MOLECULAR AND CELLULAR BIOLOGY, Oct. 1988, p. 4048–4054
0270-7306/88/104048-07$02.00/0

Vol. 8, No. 10

Identification of the Human c-*myc* Protein Nuclear Translocation Signal

CHI V. DANG[1] AND WILLIAM M. F. LEE[2]*

Hematology Division, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205,[1] and Cancer Research Institute, Department of Medicine, University of California San Francisco, San Francisco, California 94143[2]

Received 8 March 1988/Accepted 24 June 1988

We identified and characterized two regions of the human c-*myc* protein that target proteins into the nucleus. Using mutant c-*myc* proteins and proteins that fuse portions of c-*myc* to chicken muscle pyruvate kinase, we found that residues 320 to 328 (PAAKRVKLD; peptide M1) induced complete nuclear localization, and their removal from c-*myc* resulted in mutant proteins that distributed in both the nucleus and cytoplasm but retained rat embryo cell cotransforming activity. Residues 364 to 374 (RQRRNELKRSP; peptide M2) induced only partial nuclear targeting, and their removal from c-*myc* resulted in mutant proteins that remained nuclear but were cotransformationally inactive. We conjugated synthetic peptides containing M1 or M2 to human serum albumin and microinjected the conjugate into the cytoplasm of Vero cells. The peptide containing M1 caused rapid and complete nuclear accumulation, whereas that containing M2 caused slower and only partial nuclear localization. Thus, M1 functions as the nuclear localization signal of c-*myc*, and M2 serves some other and essential function.

Regulated expression of the c-*myc* proto-oncogene plays an important role in cell proliferation and differentiation, and aberrant expression is involved in the pathogenesis of a variety of neoplasms (3, 5, 32). c-*myc* expression is linked to the ability of cells to proceed through the cell cycle and perhaps to synthesize DNA (3). Its expression is downregulated in certain cells induced to differentiate, and constitutive expression interferes with the ability of murine erythroleukemia cells to differentiate normally (6, 8, 18). Alterations in the c-*myc* locus induced by proviral insertion, chromosomal translocation, or gene amplification are found in many tumors arising in chickens, mice, rats, and humans (5). The consequence of these changes appears to be the disruption of normal control mechanisms leading to augmented or inappropriate c-*myc* gene expression (5). Activation of c-*myc* oncogenic potential differs from activation of the *ras* family of proto-oncogenes in that it does not depend on the production of mutant proteins (20, 28), so that normal c-*myc* protein both mediates normal cell functions and can induce cellular changes leading to neoplasia.

The c-*myc* gene from different species contains three exons, of which only exons 2 and 3 encode the protein (1, 17, 27, 30). The human gene encodes a 439-amino-acid protein with a predicted molecular weight of 49,000 that migrates in sodium dodecyl sulfate-polyacrylamide gels with an M_r of 62,000 to 67,000 (13, 24). It is a short-lived phosphoprotein that is found in the nucleus and binds DNA in vitro (2, 9, 14, 23, 24). Within the nucleus, c-*myc* belongs to a discrete subset of nuclear proteins (10, 11) and colocalizes with small nuclear ribonucleoprotein particles, although a biochemical association has not been demonstrated (29). Its amino acid sequence has been deduced from the nucleotide sequence of c-*myc* genes and cDNAs and has led to many speculations about its properties and functions (36). However, information accumulated on c-*myc* has not permitted a critical evaluation of the functional significance of most of these properties of the c-*myc* protein.

* Corresponding author.

To address this issue, we undertook a systematic mutational analysis of the human c-*myc* gene and protein with the goal of studying how the structure, properties, and function of the c-*myc* protein are interrelated (31). Using the rat embryo cell cotransformation assay, we assessed the ability of various mutant c-*myc* genes to cooperate with a mutant *ras* gene and transform normal cells. We found that small mutations in two regions of the c-*myc* protein (amino acids 105 to 143 and 321 to 439; regions II and IV, respectively, in Fig. 6) seriously disrupt cotransforming activity, that large deletions in another portion of the protein (amino acids 1 to 104; region I in Fig. 6) are needed to abrogate activity, and that removal of the entire middle third of the protein (amino acids 144 to 320; region III in Fig. 6) still permits function. In contrast, the ability of c-*myc* protein to transform a rat-1-derived cell line (rat-1a cells) requires the integrity of this middle region. Our analysis also indicated that deletion of c-*myc* amino acids 320 to 368 resulted in mutant proteins that localize in both the nucleus and cytoplasm, suggesting that a sequence responsible for nuclear localization is present within this region. In studies described below, we defined and characterized two regions within the human c-*myc* protein that direct proteins to the nucleus. By correlating these regions with those important for c-*myc* transforming activity, we assessed the functional importance of these sequences.

MATERIALS AND METHODS

Antibodies. Polyclonal antiserum specific for human c-*myc* protein was obtained by injecting rabbits with c-*myc* protein purified from *Escherichia coli* (35). We obtained rabbit polyclonal antibody to the first *myc* homology box from Gerard Evans (22), rabbit polyclonal antibody to human c-*myc* C-terminal peptide from Gary Ramsay (24), rabbit anti-chicken muscle pyruvate kinase (PK) antibody from Bruce Roberts (15), and rabbit anti-human serum albumin (HSA) and rhodamine-conjugated goat antirabbit antibodies from Organon Teknika (Malvern, Pa.).

