Volume 14 Number 3 1986 Nucleic Acids Research

Structural analysis of type II variants within the mouse intracisternal A-particle sequence family

Kira K.Lueders and Judy A.Mietz

Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, MD 20205, USA

Received 4 October 1985; Revised and Accepted 10 January 1986

ABSTRACT

Intracisternal A-particle (IAP) elements are present in multiple copies in the mouse and other rodent genomes. The bulk of this sequence family in Mus musculus consists of 7 Kb long elements, but the majority of IAP sequences involved in known transpositions have been deleted forms. The present study describes a subset of deleted IAP sequences (type II IAP) characterized by insertion of a particular short sequence element (AIIins). AIIins are interspersed and the majority occur as part of the type II IAP elements in the mouse genome. AIIins sequences are absent or in low copy number outside Mus musculus. We have isolated clones containing AIIins from a mouse genomic DNA library and have sequenced three isolates of AIIins and their surrounding IAP sequences to define the detailed structure of type II elements. AIIins are 272, 268 and 264 bp long and 90% homologous in sequence. They are bracketed by 9 bp duplications, suggesting they may be inserted elements. A 75 bp region containing a core enhancer sequence is repeated at the 5' end in type II IAP elements. Insertion into the IAP genome, with potential to encode an integrase function, may have played a role in the amplification of AIIins.

INTRODUCTION

The intracisternal A-particle (IAP) sequence family of the mouse is composed of genetically distinctive endogenous retrovirus-like elements which are present in about 1000 copies per haploid genome of Mus musculus (1-3). These elements represent the vertebrate prototype of retrotransposons, a term proposed by Boeke, et al. (4) to describe eucaryotic transposable elements such as Ty of yeast (5) which have many of the properties of retroviral proviruses (6). Retrotransposons differ from retroviruses primarily in having no extracellular phase (7). Such elements include copia-like elements of Drosophila (8). IAPs as well as Ty and copia VLPs contain poly(A) RNAs which code for major protein components of the particles, including reverse transcriptase (9-12). Sequences complementary to these RNAs exist in the genomes of the respective organisms as repetitive DNA elements. The DNA elements have been shown to be transposable and to cause mutations.

The bulk of the IAP sequence family consists of 7 Kb long elements

(2,13) flanked by long terminal repeats (LTRs) (13). These have been designated type I IAP elements (14). These elements are colinear with the genomic RNA of the particles (13) which codes for the IAP main structural protein p73 (9). Deleted forms of type I IAP elements have also been described (3,13-15), and the majority of the IAP sequences involved in transpositions have been deleted type I elements. A number of such transposition events have been described in myeloma cells (16-23) where IAP RNA transcripts are abundant. Transposition in a leukemia cell line (24) and an instance of germ line insertion have also been reported (25).

Shen-Ong and Cole (14) have reported that approximately 10% of the IAP sequences in the BALB/c mouse genome have a characteristic 0.5 Kb insertion as well as internal deletions of various sizes; these elements have been designated type II IAP elements. These have been further subdivided into three classes IIA, IIB, and IIC on the basis of the sizes of the deletions they contain relative to the type I IAP elements. The major A-particle RNAs in 5 of 7 myelomas were found to be colinear with the type IIB IAP elements, estimated to comprise only 2% of the IAP DNA sequences. DNA of this same class of elements was shown to be amplified and transposed (23).

The deletions in type II IAP elements include most of the region expected to code for the IAP structural protein p73 and part of the putative polymerase region as well. We would like to understand how this particular subset of IAP elements is amplified and transposed and whether they have a retained or novel coding capacity, which might play a role in this process. Both of these objectives require knowledge of the sequence organization of these elements. In the present study we have, therefore, characterized the unique structural aspects of type II IAP elements of the mouse.

METHODS

Sources of DNA

High molecular weight DNA was prepared as previously described (1) from BALB/c mouse embryos and myelomas. The mouse genomic library prepared from the DNA of 14-day-old BALB/c embryos partially digested with HaeIII had an average insert size of 16 Kb (2). The type II-specific sequence was originally cloned by Shen-Ong and Cole (23). A clone (p19/62) containing this sequence along with 345 bp of pBR322 in a 850 bp BamHI fragment was generously provided by Michael Cole. The 0.85 Kb BamHI insert was isolated from p19/62 and inserted into BamHI cut pBR322. This plasmid is designated pMIAIIins, and the BamHI insert pMIAIIins probe. Subsequently a 303 bp

HaeIII fragment was isolated from the BamHI insert in pMIAIIins and designated AIIins probe. The pMIA1 plasmid has been described (2) and was used to prepare the 1.4 Kb PstI probe (see Fig. 1).

Blot hybridization

Mouse DNA blots and library plaques were hybridized with ^{32}P-labeled probes as described (2) except that hybridization and washing were at 55°C (ΔT_m -14°) instead of 65°C and salmon sperm DNA rather than calf-thymus DNA was used as carrier. Heterologous DNA spots were hybridized at 50°C and washed in 6xSSC, 0.1% SDS at 50°C (ΔT_m -48°).

