EXPLORING FUNCTIONAL GENETIC VARIANTS IN GENES INVOLVED IN MENTAL DISORDERS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Ying Zhang, M.D.

*****

The Ohio State University
2007

Dissertation Committee:

Professor Wolfgang Sadée, Advisor

Professor Laura M. Bohn

Professor Howard Gu

Professor David W. Saffen

Approved by

Advisor

Neuroscience Graduate Studies Program
Mental disorders including drug addiction are complex diseases with strong genetic components. A number of candidate genes have been intensely studied, but the effects of genetic variants for most of them remain elusive. Clinical studies have demonstrated associations of genetic variants in candidate genes with susceptibility to several common mental diseases such as schizophrenia, depression, bipolar disorders, and drug addiction, but little was known about the underlying mechanisms. Single nucleotide polymorphisms (SNPs), naturally occurring nucleotide substitutions in DNA sequence, are the most common genetic variants, sometimes reaching high frequency in target populations. Genome-wide survey studies have shown SNPs in noncoding regions are more prevalent than those SNPs that alter protein sequence (nonsynonomous SNPs). This suggests a pervasive effect of those SNPs on inter-individual variability. A large number of studies has shown that SNPs in coding and noncoding regions of a gene can modulate gene expression at the level of mRNA and protein by affecting transcription, mRNA processing, pre-mature RNA splicing, mRNA stability, or translation. Moreover, functional SNPs can affect alternative splicing, which is controlled by cis and trans regulation, another important mechanism causing inter-individual variability. My objectives were to test the hypothesis that functional SNPs regulating mRNA expression and processing of the mu opioid receptor and D2 dopamine receptor exist at frequencies
that have the potential to contribute to differential susceptibility to mental disorders in the human population.

To address this goal, I first determined whether there is mRNA expression imbalance (AEI) between two alleles of OPRM1 and DRD2 in relevant target tissues (postmortem brain samples). Primer extension assays (SNaPShot) was used to detect allelic mRNA expression in human postmortem brain tissues using indicator SNP A118G for OPRM1 and three SNPs (rs6275, rs6277, and rs6279) for DRD2, respectively. Allelic expression imbalance (AEI) was observed for both OPRM1 and DRD2. For OPRM1 (Chapter 2), in 8 heterozygous samples measured, the A118 mRNA allele was 1.5- to 2.5-fold more abundant than the G118 allele. Transfection into CHO cells of a cDNA representing only the coding region of OPRM1, carrying A, G, C, or T in position 118, resulted in 1.5-fold lower mRNA levels only for OPRM1-G118, and more than tenfold lower OPRM1 protein levels, measured by Western blotting and receptor binding assay. After transfection and inhibition of transcription with actinomycin D, analysis of mRNA turnover failed to reveal differences in mRNA stability between A118 and G118 alleles, indicating a defect in transcription or mRNA maturation. These results indicate that OPRM1-G118 is a functional polymorphism with deleterious effects on both mRNA and protein yield. Clarifying the functional relevance of polymorphisms associated with susceptibility to a complex disorder such as drug addiction provides a foundation for clinical association studies.
In chapter 3, among 68 heterozygous samples tested for AEI in DRD2, 7 of them displayed higher expression of the major allele (C allele) than T allele, while another 8 samples had the opposite trend. To search for regulatory polymorphisms, I performed SNP scanning of the gene locus, which identified a novel regulatory SNP (rs12364283) located in a conserved suppressor region, with the minor allele (~7% allele frequency) causing enhanced expression. Moreover, differences in allelic mRNA expression of the long and short DRD2 splice isoforms - showing distinct pre- and post-synaptic functions - revealed two linked intronic SNPs (rs2283265 and rs1076560, 17% minor allele frequency) strongly associated with reduced formation of DRD2-short. Exploratory clinical studies supported an association of the promoter variant with schizophrenia, while the two intronic SNPs significantly associated with enhanced cognitive processing (done by collaboration with Dr. Bertolino from the University of Bari). These results indicate that DRD2 variants appear to have strong effects on dopaminergic transmission in vivo.

In chapter 4, I investigated and characterized the direct binding of G protein coupled receptor derived peptides with calmodulin using a quantitative assay (S-Tag). Calmodulin binding to receptor regions involved in G protein coupling appears to play a role in the regulation of signaling, and at least one naturally occurring polymorphisms in the mu opioid receptor had previously been shown to affect CaM binding. CaM binding of peptides derived from the third intracellular loop (i3) of the mu opioid receptor (MOR)
was confirmed and the CaM-binding motif refined. A MORi3 peptide with a Lys>Ala substitution - shown to reduce CaM-binding of intact MOR - bound 5-fold less avidly than the wild-type peptide. Screening peptides derived from i3 loops of other GPCR families confirmed 5HT1A, and identified muscarinic receptor 3, and melanocortin receptor 1, as proteins carrying CaM-binding domains. The use of S-Tag labeling can serve for rapid screening of putative CaM-binding domains in GPCRs.

This study examines the extent and nature of regulatory polymorphisms in candidate genes, and provides a foundation for clinical association studies by clarifying the functional relevance of polymorphisms associated with susceptibility to mental disorders.
DEDICATION

Dedicated to my family and all my friends.
ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Sadée, for his excellent guidance during my graduate studies. Inspired by his gorgeous ideas of doing research, I get to know how to think of scientific questions critically, how to solve the problems step by step, and how to explain the experiment results from multiple views. He always encourage me to believe in myself and trust the judgment of myself during the process of project studies, which is critical for becoming an independent researcher and scientist. He always is there when I need help, which makes me always in the right track for pursuing my study goals. He also gives me tremendous help in scientific writing during the past five years, through which I gained great progress. Without his advises and support, I would not be able to accomplish what I have done today.

I am also thankful for my lab colleagues (Audrey C. Papp, Julia K Pinsonneault, Zunyan Dai, Gloria Smith, Andrew Johnson, and Danxin Wang); they are intelligent, warm heart, and friendly. We have learned from each other by sharing our ideas and thinking. Their kindness help and support in my study and in my life make it possible of what I have achieved, and make the whole time of my study such a pleasant journey.

My advisory committee members, Dr. David Saffen, Dr. Laura M. Bohn and Dr.
Howard Gu, continually devoted their precious time and experience to direct my research.

My family also lends abundant support and help in my life, which enable me to concentrate on my research and go this far. My husband, Yunlong Zhang, is always with me and encourages me when I feel frustrated, and during the past two years when we separate from each other, his understanding and support is very important for me.

Finally, I would like to say that all of the guidance during my Ph.D. training made the dissertation possible. All of your kindness is greatly appreciated!
VITA

1993……………….. M.D. Clinical Medicine, Shandong University
1998………………….. M.S. Cell biology Beijing University
2001- present………. PhD Neuroscience, The Ohio State University

PULICATIONS
Research Publication


FIELDS OF STUDY

Major Field: Neuroscience,

Pharmacogenomics
TABLE OF CONTENTS

ABSTRACT.................................................................................................................ii
DEDICATION..............................................................................................................vi
ACKNOWLEDGMENTS...............................................................................................vii
VITA...........................................................................................................................ix
LIST OF TABLES...........................................................................................................xiii
LIST OF FIGURES.......................................................................................................xiv

Chapters:

1. Introduction..............................................................................................................1
   1.1. OPRM1 and DRD2 in drug addiction.................................................................1
   1.2. Mental disorders and DRD2..............................................................................3
   1.3. The nature of human genetic variability.........................................................4
   1.4. The effects of \textit{cis}-acting polymorphisms on mRNA expression.................5
   1.5. The regulation of alternative splicing.............................................................6
   1.6. The effects of epigenetic DNA modification on mRNA expression..................7
   1.7. Genetic variations of mu opioid receptor.......................................................10
   1.8. Genetic variations of dopamine receptor D2..................................................13
1.9. Directly binding of calmodulin with G-protein coupled receptors ..................14
1.10. Summary of the main studies ...............................................................15

2. Allelic Expression Imbalance of Human mu Opioid Receptor (OPRM1) Caused by Variant A118G ......................................................................................20
   2.1. Introduction .........................................................................................20
   2.2. Material and Methods .......................................................................22
   2.3. Results ...............................................................................................31
   2.4. Discussion .........................................................................................38

3. Frequent polymorphisms in the human dopamine DRD2 receptor affect gene expression and splicing .................................................................57
   3.1. Introduction .........................................................................................57
   3.2. Material and Methods .......................................................................61
   3.3. Results ...............................................................................................78
   3.4. Discussion .........................................................................................87

4. Calmodulin interaction with peptides from G-protein coupled receptors measured with S-Tag labeling ..............................................................125
   4.1. Introduction .........................................................................................125
4.2. Material and Methods ................................................................. 128
4.3. Results and Discussion ............................................................. 130

BIBLIOGRAPHY ........................................................................... 144
LIST OF TABLES

2.1. Peak area ratios of mRNA and DNA for heterozygous brain samples of two SNPs analyzed by SNaPshot................................................................. 45

2.2. Peak area ratios for mRNA and plasmid DNA from CHO cells..................47

3.1. Genotyped SNPs of DRD2.................................................................98

3.2. Oligonucleotides designed for experiments.......................................99

3.3. Linkage disequilibrium analysis for 19 SNPs of DRD2 in Stanley samples....103

3.4. The association of rs12364283 with AEI..........................................105

3.5. Repeat variants in the upstream of DRD2.........................................105

3.6. The association of rs12364293 with Schizophrenia............................106

4.1. Peptide sequences.............................................................................136
LIST OF FIGURES

1.1. Human genetic variability, involving \textit{cis}- and \textit{trans}-acting factors………………..18
1.2. Schematics of quantitative allele-specific expression analysis…………………………….19

2.1. Correlation between fluorescence peak area measured by SNaPshot and the ratios of added plasmid DNAs…………………………………………………………48
2.2. mRNA expression for OPRM1 with different genotypes………………………………..50
2.3. mRNA expression of OPRM1 constructs in transfected CHO cells………………….52
2.4. Protein expression for different constructs of OPRM1 in CHO transfected cells….54
2.5. Mfold analysis to predict secondary structures for 4 possible substitutions at position 118 in human \textit{OPRM1} mRNA……………………………………………………56

3.1. Gene map and linkage blocks for DRD2…………………………………………………107
3.2. Control experiments to test PCR efficiency and calculate standard curves for DRD2S and S expression……………………………………………………………..109
3.3. The measurement of Allele-specific expression of DRD2 using three indicator SNPs…………………………………………………………………………………111
3.4. Association analysis between single SNPs and AEI using HelixTree…………………..113
3.5. \textit{cis}-Regulation of alternative splicing of DRD2………………………………………115
3.6. The correlation of rs1076560 genotype status with altered splicing
3.7. Promoter region constructs and their transcription activities tested by reporter gene assay.

3.8. Alternative splicing from DRD2 minigene constructs in HEK cells.

3.9. rs1076560 genotype-based analysis of fMRI response during the working memory task.

3.10. Alignment of DNA sequences flanking rs12364283 from several species.

4.1. Dose-dependent binding of S-Tag CaMKII CBD and S-Tag MORi3 with CaM.

4.2. Titration of S-Tag MORi3 binding to CaM-agarose beads with unlabeled MOR peptides.

4.3. Titration of S-Tag MORi3 binding to CaM-agarose beads with unlabeled peptides derived from muscarinic receptors M1-3.

4.4. Sequence alignment of M1, M2 and M3.

4.5. CaM binding of S-Tag peptides derived from various GPCRs and Dose-dependent binding of S-Tag MC1Ri3 to CaM beads.
CHAPTER 1

INTRODUCTION

1.1 OPRM1 and DRD2 in drug addiction

Drug addiction is a chronic relapsing disorder that is defined by two major characteristics: a compulsion to take a drug with a narrowing of the behavior toward excessive drug intake, and a loss of control in limiting intake despite harmful consequences. Drug abuse and addiction is “One of America's Most Challenging Public Health Problems” (NIDA homepage: http://www.nida.nih.gov/about/welcome/aboutdrugabuse/). There are various addictive drugs such as alcohol, cocaine, heroin, marijuana, methamphetamine, and nicotine. Although through distinct mechanisms, they share some common pathways to cause compulsive behaviors characteristic of drug addiction. Addiction is often (but not always) accompanied by physical dependence, a withdrawal syndrome and tolerance. Physical dependence is defined as a physiologic state of adaptation to a substance, the absence of which produces symptoms and signs of withdrawal. Withdrawal syndrome consists of a predictable group of signs and symptoms such as restlessness, insomnia, nausea, sweating, tremors, and anxiety, resulting from abrupt removal of, or a rapid decrease
in the regular dosage of, a psychoactive substance. *Tolerance* is a state in which a drug produces a diminishing biologic or behavioral response when taken over time; in other words, higher doses are needed to produce the same effect that the user experienced initially.

