Conserved amino acid sequences confer nuclear localization upon the Prophet of Pit-1 pituitary transcription factor protein

J. Chico Guy, Chad S. Hunter, Aaron D. Showalter, Timothy P.L. Smith, Kanokwan Charoonpatrapong, Kyle W. Sloop, Joseph P. Bidwell, Simon J. Rhodes

Department of Biology, Indiana University-Purdue University Indianapolis, 723 West Michigan Street, Indianapolis IN 46202-5132, USA
Endocrine Discovery, Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, IN, USA
USDA/ARS U.S. Meat Animal Research Center, Clay Center, NE, USA
Department of Anatomy and Cell Biology, Indiana University School of Medicine, Indianapolis, IN, USA

Received 5 March 2004; received in revised form 4 April 2004; accepted 22 April 2004
Available online 19 June 2004
Received by A.J. van Wijnen

Abstract

Prophet of Pit-1 (PROP1) is a homeodomain transcription factor essential for development of the mammalian anterior pituitary gland. Studies of human patients and animal models with mutations in their Prop1 genes have established that PROP1 is required for the correct development or sustained function of the hormone-secreting cells that regulate physiological pathways controlling growth, reproduction, metabolism, and the stress response. By comparative analysis of mammalian Prop1 genes and their encoded proteins, including cloning the ovine Prop1 gene and its products, we demonstrate that two conserved basic regions (B1 and B2) of the PROP1 protein located within the homeodomain are required for nuclear localization, DNA binding, and target gene activation. Interestingly, missense mutations in the human Prop1 gene causing amino acid changes in both the B1 and B2 regions have been associated with combined pituitary hormone deficiency (CPHD) diseases, suggesting that disruption of nuclear localization may be part of the molecular basis of such diseases. The ovine Prop1 gene has three exons and two introns, a different structure compared with that of the bovine gene. Two alleles of the ovine gene were found to encode protein products with different carboxyl terminal domain sequences. We demonstrate that the two alleles are distributed in different breeds of sheep. Finally, we show for the first time that the PROP1 protein is associated with the nuclear matrix.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Hormone deficiency disease; Nuclear matrix; Ovine; Alleles

1. Introduction

Following early inductive signaling events, the development of the hormone-secreting cells of the anterior pituitary gland is guided by selective combinations of transcription factors (reviewed in Keegan and Camper, 2003; Olson et al., 2003; Savage et al., 2003). The molecular targets of one of these gene regulatory proteins, Prophet of Pit-1 (or PROP1), appear to be restricted to the anterior pituitary. In animals and human patients, mutations in Prop1 cause recessively inherited diseases featuring deficiencies of all of the anterior pituitary hormones in the most severe cases (Sornson et al., 1996; Cushman et al., 2002; Mody et al., 2002).

Prop1 is expressed during early pituitary development, declining to significantly lower levels during the later stages of organogenesis and in the adult gland (Sornson et al., 1996; Sloop et al., 2000). Transgenic experiments have shown that the transient nature of Prop1 expression is critical for appropriate development of specific differentiated cells of the pituitary or the entire gland. Transgenic mice that express Prop1 before embryonic day 10 do not develop pituitaries (Dasen et al., 2001). Alternately, mice carrying transgenes containing pituitary hormone promoters that sustain Prop1

Abbreviations: CPHD, combined pituitary hormone deficiency; EMSA, electrophoretic mobility shift assay; PBS, phosphate-buffered saline; PCR, polymerase chain reaction; PROP1, Prophet of Pit-1; SNP, single nucleotide polymorphism; UTR, untranslated region.

* Corresponding author. Tel.: +1-317-278-1797; fax: +1-317-274-2846.

0378-1119/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
expression in pituitary cell types after the normal time at which Prop1 mRNA levels decrease exhibit delayed gonadotrope maturation and develop pituitary tumors (Cushman et al., 2001). Recent studies have demonstrated that Prop1, in concert with other transcription factors such as LHX4, plays important roles in the early steps of pituitary organogenesis (Raetzman et al., 2002). The genetic targets of Prop1 are poorly understood, but it is clear that Prop1-dependent outcomes include the three cell types that are dependent on the actions of the downstream Pit1 transcription factor, the thyrotropes, somatotropes, and lactotropes. The Pit1 gene itself contains Prop1 recognition elements and can be activated by Prop1 (Sorson et al., 1996; Vieira et al., 2003). Prop1 gene action is also required for appropriate expression of Notch2, a critical component of a signaling pathway implicated in pituitary cell specification (Raetzman et al., 2004).

