The EMBO Journal vol.5 no.5 pp.919–924, 1986

Sequence, topography and protein coding potential of mouse *int-2*: a putative oncogene activated by mouse mammary tumour virus

R.Moore, G.Casey[1,2], S.Brookes[1], M.Dixon, G.Peters[1] and C.Dickson

Imperial Cancer Research Fund Laboratories, PO Box 123, Lincoln's Inn Fields, London WC2A 3PX, and [1]Imperial Cancer Research Fund Laboratories, St. Bartholomew's Hospital, Dominion House, Bartholomew Close, London EC1A 7BE, UK

[2]Present address: Department of Microbiology and Molecular Genetics, University of California, Irvine, CA 92717, USA

Communicated by L.Crawford

A major proportion of carcinomas induced by mouse mammary tumour virus (MMTV) show evidence for proviral activation of a cellular gene, *int-2*, on chromosome 7. The sequence of 7869 bp of DNA spanning the transcription unit of *int-2* was determined and compared with that of a series of *int-2*-specific cDNA clones derived from mammary tumour RNA. The predicted positions of intron–exon boundaries, established by alignment of cDNA and chromosomal DNA sequences, indicate that the gene comprises at least three exons. An open reading frame capable of encoding a protein of 245 amino acids with an estimated mol. wt of 27 kd, is flanked by substantial non-coding segments at both 5′ and 3′ ends. Comparison of the chromosomal DNA sequence and the predicted amino acid sequence with available data-bases has revealed no homology to other known genes. These results are discussed in relation to the status of *int-2* as a candidate proto-oncogene.

Key words: int-2 cDNA/*int-2* sequence/MMTV/oncogene/provirus

Introduction

The oncogenic properties of retroviruses can be mediated by several distinct mechanisms, the most clearly established being the transduction or insertional mutagenesis of cellular proto-oncogenes (Bishop, 1983; Varmus, 1984). While ~50 isolates of acutely transforming retroviruses, carrying transduced cellular genes, have now been described, the vast majority of naturally occurring retroviruses do not contain recognizable oncogenes and are only tumorigenic after prolonged latency (Teich *et al.*, 1982). There is now a considerable weight of evidence to suggest that many of these latter viruses exert their tumorigenic properties by integrating within or adjacent to cellular proto-oncogenes thereby perturbing either the structure or level of expression of such a gene (Varmus, 1982). In some cases, the target oncogene is familiar (Corcoran *et al.*, 1984; Fung *et al.*, 1981; Hayward *et al.*, 1981: Nilsen *et al.*, 1985; Noorii-Daloii *et al.*, 1981: Payne *et al.*, 1982; Selten *et al.*, 1984; Shen-Ong *et al.*, 1984; Steffen, 1984), while in others, the existence of an oncogene is only inferred by the repeated finding of proviral DNA within a limited chromosomal domain in virally induced neoplasias. In carcinomas induced by mouse mammary tumour virus (MMTV) two such domains have been described, designated *int-1* and *int-2* (Nusse and Varmus, 1982; Peters *et al.*, 1983).

