Characterization of Two Age-induced Intracisternal A-particle-related Transcripts in the Mouse Liver

TRANSCRIPTIONAL READ-THROUGH INTO AN OPEN READING FRAME WITH SIMILARITIES TO THE YEAST CCR4 TRANSCRIPTION FACTOR

(Received for publication, September 13, 1996, and in revised form, November 18, 1996)

Anne Puech‡§§, Anne Dupressoir‡§, Marie-Paule Loireau‡, Marie-Geneviève Mattei‡, and Thierry Heidmann‡**

From the ‡Unité de Physicocimie et Pharmacologie des Macromolécules Biologiques, CNRS URA147, Institut Gustave Roussy, 39 rue Camille Desmoulins, 94805 Villejuif, and §§Unité de Génétique Médicale et Développement, INSERM U406, Faculté de Médecine, 27, Bd. Jean Moulin, 13385 Marseille, France

Intracisternal A-particle (IAP) sequences are endogenous retrovirus-like elements present at 1,000 copies in the mouse genome. We had previously identified IAP-related transcripts of unusual size (6 and 10 kilobases (kb)), which are observed exclusively in the liver of the aging mouse. In this report, using cDNA libraries that we have constructed from the liver mRNAs of an aged DBA/2 mouse, we have cloned and entirely sequenced the corresponding cDNAs. Both are initiated within the 5’ long terminal repeat of a type I IAP sequence, and correspond to a read-through into a unique flanking cellular sequence containing a 966-nucleotide open reading frame, located 3’ to the IAP sequence. The 6-kb IAP-related transcript corresponds to a post-transcriptional modification of the 10-kb mRNA, and is generated by a splicing event with the donor site in the IAP sequence, and the acceptor site 5’ to the open reading frame. This open reading frame is located on chromosome 3, is evolutionarily conserved, and discloses significant similarity to the yeast CCR4 transcription factor at the amino acid level. The specific expression of these age-induced transcripts, which account for more than 50% of the IAP-related transcripts in the liver of old mice, is therefore entirely consistent with the induction of a single genomic locus, thus strengthening the importance of position effects for the expression of transposable elements. Characterization of this locus should now allow studies on its chromatin and methylation status, and on the “molecular factors of senescence” possibly involved in its induction.

Intracisternal A-particle (IAP)† sequences are moderately reiterated transposable elements (approximately 1000 copies in the mouse genome), which are closely related to retroviruses and transpose via the reverse transcription of an RNA intermediate (1, 2). These sequences, 3.5–7.2 kb long, have been classified into different subgroups, each represented by a few hundred copies (3, 4), depending on the presence of various internal deletions (class I and IΔ1–4) and of one specific 0.3-kb insertion (class IIA-C). Like retroviruses, they contain two long terminal repeats (LTRs) with a U3-R-U5 organization, bordering gag-pol-like open reading frames. They are not infectious, and as expected disclose a severely truncated envelope gene. The LTRs contain the signals for the initiation and regulation of transcription (5’-LTR) and for the polyadenylation of the transcripts (3’-LTR). At least some of the IAP elements should still be functional as transposition events can be detected, either directly (2) or indirectly via the mutations that their transposition provoke, especially in tumor cells in culture where these elements are heavily transcribed and act as mutagenic agents (reviewed in Refs. 1 and 5).

The in vivo pattern of expression of IAPs is complex, owing to the large number of these elements in the genome and to possible position effects. In an extensive in vivo analysis of IAP expression in transgenic mice, using LTR-driven reporter genes, we have previously identified a specific “niche” for expression of these elements in the stem cells of the germ line (6). Expression of the IAP transgenes in these cells was found to be position-independent, and should therefore reflect an intrinsic property of these elements. It is consistent with a possible role in adaptive processes, since transposition specifically in the germ line should result in inherited mutations. However, global Northern blot analysis of endogenous IAP expression using appropriate probes have also revealed IAP transcripts in several mouse somatic tissues with, for instance, maximum expression in the thymus (reviewed in Refs. 1 and 7). The occurrence of such transcripts is intriguing, as expression of the IAP transgenes was not detected in these tissues, and a plausible interpretation could be that they simply originate from a limited number of IAP elements, which would be embedded, for instance, within genomic domains specifically activated in the corresponding tissues. An interesting situation was previously unraveled in an extensive analysis of IAP transcripts in the organs of aging mice (7). The rationale of this study was that aging is a process associated with cellular dysfunctions, for which it can be hypothesized (among other theories) that they result from an accumulation of somatic mutations (reviewed in Refs. 8–10). These could be triggered by transposable elements, acting as insertion mutagens. In this respect, we had actually revealed IAP transcripts of abnormal size (6 and 10 kb) which are induced at least 30-fold in the liver of aged mice.
whatever the mouse strain tested. A similar induction of gene transcription has only been rarely described in relation with aging, especially for potentially mutagenic transposable elements, and it was therefore of interest to characterize the IAP elements associated with these transcripts (i) to determine whether they correspond to “functional” IAP sequences, and (ii) to characterize the mechanisms of their induction upon aging.