Cell culture, transfection, microinjection, and immunoflu-

orescence microscopy. Cell culture and immunofluorescence microscopy were performed as described previously (31). Cells injected with peptide-conjugated HSA were stained with antibodies as described previously (19), except fetal calf serum was replaced with 1% gelatin to block nonspecific immunofluorescence. Transfection of COS7 cells employing DEAE-dextran was performed as described previously (31). Transfection of rat embryo cells in the cotransformation assay (using normal or mutant human c-*myc* genes and a mutant *ras* gene) was performed as previously described (20, 31). Vero cells were injected with 10 to 15-fl samples of 1 mg of protein per ml in phosphate-buffered saline and returned to 37°C until ready to be fixed with paraformaldehyde (31).

Conjugation of synthetic peptides to HSA. Synthetic peptides PM1 (CGYGPAAKRVKLDS), PM2 (CGYGLERQRR NELKRSFFAL), and SV (CGYGPKKKRKV) were obtained from New Brunswick Scientific Co., Inc. (Edison, N.J.). Synthetic peptides were conjugated with 3-maleimidobenzoyl-*N*-hydroxysuccinimidester (MBS) to HSA as described previously (12). HSA-PM1 consists of 5 to 10 peptides per HSA molecule, HSA-PM2 has 10 to 15 peptides per HSA molecule and HSA-SV contains 5 to 10 peptides per HSA molecule as estimated by molecular weight change compared with MBS-modified HSA.

Mutagenesis. Mutants generated from the plasmid pM21 and their nomenclature have been described previously (31). Additional mutants were made in exon 3 with *Bal* 31 exonuclease after pM21 was linearized with *Cla*I or after the mutant In370 was linearized with *Xho*I. *Xho*I linkers (CTCGAGCTCGAG) were attached, the deleted genes were sequenced to determine the extent of deletion and the reading frame, and in-frame c-*myc* deletion mutants were generated as previously described (31). The deletion mutants are designated D followed by the positions of the first and last c-*myc* codons deleted. The frameshift mutant that terminates translation at codon 370 was generated by linearizing the mutant In370 with *Xho*I, followed by blunting of the ends with mung bean nuclease and ligation; the frameshift mutant terminates at codon 370 with Arg substituting for Leu.

c-*myc*-PK fusion genes. Portions of c-*myc* were fused to PK cDNA (gift from Bruce Roberts; see references 15 and 25). To fuse large 5′ portions of the c-*myc* gene to PK cDNA, we subcloned *Hin*dIII-*Xho*I fragments of c-*myc* mutants In54, In105, In143, In262, In370, In412, and D312-368 (all of which include the Moloney murine leukemia virus promoter and enhancer as well as portions of c-*myc*; [31]) into the *Hin*dIII and *Xho*I sites of the PK cDNA-containing plasmid, RLPK. The *Xho*I-*Eco*RI site of PK cDNA corresponds to codon 12. To fuse portions of c-*myc* exon 3 to the 3′ end of the PK cDNA, we subcloned *Xho*I-*Nsi*I fragments from insertion and deletion mutants (which contain the c-*myc* termination codon) into a modified pSP64 plasmid (NTB1) to create *Kpn*I and *Bgl*II sites at the 5′ and 3′ ends, respectively. These *Kpn*I-*Bgl*II fragments were subcloned into the *Kpn*I and *Bgl*II sites of RLPK. The *Kpn*I site corresponds to codon 443 of PK (26). Smaller fragments of c-*myc* exon 3 were generated from the *myc*-bearing NTB1 constructs by cleavage at the *Xho*II site (codon 380) and the *Sau*3AI site (codon 336). *Kpn*I-*Xho*II or *Kpn*I-*Sau*3AI c-*myc* fragments were subcloned into the *Kpn*I and *Bgl*II site of another modified pSP64 plasmid (NTB2) which provides a termination codon and a 3′ *Bam*HI site. *Kpn*I-*Bam*HI fragments from NTB2 containing c-*myc* fragments were subcloned into the RLPK *Kpn*I and *Bgl*II sites to fuse smaller portions of c-*myc* exon 3 to the 3′ end of the PK cDNA. The NTB2 plasmid also provided a 3′ *Eco*RI site that is 5′ to the termination codon. c-*myc* *Xho*I-*Eco*RI fragments from NTB2 were subcloned into RLPK *Xho*I-*Eco*RI sites to form fusion genes that generate fusion proteins with small fragments of c-*myc* joined to the N-terminal end of PK. Finally, synthetic oligonucleotides specifying peptides within c-*myc* exon 3 were made with 5′ *Xho*I and 3′ *Eco*RI ends. These were subcloned into RLPK to insert short sequences of c-*myc* as we attempted to define the shortest peptides able to specify nuclear location.

