Vol. 187, No. 1, 1992 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
August 31, 1992 Pages 375-380

MOLECULAR CLONING OF NUCLEOBINDIN, A NOVEL DNA-BINDING PROTEIN THAT CONTAINS BOTH A SIGNAL PEPTIDE AND A LEUCINE ZIPPER STRUCTURE

Keiji Miura[1], Koiti Titani[1], Yoshikazu Kurosawa[1]
and Yoshiyuki Kanai[2*]

[1]Institute for Comprehensive Medical Science, Fujita Health University, Toyoake, Aichi 470-11, Japan

[2]Department of Molecular Oncology, Institute of Medical Science, University of Tokyo, Minato-ku, Tokyo 108, Japan

Received July 3, 1992

SUMMARY: We have previously reported that KML_1-7 cells cloned from a lupus-prone MRL/l mouse produced a soluble factor that preferentially expanded anti-DNA antibody production across the H-2 barrier. We purified this factor, a 55 kD protein that we termed nucleobindin (Nuc), and obtained its cDNA clone. Although the gene for Nuc encodes a signal peptide and, in fact, Nuc was identified as a secreted protein, Nuc had a DNA-binding property. The putative polypeptide predicted from the cDNA sequence featured a signal peptide, a leucine zipper structure and a basic amino acid-rich region. The DNA-binding property of Nuc was destroyed by deletion of either the leucine zipper structure or the basic amino acid-rich region. The amino acid sequences of Nuc are highly conserved between mouse and human. We discuss the possible role of Nuc in autoimmunity.
© 1992 Academic Press, Inc.

Sera of patients with systemic lupus erythematosus (SLE) are known to contain a variety of autoantibodies to nuclear antigens (1). In particular, IgG anti-double stranded (ds) DNA antibodies have been shown to be correlated with the clinical activity of the disease and are thought to be involved in the pathogenesis of renal lesions (2, 3). Lupus-prone MRL/l mice ($H\text{-}2^k$), developed at the Jackson Laboratory (USA) (4), are known to produce a large amount of both anti-single stranded (ss) and anti-dsDNA antibodies relatively early in the course of murine lupus (5). In previous works we established a cell line, KML_1-7, from an MRL/l mouse, which produced a soluble factor that promoted formation of anti-ss and anti-ds DNA antibodies *in vivo* and *in vitro* (6,7). To clarify the characteristics of this factor which we termed nucleobindin (Nuc), we purified it from a

*To whom correspondence should be addressed.

0006-291X/92 $4.00
Copyright © 1992 by Academic Press, Inc.
All rights of reproduction in any form reserved.

large volume of KML_1-7 cell culture supernatant (8). We report here the structure of the Nuc gene and its putative polypeptide.

MATERIALS AND METHODS

Nuc was purified from KML_1-7 cell culture supernatant as described elsewhere (8). Since the N-terminus of the purified protein was blocked, the protein was digested with *Achromobacter* protease I (9) and chromatographed by using a Synchropack RP-8 column. The amino acid sequences of two resulting fragments I and II were determined on an Applied Biosystems 470A protein sequencer with an on-line 120A PTH analyzer. The fragments I and II were QFEHLDPQNQETFEAIDLEL and LSQELDFVSHNVRTK, respectively. Polymerase chain reaction (PCR) was used to isolate cDNA against the fragment I as follows. Two primers were synthesized by an Applied Biosystem 381A DNA synthesizer. The upstream primer was 5'-CA(A/G)TT(C/T)GA(A/G)CA(C/T)CT (C/G)GA(C/T)CC corresponding to the RNA sequence encoding QFEHLDP. The downstream primer was 5'-AG(C/T)TC(C/G)AG(A/G)TC(A/G)AT(A/G)GC(C/T)TC complementary to the RNA sequence encoding EAIDLEL. Poly(A)$^+$RNA was isolated from KML_1-7 cells. Single stranded cDNA, synthesized by a copy kit (Invitrogen) using oligo(dT) as a primer, was used as a template for PCR (10). The PCR profile consisted of 40 cycles of the following treatment: 1 min at 94°C, 2 min at 37°C and 30 sec at 72°C, in a Perkin-Elmer/Cetus Thermal Cycler. The products showing 59bp were separated, cloned and sequenced. One clone, corresponding to the amino-acid sequence of the fragment I, was selected as hybridization probe. The cDNA library was constructed from poly(A)$^+$RNA of KML_1-7 cells using a cDNA synthesis kit and λgt10 (Amersham). Four positive clones were isolated and a clone termed λNuc-1, containing an insert of 2.2 kb, was sequenced. DNA sequence was determined by the dideoxy chain termination method (11). Using mouse Nuc gene as probe, a cDNA clone of human Nuc gene was also isolated from cDNA library constructed from human pheochromocytoma cells.