The Prop1 protein contains a central paired-class homeodomain that mediates the interaction of homodimeric Prop1 complexes with AT-rich consensus DNA elements (Sorson et al., 1996). Prop1 potently activates reporter genes containing these DNA binding sites, and this transcriptional response requires a conserved activation domain located within the carboxyl terminus (Sorson et al., 1996; Showalter et al., 2002). The amino termini and the homeodomain of human and bovine Prop1 can transfer repressive activity to model DNA binding proteins, suggesting that Prop1 has the potential to be both a transcriptional activator and repressor (Showalter et al., 2002).

In order to serve their gene regulatory functions, transcription factors must enter the nuclear compartment. The nuclear transport process is mediated by the importin/karyopherin family of proteins that recognize nuclear localization signals (NLSs) within nuclear proteins (Cokol et al., 2000; Jans et al., 2000; Adam, 2001). NLS motifs include short sequences rich in basic amino acids, bipartite sequences containing basic residues, sequences containing mixtures of charged and polar amino acids, and/or sequences with amino acids modified by phosphorylation (Jans et al., 2000). Nuclear proteins such as transcription factors may contain single or multiple NLSs: for example, the Lhx3 transcription factor contains four major NLSs that act in combination to guide nuclear localization of the protein (Parker et al., 2000). Nuclear localization of some transcription factors is regulated, dependent on activation of the proteins by signaling pathways, as for the SMAD and STAT proteins (Jans et al., 2000). In other cases, developmental signals may control the temporal and spatial cues that cause cytoplasmic retention of “nuclear” transcription factors, such as for the Otx1 homeodomain transcription factor in cortical neurons (Zhang et al., 2002). Once proteins have entered the nucleus, they may partition further to nuclear subcompartments. These locations include soluble fractions and insoluble structures such as the nuclear matrix, a proteinaceous structure that is resistant to nuclease digestion and high salt extraction (Fey et al., 1984). The nuclear matrix may serve a role as a scaffold for chromatin, and may participate as a mediator of some extranuclear and extracellular regulatory signals that result in altered gene expression (Bidwell et al., 2001; Alvarez et al., 2003; Stein et al., 2003).

In this study, we demonstrate that two basic regions located within the homeodomain of the Prop1 protein are conserved and required for nuclear localization. The basic sequences are essential for DNA binding and target gene activation by Prop1. We also show that the Prop1 protein is associated with the nuclear matrix.

2. Materials and methods

2.1. Cloning of genomic and cDNA sequences and plasmid construction

Genomic DNA was prepared from blood of adult Dorset sheep by a salt extraction procedure as described (Miller et al., 1988). Total RNA was extracted from the pituitary glands of adult Dorset × Romanov cross sheep using Tri-Reagent/bromochloroethane (Molecular Research Center, Cincinnati, OH).

The second intron of the ovine Prop1 (oProp1) gene was amplified from genomic DNA using primers designed via analyses of the porcine and human Prop1 gene sequences. Polymerase chain reactions (PCR) employed PfuUltra DNA polymerase (Stratagene, La Jolla, CA) and the following primers: 5'-gccagggcagagcttgtgg-3' and 5'-agggagacgtggcaggtgg-3'. The entire oProp1 gene was then amplified as two, overlapping fragments. The first fragment (5' end through intron 2) was generated using 5'-acaggagacaggtgctttgctg-3' and 5'-aagtgaagccactactctctc-3'. The second fragment (intron 2 through 3' end) was generated using 5'-aattccacagctagcgcgt-3' and 5'-ttcagacgtttgaagggaa-3'. DNA fragments were cloned into pCRII-TOPO (Invitrogen, Carlsbad, CA). To ensure sequence fidelity, products from several amplification reactions were sequenced on both strands by automated DNA sequencing (Biochemistry Biotechnology Facility, IUPUI). Sequences were analyzed using DNASIS (Hitachi Software Engineering, South San Francisco, CA) and BLAST at NCBI.

Complementary DNA fragments encompassing the full-length oProp1 ORF were amplified from single-stranded cDNA derived from adult pituitary gland total RNA using the following primers: 5'-ccggagatccagacagacagcgga-3' and 5'-cggagatccagcagacagcggga-3'. The resulting products (isoform with threonine at residue 181—see below) were digested with EcoRI and BamHI and cloned into pCDNA3.1-MyC/His expression vectors (Invitrogen) for sequence analysis and functional tests. Products from several amplification reactions were sequenced on both strands. Completed gene and cDNA
sequences were deposited in GenBank (accession numbers AY533708 [gene] and AY533709 [cDNA]).

To test whether a fourth intron (as is found in bovine Prop1 [bProp1]; Showalter et al., 2002) was present in the oProp1 gene, cDNA fragments representing the 3' end of the oProp1 mRNA were first amplified from DNase I-treated adult pituitary gland cDNA using the following primers: 5'-ttgtttaacccagcttgagagc-3' and 5'-aatctcttgttgataaggattgg-3'. Subsequently, two nested PCR reactions were performed using the following primer pairs: 5'-aagttggggtaccagtct-3' and 5'-ccagagaaccgcagagctaag-3'. DNA fragments were cloned into pCRRII-TOPO and sequenced.

