon *et al.*, 1996; Wang *et al.*, 1998).

The link between DNA replication and gene transcription was first recognised in viral systems where mutations in binding sites for transcription factors affected both transcription and replication (DeVilliers *et al.*, 1984). Since then, many transcriptional activators have been reported to activate DNA replication (reviewed in DePamphilis, 1993a; van der Vliet, 1996). In *S.cerevisiae*, the B3 auxiliary element in the ARS1 origin binds the transcription factor ABF1 and mutations that affect binding reduce the efficiency of replication initiation at ARS1, both in plasmids and in the chromosome. The defect can be corrected by substituting the B3 element with binding sites for unrelated transcription factors (Marahrens and Stillman, 1992). These studies have revealed that ORIs contain a core element at which origin recognition proteins bind, plus auxiliary sequences that are necessary for efficient initiation. This modular structure is similar to that of eukaryotic promoters, where the preinitiation complex binds to an essential core element and a number of partially redundant elements are recognised by transcriptional activators.

Mammalian ORIs have not been characterised in such detail and the link between gene transcription and DNA replication relies mainly on the correlation between early replication and transcriptional activity for a large number of genes (Goldman *et al.*, 1984; Holmquist, 1989). This correlation has been further supported by the mapping of some ORIs close to gene promoters. For example, ORIs have been detected at promoters of the human ppv1 gene (Giacca *et al.*, 1994), the Hsp70 gene (Taira *et al.*, 1994), the c-myc gene (Vassilev and Johnson, 1990), the β-globin gene (Kitsberg *et al.*, 1993) and the rat aldolase B gene (Zhao *et al.*, 1994).

CpG islands are associated with the promoters of ~50% of all mammalian genes (Larsen *et al.*, 1992; Antequera and Bird, 1993) and often contain multiple binding sites for transcription factors (Pfeifer *et al.*, 1990; Somma *et al.*, 1991) and an open chromatin conformation (Antequera *et al.*, 1989; Tazi and Bird, 1990). They are regions of ~1 kb that differ from the rest of the genome by being G+C-rich (65%) and free of methylation (Bird *et al.*, 1985; Bird, 1986). By contrast, bulk genomic DNA is comparatively G+C-poor (40% on average) and heavily methylated at CpG. We have asked in this study whether CpG islands could be associated with replication origins in mammalian chromosomes. Predictions of this hypothesis are that CpG islands will be contained in nascent DNA strands smaller than their flanking regions and that they will replicate synchronously at the beginning of S phase. Our results confirm these predictions and indicate that CpG island regions are used to initiate transcription and DNA replication. These results suggest a possible

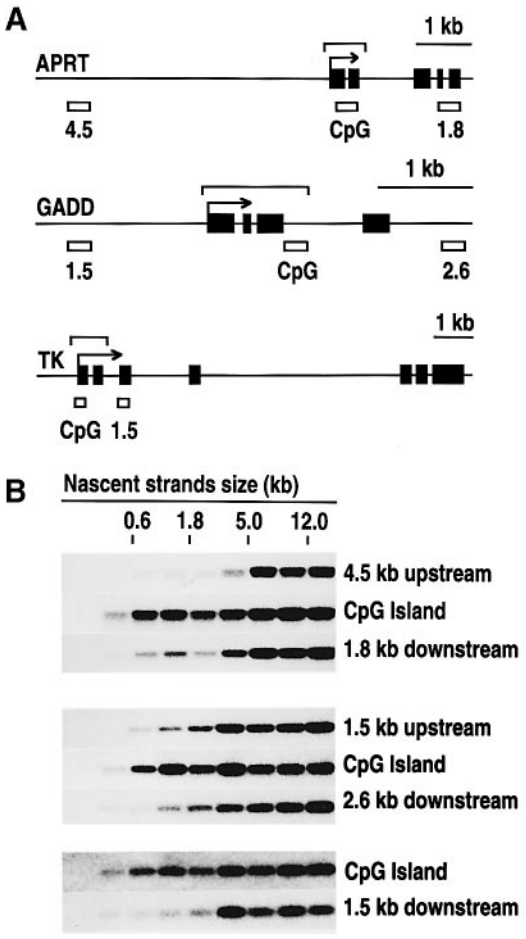

Fig. 1. Replication of CpG islands relative to their flanking regions. (**A**) Genomic maps of the APRT, GADD and TK hamster genes. Black rectangles represent exons and arrows indicate the initiation site and direction of transcription. Brackets covering the 5′ end of the genes show the position of the CpG islands. White bars represent the fragments amplified by PCR. Numbers indicate their distance in kb from the transcription initiation point. Amplified regions within the islands are labelled CpG. The 1 kb scale bar is different in each case. (**B**) BrdU-labelled nascent strands were isolated in alkaline cesium gradients and size-fractionated by agarose gel electrophoresis before being used as input DNA to PCR-amplify the fragments indicated above. Amplified products were blotted onto filters and hybridised to internal primers. The size of nascent strands is indicated. Top, middle and bottom panels show results for the APRT, GADD and TK genes respectively.

origin of the CpG islands as genomic footprints left on the chromosome by the replication initiation event.

Results

Replication order of CpG islands relative to their flanking regions

Figure 1A shows the genomic maps of the hamster Adenine Phosphoribosyl Transferase gene (APRT), the Growth Arrest and DNA Damage inducible gene (GADD) and the Thymidine Kinase gene (TK). The positions of CpG islands (brackets) and the location of sequences that were amplified by PCR are indicated. To test the replication order of CpG islands relative to their adjacent regions, nascent DNA strands derived from these CpG islands and their flanking sequences in CHO cells were examined. We followed the method of Vassilev *et al.* (1990) as modified by Virta-Pearlman *et al.* (1993), which determines the size of nascent strands emanating from a replication origin in an asynchronous population of cells. The closer a particular sequence is to an ORI, the smaller the size of the nascent strands in which it will be contained. Nascent DNA strands from unsynchronised CHO cells were labelled by a 30 min pulse with bromodeoxyuridine (BrdU). Density-labelled single strands were then isolated in alkaline cesium gradients and size-fractionated by agarose gel electrophoresis. DNA was recovered from gel slices and used as input for PCR amplification with the appropriate pairs of primers. PCR products were electrophoresed on gels, blotted onto filters and hybridised to internal primers. Appropriate controls were set up to ascertain that reactions took place under non-saturating conditions and that all primer pairs amplify with comparable efficiencies from total genomic DNA (data not shown).

The results showed that in all three cases, sequences within the CpG islands were amplified from nascent strands of small size that contained little or no sequences derived from the island flanks (Figure 1B). In fact, the island primers detected nascent strands smaller than 1 kb, which locates the origin of replication within the island region. The smallest nascent strands that contained sequences 4.5 kb upstream or 1.8 kb downstream of the APRT CpG island were ~8–9 and 3–4 kb, respectively, which is consistent with replication proceeding bidirectionally from the CpG island. In the case of the GADD gene, there were no significant differences between the sizes of the smallest nascent strands that contained sequences 1.5 kb upstream and 2.6 kb downstream of the island, but neither flank could be amplified from the small fragments that contained the CpG island. In the case of the TK gene, upstream sequences were not tested but the island region was also represented in nascent strands smaller than 1 kb, while those containing the 1.5 kb downstream region were at least 3.0 kb long.