We have therefore constructed cDNA libraries with the RNAs from the liver of aged mice, and in this report we demonstrate that the IAP transcripts induced with age in fact originate from a single locus that we have cloned, and correspond to a read-through from a non-coding IAP element into the adjacent cellular DNA, which further contains an open reading frame with homology to a yeast transcription factor. These results demonstrate the importance of position effects on the expression of highly reiterated elements, and identify a locus that now allows the characterization of domains or genes possibly directly associated with the aging process.

**EXPERIMENTAL PROCEDURES**

**Animals—**Mice from four different inbred strains (C57BL/6, C57BL/10, BALB/c, and DBA/2) were obtained from Iffa-Credo Laboratories; these mice were fed ad libitum. All mice were killed by cervical dislocation and had no evident pathology at the time of death.

**PCRs—**PCRs were performed with a Hybaid thermal cycler. 1 ng of plasmid DNA, 100 ng of genomic DNA, or 5 μl of plasmid stock (previously submitted to three cycles of liquid nitrogen freezing and thawing) were amplified with 1 unit of Taq DNA polymerase in the buffer supplied by the manufacturer (Amersham). Initial denaturation was for 5 min at 95°C, followed by 30 cycles at 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min. Final extension was at 72°C for 10 min. PCR products were gel-purified using the freeze-thaw phenol extraction method (11), ethanol-predictlicated, and resuspended in 20 μl of TE buffer. PCR products were then either used directly as a probe, or cloned after Klenow enzyme treatment into the EcoRV site of the pBluescript vector (Stratagene).

The following primers were synthesized: IAPE1, 5'-CCAATGGCAG-TTGGCCTAGTGCT; IAPE3, 5'-AGCTGACTGTTGATAGTTGCT-TA-GA; AR-B11, 5'-TTGGGACACTAACTCTGATCTTCCT; AR-B12, 5'-G-AATACCACTAATCTACTGGTC; AR-B51, 5'-GGATACCTGTC-CTACCAAATTG; AR-B52, 5'-TTGGGACACTAACTCTGATCTTCCT; AR-B5, 5'-ACAAAAATTTGAGAAGGGCTGTG; AR-B32, 5'-GCCCACTGAGGACTT; AR-U3, 5'-GCCCACTGAGGACTT; AR-U3R, 5'-TTACATGTTGACCTCCTAC.

**Probes—**DNA probes were labeled by random priming using [32P]dCTP and a commercial kit (Amersham). DNA probes were obtained by PCR amplification using primers IAPE1 and IAPE3 and a commercial kit (Amersham). IAP DNA probes were labeled with γ-[32P]ATP using T4 polynucleotide kinase (Pharmacia). The resulting PCR fragment (assayed by nucleotide extraction method) containing IAP-IL3 (12) was labeled with [32P]dCTP using T4 polynucleotide kinase (Pharmacia). The following probes were synthesized: IAPE1, 5'-CCAATGGCAG-TTGGCCTAGTGCT;IAPE3, 5'-AGCTGACTGTTGATAGTTGCT-TA-GA; AR-B11, 5'-TTGGGACACTAACTCTGATCTTCCT; AR-B12, 5'-G-AATACCACTAATCTACTGGTC; AR-B51, 5'-GGATACCTGTC-CTACCAAATTG; AR-B52, 5'-TTGGGACACTAACTCTGATCTTCCT; AR-B5, 5'-ACAAAAATTTGAGAAGGGCTGTG; AR-B32, 5'-GCCCACTGAGGACTT; AR-U3, 5'-GCCCACTGAGGACTT; AR-U3R, 5'-TTACATGTTGACCTCCTAC.

**RNA Isolation and Northern Blot Analysis—**Total cellular mRNAs from various tissues were extracted using the guanidium isothiocyanate extraction method (15). Poly(A)+ RNA was isolated from total RNA using a pre-packed spin column containing oligo(dT)-cellulose (Pharmacia Biotech Inc.). For Northern blot analysis, 10 μg of total RNA/lane were fractionated on agarose/formaldehyde gels. RNAs were transferred to a nylon membrane (Hybond N, Amersham) in 0.15 M NaHAc buffer and hybridized with riboprobes or DNA probes. Loading of equal amounts of RNA was assessed by ethidium-bromide-stained ribosomal RNAs upon UV illumination of the membrane. Prehybridization and hybridization were performed in 50% formamide with 0.75 M NaCl, 1% SDS, 50 μg/ml salmon sperm DNA, and 5% dextran sulfate when using riboprobes, or with 7% SDS, 1 mm EDTA, 0.5 M NaHPO4, pH 7 (16), when using DNA probes. Hybridized blots were first rinsed in 2× SSPE at 65°C and then washed at 65°C (time length and SSC concentration varied depending on the probe; see figures). Filters were then exposed for at least 24 h.

