

Npw38, a novel nuclear protein possessing a WW domain capable of activating basal transcription

Akihiko Komuro, Mihoro Saeki and Seishi Kato*

Kato Cytoprotein Network Project, ERATO, Japan Science and Technology Corporation (JST), c/o Sagami Chemical Research Center, Nishi-Ohnuma 4-4-1, Sagamihara, Kanagawa 229-0012, Japan

Received February 15, 1999; Revised and Accepted March 7, 1999

DDBJ/EMBL/GenBank accession no. AB016533

ABSTRACT

We have found a novel cDNA encoding a 265 amino acid protein possessing a WW domain in our full-length cDNA bank. The WW domain was sandwiched between an acidic region and an acidic-basic amino acid repetitive region. *In vitro* transcription/translation of the cDNA produced a 38 kDa product that was also found in the cell lysate by western blot analysis. Thus this protein is named the nuclear protein containing a WW domain with a molecular mass of 38 kDa, Npw38. Immunofluorescence studies and expression of a fusion protein to a green fluorescent protein revealed that this protein is localized in the nucleus. Npw38 was shown to be capable of binding to a poly(rG) resin. Interestingly, the WW domain of Npw38 was found to function as a transcriptional activator in CHO cells using the GAL4 DNA-binding fusion system. Furthermore, the WW domains of human YAP and Pin1 were demonstrated to have a similar transcription-promoting activity. Combined mutation of the conserved first and second Trp residues and a hydrophobic triplet of TyrTyrTrp in the WW domain of Npw38 abolished the transcription-promoting activity, but single mutations of these sites did not. These results suggest that some WW domains potentially possess transcription-promoting activity in mammalian cells.

INTRODUCTION

We have been constructing a Homo•Protein cDNA Bank composed of human full-length cDNA clones (1). The full-length sequenced cDNA clones were subjected to functional analysis. To find novel proteins possessing a protein–protein interaction module in this bank, a computer-assisted search was performed. We found several novel cDNAs encoding proteins possessing modules such as WW domains, an SH2 (Src homology 2) domain, an SH3 domain, a proline-rich SH3 domain-interacting module or a cysteine-rich domain. In the present study, we focus on a cDNA clone encoding the WW domain-containing protein.

The WW domain is a new protein–protein interaction module (2–5) analogous to SH2, SH3 or PH (pleckstrin homology) domains (6). It consists of 35–40 amino acids with two highly conserved tryptophans. Like other modules of intracellular

proteins, the WW domain is present in a variety of proteins including the signal transducing protein YAP65 (7), the cytoskeletal protein dystrophin (2), the ubiquitin-protein ligase NEDD4/RSP5 (4,8–10), the peptidyl-prolyl isomerases essential for regulation of mitosis, Pin1/dodo/Ess1 (11,12), the formin-binding proteins (FBP) (13,14) and the RNA-binding protein FCA (15). The WW domain binds proline-rich peptides similar to the SH3 domain (7), but the three-dimensional structure of the WW domain of YAP65 is different from that of the SH3 domain (16). The WW domains have been shown to bind peptides with the consensus sequence PPXY (where X signifies any amino acid) (7), the PPLP motif (13) or the PGM motif (17).

Recently, a few lines of evidence that the WW domains may be involved in transcriptional regulation have been reported. A hematopoietic transcription factor p45/NF-E2 has a PPXY motif capable of binding to a specific set of WW domains with which the C-terminal domain of the RNA polymerase II interacts (18). Mutation of the PPXY motif in the p45 subunit inhibited the transactivation function of NF-E2 (19), suggesting that its interaction with the WW domain is important for NF-E2-mediated transactivation. Another example is FE65, identified as a brain-enriched protein which possesses the WW domain. Its N-terminal 232 amino acid residues containing the WW domain caused transcriptional activation when expressed as a GAL4 fusion protein (20). However, the precise site necessary for transcriptional activation has not been determined. Rather, recent reports have shown that the WW domain of FE65 binds to cytosolic Mena (21), the mammalian homolog of the *Drosophila* Enabled gene, and one of the phosphotyrosine binding domains of FE65 interacts with the transcription factor CP2/LSF/LBP1 (22).

In the present paper, we report the characterization of a novel WW domain-containing protein and its functional analysis. This protein is called the nuclear protein containing a WW domain with a molecular mass of 38 kDa, or Npw38. During a search of proteins binding to the WW domain of Npw38 using a yeast two-hybrid system, we found that the WW domain of Npw38 has a transcription-promoting activity in the yeast cell. Furthermore, this WW domain and those of the other WW domain-possessing proteins such as human YAP and Pin1 were shown to activate basal transcription in CHO cells using the GAL4 DNA-binding protein fusion system. This study suggests the existence of a WW domain-mediated regulation of basal transcription in mammalian cells.

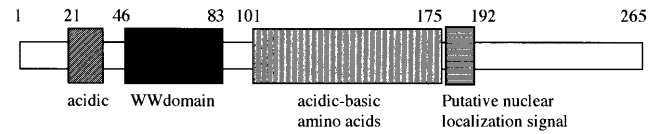
*To whom correspondence should be addressed. Tel: +81 427 42 5091; Fax: +81 427 49 7631; Email: seishi@sagami.ne.jp

A

```

AGTTACTTTCAGGCTCGGGGAGTGAAGGCCTCGTTGAGAGAAGTCTCATTCCGGTGTGTTT -111
GGGAAGAGAGTCTGTGGGCCAGGTATCGTAGCGGCACAGAGAGAGACGGGCGGTGT -51
GACAGCCTTCCACTACCTGCACAGAGTATTGGTCTGTCTGCATCAGCTATGCCCGCTGC 10
                                     M P L P 4
CCGTGCGCTGCAGACCCGCTTGGCCAGAGAGGCATCCTCAAACATCTGGAGCCGTGAAC 70
V A L Q T R L A K R G I L K H L E P E P 24
CAGAGGAAGAGATCATTGCCGAGGACTATGACGATGATCCTGTGGACTACGAGGCCACCA 130
E E E I I A E D Y D D D P V D Y E A T R 44
GGTTGAGAGGCTACCACCAAGCTGGTACAAGTGTTCGACCCCTTCCGCGGCTCCCTT 190
L E G L P P S W Y K V F D P S C G L P Y 64
ACTACTGGAATGCAGACACAGACCTTGTATCCTGGCTCTCCCAACATGACCCCAACTCCG 250
Y W N A D T D L V S W L S P H D P N S Y 84
TGGTTACCAAAATCGGCCAAGAAGCTCAGAAGCAGTAATGCAGATGCTGAAGAAAAGTGG 310
V T K S A K K L R S S N A D A E E K L D 104
ACCGGAGCCATGACAAGTCGGACAGGGCCATGACAAGTCGGACCCGACCCATGAGAAAAC 370
R S H D K S D R G H D K S D R S H E K L 124
TAGACAGGGCCAGCACAAGTCAGACCGGGCCACGACAAGTCTGACAGGGATCGAGAGC 430
D R G H D K S D R G H D K S D R D R E R 144
GTGGTATGACAAGGTAGACAGAGAGAGAGCGAGACAGGAAACGGGATCGGGACCCGCG 490
G Y D K V D R E R E R D R E R D R G R G 164
GGTATGACAAGGCAGACCGGGAAGAGGGCAAGAAGAACGGCCACCATCGCCGGGAGGAGC 550
Y D K A D R E E G K E R R H H R R E E L 184
TGGCTCCCTATCCCAAGAGCAAGAAGGCAGTAAGCCGAAAGGATGAAGAGTTAGACCCCA 610
A P Y P K S K K A V S R K D E E L D P M 204
TGGACCTAGTCACTCAGACGCCCCCGGGGCACGTGTCACAGGACTCCCAAGC 670
D P S S Y S D A P R G T W S T G L P K R 224
GGAATGAGGCCAAGACTGGCGCTGACACACAGCAGCTGGGCCCTCTTCCAGCAGCGGC 730
N E A K T G A D T T A A G P L F Q Q R G P 244
CGTATCCATCCCCAGGGCTGTGCTCGGGCCAATGCAGAGGCCTCCCGAACCAAGCAGC 890
Y P S P G A V L R A N A E A S R T K Q Q 264
AGGATTGAAGCTTCGGCCTCCCTGGCCCTGGGTTAAATAAAGCTTTCTGGTATGATCTG 950
D * 265
    
```