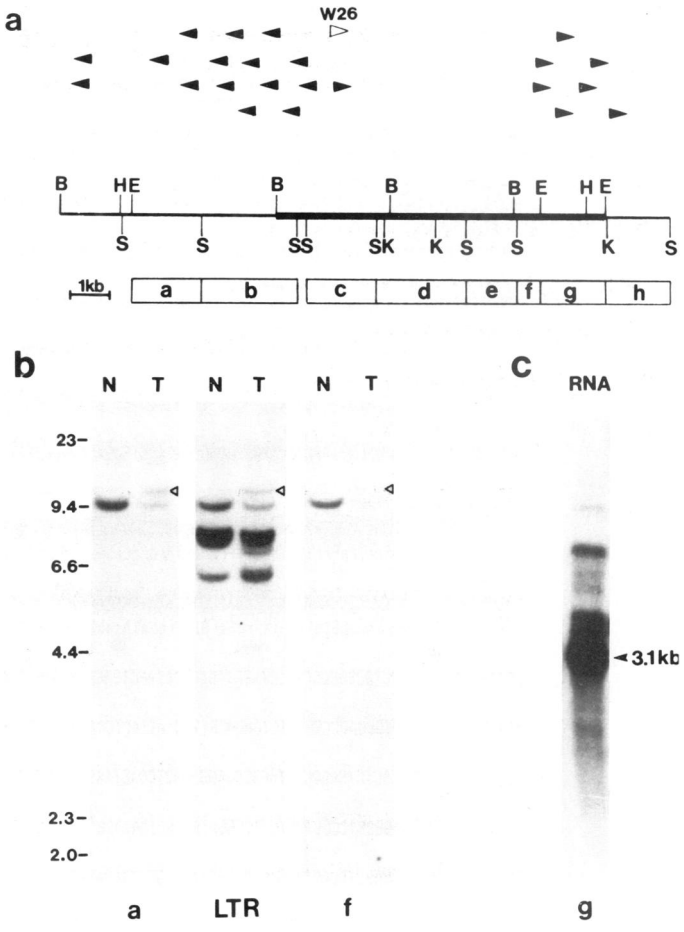
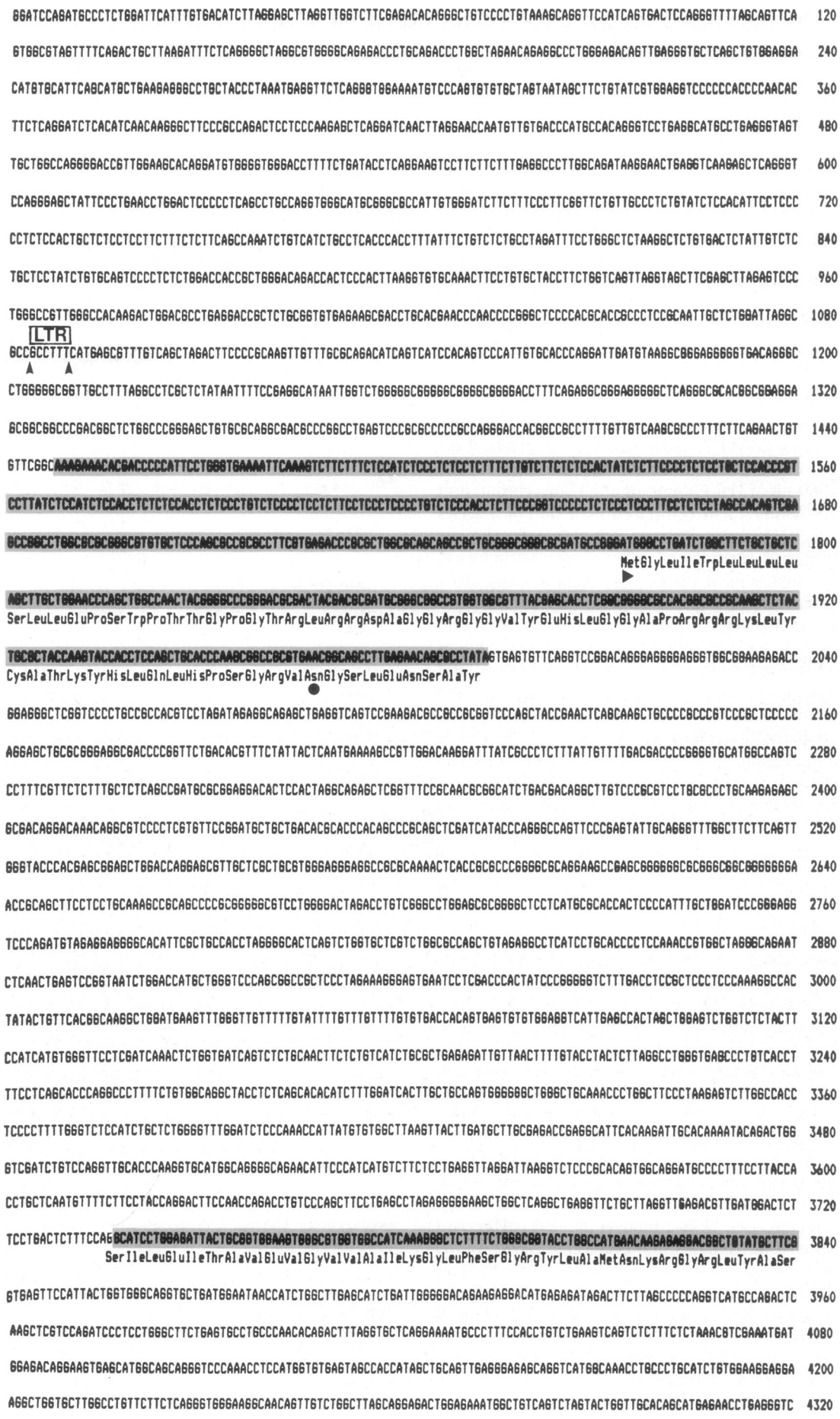