Although multiple factors contribute to the vulnerability to develop an addiction, abundant studies as reviewed in (Kreek et al., 2004) have demonstrated the importance of genetic factors in drug abuse. A growing number of genes are significantly associated with one or several kinds of drug abuse. Serving as a primary target for opioid drugs and peptides, the mu opioid receptor (OPRM1) is well known to be involved in drug abuse by mediating the effects of morphine and heroin (Basbaum and Fields, 1984; Kreek, 1996b). Mu opioid receptor knockout mice have no drug reinforcement or augmented locomotor activities mediated by morphine (Kieffer, 1999; Matthes et al., 1996). By impinging on dopaminergic pathways, OPRM1 also plays a role in addiction to other drugs of abuse, such as cocaine, nicotine, and alcohol (Herz, 1997; Kreek, 1996a). Stimulation of the mu opioid receptor induces the disinhibition of mesolimbic-cortical dopamine system that is central to the reinforcing properties caused by addiction drugs (Johnson and North, 1992). Strong evidence shows that the dopaminergic system is the major substrate of reward and reinforcement of addictive drugs (e.g., alcohol, cocaine, and heroin) (Wise and Bozarth, 1987). Addictive drugs can increase the firing rate of dopamine neurons in the ventral tegmental area (VTA), the amount of dopamine released into the nucleus accumbens
(NAc) (Di Chiara and Imperato, 1988), or activate the dopamine system through endogenous opioid pathways located in VTA or Nac (Johnson and North, 1992). Dopamine D2 receptors (DRD2) on VTA are activated by dopamine and inhibit the mesolimbic dopaminergic pathway (Di Chiara and Imperato, 1988; Koob, 1992). In the absence of DRD2 expression, the rewarding effects of opiates have been abolished (Maldonado et al., 1997).

1.2 Mental disorders and DRD2

Dopaminergic neurons located in substantia nigra, ventral tegmental area and hypothalamus project their axons to multiple brain areas including striatum, neocortex, limbic system and hypophysis. Therefore, as one of the primary targets for dopamine, DRD2 is involved in various neurophysiological activities such as cognition and memory, movement control, and drug addiction, and its altered activities are implicated in several mental disorders in addition to drug addiction such as schizophrenia, attention-deficit-hyperactivity (ADHD), and Parkinsons’ diseases (Bozzi and Borrelli, 2006; Seeman, 2006; Starr, 1995; Takahashi et al., 2006). Recently, it has been shown, that working memory impairments play an important role in the aetiopathogenesis of mental disorders such as schizophrrenia and working memory disturbances are demonstrated as a cognitive endophenotypic marker of vulnerability to these illnesses. The altered function of dopamine D2 receptor is thought to be one of the mechanisms accounting for the
impairment of working memory in schizophrenia patients. Right now one of the main treatments for schizophrenia is still focused on DRD2 blockade and modulation. To obtain an optimum treatment outcome and also minimize the adverse effects such as extrapyramidal symptoms, the optimal occupancy of D2 receptor is critical. As DRD2 located in both pre- and postsynaptic membranes with distinct functions in each site, balancing the pre and postsynaptic D2 receptor antagonism is another mechanism to prevent excessive D2 blockade (Horacek et al., 2006).

1.3 The nature of human genetic variability

Phenotypic differences can arise from genetic polymorphisms, by changing the protein coding sequence or affecting transcription and mRNA processing (splicing, mRNA maturation and stability, mRNA trafficking, other) (Figure 1.1). Recent surveys have indicated that polymorphisms located in the cis-acting regulatory regions and at sites that potentially affect mRNA processing and stability are much more common than the polymorphisms changing protein coding. Furthermore, in the process of transcription and translation, double stranded DNA is first transcribed into single strand mRNA. Using MFOLD (Zuker, 2003) for predicting mRNA secondary structures has demonstrated that most of the nucleotide acid alterations are capable of changing the RNA structure and potentially functions (Johnson A, et al., manuscript in preparation), for example affecting mRNA stability. Moreover, around half of the human genes undergo alternative splicing
(Modrek and Lee, 2002), one of the main sources for genetic diversity. An increasing number of studies has revealed the effects of genetic variants on splicing (Attaie et al., 1997; Laporte et al., 1997; Maillet et al., 1999). Lastly, epigenetic factors which can be transmitted in somatic cells without alterations in the primary DNA sequence (Grewal and Moazed, 2003) are being involved in mental disorders (Petronis, 2003) and could also play a role in addiction. Taken together, cis-acting functional polymorphisms that affect gene expression may represent a main cause of human phenotypic variability.

1.4 The effects of cis-acting polymorphisms on mRNA expression

Both cis and trans regulatory elements contribute to the regulation of mRNA expression or translation. Recent studies have shown that regulation of gene expression by cis-acting polymorphisms represents a prevalent mechanism affecting mRNA levels. In 2003, Yan et al. used a method based on the primer extension assay to compare the relative expression levels of two alleles for 13 genes. They found 6 out of 13 exhibited allelic expression variation and differences ranged from 6% to 30% (Yan et al., 2002). Bray et al. used the same method to detect the relative allelic mRNA expression of catechol-O-methyltransferase (COMT), and reported that one haplotype implicated in schizophrenia was related to lower mRNA expression in human brain. The carriers of short allele in promoter region (S-carriers) of serotonin transporter (5-HTT), which plays an important role in neurophysiology and is thought to be involved in the process of
psychiatric disorders was reported to be associated with much lower expression of mRNA and protein in vitro and in vivo (Heils et al., 1996; Lesch et al., 1996), but a study from our laboratory cast doubt over the relevance of this variants (Lim et al., 2006). Another important transporter ABCB1, which pumps out a broad range of natural and pharmacological substrates (chemotherapeutical drugs, antinociception drugs) from cells, also have allelic expression imbalance associated with one synonymous SNP, C3435T (Hitzl et al., 2004; Hoffmeyer et al., 2000; Wang et al., 2005a).

1.5 The regulation of alternative splicing

It is common to have several splice variants of one gene co-existing in eukaryotic cells, with different expression pattern of splice variants of one gene over developmental time and across specific tissues (or regions in brain) (Chinta et al., 2005; Rahman et al., 2002; Xu et al., 2001). Moreover, splice variants also can have different functions, such as physiological and pharmacological properties (Castro and Strange, 1993; Jiang et al., 2003; Perry et al., 2002). DRD2 has three splice variants in humans, the long form (D2 long), short form (D2 short) and longer form (D2 longer), showing altered pharmacological properties and expression in human. DRD2 is a target for several antipsychotic drugs, including atypical antipsychotic drugs such as clozapine. D2 short is more sensitive to benzamide antipsychotic drugs (Castro and Strange, 1993; Malmberg et al., 1993) compared to D2 long. The splice forms also display tissue-specific regulation
and expression (Autelitano and van den Buuse, 1995; Neve et al., 1991). D2 longer, a newly identified splice variant with two additional amino acids (Val and Glu) in the third intracellular loop of the receptor, has higher affinity for dopamine and higher potency to inhibit adenylyl cyclase (Liu et al., 2000). The mechanisms regulating D2 splice variant expression remain unknown, but an increasing number of studies have shown effects of genetic variants on alternative splicing (Bodzioch et al., 2001; Perez et al., 2003; Plant et al., 1999). Therefore a specific splice variant may be associated with allele specific expression when this variant is linked with a cis-acting SNP. Altered mRNA stability and turnover can also have accounted for allelic mRNA expression imbalance (Tebo et al., 2003; Wang et al., 2005a). Duan et al. have shown that the synonymous SNP C957T changes the stability of DRD2 mRNA and turnover rate of the receptor in vitro (Duan et al., 2003), with T957 (minor allele) showing faster turnover rate (two folds). MFOLD program (Zuker, 2003), a program to predict secondary structure of mRNA, predicted an obvious change of mRNA folding structure spanning the sequences 808-1212 for the T957 mutant, which may cause the mRNA stability alteration (Duan et al., 2003).

1.6 The effects of epigenetic DNA modification on mRNA expression

In addition to cis and trans regulation of mRNA expression, genomic regulation has also been ascribed to epigenetic DNA modification (Bestor et al., 1994). As a major epigenetic modification, DNA methylation is involved in several cellular processes such
as chromatin structure modification (Bird and Wolffe, 1999), X-chromosome inactivation (Brockdorff, 2002), and genomic imprinting (Brannan and Bartolomei, 1999; Constancia et al., 1998). One mechanism of methylation-related regulation is that the methylation of transcription factor binding sites causes the inhibition of transcription of specific genes (Attwood et al., 2002). CpG dinucleotides within CpG islands, mainly in promoter regions, are usually unmethylated, and the extent and pattern of methylation are tissue and species specific. Methylation of CpG island can change for particular genes with aging (Issa, 2000). Numerous studies have linked the abnormal methylation of CpG island with cancer development (Costello and Plass, 2001; Issa, 2000), but only recently, an increasing number of studies have demonstrated a correlation of methylation with mental disorders (e.g., drug addiction, schizophrenia, and bipolar disorders) (Abdolmaleky et al., 2004; Kan et al., 2004; Petronis, 2000; Petronis, 2003). In animal experiments, alcohol taken during pregnancy reduces the mRNA expression of brain-derived neurotrophic factor, which is associated with hypermethylation of the gene (Maier et al., 1999). The expression of serotonin receptor 2A (5HT2A) has been shown to be imprinted, with only maternal allele being expressed in human cell (Kato et al., 1996). However, in another study in humans, only 4 out of 18 tested brain samples displayed monoallelic expression of 5HT2A, while 14 out of 18 showed biallelic expression, suggesting a polymorphic imprinting of this gene in human population (Bunzel et al., 1998). DRD2 also has CpG islands and several methylation targets in its promoter region. Evidence showed a
significantly higher degree of methylation in the DNA from lymphocytes than from striata. Higher methylation was detected in right than in left striatum, and a greater degree of methylation was detected in older than in younger individuals (Popendikyte et al., 1999). This indicates epigenetic regulation of \textit{DRD2} expression. It remains to be determined whether DNA methylation is related to allelic expression imbalance, but methylation patterns of certain genes appear polymorphic and may also play an important role in the regulation of allele-specific expression (Bunzel et al., 1998; Jinno et al., 1994; Xu et al., 1993).

The alteration of gene expression can directly influence the functional potency of a gene by changing the amount of the encoded proteins. This is especially important for those proteins (receptors, transporters, and metabolizing enzymes) that mediate or metabolize physiologically or pharmacologically functional molecules, such as transmitters, small peptides, and chemical drugs. With the increase or decrease of those proteins, the effects of their substrates may be changed correspondingly. Most addictive drugs target G-protein coupled receptors to trigger downstream signal transductions. Thus, the expression level modulation of a receptor, e.g., mu opioid receptor, could limit or enhance the effects of drugs (opiates for example) mediated by this receptor. We hypothesize that \textit{cis}-acting polymorphisms that cause mRNA or protein expression
alteration for addiction associated receptors such as OPRM1 and DRD2 can contribute to
inter-individual differences in the vulnerability to drug abuse, and to the variability in the
outcomes of drug abuse treatment.

1.7 Genetic variations of mu opioid receptor

Single nucleotide polymorphism (SNP) is a single nucleotide base (A, T, C, or G)
change in a DNA sequence that occurs in a significant proportion of a large population.
Among human beings, 99.9% bases of genome sequence are the same, but the remaining
0.1% (11 million SNPs) make people unique (Kreek et al., 2005). As one of the main
causes for genetic variance, SNPs can account for inter-individual phenotype differences
by altering protein-coding sequence or affecting mRNA or protein expressions. It has
been hypothesized that the vulnerability of an individual to drug addictions is related to
multiple SNPs of multiple genes. As the primary target for opioid addiction, OPRM1 is
one of the most studied genes, with many variants identified and tested for association
with drug abuse (Gscheidel et al., 2000; LaForge et al., 2000). Hoehe et al. identified 43
variants and 52 haplotypes from a 7kb genome sequence around OPRM1 using multiplex
sequence comparison. After classifying those haplotypes into two main functional groups,
they revealed one group was more frequent in the substance-dependent individuals .
Among those SNPs, A118G (Asn40Asp), a nonsynonymous SNP with Asparagine (Asn)
replaced by Aspartic acid (Asp) in position 40, has been extensively studied. This SNP,
with a frequency from 2% to 30% across different populations, leads to the loss of a potential N-linked glycosylation site that is important for guiding the receptor across the cytoplasm membrane (Bart et al., 2005; George et al., 1986). G118 also appears to have altered pharmacological and physiological properties compared with A118. In 1998, Bond et al. reported that OPRM1-Asp40 had 3-fold higher affinity for β-endorphin than OPRM1-Asn40 but not for other agonists or antagonists included in the studies (Met-Enkephalin, DAMGO, morphine, methadone, and naloxone), suggesting a gain of function (Bond et al., 1998). OPRM1-Asp40 is also three-fold more potent in the activation of agonist-activated $K^+$ channels than Asn40 (Bond et al., 1998). Two groups have also reported altered Hypothalamic-Pituitary-Adrenal axis activity for G118 carriers. Their studies showed that individuals with G118 had higher basal cortisol levels and after being treated with naloxone, gained an enhanced cortisol response (Hernandez-Avila et al., 2003; Wand et al., 2002). Subjects with G118 also have smaller pupil constriction response to morphine-6-glucuronide compared with A118 homozygous individuals (Lotsch et al., 2002). More recently, Oslin et al. have reported that G118 homozygotes or heterozygotes respond better to naltrexone treatments for alcoholism (Oslin et al., 2003). However, there are inconsistent results. Befort et al. did not find any agonist binding affinity difference between the two variants in transient transfected COS-7 cells (Befort et al., 2001). Furthermore, Beyer et al. reported no apparent differences in receptor binding affinity, capacity, and desensitization of the two polymorphic forms (Beyer et al.,
Therefore, the functional significance of the A118G variant of *OPRM1* remains unresolved. A number of clinical studies, mostly case-control studies, evaluate the effects of this SNP in the vulnerability for drug abuse. Szeto et al. reported a high percentage of Asp40 in substance dependent subjects from a Chinese Han population (Szeto et al., 2001). Similarly, Tan et al. showed that G118 was highly associated with heroin dependent patients in an Indian population (Tan et al., 2003). Schinka et al. also indicated that A118G was a general risk factor for alcohol and other substance dependence (Schinka et al., 2002). Most recently, Bart et al. revealed significantly higher frequency of G118 carriers in heroin addicts from central Sweden (Bart et al., 2005). Conversely, some studies showed no association of this SNP with drug abuse. Bergen et al. revealed no association of this SNP with alcohol dependence (Bergen et al., 1997). Gscheidel et al. reported that none of the five SNPs in exon 1 of *OPRM1* including A118G contributed to vulnerability for drug abuse (Gscheidel et al., 2000). Franke and colleagues tested the allele frequencies for A118G in German Caucasians, using 287 opiate dependent and 365 control subjects, but found no differences in genotype frequencies between those two groups (Franke et al., 2001). Taken together, these results indicate that the A118G polymorphism does affect *OPRM1* function but the underlying mechanisms remain uncertain; moreover, the clinical relevance of genetic variants in *OPRM1* remain to be clarified.
1.8 Genetic variations of dopamine receptor D2

