NABP1, a novel RORγ-regulated gene encoding a single-stranded nucleic-acid-binding protein

Hong Soon KANG*, Ju Youn BEAK*, Yong-Sik KIM*, Robert M. PETROVICH†, Jennifer B. COLLINS‡, Sherry F. GRISSOM‡ and Anton M. JETTEN*¹

*Cell Biology Section, Laboratory of Respiratory Biology, Division of Intramural Research, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709, U.S.A., †Protein Expression Core Facility, Laboratory of Structural Biology, Division of Intramural Research, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709, U.S.A., and ‡Microarray Group, Division of Intramural Research, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709, U.S.A.

RORγ2 (retinoid-related orphan receptor γ2) plays a critical role in the regulation of thymopoiesis. Microarray analysis was performed in order to uncover differences in gene expression between thymocytes of wild-type and RORγ$^{-/-}$ mice. This analysis identified a novel gene encoding a 22 kDa protein, referred to as NABP1 (nucleic-acid-binding protein 1). This subsequently led to the identification of an additional protein, closely related to NABP1, designated NABP2. Both proteins contain an OB (oligonucleotide/oligosaccharide binding) motif at their N-terminus. This motif is highly conserved between the two proteins. NABP1 is highly expressed in the thymus of wild-type mice and is greatly suppressed in RORγ$^{-/-}$ mice. During thymopoiesis, NABP1 mRNA expression is restricted to CD4$^+$CD8$^+$ thymocytes, an expression pattern similar to that observed for RORγ2. These observations appear to suggest that NABP1 expression is regulated either directly or indirectly by RORγ2. Confocal microscopic analysis showed that the NABP1 protein localizes to the nucleus. Analysis of nuclear proteins by size-exclusion chromatography indicated that NABP1 is part of a high molecular-mass protein complex. Since the OB-fold is frequently involved in the recognition of nucleic acids, the interaction of NABP1 with various nucleic acids was examined. Our results demonstrate that NABP1 binds single-stranded nucleic acids, but not double-stranded DNA, suggesting that it functions as a single-stranded nucleic acid binding protein.

Key words: microarray, nucleic acid binding protein 1 (NABP1), nuclear receptor, oligonucleotide/oligosaccharide binding-fold (OB-fold), RORγ (retinoid-related orphan receptor γ), thymopoiesis.

INTRODUCTION

The ROR (retinoid-related orphan receptor) subfamily consists of three members, RORα, RORβ and RORγ (named NRF1–NRF3 and RORA–RORC respectively, by the Nuclear Receptor Nomenclature Committee and the Human Gene Nomenclature Committee) [1]. Each ROR gene yields, through alternative promoter usage and alternative splicing, several variants that differ only by their N-terminus. These variants have been implicated in the regulation of different physiological processes. The RORγ gene generates two isoforms, γ1 and γ2 (also referred to as RORγt) [1,2]. Most tissues, including liver, kidney and lung express RORγ1; however, expression of RORγ2 is restricted to CD4$^+$ CD8$^+$ [DP (double-positive)] thymocytes and lymphoid tissue inducer cells [1,3,4]. Characterization of ROR-deficient mice revealed that RORs play a critical role in the regulation of a number of physiological processes [1,3–10]. These studies have implicated RORα in the maintenance and/or differentiation of cerebellar Purkinje cells [7,11,12], and RORβ in the control of circadian rhythm [8,13]. Studies of RORγ$^{-/-}$ mice showed that these mice lack lymph nodes and Peyer's patches, suggesting that RORγ2 is essential for lymph node development [4,9,10].

In addition, these studies demonstrated that RORγ2 plays a role in the regulation of normal thymopoiesis [1,3,9,10,14]. DP thymocytes from RORγ$^{-/-}$ mice undergo accelerated apoptosis due to down-regulation of the expression of the anti-apoptotic gene Bcl-X$_L$ [9,10,15].

To obtain a greater insight into the function of RORγ in the thymus, we searched for additional changes in gene expression in RORγ$^{-/-}$ thymocytes. RNA isolated from thymocytes of wild-type and RORγ$^{-/-}$ mice were examined by microarray analysis. This analysis identified a novel gene, referred to as NABP1 (nucleic-acid-binding protein 1), that was greatly down-regulated in the thymus of RORγ$^{-/-}$ mice. Alignment of the mouse NABP1 amino acid sequence with sequences in GenBank® identified an additional novel protein closely-related to NABP1, referred to as NABP2. Both proteins are 22 kDa in size and contain one OB (oligonucleotide/oligosaccharide binding)-fold at their N-terminus. OB-folds consist of a compact structural domain of 70–150 amino acids. They do not share a strong sequence similarity but are recognized on the basis of their distinct topology. OB-fold proteins are frequently involved in nucleic acid recognition [16,17] and play a critical role in DNA replication, transcription, translation, DNA recombination, DNA repair and

the maintenance of telomeres [18–27]. The OB-fold family of proteins includes RPA (replication protein A) and the Brca2 (breast cancer susceptibility protein 2).

We further demonstrate that in wild-type mice NABP1 is most abundantly expressed in DP thymocytes and that its expression is significantly down-regulated in DP thymocytes isolated from $ROR\gamma^{-/-}$ mice. This pattern of expression is very similar to that observed for $ROR\gamma2$ [2,10]. This, in combination with its down-regulation in $ROR\gamma^{-/-}$ mice, suggests that NABP1 expression is regulated directly or indirectly by $ROR\gamma2$. NABP1 protein localizes to the nucleus, suggesting that NABP1 exerts its function in the nucleus. Finally, we demonstrate that NABP1 binds single-stranded nucleic acids but not double-stranded DNA, indicating that NABP1 functions as a single-stranded nucleic acid binding protein. Its specific expression in DP thymocytes, cells that undergo a high frequency of DNA rearrangements, raises the possibility of a role for NABP1 in DNA recombination or repair in these cells.

MATERIALS AND METHODS

Microarray analysis

Microarray analyses were carried out by the National Institute of Environmental Health Sciences, Microarray Group (NMG) using Agilent Mouse arrays (Agilent Technologies, Palo Alto, CA, U.S.A.) representing approx. 20 000 genes (http://dir.niehs.nih.gov/microarray/chips.htm). For each condition, equal amounts of total RNA from thymocytes of three individual wild-type or $ROR\gamma^{-/-}$ mice were pooled and amplified. Total RNA was amplified using the Agilent Low RNA Input Fluorescent Linear Amplification Kit protocol. Starting with 1 μg of total RNA, Cy3 (cyanine3)- or Cy5-labelled cRNA was synthesized according to the manufacturer's protocol. For each comparison, 750 ng of each Cy3- and Cy5-labelled cRNA was mixed and fragmented using the Agilent *In Situ* Hybridization Kit protocol. Hybridizations were performed for 16 h in a rotating hybridization oven using the Agilent 60-mer oligotide microarray processing protocol. Slides were washed as indicated in this protocol and then scanned with an Agilent Scanner. Data were analysed with Agilent Feature Extraction 7.1 software using default settings for all parameters.

Images and GEML files, including the error and P values, were exported from the Agilent Feature Extraction software and deposited into Rosetta Resolver (version 3.2, build 3.2.2.0.33; Rosetta Biosoftware, Kirkland, WA, U.S.A.). The resultant ratio profiles were combined into ratio experiments. Intensity plots were generated for each ratio experiment and genes were considered 'signature genes' if the P value was less than 0.001. The microarray data discussed in the present study have been deposited in the NCBI's Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) under accession number GSE4248 [28].

RNA isolation and Northern blot analysis

Total RNA from cells or tissues was isolated using mini- or midi-RNA isolation kits (Qiagen, Valencia, CA) according to the manufacturer's instructions. RNA (15 μg) was separated on a 1.2 % agarose gel, transferred to a nylon membrane (Sigma, St. Louis, MO, U.S.A.), and subsequently UV cross-linked. The membranes were hybridized to ^{32}P-labelled probes for NABP1, NABP2 or $ROR\gamma$ [9]. Hybridizations were performed at 68 °C for 3 h, the membranes were then washed twice with 2 × SSC (1 × SSC is 0.15 M NaCl/0.015 M sodium citrate)/0.1 % SDS for 20 min at room temperature, and subsequently with 0.1 × SSC/0.1 % SDS for 15 min at 50 °C. Autoradiography was

carried out at − 70 °C using Hyperfilm (Amersham Biosciences, Piscataway, NJ, U.S.A.).

Genomic structure analysis

The mouse NABP1 sequence was submitted to GenBank® under accession number AY880264. Alignment of the mouse NABP1 cDNA sequence with sequences in GenBank® was used to identify a cDNA sequence (AK028886) encoding a larger 3.1 kb NABP1 transcript. The genomic structure was deduced from the alignments of AY880264 and AK028886 with the mouse genomic DNA sequence (GenBank® AC125518.3) containing the entire NABP1 gene. This alignment indicated that the variant transcripts were generated by alternative splicing. The generation of variant transcripts was confirmed by reverse transcriptase-PCR analysis using DNA-free RNA from thymus and the following forward (F) and reverse (R) primers: 5′-CCTTTCTTCCAGGAATGTGTG-3′ (F1), 5′-TCCTTGTGACTTCTCAAAATGC-3′ (R1), 5′-AATG-GGGACCCAATCCTAGA-3′ (R2), 5′-CCTTTCCAGTCTGC-TTGGAT (F2), and 5′-TTCAGCCTTTGGGCATTTAG-3′ (R3). Based on the genomic structure, it was predicted that F1 and R1 amplify a 492 bp region specific for exon 6, F1 and R2 a 554 bp region spanning part of exon 6 and exon 7, and R2 and R3 a 503 bp region specific for exon 7. The presence of two variant transcripts was further confirmed by Northern blot analysis using probes hybridizing to full-length NABP1 (AY880264), the C-terminal half of the coding region of NABP1, and exon 7.

