Human *c-myc* gene contains a regulatory site similar to consensus of interferon response sequence (IRS)

N.M. Alexandrova, A.V. Itkes, L.R. Imamova, B.K. Chernov, E.M. Tulchinsky, N.B. Ulyanov and L.L. Kisselev

Engelhardt Institute of Molecular Biology, USSR Academy of Sciences, Moscow 117984, USSR

Received 11 April 1990

Expression of *c-myc* proto oncogene is regulated by multiple mechanisms. Here, we report that the consensus of the regulatory region of interferon-dependent genes, GGAAAN$_{1-3}$ GAAA, was found after computer search in the 5'-terminal flank of human *c-myc* gene in position (−76:−67). In vitro transcription of *c-myc* gene fragments showed that the consensus region competes with oligonucleotide GGGAAAATGAAACT for binding to specific protein(s). This oligonucleotide was shown to bind selectively the interferon-dependent positive transcription factor [1]. Transcription of *c-myc* fragments lacking 5'-terminal region up to positions −101 or +71 was initiated at two sites located in the first intron. These sites did not coincide with P_1 in vivo RNA cap-site. Binding of the protein factor(s) to the regulatory region of *c-myc* gene −76:−67 blocked the in vitro transcription initiated in the first intron.

c-myc gene; Interferon-dependent gene; Transcription

1. INTRODUCTION

At present, the gene family induced by α-, β- or γ-interferons (IFN) has been described [2]. These genes contain a similar nucleotide sequence in the 5'-terminal nontranscribing regions presumably involved in the IFN-dependent activation of transcription. Originally this sequence was detected in the HLA and metallothionein genes [3], and then in the number of IFN-dependent genes; it was termed as interferon response sequence (IRS). In addition, IRS consensus was found also in the regulatory part of β-IFN gene (Table I).

Table I

The sequence of interferon-dependent genes in the IRS (interferon response sequence) region

Gene (human)	Sequence	Ref.
F/S consensus	GA$^{G}_{A}$A GGA GAAA CT	3
Factor B	GGAAA CA- GAAA CT	4
ISG15	GGAAA CC- GAAA CT	5
ISG54	GGAAA GT- GAAA CT	5,6
IFI56	GGAAA CT- GAAA GG	6
IP-10	GGAAA GT- GAAA CT	7
6-16	GGAAA AT- GAAA CT[a]	1
2',5' oligo-A synthetase	GGAAA C- - GAAA CC	8,9
β-IFN	GGAAA ACT GAAA G	10
Consensus	GGAAA N_{1-3} GAAA	

[a] This sequence is present twice in the gene 6–16

Correspondence address: N.M. Alexandrova, Engelhardt Institute of Molecular Biology, USSR Academy of Sciences, Moscow 117984, USSR

The first goal of this investigation was a computer search of databanks to find regions of other human genes that contain the IRS consensus. The IRS consensus proposed here (Table I) had been used. These sequences were found in a few genes (see below) including *c-myc* gene. Hence, the second goal of the present work was to analyze functions of *c-myc* gene region containing IRS consensus.

2. MATERIALS AND METHODS

2.1. *Computer search for IRS consensus*

The GGAAAN$_{1-3}$GAAA consensus was searched in the 10th release of the EMBL nucleotide databank (1986) and in the first release of the USSR 'Gene-Express' databank (1988) among human sequences.

2.2. *Plasmid construction and isolation*

Plasmid pUMC1 (Fig. 1) containing human *c-myc* −101:1185 region was constructed by inserting the *SmaI/XbaI* fragment of 10 kbp human *c-myc* plasmid [11] into the pUSVLcat vector (kind gift of Dr Archipova, Inst. of Molecular Biology, Acad. Sci. USSR, Moscow). Plasmids were isolated by the standard technique [12]; the DNA isolated was purified by centrifugation in CsCl gradient.

2.3. *HeLa S3 cell extract*

The total extract of HeLa S3 cells was prepared as described [13].