RESULTS

We identified two nuclear targeting sequences within the human c-*myc* protein by studying the subcellular location of proteins produced by (i) c-*myc* deletion mutants and (ii) c-*myc*-PK fusion genes. Their nuclear transport properties were determined by microinjecting HSA conjugated to each of these peptides.

c-*myc* deletion mutants. The subcellular localization of many of our c-*myc* deletion mutants has been described previously (31). The mutant c-*myc* genes were engineered into a plasmid containing the simian virus 40 (SV40) origin of replication, which allowed their overexpression after transfection into COS7 cells. The subcellular distribution of the mutant c-*myc* proteins produced was visualized 48 h after transfection by indirect immunofluorescence microscopy with anti-*myc* antiserum. These studies revealed that most of the c-*myc* deletion mutants produced nuclear proteins and that only deletion of c-*myc* amino acids 320 to 368 resulted in mutant proteins that were both nuclear and cytoplasmic in location (31). However, in our earlier study we did not exclude a contribution from the C-terminal six amino acids toward c-*myc* protein nuclear localization. Subsequently, we removed c-*myc* amino acids 370 to the C terminus by shifting the translational reading frame at codon 370. The resulting protein adds one unrelated amino acid before terminating prematurely and is located completely in the nucleus (data not shown). Thus, the last six amino acids do not play a role in c-*myc* nuclear localization.

We focused our attention on amino acids 320 to 368 by constructing additional mutants with deletions in this region using *Bal* 31 nuclease and examining the subcellular location of their protein products. The results (Fig. 1) showed that mutant c-*myc* proteins that delete amino acids 320 to 335 distribute in both nuclear and cytoplasmic compartments (cells containing these mutant proteins are identical in appearance to the cell shown in Fig. 2B), whereas those that only delete amino acids outside of this region remain totally nuclear (cells containing these mutant proteins are identical in appearance to the cells shown in Fig. 2C). A relatively small deletion mutant (D319-341) that removed little other than this region also displayed nuclear and cytoplasmic distribution (data not shown). These data indicate that the residues between 320 and 335 are responsible for the normal nuclear location of human c-*myc* protein. Notably, however, c-*myc* mutants missing this region remain partially nuclear and are not totally cytoplasmic.

c-*myc*-PK fusion proteins. The partial nuclear distribution of c-*myc* mutants with deletions of amino acids 320 to 335 indicates that either the mutant proteins are now small enough to diffuse into the nucleus or additional nuclear targeting sequences exist elsewhere in the c-*myc* protein. The former possibility seems unlikely for the protein produced by D319-341, since it is only slightly smaller than wild-type c-*myc* protein. We addressed the latter possibility by fusing a chicken muscle PK cDNA which encodes a large

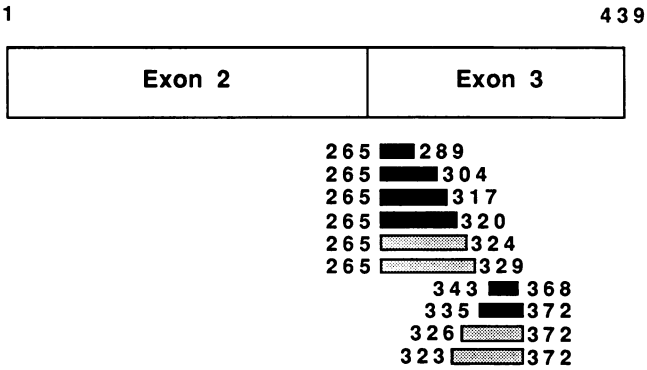

FIG. 1. Subcellular location of mutant human c-*myc* proteins. Human c-*myc* genes containing deletions were placed in a vector containing an SV40 origin of replication and transfected into COS7 cells. The subcellular distribution of the mutant proteins was determined by indirect immunofluorescent staining for c-*myc* protein after 48 h (31). The large open box represents the human c-*myc* protein with the portions encoded by exons 2 and 3 indicated. The narrow bars indicate the deleted segment for each of the mutants, with the positions of the first and last deleted amino acids indicated; black bars indicate that the mutant proteins are nuclear, and gray bars indicate that the mutant proteins are nuclear and cytoplasmic.

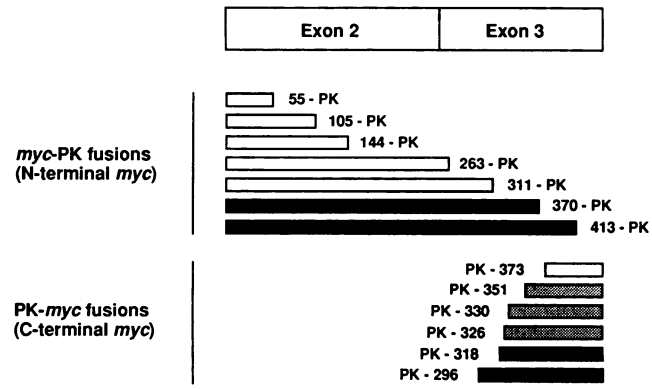

FIG. 3. Subcellular location of c-*myc*-PK fusion proteins determined by immunofluorescence microscopy. The upper set depicts the subcellular location of fusion proteins containing amino-terminal segments of c-*myc* fused to the amino terminus of PK; the lower set depicts the location of proteins containing carboxyl-terminal segments of c-*myc* fused to the carboxyl terminus of PK. The large upper box represents the c-*myc* protein, and the portions encoded by exons 2 and 3 are indicated. The bars represent the extent of the c-*myc* protein present in each of the fusion proteins, and the position of the last (for N-terminal *myc* fusions) or the first (for C-terminal *myc* fusions) c-*myc* amino acid present is indicated. ▭, Cytoplasmic distribution; ▨, nuclear and cytoplasmic distribution; ■, nuclear distribution.