Plasmids

pMAL-C2X-NABP1 and pMAL-C2X-NABP1ΔN, encoding MBP (maltose-binding protein)–NABP1 and MBP–NABP1ΔN fusion proteins, were generated by cloning cDNAs encoding full-length NABP1 and NABP1ΔN (amino acids 146–198) respectively, into the EcoRI and BamHI sites of pMAL-C2X (New England BioLabs, Ipswich, MA, U.S.A.). NABP1 was amplified with the sense, 5′-GAGAATTCATGCACGGGGTC-AACGAC-3′, and antisense, 5′-GTGGATCCTCATCATTTAAA-GGCTCTCCTC-3′, primers, NABP1ΔN was amplified with 5′-GAGAATTCGGGGGTAGAAAATGAACAGAAGG-3′ and 5′-GTGGATCCTCATCATTTAAAGGCTCTCCTC-3′. To generate pET41a(+)-NABP1, encoding the GST (glutathione S-trans-ferase)–NABP1 fusion protein, the NABP1 cDNA was amplified with 5′-CCGAATTCCACGGGGTCAACGACCCT-3′ and 5′-GT-CTCGAGTCATCTTTTAAAGGCTCTCCTC-3′ primers and then cloned into the EcoRI and XhoI sites of pET-41a(+) (EMD Biosciences, Madison, WI, U.S.A.). p3XFLAG-CMV-NABP1 was obtained by PCR using the primers 5′-GAGAATTCTATG-CACGGGGTCAACGAC-3′ and 5′-GTGGATCCTCATCATT-TAAAGGCTCTCCTC-3′. The cDNA that was obtained was then cloned into the EcoRI and BamHI sites of p3XFLAG-CMV-7.1 (Sigma). Subsequently, the 3XFLAG–NABP1 cDNA was sub-cloned into the HpaI and BamHI sites of the retroviral vec-tor pLXIN (BD Bioscience, Palo Alto, CA, U.S.A.). pLXIN-3XFLAG–NABP1ΔC (amino acids 1–114) and pLXIN-3XFLAG–NABP1ΔN (amino acids 117–198) were generated by PCR using the sense/anti-sense primer pairs: 5′-TTGCGGCC-GCGATGCACGGGGTCAACGAC-3′ and 5′-GTGGATCCTC-ACTCACTGAAATTTGGCACTTC-3′, and 5′-TTGCGGCC-GCGCCA-GATTATAGAGGACAGC-3′ and 5′-GTGGATCCTC-ATCTTTTAAAGGCTCTCCTC-3′ respectively. The sequence of all plasmids was verified by restriction enzyme analysis and DNA sequencing.

Purification of NABP1 proteins

Escherichia coli DH5α was transformed with pMAL-C2X-NABP1 or pMAL-C2X-NABP1ΔN plasmid DNA. The

transformants were grown in LB (Luria–Bertani) medium and the synthesis of the MBP–NABP1 fusion protein was induced by the addition of 0.3 mM IPTG (isopropyl β-D-thiogalactoside). After 23 h of incubation at 16 °C bacteria were harvested, washed twice with PBS, and the bacterial lysate was prepared using the Readypreps™ Protein Preparation Kit (Epicentre, Madison, WI, U.S.A.) according to the manufacturer's instructions. The crude lysate was then mixed with Amylose-Resin (NEB, Beverly, MA, U.S.A.). After 30 min of incubation at 4 °C the resin was washed 3 times with elution buffer [20 mM Tris, 200 mM NaCl and 1 mM EDTA (pH 7.4)] and subsequently with elution buffer containing 10 mM maltose to elute the MBP–NABP1 fusion protein. The MBP–NABP1 fusion protein was analysed by PAGE and identified by Coomassie Brilliant Blue staining.

E. coli Rosetta DE3 was transformed with pET41a(+)-NABP1 plasmid DNA. Transformants were grown in LB medium and the synthesis of the GST–NABP1 protein was induced by the addition of 0.2 mM IPTG. After 4 h of incubation at 30 °C bacteria were harvested and lysed by sonication in B-PER® bacterial protein extraction buffer (Pierce Biotech., Rockford, IL, U.S.A.). The soluble protein fraction obtained after centrifugation (30 min at 27 000 g) was then incubated with GST-affinity resin at room temperature for 1 h. The resin was washed in washing buffer [25 mM Hepes/150 mM NaCl (pH 8.0)] and the GST–NABP1 protein was eluted with glutathione. The eluted GST–NABP1 protein was cleaved by incubation with thrombin I (16 h at 4 °C) and GST protein was removed by passing the solution over a GST-affinity resin column. The flow-through was concentrated and further purified by HPLC using a mono-Q ion-exchange column (Amersham Biosciences). The fraction containing NABP1 was analysed by PAGE. The NABP1 protein bands were verified by MALDI-TOF (matrix-assisted laser-desorption ionization–time-of-flight).

Flow cytometry

Thymocytes were isolated from mouse thymus as described [9] and then resuspended at 4 °C in RPMI-1640 containing 10 % FBS (foetal bovine serum). Thymocytes (5×10^6/ml) were then incubated with 2 μg/ml FITC-conjugated anti-CD8 and 2 μg/ml phycoerythrin-conjugated anti-CD4 antibodies (BD Pharmingen, San Diego, CA, U.S.A.). The CD4$^+$CD8$^+$, CD4$^-$CD8$^-$, and CD4$^+$ and CD8$^+$ single-positive cells were then separated and collected using an LSR flow cytometer (Becton Dickinson, San Jose, CA, U.S.A.) as described previously [9]. The CD4$^+$ and CD8$^+$ single-positive cells were pooled. Cells that were collected from several separate runs were pooled and then processed for total RNA isolation. Data were analysed using CellQuest software (Becton Dickinson).

Real-time quantitative reverse transcriptase-PCR

Real-time reverse transcriptase-PCR reactions were carried out in triplicate in a 7300 Real-Time PCR system using the TaqMan® One-Step reverse transcriptase-PCR Master mix (Applied Biosystems, Foster City, CA, U.S.A.). The NABP1- and RORγ-specific oligonucleotides were designed with Primer Express™ 2.0 software (Applied Biosystems, Foster City, CA, U.S.A.). For NABP1, the following oligonucleotides were used: forward primer, 5′-GGGTACAAAATGAACAGAAGGATAA-AC-3′; reverse primer, 5′-GATTCAGGGCCAGTCTGATCA, and the probe 5′-FAM™-TTCCCACTGGTCCAAATGTATTGGTG-CT-TAMRA™-3′; for mouse RORγ: forward primer, 5′-CACG-GCCCTGGTTCTCAT-3′; reverse primer, 5′-CAGATGTTCCAC-TCTCCTCTTCTCT-3′; and the probe 5′-FAM™-ATGCCAAC-CGTCCTGGGCTCC-TAMRA™-3′. NABP1 oligonucleotides

were designed to amplify a region spanning parts of exon 5 and 6. A total of 25 ng of RNA/reaction was used to quantify NABP1 and RORγ mRNA levels, while 25 pg of RNA/reaction was used to quantify 18S rRNA. RNAs were reverse transcribed for 30 min at 48 °C, denatured for 10 min at 95 °C, followed by 40 cycles consisting of 15 s at 95 °C, and 60 s at 60 °C. Average threshold cycle values from triplicate PCR reactions for NABP1 and RORγ were normalized against the threshold cycle values obtained for 18S rRNA, which served as an internal control. All standard curves produced r^2 values above 0.99.

Cell culture and retrovirus infection

CV-1 cells were cultured in Dulbeco's modified Eagle's medium containing 10 % FBS. The packaging cell line pT67 was grown in RPMI-1640 plus 10 % FBS. pT67 cells were transfected with pLXIN-3XFLAG–NABP1 plasmid DNA using FuGENE™ 6 transfection reagent (Roche, Indianapolis, IN, U.S.A.) and stable cell lines were established after G418 selection. Conditioned medium containing the retrovirus was prepared and stored at −70 °C. Subsequently, CV-1 cells were infected with the retrovirus in the presence of polybrene (8 μg/ml). After G418 (1000 μg/ml) selection, CV-1 cell lines stably expressing FLAG–NABP1, FLAG–NABP1ΔN, and FLAG–NABP1ΔC were obtained.

Isolation of nuclear extracts and Sephacryl chromatography

CV-1 cells stably transfected with pLXIN-3XFLAG–NABP1 were collected, washed in PBS, and then resuspended in 10 ml of buffer A [15 mM KCl, 3.75 mM NaCl, 37.5 μM spermin, 125 μM spermidine, 0.5 mM EDTA and 3.75 mM Tris (pH 7.4)]. After centrifugation for 5 min at 2000 g, nuclear extracts were prepared as described previously [29]. This isolation procedure includes a 1 h incubation with 100 μg/ml DNase I. Nuclear proteins were separated by Sephacryl 300 HR column chromatography with elution buffer [15 mM KCl, 150 mM NaCl and 20 mM Tris (pH 7.4)]. Blue Dextran, β-amylase, yeast alcohol dehydrogenase, FBS, and carbonic anhydrase were used as molecular-mass standards. Fractions were assayed for the presence of FLAG–NABP1 protein either by dot-blot or Western blot analysis using an anti-Flag M2 antibody (Sigma). The dot-blot staining was quantified using a Fluorskan 8900 (Alpha Innotech, San Leandro, CA, U.S.A.) apparatus.

Subcellular localization

The subcellular localization of NABP1 was examined in CV-1 cells that stably expressed either the FLAG–NABP1, FLAG–NABP1ΔN, or FLAG–NABP1ΔC fusion protein. Cells were plated in culture medium in glass-bottom culture dishes (MatTek Corp., Ashland, MA, U.S.A.) and 24 h later were fixed in 4 % paraformaldehyde. Cells were washed in PBS and then incubated with a mouse anti-FLAG M2 antibody (Sigma) followed by incubation with an Alexa Fluor® 594 goat anti-mouse antibody (Molecular Probes, Eugene, OR, U.S.A.). Cells were washed with PBS and nuclei were stained with DAPI (4′-6-diamidino-2-phenylindole). Fluorescence was examined using a Zeiss confocal microscope LSM 510 NLO (Zeiss, Thornwood, NY, U.S.A.).