**DNA Isolation and Southern Blot Analysis—**High molecular weight DNA was extracted from different tissues, which were first crushed on dry ice and then incubated overnight at 55°C in a solution containing 50 mm Tris-HCl, pH 7.5, 0.1 M NaCl, 25 mm EDTA, 1% SDS, and 150 μg of proteinase K. DNA was phenol-chloroform-extracted, ethanol-predictlicated, washed twice in 70% ethanol, dried under vacuum, and resuspended in TE buffer. DNA digestion was carried out with a 4–6-fold enzyme excess. Digested DNA was electrophoresed on 0.8% agarose gels, transferred to nylon membrane (Hybond N, Amersham), and cross-linked to the membrane using a Stratalinker apparatus (Stratagene). Hybridization and washing were carried out as described for Northern blot analysis.

**Construction of Mouse Liver cDNA Libraries—**Two cDNA libraries (one oligo(dT)-primed and one random-primed) were constructed using the AZap Express cDNA synthesis kit (Stratagene), following the manufacturer’s instructions. Poly(A)+ RNA from an 18-month-old DBA2J mouse liver was used for first-strand cDNA synthesis, using either poly(dT)-unidirectional primers or random-unidirectional linker primers, which both contain a XhoI site (Stratagene). Second-strand cDNA synthesis was followed by directional cloning into the EcoRI/XhoI sites of the AZap Express arms. Ligated phage DNA was packaged in vitro using Gigapack II Gold packaging extracts (Stratagene). The libraries were subsequently amplified using XL1Blue MRF as an host strain. 3.5×1010 phage particles in 9 × 105 bacterial cells of the pBluescriptI and the random-primed cDNA libraries, respectively, were plated at near confluence and duplicate filters were prepared. Differential screening with [32P]-labeled probes was performed as described for Northern blot analysis.

**Characterization of the IAP Age-related Transcripts—**As previously demonstrated and illustrated in Fig. 1A, Northern blot analysis, using an almost full-length IAP probe, of RNAs from the liver of aged mice reveals 10- and 6-kb transcripts not present in young mice, and the former is further induced when using riboprobes instead of DNA probes.
observed in the young mouse. These IAP age-related transcripts (IAP-AR1 and -AR2) are of abnormal size as compared to the canonical 7.2- and 5.4-kb IAP transcripts classically observed in the mouse tissues, and they were first characterized by an extensive analysis using various IAP probes (see scheme in Fig. 1E).

Probing Northern blots of liver mRNA from old mice with single-stranded IAP riboprobes (data not shown) first demonstrated that both IAP-AR transcripts are transcribed from the sense strand, as expected for normal IAP transcripts. Taking into account the abnormal length of the IAP-AR1 transcript, we then questioned whether it could correspond to an envelope-containing sequence; actually, full-length IAP transcripts characterized to date (i.e. 7.2 kb) all contain a truncated env gene, but Reuss and Schaller (13) recently cloned a 4-kb cDNA coding for a related full-length envelope. As shown in Fig. 1B, hybridization with an env probe failed to detect any of the two IAP-AR transcripts even under prolonged exposure (only a minor transcript of 1.45 kb could be detected, possibly corresponding to an alternative splice of the previously characterized env transcript in Ref. 13). Finally, Northern blots were hybridized with a series of probes encompassing different IAP sub-domains (Fig. 1, B and E). Hybridization with an internal probe immediately adjacent to the 5'-LTR (IAP-5'), with the two probes in the pol region (IAP-pol1 and IAP-pol2), as well as with an LTR probe (not shown), in all cases detected the two IAP-AR transcripts (Fig. 1E). Only the IAP-3' probe, located just 5' to the 3'-LTR, failed to detect IAP-AR2, but still detected IAP-AR1. Altogether, these results indicate that the 10- and 6-kb transcripts contain closely related IAP domains, characteristic of type I, type I1A, or type IIA elements. The features of the IAP-AR2 transcript do not correspond to those of any IAP transcripts characterized to date, and could result from a 3' deletion within the progenitor IAP, or from a post-transcriptional modification; the IAP-AR1 transcript is not associated with a still hypothetical full-length gag-pol-env IAP, and most probably corresponds to an extended 5' or 3' read-through.

