The EMBO Journal vol.4 no.3 pp.667–674, 1985

# Chromosome translocation activates heterogeneously initiated, bipolar transcription of a mouse c-*myc* gene

**Franco Calabi and Michael S.Neuberger**

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Communicated by C.Milstein

**In many mouse plasmacytomas, the active c-*myc* gene has been truncated by chromosome translocation with the resultant severance of the protein-coding sequence from the normal promoter. Transcripts of such truncated c-*myc* genes were analyzed by Northern blotting, nuclease S1 mapping, primer extension assays and cDNA cloning. We conclude that transcription originates from multiple initiation sites on both c-*myc* coding and non-coding strands with the two-sets of transcripts derived from adjacent but essentially non-overlapping regions located >1 kb from the translocation junction. In X63Ag8, where c-*myc* is translocated to the immunoglobulin Cγ2b gene, the c-*myc* non-coding strand transcripts include the translocation junction and then splice directly into the γ2b CH1 exon. We propose that chromosome translocation activates a cryptic promoter in the first intron and that the heterogeneously initiated, bipolar transcription reflects the absence of a suitably placed TATA box element.**

*Key words:* chromosome translocation/c-*myc*/plasmacytoma/transcription

## Introduction

Many transformed cells exhibit an anomalous pattern of c-*myc* expression. In some cases very high levels of c-*myc* RNA are detected (Hayward *et al.*, 1981), although in most tumors the level of c-*myc* expression is not significantly different from that in proliferating normal cells (Kelly *et al.*, 1983; Keath *et al.*, 1984). However, in such tumours, there is usually either an aberrant timing of c-*myc* expression during the cell cycle or an altered pattern of c-*myc* transcription initiation (Campisi *et al.*, 1984; Adams *et al.*, 1983; Stanton *et al.*, 1983; ar-Rushdi *et al.*, 1983; Leder *et al.*, 1983). These various forms of transcriptional activation are often accompanied by a DNA rearrangement in the vicinity of c-*myc*. Thus, in chicken bursal lymphomas induced by avian leukosis virus, the provirus is found integrated close to c-*myc* (Hayward *et al.*, 1981); retroviral integrations in the vicinity of c-*myc* have also been identified in several mouse T cell leukaemias (Corcoran *et al.*, 1984). In these cases, the activation of c-*myc* expression has been attributed to transcription promotion or enhancement by viral sequences. In both Burkitt lymphomas and mouse plasmacytomas, consistent chromosomal translocations are observed which bring c-*myc* coding sequences close to one of the immunoglobulin gene loci (Taub *et al.*, 1982; Shen-Ong *et al.*, 1982). Although these translocations are assumed to cause c-*myc* activation, the mechanism remains unclear. In some Burkitt lymphomas and in most pristane-induced mouse plasmacytomas the translocation breakpoint lies within the c-*myc* transcription unit (Dalla-Favera *et al.*, 1983; Bernard *et al.*, 1983; Taub *et al.*, 1984; Shen-Ong *et al.*, 1982; Marcu *et al.*, 1983; Cory *et al.*, 1983). As a consequence, the c-*myc* protein-coding sequence is separated from its normal promoter, although – in most cases – this does not result in the gene being brought into proximity with either an immunoglobulin gene promoter or the enhancer element identified in the immunoglobulin heavy chain locus (Banerji *et al.*, 1983; Gillies *et al.*, 1983; Neuberger, 1983). Rather, both these transcription control sequences are usually found on the reciprocal product of the translocation (Cory *et al.*, 1983; Neuberger and Calabi, 1983; Dunnick *et al.*, 1983; Bernard *et al.*, 1983; Rabbitts *et al.*, 1983; Gerondakis *et al.*, 1984). To study the nature of the transcriptional activation, we have characterized the transcripts derived from the translocated c-*myc* gene in mouse plasmacytomas.

## Results

A schematic diagram of the translocated c-*myc* gene in mouse plasmacytoma X63Ag8 is shown in Figure 1. The breakpoint on chromosome 15 has fallen within exon 1 so that the protein-coding sequence, contained in exons 2 and 3, has been severed from the normal promoter. Northern blot analysis (Figure 2) reveals that c-*myc* mRNA in X63Ag8 is heterogeneous in size with major species of 1.8 and 2.0 kb; this contrasts with the 2.25–2.4 kb c-*myc* mRNAs detected in the Abelson leukaemia virus transformed pre-B cell line 18-81, which is unrearranged for c-*myc* (Adams *et al.*, 1982). Thus, as has been shown for several other plasmacytomas (Adams *et al.*, 1982; Shen-Ong *et al.*, 1982; Marcu *et al.*, 1983; Mushinski *et al.*, 1983), the DNA rearrangement correlates with the presence of altered c-*myc* mRNAs. The translocation does not, however, result in any gross alteration in the total level of c-*myc* mRNA. A quantitative com-

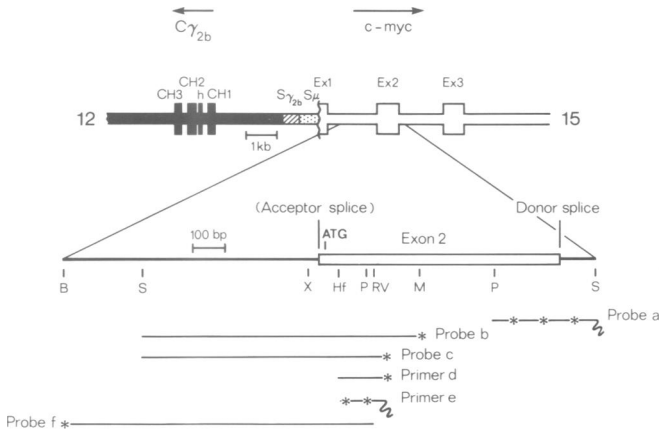


**Fig. 1.** Map of the (12;15) chromosome junction in X63Ag8. Chromosome 15 sequences are depicted by open lines, chromosome 12 by filled lines, exons by boxes, Sγ2b by cross-hatching and Sμ by dotting. A magnification of the region around the second c-*myc* exon is also provided. The ATG codon in exon 2 at which translation is assumed to start is indicated. Restriction sites mentioned in the text or in other figures are depicted; B, *Bam*HI, Hf, *Hinf*I; M, *Mbo*I; P, *Pst*I; RV, *Eco*RV; S, *Sac*I; X, *Xba*I. The derivation of probes and primers used in transcript mapping is illustrated.

