Egr-1 induces DARPP-32 expression in striatal medium spiny neurons via a conserved intragenic element.

Serene Keilani  
Departments of Neurology and Pediatrics, Mount Sinai School of Medicine, New York

Samira Chandwani  
Departments of Neurology and Pediatrics, Mount Sinai School of Medicine, New York

Georgia Dolios  
Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York

Alexey Bogush  
Weinberg Unit for ALS Research, Department of Neuroscience, Farber Institute for Neuroscience, Thomas Jefferson University

Heike Beck  
Walter Brendel Center of Experimental Medicine, Ludwig Maximilians University

Let us know how access to this document benefits you

Follow this and additional works at: https://jdc.jefferson.edu/farberneursofp

Part of the Neurology Commons

Recommended Citation
Keilani, Serene; Chandwani, Samira; Dolios, Georgia; Bogush, Alexey; Beck, Heike; Hatzopoulos, Antonis K; Rao, Gadiparthi N; Thomas, Elizabeth A; Wang, Rong; and Ehrlich, Michelle E, "Egr-1 induces DARPP-32 expression in striatal medium spiny neurons via a conserved intragenic element." (2012). Faculty papers Farber Institute for Neurosciences. Paper 16.  
https://jdc.jefferson.edu/farberneursofp/16
Authors
Serene Keilani, Samira Chandwani, Georgia Dolios, Alexey Bogush, Heike Beck, Antonis K Hatzopoulos, Gadiparthi N Rao, Elizabeth A Thomas, Rong Wang, and Michelle E Ehrlich
Egr-1 Induces DARPP-32 Expression in Striatal Medium Spiny Neurons via a Conserved Intragenic Element

Serene Keilani, 1 Samira Chandwani, 1 Georgia Dolios, 2 Alexey Bogush, 1 Heike Beck, 4 Antonis K. Hatzopoulos, 5 Gadiparthi N. Rao, 6 Elizabeth A. Thomas, 7 Rong Wang, 8 and Michelle E. Ehrlich 1,2

Departments of 1Neurology and Pediatrics and 4Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York 10029, 2Weinberg Unit for ALS Research, Department of Neuroscience, Farber Institute for Neuroscience, Thomas Jefferson University, Philadelphia, Pennsylvania 19107, 3Walter Brendel Center of Experimental Medicine, Ludwig Maximilians University, D-81377 Munich, Germany, 4Division of Cardiovascular Medicine, Department of Medicine, Vanderbilt University, Nashville, Tennessee 37240, 5Department of Physiology, University of Tennessee Health Science Center, Memphis, Tennessee 38163, and 7Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037

DARPP-32 (dopamine and adenosine 3’, 5’-cyclic monophosphate CAMP-regulated phosphoprotein, 32 kDa) is a striatal-enriched protein that mediates signaling by dopamine and other first messengers in the medium spiny neurons. The transcriptional mechanisms that regulate striatal DARPP-32 expression remain enigmatic and are a subject of much interest in the efforts to induce a striatal phenotype in stem cells. We report the identification and characterization of a conserved region, also known as H10, in intron IV of the gene that codes for DARPP-32 (Ppp1r1b). This DNA sequence forms multiunit complexes with nuclear proteins from adult and embryonic striata of mice and rats. Purification of proteins from these complexes identified early growth response-1 (Egr-1). The interaction between Egr-1 and H10 was confirmed in vitro and in vivo by super-shift and chromatin immunoprecipitation assays, respectively. Importantly, brain-derived neurotrophic factor (BDNF), a known inducer of DARPP-32 and Egr-1 expression, enhanced Egr-1 binding to H10 in vitro. Moreover, overexpression of Egr-1 in primary striatal neurons induced the expression of DARPP-32, whereas a dominant-negative Egr-1 blocked DARPP-32 induction by BDNF. Together, this study identifies Egr-1 as a transcriptional activator of the Ppp1r1b gene and provides insight into the molecular mechanisms that regulate medium spiny neuron maturation.

Introduction

Gene expression in striatal medium spiny neurons (MSNs) is altered in a number of neuropsychiatric diseases, e.g., Huntington’s disease (HD). There is dysregulation in the expression of ubiquitous and striatal-enriched transcription factors (TFs) in the caudate nucleus in multiple HD mouse models and human HD subjects, which may contribute to the downregulation in the expression of striatal-enriched proteins (Luthi-Carter et al., 2000, 2002; Zucker et al., 2005; Hodges et al., 2006; Thomas et al., 2011). However, these TFs and their targets, i.e., markers of terminally differentiated MSNs, have not been linked.

