Single-minded 1 and 2 are unique members of the basic helix-loop-helix Per-Arnt-Sim family as they are transcriptional repressors. Here we report the identification and transcriptional characterization of mouse Sim2s, a splice variant of Sim2, which is missing the carboxyl Pro/Ala-rich repressive domain. Sim2s is expressed at high levels in kidney and skeletal muscle; however, the ratio of Sim2 to Sim2s mRNA differs between these tissues. Similar to full-length Sim2, Sim2s interacts with Arnt and to a lesser extent, Arnt2. The effects of Sim2s on transcriptional regulation through hypoxia, dioxin, and central midline response elements are different than that of full-length Sim2. Specifically, Sim2s exerts a less repressive effect on hypoxia-induced gene expression than full-length Sim2, but is just as effective as Sim2 at repressing TCDD-induced gene expression from a dioxin response element. Interestingly, Sim2s bind to and activates expression from a central midline response element-controlled reporter through an Arnt transactivation domain-dependent mechanism. The differences in expression pattern, protein interactions, and transcriptional activities between Sim2 and Sim2s may reflect differential roles each isoform plays during development or in tissue-specific effects on other protein-mediated pathways.

The basic helix-loop-helix Per-Arnt-Sim (bHLH-PAS) proteins comprise a growing family of transcription factors that play key roles during development and in sensing and adapting to changes in the environment. Individual PAS proteins are known to control morphogenesis, circadian rhythmicity, responses to hypoxia and toxin metabolism. These proteins contain a bHLH motif, which mediates dimerization with other bHLH proteins and contributes to determining DNA binding specificity. The PAS domain, named after the founding members of this family (period-aryl hydrocarbon nuclear translocator-single minded), is a multifunctional protein surface responsible for such diverse activities as ligand binding, PAS protein dimerization, and non-PAS protein interactions (1).

In addition to environmental adaptation, some members of the bHLH-PAS family regulate development. In Drosophila, single-minded (sim) acts as the master regulator of central nervous system midline development by controlling expression of many genes required for differentiation. Similar to other bHLH-PAS proteins, sim functions as a heterodimer with arnt (2). This complex binds to central midline elements (CME) in the regulatory regions of target genes to activate expression of proteins required for proper central midline establishment (2, 3).

Two mammalian homologs of sim, Sim1 and Sim2, have been identified (4–6). These proteins share a high degree of similarity in their PAS domains, but little conservation is apparent in their carboxyl termini. Sim1 and Sim2 interact with Arnt, but differ from Drosophila sim, the aryl hydrocarbon receptor (AHF) and hypoxia inducible factor (HIF) by functioning as transcriptional repressors (7, 8). Sim1 and Sim2 are expressed in a variety of tissues including brain, kidney, lung, and skeletal muscle where they play important developmental roles. Human Sim2 was first identified by exon trapping of a region on chromosome 21 known to be associated with Down syndrome (9). In addition, a splice variant of human SIM2, designated SIM2 short (SIM2S), has also been identified (9). This splice variant, which is missing exon 11 and therefore lacks a portion of the region implicated in mediating the repressive effects of SIM2, is reported to be involved in cancer susceptibility (10, 11); however, functional differences between these two isoforms have not been reported.

Because the bHLH-PAS proteins share structural motifs and common binding partners, it is not surprising that cross-talk can occur between PAS-protein-mediated pathways. In the case of the HIF proteins, both Sim1 and Sim2 can compete with HIFs for Arnt binding, and interact with a prototypic hypoxia response element (HRE) to affect gene expression (7, 8, 12). Interestingly, Sim1/Arnt, but not Sim2/Arnt, can induce transcription of an HRE-controlled reporter gene via the COOH-terminal transactivation domain of Arnt (12). In contrast, Arnt-mediated transcription of a CME-controlled reporter gene is severely impaired in the presence of Sim2 and is dependent upon the Sim2 dimerization domain and carboxyl terminus, which contains two separate repressive domains (7, 8). Because repression by Sim2 is not specific for Arnt, as Sim2-Gal4 fusion constructs have repressive effects on a thymidine kinase promoter (9), it is thought that Sim2-mediated repression can also occur through direct interactions with the basal transcription machinery. Similar to Drosophila sim, Sim1 and Sim2 also bind and regulate transcription through a consensus CME, which is not surprising because the CME core sequence (5′-ACGTG-3′) is identical to that of the HRE. As was seen for the HRE, Sim1 strongly activates transcription of a CME-controlled gene through the transactivation domain of Arnt, whereas Sim2 is repressive (7).

We have isolated a splice variant of mouse Sim2 that corresponds to the human Sim2 transcript, and have characterized its expression profile and transcriptional properties. Mouse Sim2s is expressed at high levels in adult kidney and skeletal muscle but in different ratios with respect to full-length Sim2. Like Sim2, Sim2s interacts with Arnt, but is slightly less efficient at binding Arnt2. These differences in expression...
Characterization of the Mouse Sim2s Gene

pattern and binding partner preference may prove significant, as we have also determined that Sim2s and Sim2 differ dramatically in their ability to regulate expression of genes under control of hypoxia, dioxin, and central midline response elements.

MATERIALS AND METHODS

Animals and Cell Culturing—Female C57Bl/6j mice were housed under standard 12-h lights on/off conditions in a temperature and humidity controlled facility with food and water provided ad libitum. All animal housing and treatments were approved by and conformed to the Animal Use Protocols at Texas A&M University. All cell lines were maintained in a 37 °C humidified incubator in a mixture of 95% air and 5% CO2. HEK-293T and HepG2 cells were cultured in Dulbecco's modified Eagle's medium (Invitrogen) containing 10% fetal bovine serum and 1% penicillin-streptomycin.