RESULTS

Figure 1 shows the nucleotide sequence of the mouse Nuc gene. It contained a 455 amino acid-long open reading frame in which the fragments I and II were included. This polypeptide contained a highly hydrophobic sequence in the N-terminal region. Since the N-terminus of Nuc was blocked, the exact end of the secreted form could not be identified. However, the human homologue of Nuc also contains a hydrophobic region (see Fig. 4). This region appears to be the signal peptide for the secretion of Nuc. We tentatively assigned the 26th amino acid residue, valine, as the N-terminal end of the secreted form based on the rule for signal peptides summarized by Perlman and Halvorson (12). The calculated molecular mass of the secreted Nuc is 50,410 daltons. The most intriguing characteristics of Nuc is the seven-fold presence of leucine in every seventh position between the 346th and the 388th residue. Moreover, a basic amino acid-rich region is observed between the 171th and the 217th residue.

In addition to the presence of both a leucine zipper structure and a basic amino acid-rich region in Nuc, we always found polynucleotides together with Nuc's anti-DNA antibody-

```
CTGGAGAGACCAGCCACGATGCCTACCTCGGTGCCCCGCGGGGCCCCTTTTCTTCTCCTACCACCTCTACTGATGCTGTCTGCTGTGCTGGCAGTGCCCGTGGACCGCGCAGCACCTCCT  120
                  M  P  T  S  V  P  R  G  A  P  F  L  L  L  P  P  L  L  M  L  S  A  V  L  A  V  P  V  D  R  A  A  P  P    34

CAGGAGGACAGCCAGGCCACTGAGACCCCGGACACGGGCCTGTACTACCACCGGTACCTCCAGGAGGTCATCAACGTGCTAGAGACAGATGGGCACTTCCGGGAGAAGCTGCAGGCTGCC  240
Q E D S Q A T E T P D T G L Y Y H R Y L Q E V I N V L E T D G H F R E K L Q A A   74

AACGCTGAGGACATTAAGAGTGGAAAGCTGAGCCAAGAGCTGGACTTCGTCAGCCACAACGTCCGAACCAAGCTGGATGAGCTCAAGCGACAGGAGGTGTCGAGGCTGCGGATGCTGCTC  360
N A E D I K S G K L S Q E L D F V S H N V R T K L D E L K R Q E V S R L R M L L   114
                                Fragment II
AAGGCCAAGATGGATGCAAAGCAGGAGCCCAACTTGCAGGTGGACCACATGAACCTCCTGAAGCAGTTTGAGCACCTGGACCCTCAGAACCAGCACACGTTTGAGGCTCGGGACCTAGAG  480
K A K M D A K Q E P N L Q V D H M N L L K Q F E H L D P Q N Q H T F E A R D L E   154
                                                Fragment I
CTGCTGATCCAGACGGCCACCCGAGACCTCGCCCAGTATGACGCTGCACACCATGAAGAGTTCAAACGCTACGAGATGCTCAAGGAACATGAGAGGAGACGATACCTGGAGTCTCTGGGA  600
