The ⁵' Noncoding Region of the Human Leukemia-Associated Oncogene BCRIABL Is a Potent Inhibitor of In Vitro Translation

ALEXANDER J. MULLER¹ AND OWEN N. WITTE^{2,3*}

Molecular Biology Institute,¹ Department of Microbiology,² and Howard Hughes Medical Institute,³ University of California, Los Angeles, 405 Hilgard Avenue, Los Angeles, California 90024

Received 14 June 1989/Accepted 15 August 1989

The mRNA encoding the chimeric BCR/ABL oncogene, which is transcribed from the Philadelphia chromosome in human chronic myelogenous leukemia, has a ⁵' noncoding sequence greater than 500 bases in length which is highly GC rich and contains ^a short open reading frame. This untranslated sequence has ^a dramatic inhibitory effect upon translational efficiency in vitro. However, when BCR/ABL message is expressed in certain cell types such as the NIH 3T3 cell line, the ⁵' noncoding region has little inhibitory effect on translational efficiency.

A translocation between chromosomes ⁹ and 22, which is observed in more than 90% of patients with chronic myelogenous leukemia (for a review, see reference 13) as well as in some cases of acute lymphocytic and acute myelogenous leukemia (ALL and AML, respectively) (for a review, see reference 2), results in the expression of an abnormal fusion product between the c-ABL proto-oncogene and a gene of unknown function termed BCR. Transcription of the chimeric *BCR/ABL* oncogene begins more than 500 bases upstream of the translational initiation codon, as determined by sequence analysis of cDNA clones. The overall GC content of the ⁵' noncoding (NC) region is greater than 80% and remains high $($ >75%) throughout the first 500 nucleotides of coding sequence (Fig. 1) (7, 23, 28). ⁵' NC regions over ¹⁰⁰ bases long and with high GC content are rare among mammalian messages but have been identified in a large proportion of mRNAs encoding scarce cellular proteins such as growth factors, membrane receptors, GTP-binding proteins, and many proto-oncogenes (20, 21).

Sequences with the potential to form stable secondary structure in the ⁵' NC region of RNA have been demonstrated to be translationally inhibitory both in vitro and in vivo (16, 21, 32-34). Computer analysis of ⁴⁷¹ bases of ⁵' NC sequence in conjunction with the first 100 bases of coding sequence was performed to predict an overall secondary structure of minimum free energy (Fig. 2). The calculated free energy for the predicted structure is $\Delta G^{\circ}_{37} = -301.3$ kcal/mol $(1 \text{ kcal} = 4,184 \text{ J})$ (5). This analysis suggests that multiple stable stem-loop structures have the potential to form in the BCR/ABL 5' NC region. Particularly striking is a perfect 17-base inverted repeat, made up entirely of G and C residues, at 116 to 132 and 378 to 393 bases upstream of the BCR/ABL start codon (7, 28).

Another noteworthy aspect of the BCR/ABL 5' NC region is the presence of a short 57-base open reading frame beginning 105 bases upstream of the predicted start site for the initiation of translation. Upstream AUG triplets have been found in only ⁵ to 10% of eucaryotic mRNAs (14, 20). Oncogenes again are an exception in that two-thirds of those examined have AUG codons upstream of the start of the major open reading frame (18, 20). In eucaryotes, an upstream AUG will suppress correct initiation of translation to a degree corresponding to the context in which it is found $(15, 18)$. In the case of *BCR/ABL*, both AUGs are in very

good context for initiating translation, with a purine (G) in the -3 position and a G in the $+4$ position (17, 22) (Fig. 1). However, the level of initiation from ^a downstream AUG can be substantially increased by the presence of a termina-

FIG. 1. Features of the BCR/ABL ⁵' NC region. The nucleotide sequence of the BCR/ABL ⁵' NC region was previously published (28). The next-to-rightmost column gives the position of each 50-nucleotide break in the sequence relative to the junction between the ⁵' NC sequence and the major open reading frame, which is set at zero. The GC percentage of each 50-base segment is shown in the rightmost column. Boldface numbers above the sequence designate the ⁵' ends of each of the five clones referred to in the text. Also identified in the ⁵' NC region are ^a perfect, all GC 18-base inverted repeat $(-114$ to -131 and -376 to -393) and a 57-base upstream open reading frame $(-105 \text{ to } -49)$. The first 10 amino acids coded for by the major open reading frame are given.