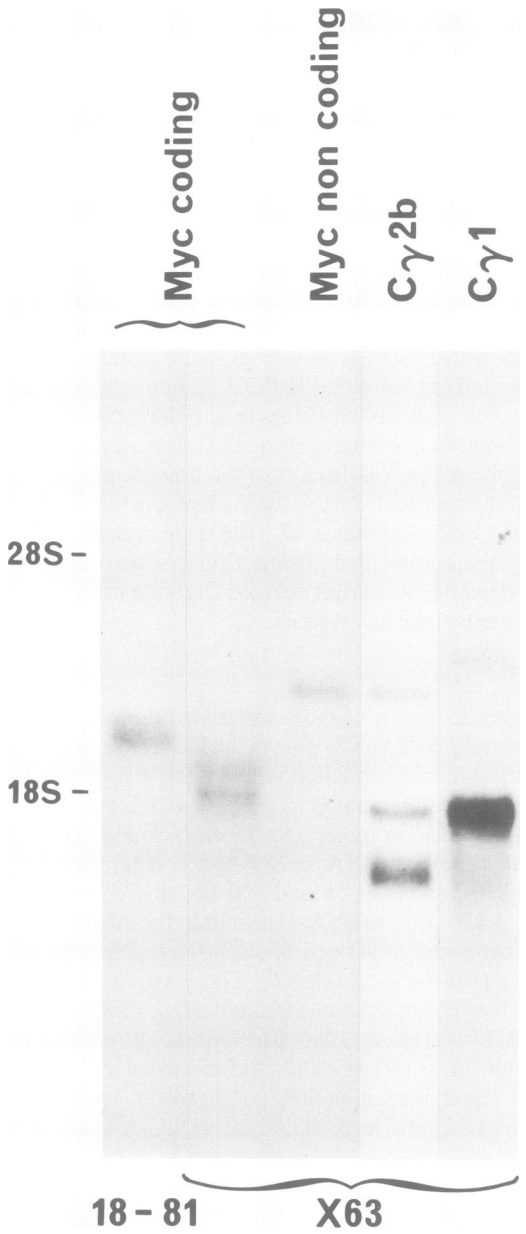


**Fig. 2.** Northern blot analysis of poly(A)$^+$ RNA from X63Ag8 and 18-81. Single-stranded probes for transcripts originating from c-*myc* coding and non-coding strand were made from M13 subclones of the *SacI-PstI* c-*myc* fragment (nucleotide 245 – 564) and the *PstI-PstI* fragment (nucleotides 967 – 1384), respectively. The Cγ2b and the Cγ1 probes were nick-translated plasmids pγ2b.1 or pγ1.13 (Rogers *et al.*, 1981). The positions of the 18S (1.8 kb) and 28S (5 kb) rRNAs are indicated.

parison of the amount of c-*myc* mRNA in the plasmacytomas MOPC 315 and X63Ag8, both of which are rearranged for c-*myc* (Shen-Ong *et al.*, 1982; Neuberger and Calabi, 1983), and in the lymphomas 18-81 and WEHI-231, both of which are unrearranged (Adams *et al.*, 1982) reveals no substantial difference (Figure 3).

*Start sites of c-myc transcription*

*Nuclease S1 mapping.* Preliminary experiments to locate the start points of the X63Ag8 c-*myc* mRNAs by nuclease S1 mapping indicated that the transcripts initiated near the *XbaI* site immediately upstream of exon 2 (data not shown). However, these experiments were carried out with internally labelled, single-stranded DNA probes and we were unable to obtain sufficiently

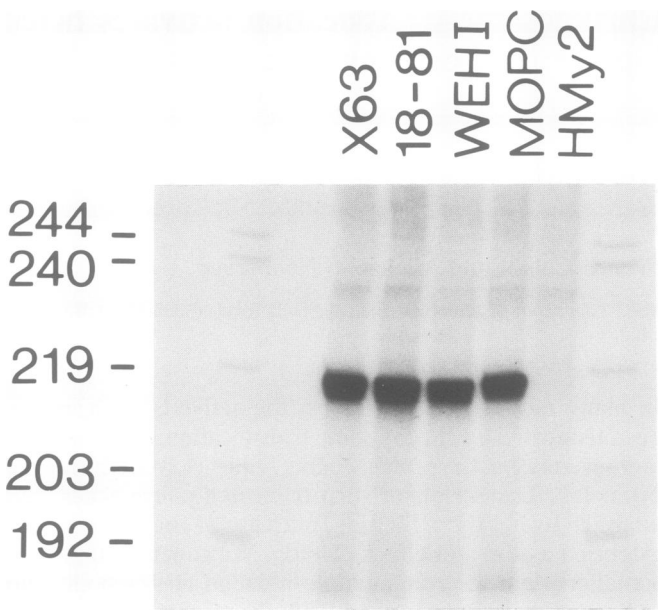


**Fig. 3.** Comparison of total c-*myc* RNA levels in cells with and without c-*myc* rearrangements. Total cytoplasmic RNA from the indicated cell lines (20 μg) was hybridized with internally labelled, single-stranded probe a (Figure 1) under conditions of probe excess prior to digestion with nuclease S1 and gel electrophoresis. The probe was synthesized from an M13mp11 clone of the *PstI-SacI* fragment spanning the 3′ end of c-*myc* exon 2. Since in all cases the probe maps downstream of the transcription initiation sites, the size of the protected band is the same, so long as correct splicing occurs at the end of exon 2. Markers are provided by an end-labelled *HpaII* digest of pBR322.

sharp bands to allow precise identification of transcription start sites. Nuclease S1 mapping performed with X63Ag8 cytoplasmic RNA using an end-labelled restriction fragment as probe (probe b in Figure 1) gave multiple nuclease-protected fragments (Figure 4A), which implies heterogeneous transcription originating or splicing into the first c-*myc* intron, just upstream of exon 2. To confirm these results and to allow a more accurate identification of the start points, we prepared a different S1 mapping probe that combined the advantages of end-labelled probes (sharper bands and unambiguous identification of end points) with the advantages of internally labelled, single-stranded probes (no probe renaturation and high specific activity). A $^{32}$P end-labelled 17-mer oligonucleotide homologous to part of the c-*myc* coding sequence was used to prime the synthesis of a probe using a c-*myc* clone in M13 as template (see Figure 4 legend). The oligonucleotide could be efficiently phosphorylated and a high specific activity, end-labelled, single-stranded S1 probe thereby obtained. As shown in Figure 4B, the same oligonucleotide was used as a primer in dideoxy-sequencing reactions. The results of this experiment confirm with a better accuracy those obtained using probe b. The major inferred transcription start sites are depicted above the sequence in Figure 5. The same pattern is obtained with RNA from mouse plasmacytomas X63Ag8, MOPC 315 (Figure 4) and J558 (data not shown), all of which harbour c-*myc* genes that have been truncated by chromosome translocation. Both WEHI 231 and 18-81 give protected fragments that align with the exon 2 splice acceptor junction, as expected for germline transcripts.

