Nucleic Acids Research, 1998, Vol. 26, No. 10 **2313–2321**

ori*GNAI3*: a narrow zone of preferential replication initiation in mammalian cells identified by 2D gel and competitive PCR replicon mapping techniques

Franck Toledo*, Bruno Baron, Maria-Aparecida Fernandez, Anne-Marie Lachagès, Véronique Mayau, Gérard Buttin and Michelle Debatisse

Unité de Génétique Somatique (URA CNRS 1960), Institut Pasteur, 25 rue du Dr Roux, 75724 Paris Cedex 15, France

Received February 16, 1998; Revised and Accepted March 28, 1998 DDBJ/EMBL/GenBank accession no. Y08232

ABSTRACT

The nature of mammalian origins of DNA replication remains controversial and this is primarily because two-dimensional gel replicon mapping techniques have identified broad zones of replication initiation whereas several other techniques, such as quantitative PCR, have disclosed more discrete sites of initiation at the same chromosomal loci. In this report we analyze the replication of an amplified genomic region encompassing the 3′-end of the *GNAI3* gene, the entire *GNAT2* gene and the intergenic region between them in exponentially growing Chinese hamster fibroblasts. These cells express *GNAI3* but not *GNAT2*. The replication pattern was first analyzed by two-dimensional neutral–alkaline gel electrophoresis. Surprisingly, the results revealed a small preferential zone of replication initiation, of at most 1.7 kb, located in a limited part of the *GNAI3–GNAT2* intergenic region. Mapping of this initiation zone was then confirmed by quantitative PCR. The agreement between the two techniques exploited here strengthens the hypothesis that preferred sites of replication initiation do exist in mammalian genomes.

INTRODUCTION

The replicon model (1), which postulates binding of a positive-acting factor (the initiator) to a specific DNA sequence (the replicator) to trigger initiation of DNA replication at a nearby site (the origin of replication), has been verified in prokaryotes and eukaryotic DNA viruses and also seems likely to apply to yeasts. To what extent this model is valid for initation of DNA replication in mammalian chromosomes remains controversial.

A large variety of techniques have been developed to study chromosomal replication initiation in mammalian cells (for a review see 2; 3,4). A first group of techniques includes methods that rely on characterization of the earliest labeled fragments when synchronized cells enter S phase (5), determination of leading strand polarities (6,7), measurement of the distribution of Okazaki fragments between the two template strands of a replication fork (8) and determination of the length or numbers of nascent strands along a genomic domain (3,9). All these methods led to the characterization of specific sites of replication initiation (for a review see 10). The other group of methods, first developed to analyze yeast DNA replication (11,12), relies on specific mobility patterns of replication intermediates during two-dimensional (2D) gel electrophoresis. Surprisingly, when 2D gels have been applied to analyze mammalian DNA replication, broad zones of replication initiation were found (13,14).

Indeed, when 2D techniques were applied to analysis of the Chinese hamster dihydrofolate reductase gene (*DHFR*) origin region, replication was found to initiate anywhere within 30–60 kb (13,15), whereas several other techniques suggested the existence of two discrete sites of initiation (5,6,16), one of which was circumscribed to only 450 bp (8). Likewise, 2D gel analysis of initiation at the human rDNA locus identified a 30 kb long broad zone of replication initiation (14), but more restricted sites of initiation were found with several other techniques (17–19). The reasons for such contradictions are poorly understood: one tentative explanation for this is that 2D gels detect all initiation events, even those occuring in regions firing very rarely, whereas techniques such as quantitative PCR (3) or Okazaki fragment polarity (8) detect only the region(s) where most initiation events occur, because they measure signal-to-noise ratios. This hypothesis implies that mammalian DNA replication origins consist of large intiation zones that may contain one or a few preferred sites of initiation (10,14,20,21), but it does not explain why oriβ, the preferred site of replication initiation identified in the Chinese hamster *DHFR* locus by many techniques, does not appear to be a preferred site by 2D gel analyses (21,22). Therefore, the existence of preferred sites of replication initiation in mammalian genomes remains a matter of controversy.

Here we describe the replication pattern of the region surrounding the Chinese hamster adenylate deaminase 2 gene (*AMPD2*). This region was chosen because its amplification can be selected for (23). In previous work this property allowed us to

*To whom correspondence should be addressed. Tel: +33 1 45 68 85 71; Fax: +33 1 40 61 31 71; Email: ftoledo@pasteur.fr

show that mechanisms relying on chromosomal rearrangements and unequal segregation at mitosis, rather than abnormal replication initiation, account for mammalian DNA amplification: cells most frequently acquire multiple copies of the *AMPD2* gene through cycles of chromatid breakage–fusion-bridges (24–26). Gene amplification was here expected to facilitate the use of a 2D gel replicon mapping technique. Indeed, application of such techniques to the study of DNA replication in mammalian cells has been hampered by the high complexity of mammalian genomes. However, 2D gels have proved adequate to analyze the replication of repeated loci, such as the *DHFR* locus in CHOC 400 cells (13) and the human rRNA genes (14), in asynchronous cell populations. Amplification *per se* does not alter the DNA replication pattern of a locus, since in the *DHFR* system amplified and unamplified cells use the same replication initiation zone (13,27) or the same preferred initiation site (5,8,16). Accordingly, we chose to analyze the *AMPD2* region in the 42 cell line, which has some 30 extra copies of this locus (28) located, like the normal copies in unamplified cells, on the long arm of chromosome 1 (29). The *AMPD2* region is a polygenic region that contains four unrelated genes within <100 kb (30). These genes (listed in their respective order) encode: the αi3 and αt2 subunits of GTP binding proteins (GNAI3 and GNAT2), the AMPD2 enzyme and a glutathione S-transferase of the μ family (GSTM). *GNAI3* is ubiquitously expressed, *GNAT2* is a retinal cell-specific gene, *AMPD2* is expressed in all tissues but muscle and the expression patterns of the different members of the *GSTM* family have not yet been precisely determined. In this report we have analyzed the replication pattern of a 30 kb long sequence that contains the 3′-end of the *GNAI3* gene, the *GNAT2* gene and the intergenic region between them by the 2D neutral–alkaline (2D NA) gel electrophoresis technique (12). Remarkably, this technique identified a narrow zone of preferential replication initiation in a limited part of the *GNAI3–GNAT2* intergenic region. We confirmed the existence and position of this initiation zone by competitive quantitative polymerase chain reaction (PCR) (3). Thus 2D gel and competitive PCR analyses show concordant results, arguing in favor of the existence, at least for some loci, of preferred sites of initiation of DNA replication in mammalian cells.

MATERIALS AND METHODS

Cell line and culture conditions

The 42 cell line is a coformycin-resistant mutant isolated after two steps of selection from the GMA32 Chinese hamster lung fibroblast cell line. Cells of line 42 were grown in Eagle's medium supplemented with adenine, azaserine, uridine and 5 μg/ml coformycin (31). Each cell of this line contains ~30 copies of the *AMPD2* region (28).