^{*} Corresponding author.

FIG. 2. Secondary structure predicted for the 5' NC region of the BCR/ABL oncogene. Secondary structure predictions were made by using the FOLD program (39) which determines ^a secondary structure of minimum free energy for an RNA sequence on the basis of published values of stacking and loop-destabilizing energies (50). The sequence analyzed is the ⁵' NC region of clone [-447]5'P210 in conjunction with ¹⁰⁰ nucleotides of downstream coding sequence. Seventeen nucleotides of pGEM4 vector sequence, which is present in in vitro synthesized transcripts, is also included at the 5' end. The calculated free energy for the entire structure is $\Delta G^{\circ}_{37} = -301.3$ kcal/mol. Nucleotide positions are numbered at 50-base intervals. The first nucleotide of coding sequence is at position $+1$, and the first upstream, noncoding nucleotide is at position - 1. The ⁵' endpoint positions for the five clones referred to in the text are bracketed and numbered relative to the junction between the noncoding and coding sequence, which is designated [0]5'. The two long stretches of inverted repeat sequence composed of the tandemly repeated triplets GGC and CCG are predicted by this model to form ^a perfect 17-base-pair structure between nucleotides -393 to -378 and -116 to -132 . The calculated free energy of this 17-base-pair duplex considered in isolation is $\Delta G^{\circ}_{37} = -41$ kcal/mol (5).

FIG. 3. Effect of the BCR/ABL 5' NC region on the efficiency of in vitro translation of synthetic message. (A) In vitro SP6-mediated transcription was carried out as previously described (27). Capping of the RNA was achieved by including the cap analog m7GpppG at 0.5 mM and lowering the concentration of GTP to 0.1 mM (3, 8, 37). Each reaction was monitored by separately incubating 5 μ l (1/10 volume) in the presence of 10 μ Ci of [α ⁻³²P]GTP. Labeled transcript generated in this way was run on a 1% agarose gel. Lanes: 1, [-447]5'P100; 2, [-113]5'P100; 3, [-94]5'P100; 4, [0]5'P100; 5, [+40]5'P100. Autoradiography was for 10 min at room temperature. (B) SP6-generated transcript was incubated in rabbit reticulocyte lysate (Bethesda Research Laboratories) as recommended by the manufacturer. Translation product was labeled by incorporation of $[35S]$ methionine (50 μ Ci/30- μ l reaction) and analyzed on a sodium dodecyl sulfate-7% polyacrylamide gel. Lanes: 1, no RNA template; 2, [-447]5'P100; 3, [-113]5'P100; 4, [-94]5'P100; 5, [0]5'P100; 6, [+40]5'P100. Exposure was for ² ^h at -70°C. (C) Lanes ¹ to 4 of panel B were exposed for 20 h.