Nucleic acid binding assay

Purified NABP1 protein was incubated with 125 ng of single-stranded viral DNA (ΦX174) (New England BioLabs), 125 ng of double-stranded DNA (1.3 Kb), or 0.5 μg of total RNA in binding buffer A [20 mM Tris (pH 7.5), 50 mM NaCl, 1 mM EDTA and 0.1 mg/ml of FBS] for 20 min at room temperature. The DNA or RNA was then analysed by 0.8 % agarose gel
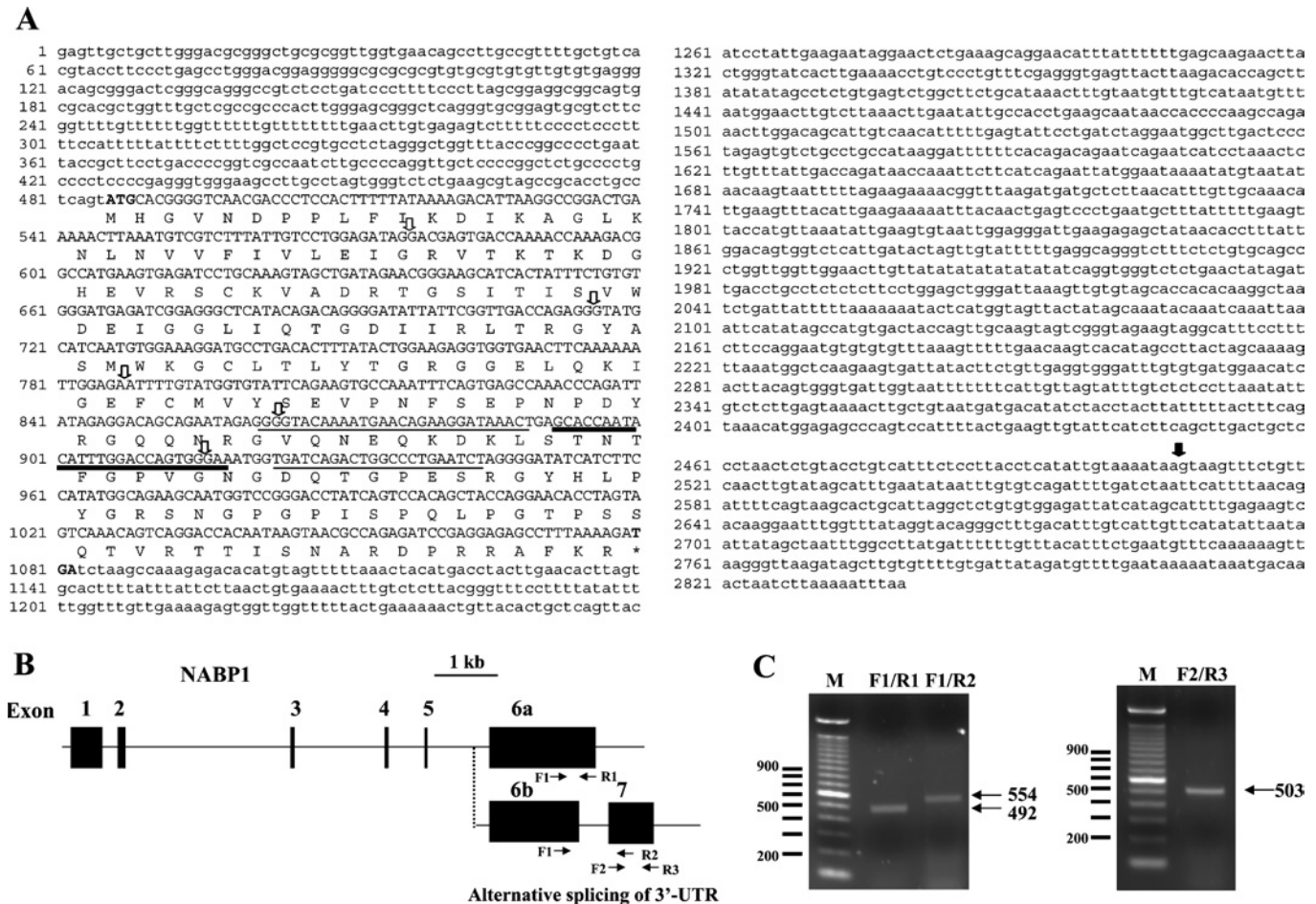
A

```
   1 gagttgctgcttgggacgcgggctgcgcggttggtgaacagccttgccgttttgctgtca
  61 cgtaccttccctgagcctgggacggaggggggcgcgcgcgtgtgcgtgtgttgtgtgaggg
 121 acagcgggactcgggcagggccgtctcctgatcccttttcccttagcggaggcggcagtg
 181 cgcacgctggttgctcgccgcccacttgggagcgggtcagggtgcggagtgcgtcttc
 241 ggttttgtttttttggttttttgttttttttgaacttgtgagagtcttttttccctcccctt
 301 ttccatttttattttcttttggctccgtgcctctagggctggtttacccggcccctgaat
 361 taccgcttcctgaccccggtcgccaatcttgccccaggttgctccccggcctctgccccctg
 421 cccctcccctgagggtgggaagccttgcctagtgggtctctgaagcgtagccgcacctgcc
 481 tcagtATGCACGGGGTCAACGACCCTCCACTTTTTATAAAAGACATTAAGGCCGGACTGA
           M  H  G  V  N  D  P  P  L  F  I  N  K  D  I  K  A  G  L  K
 541 AAAACTTAAATGTCGTCTTTATTGTCCTGGAGATAGGACGAGTGACCAAAACCAAAGACG
           N  L  N  V  V  F  I  V  L  E  I  G  R  V  T  K  T  K  D  G
 601 GCCATGAAGTGAGATCCTGCAAAGTAGCTGATAGAACGGGAAGCATCACTATTTCTGTGT
           H  E  V  R  S  C  K  V  A  D  R  T  G  S  I  T  I  S  N  V  W
 661 GGGATGAGATCGGAGGGCTCATACAGACAGGGGATATTATTCGGTTGACCAGAGGGTATG
           D  E  I  G  G  L  I  Q  T  G  D  I  I  R  L  T  R  G  Y  A
 721 CATCAATGTGGAAAGGATGCCTGACACTTTATACTGGAAGAGGTGGTGAACTTCAAAAAA
           S  M  N  W  K  G  C  L  T  L  Y  T  G  R  G  G  E  L  Q  K  I
 781 TTGGAGAATTTTGTATGGTGTATTCAGAAGTGCCAAATTTCAGTGAGCCAAACCCAGATT
           G  E  F  C  M  V  Y  S  E  V  P  N  F  S  E  P  N  P  D  Y
 841 ATAGAGGACAGCAGAATAGAGGGGTACAAAATGAACAGAAGGATAAACTGAGCACCAATA
           R  G  Q  Q  N  R  G  V  Q  N  E  Q  K  D  K  L  S  T  N  T
 901 CATTTGGACCAGTGGGGAATGGTGATCAGACTGACTGGCCCTGAATCTAGGGGATATCATCTTC
           F  G  P  V  G  N  G  D  Q  T  G  P  E  S  R  G  Y  H  L  P
 961 CATATGGCAGAAGCAATGGTCCGGGACCTATCAGTCCACAGCTACCAGGAACACCTAGTA
           Y  G  R  S  N  G  P  G  P  I  S  P  Q  L  P  G  T  P  S  S
1021 GTCAAACAGTCAGGACCACAATAAGTAACGCCCAGAGATCCGAGGAGGAGCCTTTAAAAGAT
           Q  T  V  R  T  T  I  S  N  A  R  D  P  R  R  A  F  K  R  *
1081 GAtctaagccaaagagacacatgtagttttaaactacatgacctacttgaacacttagt
1141 gcactttfattfattcttaactgtgaaaactttgtctcttacgggtttccttttatattt
1201 ttggtttgttgaaaagagtggttggttttttactgaaaaaactgttacactgctcagttac
```

```
1261 atcctattgaagaataggaactctgaaagcaggaacatttatttttttgagcaagaactta
1321 ctgggtatcacttgaaaacctgtccctgtttcgagggtgagttacttaagacaccagctt
1381 atatatagcctctgtgagtctggcttctgcataaactttgtaatgtttgtcataatgttt
1441 aatggaacttgtcttaaacttgaatattgccacctgaagcaataaccaccccaagccaga
1501 aacttggacagcattgtcaacattttgagtattcctgatctaggaatggcttgactccc
1561 tagagtgtctgcctgccataaggattttttcacagacagaatcagaatcatcctaaactc
1621 ttgtttattgaccagataaccaaattcttcatcagaattatggaataaaatatgtaatat
1681 aacaagtaattttagaagaaaacggtttaagatgatgctcttaacatttgttgcaaaca
1741 ttgaagtttacattgaagaaaaattttcaaactgagtccctgaatgctttatttttgaagt
1801 taccatgttaaatattgaagtgtaattggagggattgaagagagctataacacctttatt
1861 ggacagtggtctcattgatactagttgtattttttgaggcagggtctttctctgtgcagcc
1921 ctggttggttggaacttgttatatatatatatatatcaggtgggtctctgaactatagat
1981 tgacctgcctctctcttcctgggagctgggattaaagttgtgtagcaccacacaaggctaa
2041 tctgattatttttaaaaaaatactcatggtagttactatagcaaatacaaatcaaattaa
2101 attcatatagccatgtgactaccagttgcaagtagtcgggtagaagtaggcatttcctt
2161 cttccaggaatgtgtgttaaagttttgaacaagtcacatagcctttactagcaaaag
2221 ttaaatggctcaagaagtgattatacttctgttgaggtgggatttgtgtgatggaacatc
2281 acttacagtgggtgattggtaattttttcattgttagtatttgtctctcctaaatatt
2341 gtctcttgagtaaaacttgctgtaatgatgacatatctacctacttatttttacttttcag
2401 taaaacatggagagcccagtccatttactgaagttgtattcatcttcagcttgactgctc
2461 cctaactctgtacctgtcatttctcctacctcatattgtaaaataagtaagtttctgtt
2521 caacttgtatagcatttgaatataatttgtgtcagattttgatctaattcattttaacag
2581 attttcagtaagcactgcattaggctctgtgtggagattatcatagcattttgagaagtc
2641 acaaggaatttggtttataggtacagggctttgacatttgtcattgttcatatattaata
2701 attatatagctaatttggccttatgatttttttgtttacatttctgaatgtttcaaaaaagtt
2761 aaggggttaagatagcttgtgtgtttgtgattatagatgtttgaataaaaataaatgacaa
2821 actaatcttaaaaatttaa
```

B   **C**

Figure 1   Nucleotide and amino acid sequences of mouse NABP1

(**A**) The nucleotide and putative amino acid sequences of mouse NABP1. The start and stop codons are in bold type. Underlined nucleotide sequences indicate the location of the forward and reverse primers (thin lines), and the probe (thick lines) used in real-time reverse transcriptase-PCR analysis. (**B**) Schematic view of the genomic structure of the mouse NABP1 gene. The genomic structure was deduced from the alignment of the cDNA sequences AY880264 and AK028886 with the mouse genomic DNA sequence (GenBank® AC125518.3) containing the NABP1 gene. Black boxes indicate exons. The NABP1 gene generates two major transcripts via alternative splicing, that encode the same coding region but differ in their 3′-UTRs. The generation of the two variant transcripts was confirmed by reverse transcriptase-PCR using DNA-free RNA from thymus and three different primer sets, as described in the Materials and methods section. The location of the forward (F) and reverse (R) primers is indicated in (**B**). The primer sets F1/R1, F1/R2 and F2/R3 each generated a PCR product of the size predicted by the genomic structure (**C**). In (**A**), the location of the 5 common introns is indicated by open arrows; the start of the alternative intron 6 is indicated by the filled arrow.

electrophoresis. Gels were stained with ethidium bromide and DNA visualized by UV illumination. Oligonucleotides were end-labelled with $[\gamma\text{-}^{32}\text{P}]$-ATP by T4 polynucleotide kinase (Promega, Madison, WI, U.S.A.) and were then purified using MicroSpin™ G-25 columns (Amersham Biosciences). $[^{32}\text{P}]$-labelled single-stranded oligonucleotides (10 pM) were incubated in binding buffer B [10 mM Tris/HCl (pH 7.5), 1 mM MgCl₂, 0.5 mM EDTA, 0.5 mM DTT, 50 mM NaCl and 4% glycerol] with purified NABP1 protein for 40 min at room temperature. The reaction mixtures were then analysed by electrophoresis on 10% native polyacrylamide gels. The radiolabelled oligonucleotides were visualized by autoradiography.