RESULTS

Organization of type II IAP sequences in the genome

Type II IAP elements were initially identified (14) as characteristic 3.5, 2.8 and 2.1 Kb EcoRI/HindIII restriction fragments in mouse genomic DNA using a 0.5 Kb region immediately adjacent to the 5' LTR of a type II IAP element as probe (5' probe, Fig. 1). Although the corresponding sequence is present in type I IAP elements as well, the divergence of this region between the two IAP element types permitted its use as a type II IAP sequence specific probe under stringent hybridization conditions (ΔT_m 5°C). In that study it was estimated that type II IAP elements make up about 10% (85 copies) of the total IAP sequences in the haploid BALB/c mouse genome. Type II elements were subdivided into classes A, B, an C on the basis of the sizes of the deletions they contain; these were 2.7, 3.4 and 4.1 Kb, respectively. The majority (65 copies) of the type II IAP sequences were of the IIA subclass. The amplified IIB genes were estimated to represent only 2% (20 copies) of the total IAP sequences. No attempt was made to quantitate the least abundant IIC subclass. Maps for the three classes of type II IAP elements based on data from previous studies (14,23) as well as this one are shown in Fig. 1. In subsequent work (23), a 0.5 Kb sequence believed to be specific for type II IAP elements was isolated by Shen-Ong, and a clone containing this sequence was given to us by M. Cole. The insertion specific for the type II IAP elements is designated AIIins (A-particle type II insertion) (Fig. 1). A BamHI fragment containing this sequence is designated pMIAIIins probe (see Methods).

Isolation of AIIins clones

We used the pMIAIIins probe to isolate genomic clones containing homologous sequences from a mouse embryo DNA library in Charon 4A λphage. The number of plaques which initially reacted with this probe (0.6%) was far greater

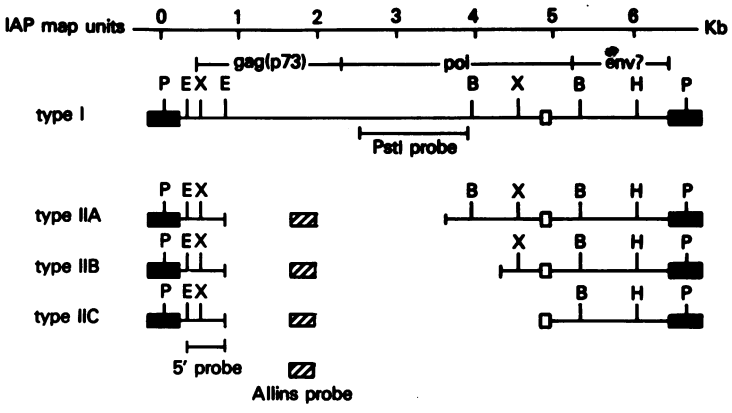

Figure 1. Physical maps of type I and type II IAP elements. The ends of the elements are delineated by LTRs indicated by black boxes. The type II IAP specific insertion (AIIins) is shown by a hatched rectangle which is placed at an arbitrarily chosen position within the deletion in the type II elements. The deletions are indicated by blank regions within the maps; sizes of the deletions are 2.7 Kb for type II A, 3.4 for II B, and 4.1 Kb for II C. Probes used are shown below the maps. The 66 bp block detected 3' to the AIIins by sequencing is indicated as an open box at its usual location on the type I IAP map at position 4.8 and in the type II elements. IAP map positions were defined in ref. 13. Restriction sites are as follows: P, PstI; E, EcoRI; X, XbaI; B, BamHI; H, HindIII.

than expected from the copy numbers reported for type II IAP sequences by Shen-Ong and Cole (14). We considered the possibility that the higher number was the result of the presence of some type I IAP sequence in the pMIAIIins probe. The same plaques were, therefore, hybridized with a type I IAP probe containing primarily sequences not present in type II elements (PstI probe, Fig. 1), and clones reacting with this probe were eliminated. With this correction, 0.4% of the library plaques contained AIIins sequences. Thus, the number of type II sequences in the BALB/c mouse genome appears to be about 4-times higher than the 85 copies previously estimated from blot hybridization (14). The PstI probe reacted with 0.9% of the total library phage plaques, in good agreement with our previous results (2).

All of the clones reacted to some extent with probes representing different parts of type I IAP elements, suggesting that the AIIins sequences in the mouse genome frequently occur in association with IAP sequences.

Characterization of AIIins clones

Eleven clones which reacted strongly with the pMIAIIins probe were selected for further analysis by restriction endonuclease digestion, blot hybridi-

zation, and heteroduplexing. We also analyzed a type IIA IAP element, λMIA80, which we had previously found near a k variable region gene in a genomic clone from a myeloma library.

The subclass of the type II IAP element in each of the clones was determined by Southern blotting and heteroduplex analysis. Three of the clones contained a type IIA IAP element, one clone contained a single type IIC IAP element, and two clones contained two copies of type IIC elements. These clones all reacted strongly with an LTR probe. Five other clones contained type II IAP sequence variants which have not previously been described. Heteroduplex analysis revealed additional insertions of 0.9 to 1.8 Kb in the IAP sequences upstream and/or downstream of AIIins in three of these clones. In two of the clones, the insertions upstream of AIIins were homologous. Two of the variants and the type IIA element in the MIA80 clone did not react with the LTR probe. Previous analysis using restriction digestion would not have detected some of these variants since the sizes of the EcoRI/HindIII fragments produced from them are atypical and variable.

All of the λphage clones containing AIIins sequence reacted with probes representing various more highly repetitive sequence families of the mouse; some clones contained sequences from several different families (Lueders, in preparation).