Because of their essential role in developing reward and reinforcement induced by addictive drugs, dopamine receptors are commonly studied targets in drug addiction research. Taq1A, a SNP 10kb downstream of DRD2, was the first shown to be associated with alcoholism (Hallikainen et al., 2003). Subsequent studies reported the association of this SNP with heroin abuse (Lawford et al., 2000), cocaine dependence (Blum et al., 1993), cigarette craving (Erblich et al., 2005) and multiple substance abuses (Comings et al., 1994). Another polymorphism in the promoter region, -141C del/Ins, has been related to heroin addiction in a Chinese Han population (Li and Villalobo, 2002), as well as to alcoholism in Mexican-Americans (Konishi et al., 2004). One intronic SNP (in intron 2), TaqIB, was reported to be associated with multiple drug abuses in Caucasians, but not in African Americans (O'Hara et al., 1993). However, for each SNP, there are also contradictory results from association-studies. Matsushita et al. did not find association of Taq1A with alcoholism in Japanese (Matsushita et al., 2001), nor in a Chinese Han population (Chen et al., 2001). Wiesbeck et al. reported no association of –141C del/Ins with primary alcohol dependence (Wiesbeck et al., 2003). One synonymous SNP, C957T in exon 7, demonstrated to be able to affect mRNA stability of DRD2 (Duan et al., 2003), was shown to have no association with drug abuse (Parsian et al., 2000; Xu et al., 2004).
Taken together, these studies indicate that allelic variants can contribute to inter-individual phenotype differences (vulnerability to drug abuse). However, a number of controversial results also suggest that it is still unclear whether and how much each SNP contributes to the phenotypic differences. Discrepant results can be due to population admixture, whereby multiple interacting genetic factors can have different frequencies in distinct ethnic populations. Moreover, linkage disequilibrium of an indicator SNP used in association studies with the functional polymorphism may vary among ethnic populations. Therefore, it is critically important that we understand the molecular genetic mechanisms underlying inter-individual variability.

1.9 Direct binding of calmodulin with G-protein coupled receptors

Calmodulin (CaM), a universal Ca\(^{2+}\) sensor, interacts with numerous signal proteins including protein kinesis, ion channels, and transcription factors (Meyer et al., 1992; Schulman et al., 1992). Recent studies have further revealed a direct interaction of CaM with G protein coupled receptors (GPCRs) (Minakami et al., 1997; Nakajima et al., 1999; Turner et al., 2004). CaM binds to the i3 loop of MOR at a domain that overlaps with a region required for G protein coupling (Wang et al., 1999). As a consequence, CaM was found to regulate G protein coupling of MOR. Moreover, upon agonist stimulation, CaM is released from the receptor and appears to serve as a signal protein, regulating CREB phosphorylation and EGF receptor transactivation (Belcheva et al.,
2001; Wang et al., 2000a). Similarly, Bofill-Cardona et al. have reported that CaM
directly binds to D2-dopamine (D2) receptors, thereby suppressing G protein activation
(Bofill-Cardona et al., 2000). They identified a CaM binding domain in the N terminus of
i3 loop, adjacent to the region required for G protein coupling.

To characterize GPCR derived peptide-CaM interactions more accurately and
expand the search for CaM binding motifs in GPCRs, I have developed a rapid
quantitative peptide-CaM interaction assay using an in vitro translation system with S-
Tag labeled peptides, coupled to a sensitive S-Tag detection method (Kim and Raines,
1993). This assay served to measure CaM-binding of peptides derived from the i3loops
of MOR, muscarinic, melanocortin, and serotonin receptors. Screening peptides derived
from i3 loops of other GPCR families confirmed MOR, 5HT1A, and identified
muscarinic receptor 3, and melanocortin receptor 1, as proteins carrying CaM-binding
domains. The use of S-Tag labeling can serve for rapid screening of putative CaM-
binding domains in GPCRs.

1.10 Summary of the main studies

My studies are to test the hypothesis that functional polymorphisms located in
transcribed or non-transcribed regions of OPRM1 and DRD2 modulate the susceptibility
to mental disorders. I first determined whether there is mRNA expression imbalance
between two alleles of OPRM1 and DRD2 in relevant target tissue (postmortem brain
samples). Primer extension assays (SNaPShot) (Figure 1.2) was used to detect allelic mRNA expression in human postmortem brain tissues using indicator SNP A118G for OPRM1 and three SNPs (rs6275, rs6277, and rs6279) for DRD2, respectively. Allelic expression imbalance (AEI) has been revealed for both OPRM1 and DRD2. The main efforts of this project were to search for functional polymorphisms or other factors responsible for the observed allelic expression imbalance of OPRM1 and DRD2. For OPRM1, in vitro studies were conducted to detect the effects of A118G on allelic expression imbalance, and its effects on mRNA transcription and protein translation in transfected CHO cells. In the case of DRD2, genotyping and haplotype analysis were used to scan for polymorphisms associated with allelic expression imbalance, and one SNP (rs12364283) in the upstream regulatory region of DRD2 was highly associated with allelic expression imbalance (AEI) and its effects were confirmed by reporter gene assay in HEK and SH-SY5Y cells. Moreover, two highly linked intronic SNPs (rs2283265 and rs1076560, 17% minor allele frequency) were associated with reduced formation of the DRD2S (short) splice variant relative to DRD2L (long) – two splice variants thought to have distinct roles in pre- and post-synaptic dopamine signaling, respectively. Consistent with this finding, we found (collaborated with Dr. Bertolino in the University of Bari, Bari, Italy) the minor allele of the two intronic SNPs was associated with greater activity of the striatum and of the prefrontal cortex in healthy humans assessed with fMRI during working memory. These results reveal novel functional genetic DRD2 variants which
impact on its expression and physiology of subcortical and cortical brain regions during working memory, a mechanism that may be relevant to several brain disorders. Another clinical case-control study (Stanley sample collections, Bethesta, MD) revealed much higher allele frequency of rs12364283 ($\chi^2 = 6.89, P = 0.009 \ (n = 95)$) in schizophrenia subjects, which need to be repeat in another cohort with larger sample size.

It is also noted that our identified functional SNPs can only explain partial AEI results, so other genetic (epigenetic) factors are also likely play a role in regulating gene expression of DRD2, which need further studies.

Clarifying the functional relevance of polymorphisms associated with susceptibility to mental disorders provides a foundation for clinical association studies. This study has the potential to improve the prevention, diagnosis and treatment of drug addiction.
Figure 1.1. Human genetic variability, involving cis- and trans-acting factors. cis-Acting regulatory polymorphisms outnumber those affecting protein sequence and function. cis-Acting polymorphisms in transcription factors cause multiple trans-acting changes. Epigenetic changes can mimic cis-acting polymorphisms, while small RNAs can also regulate activity. Measuring allelic expression imbalance (AEI) integrates the effects of all cis-acting factors determining mRNA expression and processing.
Figure 1.2. Schematics of quantitative allele-specific expression analysis. The two colored peaks (blue and green) represent two alleles of DNA (mRNA), with the peak area corresponding to the expression level of two alleles. Compared to DNA with equal amount of molecules from two alleles, mRNA may have differential expression between two alleles.
CHAPTER 2

Allelic Expression Imbalance of Human mu Opioid Receptor

(OPRM1) Caused by Variant A118G

Program in Pharmacogenomics, Department of Pharmacology, College of Medicine and Public Health, The Ohio State University, Columbus, OH, 43210

2.1 Introduction

Drug addiction is a complex disorder with a strong genetic component (Kreek, 1996a; Kreek et al., 2004). Serving as a primary target for opioid drugs and peptides, the mu opioid receptor (OPRM1) mediates the effects of morphine and heroin (Basbaum and Fields, 1984; Kreek, 1996b). By impinging on dopaminergic pathways, OPRM1 also plays a role in addiction to other drugs of abuse, such as cocaine, nicotine, and alcohol (Herz, 1997; Kreek, 1996b).

Because of its central role in drug addiction, numerous studies have addressed potential contributions of polymorphisms in the gene encoding OPRM1 to addiction susceptibility. Among multiple single nucleotide polymorphisms (SNPs) in
OPRM1, C17T (Ala6Val) and A118G (Asn40Asp) are well-studied nonsynonymous SNPs located at the N-terminus of the receptor. In particular, A118G, with an allele frequency of 10-32% in different ethnic groups (LaForge et al., 2000), has been associated with susceptibility to heroin, nicotine, and alcohol addiction (Bart et al., 2004; Schinka et al., 2002; Tan et al., 2003). However, other studies have failed to corroborate these associations (Crowley et al., 2003; Gscheidel et al., 2000).

The A118G polymorphism has also been linked to differences in pharmacological properties of OPRM1. In a population of European descent, the G118 variant was shown to correlate with better response to naltrexone in the treatment of alcoholism (Oslin et al., 2003). Moreover, OPRM1-antagonist naloxone elicited an increased cortisol response in individuals with a G118 allele (Hernandez-Avila et al., 2003; Wand et al., 2002). In transfected cells, OPRM1-Asp40 was reported to have 3-fold higher affinity for β-endorphin than OPRM1-Asn40 (Bond et al., 1998), suggesting a gain of function, but subsequent studies have failed to corroborate these results (Befort et al., 2001; Beyer et al., 2004). In vitro transfection studies however indicated that the G118 allele might be associated with lower OPRM1 protein expression than the A118 allele (20). Therefore, the functional significance of the A118G variant of OPRM1 remained unresolved.

Cis-acting functional polymorphisms that affect transcription, mRNA processing, mRNA stability and protein translation, rather than primary protein structure, may represent a main cause of human phenotypic variability (Johnson et al., 2005; Lo et al., 2005; Li et al., 2004).
Measuring allelic imbalance of mRNA expression has proven to be a powerful tool in detecting such *cis*-acting polymorphisms (Bray et al., 2004; Bray et al., 2003; Yan and Zhou, 2004). In this approach, one measures the expression of each of two alleles simultaneously in a target tissue of heterozygous individuals, thereby, eliminating the influence of *trans*-acting factors. In this study, we have measured allelic *OPRM1* mRNA expression, using *A118G* as the indicator SNP, in human brain tissues (autopsies). We have also determined mRNA and protein expression in CHO cells, after transfection of plasmids carrying the *OPRM1* coding region, with *A*, *G*, *C*, and *T* substituted at position 118. The results demonstrate that the *A118G* substitution in *OPRM1* causes significantly reduced yields of mRNA and receptor protein, indicating a loss of function.

2.2 Material and Methods

**Human brain tissues**

87 human post-mortem brain tissue samples were obtained from different sources as follows: 16 from Cooperative Human Tissue Network, PA, and 12 from Department of Pathology, Ohio State University (mainly cortical tissues samples); 10 human pons tissues from Brain and Tissue Bank, University of Maryland, Baltimore, MD, and 49 pons section slides from Stanley Research Medical Foundation (Bethesda, MD).
Genomic DNA and RNA preparation

Genomic DNA and RNA were prepared from frozen brain samples or pons section slides. Tissue samples were treated with sucrose-Triton and digested with proteinase K, followed by sodium chloride precipitation of protein and ethanol precipitation of DNA. Total RNA was extracted with TRIzol reagent (Invitrogen, CA) and purified using QIAGEN RNeasy columns according to the manufacturers’ instructions. Before reverse transcription, total RNA was treated with DNase I (Ambion, TX) at 37°C for 30 minutes. Complementary DNA (cDNA) was generated from 1 µg total RNA with Superscript II reverse transcriptase (Invitrogen, CA), using oligo[dT] and OPRM1-specific primers targeting a region 3’ of A118G (primer for cDNA synthesis: 5’ CAGGTCGCTGCCTGTTCC).

Genotyping of OPRM1

A118G and C17T were scored in genomic DNA prepared from 87 brain tissue samples using SNaPshot (Applied Biosystems, CA) as described below.

SNaPshot and quantitative analysis of allelic ratios in mRNA and genomic DNA

SNaPshot was performed as described previously (Pinsonneault et al., 2004). In brief, a section of DNA or cDNA surrounding the SNP (~100bp) was amplified by polymerase chain reaction (PCR) using amplification primers (forward: 5’
GGTTCCCTGGGTCACCTTGTC, reverse: 5’ CAGGTCGGTGCAGTTTC) under the following PCR conditions: 95°C for 2 minutes, followed by 30 cycles of 95°C for 15 seconds, 60°C for 1 minute, and 72°C 1 minute. For the analysis of C17T, the following PCR primers were used, under the same conditions. Forward: 5’ CCGTCAGTACCATGGACAGC 3’, reverse: 5’ GAGTACGCCAAGGCATCAGT 3’.

