Identification of the factors that interact with NCBP, an 80 kDa nuclear cap binding protein

Naoyuki Kataoka, Mutsuhito Ohno, Ichiro Moda[+] and Yoshiro Shimura*

Department of Biophysics, Faculty of Science, Kyoto University, Kyoto 606, Japan

ABSTRACT

It has been shown that the monomethylated cap structure plays important roles in pre-mRNA splicing and nuclear export of RNA. As a candidate for the factor involved in these nuclear events we have previously purified an 80 kDa nuclear cap binding protein (NCBP) from a HeLa cell nuclear extract and isolated its full-length cDNA. In this report, in order to obtain a clue to the cellular functions of NCBP, we attempted to identify a factor(s) that interacts with NCBP. Using the yeast two-hybrid system we isolated three clones from a HeLa cell cDNA library. We designated the proteins encoded by these clones NIPs (NCBP interacting proteins). NIP1 and NIP2 have an RNP consensus-type RNA binding domain, whereas NIP3 contains a unique domain of Arg–Glu or Lys–Glu dipeptide repeats. We also show that NCBP requires NIP1 for binding to the cap structure. Possible roles of NIPs in cap-dependent nuclear processes are discussed.

INTRODUCTION

RNAs transcribed by RNA polymerase II have a unique structure at the 5′-terminus. This structure, $m^7G(5')ppp(5')N$, is called a monomethylated cap structure and has been shown to play important roles in some cellular functions.

In the cytoplasm the cap structure enhances translation by facilitating binding of the ribosome to mRNA (1). Biochemical analyses showed that the cytoplasmic cap binding complex, termed eukaryotic translation initiation factor-4F (eIF-4F), is involved in this enhancement (for reviews see 2,3). eIF-4F consists of four polypeptide components, eIF-4E, eIF-4A, eIF-4B and p220 (3). eIF-4E is a 24 kDa protein that recognizes and binds specifically to the cap structure (3). eIF-4A is an RNA helicase and is involved in unwinding the 5′-terminal secondary structure of mRNA in an ATP-dependent manner in order to expose the initiating methionine codon (3). eIF-4B has an RNP consensus-type RNA binding domain and stimulates the RNA helicase activity of eIF-4A (3,4). One additional protein, p220, is also included in the complex (3).

It has been shown that the cap structure is also important for nuclear events such as pre-mRNA splicing and nuclear export of RNA. Formation of the spliceosome was demonstrated to be cap

dependent (5). We have shown both *in vitro* and in *Xenopus* oocyte nuclei that the cap structure stimulates excision of the proximal intron when the pre-mRNA has two introns within the molecule (6,7). It has also been suggested that the monomethylated cap structure facilitates export of RNA polymerase II transcripts from the nucleus to the cytoplasm (8–12). However, as compared with the cytoplasmic events, little information has been obtained so far about nuclear cap binding protein(s) and its associated factor(s), which may mediate the nuclear processes.

As a candidate factor involved in cap-dependent nuclear processes we have previously identified and purified an 80 kDa nuclear cap binding protein (termed NCBP) from HeLa cell nuclear extracts (13,14) and isolated its full-length cDNA (15). Independently, a cap binding protein that may mediate nuclear export of RNA was purified by others and was termed CBP80, which turned out to be identical to NCBP (9,16). Moreover, it was shown that CBP80 forms a complex with another protein, termed CBP20, and that this complex is involved in pre-mRNA splicing by facilitating spliceosome formation (16).

To obtain suggestions as to how NCBP/CBP80 is involved in pre-mRNA splicing and, possibly, in nuclear export of RNA an attempt was made to identify a factor(s) that interacts with NCBP. Using the yeast two-hybrid interaction trap system (17) we have isolated three clones from a HeLa cell cDNA library. Two of them encoded distinct proteins with an RNP consensus-type RNA binding domain (4). The third contains a unique domain of Arg–Glu or Lys–Glu dipeptide repeats. We termed these proteins NCBP interacting proteins (NIPs). We also show that NIP1 is an essential factor for NCBP binding to the cap structure.