cytoplasmic protein to portions of the c-*myc* gene to produce c-*myc*-PK fusion genes and proteins that we could localize in cells by indirect immunofluorescence microscopy with anti-PK antiserum. We observed three patterns of subcellular distribution of the fusion proteins: (i) cytoplasmic, (ii) nuclear and cytoplasmic, and (iii) totally nuclear location (Fig. 2). The subcellular distribution of the various fusion proteins is summarized in Fig. 3. When N-terminal segments of c-*myc* were fused to the N terminus of PK, we observed that c-*myc* amino acids to 311 failed to direct the fusion proteins into the nucleus and that they remained totally cytoplasmic. Inclusion of amino acids 312 to 370 made the fusion protein totally nuclear. Fusion of c-*myc* amino acids 373 to 439 to the C terminus of PK was unable to direct the fusion protein into the nucleus. However, inclusion of c-*myc* amino acids 351 to 372 resulted in a protein that was both nuclear and cytoplasmic in distribution, indicating that these residues can direct proteins partially into the nucleus and may account for the nuclear and cytoplasmic distribution of mutant c-*myc* proteins that have deleted residues 320 to 335 (Fig. 1). Only when amino acids 318 to 350 of c-*myc* were included in the fusion protein did it become totally nuclear. These results indicate that c-*myc* peptide 318 to 350 either contains an additional nuclear targeting signal or has a synergistic effect on the nuclear targeting potential of amino acids 351 to the C terminus.

To distinguish between these possibilities, we dissected this region using convenient restriction fragments or oligonucleotides to create c-*myc*-PK fusions containing smaller c-*myc* peptides. Using a c-*myc* restriction fragment generated by *Sau*3AI (which cuts at codon 336), we found that c-*myc* amino acids 317 to 336 are sufficient to target PK completely into the nucleus (data not shown). We inserted synthetic oligonucleotides encoding smaller segments of this

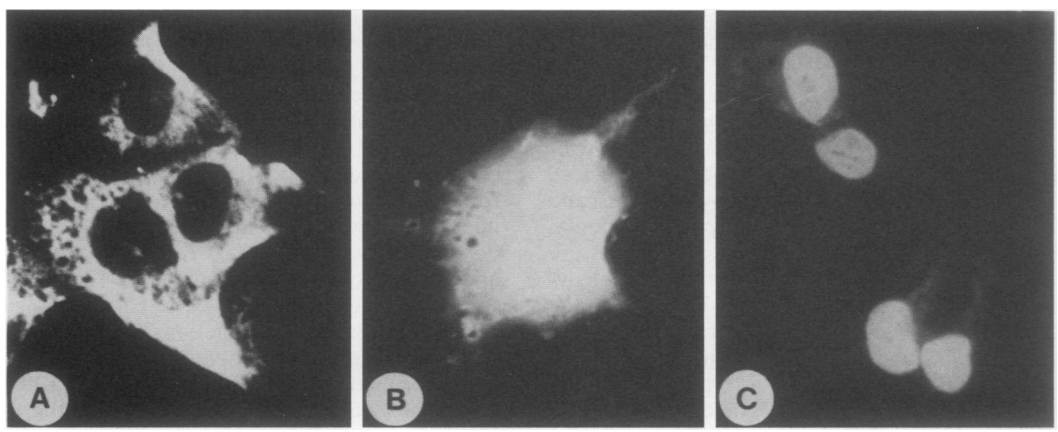

FIG. 2. Immunofluorescent staining for chicken muscle PK in COS7 cells transfected with c-*myc*-PK fusion genes. Fusion genes were made in a vector containing the SV40 origin of replication as described in the text and transfected into COS7 cells. After 48 h, the cells were stained with anti-PK antiserum to localize the fusion proteins. The photomicrographs show representative cells expressing fusion proteins that are (A) cytoplasmic, (B) nuclear and cytoplasmic, and (C) nuclear.

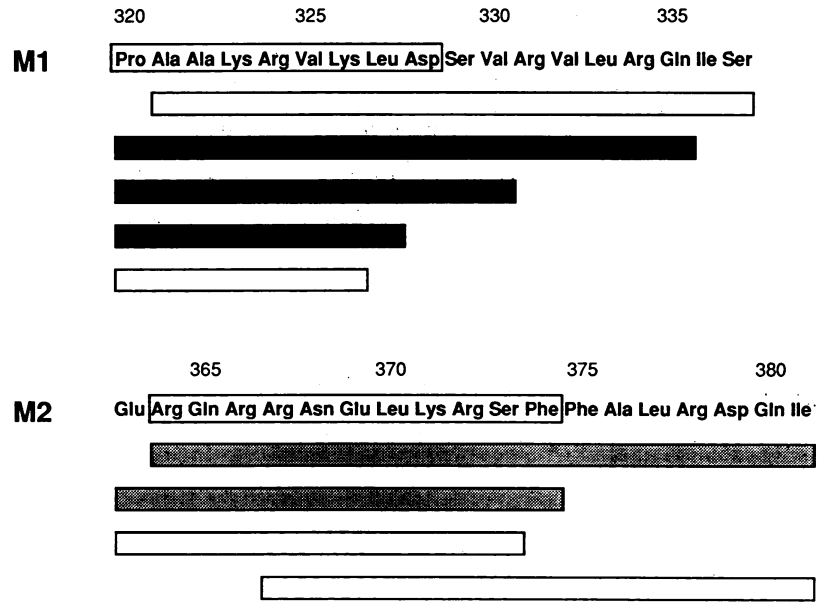

FIG. 4. Definition of the nuclear targeting sequences of human c-*myc* protein. The amino acid sequences of c-*myc* around the two nuclear localization peptides (boxed amino acids) are shown. The bars indicate the peptides encoded by oligonucleotides used to make c-*myc*-PK fusion genes that precisely define the two nuclear targeting sequences, M1 and M2. The subcellular distribution of the fusion proteins is indicated by the same convention as in Fig. 3. Note that, although Asp-328 is not necessary for M1 to function, the introduction of a Glu at exactly the same position in the c-*myc*-PK fusion protein (resulting from the *Eco*RI linker used) may result in a functionally equivalent substitution and prevents us from excluding the importance of an acidic residue at position 328.

region into the *Xho*I and *Eco*RI sites of the recombinant PK cDNA (15), and the subcellular location of the resulting fusion proteins defined a minimum c-*myc* sequence (amino acids 320 to 328; peptide M1) that enabled attached proteins to become completely nuclear (Fig. 4). Omission of either Pro-320 or Leu-327 resulted in a peptide that was incapable of nuclear targeting. Although Asp-328 is not apparently necessary for this sequence to function, the *Eco*RI cloning site at the 3′ end of the oligonucleotides contains GAA, which encodes a Glu, so that we cannot exclude the possibility that the homologous amino acid, Asp-328, is required for nuclear targeting. The region from amino acid 350 to the C terminus of c-*myc* was analyzed by a similar strategy. Using a c-*myc* restriction fragment generated by *Xho*II (which cuts at codon 381), we found that c-*myc* amino acids 351 to 381 are as effective as amino acids 351 to 439 in directing PK partially into the nucleus (data not shown). Then, by using oligonucleotides, we defined a small peptide within this region (amino acids 364 to 374; peptide M2) that is responsible for this property (Fig. 4). Note that our data do not allow us to be certain that either Arg-364 or Gln-365 is necessary for this peptide to function.

These studies demonstrate that there are two distinct and independent peptides within the c-*myc* protein that have nuclear targeting potential. We can be certain that no region of the protein outside of residues 320 to 374 has this property, but the results presented thus far cannot exclude the presence of a third nuclear targeting peptide lying between the two we already defined. To be certain that such a peptide does not exist, we created a c-*myc*-PK gene encoding PK fused to c-*myc* peptide 324 to 364 (which includes all residues lying between the first and second targeting signals), and this did not localize in nuclei (data not shown). This indicated that no additional nuclear targeting sequences reside between c-*myc* peptides 320 to 328 and 364 to 374. As confirmation that no other portion of the c-*myc* protein facilitates nuclear accumulation, introduction into COS7 cells of a c-*myc* deletion mutant (D312-412) that has removed both M1 and M2 resulted in 20% of the transfected cells displaying exclusively cytoplasmic fluorescence (this population of cells is identical in appearance to the cells shown in Fig. 2A), while the remaining cells showed predominantly cytoplasmic fluorescence. The small amount of residual nuclear fluorescence seen in the majority of the cells may be due to the considerably deleted and smaller mutant protein (predicted M_r of ~38,000) diffusing into cell nuclei (7).

Properties of c-*myc* nuclear targeting sequences. Having identified the two regions of c-*myc* able to direct proteins into the nucleus, we studied the targeting properties of these peptides. We conjugated synthetic peptides PM1 (CGYG PAAKRVKLDS) and PM2 (CGYGLERQRRNELKRSF FAL), which contain the two nuclear targeting signals of c-*myc* (underlined), to HSA. HSA and the conjugates HSA-PM1 and HSA-PM2 were microinjected into the cytoplasm of Vero cells to determine the kinetics of nuclear accumulation. HSA remained in the cytoplasm 2 h after microinjection; HSA-PM1 localized completely within nuclei in 15 min; and HSA-PM2 localized only partially within nuclei even after 90 min, with nuclear fluorescence being slightly more intense than cytoplasmic fluorescence (Fig. 5). Five hours after microinjection, there was a decrease in overall fluorescence intensity for each of the injected proteins, but the pattern remained unchanged. These results indicate that the c-*myc* peptide M1 (PAAKRVKLD) causes rapid and complete nuclear accumulation of covalently linked proteins and behaves like a nuclear transport signal, whereas peptide M2 (RQRRNELKRSF) appears to cause less complete and slower nuclear accumulation. For comparison, we also conjugated a peptide (SV) containing the SV40 T-antigen nuclear targeting sequence (PKKKRKV) to HSA and microinjected HSA-SV in parallel experiments. We found no

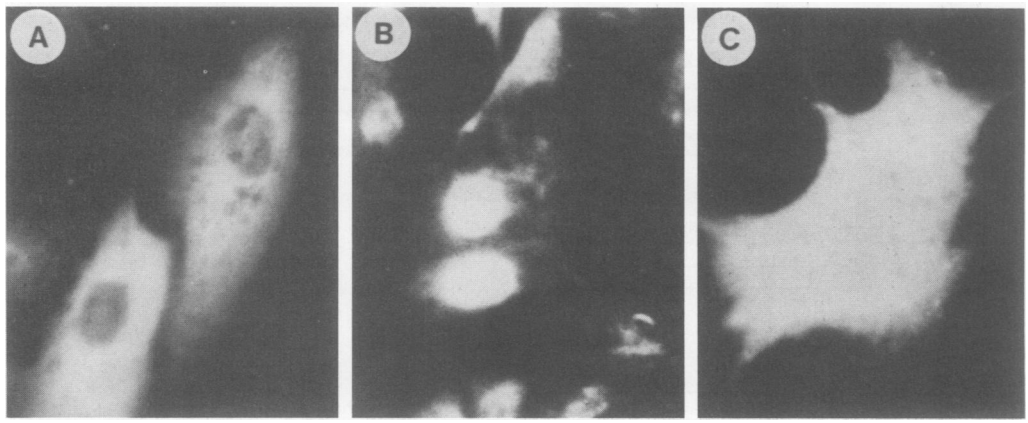

FIG. 5. Subcellular location of microinjected HSA or HSA conjugated to peptides containing human c-*myc* nuclear targeting sequences. Peptides containing c-*myc* amino acids 320 to 328 (PM1) or amino acids 364 to 374 (PM2) were conjugated to HSA. HSA or conjugated HSA was microinjected into the cytoplasm of Vero cells which were fixed and stained with anti-HSA antiserum at various times after microinjection. The photomicrographs show cells fixed and stained (A) 2 h after microinjection of unconjugated HSA, (B) 15 min after microinjection of HSA-PM1, and (C) 90 min after microinjection of HSA-PM2.

difference between the rate or degree of nuclear accumulation of HSA-PM1 and HSA-SV. Intranuclear injection of HSA and conjugated HSA resulted in retention of these proteins in the nucleus, indicating that flux of these macromolecules through the nuclear membrane is unidirectional.

c-*myc* nuclear location and transforming activity. We tested D319-341, a c-*myc* mutant that deletes the M1 targeting signal, in the rat embryo cell cotransformation assay (20, 31). The protein produced, which localized to both cytoplasm and nucleus, retained 25% of wild-type c-*myc* cotransforming activity. We interpret this to indicate that the fraction of D319-341 protein that enters the nucleus is biologically active and is responsible for the cotransforming activity of this mutant. Previous studies showed that mutant proteins that disrupt M2 (e.g., In370, In373, D371-412) are located completely in the nucleus but lack any cotransformation activity (31). Thus, peptide 320 to 328 is essential for normal c-*myc* nuclear localization but not for biological activity, whereas peptide 364 to 374 is essential for c-*myc* activity but not for its normal subcellular distribution.

DISCUSSION

We identified two regions of the human c-*myc* protein with nuclear targeting potential. The first peptide, M1 (PAAKRVKLD), can direct covalently linked cytoplasmic proteins into the nuclei of cells in a rapid and complete manner and has the characteristics expected of a nuclear transport signal peptide. Deletion of this region from human c-*myc* results in the nuclear and cytoplasmic distribution of the mutant protein. The fact that a mutant which deletes this signal (D319-341) is not totally cytoplasmic suggests that the second peptide we identified, M2 (RQRRNELKRSF), is responsible for this partial nuclear distribution. In support of this, the second peptide directs covalently linked cytoplasmic proteins partially into the nuclei of cells and appears to do so at a slower rate than M1. Because removal of M2 from c-*myc* results in no discernible disturbance of the normal nuclear distribution of the protein, and because its properties indicate that its presence is unlikely to contribute to the rate or degree of c-*myc* nuclear accumulation when M1 is present, we doubt that M2 normally serves as a nuclear translocation sequence within the c-*myc* protein. We believe that c-*myc* protein effectively has only one nuclear translocation sequence (M1) and that the karyophilic property of M2 may be epiphenomenal.

Comparisons of c-*myc*, N-*myc*, and L-*myc* sequences (Table 1) reveal that c-*myc* and N-*myc* of different species have an M1-like peptide: the basic amino acids at 323, 324, and 326, the hydrophobic amino acids at 325 and 327, and the acidic amino acid at 328 are conserved. The Pro at 320 is conserved in c-*myc* from all species except *Xenopus laevis* and in human N-*myc* and is followed by two neutral amino acids in all cases. Thus, the region homologous to M1 in different c-*myc* and N-*myc* proteins is likely to be their nuclear targeting signal. Mouse and human L-*myc* genes do not have a sequence homologous to M1 (16, 21).

The amino acid sequence of M1 resembles the C-terminal nuclear targeting signal of polyomavirus large T (PPK KARED [25]) and the solitary nuclear targeting signal present in SV40 large T (PKKKRKV [15]) (Table 1). Like M1, both these sequences contain many basic residues and are in a region of the protein predicted to be a α-helical (4). However, the polyomavirus sequence is unable to target PK into the nucleus and appears to be only one of two independent nuclear translocation signals in large T. The other sequence (VSRKRPRPA [25]) bears less of a resemblance to M1, is predicted to be non-α-helical, but is capable of