L L I Q T A T R D L A Q Y D A A H H E E F K R Y E M L K E H E R R R Y L E S L G   194

GAGGAGCAGCGGAAGGAGGCCGAGAGGAAGCTCCAAGAGCAACAGCGCAGACACCGGGAACACCCCAAAGTCAATGTTCCTGGCAGCCAAGCCCAGTTGAAGGAGGTGTGGGAGGAGTTG  720
E E Q R K E A E R K L Q E Q Q R R H R E H P K V N V P G S Q A Q L K E V W E E L   234

GATGGATTGGACCCCAACAGGTTCAACCCCAAGACCTTCTTCATACTGCATGACATCAACAGTGATGGTGTCCTAGATGAGCAAGAACTGGAAGCTCTCTTTACCAAGGAGCTGGAGAAG  840
D G L D P N R F N P K T F F I L H D I N S D G V L D E Q E L E A L F T K E L E K   274

GTGTATGACCCGAAGAATGAGGAGGATGACATGAGAGAGATGGAGGAAGAACGGCTTCGCATGCGGGAGCACGTGATGAAGAATGTGGACACCAACCAGGACCGGCTTGTGACCCTGGAG  960
V Y D P K N E E D D M R E M E E E R L R M R E H V M K N V D T N Q D R L V T L E   314

GAGTTCCTGGCATCCACCCAGAGGAAGGAGTTCGGGGACACTGGGGAGGGGTGGAAGACAGTGGAAATGTCCCCAGCCTACACAGAGGAGGAGCTGAAGCGTTTTGAAGAGGAGCTGGCT 1080
E F L A S T Q R K E F G D T G E G W K T V E M S P A Y T E E E (L) K R F E E E (L) A   354

GCCCGGGAAGCTGAGCTGAATGCCAGGGCCCAGCGCCTCAGCCAGGAGACAGAGGCCCTGGGGCGCTCCCAGGACCGCCTGGAGGCACAGAAGAGAGAGCTGCAGCAGATGGAGCAGAGG 1200
A R E A E (L) N A R A Q R (L) S Q E T E A (L) G R S Q D R (L) E A Q K R E (L) Q Q M E Q R   394

AAGCAGCAACTGCAAGAACAGAGCGCTCCGCCTTCCAAACCTGACGGGCAGCTGCAGTTCCGTGCAGACACAGATGACGCTCCTGTCCCAGCTCCAGCAGGTGACCAGAAAGATGTGCCT 1320
K Q Q L Q E Q S A P P S K P D G Q L Q F R A D T D D A P V P A P A G D Q K D V P   434

GCTTCTGAGAAGAAGGTCCCAGAGCAGCCCCCTGAGCTGCCACAGCTGGATTCCCAGCATTTATAACTCCTCAGAGACCCCAGCCCTCAGGTTCCCACAGACAGTGTGGAAGCCAGATGA 1440
A S E K K V P E Q P P E L P Q L D S Q H L *   455

AGTGGCCTTGCAGCCTCCCTGGGGACCAACGCCCTCATGGCCTCCTGGGAGCAGAACAGGGGCTACCCTGTTCCTCAGATGTGCTTCTGCTGGTCTACCATTCTCTGCTCTCAGCCTCTG 1560
GGTCCAGGTTCTGCCTCCCTCTCTCGCTCCCTCCCTCCACTGTCAGGGATTAACTACAGGAGCTGAAGTTGCAGGAGAGCTCGCTCACCTCACCTGGAGTGCACACTCCAGGCCTGCCCA 1680
CCTGCCTACTTTGCTCGTTCTCTGTTCCAGGCTCAGCCTCTCTGCTTCCTGTACCTAGTGTCACTGGATCACAGGACCTTGGCTTAGTGTCTCTGACTCTGCCTTCTGTGGAGCCAATGA 1800
TCTCAATTCCTTTTGGTGGCTAGAGGCCACCAGGATGTGGGGTGTTCTTCAAGTAGCTATCAGCTCTGCCTGGGACTCCTGGCCTTTACCACCACCCCATCAGCCTACCCTAACTCCCAG 1920
GCTCAATACCTACCCCATGGCCCCAGGAGGTACCCAGATATGGCTACACCCCTAGCTGAGGTCCAACCCTGCTCAGGGCTCCTGAAAATTGGTTCACCAACTTCCCTGGCTTCAGGCCAC 2040
ACTGCTGTGTCCTGCTCTGTGCCCCAGTTCCATCCAAATATAGCCCGGCACAAATAAACATTAGTGTATCTGTGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  2157
```