Sequencing of type II IAP elements

It was determined by Shen-Ong and Cole (14,23) that substantial portions of the type I IAP sequences were deleted in type II IAP elements and that a 500 bp insertion specific for the type II genes was acquired. Pikó and coworkers (26) also described a 500 bp insertion but distinguished small and large deletions which bracketed the insertion. Type I and II IAP elements were also found to be divergent from one another in the regions immediately adjacent to the 5' LTR (14). In this study, we were interested in determining (a) the nature and degree of similarity of AIIins in the type II IAP elements; (b) the position where AIIins was inserted; and (c) precisely which type I IAP sequences were deleted. Regions upstream of the 5' BamHI site of two type IIA IAP elements (clones 3 and 80) and the 850 bp BamHI fragment from pMIAIIins (pMIAIIins probe) were, therefore, cloned into M13 for sequencing. Figure 2 shows these sequences compared with those of a type I IAP element, MIA14 (2,13), and Figure 4 summarizes the structural organization of the type II elements relative to the type I IAP element.

AIIins sequence. The ends of AIIins, as defined by lack of homology with the type I IAP sequence, and the orientation of AIIins relative to type

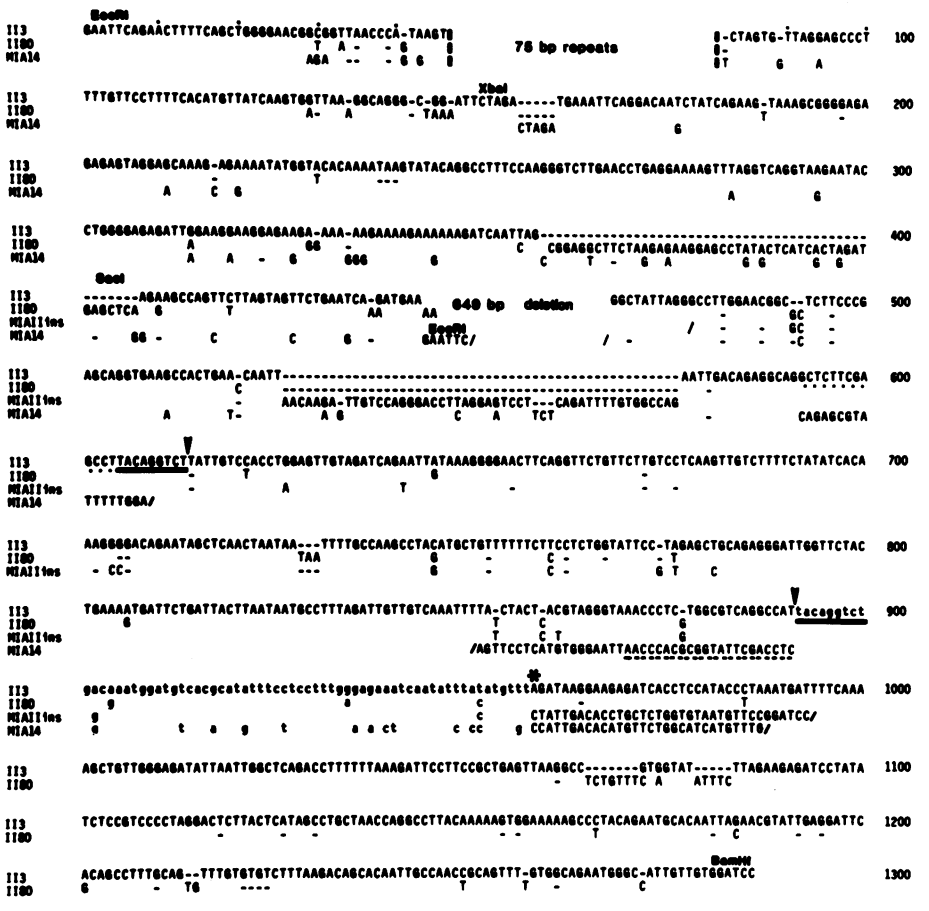

Figure 2. Nucleotide sequences of the 5' regions of type II and type I sequences. Sequences determined by the dideoxynucleotide chain termination on M13 subclones (27) are aligned beginning at the conserved EcoRI sites at IAP map position 0.3. Numbers at the right are bp given for reference and do not represent actual numbers from the EcoRI site. Sequence for the type IIA element in clone 3 is on the top line. Bases identical to those in the clone II3 sequence are represented by blanks in the type IIA element in clone 80, pMIAIIins, and the type I IAP element (MIA14). Bases absent are shown as dashes. In the region where clone 3 and 80 have deletions, the sequence of pMIAIIins is shown. Slashes indicate extent of IAP sequences in pMIAIIins and limits of MIA14 sequence presented; The EcoRI site near the 649 bp deletion is the site at map position 0.8 in MIA14. bp 887-1285 of type II IAP elements are homologous to sequences in type I elements between map positions 3.6 to 3.9, but the type I sequence is not shown here. MIA14 sequences shown from bp 850-986 are located near map position 4.8 on the type I map. The 75 bp region which is repeated is indicated as "75 bp repeats" and is shown in detail in Figure 4. AIIins is bracketed by arrowheads and the 9 bp duplications at the AIIins boundaries in all the clones are heavily

underlined only in the II3 sequence. The 66 bp block is shown in lower case letters. An asterisk marks the junction of this block with IAP sequences which differ in IIA and IIC elements. Sequences on either side of the 649 bp deletion (following the EcoRI site in the type I element) are contiguous in the type II elements.

