Proc. Natl. Acad. Sci. USA
Vol. 84, pp. 170–173, January 1987
Genetics

Evolutionarily conserved regions of the human c-myc protein can be uncoupled from transforming activity

(mutations/oncogenesis/evolution)

JACOB SARID, THANOS D. HALAZONETIS, WILLIAM MURPHY, AND PHILIP LEDER

Department of Genetics, Harvard Medical School, and Howard Hughes Medical Institute, 25 Shattuck Street, Boston, MA 02115

Contributed by Philip Leder, September 4, 1986

ABSTRACT **The *myc* family of oncogenes contains coding sequences that have been preserved in different species for over 400 million years. This conservation (which implies functional selection) is broadly represented throughout the C-terminal portion of the human c-myc protein but is largely restricted to three clusters of amino acid sequences in the N-terminal region. We have examined the role that the latter three regions of the c-myc protein might play in the transforming function of the c-*myc* gene. Several mutations, deletions and frameshifts, were introduced into the c-*myc* gene, and these mutant genes were tested for their ability to collaborate with the EJ-*ras* oncogene to transform rat embryo fibroblasts. Complete elimination of the first two N-terminal conserved segments abolished transforming activity. In contrast, genes altered in a portion of the second or the entire third conserved segment retained their transforming activity. Thus, the latter two segments are not required for the transformation process, suggesting that they serve another function related only to the normal expression of the c-*myc* gene.**

Three members of the *myc* gene family have been implicated in human malignancy. c-*myc* is associated with the development of a variety of tumor types in several species, including humans (1–3), while N-*myc* (4, 5) and L-*myc* (6) have been implicated in the development of malignancy in human cells of neural crest origin. The complete sequences of the human c- and N-*myc* genes (7–10) indicate three patches of conserved amino acid sequence in the second exon, each approximately 20 amino acids long, separated from one another by regions of either low or no sequence homology (refs. 9 and 10; Table 1 and Fig. 2). These three regions are also conserved between c-*myc* genes of mouse, chicken, and fish (11–13), which diverged over 400 million years ago. The third exon, possibly required for the DNA binding activity of the c-myc protein (14, 15), is extensively conserved between human, chicken, and fish c-*myc* and human N-*myc* (7–13).

While the function of the c-*myc* gene is not known, its transforming properties (16), its inducibility by growth factors (17), and its nuclear localization (18–20) suggest that it might be a nuclear mediator of the action of growth factors. Thus the c-myc protein might be a transducing molecule, receiving a signal and consequently carrying out an action. If the conserved segments of the c-myc protein are to be fitted to such a model, separate functions might be segregated in different conserved portions of the protein and thus be uncoupled from one another. Guided by the example of a transformation-competent c-*myc* gene with an extensive mutation in a conserved region of the second exon (21), we have systematically mutated this exon and assayed the ability of the mutant genes to participate in transformation.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Site-Directed Mutagenesis in the Second Exon of the Human c-*myc* Oncogene. The *Sma* I–*Sst* I DNA fragment of the human c-*myc* gene containing the complete second exon and part of the first and second introns was inserted into the polylinker of phage mp10 M13. Oligonucleotides that incorporated the desired mutations were hybridized to the single-stranded phage M13 DNA and were used as primers for synthesis of the complementary DNA strand, as described by Zoller and Smith (22). The oligonucleotides for the deletion mutants were 30-mers, 15 nucleotides of which were complementary to sequences upstream and 15 to sequences downstream of the region to be deleted. The oligonucleotides for the frameshift mutants were 16-mers complementary to c-*myc* but bearing a single nucleotide insertion or deletion (Fig. 1*A*). For each frameshift mutant two mutations were introduced; one was an insertion and the other a deletion. All mutations were confirmed by sequencing, which was performed by using the Sanger dideoxy chain-termination method (23) as described in the Bethesda Research Laboratories users' manual. To establish that the mutagenesis technique

A

Deletion Mutants:

Dhcc1	ATGCCCCTCAACGTT(2864-2911)TATTTCTACTGCGAC
Dcnc1	GAGCTGCAGCCCCCG(2978-3064)GGGCTCTGCTCGCCC
Dhcc2	GGTGGCGGGAGCTTC(3128-3175)ATGGTGAACCAGAGT
Dcnc2	TTCATCAAAAACATC(3218-3280)GTCTCAGAGAAGCTG
Dcnc3	TTGTACCTGCAGGAT(3374-3454)CCTCTCAACGACAGC