RNA Isolation and Reverse Transcription—Total RNA was initially isolated from mouse tissues using TRIzol reagent (Invitrogen) and was further purified using RNeasy kits (Qiagen) with on-column DNase digestion (Qiagen). One microgram of total RNA was reverse transcribed using oligo(dT) and Superscript II reverse transcriptase (Invitrogen).

3′ Rapid Amplification of cDNA Ends (RACE)—Total RNA (2 μg) was reverse transcribed using an oligo(dT)-based adapter primer (AP-dT<sub>19</sub>) (5′-GATCAGGACGTGTTAGTTTTTTTTTTTTTTTTT-3′). An initial round of PCR was performed using a 1:100 dilution of the RT reactions with the adapter primer (AP) (5′-GATCAGGACGTGTTAGTTTTTTTTTTTTTTTTT-3′) and a Sim2-specific primer located in exon 9 (P1, 5′-AGTCGAGGAGGATCTGGA-3′). Subsequent rounds of PCR were performed on 1 μl of a 1:100 dilution of the previous reaction using the adapter primer and nested Sim2 primers (P2, 5′-AAATCAGCCTAACCCAAAAACACAAA-3′; P3, 5′-AGAAGCGAACCCATATCCC-3′; P4, 5′-CTTCTCTCTGTGTAAGTC-3′). The PCR products were cloned into pCR II-TOPO (Invitrogen), and screened for inserts by restriction enzyme analyses. Clones positive for insert sequence were sequenced (DNA Technologies Lab, Texas A&M University) and compared with reported mouse Sim2 and chromosome 16 sequences. The sequence of the Sim2s 3′-RACE product was submitted to GenBank™ and was assigned accession numberAY963781.

RT-PCR—RT-PCR contained 1 μl of cDNA template with 1 μM each primer, 0.4 mM dNTPs, and 1 unit of Taq DNA polymerase (Promega, Madison, WI) per 25-μl reaction. Primers used were: Sim1 (FP, 5′-GGTCAGCAGCTGATCTTG-3′; RP, 5′-TGCTCTCGTGTGTTGTG-3′), Sim2s (FP, 5′-AAATCAGCCTAACCCAAAAACACAAA-3′; RP, 5′-TGCTCTCGTGTGTTGTG-3′), Sim2 (FP, 5′-AGAGCGGTGCACCTACA-3′), Sim2s (FP, 5′-AAACAGCTCCCGTGTTGGAC-3′; RP, 5′-ACTCTGAGGAAGCGGCACAAA-3′), and Sim2 (FP, 5′-TCTGAGCCTTCTCCTGTAATG-3′; RP, 5′-CGAGGGGCGCTTCATAGTT-3′), β-actin (FP, 5′-GCAAGAGGCGGGTCC-3′; RP, 5′-GCCAAGAGGCGGGTCC-3′).

Plasmid Construction—The full-length mouse Sim2s plasmid (pcDNA-mSim2s) was made by amplifying the 5′ most 1548-bp of Sim2 from pmSim2-GAL4-HA (kindly provided by Dr. Jerry Pelletier, McGill University, Montreal, Quebec, Canada) using HiFi Taq DNA polymerase (Roche Diagnostics) and the primers 5′-ATGAGAGGAGTAGCTCAAATAATG-3′ and 5′-AGACATTCCAAGGAGGCTCAGAA-3′. This fragment was cloned into pCR II-TOPO, and the 5′ Spel fragment was cut out and ligated to SpeI/XbaI cut pCRmSim2s-8 (original Sim2s 3′ RACE clone) to create pCRFlmSim2s. After the sequence was confirmed, the entire insert was removed by EcoRI digestion, and cloned into EcoRI-cut pcDNA3. pmSim1-HA and pmSim2-HA were a kind gift of Dr. Jerry Pelletier.

All yeast two-hybrid plasmids utilized pGBK-T7 and pGADT7-Rec as the DNA-binding and activation plasmids, respectively (Clontech, Carlsbad, CA). For mouse Sim2s yeast two-hybrid plasmids, the full-length insert of PCRFlmSim2s was removed by EcoRI digestion and ligated to EcoRI-cut vectors. Full-length Sim2L was removed from pmSim2-GAL4-HA by EcoRI digestion and cloned into EcoRI-cut vectors. Full-length Arnt and Arnt2 were amplified from mouse kidney and cloned into pCR II-TOPO. Once the sequences were verified, the inserts were removed and cloned into pGBK-T7 and pGADT7-Rec.

Reporter Plasmids—pDRE-TATA-luc, pGL2-TATA, and pβ-galactosidase were kindly given to us by Dr. Stephen Safe (Texas A&M University) (13). pHRE-TATA-luc was constructed by cloning annealed oligonucleotides corresponding to the top and bottom strands of the hypoxia-responsive region of the EPO promoter into pGL2-TATA (14). pCEM-luc was a kind gift of Dr. Jerry Pelletier.