Cloning Strategy and Sequencing of the IAP-AR cDNAs—To characterize further the IAP-AR transcripts, we constructed two cDNA libraries: a randomly primed cDNA library and an oligo(dT)-primed one. In both cases cDNAs were synthesized with poly(A) + mRNAs from the liver of an 18-month-old DBA/2 mouse, as highest expression of both transcripts was observed in this strain (7). According to the analysis above, no IAP probe could be used for a direct selective screening of the IAP-AR versus normal IAP transcripts. We therefore decided to clone...
the IAP-AR1 first, assuming that it corresponds to an IAP read-through transcript (see above), and even more precisely to a 3' read-through (its level of transcription, in the sense orientation, was shown to vary among mouse strains as the canonical IAP transcripts (7), strongly suggesting that it is initiated within an IAP LTR). Accordingly, the strategy for cloning was to screen the randomly primed cDNA library with a “border” IAP probe (IAP-3') and to select positive clones also containing non-IAP sequences. Twenty positive clones were therefore isolated and purified, and inserts subsequently excised as pBK plasmids (see “Experimental Procedures”). The fraction of IAP sequence contained in the cDNA inserts was determined after restriction of the plasmid DNAs, and hybridization of the blots on which they had been transferred with an IAP probe. Out of the 20 positive bacteriophage clones, 8 had non-IAP fragments larger than 500 bp and were sequenced. Hybridization of a XhoI-HindIII fragment (probe XH, see Fig. 1, C and E) originating from the 3' region of one of the candidate clones (pA3–13; cDNAs were cloned in an oriented way), finally allowed us to detect the 10-kb IAP-AR1 transcript. Surprisingly, this fragment also detected the 6-kb IAP-AR2, as well as a 3-kb transcript, thus suggesting that these are alternatively spliced forms of the IAP-AR1 transcript (as in Fig. 1). As illustrated in Fig. 1D using the same probe, these three transcripts could be detected in the liver mRNA originating from old mice of four different mouse strains (with varying intensity depending on the strain, as previously observed (7) for an IAP probe), and they were all induced upon aging. These results suggest that the IAP-AR transcripts are originating from the same locus and confirm that their induction is a general feature related to aging in the liver. The complete characterization of the IAP-AR1 transcript was then achieved using probe XH, which was used to screen the 18-month-old liver oligo(dT)-primed cDNA library to clone the 3' end of the 10-kb IAP-AR1 transcript (clones pB2–16 and pB4–28, Fig. 2). Screening of the randomly primed cDNA library finally allowed us to identify a clone containing the 5' end of the transcript (clone pIII-3, Fig. 2; see below). Sequencing of the subclones provided the composite cDNA sequence of the 10-kb IAP-AR1 transcript (Fig. 3).

The strategy for cloning the 6-kb IAP-AR2 and the 3-kb transcripts was then rather straightforward, taking into account that both are positive for the XH probe. According to the data in Fig. 1, IAP-AR2 should be positive for the IAP-pol probes but, unlike IAP-AR1, should be negative for probe IAP-3': out of 10 clones positive for probe XH that we isolated from the oligo(dT)-primed cDNA library, one met these combined criteria (pB2–12, Fig. 2) and was entirely sequenced. Similarly, a cDNA for the 3-kb transcript should be positive for the XH probe, negative for both the IAP-3' and the IAP-pol probes, and about 3 kb long; combination of these four criteria allowed us to identify a corresponding 3-kb cDNA (clone pB1–19, Fig. 2) that was also entirely sequenced.

Structural Features of the IAP-AR Transcripts—The IAP-AR1 cDNA discloses, within its 5' half, the characteristic features of an IAP element. Two cDNA clones have been isolated with almost exactly the same most 5' sequence, extending (within 10–15 bp) to the expected IAP transcription start site located at the 5' end of the LTR R domain (Fig. 3). The downstream IAP sequences then disclose an overall 97% identity to the sequence of a type L1 IAP element, and more precisely to the previously cloned and entirely sequenced IAP-IL3 element (12). Both carry the same internal deletion covering the 3' end of the putative gag gene and extending into the 5' end of the putative pol gene. However, due to the occurrence of six stop codons within the normally in-phase fusion gag-pol gene, the long (3060 nt) contiguous ORF in IAP-IL3 is reduced to 1776 nt in IAP-AR. Translation and homology searches, however, indicate that both coding domains are still 99% identical. The 3'-LTR contains the typical U3, R, and U5 subdomains, and is bracketed by the characteristic 4-bp inverted repeats (TGTT and AACA). Alignment of the IAP-AR and IAP-IL3 LTR sequences discloses strong similarities (97.6% identity, excluding the different number of internal repeats in the R domain, see below). Actually, the IAP-AR LTR sequence closely resembles that of the IAP elements specifically activated in the thymus or in LPS-stimulated B cells, namely the LS-type IAPs (19, 20). In the IAP R domains, internally located 13-bp direct repeats, from 1 to 6, are commonly observed (19), which are most probably generated by the retrotransposition process (2). The IAP-AR se-

---

**Fig. 2. Structure of the isolated cDNA clones.** The isolated cDNA clones (name on the right) are represented with solid lines, with the detailed structure of the chimeric IAP-AR1 transcript below (as in Fig. 1). SD, splice donor; SA, splice acceptor.
The IAP sequence within IAP-AR1 is then followed by a non-coding sequence containing 4 B1 and 1 B2 repeats. These are highly reiterated mobile elements of the mouse genome classically found in intronic domains, and which are closely related to the human Alu sequences (21, 22). All repeats except one are in an opposite transcriptional orientation as compared to IAP-AR1. The first B1 repeat is immediately adjacent to the IAP-3' end, consistent with the fact that IAP sequences are found inserted preferentially close to such repeats (23). At the 3' end of the IAP-AR1 cDNA, a 966-nt ORF (nt 7661–8632) can be found (see below), which is followed by a 1104-nt (nt 8633–9718) non-coding region ending with a poly(A) tail. As expected, a putative polyadenylation signal (AATAAA, nt 9704–9709) can be found 15 nt 5' to this poly(A) tail.