After PCR, residual primers and unincorporated dNTPs were inactivated by incubating at 37°C for 3 hrs with 5 units of Antarctic Phosphatase (BAP) and 2 units of ExoI (New England Biolabs, MA). The purified PCR products were then analyzed using a primer extension method (SNaPshot). The extension primers were designed to anneal to the amplified template DNA immediately adjacent to the A118G site. Subsequent extension with DNA polymerase added a single fluorescent dideoxyribonucleoside triphosphate (ddNTP) complementary to the nucleotide at the polymorphic site. The extended primers, labeled with different fluorescent dyes were run on an ABI 3730 capillary electrophoresis instrument, and analyzed with Gene Mapper 3.0 software (Applied Biosystems, CA). Peak area ratios were calculated to measure the relative amount of DNA (cDNA) of the two alleles. Control experiments with known mixtures of plasmids carrying different nucleotides in the A118 position demonstrated that peak areas and ratios were linearly related to added amounts of DNA (cDNA). Because of differences in fluorescent yield and terminator dye incorporation, measured peak area ratios differed from unity when a 1:1 ratio of allelic DNA was added, Because of differences in fluorescent yield and
terminator dye incorporation, measured peak area ratios differed from unity when a 1:1 ratio of allelic DNA was added, depending on the incorporated fluorescent dye. For each brain tissue, peak area ratios were measured for both genomic DNA and mRNA (cDNA). Assuming that the two alleles were present in equal amounts in genomic DNA, measured DNA and cDNA ratios were normalized to the average of genomic DNA ratios using the equation: normalized ratio = measured DNA (or cDNA) ratio / mean of genomic DNA ratio. For genomic DNA, none of the samples significantly deviated from the expected unity, indicating that the alleles were present in equal quantities in the tissues analyzed (absence of gene dosage effects on allelic expression levels). As an added control, extension primers for A118G were used in both orientations in separate analyses, and DNA and mRNA/cDNA ratios calculated in the same fashion to test for any artifacts that may have been introduced by fluorescent dye bias (forward primer: 5’ ACTGATCGACTTGTCACCACCTTAGATGGC 3’; reverse primer: 5’ ACTGACTGACTGACCATGGGTCGGACAGGT 3’). For C17T, the extension primer was 5’ ATTGCTGGCGTTTCTGGGG 3’. Allelic ratios of A118/G118 and C17/T17 were measured in all heterozygous brain tissues (8 samples for A118G and 1 sample for C17T) by determining peak area ratios. For cDNA preparations, each mRNA samples was converted to cDNA in three separate experiments, and each cDNA sample was analyzed at least twice using SNaPshot.
mRNA analysis using real-time PCR

To quantify mRNA expression of the genotyped samples, we first set up a standard curve using various amount (50fg-100pg) of plasmid DNA (OPRM1), and performed real-time PCR with the same PCR amplification primers used for SNaPshot, with 12.5μl 2× PCR SYBR master mix (3mM MgCl₂, 200μM dATP, 200μM dCTP, 200μM dGTP, 200μM dUTP, 0.625U AmpliTaq Gold, 0.25U AmpErase UNG). The mixtures were run on an ABI 7000 with the same PCR conditions, and threshold cycles measured. OPRM1 cDNA levels were calculated from the obtained standard curve. As an internal control, β-actin was amplified for the same samples, and expression levels were determined by a standard curve using different dilutions of a mixture composed of cDNAs from all analyzed samples. The cDNA expression level of each sample was expressed as the ratio of OPRM1 over β-actin. To detect any possible DNA contamination, real-time PCR was also performed in RNA control samples in the absence of reverse transcriptase during cDNA synthesis.

mRNA secondary structure prediction

mRNA secondary structure of OPRM1 (reference sequence AY521028) was predicted using Mfold (Unix and Web versions 3.1.2) (Zuker, 2003). Custom-designed Perl scripts created new sequences representing every possible transition SNP in the coding region (n=1203), including A118G. Sequences were folded in an automated
manner with Mfold (Unix Version). The maximum number-of-folding parameter was set based on the optimum of the wild type sequence (N=19). For a given RNA sequence, the resulting single-strandedness counts (ss-counts) measured the number of times each nucleotide was unpaired across all predicted secondary structures (Zuker and Jacobson, 1998). We calculated the change in ss-counts at each nucleotide in each set of SNP structures relative to wild type structures (1203^2 differences). Differences in ss-counts for each SNP sequence were analyzed regionally (in nucleotide windows surrounding the SNP) and globally (across the full sequence) as a measure of the relative amount of structural variation predicted to occur in response to a given SNP. Full-length RNA structures for A118, G118, C118 and T118 were predicted using the Mfold web server. All structures (optimal and sub-optimal) were visualized, and heuristics were considered to determine motif changes in the A118G region.

**OPRM1 expression constructs and mutagenesis**

A 1.2 kb fragment containing the human OPRM1 coding region cDNA was subcloned into pcDNA3 vector (Invitrogen, CA) using KpnI and XhoI sites. PCR was used to introduce the mutations at position 118 using the following mutagenesis primers (the mutated sites are in bold): A118G (T, C), 5’ CACTTAGATGG CG (T, C) ACCTGTCCGAC 3’. A 473bp section of OPRM1 was amplified with PCR using the mutagenesis primers and a common reverse primer (5’
CGGGATCCAGTTGCAGACATTGA 3’, with a *BamHI* site). The generated 473bp fragments, together with a common forward primer (5’ GGGGTACCCGATGGACAGCA-GCGCT 3’, with a *KpnI* site) were used to amplify a 581 bp 5'-fragment of *OPRM1*. Digestion with *KpnI* and *BamHI* and ligation into a 3'-fragment from wild-type *OPRM1* in *pcDNA3* vector yielded full length coding regions. For each construct, 2 clones were isolated. All constructs were confirmed by DNA sequencing.

**Cell culture and transient transfection**

CHO-K1 cells were cultured in F-12 nutrient medium (Invitrogen, CA) supplement with 10% fetal bovine serum, 100 U/ml penicillin and 100 µg/ml streptomycin at 37°C with 5% CO₂. Twenty-four hours before transfection, cells were seeded into 6-well or 12-well dishes. Transfections were performed using lipofectamine 2000 reagent (Invitrogen, CA) according to manufacturer’s protocol.

**mRNA stability assay**

Twenty-four hours after transfection, actinomycin D (ActD) (10 µg/ml) was added to inhibit transcription, and the cells were collected 1, 2, 6, and 12 hours after treatment. Control cells were cultured for the same time period without ActD treatment. For plasmid DNA preparation, cells were trypsinized and collected. Plasmid DNA was
prepared by QIAGEN DNA mini prep kit. For RNA preparation, cells were lysed with TRIzol reagent and prepared as described above. *OPRM1* mRNA was quantified by real-time PCR using PCR amplification primers as described above for SNaPshot.

**OPRM1 binding assay**

CHO cells in 12-well plates (poly-lysine coated) were transfected with 2 µg wild type or variant *OPRM1* plasmids. Twenty-four hours after transfection, cells were washed twice with phosphate buffer saline (PBS), incubated with 2 nM ³H-diprenorphine (Amersham Pharmacia Biotech, Piscataway, NJ, 50Ci/ mmol) in Tris-HCL buffer (50 mM, pH=7.5) for 1 hour at 37°C. Then, cells were washed three times with cold PBS and dissolved with 1N NaOH. The cell lysates were transferred to scintillation vials and radioactivity was determined with a LS-6500 Multi-purpose scintillation counter (Beckman Coulter, CA). Nonspecific binding was determined by adding naloxone (20 µM) before adding ³H-diprenorphine, which was subtracted from total binding.

**Western blotting analysis**

Twenty-four hours after the transfection of wild type and variant *OPRM1* into CHO cells plated in 6-well plates, cells were dissolved in 250 µl 2×SDS loading buffer (62.5 mM Tris-HCL, pH=6.8, 2% SDS, 25% glycerol, 0.01% Bromophenol Blue, 5% 2-mercaptoethanol (Bio-Rad Laboratories, Inc.,  *Western blotting analysis*— Twenty-four
hours after the transfection of wild type and variant *OPRM1* into CHO cells plated in 6-well plates, cells were dissolved in 250 µl 2×SDS loading buffer (62.5 mM Tris-HCL, pH=6.8, 2% SDS, 25% glycerol, 0.01% Bromophenol Blue, 5% 2-mercaptoethanol (Bio-Rad Laboratories, Inc., CA)), and incubated in 95°C for 10 minutes. Samples (40 µl) were run on an 8% SDS-polyacrylamide gel. Western blotting was performed using rabbit antibodies directed against the OPRM1 C-terminal (Incstar, Stillwater, MN), followed by detection with ECL Hyperfilm kit (Amersham Pharmacia Biotech, NJ). The membrane was then stripped with stripping buffer (100mM 2-mercaptoethonal, 2% SDS, and 62.5 mM Tris-HCL, pH=6.7), followed by incubation with mouse anti-α-tubulin antibodies (Oncogene Research Products, CA) and detection with ECL Hyperfilm kit.

**Statistical analyses**

Data are expressed as mean ± SD. Statistical analysis was performed using the program, GraphPad Prism version 3.00 for Windows, GraphPad Software (San Diego, CA).
2.3 Results

Genotyping results

We have genotyped 87 samples for A118G and C17T using SNaPshot reaction. For SNP A118G, 76 were AA, (87.4%), 8 were heterozygous AG (9.2%), and 3 were GG (3.4%). For SNP C17T, only one heterozygote was detected, all others were CC. The C17T heterozygote was homozygous for A118. Allelic frequencies of A118G were at the lower end of previously observed frequencies in various ethnic populations (2.5 % - 14%) (Bond et al., 1998; Luo et al., 2003).

Allelic expression of OPRM1 in human brain, using A118G as the indicator SNP

We identified 8 heterozygous samples for OPRM1 A118G, suitable for allelic expression analysis. Genomic DNA and total RNA were isolated from brain tissues, and mRNA was reverse-transcribed to cDNA using a gene-specific primer and oligo[dT]. cDNA synthesis using only oligo[dT] failed in brain tissue extracts because the coding region is far removed from the polyA tail of OPRM1, and because of partial mRNA degradation. After PCR amplification of the cDNA generated from the gene-specific primer (located in proximity of A118) and oligo[dT], we used SNaPshot to measure the peak area ratios of A118/G118 for genomic DNA and cDNA. Control experiments with mixtures of plasmids carrying A118 and G118-OPRM1 demonstrated that the measured ratios were proportional to amounts of each plasmid DNA added (Pearson analysis...
Genomic DNA ratios did not differ significantly between samples and were normalized to 1 (Le Hellard et al., 2002; Yan and Zhou, 2004). The SD was 0.13, indicating that the assay can detect ratio differences of ~25% (2X SD) (Table 2.1). mRNA ratios were determined in duplicate from 3 independent cDNA samples. After normalization to the genomic DNA ratio, the $A118/G118$ ratios for mRNA (or cDNA) ranged from 1.5-2.5, significantly different from the DNA ratios (P<0.01) (Table 2.1, Figure 2.2a). Repeat of the same experiment using a reverse extension primer, causing incorporation of different dideoxy terminators to test for fluorescent dye artifacts, gave similar results, showing all 8 samples have significant allelic ratio differences between cDNA and DNA (Table 2.1, Figure 2.2a). This result indicates that the level of mRNA from the wild-type $A118$ allele yielded approximately two-fold higher mRNA levels than the $G118$ allele in heterozygous brain tissues.

Next, we performed SNaPshot to determine the allelic mRNA ratio with $C17T$ as indicator SNP. Even though there was only one heterozygous sample (#9), the assay permits statistical analysis between the two alleles. The mRNA expression ratio for $C/T$ was 1.10±0.04 (normalized by DNA ratios), not significantly different from genomic DNA (Table 2.1b, Figure 2.2a). Because the heterozygous $C17T$ sample is homozygous for $A118$, this single result rules out any contribution of $C17T$, but implicates further $A118G$ as playing a role in the allelic expression imbalance.
RT-PCR of OPRM1 mRNA in brain samples

We separated the samples into two groups according to brain areas, i.e., cortical lobe and pons section samples (Stanley Foundation). Each group consisted of three genotypes for A118G (cortical sample group: AA, n=5, AG, n=3, and GG, n=2; Stanley pons section samples: AA, n=8, AG, n=4, and GG, n=1). Real-time PCR was performed to amplify the cDNA using OPRM1 and β-actin amplification primers separately, and cycle thresholds were used to calculate the mRNA expression. The overall mRNA expression of OPRM1 was lower in pons section samples, as indicated by the ratios of OPRM1/β-actin (Figure 2.2b & 2c). For each group, G118 homozygotes displayed less expression than A118 (Figure 2b & 2c). However, the relatively small number of samples, particularly of G118 homozygotes, precluded statistical analysis. Therefore, this experiment merely suggests a trend towards lower G118 mRNA levels.

mRNA expression of OPRM1 variants in transiently transfected CHO cells

We next tested whether transient expression of vectors containing only the coding region also show different mRNA expression for OPRM1-A118 and G118. Equal amounts of OPRM1-A118 and -G118 plasmids were co-transfected into CHO cells, and total RNA and plasmid DNA isolated. Real-time PCR revealed no amplification of DNA and mRNA for hamster OPRM1 in untransfected CHO cells, with cycle thresholds > 30 cycles, similar to no-template control. The same control experiments were done to
exclude the possibility of DNA contamination in the analysis of cDNA using samples without adding reverse transcriptase, again with negative results. As in brain tissues, the ratios of \textit{A118/G118} in plasmid DNA and mRNA (cDNA) were determined with SNaPshot. After normalization to the plasmid DNA ratio measured at each time point, the mRNA ratio of \textit{A/G} was determined to be 1.52±0.03 (n=6), obtained from 3 independent experiments, and 2 different \textit{OPRM1} clones for each construct. This was significantly higher than the DNA ratio at 24 hrs after transfection (P<0.01) (Figure 2.3a), again indicating that \textit{OPRM1-A118} yielded higher mRNA expression than \textit{OPRM1-G118}, even though the plasmids contained only the coding region of \textit{OPRM1}.