Fig. 1. Structure and expression of the *int-2* gene in mammary tumour W26. (**a**) A physical map of the mouse *int-2* locus is presented in which the sites of cleavage for the restriction enzymes *Bam*HI (B), *Eco*RI (E), *Hin*dIII (H), *Kpn*I (K) and *Sac*I (S) are indicated. The bold line delineates DNA sequenced in this study. Boxed segments below the map depict the *Eco*RI/*Sac*I restriction fragments, designated 'a' to 'h', which have been exploited as *int-2*-specific probes. Arrowheads identify the position and transcriptional orientation of MMTV proviruses mapped within *int-2* in 24 independent mammary tumours, with the open arrowhead corresponding to the solo LTR in tumour W26. The type of data used to map proviral insertions is illustrated with reference to tumour W26 in (**b**). Note that the physical map is shown inverted relative to previous presentations in which the direction of transcription was incorrectly assigned (Dickson *et al.*, 1984). (**b**) Southern blotting analyses were performed on *Eco*RI-digested DNA from either normal (N) or tumour (T) tissue, fractionated on a 0.8% agarose gel and hybridised sequentially with a series of *int-2*- or MMTV-specific probes. The figure shows results obtained using (from left to right), the *int-2* a, MMTV LTR and *int-2* f probes. The novel fragment (◁) detected with all three probes was estimated to be ~1.3 kb larger than the normal 10.0-kb *Eco*RI fragment from the *int-2* locus, consistent with insertion of a single LTR. Numbers to the left of the figure indicate the sizes (kb) and relative position of fragments of *Hin*dIII-digested lambda DNA employed as standards. (**c**) A typical Northern blot is shown in which 5 μg of polyadenylated RNA from tumour W26, fractionated on a 1.2% agarose–formaldehyde gel, was hybridized with the *int-2* g probe. The predominant RNA species was ~3.1 kb in length as estimated relative to rRNAs and the MMTV-specific 35S and 24S mRNAs.


LTR
MetGlyLeuIleTrpLeuLeuLeuLeu
SerLeuLeuGluProSerTrpProThrThrGlyProGlyThrArgLeuArgArgAspAlaGlyGlyArgGlyGlyValTyrGluHisLeuGlyGlyAlaProArgArgArgLysLeuTyr
CysAlaThrLysTyrHisLeuGlnLeuHisProSerGlyArgValAsnGlySerLeuGluAsnSerAlaTyr
SerIleLeuGluIleThrAlaValGluValGlyValValAlaIleLysGlyLeuPheSerGlyArgTyrLeuAlaMetAsnLysArgGlyArgLeuTyrAlaSer