B



C

name/species	pos.	sequence	acc. no.
Npw38/Human	46	EGLPESWYKVFDPSCCLPYVWADLDLWSWLSRHPDPSW	
Yap/Human	171	VPLPAGMEMAKTSSGQRNPLNHIDQTTTQDERKAMLS	P46937
Yap/Mouse-1	156	VPLPAGMEMAKTSSGQRNPLNHNDQTTTQDERKAMLS	P46938
Yap/Mouse-2	215	GPLPDCMEQAMTQDSEVYVTHHKNTTSWLDRLDRPF	P46938
Nedd4/Human-1	218	SLPPECEERQDILLRQYVWHSRRTQKRRTPQDNL	P46934
Nedd4/Human-2	375	SLPPECEERQDERSRQVWDSRRTTTRRTPQATL	P46934
Nedd4/Human-3	448	GLPPECEVYHAGSRPFTDHNKTTTDEDRKIPA	P46934
Nedd4/Human-4	500	GPLPECEERTHTDGRIFVTHNKIKRTQDEDRLEVA	P46934
Rsp5/Yeast-1	229	GRPPECEERTTNPGRVVDHNKRTTQKRRPLDQTE	P39940
Rsp5/Yeast-2	331	GPLPECEERQDTEGRVVDHNKRTTQKRRRQOYI	P39940
Rsp5/Yeast-3	387	GPLPECEERLNTARVVDHNKRTTQKRRPLDQTE	P39940
Dmd/Human	3055	TSVQGPERRAISNKRANFVDHNKRTTQKRRKTELY	P11532
IQGAP/Human	679	GDNNSKVRHWVKSGYVVDHNLEQEGGDEHPNFTVON	P46940
Pin1/Human	4	DKLPECEERMSRSTSRVVDHNINASQKRRSCNSSS	U49070
dodol/Drome	5	GPLPDCMEERTRSTGMSVVDHNKRSQKRRDQTEPAK	U35140
Ess1/Yeast	29	TGLPTEPTVYRSKSKRREFFMPEAKHSQKRRPECTNKD	P22696
FBP11-1/Mouse	1VTEHKSSESRVVDHNLEQKSTQKRRDRLKTP	U40747
FBP11-2/Mouse	36	LLSKCPKTKYKSDSQRVVDHNLEQKSTQKRRKAK.....	U40747
FBP30/Mouse	1RQVMDVETLSCVVDHNLEQKSTQKRRK.....	U40750
PE65/Rat	42	SDLPAGMEERQDTEGRVVDHNLEQKSTQKRRPEGRASPS	P46933
consensus		..LXKGExxxxxxKxVVDHNKTTTxxx.....	

D

```

99 AEKLDKDR
106 SHDKSDR
113 GHDKSDR
120 SHEKLDKDR
127 GHDKSDR
134 GHDKSDRDRER
145 GYDKVDRERDRERDRDR
164 GYDKADRE
    
```

Figure 1. Structure of Npw38 cDNA. (A) The nucleotide sequence of Npw38 cDNA and the deduced amino acid sequence. The putative polyadenylation signal is underlined. (B) Schematic representation of the domains in Npw38 protein. (C) Sequence alignment among the WW domains. The identical amino acid residues are indicated by black boxes based on the sequence of Npw38. The position of the first amino acid and the accession number of each sequence in the protein or nucleotide database are indicated. Asterisks indicate the amino acid residues conserved among all sequences. (D) Alignment of repeat sequences in the acidic-basic amino acid repetitive region.

MATERIALS AND METHODS

cDNA cloning

cDNA libraries were prepared from mRNAs of human tissues and cell lines using a DNA-RNA chimeric oligo-capping method (1). All cDNA clones in the bank have already been partially or fully sequenced. Using a consensus sequence of the WW domain [LPxGWExxxxxGxx(Y/F)(Y/F)x(N/D)HxTxT(T/S)xWxxP, where x represents an arbitrary amino acid], a computer-assisted search was performed to identify the cDNA encoding WW domains from the bank. As a result, a novel cDNA clone, HP10345, encoding protein Npw38 was found. This cDNA originated from a human gastric adenocarcinoma cDNA library.

Sequencing

The sequencing reaction was performed using a dye primer cycle sequencing kit (Perkin-Elmer). The reaction samples were electrophoresed on a DNA sequencer (Perkin-Elmer ABI, model

377). The computer analysis was done using the program GENETYX-MAC (Software Development).

In vitro transcription/translation

The *in vitro* transcription/translation was carried out in a reaction mixture containing [³⁵S]methionine (Amersham) using a T_NT coupled reticulocyte lysate system (Promega) according to the manufacturer's instructions. The translated product was analyzed by 10–20% SDS-PAGE.

Northern blot analysis

Multiple Tissue Northern Blot filters (Clontech) were hybridized with a ³²P-labeled cDNA insert prepared from the Npw38 cDNA using a random primer labeling kit (Takara). After washing with 0.1× SSC, 0.05% SDS at 50°C, the blots were exposed to an X-ray film for 36 h at –80°C with an intensifying screen.

Localization analysis of GFP-Npw38 fusion protein

To produce an enhanced green fluorescent protein (EGFP) fusion protein with Npw38, the coding region of the Npw38 cDNA was amplified by PCR with primers containing either an *EcoRI* or a *SalI* site and subcloned into pEGFP-C1 (Clontech). COS7 cells were transfected with 0.5 μ g of pEGFP-C1-Npw38 or pEGFP-C1 as a control according to the liposome-mediated method (23). After a 24 h culture, the medium was removed and the cells were washed three times with phosphate-buffered saline (PBS). The cells were fixed with 4% paraformaldehyde/PBS and viewed under a fluorescent microscope (Nikon).

