MOLECULAR AND CELLULAR BIOLOGY, May 1987, p. 1697–1709
0270-7306/87/051697-13$02.00/0

Vol. 7, No. 5

# Definition of Regions in Human c-*myc* That Are Involved in Transformation and Nuclear Localization

JAMES STONE,[1] TITIA DE LANGE,[2] GARY RAMSAY,[2, 3] EDWARD JAKOBOVITS,[2]† J. MICHAEL BISHOP,[2,3] HAROLD VARMUS,[2] AND WILLIAM LEE[4]*

*Jackson Laboratory, Bar Harbor, Maine 04609[1]; Hooper Research Foundation, San Francisco, California 94143[3]; and Department of Microbiology and Immunology,[2] and Cancer Research Institute, Department of Medicine,[4] University of California, San Francisco, California 94143*

Received 19 November 1986/Accepted 4 February 1987

**To study the relationship between the primary structure of the c-*myc* protein and some of its functional properties, we made in-frame insertion and deletion mutants of the normal human c-*myc* coding domain that was expressed from a retroviral promoter-enhancer. We assessed the effects of these mutations on the ability of c-*myc* protein to cotransform normal rat embryo cells with a mutant *ras* gene, induce foci in a Rat-1-derived cell line (Rat-1a), and localize in nuclei. Using the cotransformation assay, we found two regions of the protein (amino acids 105 to 143 and 321 to 439) where integrity was critical: one region (amino acids 1 to 104) that tolerated insertion and small deletion mutations, but not large deletions, and another region (amino acids 144 to 320) that was largely dispensable. Comparison with regions that were important for transformation of Rat-1a cells revealed that some are essential for both activities, but others are important for only one or the other, suggesting that the two assays require different properties of the c-*myc* protein. Deletion of each of three regions of the c-*myc* protein (amino acids 106 to 143, 320 to 368, and 370 to 412) resulted in partial cytoplasmic localization, as determined by immunofluorescence or immunoprecipitation following subcellular fractionation. Some abnormally located proteins retained transforming activity; most proteins lacking transforming activity appeared to be normally located.**

Regulated expression of the c-*myc* proto-oncogene is implicated in the regulation of vertebrate cell proliferation (3, 16, 17, 31); and aberrant expression contributes to the pathogenesis of a variety of malignancies, including lymphoid tumors of chickens (29, 42), mice (1, 12, 14, 49, 54, 56), and humans (1, 15, 56) and human small cell lung cancers (38). The normal c-*myc* gene of chickens (50), mice (53), and humans (5, 24, 59) contains three exons, the second and third of which encode the protein product. The c-*myc* proteins are phosphoproteins with short half-lives (15 to 30 min) which localize in the nucleus and bind DNA in vitro (18, 20, 26–28, 43, 45). The human c-*myc* protein is predicted to be 439 amino acids long, with a molecular weight of 49,000 (5, 57, 60), but specific antisera precipitate protein species with $M_r$s of 62,000 to 67,000 (27, 45).

The involvement of c-*myc* in normal and neoplastic growth makes it important to understand its function(s) and the structural basis of some of its properties. Yet, despite an abundance of sequence information on various normal and tumor-derived c-*myc* genes (2, 5, 44, 50, 53, 58, 59, 60, 62), assays for its activity in vitro (33, 36, 37), and knowledge of certain characteristics (e.g., nuclear localization) that are believed to be important for activity (18, 20, 43), little is known about how structure relates to function in this protein. In part, this is due to the absence of naturally occurring mutants of c-*myc* with discernible phenotypes. Therefore, we undertook a systematic mutational analysis of the gene and determined how mutations affect transforming activity and nuclear localization of the protein product.

* Corresponding author.
† Present address: Genentech, South San Francisco, CA 94080.

## MATERIALS AND METHODS

**Mutagenesis.** The plasmid pM21 (Fig. 1) served as the substrate for mutagenesis. This pUC8 plasmid contains a normal human c-*myc* gene from the *Xho*I site in exon 1 to the *Eco*RI site 3′ to exon 3 (5) ligated to a Moloney murine leukemia virus (MoMLV) U3 and R, modified to contain a polylinker with a *Sal*I site at the *Kpn*I site of R (37). The encoding exons of human c-*myc* have 14 *Alu*I recognition sequences (5), and pM21 contains no *Xho*I recognition sequence; therefore, we used *Alu*I to linearize the plasmid and introduced a dodecanucleotide specifying a pair of *Xho*I sites (CTCGAGCTCGAG) as the mutagen (55). After *Xho*I digestion and religation, one or more 6-base-pair (bp) insertions were found, depending on the completeness of *Xho*I digestion. The insertion mutants were designated In followed by the position of the first codon containing nucleotides derived from the inserted oligomer. These insertion mutants and the amino acid changes due to the insertions are shown in Table 1.

We used a similar strategy, using a partially self-complementary dodecanucleotide containing a *Xho*I recognition sequence and staggered ends complementary to *Cla*I-generated ends (CGACCTCGAGGT) to make an insertion mutant at the *Cla*I site near the 5′ end of exon 3. A mutant (In262m) containing one 12-bp insertion was obtained by using nonphosphorylated oligonucleotides, and a mutant (In262p) containing multiple 12-bp insertions was recovered after phosphorylated oligonucleotides were used.

For some mutants, it was desirable to determine whether the inserts were monomeric or polymeric. In the case of some, it was possible to determine this by restriction analy-

sis (Table 1). For mutants In370, In373, and In412, which were found to be totally nonfunctional in our assays, nucleic acid sequencing was performed by the dideoxy method (9); each mutant was found to have two copies of the *Xho*I hexanucleotide inserted in-frame in the expected *Alu*I site. In the case of In370 (designated here as In370d), we reduced the linker to a single *Xho*I hexamer (In370m), which was confirmed by sequence analysis.

In-frame deletions throughout the c-*myc* gene were generated by using pairs of insertion mutants. Pairs with *Xho*I linkers inserted in the same reading frame were recombined by ligating the appropriate fragments isolated after digestion of plasmids with both *Xho*I and *Eco*RI. Another large in-frame deletion was created in exon 2 by removing the 414-bp *Pst*I fragment. When in-frame deletions were wanted between two insertion mutants with *Xho*I linkers inserted in different reading frames, the *Xho*I 4-base, single-stranded ends of the appropriate fragments were made blunt by digestion with single-strand-specific mung bean nuclease or filled in with the Klenow fragment of *Escherichia coli* DNA polymerase I and ligated. Ligation of ends produced by *Xho*I digestion and treatment with mung bean nuclease generates *Hae*II sites because of the *Alu*I nucleotides that flank the former *Xho*I sites; filled in *Xho*I ends ligated together generate novel *Pvu*I sites. These maneuvers cause a +2 or a +1 base shift in the reading frame of the codons 3′ to the blunt-end ligation, respectively, and allow subsequent c-*myc* codons to be translated correctly.

Additional deletion mutants were made in exon 3 by using *Bal* 31 exonuclease. pM21 was linearized with *Cla*I and treated with *Bal* 31, and the ends were made blunt with the Klenow fragment of *E. coli* DNA polymerase I. The shortened exon 3 fragments were released by digestion with *Eco*RI, isolated, and ligated to a pM21 fragment from which the *Cla*I to *Eco*RI portion was removed by using a pair of oligonucleotides (CGATGTTCTCGAG and CTCGAG AACAT) that specify a *Xho*I site internally, a *Cla*I two-base overhang at the 5′ end, and a blunt 3′ end. Deletion mutants generated by this procedure were sequenced by the dideoxy method with plasmid DNA (9) to determine the ones that were in the correct reading frame. Deletion mutants are designated D followed by the positions of the c-*myc* codons that were deleted. The deletion mutants contain synthetic oligonucleotide-derived sequences between these codons that encode extra amino acids (Table 2). Mutants with deletions in two portions of c-*myc* were generated by using the unique *Bst*EII site in exon 2 to recombine fragments with deletions on either side (Table 3).

Many of the mutants were transferred to a vector containing the simian virus 40 (SV40) origin of replication. The *Hin*dIII to *Eco*RI fragments (containing both the MoMLV enhancer-promoter and the c-*myc* moieties) of the various mutants c-*myc* genes were excised from pUC8 and introduced into the *Hin*dIII to *Eco*RI sites of pSVSP65 (gift from N. Hay). This plasmid has the *Pvu*II to *Hin*dIII fragment of SV40 (containing the origin of replication and early promoter) subcloned into the *Pvu*II and *Hin*dIII sites of pSP65 (Promega Biotech).

**Transfection of cells.** Rat embryo cells (RECs) were harvested and transfected as described previously (36, 37). Most mutant c-*myc* genes were tested for REC cotransforming activity in at least three separate transfection experiments (Tables 1 to 3) by using plasmids derived from two separate preparations. The ability of a mutant to form foci with EJ*ras* (the mutant c-Ha-*ras*-1 gene present in the EJ/T24 bladder carcinoma cell line [41]) was normalized to the activity of pM21 in the same experiment, after accounting for foci formed after transfection with EJ*ras* alone in the same experiment (0 to 1%).

Rat-1 cells were cotransfected by the calcium phosphate technique (22) by using a plasmid containing the c-*myc* or mutant c-*myc* gene and a plasmid conferring resistance to G418 (GIBCO Laboratories, Grand Island, N.Y.) (51). Cells were selected in 400 μg of G418 per ml.

COS-7 cells (21) were transfected in serum-free medium containing 3 mg of DEAE-dextran (Pharmacia Fine Chemicals, Piscataway, N.J.) per ml and 1 μg of the appropriate plasmid per ml for 5 h. Cells were used for assays 24 to 48 h after transfection.

**Retroviruses bearing a c-*myc* cDNA.** A c-*myc* cDNA covering exons 2 and 3 was reconstituted by using a cloned partial cDNA that extends across intron 2 (a gift from C. Croce [60]) and a genomic exon 2 fragment. A portion of this cDNA, extending from the *Tha*I site directly upstream of the initiator ATG codon to the *Nsi*I site 150 bp downstream of the TAA stop codon, was inserted into the *Sma*I and *Pst*I sites of the pSP65 polylinker; this insert then became flanked by an *Eco*RI site upstream and a *Hin*dIII site downstream. Mutations were introduced into this cDNA by exchange of appropriate restriction fragments from exon 2 or 3 of mutated genomic genes. Mutant and normal c-*myc* cDNAs were subsequently inserted between the *Eco*RI and *Hin*dIII sites of the Moloney proviral construct in plasmid pMV6 (a gift from I. B. Weinstein). *myc* protein was expressed from the retroviral genomic mRNA by using the *myc* AUG codon as the translational start site. The provirus in pMV6 also carries a neomycin resistance gene linked to a thymidine kinase promoter located 3′ to the *Eco*RI and *Hin*dIII sites, in the same transcriptional orientation as the *myc* cDNA.