FIG. 4. Effect of the BCR/ABL 5' NC region on translation in cultured NIH 3T3 cells. (A) High-molecular-weight DNA (10 μ g) extracted from duplicate plates of NIH 3T3 cells 48 h after viral infection and digested with EcoRI was analyzed by blot hybridization following fractionation on a 0.8% agarose gel (36). BCR/ABL and c-ABL sequences were detected by hybridization with a 0.65 kilobase PstI fragment of ABL specific to the kinase domain (28), labeled to greater than 10^8 cpm/ μ g by random priming (Boehringer Mannheim Biochemicals). Infections were as follows: lanes ¹ and 2, [-447]5'P210; lanes ³ and 4, [0]5'P210; lanes 5 and 6, vector alone. Autoradiography was at -70° C with an intensifying screen for 6 days. (B) Total RNA (20 μ g) extracted from duplicate plates of NIH 3T3 cells 48 h after viral infection was run on a 0.8% agarose gel for blot hybridization (36). BCR/ABL RNA was detected by hybridization with a 0.65-kilobase PstI fragment of ABL specific to the kinase domain (28), labeled to greater than 10^8 cpm/ μ g by random priming (Boehringer Mannheim Biochemicals). Infections were as follows: lanes ¹ and 2, [-447]5'P210; lanes ³ and 4, [0]5'P210; lanes ⁵ and 6,

tion codon between the AUGs as is present at position -51 in the BCR/ABL ⁵' NC sequence (19, 24, 25, 30, 31) (Fig. 1).

Five clones of the *BCR/ABL* oncogene, originally derived from ^a K562 cDNA library (28) and extending to different ⁵' positions in the message (as shown in Fig. 1), were used to study the effects of the 5' NC region on translation. These clones were incorporated into the vector pGEM4 downstream of the SP6 initiation site. The DNA templates used in the transcription reaction were truncated at the HindlIl site near the BCR/ABL junction so that subsequent translation yielded a polypeptide of approximately 100 kilodaltons rather than the complete 210-kilodalton protein. Results obtained with these 3'-truncated clones were consistent with other experiments in which full-length BCR/ABL message was used (data not shown).

The efficiency of the transcription reactions was compared by adding $[\alpha^{-3}P]GTP$ to $1/10$ volume of each reaction mixture and running the labeled transcript on an agarose gel (Fig. 3A). This was done to verify that the transcripts generated were full length. In vitro translation of the [0]5'P100 transcript was very efficient (Fig. 3B, lane 5). The quantity of translation product generated was comparable to what was obtained by using c-abl as well as other, unrelated mRNAs (data not shown). The [+40]5'P100 message was translated at an efficiency comparable to that of [0]5'P100; however, a smaller protein product was generated as was predicted for initiation of translation downstream of the normal start codon (Fig. 3B, lane 6). Translation product from the $[-447]5'P100$ message was undetectable, even at a 10-fold-longer exposure (Fig. 3C, lane 2). Translation of [-113]5'P100 and [-94]5'P100 message was about 200-fold lower than that of [0]5'P100 (Fig. 3C, lanes 3 and 4) as determined by comparison to a dilution series of [0]5'P100 translation product (not shown). The similarities in translational efficiency shared between these two clones suggest that the upstream open reading frame is not contributing significantly to the inhibition observed in vitro. The $BCR/$ ABL 5' NC region did not act in trans to decrease translational efficiency and did not induce differential message stability (unpublished results). These results suggest that the ⁵' NC region acts in cis to directly inhibit translation. possibly through the formation of stable secondary structure.

Previous work has indicated that the effect of the BCR/ ABL 5' NC region on translation in vivo might be more complex than what was observed in vitro. Efficient expression of the BCR/ABL P210 oncogene was observed both in

vector alone. Autoradiography was at -70° C with an intensifying screen for ² days. (C) Protein extracted from duplicate plates of NIH 3T3 cells 48 h after viral infection was immune precipitated (12) with excess rabbit anti-pEX5 antibody (9) which is reactive with carboxyterminal sequences of the ABL protein segment. Protein was labeled by autophosphorylation in the presence of 15 μ Ci of [γ -³²P]ATP in a 45-pI reaction volume (12). Infections were as follows: lanes ¹ and 2, [-447]5'P210; lanes 3 and 4, [0]5'P210; lanes 5 and 6, vector alone. Autoradiography was at room temperature for ¹ h. (D) At 48 h after viral infection, duplicate plates of NIH 3T3 cells were incubated for ³ h in methionine-free medium containing 0.1 mCi of [³⁵S]methionine (TransLabel) per ml and were then extracted for protein (11). Lysates were normalized for trichloroacetic acidprecipitable counts, and two cycles of immune precipitation were performed (12), first with excess rabbit anti-pEX5 antibody and then with excess rabbit anti-pEX2 (9). Infections were as follows: lanes 1 and 2, [-447]5'P210; lanes 3 and 4, [0]5'P210; lanes 5 and 6, vector alone. Fluorography was at -70° C for 10 days (1).