Preparation of GST-Npw38 fusion protein

The entire coding region of the Npw38 cDNA was amplified by PCR using primers with an *EcoRI* or a *SalI* site. The PCR product was digested with restriction enzymes and subcloned into the *EcoRI* or *SalI* site of pGEX-5X-1 (Pharmacia). The constructed vector was introduced into *Escherichia coli* BL21 and expression was induced for 3 h by adding isopropyl-1-thio- β -D-galactopyranoside to a final concentration of 0.2 mM. The proteins were affinity purified with glutathione-Sepharose beads and eluted with reduced glutathione as described in the manufacturer's instruction manual (Pharmacia).

Antibody preparation

The anti-Npw38 antiserum was prepared in a rabbit using a bacterially expressed GST-Npw38 fusion protein. Anti-GST antibodies were removed from the serum by GST affinity chromatography using an NHS-activated column (Pharmacia) and the anti-Npw38 antibodies were purified using a GST-Npw38 affinity column. The absence of reactivity by the purified anti-Npw38 antibodies against GST was confirmed by western blot analysis.

Western blot analysis

HeLa cells were lysed in RIPA buffer (10 mM Tris-HCl, pH 7.4, 1% NP-40, 0.1% sodium deoxycholate, 0.1% SDS, 150 mM NaCl, 1 mM EDTA, 2 mM phenylmethanesulfonyl fluoride, 1 μ g/ml aprotinin), separated by SDS-PAGE and transferred onto a PVDF membrane. The protein was probed with the anti-Npw38 antibody and visualized by the ECL method (Amersham).

Immunofluorescence staining

HeLa cells were grown on 4-chamber culture slides (Beckton Dickinson). The cells were briefly washed with PBS and fixed with 4% paraformaldehyde in PBS for 15 min at room temperature. The fixed cells were washed, permeabilized with 0.2% Triton X-100 in PBS, incubated with the anti-Npw38 antibody for 15 min, washed with PBS and stained with rhodamine-conjugated secondary antibody and DAPI (4',6-diamidino-2-phenylindole) for 15 min.

Transactivation assay in yeast cells

The entire coding region of Npw38 or the coding region of the WW domain (position 39–87) was amplified by PCR with primers containing either an *EcoRI* or a *SalI* site and subcloned into the *EcoRI* or *SalI* site of vector pAS2-1 (Clontech) to express a GAL4 DNA-binding domain fusion gene. The resulting

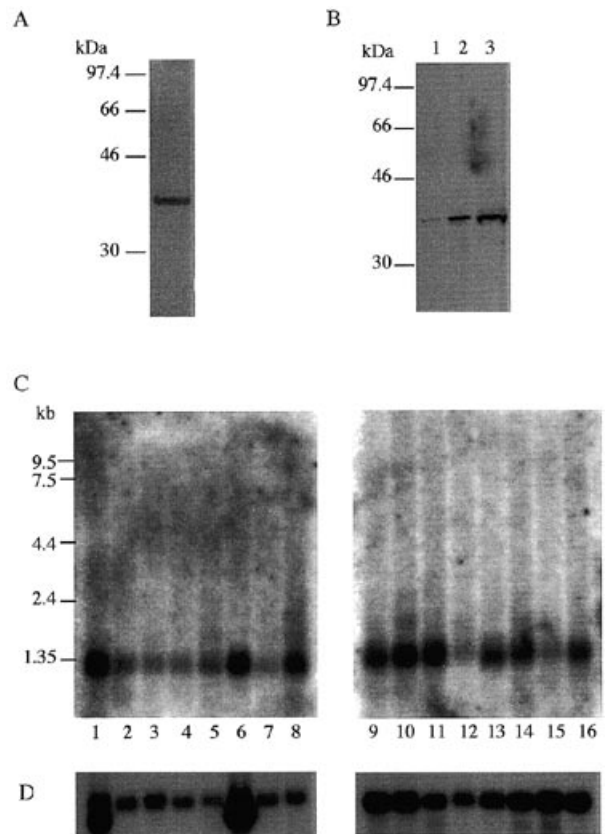


Figure 2. Expression of Npw38 protein and mRNA. (A) SDS-PAGE analysis of *in vitro* translation product. (B) Western blot analysis. Cell extracts of different volumes were applied: 1, 10 μ l; 2, 20 μ l; 3, 50 μ l. (C) Northern blot hybridization. 1, heart; 2, brain; 3, placenta; 4, lung; 5, liver; 6, skeletal muscle; 7, kidney; 8, pancreas; 9, spleen; 10, thymus; 11, prostate; 12, testis; 13, ovary; 14, small intestine; 15, colon; 16, peripheral blood leucocytes. (D) Northern blot hybridization. The same filters in experiment (C) were probed with β -actin cDNA as a control.

plasmids, pAS-WW and pAS-Npw38, were introduced into yeast strain Y190 and transformants were grown on Trp-deficient plates. The plasmid pAS-laminin supplied with the kit was used as the negative control. As the positive control, Y190 cells carrying pCL that expresses a complete GAL4 protein were grown on Leu-deficient plates. The plates were incubated at 30°C for 3 days and a β -galactosidase colony-lift assay was performed according to the manufacturer's instructions (Clontech).

Transactivation assay in mammalian cells

A TransAct™ Assay kit (Clontech) was used to detect transcriptional activation in mammalian cells. An EST clone encoding a partial coding region of human YAP (GenBank accession no. N33982) and that encoding the entire coding region of human Pin1 (accession no. H18274) were obtained from Genome Systems, Inc. The coding regions containing the test sequence of the WW domain-bearing protein were amplified by PCR using primers with either an *EcoRI* or a *SalI* site. The obtained cDNA fragments were subcloned into vector pM to express a fusion protein to a GAL4 DNA-binding domain under the control of the SV40 promoter. The substitution of amino acid residues in the WW domain were carried out by using a Quick Change™ Site-Directed