The ORF of the bProp1 cDNA (Showalter et al., 2002) was cloned in-frame into the EcoRI and BamHI sites of the pEGFP-C1 vector (Clontech, Palo Alto, CA) to create an EGFP-bPROP1 expression plasmid. The ORFs of wild type (threonine 181 isomorph) or mutated oProp1 cDNAs were cloned into the EcoRI and BamHI sites of the pEGFP-N1 vector (Clontech) to generate oPROP1-EGFP expression plasmids.

2.2. Genotype analysis of oProp1 alleles

Genotyping of a sheep breed panel was performed using a primer extension method with mass spectrometry-based analysis of the extension products on a MassArray system software (Sequenom, San Diego, CA) as described (Stone et al., 2002). Genotypes for each animal were collected in the Spectrotyper database, and the automated calls were checked by manual visualization of the spectrographs to minimize error. Marker-specific primer sequences for genotyping the single nucleotide polymorphism (SNP) predicting amino acid variation (residue 181 threonine/alanine) in oProp1 were: 5'-ccagaacgcagctaga-3' (forward amplification primer), 5'-ngaagtggtggaactagct-3' (reverse amplification primer), and 5'-tcctcctacccctc-3' (probe primer). Additional mass tag sequence (5'-gacctggtagt) was added to the 5' end of the amplification primers as recommended by the MassArray system software.

2.3. Site-directed mutagenesis

Site-directed mutagenesis was performed using Quick-Change II (Stratagene) according to the manufacturer's protocol, except that plasmids were transformed into Escherichia coli DH5alpha (Invitrogen). Mutagenic oligonucleotides were 5'-ggtgagctacggctgggtgccgagcagagctggagcgcaccctgcagc-3', 5'-gaagttggggtaccagtctagctgaggctgagcagagccgagcagcaccctgcagc-3' (oPROP1 basic region 1 [B1], changing RRRHR to RAAAA) and 5'-ttgagctacggctgggtgccgagcagagctggagcgcaccctgcagc-3', 5'-ggtgagctacggctgggtgccgagcagagctggagcgcaccctgcagc-3' (basic region 2 [B2], changing RRAKQ RK to AAAAQ RK). The resulting plasmids were confirmed by DNA sequencing.

2.4. Cell culture and DNA transfection

Mouse GHFT1 pituitary cells, human embryonic kidney 293T cells, and rat UMR 106-01 cells were cultured as described (Parker et al., 2000; Sloop et al., 2001). Cells were transiently transfected using the CalPhos system (Clontech) or with Lipofectamine (Invitrogen).

2.5. Confocal and fluorescence microscopy and immunolocalization

Cells were plated on chamber slide flskettes (Nalge Nunc, Naperville, IL) and were transiently transfected with expression vectors as described above. Twenty-four hours post-transfection, cells were fixed in 4% formalin in phosphate-buffered saline (PBS), pH 7.7. To detect EGFP, EGFP-bPROP1, and oPROP1-EGFP proteins, the nuclei of cells were counterstained with Hoechst 33258 dye and then visualized by fluorescence microscopy using a Nikon Eclipse TE 200-U inverted microscope or by confocal microscopy using a Zeiss LSM510 confocal microscope with Coherent Enterprises UV and argon lasers. To detect myc epitope-tagged PROP1 proteins, fixed cells were blocked in 10% goat serum (Sigma, St. Louis, MO), followed by the addition of an anti-myc antibody (9E10 monoclonal, Developmental Studies Hybridoma Bank, University of Iowa, Iowa City, IA, 1:500). After washing, goat anti-mouse IgG-Alexa Fluor-488 (Molecular Probes, Eugene, OR) was added and nuclear counterstaining was performed with SlowFade Light antifade containing 4', 6-diamidino-2-phenylindole dihydrochloride (DAPI, Molecular Probes).

2.6. In vitro transcription and translation

Radiolabeled PROP1 proteins were synthesized in vitro using TnT Quick Coupled rabbit reticulocyte lysate reagents (Promega, Madison, WI) and 35S-methionine (Amersham Pharmacia, Piscataway, NJ). Unlabeled proteins were synthesized via the same method by replacing the radioactive methionine with cold methionine. Substrate DNAs were 2 μg of pcDNA3 or pcDNA3.1 vectors containing human or ovine wild type or mutated Prop1 cDNAs. Radiolabeled proteins were analyzed by SDS electrophoresis using 12% polyacrylamide gels followed by treatment with Amplify fluorography reagent (Amersham Pharmacia) and exposure to Kodak BioMax MR film, or visualized using a Storm phosphorimager (Amersham Pharmacia).

