Transcriptional regulation of the human c-myc gene

Citation for published version:

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Transcriptional regulation of the human c-myc gene

J.C. Lang1, B. Whitelaw2, S. Talbot1 & N.M. Wilkie1

1Beatson Institute for Cancer Research, Garscube Estate, Switchback Road, Bearsden, Glasgow G61 1BD and 2I.A.P.G.R., King's Buildings, West Mains Road, Edinburgh EH9 3JQ, UK.

Summary The involvement of c-myc in the genesis of animal neoplasia is now well documented for several systems. In order to define the precise role played by the c-myc gene in tumorigenesis, a better understanding of the normal regulation of c-myc expression is necessary. We have begun a study of the cis-acting regulatory sequences within the 5' flanking domain of the human c-myc gene. Regions important for c-myc promoter function have been identified by linkage to the coding sequences of the bacterial chloramphenicol acetyltransferase (cat) gene. Promoter deletion studies and in vivo competition assays for c-myc/cat recombinant plasmids have allowed the identification of a proximal 'core' promoter region capable of directing high levels of CAT activity. Further upstream a negative regulatory element (NRE2) has been identified which is capable of repressing c-myc expression and which functions by interaction with a trans-acting factor(s). Preliminary data suggests detection of NRE2 is dependent on both the type and amount of carrier DNA used in transient CAT assays. Initial experiments further indicate the involvement of at least two other distal regulatory domains, a negative regulatory domain (NRE1) and a putative enhancer-type region (E). In vitro footprint analysis has allowed the identification of DNA binding proteins which interact with NRE2 and the 'core' promoter. NRE2 contains binding sites for transcription factors Sp1 and CTF. The 'core' promoter domain appears to be highly complex and possesses several Sp1 binding sites.

The c-myc gene was first described as the oncogene of certain highly transforming avian retroviruses (reviewed by Hayman, 1983). Since then a substantial body of evidence has accumulated suggesting that the c-myc gene is intimately associated with the control of cellular proliferation and differentiation. Thus experimental and spontaneous tumours in rodents and cats have been observed which contain defective retroviruses expressing transduced v-myc genes while a variety of avian, rodent and feline tumours express c-myc genes activated by proviral DNA insertion (Hayward et al., 1981; Corcoran et al., 1984; Neil et al., 1984). The c-myc locus is also interrupted by translocation in human Burkitt's lymphoma (BL) and murine plasmacytoma (MPC), (reviewed by Marcu, 1987) and by amplification both in some promyelocytic leukaemia cell lines and in some cases of human breast cancer (reviewed by Altalmo, 1985).

Further evidence for effects of myc on growth are provided by experiments which show that when the gene, linked to strong transcriptional enhancers, is transfected into cells in culture perturbations in differentiation patterns are observed (Coppola & Cole, 1986). Similarly, the c-myc gene is capable of inducing multiple neoplasms in transgenic mice when fused to immunoglobulin enhancers (Adams et al., 1985) or a mouse mammary tumour virus LTR (Leder et al., 1986).

Unlike the ras gene family, mutations within the coding sequences may not be an important feature in converting myc from a proto-oncogene to an oncogene. Thus feline v-myc coding sequences are virtually unaltered compared to the feline c-myc (Stewart et al., 1986). The bulk of the present evidence suggests that aberrant expression of c-myc is the principal mechanism for oncogenic conversion.

The current view of the structure of the c-myc gene is shown in Figure 1. Expression may be regulated under "normal" circumstances at at least three levels. Firstly, there is regulation at the level of initiation of transcription, affected by cis-acting regulatory domains 5' to the two main cap sites (P1 and P2) and perhaps by other domains within the body of the gene (reviewed by Marcu, 1987). During terminal differentiation of B cells, red cells and monocytes c-myc expression is switched off, and in the case of HL60 promyelocytic cells this partly occurs as a result of reduced initiation of transcription (Siebenlist et al., 1988). Secondly, in quiescent murine fibroblasts, non-proliferating HL60 cells, and in human mononuclear cells (MNC) the absence of full length c-myc transcripts is associated with an apparent block in elongation of already initiated RNA, located at a point close to the junction of exon 1 and the first intron (Bentley & Groudine, 1986a; Eick & Bornkamm, 1986; Nepveu et al., 1987). This block is overcome on mitogen stimulation of the cell cycle (Eick et al., 1987; Nepveu et al., 1987). Thirdly, c-myc RNA is known to have a short half life which is apparently lengthened in some murine plasmacytomas and during differentiation of F9 teratocarcinoma stem cells (Piechaczyk et al., 1985; Dean et al., 1986).