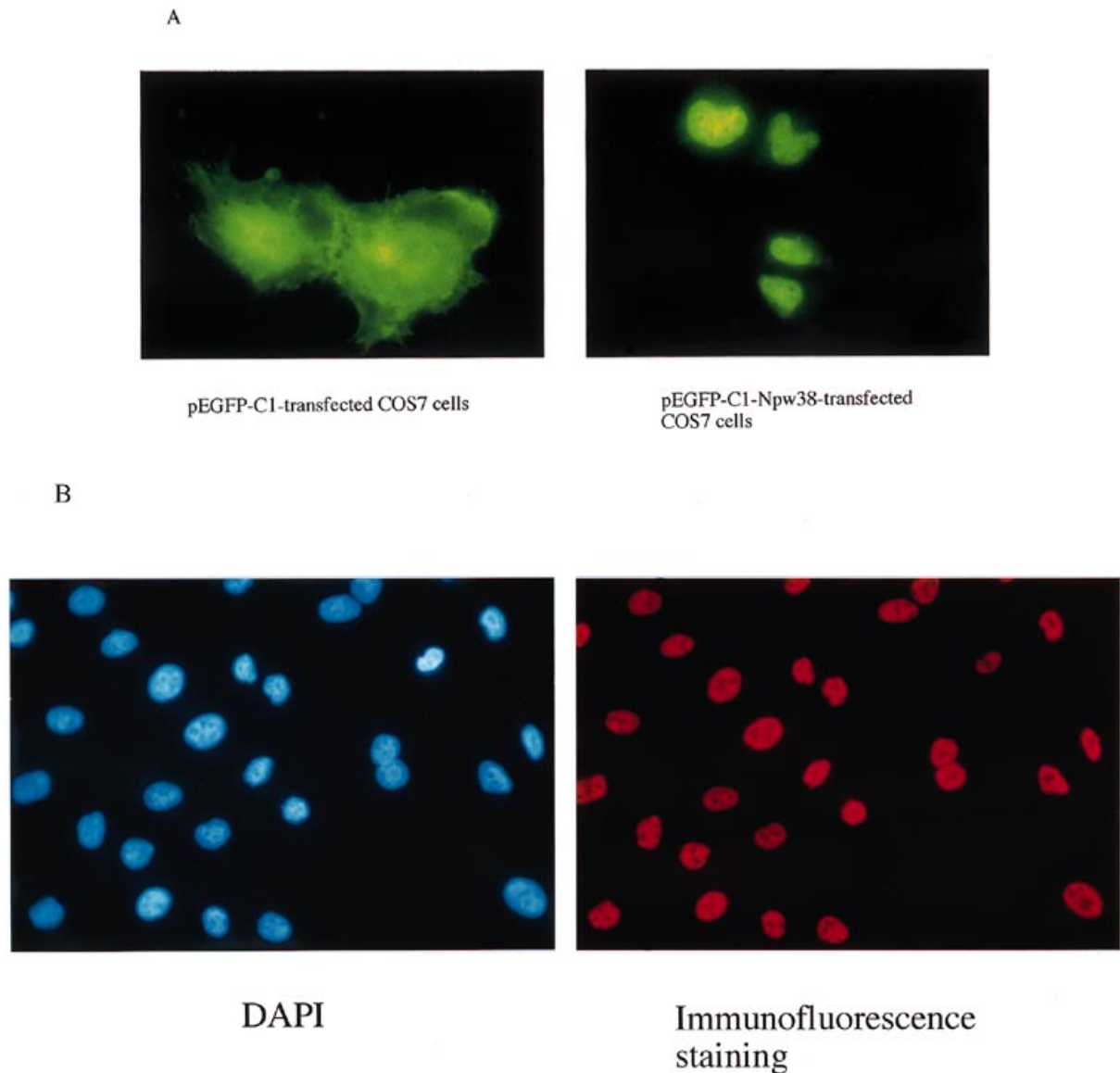


Figure 3. Localization of Npw38. (A) Localization of GFP fusion protein. COS7 cells transfected with pEGFP-C1 (left) or pEGFP-C1-Npw38 (right) were visualized by fluorescence microscopy. (B) DAPI and immunofluorescence staining. HeLa cells fixed with paraformaldehyde were incubated with anti-Npw38 antibody and subsequently rhodamine-conjugated secondary antibody and DAPI. The same area was viewed.

Mutagenesis Kit (Stratagene). A pM-VP16 vector and an empty pM vector were used as the positive and negative controls, respectively. The reporter plasmid pG5CAT has a chloramphenicol acetyltransferase (CAT) gene preceded by five GAL4 binding sites and the E1b TATA box. The pM plasmid (0.5 μ g) was introduced into CHO cells with pG5CAT (1 μ g) and pSVGal (1 μ g; Promega) by the lipofection method. The cells were harvested 24–36 h after transfection. Cell extracts were prepared and assayed for CAT activity using a CAT assay kit (Promega). The CAT activity was quantified using a BAS 1500 imaging analyzer (Fuji Film) and normalized to the β -galactosidase activity.

Nucleic acids binding assay

The Npw38 protein was prepared by *in vitro* transcription/translation. Nucleic acids binding was examined using 3 μ l of the

³⁵S-labeled translation product diluted in 500 μ l of KHN buffer (150 mM KCl, 20 mM HEPES, pH 7.9, 0.01% NP-40). Twenty microliters of poly(rA), poly(rC), poly(rG) or poly(rU) agarose beads, calf thymus single-stranded DNA or double-stranded DNA cellulose beads (all beads purchased from Sigma) were then added and rotated for 20 min at room temperature. The beads were centrifuged and then washed five times with 1 ml of KHN buffer. Bound proteins were dissolved in the SDS sample buffer and analyzed by SDS-PAGE.

RESULTS

Primary structure of Npw38

The Npw38 cDNA has an open reading frame capable of encoding a 265 amino acid protein (Fig. 1A). A characteristic

feature of the predicted amino acid sequence is the existence of an N-terminal acidic region (position 21–41) followed by a WW domain (position 46–83) and an acidic-basic repetitive region (position 101–175) as illustrated in Figure 1B. An alignment of the WW domain sequences with other known WW domains shows that the WW domain of Npw38 possesses two Trp and one Pro conserved in all members of this family (Fig. 1C). When the region inserted between the first Trp and the last Pro was compared among the WW domains, that of Npw38 has some characteristics. Similar to Pin1/Ess1/dodo and FBP30, one insertion, Cys in Npw38, was observed in this region. Npw38 possesses eight hydrophobic amino acid residues in this relatively acidic region, while other WW domains have rather hydrophilic and more basic residues. Especially noteworthy is that the second Trp in Npw38 is surrounded by three hydrophobic amino acids. This region is hydrophilic in most WW domains, except for FBP30, which also shows an acidic feature like Npw38. These characteristics may contribute to determining the ligand binding specificity of Npw38.

The acidic-basic amino acid repetitive region contains eight repeats of a consensus sequence, GHDKSDR. The first repeat sequence lacks the two N-terminal amino acid residues and the last two repeat sequences are separated by the insertion of (D/E)R repetitive sequences. The sequence GHDKSDR is not found in known protein sequences, but a (D/E)R repetitive sequence exists in several nuclear proteins, including the RD protein (24) and 70 kDa U1 small ribonuclear protein (25). The repetitive sequence was followed by two clusters of basic amino acids RRHHRREELAPYPKSKK at positions 176–192, which seem to form a bipartite nuclear localization signal (26,27).

Expression of Npw38 protein and mRNA

The cDNA was transcribed and translated *in vitro* in a rabbit reticulocyte lysate. As shown in Figure 2A, the translation product gave a main band of 38 kDa on SDS-PAGE that was larger than the 30.5 kDa expected from the open reading frame of the cDNA. Since a 38 kDa protein was also detected in a HeLa cell lysate by western blot analysis (Fig. 2B), the apparent molecular mass of Npw38 seems to be 38 kDa. This kind of increase in the apparent molecular mass has been observed in the case of various proteins possessing a highly charged region.

A major transcript of ~1.3 kb was detected by northern blot analysis in every human tissue tested. The expression of Npw38 mRNA was relatively high in heart, skeletal muscle, pancreas, spleen, thymus, prostate, ovary, small intestine and peripheral blood leucocytes (Fig. 2C). These results reveal that the transcripts are ubiquitously present in various tissues.

Localization of Npw38 in cells

To analyze the subcellular localization of Npw38, a GFP fusion-based method was used. COS7 cells were transfected with an expression plasmid coding for EGFP alone or for an EGFP-Npw38 fusion protein, and then localization of a transiently expressed product was determined using fluorescence microscopy. Cells transfected with a construct encoding EGFP alone displayed diffuse fluorescence throughout the cells. On the other hand, an intense nuclear fluorescent pattern was observed in the cells expressing EGFP-Npw38 (Fig. 3A).