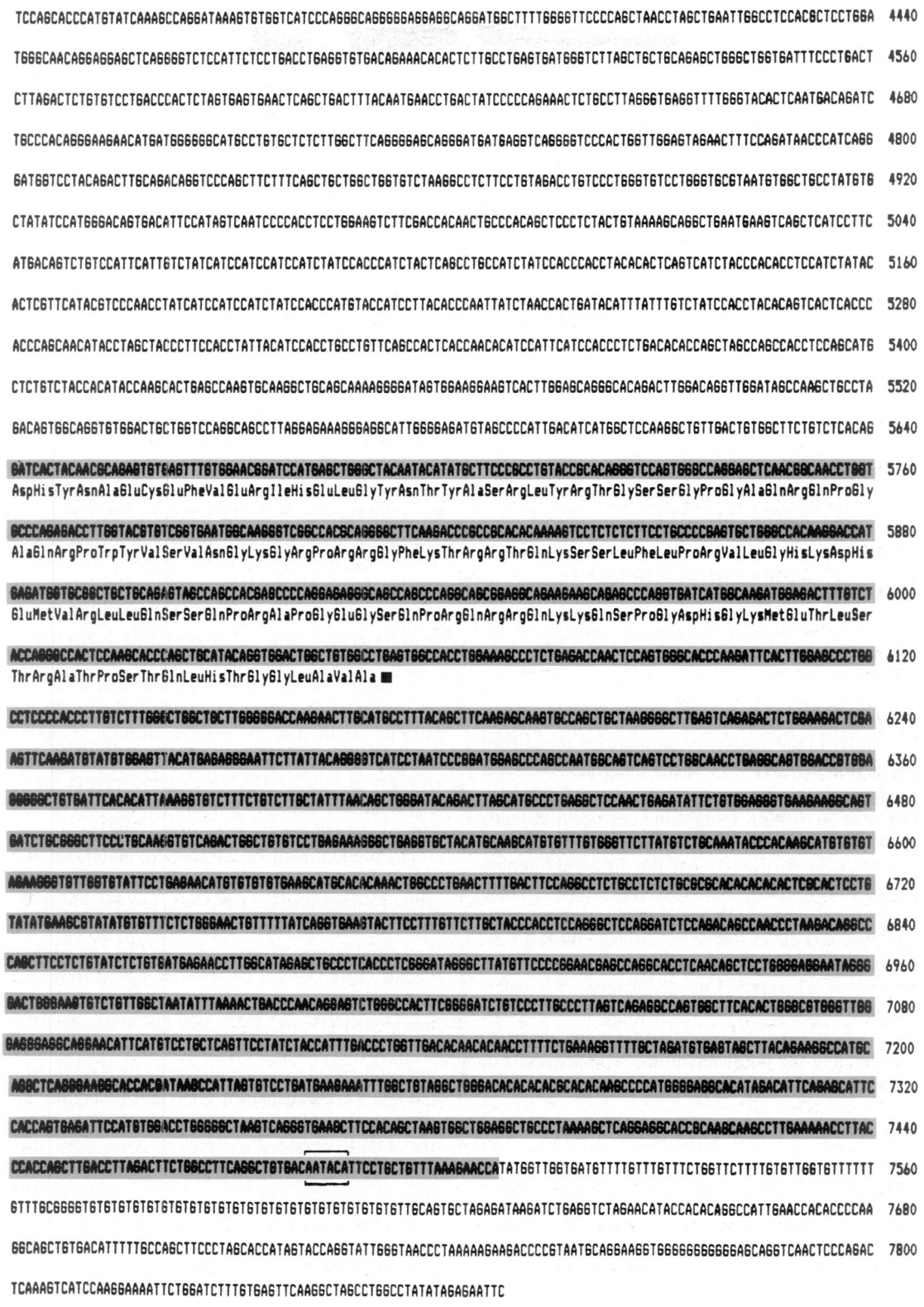


Fig. 2. The sequence of the mouse *int-2* genomic DNA, cDNA and predicted protein product. The sequence is presented of 7869 bases of BALB/c mouse DNA extending between a *Bam*HI site and *Eco*RI site in the *int-2* locus (see Figure 1). The DNA sequence was determined on both strands by M13 cloning and chain termination procedures. Shaded sections correspond to the sequences contained in cDNA clones C1, C2 and C3 (Figure 3). The predicted amino acid sequence of a potential *int-2* protein is shown below the relevant segments of cDNA. The symbols ▲ identify the position of the solo MMTV LTR in tumour W26 DNA and delineate the six-base duplication observed in the genomic sequence in this tumour. ▶ identifies the presumed initiator methionine codon, ■ the TGA stop codon and ● a potential asparagine-linked glycosylation site. The presumed polyadenylation signal AATACA (nucleotides 7482–7487) is boxed.

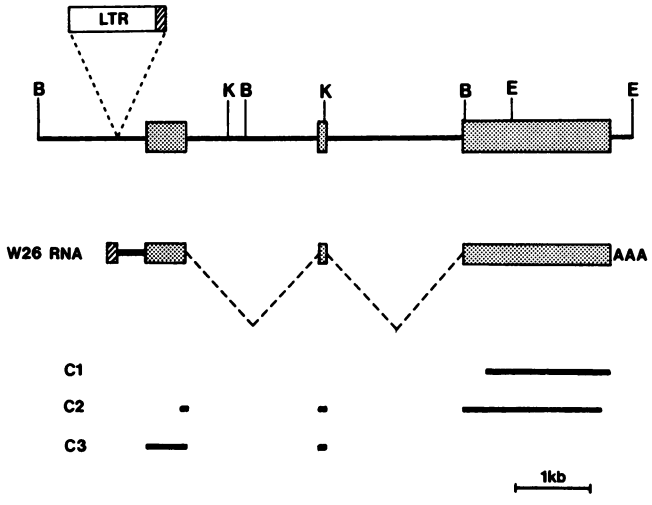


Fig. 3. Topography of the *int-2* locus and cDNA clones from tumour W26. A diagramatic representation of the proposed exon structure (stippled boxes) of the *int-2* gene is shown superimposed on the linear map of sequenced genomic DNA in which the cleavage sites for *Bam*HI (B), *Eco*RI (E) and *Kpn*I (K) are indicated. The position of the MMTV LTR is shown, together with the portion of LTR sequences (cross-hatched) normally transcribed at the 5′ end of viral RNA. The proposed structure of the major *int-2* transcript in tumour W26 therefore comprises 133 nucleotides from the LTR, 364 nucleotides of genomic DNA whose designation as intron or exon remains uncertain, and the three blocks of exon sequence represented in the C1, C2 and C3 clones. The approximate extents of these cDNAs are indicated by the bold lines.