ψ2 cells (39) were transfected with 5 μg of proviral construct by using calcium phosphate (22) and selected in 400 μg of G418 per ml. Hundreds of colonies of G418-resistant cells that were transfected with each construct were pooled, and the supernatant was harvested. Each construct yielded approximately the same virus titer (about $10^4$/ml) as measured by biological activity (G418 resistance) and by analysis of viral RNA for *myc* and neomycin resistance gene sequences by using dot blots.

**Transformation assay in Rat-1a cells.** Rat-1a cells (a subline of Rat-1 cells; $5 \times 10^5$ cells per dish [diameter, 60 mm]) were infected with about 500 infectious units of the recombinant retroviruses in the presence of 8 μg of polybrene per ml for 4 to 16 h and subsequently split 1:4 into media with or without 400 μg of G418 per ml. Foci were counted at day 7 post-infection and expressed as the percentage of the number of G418-resistant colonies.

**Analysis of *myc* RNA and protein.** Total cellular RNA was harvested from cells either by using guanidinium isothiocyanate as described previously (10, 37) or the method described by Auffray and Rougeon (4). Analysis was performed either by Northern blotting (for RNA from Rat-1a cells [11]) or RNase protection with radiolabeled RNA probe (for RNA from RECs and Rat-1 cells [23]) complementary to the portion of normal human c-*myc* mRNA derived from the *Pst*I to *Pst*I segment of exon 2 (13).

Protein was analyzed by immunoprecipitation of [$^{35}$S]methionine- or $^{32}P_i$-labeled cells by using rabbit anti-*myc* antisera. Most of the analyses were performed with antiserum raised against the C-terminal 32-amino-acid peptide (45), but other antisera (gifts from K. Moelling and G. Evan) were used in some experiments.

**Subcellular localization of c-*myc* proteins within cells.**

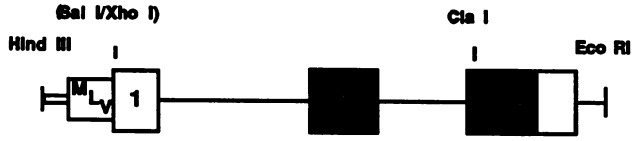


FIG. 1. The structure of the 6.4-kilobase DNA insert in pM21, the parental plasmid from which all the human c-*myc* mutants were derived. Part of the human c-*myc* gene, from the *Xho*I site in exon 1 to the *Eco*RI site 3′ to exon 3, is under the transcriptional control of the MoMLV enhancer-promoter (see text) (37). Shaded areas represent the c-*myc* coding domain, the c-*myc* exons are numbered, and landmark restriction sites are indicated.

Immunofluorescent staining was performed on COS-7 cells that were plated onto cover slips. The cells were fixed with 3.5% paraformaldehyde and made permeable by treatment with 0.1% Nonidet P-40 (NP-40). After exposure to the rabbit anti-*myc* antiserum, the cells were treated with rhodamine-conjugated goat anti-rabbit antibody (34).

Nuclear and cytoplasmic fractionation was performed on transfected COS-7 cells by using ice-cold hypotonic buffer followed by treatment with 0.5% NP-40. Following separation by centrifugation, nuclei were solubilized in lysis buffer, and the fractions were analyzed by immunoprecipitation (34).

## RESULTS

**General strategy.** We mutated the normal coding domain from a human c-*myc* gene ligated to the MoMLV promoter-enhancer (Fig. 1). The parental c-*myc* gene is efficiently expressed in rat cells and can cooperate with EJ*ras* (41) to cotransform normal RECs (37). We created insertion mutations throughout the coding domain by introducing oligonucleotides containing *Xho*I sites into *Alu*I sites and into the unique *Cla*I site (55); 11 of 14 *Alu*I sites in the human c-*myc* coding domain were mutated (Table 1). Using the unique *Xho*I sites from pairs of these linker-insertion mutants, we generated in-frame deletion mutants. We made another in-frame deletion by removing the *Pst*I-*Pst*I fragment of exon 2. Finally, we made additional deletions extending 3′ from the *Cla*I site in exon 3 by using *Bal* 31 nuclease. Deletion mutants may contain extra codons specified by synthetic linker nucleotides (Table 2). Linker-insertion mutants are designated In followed by the position of the first codon containing a synthetic linker nucleotide. Deletion mutants are designated D followed by the positions of the first and last c-*myc* codons that were deleted. We made more extensive deletions by combining two deletions in adjacent or separate regions of the gene (Table 3).

All mutant genes were tested in the REC cotransformation assay to assess their ability to cooperate with EJ*ras* in transforming normal cells. Some of the mutants were engineered into retroviruses and used to infect Rat-1a cells, which are susceptible to morphologic transformation by wild-type c-*myc* (A. Bruskin, manuscript in preparation). These cells provided an alternative assay for the transforming activity of c-*myc* mutants. Finally, we examined the subcellular distribution of the proteins produced by the mutants by using immunofluorescence and immunoprecipitation after nuclear-cytoplasmic fractionation.

**REC cotransformation activity of the mutants.** Results of cotransfection of the mutant c-*myc* genes into RECs with the EJ*ras* plasmid are presented in Tables 1 to 3. Cotransformation experiments showed some inherent variability, perhaps due to variation among embryos from which the RECs were derived. Because of this, we grouped mutants into three broad categories for the purposes of discussion. (i) Those with 25 to 100% of simultaneously determined wild-type (pM21) activity (50 to 150 foci per $10^6$ transfected RECs) were considered similar to wild-type c-*myc*. (ii) Those with 2 to 25% efficiency were considered significantly less active.

TABLE 1. Human c-*myc* linker-insertion mutants

| Mutant | Restriction site mutated | Position of codons affected[a] | Codons after mutagenesis[b] | Reading frame of *Xho*I linker[c] | Comment[d] | REC cotransforming activity[e] | Protein visualized[f] |
|---|---|---|---|---|---|---|---|
| pM21 | | Wild type | | | | 100 | IF (N), IP (N) |
| In6 | *Alu*I | 6–7 | Ser, (Ser, Ser)$_?$, Phe | 3 | | 98 ± 9 | IF (N), IP |
| In40 | *Alu*I | 39–40 | Glu, (Leu, Glu)$_?$, Leu | 1 | | 95 ± 14 | IF (N), IP |
| In55 | *Alu*I | 54–55 | Glu, (Leu, Glu)$_?$, Leu | 1 | | 74 ± 22 | IF (N), IP |
| In92 | *Alu*I | 92–93 | Ser, (Ser, Ser)$_?$, Phe | 3 | | 56 ± 30 | IF (N), IP |
| In105 | *Alu*I | 104–105 | Glu, (Leu, Glu) $_?$, Leu | 1 | | 14 ± 8 | IF (N), IP (N) |
| In144 | *Alu*I | 143–144 | Lys, (Leu, Glu)$_?$, Leu | 1 | | 57 ± 23 | IF (N), IP |
| In203 | *Alu*I | 203–204 | Ser, (Ser, Ser)$_?$, Ser | 3 | | 74 ± 23 | IF (N), IP |
| In262m | *Cla*I | 262–263 | Ile, (Asp, Leu, Glu, Val), Asp | 1 | No *Sal*I site[g] | 58 ± 33 | IF (N), IP (N) |
| In262p | *Cla*I | 262–263 | Ile, (Asp, Leu, Glu, Val)$_n$, Asp | 1 | *Sal*I site[g] | 32 ± 7 | NT |
| In370m | *Alu*I | 369–370 | Glu, (Leu, Glu), Leu | 1 | Sequenced | 0 ± 0 | NT |
| In370d | *Alu*I | 369–370 | Glu, (Leu, Glu, Leu, Glu), Leu | 1 | Sequenced | 0 ± 0 | IF (N), IP (N) |
| In373 | *Alu*I | 373–374 | Ser, (Ser, Ser, Ser, Ser), Phe | 3 | Sequenced | 0 ± 0 | IF (N), IP (N) |
| In413 | *Alu*I | 412–413 | Lys, (Leu, Glu, Leu, Glu), Leu | 1 | Sequenced | 1 ± 1 | IF (N), IP (N) |
| In434 | *Alu*I | 433–434 | Gln (Leu, Glu)$_n$, Leu | 1 | *Sac*I site[h] | 54 ± 16 | IF (N), IP (N) |

[a] Codons containing the restriction site that was mutated.

[b] Inserted codons are in parentheses. *n* indicates that two or more and ? indicates that an unknown number of codon pairs, which were derived from linker sequences, were inserted. Original codons are underlined.

[c] The number indicates the position in the codon occupied by the first nucleotide of the linker.

[d] Indicates how the number of linkers was determined (see text for details).

[e] Expressed as percentage of pM21 (wild-type c-*myc*) activity ± standard deviation. Results are based on three to seven independent experiments except for In370m activity, which is based on two experiments.

[f] Proteins detected by immunofluorescent staining of transfected COS-7 cells (IF), immunoprecipitation (IP), or both; NT, not tested. Proteins that are nuclear (N) by immunofluorescence or subcellular fractionation and immunoprecipitation are indicated.

[g] Tandem multimers of the linker used create *Sal*I sites (see text).

[h] The nucleotide preceding the *Alu*I recognition sequence is not a guanine so that after insertion of the linker, a *Sac*I site is not generated unless a tandem multimer of the linker is present.