Analysis of the 6-kb IAP-AR2 and 3-kb transcripts (clones pB2–12 and pB1–19) reveals the existence of two splicing events (Figs. 2 and 4). The first splicing event may be described as a gag (nt 378)-to-env (nt 3413) splice, which is almost similar to the gag-to-env splicing event previously described in Ref. 24. In both cases, the same splice acceptor (SA1, Fig. 4A) was used. Interestingly, the corresponding region is extremely well conserved among the different IAP elements described to date. However, the splice donor (SD1, Fig. 4A) is slightly divergent from the one described in Ref. 24. This may be related to a lesser sequence conservation observed in the corresponding region among IAP elements.

The second splicing event may be described as an env (SD2, nt 3572)-to-ORF region (SA2, nt 7528). Clone pB2–12 exhibited the same env-to-ORF splicing event as clone pB1–19; however, the SD2* site (nt 3577) used in the former was 5 bp apart from SD2 used in the latter (Fig. 4B). To ascertain that this “intronic” region is absent from both the IAP-AR2 and 3-kb transcripts, we used probes corresponding to this domain (probes A, B, and C) or to the 3' region (probe XH in the ORF and probe UTR, Fig. 1E). As indicated in Fig. 1(C and E), probes XH and UTR hybridize to the 6-kb IAP-AR2 and 3-kb transcripts but probes A, B, and C did not. This indicates that the pB2–12 and pB1–19 clones are not recombination artifacts, but are genuine reverse transcripts of the 6-kb IAP-AR2 and 3-kb transcripts, and that both are being definitely generated by splicing. Analysis of the Northern blot in Fig. 1D further strongly suggests (in addition to cDNA sequencing) that these two spliced transcripts derive from IAP-AR1, as the intensities of all three transcripts vary (both upon aging and upon the mouse strain tested) in a similar manner. Fig. 1D (right) also shows that occurrence of the IAP-AR1 and IAP-AR2 transcripts most probably does not result from a switch from an hypothetical other promoter for the ORF, as the 3-kb transcript can hardly be detected in the liver of the young mouse (and as no other ORF transcripts, with sufficient intensity, are observed). The fact that all three transcripts are expressed at a variable level depending on the mouse strain as similarly observed for the canonical IAP transcripts (see Ref. 7) is also consistent with the IAP LTR as being the common promoter for all three transcripts.
An Evolutionarily Conserved ORF with Similarity to the Yeast CCR4 Transcription Factor in the IAP-AR Locus—The 3' end of the IAP-AR transcripts contain a 966-nt ORF with the potential to encode a 322-amino acid protein. Since most important genes are evolutionarily conserved, we made a Zooblot analysis of DNAs from different species, with a probe encompassing the entire ORF together with its 3'-untranslated region (probe pB2–16, see Fig. 2). As illustrated in Fig. 5, several species (human, hamster, cat, rat, and mouse) demonstrated hybridization under moderately stringent conditions, whereas no hybridization could be detected in Saccharomyces cerevisiae or Drosophila melanogaster. DNA from monkey, dog, cow, rabbit, and chicken were also positive under similar conditions (data not shown), thus demonstrating at least a strong conservation of the ORF among vertebrates.

In parallel, using the blastX program (University of Wisconsin Genetics Computer Group), which translates both strands in all six reading frames, we searched for similarities against the entire NBRF and Swissprot protein data bases and found a significant homology to the 3' coding region of the yeast CCR4 gene (25). As illustrated in Fig. 6, alignment of the IAP-AR ORF with the yeast CCR4 protein sequence shows an overall identity of 28%, with 36% identity over amino acids 507–612 and amino acids 754–828. Similarity searches against the GenBank and the DBEST data bases using the tBlastN program (which compares a protein query sequence against a nucleotide data base translated in all reading frames) also revealed similarities to two human ESTs, clones T87026 and H45114 (Fig. 6). These two ESTs are not overlapping; clone H45114 is 35% identical (in amino acid) to the 5' end of the IAP-AR ORF, and clone T87026 is 70% identical to its 3' end (Fig. 6). Although these two ESTs are derived from two different cDNA clones, they may be part of the same gene, which, then, could be the human homologue of the yeast CCR4 gene. As illustrated in Fig. 6, it is worth mentioning that numerous residues are conserved among these three species (human, mouse, and yeast).

Finally, as illustrated in Fig. 7A, a Southern blot analysis of mouse genomic DNA restricted with three different enzymes detected essentially single bands with the ORF-containing pB2–16 DNA probe (two bands are observed with PstI, but this is expected from the nucleotide sequence). It strongly suggests that the IAP-AR ORF is part of a single copy gene, and might then represent an exon of the mouse homologue of the yeast CCR4 gene. A chromosomal localization of the corresponding locus in the mouse was finally performed by in situ hybridization using the same probe. As expected from the Southern blot analysis presented above, a single hot spot was identified, which mapped on chromosome 3 in the B–D region (Fig. 7B).