Frameshift Mutants:

	+ (2969)		− (3032)
FScnc1	GAGCGAGCTGTCAGCC	and	CCGCCCCTG_CCCCTAG
	− (3032)		+ (3145)
FSnoc1	CCGCCCCTG_CCCCTAG	and	GACCAGCTTGGAGATG
	+ (3145)		− (3262)
FScnc2	GACCAGCTTGGAGATG	and	TGGAGCGGCTT_TCGGC

B

FScnc1	41	SAPGAQRGYLEEIRAAAHPAPA	62
c-myc		QPPAPSEDIWKKFELLPTPPLS	
FSnoc1	62	PLAAAPGSARPPTLRSHPSPFGETTTAVAGASPRPTSL	99
c-myc		SPSRRSGLCSPSYVAVTPFSLRGDNDGGGGSFSTADQL	
FScnc2	99	LGDGDRAAGRRHGEPEFHLRPGRRDLHQKHHHPGLYVERL	138
c-myc		LEMVTELLGGDMVNQSFICDPDDETFIKNIIIQDCMWSGF	

FIG. 1. (*A*) Oligonucleotides used to generate the human c-*myc* mutants. Deleted and inserted nucleotides are numbered according to Battey *et al.* (8). (*B*) Predicted amino acid sequences of the mutated regions of the three frameshift mutants. The amino acid sequence of the wild-type human c-*myc* is also indicated for comparison. The standard single-letter code is used.

Table 1. Homology of human c-myc to chicken c-myc and human N-myc based on the predicted primary structures

Region*	Amino acid residues			Percent homology†	
	Human c-myc	Chicken c-myc	Human N-myc	Human/chicken c-myc	Human c-myc/N-myc
1	1–44	1–47	1–44	64	25
2	45–65	48–68	45–65	100	95
3	66–93	69–79	66–75	21	5
4	94–127	80–113	76–109	91	15
5	128–144	114–130	110–126	100	77
6	145–184	131–184	127–182	43	15
7	185–200	185–200	183–198	88	69
8	201–252	201–229	199–267	10	8

Data are based on refs. 7–11.

*The 252 amino acids of human c-myc encoded by the second exon were allocated to eight regions. Five of these regions are homologous between human and chicken c-myc and three are homologous between human c-myc and N-myc.

†The percent homology of a given region is calculated as the number of identical amino acids divided by the average number of amino acids of the two regions being compared.

did not introduce undesirable mutations, the *Sma* I–*Sst* I fragments of the three frameshift mutants were sequenced to completion in both orientations. No mutations other than the ones specifically introduced were identified.

The mutated *Sma* I–*Sst* I c-*myc* fragments from the M13 vector were inserted into the pSV7Humyc plasmid. This plasmid, which has a simian virus 40 enhancer and early region promoter driving the second and third exons of human c-*myc*, has been described previously (21).

In one of the generated deletion mutants (Dhcc1, described in *Results*), a *Sma* I site was fortuitously created within the second exon by the oligonucleotide-directed mutagenesis. This site was used to insert a truncated c-*myc* second exon into the pSV7Humyc expression vector. An in-frame ATG codon corresponding to amino acid 101 presumably allowed for initiation of translation, thus giving rise to a c-myc protein with a deletion of the first hundred amino acids. This mutant was called D100.

Transformation of Rat Embryo Fibroblasts (REF) and Focus Formation Assay. Primary cultures of REF were prepared as described by Land *et al.* (16). Secondary cells were cotransfected with the pSV7Humyc plasmids bearing either the wild-type or the mutant c-*myc* genes and plasmid pEJ6.6, which contains the EJ-Ha-c-*ras* oncogene. Foci of transformed cells were scored 2 weeks later. The transfection protocol was as described by Land *et al.* (16), except that the serum concentration in the culture media was reduced to 3% when the cells became confluent.