TABLE 2. Human c-*myc* deletion mutants

| Mutant | Mutants used in construction | No. of amino acids[a]: Deleted | No. of amino acids[a]: Added | Comment[b] | REC cotransforming activity[c] | Protein visualized[d] |
|---|---|---|---|---|---|---|
| D3-38 | N-*myc* In40 | 36 | 9 (amino acids 3-11 of N-*myc*) | *Xho*I site in codons 11 and 12 of N-*myc* | 33 ± 23 | NT |
| D3-53 | N-*myc* In55 | 51 | 9 (amino acids 3-11 of N-*myc*) | *Xho*I site in codons 11 and 12 of N-*myc* | 17 ± 13 | NT |
| D3-103 | N-*myc* In105 | 101 | 9 (amino acids 3-11 of N-*myc*) | *Xho*I site in codons 11 and 12 of N-*myc* | 1 ± 1 | NT |
| D7-38 | In6, In40 | 3 | 2 (Ser, Ile) | Klenow-treated *Xho*I ends | 61 ± 16 | NT |
| D7-91 | In6, In92 | 8 | 2 (Ser, Ser) | | 0 ± 0 | IF (N), IP (N) |
| D41-53 | In40, In55 | 5 | 2 (Leu, Glu) | | 53 ± 29 | IF (N), IP |
| D41-103 | In40, In105 | 65 | 2 (Leu, Glu) | | 8 ± 6 | IF (N), IP |
| D41-178 | | 138 | 0 | *Pst*I-*Pst*I fragment deleted | 0 ± 0 | IF (N), IP |
| D55-92 | In55, In92 | 38 | 1 (Arg) | Mung bean nuclease-treated *Xho*I ends | 26 ± 20 | NT |
| D56-103 | In55, In105 | 48 | 0 | | 11 ± 7 | IF (N), IP |
| D93-103 | In92, In105 | 11 | 2 (Ser, Ile) | Klenow-treated *Xho*I ends | 27 ± 18 | NT |
| D93-201 | In92, In203 | 109 | 0 | | 0 ± 0 | IF (N), IP (N/C) |
| D106-143 | In105, In144 | 38 | 1 (Glu) | | 0 ± 0 | IF (N), IP (N/C) |
| D145-262 | In144, In262m | 118 | 2 (Glu, Val) | | 80 ± 35 | IF (N) |
| D264-368 | In262m, In370 | 105 | 1 (Leu) | | 0 ± 0 | IF (N/C), IP (N/C) |
| D264-412 | In262m, In413 | 149 | 2 (Leu, Glu) | | NT | IF (N/C), IP (C) |
| D371-412 | In370, In413 | 42 | 1 (Glu) | | 0 ± 0 | IF (N), IP (N/C) |
| D414-433 | In413, In434 | 20 | 1 (Glu) | | 0 ± 0 | IP (N) |
| D265-289 | | 25 | 2 (Leu, Glu) | *Bal* 31 digestion | 48 ± 8 | IF (N) |
| D265-295 | | 31 | 2 (Leu, Glu) | *Bal* 31 digestion | 41 ± 3 | IF (N) |
| D265-301 | | 37 | 2 (Leu, Glu) | *Bal* 31 digestion | 62 ± 22 | IF (N) |
| D265-304 | | 40 | 2 (Leu, Glu) | *Bal* 31 digestion | 67 ± 30 | IF (N) |
| D265-317 | | 53 | 2 (Leu, Glu) | *Bal* 31 digestion | 54 ± 32 | IF (N) |
| D265-320 | | 56 | 2 (Leu, Glu) | *Bal* 31 digestion | 36 ± 20 | IF (N) |
| D265-324 | | 61 | 2 (Leu, Glu) | *Bal* 31 digestion | 14 ± 12 | IF (N/C) |
| D265-329 | | 65 | 2 (Leu, Glu) | *Bal* 31 digestion | 11 ± 11 | IF (N/C) |
| D265-341 | | 77 | 2 (Leu, Glu) | *Bal* 31 digestion | 3 ± 3 | IF (N/C) |
| D265-350 | | 86 | 2 (Leu, Glu) | *Bal* 31 digestion | 4 ± 2 | IF (N/C) |
| D265-353 | | 89 | 2 (Leu, Glu) | *Bal* 31 digestion | 10 ± 7 | IF (N/C) |
| D265-367 | | 103 | 2 (Leu, Glu) | *Bal* 31 digestion | 0 ± 0 | IF (N/C) |

[a] We assumed that *Xho*I cut all its sites in cases in which more than one linker was inserted.
[b] Some details of how the deletion mutants were made (see text).
[c] Expressed as percentage of pM21 (wild-type c-*myc*) activity ± standard deviation. Results are based on three to seven independent experiments except for D265-295, D265-301, and D265-367 activities, which are based on two experiments.
[d] Proteins detected by immunofluorescent staining of transfected COS-7 cells (IF), immunoprecipitation (IP), or both; NT, not tested. Proteins that are nuclear (N), cytoplasmic (C), or nuclear/cytoplasmic (N/C) by immunofluorescence or subcellular fractionation and immunoprecipitation are indicated.

(iii) Those with <2% efficiency were considered inactive (Fig. 2) As the data in Tables 1 to 3 indicate, however, the activities of mutants grouped within one category may differ two- to fivefold or more, whereas the activities of a few mutants in adjacent categories differ by less than twofold. Nevertheless, on the basis of the relative activities of the mutants, we can divide the normal human c-*myc* protein into four regions that differ in their susceptibility to mutational inactivation of cotransforming potential (Fig. 2).

Two regions of human c-*myc* appear to be most vulnerable to mutational inactivation (Fig. 2). Both insertion and deletion mutations in the C-terminal portion of human c-*myc*,

TABLE 3. Human c-*myc* combination deletion mutants

| Mutant | Mutants used in constuction | No. of amino acids: Deleted | No. of amino acids: Added | Comment | REC cotransforming activity[a] |
|---|---|---|---|---|---|
| D145-304 | In144, D265-304 | 160 | 2 | | 16 ± 13 |
| D145-317 | In144, D265-317 | 173 | 2 | | 14 ± 11 |
| D145-329 | In144, D265-329 | 185 | 2 | | 4 ± 4 |
| D145-341 | In144, D265-341 | 197 | 2 | | 3 ± 3 |
| D145-353 | In144, D265-353 | 209 | 2 | | 6 ± 2 |
| D3-38/145-262 | D3-38, D145-262 | 154 | 11 | Two regions deleted | 24 ± 18 |
| D3-38/145-304 | D3-38, D145-304 | 196 | 11 | Two regions deleted | 5 ± 5 |
| D3-38/145-329 | D3-38, D145-329 | 221 | 11 | Two regions deleted | 8 ± 7 |
| D3-38/145-341 | D3-38, D145-341 | 233 | 11 | Two regions deleted | 13 ± 4 |
| D3-38/145-353 | D3-38, D145-353 | 245 | 11 | Two regions deleted | 10 ± 8 |
| D41-53/145-262 | D41-53, D145-262 | 133 | 4 | Two regions deleted | 30 ± 7 |

[a] Activity is expressed as a percentage of pM21 (wild-type c-*myc*) activity ± standard deviation. Values are based on three to four independent experiments except for D3-38/145-341, which is based on two experiments.

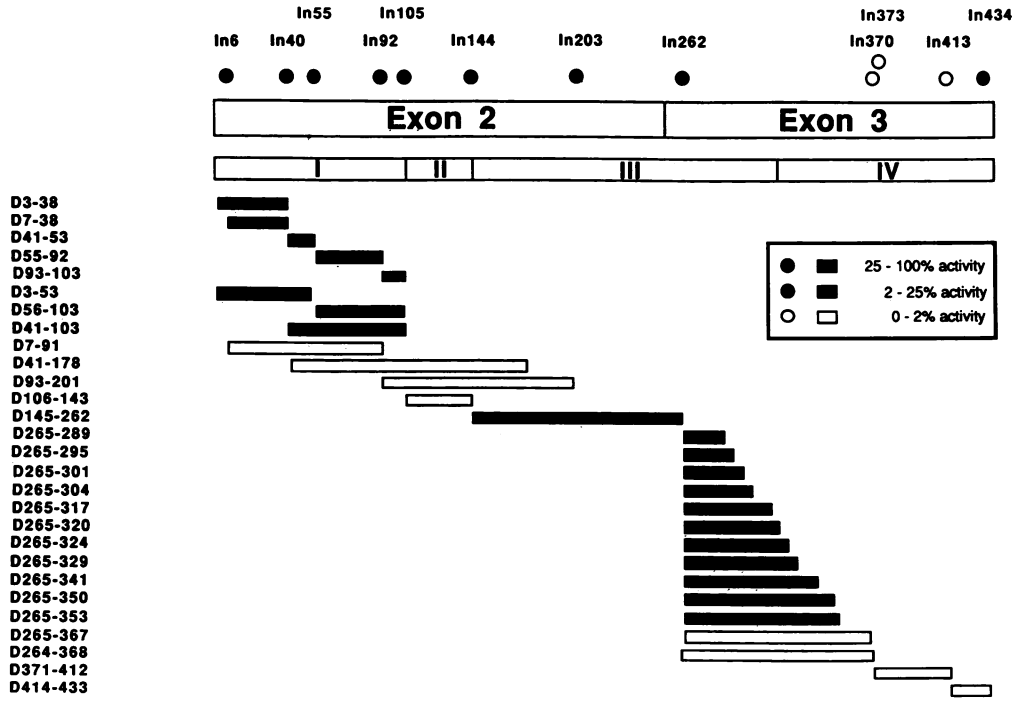


FIG. 2. REC cotransforming activity of c-*myc* insertion (In) and deletion (D) mutants. Portions of the c-*myc* protein derived from exons 2 and 3 are represented by wide rectangular boxes, and positions of insertion mutants are designated by overlying circles. Regions that are missing from proteins encoded by deletion mutants are represented by thin boxes. The level of activity of the mutant proteins is denoted by shading, as indicated in the inset. Regions of the c-*myc* protein designated I, II, III, and IV are defined in the text. For details about the mutants and their levels of activity, see Tables 1 and 2.

extending from codon 320 to near the C terminus (region IV), adversely affect and frequently destroy the cotransforming function. The integrity of the middle portion of this region appears to be absolutely essential because In370, In373, and In413 are totally inactive. We sequenced the mutated regions of these insertion mutants and confirmed that each contains an in-frame *Xho*I dodecanucleotide insert. Reduction of the dodecanucleotide in In370d to a hexanucleotide (In370m) did not increase activity. Not surprisingly, deletion mutations involving the middle portion of region IV (D264-368, D265-367, D371-412, and D414-433) were also totally inactive. Mutants with deletions that began in region III and that encroached on the 5′ end of region IV (D265-324, D265-329, D265-341, D265-350, and D265-353) were less active than mutants with deletions that came close to but did not remove any sequences from region IV (D265-317 and D265-320). The importance of sequences near the C terminus has not been adequately assessed because we only had a single mutant involving this region (In434, which is fully active).

Region II (from codons 105 through 143) also appeared to be essential for cotransforming activity. Insertion at codon 105 diminished activity, and removal of region II (D106-143, D41-178, and D93-201) abolished activity. We have not yet attempted to subdivide this region to determine whether some sequences are expendable for cotransforming activity.

In the N-terminal portion of c-*myc*, from codons 6 to 104 (region I), no single segment appeared to be essential for cotransforming activity. The insertion mutants and the four small- to moderate-sized deletions (D7-38, D41-53, D55-92, and D93-103), which together span most of region I, were in the highest activity category. We made D3-38 by attaching the first 11 codons of human N-myc (up to the *Xho*I site [52]) to the *Xho*I site of In40 and found that it has activity similar to that of D7-38, showing that differences between residues 3 and 6 of c-*myc* and 3 and 11 of N-*myc* (35, 52) are insignificant in this assay. Larger deletions, encompassing two or more segments of region I (D56-103 and D41-103), significantly diminished but did not eliminate activity. However, when region I was almost completely removed, as in D7-91, activity was abolished. Thus, region I is important for cotransforming activity, even though individual segments may be dispensable.

The middle third of human c-*myc*, from codon 144 to the vicinity of codon 320 (region III), was largely dispensable for REC cotransformation. Insertion mutations (In144, In203, In262m, and In262p) and large deletions in this region (D145-262 and D265-320) did not affect activity greatly, and even removal of the entire region (as in D145-329) permitted cotransforming function at a reduced level.

Double mutants (Table 3) that had deletions in both regions I and III were, in general, less active than mutants with deletions in only one of these regions. Only the most active of these double deletion mutants, D41-53/145-262, had activity comparable to that of its parents. Surprisingly, D3-38/145-353, which makes a protein containing 205 amino acids (of which 194 are originally from c-*myc*) and is less than half the size of the wild-type c-*myc* protein, retained significant cotransforming activity (10% of that of the wild type).