TABLE 1. M1 and M2 amino acid homology among *myc*-related and other proteins

Protein	M1 peptide	M2 peptide
c-*myc*		
Human	PAAKRVKLD	RQRRNELKRSF
Murine	PAAKRAKLD	RQRRNELKRSF
Feline	PAAKRAKLD	RQRRNELKRSF
Avian	PAAKRLKLD	RQRRNELKLSF
Trout	PAVKRLRLE	RQRRNELKLSF
Frog	VSSKRAKLE	RQRRNELKLSF
Human N-*myc*	PPQKKIKSE	RQRRNDLRSSF
Human L-*myc*		RKRRNDLRSRF
Polyomavirus large T	PPKKARED VSRKRPRPA	
SV40 large T	PKKKRKV	
Achaete-scute		
T4		RNRVKQVNNSF
T5		RNRVKQVNNGF

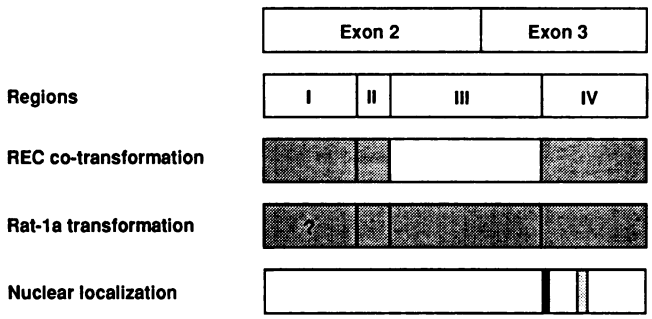

FIG. 6. Summary of regions of the human c-*myc* protein important for different activities and properties. The top box diagrams the 439-amino-acid protein, with the portions encoded by exons 2 and 3 indicated. The next box represents regions defined on the basis of the transforming activity of mutant c-*myc* genes (31): I (residues 1 to 104), II (residues 105 to 143), III (residues 144 to 320), and IV (residues 321 to 439). The two boxes below show regions (shaded) important for activity in the assays indicated (the ? indicates that region I was not analyzed in the rat-1a transformation assay). The last box indicates the domains with nuclear localization properties (residues 320 to 328 and 364 to 374). REC, Rat embryo cell.

directing PK into the nucleus. By the criteria of speed and completeness of nuclear transport of covalently linked HSA, M1 and the SV40 sequence appear to be functionally equivalent. Yet, M1 is not as basic as the SV40 T-antigen sequence, which implies that the additional basic amino acids do not enhance detectably the function of the signal. Perhaps a minimum number of basic amino acids in an α-helical array is sufficient to specify nuclear targeting function.

The second c-*myc* peptide with nuclear targeting properties, M2, shows no homology to other known nuclear targeting signals (7). Like M1, it is rich in basic amino acids and occurs in a region predicted to be α-helical. M2 is highly conserved within the *myc* family of proteins, and unlike M1, is conserved in L-*myc* (Table 1). In addition, this region shows homology to *Drosophila* achaete-scute complex proteins T4 and T5 (33) and is within the region that shows weak homology to *fos*, *jun*, and GCN4 (34). For reasons previously discussed, M2 is unlikely to function as a nuclear transport signal within the context of the c-*myc* protein, but given how important the integrity of this region of c-*myc* is for both rat embryo cell cotransforming and rat-1a transforming activities (Fig. 6) (31), it must participate in some essential aspect of c-*myc* function.

ACKNOWLEDGMENTS

We thank the following for gifts of materials and advice: Dan Cashman and Marie Dougherty (oligonucleotides); Bruce Roberts (plasmids containing PK cDNA suitable for all our manipulations and anti-PK antiserum); Rosemary Watt (pOTS*myc* and the AR68 strain of *E. coli*); David Goldfarb (advice on peptide coupling); Gary Ramsay and J. Michael Bishop (peptides, HSA, HSA-PM1, and HSA-SV; and antisera to HSA and a C-terminal c-*myc* peptide). We thank Titia de Lange and Harold Varmus for stimulating discussions and Y. W. Kan for use of his microinjection apparatus. We are grateful for the enthusiastic and expert assistance of Mary Brooks and Chin Lin in parts of this project.

This investigation was supported (in part) by California Division-American Cancer Society senior fellowship S-16-87 to C.V.D. C.V.D. is a Hubert E. and Anne E. Rogers Scholar in Academic Medicine. W.M.F.L. is a recipient of an American Cancer Society junior faculty research award. This research was supported by Public Health Service grant CA 38783 (to W.M.F.L.) from the National Institutes of Health.