DISCUSSION

We have characterized and cloned the two IAP-related transcripts specifically expressed in the liver of the aged mouse. We
previously demonstrated that expression of these two abnormally sized transcripts, called IAP-AR1 and -AR2 (7), is a common feature of all four different mouse strains tested. Sequence analysis now shows that both RNAs are chimeric transcripts from a canonical type I D1IAP element originating from a single locus. The 10-kb IAP-AR1 transcript corresponds to a transcriptional read-through, extending out of the 3’-LTR into 5 kb of cellular flanking DNA containing a 1-kb evolutionarily conserved open reading frame with similarity to the yeast CCR4 transcription factor. This large transcript gives rise to two smaller transcripts by post-transcriptional modifications, the IAP-AR2 transcript by an IAP env-to-cellular DNA splice, and a 3-kb transcript by an additional internal IAP gag-to-env splice. Due to the occurrence of transcripts from other IAPs, it was not possible to perform primer extension experiments to definitely characterize their transcription start sites. However, the overall size of the composite cDNAs as well as the sequence of the most 5’-clones that we isolated strongly suggest that

FIG. 6. Comparison of the predicted amino acid sequence of the 3’ end ORF within the IAP-AR transcripts with the yeast CCR4 and human ESTs. Sequences are represented with the single-letter code; numbers refer to the yeast CCR4 amino acid sequence (PIR accession number S36713), and dashes represent gaps introduced to optimize alignment; H45114 and T87026 are human ESTs. Amino acid identities of the IAP-AR ORF with yCCR4 are boxed, and those with H45114 or T87026 are underlined.

FIG. 7. Characterization of the ORF-containing locus. A, Southern blot analysis of DBA2 mouse genomic DNA (5 µg) restricted with the indicated enzymes and hybridized with probe pB2–16 as in Fig. 5. Plasmid DNA (pB2–16) corresponding to 1 or 2.5 gene copies were deposited as a standard. B, gene mapping by in situ hybridization with probe pB2–16 (see “Experimental Procedures”). A series of 100 metaphase cells were examined: there were 149 silver grains associated with chromosomes and 61 of these (41%) were located on chromosome 3. The distribution of grains on this chromosome was not random; 51/61 (83%) of them mapped to the B–D region, as illustrated on the figure.
their transcription is initiated, as for canonical IAP elements, in the R domain of the IAP 5'-LTR.

Read-through and Alternative Splicing to Generate Chimeric Transcripts from Retroviral-like Elements—The presently identified locus combines two previously identified situations resulting in chimeric transcripts: (i) transcriptional read-through, and (ii) alternative splicing. Transcriptional read-through has previously been documented for IAPs (19, 24), as well as for other retrotransposons (e.g. the HERV-R human endogenous retrovirus; Refs. 26 and 27). In some cases read-through results from deletions or mutations within the polyadenylation signal-containing 3'-LTRs. Mietz et al. (19) have actually described IAP cDNA from splenic B cells with deletions of the 3' R domains, and, in the course of our cloning experiments, we have also isolated two chimeric IAP transcripts with truncations at the 3' end of the IAP elements (complete deletion of the 3'-LTR, or R-US deletion) followed by 100–500 bp of non-IAP DNA. For the IAP-AR transcripts, however, their complete nucleotide sequencing has revealed a rather unique situation in which a read-through takes place without any evidence for a deletion or mutation within the IAP 3'-LTR (see “Results”). Although less well documented, sequencing of the human HERV-R element responsible for a read-through to a 3' located Krüppel-like gene (26, 27) also did not reveal evident alterations within the 3'-LTR, therefore suggesting the possible existence of mechanisms modulating polyadenylation efficiency within retroid elements, so as to account for such events.

The other process resulting in chimeric transcripts involves a subjugation of the splice sites, which are present in most retroid elements, and alternative splicing. Leslie et al. (24) described two cases of cellular gene activation generated by an IAP gag-to-IL3 and IAP gag-to-GM-CSF splicing event in WEHI-3B derived cells, and similar events have also been reported with other retrotransposable elements, such as the human HERV-H elements resulting in gag-to-calbindin and gag-to-phospholipase A₂-like chimeric transcripts. This illustrates the extent to which retroid elements can be recruited and participate to genome evolution, and emphasizes the close relationship between retroviruses and endogenous retrotransposons. In the present study, functional donor splice sites located in two distinct domains of the IAP sequence have been demonstrated: SD1, which is located in the IAP gag region close to (but distinct from) the site described in Ref. 24, and the two 5-bp-apart SD2 and SD2ₐ sites which are located at the 5' end of the truncated env gene, in a highly conserved region among IAPs. Rather surprisingly, the latter sites have never been described previously, but their position is reminiscent of internal env donor splice sites commonly found in complex infectious retroviruses (e.g. HIV and HTLV1), where they are responsible for the synthesis of small regulatory viral proteins (e.g. Tat and Rev; reviewed in Ref. 28). The presence of functional donor sites at a similar location in the IAP sequence could be an additional hint for the retroviral origin of these now strictly endogenous retroid elements.

