© 1991 Oxford University Press

Elements which stimulate gene amplification in mammalian cells: role of recombinogenic sequences/structures and transcriptional activation

James G.McArthur+, Lenore K.Beitel, John W.Chamberlain§ and Clifford P.Stanners*
Department of Biochemistry and Cancer Centre, McGill University, 3655 Drummond St., Montreal, Quebec H3G 1Y6, Canada

Received August 16, 1990; Revised and Accepted March 21, 1991

EMBL accession no. X51544

ABSTRACT

HSAG-1 is a 3.4 kb mammalian genomic element which has been shown to stimulate the amplification of the pSV_2DHFR expression vector *in cis* when transfected into a variety of cell lines (1). This amplification stimulatory activity requires the interaction of multiple positive acting elements that include sequence features associated with recombination 'hotspots', such as Alu-like repetitive sequences and A/T rich regions (2). We demonstrate here that two other members of the HSAG family of elements, HSAG-2 and HSAG-5, also stimulate vector amplification. By analysis of the HSAG-2 nucleotide sequence and of the amplification activity of HSAG-2 and HSAG-5 subfragments, we show that this activity also involves the interaction of multiple positive acting elements. The autonomous replication of the HSAG containing vectors is not responsible for this effect. We also show that the orientation of HSAG elements in pSV_2DHFR has a profound effect on their amplification stimulatory activity, and present evidence that the transcription of these elements in pSV_2DHFR is necessary for the effect.

INTRODUCTION

Gene amplification, an increase in the copy number of genes, is a strictly controlled process in certain multicellular organisms (for review see 3) and, although rare in normal mammalian cells (4,5), occurs with frequencies between 10^{-3} and 10^{-7} in transformed cells (4,6). A number of different mechanisms for gene amplification have been proposed (for review see 7), which may be affected by nucleotide sequences which are *in trans* (8,9) as well as *in cis* (10–15). Direct evidence for *cis* acting elements in mammalian cells include a fragment of DNA that induces amplification in response to 5-/bromo-deoxyuridine (16), plasmid amplification promoting sequences isolated from mouse rDNA (17), and the HSAG family of sequences, the subject of this paper.

HSAG-1, a 3.4 kb genomic element, cloned from a human CLL-CHO (chronic lymphocytic leukemia-Chinese hamster ovary) hybrid cell genomic library (18), was subcloned into pSV_2DHFR (19) and found to stimulate vector amplification following transfection into a variety of mammalian cell lines (1). This activity required the interaction of multiple positive acting elements (1,2). In order to further investigate the question of the mechanism of this phenomenon, we decided to examine the generality of the sequence features found in HSAG-1 in other active elements.

HSAG-1 is repeated 20–50 times in the CHO genome, and is a member of a family of middle repetitive elements with a copy number in the thousands (20). This paper demonstrates that two other members of the HSAG family stimulate gene amplification of the pSV_2DHFR vector: HSAG-2, a 9.3 kb element, which like HSAG-1, was cloned from a human CLL-CHO hybrid cell genomic library (18), and HSAG-5, a 7.8 kb element cloned from a human CLL genomic library (21). We demonstrate the requirement, as with HSAG-1, for the interaction of multiple positive acting elements in HSAG-2 and HSAG-5 and suggest that the transcription of the HSAG insert is correlated with the amplification stimulatory activity of HSAG-1 in pSV_2DHFR.

MATERIALS AND METHODS

pSV_2DHFR constructs

HSAG-1, HSAG-2 and HSAG-5 (18,21) and their subfragments were subcloned in the same orientation, unless otherwise noted, into the EcoRI site of pSV_2DHFR (19) (Fig. 1) using standard techniques (22). The pSV_2DHFR vector will be referred to as D; HSAG-1 in the 'left-hand' orientation in pSV_2DHFR as HLD (Fig. 1); and HSAG-1 in the 'right-hand' orientation as HRD (Fig. 1).

Cell culture, transfection and survival curve assay

The culture and transfection of the CHO cell line, LR73, (23) was carried out as described previously (1). Briefly, LR73 cells were co-transfected by the calcium phosphate precipitation technique (24) with 0.5 μg of the dominant selectable marker,

* To whom correspondence should be addressed
Present addresses: +Addition, Life Sciences, University of California in Berkeley, Berkeley, CA 94720, USA and §Department of Surgical Research, Hospital for Sick Children, Toronto, Ontario, Canada M5G 1X8

pSV$_2$AS (25), 5 μg of the pSV$_2$DHFR test construct and 10 μg of genomic carrier DNA. Transfectants were selected for asparagine synthetase (AS) expression in asparagine-free medium containing 2 mM albizziin, an inhibitor of the mammalian AS gene. The resulting transfectant colonies were trypsinized, pooled, grown for several days in non-selective nucleoside-free medium, and then plated in selective media containing different concentrations of methotrexate (MTX); the resulting colonies were fixed, stained, and counted microscopically. The frequency of surviving colonies was normalized to the frequency of colonies transfected with the parental vector, scaled according to a value of 50 given to the HLD construct, and averaged for two different MTX concentrations (0.6 and 2.0 μM); amplification activities for a minimum of two separate experiments were averaged. HLD usually gave frequencies of colony survival 50 times greater than D; this procedure allowed comparison between sets of transfections where the HLD normalized frequency differed significantly from this value.

Nucleotide sequencing and computer analysis

The sequence presented was compiled from sequencing both strands or multiple sequencing of the same strand using the dicleoxynucleotide chain termination method of Sanger et al. (26).