Yeast Two-hybrid Liquid Culture Assay—One hundred ml of YPD was inoculated with 10 ml of an overnight AH109 yeast culture (Clontech) and grown to an A<sub>600</sub> of ~0.6. Cells were spun down and washed two times in 50 ml of sterile water and once in 2 ml of 1 M cold sorbitol. Cell pellets were resuspended in 2 ml of 100 mM LiAc/TE containing 25 mM dithiothreitol and incubated at room temperature for 1 h. Following an additional 1 M sorbitol wash, cell pellets were resuspended in 1 mM sorbitol and mixed with plasmids on ice (50 μl of cells plus 250 ng of each plasmid). Cells were transfected by electroporation with a 5-ms pulse at 1.5 kV, 50 microfarads, and 100 ohm. One ml of YPD was added to transformed yeast, which were then incubated at 30 °C for 1 h. Cells were spun, and resuspended in 100 μl of 1 M sorbitol and plated onto DOB/-Leu/-Trp plates and incubated at 30 °C until colonies appeared. Double dropout broth liquid cultures of three colonies from each plate were grown overnight at 30 °C with shaking. Fresh media was inoculated with 1 ml of overnight culture and incubated at 30 °C with shaking until cells reached log phase (A<sub>600</sub> of the cultures between 0.5 and 0.8). At this point, the A<sub>600</sub> of each culture was recorded and 1.5 ml of yeast were spun and washed in Z buffer (10 mM NaH<sub>2</sub>P<sub>4</sub>O<sub>7</sub>, 5 mM Na<sub>2</sub>HPO<sub>4</sub>, 10 mM KCl, 1 mM MgSO<sub>4</sub> pH 7.0). Cells resuspended in 100 μl of Z buffer were subjected to three cycles of freezing/thawing for 60 s each in liquid nitrogen and a 37 °C water bath. Seven hundred μl of Z buffer containing 4 mM β-mercaptoethanol and 160 μl of Z buffer containing 4 mg/ml o-nitrophenyl β-D-galactopyranoside was added to each tube and the reactions were placed at 30 °C for 60 min. Reactions were stopped by addition of 400 μl of 1 M Na<sub>2</sub>CO<sub>3</sub> and the time of reaction recorded. The OD<sub>420</sub> of each reaction was recorded and β-galactosidase activity calculated using the equation β-galactosidase units = 1000 × OD<sub>420</sub>/t × V × A<sub>600</sub>, where t = elapsed time (in min) of incubation, V = 0.1 ml × concentration factor, and A<sub>600</sub> = A<sub>600</sub> of 1 ml of culture. Data are expressed as the average of three separate colonies ± S.E.
Co-immunoprecipitation—Full-length Arnt, Arnt2, Sim2, and Sim2s were made from pcDNA3-based expression plasmids using an in vitro translation kit (Promega). Sim2 and Sim2s were made in the presence of [35S]methionine (Amersham Biosciences). Five μl of Arnt or Arnt2 product were mixed with 15 μl of radiolabeled Sim2 or Sim2s and allowed to incubate at room temperature for 2 h. Five μg of anti-Arnt (Upstate), anti-Arnt2 (Santa Cruz Biotechnology, Inc.), or rabbit IgG (Upstate) were added, and the volumes were increased to 150 μl with water and 2× co-immunoprecipitation buffer to make a 1× solution (10 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100 and 0.5% IGEPEAL). Following a 2-h incubation at room temperature, 20 μl of agarose A bead slurry was added and samples were incubated at room temperature for 1.5 h with gently mixing. Beads were recovered at 5000 rpm for 1 min, and washed 3 times in 500 μl of 1× co-immunoprecipitation buffer. Final pellets were resuspended in loading buffer, boiled, and separated on polyacrylamide gels. Gels were vacuum-dried and exposed to film for 1 week.

Transient Transfection—Sim2 and Sim2s expression was confirmed by Western analyses before large-scale experiments were conducted. For each transfection, cells were seeded at 4 × 10^4 cells per well in 24-well plates the day before transfection. The following morning, 200 ng of the appropriate reporter plasmid was co-transfected with 100 ng of internal control (pβ-galactosidase) and various amounts of test plasmids using Lipofectamine and Plus reagent (Invitrogen). Twenty-four hours later, cells were incubated under hypoxic (1% O2) or normoxic (21% O2) conditions (pHRE-TATA-luc transfected HEK-293T cells) or in the presence of vehicle (MeSO) or 10 mM 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) (pDRE-TATA-luc transfected HepG2 cells) for 40 h. Cells were harvested, and the luciferase and β-galactosidase activities of the cell lysates were determined by dual luciferase assay using Luciferin (Molecular Probes) and Galacto-Light (Tropix, Applied Biosystems). Luciferase activities were normalized to the internal control values and are represented as the mean ± S.E. for three wells per condition. Significant differences were determined using Student’s t test.