human leukemia cells and in established Ph-positive, chronic myelogenous leukemia-derived cell lines such as K562 (10- 12). Fibroblast and lymphoid cells infected with recombinant retrovirus carrying the BCR/ABL gene were also found to express P210 despite the presence of ⁵' NC sequence (4, 26). However, attempts to overexpress the *BCR/ABL* oncogene in the insect cell line S_F9 by using a baculovirus expression system indicated that in this cell type the BCR leader sequence strongly inhibited translation. In S_F9 cells expressing the [-113]5'P210 and [0]5'P210 constructs, the presence of ⁵' NC sequence resulted in ^a dramatic (greater than 100-fold) decrease in the amount of recoverable BCR/ABL P210 (35). These data suggested that the effect of the BCR 5' NC region on in vivo translation could vary markedly between cell types.

To quantitatively evaluate the effect of the BCR 5' NC region on the efficiency of translation in NIH 3T3 cells, recombinant retroviruses harboring [-447]5'P210 and [0]5'P210 were employed in an acute infection assay (26). Normalized levels of each virus were used to set up six identical infections. Two of these plates were separately extracted for duplicate analysis of both DNA and RNA. The level of integrated *BCR*/*ABL* sequence was nearly equivalent for the two viruses (Fig. 4A). RNA analysis revealed that the relative amounts of message produced were also nearly equivalent for the two viruses (Fig. 4B). The four remaining plates in each set were extracted in duplicate for determination of the steady-state level of P210 protein both by autophosphorylation of the extracted protein in the presence of $[\gamma^{-32}P]ATP$ and following in vivo labeling with $[^{35}S]$ methionine. In both cases, the level of P210 protein expressed from the $[-447]5'P210$ construct was approximately twofold lower than from the [0]5'P210 construct (Fig. 4C and D). Although reproducible, this effect of the BCRIABL ⁵' NC sequence in mass populations of NIH 3T3 cells is much less dramatic than the suppression of translation observed in vitro.

The failure of BCR/ABL ⁵' NC sequence to effectively suppress translation in NIH 3T3 cells suggests that these cells possess the capacity to actively relieve translationally inhibitory secondary structure in the ⁵' NC region. Control over this capacity would provide a means of regulating gene expression. Little data is currently available on the expression of the normal BCR protein in different cell types; however, the expression of c-*myc*, which has a similar long, GC-rich ⁵' NC region, appears to be translationally regulated during Xenopus oocyte development. Two separate studies have demonstrated that the rate of synthesis of c-mycencoded protein varies substantially during oogenesis without a corresponding change in the level of message (6, 38). However, the c-myc ⁵' NC region also has little effect on protein expression in a variety of established cell lines, including NIH 3T3 cells (29). To better understand the role, then, that the BCR/ABL 5' NC region might play in the progression of leukemia, it will be important to study the effects of this region on translational expression in specific hematopoietic cell types.

ACKNOWLEDGMENTS

We gratefully acknowledge Jami McLaughlin for important technical contributions, Carol Crookshank for secretarial assistance, and Mike Cohan for photographic work. We also thank A. Berk, A. Dasgupta, R. Simons, and W. Wickner for critical reading of the manuscript.

This work was supported by Public Health Service grants from the National Cancer Institute to O.N.W. and by U.S. Public Health Service National Research Service Award GM07185 from the National Institutes of Health to A.J.M. O.N.W. is an Investigator of the Howard Hughes Medical Institute.