2D NA gel electrophoresis analysis

Typical experiments started from eight 23 cm^2 dishes, each containing 2–3 × 10^7 exponentially growing cells of line 42. Nuclear halos (prepared as described by Dijkwel *et al*., method E; 32) were collected by centrifugation at 2500 r.p.m. for 5 min at 4°C. Pellets were washed with appropriate cold restriction buffer and digested with 10 000 U restriction enzyme in 20 ml restriction buffer at 37°C for 90 min. Two milligrams of RNase A were added and incubation was continued for another 30 min at 37°C. The loop and the matrix fractions were separated by centrifugation at 3000 r.p.m. for 5 min. The pellet containing the matrix and bound DNA was incubated for 15 min at 37°C with 2 mg/ml proteinase K in PK buffer (10 mM Tris–HCl, pH 8.0, 1 mM EDTA, 0.3 M NaCl). NaCl concentration was similarly raised to 0.3 M in the loop fraction. Replication intermediates from both fractions were independently purified by benzoyl-naphtoyl-DEAE (BND)–cellulose columns as described by Dijkwel *et al.* (32).

The protocol used for 2D NA gel electrophoresis was modified from Nawotka and Huberman (12). Restriction fragments were separated in the first dimension on a 0.4% agarose gel run in a cold room for 40 h at 0.5 V/cm in TBE (89 mM Tris–HCl, pH 8.0, 89 mM boric acid, 2 mM EDTA). The lanes containing the matrix and loop fractions were cut out and each was placed at the top of a second gel tray and a 0.6% agarose gel (in H_2O) was cast around it. This gel was soaked in alkaline electrophoresis buffer (50 mM NaOH, 2.5 mM EDTA) for 1 h and the second dimension electrophoresis was run at 1.2–1.5 V/cm overnight in a cold room with buffer circulation. The sizes of fragments and nascent strands were determined using DNA molecular weight marker X (Boehringer Mannheim) run in lanes adjacent to the experimental samples in both dimensions. For restriction fragments the observed sizes were further verified with the nucleotide sequence data. After electrophoresis DNA was blotted, cross-linked to nylon membranes (Nytran 13 N; Schleicher and Schuell) by UV irradiation and hybridized with [α-^{32}P]dCTP- and [α-^{32}P]dGTP-labeled probes (devoid of repetitive sequences). Membranes were washed in 2× SSC, 0.1% SDS at room temperature, then 0.2× SSC, 0.1% SDS at 37°C and exposed on storage phosphor screens (Molecular Dynamics). After 10 days exposure screens were scanned by a Phosphorimager (Molecular Dynamics). Most replication intermediates were reproducibly recovered in matrix DNA preparations. All the photographs presented were obtained with ImageQuant (Molecular Dynamics) and Photoshop (Adobe) software.

Competitive PCR analysis

In a first series of experiments, DNA extraction and size fractionation by sedimentation through neutral sucrose gradients was adapted from Dijkwel *et al*., Biamonti *et al*. and Yoon *et al*. (4,32,33). Cells, grown in eight 23 cm^2 dishes as for 2D experiments, were labeled for 4 min with 10^{-2} M BrdUrd. This and all subsequent steps were done in the dark. The labeling medium was removed from the dishes, cells were washed, scraped and nuclei were recovered as described by Dijkwel *et al*. (32). Each pellet was resuspended in 1 ml 10 mM Tris, pH 8.0, 50 mM EDTA, then DNA preparations were denatured by addition of 1 ml 1 M NaCl, 0.5 M NaOH, 50 mM EDTA, 0.4% Sarkosyl, followed by incubation at 45°C for 30 min. The resulting 2 ml were poured on top of a 35 ml neutral sucrose gradient (linear, 5–30% sucrose in 10 mM Tris, pH 7.4, 5 mM EDTA, 250 mM NaCl). DNA strands were separated according to size by centrifugation in a Beckman SW28 rotor (19 000 r.p.m. for 16 h at 21°C) and fractions (1 ml) were collected from the gradient. The 10–12 fractions corresponding to the upper portion of the gradient were neutralized by addition of 100 μl 1 M Tris, pH 8 (the other fractions were neutral after centrifugation). Aliquots (20 μl) of each fraction were loaded on an alkaline agarose gel and the size distribution of DNA strands was determined by hybridization with labeled Chinese hamster total genomic DNA. Fractions containing 1–2 kb DNA strands were

pooled and further purified on anti-BrdUrd immunoaffinity chromatography columns as previously described by Contreas *et al.* (34). Eluates from immunoaffinity chromatography were then quantified by competitive PCR as described in detail by Diviacco *et al.* (35). In a second series of experiments we applied the simplified procedure of Kumar *et al.* (36), omitting the labeling of cells with BrdUrd. Molecules (0.6–1.6 kb in length) were selected from sucrose gradients and quantified by competitive PCR. In the experiment that compares exponentially growing cells and confluent cells, cells were seeded in ten 23 cm^2 dishes (5.5×10^6 cells/dish). Growth rates and thymidine (dT) incorporation were determined after 48, 72 and 96 h: cells from one dish were labeled for 15 min with [^{3}H]T (1 μCi/ml) and T (2×10^{-6} M), trypsinized, counted and aliquots of 2×10^6 cells were used to determine [^{3}H]T incorporation. The cells of each aliquot were treated for 15 min at 80 °C in 0.4 M NaOH, spun briefly at 10 000 r.p.m. in an Eppendorf centrifuge, the DNA precipitated with trichloroacetic acid and the incorporated radioactivity determined. After 48 h cells were exponentially growing (1.6×10^7 cells/dish); after 72 h (5.3×10^7 cells/dish) they still incorporated about half as much [^{3}H]T as after 48 h; after 96 h cells stopped proliferating (5.8×10^7 cells/dish) and incorporation fell to one tenth of that observed after 48 h. Exactly the same number (6×10^7) of exponentially growing (48 h) or confluent (96 h) cells from parallel dishes were recovered and treated according to the simplified procedure.

The prerequisite for quantification by PCR was construction of competitors, which required the synthesis of four primers for each locus to be tested: two were chosen in order to amplify DNA fragments of 200–300 bp [ef, external forward; er, external reverse); the two others (if, internal forward; ir, internal reverse) consisted of two complementary 5′-tails of 20 nt unrelated to genomic sequences linked to specific sequences on the 3′-end complementary to the genomic targets (tail1, 5′-GTCGACGGATCCGAATTCGT-3′; tail2, 5′-ACGAATTCGGATCCGTCGAC-3′). This allows competitors to differ from the genomic target by an addition of 20 nt only. Sequences (5′→3′) of the primers were as follows. Primer set α: ef, CCATTAGAGCCTTTGGTTTCC; er, ACCCCCTTACTCTGAACAGATG; if, tail1+GCCTGTCTCCAGAGCGAGT; ir, tail2+TGGCCTTGAACTCACAGAGC. Primer set β: ef, AACCATAACCTTTGTTGTTGGC; er, TTTGGGGGGTAGGTGTGTAA; if, tail1+ATAGGAGATGGTCTCTCTTTCCTT; ir, tail2+CTGCTCAAGTTTCCACTCAGG. Primer set γ: ef, TTCGTTGGGCTACTCAGACAT; er, AGGGAACATACCAGTGTGAACA; if, tail1+TTCCTTTTGTGTTGTCCAAAGAT; ir, tail2+TGTAAGACTGAAAGCGACCTGTT. Primer set δ: ef, TAAACCATGGATGCCAAGTG; er, TCATTCCAGCTGCTGTATGTTT; if, tail1+ATAATTTACCCCTGGTTTGCTT; ir, tail2 +TCCTTTGAACTGCTCTGAGG. Primer set ε: ef, TGGGGAGGATATAAGGGTCA; er, ACTGACAGCATGGACATTCC; if, tail1+CCACTGACTCCCACACCTTG; ir, tail2 +AGCGGGTCATGATGTTGAG.

