Functional Analysis of Upstream Common Polymorphisms of the Dopamine Transporter Gene

Mikhil N. Bamne², Michael E. Talkowski²,³, Kodavali V. Chowdari², and Vishwajit L. Nimgaonkar¹,³

²Department of Psychiatry, Western Psychiatric Institute and Clinic, University of Pittsburgh School of Medicine, 3811 O’Hara Street, Pittsburgh, PA 15213; ³Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15213

The human dopamine transporter (DAT, SLC6A3) has been extensively investigated because of its potential involvement in neuropsychiatric disorders. The core elements responsible for its transcription have been identified. A regulatory role for certain genomic variants upstream to the core promoter is known. Recently, other single-nucleotide polymorphisms (SNPs) have been identified in this region and are thought to be associated with schizophrenia and bipolar I disorder. Hence, we have investigated the impact of common SNPs in a 2.8-kilobase region flanking the core promoter region (−2.7 to +63 base pair) in the neuroblastoma cell line SH-SY5Y. Haplotypes generated by site-directed mutagenesis revealed varying impact of individual SNPs on promoter activity using dual luciferase assays. In silico analyses also predicted allele-specific binding of transcription factors for some of these SNPs. Though electrophoretic mobility shift assays indicated several factors that appeared to bind to specific sites within this region, allele-specific binding was not detected for any SNP apart from rs3756450. We have thus identified novel putative regulatory domains flanking the core promoter of DAT that merit further investigation.

Key words: dopamine transporter/single-nucleotide polymorphism/electrophoretic mobility shift assay/reporter assay/schizophrenia/bipolar disorder

Introduction

There is growing evidence for intricate molecular mechanisms that regulate dopamine (DA) levels in the brain. DA released into the synaptic space can undergo enzymatic degradation and dilution by diffusion. DA is primarily metabolized by 2 mechanisms, namely, oxidative deamination by monoamine oxidase and O-methylation by catechol-O-methyltransferase. The primary mechanism controlling extracellular DA levels in most brain regions is likely to be reuptake by presynaptic neurons via the plasma membrane dopamine transporter (DAT/SLC6A3). DAT is a sodium-dependent, membrane-bound transporter. Alterations in DAT expression can have a direct effect on DA reuptake and availability of free DA in the synapses. DAT is localized to chromosome 5p15.3 and spans approximately 60 kilobases (kb). It incorporates 15 exons, with the protein-coding portion of the gene being encoded by exons 2–15. The protein product spans 12 transmembrane domains, cytoplasmic amino- and carboxy terminals, and a glycosylated second extracellular loop. Genetic association studies have implicated DAT in the etiology of several psychiatric disorders, including attention deficit/hyperactivity disorder, schizoid/avoidant behavior, bipolar disorder, and schizophrenia.

The DAT core promoter region has been identified. It is unusual because promoter sequence analysis did not reveal conventional “TATA” or “CAT” boxes upstream of the transcription initiation site. A 180–base pair (bp) GC-rich sequence incorporating multiple Sp1 sites may direct transcription. Several other factors regulating DAT transcription have been identified, including NURR1, HEY1/HESR, and Stp1/Spt11,13–15

Apart from the core promoter region, transcriptional regulation has also been reported in other genomic sequences. A downstream variable number of tandem repeat element in exon 15 has long been thought to regulate transcription, but the regulatory activity appears to vary for individual studies. Additional sequences in the 5′ untranslated region (5′ UTR) are also known to modulate DAT transcription. Based on positron emission tomography and postmortem [³H]carboxyfluoropropyl-binding assays, Drögn et al²⁰ suggested that cis-acting polymorphisms in the 5′ UTR influence DAT expression. Kelada et al²¹ sequenced approximately 7.4-kb region in relation to the transcription start site (−5273 to +2118 bp). The promoter activity of several common haplotypes that encompassed single-nucleotide polymorphisms
(SNPs) in this region was demonstrated, using luciferase reporter assays. Thus, several lines of evidence indicate a regulatory role for upstream SNPs in DAT transcriptional regulation. However, it is uncertain whether individual SNPs independently affect transcriptional activity or whether the haplotype data indicate synergistic effects of several cis-acting elements. This question has become relevant because of recent genetic association studies. A study involving Iranian cases and controls revealed association between a SNP in this region (rs2975226, −67A/T) and bipolar I disorder (odds ratio, OR = 2.25).22 We recently detected associations between DAT SNPs and schizophrenia.10 Following comprehensive analysis of a 2.8-kb region spanning the core promoter of DAT SNPs and suggested disorder-related associations within 2.8-kb region encompassed by a common haplotype (−2783 to +63 base pair) used for reference in the present study and a previous analysis (Kelada et al21). Vertical arrows indicate single-nucleotide polymorphisms (SNPs) analyzed in the present study. **SNP associated with bipolar I disorder (Khodayari et al22). **SNP (T allele) associated with schizophrenia (Talkowski et al). We re-estimated the majority of the replicable interactions in vitro. Electrophoretic mobility shift assays (EMSA) also indicated allele-specific DNA-protein interactions. The present study was intended to extend these analyses. We have evaluated the role of additional individual SNPs on transcriptional regulation of DAT using in vitro promoter assays. The promoter assays were followed by transcription factors (TFs) binding analyses, which included in silico analyses, as well as EMSA.