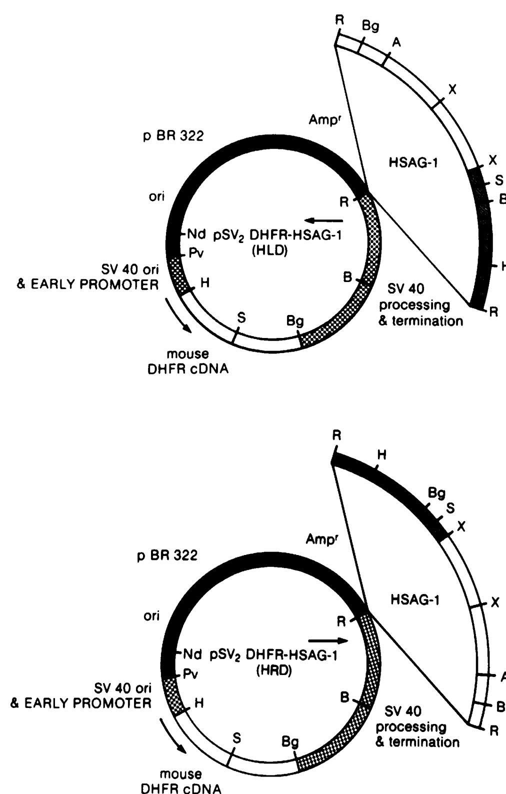
Sequence analysis was carried out using DNASIS (Hitachi Software Engineering Co. Ltd.) and standard computer programs for nucleic sequence analysis adapted for use by the Clinical Research Institute of Montreal. The ANALYSEQ (27) and DIAGON (28) programs were used extensively in this analysis.

RESULTS

We have previously shown that HSAG-1 promotes pSV$_2$DHFR vector amplification *in cis* (1). To determine if other members of the HSAG family of middle repetitive elements had similar activity, HSAG-2 (18) and HSAG-5 (21), were subcloned into the unique EcoRI site of pSV$_2$DHFR (Fig. 1) (19) and their amplification activity assayed using a methotrexate (MTX) survival curve assay (1). It can be seen in Figure 2 that LR73 cell populations transfected with the HSAG-2 and HSAG-5 containing vectors, as with HLD, had an increased frequency of MTX resistant cells compared to cells transfected with the parental D vector. Previous work has shown that this increased MTX resistance results from a higher frequency of cells with elevated copy numbers of the DHFR containing vector (1). Five of five HSAG family members tested to date including HSAG-1 (1), HSAG-2 and HSAG-5 (Fig. 2), and LRH-1 and LRH-2 (2) markedly stimulate gene amplification, as opposed to one of 24 random CHO genomic clones (2).

HSAG-1 and HSAG-2 sequence analysis

A computer assisted analysis of the nucleotide sequences of HSAG-2 (Fig. 3, EMBL Accession No. X51544) and HSAG-1 (29, EMBL Accession No. X03822) was conducted. HSAG family members are known to possess clusters of diverged Alu-like elements (20) and HSAG-1 contains 12 Alu-like sequences within 3.4 kb (29); these are less than 80% homologous to the corresponding Alu consensus sequences (30). HSAG-2 was found to have 9 diverged Alu-like elements within 9.3 kb (Fig. 3), including a CHO Type I Alu at position 8434 which has apparently transposed into a CHO Type II Alu with the generation of flanking 10 bp direct repeats. Analysis of sequenced portions of HSAG-5 (unpublished data), showed that this element has two



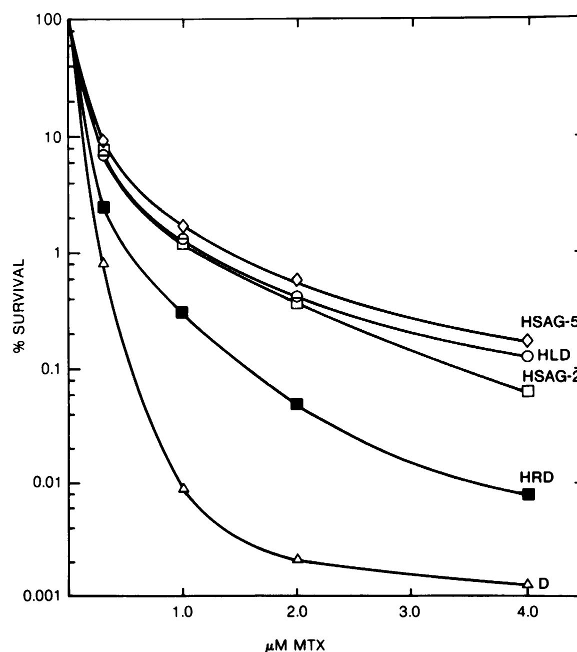


Figure 1. Map of pSV$_2$DHFR-HSAG-1 in the 'lefthand' (HLD) and 'righthand' (HRD) orientations. The highly active 1.45 kb EcoRI-XbaI fragment of HSAG-1 is indicated by hatching. Restriction enzyme sites are: R, EcoRI; Bg, BglII; A, AvaI; X, XbaI; S, SstI; H, HindIII; B, BamHI; Pv, PvuII; Nd, NdeI.

Figure 2. Semilogarithmic plot of the percentage of surviving colonies of pSV$_2$DHFR-HSAG-5 (◇), pSV$_2$DHFR-HSAG-2 (□), HLD (○), HRD (■), and D (△)—transfected LR73 cells plated in media containing different concentrations of MTX. See Materials and Methods for experimental details.

Figure 3. Nucleotide sequence of HSAG-2 (EMBL Accession No. X51544). Certain restriction sites, repetitive sequences, and purine-pyrimidine tracts are indicated. Sequences related to hamster type I Alu consensus sequence are shaded, hamster type II Alu consensus sequences are shaded and outlined by dotted lines, and purine-pyrimidine tracts are boxed. The orientation of the Alu-like elements compared to the 5′ to 3′ orientation of the corresponding consensus sequence is indicated by the large underlying black arrows; direct repeats are indicated by overlying dashed arrows.