We have previously established the existence of the *int-2* region by showing that >50% of the mammary tumours arising in BR6 mice contain an acquired MMTV provirus integrated within a defined 25-kb domain of DNA on mouse chromosome 7 (Dickson *et al.*, 1984; Peters *et al.*, 1983). Integration can occur at many different sites within this region, but the distribution and transcriptional orientation of the proviruses detected in different tumours is non-random, generally falling into two oppositely oriented clusters separated by a region of DNA which is transcriptionally active in these tumours (Dickson *et al.*, 1984). Working on the premise that this region of DNA must contain a cellular gene and that it may represent a potentially novel oncogene, we undertook its complete structural characterization. Here we report the sequence of ~8 kb of cellular DNA spanning the locus, the isolation of cDNA clones and the predicted amino acid sequence of a protein encoded by the *int-2* gene.

Results

DNA sequence of the mouse int-2 locus

In a previous report, we employed RNA blot hybridisation to establish the approximate limits of the *int-2* transcription unit as expressed in tumours bearing an MMTV provirus within the *int-2* domain (Dickson *et al.*, 1984). Such analyses suggested that RNA was transcribed from sequences within an ~8-kb segment of chromosomal DNA as depicted by the bold line in Figure 1a. As a first step in the more detailed characterisation of the locus, we determined the sequence of this stretch of genomic DNA. A series of overlapping bacteriophage clones had been recovered from libraries of BALB/c mouse DNA and the region of interest was present in a continuous insert in one such clone, λ3. However, to facilitate the initial phases of the sequencing, the DNA was subdivided into suitably sized pieces by digestion with either *Sac*I, *Eco*RI or *Bam*HI. These segments were randomly fragmented by sonication or partial digestion with DNase I, blunt-ended, and ligated into the M13 mp8 or mp9 vectors (Messing and Vieira, 1982). In some instances, specific restriction fragments were cloned to generate sequence overlaps or to clarify regions which proved refractory to random cloning. All sequences were determined using dideoxynucleotide chain termination methods (Bankier and Barrell, 1983). In this way, a continuous sequence of 7869 bases was established for both DNA strands (Figure 2).

cDNA cloning of int-2 RNA

To establish the detailed architecture of the *int-2* gene, two avenues were open to use: to determine the intron–exon boundaries by S1 mapping or to isolate cDNA clones. Both of these strategies were likely to be compromised by the low levels of *int-2* RNA obtainable, estimated to be of the order of 10 copies per cell in primary tumours and not detectable in normal tissues (Dickson *et al.*, 1984). At the inception of these studies, the most abundant source of *int-2* RNA was a particular BR6 mouse mammary tumour, designated W26, which not only expressed higher levels of *int-2* than other tumours then available, but maintained these levels after transplantation into syngeneic mice. The probable reason for elevated transcription became apparent when DNA from tumour W26 was subjected to detailed Southern blotting analysis (Figure 1b). A discontinuity in one of the two *int-2* alleles was found to be consistent with the presence of a single MMTV long terminal repeat (LTR) in what is presumed to be a promoter insertion mode. This has subsequently been verified by cloning the appropriate 11.3-kb *Eco*RI fragment (indicated ◁ in Figure 1b) and sequencing through the LTR to establish both its orientation and precise location relative to the genomic DNA sequence (Figure 2 and data not shown). A duplication of 6 bp of cellular DNA was generated at the site of insertion as previously reported for the integration of complete MMTV proviruses (Majors and Varmus, 1981).

The predominant *int-2* transcript in tumour W26 was a 3.1-kb species (Figure 1c) though several other species were also detectable. Complementary DNA was prepared from W26 polyadenylated RNA, primed with oligo(dT), and introduced by conventional procedures into the *Pst*I site of the pUC9 plasmid vector. From ~10^5 colonies screened using an *int-2* f probe, six positive recombinants were identified, the longest of which, designated C1 and C2, had cDNA inserts of 1.5 and 2.0 kb, respectively. Both these inserts were excised from the plasmid and sequenced via a combination of random and directed M13 clones. The sequence of C1 could be readily aligned with that of genomic DNA from nucleotides 5983 to 7508, at which point the cDNA contained a stretch of A residues, presumably derived from the poly(A) tract at the 3′ end of the mRNA. A variant polyadenylation signal, AATACA, located 20 bases upstream of the poly(A) tract, is consistent with this interpretation (Figure 2). Clone C2 on the other hand terminated 46 bases from the poly(A) tract but extended the homology with genomic DNA to map position 5641. In addition, C2 showed two further blocks of homology with the chromosomal sequence, from nucleotides 1953 to 1993 and 3737 to 3840 (see Figure 3). Since the points of discontinuity between the genomic and cDNA sequences were consistent with the consensus signals for splice donor and acceptor sites, we concluded that the 2 kb of mRNA encompassed by C1 and C2 was derived from three discrete exons.