None of the mutants tested demonstrated greater REC cotransforming activity than the normal gene under the conditions of our assay. However, because we did not search for activating mutations, e.g., by using weaker promoters or by using less plasmid DNA in our REC cotransfections, we cannot be certain that none of our mutants is more active than the normal gene. We examined some of our

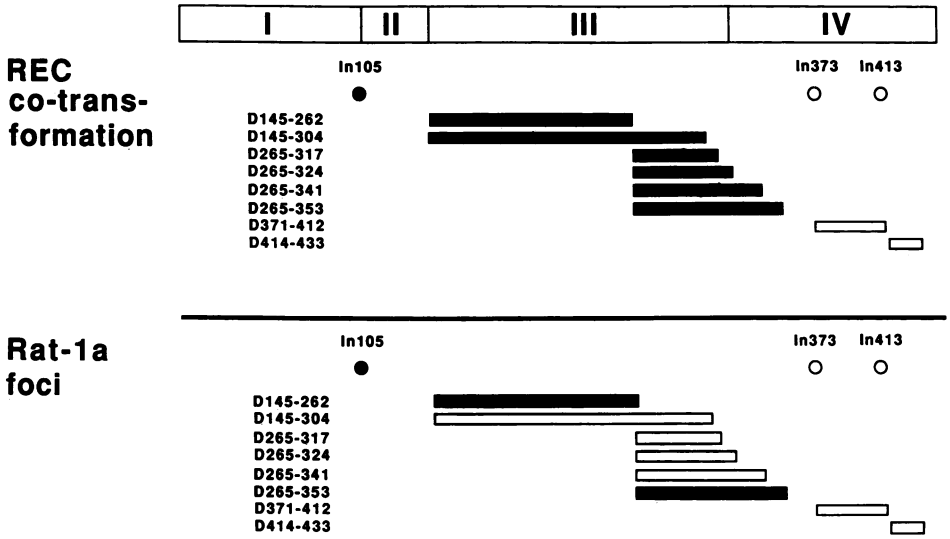


FIG. 3. Comparison of the activity of selected c-*myc* mutants in the Rat la focus-forming assay and the REC cotransformation assay. The display follows the conventions used in Fig. 2. Calculation of activity in the two assays was performed as described in the text.

c-*myc* mutants for temperature-sensitive cotransforming activity by incubating the cotransfected RECs at 32, 37, and 39°C; but none had differential activity at the lower or higher temperature.

The RECs that were cotransformed by the c-*myc* mutants and EJ*ras* appeared to be similar to those cotransformed by the normal gene and EJ*ras*. All were round, refractile, and variable in size. They overgrew the surrounding monolayer and became cell lines, although RECs cotransformed by some of the very large deletion mutants (e.g., D145-341 and D145-353) appeared to develop into cell lines less efficiently. We tested RECs that were cotransformed by each of the active insertion mutants and by D41-53, D56-103, D41-103, and D145-262 for tumorigenic potential by injecting $10^5$ to $10^6$ transformed cells cloned from individual foci into Fisher rats. All tested cells formed progressively growing solid tumors that were detectable after a 2- to 4-week latency period.

**Transformation of an established cell line with mutant c-*myc* genes.** Having defined regions of c-*myc* by their relative importance for cotransforming activity, we wished to determine whether this map applies to c-*myc* activity measured with a different assay. Recently Rat-1a cells were found to undergo morphological transformation after infection with a retrovirus carrying a normal c-*myc* cDNA, whereas parental Rat-1 cells did not show gross morphologic change after infection with a similar virus. Furthermore, *myc*-transformed Rat-1a cells form foci, grow in soft agar, and form tumors in nude mice (A. Bruskin, manuscript in preparation). We used Rat-1a cells to test the effect of some of our c-*myc* mutants.

Eleven mutations were separately engineered into a c-*myc* cDNA, which was then introduced into a retroviral vector (gift from I. B. Weinstein) that also carries the neomycin resistance gene (51). Helper-free retroviruses carrying the mutant *myc* cDNAs were produced by using the ψ2 packaging cell line (39) and were used to infect Rat-1a cells. The number of foci of transformed cells was compared with the number of G418-resistant colonies. The transforming efficiencies of the mutant *myc* cDNAs are expressed as a percentage of the activity of normal human c-*myc* cDNA. Viruses bearing normal c-*myc* induced equivalent numbers of foci and G418-resistant colonies; retroviruses without c-*myc* or one with a termination codon introduced at position 100 in c-*myc* had no detectable transforming activity.

The Rat-1a-transforming activity of each mutant *myc* cDNA tested in this assay is shown in Fig. 3. The inactivity of mutants with lesions in region IV is similar to their inactivity in the REC cotransformation assay. However, several mutants behaved in a different manner in the two assays: In105 was impaired (14% of normal activity) in REC cotransformation but was quite active (50% of normal activity) in Rat-1a transformation; conversely, deletion mutants involving region III, which are significantly active in REC cotransformation, were less active or totally inactive in Rat-1a transformation. Thus, the regions of c-*myc* that are important for cotransformation of RECs are not the same as those that are important for transformation of the established cells. Curiously, D265-353 reproducibly showed 10% activity in the Rat-1a assay, while smaller deletion mutants covering the same region (D265-317, D265-324, and D265-341) had no detectable activity. Thus, the effects of mutations on protein function are, to a certain extent, unpredictable, such that a small deletion may have a more deleterious effect than a larger deletion affecting the same sequences.

**Analysis of mutant c-*myc* gene products.** The ability of c-*myc* mutants to cotransform RECs or to transform Rat-1a cells may reflect the transforming capability of their protein products, but it could also be due to differences in the amount of mutant proteins in the cells. Therefore, we measured the amount of *myc* RNA and protein made by the mutant genes, employing a RNA protection assay (13, 23) and Northern blotting (11) to measure the level of mutant c-*myc* RNA and immunoprecipitation to measure the level of mutant c-*myc* protein (45).

Transcripts of mutant c-*myc* genes that were active in the REC cotransformation assay were studied in individual clones of transformed RECs (Fig. 4a), and those of inactive genes were studied in individual clones of Rat-1 cells transfected with the mutants (Fig. 4b). Because transcripts from the transfected and endogenous *myc* genes are generally similar in size (2.3 kbp), we used a protection assay (13,

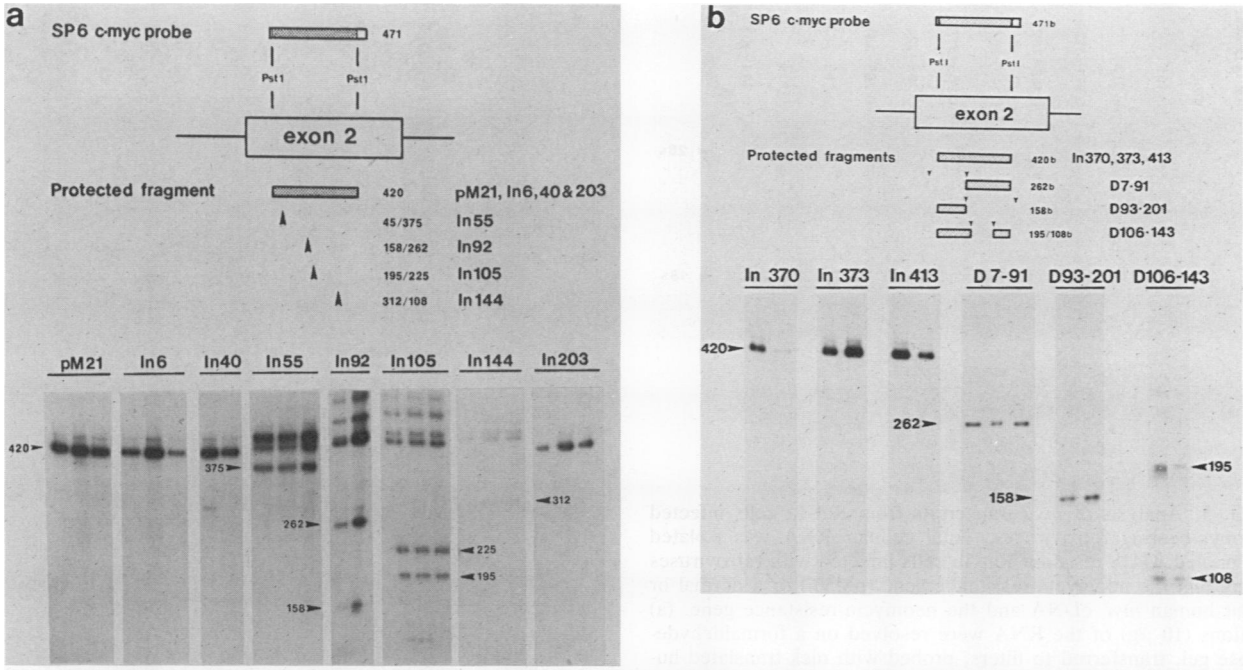


FIG. 4. Protection assay of human c-*myc* transcripts from transfected REC or Rat-1 cells. SP6-generated radiolabeled RNA probe, complementary to the *Pst*I-*Pst*I (exon 2) portion of human c-*myc* mRNA, was hybridized to total cellular RNA isolated from two or three individual clones of REC, cotransformed by EJ*ras* and various *myc* genes (a) or Rat-1 cells transfected with mutant *myc* genes inactive in the cotransformation assay (b). After digestion of RNase A and T1, the protected fragments were resolved on 7 M urea–6% polyacrylamide gels and autoradiographed for 24 h. Lengths of probe and protected fragments are indicated.

23) that only detects transcripts from the transfected human genes. The examples shown in Fig. 4a are from RECs cotransformed by mutants with linker insertions in exon 2. In the case of certain mutants, protected fragments with characteristic sizes were generated, verifying that the transcripts are from the transfected mutant genes. The amount of human *myc* RNA found among these cells can vary up to fivefold, but all these levels are adequate to cotransform RECs. For instance, each of the three REC lines cotransformed by In144 had significantly less human *myc* RNA than the three lines cotransformed by pM21. Curiously, transcripts from insertion mutants that protect novel-sized fragments also generate bands larger than that protected by the wild-type human c-*myc* transcript (420 bases); the origin of these larger fragments is unclear. Expression of the cotransformationally inactive mutants was studied in transfected Rat-1 cells (Fig. 4b). By employing the same probe and assay procedure used for RECs (the band intensities in Figure 4a and 4b are directly comparable), we found that a similar range of *myc* RNA levels was produced by these inactive mutants.

We measured the level of *myc* transcripts produced by the integrated proviruses in the infected Rat-1a cells by using Northern analysis (Fig. 5a). The proviral-*myc* transcripts were considerably larger than transcripts from the endogenous gene and were easily distinguished. We found less than twofold variation in proviral-*myc* transcript levels among pooled Rat-1a cells that were infected by any of the 12 *myc*-bearing retroviruses, with the possible exception of those bearing D145-304, D145-262, or D265-317. In the case of transcripts from proviruses bearing certain of the *myc* mutants (e.g., D145-304 and D145-262), their size caused them to comigrate with 28S rRNA, which may interfere with hybridization with *myc* probe. Therefore, we performed dot blot analysis on the same RNA samples (Fig. 5b) and found no significant variation in the amount of *myc* transcripts. Inspection of Fig. 5a further reveals that proviral-*myc* expression does not downregulate the level of c-*myc* transcripts from the endogenous rat c-*myc* gene.