RESULTS

Generation of Mutations in the Second Exon of the Human c-*myc* Oncogene. Using oligonucleotide-directed mutagene-

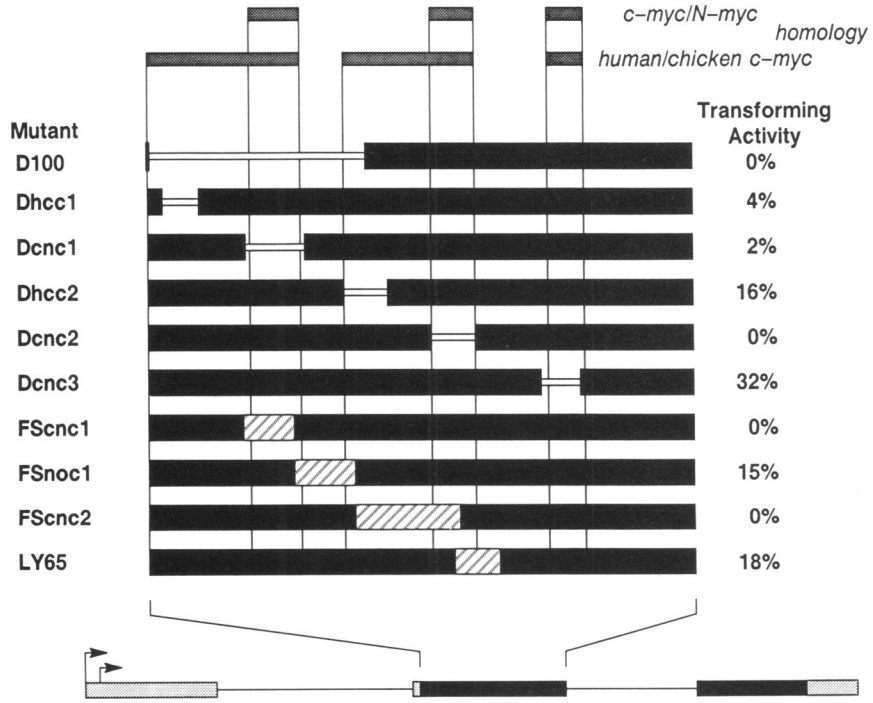

FIG. 2. Transforming efficiency of the constructed mutant c-*myc* genes and the translocated c-*myc* gene from the Burkitt lymphoma cell line Ly65 (21). Double lines indicate deleted regions, whereas the hatched boxes indicate the mutated regions of the frameshift mutants. The conserved regions between human c-*myc* and N-*myc* and between human and chicken c-*myc* are also shown. The map is drawn to scale. All mutations were within the coding sequence of the second exon of the c-*myc* gene. The structure of the gene is shown at the bottom of the figure. It has three exons (boxes), of which only the latter two contain coding sequences (solid areas). The two c-*myc* promoters are indicated by arrows.

Table 2. Transforming activity of the *myc* mutants as determined by focus formation of REF cotransfected with *ras*

Transfected DNA	Exp. 1		Exp. 2		Exp. 3		Exp. 4		Exp. 5		Exp. 6	
	Plates	Foci	Plates	Foci	Plates	Foci	Plates	Foci	Plates	Foci	Plates	Foci
pSV7PL-wt *myc* + pEJ.6.6 (*ras*)	6/6	12	6/6	21	6/6	25	6/6	18	5/6	12	6/6	17
pSV7PL-D100 *myc* + pEJ.6.6 (*ras*)	ND	ND	0/6	0	0/6	0	0/6	0	0/6	0	0/6	0
pSV7PL-Dhcc1 *myc* + pEJ.6.6 (*ras*)	ND	ND	1/6	0.5	1/6	0.5	2/6	1	0/6	0	1/6	1
pSV7PL-Dcnc1 *myc* + pEJ.6.6 (*ras*)	1/6	1	1/6	0.5	0/6	0	0/6	0	0/6	0	0/6	0
pSV7PL-Dhcc2 *myc* + pEJ.6.6 (*ras*)	ND	ND	4/6	4	3/6	3	3/6	3	2/6	2	3/6	3
PSV7PL-Dcnc2 *myc* + pEJ.6.6 (*ras*)	0/6	0	0/6	0	0/6	0	0/6	0	0/6	0	0/6	0
pSV7PL-Dcnc3 *myc* + pEJ.6.6 (*ras*)	ND	ND	5/6	8	4/6	8	4/6	6	3/6	4	4/6	4
pSV7PL-FScnc1 *myc* + pEJ.6.6 (*ras*)	0/6	0	0/6	0	0/6	0	0/6	0	0/6	0	0/6	0
pSV7PL-FSnoc1 *myc* + pEJ.6.6 (*ras*)	0/6	0	3/6	4	3/6	2	2/6	2.5	2/6	2	5/6	5
pSV7PL-FScnc2 *myc* + pEJ.6.6 (*ras*)	ND	ND	0/6	0	0/6	0	0/6	0	0/6	0	0/6	0
pSV7PL-wt *myc* + pBR327	0/6	0	0/6	0	0/6	0	0/6	0	0/6	0	0/6	0
pSV7neo + pEJ.6.6 (*ras*)	0/6	0	0/6	0	0/6	0	0/6	0	0/6	0	0/6	0