Measurement of nascent strands abundance by competitive PCR

The above results were confirmed by an independent approach using alkaline sucrose gradients to fractionate DNA. Two parallel gradients were run containing identical amounts of total CHO DNA derived from asynchronous exponentially growing cells and from non-cycling cells arrested at the G_1/S phase with mimosine (Mosca *et al.*, 1992). This inhibitor effectively prevents entry into S phase (Figure 6A). Six fractions were collected from the gradients and an equal volume from each fraction was electrophoresed in an agarose gel, blotted and hybridised to total DNA to monitor fractionation. Under these conditions the great majority of DNA pelleted in fraction 6 and therefore only 1/100 of the volume of this fraction relative to fractions 1–5 was used both in the gel and as input for PCR (see below). Figure 2A shows that fractions 1–5 of the gradient derived from cycling cells contained DNA strands that were undetectable in the fractions of mimosine-arrested cells indicating that they are nascent strands derived from DNA replication intermediates. When fractions from the gradient were assayed with the primer pairs within and flanking the APRT CpG island (Figure 1), fraction 1 was found to contain the island region, but not sequences upstream and downstream of the island. The

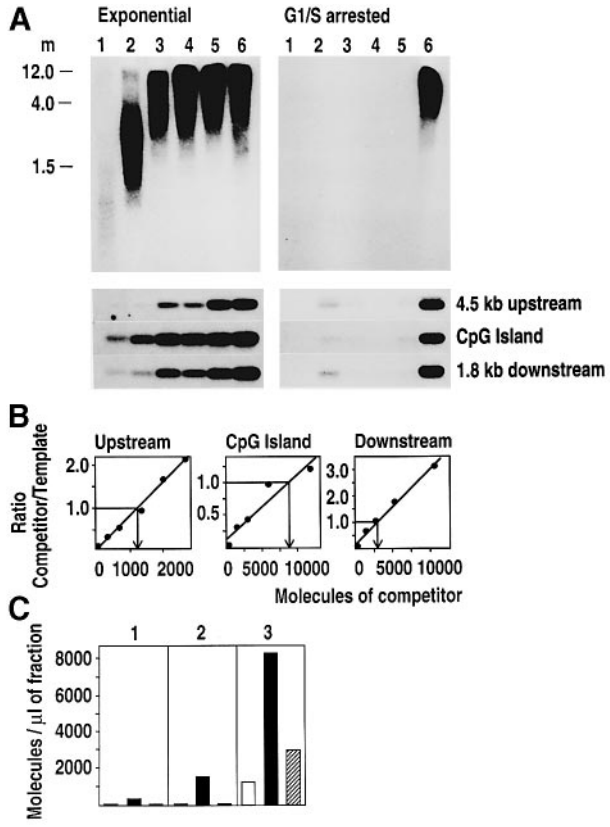

Fig. 2. Measurement of nascent strand abundance by competitive PCR. (**A**) Alkaline sucrose gradient fractionation of total CHO DNA from asynchronous exponentially growing cells and non-cycling cells arrested at the G_1/S boundary. Six fractions were collected and an equal volume from each was electrophoresed in an alkaline agarose gel, except for fraction 6 which contained parental DNA, where only 1/100 of the volume relative to fractions 1–5 was loaded and used as input for PCR. Electrophoresed DNA was blotted and visualised by hybridisation using total CHO DNA as a probe. Panels below show amplified PCR products from each fraction hybridised to internal primers as in Figure 1. (**B**) Quantitative measurement by competitive PCR of the CpG island, upstream and downstream flanking regions in gradient fraction 3. The ratio between the amount of competitor and genomic target is linearly related to the number of competitor molecules. The correlation coefficient of the fitted straight line was 0.991 for the upstream, 0.904 for the CpG island and 0.992 for the downstream regions, respectively. Arrows point to the number of molecules of competitor at which the ratio competitor/template is 1.0. (**C**) Histograms show the number of molecules for the CpG island (black), upstream (white) and downstream (striped) regions in gradient fractions 1, 2 and 3. The number of molecules containing the upstream and downstream regions was below the level of detection in fractions 1 and 2.

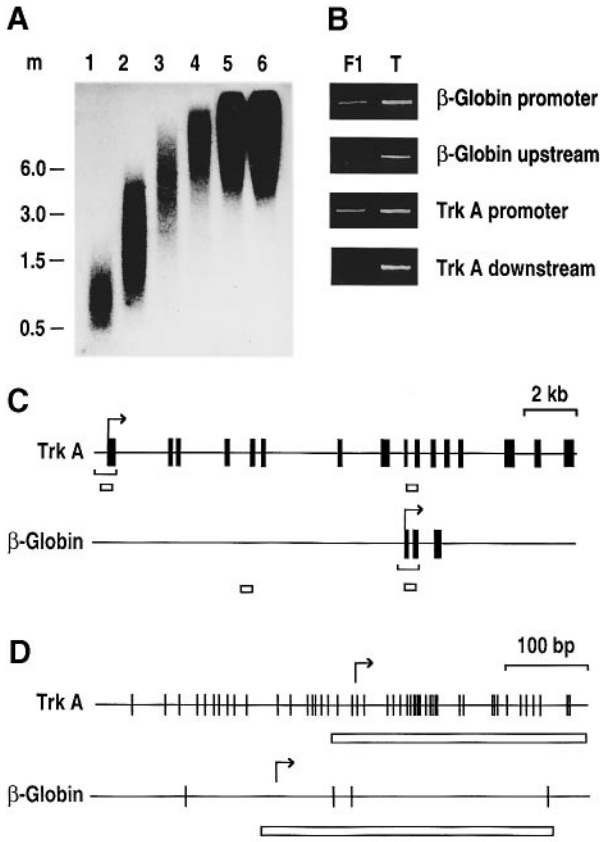

Fig. 3. Small nascent strands contain CpG islands. (**A**) Alkaline sucrose gradient fractionation of total DNA from asynchronous exponentially growing human K562 cells. Six fractions were collected, fractionated in an alkaline agarose gel, blotted and visualised by hybridisation using total K562 DNA as a probe. m, size markers in kb. (**B**) BrdU-labelled nascent strands were immunoprecipitated from gradient fraction 1 (F1) and used as input to PCR the regions of the genes shown below in (C). T, control PCR using total K562 DNA as input. (**C**) Genomic maps of the human Trk A and β-globin genes. Black boxes represent exons and white bars below indicate the amplified regions shown in (B). Arrows indicate initiation of transcription. (**D**) Vertical lines represent CpG dinucleotides across the regions bracketed in (C). White bars represent the amplified promoter regions shown in (C) and arrows indicate initiation of transcription.

downstream primer pair, which is closer to the island, amplified fractions that were somewhat smaller than those amplified by the upstream primers. This agrees with the data in Figure 1, and again suggests that the replication origin is located at the CpG island. No signal was detected in non-cycling cells indicating that PCR products were amplified from replication intermediates.

The abundance of the nascent strands containing the APRT CpG island and flanks was accurately measured by competitive PCR as reported by Diviacco *et al.* (1992) (Figure 2B and C). Competitors were constructed as described in Materials and methods. Arrows in Figure 2B point to the number of molecules of competitor in gradient fraction 3 at which the competitor/template ratio is 1.0. Figure 2C shows the number of molecules for each of the three APRT regions in gradient fractions 1, 2 and 3. The differences in the relative number of molecules agree with the intensities of the PCR bands in Figure 2A.

Short DNA nascent strands are enriched in CpG islands

The finding that three CpG islands selected at random were associated with ORIs raised the possibility that CpG island regions could represent a significant fraction of all chromosomal ORIs. To test this we isolated a population of short DNA nascent strands that should be enriched in ORIs from human K562 erythroleukaemia cells and asked whether fragments with the characteristics of CpG islands were also enriched. Asynchronous exponentially growing cells were pulsed with BrdU for 30 min and total genomic DNA was fractionated in an alkaline sucrose gradient (Figure 3A). Fraction 1 contained fragments up to 1.5 kb long and DNA nascent strands were further purified from fraction 1 by two rounds of immunoprecipitation with a monoclonal anti-BrdU antibody (Vassilev *et al.*, 1990).