Fig. 1. Nucleotide and amino acid sequences of mouse Nuc gene. The sequences of the fragments I and II were included in this sequence with two amino-acid discrepancies, which were boxed. The underlined sequences were confirmed by amino acid sequence analysis of rNuc protein. Leucine residues in the putative leucine zipper structure were circled. (GenBank accession number of the sequence is M96823)

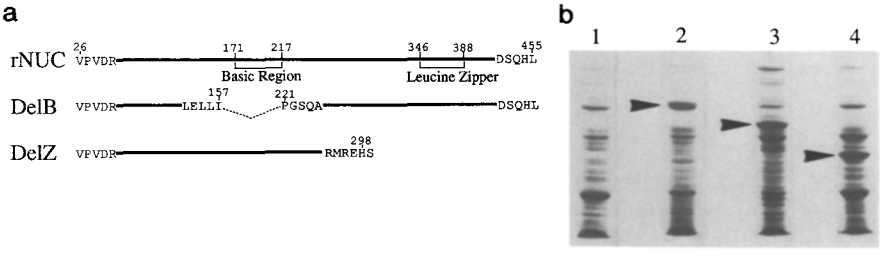

Fig. 2. Production of rNuc and its deletion mutants in *E.coli*. (a) The expression vector λLc1 containing the pelB sequence was used for Nuc production in *E. coli* (15). Two kinds of deletion mutants, DelB lacking the 63 amino acids (from the 158th to 220th) and DelZ lacking the 157 amino acids (from the 299th to 455th) and containing one extra serine at the C terminus, were also constructed. These three recombinant proteins were produced in *E. coli* in periplasmic forms. Sequencing of the N-terminal regions of the isolated proteins showed that the pelB sequences were removed and their sequences were VPVD as expected . (b) SDS-PAGE of crude fraction of *E.coli* harboring the vector alone (*lane 1*). λLc1-rNuc (*lane 2*), λLc1-DelB (*lane 3*), and λLc1-DelZ (*lane 4*). Positions of rNuc, DelB and DelZ were indicated by arrowheads.

boosting activity in the same fractions during the process of purifying Nuc from culture supernatants of KML_1-7 cells (8). Based on these observations, we suspected Nuc to have DNA-binding property, and thus induce anti-DNA antibodies in an antigen-specific manner. To test this possibility, a recombinant Nuc (rNuc) protein was produced in *E. coli*. Two kinds of deletion mutants, DelB and DelZ, were also constructed and produced in *E. coli* (Fig. 2). As shown in Fig.3, rNuc protein showed a DNA-binding property. DNA fragments used in this experiment consisted of a mixture which was prepared from KML_1-7 cell culture supernatants. A significant portion of DNA was shifted on the gel by addition of rNuc protein (Fig. 3, lane 3). When the basic amino acid-rich region of rNuc was deleted, this property was greatly reduced (Fig. 3, lane 4) and when the leucine zipper structure of rNuc was deleted, DNA mobility shift was not observed at all (Fig. 3, lane 5). Thus, we concluded that rNuc protein has a DNA-binding property, the leucine zipper structure should be requisite for DNA binding and that the basic amino acid-rich region may form a DNA-binding site. Importance of the leucine-zipper structure in Nuc was further supported by comparison of amino acid sequences of Nuc between mouse and