The sequence used to search for primer sets results from linking the sequence for the *GNAI3–GNAT2* intergenic region (this work; sequence data submitted to the EMBL database under accession no. Y08232, sequencing performed as previously described; 37) with previously published sequences for the *GNAI3* (37,38) and *GNAT2* genes (39). The localizations of the amplification products obtained with the different primer sets are: α (750–2057), β (6103–6286), γ (7270–7600), δ (10506–10763), ε (16897–17153). [The positions of probes used for 2D gel analysis are: a (2578–3196), b (4269–5151), c (5146–6342), d (10764–12992), e (16891–17806)]. Once constructed, the competitors were then precisely quantified according to Pelizon *et al.* (16).

In a typical competitive PCR experiment quantification of each genomic target sequence was performed by co-amplifiying known and increasing amounts of the corresponding competitor with a fixed amount of a nascent strand preparation. To enhance reliability of quantification all sets of external primers allowed amplification reactions to occur under the same temperature cycling conditions, i.e. 5 min at 95 °C, 40 cycles of 1 min at 95 °C, 1 min at 60 °C, 2 min at 72 °C, 5 min at 72 °C, end at 4 °C. TaqStart antibody (Clontech) and Taq polymerase (Roche) were used according to the suppliers procedures. PCR products were run on a 2.2% Metaphor agarose gel (FMC Bioproducts), images of the gels were captured by the Bioprint hardware system (Vilber Lourmat), the ratio between upper and lower bands was precisely quantified by using NIH Image (W.Rasband, NIH) and results were plotted with Excel (Microsoft).

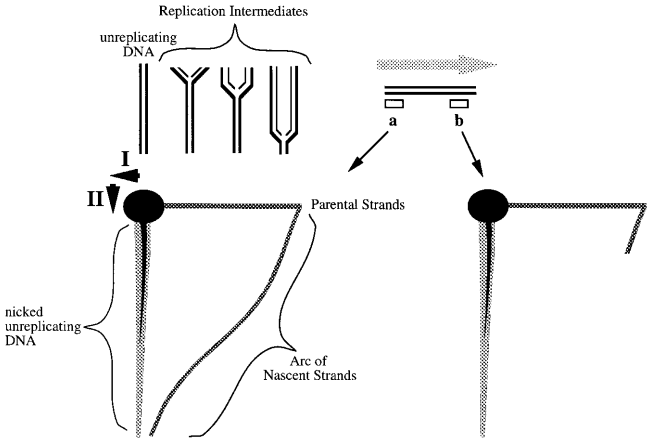

Figure 1. Principle of the 2D NA gel electrophoresis. In the first dimension (I) a size-based separation of the fragments is achieved by standard neutral gel electrophoresis, leading to retardation of fork-containing molecules according to their extent of replication. Electrophoresis in the second dimension (II) is performed perpendicularly to the first one under alkaline conditions, thus separating newly synthesized ('nascent') strands from parental strands. The direction of progression of DNA replication forks is deduced from the pattern of nascent strands detected by probes a and b: in the depicted example DNA replication forks progress from left to right in the fragment (large grey arrow), since all nascent strands are detected with probe a, whereas probe b only allows detection of the longest ones (adapted from 12).

RESULTS

2D NA gel analysis reveals DNA replication initiation events in the *GNAI3–GNAT2* region

The principle of the 2D NA gel electophoresis technique is represented in Figure 1. We chose to use this technique because it enables determination of the direction of replication fork progression and is thus particularly well adapted to analysis of large genomic regions. We studied the replication pattern of a 30 kb domain encompassing 6 kb of the 3′-end of the *GNAI3* gene, the entire *GNAT2* gene and the intergenic region in asynchronous cells of line 42. A narrow replication initiation zone was detected within the *GNAI3–GNAT2* intergenic region.

The experimental results sustaining this conclusion are shown in Figure 2A. When *Hin*dIII-digested DNA was hybridized with

probes b+c, a 6.3 kb long fragment (referred to as H1) was revealed and the large majority of detected nascent strands were 4.5–6.3 kb long (Fig. 2A, H1/b+c, left), indicating that most replication forks proceed from right to left in this fragment. However, image intensification and contrast enhancement of the same image revealed the presence of less abundant 2.6–4.5 kb long nascent strands (Fig. 2A, H1/b+c, right). These nascent strands did not reflect the presence of replication forks entering the fragment from its left nor emanating from the region homologous to probes b+c, since strands <2.6 kb should have been observed in these cases (indeed, the same blot hybridized with probe d revealed nascent strands as small as 1.4 kb; see Fig. 2A, H2/d). One hypothesis that accounts for both visualization of 2.6–4.5 kb long nascent strands and the absence of strands <2.6 kb is that replication initiation occurs close to the region homologous to probe c, on its right. Assuming that replication initiation is bidirectional, the existence of 2.6 kb long nascent strands suggests that initiation events occur 1.3 kb away from probe c. This distance is slightly overestimated, since nascent strands need to share some homology with probe c in order to be detected. For example, if the required amount of homology is ~500 bp, the initiation events closest to probe c would occur ~0.8 kb, rather than 1.3 kb, from this probe (in future estimations the homology required for strand detection will also be taken as 500 bp). Hybridization of *Hin*dIII-digested DNA with probes d and e revealed the adjacent fragment H2. The shortest nascent strands detected with probe d were 1.4 kb long, a length that can be considered as the limit of sensitivity in this experiment (the shortest detectable nascent strands were, depending on the blot, 1–1.6 kb in length). The observation of strands of this size suggested either that some replication forks enter fragment H2 from its left or that initiation events occur at or very near the sequence homologous to probe d. This ambiguity is resolved below. In any case, these forks, as well as forks entering the fragment from its right, should generate nascent strands of at least 7 kb upon hybridization with probe e; Figure 2A, H2/e shows that this is indeed the case. Considering that replication initiation is bidirectional and that 500 bp of homology are required between the probe and nascent strands, the absence of detectable nascent strands <7 kb suggested that initiation events occur neither in the sequence homologous to probe e nor in the 3 kb long regions on the right and on the left of this sequence. Thus, replication initiation events do not occur at a detectable rate in a region of ~7 kb centered around probe e (Fig. 2B). In this and in the experiments described below, the background of the blots represents ~15–20% of the faintest detectable signal. Thus, it can be estimated that if replication initiation events occur in this 7 kb long region, they are at least five times less frequent than in the region for which initiation events were successfully detected.

These results, obtained with *Hin*dIII-digested DNA, were confirmed and extended with DNA digested with *Pst*I, *Bam*HI and *Eco*RV. In the case of the 8.2 kb long fragment P nascent strands 5–8.2 kb in length were detected with probe a (Fig. 2A, P/a), whereas 3.5–8.2 kb long nascent strands were detected with probe b (Fig. 2A, P/b), indicating that replication forks proceed from right to left in this fragment. As above, the absence of nascent strands <5 kb upon hybridization with probe a indicated that detectable initiation events occur neither in the sequence homologous to probe a nor in the 2 kb long regions on the right and on the left of this sequence. Thus replication initiation does not occur at a detectable level in a 4.7 kb long region centered around probe a (Fig. 2B). Likewise, the absence of nascent strands <3.5 kb with probe b indicates that initiation events do not occur at a detectable rate in a 3.3 kb long region centered around probe b (Fig. 2B); this region includes the sequence homologous to probe c, confirming the *Hin*dIII analysis (Fig. 2A, H1/c). The results with both *Hin*dIII and *Pst*I indicate that detectable replication initiation does not occur in a 6.8 kb long region that encompasses the 3′-end of the *GNAI3* gene (Fig. 2B).