MSNs are GABAergic and contribute to regulation of movement, organization of motor behavior, and cognition (Gerfen, 1992; Jain et al., 2001; Graybiel, 2005). They project to the globus pallidus or substantia nigra (Gerfen and Young, 1988), and they receive glutamatergic synaptic inputs from the cerebral cortex and dopaminergic afferents from the substantia nigra and ventral tegmental area (Parent and Hazrati, 1995). DARPP-32 (dopamine and adenosine 3’, 5’-cAMP-regulated phosphoprotein, 32 kDa), a central mediator of signal transduction pathways in the dopaminceptive MSNs, is expressed in >90% of striatal MSNs and is the most commonly used marker of this mature neuronal subtype (Gustafson et al., 1992; Ouimet et al., 1998; Greengard et al., 1999; Svenningsson et al., 2004). Identification of the molecular mechanisms that regulate DARPP-32 expression will aid in a rational design of therapeutic strategies to modulate expression of DARPP-32 and likely other striatal-enriched genes, either in neuropathological conditions of the striatum or for the induction of a striatal phenotype in stem cells.

A number of TFs that regulate neurogenesis, migration, and differentiation of MSNs have been identified, including Dlx1, Dlx2, Dlx5, Dlx6, Nolz-1, Isl-1, Mash1, bcl11b/CTIP2, and Gsh1/2. They are present during embryonic development and are essentially undetectable after the first postnatal week, with the exception of CTIP2 (Jain et al., 2001; Chang et al., 2004; Arlotta et al., 2008). None of these has been demonstrated to bind to Ppp1r1b sequences or to be sufficient for induction of DARPP-32 during development or in the adult.

We demonstrated that the promoter and intragenic regions of DARPP-32, collectively termed D9 (genomic elements from DARPP-32 encompassing 9 kb), selectively direct transgene expression to MSNs (Bogush et al., 2005; Brown et al., 2008), whereas the 2.1 kb of 5’ UTR sequence alone does not direct expression to any region of the CNS (Blau et al., 1995). In this study, we sought to characterize the function of a region in intron
IV of the Ppp1r1b gene, also known as H10. Our interest in this particular sequence arose from its high conservation among species and its location in a region of the DARPP-32 gene that is included in D9 (Bogush et al., 2005). After the identification of multiunit complexes between H10 and striatal nuclear proteins, the protein-binding site was determined. We found that the TF early growth response-1 (Egr-1) binds to H10 in vitro and in vivo. Importantly, we report that Egr-1 is sufficient to induce DARPP-32 expression in vitro and is required for the induction of DARPP-32 by brain-derived neurotrophic factor (BDNF).

Materials and Methods

Phylogenetic analysis of Ppp1r1b intronic sequences. The mouse and human genomic sequences of Ppp1r1b were retrieved from the National Center for Biotechnology Information (NCBI) GenBank database. These gene sequences were aligned pairwise (mouse vs human) using BLAST (basic local alignment search tool; b2seq) at NCBI website with the following parameters: word size = 11; reward for match = 1; penalty for mismatch = −2; gap open penalty = 5; gap extension penalty = 2; gap × dropoff = 50; and low-complexity filter checked. Only those alignments in the intronic regions (mouse) with >70% sequence homology were considered for additional analysis. The mVISTA program (Frazer et al., 2004) was used for comparative sequence analysis of the introns. The mVISTA program (Frazer et al., 2004) was used for comparative sequence analysis of the introns. The mVISTA program (Frazer et al., 2004) was used for comparative sequence analysis of the introns.

Preparations of nuclear and cytoplasmic extracts. For preparation of extracts from brain tissue, mice of either sex were killed by CO2 asphyxiation, and their brains were removed and placed in ice-cold Tris-buffered saline (pH 7.4; containing 30 mM Tris, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, and 0.1% Nonidet P-40). The brains were homogenized for 10 minutes on ice with a Dounce homogenizer, and the homogenate was centrifuged at 10,000 × g for 15 minutes at 4°C. The supernatant was collected as the cytoplasmic fraction and the pellet was washed twice with ice-cold Tris-buffered saline, and the final pellet was resuspended in 50 mM MES running buffer (pH 7.5) and used as the nuclear fraction.

Figure 1. H10 and F10A sequences are highly conserved across species (mouse, rat, human, orangutan, dog, and horse). Nucleotides that are not conserved, compared with mouse, are underlined. The numbers depict the location on mouse chromosome 11, TSS, Transcription start site; I–VI, Ppp1r1b introns.
buffer (10 mM HEPES, pH 7.9, 30 mM KCl, 1.2% glycerol, 0.5 mM DTT, 5 mM MgCl₂, 0.2 mM PMSF, and 0.1 mM EDTA, pH 8) to allow the formation of the protein–DNA complexes in solution. As controls, we combined biotinylated FI0A oligonucleotide with striatal tissue and H10 oligonucleotide with brainstem tissue. The reactions were incubated with gentle mixing for 15 min at 4°C, 15 min at room temperature, and then applied to the capture columns, which contained 250 µl of streptavidin agarose resin (50 µl of 197 pmol biotin/1 µl resin). The column was washed five times with 1× binding buffer to remove unbound DNA and protein and was then eluted with high-salt buffer at room temperature. The optimal elution condition for the H10-binding proteins as determined by EMSA was 2 M NaCl buffer (12% glycerol, 20 mT Tris, pH 6.8, 2 M NaCl, 5 mM MgCl₂, 1 mM EDTA, and 1 mM EGTA) at room temperature. Eluted proteins were analyzed by mass spectrometry.