To determine whether the functional change is specific to the \textit{A118G} transition, we constructed \textit{OPRM1} expression vectors with \textit{A118C} and \textit{A118T} substitutions. CHO cells were co-transfected with equal amounts of \textit{OPRM1-A118} and either \textit{OPRM1-C118} or \textit{OPRM1-T118}. Twenty-four hours after transfection, DNA and total RNA were isolated, and the ratios of \textit{A118/C118} and \textit{A118/T118} were measured in DNA and RNA samples. Control experiments with mixtures of two plasmids (\textit{A118+C118} or \textit{A118+T118} with 5 different ratios for A/C or A/T, 2:8, 4:6, 5:5, 6:4, and 8:2) showed the peak area ratios of \textit{A118/C118} and \textit{A118/T118} measured by SNaPshot were proportional to the amount of plasmid DNA added. (A/C: Pearson r=0.997, p<0.001; A/T: Pearson r=0.998, p<0.0001) (Figure 1b, c). Measured DNA peak area ratios for A/C and A/T differed from unity (A/C: 3.82±0.2, n=4; A/T: 0.35±0.05, n=4) (Table 2.2), owing to different
incorporation efficiency and fluorescence yield of different dyes (the ABI3730 uses only one excitation wave-length). After normalization to the DNA ratios (adjusted to unity), the mRNA (cDNA) ratios of \(A118/C118\) and \(A118/T118\) did not differ significantly from that in DNA samples, in contrast to the \(A118/G118\) ratio (Figure 2.3a). This demonstrates that the effect at position 118 on \(OPRM1\) mRNA expression is specific to \(G118\).

**mRNA turnover of \(OPRM1-A118\) and \(OPRM1- G118\)**

To test the mRNA turnover rate of \(OPRM1\), cells transfected with wild-type \(OPRM1\) were treated with 10 µg/ml ActD for different time periods, and the amount of \(OPRM1\) mRNA was determined with real-time PCR. \(OPRM1\) mRNA levels declined after ActD treatment with a half-life of \(~1.2\) hrs (Figure 2.3b), determined by non-linear regression analysis. To test whether \(OPRM1-A118\) and -\(G118\) mRNA decay differently, we treated cells with ActD 24 hrs after co-transfection with same amounts of \(OPRM1-A118\) and -\(G118\) plasmids, then collected the cells and isolated plasmid DNA and total RNA at different times after treatment (0, 1, 2, 6, and 12 hrs). cDNA ratios were determined with SNaPshot. If the differential expression between \(OPRM1-A118\) and -\(G118\) were associated with altered mRNA stability and turnover, one would expect to see increasing \(A118/G118\) ratios after ActD treatment. However, compared to untreated RNA
samples, ActD treatment did not increase the ratio of A118/G118 in mRNA (compared to untreated control, P> 0.05, t-tests) (Figure 2.3c). Therefore, altered mRNA turnover does not appear to account for the differential expression of OPRM1-A118 and -G118.

OPRM1 protein expression measured with receptor binding and Western blotting

Twenty-four hours after transfection with OPRM1 plasmids carrying the four nucleotides substitutions at position 118, receptor binding sites were measured with ³H-diprenorphine (2 nM). Shown in Figure 4a, ³H-diprenorphine binding was significantly reduced only for the G118 variant (P<0.01) (G118: 13.5±2.4 fmol/mg; A118: 174±33 fmol/mg; C118: 182±40 fmol/mg; T118, 161±49 fmol/mg). The more than tenfold difference in protein level between OPRM1-Asp40 (G118) and three other OPRM1 variants with different nucleotides at position 118 (A, C, T) is considerably greater than the observed mRNA ratio (1.5-fold). This result is consistent with results on protein expression obtained with Western blotting. With similar total protein loading (represented by α-tubulin, Figure 2.4b), OPRM1-Asp40 yielded barely detectable expression, in contrast to the other three variants (Figure 2.4b). To determine whether OPRM1-G118 has dominant effects on the expression or degradation of the expressed protein, we measured ³H-diprenorphine binding in CHO cells co-transfected with OPRM1-A118 and -G118. Co-transfected cells displayed approximately half the tracer
binding (137±27 fmole/mg) compared to cells transfected with wild-type OPRM1-A118 only (251±51 fmole/mg) (Figure 2.4c), while OPRM1-G118 transfected cells again showed less than 10% of binding.

**mRNA secondary structure predictions**

Using mRNA secondary structure predictions with Mfold (Zuker, 2003), we modeled predicted changes in base-pairing behavior, for all possible transitions in each nucleotide position of OPRM1, by calculating ss-count differences from wild-type both locally (varying nucleotide windows around each SNP) and globally (across all nucleotides). The results showed G118 trending toward greater predicted structural difference than the average transition (3 nt local difference = 29, 3 nt local difference mean = 15.1 ± 11.7, skew = 0.73; global value=1593, global value mean = 1218 ± 628, skew = 0.22). Although this approach can reveal SNP positions predicted to be more likely to alter structure, it has limited descriptiveness because it employs linear windows to describe secondary changes. Thus, examination of other heuristics as well as visualization of optimal and suboptimal structures is necessary at positions of interest (Zuker and Jacobson, 1998). Doing such analysis for variations at the OPRM1-G118 position indicated a well-predicted helix (nt 102-107; 114-120) in the G118 mRNA that did not exist in predictions for A118, C118, and T118 structures (Figure 2.5). Moreover, optimal and suboptimal structures for the A118, C118 and T118 mRNA commonly
contained a loop motif in this region that did not appear in any predicted G118 structures (Figure 2.5). These results indicate that the G118 transition has the potential to affect local folding, and hence mRNA functions.

2.4 Discussion

The present study documents a significant functional change caused by the substitution of A118 with G118 in OPRM1, a common SNP in human populations. This substitution appears to affect at least two distinct mechanisms, involving mRNA expression in human autopsy brain tissues and in transfected cells, and translation into functional protein, observed in vitro. The collection of brain tissue samples analyzed was not designed to permit comparative analysis of OPRM1 protein content, because of the quality of the tissues, the different anatomical locations, and the rather low number of homozygous G118 samples. Allelic mRNA ratio analysis was successful because we utilized a gene-specific primer close to the allelic site for cDNA generation. For OPRM1 protein analysis as a function of A118G genotype, a larger prospective trial with anatomically well-defined brain sections will be required.

Allele-specific expression analysis revealed a two-fold difference in expression of the wild-type A118 variant over the G118 variant in autopsy brain samples of heterozygous individuals. Because environmental factors are canceled out by measuring allelic mRNA expression ratios, with one allele serving as the control for the other in
individual target tissues, this result provides compelling evidence for the existence of a cis-acting factor determining mRNA levels. We also developed a SNaPshot assay for *OPRM1-C17T* as an indicator SNP. Even though only one brain tissue was heterozygous for *C17T*, this single sample was informative because it was homozygous for *A118*. Lack of allelic expression imbalance in this sample indicates that *C17T* does not contribute, and moreover, is consistent with the notion that *A118G* is causative.

We also attempted to quantify the absolute levels of *OPRM1* mRNA, known to vary between different brain areas (Peckys and Landwehrmeyer, 1999), presumably via trans-acting regulatory factors. Real-time PCR results yielded highly variable levels for samples taken from different cortical lobes. To minimize the effects of tissue-specific expression, we also used samples from a well-defined region (pons) and observed less variation within genotypes. This analysis suggested a trend in mRNA levels in the following order AA > AG > GG, but the sample size was insufficient for statistical analysis, in consideration of the quality of the available tissue samples, and additional variability introduced by potential trans-acting factors. These results demonstrate that the analysis of allelic expression imbalance can be performed successfully in small numbers of target tissues because one allele serves as the control for the other. Therefore, the analytical SNaPshot procedure must be repeated independently so that statistical analysis can be performed for each individual sample.
To permit interpretation of clinical association studies, it is critically important to identify the functional polymorphism underlying the observed allelic expression imbalance of *OPRM1*. Because each of the 8 heterozygous samples measured showed a similarly elevated *A118/G118* ratio in human brain tissues, the functional polymorphism must be in linkage disequilibrium with *A118G*, or *A118G* itself is responsible for this functional difference. Previous genotype and haplotype studies had failed to detect any known candidate SNP of sufficient frequency in linkage disequilibrium with *A118G* that could account for this result. Therefore, we tested the role of *A118G* in transcription and translation, regardless of its proposed effect on β-endorphin ligand binding and G protein coupling (Bond et al., 1998). To study the mechanism for the allelic expression imbalance, we constructed expression vectors carrying only the coding region of *OPRM1*, excluding effects of untranslated regions and promoter regions, with A, G, T, and C in position 118. Measurements of allelic mRNA ratios after co-transfection of *OPRM1-A118* with the *G118, T118* and *C118* variants revealed that only the *A118/G118* ratio differed significantly from unity. This ratio is somewhat lower than what had been observed in brain tissue, probably reflecting different tissues, and confirming that the A to G transition affects mRNA levels. Inserting *C118* or *T118* into *OPRM1* failed to affect mRNA expression compared to the wild-type *A118*. These results demonstrate that substitution with *G118* alone, in the absence of any other regulatory regions, causes a significant change in mRNA expression. Altered mRNA stability and turnover could
have accounted for these results, as demonstrated for a synonymous SNP, C957T, in the D2 dopamine receptor (Duan et al., 2003). However, analysis of mRNA decay after treatment with ActD failed to reveal any significant difference in turnover between the A118 and G118 variants. Therefore, mRNA stability does not appear to cause differential mRNA expression. Alternatively, the A118G substitution could have affected transcription or processing of heterogeneous nuclear RNA (hnRNA) into mature mRNA (Hirota et al., 2004). Polymorphisms could induce mRNA decay at several steps during hnRNA maturation and pioneer protein synthesis in the nucleus (Arraiano and Maquat, 2003; Byers, 2002; Maquat, 2004). Further studies are required to clarify the underlying mechanisms of the observed allelic mRNA expression imbalance.

To determine whether A118G affects translation and receptor protein yield, we measured opioid receptor binding sites for the A118 and G118 OPRM1 variants expressed in CHO cells. The wild type OPRM1-Asn40 (A118) yielded tenfold more binding sites than OPRM1-Asp40 (G118). Results from Western blotting also demonstrated a much lower protein yield for OPRM1-Asp40, indicating that the low number of binding sites was not caused by changes in tracer affinity (3H-diprenorphine was added at saturating concentrations). Since this allelic difference is substantially greater than the allelic mRNA ratio, the G118 substitution, substituting Asn with Asp, appears also to affect translation, or post-translational processing and turnover of OPRM1 protein. Similarly, Beyer et al. had reported considerably greater protein expression after
stable transfection of OPRM1-A118 compared to -G118 in HEK293 cell clones (~7-fold) but attributed this difference to variations in selecting single clones (Beyer et al., 2004). The two other possible substitutions at position 118 (A→C or A→T) failed to affect protein expression and binding, similar to what was observed with mRNA expression. These results support the notion that the nonsynonymous SNP A118G, changing the codon from AAC to GAC, has a substantial effect on both mRNA and protein levels of OPRM1. Whether translation is affected by A118G, or the amino acid substitution Asn40 to Asp40 affects protein turnover, remains to be clarified. Since opioid receptors are known to dimerize (Wang et al., 2005b), we further tested whether cotransfection of wild-type with the G118 variant would affect protein levels by conveying instability to a potential heterodimer formed between them. As the functional protein levels (measured with tracer binding) of cotransfected cells were exactly in-between those of A118 and G118 transfected cells, we were unable to detect any negative dominant effects of G118 in this experiment, indicating that the defect either affects translation, or a potential heterodimer has normal stability.

We were unable to ascertain whether in human brain tissues, allelic differences in OPRM1 protein levels are more pronounced than would have been expected solely from the twofold difference in allelic mRNA levels. This requires analysis of a sufficient
number of homozygous tissues of the same brain region for \textit{OPRM1-A118} and \textit{-G118}. \textit{In vivo} image analysis of OPRM1 expression (for example, PET) would be a useful tool to test these results in human brain.

Previous studies have suggested possible mechanisms affecting translation. Specific sequences in exons can control translation through formation of hairpin structures as shown for the angiotensin type 1a receptor (Zhang et al., 2004). Predicted differences in mRNA structure as a result of SNPs have been associated with altered RNA and protein levels (Duan et al., 2003). For \textit{OPRM1}, the \textit{A118G} transition could similarly change the secondary structure of the mRNA and affect its maturation or translation. Our analysis of global and local mRNA structures of wild-type and variant \textit{OPRM1} indicate that the \textit{A118G} transition likely causes secondary structure alteration, which could have resulted in the observed difference in transcription, mRNA maturation, and translation.