Importance of “Position Effects” for IAP Expression—This study also demonstrates, de facto, the importance of position effects for the expression of IAP elements. Actually, we have shown that the IAP-AR transcripts, which account for >50% of the IAP-related transcripts in the liver of the aged mice, unambiguously originate from a single IAP element, in a definite genomic locus that we have mapped. Furthermore, the identified IAP has no special features; its sequencing has revealed that it is a type ΔI IAP element, closely related to the previously cloned and characterized IAP-IL3 (12). The absence of a generalized induction of the IAPs in the liver of the aged mice and the rather “localized” induction of a definite and canonical element, therefore, strongly suggest that it is not the nature of the IAP element per se but rather its chromosomal position and environment (chromatin structure, DNA methylation, presence of adjacent age-specific genes or regulatory sequences, see below) that is determinant for the expression of these otherwise severely repressed mutagenic elements. This conclusion is in fact fully consistent with previous experiments in transgenic mice (6) with marked IAP elements (among which, precisely, IAP-IL3), where IAP expression was found to be restricted to the male germline, and was never observed in somatic tissues (even in those of aged mice). Additionally, these two series of results also strongly suggest that the transcripts detected by global Northern blot analysis in most mouse somatic tissues (reviewed in Refs. 1 and 7) most probably originate from a very limited number of IAP elements and not from the overall IAP population (see also Refs. 19 and 20), and are therefore “singularities” reflecting DNA site-specific, tissue-dependent, and (in the present case) age-dependent, IAP inductions.

IAP Expression and Aging—Several theories have been proposed to account for the aging process (reviewed in Refs. 8–10). Deterministic theories have proposed that aging is a programmed process, whereas stochastic theories have involved accumulation of errors, of various origin. In the latter case, activation of transposable elements and endogenous retroviruses have been invoked and in some cases documented (35, 36) as a possible event participating in somatic mutations. Along the same line, demethylation upon aging of the LINE murine mobile elements (and to some extent of the IAPs) has been reported (37, 38). In fact, the present study (see also Refs. 7 and 39) clearly shows that there is not a generalized de-repression of IAP expression upon aging, which would result in an enhanced IAP-mediated mutagenesis; IAP induction is limited to a single element, which further is a defective ΔI IAP element, therefore not autonomous for transposition. However, this specific induction might not be “neutral” for the aging mouse. Actually, the age-induced transcripts include a 3' adjacent cellular sequence containing a large ORF with strong phylogenetic conservation and similarity to the yeast CCR4 protein. Although we do not know yet whether the age-induced transcription results in the actual production of CCR4-like proteins (antibodies are presently being raised to a recombinant protein to test this point), it is possible that this ORF, as well as in fact part of the defective IAP, is a candidate “auto-gene” as defined in Ref. 40, and is therefore involved in auto-immune reactions in the aging mice. It could also encode a factor involved in transcriptional regulations: actually, the CCR4 protein is a transcription factor for the glucose-repressible genes in yeast (41), and recently Draper et al. (42) identified a mouse protein that binds to the yeast CCR4 protein, resulting in a protein complex with evolutionarily conserved transcriptional functions. The presently identified ORF could be an exon of the mouse homologue (which still remains to be characterized) of the yeast CCR4 gene, and the IAP sequence be inserted within an intron of the corresponding gene. Accordingly, induction of IAP transcription could result in the synthesis of a truncated CCR4-like protein with distinct functional properties, which as such might play a role in aging.

The reasons for the specific induction of the IAP-AR tran-

---

2 A. Puech, unpublished data.

3 A. Dupressoir, unpublished data.
scripts with aging are still unknown. Spreading of an inactive chromatin conformation from a methylated domain to an adjacent non-methylated region was previously shown to inhibit the expression of a reporter gene in mammalian cells (43), and position effect variegation involving heterochromatin spreading and resulting in gene silencing has been well characterized in Drosophila (44). By analogy, the age-dependent induction of the IAP-AR transcripts could therefore be associated with the spreading of an active chromatin conformation from an adjacent gene that would be specifically expressed upon aging. It spreading of an active chromatin conformation from an adjacent gene that would be specifically expressed upon aging. It could also be more directly associated with the presence of enhancer sequences responsive to age-dependent factors (e.g. the ubiquitous Age-Dependent Factor of the rat; Ref. 45). Whatever the case, the identification of the present locus should allow the characterization of age-responsive elements and/or genes, and therefore be important for the study of the aging process in concrete and molecular terms.

Acknowledgments—We are grateful to L. Cavarec and C. Delaporte for their help in cDNA libraries construction, D. Depetris for technical assistance for the in situ hybridizations, and C. Lavialle and M. Gunther for helpful discussions and critical reading of the manuscript.