Alu-like elements in the 1.2 kb at its 5′ end, confirming that a common feature of HSAG family members is clusters of Alu-like elements.

HSAG-1 and HSAG-2 also have in common long alternating adenine-thymidine sequences [$d(AT)_n$ tracts] of 54 and 46 nucleotides at positions 3204 and 7085, respectively. Unique to HSAG-2 however, are two other alternating purine-pyrimidine tracts: $d(AC)_{39}$ at position 2633 [actual sequence: $d(AC)_{21}TC(AC)_6TC(AC)_6TC(AC)_3$] and $d(GT)_{22}$ at position 5952.

A comparison of the nucleotide sequences of HSAG-1 and HSAG-2 using OVRLAP and DIAGON programs (28) found no extensive homology aside from Alu-like elements and $d(AT)_n$ tracts, using a criterion of 12 bp with 1 mismatch or 10 bp with no mismatches (data not shown). Comparison of HSAG-1 and HSAG-2 with GENBANK did not reveal significant homologies to other sequences apart from those with repetitive elements or A/T rich regions (data not shown). Further analysis of HSAG-1 and HSAG-2 with the ANALYSEQ programs (27) found these sequences to have numerous stop codons in all possible reading frames, such that the longest open reading frames were 465 and 372 bp respectively; when subjected to tests for protein coding capability using several algorithms, including Shepherd's (31) and Fickett's (32), these open reading frames did not rank as high as true protein coding cDNA sequences. This low protein coding potential, coupled with the lack of sequence homology between HSAG-1 and HSAG-2, made it unlikely that the stimulation of gene amplification by these elements is mediated by a common polypeptide coded for directly by the HSAG elements themselves.

Activity of HSAG-1, HSAG-2, and HSAG-5 subfragments

To identify the sequence features required for the amplification stimulatory activity, various fragments of HSAG-1, HSAG-2, and HSAG-5 were subcloned into the EcoRI site of pSV_2DHFR, transfected into LR73 cells, and the activity of the resulting transfectants determined with the MTX survival curve assay. The results of this analysis are shown in Figure 4. Like HSAG-1 (top of Figure 4; ref. 2), the amplification activity of HSAG-2 and HSAG-5 is not all or none but graded in magnitude, with most subfragments having a fraction of the activity of the parent element. Detailed analysis of HSAG-1 subfragments and deletion mutants of the highly active HSAG-1 EcoRI-XbaI fragment (H1 RX) led to the suggestion that multiple positive acting elements were involved in the amplification effect (2). Similarly, more than one sequence or structural feature appears to be necessary for maximum activity of HSAG-2 and HSAG-5. One example of this is the H5 RH and H5 HR fragments which have 15% or less of the activity of HSAG-5 individually, yet combined give HSAG-5 an activity greater than that of HSAG-1. These multiple positive elements also appear to be redundant as both H2 RPv, the left end of HSAG-2, and H2 PvR, the right end, are as active as the entire HSAG-2 for amplification stimulatory activity. The presence of elements that negatively affect activity may also be postulated, as the EcoRI-XbaI subfragment of HSAG-1 (H1 RX) and the Sst-Sst subfragment of HSAG-2 (H2 SS) were more active than the complete HSAG-1 and HSAG-2 elements respectively; the additional sequences appear to dampen activity.

Role of Alu-like elements

Alu-like sequences were found in all HSAG-1, HSAG-2, and HSAG-5 fragments with an amplification activity greater than