The pattern of nascent strands detected for the 8.5 kb long fragment B1 also confirmed the results obtained for fragment H1; a strong arc of 1.6–8.5 kb long nascent strands was observed with probe d (Fig. 2A, B1/d), whereas, at a similar exposure, no nascent strand could be detected with probe c (Fig. 2A, B1/c, left), indicating that most replication forks proceeded from right to left in this fragment. Image intensification and contrast enhancement of the latter blot (Fig. 2A, B1/c, right) revealed a faint arc of nascent strands 2.4–8.5 kb in length (whereas molecules of 1.6 kb can be seen with the same blot and probe d). As above (see H1/c), this pattern suggests that initiation events occur ~0.7 kb from probe c, an estimation that is in agreement with the 0.8 kb measurement deduced from the *Hin*dIII results (Fig. 2B). Furthermore, hybridization of fragments H2 and B1 with probe d revealed very short nascent strands. Given the position of this probe in the fragments, such a result means either that this sequence is passively replicated by forks travelling in both directions or that replication initiates in or very close to the sequence homologous to d. Hybridization of *Eco*RV-digested DNA with probe d clarified this question (Fig. 2A, EV/d); no nascent strands <5 kb were detected. This result indicates that initiation events do not occur in the region homologous to probe d nor in the 2 kb long regions to its right and left, thus in an ~6.3 kb long region centered around probe d (Fig. 2B). Moreover, hybridization of *Bam*HI-digested DNA with probe e (Fig. 2A, B2/e) revealed the existence of replication forks entering fragment B2 from its right end and moving from right to left. The existence of forks travelling from right to left was also confirmed the hybridization of *Eco*RV-digested DNA with probe e (Fig. 2A, EV/e), which allowed detection of 2.8–14.5 kb long nascent strands. The absence of strands <2.8 kb also confirmed that no detectable replication initiation events occur near the sequence homologous to probe e, as first suggested by *Hin*dIII-digested DNA hybridized with this probe (Fig. 2B).

Thus, the entire set of results obtained with probes d and e show that no detectable replication initiation events occur in a region of ~12 kb encompassing the right-most part of the *GNAI3–GNAT2* intergenic region and the entire *GNAT2* gene (Fig. 2B). This region appears to be passively replicated by forks moving in opposite directions; several hypotheses may account for this situation (see Discussion below). Initiation events were detected with probe c independently and from the results presented here it is possible to roughly estimate the maximal size of the identified initiation zone. As discussed before, analyses of *Hin*dIII- and *Bam*HI-digested DNA place the left-most boundary of the initiation zone 0.7–0.8 kb from probe c. Analysis of *Eco*RV-digested DNA with probe d does not permit mapping of the right-most limit of the initiation zone, but indicates that initiation events are not detected in the 2 kb long region contiguous with probe d. Taken together, these results show that the initiation zone spans at most 1.7 kb. As mentioned above, since the background of the different experiments is at most 20% of the faintest detectable signal, initiation events in the initiation zone defined here occur