Mass spectrometry. The pH of all three protein samples was adjusted to 8.5 with 100 mM NH₄HCO₃. Proteins were reduced with 5 mM tris(2-carboxyethyl)phosphine hydrochloride at 37°C for 20 min and alkylated with 10 mM iodoacetamide for 30 min in the dark at room temperature. A Lys-C digestion (50 ng/sample) was performed for 4 h, followed by overnight trypsin digestion at 37°C (1:50 enzyme:substrate ratio). The digestion was quenched by adding trifluoroacetic acid (5%) to achieve a pH of 2–4. The samples were desalted with ZipTip C18 (Millipore). The eluate was dried down, and peptides were reconstituted with 0.1% formic acid in water as solvent A and in ACN as solvent B, with solvent B raised from 1% to 50% in 90 min, then 50% to 85% in the next 10 min at a flow rate of 0.6 µl/min. The mass spectrometer was operated in positive mode with spray voltage at 2.5 kV, ion transfer tube voltage at 45 V, and ion transfer tube temperature at 170°C. A normalized collision energy of 35% and activation time of 30 ms were applied in MS/MS acquisitions. The top eight most intense ions were selected for fragmentation in the LTQ-Orbitrap mass spectrometer. The following dynamic exclusion settings were applied to precursor ions chosen for MS/MS analysis: repeat count, 1; repeat duration, 60 s; and exclusion duration, 240 s. MS/MS spectra were searched against the UniProtKB rat database (downloaded February 2011) using Sorcerer-SEQUEST (version 27, revision 11; Sage-N Research), Mascot (version 2.3.02; Matrix Science), X! Tandem (version 2007.01.01.1; The Global Proteome Machine Organization), and Scaffold (version 1.6.3.1; Proteome Software) algorithms (Perkins et al., 1999; Severson et al., 1999; Searle, 2010). Searches were performed with full tryptic specificity (two missed cleavages): carbamidomethylated cysteine residues as static modifications, deamidated asparagine and glutamine (+0.9840 Da), and oxidized methionine, histidine, and tryptophan (+15.9949 Da) as differential modifications. A precursor mass error tolerance of 10 ppm and default product ion mass error tolerance of ASTRO searching algorithms were used.

Chromatin immunoprecipitation assay. Chromatin immunoprecipitation (ChIP) assays were performed as described previously (Thomas et al., 2008), with some modifications. Mouse striatal equal to 100 mg wet weight were combined for each ChIP reaction. The tissue was fixed with 15% formaldehyde in PBS for 10 min at room temperature, and then applied to the capture columns, which contained 250 µl of streptavidin agarose resin (50 µl of 197 pmol biotin/1 µl resin). The column was washed five times with 1× binding buffer to remove unbound DNA and protein and was then eluted with high-salt buffer at room temperature. The optimal elution condition for the H10-binding proteins as determined by EMSA was 2 M NaCl buffer (12% glycerol, 20 mT Tris, pH 6.8, 2 M NaCl, 5 mM MgCl₂, 1 mM EDTA, and 1 mM EGTA) at room temperature. Eluted proteins were analyzed by mass spectrometry.

**Figure 2.** H10 and FI0A oligonucleotides interact in vitro with embryonic and adult striatal nuclear proteins. A, EMSA shows that 32P-labeled F10A and H10 oligonucleotides bind with proteins in embryonic (E16–17) and adult striatal nuclear extracts. The specificity of the complexes was determined by the addition of 100-fold excess of unlabeled FI0A or H10 oligonucleotide as a competitor. *Refers to nonspecific complexes, i.e., not quenched by competitor. B, EMSA reveals the formation of complexes between H10 and nuclear extracts from adult striatum, hippocampus, cortex, cerebellum, and brainstem.
then boiled for 15 min. DNA was purified for quantitative PCR analysis by QIAquick PCR purification kit (Qiagen). Quantitative reverse transcription-PCR analysis. Primers were designed with PrimerExpress software (forward, 5’-CCCAGCTTAACCCAGTACTGTTC-3’; and reverse, 5’-GGACAGGAGAGAGCGTTTGGG-3’ and analyzed with BLAT (basic local alignment tool) of the University of California, Santa Cruz genome browser. Quantitative reverse transcription-PCR was performed by SYBR Green-based real-time PCR using RT2 SYBR Green qPCR Master Mixes (catalog #PA-012; SABiosciences) according to the instructions of the manufacturer. PCR products were detected by ethidium bromide on a Fujiﬁlm LAS-3000 developer after electrophoresis on 1% Tris borate–EDTA agarose gel for 1 h at 60 V. For data analysis, the threshold cycles (Ct) of all replicates were averaged after dropping reactions with PCR inhibitors or a varying slope as determined from the melt and amplification curves. Each “cold” mutant was reacted with striatal nuclear proteins before the addition of the 32P-labeled native H10. Oligonucleotides containing the sequence 5’-CGCCCACA-3’ with the unique ﬂanking sequences from each gene in the 5’ UTR of Drd3, Baiap2, Ppp1r16b, and Cpne2 genes interact in vitro with striatal nuclear extract (NE) but not with brainstem. The band pattern is similar to that seen with H10 except for an additional, higher-molecular-weight complex (complex A). Complex B forms between complexes I and II with the Drd3, Baiap2, Ppp1r16b, and Cpne2 sequences and also intermittently with H10. E, Southwestern blotting (SW) determines that H10 striatal nuclear binding proteins range from 45 to 130 kDa.