RESULTS

Identification of NABP1

Previous studies have demonstrated that RORγ2 exhibits several important regulatory functions in the thymus [9,10,15,30]; however, little is known about the genes regulated directly or indirectly by RORγ2 in this tissue. To identify genes regulated

by RORγ2 in this tissue, microarray analysis was performed using RNA derived from thymocytes of wild-type and ROR$\gamma^{-/-}$ mice. This analysis identified many mRNAs that were either induced or repressed in the thymus of ROR$\gamma^{-/-}$ mice (GEO, http://www.ncbi.nlm.nih.gov/geo/). A partial list of differentially expressed mRNAs is shown in Table 1 in the Supplementary material (http://www.BiochemJ.org/bj/397/bj3970089.add.htm). The list of RNAs down-regulated in ROR$\gamma^{-/-}$ thymocytes included RORγ (RORC), the transcription factor EGR1 (early growth response gene product 1), GLCCI1 (glucocorticoid-induced transcript 1), and a novel mRNA not reported previously. This mRNA encoded a 22 kDa protein containing one OB-fold motif at its N-terminus. Since an OB-fold is frequently implicated in nucleic acid binding, we named it mouse NABP1 (mNABP1). The nucleotide and putative amino acid sequences are shown in Figure 1.

Alignment of our mouse NABP1 cDNA sequence (AY880264) with sequences in GenBank® identified a cDNA sequence (AK028886) encoding a larger 3.1 kb NABP1 transcript. This variant transcript encoded the same coding region as AY880264; however, the two transcripts differed in their 3′-UTRs
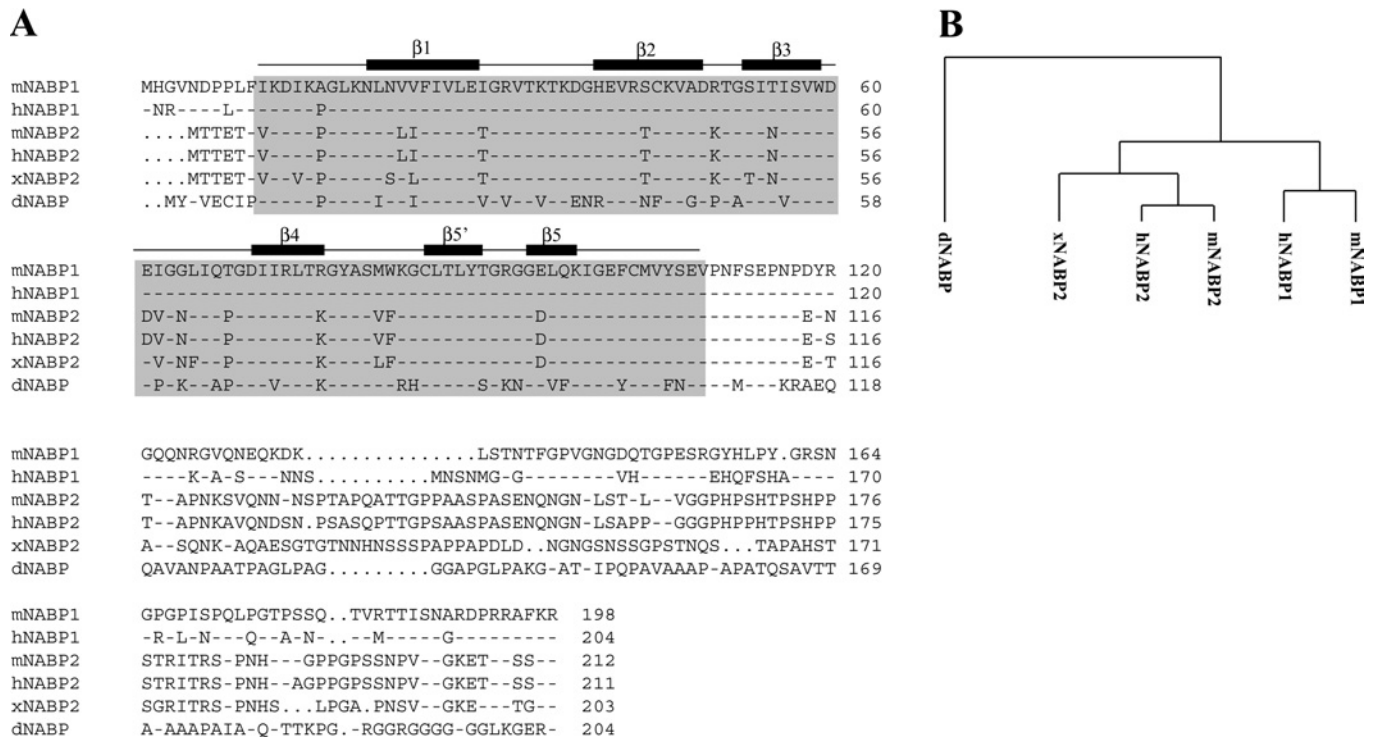
A

```
                               β1                    β2            β3
mNABP1   MHGVNDPPLFIKDIKAGLKNLNVVFIVLEIGRVTKTKDGHEVRSCKVADRTGSITISVWD   60
hNABP1   -NR----L--------P-------------------------------------------   60
mNABP2   ....MTTET-V----P------LI-----T--------------T-----K----N-----   56
hNABP2   ....MTTET-V----P------LI-----T--------------T-----K----N-----   56
xNABP2   ....MTTET-V--V-P------S-L-----T--------------T-----K--T-N-----   56
dNABP    ..MY-VECIP-----P----I--I------V-V--V--ENR---NF--G-P-A----V----   58

                        β4            β5'        β5
mNABP1   EIGGLIQTGDIIRLTRGYASMWKGCLTLYTGRGGELQKIGEFCMVYSEVPNFSEPNPDYR   120
hNABP1   ----------------------------------------------------------   120
mNABP2   DV-N---P-------K----VF------------D-----------------E-N   116
hNABP2   DV-N---P-------K----VF------------D-----------------E-S   116
xNABP2   -V-NF--P-------K----LF------------D-----------------E-T   116
dNABP    -P-K--AP---V---K------RH-----S-KN--VF----Y---FN---M---KRAEQ   118

mNABP1   GQQNRGVQNEQKDK..............LSTNTFGPVGNGDQTGPESRGYHLPY.GRSN   164
hNABP1   ----K-A-S---NNS..........MNSNMG-G--------VH------EHQFSHA----   170
mNABP2   T--APNKSVQNN-NSPTAPQATTGPPAASPASENQNGN-LST-L--VGGPHPSHTPSHPP   176
hNABP2   T--APNKAVQNDSN.PSASQPTTGPSAASPASENQNGN-LSAPP--GGGPHPPHTPSHPP   175
xNABP2   A--SQNK-AQAESGTGTNNHNSSSPAPPAPDLD..NGNGSNSSGPSTNQS...TAPAHST   171
dNABP    QAVANPAATPAGLPAG.........GGAPGLPAKG-AT-IPQPAVAAAP-APATQSAVTT   169

mNABP1   GPGPISPQLPGTPSSQ..TVRTTISNARDPRRAFKR    198
hNABP1   -R-L-N---Q--A-N-...-M-----G---------    204
mNABP2   STRITRS-PNH---GPPGPSSNPV--GKET--SS--   212
hNABP2   STRITRS-PNH--AGPPGPSSNPV--GKET--SS--   211
xNABP2   SGRITRS-PNHS...LPGA.PNSV--GKE---TG--   203
dNABP    A-AAAPAIA-Q-TTKPG.-RGGRGGGG-GGLKGER-   204
```

B

Figure 2 **Multiple amino acid sequence alignment of mouse (m) and human (h) NABP1 and NABP2, the *Xenopus* (x) homologue of NABP2, and the putative *Drosophila* (d) NABP homologue**

(**A**) The OB-fold domain is shaded. (-) indicates sequence identity with mNABP1; (.) indicates a gap. The β-strands are indicated by black bars. (**B**) Phylogenetic relationship between NABP1 and NABP2 from different species.

(untranslated regions). By aligning the two cDNA sequences with the mouse genomic sequence we deduced that this variant transcript is generated by alternative splicing. It was concluded that the smaller transcript is encoded by exons 1–6a, whereas the larger transcript is encoded by exons 1–6b and 7 (Figure 1B). This conclusion is supported by reverse transcriptase-PCR analysis with the primer sets F1/R1, F2/R3 and F1/R2 (Figure 1C). These primer sets specifically amplify exon 6a, exon 7, and a region spanning parts of exon 6 and 7 respectively. Each of these three reactions produced a PCR product of the predicted size (492, 503 and 554 bp respectively) (Figure 1C). A schematic view of the genomic structure and the alternative splicing is shown in Figure 1(B). The location of the five shared introns and the start of the alternative exon 6 is indicated in Figure 1(A).

Figure 2(A) compares the amino acid sequence of mouse NABP1 with that of human (GenBank® BC017114) and rat NABP1 (GenBank® BC089852). Mouse NABP1 exhibits 97% and 86% homology with rat and human NABP1 respectively (Figure 2A). In addition, the sequence of a previously unreported protein, closely related to that of NABP1, was identified and referred to as NABP2 (GenBank® NP_076973). NABP2 consists of 212 amino acids, and like NABP1, contains one OB-fold motif at its N-terminus. NABP1 and NABP2 exhibit an overall homology of 52% but show an 83% identity within their OB-fold domain (Figure 2A). The OB-fold of NABP1 and NABP2 is 80 and 94% identical respectively, to the sequence of an unpublished *Xenopus laevis* protein (GenBank® AAH70793) and shows a 62–69% identity with a previously unpublished *Drosophila melanogaster* protein (NP_609115), suggesting that they may be the homologues of *Xenopus* and *Drosophila* NABPs. The phylogenetic relationship between these proteins is shown in Figure 2(B). Based on sequence simi-

larities, the *Xenopus* protein is more closely related to NABP2 than NABP1, and therefore was named xNABP2. The mouse and human NABP1 genes map to chromosome 1C1 and 2q32.3 respectively, whereas the NABP2 genes map to mouse chromosome 10D3 and human chromosome 12q13.3.

The sequence of the OB-folds identified in NABPs exhibits little homology with those of other OB-fold proteins. OB-folds can be recognized on the basis of their distinct topology and consist of a β-barrel formed by five β-sheets.