To monitor for enrichment in ORIs relative to bulk DNA, we amplified by PCR a fragment at the 5′ end of the β-globin gene (Figure 3C and D). An active ORI has been found to overlap the β-globin promoter in this cell line (Kitsberg *et al.*, 1993) and, therefore, this region would be expected to be present in fraction 1. As a control, we used primers distal to the ORI corresponding to a region 4 kb upstream of the gene (Figure 3C). These primers were unable to amplify from the same fraction, thereby confirming the enrichment for the β-globin ORI. Both pairs of primers amplified the correct fragments with equal efficiency from total genomic DNA (Figure 3B) and from fractions 2–6 of the gradient after immunoprecipitation (data not shown). We also tested whether the CpG island spanning the promoter of the TrkA proto-oncogene (Figure 3C and D), which is actively transcribed in this cell line (Martín-Zanca *et al.*, 1986), was contained among the immunoprecipitated strands from fraction 1. As shown in Figure 3B, primers from the island region amplified the expected fragment, but primers corresponding to a downstream region across exons 9 and 10 did not. Thus, the TrkA CpG island is contained within DNA nascent strands <1.5 kb in this cell line. These results extend and confirm those obtained with the three CHO islands described above.

After verifying the enrichment of ORI sequences in fraction 1, we cloned the nascent strands to assess the representation of CpG islands. Nascent strands were made double-stranded as indicated in Materials and methods, and cloned into an appropriate vector. A total of 35 inserts, ranging in size from 67 to 763 bp (Table I), were selected at random and sequenced. All 35 fragments were hybridised to filters with K562 genomic DNA to rule out possible contamination with adventitious DNA during the cloning procedure. With the exception of three (see below), all inserts detected single copy sequences. To ensure that fragments isolated from the nascent strand library were reproducibly present among short nascent strands, two island clones (numbers 3 and 79) and two non-island clones (numbers 8 and 64) were tested for their presence in the fraction 1 of an independent gradient (data not shown). Both of the islands clones and one of the non-island clones (number 64) were again found to be present in fraction 1, supporting the view that these fragments lie close to ORIs. Clone 8 was not present in fraction 1, but was present in fractions that contained larger fragments. This fragment probably represents contamination by bulk genomic DNA.

Table I shows the G+C content and CpG frequency of the clones. The fragments had a higher G+C content than bulk genomic DNA (which averages 40% G+C). Six clones (clones 3, 28A, 47A, 74A, 77 and 79) also had an observed/expected CpG ratio >0.6 (compared with an average of 0.25 for bulk DNA) and therefore qualify as candidate CpG islands (Bird, 1986; Gardiner-Garden and Frommer, 1987). Another property of CpG islands is their lack of methylation. To assess the methylation status of the clones we hybridised them to filters of K562 genomic DNA digested with *Msp*I and *Hpa*II. Both enzymes recognise the same sequence (CCGG), but *Hpa*II is inhibited when the internal cytosine is methylated whereas *Msp*I is insensitive to methylation. Figure 4A shows examples of clones derived from methylated or unmethylated genomic sequences (clones 64 and 77, respectively). The methylation data (Table I) confirmed the CpG island-like character of the six G+C-rich clones with an observed/expected CpG ratio >0.6. In addition, database searching revealed that clone 74A was derived from the CpG island associated with the human Insulin-like Growth Factor II gene (IGF II). Clone 79 was 100% homologous to the 5′ end of a basic type II keratin mRNA, suggesting that it is derived from its promoter region. Several other members of the keratin gene family have CpG islands at their promoters (Larsen *et al.*, 1992), and it is therefore likely that this non-methylated, CpG-rich sequence is also included in a CpG island. Altogether, six out of 35 clones (17%) isolated from a genomic fraction enriched in DNA nascent strands <1.5 kb appear to be derived from CpG islands. Since CpG islands represent ~1.5% of the human genome (Antequera and Bird, 1993), they are enriched >10-fold in the nascent strand library. It is likely that the observed proportion of clones that are derived from CpG islands represents a minimum estimate (see also below), as any contamination from bulk genomic DNA will usually contribute to the non-island category (for example, clone 8).

Table I. G+C content, CpG frequency and methylation analysis of short nascent DNA strands

Clone	Length (bp)	%G+C	Obs/exp CpG	Methylation[a]
2	211	64.0	0.27	nt
4	180	62.7	0.38	nt
4A	400	44.2	0.31	nt
7A	471	59.9	0.16	nt
8	294	61.2	0.43	methylated
9	404	57.4	0.15	nt
18	254	55.5	0.30	nt
19	292	49.3	0.00	nt
20A	184	58.7	0.19	methylated
22	159	64.8	0.42	methylated
22A	393	64.9	0.36	nt
25	178	66.8	0.15	nt
26	196	64.5	0.16	methylated
29A	292	55.0	0.44	nt
31A	293	59.8	0.16	nt
37A	361	63.1	0.14	methylated
39A	234	59.0	0.15	nt
45A	130	63.0	0.08	nt
64	763	53.4	0.24	methylated
65	335	55.0	0.39	methylated
69	218	56.8	0.22	methylated
73	431	56.1	0.23	methylated
76	473	55.0	0.40	methylated
78[b]	287	62.7	0.18	methylated
80	329	55.0	0.20	nt
85	355	53.8	0.19	nt
3	150	62.0	0.69	non-methylated
28A	67	59.7	0.68	non-methylated
30	201	58.0	0.12	non-methylated
31	229	55.0	0.23	non-methylated
38	190	60.5	0.29	non-methylated
47A	226	65.6	0.81	non-methylated
74A	70	75.0	0.83	non-methylated
77	616	68.5	0.80	non-methylated
79	292	55.8	0.66	non-methylated

[a]Clones without internal *Msp*I/*Hpa*II sites were not tested for methylation (nt).
[b]Clone 78 is partially methylated.

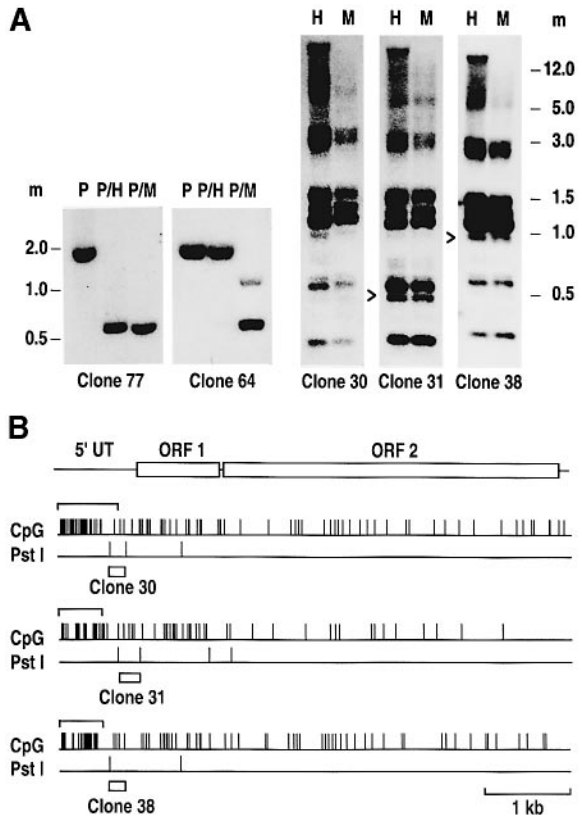

Fig. 4. Methylation analysis of short nascent DNA strands. (**A**) Genomic DNA of K562 cells was digested with *Pst*I (P), *Pst*I + *Hpa*II (P/H), *Pst*I + *Msp*I (P/M), *Hpa*II (H) and *Msp*I (M). Restriction fragments were electrophoresed, blotted and hybridised with clones 77 and 64 and with LINE-1 clones 30, 31 and 38 as probes. Arrowheads point to bands specifically detected with each LINE-1 clone. m, size markers in kb. (**B**) Genomic map of a full length human LINE-1 element. The two open reading frames (ORF1 and ORF 2) and the 5′ untranslated region containing the promoter (5′ UT) are indicated. Diagrams below represent the three LINE elements from which clones 30, 31 and 38 (white rectangles) are derived. Vertical lines represent CpG dinucleotides and *Pst*I sites. Brackets indicate the approximate position of the CpG island.