Chromatin Immunoprecipitation Assay—Twenty-four h after transfection, formalin (270 μl/10 ml of medium) was added to HEK-293 cells and allowed to incubate at 37 °C for 10 min. Cross-linking was stopped by addition of glycerol to a final concentration of 125 mM. Following a series of phosphate-buffered saline washes, cells were scraped and recovered by centrifugation. Cells (500,000 cells per ChIP assay) were resuspended in SDS lysis buffer (50 mM Tris-HCl, pH 8.1, 1% SDS, 10 mM EDTA) and sonicated on ice. Chromatin was recovered, and mixed with ChIP dilution buffer (16.7 mM Tris-HCl, pH 8.1, 0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 167 mM NaCl) containing 1× Complete Protease Inhibitor Mixture (Roche). Chromatin was cleared twice by incubation with agarose bead slurry at 4 °C for 30 min, followed by centrifugation. Antibody was added (10 μg for Sim2, 1 μg for IgG) and incubated at 4 °C overnight with mixing. Agarose bead slurry was added and reactions were incubated at 4 °C for 1 h with mixing. The agarose beads were recovered and washed successively for 5 min at 4 °C in low salt (20 mM Tris-HCl, pH 8.1, 150 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA), high salt (20 mM Tris-HCl, pH 8.1, 500 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA), LiCl (10 mM Tris-HCl, pH 8.1, 250 mM LiCl, 1% IGEPEAL-Ca630, 1% deoxycholic acid, 1 mM EDTA), and TE (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) buffers. Chromatin was eluted at room temperature by incubation in freshly prepared elution buffer (1% SDS, 0.1 mM NaHCO3) for 15 min. Samples were then treated with RNase A in the presence of 200 mM NaCl and cross-linking was reversed by incubation at 65 °C for 4 h. Chromatin was ethanol precipitated and resuspended in buffer containing 40 mM Tris-HCl, pH 8.1, 10 mM EDTA, and 200 μg/ml protease K. DNA was purified using Qiagen PCR purification columns (Qiagen) and eluted in 50 μl of sterile water. PCR detection of pGL2-CME immunoprecipitation targets was performed in 25-μl reactions containing 2 μl of chromatin and the primers pGL2-ChIP-F1 (5′-CCCATGAGACCT-

FIGURE 1. Isolation of mouse Sim2s by 3′ rapid amplification of cDNA ends (3′ RACE). A, structure of mouse Sim2 and Sim2s cDNAs. Numbered squares indicate exons. Numbered arrows indicate relative positions of mouse Sim2-specific primers used in 3′ RACE PCR. Dotted lines indicate relative positions of restriction sites used to verify the identity of 3′ RACE products. B, restriction enzyme analyses of mouse kidney and liver 3′ RACE products. Two micrograms of total RNA from mouse kidney or liver was reverse transcribed using an oligo(dT)-based adapter primer (AP-dT17). An initial round of PCR was performed on 1 μl of a 1:100 dilution of the RT reaction using the adapter primer and a mouse Sim2-specific forward primer (P1). The products of this initial PCR were not visible (data not shown). Subsequent PCR were performed on 1 μl of a 1:100 dilution of the previous reaction using the adapter primer and nested Sim2-specific primers (P2, P3, and P4 in Fig. 2A). For visualization and confirmation of identity, 10 μl of PCR product plus (+) or minus (−) restriction enzyme was analyzed on a 1.2% agarose gel. Mouse Sim2s 3′ RACE products are predicted to have SpeI (P2/AP), BamHI (P3/AP), or HaeIII (P4/AP) sites. Sim2-specific products were only detected in kidney (Kid) RNA samples, and only Sim2 was detected. Liv, liver.
GAAACATA-3') and pGL2-ChIP-R1 (5'-GCCTTATGCAGTT-GCTTCC-3'). Thermocycling was done with an initial denaturation at 95 °C for 5 min followed by 27 cycles of 95 °C for 30 s, 50 °C for 30 s, and 72 °C for 30 s followed by a final elongation step at 72 °C for 5 min.

RESULTS

Previous studies have identified a splice variant of human SIM2; however, a similar isoform has not been reported in mice, nor have functional differences between Sim2 and Sim2s been reported. To better characterize biochemical and physiological differences

FIGURE 2. Structure and comparison of mouse and human Sim2s coding sequences. A, genomic structure of the mouse Sim2 locus surrounding exons 9–11. Exons are indicated by numbered boxes and the s exon is represented by a shaded box A, comparison of Sim2 with Sim2 and mouse chromosome 16. The 3' boundary of exon 10 is indicated by the arrow and vertical line. Stop codons are bold and underlined. The mouse Sim2 sequence used (accession number U42554) is that of Moffett et al. (5). C, comparison of mouse and human Sim2s coding and predicted amino acid sequences. The 3' boundary of exon 10 is indicated by the arrow and vertical line. Stop codons are bold and underlined and represented as asterisks in the amino acid sequences. Regions of amino acid identity are boxed.
between Sim2 and Sim2s in mice, we set out to identify the mouse Sim2s transcript. Total RNA from mouse kidney and liver tissues were used for 3′ RACE experiments because Sim2 is expressed at high levels in kidney, but not in liver. A unique adapter primer (AP-dT17) was used to reverse transcribe total RNA from the tissues. This reaction was then used as a template for PCR using a primer directed to the AP region of the 3′/H11032 adapter and a Sim2-specific forward primer (P1, Fig. 1A). Subsequent PCR were performed on diluted PCR products using the AP and nested Sim2-specific reverse primers (P2–P4, Fig. 1A). The resulting PCR products from each reaction were submitted to restriction enzyme analyses for predicted sites to support the identity of the PCR products as Sim2 (Fig. 1B). Potential Sim2 and Sim2s expressed in these tissues differed with kidney expressing more Sim2s than Sim2 and vice versa in skeletal muscle. C–E, quantitative real time PCR analyses of mouse kidney, liver, and skeletal muscle mRNA for total Sim2 (C), Sim2s (D), and full-length Sim2 (E).