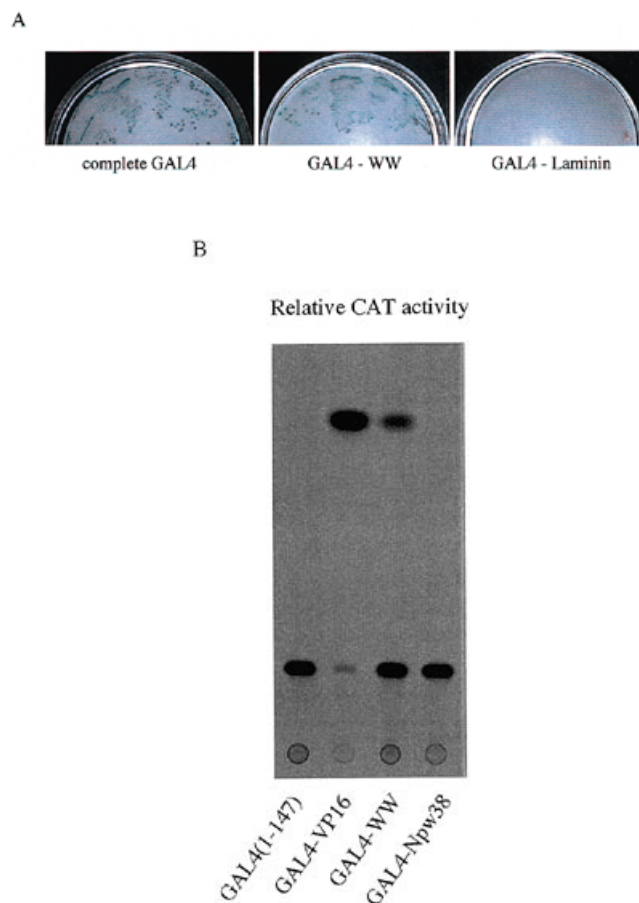


Figure 4. Activation of transcription by the WW domain of Npw38 in yeast and mammalian cells. (A) β -Galactosidase colony-lift assay in *S.cerevisiae*. pAS-WW, pAS-Npw38, pAS-laminin as a negative control or pCL (complete GAL4) as a positive control was introduced into yeast strain Y190 and transformants were plated onto selection plates. The plates were incubated at 30°C for 3 days and a β -galactosidase colony-lift assay was performed. (B) Transactivation assay in mammalian cells. CHO cells were transfected with the corresponding plasmids. Cells were collected 36–60 h after transfection and assayed for CAT activity.

The above-mentioned results were confirmed by immunofluorescence staining. When HeLa cells were incubated with anti-Npw38 antibody followed by a rhodamine-conjugated secondary antibody, an intense fluorescence was observed only in nuclei, as well as by DAPI staining (Fig. 3B). These data are in agreement with the result that a 38 kDa protein was detected only in the nuclear extract but not in the cytosolic fraction by the western blot (data not shown). All these results reveal that Npw38 is localized in nuclei, as expected from the existence of the nuclear localization signal in the deduced sequence.

The WW domain of Npw38 functions as a transcriptional activator *in vivo*

To identify the proteins interacting with the WW domain of Npw38, we planned a yeast two-hybrid screening of a HeLa-S3 cDNA library using the WW domain as bait. Before performing the two-hybrid assay, autonomous activation of the β -galactosidase reporter gene by the hybrid construct was verified as suggested by

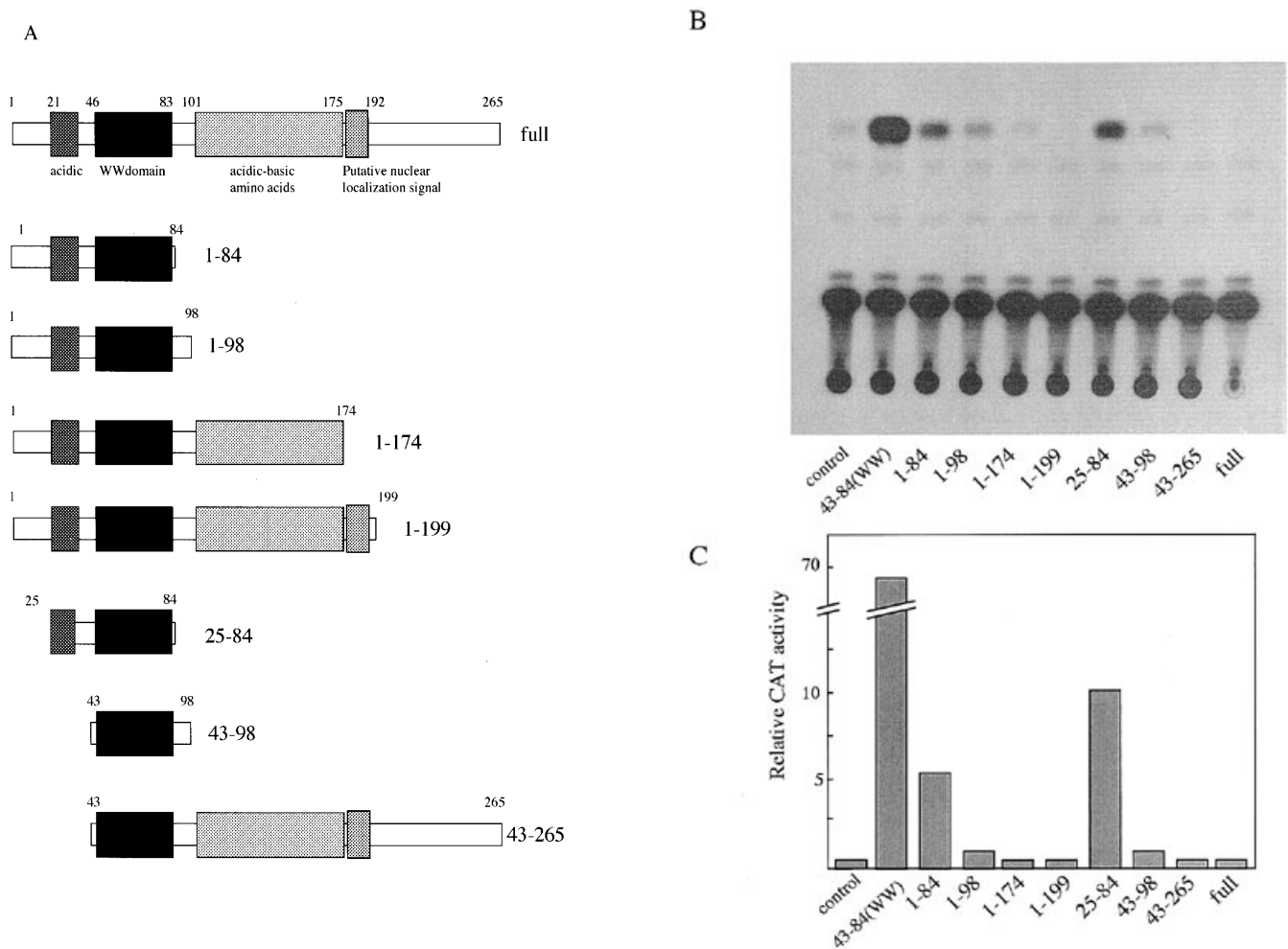


Figure 5. Transactivation assay of the Npw38 deletion mutants. (A) Schematic representation of the Npw38 deletion mutants used in this study. Each fragment was fused to the GAL4 DNA-binding domain (1–147) in the pM plasmid. (B) CAT assays were performed as described in Figure 4B with extracts from CHO cells co-transfected with pM plasmids encoding fusion proteins. (C) The relative CAT activities normalized to β -galactosidase activity are given as the average values of three experiments. The CAT activity of CHO cells carrying pM is defined as 1.0 (control).