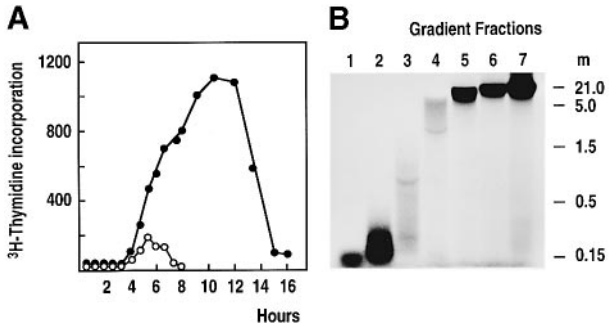

Fig. 5. Replication of bulk DNA and CpG islands during S phase in CHO cells. (**A**) Cells were synchronised by isoleucine starvation. ^{3}H-thymidine incorporation (in thousands of c.p.m.) in total DNA (black circles) and in CpG islands (white circles) was measured at the indicated times. Incorporation values in the CpG island fraction have been scaled up by a factor of 50 to fit within the same diagram. As isoleucine starvation arrests cells in G_1, no ^{3}H-thymidine incorporation was detected during the first 3 h upon addition of the amino acid. (**B**) Representative neutral sucrose gradient used to estimate ^{3}H-thymidine incorporation at CpG islands. The figure shows one of the time samples and illustrates how correct fractionation was monitored by end-labelling a small amount of DNA in each of the seven fractions with [α-^{32}P]dCTP and visualisation by autoradiography after agarose gel electrophoresis. m, size markers (kb).

Short nascent strands derived from the promoter of LINE elements

Inserts of clones 30, 31 and 38 were not obviously CpG island-like, but surprisingly gave similar patterns of multiple bands on Southern blots (Figure 4A; see below). DNA sequence comparisons showed that they were 85–90% identical to one another. Database analysis further showed that clones 30, 31 and 38 were derived from equivalent positions in LINE-1 (Long Interspersed Elements) repeats. Initially, we did not consider that they were derived from CpG islands because their observed/expected CpG frequencies were well below 0.6 (Table I), but when aligned relative to a full-length LINE element their close proximity to the promoter CpG island became evident (Figure 4B). Most restriction fragments generated by *Msp*I and *Hpa*II were identical, confirming that they were derived from non-methylated DNA. In addition to common bands, each probe hybridised to a specific fragment (arrowheads in Figure 4A) which probably corresponds to the genomic source for that clone. The clone-specific bands were also non-methylated. LINE-1 elements are often heavily methylated and functionally inactive due to truncations and nonsense mutations (Hutchinson *et al.*, 1989). Some full-length elements, however, have been reported to have a non-methylated CpG island at their 5′ end (Nur *et al.*, 1988). Our data indicate that clones 30, 31 and 38 are close to origins of DNA replication and appear to be derived from the non-methylated CpG islands of intact elements which are potentially active (Sassaman *et al.*, 1997). Including the three LINE sequences, it follows that nine out of 35 clones (26%) in the nascent strand library are within or immediately adjacent to CpG islands.

Replication of CpG islands during S phase

If many CpG islands are associated with ORIs, they would be expected to replicate coordinately in early S phase. To test this prediction we used CHO cells because they can be easily and tightly synchronised. CHO cells were arrested in G_1 by isoleucine deprivation and then released by addition of the amino acid (Johnson *et al.*, 1993). Samples were labelled at 40 min intervals during the first 8.0 h upon release and synchrony was monitored by ^{3}H-thymidine incorporation (Figure 5A, black circles). To quantitate incorporation at CpG islands relative to total DNA, genomic DNA from each sample was digested to completion with *Hpa*II and the resulting restriction fragments were fractionated in neutral sucrose gradients. A representative gradient profile is shown in Figure 5B. As *Hpa*II is inhibited by methylation of CpG, it generates a population of very small fragments derived mainly from CpG islands (Figure 5B, fractions 1 and 2) and another population of very large fragments comprising the highly methylated remainder of the genome (Figure 5B, fractions 5–7). We have previously shown that fragments in fraction 1 are derived almost exclusively from CpG islands (Antequera and Bird, 1993). The proportion of total genomic radioactivity that was in fraction 1 allowed measurement of the proportion of all CpG islands that were replicated at each time interval. The result indicates that CpG island

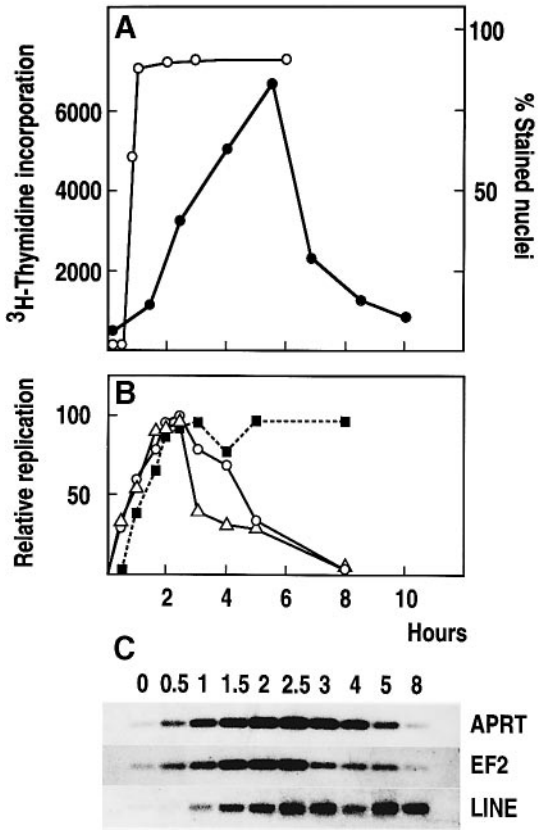

Fig. 6. Replication of individual CpG islands during S phase. (**A**) CHO cells were synchronised with mimosine and progress through S phase was followed by ^{3}H-thymidine incorporation (black circles). Synchronous entry into S phase was monitored by BrdU incorporation and immunofluorescence of nuclei with anti-BrdU antibodies (white circles). (**B**) Cells were pulsed with BrdU for 30 min prior to the indicated times and immunoprecipitated nascent strands were used as input for PCR. The replication profile of the APRT (circles), EF2 (triangles) CpG island regions and a 3′ region of a LINE repetitive sequence (squares) corresponds to the measurement of radioactivity in the autoradiographs shown below. (**C**) APRT, EF2 and LINE PCR-amplified fragments were hybridised with internal primers as a control for specificity prior to quantitation.

replication is essentially complete within the first third of the S phase, reaching a maximum at ~2.0 h into S phase (Figure 5A, white circles).