2.7. Electrophoretic mobility shift analysis

Electrophoretic mobility shift assays (EMSA) were performed using a 32P-labeled oligonucleotide containing a high-affinity binding site for paired class proteins with a glutamine at position 9 of the third helix of the homeodomain (PRDQ9) as previously described (Sloop et al., 2001).
Equivalent amounts (~7 µl) of in vitro translated oPROP1 proteins or human PROP1 protein were added per binding reaction. Bound complexes were visualized by exposure to Kodak BioMax MR film, or using a Storm phosphorimager.

2.8. Gene regulation assays

Typically, 1.5 × 10^5 293T cells in a 35 mm well were transfected with 1 µg of reporter gene plasmid and 50 ng of expression vectors. Control parallel samples received empty expression vector DNA. All assay groups were performed in triplicate. Forty-eight hours following transfection, cells were lysed in 25 mM Tris (pH 7.8), 2 mM dithiothreitol, 1% Triton X-100, 2 mM ethylenediaminetetraacetic acid (pH 8.0), 10% glycerol, and luciferase activity was measured as described (Sloop et al., 2001). Protein assays were performed by Bradford method (BioRad, Hercules, CA), and luciferase activity was normalized to protein concentration.

2.9. In situ nuclear matrix extraction

Twenty-four hours prior to transfection, UMR 106-01 cells were plated at 4 × 10^4 cells/10 cm² on coverslips in six-well plates. Cells were transfected using 1 µg plasmid DNA. After 48 h, cells either were directly fixed in 3.7% formaldehyde or subjected to sequential extraction, stained, and mounted for immunofluorescence. Soluble cytoskeletal and nuclear proteins were sequentially extracted from cells as described (Parker et al., 2000). Cells then were fixed and processed for immunodetection of nuclear antigens and chromatin staining. Nuclear matrix filaments were detected with an anti-NuMA antibody and removal of nuclear DNA was verified by staining with DAPI as described (Parker et al., 2000). Phase contrast and fluorescent digital images were captured using a Nikon inverted epifluorescence microscope (Nikon, Garden City, NY) with a phase contrast 100× (NA1.3) oil-immersion lens. Fluorescent images of triple-labeled cells were obtained using narrow band-pass rhodamine, fluorescein, and DAPI filters and captured using a CCD camera (Photometrics, Tucson, AZ).

3. Results and discussion

3.1. Cloning of the ovine Prop1 gene and pituitary cDNAs

To examine conservation of mammalian PROP1 proteins, we deduced the primary sequence of the ovine PROP1 (oPROP1) protein from genomic clones and from cDNAs derived from pituitary gland RNA. The putative second intron of the gene was first amplified from genomic DNA using primers based upon porcine and human sequences (Wu et al., 1998; Sloop et al., 2000). This sequence information was then used to design primer pairs allowing amplification of the entire gene in two fragments that overlapped in the second intron region. Primers for the 5′ and 3′ ends of the gene were based upon bovine gene sequences (Showalter et al., 2002). The 3510 bp oProp1 gene has three exons and two introns (Fig. 1A), a structure similar to that of some other mammals such as mice and pigs (Sornson et al., 1996; Sloop et al., 2000). The exon/ intron junctions of the gene exhibit consensus RNA splice Fig. 1. Cloning and analysis of the ovine Prop1 gene. (A) Structure of the oProp1 gene. The gene was amplified from genomic DNA and sequenced. Exons are labeled in Roman numerals and introns are labeled in Arabic numerals. Black boxes denote translated regions. A diagram of the oPROP1 protein is shown indicating the regions of the protein encoded by each exon. N = amino terminus, and C = carboxyl terminus. The GenBank accession number of the oProp1 gene is AY533708. (B) Features of oProp1 exons, introns, and exon–intron junctions. Protein-coding sequences are in upper case, intronic sequences are in lower case, and splice donor and acceptor consensus RNA splice

### Table: Exon Identity Size (bp)
<table>
<thead>
<tr>
<th>Exon</th>
<th>Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>≥109</td>
</tr>
<tr>
<td>II</td>
<td>233</td>
</tr>
<tr>
<td>III</td>
<td>≥1049</td>
</tr>
</tbody>
</table>

### Table: Sequence at Exon-Intron Boundary

<table>
<thead>
<tr>
<th>Splice Acceptor</th>
<th>Splice Donor</th>
</tr>
</thead>
</table>
| TCTGGGTTGGgtgag | TTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT...
acceptor/donor sequences (Fig. 1B). Interestingly, the bovine gene has a third intron and an untranslated fourth exon (Showalter et al., 2002). Because of the overall similarities between the bovine and ovine genomes, we performed additional experiments to confirm that the oProp1 gene indeed lacked a third intron. Using primers representing exon II and putative 3′ untranslated sequences, DNA fragments were amplified from pituitary gland cDNA. These experiments demonstrated that intron 2 sequences were appropriately spliced (providing a control for genomic contamination) but that sequences in the region encompassed by intron 3 of the bovine gene were retained as 3′ UTR, confirming the ovine gene structure (data not shown).