LITERATURE CITED

1. **Battey, J., C. Boulding, R. Taub, W. Murphy, T. Stewart, H. Potter, G. Lenoir, and P. Leder.** 1983. The human c-*myc* onogene: structural consequences of translocation into the IgH locus in Burkitt lymphoma. Cell **34**:779–787.
2. **Beimling, P., T. Benter, T. Sander, and K. Moelling.** 1985. Isolation and characterization of the human cellular *myc* gene product. Biochemistry **24**:6349–6355.
3. **Bishop, J. M.** 1987. The molecular genetics of cancer. Science **235**:305–311.
4. **Chou, P. Y., and G. D. Fasman.** 1978. Empirical predictions of protein conformation. Annu. Rev. Biochem. **47**:251–276.
5. **Cole, M. D.** 1986. The *myc* oncogene: its role in transformation and differentiation. Annu. Rev. Genet. **20**:361–384.
6. **Coppola, J. A., and M. D. Cole.** 1986. Constitutive c-*myc* oncogene expression blocks mouse erythroleukaemia cell differentiation but not commitment. Nature (London) **320**:760–763.
7. **Dingwall, C., and R. A. Laskey.** 1986. Protein import into the cell nucleus. Annu. Rev. Cell Biol. **2**:367–390.
8. **Dmitrovsky, E., W. M. Kuehl, G. F. Hollis, I. R. Kirsch, T. P. Bender, and S. Segal.** 1986. Expression of a transfected human c-*myc* oncogene inhibits differentiation of a mouse erythroleukemia cell line. Nature (London) **322**:748–750.
9. **Donner, P., T. Bunte, I. Greiser-Wilke, and K. Moelling.** 1983. Decreased DNA-binding ability of purified transformation-specific proteins from deletion mutants of the acute avian leukemia virus MC29. Proc. Natl. Acad. Sci. USA **80**:2861–2865.
10. **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.
11. **Evan, G. I., and D. C. Hancock.** 1985. Studies on the interaction of the human c-*myc* protein with cell nuclei: p62 c-*myc* as a member of a discrete subset of nuclear proteins. Cell **43**:253–261.
12. **Goldfarb, D. S., J. Gariepy, G. Schoolnik, and R. D. Kornberg.** 1986. Synthetic peptides as nuclear localisation signals. Nature (London) **322**:641–644.
13. **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.
14. **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.
15. **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.
16. **Kaye, F., J. Battey, M. Nau, B. Brooks, E. Seifter, J. deGreve, M. Birrer, E. Sausville, and J. Minna.** 1988. Structure and expression of the human L-*myc* gene reveal a complex pattern of alternative mRNA processing. Mol. Cell. Biol. **8**:186–195.
17. **King, M. W., J. M. Roberts, and R. N. Eisenman.** 1986. Expression of the c-*myc* proto-oncogene during development of *Xenopus laevis*. Mol. Cell. Biol. **6**:4379–4386.
18. **Lachman, H. M., G. H. Cheng, and A. I. Skoultchi.** 1986. Transfection of mouse erythroleukemia cells with *myc* sequences changes the rate of induced commitment to differentiate. Proc. Natl. Acad. Sci. USA **83**:6480–6484.
19. **Lanford, R. E., P. Kanda, and R. C. Kennedy.** 1986. Induction of nuclear transport with a synthetic peptide homologous to the SV40 T antigen transport signal. Cell **46**:575–582.
20. **Lee, W. M., M. Schwab, D. Westaway, and H. E. 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.
21. **Legouy, E., R. DePinho, K. Zimmerman, R. Collum, G. Yancopoulos, L. Mitsock, R. Kriz, and F. Alt.** 1987. Structure and expression of the murine L-*myc* gene. EMBO J. **6**:3359–3366.
22. **Moore, J. P., D. C. Hancock, T. D. Littlewood, and G. I. Evan.** 1987. A sensitive and quantitative enzyme-linked immunosorbence assay for the c-*myc* and N-*myc* oncoproteins. Oncogene Res. **2**:65–80.
23. **Persson, H., and P. Leder.** 1984. Nuclear localization and DNA binding properties of a protein expressed by human c-*myc*

oncogene. Science **225**:728–731.
24. **Ramsay, G. R., 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.
25. **Richardson, W. D., B. L. Roberts, and A. E. Smith.** 1986. Nuclear location signals in polyoma virus large-T. Cell **44**:77–85.
26. **Roberts, B. L., W. D. Richardson, and A. E. Smith.** 1987. The effect of protein context on nuclear location signal function. Cell **50**:465–475.
27. **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.
28. **Sinn, E., W. Muller, P. Pattengale, I. Tepler, R. Wallace, and P. Leder.** 1987. Coexpression of MMTV/v-Ha-*ras* and MMTV/c-*myc* genes in transgenic mice: synergistic action of oncogenes in vivo. Cell **49**:465–475.
29. **Spector, D. L., R. A. Watt, and N. F. Sullivan.** 1987. The v- and c-*myc* oncogene proteins co-localize in situ with small nuclear ribonucleoprotein particles. Oncogene **1**:5–12.
30. **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.
31. **Stone, J., T. deLange, G. Ramsay, E. Jakobovits, J. M. Bishop, H. E. Varmus, and W. Lee.** 1987. Definition of regions in human c-*myc* that are involved in transformation and nuclear localization. Mol. Cell. Biol. **7**:1697–1709.
32. **Varmus, H.** 1987. Cellular and viral oncogenes, p. 271–346. *In* G. Stamatoyannopoulos, A. W. Nienhius, P. Leder, and P. W. Majerus (ed.), The molecular basis of blood diseases. The W. B. Saunders Co., Philadelphia.
33. **Villares, R., and C. V. Cabrera.** 1987. The achaete-scute gene complex of D. melanogaster: conserved domains in a subset of genes required for neurogenesis and their homology to *myc*. Cell **50**:415–424.
34. **Vogt, P. K., T. J. Bos, and R. F. Doolittle.** 1987. Homology between the DNA-binding domain of the GCN4 regulatory protein of yeast and carboxyl-terminal region of a protein coded for by the oncogene jun. Proc. Natl. Acad. Sci. USA **84**:3316–3319.
35. **Watt, R. A., A. R. Shatzman, and M. Rosenberg.** 1985. Expression and characterization of the human c-*myc* DNA-binding protein. Mol. Cell. Biol. **5**:448–456.
36. **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.