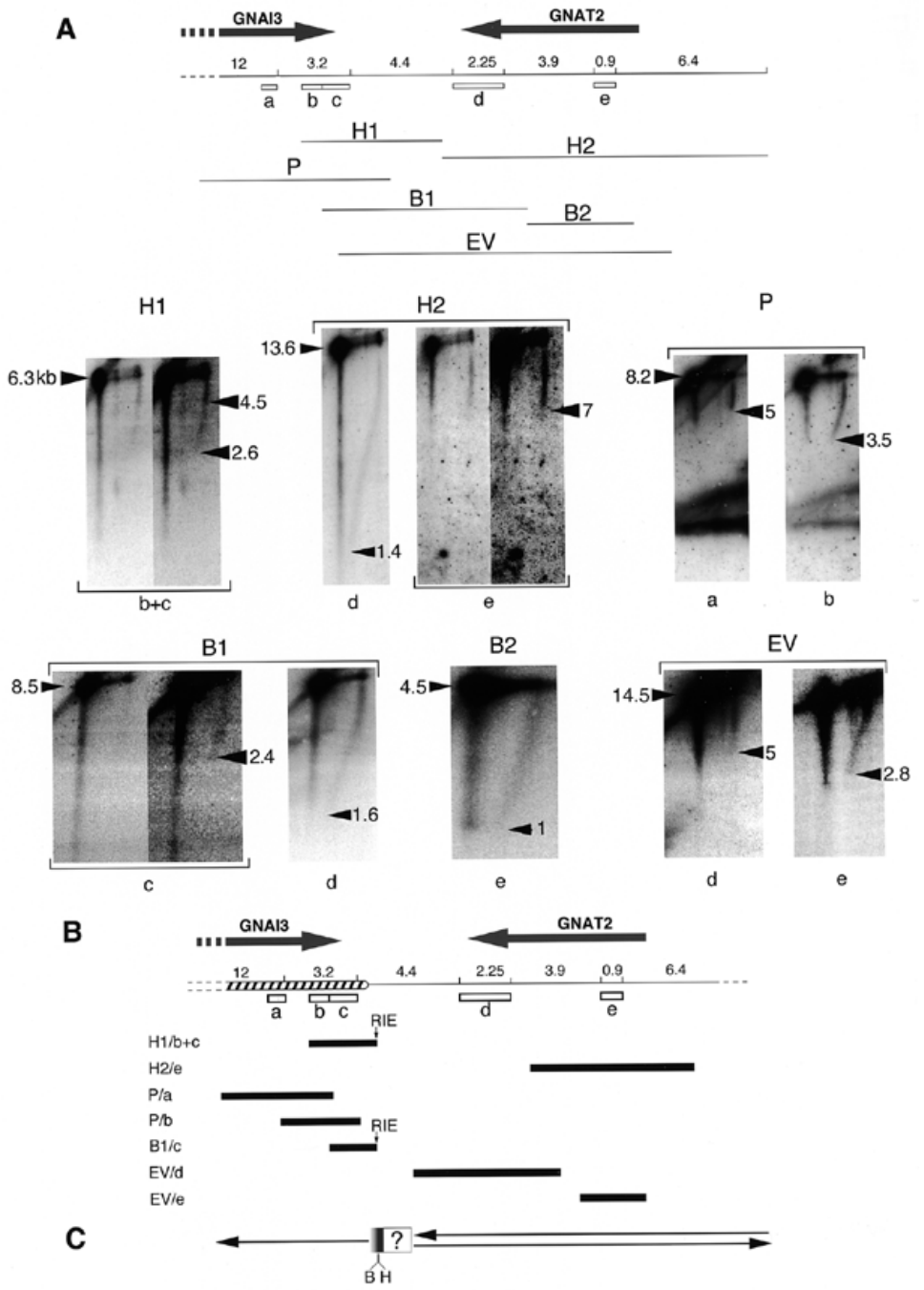

Figure 2. Analysis of the replication pattern in the *GNAI3–GNAT2* region by 2D NA gel electrophoresis. (**A**) Map of the *GNAI3–GNAT2* region and 2D NA analysis. On the map the *GNAT2* gene and part of the *GNAI3* gene are shown (arrows indicate directions of transcription). Vertical bars, *Eco*RI sites; numbers, size of *Eco*RI restriction fragments; rectangles, probes a–e. Below the map are the fragments analyzed (H1, H2, P, B1, B2 and EV); the corresponding restriction sites are: H, *Hin*dIII; P, *Pst*I; B, *Bam*HI; EV, *Eco*RV. The sizes of restriction fragments and positions of the probes were verified with the nucleotide sequence, which is available for the entire region. The names of the fragments are indicated above the photographs and probes are reported below. Arrowheads pointing to the right indicate the sizes of the studied fragments and arrowheads pointing to the left focus on the sizes of nascent strands. Photographs corresponding to analyses of fragments H1 and H2 represent different hybridizations of the same blot; a single blot was used for P, as well as for B1. (**B**) Localization of initiation events in the *GNAI3–GNAT2* region. Map of the region as above; below the map are summarized the hybridization experiments that allowed mapping of the initiation zone. Black boxes, regions in which initiation events were not detected; RIE, detected replication initiation events. (**C**) Conclusions of the analysis. Thin arrows, direction of progression of replication forks; grey rectangle, localization of the left-most limit of the replication initiation region (the gradient represents the uncertainty resulting from the required shared homology between probe and nascent strands; see text). The right-most limit of this region lies somewhere within the box containing a question mark.

at least five times more frequently than in the surrounding regions. Finally, *Hin*dIII and *Bam*HI analyses with probe c and *Pst*I analysis with probes a and b show that no replication forks enter fragments H1, B1 or P from their left end, which indicates that the *GNAI3* gene, unlike the *GNAT2* gene, is replicated only

by forks moving from right to left. These conclusions are summarized in Figure 2C.

Analysis of the *GNAI3–GNAT2* region by competitive PCR

Since repetitive elements increase the background noise dramatically, the use of the 2D NA technique is limited by the availability of probes that are devoid of such elements and that are at least 500 bp in length, to ensure adequate sensitivity. We were unable to use such a probe closer to the initiation zone than probes c and d, because of the repetitive sequences present at or very close to the initiation zone, a situation similar to the one described for the oriβ region of the *DHFR* locus (5,13). In order to confirm and refine mapping of the initiation zone we chose to study the *GNAI3–GNAT2* region using the competitive PCR technique (3,33), which relies on quantification of purified short nascent strands. Quantification of these strands at different loci of a chosen region is performed with a set of two 20 bp long primers located some 200–300 bp apart; this technique can be used even for sequences enriched in repetitive elements.

Primer sets were designed to quantify short nascent strands at five different loci in the *GNAI3–GNAT2* region (Fig. 3, top); one (γ) maps within the initiation zone defined by 2D NA results, two (β and δ) within the intergenic region but outside the initiation zone and the last two within the *GNAI3* (α) or the *GNAT2* (ε) genes. In a first series of experiments, exponentially growing cells of line 42 were labeled with a short pulse of BrdUrd and 1–2 kb long BrdUrd-labeled molecules were recovered (see Materials and Methods). Quantification was achieved by primer competition between the genomic target and a competitor molecule differing from the genomic sequence only by a 20 bp insertion (35). In order to determine the amount of BrdUrd-labeled molecules at each of the tested loci, increasing amounts of the appropriate competitor molecules were mixed with a fixed unknown amount of a preparation of 1–2 kb long BrdUrd-labeled molecules and submitted to PCR. A typical example of this analysis is shown in Figure 3, with peaks of short BrdUrd-labeled molecules being observed at loci amplified with primer sets γ and β. The results of five experiments carried out with independent DNA preparations are presented in Figure 4A. These results show the following. (i) A peak of molecules is detected with primer set γ, the only primer set that maps within the initiation region previously defined by 2D NA. (ii) Primer set β, which maps outside the initiation region defined by the 2D experiments, allows detection of nearly as many molecules as primer set γ. Primer set β lies ~700 bp from the left-most limit of the initiation zone defined by 2D NA, thus the nascent strands emanating from this initiation region should be ~1.4 kb long when they reach the sequence amplifiable with primer set β. Since 1–2 kb long nascent strands were selected for quantification, PCR results are consistent with the hypothesis that molecules quantified with primer set β correspond to initiation events occurring within the zone delimited by 2D NA. In order to test this hypothesis we also performed experiments with 0.6–1.6 kb long molecules (see below); quantification of smaller molecules was indeed expected to significantly lower the quantity of molecules amplifiable with primer set β. (iii) Primer sets α, δ and ε disclose comparable amounts of molecules, about half as many molecules as those found at locus γ. The 2D NA analysis did not detect any initiation events in the regions containing primer sets α (Fig. 2A, fragment P, probe a), δ (Fig. 2A, fragment EV, probe d) and ε (Fig. 2A, fragments H2 and EV, probe e), indicating that replication initiation events are at least five times less frequent at these loci than at the region containing primer set γ. In striking contrast, the PCR technique disclosed only a 2-fold difference between locus γ and any of the three loci α, δ and ε, suggesting that this low bias could result from technical artifacts. Since BrdUrd-labeled DNA is prone to breakage following exposure to light or alkali, a fraction of the quantified molecules could correspond to DNA fragments generated during extraction. To verify this we used the simplified extraction procedure of Kumar *et al*., in which the steps of BrdUrd labeling and anti-BrdUrd chromatography are omitted (36).

The results of this second series of experiments, performed using unlabeled DNA molecules 0.6–1.6 kb in length, are reported in Figure 4B. As expected, molecules quantified at locus γ were now about four times more abundant than at locus α, δ or ε, and a significant decrease in the number of molecules was observed at locus β. This decrease at locus β is accounted for by the size of quantified molecules, since the quantification of preparations containing longer unlabeled molecules disclosed higher amounts of strands at locus β (data not shown). In order to demonstrate definitively that the peak of molecules detected with primer set γ corresponds to replication initiation events an experiment designed to rigorously compare the number of molecules present at each of the five loci in exponentially growing cells and in confluent cells was performed (see Materials and Methods). Figure 4C shows the results of quantification of unlabeled molecules 0.6–1.6 kb in length recovered from 6×10^7 exponentially growing cells or from exactly the same number of confluent arrested cells. With exponentially growing cells the results were similar to those reported in Figure 4B; a 5-fold enrichment of molecules was detected with primer set γ when compared to molecules quantified with primer sets α, δ and ε and an ~2.5-fold enrichment when compared with molecules amplifiable with primer set β. In striking contrast, a statistically equivalent number of molecules was found at all five loci in confluent cells, indicating that the peak of molecules observed at locus γ in exponentially growing cells corresponds to replication initiation. Furthermore, the number of molecules found at all loci in confluent cells, which most likely result from *in vivo* DNA repair and/or DNA fragmentation during extraction events, is not significantly different from the number of molecules found at loci α, δ and ε in exponentially growing cells. This suggests that if replication initiation events occur at these loci in growing cells, they are very infrequent. Thus, the results of PCR quantification and of 2D NA are in remarkable agreement and both support the existence of a narrow initiation zone at or close to locus γ.