Secondary structure analyses (http://cubic.bioc.columbia.edu/predictprotein) predicted the presence of five β-sheets in the OB-fold of NABP1 and NABP2 (Figure 2A) that aligned with OB-folds present in SSB (single-stranded DNA-binding protein) and RPA-like proteins (results not shown) [17,31]. In contrast with the SSB and RPA-like proteins NABP1 and NABP2 do not appear to contain an α-helix between their third and fourth β-sheets.

Tissue-specific pattern of expression of NABPs

To analyse the tissue-specific expression of NABP1 and 2, we performed Northern blot analysis using RNA from multiple mouse tissues. Although NABP1 mRNA was present in many tissues, the level of expression in these tissues differed significantly. NABP1 mRNA was expressed at relatively high levels in heart, kidney, thymus, testis and placenta and at low levels in all other tissues (Figure 3A). In most tissues, the NABP1-specific probe hybridized to two transcripts, 2.8 and 3.1 kb, whereas in testis the probe hybridized to two additional transcripts of 1.8 and 0.81 kb. By contrast, the NABP2-specific probe hybridized to only one transcript of 3.0 kb that was expressed at similar levels in all tissues analysed.

We next determined whether the 2.8 and 3.1 kb transcripts represented the variant transcripts described in Figure 1(B).
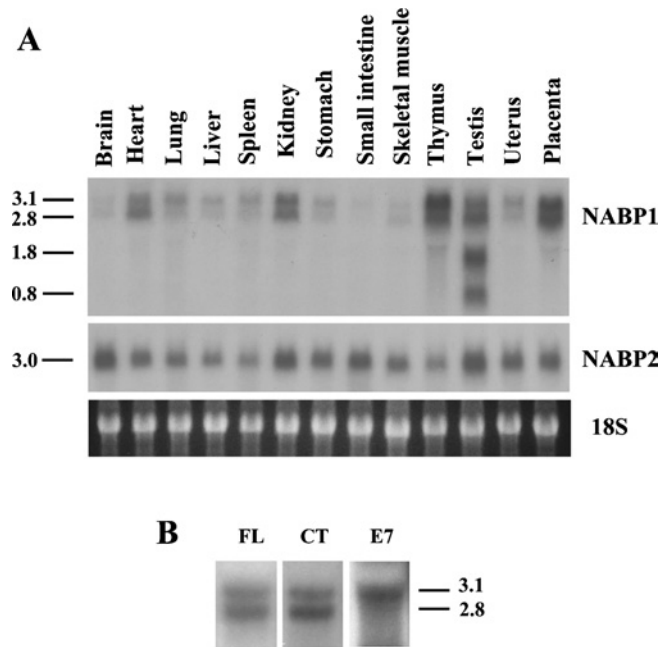
Figure 3   Tissue-specific expression of NABP1 and NABP2 mRNA

(**A**) Total RNA from various mouse tissues was examined by Northern blot analysis using radiolabelled probes for NABP1 and NABP2. The sizes (kb) of the NABP transcripts are indicated on the left. (**B**) The NABP1 gene generates two major transcripts via alternative splicing. The generation of these two transcripts was confirmed by Northern blot analysis using probes encoding the full-length (FL) NABP1 cDNA AY880264, the C-terminal (CT) half of the coding region of NABP1, and exon 7 (E7). The CT probe hybridizes to both transcripts whereas the E7 probe detects only the larger transcript, which is in agreement with the alternative splicing shown in Figure 1(B).

Northern blot analysis was performed using probes for full-length NABP1, the region encoding the C-terminus of NABP1 (amino acids 117–198), and exon 7 (E7). As shown in Figure 3(B), the full-length and C-terminal probes hybridized to both transcripts, whereas E7 hybridized only to the larger transcript. These results support the conclusion that the NABP1 gene generates two variant transcripts, by alternative splicing, that differ in their 3′-UTRs (Figure 1B).

Differential regulation of NABP1 in thymus

NABP1 was initially identified by microarray analysis as one of the genes that was down-regulated in the thymus of $ROR\gamma^{-/-}$ mice. To validate this observation, we compared the level of NABP1 mRNA expression in the thymus of wild-type and $ROR\gamma^{-/-}$ mice by Northern blot analysis. As shown in Figure 4(A), NABP1 mRNA was highly expressed in the thymus of wild-type mice but its expression was dramatically decreased in the thymus of $ROR\gamma^{-/-}$ mice. The lack of $ROR\gamma$ expression did not affect the expression of NABP1 mRNA in other tissues, including heart, kidney and testis (Figure 4B). In contrast with NABP1, NABP2 expression was unchanged in the thymus of $ROR\gamma^{-/-}$ mice (Figure 4A). In addition, the expression of both NABP1 and NABP2 was unaltered in the thymus of staggerer ($ROR\alpha^{sg/sg}$) mice which are deficient in $ROR\alpha$ expression (results not shown). These observations suggest that repression of NABP1 in $ROR\gamma^{-/-}$ mice is restricted to the thymus. This is interesting when compared with the previously reported expression pattern of $ROR\gamma2$. These studies showed that expression of $ROR\gamma2$ in the thymus is restricted to CD4$^+$CD8$^+$ (DP) thymocytes [1,3,10,30]. To further examine the possible correlation between
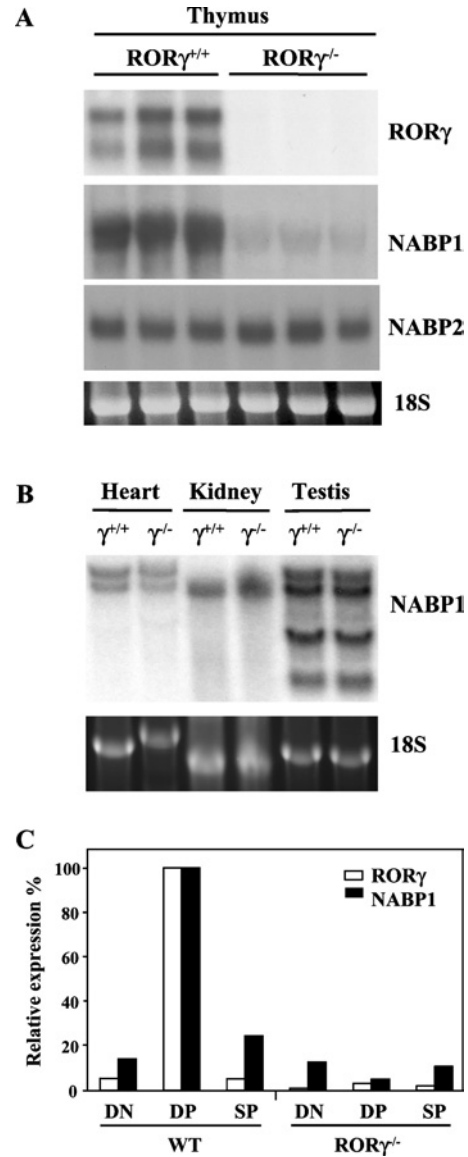
Figure 4   Expression of NABP1 mRNA is suppressed in the thymus of $ROR\gamma^{-/-}$ mice

(**A**) Total RNA was isolated from thymi of wild-type ($ROR\gamma^{+/+}$) and $ROR\gamma^{-/-}$ mice, and then examined by Northern blot analysis using radiolabelled probes for $ROR\gamma$, NABP1 and NABP2. In each case tissues from three different mice were analysed. (**B**) $ROR\gamma$ does not regulate NABP1 expression in heart, kidney and testis. Total RNA was isolated from heart, kidney and testis of wild-type ($ROR\gamma^{+/+}$) and $ROR\gamma^{-/-}$ mice, and then examined by Northern blot analysis using radiolabelled probes for $ROR\gamma$. (**C**) NABP1 mRNA is abundant in CD4$^+$CD8$^+$ (DP) thymocytes. Total RNA was isolated from CD4$^-$CD8$^-$ (double negative, DN), DP, and CD4$^+$ and CD8$^+$ (single positive, SP) thymocytes from wild-type and $ROR\gamma^{-/-}$ mice. The expression of NABP1 and $ROR\gamma$ mRNA was quantified by real-time reverse transcriptase-PCR as described in the Materials and methods section. The locations of the primers and probe are indicated in Figure 1(A).

the expression of NABP1 and $ROR\gamma2$, we compared the expression of NABP1 in several thymocyte subpopulations from wild-type and $ROR\gamma^{-/-}$ mice. As shown in Figure 4(C), real-time reverse transcriptase-PCR analyses demonstrated that in wild-type mice the level of NABP1 mRNA was much higher in DP thymocytes compared to double-negative and single-positive thymocytes. Moreover, the expression of NABP1 mRNA was greatly decreased in DP thymocytes from $ROR\gamma^{-/-}$ mice. These results indicate that during thymopoiesis NABP1 and $ROR\gamma2$
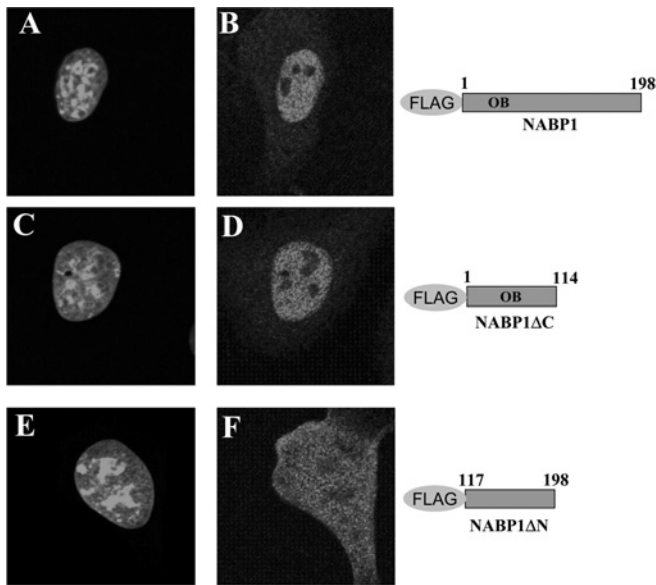
Figure 5   Subcellular localization of NABP1 protein. 3XFLAG–NABP1, 3XFLAG–NABP1ΔN, and 3XFLAG–NABP1ΔC fusion proteins were stably expressed in CV-1 cells

The subcellular localization of these proteins was examined by confocal microscopy using a mouse anti-Flag M2 antibody and an AlexaFluor® 594 goat anti-(mouse IgG) antibody (**B**, **D** and **F**) as described in the Materials and methods section. Nuclei were identified by DAPI staining (**A**, **C** and **E**). (**A**) and (**B**), 3XFLAG–NABP1; (**C**) and (**D**), 3XFLAG–NABP1ΔC; (**E**) and (**F**), 3XFLAG–NABP1ΔN.

mRNA are expressed in a very similar pattern. This finding, together with the observed repression of NABP1 in the thymus of $ROR\gamma^{-/-}$ mice, is in agreement with the hypothesis that NABP1 expression is directly or indirectly controlled by $ROR\gamma2$.