LITERATURE CITED

- 1. Chamberlain, J. P. 1979. Fluorographic detection of radioactivity in polyacrylamide gels with the water soluble fluor, sodium salicylate. Anal. Biochem. 98:132-135.
- 2. Clark, S. C., W. M. Crist, and 0. N. Witte. 1989. Molecular pathogenesis of Ph-positive leukemias. Annu. Rev. Med. 40: 113-122.
- 3. Contreras, R., H. Cheroutre, W. Degrave, and W. Fiers. 1982. Simple, efficient in vitro synthesis of capped RNA useful for direct expression of cloned eukaryotic genes. Nucleic Acids Res. 10:6353-6362.
- 4. Daley, G. Q., J. McLaughlin, 0. N. Witte, and D. Baltimore. 1987. The CML-specific P210 berlabl protein, unlike v-abl, does not transform NIH3T3 fibroblasts. Science 237:532-535.
- 5. Freier, S. M., R. Kierzek, J. A. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson, and D. H. Turner. 1986. Improved free-energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. USA 83:9373-9377.
- 6. Godeau, F., H. Persson, H. E. Gray, and A. B. Pardee. 1986. c-myc expression is dissociated from DNA synthesis and cell division in Xenopus oocyte and early embryonic development. EMBO J. 5:3571-3577.
- 7. Hariharan, I. K., and J. M. Adams. 1987. cDNA sequence for human bcr, the gene that translocates to the *abl* oncogene in chronic myeloid leukaemia. EMBO J. 6:115-119.
- 8. Konarska, M. M., R. A. Padgett, and P. A. Sharp. 1984. Recognition of cap structure of splicing in vitro of mRNA precursors. Cell 38:731-736.
- 9. Konopka, J. B., R. Davis, S. M. Watanabe, R. Ponticelli, L. Schiff-Maker, and 0. N. Witte. 1984. Only site-directed antibodies reactive with the highly conserved src homologous region of the v-abl protein neutralize kinase activity. J. Virol. 51:223-232.
- 10. Konopka, J. B., S. M. Watanabe, J. W. Singer, S. J. Collins, and 0. N. Witte. 1985. Cell lines and clinical isolates derived from Ph'-positive chronic myelogenous leukemia patients express c-abl proteins with ^a common structural alteration. Proc. Natl. Acad. Sci. USA 82:1810-1814.
- 11. Konopka, J. B., S. M. Watanabe, and 0. N. Witte. 1984. An alteration of the human c-abl protein in K562 leukemia cells unmasks associated tyrosine kinase activity. Cell 37:1035-1042.
- 12. Konopka, J. B., and 0. N. Witte. 1985. Detection of c-abl tyrosine kinase activity in vitro permits direct comparison of normal and altered abl gene products. Mol. Cell. Biol. 5: 3116-3123.
- 13. Konopka, J. B., and 0. N. Witte. 1985. Activation of the abl oncogene in murine and human leukemias. Biochim. Biophys. Acta 823:1-17.
- 14. Kozak, M. 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol Rev. 47: $1 - 45.$
- 15. Kozak, M. 1984. Selection of initiation sites by eucaryotic ribosomes: effect of inserting AUG triplets upstream from the coding sequence for preproinsulin. Nucleic Acids Res. 12: 3873-3893.
- 16. Kozak, M. 1986. Influences of mRNA secondary structure on initiation by eukaryotic ribosomes. Proc. Natl. Acad. Sci. USA 83:2850-2854.
- 17. Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44:283-292.
- 18. Kozak, M. 1986. Bifunctional messenger RNAs in eukaryotes. Cell 47:481-483.
- 19. Kozak, M. 1987. Effects of intercistronic length on the efficiency of reinitiation by eukaryotic ribosomes. Mol. Cell. Biol. 7: 3438-3445.
- 20. Kozak, M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. 15:8125- 8148.
- 21. Kozak, M. 1988. Leader length and secondary structure modu-

late mRNA function under conditions of stress. Mol. Cell. Biol. 8:2737-2744.