While this study documents strong effects of \textit{A118G} on OPRM1 receptor functions, we have confirmed an allelic imbalance only at the level of mRNA in human brain tissues. Nevertheless, the magnitude of the functional changes observed \textit{in vitro} and \textit{in vivo} suggests that the \textit{A118G} SNP is relevant to OPRM1 function in human subjects. Because we were able to assess allelic expression in only a limited number of subjects in the present study, we cannot exclude the possibility that additional regulatory polymorphisms may exist that contribute to genetic differences in OPRM1 expression.
Taken alone, clinical association studies using single polymorphisms as indicators are often confounded by the contributions of multiple additional factors in different subject populations. While the evidence has been growing that A118G is associated with substance abuse and response to naltrexone therapy (Bart et al., 2004; Oslin et al., 2003; Szeto et al., 2001; Tan et al., 2003), our finding of a clear functional effect strengthens the conclusion that A118G does play a role in susceptibility to substance abuse. However, previous clinical association studies showing association of G118 with multiple drug addiction and treatment outcomes, assuming a gain of function mechanism, which opposite to our findings; therefore, the interpretation of physiological implications needs to be revisited, from a previously suggested gain of function (Bond et al., 1998) to a possible loss of OPRM1 function as shown here. While this research clarifies some aspect of the molecular genetic mechanism associated with A118G in OPRM1, significant question remain unsolved. Specifically, it will be essential to measure OPRM1 expression and function in human brain tissues, in relationship to genotype. It will further be important to link the molecular genetic mechanisms to neurophysiological processes and then clinical associations of A118G, to better understand how genetic variations affect drug addition and response to therapy.
Table 2.1. Peak area ratios for mRNA and DNA for heterozygous brain samples analyzed by SNaPshot. a. Peak area ratios for mRNA and DNA for 8 \textit{A118G} heterozygous brain samples analyzed by SNaPshot. Measured DNA peak area ratios differ from unity, reflecting different fluorophore incorporation and fluorescence yield. Reverse extension primers were used to confirm the data obtained with forward primer. When using the reverse primer, the observed peak area ratios are for \textit{T/C}, corresponding to the ratios of \textit{A/G}. All measured data are then normalized, by dividing observed RNA ratios with average DNA ratios of all DNA samples, assuming a 1:1 allele ratio in genomic DNA. Data are mean ± SD. b. Peak area ratios for mRNA and DNA for one \textit{C17T} heterozygous brain sample analyzed by SNaPshot. Observed and normalized ratios are shown as described in 1a. Data are mean ± SD.

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>DNA (n=3)</th>
<th>mRNA (cDNA) (n=6)</th>
<th>DNA</th>
<th>mRNA (cDNA) (n=3)</th>
<th>DNA</th>
<th>mRNA (cDNA) (n=3)</th>
<th>DNA Ratios</th>
<th>mRNA (cDNA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.03±0.25</td>
<td>2.13±0.41</td>
<td>0.96±0.27</td>
<td>2.02±0.38</td>
<td>3.59±0.40</td>
<td>5.44±0.30</td>
<td>1.03±0.11</td>
<td>1.56±0.30</td>
</tr>
<tr>
<td>2</td>
<td>2.22±0.33</td>
<td>2.33±0.27</td>
<td>1.13±0.36</td>
<td>2.18±0.25</td>
<td>3.59±0.16</td>
<td>9.22±0.60</td>
<td>1.03±0.05</td>
<td>2.78±0.17</td>
</tr>
<tr>
<td>3</td>
<td>3.26±0.26</td>
<td>2.72±0.23</td>
<td>1.17±0.17</td>
<td>2.53±0.21</td>
<td>3.27±0.09</td>
<td>8.96±0.04</td>
<td>0.93±0.03</td>
<td>2.56±0.01</td>
</tr>
<tr>
<td>4</td>
<td>4.00±0.11</td>
<td>2.14±0.02</td>
<td>0.92±0.12</td>
<td>1.98±0.01</td>
<td>3.40±0.35</td>
<td>7.17±0.28</td>
<td>0.97±0.10</td>
<td>2.05±0.08</td>
</tr>
<tr>
<td>5</td>
<td>5.08±0.21</td>
<td>1.56±0.48</td>
<td>0.81±0.22</td>
<td>1.51±0.44</td>
<td>3.34±0.16</td>
<td>6.52±0.18</td>
<td>0.95±0.05</td>
<td>1.86±0.05</td>
</tr>
<tr>
<td>6</td>
<td>6.21±0.09</td>
<td>1.65±0.40</td>
<td>1.12±0.08</td>
<td>1.57±0.37</td>
<td>3.54±0.60</td>
<td>6.42±1.01</td>
<td>1.01±0.17</td>
<td>1.83±0.29</td>
</tr>
<tr>
<td>7</td>
<td>7.04±0.13</td>
<td>1.86±0.01</td>
<td>0.96±0.14</td>
<td>1.72±0.00</td>
<td>4.45±0.13</td>
<td>5.40±0.35</td>
<td>1.27±0.04</td>
<td>1.54±0.10</td>
</tr>
<tr>
<td>8</td>
<td>8.22±0.07</td>
<td>1.93±0.21</td>
<td>1.13±0.08</td>
<td>1.80±0.19</td>
<td>3.10±0.18</td>
<td>5.43±0.51</td>
<td>0.89±0.05</td>
<td>1.55±0.15</td>
</tr>
<tr>
<td>Mean</td>
<td>1.11</td>
<td>1.00</td>
<td>3.50</td>
<td>1.00</td>
<td>5.43</td>
<td>1.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>0.13</td>
<td>0.13</td>
<td>0.18</td>
<td>0.12</td>
<td>0.18</td>
<td>0.12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### C/T peak area ratios (C/T)

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Observed ratios</th>
<th>Normalized ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA (n=3)</td>
<td>mRNA (cDNA) (n=3)</td>
<td>DNA</td>
</tr>
<tr>
<td>9 1.03±0.03</td>
<td>1.13±0.04</td>
<td>1.00±0.03</td>
</tr>
</tbody>
</table>
Table 2.2. Peak area ratios for mRNA and plasmid DNA from CHO cells transfected with equal amount of A118-OPRM1 and one of the three mutant constructs (G118-, C118- and T118-OPRM1). The observed DNA ratios deviated significantly from unity for A/C and A/T, because of differences in fluorescence yield and incorporation efficiency. The results were normalized as described in Table 1. Data are mean ± SD.

<table>
<thead>
<tr>
<th></th>
<th>Observed ratios</th>
<th>Normalized ratios</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DNA (n=4)</td>
<td>mRNA (cDNA) (n=6)</td>
</tr>
<tr>
<td>A/G</td>
<td>0.94±0.02</td>
<td>1.52±0.03</td>
</tr>
<tr>
<td>A/C</td>
<td>3.82±0.21</td>
<td>3.94±0.38</td>
</tr>
<tr>
<td>A/T</td>
<td>0.35±0.05</td>
<td>0.29±0.04</td>
</tr>
</tbody>
</table>
Figure 2.1. Correlation between fluorescence peak area measured by SNaPshot and the ratios of added plasmid DNAs. Different amounts of \textit{A118-OPRM1} plasmid were mixed with \textit{G118-OPRM1} (Panel a), \textit{C118-OPRM1} (Panel b), or \textit{T118-OPRM1} (Panel c) to obtain ratios of 2:8, 4:6, 5:5, 6:4, and 8:2. The plasmid mixtures were amplified and detected with SNaPshot using different nucleotides at position 118 as indicator. The measured relative amounts of each \textit{OPRM1} variant present in the mixture were expressed as the peak area ratio of \textit{A/G}, \textit{A/C} or \textit{A/T}. The peak area ratios for each combination were plotted against the plasmid DNA ratio. Pearson r for \textit{A/G}, \textit{A/C} and \textit{A/T} were 0.992, P<0.001, 0.997, P<0.001, and 0.998, P<0.0001, respectively. Data are Mean ± SD, n=3.
Plasmid DNA ratios (A118/G118)

Plasmid DNA ratios (A118/C118)

Plasmid DNA ratios (A118/T118)
Figure 2.2. mRNA expression for OPRM1 with different genotypes. Panel a. Allele specific mRNA expression of OPRM1 in heterozygous human brain samples using indicator SNPs A118G (8 samples) and C17T (1 sample). SNaPshot was employed to measure peak areas of A over G, or C over T alleles in both genomic DNA and cDNA samples. cDNA ratios were normalized to the ratio for genomic DNA. Three independent cDNA samples were prepared and SNaPshot assays performed twice for each cDNA preparation. Data are mean ± SD. Both forward and reverse extension primer were used for A118G as a marker. Compared to the DNA ratio, cDNA ratios are significantly higher, P<0.01, in 8 A118G heterozygote samples (both primers) but not in the single C17T heterozygous sample, P>0.05 (one-way ANOVA with Dunnett’s post test). Panels b & c. mRNA expression for OPRM1 in samples with different genotypes. Real-time PCR was applied to quantify mRNA, with β-actin as an internal control. Data are mean ± SD. n=5, 3, 2 for AA, AG, and GG in cortical lobes samples; n=8, 4, and 1 for pons section samples.
**A118G or C17T mRNA allele ratios (A/G or C/T)**

**Cortical lobes**

OPRM1 mRNA level (OPRM1/β-actin)

**Pons section**

OPRM1 mRNA level (OPRM1/β-actin)
Figure 2.3. mRNA expression of OPRM1 constructs in transfected CHO cells. Panel a. Differential mRNA expression of OPRM1 variants in transiently transfected CHO cells. Equal amounts of OPRM1 plasmids containing A118 were co-transfected into CHO cells with plasmids containing G118, C118, or T118. Cells were collected at 24 hrs after transfection and plasmid DNA and total RNA isolated. The ratios of A118/G118, A118/C118 or A118/T118 in cDNA samples were measured with SNaPshot and normalized with the respective plasmid DNA ratios (Table 2). Data are from three independent experiments with duplicate cDNA preparation and SNaPShot assays. Data are mean ± SD. Compared to DNA ratio, ** P<0.01 (one-way ANOVA with Bonferroni post test). Panel b. Time course of OPRM1 mRNA levels after actinomycin D (10 µg/ml) treatment. Amount of mRNA remaining after ActD is shown as % control, yielding a half-life of 1.2 hrs (one-phase exponential decay); data are mean ± SEM, n=4. Panel c. mRNA ratios of A118-OPRM1 /G118-OPRM1 at different times after ActD treatment. 24 hrs after transfecting equal amounts of A118-OPRM1 and G118-OPRM1 into CHO cells, ActD (10µg/ml) was added, and RNA/cDNA and plasmids DNA were isolated at different time point (1h, 2h, 6h, 12h). The mRNA ratios were detected by SNaPShot, normalized to DNA ratios measured at each time point. Data are mean ± SD, n=4.
Panel a: mRNA (cDNA) ratios

Panel b: OPRM1 mRNA level (% of control) over time after ActD treatment (h)

Panel c: mRNA ratio (A118G/118-OPRM1) over time after ActD treatment (h)
Figure 2.4. Protein expression for different constructs of OPRM1 in CHO transfected cells. Panel a. OPRM1 expression tested with $^3$H-diprenorphine binding. Cells were transfected with wild type and variant OPRM1 expression vectors, carrying OPRM1-A118, -G118, -C118, and -T118. 24 hrs after transfection, $^3$H-diprenorphine (2 nM) binding was determined in intact cells. Data are Mean ± SD, n=4. Compared to A118, ** P<0.01 (One-way ANOVA with Dunnett’s post test). Panel b. OPRM1 protein expression detected by Western blotting. As an internal loading control, α-tubulin was also detected with Western blotting. Panel c. Co-transfected OPRM1 expression detected with $^3$H-diprenorphine binding. Cells were transfected with OPRM1-A118, OPRM1-G118 or co-transfected with equal amounts of these two constructs. 24 hrs after transfection, $^3$H-diprenorphine (2 nM) binding was determined in intact cells. Data are Mean ± SD, n=3. Compared to A118, * P<0.05, ** P<0.01 (One-way ANOVA with Dunnett’s post test).
**Figure 1:**

(a) 

![Bar graph showing 3H-Diprenorphine binding (fmol/mg) for A, G, C, T genotypes.](#)

(b) 

![Bar graph showing 3H-Diprenorphine binding (fmol/mg) for AA, AG, GG genotypes.](#)

(c) 

![Western blot showing 75 kD protein band labeled OPRM.](#)
Figure 2.5. Mfold analysis to predict secondary structures for 4 possible substitutions at position 118 in human OPRM1 mRNA. Each variant sequence generated 1 optimal and 16 sub-optimal secondary structures. G118 displays a consistent 6bp helix in all but one of its predicted structures. The most common predicted motif in this region in all other structures (A118, C118, and T118) is an internal loop, while the 6bp helix is not observed.
CHAPTER 3

Frequent polymorphisms in the human dopamine DRD2 receptor affect gene expression and splicing

3.1 Introduction

The G-protein coupled dopamine D2 receptor (DRD2) is a key target of antipsychotic drugs (GPCR) (Creese et al., 1976; Seeman et al., 1976). Aberrant DRD2 signaling has been linked to mental disorders including drug addiction (Di Chiara and Imperato, 1988; Wise and Bozarth, 1987), schizophrenia, and Parkinson’s diseases (Seeman, 2006; Starr, 1995; Takahashi et al., 2006). Extensive genetic studies have associated a single nucleotide polymorphism (SNP) termed Taq1A with schizophrenia (Dubertret et al., 2004), alcoholism (Hallikainen et al., 2003), heroin abuse (Lawford et al., 2000), and cigarette craving (Erblich et al., 2005). Located 10kb downstream of DRD2 in a distinct gene, Taq1A was suggested to be in linkage disequilibrium with a functional DRD2 variant. Moreover, a promoter region variant (-141C Ins/Del),
shown to affect DRD2 transcription \textit{in vitro} (Arinami et al., 1997), has been associated with schizophrenia (Arinami et al., 1997), heroin addiction in a Chinese Han population (Li et al., 2002), and alcoholism in Mexican-Americans (Konishi et al., 2004), but its \textit{in vivo} relevance remains to be established. Similarly, the \textit{in vivo} role remains uncertain for a synonymous SNP (C957T) in exon 7 of \textit{DRD2}, shown to affect mRNA stability of DRD2 in vitro (Duan et al., 2003). Here, we address the question whether functional genetic variants affect DRD2 activity in human brain.