Figure 3. Mutation analysis of H10 identifies the sequence 5’-CGCCCACA-3’, or its complement, as the TF binding site. A, Sequences of 11 mutant (Mut) oligonucleotides generated from the H10 oligonucleotide by sequentially mutating 2 nt (underlined). B, EMSA using native H10 and 11 mutated H10 oligonucleotides with striatal nuclear proteins shows that mutated nucleotides 2–5 are required for TF binding. C, Unlabeled mutant oligonucleotides 2–5 are unable to compete for protein binding with the native radioactive H10 oligonucleotide. Each “cold” mutant was reacted with striatal nuclear proteins before the addition of the 32P-labeled native H10. D, Oligonucleotides containing the sequence 5’-CGCCCACA-3’ with the unique ﬂanking sequences from each gene in the 5’ UTR of Drd3, Baiap2, Ppp1r16b, and Cpne2 genes interact in vitro with striatal nuclear extract (NE) but not with brainstem. The band pattern is similar to that seen with H10 except for an additional, higher-molecular-weight complex (complex A). Complex B forms between complexes I and II with the Drd3, Baiap2, Ppp1r16b, and Cpne2 sequences and also intermittently with H10. E, Southwestern blotting (SW) determines that H10 striatal nuclear binding proteins range from 45 to 130 kDa.
Transfection was performed using electroporation (Amaxa Biosystems). Neurons were transfected with Egr-1–pIRES2–EGFP vector (Beck et al., 2008) immediately after dissociation, and electroporation was conducted using electroporation (Amaxa Biosystems). Transfection was performed using electroporation (Amaxa Biosystems).

To begin to identify the cis-regulatory elements in the Ppp1r1b gene, we searched each sequence for TF binding sites using MatInspector from Genomatix. Putative sites in H10 and F10A were located in intron IV. In Figure 1, we show the location of two of the most highly conserved sequences, also known as F10A (75%) and H10 (91%).

**Table 1. List of H10-specific binding proteins identified in affinity chromatography eluate by MS**

<table>
<thead>
<tr>
<th>Protein name</th>
<th>Molecular weight (kDa)</th>
<th>UniPortKB accession number</th>
<th>Number of unique peptides Stem-H10</th>
<th>Striatum-F10A</th>
<th>Striatum-H10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncharacterized protein (Tpr)</td>
<td>267</td>
<td>D3ZNM2</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Myosin-9 (Myh9)</td>
<td>226</td>
<td>Q6Z812</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Uncharacterized protein (Rock2)</td>
<td>167</td>
<td>D3ZU53 (+1)</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>α-Actinin-4 (Actn4)</td>
<td>105</td>
<td>Q900X0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Clathrin coat assembly protein AP180 (Snap91)</td>
<td>94</td>
<td>Q05540 (+1)</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Uncharacterized protein (Ranbp2)</td>
<td>341</td>
<td>D4A054</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Filamin α (Fna)</td>
<td>280</td>
<td>C0F7P7 (+2)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Uncharacterized protein (Map2)</td>
<td>199</td>
<td>D3ZV41 (+3)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Uncharacterized protein (Dxh9)*</td>
<td>150</td>
<td>D3ZG09 (+1)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Cullin-associated NEDD8-dissociated protein 1 (Cand1)*</td>
<td>136</td>
<td>P9T7536</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Uncharacterized protein (Atp1a3)</td>
<td>112</td>
<td>D3ZEB8 (+3)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Uncharacterized protein (Mcm2)</td>
<td>102</td>
<td>D3ZPF6</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>SDF3 protein (Fragment) (Sdf3)</td>
<td>101</td>
<td>B5DF12</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Transcription intermediary factor 1-β (Trim28)*</td>
<td>89</td>
<td>O08629</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Uncharacterized protein (App)</td>
<td>83</td>
<td>D3ZPF7 (+3)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Nucleolin (Ncl)*</td>
<td>77</td>
<td>P13383 (+1)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ATP-binding cassette, subfamily E (OABP), member 1 (Abce1)</td>
<td>67</td>
<td>D3ZD23</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Uncharacterized protein (Ythdf1)</td>
<td>64</td>
<td>D3ZKJ3 (+5)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ATP synthase subunit α, mitochondrial (Atp5a1)</td>
<td>60</td>
<td>P15999</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Acn1 protein (Fragment) (Acn1)</td>
<td>60</td>
<td>Q8BU9</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>NADPH:adenodoxin oxidoreductase, mitochondrial (Fdxn)</td>
<td>54</td>
<td>P65522</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Early growth response-1 (Egr-1)*</td>
<td>54</td>
<td>P0B154</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Uncharacterized protein (Rbm59)*</td>
<td>50</td>
<td>D4A55G (+1)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Ferrochelatase (Predicted) (Fech)</td>
<td>48</td>
<td>D3ZBF3</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Uncharacterized protein (Rgd1560464)</td>
<td>45</td>
<td>D4A063 (+1)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Uncharacterized protein (Psip1)*</td>
<td>37</td>
<td>D4A7U3 (+3)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>β-Soluble NSF attachment protein (Nabp)</td>
<td>33</td>
<td>RPS6976</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Myristoylated alanine-rich C-kinase substrate (Marck3)</td>
<td>30</td>
<td>P30009</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>14-3-3 protein epsilon (Ywhae)</td>
<td>29</td>
<td>P62260</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>SAP domain-containing ribonucleoprotein (Sarp)*</td>
<td>24</td>
<td>O498U4</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Uncharacterized protein (Ppp1r1b)</td>
<td>22</td>
<td>D4A4B8 (+1)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>ADP-riboseylation factor 3 (Arf3)</td>
<td>21</td>
<td>P61206 (+2)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Calmodulin subunit B type 1 (Psip1r1)</td>
<td>19</td>
<td>P63100 (+1)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Uncharacterized protein</td>
<td>14</td>
<td>D3ZKLO</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Uncharacterized protein</td>
<td>12</td>
<td>D4AD8</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>RGC72700 (Timm8a2)</td>
<td>11</td>
<td>D3ZV6 (+1)</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Crosswise comparison of the proteins identified from the Striatum–H10 reaction to the ones identified from the two control columns, Striatum–F10A and Stem–H10, to eliminate the common proteins. This refinement reduced the number of the proteins of interest down to 36. * Indicates proteins with a reported function in gene transcription.