Sim expression was analyzed in mouse tissues by RT-PCR and real time RT-PCR (Fig. 3). Sim1 was expressed at high levels in kidney, brain, lung, and skeletal muscle (Fig. 3A). In contrast, Sim2 and Sim2s were expressed at high levels in kidney and skeletal muscle. The ratio of Sim2s and Sim2 mRNA expressed in kidney and skeletal muscle appeared to
Characterization of the Mouse Sim2s Gene

TABLE 1

|          | pGADT7 | Arnt | Arnt2 | Sim2 | Sim2s |
|----------|--------|------|-------|------|-------|
| pGBK?    | 0.20 ± 0.09 | ND   | ND    | ND   | ND    |
| Arnt     | 0.56 ± 0.38 | 18.31 ± 0.44 | 9.648 ± 0.42 | 8.14 ± 1.61 | 10.94 ± 0.13 |
| Arnt2    | 2.59 ± 0.73 | 0.77 ± 0.01 | 2.33 ± 0.26 | 2.02 ± 0.33 | 3.22 ± 0.86 |
| Sim2     | 0.89 ± 0.09 | 2.09 ± 0.07 | 0.85 ± 0.16 | 1.07 ± 0.06 | 1.43 ± 0.42 |
| Sim2s    | 0.48 ± 0.21 | 1.55 ± 0.16 | 0.77 ± 0.07 | 0.91 ± 0.09 | 0.89 ± 0.2 |

ND, not determined.

FIGURE 4. Sim2s preferentially interacts with Arnt over Arnt2. In vitro translated Arnt and Arnt2 were mixed with radiolabeled Sim2 or Sim2s and co-immunoprecipitated using antibodies directed at either Arnt or Arnt2. Anti-Arnt antibody was highly efficient at pulling down Sim2s (lane 1) and Sim2 (lane 2) in the presence of in vitro translated Arnt. Similarly, anti-Arnt2 was able to pull down Sim2s (lane 3) and Sim2 (lane 4) in the presence of in vitro translated Arnt2, although Sim2s was not pulled down as well as Sim2. IgG pulled down Arnt to a greater extent than Arnt2 (lane 5), but was unable to pull down either Sim2 or Sim2s (lane 6). A diluted mixture of radiolabeled Arnt and Arnt2 (lane 7) as well as Sim2 and Sim2s (lane 8) were run on the gel to verify translation efficiency.

differ, suggesting that their expression is controlled in a tissue-specific manner. Therefore, expression of both isoforms of Sim2 were determined in kidney, liver, and skeletal muscle by RT-PCR using a common forward primer located in exon 9, and both Sim2- and Sim2s-specific reverse primers located in exons 11 and s, respectively. As expected, neither Sim2 isoform was detected in liver, but kidney and skeletal muscle expressed both Sim2 and Sim2s in different relative amounts (Fig. 3B). In agreement with our initial analyses (3A), kidney expressed slightly more Sim2s than Sim2, whereas skeletal muscle expressed significantly more Sim2 than Sim2s. Quantitative real time RT-PCR analyses of kidney, liver, and skeletal muscle RNA for total Sim2 (Fig. 3C), Sim2s (Fig. 3D), and full-length Sim2 (Fig. 3E) corroborated our PCR analyses and found that, overall, skeletal muscle contains significantly higher levels of full-length Sim2 mRNA than kidney.

The utility of the yeast two-hybrid system in addressing functional interactions between Sim genes and other PAS proteins has been well established (6). To begin assessing if Sim2 and Sim2s differ in their interactions between Sim genes and other PAS proteins has been well established (6). To begin assessing if Sim2 and Sim2s differ in their interactions with other PAS proteins, we took advantage of the two-hybrid liquid culture assay. Chimeras containing the GAL4 DNA binding domain (pGAD-based plasmids) and the DBD of Arnt, Arnt2, Sim2, or Sim2s were transfected into AH109 yeast and plated onto selective medium. Three colonies per plate were grown in liquid medium and analyzed for ß-galactosidase activity as described under “Materials and Methods.” Values presented are the mean ± S.E.

Direct targets of mammalian Sim proteins have not been determined; however, indirect transcriptional activities of Sim1 and Sim2 have been reported. In those studies, it was found that Sim2 can inhibit the actions of other bHLH-PAS proteins through active repression or competition for Arnt binding (7, 8, 12). Overexpression of Sim2 blocked hypoxia-induced expression from HRE-controlled genes by directly binding to the HRE and actively repressing expression of adjacent genes. This effect was dependent upon the Sim2 dimerization domain and carboxyl terminus. Sim2 can also inhibit TCDD-induced expression from XRE-controlled genes, but does so by competing with AHR for Arnt binding. In contrast, Sim2 represses expression of a CME-driven reporter, whereas Sim1 can activate CME-mediated gene expression via the Arnt transactivation domain (7). To investigate if Sim2s can affect the function of other bHLH-PAS protein-mediated pathways, we performed co-transfection assays using various mouse Sim and Arnt expression plasmids with HRE-, XRE-, and CME-controlled luciferase reporter constructs.

Expression of Sim2 and Sim2s in transfected HEK293T cells was confirmed by Western blotting (Fig. 5A). HEK293T cells transfected with an HRE-controlled luciferase reporter and the pcDNA empty vector showed a significant increase in luciferase activity following a 40-h incubation under hypoxic conditions (Fig. 5B). Introduction of increasing amounts of mouse Sim2 or Sim2s significantly repressed hypoxia-induced reporter gene expression. Inclusion of 100 ng of mouse Arnt expression vector increased the response to hypoxia and slightly overcame both Sim2 and Sim2s-mediated repression. Addition of mouse Arnt2 expression vector did not alter the response to hypoxia and both Sim2 and Sim2s repressed hypoxia-induced expression of the HRE-controlled reporter in the presence of Arnt2. Increasing amounts of Sim or Sim2s did not repress in a concentration-dependent manner suggesting that both Sim2 and Sim2s repress hypoxia-induced gene expression through direct interaction with the HRE and not by competing with HIF factors for Arnt binding.