Proteins from cotransformationally active mutant c-*myc* genes were studied in transformed RECs, in which we found that levels of mutant proteins are comparable to that of normal human c-*myc* protein (data not shown). We examined the proteins that were produced by the normal and cotransformationally inactive mutant human c-*myc* genes in transfected Rat-1 cells but found that levels of *myc* proteins were too low for convenient analysis (data not shown). To obtain higher levels of expression, we transferred these genes into a vector containing the SV40 origin of replication and transfected COS-7 cells. Immunoprecipitation of radiolabeled *myc* proteins demonstrated that the transfected normal and inactive mutant c-*myc* genes produce comparable steady-state levels of protein (Fig. 6a). Similarly, the level of protein produced by the normal and 11 mutant *myc* genes introduced into Rat-1a cells by retroviral infection was similar, with the exception of In373, which produced significantly less *myc* protein (Fig. 6b). These results suggest that the ability of the mutant c-*myc* genes to cotransform RECs or to transform Rat-1a cells is due to the activity of their protein products in these assays and not to major differences in the level of protein produced.

The *myc* proteins in cells transfected with insertion mutants had gel mobilities similar to that of the normal protein ($M_r$s of 62,000 to 67,000), whereas the proteins produced by deletion mutants had lower $M_r$s (Fig. 6). However, the $M_r$s of most of the deletion mutants remained significantly greater than their deduced molecular weights (like that of wild-type c-*myc*). The two exceptions in Fig. 6 are D145-262 and D145-304, which migrated in accordance with their calculated size; this suggests that sequences between resi-

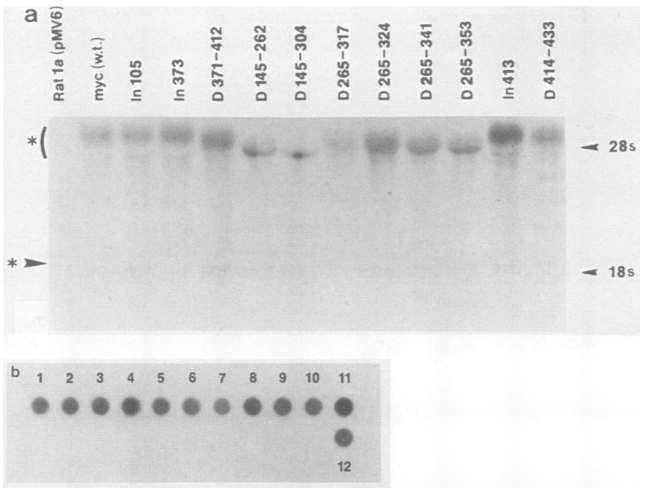


FIG. 5. Analysis of *myc* transcripts from Rat-1a cells infected with *myc*-bearing retroviruses. Total cellular RNA was isolated from pooled, G418-resistant Rat-1a cells infected with retroviruses bearing just the neomycin-resistance gene (pMV6) or a normal or mutant human *myc* cDNA and the neomycin-resistance gene. (a) Fractions (10 μg) of the RNA were resolved on a formaldehyde-agarose gel, transferred to filters, probed with nick-translated human c-*myc* cDNA probe (exons 2 and 3), and autoradiographed. The *myc* allele carried by the retroviruses that were used to infect the cells are indicated above the lanes, and the positions of the 18S and 28S rRNA bands are indicated (w.t., wild type). Proviral-*myc* transcripts migrated near the 28S band, and the endogenous rat *myc* transcripts migrated slightly above the 18S band (*). (b) Fraction (2 μg) of the RNA were dot blotted onto filters, probed as described above and autoradiographed. RNA were from Rat-1a cells infected with retroviruses bearing the following: wild-type c-*myc* (lane 1), In105 (lane 2), In373 (lane 3), D371-412 (lane 4), D145-262 (lane 5), D145-304 (lane 6), D265-317 (lane 7), D265-324 (lane 8) D265-341 (lane 9), D265-353 (lane 10), In413 (lane 11), D414-433 (lane 12).

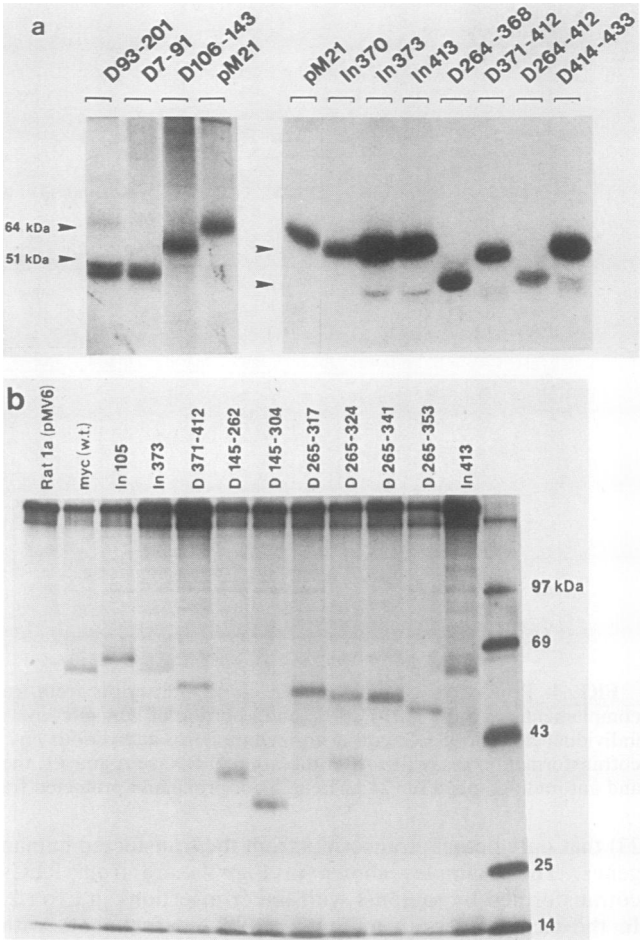


FIG. 6. Immunoprecipitation of proteins encoded by mutant c-*myc* genes. (a) COS-7 cells were transfected with plasmids containing a SV40 origin of replication and an inactive mutant or normal human c-*myc* gene and labeled with $^{32}P_i$. Proteins were immunoprecipitated with rabbit antiserum raised to the C-terminal 32-amino-acid peptide of human c-*myc*, the complexes were precipitated with inactivated *Staphylococcus aureus*, and the proteins were resolved by sodium dodecyl sulfate-10% polyacrylamide gel electrophoresis. Each lane demonstrates the protein produced by the transfected c-*myc* gene indicated above the lane; the normal human c-*myc* protein, produced by transfection of pM21, is provided as control. The two sets of autoradiographs represent separate experiments. Molecular size markers (in kilodaltons [kDa]) are indicated by arrows. (b) Rat-1a cells infected with *myc*-bearing retroviruses were obtained as described in the legend to Fig. 5. Cells were labeled with [$^{35}$S]methionine, and *myc* protein was precipitated with rabbit antiserum raised against a c-*myc* exon 2 peptide (a gift from G. Evan) (w.t., wild type). Immune complex precipitation and electrophoresis were carried out as described above for panel a. In both panels a and b, some lanes had a minor second band that migrated faster than the major *myc* band and probably resulted from degradation.

dues 145 and 262 may be responsible for retarding the mobility of *myc* protein in sodium dodecyl sulfate-polyacrylamide gels. Finally, we did not observe *myc* doublets (27, 45) in the infected Rat-1a cells (Fig. 6b; we did not expect to be able to see doublets in Fig. 6a, because the proteins were labeled with $^{32}P$, which gives broad, less well defined bands). In other experiments, we observed the characteristic *myc* doublets in RECs that were cotransformed by *myc* and EJ*ras* (data not shown).

**Subcellular localization of mutant c-*myc* proteins.** Normal c-*myc* protein accumulates in the nucleus (18, 20, 43), and we used the mutants to search for the regions that are responsible for this property. We determined the subcellular distribution of the mutant c-*myc* proteins within transfected COS-7 cells both by immunofluorescent staining and by nuclear-cytoplasmic fractionation of radiolabeled cells followed by immunoprecipitation. The protein product of all mutants with alterations within exon 2 showed immunofluorescence only in nuclei (Fig. 7). Similarly, most of the mutants of exon 3 appeared to be nuclear, with the notable exceptions of D264-368 and D264-412, which produced proteins that were both nuclear and cytoplasmic. By using mutants with progressively larger deletions starting from the *Cla*I site at the 5′ end of exon 3 (codons 262 and 263), partial cytoplasmic localization occurred only when sequences 3′ to codon 320 were deleted. None of the tested mutants produced a protein that was predominantly cytoplasmic in this assay.

We fractionated cells to confirm the results obtained by immunofluorescence (Fig. 8). Indeed, following nuclear-cytoplasmic fractionation the normal c-*myc* protein was predominantly (80 to 95%) in the nuclear fraction, and the D264-368 protein was in both the nuclear and cytoplasmic fractions (varying between 30 and 70% in either fraction; Fig. 8a). Most of the other mutants tested (Fig. 8b) showed good correlation between the fractionation and immunofluorescence results, and D414-433 was predominantly in the