2.4. *Transcription*

Transcription of linearized plasmid DNA in vitro was carried out as described by Hames and Higgins [14]. RNA transcripts were isolated using guanidine isothiocyanate [15] and analyzed by electrophoresis in agarose gel in the presence of glyoxal [16]. The incubation mixture contained (if mentioned) the double-stranded oligonucleotides GGGAAAATGAAACT ('normal') or GGGAAAATGGACACT ('mutant'). Both oligonucleotides were added into

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

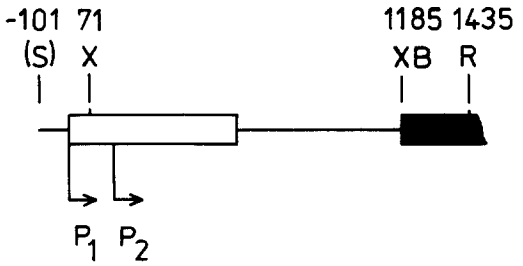

Fig. 1. Scheme of hybrid *c-myc*/CAT plasmid pUMC1. Empty box indicates the first exon of the human *c-myc* gene, and solid box stands for CAT gene (fragment of pUSVLcat vector). Thin lines indicate the first *c-myc* intron and the upstream noncoding region of the gene. Restriction sites are marked as follows: S, *Sma*I; X, *Xho*I; XB, *Xba*I; R, *Eco*RI. Arrows indicate the initiation sites of transcription controlled by P_1 and P_2 promoters. The nucleotide positions for *c-myc* are defined relative to P_1 site.

the mixture at 10- or 100-fold molar excess over plasmid DNA content.

2.5. *Synthesis of oligonucleotides*

Oligonucleotides were synthesized by solid phase phosphoramidite method using Beckman System I plus DNA synthesizer at 0.2 μmol scale and purified by reversed-phase HPLC on Ultrasphere-ODS column.

3. RESULTS AND DISCUSSION

3.1. *Computer search for IRS consensus among the human gene sequences*

We have found 9 sites matching the GGAA-AN_{1-3}GAAA consensus in the human sequences available to us; these sites have been located in regulatory regions of the following genes: *c-myc* [17], thyroglobulin [18], IGF-1 [19], prointerleukin-1β [20], ϵ-globin [21], preproencephalin [22], β-globin [23], H_χ immunoglobulin chain [24], prolactin [25]. Besides, these sites were found in a number of genes of IFN-inducible proteins and β-IFN (this result is in good agreement with the data of Table I).

3.2. *Location of IRS consensus sequence in the human* c-myc *gene*

The IRS consensus in human *c-myc* was positioned in nontranscribed part of the gene, −76: −67 (+1 position means the cap-site corresponding to P_1 promoter; see [26]). The sequence of this region is the following:

```
            −80                    −60
gene:       TGGGGGAAAAGAAAAAAGATC
consensus:      GGAAANGAAA
```

In addition, 3 similar sites with one mismatch were found in the transcribing part of the gene, namely, in first exon and first intron; their positions are 100:110, 766:777 and 1574:1585.

3.3. *In vitro transcription of the* c-myc-*containing plasmids*

To investigate a potential regulatory function(s) of the IRS-like sequences of the *c-myc* gene we have used a technique of in vitro transcription with HeLa S3 cell extract; plasmid pUMC1 (Fig. 1) was a template DNA.

We run in vitro transcription either with or without the competing oligonucleotide, GGGAAAATGAAA-CT (IRS consensus is underlined). As it was shown, this oligonucleotide binds to the human IFN-dependent transcription factor [1]; moreover, similar GGGAA-AATGA*CACT oligonucleotide with the single mismatch (shown by asterisk) does not bind to this factor. We have used both oligonucleotides referred here as 'normal' and 'mutant', respectively.

Plasmid pUMC1 was cleaved with *Xba*I restriction endonuclease to linearize DNA for transcription. In the absence of the competing oligonucleotide in incubation mixture the RNA synthesized was nonspecific, probably due to random initiation of RNA synthesis. Contrary to this, addition of the oligonucleotide to the incubation mixture, resulted in appearance of specific RNA products (Fig. 2); when 100-fold excess of the oligonucleotide over plasmid DNA was added, only the specific initiation of RNA synthesis was observed (Fig. 2): two main products were detected of about 410 and 300 nucleotides long. It is noteworthy that the activity of the 'mutant' oligonucleotide was about 10-fold lower than the 'normal' one.

Thus, addition of the specific oligonucleotide

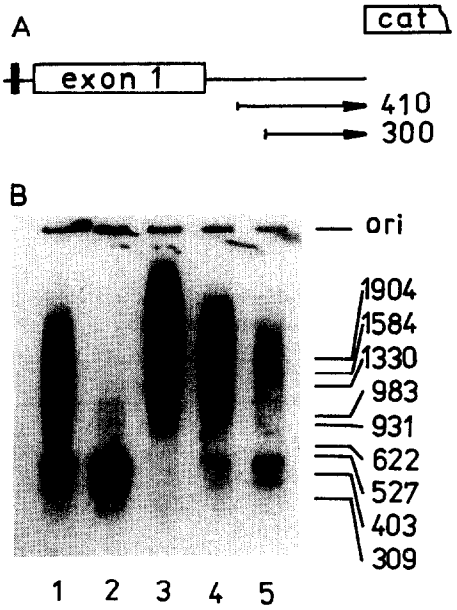

Fig. 2. In vitro transcription of the pUMC1 plasmid cut with *Xba*I. (A) Scheme of the transcribed DNA and RNA (length of RNA is marked); solid box means the IRS consensus, arrows show the RNA obtained. (B) Electrophoresis of RNA synthesized (in 1.5% agarose gel); 10-fold (line 1), 100-fold (line 2) excess of the ('normal') oligonucleotide, without added oligonucleotide (line 3), 10-fold (line 4), 100-fold (line 5) excess of 'mutant' oligonucleotide. Positions of marker DNA fragments (*Eco*RI-*Hin*dIII digest of λ phage DNA and *Msp*I one of pBR-322 plasmid DNA) are shown.

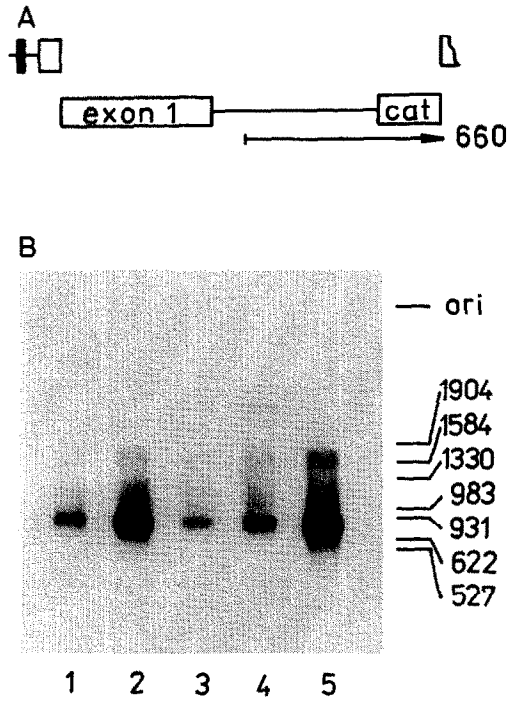

Fig. 3. In vitro transcription of the *c-myc*/CAT *Xho*I-*Eco*RI fragment of pUMC1 plasmid. (A) Scheme of the transcribed DNA and RNA (length of the RNA is marked); solid box is the IRS consensus, arrow shows the RNA obtained. (B) Electrophoresis of RNA synthesized (in 1.5% agarose gel); 10-fold (line 1), 100-fold (line 2) excess of the ('normal') oligonucleotide, without added oligonucleotide (line 3), 10-fold (line 4), 100-fold (line 5) excess of 'mutant' oligonucleotide. Marker DNA is shown (see legend to Fig. 2).

presumably causes the redistribution of the transcription factor(s) that leaves the *c-myc* regulatory regions and binds to the oligonucleotide competitor. 'Opening' of the *c-myc* regulatory regions causes initiation of RNA transcription at the specific initiation sites. However, these sites do not coincide with the known transcription starts, P_1 and P_2 [26]. These new sites, referred as I_1 and I_2, are situated in the first intron of human *c-myc* gene ca. 410 and 300 nucleotides upstream of the *Xba*I site (Fig. 2).

To check the assumption that binding of a transcription factor(s) to −76: −67 site of *c-myc* takes place, and to confirm that the synthesized RNA was initiated at the I_1 site rather than being a product of a premature termination, we have also used the pUMC1 fragment 71:1435 (*Xho*I/*Eco*RI). In this experiment it turned out that the main RNA product is ca. 660 nucleotides long (Fig. 3); hence, the RNA is in fact initiated at I_1 site (its position is ca. 775) and terminated in a standard run-off manner. The oligonucleotide addition caused an enhancement of transcription (Fig. 3), but RNA product was synthesized both with and without competitive oligonucleotides. Thus, if the −76: −67 fragment is present, the specific transcription becomes possible after dissociation of negative regulatory factor(s). If this fragment is deleted, the transcription becomes constitutive. This finding has confirmed the interpretation proposed above our data. An enhancement of transcription of the DNA fragment lacking the −76: −67 sequence (Fig. 3 lines 1, 2, 4, 5) may be caused by 100:110 region of the gene; factor(s) that associates with the IRS consensus and affects the transcription could presumably bind (see above) to this region.

Finally, we believe that the data presented provide good evidence that the IRS-like sequence of human *c-myc* gene makes it possible to affect its expression via binding of regulatory proteins including IFN-dependent factor(s).

Acknowledgements: We express our gratitude to Dr L.G. Nikolaev for providing some HeLa S3 cell extract and helpful discussion. We thank the Center of Information of the Institute of Molecular Genetics of the USSR Academy of Sciences for help on the search of the 'Gene Express' data bank.

REFERENCES

[1] Porter, A.C.G., Chernajovsky, Y., Dale, T.C., Gilbert, C.S., Stark, G.R. and Kerr, I.M. (1988) EMBO J. 7, 85–92.
[2] Revel, M. and Chebath, J. (1986) Trends Biochem. Sci. 11, 166–170.
[3] Friedman, R.L. and Stark, G.R. (1985) Nature 314, 637–639.
[4] Wu, L.-C., Morley, B.J. and Campbell, K.D. (1987) Cell 48, 331–342.
[5] Reich, N., Evans, B., Levy, D., Fahey, D., Knight, E. and Darnell, J.E., Jr. (1987) Proc. Natl. Acad. Sci. USA 84, 6394–6398.
[6] Wathelet, M.G., Clauss, I.M., Content, J. and Huez, G.A. (1988) FEBS Lett. 231, 164–171.
[7] Luster, A.D. and Ravetch, J.V. (1987) Mol. Cell. Biol. 7, 3723–3731.
[8] Rutherford, M.N., Hannigan, G.E. and Williams, B.R.G. (1988) EMBO J. 7, 751–759.
[9] Cohen, B., Peretz, D., Vaiman, D., Benech, P. and Chebath, J. (1988) EMBO J. 7, 1411–1419.
[10] Goodbourn, S., Zinn, K. and Maniatis, T. (1985) Cell 41, 509–520.
[11] Dalla-Favera, R., Gelmann, E.P., Martinotti, S., Franchini, G., Papas, T.S., Gallo, R.C. and Wong-Staal, F. (1982) Proc. Natl. Acad. Sci. USA 79, 6497–6501.
[12] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, pp. 100–102, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
[13] Sugden, B. and Keller, W. (1973) J. Biol. Chem. 248, 3777–3788.
[14] Hames, B.D. and Higgins, S.J. (1985) Transcription and Translation, A Practical Approach, p. 91, IRL Press, Oxford.
[15] Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156–159.
[16] McMaster, G.K. and Carmical, G.C. (1977) Proc. Natl. Acad. Sci. USA 74, 4835–4838.
[17] Watt, R., Nishikura, K., Sorrentino, J., Ar-Rushdi, A., Croce, C.M. and Rovera, G. (1983) Proc. Natl. Acad. Sci. USA 80, 6307–6311.
[18] Christophe, D., Cabrer, B., Bacolla, A., Targovnik, H., Pohl, V. and Vassart, G. (1985) Nucleic Acids Res. 13, 5127–5144.
[19] Dull, T.J., Gray, A., Hayflick, J.S. and Ullrich, A. (1984) Nature 310, 777–781.

[20] Clark, B.D., Collins, K.L., Gandy, M.S., Webb, A.C. and Auron, P.E. (1986) Nucleic Acids Res. 14, 7897–7914.
[21] Li, Q., Powers, P.A. and Smithies, O. (1985) J. Biol. Chem. 260, 14901–14910.
[22] Gahmberg, C.G. (1983) EMBO J. 2, 223–227.
[23] Prchal, J.T., Cashman, D.P. and Kan, Y.W. (1986) Proc. Natl. Acad. Sci. USA 83, 24–27.
[24] Rabbits, T.H., Forster, A., Baer, R. and Hamlyn, P.H. (1983) Nature 306, 806–809.
[25] Truong, A.T., Duez, C., Belayew, A., Renard, A., Pictet, R., Bell, G.I. and Martial, J.A. (1984) EMBO J. 3, 429–437.
[26] Leder, P., Battey, J., Lenoir, G., Moulding, G., Murphy, W., Potter, H., Stewart, T. and Taub, R. (1983) Science 222, 765–771.