It is not yet fully understood how the changes in the c-myc locus observed in tumour cells affect regulation. The translocations observed in BL and MPC often, but not always, involve breakpoints within the transcriptional control region of c-myc. This might disrupt normal regulatory sequences (Yang et al., 1985). Alternatively, transcriptional regulatory sequences within the immunoglobulin locus translocated to the c-myc gene may exert a dominant effect on c-myc as may retroviralLTRs found adjacent to c-myc in lymphomas of chickens, mice and cats (Hayward et al., 1981; Payne et al., 1984; Corcoran et al., 1984; Neil et al., 1984; O'Donnell et al., 1985). In many cases of BL, however, aberrant myc expression may be due to mutation of sequences located at the 3' end of exon 1 thought to be involved in an elongation block to transcription (Cesarman et al., 1987). Alternatively, mutations located at this site might affect the expression of the proposed MYC1 protein (Hann et al., 1988). The mechanisms of alteration in cases of amplification of c-myc,
and in those cases where no rearrangement in the myc locus is apparent, remain to be investigated.

It seems obvious that greater understanding of the normal regulation of myc expression will be necessary for analysis of the alterations which occur during oncogenic conversion. In this paper we summarise our own studies of cis-acting regulatory sequences within the 5′ flanking domain of the human c-myc gene. Such sequences were identified by linkage to the bacterial chloramphenicol acetyltransferase gene (cat), and in some cases further characterised by analysis of DNA binding proteins using \textit{in vitro} DNAase I footprint analysis.

\textbf{Identification of transcriptionally active domains within the 5′ flanking sequence of c-myc}

We linked the 5′ flanking sequence and most of the first exon sequence of c-myc (nucleotides −2319 to +511, Figure 1) to the bacterial cat gene. This interrupts the myc sequences at which an elongation block to transcription has been suggested and also removes all known coding sequences for MYC 1 and 2 (see Figure 1). The cat gene is then fused to untranslated myc leader sequences and placed under transcriptional control of the myc promoter. A series of 5′ terminal deletions and internal deletions were also constructed. Following transfection into murine fibroblasts (LATK +) and human epithelial cells (HeLa), relative levels of CAT activity were measured using transient assays. In other experiments, various domains of the 5′ flanking sequence were transferred to cat under control of heterologous promoters and CAT activity similarly measured. This approach permitted identification of cis-acting regulatory domains affecting the expression of cat.

The results have been reported in detail elsewhere (Whitelaw et al., 1988) but are summarised in Figure 2. Evidence was obtained for a positive regulatory element, perhaps an enhancer, located between −2319 and −1980 relative to P1. From Figure 2 it can be seen that this domain does not encompass \textit{in vitro} DNAase hypersensitive site I. Closer to the cap site, a cis-acting negative regulatory element (NRE1) encompassing hypersensitive site IIₕ was identified, which mapped between −1527 to −1246. Between −1052 and −607 another strong negative element, NRE2 encompasses hypersensitive site II₂. Much closer to the cap site a positive acting region was found which is in the domain containing hypersensitive sites III₁ and III₂, and we presume this corresponds to the ‘core’ promoter.

\textit{Cis-acting} negative regulatory sequences have also been reported for the human and murine c-myc genes by others (Chung et al., 1986; Remmers et al., 1986; Hay et al., 1986; Lipp et al., 1987). Comparison of the results reported from various laboratories suggests that the situation is complex. Remmers et al. and Lipp et al. find negative regulatory domains located between −1188 and −428 and −1052 and −511 respectively. These results are consistent with our own which locate NRE2 to within a 445bp sequence (−1052 to −607). In contrast, Chung \textit{et al.} and Hay \textit{et al.} describe negative elements between −607 to −407 and −353 to 293 respectively, outside the map location of NRE2. Chung \textit{et al.} (1986) also describe two additional negative regions located in the first exon and the first intron, domains which we have not yet examined. It is suggested that the latter element may be involved in the elongation block proposed by Bentley and Groudine (1986).