MATERIALS AND METHODS

Yeast two-hybrid interaction trap cloning

Yeast two-hybrid interaction trap cloning was performed according to Zervos *et al.* (17). The bait plasmid was constructed by inserting the NCBP full-length cDNA in-frame with the *LexA* gene in the vector pEG202 (17). The vectors, HeLa cDNA library and the yeast host strain EGY48 were kindly provided by Dr Roger Brent (Harvard Medical School, Boston, MA). The positive cDNAs were subcloned into pSP73 (Promega) and sequenced. The cDNA sequence of NIP1 will appear in the DDBJ, EMBL and
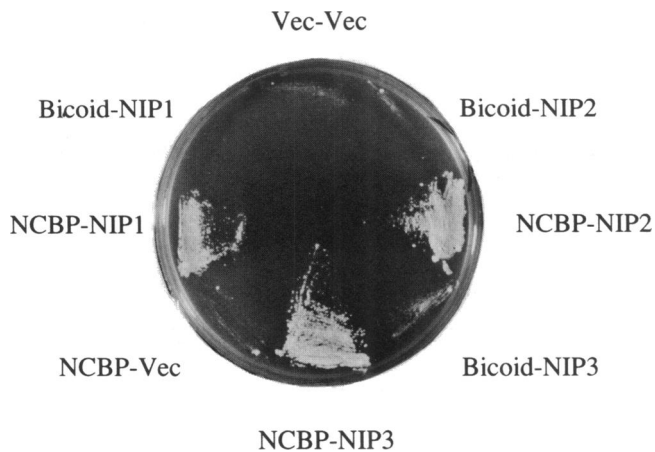

Figure 1. Specific interactions between NCBP and NIPs in the yeast two-hybrid system. Each NIP plasmid was introduced into the yeast host strain EGY48 with either the LexA–NCBP or LexA–bicoid plasmid and growth of each transformant on a leucine drop-out/galactose plate was examined. As negative controls library plasmid pJG4-5 was used with either the LexA–NCBP or LexA vector (indicated as NCBP–Vec or Vec–Vec, respectively).

Figure 2. The amino acid sequence of NIP1. The amino acid sequence is shown in the single letter amino acid code. The conserved amino acid sequence of the RNA binding domain is boxed and the two peptide sequences obtained from the purified NCBP fraction are underlined (see text).

GenBank nucleotide sequence databases under accession no. D59253.

Northern blot analysis

HeLa cell poly(A)$^+$ RNA (2 µg/lane) was electrophoresed in a 1% formaldehyde–agarose gel and blotted onto Hybond N nylon membrane (Amersham). ^{32}P-Labeled cDNA fragments derived from the positive clones were used as probes. Hybridization and the following procedures were carried out as described previously (15).

Preparation of yeast whole cell extracts

Yeast cells were cultured in galactose/His, Ura, Trp drop-out liquid medium at 30°C to an OD$_{600}$ of 1.0. The yeast cells from 1 ml culture were spun down and resuspended in 400 µl buffer D (18) containing 5 mg/ml each of leupeptin, peptastatin and aprotinin (Boehringer Mannheim). Glass beads were added up to the level of the solution and vortexed for 30 s, followed by cooling on ice. These steps were repeated twice. After centrifugation at 5000 r.p.m. at 4°C the supernatant was recovered and centrifuged again at 14 000 r.p.m. at 4°C. The final supernatant was saved as a whole cell extract.

Gel mobility shift assay

The gel mobility shift assay was carried out as described previously (13) using yeast whole cell extracts. The extracts were used without dilution and 5 µg carrier RNA was added to each reaction mixture.

RESULTS

Identification of NIPs

To obtain a clue as to how NCBP participates in nuclear events such as pre-mRNA splicing and, possibly, RNA export we used the yeast interaction trap cloning system to identify cDNA clones that encode proteins interacting with NCBP (17). As a bait plasmid the full-length NCBP cDNA was inserted into the bait expression vector EG202 (17). This bait plasmid is designated EG-NCBP. A HeLa cell cDNA library, into which the individual cDNAs were inserted in the prey expression vector pJG4-5 (17), was used for screening. Transcription from two reporter genes, *LacZ* and *Leu2*, is stimulated if the protein derived from the library forms a complex with the bait protein (NCBP in this case), which allows growth in the absence of leucine and enables color development on X-Gal plates.