To obtain more information on sequences 5′ of the C2 clone, a second library was constructed by priming cDNA synthesis from a synthetic oligonucleotide corresponding to the final 18

bases of the small exon (nucleotides 3823–3840). This oligonucleotide was chosen because of its proximity to a unique *Kpn*I site which would facilitate reconstruction of cDNA sequences from different clones (Figure 3). Screening with the *int-2* c probe identified four positive recombinants, all of which began at the expected primer site and excluded the intron sequences between nucleotides 1994 and 3736. The longest of these cDNA clones, designated C3, extended the cDNA sequence a further 506 bases upstream of C2 with no further discontinuities relative to genomic DNA.

Presence of a long open reading frame in the int-2 cDNA sequence

Although the cDNA clones obtained did not extend to the precise 5′ end of the major *int-2* RNA transcript, the composite sequence from C1, C2 and C3 was examined for protein coding potential. A single long open reading frame was identified, beginning at a methoinine codon at position 1774 in the 5′ exon and terminating at a TGA codon at position 6052 in the 3′ exon (Figure 2). The continuous stretch of 735 bases in-frame would code for a 245-amino acid product with an estimated mol. wt of 27 000 daltons. The predicted sequence of this protein is shown aligned with the DNA sequence in Figure 2.

Discussion

The continuous *int-2* cDNA sequence reconstructed from the C1, C2 and C3 clones spans ~2.5 kb of an RNA believed to represent the major 3.1-kb transcript in tumour W26. Assuming that poly(A) at the 3′ end could contribute a further 150–200 nucleotides, we can predict the nature of the remaining sequences with some confidence. The single MMTV LTR in W26 tumour DNA is located 364 nucleotides upstream of the 5′ end of the cDNA sequence, and preliminary data from primer extension analyses are consistent with a major transcription start site at the known MMTV promoter 133 nucleotides into the LTR (Majors and Varmus, 1981). Predictably, tumour W26 expresses hybrid mRNAs, comprising both MMTV LTR and *int-2* sequences, as demonstrated by sandwich blotting techniques (data not shown; Payne *et al.*, 1982). We therefore conclude that the major *int-2* RNA transcript expressed in this tumour initiates within the LTR, extends continuously through the three exons defined by cDNA cloning, and terminates at the polyadenylation site defined by the end of clone C1 (Figure 3).

An RNA with this structure would have an estimated size of ~3.1 kb, in close agreement with predictions from Northern blotting analysis (Figure 1). Based on a theoretical translation of these sequences, the longest open reading frame identified is only 735 nucleotides, made up of 220, 104 and 411 nucleotides, respectively from each of the three exons (Figure 2). The potential product of translation would be a relatively basic protein with a calculated mol. wt of 27 000 daltons. Neither the amino acid nor the nucleic acid sequences appear related to any previously characterised entry in the available data bases (EMBL and GENBANK). Nevertheless, some aspects of the primary amino acid sequence are worthy of comment: a high proportion of arginine and lysine residues, frequently in pairs or clusters, only two cysteine residues, a single potential asparagine-linked glycosylation site and a hydrophobic amino terminus (Figure 2). However, it is debatable whether the hydrophobic domain is extensive enough to serve as a transmembrane signal sequence (Watson, 1984). Resolution of such issues awaits the preparation of *int-2*-specific antisera, raised against either synthetic peptides or bacterially expressed products, and experiments are in progress along these lines.

Although many of the details of *int-2* gene expression are yet to be resolved, the cDNA structure and predictions regarding a potential product are consistent with some additional observations on the locus. The most compelling is that while several MMTV proviruses have now been mapped within what appear to be 5′- and 3′-untranslated regions leading to abnormally terminated RNA transcripts (Figure 1a and data not shown) in no case does a provirus interrupt the protein coding domain. We would argue therefore that the integrity of the product may be essential for tumorigenicity, athough we cannot at this stage exclude point mutations or small alterations. Two single base changes were observed between the genomic and cDNA sequences (not shown) but these were presumed to reflect either strain differences (BALB/c versus BR6) or cloning artifacts. At the gross level therefore, activation of *int-2* does not require the formation of a truncated product, as described for example in the insertional mutagenesis of *erb-B* and *myb* (Nilsen *et al.*, 1985; Shen-Ong *et al.*, 1984).

The assignment of the initiator methionine codon at nucleotide 1774 was based on a number of considerations. In the first place, other potential initiator codons in the genomic DNA sequence upstream of this ATG are followed by stop codons in all three reading frames; secondly, the DNA sequence in this region includes substantial CT and GC-rich elements which would be unusual in a coding domain; and thirdly, one example of provirus insertion maps even closer to the proposed initiation site than the solo LTR (see Figure 1a). However, while tumour W26 was chosen to facilitate cDNA cloning, it does represent an unusual situation and we cannot exclude the possibility that proviruses integrated in a promoter insertion mode may have created pseudo-exon sequences by subverting the normal RNA splicing patterns.