Previous studies have shown that $ROR\gamma$ mediates its transcriptional regulation by binding to ROREs (ROR-response elements) consisting of the core motif, AGGTCA, preceded by an AT-rich region [32]. Examination of the 10 kb promoter-regulatory-region of the mouse NABP1 gene showed the presence of several DNA sequences similar to the consensus RORE. However,

EMSAs (electrophoretic mobility shift assays) demonstrated that $ROR\gamma2$ was unable to bind any of these potential ROREs (results not shown). These results suggest that $ROR\gamma$ regulates NABP1 either through an RORE further up- or down-stream of this 10 kb promoter-regulatory-region, or that it regulates NABP1 expression indirectly.

Subcellular localization of NABP1

To examine the subcellular localization of NABP1, CV-1 cell lines were generated that stably expressed 3XFLAG–NABP1, or the deletion mutants 3XFLAG–NABP1ΔC or 3XFLAG–NABP1ΔN. The subcellular localization of NABP1 was then analysed by confocal microscopy using an anti-FLAG M2 antibody. As shown in Figure 5(B), the expression of the full-length NABP1 protein was largely restricted to the nucleus where it was localized in a speckled pattern. 3XFLAG–NABP1ΔC, containing the N-terminal half of NABP1, including the OB-fold, was also associated with the nucleus (Figure 5D). However, 3XFLAG–NABP1ΔN, containing only the C-terminal half, localized to both the cytoplasm and nucleus (Figure 5F). These observations suggest that NABP1 is a nuclear protein and indicate that the N-terminus plays an important role in retaining NABP1 in the nucleus.

NABP1 is part of a high-molecular-mass protein complex

Since most nuclear proteins do not function alone but interact with other nuclear proteins, we examined whether NABP1 is part of a larger nuclear protein complex. Nuclear lysates were prepared from CV-1 cells stably expressing FLAG–NABP1 and nuclear proteins fractionated by Sepharose S-300 size-exclusion chromatography. The elution of the FLAG–NABP1 protein was monitored by dot-blot and Western blot analysis. As shown in Figure 6, the majority of the eluted FLAG–NABP1 protein was approx. 2000 kDa, suggesting that NABP1 may be part of a high-molecular-mass complex.

NABP1 binds single-stranded nucleic acids

Although OB-folds have been implicated in protein–protein interactions, they are frequently involved in nucleic acid recognition [17,33,34]. Since NABP1 contains an OB-fold motif at
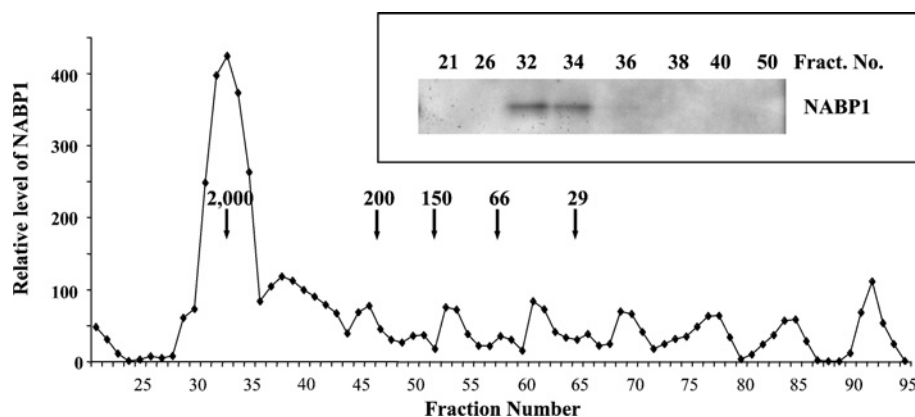

Figure 6   NABP1 is part of a high-molecular-mass protein complex

Nuclear lysates were isolated from CV-1 cells stably expressing the FLAG–NABP1 fusion protein. Proteins were then separated by Sepharcyl 300HR size-exclusion chromatography. The column was calibrated using Blue Dextran (2000 kDa), β-amylase (200 kDa), yeast alcohol dehydrogenase (150 kDa), FBS (66 kDa) and carbonic hydrase (29 kDa) as molecular-mass standards. The arrows indicate the fraction number at which these standards were eluted. Samples of each fraction were examined by immuno dot-blot analysis for the presence of the FLAG–NABP1 protein. The level of immunostaining was quantified and plotted as the relative level of FLAG–NABP1 protein. In addition, samples from several fractions were examined by Western blot analysis with the anti-Flag M2 antibody as shown in the inset.
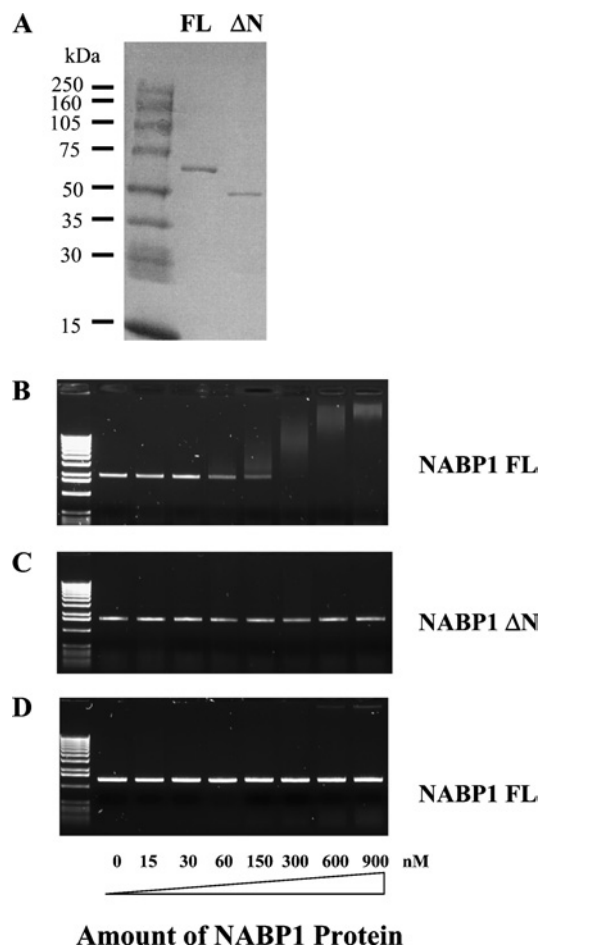
A

B

NABP1 FL

C

NABP1 ∆N

D

NABP1 FL

Amount of NABP1 Protein

Figure 7   NABP1 protein binds single-stranded, but not double-stranded, DNA

(**A**) PAGE of purified MBP–NABP1(full-length, FL) (lane 2) and MBP–NABP1∆N (lane 3) fusion proteins. Molecular-mass markers were run in lane 1. Protein was identified by Coomassie Brilliant Blue staining. Both proteins migrated as single bands of the expected molecular mass. (**B**)–(**D**) Single-stranded ΦX174 viral DNA (**B** and **C**), and double-stranded DNA (**D**) were incubated with increasing amounts of MBP–NABP1(FL) or MBP–NABP1∆N fusion proteins at the concentrations indicated and were then analysed by 0.8% agarose gel electrophoresis as described in the Materials and methods section.

its N-terminus, we analysed whether NABP1 was able to interact with nucleic acids. To investigate this, MBP–NABP1 (full-length) and MBP–NABP1∆N fusion proteins were purified (Figure 7A), incubated with single-stranded viral DNA (ΦX174 DNA) containing 5386 nt, and subsequently DNA–protein complexes were analysed by 0.8% agarose gel electrophoresis. As shown in Figure 7(B), NABP1 caused the single-stranded DNA band to shift to a band with less mobility. The shifted band migrated increasingly slowly with larger amounts of NABP1 protein. This was probably due to the binding of an increasing number of NABP1 molecules to one DNA strand. By contrast, incubation of NABP1 with double-stranded DNA did not cause a significant shift in the mobility of the DNA (Figure 7D). Full-length NABP1 was also able to bind RNA (Figure 8A). Incubation of ΦX174 DNA or RNA with the C-terminus of NABP1 did not cause a change in the mobility of the nucleic acids (Figures 7C and 8B), indicating that as expected the OB-fold domain is required for nucleic acid recognition. Our results demonstrate that NABP1 binds single-stranded DNA and RNA but not double-stranded
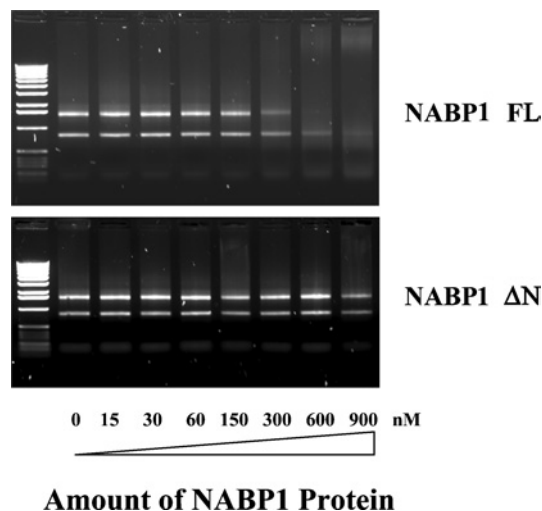

NABP1 FL

NABP1 ∆N

Amount of NABP1 Protein

Figure 8   NABP1 protein binds to RNA

Total cellular RNA (0.5 μg) was incubated for 20 min at room temperature with MBP–NABP1 or MBP–NABP1∆N fusion proteins at the concentrations indicated. The reaction mixture was separated on a 0.8% agarose gel and then visualized by ethidium bromide staining. FL, full-length; CT, C-terminal.

DNA. Therefore NABP1 appears to function as a single-stranded nucleic acid binding protein.

Next, we examined the binding of NABP1 to random oligonucleotides, 10, 20, 30, 50 and 80 nt in length (dN_{10}–dN_{80}). The ^{32}P-labelled oligonucleotides were incubated with increasing amounts of purified NABP1 protein, and the formation of NABP1–oligonucleotide complexes was analysed by EMSA. As shown in Figure 9(A), even in the presence of 2 μM NABP1 protein, very little NABP1 protein bound to oligonucleotide dN_{10} suggesting that NABP1 has a low affinity for dN_{10}. When dN_{20} was incubated with NABP1, a single retarded band was observed at the higher NABP1 concentrations (Figures 9A and 9B). Almost 50% of the probe was shifted in the presence of 2 μM NABP1. We do not yet know whether NABP1, as has been reported for other OB-fold proteins [26,35], binds to nucleic acids as a monomer or multimer. Incubation of dN_{30} and dN_{50} with NABP1 yielded one retarded band at low NABP1 concentrations, whereas at higher NABP1 concentrations an additional band with even lower mobility appeared. The second band probably represents an oligonucleotide that is bound to two single NABP1 proteins, or a multimeric NABP1 complex. With dN_{80}, most of the labelled nucleotides migrated as the second shifted band, suggesting that most oligonucleotides are in complex with two or more NABP1 molecules. Examination of the binding of NABP1 to oligonucleotides $(dT)_{35}$, $(dC)_{35}$, $(dG)_{35}$ and $(dA)_{35}$ demonstrated that NABP1 was able to bind $(dT)_{35}$ and $(dC)_{35}$, but exhibited weak affinity for $(dG)_{35}$ and $(dA)_{35}$ (Figure 9C).