Because the \textit{DRD2} locus lacks frequent nonsynonomous SNPs that could affect receptor protein function, we searched for regulatory polymorphisms that affect gene transcription and mRNA processing and splicing. Two main splice isoforms of DRD2 are detectable in human post-mortem brain tissues, a long form (D2 long) with an insertion of 29 amino acids (exon6), and a short form (D2 short) lacking exon 6 (a rare longer form (D2 longer) (Liu et al., 2000) was not detected in this study). These splice isoforms display distinct physiological and pharmacological characteristics (Castro and Strange, 1993; Guiramand et al., 1995; Khan et al., 1998; Malmberg et al., 1993; Usiello et al., 2000). Tissue-specific expression of splice isoforms indicates \textit{trans} regulation (Lorson et al., 1999; Rau et al., 2006), but except this, cis-acting factors could also play a role. We have used allelic expression analysis to determine whether one allele produces different levels, or different splice variants, of DRD2 in the same individual, using autopsy tissues from human prefrontal cortex and striatum. Detection of allelic
expression imbalance (AEI) indicates the presence of cis-acting regulatory factors, a powerful tool in search of genetic variants (Lim et al., 2006; Pinsonneault et al., 2006; Wang et al., 2005a; Zhang et al., 2005). I demonstrate the presence one frequent promoter SNP and two intronic SNPs that affects DRD2 splicing, whereas the previously proposed polymorphisms had no apparent or only limited effects.

The in vivo role of these novel functional SNPs was then tested in human subjects, using a working memory paradigm. The ventral striatal complex forms a crossroads between the cortex and basal ganglia structures, receiving substantial input from the dopaminergic transmitter system. Within the cortico-striato-thalamo-cortical circuitry, the striatum is specifically involved in cognitive processes (Alexander et al., 1986; Graybiel, 1997) such as working memory (WM) (Levy et al., 1997; Lewis et al., 2004; Owen et al., 1996; Postle and D'Esposito, 1999; Postle and D'Esposito, 2003). Consistent with known neuroanatomic projections, the caudate and the pallidum within the basal ganglia are more specifically involved in WM processing (Levy et al., 1997; Postle and D'Esposito, 1999), integrating perceptual, mnemonic, and affective information along the cortico-striato-thalamo-cortical re-entrant system (Newman and Grace, 1999). A primary purpose of striatal circuitry may be to increase the signal-to-noise ratio in the cortex, facilitating binding and suppressing cortical activity not directly contributing to the focus of working memory. Dopamine is a major modulator within these structures, affecting cortical glutamate signals and impinging upon striatal medium spiny neurons (Surmeier et al.,
Whereas D2 receptor signaling reduces dendritic excitability at striatal neurons, D1 signaling exerts the opposite effect (West and Grace, 2002). Genetic or pharmacological manipulation of striatal D2 receptors is associated with deteriorated working memory performance in animal models (Arnsten et al., 1995; Kellendonk et al., 2006), while radiotracer imaging techniques in humans have reported tight relationships between striatal D2 receptor availability and working memory (for review, see (Cropley et al., 2006)). Consistent with these results, D2 agonists and antagonists modulate working memory performance in humans (for review, see (Mehta and Riedel, 2006)).

Therefore, we hypothesized that genetic variation affecting expression of DRD2, and specifically the presynaptically expressed DRD2S autoreceptor (Khan et al., 1998; Usiello et al., 2000), might be associated with differential striatal activity measured with fMRI during working memory, potentially reflecting variation in neuronal excitability. Because the net effect of dopamine in the striatum is to increase activity of thalamo-cortical pathways (Tisch et al., 2004), we also hypothesized that variation in DRD2 expression in the striatum would alter response of the thalamo-prefrontal pathway.
3.2 Material and methods

**Postmortem human brain tissues**

105 DNA and RNA samples, extracted from prefrontal cortex autopsy tissues, were obtained from The Stanley Medical Research Institute's brain collection, courtesy of Drs. Michael B. Knable, E. Fuller Torrey, Maree J. Webster, and Robert H. Yolken (Chevy Chase, MD). Among them, 35 are from bipolar disorder patients, 35 are from schizophrenia patients and 35 are from control population. The average post-mortem interval (PMI) for those samples is 32.9 hrs, ages are from 19 to 64 years old, races for samples population are: 97.1% are white (n = 102), 0.95% are black (n = 1), 0.95% are Hispanic (n = 1), and 0.95% are Native Americans (n = 1). Twenty-five frozen striatum brain samples were from the Brain and Tissue Bank for Developmental Disorders at the University of Maryland (Baltimore, MD). All of them are from control Caucasian population, ages from 18 – 53 years old, with PMI less than 16 hrs.

**Genomic DNA and RNA isolation, and cDNA synthesis**

As described previously (Zhang et al., 2005), for DNA preparation, frozen tissues were treated with sucrose-Triton followed by proteinase K digestion overnight, and then sodium chloride precipitation of protein and ethanol precipitation of DNA. RNA was isolated with Trizol and purified with Qiagen RNeasy columns. 1 µg total RNA then was treated with DNase I in 37°C for 15 minutes, followed by cDNA synthesis using
reverse transcriptase II (Invitrogen, Carlsbad, CA) with gene specific primers (located close to the marker SNPs used for AEI analysis) and Oligo (dT), and control reactions without reverse transcriptase were set up to test for any DNA contamination.

**Genotyping with GC-clamp (allele-specific PCR) and SNaPshot assay**

To pick up SNPs in different linkage blocks so that the coverage of genetic information of the gene will be increased, it is essential to know the linkage disequilibrium blocks (LD) of DRD2. Right now, a powerful tool for examining linkage blocks is HapMap, an international project to develop a haplotype map of the human genome (Hinds et al., 2005). It is freely accessible and describes the common patterns of human DNA sequence variations. Figure 3.1b shows a haplotype map for *DRD2* in CEU population (Utah residents with ancestry from northern and western Europe), the different colors represent the D prime scores (D’) between SNP pairs with red representing a higher D’ and more LD between two SNPs. Choosing SNPs from different LDs along the gene increases the likelihood of identifying a SNP that is highly associated with the allelic expression imbalance of *DRD2*.

Spanning the DRD2 locus, 19 SNPs were genotyped in initial step, and an additional intronic SNP (rs2283265) was genotyped in the second stage for 37 Stanley samples. Three SNPs (rs12364283, rs1076560 and rs2283265) were genotyped for DNA samples (n = 100) from University of Bari (Bari, Italy)(Table 3.1). Of these, 15 were
analyzed with a GC-clamp assay described previously (Papp et al., 2003) and 4 were genotyped by SNaPshot assay (three marker SNPs, rs6275, rs6277, and rs6279; one Ins/Del polymorphism rs1799732) (Pinsonneault et al., 2004; Zhang et al., 2005). Briefly, allele-specific primers were designed with a GC-clamp added to one of the primers, and following polymerase chain reactions, melting curves of the amplicons were analyzed to produce genotyping data for each SNP. The specificity of allele-specific PCR depends on the allele-specific primers that are designed to amplify only one specific allele over the other (Papp et al., 2003). By adding GC-rich sequence to one of the allele-specific primer, the melting temperatures for two PCR products differ from each other by 3-5 degrees, which can be detected easily using DNA melting analysis with fluorescent SYBR Green. To discriminate the two PCR amplification products from the amplification of only one allele over the other, 10-15 bases of GC rich sequence is added onto the 5’ end of one allele-specific primer to increase the melting temperature of one amplicon (up to 5°C). SYBR green I dye binds in the minor groove of double-stranded DNA and fluorescence are enhanced dramatically after binding. Increasing the amount of PCR products will subsequently increase fluorescence during the amplification. As the temperature rises to the melting temperature (Tm), at which the DNA strands are denatured and become single stranded, the fluorescence intensity drops, which results in the maximal fluorescence change. After 30-40 cycles of PCR, fluorescence melting curve analysis is performed, which displays the releasing rate of SYBR green fluorescence from the
double-stranded products by plotting the intensity of fluorescence against temperature. This allows clear discrimination between two alleles because the Tm values for two allele-specific PCR primers are differentiated by adding of GC-rich sequences in the 5’ end of one primer. Primers used for the assay were listed in table 3.2a.

Determining DRD2 mRNA levels by real-time RT-PCR

RT-PCR was performed with β-actin as an internal control. 1 µl of cDNA (~50ng) from each sample, mixed with 200nM primers (forward and reverse) and 7.5 µl 2× power PCR SYBR master mix (3mM MgCl$_2$, 200µM dATP, 200µM dCTP, 200µM dGTP, 200 µM dUTP, 0.625U AmpliTaq Gold, 0.25U AmpErase UNG), was amplified using an ABI 7000 (Pinsonneault et al., 2006). The amplification program consists of one single cycle of 95 °C to denature the DNA strands, followed by 30 cycles of 95°C for 15s, 60°C for 1min to anneal primers and extend the new complement strand. Cycle thresholds of DRD2 and β-actin were compared to test relative mRNA expression levels of DRD2.

Allele-specific expression measurements (SNaPshot)

The detection of the allelic mRNA expression relies on the indicator or marker SNP. Two criteria should be considered to choose an indicator SNP. First, the SNP should be frequent (>10%), so we have enough heterozygous samples. According to Hardy-Weinberg equation, the estimated heterozygosity is equal to 1-(p$^2$+q$^2$) (p, q
are frequencies for two alleles). Therefore the number of heterozygous samples depends on the frequency of the SNP. The second criteria is to have the SNP located in transcribed region, including 5’, 3’ untranslated region and coding region, since we need to detect the allele expression for mature mRNA. Two SNPs in exon 7, rs6275 and rs6277 (C957T), and one SNP in 3’UTR region, rs6279, with minor allele frequency arranged from 0.28 – 0.46, were selected as indicator SNPs for DRD2. There are two reasons for choosing more than one indicator SNP, one is to replicate the AEI measurements using SNPs belonging to different regions of mRNA so that bias rose from one SNP could be detected, the other reason is to increase the heterozygous sample pool to better serve statistical analysis, supporting the validity of the allelic expression analysis.

SNaPshot analysis was conducted as described before (Yan et al., 2002; Zhang et al., 2005). Three marker SNPs located in transcribed regions (rs6277 and rs6275 in exon 7, and rs6279 in 3’UTR.) were used for measuring allelic ratios of genomic DNA and mRNA (after conversion to cDNA). In brief, a fragment of DNA or cDNA (generated by reverse transcriptase) flanking the polymorphic site (∼ 100 bp) was amplified by PCR with forward primer, 5’ CCAGCTGACTCTCCCCGAC 3’ (rs6277 and rs6275), 5’ AGCCTGAGTCAGGGCCC 3’ (rs6279) and reverse primer, 5’ GCATGCCCATTTCTCTCTGG 3’ (rs6277 and rs6275) and 5’ ACCGCCTGCTCCACG 3’ (rs6279) (Table 3.2a) with the following program: one cycle of 95 °C for 2 minutes, 30
cycles of 95°C, 15 seconds, 60°C, 1 minute, and 72 °C, 1 minute, and then final extension at 72 °C for 3 minutes. 2 µl of the PCR products purified by 5 units of Antarctic Phosphatase and 2 units of ExoI (New England Biolabs) were added into the primer extension reaction to add one fluorescently labeled deoxyribonucleoside triphosphate (ddNTP) complementary to the polymorphic site to the 3’ end of the primer. After running capillary electrophoresis (ABI 3730), the extension primers labeled with distinct fluorescent dyes were analyzed by Gene Map software (Applied Biosystems, Foster City, CA). Peak area ratios were calculated to represent relative amounts of the two alleles in genomic DNA and cDNA. Allelic DNA ratios varied within a narrow range (0.96 ± 0.05) were normalized to unity and allelic cDNA ratios adjusted accordingly to account for differential fluorescent yield and incorporation rates for each dye. To confirm the SNaPshot results from one marker SNP, additional marker SNPs in the transcribed region of DRD2 were chosen, and the resultant allelic mRNA ratios were compared in individuals where at least two marker SNPs were heterozygous.

**Splice isoform detection using fluorescently-labeled primers and splice-specific PCR**

We measured relative expression of three DRD2 splice isoforms (DRD2L, DRD2S, DRD2longer) in prefrontal cortex and striatum brain tissues. Fluorescent-dye (Fam) labeled forward primer 5’ ACATTGTCCTCCGCAGACG 3’ (in exon 5) and
reverse primer (in exon 7) 5' GCATGCCCATTCTTCTCTGG 3' (Table 3.2b) located in common regions of DRD2L and DRD2S were used to amplify DRD2L, DRD2S, and DRD2longer (Wang et al., 2006). Relative expression of the three isoforms was measured by capillary electrophoresis (ABI3730). To test PCR amplification efficiency of those isoforms, real-time PCR was used to amplify cDNA constructs of long or short isoforms separately (Figure 3.2a). Mixtures with different ratios of the splice variants were used to construct standard curve for calculating relative abundances (Figure 3.2b).