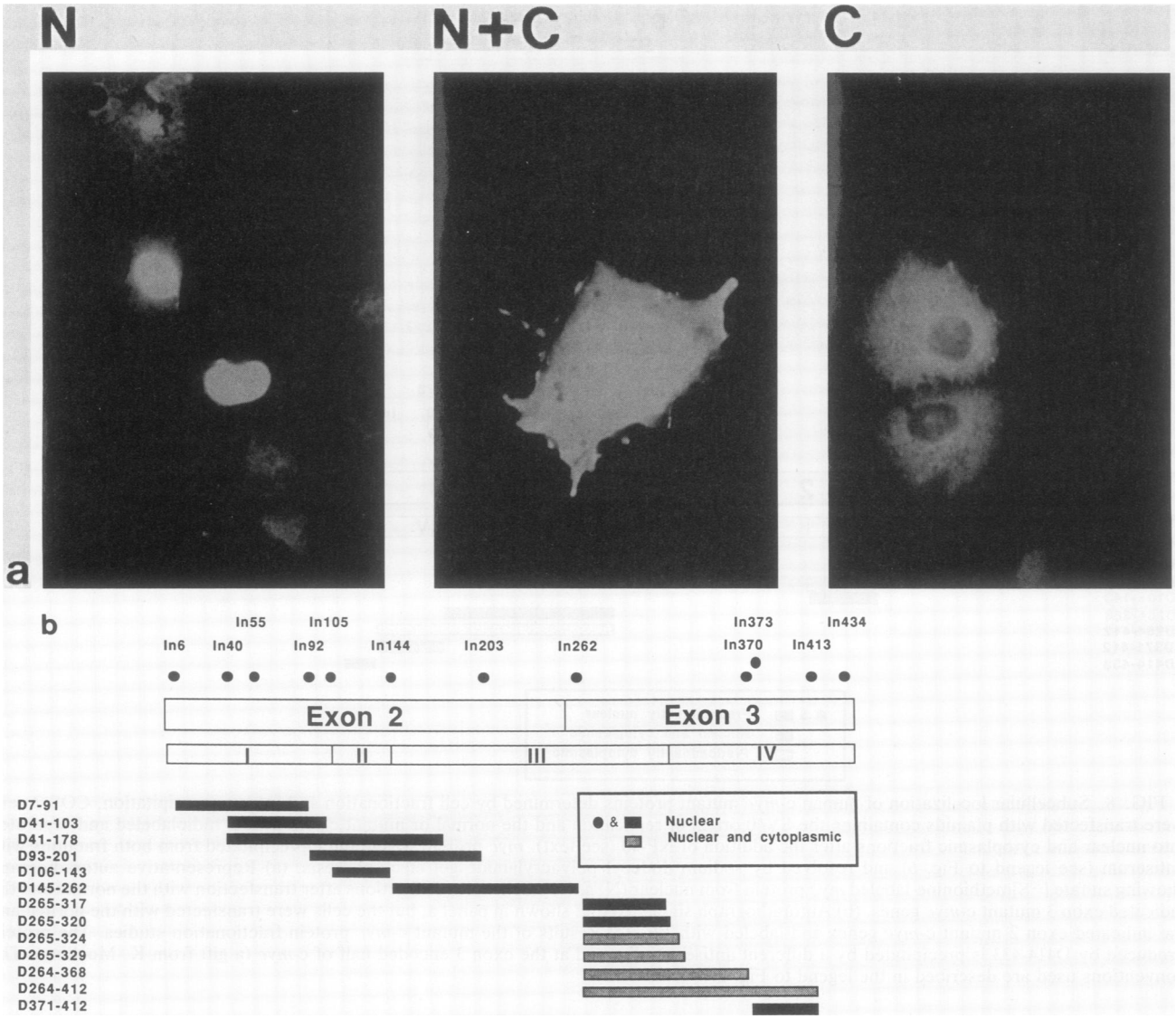


FIG. 7. Subcellular localization of human c-*myc* mutant proteins determined by immunofluorescence. COS-7 cells were transfected with plasmids containing a SV40 origin or replication and the normal or mutant c-*myc* genes and examined after 24 to 48 h for the location of the protein products. (a) Representative results for different mutant c-*myc* proteins and for the chicken muscle pyruvate kinase protein. Most c-*myc* mutant proteins (D265-317 in the cells in this photograph) give only nuclear (N) fluorescence, but some (D264-368 in the cell in this photograph) give both nuclear and cytoplasmic (N+C) fluorescence. In comparison, staining for the chicken muscle pyruvate kinase protein gives only cytoplasmic (C) fluorescence, a pattern that we have not seen with any of our *myc* mutants. (b) Results of mutant c-*myc* protein localization studies. The conventions used are described in the legend to Fig. 2.

nuclear fraction (Fig. 8c). Surprisingly, proteins encoded by D371-412, 106-143, and 93-201 were found in both subcellular fractions (similar to the protein produced by D264-368), which is in contrast to their nuclear location shown by immunofluorescence (Fig. 7b). This indicates that although these proteins accumulate in the nucleus of intact cells, they do not possess some other property that allows the normal c-*myc* protein to be retained in the nucleus during fractionation. Furthermore, the protein produced by D264-412, which was nuclear and cytoplasmic by immunofluorescence and indistinguishable from the D264-368 protein by this technique, was consistently more cytoplasmic than either the D264-368 or the D371-412 protein by fractionation (Fig. 8a). This indicates that nuclear localization functions inactivated by the deletion of amino acids 264 to 368 and 371 to 412, in addition to being different, may also have complementary properties.

## DISCUSSION

Linker insertion mutagenesis (55) proved to be a useful way to begin the examination of functional regions within the human c-*myc* gene. It introduced mildly disruptive mutations and unique restriction sites at scattered points throughout the gene, allowing subsequent manipulations of the gene to be performed with relative ease. By using mutants generated by this method, we mapped regions of the human c-*myc* protein that are important for transformation and nuclear localization.

**Partial cotransforming activity of certain mutants.** Some of the mutant c-*myc* genes cotransformed RECs with decreased

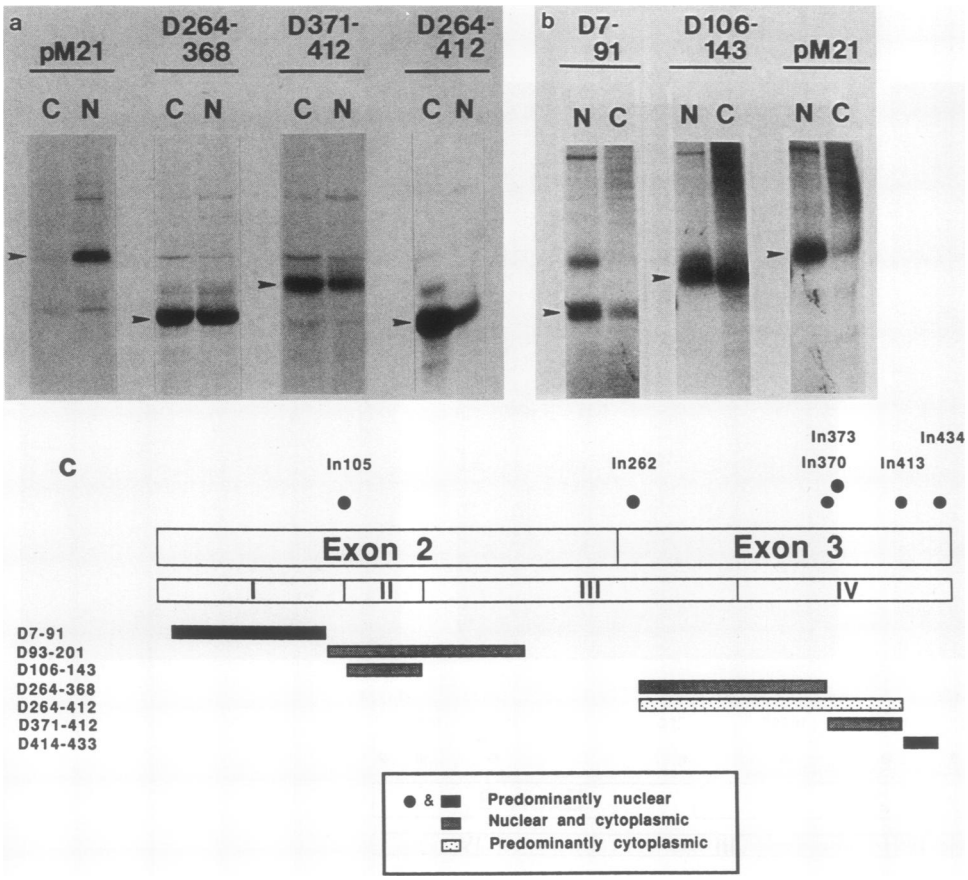


FIG. 8. Subcellular localization of human c-*myc* mutant proteins determined by cell fractionation and immunoprecipitation. COS-7 cells were transfected with plamids containing the SV40 origin of replication and the normal or mutant c-*myc* genes, radiolabeled and separated into nuclear and cytoplasmic fractions after the addition of NP-40 (see text). *myc* protein was immunoprecipitated from both fractions with antiserum (see legend to Fig. 6) and resolved by sodium dodecyl polyacrylamide gel electophoresis. (a) Representative autoradiograph showing sulfate [$^{35}$S]methionine-labeled *myc* protein from nuclear (N) and cytoplasmic (C) fractions after transfection with the normal and the indicated exon 3 mutant c-*myc* genes. (b) Autoradiograph similar to that shown in panel a, but the cells were transfected with the normal and the indicated exon 2 mutant c-*myc* genes and labeled with $^{32}P_i$. (c) Results of the mutant c-*myc* protein fractionation studies. The protein produced by D414-433 is precipitated by a different antiserum directed at the exon 3 encoded half of c-*myc* (a gift from K. Moelling). The conventions used are described in the legend to Fig. 2.

efficiency. The cells transformed by these genes and EJ*ras* were indistinguishable by several different criteria (e.g., morphology, ability to form foci, and tumorigenicity), so that transformation by these mutants was not incomplete. We have shown the c-*myc* expression in the RECs transformed by such mutants is comparable to that in RECs transformed by wild-type c-*myc*. In addition, we found similar levels of EJ*ras* RNA in these cells (data not shown). Thus, neither a requirement for higher *myc* nor EJ*ras* expression can explain the decreased activity of certain mutants. RECs form a heterogeneous population of cells and may represent a heterogeneous target for *myc* cotransformation with EJ*ras*, so that partially defective c-*myc* mutants may only transform a subset of the REC population that is susceptible to transformation by wild-type c-*myc*. Indeed, the inefficiency shown by In105 in the REC assay is not apparent when activity is measured with the Rat-1a cell line. This situation is reminiscent of the ability of partially transformation-defective mutants of MC29, which have deleted portions of the v-*myc* gene, to transform chick fibroblasts but not macrophages (7, 19, 30, 46).

**Regions of human c-*myc* protein required for activity.** We identified three areas (regions I, II and IV) that are essential for REC cotransforming activity. Both insertion and deletion mutations in regions II and IV adversely affect or abolish activity, indicating that they probably serve functions that are essential for cotransformation. Region I is important, but only larger deletions severely affect activity, and insertion or smaller deletion mutations have lesser or no effects. Explanations for these results include the possibility that region I plays a secondary role in c-*myc* activity (e.g., in protein folding, in optimizing the function of essential domains, or both). Alternatively, the substantial activity of the insertion and smaller deletion mutants in region I may result from our inability to ascribe significance to smaller decrements in activity; the variability inherent in the cotransformation assay dictates that quite large differences in activity must be seen before significance is attached. Finally, region III in the middle third of the protein appears to be largely dispensable for cotransformation of RECs.