**Methods**

**Region of Interest.** The earlier association studies suggested disorder-related associations within 2.8-kb region spanning the core promoter of DAT.10,22 Earlier sequencing assays by Kelada et al21 in the panel of 40 Caucasian participants revealed 7 SNPs (rs3756450, rs2652510, rs28362317, rs2550956, rs2617596, rs2652511, and rs2975226) in this region, documented in public databases (Hapmap.org). These SNPs were evaluated as follows.

**Cloning and Dual Luciferase Assays.** Initially, an approximately 2.8-kb region including a 5′ UTR genomic region of DAT (−2783 to +63 bp) from a Centre d’Etude du Polymorphisme Humain (CEPH) individual (CEPH: 1340–2) homozygous for rs3756450 was polymerase chain reaction (PCR) amplified with Expand High Fidelity PCR System (Roche Applied Science). The haplotype thus generated served as the reference haplotype.10,21 (see figure 1). The genomic fragments amplified with Kpn1 and HindIII sites were cloned into a pGL3 Basic Vector (Promega, Madison, WI). The inserted sequences were confirmed by sequencing. Each of the 6 known common SNPs (rs2652510, rs28362317, rs2550956, rs2617596, rs2652511, rs2975226) between −2783 and +63 bp were then sequentially altered using a QuikChange II site-directed mutagenesis kit (Stratagene, La Jolla, CA). All the sequence variations generated by site-directed mutagenesis were confirmed by sequencing. Transfections and luciferase assays were conducted as described earlier.10 Briefly, the constructs (1.0 μg) were cotransfected with Renilla luciferase (20 ng)–expressing vector pRL-TK (Promega, Inc) into SH-SY5Y, a neuroblastoma cell line (American Type Culture Collection) (1 × 105 cells per well) using lipofectamine reagent (Invitrogen, Carlsbad, CA). Transfections were conducted in triplicate in 24-well culture plates. Six hours after transfection, fresh medium with 2× fetal bovine serum (Invitrogen) was added to each well. After incubation for 30 hours, cells were harvested. Firefly luciferase activities from cell lysates for individual constructs were normalized to Renilla luciferase activity. Normalized activities of all the test clones were compared with those of clones representing the common haplotype (incorporating rs3756450−T, see figure 2). Assays were read on a Victor 3 multilabel plate reader (Perkin Elmer, Turku, Finland) according to the manufacturer’s protocol (Promega, Inc). Each assay was conducted in triplicate in 2 independent experiments.

**Transcription Factor Binding Prediction.** We used the in silico prediction software MatInspector to identify putative binding sites for TFs at specific SNPs in the region of interest.23 TFs were filtered for vertebrate specificity. For a given genomic region, this software evaluates putative binding sites for TFs in its directory and yields a matrix similarity score. We evaluated the impact of allelic changes for each of the SNPs of interest across the 2.8-kb region used in the promoter assays. A 100% match to the matrix yields a maximum score of 1.00 (each base position has the highest conserved nucleotide at that position in the matrix), and a matrix similarity score greater than 0.80 is considered a “good” match.
FIG. 2. Dual Luciferase Promoter Assay for 5' Untranslated Region Dopamine Transporter Single-Nucleotide Polymorphisms (SNPs). All assays were conducted using SH-SY5Y cell lines. Relative luciferase activities for individual constructs are shown, with induced variations. Values for individual SNPs are standardized against the reference haplotype (*common haplotype*) and are represented as means of triplicate assays for each construct. The error bars represent SDs. Relative positions of SNPs with respect to the transcription start point are given in brackets. The rs3756450T/C data were published earlier (Talkowski et al10) but are presented here in terms of change relative to the common haplotype.21 The analyses included differences in luciferase activity between each SNP, compared with the reference haplotype (P value less than .005).