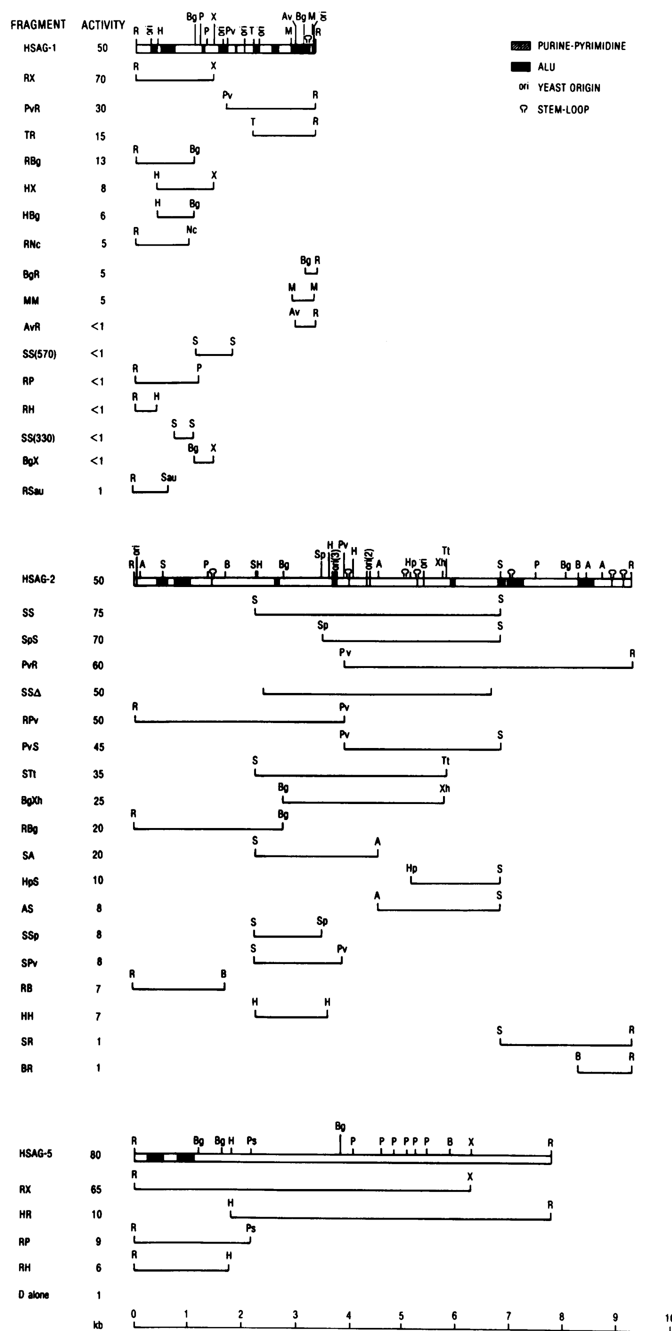


Figure 4. Amplification activity of HSAG-1, HSAG-2, and HSAG-5 subfragments as determined by the MTX survival curve assay. At the top of each set of subfragments is the restriction map (not all sites are shown) of the corresponding HSAG element with the positions of Alu-like elements (black boxes), purine-pyrimidine tracts (shaded boxes), stem-loop structures (with a minimum of 80% stem homology), and yeast ARS consensus sequences (ori) indicated. The HSAG-2 SS fragment lacking the 30 bp of Alu-like element is indicated as SSΔ. The restriction sites are: A, AccI; Av, AvaI; B, BamHI; Bg, BglII; H, HindIII; Hp, HpaI; M, MspI; Nc, NcoI; Nd, NdeI; Ps, PstI; Pv, PvuII; R, EcoRI; S, SstI; Sau, Sau3A; Sp, SphI; T, TaqI; Tt, Tth111I; X, XbaI; Xh, XhoI.

20 (except H2 BgXh), although fragments containing them were not necessarily active e.g. H2 SR. While all highly active HSAG-1 subfragments contain numerous Alu-like elements, HSAG-2 had two large active fragments, H2 SS and H2 PvS, that possess only 30 bp of the CHO Type II Alu that straddles

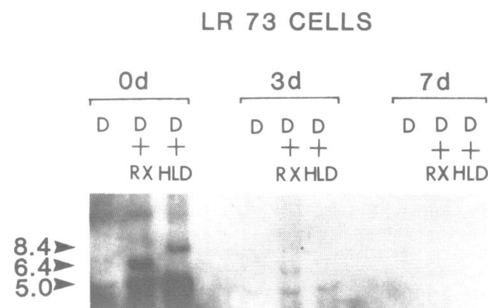


Figure 5. Assay for autonomous replication in LR73 cells. Hirt supernatants (54) were prepared from cells transfected with equivalent quantities of D and H1 RXD, D and HLD, or D alone. 18 hours (0d), 3 days (3d), and 7 days (7d) following DNA exposure, the DNA was digested with BamHI (which cuts each plasmid at a unique site) and DpnI (which cuts plasmids that have not replicated (39) in the LR73 cells), and analyzed by the Southern blot technique with a ^{32}P-labelled (55) DHFR probe. The positions of the unit length 5.0 kb D, 6.4 kb RXD, and 8.4 kb HLD bands are indicated.

the SstI site (6885 bp). Removal of 65 bp from the right hand end of H2SS, including this Alu-like element, reduced activity 30%; however, the resulting construct, SSΔ, was still as active as the full length HSAG-2 element (Fig. 4). Previous work has shown that the presence of repetitive sequences in a series of random CHO genomic clones was insufficient for amplification activity (1), and that BLUR-8, the prototype human Alu element (33), in pSV$_2$DHFR, was itself inactive (unpublished results). These results and the high activity of H2 SSΔ indicates that, while Alu-like elements may increase the amplification stimulatory activity of HSAG elements, they are neither sufficient nor essential for this activity.

Role of purine-pyrimidine tracts

Both HSAG-1 and HSAG-2 possess large d(AT)$_n$ tracts (Figs. 2, 3); the d(AT)$_n$ tract of HSAG-1 (3204 bp) was implicated in the activity of the diminutive H1 BgR fragment (2). The d(AT)$_n$ tract (7085 bp) of HSAG-2, however, is located outside of the very active H2 SS fragment, and is present in the inactive H2 SR fragment. At least one of the two other long alternating purine-pyrimidine tracts in HSAG-2, d(AC)$_{39}$ (2633 bp) and d(GT)$_{22}$ (5952 bp), is present in all highly active fragments of HSAG-2, except for H2 BgXh. The presence of a d(GT)$_n$ tract however, does not assure activity (see H2 SPv and H2 AS). Like Alu-like elements, purine-pyrimidine tracts thus appear to be neither sufficient nor essential for activity.

Role of stem loop structures

Potential hairpins or stem-loop structures have been found in the vicinity of novel joints formed during gene amplification (34,35). HSAG-1 possesses two stem-loops with a minimum 12 bp stem and 80% stem homology (Fig. 4), just before (centered at position 3186) and at the d(AT)$_n$ tract (3204 bp) while HSAG-2 has 7 distributed throughout its length. Stem-loop structures, however, do not occur with a greater frequency in active, as opposed to inactive HSAG-2 subfragments (compare the inactive H2 SR with 3 stem-loops to the larger active H2 SS fragment also with 3 stem-loops).

A/T-rich regions

HSAG-2 possesses a number of regions which are greater than 70% A/T and over 30 bp long distributed throughout its length

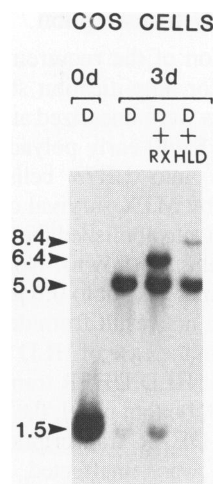


Figure 6. Assay for autonomous replication in COS cells (56). Procedures are indicated in the legend of Fig. 5. The positions of the unit length 5.0 kb D, 6.4 kb RXD, and 8.4 kb HLD DpnI insensitive bands, and 1.5 kb DHFR-pBR322 DpnI fragment are indicated. The latter band decreases in intensity with time as the vectors replicate and are demethylated becoming DpnI insensitive.