These findings are consistent with results of other published studies on domains of *myc* that are important for activity. Heaney et al. (30) recently reported that regions of Mc29 *gag-myc* corresponding to an area covered by region III are dispensable for chick fibroblast transforming activity. Interestingly, some of these sequences are required for

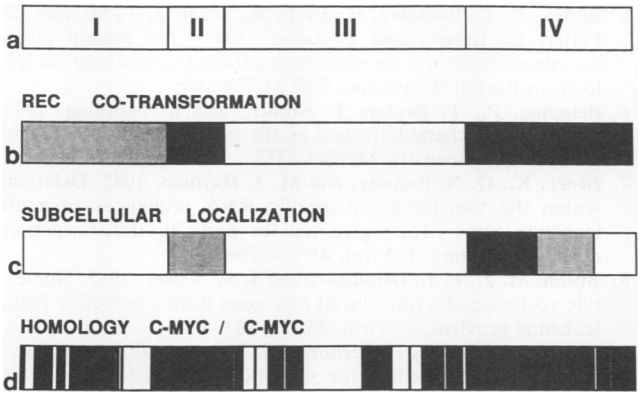


FIG. 9. Functional regions of the human c-*myc* protein and regions of conserved amino acids among c-*myc* proteins of chickens, mice, and humans. (a) Regions I through IV are defined in the text and in Fig. 2. (b) The darker shading indicates regions essential for activity; lighter shading indicates the region that only tolerates insertions and partial deletions; the clear area indicates the region that is not essential for activity (see text for details). (c) The shaded areas indicate regions important for normal subcellular distribution, with the darker shading representing the area determined by immunofluorescence studies and the lighter shading representing additional areas determined by fractionation and immunoprecipitation studies. (d) Blackened areas indicate regions of amino acid identity among normal c-*myc* proteins from the three species (8).

transformation of macrophages (see below). Sarid et al. (48) found that small deletions in region I allow residual activity in the REC cotransformation assay, albeit at a lower level than we obtained, whereas deletion of almost the entirety of region I abolishes activity; deletion of the 3′ half of region II abolishes activity; and deletion of a part of region III has the least effect on activity.

We caution that the boundaries of the functionally important regions mapped by us are defined by the insertion and deletion mutants tested and cannot be considered exact. For instance, region II, which we found to be critical for cotransformation of RECs, is defined by In105 and the endpoints of D106-143. However, the findings of Sarid et al. (48) (see above) and Murphy et al. (40) indicate that not all residues between amino acids 106 and 143 may be critical. Finer mutational analysis is required to more precisely delineate the boundaries of important regions.

It is premature to assume that the regions of c-*myc* that are sensitive to mutational inactivation of activity are directly responsible for essential c-*myc* functions. The lack of insight into the *myc* protein beyond its primary structure makes it impossible to predict the effects of mutations on protein structure. Therefore, inactivating mutations may not occur at active sites, but may adversely affect function through unfavorable effects on conformation.

**Activity of some mutant c-*myc* genes depends on the target cell.** Rat-1a cell transformation provides an alternative assay for *myc*-transforming activity. It differs from the REC cotransformation assay, because we used retroviral vectors to deliver c-*myc* cDNA into an established cell line in the absence of a second oncogene. Our results indicate that mutation in one area (In105) severely diminishes REC cotransforming activity, but only mildly diminishes Rat-1a transforming activity; the opposite is true of several deletions in region III. The presence of similar amounts of *myc* RNA and protein in Rat-1a cells infected with mutant and normal cDNAs indicates that these cells are resistant to morphological transformation by some of the mutants that are capable of cotransforming RECs. A parallel may be found in the study of Heaney et al. (30), who found that deletion mutants involving the acidic region of v-*myc* (in region III near the *Cla*I site) transform chick fibroblasts but not bone marrow cells. We hypothesize that the c-*myc* protein participates in more than one interaction, reaction, or both in the nucleus and that some of these events may be required for transformation of Rat-1a cells but not RECs and vice versa. Consequently, a region of the c-*myc* protein that is involved in a particular interaction may be necessary only in certain assays (7, 19, 30, 46). It follows that our mapping of functional regions by the use of assays that measure c-*myc* neoplastic activity may not accurately reflect regions of the protein that are important for normal physiologic activity.

**Relationship of conserved c-*myc* domains to activity.** Comparison of the amino acid sequence of c-*myc* from humans, mice, and chickens (53, 57) reveal regions of conservation and regions of variability (Fig. 9). Sequences in regions II and IV of human c-*myc*, which we found to be most critical for REC cotransformation, are the most highly conserved. Region I, which tolerates insertion and small deletion mutations, has blocks of highly conserved residues, but it also contains stretches of nonidentity. Region III, which we found to be insensitive to insertion or deletion inactivation, had the greatest variability; yet, even this region contained blocks of conserved sequences. Thus, despite the correlation between regions that are important for REC cotransformation and regions of interspecies sequence conservation, the dispensibility of large portions of c-*myc* may indicate that this assay does not require all of the functional capabilities of c-*myc*. The requirement for much of region III in the Rat-1a transformation assay supports this view.

**Nuclear accumulation of mutant c-*myc* proteins.** Many nuclear proteins, such as the SV40 T antigen (32) and the polyomavirus large T antigen (47), have been found to have amino acid sequences that confer nuclear localization. The SV40 T antigen has one signal that has been proven to be necessary for proper nuclear localization (32), although another region may contribute to this process (61). Polyomavirus large T antigen appears to have two signals, both of which contribute to nuclear localization (47).

By using immunofluorescent staining, we identified one region (between residues 320 and 368) that when removed resulted in a protein that was both nuclear and cytoplasmic in intact cells. This region may contain a nuclear localization signal or may be involved in the proper presentation of such a signal located elsewhere, e.g., in the C-terminal six amino acids, the contribution of which to subcellular distribution has not been determined yet. Additional tests of the karyophilic character of the region between residues 320 and 368, such as the demonstration of its ability to redirect a cytoplasmic protein to the nucleus, will be necessary to clarify this issue.

Deletion of two other regions of human c-*myc* results in poor nuclear retention of the protein during separation of nuclear and cytoplasmic fractions after lysis of cells with nonionic detergent. Deletion of either of these two regions (in D106-143 and D371-412) does not affect localization, as determined by immunofluorescent staining of cells; and therefore, they are not needed to facilitate entry of the protein into nuclei. However, mutant proteins lacking these sequences redistribute during fractionation, leading to their appearance in the cytoplasmic fraction. Since NP-40 treatment disrupts the nuclear membrane (25), we speculate that these mutants fail to interact or associate normally with some intranuclear component(s), are consequently less av-

idly retained within nuclei, and can more readily leave the nuclear fraction during isolation. It is probably significant that both of these regions of c-*myc* are not only highly conserved among c-*myc* genes of different species but also between N-*myc* and c-*myc* (35, 52).

**Correlation of functionally important regions with properties of the protein.** Abnormal nuclear localization may account for the loss or diminution of activity of some of our mutants, but it cannot account for the inactivity of most of the mutants. For instance, the three inactive insertion mutants in region IV and the large deletion mutants in region I all behaved normally in our localization assays, so that their disruptive effect on c-*myc* activity resulted from some as yet undefined disturbance(s) of function. Nucleophilic properties can be assigned to three functionally important regions defined by the REC cotransformation assay. Two of these regions (amino acids 106 to 143 and 371 to 412) appear to be important for proper c-*myc* retention within nuclei, and neither can be deleted without destroying activity. Amino acids between 320 and 368 may function to direct the protein to nuclei of intact cells. Curiously, loss of this property, which results in a nuclear and cytoplasmic distribution of the mutant protein, only diminishes but does not abolish activity. We assume that the residual cotransforming activity of some of these mutants can be attributed to the fraction that finds its way to the nucleus, and once there it is capable of performing some or all of its functions.

## ACKNOWLEDGMENTS

J.S. thanks L. Siminovitch and his laboratory at the Hospital for Sick Children, Toronto, Canada, for support. We thank the following for gifts of reagents: Tom Atkinson (oligonucleotides), Dan Cashman (oligonucleotides), I. Bernard Weinstein (pMV6), Carlo Croce (human c-*myc* cDNA), Bruce Roberts (SV40-chicken muscle pyruvate kinase plasmid and antiserum to the protein), Karin Moelling (antiserum to c-*myc*), Gerard Evan (antiserum to c-*myc*), Richard Mulligan (ψ2 cells), Nissim Hay (pSVSP65), and Arthur Bruskin (Rat-1a cells). We gratefully acknowledge Cammie Lesser and John Chant, who contributed to parts of this project during student rotation projects, and Janine Marinos, who helped to prepare the manuscript. Most of all, we are indebted to Chin Lin for expert technical assistance.

This study was supported by Public Health Service grants CA38783 (to W.L.), CA39832 (to H.V.), CA12705 (to J.M.B.), and SO7 RR05355 (to J.M.B.) from the National Institutes of Health and by funds from the G. W. Hooper Research Foundation. T.dL. is a Constantijn and Christiaan Huygens Fellow supported by The Netherlands Organization for the Advancement of Pure Research, G.R. is a fellow of the Weigart Foundation, H.V. is an American Cancer Society Research Professor, and W.L. is a recipient of the Junior Faculty Research Award from the American Cancer Society.