*Primer extension.* We have assumed that the results of the X63Ag8 S1 mapping reflect heterogeneity of transcription initiation rather than a complex splicing pattern. Primer extension assays were performed in order to confirm this. The same c-*myc*

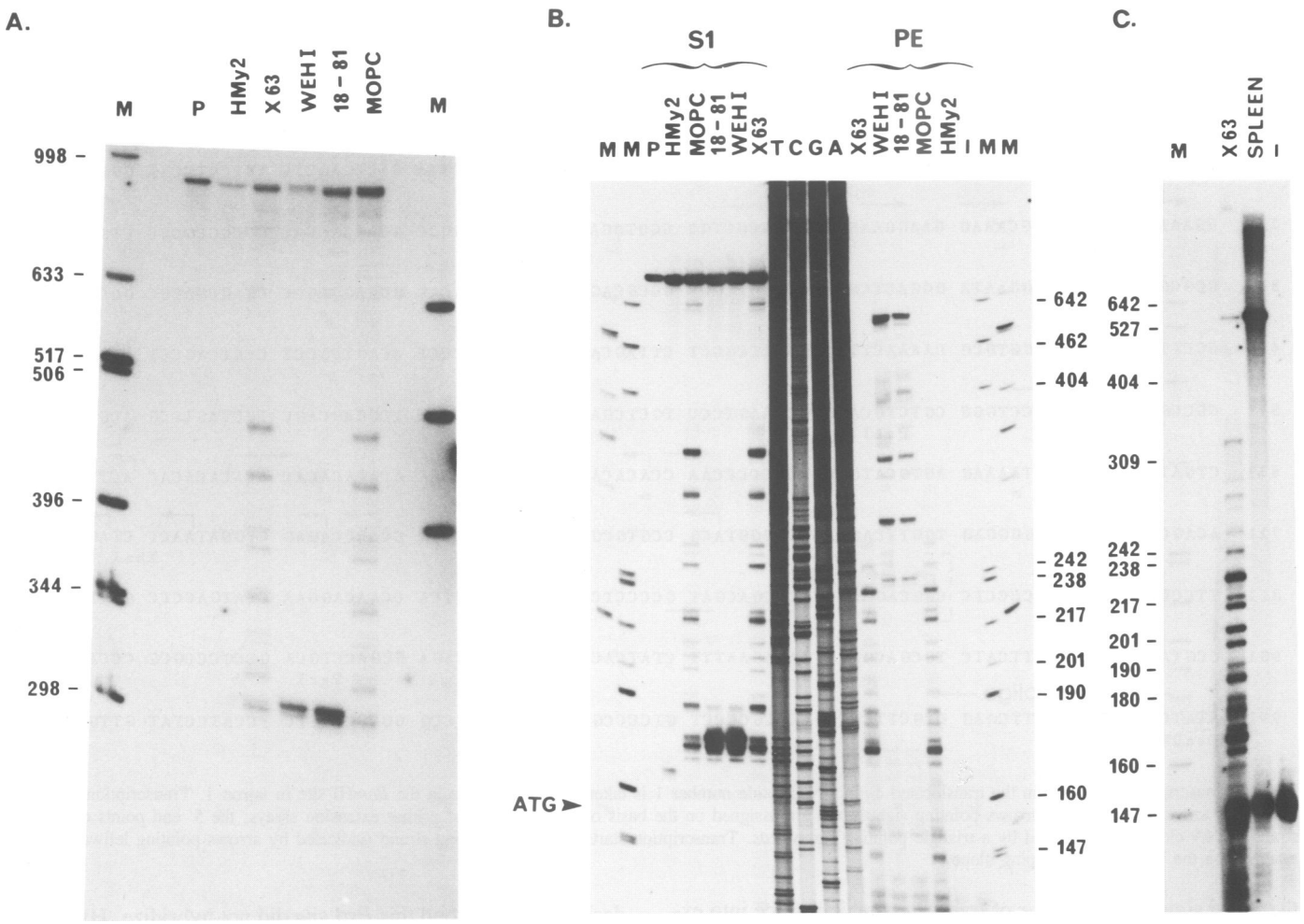


**Fig. 4.** S1 mapping and primer extension assays of c-*myc* coding strand transcripts. (**A**) S1 mapping with probe b (Figure 1), an *Mbo*I-*Sac*I (nucleotide 1141−245) fragment 5′ end-labelled at the *Mbo*I site. (**B**) **left**: S1 mapping with probe c (Figure 1), a 5′ end-labelled, single-stranded DNA fragment extending from the 5′ end of the c-*myc* 17-mer (nucleotide 1012) to the *Sac*I site in intron 1 (nucleotide 245); (**B**) **centre**: dideoxy sequencing tracks of mouse c-*myc* exon 2; the sequencing reaction was primed with the c-*myc* 17-mer on M13myc-1 template; (**B**) **right**: primer extension using primer d (Figure 1), a 5′ end-labelled, single-stranded DNA fragment extending from the 5′ end of the c-*myc* 17-mer to the *Hinf*I site (nucleotide 898). (**C**) Primer extension using primer e (Figure 1). This primer was made by extending the universal sequencing 17-mer on an M13mp8 clone of the c-*myc* *Pst*I fragment (nucleotide 967−564) and subsequent digestion with *Hinf*I. Size markers (M) are provided by digestion of pBR322 or pAT153 with *Hinf*I, *Hinf*I + *Eco*RI or *Hpa*II. In the S1 mapping, lanes (P) were run with probe alone and, in the primer extensions, incubations (−) were performed without added template. Bands due to fragments of ~650 nucleotides and which could correspond to the initiations described by Keath *et al.* (1984) are seen in the S1 mapping in **panel B**.

17-mer used to prepare the S1 probe was also used to prime synthesis of an end-labelled, single-stranded DNA fragment extending up to the *Hinf*I site at nucleotide position 898 in the c-*myc* sequence. This 111 nucleotide fragment was used as a primer in reverse transcriptase catalysed extension assays performed in parallel with the S1 mapping (Figure 4B). Most of the major S1-protected fragments coincide with reverse transcriptase stops. However it is clear, particularly from the extensions carried out using RNA from WEHI-231 and 18-81, that there is some stopping short giving rise to artefactual bands and loss in intensity of the longer cDNAs. For example, the primer extension band corresponding to the S1-protected fragment of 339 nucleotides (Figure 4B) is very weak in the original autoradiograph. Less stopping short was obtained in a subsequent experiment (Figure 4C) using a different batch of reverse transcriptase and a primer (primer e in Figure 1) that extended from the *Pst*I site in exon 2 (nucleotide 966) back to the same *Hinf*I site as used with primer d. The size of the cDNA obtained with a BALB/c spleen RNA sample corresponds to that expected for germline transcripts initiated at the most 3′ of the two exon 1 cap sites proposed by Bernard *et al.* (1983). With X63Ag8 RNA a very good correlation is again observed between the sizes of the species obtained by primer extension and the S1-protected fragments; one major anomalous band of 150 nucleotides should be ignored as it is seen to be due to self-priming. A minor band is observed with the X63Ag8 sample that co-migrates with the cDNA obtained using spleen RNA template. This suggests that, in X63Ag8, there are some mRNAs derived from the unrearranged c-*myc* allele. Consistent with this is the observation that in the S1 mapping with X63Ag8 and MOPC 315 RNAs, protected bands are obtained which could correspond to transcripts splicing into exon 2 (Figure 4A and B). However, such bands must also partly be due to transcripts initiating around the exon 2 splice acceptor site, since they co-migrate with bands obtained by primer extension. Thus the S1 mapping with the probes used cannot provide an estimate of the level of germline c-*myc* transcription in X63Ag8.