DISCUSSION

Mapping of ori*GNAI3* by 2D NA

Using exponentially growing cells of line 42 we have analyzed the replication pattern of the *GNAI3–GNAT2* region using two experimental approaches. We first used the 2D NA gel electrophoresis technique, which permits determination of the direction of fork progression and is well adapted for screening of large DNA regions. Results indicate the existence of an initiation region that spans at most 1.7 kb within the 5.2 kb long *GNAI3–GNAT2* intergenic region (Fig. 2C).

Initiation events were detected neither in the unexpressed *GNAT2* gene nor in the analyzed 6 kb of the 3′-end of the *GNAI3*

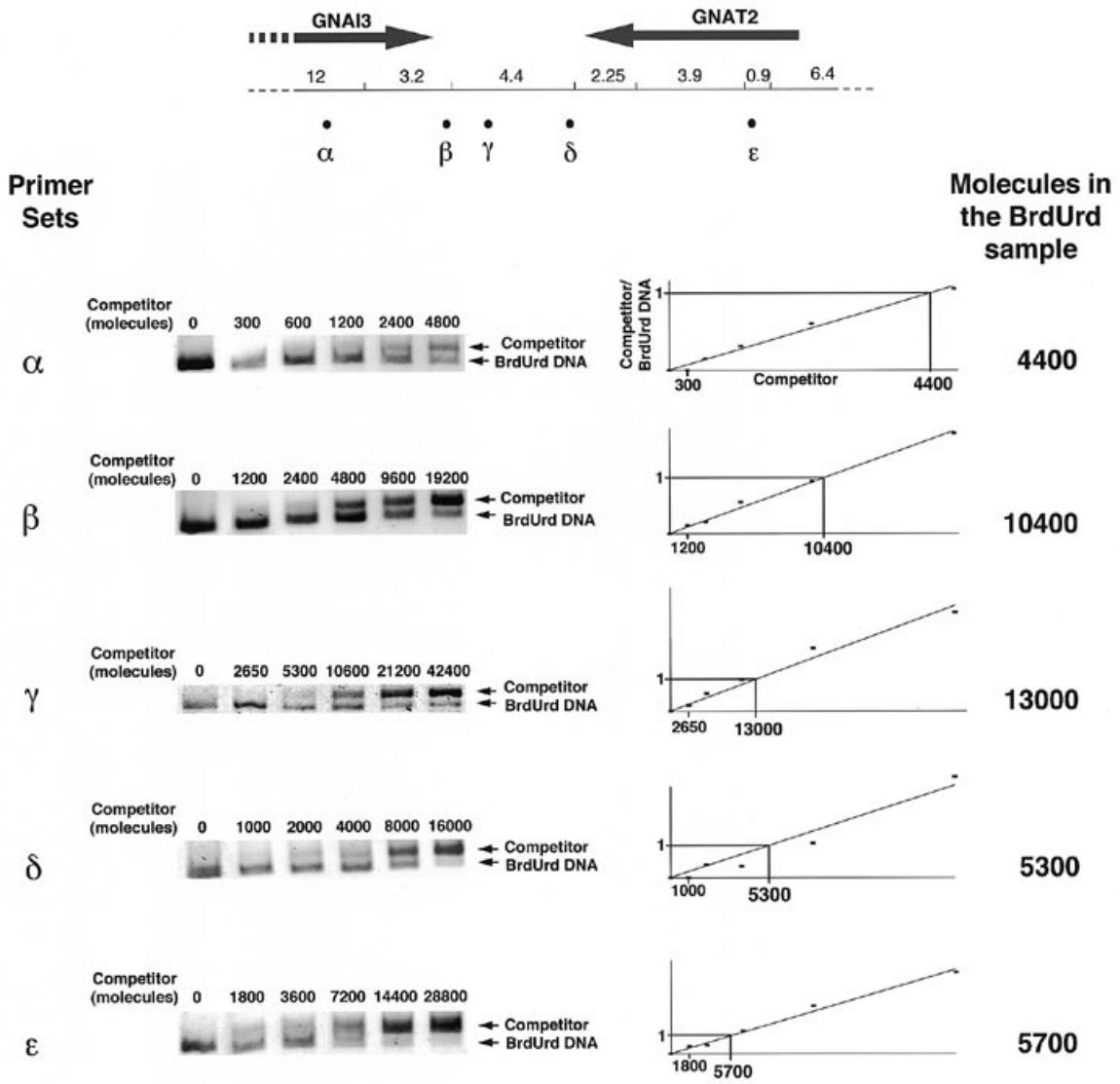

Figure 3. Example of quantification of nascent molecules in the *GNAI3–GNAT2* region by competitive PCR. (Top) Map of the region (as in Fig. 2). Black circles, location of the five primer sets used in this study. (Bottom) Results of a typical quantification experiment performed with a preparation of 1–2 kb long BrdUrd-labeled molecules. For each primer set mixtures of a fixed amount of BrdUrd-labeled molecules and of increasing quantities of molecules of the appropriate competitor were co-amplified. After electrophoresis the ratio between the upper (competitor) and lower (BrdUrd-labeled DNA) bands was plotted against the number of competitor molecules and the best fit line was drawn. The number of molecules in the BrdUrd sample was then deduced from the number of competitor molecules corresponding to 1:1 ratios with the BrdUrd-labeled DNA preparation. In this experiment a peak of molecules was observed at loci γ and β.

gene (Fig. 2B), but the replication patterns in these two genes were found to be very different; in the *GNAI3* gene replication forks travel in a single direction, while replication forks progress in both directions within the *GNAT2* gene. Two hypotheses may account for the latter situation. The first considers that the *GNAT2* gene is passively replicated and the existence of forks progressing in opposite directions results from the alternative use (in different cells or in different amplified units of the same cell) of two replication origins flanking this gene; one would be the initiation region defined here on the left and the other one on the right could be the same region in an adjacent amplified unit or another unidentified origin located within the same unit. The second hypothesis may be inferred from the results reported for the *DHFR* locus in Chinese hamster cells (13,27) and the rDNA locus in human cells (14). In both cases broad zones of replication initiation in which replication forks move in both directions have been observed; this was shown to result from alternative use of multiple potential replication start sites scattered across the broad zone. The *DHFR* and rDNA loci were studied either with cells containing hundreds of copies of the locus of interest or with unamplified but synchronized cells, allowing detection of all initiation events. In the study we present here, unsynchronized cells with only 30 copies of the amplified region were used and we may have revealed only the region firing most frequently. According to this hypothesis less frequent initiation events would have escaped detection, but the replication forks they generate would be detectable as arcs of replication forks moving in both directions. The 1.7 kb initiation region identified here would correspond to a preferred region of replication initiation within a broad zone where initiation events occur at least five times less frequently, as indicated by the signal-to-noise ratios in our experiments.