the manufacturer. Unexpectedly, autonomous activation of the reporter gene was observed on expression of the WW domain fusion gene (Fig. 4A) as well as a complete GAL4 construct, while not on that of a laminin construct. When the intact Npw38 was used as bait, only faint activation of the reporter gene was observed during long-term incubation, suggesting that some part of Npw38 represses the activity. These results imply that the WW domain of Npw38 can activate transcription directly or via association with the transcription machinery.

Next, activation of transcription by the WW domain in a mammalian system was examined using a reporter plasmid carrying five GAL4 binding sites followed by the adenovirus E1b TATA box upstream of the CAT gene. The cDNA fragment encoding the entire Npw38 or the WW domain was inserted into vector pM carrying the GAL4 DNA-binding domain gene under the transcriptional regulation of the SV40 early promoter. After the reporter plasmid and pM fusion vector were introduced into CHO cells, the CAT activity of the CHO cell lysate was measured. The GAL4–WW hybrid could activate transcription of the CAT gene ~50–70-fold compared with the GAL4 DNA-binding domain alone (Fig. 4B). In contrast, no activation of the reporter gene was

observed when the GAL4–Npw38 hybrid expression vector was introduced into the cells.

To investigate which region in Npw38 represses the WW domain-mediated transactivation, we prepared mammalian expression vectors for fusion proteins between the GAL4 DNA-binding domain and various deletion mutants of Npw38 as described in Figure 5A. Only fusion proteins having a short region containing the WW domain (positions 1–84, 1–98, 25–98 and 43–98) could weakly activate transcription, whereas no activation was observed in other longer constructs (Fig. 5B and C). The activation seems to be repressed by both flanking regions of the WW domain and strongly by the short region of positions 84–98. These results suggest that some factor of the sequences surrounding the WW domain could be involved in the regulation of the transcriptional activity of Npw38.

Transcriptional activation by other WW domains

To examine whether the transcription-promoting activity was common to WW domains, other WW domains were tested by this assay. The WW domain-coding region of human YAP or human

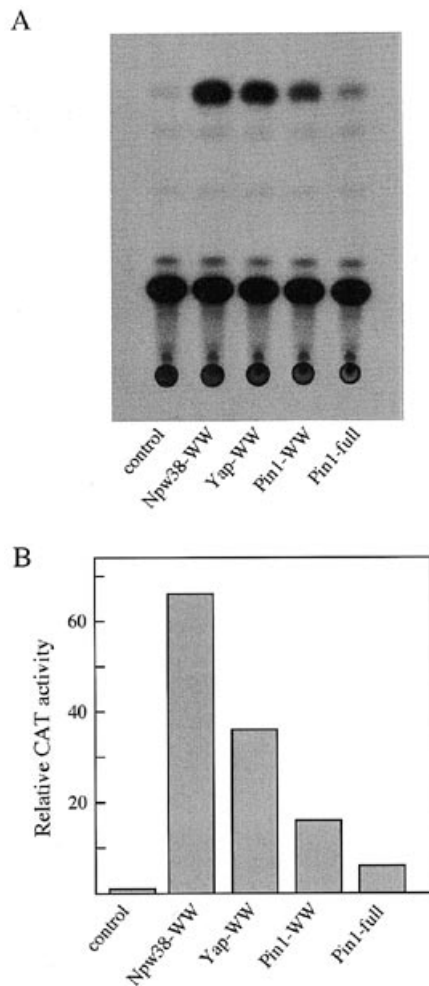


Figure 6. Transactivation assay of the WW domains of YAP and Pin1. (A) CAT assays were performed as described in Figure 4B with extracts from CHO cells co-transfected with pM plasmids fused to the WW domain coding region of human YAP (position 162–216), that of human Pin1 (position 1–49) or the entire coding region of human Pin1 (position 1–163). (B) The relative CAT activities normalized to β -galactosidase activity are given as the average values of three experiments. The CAT activity of CHO cells carrying pM is defined as 1.0 (control).

Pin1 was cloned into the pM vector and the CAT assay was performed. GAL4-YAPWW and GAL4-Pin1WW could activate transcription of the CAT gene ~40- and 20-fold, respectively, compared with the GAL4 DNA-binding domain control (Fig. 6). Furthermore, in contrast to Npw38, the entire region of Pin1 could also activate transcription ~6-fold. These results imply that some WW domains potentially have a basal transcription-promoting activity.

Mutation study of the WW domain

We next explored the role of conserved amino acid residues of the WW domain in transcriptional activation. The two Trp and the Pro in the WW domain of Npw38 were replaced with Ala or Gly using site-directed mutagenesis and the CAT assay was performed. Figure 7 shows that neither single substitution of the first Trp to Ala (W52A), the second Trp to Ala (W75A) nor the Pro to Gly

(P78G) abolished the activity. Rather, the W52A mutant enhanced the activity ~2-fold compared with the wild-type WW domain. Double mutation (W52A, W75A) reduced the activity to 60% of the wild-type. Single and multiple mutations in a hydrophobic amino acid triplet (Y64A, Y65A, W66A) did not abolish the activity (data not shown about single mutation). The combination of mutations (W52A, Y64A, Y65A, W66A) reduced the activity to 20%. Furthermore, when all these residues were simultaneously substituted (W52A, Y64A, Y65A, W66A, W75A), the transcription-promoting activity was lost to the control level. These results imply that single mutation of the conserved amino acid residues did not affect the entire structure of the WW domain which may be involved in transactivation.

Nucleic acid binding ability of Npw38

Since the RGH repeat sequence in the acidic-basic amino acid repetitive region is similar to an RGG box sequence which is one of the RNA binding motifs (28), we examined the nucleic acid binding ability of Npw38 using a semi-specific assay which has been extensively used to study the RNA-binding properties of hnRNP and other proteins (15,29,30). When *in vitro* translated protein was incubated with nucleic acid resins, Npw38 bound preferentially to a poly(rG) sequence (Fig. 8), implying that Npw38 is an RNA-binding protein. The binding of Npw38 to a single-stranded or double-stranded DNA was not detected under the same conditions. The result that GST-Npw38 (position 81–265) is bound to poly(rG) (data not shown) suggests that the acidic-basic repetitive region is a putative RNA-binding site.