It is clear that while various laboratories have detected negative regulatory elements, their reported locations (and the strength of their effects) appear to vary. The reasons for this are not yet known but one factor may be the different cell types and the different conditions used in the transient assays which most laboratories utilize.

Preliminary experiments from our laboratory indicate that one such variable may be the type and amount of carrier DNA used. We measured transient CAT expression from equimolar amounts of \textit{cat/myc} recombinants containing the entire myc 5′ regulatory region or with a terminal deletion to position −290, which removes hypersensitive sites I, II₁, and II₂ and also E, NRE1 and NRE2. Mouse embryo, salmon sperm or pUC18 plasmid DNA was used as carrier in varying amounts. The results are shown in Figure 3 and reveal two important effects. Firstly, no significant difference in CAT activity between the full length and deleted plasmids is obtained at carrier DNA concentrations of less than 10μg per plate. Secondly, the deleted plasmid directs substantially more CAT activity than the full length construction at mouse embryo and salmon sperm carrier DNA concentrations of between 10 and 60μg per plate. In contrast, using pUC18 DNA as carrier, no clear difference in CAT levels between the full length and deleted plasmids was observed at all carrier DNA concentrations measured. However, these effects should not be confused with the general decrease in CAT levels observed as the amount of carrier DNA increases. This additional factor is probably due to reduced efficiency of DNA transfection at higher carrier DNA concentrations and can equally be observed with recombinants containing the cat gene linked to other non-myc promoter sequences (data not shown).

We tentatively propose the following explanation for these effects. The 5′ flanking sequence of myc contains at least four detectable \textit{cis-acting} regulatory elements, an upstream positive element (E), two negative elements (NRE1 and NRE2) and a cap-proximal ‘core’ promoter (see Figure 2), which will all interact with trans-acting factors in the cell. Eukaryote DNA may contain specific sequences which act as targets for some of these trans-acting factors, and can titrate these factors if supplied at sufficiently high concentrations. There is evidence that this can occur in several systems and also in the myc regulatory sequence (see following section).

At low carrier DNA concentrations the amount of target sequence may be too low to titrate trans-acting factors, and the effect of E balances the effects of NREs in LATK + cells. As the concentration of carrier increases, specific sequences in the DNA compete with the upstream E sequences for trans-acting factors which modulate its activity, thus unveiling the negative effects of the NRE sequences which can then be detected by the terminal deletion. pUC18 plasmid DNA is unlikely to have specific target sites for eukaryote trans-acting factors (e.g., TATAA) and might be expected to have no significant effect on CAT levels.

\textbf{Further characterisation of NRE2}

The magnitude of the negative effect of plasmids containing NRE1 and NRE2 vary to some extent from experiment to experiment (Whitelaw \textit{et al.}, 1988). We considered the possibility that the \textit{cis-acting} NREs interact with cellular trans-acting factors present in limiting amounts. Increasing concentrations of DNA containing \textit{myc} promoter sequences might then titrate out such factors leading to derepression of unbound DNA molecules which remain.
In order to test this we performed plasmid competition assays in which the 5' sequences containing only hypersensitive sites I, II, and III (−2319 to −350; Figure 2) are used as competitor in CAT assays where cat is driven by the entire myc promoter region (plasmid pB12, Figure 3). As a control the level of CAT activity obtained with the truncated promoter (position −350) in which cat is driven by only the core promoter containing hypersensitive sites III, and III (see Figures 1 & 2) was also measured. The results presented in Figure 4 show that the competitor DNA can titrate out factors responsible for repression of expression from pB12 and that levels of expression equal to or greater than that obtained with cat driven by the truncated promoter can be obtained.