If additional exon sequences do exist 5′ of the LTR in tumour W26, they are almost certainly included within the genomic DNA sequenced in this study. As illustrated in Figure 1a, there is a rather sharp cut off in the mapped positions of proviruses in the so-called 'enhancer mode', in the opposite orientation to that of the *int-2* gene. Since all of these insertions lead to expression of the same *int-2* RNA transcripts, the proviruses presumably act on the same transcriptional control elements. However, it is now clear that these tumours express several discrete *int-2* transcripts, raising the possibility of alternative start sites. Although there are no obvious TATA boxes, other sequence motifs, such as GC-rich clusters could conceivably serve this function (Melton *et al.*, 1984; Reynolds *et al.*, 1984). With hindsight, therefore, the choice of tumour W26 with a single predominant start site in the MMTV LTR may have been advantageous in simplifying the pattern of cDNA clones expected.

Other opportunities exist for creating multiple RNA transcripts, such as alternate exon usage, or alternate usage of polyadenylation sites. The AATACA sequence identified here is a rare but apparently functional variant of the consensus (Mason *et al.*, 1985) but no obvious alternative is apparent in >1.5 kb downstream of the *int-2* termination codon. However, the situation is further confused by the finding that the choice of probe can have a bearing on the pattern of *int-2* transcripts observed. At present, we have insufficient information to provide a satisfactory explanation. By using primer extension analyses to define the 5′ ends of the various messages and by identifying a source of *int-2* RNA in which the normal gene is expressed without the

influence of an MMTV provirus, we hope to establish a more detailed rationale for the control of mouse *int-2* gene expression.

Materials and methods

Analysis of mammary tumour DNA

The methods employed in the extraction of high mol. wt DNA, its analysis by restriction enzyme digestion and gel electrophoresis, and conditions for hybridization and washing of Southern blots have been described in detail in previous publications (Peters *et al.*, 1983). Cloned restriction fragments derived by digestion of *int-2* DNA with *Eco*RI and *Sac*I were used as hybridization probes, designated *a*–*h* in Figure 1. Alternatively, MMTV-specific sequences were visualized using a 1.4-kb *Pst*I fragment encompassing the proviral LTR. Probes were labeled by nick translation with [α-^{32}P]dCTP (Amersham).

Extraction and analysis of mammary tumour RNA

Total RNA was extracted from frozen tumour tissue by homogenization in guanidine thiocyanate and centrifugation through CsCl (Chirgwin *et al.*, 1979). The polyadenylated RNA fraction was recovered by affinity chromatography on poly(U)–Sepharose as previously described (Dickson *et al.*, 1984). Northern blotting analysis was performed on 5 μg aliquots, resolved by electrophoresis through formaldehyde–agarose gels and transferred to nitrocellulose in 20 × SSC (Thomas, 1980). Conditions for hybridization and washing of filters were as for DNA blots, except that hybridization was at 50°C in 50% formamide.

M13 cloning and DNA sequencing

The procedures used for cloning into M13 vectors and DNA sequencing by the dideoxynucleotide chain termination method were as described by Bankier and Barrell (1983). Essentially, purified DNA fragments were self-ligated to form concatemers, randomly cleaved by sonication or partial digestion with DNaseI to generate pieces of between 50 and 500 bp, and blunt ended by repairing the ends with T4 DNA polymerase. Size-selected fragments of 200–400 nucleotides in length were then ligated into the *Sma*I site of the M13 mp8 or mp9 vectors. Alternatively, specific restriction fragments or DNA cleaved by combinations of frequent cutter restriction enzymes were cloned into appropriate sites in the vectors. Recombinant phage plaques were propagated on *Escherichia coli* JM101 and single-stranded DNA extracted from the progeny phage by conventional procedures. DNA synthesis was primed from a synthetic 15 nucleotide primer (BRL Laboratories) using [^{35}S]dATP (Amersham) as labeled precursor. Sequences were compiled using the DBUTIL programme (Staden, 1980) and subsequently compared with both the EMBL and GENBANK data-bases.

cDNA cloning of mammary tumour RNA

Approximately 10 μg of polyadenylated RNA from tumour W26 was used to direct oligo(dT)-primed complementary DNA synthesis using the conditions described by Gubler and Hoffmann (1983). Following treatment of the RNA–DNA hybrid with RNase H, the second DNA strand was synthesized using a combination of DNA polymerase I and reverse transcriptase. Double-stranded DNA products of appropriate length were isolated by preparative gel electrophoresis and introduced into the *Pst*I site of the plasmid vector pUC8 by G-C tailing. Approximately 15–20 dG residues were added to the synthetic DNA and a comparable number of dC residues to the vector using terminal transferase. The recombinant plasmids were then used to transform competent *E. coli* DH1. The transformation efficiency was estimated at 5 × 10^7 colonies per μg of DNA for supercoiled plasmid DNA. Approximately 10^5 recombinant colonies were screened by hybridization of *int-2*-specific probes, essentially as described by Grunstein and Hogness (1975).