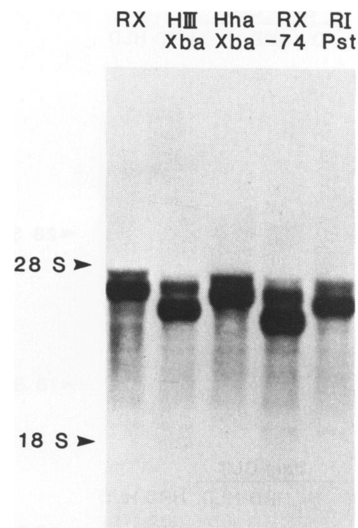


Figure 7. Transcription of various HSAG-1 subfragments subcloned into pSV_2DHFR. LR73 cells were transfected with the EcoRI-XbaI (RX), HindIII-XbaI (HIII Xba), HhaI-XbaI (Hha Xba), EcoRI-XbaI with a 74 bp deletion at the HindIII site (RX-74) and EcoI-PstI (RI Pst) fragments of HSAG-1 in pSV_2DHFR as described in Materials and Methods. Total RNA from transfectants resistant to 0.6 μM MTX was analyzed by the Northern technique (57) and probed with a ^{32}P-labelled HSAG-1 specific fragment (SS330). The positions of the 18S and 28S rRNA are indicated. The locations of the fragments in HSAG-1 may be seen in Figure 4; the HhaI site (not shown) is at 171 bp.

(Fig. 3). Again the presence of A/T-rich DNA alone is not sufficient for amplification activity; for example, H2 SR contains a 92 bp segment of DNA which is 88% A/T and is inactive. This, however, does not preclude a role for A/T-rich regions in stimulating the amplification activity of HSAG-2; thus the addition of a 297 bp sequence which is 76% A/T to H2 AS with an activity of 8 produces the H2 PvS fragment which has an activity of 45. Further extension of the H2 PvS sequence to the SphI site (3521) adds a 92 bp region which is 75% A/T and increases the activity to 70, the level of H2 SS.

Autonomous replication of HSAG containing vectors

HSAG-1 and HSAG-2 were found to contain 5 and 7 regions respectively (Fig. 4) with a minimum 10/11 match to the yeast ARS (autonomously replicating sequence, 36) consensus (A/T)TTTAT(A/G)TTT(A/T). Although the role of such sequences in mammalian DNA replication has not been established, they have been found in 7 of 8 monkey origin rich sequences (ors) (37) and recently in the initiation region of the amplified DHFR replicon from CHO400 cells (38). The DpnI sensitivity assay (39) was used to determine directly if HLD or RXD plasmids replicated autonomously in LR73 cells. As an internal control, the cells were cotransfected with the D plasmid, which lacks amplification stimulatory sequences. At day 0 (Fig. 5) bands corresponding to the unit length 5.0 kb D, 6.4 kb RX, and 8.4 kb HLD vectors were seen. The vast majority of transfected plasmid DNA was not replicated and was digested by DpnI, migrating lower on the gel. Unexpectedly, a significant fraction of the D vector DNA was DpnI resistant at day O (18 hours after transfection). This DpnI insensitivity cannot be accredited to vector replication in LR73 cells, however, as D does not possess an origin functional in mammalian cells which do not contain SV40 T antigen. Furthermore, the decreasing quantity of unit length D, RXD, and HLD plasmid bands with time of incubation (days 3 and 7) indicates that only a negligible amount of plasmid replicated autonomously and that none of these plasmids possessed origin sequences functional in LR73 cells. While the plasmid extraction efficiencies varied with each preparation, equivalent weights of D and the test plasmid were used in each transfection. At each time point the quantity of HLD or RXD was lower than D indicating that, while D did not replicate well in LR73 cells, HSAG-1 and the H1 RX fragment did not boost this low level of autonomous replication.

Autonomous replication was also tested in SV40 T antigen-producing COS cells (Fig. 6). While the HLD and RXD constructs replicated well due to the presence of the SV40 origin present in pSV_2DHFR, they did not replicate as well as the parental vector D. These experiments suggest that HSAG-1 and H1 RX do not possess origins of replication functional in LR73 cells nor do they stimulate the SV40 origin of replication in COS cells. The potential effects of HSAG sequences integrated into the genome on replication *in vivo*, however, remain to be investigated.

Orientation dependence of HSAG sequences and transcription

The orientation of the HSAG insert in the EcoRI site of pSV_2DHFR has a profound effect on the activity of a fragment. All experiments described above were done with the inserts in the 'left-hand' orientation (as in HLD, Fig. 1). When HSAG-1 is in the 'right-hand' (HRD) orientation (Fig. 1), the activity was half that of HLD, as shown in the MTX survival curve of Figure 2. This orientation dependence of the insert was also observed with all highly active fragments of HSAG-2 and HSAG-5 tested, without exception (data not shown). A possible explanation for the orientation dependence is the involvement of strand specific transcription.