I sequence were the same for the three type II elements sequenced. Thirteen bp downstream from the point where apparent homology with the type I sequence stops, there is a 9 bp sequence TACAGGTCT (underlined in Fig. 2) which is repeated precisely at the 3' junction of AIIins with type I sequences. Assuming that this repeat represents a target site duplication, then AIIins (delineated by arrowheads, Fig. 2) has sizes of 272, 268, and 264 bp in these elements and has at its ends the inverted repeats TA...AT. This structure suggests that AIIins is an inserted element. The origin of the remaining non-homologous 13 bp (bp 592-604 dotted line Fig. 2) is not known. The three AIIins are homologous in 90% of the bases. Most of the differences are single base changes or deletions/insertions; AIIins in clone 80 contains a 3 bp duplication. The sizes for AIIins are appreciably smaller than the previous estimate of 500 bp, which was derived from measurements of heteroduplexes (3,26, our unpublished data).

5' Flanking IAP sequences. A feature of both type IIA elements sequenced is repetition of 75 bp found near the 5' end (Figs. 3 and 4). This 75 bp region is found four times in clone II3 and twice in clone II80. It includes a concensus enhancer sequence (28) GTGGAAA (enclosed in brackets, Fig. 4), which is present in three of the 75 bp repeat units in clone II3 and in one repeat in clone II80. The ends of the 75 bp sequence are repeated at the beginning and end of the duplicated regions, and a 29 bp segment homologous to an internal portion of the 75 bp region is repeated 71-75 bp downstream in the arrangement shown in Fig. 4. An IAP element, MIARN, located downstream from the ren-2 gene (25) also has this 75 bp region repeated, but the enhancer core is absent. The 75 bp region is present as a single copy in the type I MIA14 element although with a 13 bp deletion and the enhancer core sequence is modified (A to G at position 6). Variability in the number of such repeats may be a factor contributing to restriction fragment length polymorphism in the type II elements (23).

There is a 649 bp deletion of type I IAP sequences 5' to AIIins. The position of this deletion relative to type I sequences differed by only 2 bp in two type II elements (bases 444 and 446 in Fig. 2). This deletion begins very close to the EcoRI site at map position 0.8 in the type I IAP element.

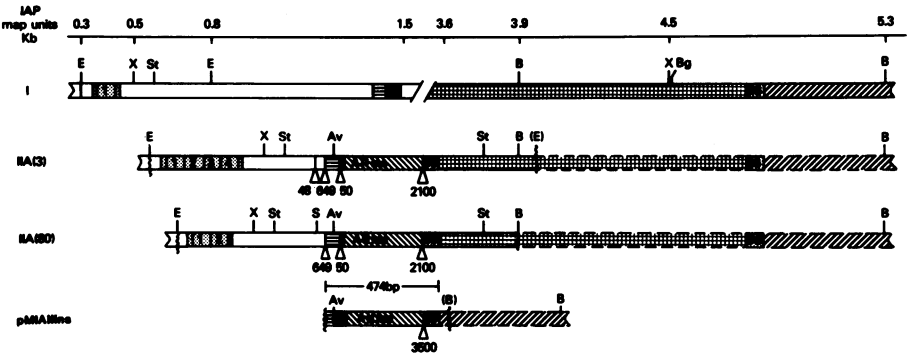

Figure 3. Diagram of the sequence organization of type II IAP elements relative to a type I element. The 5' regions of two type IIA (clone 3 and 80), pMIAIIIins probe and a type I IAP element (MIA14) are compared using sequence data from Fig. 2, to define deletions and the position and size of AIIIins. The type I element is shown at the top from the EcoRI site at map position 0.3 to the 3' BamHI site at 5.3 with approximately 2 Kb between 1.5 and 3.6 omitted for space; homologous regions in type II elements have identical graphic designations. Regions of type II elements sequenced in this study are delineated between vertical wavy lines on the maps. Regions determined to be present by other means (heteroduplex analysis and restriction mapping, as described in ref. 13) are enclosed by dashed lines. Deletions in type II elements relative to the type I element are indicated by triangles with the number of bp deleted. Repeated 75 bp units are numbered 1-4. The apparent size of the AIIIins (474 bp) as estimated from heteroduplex analysis is shown below the map of II80. The 300 bp region upstream of the BamHI site at 3.9 in type I elements has not yet been sequenced in our laboratory. The sequence of another IAP element, clone 10.2, (determined by M. Trounstine, DNAX) was used to establish that the 66 bp block does not occur at 3.6 on the IAP map in type I IAP elements. This block is present in both MIA14 and clone 10.2 at map position 4.8. Clone 10.2 was colinear with MIA14 throughout this region by heteroduplex analysis (J. Mietz and E. Kuff, unpublished data) and homologous by sequence throughout the 1.4 Kb BamHI region. Restriction sites are as follows: E, EcoRI; X, XbaI; St, StuI; S, SacI; Av, AvaI; B, BamHI; and Bg, BglII. Sites created in cloning are in parentheses.

As shown in Fig. 3, this deletion is not contiguous with the large deletions of variable length described earlier (14,23). The region 5' to AIIIins also contains a number of small deletions. This sequence organization suggests that type II elements were derived by a series of complex events.