Statistical analysis. Each data point represents triplicate experiments presented as mean ± SEM. One-way ANOVA with Bonferroni’s post hoc test was performed using GraphPad Prism version 5.00 for Windows (GraphPad Software). Significance are reported with a p < 0.05.

**Results**

**Intron IV of the Ppp1r1b gene contains seven conserved sequences**

To begin to identify the cis-regulatory elements in the Ppp1r1b genomic elements (D9) that direct gene expression to MSNs (Bo-gush et al., 2005), we performed an in silico analysis of D9. We identified 30 sequences longer than 11 nt that were >70% conserved across mouse, rat, human, orangutan, dog, and horse (data not shown). Seven of these non-overlapping conserved regions were located in intron IV. In Figure 1, we show the location of two of the most highly conserved sequences, also known as F10A (75%) and H10 (91%).

**H10 and F10A form complexes with striatal nuclear proteins in vitro**

To determine whether H10 and F10A participate in the regulation of the Ppp1r1b gene, we searched each sequence for TF binding sites using MatInspector from Genomatix. Putative sites in F10A matched those for GCM1 (glial cells missing homolog 1) and AP2 (activator protein 2), whereas putative sites in H10
expression in both embryonic and adult striatum, we focused this study on H10, with F10A as a control in some assays.

To determine whether the H10–protein complexes are striatal specific, we repeated the EMSA with nuclear extracts from adult striatum, hippocampus, cortex, cerebellum, and brainstem (Fig. 2B). Complexes I–IV were of markedly greater intensity with nuclear proteins from forebrain regions relative to brainstem. Unlike in most H10 gel shifts with striatal extract, complex I was of lower intensity relative to other forebrain regions. It was the only complex visible with cerebellar extract (Fig. 2B, lane 12). We concluded that the TF(s) complexing with H10 are widely expressed throughout the forebrain. At least one of them is present in cerebellum, and, if they are expressed in brainstem, they are undetectable by EMSA.

**H10 contains a novel GC-rich protein-binding site**

To determine the precise protein-binding nucleotides in H10, we used EMSA to perform a mutation analysis. H10 variants were created by sequentially mutating 2 nt, resulting in a total of 11 novel sequences (Fig. 3A). Mutant 2 showed lower affinity for the H10-binding proteins relative to the endogenous sequence (Fig. 3B, lane 6), whereas mutants 3–5 failed to complex with striatal nuclear extract (Fig. 3B, lanes 7–9). The new complex with mutant 3 likely resulted from the generation of a non-physiologic binding site and was not further investigated (Fig. 3B, lane 7). Conversely, mutant 2 partially competed with 32P-labeled native H10 oligonucleotide (Fig. 3C, lane 5), whereas mutants 3–5 were totally unable to compete (Fig. 3C, lanes 6–8). We concluded that the sequence 5′-CGCCCAACA-3′ or 5′-TGTGGGCG-3′ is the protein-binding site in H10. The presence of multiple complexes derived from TF binding to such a short, single sequence within H10 suggests that multiple TFs likely compete for this binding site.