To investigate this question, the nature and quantity of HSAG homologous transcripts in HLD and HRD transfected cells was analyzed. The equal level of DHFR transcripts in total RNA from HRD and HLD transfectants resistant to 0.6 μM MTX (Fig. 8; lower panel; HRD and HLD lanes at right) indicates that the

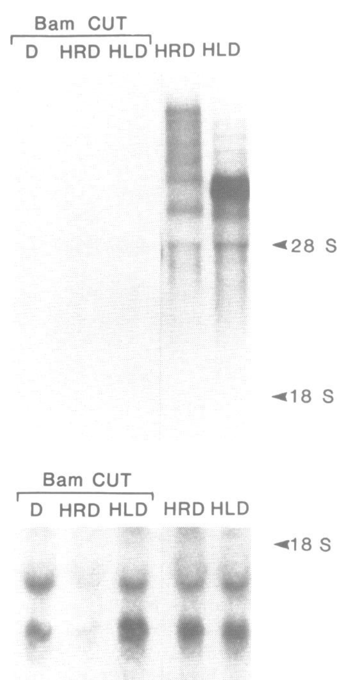


Figure 8. Effect on HSAG-1 transcription of linearizing the HRD and HLD vectors with BamHI. Total RNA from LR73 cells transfected with BamHI cut D, HRD, or HLD vectors or uncut HRD or HLD vectors, resistant to 0.6 μM MTX was analyzed by the Northern technique and probed with a ^{32}P-labelled HSAG-1 specific fragment (SS330) (upper panel) or DHFR cDNA (lower panel). The positions of the 18S and 28S rRNA are indicated.

difference in the frequency of HRD relative to HLD MTX-resistant transfectants cannot be attributed to a difference in the level of DHFR transcription in these constructs. The two DHFR transcripts observed result from the use of alternative vector polyadenylation signals. A striking difference between the two plasmids in the quantity and nature of HSAG-1 homologous transcripts, however, was observed (Fig. 8, upper panel, HRD and HLD lanes at right). A 6.5 kb transcript that includes flanking pBR322 and SV40 vector sequences is transcribed from the HLD vector and, while the site of initiation of this transcript is unknown, it is not a read-through of the DHFR transcript (unpublished data). The HRD vector, in contrast to HLD, produces a series of transcripts of different sizes, with none as abundant as the HLD HSAG-1 homologous transcript. These results show a correlation between the orientation dependence of the amplification phenomenon and the quantity of vector-HSAG transcription.

Northern analysis of RNA preparations from cells transfected with several HSAG-1 and HSAG-2 subfragments indicates that all pSV_2DHFR EcoRI inserts in the 'left-hand' orientation are transcribed, regardless of their activity (data not shown). The results of this analysis for several subfragments of the highly active H1 RX fragment are shown in Fig. 7. A similar analysis of a series of HSAG-2 subfragment constructs with different levels of activity demonstrated that they also were all transcribed at approximately the same level regardless of their activity (data not shown). These studies indicate therefore that, while HSAG transcription may be necessary for activity and may influence the level of activity of a fragment, transcription of inactive elements is insufficient for activity; certain sequence or structural features may still be lacking.

Significance of HSAG transcription

To address the question of the requirement of vector-initiated HSAG transcription for amplification stimulatory activity, D, HRD and HLD vectors were linearized at the unique BamHI site downstream from the SV40 early polyadenylation site (Fig. 1) prior to transfection into LR73 cells and assessment of amplification activity by MTX survival curves. Linearizing the plasmids at the BamHI site abolished the amplification activities of HRD and HLD (data not shown). Northern analysis of total RNA from transfectants resistant to 0.6 μM MTX indicated that this loss in activity did not result from decreased vector DHFR transcription, at least in the case of HLD transfectants (compare the level of BamHI-cut HLD DHFR transcripts to that of uncut HRD and HLD, Fig. 8, bottom panel; the reason for the decrease in BamHI-cut HRD DHFR transcription is unclear). While DHFR transcription was unaffected in BamHI-cut HLD transfectants, linearizing the plasmids at the BamHI site did abolish transcription of HSAG-1 sequences in the HRD and HLD transformants (Fig. 8, upper panel). Thus in this case, there was a strong correlation between lack of transcription and loss of amplification activity.

DISCUSSION

The frequency of gene amplification can vary a hundred fold between different cell lines (4,8,9) and by the same factor at different loci within a particular cell line (13). Sequences that promote illegitimate recombination have been implicated in gene amplification and have been suggested to function as novel joint 'hotspots' (34,40,41). Other sequences that stimulate gene amplification *in cis* have also been described (10−12,16,17). HSAG-1 (18) promotes the amplification of the expression vector pSV_2DHFR *in cis* when transfected into different mammalian cell lines (1). This effect was dependent on the interaction of multiple positive acting elements in HSAG-1, with the native genomic configuration of these elements as found in HSAG-1 producing the greatest effect (2).

In this paper we describe two other HSAG family members that promote pSV_2DHFR amplification, HSAG-2 and HSAG-5. As was observed with HSAG-1, several subfragments of HSAG-2 and HSAG-5 had partial amplification activity, suggesting that the stimulation of vector amplification by all HSAG elements involves the interaction of multiple positive acting elements. While the comparison of the nucleotide sequences of HSAG-1 and HSAG-2 (this paper) indicated little intersequence homology, both family members possessed several sequence features associated with gene amplification-recombination 'hotspots', including repetitive elements, stem-loop structures, and alternating purine-pyrimidine tracts (34,40) as discussed below. The presence or absence of any one of these sequence features in a particular HSAG fragment did not assure or preclude activity. We submit, however, that several of these features are required in a particular organization for maximal activity, and that different types of positive acting elements may be interchangeable.

Both HSAG-1 and HSAG-2 were found to contain clusters of Alu-like sequences which could contribute to their ability to stimulate amplification. The novel amplification joint 'hotspot' in the adenylate deaminase amplified domain contains four Alu equivalent sequences (40); clusters of human Alu sequences have also been associated with genomic instability (40,41,42). In a similar fashion, the poly purine-pyrimidine tracts in HSAG family members may promote gene amplification as both poly

[d(GT)·d(AC)] sequences and $d(TG)_n$ repeats have been found to increase recombination in various systems (44–47), perhaps by influencing chromatin structure and allowing greater accessibility to recombination enzymes.

A/T-rich regions in HSAG-1 and HSAG-2 may also influence amplification activity (see Results). A/T-rich regions have been implicated in the amplification process in other mammalian systems (40,41) and are also required for yeast ARS function (36). The A/T-rich regions in HSAG may therefore promote DNA unwinding (see additional references in 40) and allow entry of recombinases or other proteins that promote amplification.