Previous studies have suggested that Sim2 can interfere with AHR-mediated gene expression through competition with AHR for Arnt binding (8). To determine whether Sim2s exerts a similar effect on AHR-mediated signaling, HepG2 cells were co-transfected with a DRE-controlled reporter gene with various combinations of Sim2 or Sim2s and Arnt expression plasmids. AHR activation was accomplished by addition of 10 nM TCDD to growth medium 24 h before harvest. A robust TCDD response seen in pcDNA-transfected cells was significantly repressed (p < 0.0001) by inclusion of either Sim2 or Sim2s expression plasmids (Fig. 5C). The degree of this repression was not significantly different between Sim2 and Sim2s; however, increasing amounts of Arnt attenuated the repression (p < 0.01) in both Sim2 and Sim2s-transfected cells indicating that the mechanism of
this repression involves competition for Arnt. These results are consistent
with previous studies (8) and suggest that Sim2 and Sim2s do not differ in
their ability to repress AHR-mediated gene expression, which occurs
through direct competition for Arnt binding.

In Drosophila, sim regulates transcription through the CME, which
contains the core sequence 5’-ACGTG-3’ also found in an HRE. Previous
studies have shown that murine Sim1 and Sim2, in concert with
Arnt, can bind and regulate expression of a CME reporter construct (7).
Sim1 activated expression of a CME-controlled gene through the Arnt
transactivation domain, whereas full-length Sim2 was repressive. Activi-
tion of a CME-controlled gene by Sim2 was accomplished when por-
tions of the C-terminal repression domain were deleted. Because Sim2s
is missing part of this repressive region, we anticipated that Sim2s would
be less repressive than Sim2 on CME-mediated gene expression.

Cotransfection of HEK293 cells with a CME-controlled reporter gene
and increasing amounts of Arnt or Arnt2 had no effect on reporter
expression (Fig. 6A). Contrary to previous reports, we found that intro-
duction of Sim2 with Arnt, but not Arnt2, resulted in a slight increase in
luciferase activity (Fig. 6A). Co-expression of Sim2s with Arnt, but not
Arnt2, resulted in significantly increased reporter gene expression. This
effect appears to be Arnt-dependent as luciferase activity increased with
increasing amounts of Arnt. This conclusion was further supported by
experiments utilizing constant Arnt and increasing amounts of Sim2
expression plasmids (Fig. 6B). Although reporter gene expression
increased with increasing amounts of Sim2 and Sim2s expression plas-
mids, this effect was only significant in cells receiving the highest
amount of Sim2 and Sim2s expression plasmid. Confirmation of Arnt-
dependent Sim2s transcriptional activation from a CME-controlled
FIGURE 5. Transcriptional activity of mouse Sim2 and Sim2s on HRE- and DRE-controlled reporter genes.
A, confirmation of Sim2 expression in HEK293T cells by Western blot.
HEK293T cells transfected with pcDNA (Con) or increasing (3, 9, and 15 μg of plasmid in 100-mm plate format) amounts of Sim2 or Sim2s expression plasmids were harvested and
analyzed by Western blot to confirm Sim2 gene expression. B, the effects of increasing amounts of ARNT and Arnt2 on mouse Sim2- and Sim2s-mediated repression of an
HRE-controlled reporter gene. HEK-293T cells were transfected with Arnt or Arnt2 expression plasmids in the presence of increasing amounts of Sim2 or Sim2s expression plasmids
plus an HRE-controlled luciferase reporter vector and a β-galactosidase expression plasmid for normalization. The amount of each expression vector (in nanograms) is indicated
under the figure. Following 40 h incubation under normoxic (white bars) or hypoxic (black bars) conditions, cell lysates were collected and analyzed for β-galactosidase and luciferase
activities. Data are expressed as the mean ± S.E. for three plates per condition. Asterisk, luciferase expression under hypoxic conditions is significantly higher than control (p < 0.002).
Double asterisk, luciferase expression under hypoxic conditions is significantly lower in Sim- and Sim2s-transfected compared with pcDNA-transfected cells (p < 0.002). Triple asterisk,
luciferase expression in the presence of ARNT during hypoxia is significantly less in Sim2-transfected cells than in Sim2s-transfected cells (p < 0.01). C, effects of increasing Sim2 and
Sim2s on DRE-controlled gene expression. HepG2 cells were transfected with varying amounts of either Arnt or Arnt2 expression plasmids in the presence of increasing amounts of Sim2 or Sim2s expression plasmids plus a DRE-controlled luciferase reporter vector and a β-galactosidase expression plasmid for normalization. The amount of each expression vector
(in nanograms) is indicated under the figure. Following a 40-h incubation with vehicle (Me2SO, white bars) or 10 ng TCDD (black bars) cell lysates were collected and analyzed for
β-galactosidase and luciferase activities. Data are expressed as the mean ± S.E. for three plates per condition. Asterisk, significantly higher due to TCDD treatment (p < 0.0001). Double
asterisk, TCDD-induced luciferase expression in the presence of Arnt is significantly less in Sim2-transfected cells than in Sim2s-transfected cells (p < 0.01).
Characterization of the Mouse Sim2s Gene