Complementary DNAs encompassing the full-length ORF of oProp1 were amplified from pituitary gland cDNA, and the sequence of the oProp1 protein was deduced (Fig. 2A). The predicted oPROP1 protein is 226 amino acids with a proline-rich carboxyl terminus. The amino terminus, which can act as a transferable repression domain based on studies of the human protein (Showalter et al., 2002), is poorly conserved with only 48–50% identity to rodent PROPI sequences (Fig. 2B). The DNA binding homeodomain is strikingly conserved with greater than 90% identity shared between the sheep and other species. Notably, two basic sequences at the ends of the homeodomain are perfectly conserved in all known PROPI proteins (Fig. 2B). The carboxyl terminus, the location of the major trans-activation domain of PROPI (Sornson et al., 1996; Showalter et al., 2002), is also strongly conserved (Fig. 2B). Phylogenetic studies place oPROP1 close to bPROP1, with less relationship to rodent PROPI sequences (Fig. 2C).

### 3.2. Two alleles of ovine Prop1 encode distinct proteins

DNA sequence analyses of multiple cDNA and genomic clones identified an apparent polymorphism in the protein coding region predicted to change the amino acid at position 181 from a threonine (ACA codon) to an alanine (GCA codon) (Fig. 2A). The equivalent residue is a threonine in all other sequenced species. To test whether the putative amino

---

**A**

<table>
<thead>
<tr>
<th></th>
<th>MDTEGRSEQAQEQKAEQCVSSLWPEGYPAAETVSSVDMNTQPYRNSGVRVGRPKLSLQGGQRGRPHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>69</td>
<td>RRRRRRTFSPAQLEQLESAGKNGQYPDINARESQAQTGLSAREITQVFOWFRRAKQRKQLRSSQQLA</td>
</tr>
<tr>
<td>137</td>
<td>HLSPATFSGFLPEPPSCPYPTPPPPMTCFPHYPYNHALPSQPSTGSSFARPYQSEDWYPNLHP</td>
</tr>
<tr>
<td>205</td>
<td>HLPFFFFPPMLPLSLFEPPKSNW</td>
</tr>
</tbody>
</table>

**B**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>B1</th>
<th>HD</th>
<th>B2</th>
<th>C</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cow</td>
<td>94</td>
<td>100</td>
<td>98</td>
<td>100</td>
<td>98</td>
<td>97</td>
</tr>
<tr>
<td>Pig</td>
<td>71</td>
<td>100</td>
<td>95</td>
<td>100</td>
<td>91</td>
<td>86</td>
</tr>
<tr>
<td>Human</td>
<td>50</td>
<td>100</td>
<td>95</td>
<td>100</td>
<td>83</td>
<td>76</td>
</tr>
<tr>
<td>Dog</td>
<td>65</td>
<td>100</td>
<td>90</td>
<td>100</td>
<td>91</td>
<td>83</td>
</tr>
<tr>
<td>Mouse</td>
<td>48</td>
<td>100</td>
<td>95</td>
<td>100</td>
<td>77</td>
<td>73</td>
</tr>
<tr>
<td>Rat</td>
<td>50</td>
<td>100</td>
<td>92</td>
<td>100</td>
<td>78</td>
<td>73</td>
</tr>
</tbody>
</table>

**C**

- Rat: 86%
- Mouse: 85%
- Sheep: 97%
- Cow: 66%
- Pig: 72%
- Dog: 81%
- Human: 66%