Electrophoretic Mobility Shift Assays. Nonradioactive EMSA was conducted using published procedures for the following SNPs: rs2652510, rs28362317, rs2550956, rs2617596, rs2652511, rs2975226.10 Briefly, allele-specific single-stranded primers (IDT DNA Inc., Coralville, Iowa) encompassing SNPs were annealed to their complementary fragments to generate double-stranded probes and labeled using protocol suggested by DIG Gel Shift Kit (Roche Applied Science Mannheim, Germany) (see sequences in Supplementary Table 1). Labeled double-stranded probes were incubated at room temperature for 30 minutes with SH-SY5Y cell line nuclear extracts (5 μg) in 20 μl reaction (5× binding buffer, poly-L-lysine, poly[d(I-C)]). Appropriate control reactions comprising 50× excess unlabeled probe and nuclear extract were included for each SNP/allele. Labeled oligonucleotide-nuclear extract complexes separated on 6% nondenaturing polyacrylamide gel were electroblotteded on positively charged nylon membranes (Boehringer Mannheim-Roche Applied Science, Inc). Blots were treated by an enzyme chemiluminescent method (Roche Applied Science, Inc), and the resolved bands were visualized on the x-ray film. OCT-1 protein (pOCT-1) provided with the DIG Gel Shift Kit was used as control to probe the effect of pOCT-1 to allele-specific changes in band shift pattern for rs3756450. The primer sequences used for generating allele-specific probes are listed in the Supplementary Table 1. The altered bases are highlighted. Cell line nuclear extracts were prepared as described elsewhere.24

Checking Endogenous DAT Expression by Reverse Transcription-PCR. Total RNA from the SH-SY5Y cell line and control human postmortem substantia nigra tissue was extracted according to the manufacturer’s specification (Invitrogen, Inc). The cDNA was synthesized using Random hexamer primers and Superscript III First Strand Synthesis kit (Invitrogen, Inc). Reverse transcription-PCR (RT-PCR) amplification utilized primers hDAT, E2F + hDAT.E4R and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (see sequences in Supplementary Table 2). The 10-μl PCR mixture contained 0.2 mM each of dATP, dCTP, dGTP, and dTTP; 10× PCR buffer (15 mM MgCl2, 500 mM KCl, 1.5 M, and Tris–HCl [pH 8.7]); 200 nM of the appropriate paired primers; and 0.25 U of AmpliTaq Gold DNA polymerase (Applied Biosystems, Branchburg, NJ). The PCR conditions were 95°C for 2 minutes; 40 cycles of 95°C for 45 seconds, 60°C for 30 seconds, and 72°C for 45 seconds; followed by final extension at 72°C for 7 minutes. The PCR products were electrophoresed on 2% agarose.

Statistics. Paired t tests were used to determine differences in luciferase activity between each SNP, compared with the reference haplotype.

Results

Reporter Assays for Individual SNPs

We evaluated promoter activity for each of 7 SNPs flanking the core promoter of DAT against the background of the most frequent haplotype.21 The analyses included rs3756450, which has been reported on earlier.10 The SNPs had varying effects on promoter activity. The most substantial impact was noted at rs2550956, where a C→T change caused 4-fold reduction in the reporter expression compared with the common haplotype. Two-fold changes were observed at rs2652511 (T→C) and rs2975226 (T→A). Less substantial changes were noted at the other SNPs, though all the SNPs appeared to have significant impact (all P values were less than .005) (figure 2).