A gene is presented in Fig. 6C. In these experiments, contrasfection of a mutant Arnt (ArntΔTAD), which is missing the transactivation domain, abolished the ability of Sim1, Sim2, and Sim2s to increase expression from a CME. Expression of ArntΔTAD repressed basal CME-mediated gene expression in the absence of external Sim proteins (Fig. 6C). Co-expression of Sim1 and Arnt resulted in significant reporter gene expression that was abolished when Arnt was replaced with the ArntΔTAD expression vector. A similar effect was seen with Sim2, but the degree of gene activation was significantly lower than was seen with Sim1. Surprisingly, Sim2s was almost as potent as Sim1 in activating CME-controlled gene expression (Fig. 6C). As was seen with Sim1, this effect was completely abolished when ArntΔTAD was substituted for Arnt, implying that this effect is entirely mediated by the activation domain of Arnt.
In contrast, Sim2s can activate expression from a CME apparently by repressing gene expression to an equal extent suggesting that only the observation that Sim2s, which is missing the Pro/Ala-rich region, is less potent than full-length Sim2 at repressing Hif1 expression through a DRE (Fig. 5).

To confirm the interactions between Sim2 and Sim2s on a CME, ChiP were performed on CME-luc-transfected control and Sim2/Arnt or Sim2s/Arnt cells. Chromatin was immunoprecipitated with an anti-Sim2 antibody that recognizes both Sim2 and Sim2s, and was analyzed for CME binding using a set of PCR primers specific for the CME reporter plasmid. The presence of Sim2 on the CME was detectable in control cells, most likely due to endogenous Sim2 (data not shown). More importantly, the presence of both Sim2 and Sim2s on the CME was elevated in Sim2-transfected cells (Fig. 7). These data suggest that the differential outcomes of Sim2 isoform binding to a CME are not because of changes in DNA binding.

**DISCUSSION**

Transcriptional regulation occurs through multiple distinct mechanisms involving negative as well as positive interactions between regulatory factors. The mammalian Sim proteins are unique members of the bHLH-PAS family because they can exert negative effects on transcription. It has been determined that the repressive effects of Sim2 are mediated by two domains in its carboxyl terminus. One of these domains is rich in Pro and Ser residues, whereas the other is Pro and Ala rich. These repressive domains appear to be nonspecific as Sim2 can suppress activation of a Gal4 activation domain on a thymidine kinase promoter (6, 7) as well as Arnt-mediated transactivation. Similar hydrophobic domains are present in other transcriptional repressors including the Drosophila Kruppel transcription factors and Even-skipped, which inhibit transactivation by competing with TBP for TATA box binding thus preventing assembly of the preinitiation complex (16).

In this paper, we have shown that mouse Sim2s, a splice variant of Sim2, has differential effects on CME- and HRE-mediated gene expression. Sim2s is less potent than full-length Sim2 at repressing Hif1α (Fig. 5B), and can activate expression of a gene controlled by the Drosophila toll gene CME via the transactivation domain of Arnt (Fig. 6).

The hypo-suppressive effects of Sim2s are not surprising given that Sim2s is missing the Pro/Ala-rich repression domain. What is surprising is that Sim2s is just as repressive as Sim2 on TCDD-mediated gene expression through a DRE (Fig. 5C) and is able to increase expression of a CME-controlled gene (Fig. 6). These data suggest a model in which the response element dictates transcription factor domain-dependent suppression or activation. On an HRE, both the Pro/Ser-rich and Pro/Ala-rich domains of Sim2 appear to exert repressive effects. This is based on the observation that Sim2s, which is missing the Pro/Ala-rich region, still exerts a repressive effect although it is not as strong as that observed with Sim2 (Fig. 5B). In the case of the DRE, both Sim2 and Sim2s can repress gene expression to an equal extent suggesting that only the Pro/Ser-rich domain mediates Sim2-mediated repression from a DRE. In contrast, Sim2s can activate expression from a CME apparently by acting as a docking protein for Arnt (Figs. 6 and 7). This implies that the Pro/Ala-rich sequence present in Sim2, but not Sim2s, exerts a negative effect on CME-mediated gene expression. As Sim2s lacks this domain, interactions between Sim2s and Arnt on a CME result in Arnt-mediated activation of Sim2s targets.

Interactions between transcription factors and their cognate response elements are influenced by sequence flanking the core binding motif. Whitelaw *et al.* (17) reported that a region of the AHR ligand-binding domain exerted different degrees of repression on different DNA targets thus, providing an example of such promoter-specific influence on transcription factor function in the bHLH-PAS family (17). These response element-specific effects are most likely due to DNA-dependent conformational changes in the interacting factor, which may influence the ability of the transcription factor to recruit co-regulatory proteins to the promoter. For example, the POU domain-containing transcription factor POU1F1 (e.g. PIT1) represses transcription when bound to its response element in the growth hormone gene, but induces expression through a similar element in the prolactin promoter. This was shown by crystallography to be due to DNA-induced allosteric changes in PIT1 confirmation resulting in differential coregulator recruitment (18). In the case of Sim2s, such dynamic, DNA-mediated changes in transcriptional outcome may reflect the ability of Sim2s to exert differential effects on similar response elements. Further studies are necessary to elucidate the mechanisms governing Sim2s-mediated gene repression and activation.