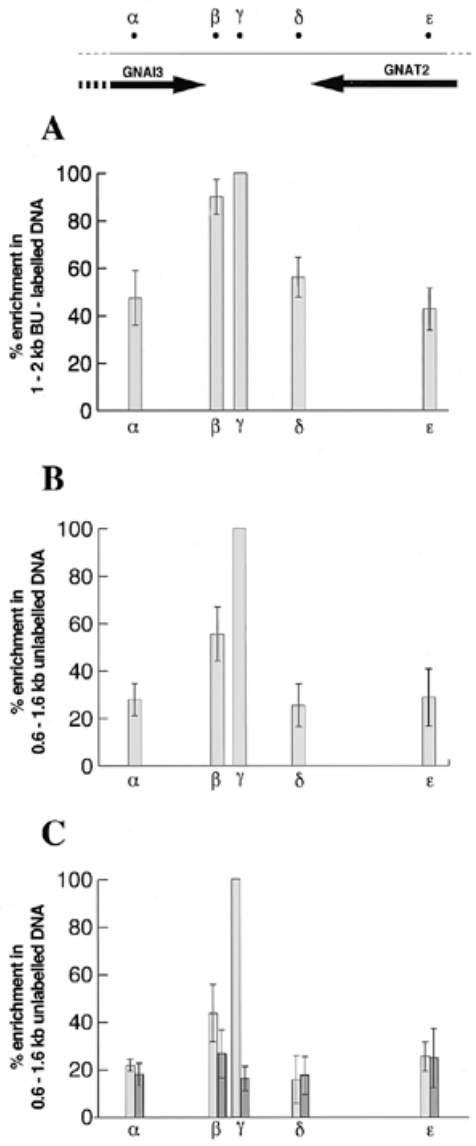

Figure 4. Summary of the PCR quantification analyses. (**A**) Quantification of 1–2 kb long BrdUrd-labeled molecules. Results are the average of quantification experiments from five independent BrdUrd (BU) DNA preparations, plotted against the map of the region (top). In order to eliminate fluctuations resulting from variable efficiencies in the recovery of BrdUrd-labeled molecules the results are expressed as a percentage of the number of molecules determined with primer set γ. Bars indicate confidence limits at the 0.95 significance level. (**B**) Quantification of 0.6–1.6 kb long unlabeled molecules. Represented as above, the results are the average of six quantification experiments from three independent DNA preparations (**C**) Comparison of exponentially growing and confluent cells. Unlabeled molecules, 0.6–1.6 kb in length, were recovered from 6×10^7 exponentially growing (light grey histograms) or confluent (dark grey histograms) cells. Each DNA preparation was quantified at least three times. Results are expressed as a percentage of the number of molecules found with primer set γ in exponentially growing cells.

In conclusion, the 2D NA results reported here identified either a narrow region of replication initiation or at least a narrow zone of preferential intiation within a broad zone. In any case, this 2D analysis argues in favor of the existence of preferred sites of replication initiation in mammalian cells.

Mapping of ori*GNAI3* by competitive PCR

The second approach we used to study replication initiation relies on PCR quantification of short newly synthesized strands recovered from exponentially growing cells. DNA preparations were obtained either after BrdUrd pulse labeling, size separation through a sucrose gradient and anti-BrdUrd affinity chromatography according to Giacca *et al*. (3) or by size separation alone, according to the simplified procedure of Kumar *et al*. (36). Both extraction techniques allowed detection of a peak of molecules with primer set γ, the only one that maps within the preferential initiation zone defined by 2D NA. Moreover, the signal-to-noise ratio improved strikingly when we used the simplified technique, probably because the studied region is highly AT-rich (40) and prone to breakage when labeled with BrdUrd. Comparison of the results obtained with exponentially growing or confluent cells shows that the peak of short molecules is specific to replicating cells. This confirmed both the existence of the initiation region and its location.

Furthermore, the PCR results reported here may help to refine mapping of the initiation zone. The first PCR-based replicon mapping technique, developed by Vassilev and Johnson, relied on the use of nascent strands of increasing length for PCR amplification with primer sets distributed along a region of interest. The rationale of this technique is that if initiation events occur in the region, then lengthening of nascent strands will progressively include the loci amplifiable with the different primer sets. Thus the locus closest to the initiation zone is amplifiable with the shortest nascent molecules (9,41). We analyzed the *GNAI3–GNAT2* region with preparations of nascent molecules that were either 1–2 kb or 0.6–1.6 kb in length. While similar amounts of molecules were found with primer sets β and γ in preparations of 1–2 kb long DNA strands (Figs 3 and 4A), the number of strands amplifiable with primer set β was significantly lowered upon quantification of smaller molecules (Fig. 4B and C). This indicates that locus γ lies within or close to the initiation region, whereas locus β is outside this region. According to 2D NA results, which mapped the initiation zone some 700 bp from primer set β (see Materials and Methods for the exact positions of primer sets and probes), and assuming a bidirectional propagation of replication forks, nascent molecules emanating from the initiation region would be amplifiable with primer set β only when their size exceeds 1.4 kb. Such nascent molecules are expected to be more abundant in preparations of strands that are 1–2 kb in length, rather than in 0.6–1.6 kb long strands. Thus, PCR results with primer set β agree remarkably with mapping of the initiation region by 2D NA. Moreover, detection of similar amounts of 1–2 kb long molecules with primer sets β and γ suggests that most intiation events occur <1 kb from primer set β. In conclusion, when PCR and 2D NA data are taken together most initiation events appear to occur in a very narrow region of ~300 bp, located 700 bp to the right of primer set β.

Comparison with other mammalian initiation zones analyzed by 2D gels

A remarkable conclusion of this study is that a narrow zone of active replication initiation, referred to as ori*GNAI3*, has been identified both by a 2D gel electrophoresis and a competitive PCR technique. This differs from the data obtained for the Chinese hamster *DHFR* and human rDNA loci, in which broad zones of initiation, at least 30 kb in length, were disclosed by 2D gels (13,14), whereas narrow regions were observed with other replicon mapping techniques (5,6,8,16–19).

In the case of the *DHFR* locus, the conflicting results of the different approaches are difficult to reconcile. Two sites of preferential replication initiation, termed oriβ and oriγ, were identified by several techniques in the *DHFR–2BE2121* intergenic region. Repeated analyses of this locus by 2D gel electrophoresis showed that initiation events occur at multiple sites in the 55 kb long intergenic region and in the *2BE2121* gene (15). Although a 30 kb long central region containing both oriβ and oriγ was reported to be preferred (22), oriβ itself did not appear as a preferential site of replication initiation when analyzed by 2D neutral–neutral (NN) (22) or by a 3D gel technique that combines 2D NN and 2D NA procedures (21). In this regard ori*GNAI3* and oriβ appear to be different. Whether or not this reflects differences in the structure or regulation of these origins remains to be determined.

On the contrary, the differences between ori*GNAI3* and the origin at the human rDNA locus may be less pronounced than may appear at first. Indeed, 2D NN gel analysis of the human rDNA locus has suggested that replication initiation occurs preferentially in a limited part of the 31 kb long initiation zone. While precise mapping of this preferential region was not attempted, Little *et al*. observed that the strongest bubble arc signals were located in an ~8 kb long region of the broad initiation zone (14). Interestingly, alternative techniques later mapped a 1.5 kb long (17) or a 4 kb long (18) initiation region within the preferential region identified by 2D gels (14). Thus ori*GNAI3* and the origin at the rDNA locus appear to share similar properties, if one considers either that infrequent initiation events were below the threshold of detection in our 2D NA study or that the degree of preferential initiation is higher at the initation region described here.