- 22. Kozak, M. 1989. The scanning model for translation: an update. J. Cell Biol. 108:229-241.
- 23. Lifshitz, B., E. Fainstein, C. Marcelle, E. Shtivelman, R. Amson, R. P. Gale, and E. Canaani. 1988. bcr genes and transcripts. Oncogene 2:113-117.
- 24. Liu, C.-C., C. C. Simonsen, and A. D. Levinson. 1984. Initiation of translation at internal AUG codons in mammalian cells. Nature (London) 309:82-85.
- 25. Marth, J. D., R. W. Overall, K. E. Meier, E. G. Krebs, and R. M. Perlmutter. 1988. Translational activation of the lck proto-oncogene. Nature (London) 332:171-173.
- 26. McLaughlin, J., E. Chianese, and 0. N. Witte. 1987. In vitro transformation of immature hematopoietic cells by the P210 BCR/ABL oncogene product of the Philadelphia chromosome. Proc. Natl. Acad. Sci. USA 84:6558-6562.
- 27. Melton, D. A. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. 12:7035-7056.
- 28. Mes-Masson, A.-M., J. McLaughlin, G. Q. Daley, M. Paskind, and 0. N. Witte. 1986. Overlapping cDNA clones define the complete coding region for the $P210^{\circ}$ able product associated with chronic myelogenous leukemia cells containing the Philadelphia chromosome. Proc. Natl. Acad. Sci. USA 83:9768- 9772.
- 29. Parkin, N., A. Darveau, R. Nicholson, and N. Sonenberg. 1988. cis-acting translational effects of the ⁵' noncoding region of c-myc mRNA. Mol. Cell. Biol. 8:2875-2883.
- 30. Peabody, D. S., and P. Berg. 1986. Termination-reinitiation occurs in the translation of mammalian cell mRNAs. Mol. Cell.

Biol. 6:2695-2703.

- 31. Peabody, D. S., S. Subramani, and P. Berg. 1986. Effect of upstream reading frames on translation efficiency in simian virus 40 recombinants. Mol. Cell. Biol. 6:2704-2711.
- 32. Pelietier, J., G. Kaplan, V. R. Racaniello, and N. Sonenberg. 1988. Translational efficiency of poliovirus mRNA: mapping inhibitory cis-acting elements within the ⁵' noncoding region. J. Virol. 62:2219-2227.
- 33. Pelietier, J., and N. Sonenberg. 1985. Insertion mutagenesis to increase secondary structure within the ⁵' noncoding region of a eukaryotic mRNA reduces translational efficiency. Cell 40: 515-526.
- 34. Pelletier, J., and N. Sonenberg. 1987. The involvement of mRNA secondary structure in protein synthesis. Biochem. Cell Biol. 65:576-581.
- 35. Pendergast, A. M., R. Clark, E. S. Kawasaki, F. P. McCormick, and 0. N. Witte. 1989. Baculovirus expression of functional P210 BCR-ABL oncogene product. Oncogene 4:759-766.
- 36. Schwartz, R. C., L. W. Stanton, S. C. Riley, K. B. Marcu, and 0. N. Witte. 1986. Synergism of v-myc and v-Ha-ras in the in vitro neoplastic progression of murine lymphoid cells. Mol. Cell. Biol. 6:3221-3231.
- 37. Shatkin, A. J. 1978. Capping of eucaryotic mRNAs. Cell 9: 645-653.
- 38. Taylor, M. V., M. Gusse, G. I. Evan, N. Dathan, and M. Mechali. 1986. Xenopus myc proto-oncogene during development: expression as ^a stable maternal mRNA uncoupled from cell division. EMBO J. 5:3563-3570.
- 39. Zuker, M., and P. Stiegler. 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. 9:133-148.