3' Flanking IAP sequences

Sequences immediately downstream of AIIIins in pMIAIIIins were homologous to type I sequences beginning at map position 4.8. This junction is that expected between AIIIins and type I sequence for type IIC elements (see map Fig. 1). The first 9 bp of the type I sequences at the junction correspond

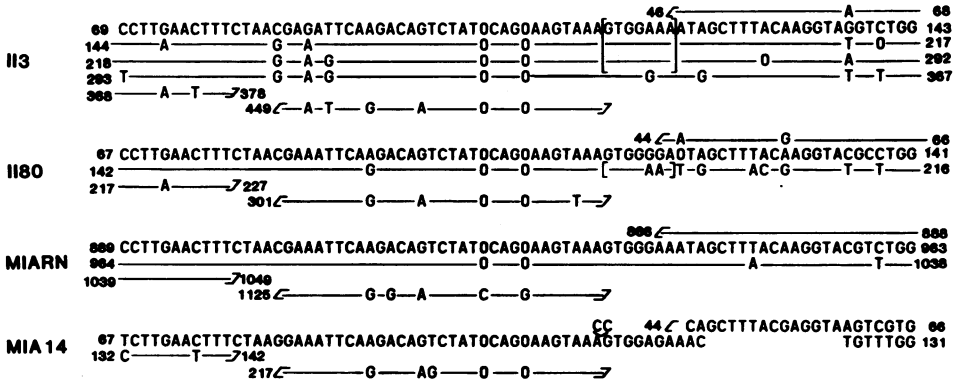

Figure 4. Sequence organization of the repeated 75 bp region. The repeats from two type IIA elements (clones 3 and 80) are compared with those in a deleted type I element (MIARN, see ref. 25) and a type I element (MIA14). Lines indicate bases identical to those on the full line shown, blanks indicate deletions and 0 indicates no base is present. Numbers for II3, II80 and MIA14 sequences indicate bps from the EcoRI site at IAP map position 0.3; numbers for MIARN indicate bps from the sequence in ref. 25. Enhancer core sequences are enclosed in brackets.

to the target site duplication flanking AIIins. Type I (MIA14) sequences upstream of the junction (underlined with broken line in Fig. 2) are 60% homologous to sequences at the 3' end of AIIins for 22 bp (Fig. 2, bp 869-891). This homology may have targeted AIIins to this position in the IAP genome. This same type I sequence is also 60% homologous to the 13 bp present in the type II clones at the AIIins 5' junction (underlined with dots in Fig. 2). The organization of type I IAP sequences downstream from AIIins in type IIA elements is more complex than in type IIC elements. For IIA elements such as those in clones II3 and II80, the junction between AIIins and type I sequences is expected at map position 3.6 (see map Fig. 1). Surprisingly, sequences immediately downstream of AIIins in these clones are identical to those in pMIAIIins for 66 bp (hereafter called the 66 bp block and shown in lower case letters in Fig. 2). Downstream from this 66 bp block the type IIA element sequences are homologous to type I sequences at the expected junction between AIIins and type I sequences for IIA elements near map position 3.6. This junction is indicated by an asterisk in Fig. 2. The type IIA elements, thus, contain a duplication of the 66 bp block, with a new copy present at the junction of AIIins with type I sequences at IAP map position 3.6 and the original copy at map position 4.8, as shown in Fig. 3.

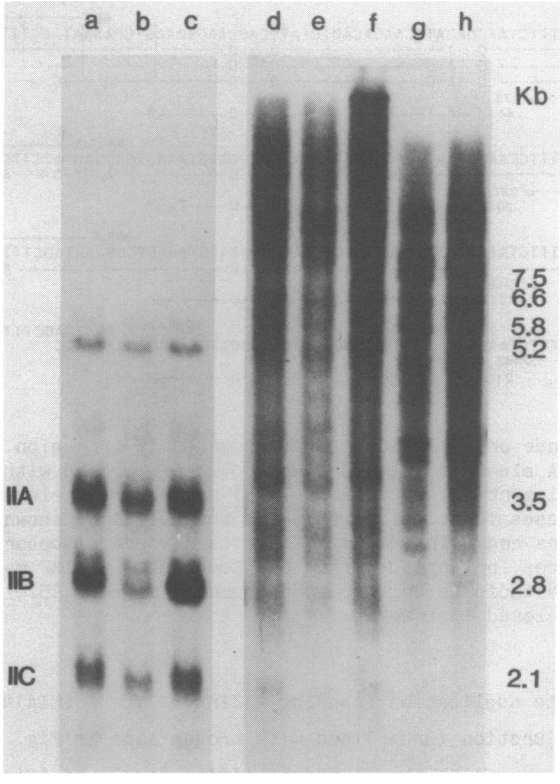

Figure 5. Hybridization of the AIIins probe with mouse genomic DNA digests. 10 µg of BALB/c DNAs cut with EcoRI/HindIII: a. myeloma MOPC-104E, b. embryo, c. myeloma MOPC-315; cut with BamHI: d. MOPC-104E, e. embryo, f. MOPC-315; or cut with EcoRI: g. embryo, h. MOPC 315. The AIIins probe consists of the fragment between the HaeIII sites at 571 and 888 in Fig. 2.