DISCUSSION

We have found a cDNA encoding a novel nuclear protein possessing a WW domain, Npw38, in the Homo•Protein cDNA bank. Interestingly, the WW domain of Npw38 was able to activate transcription in both yeast and mammalian cells *in vivo*, when the WW domain was recruited to the promoter region of a vector plasmid in a transactivation assay. However, intact Npw38 caused no activation because of the existence of the flanking inhibitory sequence. Thus the WW domain of Npw38 may ordinarily be masked and inactivated by a factor capable of associating with Npw38 or by the remaining part of Npw38 itself. Furthermore, we demonstrated that the WW domains of YAP and Pin1 possess similar transcription-promoting activity. These results imply that the transcriptional machinery consists of a protein capable of interacting with these WW domains. One candidate for the interacting protein may be RNA polymerase II in which the C-terminal domain has been shown to bind to various WW domains (18). Alternatively, the interaction may be mediated by a transcriptional co-activator possessing a Pro-rich motif. At present, however, it remains unclear whether these interactions have a physiological meaning.

The mutation analysis of the amino acid residues in the WW domain of Npw38 revealed that any single mutation of the conserved sequence in the WW domain did not abolish the transactivation-promoting activity. These results are partly in agreement with the results obtained with *in vitro* systems such as far western blotting analysis or co-precipitation of GST-YAP WW domain fusion protein and PY motif peptides. Chen *et al.* reported that a single mutation of the first Trp (W177F) or that in the hydrophobic triplet region (Y188F, F189Y) did not abolish

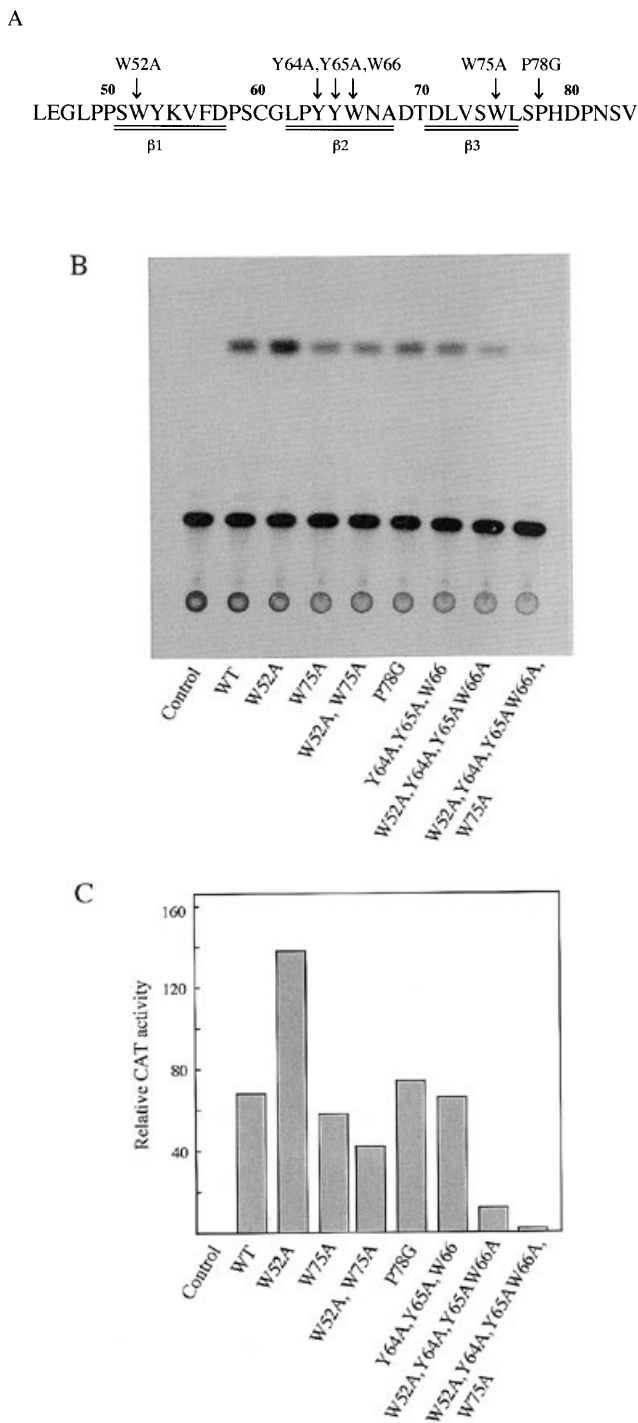


Figure 7. Effects of substitution of the conserved amino acid residues in the WW domain on the transcription-promoting activity. (A) Mutant WW domains used in the assay. Selected conserved residues were substituted using site-directed mutagenesis of the WW domain of Npw38 subcloned in the pM vector. The underlining indicates the β -sheet strands expected from the solution structure of a YAP WW domain. (B) CAT assays were performed as described in Figure 4B. (C) The relative CAT activities normalized to β -galactosidase activity are given as the average values of three experiments. The CAT activity of CHO cells carrying pM is defined as 1.0 (control).

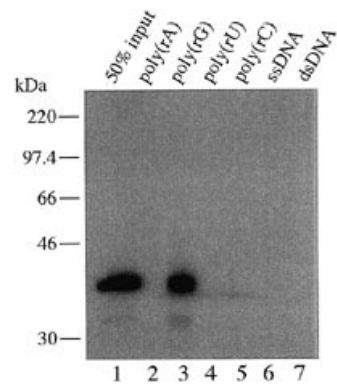


Figure 8. Binding of Npw38 to nucleic acids. The 35 S-labeled product obtained by *in vitro* transcription/translation of Npw38 cDNA was incubated with agarose beads carrying various nucleic acids including poly(rA) (lane 2), poly(rG) (lane 3), poly(rU) (lane 4), poly(rC) (lane 5), single-stranded DNA (lane 6) and double-stranded DNA (lane 7). Bound proteins were analyzed by SDS-PAGE. An amount corresponding to 50% input of each assay was loaded on lane 1.

the binding ability to the PY peptide in solution, but mutation of the second Trp (W199F) or the Pro (P202A) did (31). The W177F mutant was bound at a slightly higher affinity than the wild-type on the filter. A similar increase in activity was observed in the corresponding W52A mutant of the Npw38 WW domain. No effect of mutation of the second Trp or the Pro on the activity in our case may be attributed to the differences in the WW domain and that of the assay systems. Our system observed an interaction between the Npw38 WW domain and an unknown ligand protein in the cell under physiological conditions. On the other hand, Chen *et al.* observed an interaction between the YAP WW domain and PY motif peptides on a filter or in solution. A discrepancy in the effect of the mutation was reported in an *in vitro* system as well. The Y188F mutant of the YAP WW domain could bind to the PY peptides on the filter but not in solution (31). The decrease in the activation by combined mutation of the conserved residues may result from a conformation change in the WW domain. According to the structure of the WW domain of YAP, the hydrophobic triplet is adjacent to the first and second Trp (16), so interactions between the sidechains of these residues may contribute to the stability of the domain structure.

The *in vitro* translated Npw38 was specifically bound to poly(rG) resin, suggesting that Npw38 is an RNA-binding protein. Some transcription-related factors such as WT1 (32,33) or heterogeneous nuclear ribonucleoprotein K (34,35) have been reported to bind RNAs and defined DNAs with high affinity. Since Npw38 itself could not bind to a single- or double-stranded DNA, it may bind to DNA via associated proteins and act as a transcriptional regulator, if Npw38 acts as a transcriptional regulator. Alternatively, Npw38 may be involved in RNA processing such as splicing, poly(A) adenylation or capping via its RNA binding ability. Recently, McCracken *et al.* have proposed the concept of the 'mRNA factory' (36) in which the transcriptional, splicing (37,38) and cleavage-polyadenylation (39) machineries interact through contacts with the C-terminal domain of RNA polymerase II. Possibly, Npw38 may be a component of the 'mRNA factory'. We are continuing the further characterization of Npw38 including the identification of binding proteins of Npw38 which will help elucidate its physiological role.