Based on the knowledge that each striatal TF participates in the regulation of multiple genes, we searched for the sequence 5′-CGCCCAACA-3′ in the 1 kb 5′ UTR upstream from the transcription start site of 49 striatal-enriched genes and 49 non-striatal genes (Desplats et al., 2008). We found that this sequence is found significantly more frequently in the promoters of striatal-enriched genes (8 of 49) relative to genes not enriched in the striatum (2 of 49) (Fisher’s test, one-tailed, \( p = 0.045 \)). Although we did not search in the introns, we hypothesize that we would have found similar results. We generated oligonucleotides that included 7 nt upstream and 7 nt downstream of the core sequence (sequences are detailed in Materials and Methods) and

**matched those for GREs (glucocorticoid responsive and related elements) (Fig. 1). On EMSA, the H10 oligonucleotide formed DNA–protein complexes in equal levels with both embryonic and adult striatal nuclear extract (Fig. 2A, lanes 11, 13). These complexes did not form with the embryonic (Fig. 2A, lane 9) or adult (Fig. 2B, lane 2) cytoplasmic extract and were depleted on competition with 100-fold molar excess of unlabeled H10 oligonucleotide (Fig. 2A, lanes 10, 12, 14). Conversely, F10A formed specific, high-intensity protein complexes with embryonic striatal nuclear proteins but at very low intensity with adult striatal nuclear proteins (Fig. 2A, lanes 4, 6). Because we are interested in understanding the regulation of DARPP-32 and MSN-specific

**Figure 4.** Super-shift assay demonstrates that Egr-1 binds to H10 and to oligonucleotides derived from the 5′ UTR of Baiap2, Ppp1r16b, Klf16, and Cpne2 and containing the H10 binding site 5′-CGCCCAACA-3′. A, The addition of α-Egr-1, but not α-SRF, super-shifts H10 complex I. SS, Super-shift, indicated by white arrowheads. Complexes II–IV are unaltered after immunodepletion of Egr-1 from striatal nuclear extracts (NE). B, α-Egr-1 super-shifts complex I from oligonucleotide derived from the 5′ UTR of striatal-enriched expressed genes, Baiap2, Ppp1r16b, and Klf16, and of Cpne2, a ubiquitously expressed gene. SS, Super-shift, indicated by white arrowheads. C, Two Sp1 binding sites were identified by Genomatix, the first of which is 15 nt upstream of the transcription start site of 49 striatal-enriched genes and 49 non-striatal genes (Desplats et al., 2008). We found that this sequence...
Egr-1 binds to H10 in vitro and in vivo

We used Southwestern blotting to determine the approximate molecular weights of the H10-binding proteins. Brainstem nuclear extracts were used as a negative control. Relative to brainstem, at least four striatal-enriched proteins were found to bind H10 (Fig. 3E), with estimated molecular weights at 130, 80, 60, and 45 kDa.

The H10-binding proteins were purified by affinity chromatography. The identities of the purified proteins were determined by M5 (Table 1). The sequences included two TFs (Egr-1 and Sarnp), four transcription cofactors (Dhx9, Trim28, Nucleolin, and Rbm39), and two transcriptional activators (Psip1 and CAND1). Notably, the glucocorticoid receptor was not identified. We considered Egr-1 to be a strong candidate as a factor that binds to H10 in vitro based on its consensus binding site (5’-GGGC/GGGGGC-3’) and its molecular weight of 82 kDa. We confirmed binding of Egr-1 to H10 by the super-shift of complex I (Fig. 4A, lanes 4, 10). Immunodepletion of Egr-1 from striatal nuclear extract abolished formation of complex I but did not disrupt complexes II–IV, indicating that, although they bind the same site, they form independently of Egr-1 (Fig. 4A, lane 14). Finally, super-shift assay with addition of α-Egr-1 to the EMSA striatal mixes with sequences from Baiap2, Ppp1r16b, Klf16, and Cpnell also resulted in a super-shift of complex I (Fig. 4B). This indicates that Egr-1 binds to the 5’ UTR of these genes in vitro.

To identify potential Egr-1 interacting proteins, we searched 30 nt upstream and downstream of H10. Worth noting, the addition of only 1 nt upstream of H10 identified an Egr binding site that falls on the binding site we identified by the mutational analysis. However, this nucleotide is clearly not required for that falls on the binding site we identified by the mutational analysis. The sequences derived from genes other than Ppp1r16b formed an additional complex (complex A) with nuclear extracts from both striatum and brainstem. Also, several of the sequences, including H10 from Ppp1r16b, formed a complex between I and II (complex B).

To determine whether Egr-1 binds to H10 in vivo, we performed ChIP followed by quantitative real-time PCR analysis with mouse striatal chromatin. Chromatin from NIH-3T3 cells, which do not express DARPP-32, was used as a negative control. The primers were designed to amplify a 252 bp product that contains the identified Egr-1 binding site (sequences are in Materials and Methods). The percentage of the DNA recovered from the Egr-1 IP reaction was higher than with the IgG IP reaction from striatal chromatin (*p < 0.05) or the Egr-1 IP reaction from DARPP-32-negative NIH-3T3 cells (Fig. 5).