Conclusion

In this work a narrow zone of preferential replication initiation was identified and mapped by 2D NA gel electrophoresis near the 3′-end of the Chinese hamster *GNAI3* gene. The existence and location of this initiation region, termed ori*GNAI3*, was confirmed by competitive PCR. Our results support the notion that preferred sites of replication initiation do exist in mammalian genomes and that they may be identified by 2D gels in some cases. Furthermore, since 2D gels and competitive PCR here disclosed a narrow initiation region, ori*GNAI3* appears as a new attractive model system to identify the distribution of sequences involved in mammalian DNA replication initiation.

ACKNOWLEDGEMENTS

We would like to thank Drs M.C.Weiss and E.Heard for critical reading of the manuscript. This work was supported in part by the Université Pierre et Marie Curie, the Ligue Nationale Française contre le Cancer (Comité de Paris) and the Association pour la Recherche sur le Cancer.

REFERENCES

1 Jacob,F., Brenner,S. and Cuzin,F. (1963) *Cold Spring Harbor Symp. Quant. Biol*., **28**, 329–348.
2 Vassilev,L.T. and De Pamphilis,M.L. (1992) *Crit. Rev. Biochem. Mol. Biol*., **27**, 445–472.
3 Giacca,M., Zentilin,L., Norio,P., Diviacco,S., Dimitrova,D., Contreas,G., Biamonti,G., Perini,G., Weighardt,F., Riva,S. and Falaschi,A. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 7119–7123.
4 Yoon,Y., Sanchez,A., Brun,C. and Huberman,J.A. (1995) *Mol. Cell. Biol*., **15**, 2482–2489.
5 Anachkova,B. and Hamlin,J.L. (1989) *Mol. Cell. Biol*., **9**, 532–540.
6 Handeli,S., Klar,A., Meuth,M. and Cedar,H. (1989) *Cell*, **57**, 909–920.
7 Burhans,W.C., Vassilev,L.T., Wu,J., Sogo,J.M., Nallaseth,F. and DePamphilis,M.L. (1991) *EMBO J*., **10**, 4351–4360.
8 Burhans,W.C., Vassilev,L., Caddle,M.S., Heintz,N.H. and DePamphilis,M.L. (1990) *Cell*, **62**, 955–965.
9 Vassilev,L. and Johnson,E.M. (1989) *Nucleic Acids Res*., **17**, 7693–7705.
10 DePamphilis,M.L. (1996) In DePamphilis,M.L. (ed.), *DNA Replication in Eukaryotic Cells*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 45–86.
11 Brewer,B.J. and Fangman,W.L. (1987) *Cell*, **51**, 463–471.
12 Nawotka,K.A. and Huberman,J.A. (1988) *Mol. Cell. Biol*., **8**, 1408–1413.
13 Vaughn,J.P., Dijkwel,P.A. and Hamlin,J.L. (1990) *Cell*, **61**, 1075–1087.
14 Little,R.D., Platt,T.H.K. and Schildkraut,C.L. (1993) *Mol. Cell. Biol*., **13**, 6600–6613.
15 Dijkwel,P.A., Vaughn,J.P. and Hamlin,J.L. (1994) *Nucleic Acids Res*., **22**, 4989–4996.
16 Pelizon,C., Diviacco,S., Falaschi,A. and Giacca,M. (1996) *Mol. Cell. Biol*., **16**, 5358–5364.
17 Coffman,F.D., Georgoff,I., Fresa,K.L., Sylvester,J., Gonzalez,I. and Cohen,S. (1993) *J. Cell Biochem*., **51**, 157–164.
18 Gencheva,M., Anachkova,B. and Russev,G. (1996) *J. Biol. Chem*., **271**, 2608–2614.
19 Scott,R.S., YoungTruong,K. and Vos,J.M.H. (1997) *Nucleic Acids Res*., **25**, 4505–4512.
20 Burhans,W.C. and Huberman,J.A. (1994) *Science*, **263**, 639–640.
21 Kalejta,R.F., Lin,H.B., Dijkwel,P.A. and Hamlin,J.L. (1996) *Mol. Cell. Biol*., **16**, 4923–4931.
22 Dijkwel,P.A. and Hamlin,J.L. (1992) *Mol. Cell. Biol*., **12**, 3715–3722.
23 Debatisse,M., Toledo,F., Robert de Saint Vincent,B. and Buttin,G. (1992) In Kellems,R.E. (ed.), *Gene Amplification in Mammalian Cells—A Comprehensive Guide*. Marcel Dekker, New York, NY, pp. 173–183.
24 Toledo,F., LeRoscouet,D., Buttin,G. and Debatisse,M. (1992) *EMBO J*., **11**, 2665–2673.
25 Toledo,F., Buttin,G. and Debatisse,M. (1993) *Curr. Biol*., **3**, 255–264.
26 Coquelle,A., Pipiras,E., Toledo,F., Buttin,G. and Debatisse,M. (1997) *Cell*, **89**, 215–225.
27 Dijkwel,P.A. and Hamlin,J.L. (1995) *Mol. Cell. Biol*., **15**, 3023–3031.
28 Debatisse,M., Hyrien,O., Petit-Koskas,E., Robert de Saint Vincent,B. and Buttin,G. (1986) *Mol. Cell. Biol*., **6**, 1776–1781.
29 Toledo,F., Smith,K.A., Buttin,G. and Debatisse,M. (1992) *Mutat. Res*., **276**, 261–273.
30 Baron,B., Fernandez,M.A., Carignon,S., Toledo,F., Buttin,G. and Debatisse,M. (1996) *Mamm. Genome*, **7**, 429–432.
31 Debatisse,M., Robert de Saint Vincent,B. and Buttin,G. (1984) *EMBO J*., **3**, 3123–3127.
32 Dijkwel,P.A., Vaughn,J.P. and Hamlin,J.L. (1991) *Mol. Cell. Biol*., **11**, 3850–3859.
33 Biamonti,G., Perini,G., Weighardt,F., Riva,S., Giacca,M., Norio,P., Zentilin,L., Diviacco,S., Dimitrova,D. and Falaschi,A. (1992) *Chromosoma*, **102**, S24–S31.
34 Contreas,G., Giacca,M. and Falaschi,A. (1987) *BioTechniques*, **12**, 824–826.
35 Diviacco,S., Norio,P., Zentilin,L., Menzo,S., Clementi,M., Biamonti,G., Riva,S., Falaschi,A. and Giacca,M. (1992) *Gene*, **122**, 313–320.
36 Kumar,S., Giacca,M., Norio,P., Biamonti,G., Riva,S. and Falaschi,A. (1996) *Nucleic Acids Res*, **24**, 3289–3294.
37 Baron,B., Fernandez,M.A., Toledo,F., LeRoscouet,D., Mayau,V., Martin,N., Buttin,G. and Debatisse,M. (1994) *Genomics*, **24**, 288–294.
38 Hyrien,O., Debatisse,M., Buttin,G. and Robert de Saint Vincent,B. (1987) *EMBO J*., **6**, 2401–2408.
39 Baron,B., Fernandez,M.A., Carignon,S., Toledo,F., Buttin,G. and Debatisse,M. (1996) *Mamm. Genome*, **7**, 922–923.
40 Fernandez,M.A., Baron,B., Prigent,M., Toledo,F., Buttin,G. and Debatisse,M. (1997) *J. Cell Biochem*., **67**, 541–551.
41 Vassilev,L. and Johnson,E.M. (1990) *Mol. Cell. Biol*., **10**, 4899–4904.