Results are presented as number of plates with foci/total number of plates and number of foci per million transfected cells. ND, not done.

sis, we generated six deletion and three frameshift mutations in the second exon of the human c-*myc* gene. Mutants Dcnc1 (*d*eletion *c-myc N-myc c*onserved no. *1*), Dcnc2, and Dcnc3 have deletions corresponding to amino acids 44–67, 129–144, and 181–197, respectively. These, in turn, correspond to the three conserved regions between human c-*myc* and N-*myc* (see Fig. 1*A* and diagram in Fig. 2). Mutants Dhcc1 (*d*eletion *h*uman c-*myc c*hicken c-*myc c*onserved no. *1*) and Dhcc2 have deletions in regions conserved between human and chicken c-*myc*. Amino acids 6–21 and 94–109, respectively, are deleted in these mutants. Finally, mutant D100 is a deletion mutant of the first hundred amino acids of c-*myc*, which includes the first conserved region between c-*myc* and N-*myc*. Mutants FScnc1 (*f*rame*s*hift *c-myc N-myc c*onserved no. *1*) and FScnc2 have frameshift mutations of amino acids 41–62 and 99–138, respectively. Mutant FSnoc1 (*f*rame*s*hift *non*conserved no. *1*) is a frameshift of amino acids 62–99. These amino acids are not conserved between human c-*myc*, N-*myc*, or chicken c-*myc*. The predicted amino acid sequences of the mutated regions of the three frameshift mutants are shown in Fig. 1*B*, and all the mutant genes are diagrammatically represented in Fig. 2.

***In Vitro* Transforming Ability of the Human c-*myc* Mutants.** The mutated c-*myc* genes were cotransfected with a plasmid containing the EJ-Ha-c-*ras* gene into rat embryo fibroblasts. From the numbers of transformed foci (Table 2) we calculated the percent transforming efficiency of each mutant in relation to the wild-type gene (Fig. 2). Mutants D100, Dcnc1, Dcnc2, FScnc1, and FScnc2 were virtually nonfunctional in the transformation assay. All these mutants involved the first or second conserved regions between c-*myc* and N-*myc*. In contrast Dcnc3, the deletion mutant of the third conserved region between human c-*myc* and N-*myc*, had the highest transforming activity of all mutants (32% of wild type). Mutants Dhcc1 and Dhcc2, bearing deletions of regions conserved between human and chicken c-*myc*, but not N-*myc*, had transforming activities that were 4% and 16% of the wild type, respectively. Finally, the frameshift mutant of a nonconserved sequence, FSnoc1, had 15% of the wild-type transforming efficiency. The results are shown in comparison to our published studies of the transforming activity of the mutated c-*myc* gene of the human Burkitt lymphoma cell line Ly65 (21).

DISCUSSION

It is difficult to draw conclusions regarding mutations that eliminate transforming activity of the c-myc protein. Loss of function could come about for a variety of reasons unrelated to a requirement for a specific sequence: for example, the steric fit of the regions juxtaposed by a deletion or substituted by a pair of frameshift mutations. Thus the loss of transforming activity by mutations in the first two of the three N-terminal conserved segments does not prove that these segments are required for transforming activity. For example, mutant Dcnc2, which has a deletion of the second conserved region between human c-*myc* and N-*myc*, has lost transforming activity. On the other hand, the translocated c-*myc* gene derived from the Burkitt lymphoma cell line Ly65 has a frameshift mutation in a portion of the same conserved segment yet retains transforming activity. This discrepancy may relate to different long-range effects of the mutations on the protein tertiary structure.

Retention of activity in the face of amino acid loss or substitution has significance. When a significant mutation, deletion, or dramatic alteration of primary amino acid sequence does *not* lead to loss of function, one can argue that the mutated region does not mediate this function. Mutations of nonconserved regions or regions conserved between c-*myc* genes of different species, but not between c- and N-*myc*, generally retained transforming activity, confirming their tolerance of evolutionary drift. However, the deletion mutant of the third conserved region between c-*myc* and N-*myc* also retained transforming activity, suggesting that it does not directly contribute to the transformation process.