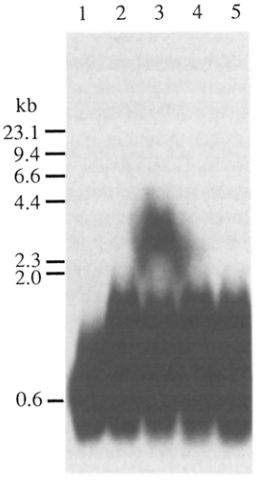

Fig. 3. Mobility shift in DNA bound to rNuc. A gel electrophoresis DNA binding assay was performed according to the standard protocol (16). Briefly, reaction mixtures (19 µl) were made up of 10 mM Hepes-KOH (pH 7.9), 50 mM KCl, 1 mM $MgCl_2$, 0.25 mM DTT, 10% glycerol, 0.25 mM PMSF, 12.5 mM spermidine, 0.1 mg poly(dI-dC) /ml and 64 µg of crude fraction. They were incubated at room temperature for 10 min. Then 1 µl (1 ng) of DNA (~600 bp), prepared from supernatants of KML_1-7 cell cultures and labeled at the 5' end with [γ-^{32}P]ATP and T4 polynucleotide kinase, was added to the reaction mixtures and incubated at 30°C for 30 min. Finally, they were electrophoresed on 1% agarose gel in low ionic buffer (6.7 mM Tris, pH 7.9, 3.3 mM sodium acetate, 1 mM EDTA). The gel was dried and subjected to autoradiography. DNA size markers shown on the left of the photograph are *Hin*dIII digested λDNA and given in kb. *lane 1*, ^{32}P-DNA alone; *lane 2*, addition of extracts of *E.coli* harboring only the vector; *lane 3*, addition of rNuc protein; *lane 4*, addition of DelB protein; *lane 5*, addition of DelZ protein.

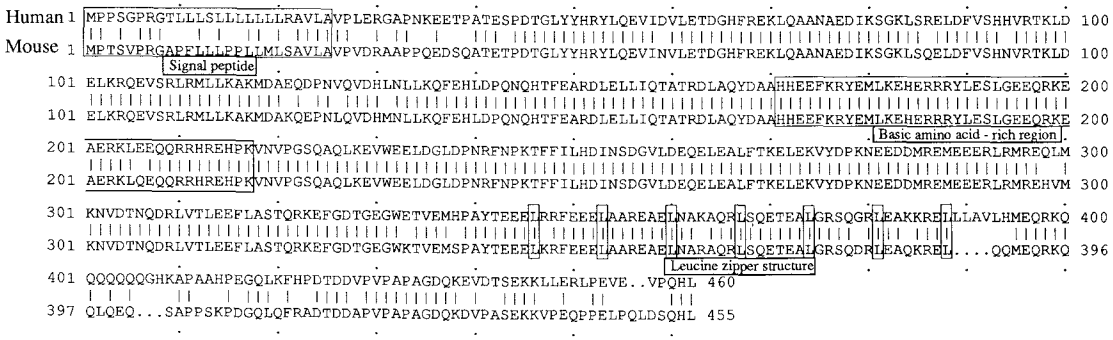

Fig. 4. Comparison of amino acid sequences of Nuc between human and mouse. (GenBank accession number of the human sequence is M96824)

human. The major parts of Nuc from the 40th residue to the 388th residue which corresponded to the end of the leucine zipper showed 95% identities of amino acids between two species as shown in Fig.4.

DISCUSSION

Molecular cloning of Nuc revealed presence of a new kind of molecule which contains a signal peptide and a leucine zipper. The leucine zipper structure is assumed to be a dimerization motif prevailing in various transcription factors such as Fos and Jun (13). In contrast with the transcription factors, the presumed DNA-binding site, a basic amino acid-rich region, is not located immediately upstream from the leucine zipper but further upstream. In fact, however, rNuc has a DNA-binding property and deletion of either the leucine zipper or the basic amino acid-rich region destroyed the activities. Nuc is the known first example of a polypeptide apart from transcription factors that has both leucine zipper and DNA-binding property.