AIIins sequences occur primarily in association with IAP elements in the mouse genome

Sequencing of the pMIAIIins probe revealed that in addition to AIIins, it contained 107 bp of type I IAP sequence 5' to it and 103 bp of type I IAP sequence 3' to it. A 303 bp HaeIII fragment was isolated from this fragment to provide a probe which would be more specific for AIIins. This fragment contained only 27 bp of type I sequence and is designated AIIins probe.

Southern blot hybridization of the AIIins probe to mouse genomic DNA digested with EcoRI/HindIII reveals three prominent discrete bands of 3.5, 2.8 and 2.1 Kb (Fig. 5) as previously described by Sheng-Ong and Cole (14,23).

However, the distribution of the three type II IAP subclasses is different with the AIIins probe than described by Shen-Ong and Cole who estimated that the BALB/c mouse genome contained 65 copies of type IIA IAP elements and 20 copies of type IIB elements; the number of IIC elements was not determined. Scanning of the autoradiographs in Fig. 5 showed that both type IIB and IIC elements are present at levels about one half of the type IIA elements in embryo DNA. It is likely that the copy numbers originally determined for the three type II subclasses on the basis of reaction with the 5' probe (14, see Fig. 1) underestimated the number of type IIB and IIC elements because of differences in the number of 75 bp repeat units. The blot in Fig. 5 shows a 6-fold amplification of type IIB elements in myeloma MOPC-315 and 2-fold amplification in MOPC-104E relative to embryo DNA. The weak band at 5.2 Kb is from reaction of remaining type I sequence in the probe with the approximately 330 copies (14) of type I elements which contain an EcoRI/HindIII fragment of this size. This blot also confirms that the majority of AIIins sequences occur in the mouse genome as part of type II IAP elements.

EcoRI/HindIII digestion produces the simple pattern seen in Fig. 5 because both sites are within the type II elements (Fig. 1). When the DNA is digested with an enzyme which cuts only once within the type II elements, there is no evidence for tandem repeats of the AIIins sequences. As illustrated with the enzyme BamHI which cuts type II elements 3' to AIIins and at variable distances in the 5' flanking DNA, a pattern characteristic of interspersed repetitive DNA is seen. In an EcoRI digest, which cuts IAP elements 5' to AIIins, prominent hybridizing bands larger than expected for the type II elements are superimposed on the background smear of hybridization, indicating that many type II elements are parts of repeated sequence units larger than themselves. The common flanking sequences have been characterized; they include members of the highly repeated L1 family of the mouse (29) and are described in detail elsewhere (Lueders, in preparation).

Repetitive genomic AIIins sequences are mouse specific

Heterologous DNAs bound to nitrocellulose filters as spots were hybridized with the AIIins probe under conditions designed to permit association of divergent sequences. Even with non-stringent washing (T_m-48°C), the only strong positive reaction was with Mus musculus DNA. The DNAs from the Asian mouse species M. cervicolor and M. caroli gave very weak reactions. These same DNAs reacted well with a type I IAP probe and the hybrids had a ΔT_m of 6-8°C (30).

DISCUSSION

In this study we have characterized a particular subset of deleted IAP elements, the type II elements, which contain a characteristic insertion (AIIins). This subset of IAP elements codes for abundant poly(A) RNA transcripts in a number of myelomas. General features of the type II IAP elements which have been defined by our sequencing of three examples and comparison with a type I IAP element are as follows: a) AIIins from three independent isolates have sizes of 272, 268, and 264 bp and are 90% homologous; b) AIIins has the properties of an inserted element; c) the type I sequences deleted represent several non-continguous regions; d) sequences immediately downstream from AIIins in IIA elements represent duplication of a 66 bp block present 1.2 Kb downstream. e) a region of 75 bp found at the 5' end of type I elements is repeated 2 to 4-times in the type II elements.

Recently an IAP element located downstream from the ren-2 gene has been sequenced (25). This element has a large deletion of internal sequences similar to type IIC elements but does not contain AIIins sequences. However, the same 649 bp deletion present in type II elements is also present in this transposed IAP element (base no. 1523 in ref. 25) and this element also has the 75 bp region duplication as shown in Fig. 4. It has been observed that duplication of regions of variable length is a characteristic of a variety of viral enhancer sequences (see 31,32). It will be of interest to determine whether repetition of these 5' sequences is a general feature of IAP elements which contain internal deletions and whether it plays any role in their transcription.

It was surprising that the AIIins sequences are present as a repetitive sequence family only in the genome of Mus musculus. In contrast, the surrounding type I IAP sequences could be easily detected as divergent multigene families in evolutionarily distant rodent genomes (33-37). A computer search for homology between a composite of the three AIIins sequences and those in the rodent category of the GenBank (November 1985) failed to detect any significant homologies.