Fig. 2. Differential conservation of mammalian PROPI protein domains. (A) Protein sequence of ovine PROPI. The homeodomain is boxed. Two basic regions (B1 and B2) are shown in reverse type. A variant amino acid at residue 181 is indicated (T181 or A181). The GenBank accession number of the oProp1 cDNA is AY533709. (B) Comparison of ovine PROPI with other mammalian PROPI proteins. The percent amino acid identities of individual domains and the overall primary sequence identity compared to ovine PROPI are indicated. N = amino terminus, HD = homeodomain, and C = carboxyl terminus. (C) Dendrogram depicting the phylogenetic relationships between mammalian PROPI proteins. Numbers are percent identities between the indicated pairs.
acid substitution at codon 181 represented distinct alleles of the oProp1 gene rather than a product of RNA editing or a cloning artifact, the corresponding region of oProp1 was amplified from genomic DNA in a panel of animals representing 8 production sheep breeds and one interbreed composite population (Freking et al., 2002) and genotyped for the nucleotide substitution responsible for the two predicted isoforms. A sample of four Bighorn sheep (Ovis canadensis) was also successfully amplified and genotyped with the same primer set. Both alternative homozygous oProp1 genotypes (AA and GG) and heterozygous (AG) individuals were observed, confirming that the two isoforms are present in normal populations of sheep (Table 1). The A allele encoding a conserved threonine residue is the minor allele in this multibreed panel. There is no indication of active selection for one or the other allele, as they are in Hardy–Weinberg equilibrium in the panel ($\chi^2 = 0.0014$). The A allele frequency in production sheep varies from 0.15 to 0.70 depending on breed, although the number of animals tested per breed is too low to determine if this variation is significant or could be related to phenotypic differences among breeds. All four Bighorn sheep were homozygous for the A allele, which suggests that the alternative isoform is a relatively recent variation. However, in the absence of selection, the presence of both alleles in all of the production sheep breeds runs counter to this hypothesis. A much larger sample of Bighorn sheep and investigation in related species such as goat would be required to adequately address the natural history of this variation. Alleles encoding similar amino acid substitutions in the carboxyl terminus of the PROP1 protein (and therefore the activation domain) have also been reported for humans (Duquesnoy et al., 1998) and cattle (Showalter et al., 2002). Such differences could account for phenotypic variation in “normal” populations. Indeed, the two bovine Prop1 alleles encode proteins with different DNA binding and gene activation capacities (Showalter et al., 2002). We predict that the observed isoforms of the oPROP1 protein may similarly display differences in their transcriptional activities.

### 3.3. Two basic sequences are required for nuclear localization of the PROP1 transcription factor protein

The oPROP1 protein contains two perfectly conserved basic regions (B1 and B2, encompassing amino acids 69–73 and 120–126, respectively; Fig. 2) that are potential NLSs. Prediction programs such as PredictNLS (Cokol et al., 2000) suggest B2 as a likely NLS within the PROP1 protein. However, experimental tests of the sequences required for the nuclear localization of other homeodomain proteins have demonstrated that some require NLSs located only in the B1 or B2 equivalent regions, whereas others utilize combinatorial NLS functions requiring at least the B1 and B2 equivalent sequences (e.g. Parker et al., 2000; Furukawa et al., 2002 and references therein).

To examine the mechanisms by which PROP1 proteins achieve nuclear localization, we first created expression vectors encoding fusion proteins containing the enhanced green fluorescent protein (EGFP) and PROP1. These plasmids were transfected into heterologous 293T and pituitary GHFT1 cultured cells, and the intracellular location of the proteins was monitored using confocal or conventional fluorescence microscopy. Bovine and ovine PROP1-EGFP fusion proteins were restricted to the nuclei of all tested cell types (Fig. 3). Fusion proteins with the EGFP moiety placed in the cytoplasm (Fig. 3I). Combined mutation of both the activation domain and B1 regions resulted in an equivalent detection of the fusion protein with some cytoplasmic staining apparent (Fig. 3G). Mutation of the B2 region caused a similar distribution pattern with slightly more protein found in the cytoplasm (Fig. 3I). Combined mutation of both the B1 and B2 regions resulted in an equivalent detection of the protein in both the cytoplasmic and nuclear compartments, indicating loss of nuclear localization (Fig. 3K). We conclude that both B1 and B2 regions contribute to nuclear localization of PROP1. Others have studied the mechanisms underlying nuclear localization of different paired class homeodomain proteins. For example, the regions of the CART1 homeodomain protein that are equivalent to the B1 and B2 sequences of PROP1 are similarly required in

### Table 1

Analysis of the Prop1$^{181T}$ and Prop1$^{181A}$ alleles of oProp1 in populations of production sheep breeds (Ovis aries) and other species (Bighorn sheep, Ovis canadensis)

<table>
<thead>
<tr>
<th>Breed</th>
<th>Prop1 genotype</th>
<th>AA</th>
<th>AG</th>
<th>GG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bighorn sheep</td>
<td></td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Composite</td>
<td></td>
<td>0</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Dorper</td>
<td></td>
<td>2</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Dorset</td>
<td></td>
<td>4</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Finn</td>
<td></td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Katahdin</td>
<td></td>
<td>0</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Rambouillet</td>
<td></td>
<td>5</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Romanov</td>
<td></td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Suffolk</td>
<td></td>
<td>1</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Texel</td>
<td></td>
<td>0</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Total in production sheep</td>
<td></td>
<td>15</td>
<td>42</td>
<td>33</td>
</tr>
<tr>
<td>Overall frequency</td>
<td></td>
<td>17%</td>
<td>47%</td>
<td>37%</td>
</tr>
</tbody>
</table>