A pool of yeast cells, representing 1.5×10^7 primary library transformants, was plated out onto galactose/Leu drop-out plates and 83 *Leu*$^+$ colonies were isolated. Among them, 16 showed galactose-dependent blue color on X-gal containing plates (data not shown). To characterize these clones cDNA plasmids from the colonies were recovered by transformation into *Escherichia coli* strain KC8, which is *trp*$^-$ and thereby enables the recovery of only the library plasmids (*Trp*$^+$) on the minimal plates (17). The recovered cDNAs were re-introduced into the yeast strain EGY48 with either EG-NCBP or the vector EG202 to confirm that transcription activation of the reporter genes was dependent on the presence of NCBP. By these selections we obtained four positive clones. DNA sequence analyses revealed that two of them were identical and, therefore, we finally obtained three different cDNA clones. The factors encoded by the cDNAs were designated NCBP interacting proteins (NIPs). As shown in Figure 1, NIP1–NIP3 can interact with NCBP, but not with bicoid, an unrelated *Drosophila* protein, confirming the specificity of interaction between NIPs and NCBP.

Primary sequence of NIP1–NIP3

The cDNA inserts of NIP1–NIP3 were sequenced and their reading frames, in-frame from the vector sequence, were assigned.

The NIP1 cDNA insert was 675 bases long and its predicted reading frame contained an RNP consensus-type RNA binding domain (RBD) (Fig. 2) (4). Northern blot analysis showed that the major transcript of NIP1 is 2.0 kb long and that the minor one is 0.8 kb (Fig. 3). A λgt10 HeLa cDNA library was screened using

Figure 3. Northern blot analyses of NIPs transcripts. Cytoplasmic poly(A)$^+$ RNA from HeLa cells was analyzed using NIPs and β-actin cDNA probes. The RNA size markers are shown on the left of the figure. The size of each transcript is also indicated on the right of each lane.

Figure 4. Alignment of the amino acid sequences of the NIP3 RE domain and the epitope of mAb 16H3. Identical amino acids are boxed.

Figure 5. Gel mobility shift assay using yeast whole cell extracts. Whole cell extracts from yeast cells expressing LexA–NCBP alone (NCBP) or NIP1 alone (NIP1) or both LexA–NCBP and NIP1 (NCBP-NIP1) were prepared and used for the gel mobility shift assay with either an m^{7}GpppG- or ApppG-primed RNA probe (indicated as cap + or –, respectively) in the absence or presence of a cap analog, m^{7}GpppG. As controls extracts from yeast cells carrying no plasmid (EGY48) and from yeast cells expressing both LexA–bicoid and NIP1 (Bicoid-NIP1) were used. The band of free RNA probes and the band shifted specifically with the m^{7}GpppG-primed probe are indicated by F and B, respectively.

NIP1 cDNA as a probe and two cDNA clones were obtained. One is nearly 2.0 kb long and the other 0.8 kb. It is most likely that they are derived from the two distinct transcripts detected in the Northern analysis. It turned out that they have identical coding regions, but have different 3′ untranslated regions (data not shown). Moreover, an in-frame stop codon was found upstream of the first methionine, so the cDNA fragments are thought to contain the whole coding region of NIP1. The calculated molecular mass of NIP1 is 17.997 kDa. The whole coding region was contained in the original cDNA insert obtained from the two-hybrid screening.

The original NIP2 cDNA insert was 811 bp long and also encoded an RBD. Moreover, this protein contains, near the C-terminus, a putative ATP binding motif, AXXXXGK(S/T) in the single letter amino acid code (19). NIP2 is presumably a large protein as judged by the size of the mRNA, ~9 kb (Fig. 3). We are now in the process of isolating the NIP2 full-length cDNA. The NIP3 cDNA insert was 488 bases long and encoded a polypeptide that had a unique domain. The domain consists of Arg–Glu (R–E) or Lys–Glu (K–E) dipeptide repeats. This domain is very similar to the epitope of the monoclonal antibody 16H3 (20), as will be discussed later (Fig. 4). The size of the NIP3 mRNA is 3.6 kb (Fig. 3). Isolation of the full-length NIP3 cDNA is also in progress.