This result was confirmed by monitoring the replication profile of two individual hamster CpG islands, one associated with the APRT gene and another with the Elongation Factor 2 (EF2) gene. For this experiment CHO cells were synchronised with mimosine which arrests close to the G_1/S phase boundary (Figure 6A, black circles) (Mosca *et al.*, 1992). Cells were labelled for 30 min with BrdU at various times after release from the mimosine block and nascent DNA strands were isolated by two rounds of immunoprecipitation (Vassilev *et al.*, 1990). The resulting DNA was used as input for PCR amplification of regions of the APRT and EF2 CpG islands (Figure 6C). Quantitative results in Figure 6B show that the two islands were amplified preferentially from samples that were pulse-labelled early after release of the mimosine block, and gave a peak of replication at 2.5 h after the onset of S phase. Replication profiles closely mimic the pattern of the bulk CpG island population shown in Figure 5A. As a control for equal recovery of immunoprecipitated nascent strands, from the same samples we amplified a fragment derived from the 3′ region of a hamster LINE element. Because of truncation, the 3′ end is ~10 times more highly repeated in the genome than the 5′ end (Hutchison *et al.*, 1989). Replication kinetics of this LINE region were clearly different from those of the two CpG islands, as nascent strands remained abundant during the second half of S phase in agreement with previous observations (Taljanidisz *et al.*, 1989). Synchrony was monitored at 15 min intervals upon release from the mimosine block by incorporation of BrdU as detected by nuclei staining with fluorescent anti-BrdU antibodies (Leonhardt *et al.*, 1992). No staining was detected earlier than 45 min after release and 90% of nuclei showed positive staining in the time window between 45 and 60 min. The remaining 10% of cells probably never entered S phase, as 6 h later the fraction of stained nuclei was still 90% (Figure 6A, white circles). This indicates that the wide window of CpG island replication cannot be attributed to an asynchronous entry into S phase but probably reflects an intrinsic scattering in the activation of ORIs at the beginning of S phase.

Discussion

CpG islands as replication origins

The evidence presented in this paper argues that DNA replication often initiates at CpG island regions *in vivo*. This is deduced from the observation that each of four CpG island sequences selected at random is present in a population of very short nascent DNA strands, whereas flanking sequences are absent from the same population. The length of the nascent fragments that contain island sequences are comparable with the lengths of the CpG islands themselves. Although the ORIs cannot be located precisely because of limits to the resolution of the method, the data indicate that they all map within or immediately adjacent to the CpG island. Their association with ORIs is further argued by the enrichment of CpG island-like sequences in a library of short nascent DNA strands. Taken together, our data suggest that replication origins are present in many CpG islands. The finding that they replicate synchronously at the beginning of S phase is compatible with this role and provides an immediate explanation for the early replication of housekeeping genes, all of which are associated with CpG islands (Larsen *et al.*, 1992). The results can also account for the early replication of metaphase G-light bands (Holmquist, 1989), as CpG islands have been found to be concentrated there (Craig and Bickmore, 1994).

The genomic location of a number of mammalian ORIs has been previously reported and several are located at CpG islands. Analysis of CpG frequency reveals that ORIs located at the 5′ region of the human c-Myc gene (Vassilev and Johnson, 1990), at the promoter of the human hsp70 gene (Taira *et al.*, 1994) and at the promoter of the human ppv1 gene (Giacca *et al.*, 1994) are close to or within CpG islands. In contrast, other mammalian ORIs clearly belong to a non-island category as judged by their deficiency of CpGs. This category includes the ORIs at the promoter regions of the human β-globin gene (Kitsberg *et al.*, 1993; Figure 3), at the rat aldolase B gene (Zhao *et al.*, 1994), at the hamster rhodopsin locus (Gale *et al.*,

1992), at the mouse adenosine deaminase region (Virta-Pearlman *et al.*, 1993) and downstream of the hamster DHFR gene (Dijkwel and Hamlin, 1995). The latter ORI is the best studied in mammalian systems and long range analysis has shown that initiation activity at this locus originates from a 55 kb region downstream of the DHFR gene with two preferred initiation regions (DePamphilis, 1993b; Pelizon *et al.*, 1996; Wang *et al.*, 1998). A cluster of methylated CpGs has been reported close to one of the DHFR regions of preferred initiation, although its role in ORI activation has not yet been defined (Rein *et al.*, 1997). The latter region does not qualify as a CpG island by the criteria of G+C-richness, lack of methylation and promoter association and should therefore be distinguished from the regions under study here.

Replication analysis by polarity switching across 60 kb of the hamster APRT locus failed to detect ORI activity at the CpG island region (Handeli *et al.*, 1990). A possible explanation could be that the method used required long periods of protein synthesis inhibition that presumably prevented new initiations and could therefore discriminate against short nascent strands. Our approach is better suited for detection of ORIs within regions a few kb long and suggests that both long range and more detailed local analysis are necessary to identify all the ORIs across a genomic region.

If most CpG islands act as replication origins, it is possible to trace the kinetics of activation of a subset of chromosomal ORIs by tracing the profile of total CpG island replication. The combination of synchronous entry of the cell population into S phase with probabilistic activation of ORIs would result in the same replication profile of total and individual CpG islands as shown in Figures 5 and 6. An extended period of firing could apply also to non-island origins as exemplified by the DHFR origin which is active during at least the first 3 h of S phase in synchronised CHO cells (Dijkwel and Hamlin, 1995). Stochastic activation of origins may be more protracted in the large mammalian nucleus than in its more compact yeast counterpart. This could be one reason why S phase takes so much longer in mammalian cells than in yeast, despite the similar density of replication origins along their chromosomes (Huberman and Riggs, 1968; Newlon, 1988).

DNA replication and gene transcription

The co-localisation of CpG islands and ORIs suggests that CpG islands are sites of initiation for both transcription and DNA replication. The activity of CpG island promoters depends on their lack of methylation and there is evidence that the repressive effect exerted by methylation upon transcription is paralleled by a similar effect on replication. An example is the CpG island associated with the human X-linked hypoxanthine phosphoribosyltransferase gene (HPRT) which is unmethylated in the active allele and methylated in the transcriptionally inactive homologue (Yen *et al.*, 1984). The two alleles replicate early and late respectively during S phase and DNA demethylation with 5-azacytidine reactivates transcription of the inactive allele (Venolia *et al.*, 1982) and shifts its replication time into early S phase at approximately the same time as the allele on the active X chromosome (Schmidt and Migeon, 1990).

Several studies of CpG island promoters have shown that they are bound *in vivo* by many transcription factors, some of which are functionally redundant for transcription (Pfeifer *et al.*, 1990; Somma *et al.*, 1991; Tommasi and Pfeifer, 1995). It may be that some are dedicated to either transcription or replication, whereas others might play a role in both processes. *In vivo* footprints across CpG islands at different phases of the cell cycle may reveal the presence of factors involved in the activation of DNA replication, as has been shown for the origin binding complex (ORC) bound to the ARS1 ORI in *S.cerevisiae* (Diffley *et al.*, 1994). The footprint of a candidate for a mammalian ORC has recently been identified at the replication origin located adjacent to the CpG island promoter region of the human ppv-1 gene (Dimitrova *et al.*, 1996). Alterations in the binding of transcription factors *in vivo* at the G_1/S phase transition have also been reported at the CpG island spanning the promoter of the human cdc2 gene (Tommasi and Pfeifer, 1995).

Maintenance of the unmethylated state of CpG islands

Our results suggest a possible mechanism whereby CpG islands could be spared from methylation within an otherwise heavily methylated genome. The DNA methyltransferase (MTase) has a diffuse nuclear localisation in non-S phase cells, but is targeted to the replication foci during S phase as a component of the replication machinery (Leonhardt *et al.*, 1992). It is conceivable that the MTase is not assembled onto the replication preinitiation complex or that steric hindrance prevents its access to its DNA substrate at the initiation stage. Recently it has been suggested that human MTase binds to PCNA and that its binding might be negatively regulated by the cell cycle regulator $p21^{WAF1}$ (Chuang *et al.*, 1997). High levels of $p21^{WAF1}$ could prevent recruitment of the MTase to the replication machinery at the beginning of S phase leading to the effective replication of CpG islands in the absence of MTase. By consistently denying access to MTase during their replication, CpG islands would avoid its *de novo* methylation activity (Bestor and Verdine, 1994) and would be assured of a methylation-free status.