**Transcription factor binding sites in the regulatory domains**

It is likely that the trans-acting factors which can be titrated in the competition experiments are DNA-binding proteins which affect the activity of the cis-acting regulatory DNA sequences. These might be proteins which recognise and bind to specific sequences, or proteins which interact through protein–protein interactions. In order to begin analysis of the regulatory domain, we examined the interaction of specific DNA-binding proteins with NRE2 and the ‘core’ promoter using partially and highly purified preparations of nuclear proteins in in vivo DNase I protection assays (Jones et al., 1985). The results have been presented in detail elsewhere (Whitelaw et al., 1988) but the footprints obtained using partially purified nuclear extracts binding to NRE2 are shown in Figure 2. Four distinct ‘footprints’ are obtained. Using DNA-affinity purified preparations of transcription factors Sp1 and CTF (CAAT Transcription Factor) (Kadonaga & Tjian, 1985), it was possible to show that of these one binding site was due to Sp1 and two adjacent binding sites to CTF. Examination of the DNA sequence confirms typical consensus signals for these DNA-binding proteins. The fourth footprint encompasses the reported location of hypersensitive site II (Siebenlist et al., 1984) and although it contains an Sp1-like sequence it bound neither purified Sp1 nor CTF. This binding site may represent a target for a so far unknown DNA-binding protein.

A summary of this study and a similar analysis of the ‘core’ promoter domain is included in Figure 2. This shows the location and orientation of the DNA binding sites relative to the other known features of the regulatory region. We have not yet analysed NRE1 or E, or the NRE domains suggested by the reports of Chung et al. (1986) or Hay et al. (1987).
Discussion
Cumulative experimental evidence to date suggests that transcriptional control of the c-myc oncogene is particularly complex. Despite efforts from a number of research groups, the domains responsible for transcriptional regulation have yet to be fully characterized. The emerging picture is one of intricate interactions between multiple positive and negative control elements. C-myc is expressed in many different cell types. It appears to encode a protein(s) intimately involved in the regulation of cell growth and differentiation and as such may itself require to be regulated very precisely. It is plausible that varying interactions occur between control domains and trans-acting factors depending on the cell type and stage of differentiation. That negative regulatory elements identified by transient assays have in vivo significance is suggested by the observations that such an element detected in murine myc (which may correspond to our NRE2) correlates with the position of chromosomal breakpoints and proviral insertion sites in B- and T-cell neoplasia (Remmers et al., 1986). Similarly mutations in the 3' region of exon 1 in BL cells correlates with aberrant patterns of myc transcription and may affect elongation block, or perhaps the expression of MYC I (Cesarman et al., 1987; Hahn et al., 1988).

Further efforts to identify the in vivo regulatory function of these and the other elements described to date are obviously crucial. It is also important to determine whether alterations in the DNA sequence of the elements themselves, or alteration in the trans-acting factors which regulate their activity, are involved in the alteration of c-myc expression associated with various types of cancer apart from BL.

From the summary shown in Figure 2, we suggest a role for specific DNA-binding proteins Sp1 and CTF in modulating the activity of NRE2. Binding of these general transcription factors have been associated so far with up-regulation of promoter elements. If they are also involved in negative regulation, the specificity of their effect on transcription remains to be explained. Specificity could be determined by the structure of the local DNA sequence itself, by the orientation or spacing of binding or the nature of other specific DNA binding proteins which interact with the domain. Alternatively, specificity could be determined by protein–protein interactions with other non-DNA binding proteins, or a combination of all of these. Regulation of cis-acting DNA sequences in transcription by the interaction of accessory proteins with specific protein-DNA complexes has been reported for other biological systems including the interaction of herpes simplex virus with host cellular proteins (Preston et al., 1988). Another possible mode of action of a negative element might be to act as a promoter and lead to 'promoter occlusion effects' on downstream promoters as previously suggested for bacteriophage lambda gene expression (Adhya & Gottesman, 1982). In this regard, the presence of an additional cap site (PO) has been reported adjacent to NRE2 (Bentley & Groudine, 1986a).

It is obvious that the mechanisms by which expression of the myc gene is regulated is both complex and little understood at the present time. We are currently investigating the role of Sp1, CTF and the unknown protein which binds at hypersensitive site II*, and the possible role of accessory proteins, using a combination of site directed mutagenesis, DNase I protection studies, and analysis of protein–DNA complexes.

We thank the Cancer Research Campaign of Great Britain for support.

References