The fact that regions in the N-terminal half of the c-myc protein are conserved in evolution but not required for transformation offers a hint as to how function might be distributed through the structure of the c-myc protein. If c-myc is involved in cell division, it is reasonable to imagine that it acts in concert with other elements that modulate its function. If it is a regulated molecule, its N-terminal exon, for example, might encode a primarily regulatory region. Disturbing such a region might perturb the regulation of c-myc, but not necessarily its subsequent function, which could be more closely related to its action in transformation. Our experiments suggest such a division of function.

We thank Patricia Thistlethwaite for helpful discussions, Richard Woychik for reading the manuscript, and Fred Alt for making available to us the human N-*myc* sequence in advance of publication. This work was supported in part by grants from the American Business Cancer Research Foundation and E.I. Du Pont de Nemours and Co. J.S. was supported by a fellowship from the American Committee for the Weizmann Institute.

1. Collins, S. & Groudine, M. (1982) *Nature (London)* **298**, 679–681.
2. Alitalo, K., Schwab, M., Lin, C. C., Varmus, H. E. & Bishop, J. M. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 1707–1711.
3. Leder, P., Battey, J., Lenoir, G., Moulding, C., Murphy, W., Potter, H., Stewart, T. & Taub, R. (1983) *Science* **222**, 765–771.
4. Kohl, N. E., Kanda, N., Schreck, R. R., Bruns, G., Latt, S. A., Gilbert, F. & Alt, F. W. (1983) *Cell* **35**, 359–367.
5. Schwab, M., Alitalo, K., Klempnauer, K. H., Varmus, H. E.,

Bishop, J. M., Gilbert, F., Brodeur, G., Goldstein, M. & Trent, J. (1983) *Nature (London)* **305,** 245–248.
6. Little, C. D., Nau, M. M., Carney, D. N., Gazdar, A. F. & Minna, J. D. (1983) *Nature (London)* **306,** 194–196.
7. Watson, D. K., Psallidopoulos, M. C., Samuel, K. P., Dalla-Favera, R. & Papas, T. S. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3642–3645.
8. Battey, J., Moulding, C., Taub, R., Murphy, W., Stewart, T., Potter, H., Lenoir, G. & Leder, P. (1983) *Cell* **34,** 779–787.
9. Kohl, N. E., Legouy, E., DePinho, R. A., Nisen, P. D., Smith, R. K., Gee, C. E. & Alt, F. W. (1986) *Nature (London)* **319,** 73–77.
10. Stanton, L. W., Schwab, M. & Bishop, J. M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 1772–1776.
11. Watson, D. K., Reddy, E. P., Duesberg, P. H. & Papas, T. S. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 2146–2150.
12. Bernard, O., Cory, S., Gerondakis, S., Webb, E. & Adams, J. M. (1983) *EMBO J.* **2,** 2375–2383.
13. Van Beneden, R. J., Watson, D. K., Chen, T. T., Lautenberger, J. A. & Papas, T. S. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 3698–3702.
14. Beimling, P., Benter, T., Sander, T. & Moelling, K. (1985) *Biochemistry* **24,** 6349–6355.
15. Watt, R. A., Shatzman, A. R. & Rosenberg, M. (1985) *Mol. Cell. Biol.* **5,** 448–456.
16. Land, H., Parada, L. F. & Weinberg, R. A. (1983) *Nature (London)* **304,** 596–602.
17. Kelly, K., Cochran, B. H., Stiles, C. D. & Leder, P. (1983) *Cell* **35,** 603–610.
18. Persson, H. & Leder, P. (1984) *Science* **225,** 718–721.
19. Alitalo, K., Ramsay, G., Bishop, J. M., Pfeifer, S. O., Colby, W. W. & Levinson, A. D. (1983) *Nature (London)* **306,** 274–277.
20. Hann, S. R., Abrams, H. D., Rohrschneider, L. R. & Eisenman, R. N. (1983) *Cell* **34,** 789–798.
21. Murphy, W., Sarid, J., Taub, R., Vasicek, T., Battey, J., Lenoir, G. & Leder, P. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 2939–2943.
22. Zoller, M. J. & Smith, M. (1984) *DNA* **3,** 479–488.
23. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.