The picture that emerges from this model is that CpG islands could be genomic footprints due to the replication initiation event. According to this view, initiation of replication would itself be a consequence of the promoter activity at CpG islands, in agreement with the observation that binding of transcription factors is essential for establishment of a CpG island (Brandeis *et al.*, 1994; Macleod *et al.*, 1994). Lack of methylation would keep promoters free of methyl-CpG-binding proteins (Meehan *et al.*, 1989; Lewis *et al.*, 1992) and this would in turn maintain their accessibility to the transcription and replication machinery. The result would be a stable cycle in which transcription factors would contribute to recruitment of the replication machinery, leading to methylation-avoidance, early replication and consequent persistence of gene activity. Conversely, a stable cycle of transcriptional inactivity, methylation and passive replication can be constructed. It has previously been proposed that DNA methylation states and DNA replication timing can form alternative cycles (Riggs and Pfeifer, 1992). In the present scheme, replication initiation and transcription would reinforce one another to provide a potentially more robust cycle. It is

possible that perturbation of the active cycle could lie behind the unscheduled *de novo* methylation of CpG islands found in many cell lines (Antequera *et al.*, 1990) and some human tumours (Laird and Jaenisch, 1996).

Materials and methods

Cesium chloride gradients and nascent strand fractionation

Non-synchronised and exponentially growing CHO cells were labelled for 20 min with 50 μM BrdU and 5 μCi of ^{3}H-thymidine per ml of culture (specific activity 5 Ci/mmol). Total DNA was isolated, denatured in NaOH 0.4 N and fractionated in alkaline cesium gradients as described by Epner *et al.* (1981). Fractions containing the heavy peak were pooled and dialysed against TE (10 mM Tris–HCl, pH 8.0/1 mM EDTA) with 0.05% sodium pyrophosphate, ethanol-precipitated and resuspended in TE. Single-stranded DNA nascent strands were electrophoresed in 1% alkaline low melting agarose with size markers. The gel was cut into slices and DNA was recovered from each to be used as input for PCR.

CHO primers and PCR conditions

Positions of primer pairs are given according to genomic sequences in the DDBJ/EMBL/GenBank database. APRT (accession No. X03603) CpG island primers: 401–420 and 661–681, internal primer 604–623; 1.8 kb downstream primers: 2010–2029 and 2290–2309, internal primer 2231–2250; for the 4.5 kb upstream region we amplified a region of 301 bp with the primers 5′-AGAGAGGGTCATGAGTCACC and 5′-CAGGGACCCAAGGACTTCTG. The primer 5′-ACAGCCCACTCTG-ACTTCAG was used as an internal hybridization probe. GADD (accession No. L25339) CpG island primers: 2449–2468 and 2651–2670, internal 2491–2510; 1.5 kb upstream primers: 142–161 and 452–471, internal 279–298; 2.6 kb downstream primers: 4141–4160 and 4411–4430, internal 4217–4236. Thymidine kinase (accession Nos. L29424 and L00365) CpG island primers: 180–199 and 426–445, internal 215–234; exon 3 primers: 1–20 and 109–128, internal 21–40. Elongation Factor 2 (accession No. J03200) CpG island primers: 321–340 and 621–640, internal 391–410. PCR conditions were 30 cycles at 94°C 1 min, 58°C 1 min and 72°C 2 min. One-third of the products was electrophoresed, blotted and hybridised to internal primers.

Three competitor fragments with the same sequence as their genomic targets within the APRT CpG island and their flanks but including 20 internal additional nucleotides (between brackets) were synthesised by using the following primers. Conditions for construction of the competitors and for their calibration were as described by Diviacco *et al.* (1992).
CpG island region: 5′-(GTCGACGGATCCCTGCAGGT)AGGGATAT-CTCGCCCCTCCT and 5′-(ACCTGCAGGGATCCGTCGAC)GGGAA-CAAGGAGAGGCTGGT.
4.5 kb upstream region: 5′-(GTCGACGGATCCCTGCAGGT)GCTA-GACCATTTCATGTCCT and 5′-(ACCTGCAGGGATCCGTCGAC)CT-GGAAAGCCTGAGACCTTG.
1.8 kb downstream region: 5′-(GTCGACGGATCCCTGCAGGT)G-AGCTGCTGGGCCAGCTACA and 5′-(ACCTGCAGGGATCCGTCG-AC)ACAGGCAGCGCACATGGTTC.

Human primers and cloning of short DNA nascent strands

Human K562 erythroleukaemia cells were grown in RPMI medium (Biowhittaker) with 10% fetal calf serum. DNA nascent strands were isolated by alkaline sucrose centrifugation essentially as described by Zhao *et al.* (1994). Asynchronous exponentially growing cells were pulsed with 50 μM BrdU for 30 min, harvested in PBS and immediately lysed by addition of an equal volume of 0.9 M NaOH, 1.1 M NaCl, 20 mM EDTA and 2% sodium *N*-lauroyl sarcosinate. The cell lysate was overlaid directly on top of a seven-step sucrose gradient (5–20% in steps of 2.5%) made up in 0.1 M NaOH, 0.9 M NaCl and 50 mM EDTA and centrifuged at 24 000 rpm for 20 h at 4°C in a Beckman SW-40 rotor. Six fractions were collected after centrifugation, dialysed against TE, ethanol precipitated and resuspended in TE. An equal volume of each sample was electrophoresed in an alkaline agarose gel, blotted and hybridised to total human DNA to monitor correct fractionation. Human β-globin (accession No. HSHBB) promoter primers are located at positions 62115–62134 and 62511–62530; upstream primers 54759–54778 and 55209–55228. A region 326 bp long from the human Trk A CpG island was amplified with the primers 5′-TTTCCTGGCGGCT-GGGTCTT and 5′-GCTCAGTCAGGTTCTCTGCG. The region 370 bp long spanning exons 9 and 10 was amplified with the primers 5′-AGGACGAAACACCTTTTGG and 5′-ACATTTGTTGAGCACAA-GGAGCAGCG. Nascent strands from fraction 1 were purified from potential non-replicating contaminating genomic DNA by two rounds of immunoprecipitation as described below. After that they were tailed with poly C at the 3′ end with Terminal Transferase following the supplier's instructions (Boehringer Mannheim). Complementary strand was synthesised by extending an 8mer poly G primer annealed to the poly C tail with T4 DNA polymerase. A second tail of poly C was added to this newly synthesised strand and the final product was amplified by 35 cycles of PCR (94°C 1 min, 48°C 1 min, 72°C 2 min) with the primer 5′-AACTGCAGGGGGGGGGGG. Electrophoresis of the PCR products gave a homogeneous distribution of fragments up to 1.5 kb identical to that of the departure material in fraction 1 of the sucrose gradient. They were digested with *Pst*I and cloned into the same site of vector pTZ18R for sequencing.

Sucrose gradients of HpaII-digested genomic DNA

Total DNA was isolated after ^{3}H-thymidine labelling and digested to completion with *Hpa*II to release the CpG island fraction. Restriction fragments were loaded on a 4 ml seven-step sucrose gradient (5–20% in steps of 2.5%) made up in 10 mM Tris, pH 7.4/50 mM NaCl/1 mM EDTA. Gradients were centrifuged at 50 000 r.p.m. for 3 h at 20°C in a Beckman SW-60 rotor. Seven fractions were collected from each gradient and 20 μl of each were used to monitor correct gradient fractionation by end-labelling. The remaining DNA fragments were precipitated with an equal volume of 10% trichloroacetic acid, collected onto glass fibre filters and the incorporated ^{3}H-thymidine was measured by scintillation counting.