*Isolation of cDNA clones.* We have constructed an X63Ag8 cDNA library enriched for c-*myc* sequences to determine whether

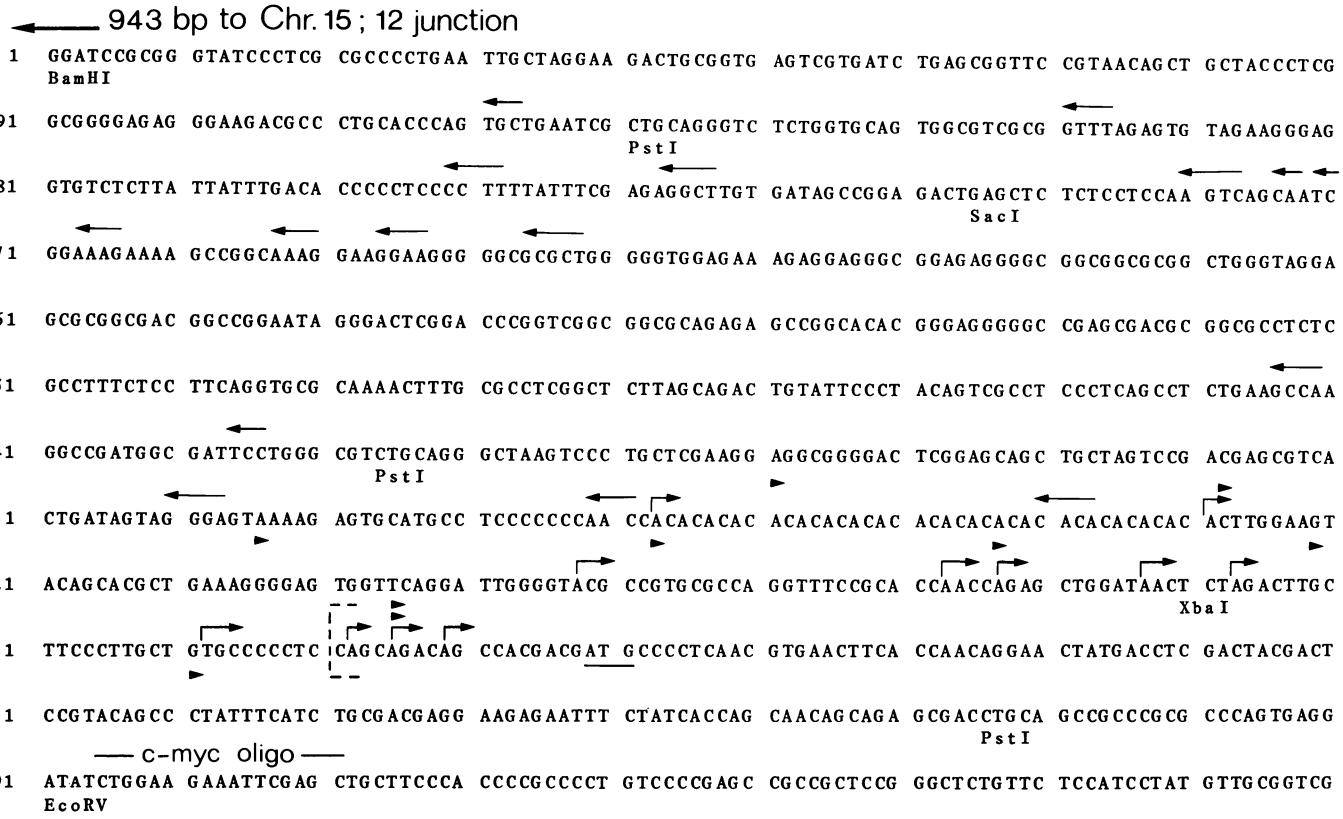


**Fig. 5.** Transcription start sites in the translocated c-*myc*. Nucleotide number 1 is taken as the first base in the *Bam*HI site in intron 1. Transcription starts of the coding strand (indicated by arrows pointing rightwards) are assigned on the basis of S1 mapping and primer extension assays; the 5′ end points of coding strand cDNA clones are indicated by a triangle pointing rightwards. Transcription starts of the non-coding strand (indicated by arrows pointing leftwards) are assigned on the basis of S1 mapping alone.

there is a significant number of transcripts which splice into exon 2. Such transcripts could either derive from the unrearranged allele or could initiate at unidentified sites upstream of the protein-coding sequence of the translocated c-*myc*. The cloning strategy is illustrated in Figure 6. cDNA first strand synthesis was primed with the c-*myc* 17-mer. Double-stranded molecules were made in such a way as to favour the subsequent isolation of full length molecules and cloned into M13mp11. Twelve c-*myc* cDNA clones were isolated. Three of these clones contained aberrant cDNA structures analogous to those obtained in a different system (Fields and Winter, 1981); the end-points of the other nine clones are presented in Figure 5. No clone was obtained that was attributable to transcripts splicing into exon 2. The 5′ ends of over half the cDNA clones coincide with positions which had been designated transcription start sites on the basis of S1 protection and primer extension assays (Figure 5). Thus the germline transcripts observed in the primer extension experiments can only account for a small proportion of the total and the bulk of X63Ag8 c-*myc* mRNA originates in the intron upstream of exon 2.

*Transcription of the c-myc non-coding strand*

Northern blot analysis performed using single-stranded DNA probes revealed the presence in X63Ag8 of polyadenylated transcripts which hybridize to the c-*myc* non-coding strand; analogous transcripts have previously been detected in other mouse plasmacytomas (Dean *et al.*, 1983). The most abundant species in X63Ag8 is 2.7 kb, although other minor bands are also observed (Figure 2). These transcripts were detected using a set of single-stranded, non-overlapping probes which span the segment stretching from the translocation junction to the *Pst*I site at nucleotide position 564 in the c-*myc* coding sequence; probes derived from beyond this *Pst*I site did not hybridize. Hybridization of a Cγ2b probe to the Northern blot (Figure 2) reveals that the 2.7-kb c-*myc* non-coding strand transcript contains Cγ2b sequences and that this transcript therefore extends across the chromosome translocation junction. However, in addition, two stronger Cγ2b positive, c-*myc* negative bands (1.7 and 1.3 kb) are also observed. Hybridization with a Cγ1 probe demonstrates that the 1.7-kb band represents the X63Ag8 secreted γ1 mRNA, which cross-hybridizes with the Cγ2b probe. The 1.3-kb band was not detected with any probe containing chromosome 15 sequences and may reflect a sterile γ2b transcript analogous to those identified in the germline immunoglobulin μ heavy chain locus (Alt *et al.*, 1982; Nelson *et al.*, 1983).

*Nuclease S1 mapping.* Nuclease S1 mapping using probe f was carried out to locate the start points of the c-*myc* non-coding strand transcripts (Figure 7). Both X63Ag8 and J558 (as well as MOPC 315, data not shown) give an identical pattern of multiple protected fragments; the major inferred initiation sites are indicated on the sequence in Figure 5. Transcription of the non-coding strand is only observed in lines harbouring a translocated c-*myc*; no specific S1-protected fragment is detected in WEHI 231 and 18-81 – the single 160 nucleotide band also being present in the human line HMy2 which serves as a negative control.

Thus, transcription of the non-coding strand of the translocated c-*myc* gene originates from multiple sites in the first intron close to, but essentially non-overlapping with the multiple initiation sites on the coding strand. The assumption that the multiplicity of S1-protected fragments is due to heterogeneity of transcription initiation rather than some byzantine splicing pattern is supported by the analysis of cDNA clones described below.