Egr-1 is sufficient to induce DARPP-32 expression in primary striatal neurons and is necessary for induction of DARPP-32 by BDNF

We next sought to determine whether Egr-1 is necessary and/or sufficient for expression of DARPP-32 in embryonic striatal neurons in vitro. Egr-1–IRES–EGFP plasmid (Beck et al., 2008) was transfected into primary striatal neurons via nucleofection and increased the percentage of DARPP-32-immunopositive cells from 6% in neurons transfected with empty vector to 15% in neurons transfected with the Egr-1–IRES–EGFP plasmid. Although only 8% of the neurons were transfected with Egr-1–IRES–EGFP, 79% (***p < 0.001) of the neurons expressing EGFP were also DARPP-32 immunopositive. Of the neurons that were GFP negative in the same wells, only 10% were DARPP-32 immunopositive. In comparison, in empty vector GFP-transfected neurons, only 15% of the EGFP-positive cells expressed DARPP-32, and 5% of the EGFP-negative neurons were DARPP-32 immunopositive (Fig. 6).

BDNF induces expression of both DARPP-32 (Ivkovic et al., 1997, 1999) and Egr-1 (Minichiello et al., 2002; Alder et al., 2003; Calella et al., 2007), and addition of BDNF to primary striatal neurons led to an increase in Egr-1 protein after 90 min and an increase in DARPP-32 protein after 24 h (Fig. 7A). We hypothesized that BDNF induces DARPP-32 expression via inducing Egr-1 expression, thereby increasing its binding to intron IV of Ppp1r16b gene, as shown in Figure 7B.

To determine whether Egr-1 is essential for induction of DARPP-32 by BDNF, we transduced embryonic striatal primary neuronal cultures with either an empty GFP adenovirus (AdGFP)
or adnEgr-1. The adnEgr-1 construct contained the transcriptional repressor domain of the Wilms tumor protein WT1 (amino acids 1–307) at the N terminus of the Egr-1 DNA binding domain (amino acids 327–427) (Kundumani-Sridharan et al., 2010), and Western blotting of WT1 confirmed the expression of adnEgr-1 (Fig. 7C). Transduction of primary striatal neurons with adnEgr-1 slightly reduced DARPP-32 protein level relative to control (AdGFP). However, transduction of neurons with adnEgr-1 markedly reduced induction of DARPP-32 by BDNF (*p < 0.05) (Fig. 7C).

To confirm that BDNF increases Egr-1 binding to Ppplrb gene, we tested the formation of the H10 complexes before and after treating primary striatal cultures with BDNF for 3 h. The intensity of all H10 complexes, including the Egr-1 complex, increased as early as 3 h after addition of BDNF (Fig. 7D, lanes 5, 6). In addition, a slower migrating complex became apparent, which appeared similar to complex A in Figure 4B. Complex B also appeared between I and II and became much more prominent during BDNF treatment. Egr-1 is a member of a family that includes Egr-2, Egr-3, and Egr-4, all of which are expressed in the striatum. Of these, Egr-2 is the likely candidate to be redundant with Egr-2, Egr-3, and Egr-4, all of which are expressed in the striatum. The latter pathway is ERK dependent. BDNF also transiently increases phosphorylation of CREB via phosphatidylinositol 3-kinase/Akt and Cdk5 pathways and is independent of ERK (Stroppolo et al., 2001; Bogush et al., 2007). The Egr-1 promoter contains both an SRE (serum response element) and a CRE (cAMP response element) (Knapska and Kaczmarek, 2004). In the hippocampus, BDNF induces the expression of Egr-1 via the activities of TrkB, MEK, CaMK, PLCγ, and CREB (Minichelli et al., 2002; Alder et al., 2003). In the striatum, BDNF induces the phosphorylation of Elk-1, which targets the SRE site on Egr-1 (Gokce et al., 2009). Of note, however, the latter pathway is ERK independent. BDNF also transiently increases phosphorylation of CREB in MSNs (Stroppolo et al., 2001), and it is possible that DARPP-32 induction by BDNF is in part dependent on Egr-1 induction by phosphorylated CREB. This process, however, must be context specific because other inducers of CREB phosphorylation, most notably dopamine and its agonists, do not alter expression of DARPP-32 in MSNs (Ehrlich et al., 1990; Greb et al., 1990).