NCBP is active in binding to the cap structure only in the presence of NIP1

Two lines of evidence have suggested the possibility that NIP1 is an essential subunit for binding of NCBP to the cap structure: (i) repeated attempts to detect cap binding activity with recombinant NCBP, expressed in *E.coli* or in rabbit reticulocyte lysates, were completely unsuccessful (data not shown); (ii) two peptide sequences obtained by proteolytic digestion of the highly purified NCBP fraction but that could not be assigned in the coding sequence of NCBP (15) were both found in NIP1 (Fig. 2).

To test this possibility we carried out a gel mobility shift assay using extracts from yeast cells expressing NCBP and/or NIP1.

The result is shown in Figure 5. If an extract from host yeast strain EGY48 was used some shifted bands were observed due to yeast endogenous proteins, however none of them were strictly dependent on the cap structure of the probe (lanes 1–3). We could not detect any additional band with extracts from yeast cells expressing only NCBP or NIP1 (lanes 7–12). It is worth noting that NIP1, although containing an RBD, does not seem to bind to RNA (compare lanes 2 and 11). In contrast, when extract from yeast cells expressing both NCBP and NIP1 was used a shifted band could be detected with a capped probe RNA, but not with an uncapped probe (lanes 4 and 5). This band disappeared on addition of a cap structure analog, m^{7}GpppG (lane 6). No cap structure-dependent band could be seen with the extract containing both bicoid and NIP1 (lanes 13–15). Since some shifted bands are observed due to yeast endogenous proteins, one might think that NIP1 alone can bind to the capped and/or uncapped RNAs and that the shifted band due to NIP1 is hidden in the background from yeast proteins. However, such a possibility is unlikely because when RNA gel shift experiments were performed using purified GST–NIP1 fusion protein produced in *E.coli* the protein alone did not bind to either capped or uncapped RNA (data not

shown). These results strongly suggest that NCBP binds to the cap structure only as a complex with NIP1.

DISCUSSION

Using the yeast two-hybrid system we have identified three candidate factors that interact with NCBP and designated them NIP1–NIP3. NIP1 is an 18 kDa protein with an RBD. Two peptide sequences obtained from the highly purified NCBP fraction but that could not be assigned in the coding sequence of NCBP (15) were found in NIP1. This means that NIP1 is associated with NCBP during extensive purification procedures, indicating that NCBP is in a stable complex with NIP1. It is highly likely that NIP1 is identical to CBP20, previously described by Izaurralde *et al.* (16). They have shown that NCBP/CBP80 is stably complexed with CBP20, a 20 kDa protein, and that the protein complex is involved in pre-mRNA splicing (16). They have also shown by UV-induced crosslinking experiments that CBP20, but not CBP80, is efficiently crosslinked to (or near) the cap structure (16). This result indicates that CBP20 is in close contact with the cap structure in the CBP80–CBP20 complex, thereby raising the possibility that CBP20 alone may be active in binding to the cap structure, in spite of the fact that it is complexed with NCBP/CBP80. In this study we have obtained evidence indicating that this is not the case and that both NCBP/CBP80 and NIP1 (presumably identical to CBP20) are essential for binding to the cap structure (Fig. 5). This is in contrast to the case of the cytoplasmic cap binding protein complex eIF-4F. In this case one of the protein components, eIF-4E, is fully active in binding to the cap structure by itself (3). It is interesting that the cap recognition mechanisms of the two cellular compartments are very different. In this respect it is noteworthy that some tryptophan residues in eIF-4E, conserved among species, are important for cap binding activity (3), whereas no similar sequence is found in either NCBP or NIP1. Although essential domains of NCBP and NIP1 for cap binding activity remain to be elucidated, it is likely that the RBD in NIP1 plays some role in binding activity.