Using a primer extension/mass spectrometry assay, the genotypes of representative individuals were determined. AA indicates animals homozygous for alleles encoding a threonine at amino acid 181; GG indicates animals homozygous for the alanine-encoding allele; AG indicates heterozygous individuals. Numbers indicate the number of animals observed with the listed genotype. Composite animals are crosses of multiple breeds.
concert for nuclear localization (Furukawa et al., 2002). Some transcription factors exhibit regulated movement between cellular compartments. In our experiments, PROP1 was nuclear but it is possible that, under specific conditions, the PROP1 protein is exported from the nucleus. Nuclear export signals vary in nature but they can overlap with NLS motifs or can be separate, leucine-rich sequences. Several leucine-rich regions are found in the carboxyl terminus of the PROP1 protein (Fig. 2) but the role of these sequences in cellular localization is not yet known.

Interestingly, missense mutations in the human PROP1 gene altering amino acid sequences in the B1 and B2 regions of the homeodomain (e.g. R71C, R71H, R73C, R73H, R120C, R120H) have been reported to be associated with varying types of combined pituitary hormone deficiency (CPHD) diseases (reviewed in Cushman et al., 2002; Mody et al., 2002). Molecular analyses of some of these mutant proteins have demonstrated that they have impaired DNA binding and gene activation functions (reviewed in Cushman et al., 2002). Our data suggest these mutations in the human PROP1 regions may also disable protein function by affecting nuclear localization. By contrast, a disease-associated mutation altering the protein in the middle of the homeodomain (R99Q) affects the DNA binding and trans-activation properties of the protein but not nuclear localization in nonpituitary cells (Vieira et al., 2003).

3.4. The B1 and B2 basic sequences are critical for PROP1 DNA binding and gene activation

The importance of the B1 and B2 regions to DNA binding and gene activation by PROP1 was determined. Expression vectors encoding wild type and mutant forms of the oPROP1 protein were constructed. The DNA binding capacities of the wild type and mutant PROP1 proteins were assessed by EMSA analysis. Proteins of the expected sizes were generated by in vitro transcription/translation and confirmed by denaturing acrylamide gel electrophoresis (Fig. 4A). Equivalent amounts of PROP1 proteins then were tested for their abilities to bind to the PRDQ9 DNA element. Ovine PROP1 and human PROP1 (included as a positive control) bound strongly to the PRDQ9 DNA probe (Fig. 4B). By contrast, mutation of the B1 or B2 region

Fig. 3. The B1 and B2 basic sequences are required for nuclear localization of PROP1. (Panels A–N) Cultured cells were transiently transfected with expression vectors for the indicated EGFP-containing proteins. Cells were fixed 24–48 h later, the nuclei were stained with Hoechst 33258, and the cells were observed by krypton–argon laser scanning confocal microscopy or by conventional fluorescence microscopy. EGFP is localized to both cytoplasmic and nuclear compartments of human 293T cells in control transfections (panels A, C). In panels C–N, images on the left show EGFP fluorescence; right panels show Hoechst 33258 staining. tPROP1-EGFP and oPROP1-EGFP are restricted to the nuclei of transfected human embryonic kidney 293T cells (panels B, E) or mouse GHFT1 pituitary cells (panel M). Mutation of the B1 or B2 region of oPROP1 results in some cytoplasmic location (panels E–H). Mutation of both basic regions (B1 + B2) results in equivalent detection of oPROP1-EGFP protein in the nuclear and cytoplasmic compartments (panels K, L). (Panels O, P) 293T cells were transiently transfected with myc epitope-tagged oPROP1 expression vectors, and fixed cells were stained with an anti-myc antibody and visualized with a fluorescent secondary antibody (panel O). Nuclei were counterstained with DAPI (panel P). PROP1 is localized to the nuclei of transfected cells.
abolished detectable DNA binding (Fig. 4B). Because the B1 and B2 sequences are located within crucial regions of the homeodomain, this result might seem to be an expected outcome. However, similar experiments examining the CART1 paired class homeodomain protein revealed that mutation of the “B2” region of CART1 did not prevent DNA interaction, but resulted in binding by protein monomers rather than dimers (Furukawa et al., 2002). This difference may be a result of inherent amino acid differences between the CART1 and PROP1 B2 regions (the CART1 B2 sequence is longer and differs at two positions), or may be due to the nature of the introduced mutations (glycine residues were substituted in the CART1 study (Furukawa et al., 2002); alanines in these experiments with PROP1). Further, although we have shown that the B1 and B2 basic regions are important for DNA binding affinity to the PRDQ₉ class of site, the role of these basic sequences in the specificity of PROP1 DNA binding is not yet known.