## LITERATURE CITED

1. **Adams, J. M., S. Gerondakis, E. Webb, L. M. Corcoran, and S. Cory.** 1983. Cellular *myc* oncogene is altered by chromosome translocation in murine plasmacytomas and is rearranged similarly in human Burkitt's lymphomas. Proc. Natl. Acad. Sci. USA **80**:1982–1986.
2. **Alitalo, K., G. Ramsay, J. M. Bishop, S. O. Pfeifer, W. Colby, and A. D. Levinson.** 1983. Identification of nuclear proteins encoded by viral and cellular *myc* oncogenes. Nature (London) **306**:274–277.
3. **Armelin, H. A., M. C. S. Armelin, K. Kelly, T. Stewart, P. Leder, B. H. Cochran, and C. D. Stiles.** 1984. Functional role for c-*myc* in mitogenic response to platelet-derived growth factor. Nature (London) **310**:655–660.
4. **Auffray, C., and F. Rougeon.** 1980. Purification of mouse immunoglobulin heavy-chain messenger RNAs from total myeloma tumor RNA. Eur. J. Biochem. **107**:303–314.
5. **Battey, J., C. Boulding, R. Taub, W. Murphy, T. Stewart, H. Potter, G. Lenoir, and P. Leder.** 1983. The human c-*myc* oncogene: structural consequences of translocation into the IgH locus in Burkitt lymphoma. Cell **34**:779–787.
6. **Beimling, P., T. Benter, T. Sander, and K. Moelling.** 1985. Isolation and characterization of the human cellular *myc* gene product. Biochemistry **24**:6349–6355.
7. **Bister, K., G. N. Ramsay, and M. J. Hayman.** 1982. Deletions within the transformation-specific RNA sequences of acute leukemia virus MC29 give rise to partially transformation-defective mutants. J. Virol. **41**:754–766.
8. **Braun, M. J., P. L. Deininger, and J. W. Casey.** 1985. Nucleotide sequence of a transduced *myc* gene from a defective feline leukemia provirus. J. Virol. **55**:177–183.
9. **Chen, E. Y., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**:165–170.
10. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
11. **Church, G. M., and W. Gilbert.** 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA **81**:1991–1995.
12. **Corcoran, L. M., J. M. Adams, A. R. Dunn, and S. Cory.** 1984. Murine T lymphomas in which the cellular *myc* oncogene has been activated by retroviral insertion. Cell **37**:113–122.
13. **Coughlin, S. R., W. M. F. Lee, P. W. Williams, G. M. Giels, and L. T. Williams.** 1985. c-*myc* gene expression is stimulated by agents that activate protein kinase C and does not account for the mitogenic effect of PDGF. Cell **43**:243–251.
14. **Crews, S., R. Barth, L. Hood, J. Prehn, and K. Calame.** 1982. Mouse c-*myc* oncogene is located on chromosome 15 and translocated to chromosome 12 in plasmacytomas. Science **218**:1319–1321.
15. **Dalla-Favera, R., M. Bregni, J. Erikson, D. Patterson, R. C. Gallo, and C. M. Croce.** 1982. Human c-*myc* oncogene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc. Natl. Acad. Sci. USA **79**:7824–7827.
16. **Dean, M., R. A. Levine, and J. Campisi.** 1986. c-*myc* regulation during retinoic acid-induced differentiation of F9 cells is post-transcriptional and associated with growth arrest. Mol. Cell. Biol. **6**:518–524.
17. **Einat, M., D. Resnitzky, and A. Kimchi.** 1985. Close link between reduction of c-*myc* expression by interferon and $G_0/G_1$ arrest. Nature (London) **313**:597–600.
18. **Eisenman, R. N., C. Y. Tachibana, H. D. Abrams, and S. R. Hann.** 1985. v-*myc* and c-*myc* encoded proteins are associated with the nuclear matrix. Mol. Cell. Biol. **5**:114–126.
19. **Enrietto, P. J., and M. J. Hayman.** 1982. Restriction enzyme analysis of partially transformation-defective mutants of acute leukemia virus MC29. J. Virol. **44**:711–715.
20. **Evan, G. I., and D. C. Hancock.** 1985. Studies on the interaction of the human c-*myc* protein with cell nuclei: $p62^{c\text{-}myc}$ as a member of a discrete subset of nuclear proteins. Cell **43**:253–261.
21. **Gluzman, Y.** 1981. SV40-transformed simian cells support the replication of early SV40 mutants. Cell **23**:175–182.
22. **Graham, F. L., and A. van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**:456–467.
23. **Green, M. R., T. Maniatis, and D. A. Melton.** 1983. Human β globin pre-mRNA synthesized *in vitro* is accurately spliced in Xenopus oocyte nuclei. Cell **32**:681–694.
24. **Hamlyn, P. H., and T. H. Rabbitts.** 1983. Translocation joins the c-*myc* and the immunoglobulin 1 genes in a Burkitt lymphoma revealing a third exon in the c-*myc* oncogene. Nature (London) **304**:135–139.
25. **Hancock, R., and T. Boulikas.** 1982. Functional organisation in the nucleus. Int. Rev. Cytol. **79**:165–214.
26. **Hann, S. R., H. D. Abrams, L. R. Rohrschneider, and R. N. Eisenman.** 1983. Proteins encoded by the v-*myc* and c-*myc* oncogenes: identification and localization in acute leukemia virus transformants and bursal lymphoma cell line. Cell

**34**:789–798.
27. **Hann, S. R., and R. N. Eisenman.** 1984. Proteins encoded by the human c-*myc* oncogene: differential expression in neoplastic cells. Mol. Cell. Biol. **4**:2486–2497.
28. **Hann, S. R., C. B. Thompson, and R. N. Eisenman.** 1985. c-*myc* oncogene protein synthesis is independent of the cell cycle in human and avian cells. Nature (London) **314**:366–369.
29. **Hayward, W. S., B. G. Neel, and S. M. Astrin.** 1981. Activation of a cellular *onc* gene by promoter insertion in ALV-induced lymphoid leukosis. Nature (London) **209**:475–479.
30. **Heaney, M. L., J. Pierce, and J. T. Parsons.** 1986. Site-directed mutagenesis of the *gag-myc* gene of avian myelocytomatosis virus 29: biological activity and intracellular localization of structurally altered proteins. J. Virol. **60**:167–176.
31. **Kaczmarek, L., J. K. Hyland, R. Watt, M. Rosenberg, and R. Baserga.** 1985. Microinjected c-*myc* as a competence factor. Science **228**:1313–1315.
32. **Kalderon, D., B. L. Roberts, W. D. Richardson, and A. E. Smith.** 1984. A short amino acid sequence able to specify nuclear location. Cell **39**:499–509.
33. **Keath, E. J., P. G. Caimi, and M. D. Cole.** 1984. Fibroblast lines expressing activated c-*myc* oncogenes are tumorigenic in nude mice and syngeneic animals. Cell **39**:339–348.
34. **Klempnauer, K.-H., G. Symonds, G. Evans, and J. M. Bishop.** 1984. Subcellular location of proteins encoded by viral and cellular *myb* genes. Cell **37**:537–547.
35. **Kohl, N. E., R. A. De Pinho, P. D. Nisen, R. K. Smith, C. E. Gee, and F. W. Alt.** 1986. Human N-*myc* is closely related in organization and nucleotide sequence to c-*myc*. Nature (London) **319**:73–77.
36. **Land, H., L. F. Parada, and R. A. Weinberg.** 1983. Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. Nature (London) **304**:596–601.
37. **Lee, W. M. F., M. Schwab, D. Westaway, and H. Varmus.** 1985. Augmented expression of normal c-*myc* is sufficient for cotransformation of rat embryo cells with a mutant *ras* gene. Mol. Cell. Biol. **5**:3345–3356.
38. **Little, C. D., M. M. Nau, D. N. Carney, A. F. Gazdar, and J. D. Minna.** 1983. Amplification and expression of the c-*myc* oncogene in human lung cancer cell lines. Nature (London) **306**:194–196.
39. **Mann, R., R. C. Mulligan, and D. Baltimore.** 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell **33**:153–159.
40. **Murphy, W., J. Sarid, T. Vasicek, J. Battey, G. Lenoir, and P. Leder.** 1986. A translocated human c-*myc* oncogene is altered in a conserved coding sequence. Proc. Natl. Acad. Sci. USA **83**:2939–2943.
41. **Parada, L. F., C. J. Tabin, C. Shih, and R. A . Weinberg.** 1982. Human EJ bladder carcinoma oncogene is a homologue of Harvey sarcoma virus *ras* gene. Nature (London) **295**:474–478.
42. **Payne, G. S., J. M. Bishop, and H. E. Varmus.** 1982. Multiple arrangements of viral DNA and an activated host oncogene (c-*myc*) in bursal lymphoma. Nature (London) **295**:209–217.
43. **Persson, H., and P. Leder.** 1984. Nuclear localization and DNA binding properties of a protein expressed by human c-*myc* oncogene. Science **225**:718–721.
44. **Rabbitts, T. H., A. Forster, P. Hamlyn, and R. Baer.** 1984. Effect of somatic mutation within translocated c-*myc* genes in Burkitt's lymphoma. Nature (London) **309**:592–597.
45. **Ramsay, G., G. I. Evan, and J. M Bishop.** 1984. The protein encoded by the human proto-oncogene c-*myc*. Proc. Natl. Acad. Sci. USA **81**:7742–7746.
46. **Ramsay, G., and M. J. Hayman.** 1982. Isolation and biochemical characterization of partially transformation-defective mutants of avian myelocytomatosis virus strain MC29: localization of the mutation to the *myc* domain of the 110,000-dalton *gag-myc* polyprotein. J. Virol. **41**:745–753.
47. **Richardson, W. D., B. L. Roberts, and A. E. Smith.** 1986. Nuclear location signals in polyoma virus large-T. Cell **44**:77–85.
48. **Sarid, J., T. D. Halazonetis, W. Murphy, and P. Leder.** 1987. Evolutionarily conserved regions of the human c-*myc* protein can be uncoupled from transforming activity. Proc. Natl. Acad. Sci. USA **84**:170–173.
49. **Shen-Ong, G. L., E. J. Keath, S. P. Piccoli, and M. D. Cole.** 1982. Novel *myc* oncogene RNA from abortive immunoglobulin-gene recombination in mouse plasmacytomas. Cell **32**:443–452.
50. **Shih, C. K., M. Linial, M. M. Goodenow, and W. S. Hayward.** 1984. Nucleotide sequence 5′ of the chicken c-*myc* coding region: localization of a noncoding exon that is absent from *myc* transcripts in most avian leukosis virus-induced lymphomas. Proc. Natl. Acad. Sci. USA **81**:4697–4701.
51. **Southern, P. J., and P. Berg.** 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. J. Mol. Appl. Genet. **1**:327–341.
52. **Stanton, L. W., M. Schwab, and J. M. Bishop.** 1986. Nucleotide sequence of the human N-*myc* gene. Proc. Natl. Acad. Sci. USA **83**:1772–1776.
53. **Stanton, L. W., R. Watt, and K. B. Marcu.** 1983. Translocation breakage and truncated transcripts of c-*myc* oncogene in murine plasmacytomas. Nature (London) **303**:401–405.
54. **Steffen, D.** 1984. Proviruses are adjacent to c-*myc* in some murine leukemia virus-induced lymphomas. Proc. Natl. Acad. Sci. USA **81**:2097–2101.
55. **Stone, J., T. Atkinson, M. Smith, and T. Pawson.** 1984. Identification of functional regions in the transforming protein of Fujinami sarcoma virus by in-phase insertion mutagenesis. Cell **37**:549–558.
56. **Taub. R., I. Kirsch, C. Morton, G. Lenoir, D. Swan, S. Tronick, S. Aaronson, and P. Leder.** 1982. Translocation of the c-*myc* gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. Proc. Natl. Acad. Sci. USA **79**:7837–7841.
57. **Watson, D. K., M. C. Psallidopoulos, K. P. Samuel, R. Dalla-Favera, and T. S. Papas.** 1983. Nucleotide sequence analysis of human c-*myc* locus, chicken homologue, and myelocytomatosis virus MC29 transforming gene reveals a highly conserved gene product. Proc. Natl. Acad. Sci. USA **80**:3642–3645.
58. **Watson, D. K., E. P. Reddy, P. H. Duesberg, and T. S. Papas.** 1983. Nucleotide sequence analysis of the chicken c-*myc* gene reveals homologous and unique coding regions by comparison with the transforming gene of avian myelocytomatosis virus MC29, *gag-myc*. Proc. Natl. Acad. Sci. USA **80**:2146–2150.
59. **Watt, R., K. Nishikura, J. Sorrentino, A. ar-Rushdi, C. M. Croce, and G. Rovera.** 1983. The structure and nucleotide sequence of the 5′ end of the human c-*myc* oncogene. Proc. Natl. Acad. Sci. USA **80**:6307–6311.
60. **Watt. R., L. W. Stanton, K. B. Marcu, R. C. Gallo, C. M. Croce, and G. Rovera.** 1983. Nucleotide sequence of cloned cDNA of human c-*myc* oncogene. Nature (London) **303**:725–728.
61. **Welsh, J. D., C. Swimmer, T. Cocke, and T. Shenk.** 1986. A second domain of simian virus 40 T antigen in which mutations can alter the cellular localization of the antigen. Mol. Cell. Biol. **6**:2207–2212.
62. **Westaway, D., G. Payne, and H. E. Varmus.** 1984. Deletions and base substitutions in provirally mutated c-*myc* alleles may contribute to the progression of B-cell tumors. Proc. Natl. Acad. Sci. USA **81**:843–847.