It is likely that the initial insertion of AIIins occurred upstream of the 66 bp block on the basis of sequence homology between the 3' end of AIIins and the type I sequence. The junction between type I sequences and the 5' end of AIIins is at the same position in 3 elements, making it unlikely that multiple independent insertions occurred. We do not know how the subsequent deletions of type I sequence, which differ by 700 bp increments, led to the the three classes of type II IAP elements. The divergence between

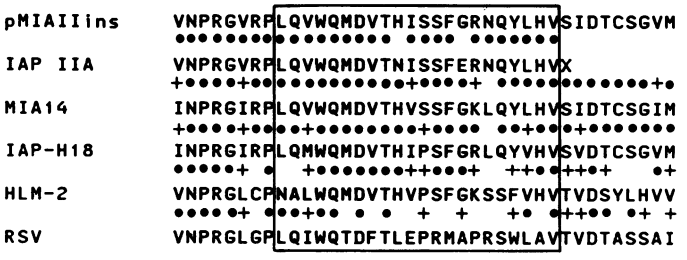

Figure 6. Alignment of homologies between the type II IAP element sequences, the putative pol gene products of the mouse type I IAP element MIA14 (2), the Syrian hamster element H18 (37), the human element HLM-2 (45), and the known pol gene product of RSV (46). Standard amino acid abbreviations as in Ono, et al. (37) are used; x indicates a stop codon. Dots indicate residues identical to those in type II elements and + conservative substitutions (47). The 66 bp block is boxed; these amino acids correspond to positions 630-651 as shown by Ono, et al. in Fig. 4 (37) of reference 37 and to bases 892 to 958 in Fig. 2.

type II IAP elements is similar (~10%) throughout the element, indicating that AIIins and the flanking type I IAP sequences were probably amplified as a unit. Several endogenous retroviral element families and the copia-like element families include deleted forms (7,38,39) but these have not been shown to contain insertions of other repetitive sequences. An exception may be the VL30 sequence family of the mouse, which has many of the properties of retrotransposons. VL30 elements are internally heterogeneous and grossly species specific (40,41). The Ty element family (5) contains no deletion variants analogous to the type II IAP elements.

The best homologies between various retroviral genomes as well as Ty and copia-like elements, have been found in the 3' region of the polymerase gene considered to code for endonuclease (42-44). Recent publication of sequences of IAP related elements from a Syrian hamster (37) and human (45) has extended these homologies to IAP elements also. A short sequence with good homology to the conserved pol sequence defined by Chiu, et al. (42) is located at map position 4.8 in type I IAP elements (M. Trounstine, personal communication), and lies in a region that is present in all the type II elements. We found a more extensive stretch of homology at the amino acid sequence level between the type II IAP elements and the putative polymerase region of IAP-related elements of other species. This homology includes the 66 bp block and is shown in Figure 6. There is only one non-conservative change in 22 amino acids between the mouse type II IAP elements and the hamster IAP

element and four changes between the type II elements and the human element. Amino acid homology with Rous sarcoma virus in this region was about 50% (46). This region lies within the integrase gene of retroviruses and retrotransposons (42-44). The polymerase region expected to code for reverse transcriptase is deleted in all classes of type II IAP elements.

We assume that amplification (transposition) of IAP elements involves reverse transcription of the RNA. Using a yeast Ty element under GAL1 promoter control, Boeke, et al. (4) have demonstrated a link between RNA synthesis and transposition frequency. Correlation between the abundant type IIB element RNA and amplification of the same class of DNA elements in myeloma cells, shows that such a link exists in vertebrate cells as well. Boeke, et al. also found that reverse transcriptase produced from one expressed element can act in trans to cause transposition of other transcripts. It is likely that low levels of 7.2 Kb type I IAP RNA are the source of reverse transcriptase in the myelomas in which type II elements are amplified. The region expected to code for integrase is retained in both type IIA and IIB elements. The LTR promoter is in closer proximity to the integrase sequences in these deleted type II elements than in type I elements, and if expressed, the integrase function could play a role in amplification of the IAP elements. It remains to be determined whether any type II transcripts actually have protein coding activity. This question is of particular interest because of the recent finding that other types of deleted IAP elements can encode protein products with novel physiological functions (48).

ACKNOWLEDGEMENTS

We thank Kirsten Meyer for expert technical assistance. We are especially grateful to Dr. Edward Kuff for many invaluable discussions and for his encouragement. We also thank Dr. Michael Cole for so generously providing the type II IAP probe and Dr. Kevin Moore and Mary Trounstine for providing unpublished sequence data.

REFERENCES

1. Lueders, K.K. and Kuff, E.L. (1977) Cell 12, 963-972.
2. Lueders, K.K. and Kuff, E.L. (1980) Proc. Natl. Acad. Sci. 77, 3571-3575.
3. Ono, M., Cole, M.D., White, A.T. and Huang, R.C. (1980) Cell 21, 465-473.
4. Boeke, J.D., Garfinkel, D.J., Styles, C.A. and Fink, G.R. (1985) Cell 40, 491-500.
5. Roeder, G.S. and Fink, G.R. (1983) in Mobile Genetic Elements, Shapiro, J. Ed, pp 299-328, Academic Press, Inc., Orlando, FL.