Putative Factors Modulating Transcription of the DAT Gene

MatInspector identified 15 different TFs that would be predicted to have significant binding in the genomic region spanning these SNPs. Binding for 9 TFs would be abolished by allelic variation at particular SNPs. MatInspector software analysis suggested putative binding sites for domains encompassing all the SNPs except rs2652510 (Table 1). Consistent with our prior report, allele-specific differences in affinities for 5 different TFs were noted at SNP rs3756450, with each allele interacting with a unique set of binding partners. SNP rs2550956 shows allele-specific binding profile. C→T variation abolishes binding
Table 1. Predicted Transcription Factors Binding Sites at Individual SNPs

<table>
<thead>
<tr>
<th>SNP (MAF)</th>
<th>Alleles</th>
<th>Putative binding sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs3756450 (0.158)</td>
<td>C</td>
<td>GATA1.04 (0.974), LEF1.01 (0.885), OCT1.04 (0.825), GATA2.01 (0.945), HOXC13.01 (0.913)</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>rs2652510 (0.44)</td>
<td>A</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>G</td>
<td></td>
</tr>
<tr>
<td>rs28362317 (ND)</td>
<td>A</td>
<td>AHRARNT.03*</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>rs2550956 (0.22)</td>
<td>C</td>
<td>AHRARNT.03 (0.954)</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>rs2617596 (0.46)</td>
<td>G</td>
<td>WT1.01 (0.967), KKL1.01 (0.932)</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>KKL1.01 (0.930)</td>
</tr>
<tr>
<td>rs2652511 (0.46)</td>
<td>T</td>
<td>AP2.01 (0.919), ZID.01 (0.856), TEAD.01 (0.935)</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>ZID.01 (0.856), TEAD.01 (0.935).</td>
</tr>
<tr>
<td>rs2975226 (0.44)</td>
<td>T</td>
<td>EGR3.01 (0.800), CKROX.01 (0.984), CTTF.01 (0.826), GC.01 (0.958)</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>CKROX.01 (0.984), CTTF.01 (0.826), GC.01 (0.958).</td>
</tr>
</tbody>
</table>

Note: All analyses were conducted using MatInspector, software, Genomatix Inc. Transcription factors (TFs) in bold show allele-specific variations for the corresponding single-nucleotide polymorphism (SNP). Percentage matrix similarity score (in parentheses) accompany each TF for each motif encompassing corresponding SNP (100% match to the matrix yields a maximum score of 1.00). MAF, minor allele frequency (based on National Center for Biotechnology Information [NCBI] SNP browser—http://www.ncbi.nlm.nih.gov/sites/entrez?db=snp); ND, not available based on NCBI SNP browser. GATA1.04, GATA-binding factor 1; LEF1.01, TCF/LEF-1, involved in the Wnt signal transduction pathway; OCT1.04, octamer-binding factor 1; GATA2.01, GATA-binding factor 2; HOXC13.01, homeodomain TF HOXC13; AHRARNT.03, dioxin response elements, xenobiotic response elements bound by AHR/ARNT heterodimers; WT1.01, Wilms tumor suppressor; KKL1.01, kidney-enriched kruppel-like factor; AP2.01, activator protein 2; ZID.01, zinc finger with interaction domain; TEAD.01, TEA domain–containing factors, transcriptional enhancer factors 1,3,4,5; EGR3.01, early growth response gene 3 product; CKROX.01, collagen krox protein (zinc finger protein 67—zfp67); CTTF.01, CCCTC-binding factor; GC.01, GC box elements.

*Binding for AHRARNT0.3 is dependent on allelic change at neighboring SNP.

Electrophoretic Mobility Shift Assays

To evaluate allele-specific effects, we conducted EMSA using allele-specific probes (figure 3A–F). We included unlabeled probes in the EMSA experiments to reduce nonspecific binding. We observed DNA-protein interaction for all SNPs. With the exception of rs3756450, no allele-specific DNA-protein gel shift bands were observed.10

When we used pOCT-1 instead of tissue extract in our assays involving rs3756450, no band shifts were observed (data shown in Supplementary Figure 1).

Endogenous Expression of DAT in SH-SY5Y Cells

We investigated endogenous expression of DAT in SH-SY5Y cell line. We targeted the regions overlapping exon 2 to exon 4 of DAT cDNA. RT-PCR analysis suggested very low endogenous expression of DAT. RNA isolated from human postmortem substantia nigra tissue was investigated as a positive control for DAT gene expression, and GAPDH expression was used as internal control for RT-PCR in these experiments (data shown in Supplementary Figure 2).