The striatum develops from the lateral ganglionic eminence, and the first MSNs to mature are located in patches and the lateral streak, whereas later-born neurons populate the matrix (van der Kooy and Fishell, 1987). The mature, terminal phenotype of the MSN is characterized by DARPP-32 and other markers (Desplats et al., 2006; Lobo et al., 2006). Many of these, including DARPP-32, are enriched in patch neurons during the embryonic period and become more diffusely expressed as matrix neurons mature. Others, of course, remain confined to the patch compartment. Relative to other embryonic brain regions, Egr-1 is enriched in the striatum. In the adult, it is expressed in multiple brainstem nuclei and in the striatum (Wells et al., 2011), and it is obviously induced in multiple regions after neuronal activation. Within the embryonic striatum, Egr-1 is further enriched in patches, perhaps suggesting a developmental, regulatory role (Snyder-Keller et al., 2002). Beckwith et al. (1966)
hypothesizes that a mature neuronal subtype is determined by a group of TFs that induce a specific set of genes, which together comprise a "regulon," a term previously restricted to prokaryotes. Furthermore, cis-motifs may be bound by multiple TFs, creating a system for coordinated, master, regulation of phenotype (Flames and Hobert, 2011). Thus, we show that the H10 binding sequence is more frequently found in striatal-enriched genes and that Egr-1 binds to these sequences in vitro in the presence of distinct flanking sequences. Thus, Egr-1 may induce additional markers of the mature MSN.

Egr-1 is a ubiquitously expressed immediate early gene, usually associated with neuronal plasticity after stimuli that alter intracellular calcium stores (Knapski and Kaczmarek, 2004; Thiel et al., 2010). Despite its widespread expression, many examples exist of regulation of cell-specific markers by Egr-1, both within and outside the nervous system (Boyle et al., 2009). Tissue-specific expression of proteins may be mediated by cell-specific TFs but also by unique combinations of ubiquitously expressed TFs but also by unique combinations of ubiquitously expressed TFs (Hobert et al., 2010). For example, Egr-1 acts synergistically with NFATc to regulate expression of the IL-2 gene in T lymphocytes by binding independently to adjacent sites within the IL-2 gene promoter (Decker et al., 1998).

Thus, it is unlikely that Egr-1 alone can induce the DARPP-32 phenotype in precursors that have not yet committed to the MSN lineage but may be sufficient once the neurons are committed. In both cases, therefore, Egr-1 almost certainly acts in concert with other TFs, because induction of Egr-1 in most cell subtypes does not induce DARPP-32. For example, Egr-1 is required in part for BDNF induction of the tyrosine hydroxylase promoter (Fukuchi et al., 2010), but DARPP-32 is not expressed in aminergic neurons.

Conversely, the adult Egr-1-null mouse has a normal level of striatal DARPP-32 (Valjent et al., 2006), indicating that Egr-1 is ultimately not necessary for expression of DARPP-32 in vivo. Importantly, development of the striatum in the absence of Egr-1 has not been examined, and we know that, in the BDNF-null mouse, DARPP-32 expression is reduced early but normalizes during the second postnatal week (Ivkovic et al., 1997). Also, because we show that Egr-2 binds to the same H10 binding site as does Egr-1, Egr-2 may substitute for Egr-1. An Egr-2 regulated transcription network has been suggested to exist in the striatum, downstream of the adenosine A1 receptor. However, the involvement of Egr-2 in regulating DARPP-32 transcription in vivo remains to be determined.

Our data are highly relevant to the pathophysiology of diseases in which the striatum plays a role and their potential therapies. Notably, Egr-1 levels are specifically decreased in the dorsal sub-

region of the striatum in R6/2 mouse model of HD (Spektor et al., 2002) and are elevated in the striatum after exposure to drugs of abuse (Valjent et al., 2000, 2006). More important perhaps is the potential contribution to cell replacement therapy and to the use of induced pluripotent stem (iPS) cells for the study of disease mechanisms. At this juncture, these approaches are hampered by the inability to induce large numbers of striatal subtype-specific neurons that are physiologically active and able to properly integrate into the CNS. Evidence is accumulating that the forced expression of neuronal subtype-specific TFs is able to re-program cells, e.g., fibroblasts, and to convert embryonic stem cells and iPS

![Figure 7](image-url)
to a specific phenotype (Caiazzo et al., 2011; Son et al., 2011; Sánchez-Danés et al., 2012). Perhaps the most successful MSN differentiation program to date is that of Aubry et al. (2008) in which neural precursors are treated with BDNF and then with BDNF and valproic acid for terminal differentiation into MSN-like neurons. Multiple markers were assayed, although not their colocalization, and only extrinsic signals were used to induce differentiation. In addition to our findings, Egr-1 was recently identified in a novel SOX2-positive neuronal progenitor cell population in developing NeuN-positive striatal neurons, indicating a role for Egr-1 in both intermediate and terminal differentiation of MSNs and perhaps in maintenance of the differentiated phenotype (Wells et al., 2011).

In summary, our study is the first to identify a TF, Egr-1, that directly interacts with and regulates Ppp1r1b gene expression in MSNs. In addition, we demonstrate that it acts downstream of BDNF, a major regulator of DARPP-32 and other markers of the differentiated MSN. As we continue to identify specific TFs that can directly induce DARPP-32, it will be critical to test them for their utility in driving an MSN phenotype, the identification of which should include multiple, colocalized markers and physiologically

References
Gerfen CR (1992) The neostriatal mosaic: multiple levels of compartmental organization in the basal ganglia. Annu Rev Neurosci 15:285–320.
Egr-1 Induces DARPP-32 Expression


Keilani et al.