We could identify two more factors, NIP2 and NIP3, which interact with NCBP in the yeast two-hybrid system. Although we do not have any biochemical or functional evidence that NIP2 and NIP3 truly interact with NCBP in the cell, it is possible that these proteins may be involved in nuclear events mediated by the NCBP–NIP1 complex. The fact that both NIP1 and NIP2 have an RBD raises an intriguing possibility that NCBP may change partners from NIP1 to NIP2 in a certain step of the cellular process to modulate cap binding activity. For example, in the process of nuclear export of RNA the NCBP–NIP1 complex may be bound to the cap structure of RNA in the nucleus, but NIP1 might be replaced by NIP2 at the nuclear surface or in the cytoplasm to allow the RNA to enter the next step of gene expression. In this

respect it would be of interest to investigate the localization of NIP2 in the cell.

NIP3 has a unique domain that consists of Arg–Glu (R–E) or Lys–Glu (K–E) dipeptide repeats. We termed it the RE domain. It has recently been shown that the monoclonal antibody 16H3, which recognizes an 'alternating arginine' domain, inhibits pre-mRNA splicing *in vitro* (20). One of the epitopes recognized by mAb 16H3 contains alternating arginine and glutamate in a sequence (20) which is extremely similar to the RE domain of NIP3 (Fig. 4). This domain might be functionally related to the Arg–Ser-rich domain (RS domain) characteristic of some metazoan splicing factors (20). It is possible that NIP3 may mediate enhancement of pre-mRNA splicing by the NCBP–NIP1 complex. However, the significance of NIP3 (and also of NIP2) in the cellular functions of NCBP remains to be clarified.

ACKNOWLEDGEMENTS

REFERENCES

1   Shatkin,A.J. (1985) *Cell*, **40**, 223–224.
2   Rhoads,R.E. (1988) *Trends Biochem. Sci.*, **13**, 52–56.
3   Sonenberg,N. (1988) *Prog. Nucleic Acid Res. Mol. Biol.*, **35**, 174–207.
4   Burd,C.G. and Dreyfuss,G. (1994) *Science*, **265**, 615–621.
5   Patzelt,E., Thalmann,E., Hartmuth,K., Blaas,D. and Kuechler,E. (1987) *Nucleic Acids Res.*, **15**, 1387–1399.
6   Ohno,M., Sakamoto,H. and Shimura,Y. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 5187–5191.
7   Inoue,K., Ohno,M., Sakamoto,H. and Shimura,Y. (1989) *Genes Dev.*, **3**, 1472–1479.
8   Hamm,J. and Mattaj,I.W. (1990) *Cell*, **63**, 109–118.
9   Dargemont,C. and Kühn,L.C. (1992) *J. Cell Biol.*, **118**, 1–9.
10 Izaurralde,E., Stepinski,J., Darzynkiewicz,E. and Mattaj,I.W. (1992) *J. Cell Biol.*, **118**, 1287–1295.
11 Jarmolowski,A., Boelens,W.C., Izaurralde,E. and Mattaj,I.W. (1994) *J. Cell Biol.*, **124**, 627–635.
12 Izaurralde,E. and Mattaj,I.W. (1995) *Cell*, **81**, 153–159.
13 Ohno,M. and Shimura,Y. (1990) *Methods Enzymol.*, **181**, 209–215.
14 Ohno,M., Kataoka,N. and Shimura,Y. (1990) *Nucleic Acids Res.*, **18**, 6989–6995.
15 Kataoka,N., Ohno,M., Kangawa,K., Tokoro,Y. and Shimura,Y. (1994) *Nucleic Acids Res.*, **22**, 3861–3865.
16 Izaurralde,E., Lewis,J., McGuigan,C., Jancowska,M. Darzynkiewicz,E. and Mattaj,I.W. (1994) *Cell*, **78**, 657–668.
17 Zervos,A.S., Gyuris,J. and Brent,R. (1993) *Cell*, **72**, 223–232.
18 Dignam,J.D., Lebovitz,R.M. and Roeder, R.G. (1983) *Nucleic Acids Res.*, **11**, 1475–1489.
19 Saraste,M., Sibbald,P.R. and Wittinghofer,A. (1990) *Trends Biochem. Sci.*, **15**, 430–434.
20 Neugebauer,K.M., Stolk,J.A. and Roth,M.B. (1995) *J. Cell Biol.*, **129**, 899–908.