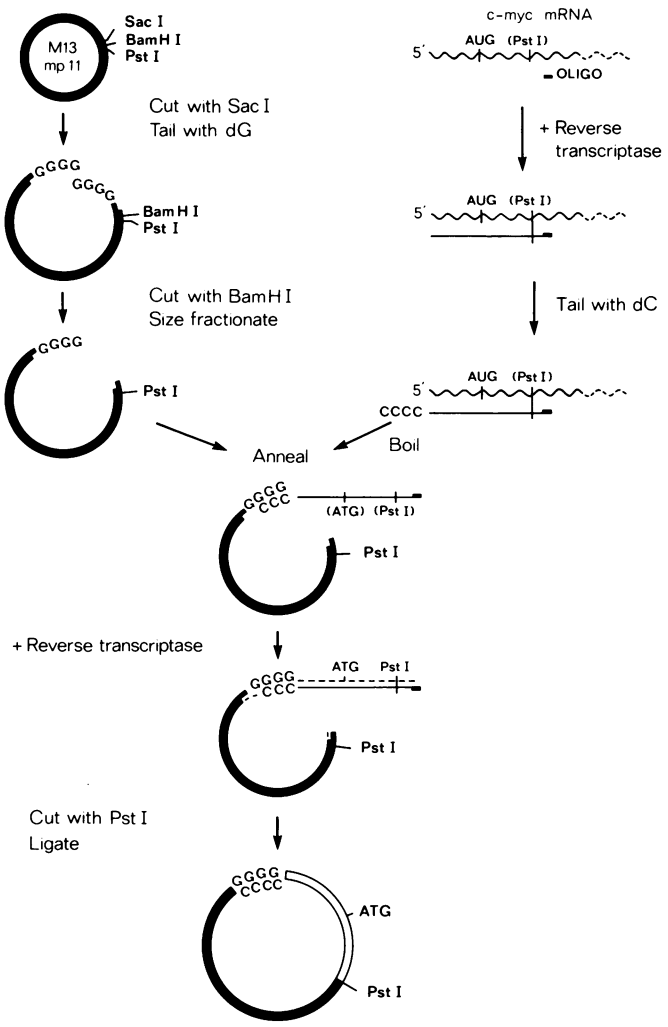


**Fig. 6.** Procedure used for the isolation of cDNA clones containing the 5′ ends of c-*myc* transcripts in X63Ag8. Details are provided in Materials and methods.

*Isolation of cDNA clones.* We isolated cDNA clones to confirm that transcripts derived from the c-*myc* non-coding strand contain Cγ2b sequences as well as to clarify their intron/exon structure. An oligonucleotide homologous to the 5′ end of the γ2b CH1 exon was used to prime DNA first strand synthesis on X63Ag8 poly(A)$^+$ RNA; double-stranded cDNA was then blunt-end cloned into M13mp11. In this way, 38 cDNA clones were isolated which hybridized with a c-*myc* probe. The structure of a representative cDNA clone is presented in Figure 8. All of 24 cDNA clones analyzed revealed an identical splicing pattern. Thus, starting in the intron upstream of exon 2, transcription from the non-coding strand of c-*myc* continues uninterruptedly up to the translocation junction. Eight bases across this point the RNA is spliced removing the bulk of the heavy chain switch region and continues directly into the γ2b CH1 exon. Presumably, such transcripts contain the entire Cγ2b coding sequence terminating at one of the γ2b polyadenylation signals as S1 mapping reveals the presence of correctly spliced Cγ2b exons in X63Ag8 poly(A)$^+$ RNA (data not shown). The sizes of the bands seen in the Northern blotting (Figure 2) are consistent with such an interpretation.

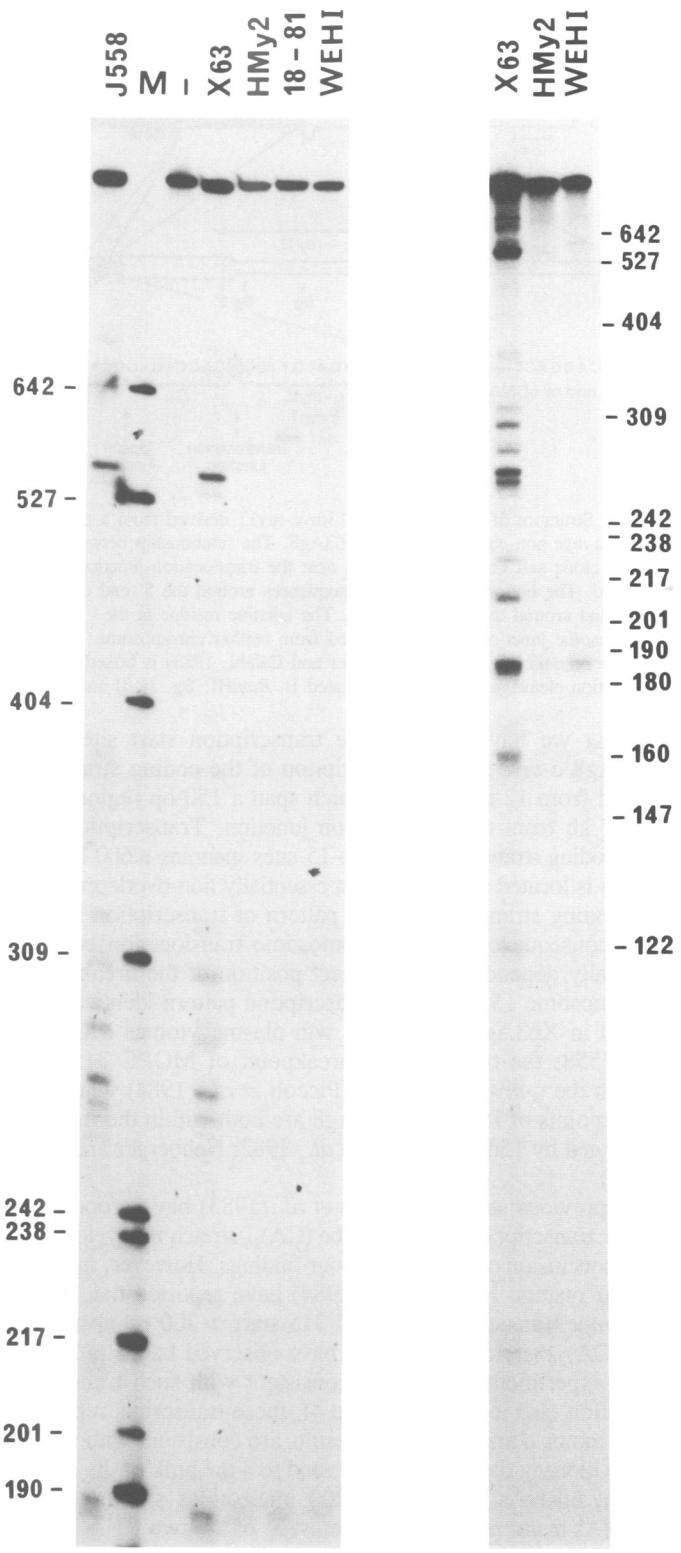


**Fig. 7.** S1 mapping of transcripts from the c-*myc* noncoding strand. The probe used, probe f (Figure 1), is a *Bam*HI-*Eco*RV fragment (nucleotides 1–991) 5′ end-labelled at the *Bam*HI site. The two panels present long and short runs of the electrophoretic separation of S1 protected fragments. Lanes with size markers (M) or probe alone (P) are indicated.

## Discussion

### *Transcription of a translocated, truncated c-myc gene*

By a combination of S1 mapping, primer extension and cDNA

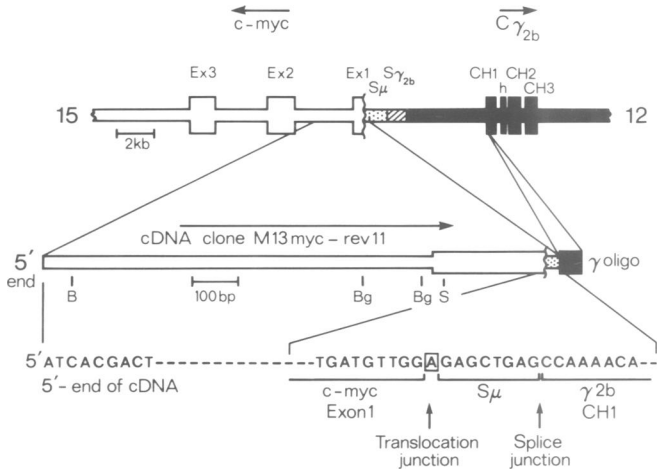


**Fig. 8.** Structure of cDNA clone M13myc-rev11 derived from a transcript of the c-*myc* non-coding strand in X63Ag8. The relationship between the cDNA clone and genomic sequences near the translocation junction is depicted. The bottom line gives the sequences around the 5′ end of the clone and around the splice junction. The adenine residue at the chromosome junction which is derived from neither chromosome 15 nor chromosome 12 sequences (Neuberger and Calabi, 1983) is boxed. Restriction cleavage sites are abbreviated B, *Bam*HI; Bg, *Bgl*II and S, *Sac*I.

cloning we have mapped the transcription start sites of the X63Ag8 c-*myc* gene. Transcription of the coding strand is initiated from 12 major sites which span a 150-bp region located ~1.7 kb from the translocation junction. Transcription of the non-coding strand starts from ~15 sites spanning a 600-bp region which is located adjacent to, but essentially non-overlapping with, the coding strand starts. This pattern of transcription initiation is a consequence of the chromosome translocation but is not critically dependent on the exact position of the breakpoint on chromosome 15. A c-*myc* transcription pattern identical to that found in X63Ag8 is obtained wth plasmacytomas MOPC 315 and J558; the translocation breakpoint of MOPC 315 occurs within the c-*myc* first intron (Piccoli *et al.*, 1984) whereas the breakpoints of J558 and X63Ag8 are both within the first exon, separated by 156 bp (Adams *et al.*, 1982; Neuberger and Calabi, 1983).

In previous studies, Stanton *et al.* (1983) have proposed that c-*myc* transcription starts near the $(CA)_{20}$ stretch in myeloma J558 – a conclusion consistent with our findings. However, in contrast to our results, Keath *et al.* (1984) have reported that the bulk of c-*myc* transcripts in MOPC 315 start >300 bp upstream of the $(CA)_{20}$ stretch. Whilst we have observed bands in S1 mapping experiments which are consistent with such transcription initiation (see legend to Figure 4), these transcripts represent a very minor component. Our results are consistent with the Northern blotting data where it is found that the bulk of the X63Ag8 c-*myc* mRNAs are at least 200 nucleotides shorter than the mRNAs transcribed from the most 3′ of the two RNA cap sites present in the germline gene (Figure 2 and Adams *et al.*, 1983; Bernard *et al.*, 1983). Given the length of the first exon of the germline c-*myc* gene and the splicing pattern of transcripts from the translocated allele, we can infer that X63Ag8 c-*myc* transcripts must initiate 3′ of nucleotide 630 in the sequence presented in Figure 5. It is also notable that the initiations proposed by Keath *et al.* (1984) do not include the major start sites identified in our analysis. Although we have used a subline of MOPC 315 (MOPC 315.26), this is not likely to be the cause of the disagreement, as neither we nor Keath *et al.* find that the results obtained depend critically upon the plasmacytoma used, so long as c-*myc* is rearranged. Technical differences provide a far more likely explanation. The only S1-probe used by Keath *et al.* that would detect some of the c-*myc* transcripts that we identify in Figure 4 is a uniformly labelled fragment extending upstream from the *Xba*I site (nucleotide position 801). Such a probe would contain either too short a stretch of homology or no homology with the transcripts we observe. In this context it is significant that in other work we have noted that we have been unable to detect an S1 protected fragment of 90 nucleotides using an internally labelled single-stranded probe and hybridization conditions substantially less stringent than those employed by Keath *et al.*; however, other experiments revealed the presence of the appropriate RNA species in abundance (Neuberger, 1983).

*Transcription of the c-myc non-coding strand*

The pattern of transcription initiation from the c-*myc* non-coding strand is similar in X63Ag8, MOPC 315 and J558; the non-coding strand transcripts in MOPC 315 have previously been analyzed by Keath *et al.* (1984) and the conclusions that should be drawn from their data are consistent with our own. The analysis of cDNA clones reveals that the non-coding strand mRNAs continue through the translocation junction; eight nucleotides across the junction the sequence GAG′GTGAGCT is encountered, and this acts as a splice donor. GAGGT is a variant μ switch region (Sμ) repeat (Nikaido *et al.*, 1981) and the sequence GAG′GTGAGCT includes the consensus for 5′ splice borders, AG′GTRAG (Mount, 1982). Is the Sμ unit repeat itself sufficient to act as a splice donor or has it only become such a site by being brought into proximity with novel sequences by chromosome translocation? A pointer to the answer to this question is provided by an analysis of sterile μ transcripts (Alt *et al.*, 1982; Nelson *et al.*, 1983). These transcripts appear to originate near the $J_H$ segments in the immunoglobulin heavy chain locus and splice into Cμ. Whilst the splice junctions of these transcripts have not been identified, the gross structure of the RNAs is consistent with sequences near the 5′ end of Sμ acting as a splice site. Thus, by analogy with our data, splicing may well occur at the first GAG′GTGAGCT motif of Sμ that is encountered.

The c-*myc* non-coding strand transcripts cannot be translated into Cγ2b containing polypeptides as examination of the cDNA clones reveals multiple translation stop codons in phase with the γ2b reading frame and in the region immediately 5′ of it.

*Transcription activation*

The most striking features of the transcription of the X63Ag8 c-*myc* gene are the bipolarity and extreme heterogeneity of start sites. With the possible exceptions of one or two of the non-coding strand initiations, none of the transcription start sites is preceded by TATA-like sequences. In fact, the heterogeneity of initiation can be reasonably attributed to the lack of suitably placed TATA box elements; certainly, mutation of the TATA box in the SV40 early promoter has been shown to result in the use of multiple sites of transcription initiation (Benoist and Chambon, 1981). In the case of cellular genes, human ε-globin uses multiple transcription start sites, although 85% of the transcripts are derived from the two canonical cap sites (Allan *et al.*, 1983). However, a more intriguing comparison with X63Ag8 c-*myc* is provided by the gene for the 'housekeeping' enzyme HMG CoA reductase. Reynolds *et al.* (1984) have shown that transcription of the HMG CoA reductase gene starts from at least five sites scattered over a 100 nucleotide region, there being no apparent TATA box in the region. These authors draw attention to the sequence CCGCCC which is repeated three times in the region

immediately upstream of the transcription start sites; its complement, GGGCGG, is also present twice in the same area. Not only have the sequences CCGCCC and GGGCGG been shown to be components of the upstream promoter element of the herpes virus thymidine kinase gene (McKnight *et al.*, 1984), but they are also found within the 21-nucleotide repeats that constitute part of the SV40 early and late promoters (Fromm and Berg, 1982). It may therefore be significant that the GGGCGG motif is repeated twice immediately upstream of the major group of c-*myc* non-coding strand starts (nucleotides 317 and 326). Interestingly, although mouse and human c-*myc* show little homology in this region (Neuberger and Calabi, 1983), examination of the human c-*myc* sequence (Colby *et al.*, 1983) reveals two closely linked GGGCGG elements in an analogous region.

The bipolarity of the X63Ag8 c-*myc* transcription might also be attributed to the absence of a suitably placed TATA box. Several promoter upstream elements have been shown to be active in both orientations (Everett, 1984; McKnight *et al.*, 1984; Vigneron *et al.*, 1984) and it therefore seems reasonable to speculate that it is the position of the TATA box with respect to the upstream element that confers directionality on transcription.

In summary, we would like to propose that, located between the major start sites of the X63Ag8 c-*myc* rightward and leftward transcripts, lies a promoter upstream element. The mechanism by which this promoter is activated by chromosome translocation remains obscure. Whilst the direct involvement of the identified immunoglobulin heavy chain locus enhancer element can be excluded, an indirect involvement cannot. Certainly, in the case of mouse mammary tumour virus, it has been shown that – following activation of the enhancer – irreversible changes in chromatin structure result (Zaret and Yamamoto, 1984). Alternatively, the promoter may simply be activated by gene truncation.

## Materials and methods

*Cell lines, DNA clones and oligonucleotides*

Mouse lymphoid cell lines X63Ag8 (IgG1, $\kappa$, Köhler and Milstein, 1975), MOPC 315.26 (an $\alpha$-heavy chain-loss variant of MOPC 315; Winberry *et al.*, 1980), J558L (an $\alpha$-heavy chain-loss variant of J558; Oi *et al.*, 1983), WEHI 231 (Gutman *et al.*, 1981) and 18-81 (Siden *et al.*, 1979) as well as human lymphoblastoid line HMy2 (Edwards *et al.*, 1982) were all grown in DMEM medium containing 10% foetal calf serum. All c-*myc* DNA clones were ultimately derived from phage λMYG2 or λMYG9 (Neuberger and Calabi, 1983). M13myc-1 contains the *Sac*I fragment including c-*myc* exon 2 cloned into M13mp11. An oligonucleotide (5′-GCTCGAATTTCTTCCAG) complementary to a sequence in c-*myc* exon 2 was made using phoshotriester chemistry and purified by high pressure liquid chromatography (Gait *et al.*, 1982). The $\gamma$ oligonucleotide is described by Kaartinen *et al.* (1983).

*Northern blot analysis*

Poly(A)$^+$ RNA (5 $\mu$g) was glyoxylated and subjected to electrophoresis through a 1.2% agarose gel. After transfer to nitrocellulose (Thomas, 1983), the RNA was hybridized with radioactive probes. Internally labelled, single-stranded DNA probes were prepared from M13 clones as described previously (Neuberger, 1983).

*Nuclease S1 mapping and primer extension*

For the preparation of total cytoplasmic RNA, cells were lysed in 10 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1 mM $MgCl_2$, 0.5% Nonidet P40, the nuclei pelleted and RNA purified from the supernatant by phenol extraction in the presence of 1% SDS and subsequent ethanol precipitation. 5′ End-labelled restriction fragment probes were made using calf intestinal alkaline phosphatase and T4 polynucleotide kinase with [$\gamma$-$^{32}$P]ATP. Internally labelled, single-stranded probes were prepared as described previously (Neuberger, 1983); end-labelled, single-stranded probes were made in the same way except that the 5′ end-lablled c-*myc* oligonucleotide was used to prime probe synthesis on a M13myc-1 template and unlabelled dATP was used. Hybridization prior to S1 digestion or primer extension was performed for 16 h at 46°C in 80% formamide, 0.4 M NaCl, 10 mM Pipes (pH 6.4), 1 mM EDTA. Digestion with nuclease S1 (BRL; 250 units) was carried out for 30 min at 15°C in 30 mM sodium acetate (pH 4.5), 0.25 M NaCl, 1 mM $ZnSO_4$, 5% glycerol (Berk and Sharp, 1977). Primer extension assays using reverse transcriptase (either a gift from J.Beard or from Anglia Biotechnology) were done as previously described (Neuberger, 1983). Samples were analyzed on 4% or 6% polyacrylamide/urea sequencing gels.

*Cloning of cDNA*

The strategy for cloning c-*myc* cDNA is illustrated in Figure 6. M13mp11 (Messing, 1983) was digested with *Sac*I and tailed with dGTP using terminal transferase to give an average tail length of 20, essentially using the procedure of Deng and Wu (1981). The vector was then digested with *Bam*HI and purified over a column of Sepharose CL-4B. cDNA synthesis using X63Ag8 poly(A)$^+$ RNA (5 $\mu$g) as template and the c-*myc* 17-mer as primer was performed under the conditions described by Buell *et al.* (1978) and the resulting cDNA-mRNA hybrids were tailed with dCTP to give an average tail length of 20. This follows the strategy of Land *et al.* (1983) for the preferential cloning of full-length cDNA. Template mRNA was removed by alkaline hydrolysis and the dC-tailed cDNA annealed to the dG-tailed vector. Reverse transcriptase was used for the synthesis of cDNA second strand. After digestion with *Pst*I and incubation with T4 DNA ligase (gift of K.Nagai), the sample was transfected into *Escherichia coli* TG1 by the procedure of Hanahan (1983). Plaques containing c-*myc* sequences were screened for by hybridization with the c-*myc*-containing *Eco*RI fragment of phage λMYG2. Twelve *myc*-positive plaques were obtained out of a total of ~2500.

Non-coding strand cDNAs were made by using an oligonucleotide homologous to the 5′ end of the $\gamma$2b CH1 exon to prime cDNA first strand synthesis. After synthesis of the second strand using the Klenow fragment of DNA polymerase I, digestion with nuclease S1 and end-repair, cDNA molecules were cloned into M13mp11 by blunt-end ligation.

## Acknowledgements

We are grateful to Gareth Williams for expert technical assistance in the latter part of the work, to Sue Insole and Philip McArdle for preparing the illustrations, to Greg Winter for help with synthesis of the c-*myc* oligonucleotide and to Gillian Griffiths and Cesar Milstein for the $\gamma$ oligonucleotide. F.C. was supported by fellowships from EMBO and the Leukemia Society of America.

## References

Adams,J.M., Gerondakis,S., Webb,E., Mitchell,J., Bernard,O. and Cory,S. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 6966-6970.

Adams,J.M., Gerondakis,S., Webb,E., Corcoran,L.M. and Cory,S. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 1982-1986.

Allan,M., Lanyon,W.G. and Paul,J. (1983) *Cell*, **35**, 187-197.

Alt,F.W., Rosenberg,N., Eenea,V., Siden,E. and Baltimore,D. (1982) *Mol. Cell. Biol.*, **2**, 386-400.

ar-Rushdi,A., Nishikura,K., Erikson,J., Watt,R., Rovera,G. and Croce,C.M. (1983) *Science (Wash.)*, **222**, 390-393.

Banerji,J., Olson,L. and Schaffner,W. (1983) *Cell*, **33**, 729-740.

Benoist,C. and Chambon,P. (1981) *Nature*, **290**, 304-310.

Berk,A.J. and Sharp,P. (1977) *Cell*, **12**, 721-732.

Bernard,O., Cory,S., Gerondakis,S., Webb,E. and Adams,J.M. (1983) *EMBO J.*, **2**, 2375-2383.

Buell,G.N., Wickens,M.P., Payvar,F. and Schimke,R.T. (1978) *J. Biol. Chem.*, **253**, 2471-2482.

Campisi,J., Gray,H.E., Pardee,A.B., Dean,M. and Sonenshein,G.E. (1984) *Cell*, **36**, 241-247.

Colby,W.W., Chen,E.Y., Smith,D.H. and Levinson,A.D. (1983) *Nature*, **301**, 722-725.

Corcoran,L.M., Adams,J.M., Dunn,A.R. and Cory,S. (1984) *Cell*, **37**, 113-122.

Cory,S., Gerondakis,S. and Adams,J.M. (1983) *EMBO J.*, **2**, 697-703.

Dalla-Favera,R., Martinotti,S., Gallo,R., Erikson,J. and Croce,C.M. (1983) *Science (Wash.)*, **219**, 963-967.

Dean,M., Kent,R.B. and Sonenshein,G.E. (1983) *Nature*, **305**, 443-446.

Deng,G. and Wu,R. (1981) *Nucleic Acids Res.*, **9**, 4173-4188.

Dunnick,W. Shell,B.E. and Dery,C. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 7269-7273.

Edwards,P.A.W., Smith,C.M., Neville,A.M. and O'Hare,M.J. (1982) *Eur. J. Immunol.*, **12**, 6412-6418.

Everett,R.D. (1984) *Nucleic Acids Res.*, **12**, 3037-3056.

Fields,S. and Winter,G. (1981) *Gene*, **15**, 207-214.

Fromm,M. and Berg,P. (1982) *J. Mol. Appl. Genet.*, **1**, 457-481.

Gait,M.J., Matthes,H.W.D., Singh,M., Sproat,B.S. and Titmas,R.C. (1982) *Nucleic Acids Res.*, **10**, 6243-6254.

Gerondakis,S., Cory,S. and Adams,J.M. (1984) *Cell*, **36**, 973-982.

Gillies,S.D., Morrison,S.L., Oi,V.T. and Tonegawa,S. (1983) *Cell*, **33**, 717-728.

Gutman,G.A., Warner,N.L. and Harris,A.W. (1981) *Clin. Immunol. Immunopathol.*, **18**, 230-244.

Hanahan,D. (1983) *J. Mol. Biol.*, **166**, 557-580.
Hayward,W.S., Neel,B.G. and Astrin,S.M. (1981) *Nature*, **290**, 475-480.
Kaartinen,M., Griffiths,G.M., Hamlyn,P.H., Markham,A.F., Karjalainen,K., Pelkonen,J.L.T., Mäkelä,O. and Milstein,C. (1983) *J. Immunol.*, **130**, 937-945.
Keath,E.J., Kelekar,A. and Cole,M.D. (1984) *Cell*, **37**, 521-528.
Kelly,K., Cochran,B.H., Stiles,C.D. and Leder,P. (1983) *Cell*, **35**, 603-610.
Köhler,G. and Milstein,C. (1975) *Nature*, **256**, 495-497.
Land,H. Grez,M., Hauser,H., Lindenmaier,W. and Schütz,G. (1983) *Methods Enzymol.*, **100**, 285-292.
Leder,P., Battey,J., Lenoir,G., Moulding,C., Murphy,W., Potter,H., Stewart,T. and Taub,R. (1983) *Science (Wash.)*, **222**, 765-771.
Marcu,K.B., Harris,L.J., Stanton,L.W., Erikson,J., Watt,R. and Croce,C.M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 519-523.
McKnight,S.L., Kingsbury,R.C., Spence,A. and Smith,M. (1984) *Cell*, **37**, 253-262.
Messing,J. (1983) *Methods Enzymol.*, **101**, 20-78.
Mount,S.M. (1982) *Nucleic Acids Res.*, **10**, 459-472.
Mushinski,J.F., Bauer,S.R., Potter,M. and Premkumar Reddy,E. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 1073-1077.
Nelson,K.J., Haimovich,J. and Perry,R.P. (1983) *Mol. Cell. Biol.*, **3**, 1317-1332.
Neuberger,M.S. (1983) *EMBO J.*, **2**, 1373-1378.
Neuberger,M.S. and Calabi,F. (1983) *Nature*, **305**, 240-243.
Nikaido,T., Nakai,S. and Honjo,T. (1981) *Nature*, **292**, 845-848.
Oi,V.T., Morrison,S.L., Herzenberg,L.A. and Berg,P. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 825-829.
Piccoli,S.P., Caimi,P.G. and Cole,M.D. (1984) *Nature*, **310**, 327-330.
Rabbitts,T.H., Forster,A., Baer,R. and Hamlyn,P.H. (1983) *Nature*, **306**, 806-809.
Reynolds,G.A., Basu,S.K., Osborne,T.F., Chin,D.J., Gil,G., Brown,M.S., Goldstein,J.L. and Luskey,K.L. (1984) *Cell*, **38**, 275-285.
Rogers,J., Choi,E., Souza,L., Carter,C., Word,C., Kuehl,M., Eisenberg,D. and Wall,R. (1981) *Cell*, **26**, 19-28.
Shen-Ong,G.L.C., Keath,E.J., Piccoli,S.P. and Cole,M.D. (1982) *Cell*, **31**, 443-452.
Siden,E.J., Baltimore,D., Clark,D. and Rosenberg,N.E. (1979) *Cell*, **16**, 389-396.
Stanton,L.W., Watt,R. and Marcu,K.B. (1983) *Nature*, **303**, 401-406.
Taub,R., Kirsch,I. Morton,C., Lenoir,G., Swan,D., Tronick,S., Aaronson,S. and Leder,P. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7837-7841.
Taub,R., Moulding,C., Battey,J., Murphy,W., Vasicek,T., Lenoir,G.M. and Leder,P. (1984) *Cell*, **36**, 339-348.
Thomas,P.S. (1983) *Methods Enzymol.*, **100**, 255-266.
Vigneron,M., Barrera-Saldana,H.A., Baty,D., Everett,R.E. and Chambon,P. (1984) *EMBO J.*, **3**, 2373-2382.
Winberry,L., Marks,A. and Baumal,R. (1980) *J. Immunol.*, **124**, 1174-1182.
Zaret,K.S. and Yamamoto,K.R. (1984) *Cell*, **38**, 29-38.

*Received on 30 November 1984*

**Note added in proof**

Since submission of this paper, a detailed analysis of the transcription of a translocated c-*myc* gene in another mouse plasmacytoma, M603, has been carried out by J.Prehn, M.Mercola and K.Calame and is published in *Nucleic Acids Res.* **12**, 8987-9007 (1984).