At this point, it is not known whether any of the predicted stem-loops (Fig. 4) or palindromic sequences found in the HSAG elements are actually involved in recombination events during the amplification of the pSV_2DHFR vector. Hairpin structures have been hypothesized to be involved in illegitimate recombination during amplification (34) and novel joints have been shown to occur within predicted stem-loop structures (35,40).

Autonomous replication of pSV_2DHFR in LR73 or COS cells was not stimulated by the presence of HSAG-1 sequences and therefore does not seem to be a factor in stimulation of amplification by HSAG sequences. It is nevertheless possible that the integrated HSAG elements may interact with the cellular replication machinery to influence amplification.

The orientation of HSAG-1 and HSAG-2 subfragments in pSV_2DHFR affects their amplification stimulatory activity, with the 'left' orientation producing the greatest effect. Several pieces of evidence indicate that transcription of HSAG-1 inserts is involved in this orientation dependence: (i) HSAG-1 transcription in HRD is a fraction of that of the more active HLD construct and (ii) linearizing the vector with the restriction enzyme BamHI abolishes HSAG transcription and also abolishes the amplification stimulatory activity of HLD and HRD. In both cases, alterations in the level of DHFR transcription are not responsible for the differences in MTX resistance. Transcription alone, however, is not sufficient for activity as several highly transcribed subfragments of HSAG-1 and HSAG-2 have little activity. The sequence and structural features of the transcribed sequences may thus also be important. We hypothesize that transcription may increase the activity of an HSAG fragment altering conformation of the insert. Transcription increases the superhelicity of DNA (48), which may increase its recombinogenicity (49). It has been suggested that actively transcribed genes are also more accessible to sequences and proteins involved in recombination (50,51,52).

The effect of the HSAG elements is therefore likely to be due to a combination of sequence and structural features, such as repetitive sequences, purine-pyrimidine tracts, stem-loops and A/T-rich regions, and the transcription of these elements, which all work in concert to enhance recombination, a necessary process in the generation and resolution of amplified structures (for review see 7). Multiple positive acting elements are involved, which, as HSAG-1 and HSAG-2 are not generally homologous, are unlikely to be highly sequence specific and may each contribute in a quantitative fashion to amplification activity. A strong precedent for multiple interacting elements which affect gene amplification has previously been set with the identification of the functionally redundant amplification enhancing regions in the *Drosophila* chorion gene cluster (12). While there are thousands of copies of HSAG family members in the genome, these results suggest that unless these elements are transcribed they may not stimulate gene amplification. It is appealing to propose that alterations in the sequences flanking one of these elements or its transposition from a transcriptionally inactive region into one that is more active would create a 'hotspot' for gene amplification. In the relatively genetically unstable environment of tumorigenic cells (4,53), such changes could activate endogenous HSAG elements and through increased gene amplification contribute to the malignant phenotype. It remains to be seen if these elements function in this fashion *in vivo*.

ACKNOWLEDGEMENTS

We thank Daniel Dignard for sequencing portions of HSAG-2 and HSAG-5. This work was supported by grants to C.P.S. from the Medical Research Council of Canada (MRC of Canada) and the National Cancer Institute of Canada (NCI of Canada). J.G.M. was supported by a studentship from the Cancer Research Society of Quebec, L.K.B. was supported by studentships from the MRC of Canada and the NCI of Canada, and J.W.C. was supported by a studentship from the MRC of Canada.

REFERENCES

1. McArthur,J.G. and Stanners,C.P. (1991) J. Biol. Chem. (in press).
2. Beitel,L.K., McArthur,J.G. and Stanners,C.P. (1991) Gene (in press).
3. Stark,G.R. and Wahl, G.M. (1984) Ann. Rev. Biochem., **53**, 447–491.
4. Tlsty,T.D. (1990) Proc. Natl. Acad. Sci. USA, **87**, 3132–3136.
5. Wright,J.A., Smith,H.S., Watt,F.W., Hancock,M.C., Hudson,D.L. and Stark,G.R. (1990) Proc. Natl. Acad. Sci. USA, **87**, 1791–1795.
6. Schimke,R.T. (1988) J. Biol. Chem., **263**, 5989–5992.
7. Stark,G.R., Debatisse,M., Giulotto,E. and Wahl,G.M. (1989) Cell, **57**, 901–908.
8. Giulotto,E., Knights,C. and Stark,G.R. (1987) Cell, **48**, 837–845.
9. Rice,G.C., Ling,V. and Schimke,R.T. (1987) Proc. Natl. Acad. Sci. USA, **84**, 9261–9264.
10. Spradling,A.C., de Cicco,D.V., Wakimoto,B.T., Levine,J.F., Kalfayan,L.J. and Cooley,L. (1987) EMBO J., **6**, 1045–1053.
11. Orr-Weaver,T.L., Johnston,C.G. and Spradling,A.C. (1989) EMBO J., **8**, 4153–4162.
12. Swimmer, C., Delidakis,C. and Kafatos,F.C. (1989) Proc. Natl. Acad. Sci. USA, **86**, 8823–8827.
13. Wahl,G.M., de Saint Vincent,B.R. and DeRose,M.L. (1984) Nature, **307**, 516–520.
14. Carroll,S.M., Gaudray,P., DeRose,M.L., Emery,J.F., Meinkoth,J.L., Nakkim,E., Subler,M., Von Hoff,D.D. and Wahl,G.M. (1987) Mol. Cell. Biol., **7**, 1740–1750.
15. Ruiz,J.C. and Wahl,G.M. (1988) Mol. Cell. Biol., **8**, 4302–4313.
16. Pasion,S.G., Hartigan,J.A., Kumar,V. and Biswas,D.K. (1987) DNA, **6**, 419–428.
17. Wegner,M., Zastrow,G., Klavinius,A., Schwender,S., Müller,F., Luksza,H., Hoppe,J., Wienberg,J. and Grummt,F. (1989) Nucleic Acids Res., **17**, 9909–9932.
18. Stanners,C.P., Lam,T., Chamberlain,J.W., Stewart,S.S. and Price,G.B. (1981) Cell, **27**, 211–221.
19. Subramani,S., Mulligan,R. and Berg,P. (1981) Mol. Cell. Biol., **1**, 854–864.
20. Beitel,L.K., Chamberlain,J.W., Benchimol,S., Lam,T., Price,G.B. and Stanners,C.P. (1986) Nucleic Acids Res., **14**, 3391–3408.
21. Stanners,C.P., Chamberlain,J.W., Lam,T. and Price,G.B. (1983) In O'Connor, T.C. (ed.), Oncogenes: Evaluation of Basic Findings and Clinical Potential. Allan R. Liss, Inc., New York, pp. 173–183.
22. Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor University Press, Cold Spring Harbor.
23. Pollard,J.W. and Stanners,C.P. (1979) J. Cell. Physiol., **98**, 571–586.
24. Graham,F.L., Bacchetti,S., McKinnon,R., Stanners,C., Cordell,B. and Goodman,H.M.. (1980) In Baserga,R., Croce,C., and Rovera,G. (eds.), Introduction of Macromolecules Into Viable Mammalian Cells. Alan R. Liss, Inc., New York, pp. 3–25.
25. Cartier,M., Chang,M.W.-M. and Stanners,C.P. (1987) Mol. Cell. Biol., **7**, 1623–1628.
26. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) Proc. Natl. Acad. Sci. USA, **74**, 5463–5467.