DISCUSSION

In the present study, we describe the identification of two novel, closely related proteins, referred to as NABP1 and NABP2. Both NABPs contain one OB-fold motif at their N-terminus. The OB-folds of NABP1 and NABP2 are 83% identical and are highly conserved between species. The OB-fold of the *Xenopus* homologue of NABP2 exhibits 94% identity with mouse and human NABP2. Outside their OB-fold, NABP1 and NABP2 share little homology. Although the C-terminal halves of human NABP1 and mouse NABP1 are moderately conserved, they contain several
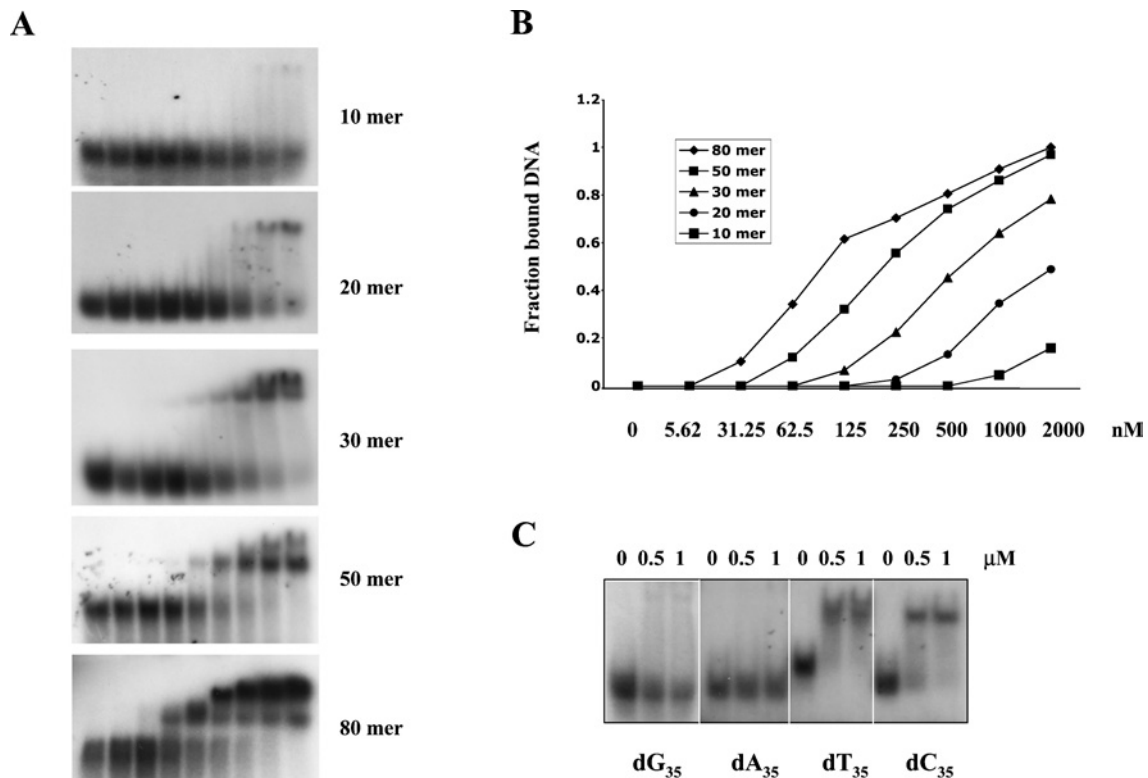
Figure 9   Binding of NABP1 to single-stranded oligonucleotides

(**A**) Single-stranded ^{32}P-labelled oligonucleotides, 10, 20, 30, 50 and 80 nt in length (dN$_{10}$–dN$_{80}$), were incubated with increasing concentrations of NABP1 and then analysed by EMSA. (**B**) The fraction of bound (shifted) DNA was measured by a Fluorskan 8900 apparatus. (**C**) Binding of NABP1 to ^{32}P-labelled single-stranded dG$_{35}$, dA$_{35}$, dT$_{35}$ and dC$_{35}$.

short sequences that are highly homologous. These regions of homology may have functional importance and may play a role in the interaction of NABP1 with other proteins. The OB-folds of NABPs share little sequence homology with those of other known OB-fold proteins. The latter is in agreement with previous observations showing a low level of sequence homology between OB-folds in different proteins [17,34]. Sequence alignment analysis with Meta-BASIC indicated that the OB-folds of NABPs are most closely related to those of archaeal SSB proteins from *Sulfolobus*, *Methanococcus jannaschii* (mjaSBB) and *Thermus thermophilus* [26,31,36].

OB-folds usually span a sequence of 70–150 amino acids and are comprised of 5 to 6 β-strands that form a closed β-barrel that is often capped by an α-helix located between the 3rd and 4th β strands [17,34]. Secondary structure analyses (http://cubic. bioc.columbia.edu/predictprotein) predicted that the OB-folds of NABP1 and NABP2 also contain five β-sheets (Figure 2A) that align with those present in archaeal SBB proteins [26,31,36]. However, the OB-folds of NABPs do not appear to contain an α-helix in the loop between their third and fourth β-strands.

Confocal microscopy demonstrated that the NABP1 protein localized largely to the nucleus. Because NABP1 is small and does not contain a nuclear localization signal, it potentially enters the nucleus by itself through nuclear pores. However, we cannot rule out the possibility that NABP1 enters the nucleus as part of a protein complex that is assembled in the cytoplasm. The accumulation of NABP1 in the nucleus is dependent on the OB-fold since NABP1 that lacks the OB-fold was equally distributed between the cytoplasm and nucleus. Moreover, an NABP1 mutant consisting of only the OB-fold also localized mainly to the nucleus. In the nucleus, NABP1 may become associated with

other nuclear proteins or bind nucleic acids which help to retain it in the nucleus. Analysis of nuclear extracts by Sephacryl 300 HR column chromatography indicated that NABP1 eluted at approx. 2000 kDa supporting the hypothesis that NABP1 may be part of a high-molecular-mass complex. A number of OB-fold proteins have been shown to bind nucleic acids as part of a multimeric complex [19,20]. RPA, which plays a critical role in replication, recombination and repair, consists of a complex of three OB-fold proteins: RPA70, RPA32 and RPA14 [20,27]. Our analysis of nuclear NABP1 complexes by Sephacryl 300 HR column chromatography suggests that this may also be the case for NABP1.

Although OB-folds can be involved in protein–protein inter-actions, they have been largely implicated in the binding of single-stranded nucleic acids (DNA/RNA) [17,33,34]. Our results demonstrate that NABP1 binds single-stranded DNA and RNA, but not double-stranded DNA, suggesting that NABP1 functions as a single-stranded nucleic binding protein. NABP1 exhibited low affinity for dN$_{10}$ oligonucleotides. The affinity of NABP1 for oligonucleotides increased with the length of oligonucleotides, as did the number of NABP1 molecules interacting with the oligonucleotide. SSBs from *S. solfataricus* and *Streptococcus pneumoniae* have been reported to bind nucleic acids as a tetramer [26,35]. Binding of SSB monomers to single-stranded DNA requires 4–5 nt per monomer, whereas the binding of SSB tetramers requires a binding site of 20–25 nt [26]. Based on our observations that NABP1 effectively binds oligonucleotides only when their length exceeds 10 nt, one might conclude that NABP1 proteins bind nucleic acids as a multimer rather than as a monomer. Future studies will determine whether NABP1 is able to form homodimers or tetramers.

Many OB-fold proteins, including SSB proteins, interact with single-stranded DNA in a sequence independent manner and prevent degradation by nucleases or the formation of secondary structures [17]. However, some OB-fold proteins, including Pot1 (protection of telomeres 1) [22], bind specific nucleic acid sequences. We show that NABP1 binds oligo(dT) and oligo(dC) with much higher affinity than oligo(dG) and oligo(dA). The weak affinity for oligo(dG) and oligo(dA); however, might be due to their high propensity to form secondary structures as has been reported for the oligonucleotide binding of human BRCA2 [25]. At present, we cannot rule out whether, *in vivo*, NABP1 interacts with a specific nucleic acid sequence with a higher affinity than with random oligonucleotides.

NABP1 and NABP2 are expressed in many tissues (Figure 3). Although NABP2 mRNA is expressed in all tissues almost equally, NABP1 is most highly expressed in testis and thymus. Through alternative splicing, the NABP1 gene generates two variant transcripts that encode the same protein but differ in their 3′-UTRs. Real-time reverse transcriptase-PCR analyses demonstrated that in the thymus NABP1 expression is differentially regulated and is most highly expressed in DP thymocytes. These results suggest that NABP1 expression is induced when double-negative thymocytes differentiate into DP thymocytes and is repressed when DP thymocytes mature into single-positive T lymphocytes. This expression pattern is very similar to that reported for the nuclear receptor, $ROR\gamma 2$ [2,10]. DP thymocytes are only one of two cell types in which $ROR\gamma 2$ ($ROR\gamma t$) has been reported to be expressed, whereas $ROR\gamma 1$ is expressed in the remaining tissues. Northern blot and real-time reverse transcriptase-PCR analyses further demonstrated that NABP1 expression was greatly repressed in the thymus and DP thymocytes of $ROR\gamma^{-/-}$ mice. In contrast with NABP1, NABP2 expression did not change in the thymus of $ROR\gamma^{-/-}$ mice. Moreover, the lack of $ROR\gamma$ did not affect NABP1 expression in other tissues. Thus the expression of NABP1 in $ROR\gamma^{+/+}$ DP thymocytes and the decreased expression in $ROR\gamma^{-/-}$ DP thymocytes correlates with that of $ROR\gamma 2$. These observations suggest that the expression of NABP1 is dependent on $ROR\gamma 2$ and that transcription of the NABP1 gene may be directly or indirectly regulated by $ROR\gamma 2$. The regulation of NABP1 in DP thymocytes resembles that of Bcl-X$_L$, another gene controlled by $ROR\gamma 2$ [9,10]. Like NABP1 and $ROR\gamma 2$, Bcl-X$_L$ expression is restricted to DP thymocytes and is repressed in $ROR\gamma^{-/-}$ DP thymocytes. Repression of Bcl-X$_L$ in $ROR\gamma^{-/-}$ DP thymocytes is the main cause for the accelerated apoptosis observed in these cells [9,10]. RORs regulate gene transcription by binding to ROREs in the regulatory region of target genes [1,5,32,37]. Whether $ROR\gamma 2$ regulates the transcription of NABP1 and Bcl-X$_L$ by a direct or indirect mechanism has yet to be established. Analysis of a 10 kb promoter-regulatory-region of NABP1 identified several potential ROREs. However, EMSA analyses showed that $ROR\gamma 2$ did not bind to any of these potential ROREs. These findings suggest that NABP1 might be regulated by an RORE either up- or downstream of this 10 kb regulatory region or is regulated indirectly by $ROR\gamma 2$, possibly through another transcription factor the expression of which is controlled by $ROR\gamma 2$.