A second cDNA library was constructed in which a synthetic oligonucleotide (nucleotides 3823–3840 in Figure 2) was used as a primer. Conditions were otherwise the same except that the *int-2* c probe was used in order to identify clones extending towards the 5′ end of the gene. DNA from positive cDNA clones was characterised by restriction enzyme mapping and selected fragments were subjected to both random and directed M13 cloning and DNA sequencing as described above.

Acknowledgements

We thank L.Crawford, I.Kerr, M.Parker, J.Witkowski and J.Wyke for helpful comments on the manuscript and A.Sterlini for assistance in its preparation.

References

Bankier,H.T. and Barrell,B.G. (1983) in Flavell,R.A. (ed.), *Techniques in the Life Science*. Elsevier Press, Ireland, B508, pp. 1–34.
Bishop,J.M. (1983) *Annu. Rev. Biochem.*, **52**, 301–354.
Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J. (1979) *Biochemistry*, **18**, 5294–5299.
Corcoran,L.M., Adams,J.M., Dunn,A.R. and Cory,S. (1984) *Cell*, **37**, 113–122.
Dickson,C., Smith,R., Brookes,S. and Peters,G. (1984) *Cell*, **37**, 529–536.
Fung,Y.-K.-T., Fadly,A.M., Crittenden,L.B. and Kung,H.-J. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 3418–3422.
Grunstein,M. and Hogness,D. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 3961–3965.
Gubler,U. and Hoffmann,B.J. (1983) *Gene*, **25**, 263–269.
Hayward,W.S., Neel,B.G. and Astrin,S.M. (1981) *Nature*, **209**, 475–479.
Majors,J.E. and Varmus,H.E. (1981) *Nature*, **289**, 253–258.
Mason,P.J., Jones,M.B., Elkington,J.A. and Williams,J.G. (1985) *EMBO J.*, **4**, 205–211.
Melton,D.W., Konecki,D.S., Brennand,J. and Caskey,C.T. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2147–2151.
Messing,J. and Vieira,J. (1982) *Gene*, **19**, 269–276.
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.
Noorii-Daloii,M.R., Swift,R.A., Kung,H.-J., Crittenden,L.B. and Witter,R.L. (1981) *Nature*, **294**, 574–576.
Nusse,R. and Varmus,H.E. (1982) *Cell*, **31**, 99–109.
Payne,G.S., Bishop,J.M. and Varmus,H.E. (1982) *Nature*, **295**, 209–214.
Peters,G., Brookes,S., Smith,R. and Dickson,C. (1983) *Cell*, **33**, 369–377.
Peters,G., Kozak,C. and Dickson,C. (1984) *Mol. Cell Biol.*, **4**, 371–378.
Reynolds,G.A., Basu,S.K., Osborne,T.F., Chin,D.J., Gil,G., Brown,M.S., Goldstein,J.L. and Laskey,K.L. (1984) *Cell*, **38**, 275–285.
Selten,G., Cuypers,H.T., Zijlstra,M., Melief,C. and Berns,A. (1984) *EMBO J.*, **3**, 3215–3222.
Shen-Ong,G.L.C., Potter,M., Mushinski,J.F., Lavu,S. and Reddy,E.P. (1984) *Science*, **226**, 1077–1080.
Staden,R. (1980) *Nucleic Acids Res.*, **8**, 3673–3694.
Steffen,D. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2097–2101.
Teich,N., Wyke,J., Mak,T., Bernstein,A. and Hardy,W. (1982) In Weiss,R., Teich,N., Varmus,H. and Coffin,J. (eds), *Molecular Biology of Tumor Viruses. Part III. RNA Tumor Viruses*. Cold Spring Harbor Laboratory Press, NY, pp. 785–998.
Thomas,P.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5201–5205.
Varmus,H.E. (1982) *Cancer Surv.*, **2**, 301–319.
Varmus,H.E. (1984) *Annu. Rev. Genet.*, **18**, 553–612.
Watson,M.E.E. (1984) *Nucleic Acids Res.*, **12**, 5145–5164.

Received on 27 February 1986