6. Varmus, H.E. (1983) in Mobile Genetic Elements, Shapiro, J. Ed, pp 411-503, Academic Press, Inc., Orlando, FL.
7. Baltimore, D. (1985) Cell 40, 481-482.
8. Rubin, G.M. (1983) in Mobile Genetic Elements, Shapiro, J. Ed, pp 329-361, Academic Press, Inc., Orlando, FL.
9. Paterson, B.M., Segal, S., Lueders, K.K. and Kuff, E.L. (1978) J. Virol. 27, 118-126.
10. Shiba, T. and Saigo, K. (1983) Nature 302, 119-124.
11. Garfinkel, D.J., Boeke, J.D. and Fink, G.R. (1985) Cell 42, 507-517.
12. Wilson, S.H., Bohn, E.W., Matsukage, A., Lueders, K.K. and Kuff, E.L. (1974) Biochemistry 13, 1087-1094.
13. Kuff, E.L., Smith, L.A. and Lueders, K.K. (1981) Mol. Cell. Biol. 1, 216-227.
14. Sheng-Ong, G.L.C. and Cole, M.D. (1982) J. Virol. 42, 411-421.
15. Lueders, K., Leder, A., Leder, P. and Kuff, E. (1982) Nature 295, 426-428.
16. Hawley, R.G., Shulman, M.J., Murialdo, H., Gibson, D.M. and Hozumi, N. (1982) Proc. Natl. Acad. Sci. USA 79, 7425-7429.
17. Rechavi, G., Givol, D. and Canaani, E. (1982) Nature 300, 607-611.
18. Kuff, E.L., Feenstra, A., Lueders, K., Smith, L., Hawley, R., Hozumi, N. and Shulman, M. (1983) Proc. Natl. Acad. Sci. USA 80, 1992-1996.
19. Canaani, E., Dreazen, O., Klar, A., Rechavi, G., Ram, D., Cohen, J.B. and Givol, D. (1983) Proc. Natl. Acad. Sci. USA 80, 7118-7122.
20. Kuff, E.L., Feenstra, A., Lueders, K., Rechavi, G., Givol, D. and Canaani, E. (1983) Nature 302, 547-548.
21. Cohen, J.B., Unger, T., Rechavi, G., Canaani, E. and Givol, D. (1983) Nature 306, 797-799.
22. Gattoni-Celli, S., Hsiao, W-L.W., Weinstein, I.B. (1983) Nature 306, 795-796.
23. Sheng-Ong, G.L.C. and Cole, M.D. (1984) J. Virol. 49, 171-177.
24. Ymer, S., Tucker, W.Q.J., Sanderson, C.J., Hapel, A.J., Campbell, H.D. and Young, I.G. (1985) Nature 317, 255-258.
25. Burt, D.W., Reith, A.D. and Brammar, W.J. (1984) Nucleic Acids Res. 12, 8579-8593.
26. Pikő, L., Hammons, M.D. and Taylor, K.D. (1984) Proc. Natl. Acad. Sci. USA 81, 488-492.
27. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
28. Laimins, L.A., Kessel, M., Rosenthal, N. and Khoury, G. (1983) in: Enhancers and Control Elements, Gluzman, Y. and Shenk, T. (Eds.), pp. 28-37 Cold Spring Harbor, New York.
29. Voliva, C.F., Jahn, C.L., Comer, M.B., Hutchinson, C.A. and Edgell, M.H. (1983) Nucleic Acids Res. 24, 8847-8859.
30. Kuff, E.L., Lueders, K.K. and Scolnick, E.M. (1978) J. Virol. 28, 66-74.
31. Herr, W. and Gluzman, Y. (1985) Nature 313, 711-714.
32. Michalides, R., Wagenaar, E. and Weijers, P. (1985) Mol. Cell. Biol. 5, 823-830.
33. Lueders, K.K. and Kuff, E.L. (1981) Nucleic Acids Res. 9, 5917-5930.
34. Suzuki, A., Kitasato, H., Kawakami, M and Ono, M. (1982) Nucleic Acids Res. 10, 5733-5745.
35. Lueders, K.K. and Kuff, E.L. (1983) Nucleic Acids Res. 11, 4391-4408.
36. Ono, M., Kitasato, H., Ohishi, H., Motobayashi-Nakajima, Y. (1984) J. Virol. 50, 352-358.
37. Ono, M., Toh, H., Myata, T. and Tomoe, A. (1985) J. Virol. 55, 387-394.
38. Sterle, P.E., Rabson, A.B., Bryan, T. and Martin, M.A. (1984) Science 225, 943-947.

39. Soe, L.H., Devi, B.G., Mullins, J.I., and Roy-Burman, P. (1983) J. Virol. 46, 829-840.
40. Keshet, E., Shaul, Y., Kaminchik, J. and Aviv, H. (1980) Cell 20, 431-439.
41. Keshet, E. and Itin, A. (1982) J. Virol. 43, 50-58.
42. Chiu, I.M., Callahan, R., Tronick, S.R., Schlom, J. and Aaronson, S.A. (1984) Science 223, 364-370.
43. Mount, S.M. and Rubin, G.M. (1985) Mol. Cell Biol. 5, 1630-1638.
44. Toh, H., Kikuno, R., Hayashida, H., Miyata, T., Kugimiya, W., Inouye, S., Yuki, S. and Saigo, K. (1985) EMBO J. 4, 1267-1272.
45. Callahan, R., Chiu, I-M., Wong, J.F.H., Tronick, S.R., Roe, B.A., Aaronson, S.A. and Schlom, J. (1985). Science 228, 1208-1211.
46. Schwarts, D.E., Tizard, R. and Gilbert, W. (1983) Cell 32, 853-869.
47. Schwartz, D.M. and Dayhoff, M.O. (1978) Altas of Protein Sequence and Structure, Vol. 5, suppl. 3 (ed. Dayhoff, M.O.) 353-358.
48. Martens, C.L., Huff, T.F., Jardieu, P., Trounstine, M.L., Coffman, R.L., Ishizaka, K. and Moore, K. (1985) Proc. Natl. Acad. Sci. USA 82, 2460-2464.