Discussion

Our analyses complement earlier studies, which reported transcription regulatory effects for haplotypes in the region.16,21 Here we show that SNPs encompassing these haplotypes have significant individual impacts. As reported earlier, we find that allele C for rs3756450 appeared to have lower promoter activity compared with allele T against the same haplotypes background. The T allele was found to be associated with schizophrenia in our previous association study.10 We observed similar differences in the expression of the T vs C alleles (figure 2). The results differ in Kelada et al21 and may reflect the impact of the larger genomic construct used in that study. Two SNPs have been reported to be associated with bipolar disorder/schizophrenia in this region (rs2975226 and rs3756450).10,22 Although these SNPs appear to have impact on promoter activity, other SNPs such as rs2652510, rs28362317, rs2550956, rs2617596, and rs2652511 have greater activity compared with rs2975226 but lower than rs3756450. Kelada et al21 ascertained haplotype pattern for these 7 SNPs along with other SNPs flanking this region. Based on their genotyping results, SNPs rs2652510, rs2617596, rs2652511, and rs2975226 are in linkage disequilibrium (LD) with each other and with intron 1 SNPs (rs2963288, rs2937639, rs2937640, and rs2975223—not investigated in this study). Other SNPs rs2550956 and rs28362317 were not in LD with any of the SNPs in the vicinity. SNP rs3756450 was in LD with one of the intron 1 SNP (rs2937638) but not with the remaining 6 SNPs assayed...
in this study. Thus, it would be important to evaluate associations in those SNPs.

Our prior analyses suggested allele-specific binding with putative TFs at rs3756450. Therefore, we investigated similar effects at the flanking SNPs. MatInspector software revealed putative binding sites for domains encompassing all the SNPs, except rs2652510. It should also be noted that the SNPs might have cooperative effects. For example, we found that allelic changes at SNP rs2550956 abolished binding of TF AHRARNT.03 at neighboring SNP rs28362317. Allele-specific effects were predicted for several SNPs. These analyses motivated further EMSA analyses. DNA-protein binding was observed for all these SNPs (figure 3A–F). However, we did not observe any allele-specific band shift for other SNPs, as observed in our earlier study. The inconsistency between the in silico and in vitro analyses may be due to a number of factors. First, additional trans-acting factors may required for the formation of stable DNA-protein complexes. Based on our EMSA experiments, the band shift pattern observed for rs3756450 is unlikely to be due to TF OCT-1 alone. Alternately, the significant changes observed in the reporter assay could reflect subtle changes in these DNA-protein complex formation not detectable with EMSA. But, the conditions employed for EMSA in our study may not be sensitive enough to detect these minute band shifts on gel. It should be noted that though MatInspector does not

---

**Fig. 3.** Electrophoretic Mobility Shift Assays (EMSA) for SLC6A3 Single-Nucleotide Polymorphisms (SNPs). Representative EMSA for SNPs A) rs2652510, B) rs28362317, C) rs2550956, D) rs2617596, E) rs2652511, and F) rs2975226. Nuclear factors extracted from SH-SY5Y cell line were incubated with allele-specific labeled probes. Reaction conditions for each lane are presented over each lane. Arrows indicate binding of unknown nuclear factors to allele-specific labeled probes. To ascertain specific binding, 50× unlabeled probes were used as competitor. (Lanes 3 and 6).
predict any TF binding to the genomic region flanking SNP rs2652510, we do observe DNA-protein interaction for this region (figure 3B). It is possible that a TF not reported in MatInspector database might be interacting with the genomic probe representing rs2652510 and warrants further investigation. Additionally, though we used appropriate controls, an EMSA-related artifact could not be ruled out.

In conclusion, we suggest that genomic variations flanking the core promoter of the DAT gene may contribute to its transcriptional control. We observed DNA-protein interactions in this region, indicating the presence of putative transcriptional domains. Further analyses to identify TFs interacting with the domains encompassing these SNPs are needed in order to understand transcriptional regulation of DAT.

Supplementary Material

Supplementary Tables 1 and 2 and Figures 1 and 2 are available at http://schizophreniabulletin.oxfordjournals.org.

Funding

National Institute of Mental Health (MH56242 and MH63480 to V.L.N.).

References