Immunoprecipitation with anti-BrdU antibodies

We have followed the method described by Vassilev *et al.* (1990) with minor modifications. Total BrdU-substituted DNA (3 μg) was mixed with 15 μg of tRNA and boiled for 3 min to denature DNA. Tubes were immediately cooled on ice to room temperature and adjusted to a final volume of 200 μl with 10 mM sodium phosphate, pH 7.2/140 mM NaCl/0.05% Triton X-100. Anti-BrdU monoclonal antibody (10 μl of 25 mg/ml stock; Becton Dickinson) were added and the mixture was incubated for 1 h at room temperature with gentle shaking. Rabbit anti-mouse immunoglobulin G (2.5 ml of 14.0 mg/ml stock, Sigma) was then added and incubation proceeded an additional 1 h under the same conditions. Immunoprecipitates were collected by centrifugation for 15 min in a bench centrifuge, resuspended in 200 μl of 50 mM Tris, pH 8.0/10 mM EDTA/0.5% SDS/250 μg/ml Protease K and incubated for 2 h at 37°C. Carrier tRNA (10 μg) was added before phenol/chroloroform extraction and ethanol precipitation. DNA pellets were resuspended in TE (10 mM Tris, pH 8.0/1 mM EDTA) and submitted to a second round of immunoprecipitation before being used as input for PCR.

Cell synchronisation and ^{3}H-thymidine labelling

CHO cells were grown in α-MEM medium (Biowhittaker) with 10% fetal calf serum. They were synchronised by isoleucine starvation as described by Johnson *et al.* (1993). For synchronisation with mimosine, cells were grown to confluence and then split at 50% confluency into α-MEM medium with 10% fetal calf serum and 400 μM mimosine to arrest them at the G_1/S boundary (Mosca *et al.*, 1992). Synchronous progression through S phase was monitored by pulsing the cultures for 30 min before processing of the samples with 5 μCi of ^{3}H-thymidine per ml of culture (specific activity 5.0 Ci/mmol). DNA was isolated from each sample and the incorporated radioactivity measured by scintillation counting. Synchronous entry into S phase was monitored by immunofluorescence with anti-BrdU antibodies as described by Leonhardt *et al.* (1992).

Acknowledgements

We thank Avelino Bueno, Sergio Moreno and all members of the ASP laboratory for advice and helpful criticism, and Annette Düwel, Atanasio Pandiella, María Sacristán and Mercedes Tamame for comments on the manuscript. We are also grateful to Vicky Clark for DNA sequencing. This work was supported by grants from the ICRF and the Wellcome Trust (to A.B.) and DGICYT and Fundación Ramón Areces (to F.A.). S.D. and M.G. are recipients of CSIC/Glaxo and Ministerio de Educación y Cultura postgraduate fellowships.

References

Antequera,F. and Bird,A. (1993) Number of CpG islands and genes in human and mouse. *Proc. Natl Acad. Sci. USA*, **90**, 11995–11999.

Antequera,F., Macleod,D. and Bird,A. (1989) Specific protection of methylated CpGs in mammalian nuclei. *Cell*, **58**, 509–517.

Antequera,F., Boyes,J. and Bird,A. (1990) High levels of *de novo* methylation and altered chromatin structure at CpG islands in cell lines. *Cell*, **62**, 503–514.

Bestor,T.H. and Verdine,G.L. (1994) DNA methyltransferases. *Curr. Opin. Cell Biol.*, **6**, 380–389.

Bird,A. (1986) CpG-rich islands and the function of DNA methylation. *Nature*, **321**, 209–213.

Bird,A., Taggart,M., Frommer,M., Miller,O.J. and Macleod,D. (1985) A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell*, **40**, 91–99.

Brandeis,M., Frank,D., Keshet,I., Siegfried,Z., Mendelsohn,M., Nemes,A., Temper,V., Razin,A. and Cedar,H. (1994) Sp1 elements protect a CpG islands from *de novo* methylation. *Nature*, **371**, 435–438.

Chuang,L.S., Ian,H.I., Koh,T.W., Ng,H.H., Xu,G. and Li,B.F. (1997) Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for $p21^{WAF1}$. *Science*, **277**, 1996–2000.

Craig,J.M. and Bickmore,W.A. (1994) The distribution of CpG islands in mammalian chromosomes. *Nature Genet.*, **7**, 376–382.

DePamphilis,M.L. (1993a) How transcription factors regulate origins of DNA replication in eukaryotic cells. *Trends Cell Biol.*, **3**, 161–167.

DePamphilis,M.L. (1993b) Eukaryotic DNA replication: Anatomy of an origin. *Annu. Rev. Biochem.*, **62**, 29–63.

DeVilliers,J., Schaffner,W., Tyndall,C., Lupton,S. and Kamen,R. (1984) Polyoma virus DNA replication requires an enhancer. *Nature*, **312**, 242–246.

Diffley,J.F.X., Cocker,J.H., Dowell,S.J. and Rowley,A. (1994) Two steps in the assembly of complexes at yeast replication origins *in vivo*. *Cell*, **78**, 303–316.

Dijkwel,P.A. and Hamlin,J.L. (1995) The Chinese hamster dihydrofolate reductase origin consists of multiple potential nascent-strand sites. *Mol. Cell. Biol.*, **15**, 3023–3031.

Dimitrova,D.S., Giacca,M., Demarchi,F., Biamonti,G., Riva,S. and Falaschi,A. (1996) In vivo protein–DNA interactions at a human DNA replication origin. *Proc. Natl Acad. Sci. USA*, **93**, 1498–1503.

Diviacco,S., Norio,P., Zentilin,L., Menzo,S., Clememti,M., Biamonti,G., Riva,S., Falaschi,A. and Giacca,M. (1992) A novel procedure for quantitative polymerase chain reaction by coamplification of competitive templates. *Gene*, **122**, 313–320.

Epner,E., Rifkind,R.A. and Marks,P.A. (1981) Replication of α and β-globin DNA sequences occurs during early S phase in murine erythroleukemia cells. *Proc. Natl Acad. Sci. USA*, **78**, 3058–3062.

Gale,J.M., Tobey,R.A. and D'Anna,J.A. (1992) Localization and DNA sequence of a replication origin in the rhodopsin gene locus of Chinese hamster cells. *J. Mol. Biol.*, **224**, 343–358.

Gardiner-Garden,M. and Frommer,M. (1987) CpG islands in vertebrate genomes. *J. Mol. Biol.*, **196**, 261–282.

Giacca,M. *et al.* (1994) Fine mapping of a replication origin of human DNA. *Proc. Natl Acad. Sci. USA*, **91**, 7119–7123.

Goldman,M.A., Holmquist,G.P., Gray,M.C., Caston,L.A. and Nag,A. (1984) Replication timing of genes and middle repetitive sequences. *Science*, **224**, 686–692.

Hamlin,J.L. and Dijkwel,P.A. (1995) On the nature of the replication origins in higher eukaryotes. *Curr. Opin. Genet. Dev.*, **5**, 153–161.

Handeli,S., Klar,A., Meuth,M. and Cedar,H. (1990) Mapping replication units in animal cells. *Cell*, **57**, 909–920.

Heintz,N.H. and Hamlin,J.L. (1982) An amplified chromsomal sequence that includes the gene for dihydrofolate reductase initiates replication within specific rectriction fragments. *Proc. Natl Acad. Sci. USA*, **79**, 4083–4087.

Heinzel,S.S., Krysan,P.J., Tran,C.T. and Calos,M.P. (1991) Autonomous DNA replication in human cells is affected by the size and the source of the DNA. *Mol. Cell. Biol.*, **11**, 2263–2272.

Holmquist,G.P. (1989) Evolution of chromosome bands: Molecular ecology of noncoding DNA. *J. Mol. Evol.*, **28**, 469–486.

Huberman,J.A. and Riggs,A.D. (1968) On the mechanism of DNA replication in mammalian chromosomes. *J. Mol. Biol.*, **32**, 327–337.