27. Staden,R. (1984) Nucleic Acids Res., **12**, 521–538.
28. Staden,R. (1982) Nucleic Acids Res., **10**, 2951–2961.
29. Chamberlain,J.W., Henderson,G., Chang,M.W.-M., Lam,T., Dignard,D., Ling,V., Price,G.B. and Stanners,C.P. (1986) Nucleic Acids Res., **14**, 3409–3424.
30. Jelinek,W.R. and Schmid,C.W. (1982) Ann. Rev. Biochem., **51**, 813–844.
31. Shepherd,J.C.W. (1981). Proc. Natl. Acad. Sci. USA, **78**, 1596–1600.
32. Fickett,J.W. (1982) Nucleic Acids Res., **10**, 5303–5318.
33. Deininger,P.L., Jolly,D.J., Rubin,C.M., Friedmann,T. and Schmid,C.W. (1981) J. Mol. Biol., **151**, 17–33.
34. Passananti,C., Davies,B., Ford,M. and Fried,M. (1987) EMBO J., **6**, 1697–1703.
35. Hyrien,O., Debatisse,M., Buttin,G. and de Saint Vincent,B.R. (1988) EMBO J., **7**, 407–417.
36. Umek,R.M. and Kowalski,D. (1988) Cell, **52**, 559–567.
37. Rao,B.S., Zannis-Hadjopoulos,M., Price,G.B., Reitman,M. and Martin,R.G. (1990) Gene, **87**, 233–242.
38. Caddle,M.S., Lussier,R.H. and Heintz,N.H. (1990) J. Mol. Biol., **211**, 19–33.
39. Peden,K.W.C., Pipas,J.M., Pearson-White,S. and Nathans,D. (1980) Science, **209**, 1392–1396.
40. Hyrien,O., Debatisse,M., Buttin,G. and de Saint Vincent,B.R. (1987) EMBO J., **6**, 2401–2408.
41. Legouy,E., Fossar,N., Lhomond,G. and Brison,O. (1989) Som. Cell Mol. Gen., **15**, 309–320.
42. Lehrman,M.A., Russell,D.W., Goldstein,J.L. and Brown,M.S. (1987) J. Biol. Chem., **262**, 3354–3361.
43. Chen,S.J., Chen,S., Font,M.-P., d'Auriol,L., Larsen,C.-J. and Berger,R. (1989) Nucleic Acids. Res., **17**, 7631–7642.
44. Stringer,J.R. (1985) Mol. Cell. Biol., **5**, 1247–1259.
45. Treco,D. and Arnheim,N. (1986) Mol. Cell. Biol., **6**, 3934–3947.
46. Wahls,W.P., Wallace,L.J. and Moore,P.D. (1990) Mol. Cell. Biol., **10**, 785–793.
47. Boehm,T., Mengle-Gaw,L., Kees,U.R., Spurr,N., Lavenir,I., Forster,A. and Rabbitts,T.H. (1989) EMBO J., **8**, 2621–2631.
48. Giaever,G.N. and Wang,J.C. (1988) Cell, **55**, 849–856.
49. Frank-Kamenetskii,M. (1989) Nature, **337**, 206.
50. Blackwell,T.K., Moore,M.W., Yancopoulus,G.D., Suh,H., Lutzker,S., Selsing,E. and Alt,F.W. (1986) Nature, **324**, 585–589.
51. Voelkel-Meiman,K., Keil,R.L. and Roeder,G.S. (1987) Cell, **48**, 1071–1079.
52. Thomas,B.J. and Rothstein,R. (1989) Cell, **56**, 619–630.
53. Sager,R., Gadi,I.K., Stephens,L. and Grabowy,C.T. (1985) Proc. Natl. Acad. Sci. USA, **82**, 7015–7019.
54. Hirt,B. (1967) J. Mol. Biol., **26**, 365–369.
55. Feinberg,A.P. and Vogelstein,B. (1984) Anal. Biochem., **137**, 266–267.
56. Gluzman,Y. (1981) Cell, **23**, 175–182.
57. Chirgwin,J.M., Pryzbla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) Biochem., **18**, 5294–5299.