In Drosophila, several sim targets have been identified including slit, engrailed, breathless, and spitz (2, 19). Not surprisingly, Hif1α regulates many of the mammalian homologs of these genes because Hif1α can bind the same core response element. Although definitive targets of Sim2 have not been identified in mammals, this study and others have shown that Sim2 can affect the actions of other bHLH-PAS proteins by active repression and interference (8, 12).

Human SIM2 is expressed in normal kidney and tonsil as well as lung and testes (11, 20). We have found that Sim2s expression in mice is comparable, with high levels of Sim2s mRNA detected in kidney and skeletal muscle (Fig. 3A). Interestingly, the ratio of Sim2 to Sim2s differs between these tissues with Sim2s being expressed at higher levels than Sim2 in kidney and vice versa in skeletal muscle (Fig. 3B). The significance of Sim2 isoform predominance in these tissues is unknown, but presumably could have substantial effects as we have demonstrated differences between binding partner specificity (Table 1 and Fig. 4) and transcriptional activities between Sim2 and Sim2s.

Due to the complex interaction potential and overlapping expression patterns of Sim proteins and Hif1α, a hypoxic switch has been proposed to operate in cells expressing both genes (12). Such a switch could have profound implications for environmental regulation of developmental signaling pathways as developmental stage and organ-specific differences in response to systemic hypoxia have been reported (21, 22). Presently, little is known about the functions of Sim2 proteins and direct targets of mammalian Sim2 have not been reported. We are actively investigating the role of Sim2 and Sim2s in development and characterizing downstream target genes. With a better understanding of the biochemical properties of Sim2s, and identification of *bona fide* Sim2s target genes, a better comprehension of its role in development will be achieved.

Acknowledgments—We thank Dr. Stephen Safe (Texas A&M University, College Station, TX) for the pG2L-TATA, β-galactosidase and pDRE-luc plasmids and Dr. Jerry Pelletier (McGill University, Montreal, Quebec, Canada) for the pCME-luc reporter and full-length mouse Sim1 and Sim2 expression plasmids.
REFERENCES

1. Kewley, R. J., Whitelaw, M. L., and Chapman-Smith, A. (2004) Int. J. Biochem. Cell Biol. 36, 189–204
2. Crews, S., Franks, R., Hu, S., Matthews, B., and Nambu, J. (1992) J. Exp. Zool. 261, 234–244
3. Estes, P., Mosher, J., and Crews, S. T. (2001) Dev. Biol. 232, 157–175
4. Ema, M., Suzuki, M., Morita, M., Hirose, K., Sogawa, K., Matsuda, Y., Gotoh, O., Saijo, Y., Fujii, H., Hamada, H., and Fujii-Kuriyama, Y. (1996) Biochem. Biophys. Res. Commun. 218, 588–594
5. Moffett, P., Dasyo, M., Reece, M., McCormick, M. K., and Pelletier, J. (1996) Genomics 35, 144–155
6. Probst, M. R., Fan, C. M., Tessier-Lavigne, M., and Hankinson, O. (1997) J. Biol. Chem. 272, 4451–4457
7. Moffett, P., and Pelletier, J. (2000) FEBS Lett. 466, 80–86
8. Moffett, P., Reece, M., and Pelletier, J. (1997) Mol. Cell. Biol. 17, 4933–4947
9. Chrast, R., Scott, H. S., Chen, H., Kudoh, J., Rossier, C., Minoshima, S., Wang, Y., Shimizu, N., and Antonarakis, S. E. (1997) Genome Res. 7, 615–624
10. DeYoung, M. P., Scheurlle, D., Damania, H., Zylberberg, C., and Narayanan, R. (2002) Anticancer Res. 22, 3149–3157
11. DeYoung, M. P., Tress, M., and Narayanan, R. (2003) Cancer Lett. 200, 25–31
12. Woods, S. L., and Whitelaw, M. L. (2002) J. Biol. Chem. 277, 10236–10243
13. Morrow, D., Qin, C., Smith, R., 3rd, and Safe, S. (2004) J. Steroid Biochem. Mol. Biol. 88, 27–36
14. Chun, Y. S., Yeo, E. J., Choi, E., Teng, C. M., Bae, J. M., Kim, M. S., and Park, J. W. (2001) Biochem. Pharmacol. 61, 947–954
15. Sogawa, K., Nakano, R., Kobayashi, A., Kikuchi, Y., Ohe, N., Matsushita, N., and Fujii-Kuriyama, Y. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 1936–1940
16. McKay, L. M., Carpenter, R., and Roberts, S. G. (1999) Nucleic Acids Res. 27, 3064–3070
17. Whitelaw, M. L., Gustafsson, J. A., and Poellinger, L. (1994) Mol. Cell. Biol. 14, 8343–8355
18. Scully, K. M., Jacobson, E. M., Jepsen, K., Lunyak, V., Viadiu, H., Carriere, C., Rose, D. W., Hooshmand, F., Aggarwal, A. K., and Rosenfeld, M. G. (2000) Science 290, 1127–1131
19. Crews, S. T., and Fan, C. M. (1999) Curr. Opin. Genet. Dev. 9, 580–587
20. DeYoung, M. P., Tress, M., and Narayanan, R. (2003) Proc. Natl. Acad. Sci. U. S. A. 100, 4760–4765
21. Minchenko, O., Opentanova, I., and Caro, J. (2003) FEBS Lett. 554, 264–270
22. Stroka, D. M., Burkhardt, T., Desbaillets, L., Wenger, R. H., Neil, D. A., Bauer, C., Gassmann, M., and Candinas, D. (2001) FASEB J. 15, 2445–2453