The trans-activation capacities of wild type and mutant oPROP1 proteins were assessed by transfection assays. Human 293T cells were transiently cotransfected with oPROP1 expression vectors and a luciferase reporter gene containing three copies of the PRDQ₉ element. In these experiments, oPROP1 activated transcription of the reporter gene by 50- to 100-fold (Fig. 5). Human PROP1 also strongly induced transcription of the reporter (Fig. 5). Mutation of either B1 or B2, or both sequences, inacti-

Fig. 4. The B1 and B2 sequences are required for DNA binding by PROP1. (A) Radiolabeled ovine (oPROP1) and oPROP1 proteins with mutations in the B1, B2, or B1 + B2 regions were generated by in vitro translation of RNAs in the presence of 35S methionine. Reaction products were separated by SDS gel electrophoresis, and the resulting products were visualized by autoradiography. Control reactions using unprogrammed lysates did not produce labeled protein products. The migration positions of molecular markers are shown (in kDa). (B) Assay of oPROP1 DNA binding by EMSA using an oligonucleotide probe containing a paired class homeodomain binding site (PRDQ₉). Radiolabeled probe was incubated with the indicated in vitro translated proteins, and the resulting complexes were separated from free probe (F) by electrophoresis. Probe alone and probe plus unprogrammed lysate (lysate control) were used as negative controls. Human PROP1 (hPROP1) served as a positive control. Square bracket = bound complex.
vated the transcriptional capacity of oPROP1. The tran-
scriptional inactivity of the mutated oPROP1 proteins
likely reflects reduced or absent nuclear entry, as well as
a lack of DNA binding. We conclude that the B1 and B2
regions are required for DNA binding and gene activation
by PROP1.

3.5. PROP1 is associated with the nuclear matrix

Once proteins have entered the nucleus, they may
partition further to nuclear subcompartments. To test
whether PROP1 exhibits intranuclear association with the
nuclear matrix, we performed in situ nuclear matrix
extraction assays. PROP1-EGFP was transiently trans-
ferred into cultured cells. Forty-eight hours after transfec-
tion, cells were treated by sequential extraction with
detergent, deoxyribonuclease, and high salt to remove
soluble DNA and protein nuclear components. Fluores-
cence was then observed by microscopy. The extraction
protocol treatment removed detectable fluorescence from
control cells transfected with EGFP alone confirming that
nuclear-located EGFP is not associated with the nuclear
matrix (Fig. 6). However, significant levels of nuclear
PROP1-EGFP fluorescence remained following all extrac-
tions (Fig. 6N). Cells were counterstained with DAPI to
detect chromatin, and NuMA, a known nuclear matrix
protein, was co-visualized by indirect immunofluorescence.
Following extraction, PROP1-EGFP was observed in the
nuclei of transfected cells displaying NuMA fluorescence
(Fig. 6). We conclude that PROP1 is associated with the
nuclear matrix. The nuclear matrix has been proposed to
serve both structural and regulatory functions in gene
transcription in specific tissues (e.g. Bidwell et al., 2001;
Alvarez et al., 2003; Stein et al., 2003) and other pituitary-
acting transcription factors such as PIT1 and LHX3 are
known to have nuclear matrix association (Mancini et al.,
1999; Parker et al., 2000). The association of PROP1 with
the nuclear matrix may imply a targeted or architectural
role for this transcription factor in pituitary gene regulation
during development, or the partitioning of PROP1 in and
out of the nuclear matrix may be a component of its post-
translational regulation.

3.6. Conclusions

(1) The ovine Prop1 gene has three exons and two introns—a
different structure from the bovine gene.
(2) Two alleles of the ovine Prop1 gene encode protein
products with different carboxyl terminal domain
sequences.
Mammalian PROP1 proteins have conserved carboxyl termini and DNA binding (homeodomain) domains, but their amino termini are diverged.

Two conserved, basic amino acid sequences (B1 and B2) within the homeodomain are required for DNA binding, gene activation, and exclusive nuclear localization of the PROP1 protein: mutations causing changes in the B1 and B2 regions of human PROP1 in pituitary hormone deficiency diseases may also disable PROP1 function by affecting nuclear localization.

The PROP1 protein is associated with the nuclear matrix.

Acknowledgements

We thank Dr. Ellen Chernoff, Surilda Clark, Linda Flathman, Jesse Savage, Kevin Tennill, and Sean Werner for reagents and assistance. Supported by grants to SJR (USDA/NRRCGP/CSREES; NIH HD42024; NSF 0131702) and JPB (NIH PO1 AR45218; NIH DK53769).

References