Hutchison,C.A., Hardies,S.C., Loeb,D.D., Shehee,W.R. and Edgell,M.H. (1989) LINEs and related retroposons: Long interspersed repeated sequences in the eucaryotic genome. In Berg,D.E. and Howe,M.M. (eds), *Mobile DNA*. ASM Press, Washington DC, pp. 593–617.

Johnson,R.T., Downes,C.S. and Meyn,R.E. (1993) The synchronization of mammalian cells. In Fantes,P. and Brooks,R. (eds), *The Cell Cycle. A Practical Approach*. Oxford University Press, Oxford, UK, pp. 1–24.

Kelman,Z. and O'Donnell,M. (1994) DNA replication: enzymology and mechanisms. *Curr. Opin. Genet. Dev.*, **4**, 185–195.

Kitsberg,D., Selig,S., Keshet,I. and Cedar,H. (1993) Replication structure of the human β-globin gene domain. *Nature*, **366**, 588–590.

Kornberg,A. and Baker,T. (1992) *DNA Replication*. 2nd edn. W.H. Freeman, New York, NY.

Laird,P.W. and Jaenisch,R. (1996) The role of DNA methylation in cancer genetics and epigenetics. *Annu. Rev. Genet.*, **30**, 441–464.

Larsen,F., Gundersen,G., Lopez,R. and Prydz,H. (1992) CpG islands as gene markers in the human genome. *Genomics*, **13**, 1095–1107.

Leonhardt,H., Page,A.W., Weier,H.U. and Bestor,T.H. (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell*, **71**, 865–873.

Lewis,J.D., Meehan,R.R., Henzel,W.J., Maurer-Fogy,I., Jeppesen,P., Klein,F. and Bird,A. (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell*, **69**, 905–914.

Macleod,D., Charlton,J., Mullins,J. and Bird,A. (1994) Sp1 sites in the mouse aprt gene promoter are required to prevent methylation of the CpG island. *Genes Dev.*, **8**, 2282–2292.

Marahrens,Y. and Stillman,B. (1992) A yeast chromosomal origin of DNA replication defined by multiple functional elements. *Science*, **255**, 817–823.

Martín-Zanca,D., Mitra,G., Long,L.K. and Barbacid,M. (1986) Molecular characterization of the human Trk oncogene. *CSH Symp. Quant. Biol.*, **51**, 983–992.

Meehan,R.R., Lewis,J.D., McKay,S., Kleiner,E. and Bird,A. (1989) Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell*, **58**, 499–507.

Mosca,P.J., Dijkwel,P.A. and Hamlin,J.L. (1992) The plant aminoacid mimosine may inhibit initiation at origins of replication in Chinese hamster cells. *Mol. Cell. Biol.*, **12**, 4375–4383.

Newlon,C.S. (1988) Yeast chromosome replication and segregation. *Microbiol. Rev.*, **52**, 568–601.

Nur,I., Pascale,E. and Furano,A.V. (1988) The left end of rat L1 (L1Rn, long interspersed repeated) DNA which is a CpG island can function as a promoter. *Nucleic Acids Res.*, **16**, 9233–9251.

Pelizon,C., Diviacco,S., Falaschi,A. and Giacca,M. (1996) High-resolution mapping of the origin of DNA replication in the hamster dihydrofolate reductase gene domain by competitive PCR. *Mol. Cell. Biol.*, **16**, 5358–5364.

Pfeifer,G.P., Tanguay,R.L., Steigerwald,S.D. and Riggs,A.D. (1990) In vivo footprint and methylation analysis by PCR-aided genomic sequencing: comparison of active and inactive X chromosomal DNA at the CpG island and promoter of human PGK-1. *Genes Dev.*, **4**, 1277–1287.

Rein,T., Zorbas,H. and DePamphilis,M.L. (1997) Active mammalian replication origins are associated with a high-density cluster of mCpG dinucleotides. *Mol. Cell. Biol.*, **17**, 416–426.

Riggs,A.D. and Pfeifer,G.P. (1992) X-chromosome inactivation and cell memory. *Trends Genet.*, **8**, 169–174.

Sassaman,D.M., Dombroski,B.A., Moran,J.V., Kimberland,M.L., Naas,T.P., DeBerardinis,R.J., Gabriel,A., Swergold,G.D. and Kazazian,H.H. (1997) Many human L1 elements are capable of retrotransposition. *Nature Genet.*, **16**, 37–43.

Schmidt,M. and Migeon,B.R. (1990) Asynchronous replication of homologous loci on human active and inactive chromosomes. *Proc. Natl Acad. Sci. USA*, **87**, 3685–3689.

Somma,M.P., Pisano,C. and Lavia,P. (1991) The housekeping promoter from the mouse CpG island HTF9 contains multiple protein-binding elements that are functionally redundant. *Nucleic Acids Res.*, **19**, 2817–2824.

Taira,T., Iguchi-Ariga,S.M. and Ariga,H. (1994) A novel DNA replication origin identified in the human heat shock protein 70 gene promoter. *Mol. Cell. Biol.*, **14**, 6386–6397.

Taljanidisz,J., Popowski,J. and Sarkar,N. (1989) Temporal order of gene replication in Chinese hamster ovary cells. *Mol. Cell. Biol.*, **9**, 2881–2889.

Tazi,J. and Bird,A. (1990) Alternative chromatin structure at CpG islands. *Cell*, **60**, 909–920.

Tommasi,S. and Pfeifer,G.P. (1995) *In vivo* structure of the human *cdc2* promoter: Release of a p130-E2F-4 complex from sequences immediately upstream of the transcription initiation site coincides with induction of *cdc2* expression. *Mol. Cell. Biol.*, **15**, 6901–6913.

van der Vliet,P.C. (1996) Roles of transcription factors in DNA replication. In DePamphilis, M. (ed.), *DNA Replication in Eukaryotic Cells*. CSH Laboratory Press, New York, NY, pp. 87–118.

Vassilev,L. and Johnson,E.M. (1990) An initiation zone of chromosomal DNA replication located upsteam of the c-myc gene in proliferating HeLa cells. *Mol. Cell. Biol*., **10**, 4899–4904.

Vassilev,L.T., Burhans,W.C. and DePamphilis,M.L. (1990) Mapping an origin of DNA replication at a single-copy locus in exponentially proliferating mammalian cells. *Mol. Cell. Biol*., **10**, 4685–4689.

Venolia,L., Gartler,S.M., Wassman,E.R., Yen,P., Mohandas,T. and Shapiro,L.J. (1982) Transformation with DNA from 5-azacytidine-reactivated X chromosomes. *Proc. Natl Acad. Sci. USA*, **79**, 2352–2354.

Virta-Pearlman,V.J., Gunaratne,P.H. and Chinault,A.C. (1993) Analysis of a replication initiation sequence from the adenosine deaminase region of the mouse genome. *Mol. Cell. Biol*., **13**, 5931–5942.

Wang,S., Dijkwel,P.A. and Hamlin,J.L. (1998) Lagging-strand, early-labelling, and two-dimensional gel assays suggest multiple potential initiation sites in the Chinese hamster dihydrofolate reductase origin. *Mol. Cell. Biol*., **18**, 39–50.

Yen,P.H., Patel,P., Chinault,A.C., Mohandas,T. and Shapiro,L.J. (1984) Differential methylation of hypoxanthine phosphoribosyltransferase genes on active and inactive human chromosomes. *Proc. Natl Acad. Sci. USA*, **81**, 1759–1763.

Zhao,Y., Tsutsumi,R., Yamaki,M., Nagatsuda,Y., Ejiri,S. and Tsutsumi,K. (1994) Initiation zone of DNA replication at the aldolase B locus encompasses transcription promoter region. *Nucleic Acids Res*., **22**, 5385–5390.

Received September 30, 1997; revised February 23, 1998; accepted February 25, 1998