REFERENCES
1. Kuff, E. L., and Lueders, K. K. (1988) Adv. Cancer Res. 51, 183–276
2. Heidmann, O., and Heidmann, T. (1991) Cell 64, 159–170
3. Shen-Ong, G. L. C., and Cole, M. D. (1982) J. Virol. 42, 411–421
4. Lueders, K. K., and Mietz, J. A. (1986) Nucleic Acids Res. 14, 1495–1510
5. Kuff, E. L. (1990) Cancer Cells 2, 398–400
6. Dupressoir, A., and Heidmann, T. (1996) Mol. Cell. Biol. 16, 4495–4503
7. Dupressoir, A., Puech, A., and Heidmann, T. (1995) Biochim. Biophys. Acta 1264, 397–402
8. Voja, P. J., and Barrett, J. C. (1995) Biochim. Biophys. Acta 1242, 29–41
9. Jazwinski, S. M. (1996) Science 273, 54–59
10. Smith, J. R., and Pereira-Smith, O. (1996) Science 273, 63–67
11. Benson, S. (1984) Biotechniques 2, 66–67
12. Ymer, S., Tucker, W. Q. J., Campbell, H. D., and Young, I. G. (1986) Nucleic Acids Res. 14, 5901–5918
13. Reuss, F. U., and Schaller, H. C. (1991) J. Virol. 65, 5702–5709
14. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
15. Chirgwin, J.-M., Przybyla, A. E., McDonald, R. J., and Rutter, W. J. (1979) Biochemistry 18, 5294–5298
16. Church, G. M., and Gilbert, W. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 1991–1995
17. Dessen, P., Fondrath, C., Valencien, C., and Mugnier, C. (1990) Compt. Appl. Biosec. 6, 355–356
18. Mattei, M.-G., Philip, N., Passage, E., Maoan, J.-P., Mandel, J.-L., and Mattei, J.-F. (1985) Hum. Genet. 69, 268–271
19. Mietz, J. A., Fewell, J. W., and Kuff, E. L. (1992) Mol. Cell. Biol. 12, 220–228
20. Lueders, K. K., Fewell, J. W., Morozov, V. E., and Kuff, E. L. (1993) Mol. Cell. Biol. 13, 7439–7446
21. Krayev, A. S., Kramerov, D. A., Skryabin, K. G., Ryskov, A. P., Bayev, A. A., and Georgiev, G. P. (1980) Nucleic Acids Res. 8, 1201–1215
22. Krayev, A. S., Markushova, T. V., Kramerov, D. A., Ryskov, A. P., Skryabin, K. G., Bayev, A. A., and Georgiev, G. P. (1982) Nucleic Acids Res. 10, 745–746
23. Lueders, K. K. (1987) Gene (Amst.) 52, 139–146
24. Leslie, K. B., Lee, F., and Schrader, J. W. (1991) Mol. Cell. Biol. 11, 5562–5570
25. Malvan, T., Biron, R. W., Kaback, D. B., and Denis, C. L. (1992) Genetics 132, 951–962
26. Kato, N., Pfeifer-Olsson, S., Kato, M., Larsson, E., Rydnert, J., Olsson, R., and Cohen, M. (1987) J. Virol. 61, 2182–2191
27. Kato, N., Shimotohno, K., VanLeeuwen, E., and Cohen, M. (1990) Mol. Cell. Biol. 10, 4401–4405
28. Coffin, J. M. (1992) in The Retroviridae (J. A. Levy, ed) pp. 19–50, Plenum Press, New York
29. Liu, A. Y., and Abraham, B. A. (1991) Cancer Res. 51, 4107–4110
30. Feuchter-Murthy, A. E., Freeman, J. D., and Mager, D. L. (1993) Nucleic Acids Res. 21, 135–143
31. Shen-Ong, G. L. C., Morse, H. C., III, Potter, M., and Moshinski, J. F. (1986) Mol. Cell. Biol. 6, 380–392
32. Gonda, T. J., Cory, S., Sobieszczuk, P., Holtzman, D., and Adams, J. (1987) J. Virol. 61, 2754–2763
33. Nilsen, T. W., Maroney, P. A., Goodwin, R. G., Rottman, F. M., Crittenden, L. B., Raines, M. A., and Kung, H.-J. (1985) Cell 41, 719–726
34. Goodwin, R. G., Rottman, F. M., Callaghan, T., Kung, H.-J., Maroney, P. A., and Nilsen, T. W. (1986) Mol. Cell. Biol. 6, 3128–3133
35. Flaynor, P. D., Ono, T., Cutler, N., and Geit, M. J. (1988) Cancer Res. 48, 519–523
36. Dean, R. G., Socher, S. H., and Cutler, R. G. (1985) Arch. Gerontol. Geriat. 4, 43–51
37. Mays-Hoopes, L., Brown, A., and Huang, R. C. C. (1983) Mol. Cell. Biol. 3, 1371–1380
38. Mays-Hoopes, L., Chao, W., Butcher, H. C., and Huang, R. C. C. (1986) Dev. Genet. 7, 65–73
39. Gaubatz, J. W., Arcement, B., and Cutler, R. G. (1991) Mech. Ageing Dev. 57, 71–85
40. Mountz, J. D., and Talal, N. (1983) Immunol. Today 14, 532–536
41. Draper, M. P., Liao, H., Nelsbach, A. H., Mosley, S. P., and Denis, C. L. (1994) Mol. Cell. Biol. 14, 4522–4531
42. Draper, M. P., Salvador, C., and Denis, C. L. (1995) Mol. Cell. Biol. 15, 3487–3495
43. Kass, R. J., Goddard, J. P., and Adams, R. L. (1993) Mol. Cell. Biol. 13, 7372–7377
44. Karpen, G. H. (1994) Curr. Opin. Genet. Dev. 4, 281–291
45. Supakur, P. C., Song, C. S., Jung, M. H., Slomczynska, M. A., Kim, J.-M., Vellanoweth, R. L., Chatterjee, B., and Roy, A. K. (1993) J. Biol. Chem. 268, 2640–26408