ACKNOWLEDGEMENTS

We thank all our colleagues in our project for their helpful discussions and Dr Ryll for critical reading of the manuscript.

REFERENCES

- 1 Kato,S., Sekine,S., Oh,S.-W., Kim,N.-S., Umezawa,T., Abe,N., Yokoyama-Kobayashi,M. and Aoki,T. (1994) *Gene*, **150**, 243–250.
- 2 Bork,P. and Sudol,M. (1994) *Trends Biochem. Sci.*, **19**, 531–533.
- 3 André,B. and Springael,J.-Y. (1994) *Biochem. Biophys. Res. Commun.*, **205**, 1201–1205.
- 4 Hofmann,K. and Bucher,P. (1995) *FEBS Lett.*, **358**, 153–157.
- 5 Sudol,M., Chen,H.I., Bougeret,C., Einbond,A. and Bork,P. (1995) *FEBS Lett.*, **369**, 67–71.
- 6 Pawson,T. (1995) *Nature*, **373**, 573–580.
- 7 Chen,H.I. and Sudol,M. (1995) *Proc. Natl Acad. Sci. USA*, **92**, 7819–7823.
- 8 Staub,O., Dho,S., Henry,P.C., Correa,J., Ishikawa,T., McGlade,J. and Rotin,D. (1996) *EMBO J.*, **15**, 2371–2380.
- 9 Schild,L., Lu,Y., Gautschi,I., Schneeberger,E., Lifton,R.P. and Rossier,B.C. (1996) *EMBO J.*, **15**, 2381–2387.
- 10 Huijbregetse,J.M., Yang,J.C. and Beaudenon,S.L. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 3656–3661.
- 11 Lu,K.P., Hanes,S.D. and Hunter,T. (1996) *Nature*, **380**, 544–547.
- 12 Maleszka,R., Hanes,S.D., Hackett,R.L., de Couet,H.G. and Miklos,G.L.G. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 447–451.
- 13 Chan,D.C., Bedford,M.T. and Leder,P. (1996) *EMBO J.*, **15**, 1045–1054.
- 14 Bedford,M.T., Chan,D.C. and Leder,P. (1997) *EMBO J.*, **16**, 2376–2383.
- 15 Macknight,R., Bancroft,I., Page,T., Lister,C., Schmidt,R., Love,K., Westphal,L., Murphy,G., Sherson,S., Cobbett,C. and Dean,C. (1997) *Cell*, **89**, 737–745.
- 16 Macias,M.J., Hyvönen,M., Baraldi,E., Schultz,J., Sudol,M., Saraste,M. and Oschkinat,H. (1996) *Nature*, **382**, 646–649.
- 17 Bedford,M.T., Reed,R. and Leder,P. (1998) *Proc. Natl Acad. Sci. USA*, **95**, 10602–10607.
- 18 Gavva,N.R., Gavva,R., Ermekova,K., Sudol,M. and Shen,C.-K.J. (1997) *J. Biol. Chem.*, **272**, 24105–24108.
- 19 Mosser,E.A., Kasanov,J.D., Forsberg,E.C., Kay,B.K., Ney,P.A. and Bresnick,E.H. (1998) *Biochemistry*, **37**, 13686–13695.
- 20 Duilio,A., Zambrano,N., Mogavero,A.R., Ammendola,R., Cimino,F. and Russo,T. (1991) *Nucleic Acids Res.*, **19**, 5269–5274.
- 21 Ermekova,K.S., Zambrano,N., Linn,H., Minopoli,G., Gertler,F., Russo,T. and Sudol,M. (1997) *J. Biol. Chem.*, **272**, 32869–32877.
- 22 Zambrano,N., Minopoli,G., de Candia,P. and Russo,T. (1998) *J. Biol. Chem.*, **273**, 20128–20133.
- 23 Yokoyama-Kobayashi,M. and Kato,S. (1994) *Anal. Biochem.*, **223**, 130–134.
- 24 Cheng,J., Macon,K.J. and Volanakis,J.E. (1993) *Biochem. J.*, **294**, 589–593.
- 25 Spritz,R.A., Strunk,K., Surowy,C.S., Hoch,S.O., Barton,D.E. and Francke,U. (1987) *Nucleic Acids Res.*, **15**, 10373–10391.
- 26 Silver,P.A. (1991) *Cell*, **64**, 489–497.
- 27 Nigg,E.A., Beauerle,P.A. and Lührmann,R. (1991) *Cell*, **66**, 15–22.
- 28 Burd,C.G. and Dreyfuss,G. (1994) *Science*, **265**, 615–621.
- 29 Swanson,M. and Dreyfuss,G. (1988) *Mol. Cell. Biol.*, **8**, 2237–2241.
- 30 Dreyfuss,G., Matunis,M.J., Pinol-Roma,S. and Burd,C.G. (1993) *Annu. Rev. Biochem.*, **62**, 289–321.
- 31 Chen,H.I., Einbond,A., Kwak,S.-J., Linn,H., Koepf,E., Peterson,S., Kelly,J.W. and Sudol,M. (1997) *J. Biol. Chem.*, **272**, 17070–17077.
- 32 Caricasole,A., Duarte,A., Larsson,S.H., Hastie,N.D., Little,M., Holmes,G., Todorov,I. and Ward,A. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 7562–7566.
- 33 Larsson,S.H., Charlier,J.-P., Miyagawa,K., Engelkamp,D., Rassoulzadegan,M., Ross,A., Cuzin,F., van Heyningen,V. and Hastie,N.D. (1995) *Cell*, **81**, 391–401.
- 34 Michelotti,E.F., Tomonaga,T., Krutzsch,H. and Levens,D. (1995) *J. Biol. Chem.*, **270**, 9494–9499.
- 35 Michelotti,E.F., Michelotti,G.A., Aronsohn,A.I. and Levens,D. (1996) *Mol. Cell. Biol.*, **16**, 2350–2360.
- 36 McCracken,S., Fong,N., Yankulov,K., Ballantyne,S., Pan,G., Greenblatt,J., Patterson,S.D., Wickens,M. and Bentley,D.L. (1997) *Nature*, **385**, 357–361.
- 37 Yuryev,A., Patturajan,M., Litingtung,Y., Joshi,R.V., Gentile,C., Gebara,M. and Corden,J.L. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 6975–6980.
- 38 Mortillaro,M.J., Blencowe,B.J., Wei,X., Nakayasu,H., Du,L., Warren,S.L., Sharp,P.A. and Berezney,R. (1996) *Proc. Natl Acad. Sci. USA*, **93**, 8253–8257.
- 39 Dantoni,J.-C., Murthy,K.G.K., Manley,J.L. and Tora,L. (1997) *Nature*, **389**, 399–402.