As concerns autoimmunity, it has been one of the major objectives to identify molecules that help produce IgG anti-DNA antibodies. However, analyses of the cytokines identified so far did not indicate any antigen specificity. We wonder why anti-DNA antibodies are specifically amplified in SLE patients. The DNA-binding property of the rNuc protein might explain why a soluble factor from KML_1-7 cells could preferentially induce IgG anti-DNA antibody production across the H-2 barrier. The biological activities of Nuc is now under investigation and will be published elsewhere. In brief, addition of rNuc induced IgG anti-DNA antibody production in cultures of B cells from MRL/l mice. However, mutant DelB and DelZ have lost this activity.

In the culture supernatants of KML_1-7 cells, a large amount of DNA coexisted with Nuc. Size analysis of the DNA revealed its multiple forms of nucleosome units shaping a so-called nucleosome ladder. A part of KML_1-7 cells may undergo apoptosis (14). The message of mouse Nuc gene was also detected in normal lymphoid cells and in other somatic cells such as liver and kidney cells. In normal immune systems, the circulation of secreted DNA-Nuc complexes might rapidly be terminated, but in lupus-prone mice, B cells could be over-stimulated, resulting in clonal expansion of anti-DNA antibody-producing cells.

ACKNOWLEDGMENTS

We thank Dr. A. Awaya for promotion on this collaboration work, Dr. W. Ito for introducing the λLc1 vector, Mr. M. Suzuki and Ms. R. Oyama for protein sequence analysis, and Ms. A. Nagata for preparing the manuscript.

REFERENCES

1. Tan, E.M. (1989) *Adv. Immunol.* **44**, 93-151.
2. Schur, P.H. & Sandson, J. (1968) *N. Engl. J. Med.* **278**, 533-538.
3. Pincus, T., Schur, P.H., Rose, J.A., Decker, J.L. & Talal, N. (1969) *N. Engl. J. Med.* **281**, 701-705.
4. Murphy, E.D & Roths, J.B. (1978) Autoimmunity and lymphoproliferation: Induction by mutant gene lpr, and acceleration by a male-associated factor in strain BXSB mice. In Genetic Control of Autoimmune disease, N.R. Rose, P.E.Bigazzi & N.L. Warner eds. (Elsevier North Holland Inc.), pp. 207-221.
5. Theofilopoulos, A.N. & Dixon, F.J. (1985) *Adv. Immunol.* **37**, 269-390.
6. Kanai, Y., Katagiri, T., Mori, S. & Kubota, T. (1986) *Int. Arch. Allergy appl. Immunol.* **81**, 92-94.
7. Kanai, Y., Isonishi, S., Katagiri, T., Koizumi, T., Miura, K. & Kurosawa, Y. (1990) *Immunol. Lett.* **24**, 49-56.
8. Kanai, Y. & Tanuma, S. (1992) *Immunol. Lett.* **32**, 43-48.
9. Masaki, T., Tanabe, M., Nakamura, K. & Soejima, M. (1981) *Biochim. Biophys. Acta* **660**, 44-50.
10. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J.,Higuchi, R., Horn, G.T., Mullis, K.B. & Erlich, H.A. (1988) *Science* **239**, 487-491.
11. Sanger, F., Nicklen, S. & Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.
12. Perlman, D. & Halvorson, H.O. (1983) *J. Mol. Biol.* **167**, 391-409.
13. Landschulz, W.H., Johnson, P.F. & McKnight, S.L. (1988) *Science* **240**, 1759-1764.
14. Wyllie, A.H., Morris, R.G., Smith, A.L. & Dunlop, D. (1984) *J. Pathol.* **142**, 67-77.
15. Huse, W.D., Sastry, L.,Iverson, S.A., Kang, A.S., Alting-Mees, M., Burton, D.R., Benkovic, S.J. & Lerner, R.A. (1989) *Science* **246**, 1275-1281.
16. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. & Struhl,K. (1990). Current Protocols in Molecular Biology. (John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore).