Volume 321, number 2,3, 233–236 FEBS 12357 April 1993

The ZNF35 human zinc finger gene encodes a sequence-specific DNA-binding protein

Gina Pengue, Paola Cannada-Bartoli and Luigi Lania

Dipartimento di Genetica, Biologia Generale e Molecolare, University of Naples, Via Mezzocannone 8, 80134 Naples, Italy

Received 28 January 1993

We developed a rapid method to determine DNA-binding sites for putative DNA-binding proteins. This procedure has been successfully used to define a specific consensus site for the human ZNF35 zinc finger gene. ZNF35 encodes a 58-kDA polypeptide containing 11 consecutive finger motifs located at the amino terminus, and an acidic domain located at the carboxy terminus. These features suggest that ZNF35 is a site-specific DNA-binding protein involved in the regulation of gene expression. We have expressed the ZNF35 protein from *E. coli* and have employed a Southwestern-polymerase chain reaction method using random oligonucleotides to identify its high-affinity binding site. The core sequence for the ZNF35 protein-binding site is 5′-C/GC/GAAG/TA-3′.

DNA-binding protein; Zinc finger protein; DNA-binding site

1. INTRODUCTION

In recent years a large number of DNA-binding proteins have been cloned based on evolutionary conserved homologies in their DNA binding domain, e.g. homeo domains [1], pou-boxes [2], paired boxes [3,4] and zinc finger regions [5–9]. However, the DNA sequences to which the various domains bind have been determined only for relatively few proteins [10–14]. Our group has cloned a number of human genes encoding zinc finger domains from both human placenta and T cell cDNA libraries [8,15]. Some of these genes have been shown to be involved in regulation of myeloid cell differentiation [8,15].

We designed the Random Oligonucleotide Blot (ROB) procedure in order to determine the DNA recognition sites for these zinc finger proteins. The method required the production of a bacterial synthesized human zinc finger protein and a pool of randomized oligonucleotides. Template oligonucleotides were chemically synthesized to contain random nucleotides surrounded by specific sequences with appropriate restriction sites on either end. The primers (5′, primer A and 3′, primer B), complementary to the sequences at the ends of the template oligonucleotides, were used for PCR amplification. In this report we describe a simple method with which the double-stranded oligonucleotides containing random nucleotides can be used in a Southwestern experiment for determining high-affinity binding sites for the human ZNF35 zinc finger protein.

Correspondence address. L. Lania, Dipartimento di Genetica, Biologia Generale e Molecolare, University of Naples, Via Mezzocannone 8, 80134 Naples, Italy. Fax: (39) (81) 5527950.

2. MATERIALS AND METHODS

2.1. *Plasmid construction and expression in E. coli*

To express a ZNF35 zinc finger containing fragment in *E. coli* under the control of the T7 RNA polymerase-dependent φ promoter [16], an expression plasmid was constructed by inserting a 2230 bp *Nco*I fragment derived from the full-length ZNF35 cDNA into the *Sma*I site of the plasmid vector pT7-7, via blunt ligation. The expression vector was transformed into *E. coli* strain BL21 (LysS) [16]. Cells were grown in LB medium to an OD_{600} of 0.5 and expression was induced by adding isopropyl-β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. After 2 h at 37°C, the bacteria were harvested and extracts prepared according to the procedure described for Sp1 [17].

2.2. *Random Oligonucleotide Blot (ROB) analysis*

The following oligonucleotides were used: random sequence oligonucleotide N; 5′-AGACGGATCCATTGCA-$(N)_{20}$-CTGTAGGAATTCGGA-3′; primer A; 5′-TCCGAATTCCTACAG-3′; primer B, 5′-AGACGGATCCATTGCA-3′. The oligonucleotide N was rendered double-stranded by primed synthesis using primer A and Klenow fragment, in the presence of ^{32}P-nucleotides. The labeled oligonucleotides were used in Southwestern blot hybridization as described [18]. Briefly, the bacterially produced proteins (15 μg) were electrophoresed on SDS polyacrylamide gels and electro-transfered to nitrocellulose. The nitrocellulose filters were rinsed in 5% non-fat milk in buffer A (10 mM Tris-HCl pH 8, 50 mM NaCl, 1 mM EDTA pH 8) to denature the proteins. The nitrocellulose blot and the ^{32}P random oligonucleotides were incubated overnight at 4°C in buffer A containing 50 mM NaCl. The filters were washed twice at 4°C in buffer A containing 100 mM NaCl and exposed to film for 16 h. Identical results were observed when filters were washed with 100 mM and 300 mM NaCl. The radiolabeled band was cut from the filter, the DNA was eluted, then was subjected to PCR using the A and B primers for 20 cycles with 1 U Amplitaq polymerase (Cetus). Amplification was performed at 94°C for 1 min, 42°C for 1 min, and 72°C for 2 min. The probe was purified by gel electrophoresis and quantitated by Cerenkov counting. After four cycles the PCR products were cloned in pUC8 by taking advantage of the presence of the *Eco*RI and *Bam*HI sites present in the PCR products. 26 independent plasmid inserts were sequenced.

2.3. *Gel retardation experiments*

Gel retardation assays (EMSA) were essentially performed as described [19]. Bacterial extracts from induced cells were mixed with radiolabeled 45 bp *Eco*RI–*Bam*HI fragments derived from the indicated ROB selected clones; competition experiments were performed using 100-fold molar excess of the same unlabeled fragments. Oligonucleotides for SP1 and AP1 binding sites were used as unspecific competitors.

3. RESULTS

The predicted amino acid sequence of ZNF35 contains an acidic region at the amino terminus and eleven contiguous zinc fingers of the Cys_2–His_2 class at the carboxy terminus [9,20]. The presence of these two structural motifs implies that the ZNF35 protein is a sequence-specific nucleic acid-binding protein that may function as a transcriptional regulator. In this report we have initiated a biochemical characterization of the ZNF35 protein by identifying a DNA target sequence which it recognizes. Our strategy was to express the ZNF35 protein in *E. coli* and to use this protein in Southwestern blot experiments for the isolation of a DNA target sequence. The ZNF35 cDNA was cloned in the bacterial expression vector pT7-7 as described in section 2, and the bacterial protein (termed bZNF35) migrated with an apparent molecular size of 58,000 kDa on SDS-polyacrylamide gels (Fig. 1A). Southwestern

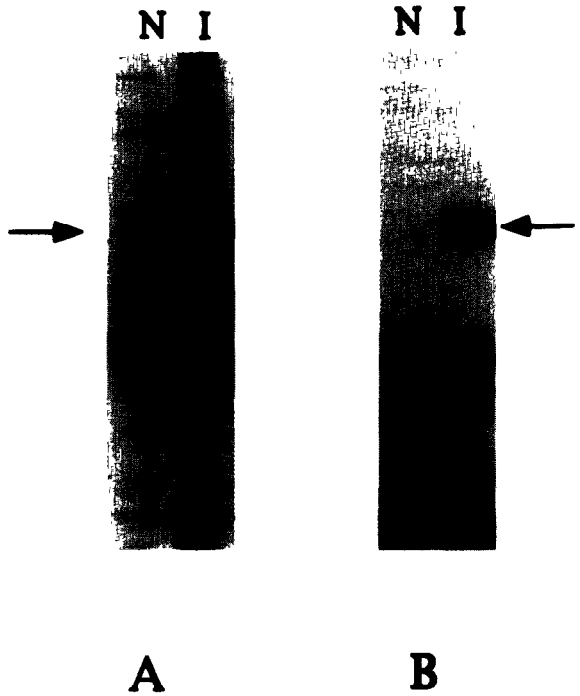

Fig. 1. SDS-PAGE and Southwestern analysis of bZNF35 protein. (A) Coomassie blue-stained SDS-PAGE; lane N, non-induced bacterial extract; lane I, extract from IPTG induced cells. (B) Southwestern of non-induced and induced *E. coli* cell extracts probed with ^{32}P-labeled ds random oligonucleotides. The arrows indicate the position of the bZNF35 protein.

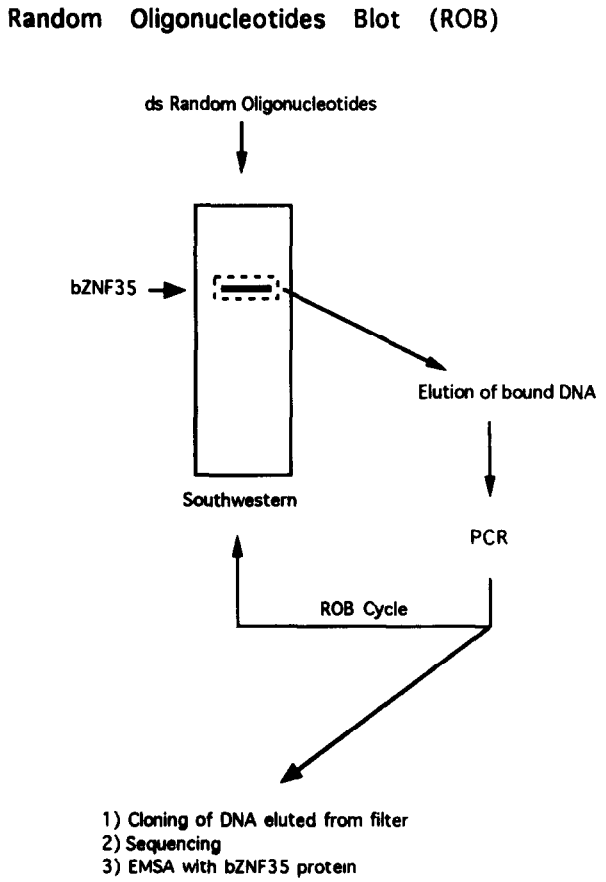

Fig. 2. Random oligonucleotide blot (ROB). The ROB assay requires a bacterial synthesized DNA-binding protein and a pool of random oligonucleotides. Selected oligonucleotides are subjected to further ROB cycles. Finally, the ROB-selected oligonucleotides are cloned, analyzed and sequenced.

experiments of the bacterial proteins indicated that the bZNF35 is capable of binding to DNA (Fig. 1B).

A modification of the method developed by Treacy et al. [18] was used to isolate DNA sequences that were recognized specifically by the bZNF35 (Fig. 2). A degenerate oligonucleotide that contained random nucleotides at 20 consecutive positions flanked by a defined sequence was chemically synthesized. The 50-base single-strand synthetic oligonucleotide (N) was used as a template for second-strand synthesis, which was performed with primer A and DNA polymerase in the presence of deoxynucleotide triophosphates, and a complementary degenerate library of double stranded oligonucleotides was obtained. To identify putative target sequences for ZNF35, we utilized a procedure that involved immobilizing bacterially produced ZNF35 protein on nitrocellulose membrane, renaturation and incubation with the pool of ^{32}P-labeled double-stranded oligonucleotides containing 20 consecutive random base pairs. The specific bound DNA was then eluted and amplified by using the defined flanks as primer sites in a PCR. The amplified DNA was then subjected to

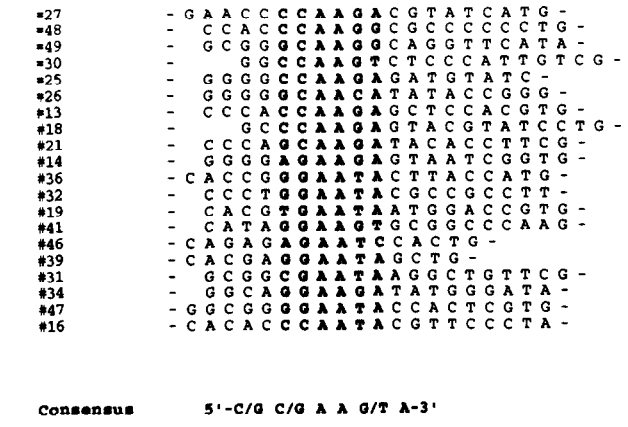

Fig. 3. Selected binding sites for bZNF35 protein. The sequences of 20 cloned ZNF35 binding sites derived from four rounds of ROB selection are shown and aligned. The limits of the central 6 bases are indicated by boldface letters. At the bottom is the consensus sequence.

three additional cycles of binding and amplification. The DNA recovered after four rounds of selection was analyzed following purification by molecular cloning and DNA sequencing, and results are shown in Fig. 3. Of 20 cloned oligonucleotides analyzed, a 6-bp long consensus sequence was identified in all of them. The consensus sequence was 5'-C/GC/GAAG/TA-3'.

To further show the binding specificity of bZNF35 to the 5'-C/GC/GAAG/TA-3' sequence, we performed gel retardation assays (EMSA) in the presence of an unlabeled specific competitor (Fig. 4). As probes we used DNA fragments derived from the no. 36 and no. 18 clones after digestion with *Eco*RI and *Bam*HI, respectively. The presence of a 100-fold excess of the specific competitor inhibited the binding of the radiolabeled DNA fragment to the bacterial bZNF35. In addition, we found that the presence of non-specific competitors, such as the binding site for SP1 and for AP1 proteins, did not have any effect on the binding (Fig. 4).

4. DISCUSSION

The identification of high-affinity binding sites in DNA for transcription factors is an important step in their characterization. In this paper, we present evidence that the human ZNF35 gene encodes a sequence-specific DNA-binding protein, and the core sequence of the binding site is 5'-C/GC/GAAG/TA-3'. This binding activity most likely resides in the carboxy terminus which contains 11 zinc fingers. To identify the ZNF35-binding site, we have used a modification of a technique described recently by Treacy et al. [18]. Definition of target sequences of various DNA proteins by random selection and PCR have been reported by several groups [18,21–24]. We found that the major advantage of the technique is that the bacterially expressed protein requires no purification. In addition, it is possible to visualize directly the binding of the recombinant protein from the binding of the contaminant bacterial proteins (Fig. 1B).

We noticed that the ZNF35 binding site is also present within the Evi-1-binding site. The Evi-1 is a finger gene originally identified as a common site of viral integration in murine myeloid tumors [25]. The binding site for Evi-1, identified by the gel shift PCR method using random oligonucleotides, was found to be TGACAAGATAA, which contains the core sequence of the ZNF35-binding site (CAAGA). Therefore, it appears that both ZNF35 and Evi-1 zinc-finger proteins might bind to a similar DNA sequence. It will be interesting to determine the transcription regulatory potential of both ZNF35 and Evi-1 proteins on a reporter gene containing the ZNF35-binding site. It is pertinent to note that two human zinc finger genes, WT-1 and EGR-1, have been recently shown to bind to the same DNA sequences [26]. However, co-transfection experiments demonstrated that each protein has an opposite effect on transcription: the WT-1 acts a repressor, while the EGR-1 acts as an activator [27].

Several transcriptional activators contain acidic domains which most likely interact with other proteins in the transcriptional complex [28]. The presence of an acidic domain in the amino part of ZNF35 suggests that ZNF35 can function in vivo as a transcriptional activator. We are currently investigating this possibility by constructing promoter sequences with the ZNF35-binding site linked to reporter genes and assaying for ZNF35 trans-activation by cotransfection.

Acknowledgements: We gratefully acknowledge Agnes Stacia for editorial assistance, and we thank Miss R. Terracciano for technical help. This work was paid for by a grant from the Italian Association for Cancer Research (AIRC) L.L.

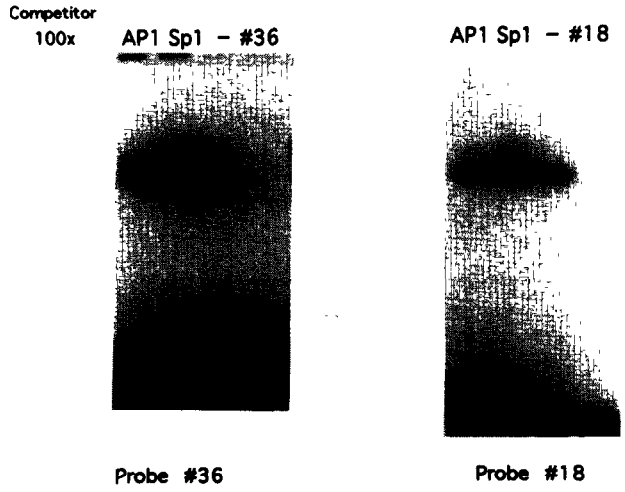

Fig. 4. Characterization of ZNF35-binding sites. EMSA were performed using recombinant ZNF35 protein and the *Eco*RI–*Bam*HI inserts derived from clones no. 36 (left panel) and no. 18 (right panel) were used as probes. 100× molar excess of specific and non-specific unlabeled competitors (Sp1 and AP1) were used as indicated. Lane (−), no competitor was added to the binding reactions.

REFERENCES

[1] Gehring, W.J. (1987) Science 236, 1245–1252.
[2] Frigerio, G., Burri, M., Bopp, D., Baumgartener, S. and Noll, M. (1986) Cell 47, 735–746.
[3] Herr, W., Strum, R.A., Clerc, R.G., Corcoran, L.M., Baltimore, D., Sharp, P.A., Ingram, H.A., Rosenfeld, M.G., Finney, M., Ruvkin, G. and Horvit, H.R. (1988) Genes Dev. 2, 1513–1516.
[4] Deutsch, U., Dressler, G.R. and Gruss, P. (1988) Cell 53, 617–625.
[5] Schuh, R., Archer, W., Gaul, U., Cote, S., Preiss, A., Maier, D., Seifert, E., Nauber, U., Schroder, C., Kernier, R. and Jackie, H. (1986) Cell 47, 1025–1032.
[6] Chowdhury, K., Deutsch, U. and Gruss, P. (1987) Cell 48, 771–778.
[7] Ruppert, J.M., Kinzler, K.W., Wong, A.J., Bigner, S.H., Kao, F.-T., Law, M.L., Seunaez, H.N., O'Brien, S.J. and Volgelstein, B. (1988) Mol. Cell Biol. 8, 3104–3113.
[8] Pannuti, A, Lanfrancone, L., Pascucci, A., La Mantia, G., Pelicci, P.-G. and Lania, L. (1988) Nucleic Acids Res. 16, 4227–4237.
[9] Bellefroid, E.J., Lercocq, P.J., Benhida, A., Poncelet, D.A., Belayew, A. and Martial, J.A. (1989) DNA 8, 377–387.
[10] Biedenkapp, H., Borgmeyer, U., Sippel, A.E. and Klempnauer, K.H. (1988) Nature 335, 835–837.
[11] Odenwald, W.F., Garbern, J., Arnheiter, H., Tournier-Lasserve, E. and Lazzarini, R.A. (1989) Genes Dev. 3, 158–172.
[12] Stanojevic, D., Hoey, T. and Levine, M. (1989) Nature 341, 331–335.
[13] Treisman, J. and Desplan, C. (1989) Nature 341, 335–337.
[14] Faisst, S. and Meyer, S. (1992) Nucleic Acids Res. 20, 3–26.
[15] Lania, L., Donti, E., Pannuti, A., Pascucci, A., Pengue, G., Feliciello, I., La Mantia, G., Lanfrancone, L. and Pelicci, P.G. (1990) Genomics 6, 333–340.
[16] Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1985) Methods Enzymol. 185, 60–89.
[17] Courey, A.J. and Tjian, R. (1988) Cell 55, 887–898.
[18] Treacy, M.N., He, X. and Rosenfeld, M.G. (1991) Nature 350, 577–584.
[19] La Mantia, G., Majello, B., Di Cristofano, A., Strazzullo, M., Minchiotti, G. and Lania, L. (1992) Nucleic Acids Res. 20, 4129–4136.
[20] Lanfrancone, L., Pengue, G., Pandolfi, P.-P., Salcini, A.-E., Giacomuzzi, A., Longo, L., Donti, E., De Luca, P., La Mantia, G., Pelicci, P.-G. and Lania, L. (1992) Genomics 12, 720–728.
[21] Thisen, H.J. and Bach, C. (1990) Nucleic Acids Res. 18, 3203–3208.
[22] Kinzler, K.W. and Vogelstein, B. (1990) Mol. Cell. Biol. 10, 634–642.
[23] Mavrothalassitis, G., Beal, G. and Papas, T.S. (1990) DNA Cell Biol. 9, 783–788.
[24] Pollock, R. and Treisman, R. (1990) Nucleic Acids Res. 18, 6197–6204.
[25] Perkins, A.S., Fishel, R., Jenkins, N.A. and Copeland, N.G. (1991) Mol. Cell. Biol. 11, 2665–2674.
[26] Rauscher, F.J., Morris, J.F., Tounay, O.E., Cook, D.M. and Curran, T. (1990) Science 250, 1259–1262.
[27] Madden, S.L., Cook, D.M., Morris, J.F., Gashler, A., Sukhatme, V.P. and Rausher, F.J. (1991) Science 1550–1552.
[28] Ptashne, M. and Gann, A.A. (1990) Nature 346, 329–331.