In summary, in the present study we have identified two novel nucleic acid binding proteins, NABP1 and NABP2. We show that in the thymus the expression of NABP1, but not that of NABP2, is dependent upon the expression of $ROR\gamma 2$. These observations suggest that transcription of the NABP1 gene is directly or indirectly regulated by $ROR\gamma 2$. We further demonstrate that NABP1 is a nuclear protein that binds single-stranded nucleic acids. Although the precise function of NABP1 has yet to be established, its specific expression in DP thymocytes, cells that undergo a high frequency of DNA rearrangements, raises the possibility of a role in DNA recombination or repair in these cells.

REFERENCES

1   Jetten, A. M. (2004) Recent Advances in the Mechanisms of Action and Physiological Functions of the Retinoid-related Orphan Receptors (RORs). Current Drug Targets. Inflammation and Allergy **3**, 395–412

2   He, Y. W., Deftos, M. L., Ojala, E. W. and Bevan, M. J. (1998) RORγ t, a novel isoform of an orphan receptor, negatively regulates Fas ligand expression and IL-2 production in T cells. Immunity **9**, 797–806

3   He, Y. W. (2002) Orphan nuclear receptors in T lymphocyte development. J. Leukoc. Biol. **72**, 440–446

4   Eberl, G. and Littman, D. R. (2003) The role of the nuclear hormone receptor RORγ t in the development of lymph nodes and Peyer's patches. Immunol. Rev. **195**, 81–90

5   Jetten, A. M., Kurebayashi, S. and Ueda, E. (2001) The ROR nuclear orphan receptor subfamily: critical regulators of multiple biological processes. Prog. Nucleic Acid Res. Mol. Biol. **69**, 205–227

6   Steinmayr, M., André, E., Conquet, F., Rondi-Reig, L., Delhaye-Bouchaud, N., Auclair, N., Daniel, H., Crépel, F., Mariani, J., Sotelo, C. and Becker-André, M. (1998) *Staggerer* phenotype in retinoid-related orphan receptor α-deficient mice. Proc. Natl. Acad. Sci. U.S.A. **95**, 3960–3965

7   Hamilton, B. A., Frankel, W. N., Kerrebrock, A. W., Hawkins, T. L., FitzHugh, W., Kusumi, K., Russell, L. B., Mueller, K. L., van Berkel, V., Birren, B. W. et al. (1996) Disruption of the nuclear hormone receptor RORα in staggerer mice. Nature (London) **379**, 736–739

8   Andre, E., Conquet, F., Steinmayr, M., Stratton, S. C., Porciatti, V. and Becker-Andre, M. (1998) Disruption of retinoid-related orphan receptor β changes circadian behavior, causes retinal degeneration and leads to vacillans phenotype in mice. EMBO J. **17**, 3867–3877

9   Kurebayashi, S., Ueda, E., Sakaue, M., Patel, D. D., Medvedev, A., Zhang, F. and Jetten, A. M. (2000) Retinoid-related orphan receptor γ (RORγ) is essential for lymphoid organogenesis and controls apoptosis during thymopoiesis. Proc. Natl. Acad. Sci. U.S.A. **97**, 10132–10137

10 Sun, Z., Unutmaz, D., Zou, Y. R., Sunshine, M. J., Pierani, A., Brenner-Morton, S., Mebius, R. E. and Littman, D. R. (2000) Requirement for RORγ in thymocyte survival and lymphoid organ development. Science (Washington DC) **288**, 2369–2373

11 Dussault, I., Fawcett, D., Matthyssen, A., Bader, J. A. and Giguere, V. (1998) Orphan nuclear receptor ROR α-deficient mice display the cerebellar defects of staggerer. Mech. Dev. **70**, 147–153

12 Gold, D. A., Baek, S. H., Schork, N. J., Rose, D. W., Larsen, D. D., Sachs, B. D., Rosenfeld, M. G. and Hamilton, B. A. (2003) RORα coordinates reciprocal signaling in cerebellar development through sonic hedgehog and calcium-dependent pathways. Neuron **40**, 1119–1131

13 Schaeren-Wiemers, N., Andre, E., Kapfhammer, J. P. and Becker-Andre, M. (1997) The expression pattern of the orphan nuclear receptor RORβ in the developing and adult rat nervous system suggests a role in the processing of sensory information and in circadian rhythm. Eur. J. Neurosci. **9**, 2687–2701

14 Guo, J., Hawwari, A., Li, H., Sun, Z., Mahanta, S. K., Littman, D. R., Krangel, M. S. and He, Y. W. (2002) Regulation of the TCRα repertoire by the survival window of CD4(+)CD8(+) thymocytes. Nat. Immunol. **3**, 469–476

15 Ueda, E., Kurebayashi, S., Sakaue, M., Backlund, M., Koller, B. and Jetten, A. M. (2002) High incidence of T-cell lymphomas in mice deficient in the retinoid-related orphan receptor RORγ. Cancer Research **62**, 901–909

16 Arcus, V. (2002) OB-fold domains: a snapshot of the evolution of sequence, structure and function. Current Opin. Struc. Biol. **12**, 794–801

17 Theobald, D. L., Mitton-Fry, R. M. and Wuttke, D. S. (2003) Nucleic acid recognition by OB-fold proteins. Annu. Rev. Biophys. Biomol. Struct. **32**, 115–133

18 Bochkarev, A., Bochkareva, E., Frappier, L. and Edwards, A. M. (1999) The crystal structure of the complex of replication protein A subunits RPA32 and RPA14 reveals a mechanism for single-stranded DNA binding. EMBO J. **18**, 4498–4504

19 Callebaut, I. and Mornon, J. P. (1997) The human EBNA-2 coactivator p100: multidomain organization and relationship to the staphylococcal nuclease fold and to the tudor protein involved in Drosophila melanogaster development. Biochem. J. **321**, 125–132

20 Iftode, C., Daniely, Y. and Borowiec, J. A. (1999) Replication protein A (RPA): the eukaryotic SSB. Crit. Rev. Biochem. Mol. Biol. **34**, 141–180

21 Kerr, I. D., Wadsworth, R. I., Cubeddu, L., Blankenfeldt, W., Naismith, J. H. and White, M. F. (2003) Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of Sulfolobus SSB protein. EMBO J. **22**, 2561–2570

22 Lei, M., Podell, E. R., Baumann, P. and Cech, T. R. (2003) DNA self-recognition in the structure of Pot1 bound to telomeric single-stranded DNA. Nature (London) **426**, 198–203

23 Li, W. and Hoffman, D. W. (2001) Structure and dynamics of translation initiation factor aIF-1A from the archaeon Methanococcus jannaschii determined by NMR spectroscopy. Protein Sci. **10**, 2426–2438

24 Theobald, D. L., Cervantes, R. B., Lundblad, V. and Wuttke, D. S. (2003) Homology among telomeric end-protection proteins. Structure **11**, 1049–1050

25 Yang, H., Jeffrey, P. D., Miller, J., Kinnucan, E., Sun, Y., Thoma, N. H., Zheng, N., Chen, P. L., Lee, W. H. and Pavletich, N. P. (2002) BRCA2 function in DNA binding and recombination from a BRCA2-DSS1-ssDNA structure. Science (Washington DC) **297**, 1837–1848

26 Wadsworth, R. I. and White, M. F. (2001) Identification and properties of the crenarchaeal single-stranded DNA binding protein from *Sulfolobus solfataricus*. Nucleic Acids Res. **29**, 914–920

27 Wold, M. S. (1997) Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. Annu. Rev. Biochem. **66**, 61–92

28 Edgar, R., Domrachev, M. and Lash, A. E. (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucl. Acids Res. **30**, 207–210

29 Daly, A. K., Redfern, C. P. F. and Martin, B. (1990) Identification and analysis of retinoic acid-binding proteins and receptors from nuclei of mammalian cells. Meth. Enzymol. **189**, 239–247

30 Eberl, G., Marmon, S., Sunshine, M. J., Rennert, P. D., Choi, Y. and Littman, D. R. (2004) An essential function for the nuclear receptor RORγt in the generation of fetal lymphoid tissue inducer cells. Nat. Immunol. **5**, 64–73

31 Kelly, T. J., Simancek, P. and Brush, G. S. (1998) Identification and characterization of a single-stranded DNA-binding protein from the archaeon Methanococcus jannaschii. Proc. Natl. Acad. Sci. U.S.A. **95**, 14634–14639

32 Medvedev, A., Yan, Z. H., Hirose, T., Giguere, V. and Jetten, A. M. (1996) Cloning of a cDNA encoding the murine orphan receptor RZR/ROR γ and characterization of its response element. Gene **181**, 199–206

33 Agrawal, V. and Kishan, K. V. (2003) OB-fold: growing bigger with functional consistency. Curr. Protein Pept. Sc. **4**, 195–206

34 Bochkarev, A. and Bochkareva, E. (2004) From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold. Cur. Opin. Struc. Biol. **14**, 36–42

35 Grove, D. E., Willcox, S., Griffith, J. D. and Bryant, F. R. (2005) Differential single-stranded DNA binding properties of the paralogous SsbA and SsbB proteins from *Streptococcus pneumoniae*. J. Biol. Chem.**280**, 11067–11073

36 Dabrowski, S., Olszewski, M., Piatek, R., Brillowska-Dabrowska, A., Konopa, G. and Kur, J. (2002) Identification and characterization of single-stranded-DNA-binding proteins from *Thermus thermophilus* and *Thermus aquaticus* – new arrangement of binding domains. Microbiology **148**, 3307–3315

37 Giguere, V., Tini, M., Flock, G., Ong, E., Evans, R. M. and Otulakowski, G. (1994) Isoform-specific amino-terminal domains dictate DNA-binding properties of ROR α, a novel family of orphan hormone nuclear receptors. Genes Dev. **8**, 538–553