To determine whether allelic expression differs between the splice variants, forward primers specific to long (5’ GCTCCACTAAAGGCAACTGTA 3’) or short isoform (5’ TGAGGGCTCCACTAAAGGAGGC 3’) were designed to amplify only one isoform. Specificity of the primers was tested with DRD2L and DRD2S cDNA constructs (DRD2L from UMR cDNA Resource Center, Rolla, MO). Under the PCR condition of initial denaturing at 95 °C for 5 min, then 30 or 23 cycles for samples from prefrontal cortex or striatum, at 95°C for 5 min (within linear amplification range for samples), 60 °C for 1 min, and 72 °C for 1 min, and final extension step of 72 °C for 5 min, the splice-specific PCR products were generated using either Taq DNA polymerase (Invitrogen, Carlsbad, CA) for DRD2L or Sigma PCR master mix (Sigma, MO) for DRD2S, using the same PCR conditions as for SNaPShot (Table 3.2a). Primer extension reaction and capillary electrophoresis then generates two peaks, representing the relative mRNA expression of the two alleles in each splice forms.
Linkage disequilibrium (LD) and haplotype analysis

Expressed by D’ score, LD between each pair of SNPs, and the main haplotypes were calculated by HelixTree (Golden Helix, Inc., Bozeman, MT) (Lambert, 2004), a commercially available genetic statistical analysis program that utilizes Formal Inference-Based Recursive Modeling (FIRM). HelixTree is capable of scanning across large SNP maps to locate markers that correlate with a phenotype/response/disease status (binary and continuous response variables are supported). A Two-Loci genetic plot feature calculates the statistical significance of associations of pairs of genetic markers with the response variable(s).

Reporter gene assays with promoter constructs of different length

We constructed several promoter fragments in a reporter gene plasmid to test for regulatory activity of rs12364283 in vitro. One common reverse primer 5’ GAAGATCTTCGGGGCAGAGACGGCGCCGGCTGCTT 3’ with a Bgl II cutting site, and three forward primers harboring a Kpn I cutting site, D2upstreamFS: 5’ GGGGTACCCCACTGGCGAGACGGTGAGGACCC 3’, D2upstreamFM: 5’ GGGGTACCCCCTGGGCAGGGTAGCAGCGGAACACC 3’, and D2upstreamFL: 5’ GGGGTACCCCCTTCACAGCACCTGTTTAAGCCTCAGT 3’ (Table 3.2b) were designed to amplify three pieces of DNA fragments (D2Pro_S, D2Pro_M, and D2Pro_L)
with different lengths of promoter regions (D2Pro_S, –283 to +292 as used by (Arinami et al., 1997), D2Pro_M, –600 to +292, and D2Pro_L, –963 to +292, numbers are calculated from the transcription start site (+1), (Gandelman et al., 1991)). Genomic DNA from two heterozygous samples of rs12364283 but homozygous for all other SNPs within the amplified regions was amplified by platinum Taq DNA polymerase high fidelity and 0.5 × enhancer (Invitrogen, Carlsbad, CA) in a total volume of 20 µl. PCR conditions were determined to be one cycle at 95 °C for 5 minutes, 30 cycles at 95°C, 30 seconds, 62°C, 1 minute, and 72 °C, 2 minutes, and then the final extension at 72 °C for 5 minutes. Purified DNA fragments were inserted into PGL3_basic vector upstream of luciferase gene using Kpn I and Bgl II cloning sites (Promega Biosciences, San Luis Obispo, CA).

Cell culture and transient transfection

Human embryonic kidney cells (HEK-293) and SH-SY5Y were cultured in DMEM/F12 (50/50 mix) media supplemented with 10% fetal bovine serum (FBS), penicillin (10 units/ml), streptomycin (10 µg/ml) and incubated in 37 °C with 5% CO₂. For report gene assay, 24-hours before transfection, 1 × 10⁵ (HEK) or 2 × 10⁵ (SH-SY5Y) cells were planted into 24-well plates in DMEM/F12 media, and transient transfection was performed using FuGENE HD Transfection Reagent (Roche Applied Science, Indianapolis, IN) with 1: 3 ratios of DNA (0.5 µg) and reagent (1.5 µl) per well in serum free medium for 5 hours. As a transfection control, Renilla luciferase constructs were
cotransfected with PGL3 fused constructs with a ratio of 1:20. Forty-eight hours later, cells were harvested and transferred into a 96-well plate, luciferase activity was detected by Dual-Glo luciferase Assay system (Promega, Madison, WI) on a fluorescence plate reader (PerkinElmer Life and Analytical Sciences, Shelton, CT). For both cell lines, three independent transfections and duplicate luciferase assays for each construct have been done. For minigene constructs, $5 \times 10^5$ HEK cells were planted into each well for a 6-well plate before transfection. The transfection was done with the same reagent as we did for reporter gene assay except that 2 µg DNA and 6 µl transfection reagent were added to each well.

**MiniGene constructs making and detecting for alternative splicing in transfected cells**

To generate the mini-gene construct, forward and reverse primers with a HindIII and XhoI cutting site respectively were designed to amplify a DNA fragment including exon 5, intron 5, exon 6, intron 6 and exon 7 (~3kb) from genomic DNA that is heterozygous for two intronic SNPs (rs2283265 and rs1076560). HindE5F: 5’ CCCAAGCTTACCAGAACGAGTGCATCATTGCC 3’; XhoE7R: 5’ CCGCTCGAGCGAGAACAATGGCGAGCATCTCTGA 3’. The transfection was done as described above, except that 6-well plate was used. Platinum high fidelity Taq polymerase (Invitrogen, Carlsbad, CA) was used under PCR program of one cycle at 95
°C for 2 minutes, 29 cycles at 95°C, 15 seconds, 62°C, 1 minute, and 72 °C, 3.5 minutes, and then the final extension at 72 °C for 10 minutes (Table 3.2b). The amplified PCR products then were inserted into pcDNA3 vector, downstream of a T7 promoter.

Two constructs carrying two haplotypes (one allele carrying C (major) for both intronic SNPs, the other carrying A (minor) for both intronic SNPs) for DRD2 minigene (exon 5, intron 5, exon 6, intron 6, and exon 7) were transfected into HEK cells separately. 24 hours and 48 hours after transfection, cells were collected and total RNA were isolated with Trizol (Invitrogen, Carlsbad, CA). For cDNA synthesis, gene specific primer SP6 (5’ CATTTAGGTGACACTATAG 3’) was added to avoid the synthesis of endogenous DRD2 cDNA. Fam-labeled forward primer (5’ ACATTTGTCCCTCCGCAGACG 3’) and SP6 were added to amplify both long and short isoforms produced from the minigene contracts in transfected cells, and the relative expression of long and short forms was determined as described in previous section.

Clinical studies

The association of mental disorders with SNPs in DRD2

Subjects

105 DNA and RNA samples, extracted from prefrontal cortex autopsy tissues, were obtained from The Stanley Medical Research Institute's brain collection, courtesy of Drs. Michael B. Knable, E. Fuller Torrey, Maree J. Webster, and Robert H. Yolken
(Chevy Chase, MD). Among them, 35 (17 males and 18 females) are from bipolar disorder patients, 35 (26 males and 9 females) are from Schizophrenia patients and 35 (26 males and 9 females) are from control population. The average post-mortem interval (PMI) for those samples is 32.9 hrs, ages are from 19 to 64 years old, races for samples population are 103 white, one black, and one Native Americans.

The association of cognition and memory with SNPs in DRD2 (collaborated with Dr. Bertolino from University of Bari)

Subjects

To examine the effect of genotypes on working memory associated brain activity independent of sample size, demographic or behavioral variation, we selected subjects to control for these variables.

Effects of rs1076560 genotype were studied in a sample of forty-four subjects: twenty-two CC [12 males; mean age ± SD: 30 ± 9, IQ: 109.9 ± 14.6, socioeconomic status index: 36 ± 13.4 (Hollingshead and Redlich, 1958), Edinburgh handedness score: 0.8 ± 0.2 (Oldfield, 1971)] and twenty-two CA [12 males; mean age ± SD: 28.8 ± 7.2, IQ: 113.9 ± 15, socioeconomic status index: 41 ± 17.7, Edinburgh handedness score: 0.7 ± 0.4].
Effects of rs12364283 genotype were evaluated in a sample of thirty-four subjects: seventeen AA [10 males; mean age ± SD: 30 ± 9.1, IQ: 115.2 ± 10.7, socioeconomic status index: 35.1 ± 17.4, Edinburgh handedness score: 0.8 ± 0.4] and seventeen AG [10 males, 7 females; mean age ± SD: 29.5 ± 7.1, IQ: 115.1 ± 14.2, socioeconomic status index: 35.3 ± 15.4, Edinburgh handedness score: 0.8 ± 0.3].

Thirty-four subjects were studied to evaluate potential effect of rs2283265 genotype: seventeen C/C [9 males, 8 females; mean ± SD: 30.5 ± 10.1, IQ: 111.2 ± 14.1, socioeconomic status index: 35.5 ± 14.3, Edinburgh handedness score: 0.8 ± 0.3] and seventeen CA [9 males, 8 females; mean age ± SD: 29.8 ± 7.7, IQ: 110.5 ± 14.3, socioeconomic status index: 39.3 ± 18.7, Edinburgh handedness score: 0.7 ± 0.4]. Subjects heterozygote for rs1076560 and rs2283265 are substantially the same, but for five of them. All subjects had normal or corrected-to-normal visual acuity. Exclusion criteria were presence of any neurological or psychiatric disorders, presence of any pharmacological treatment or medical condition that might have influence cerebral metabolism or blood flow, presence of drug abuse and past head trauma with loss of consciousness.

All subjects gave written informed consent to the study after the procedure was fully explained to them. The protocol was approved by the local IRB (Comitato Etico Locale Indipendente Azienda Ospedaliera”Ospedale Policlinico Consorziale” Bari)
**Working Memory task**

During fMRI, all subjects completed a blocked paradigm of the N-Back task (Bertolino et al., 2006a; Bertolino et al., 2004). Briefly, “N-back” refers to how far back in the sequence of stimuli the subject had to recall. The stimuli consisted of numbers (1-4) shown in random sequence and displayed at the points of a diamond-shaped box. There was a visually paced motor task which also served as a non-memory guided control condition (0-Back) that presented the same stimuli, but simply required subjects to identify the stimulus currently seen. In the working memory condition, the task required the recollection of a stimulus seen two stimuli (2-Back) previously while continuing to encode additionally incoming stimuli. Performance data were recorded as the number of correct responses (accuracy) and as reaction time.

**Acquisition of fMRI data**

Each subject was scanned using a GE Signa 3T scanner with a standard head-coil (Milwaukee, WI). Echo planar imaging BOLD fMRI data were acquired as described previously (TE=30 msec, TR= 2 seconds, 20 contiguous slices, voxel dimensions=3.75x3.75x5 mm) (Bertolino et al., 2004). We used a simple block design in which each block consisted of eight alternating 0-Back and 2-Back conditions (each
lasting 30 seconds), obtained in 4 min and 8 sec, 120 whole-brain scans. The first four scans at the beginning of the time series were acquired to allow the signal to reach a steady state and were not included in the final analysis.

**Demographic and behavioral data analysis**

One-way ANOVAs and $\chi^2$ statistical tests were used to evaluate the effects of genotype groups on demographics as well as on behavioral performance at the N-back (accuracy and response time).

**fMRI data analysis**

Data analysis was performed using SPM2 (http://www.fil.ion.ucl.ac.uk/spm/software/spm2). All fMRI data were reconstructed, registered, linear detrended, globally normalized, and then smoothed (10 mm Gaussian kernel) before analysis within SPM2. fMRI data were analyzed as a time series modeled by a sine wave shifted by an estimate of the hemodynamic response. Individual subject maps were created using $t$ statistics (2-Back>0-Back). These individual contrast images were then used in second-level random effects models to determine task-specific regional responses at the group-level with one-sample t-tests (main effects of task). As we were not interested in differences in anatomical areas that were not activated in the main effect of task, we restricted the subsequent second level random effects analysis to only areas...
that were activated during the task. To facilitate this, a functional mask was created by using the activation maps from 2-Back>0-Back contrasts (p<0.05, k=3) limiting the analysis to the working memory cortical and subcortical network. Therefore, this procedure controls for the possibility that potential differences between the groups arise from areas that are engaged by only one of the groups. Using this mask, separate ANOVAs with genotype at the different three SNPs as a grouping factor was performed on 2-back > 0-back contrasts. Because of our strong \textit{a priori} hypothesis regarding the differential response of striatal regions and of the working memory cortical network and our use of a rigorous random effects statistical model, we chose a statistical threshold of p<0.001, k=3. Moreover, since areas within the working memory cortical network represented \textit{a priori} regions of interest, we corrected for multiple comparisons the statistical threshold with a FWE small volume correction (using a 10mm radius sphere centered on prefrontal and cingulate coordinates published in previous studies, p = 0.01 (Bertolino et al., 2006a; Bertolino et al., 2004; Bertolino et al., 2006b; Callicott et al., 2000; Callicott et al., 1999; Callicott et al., 2003; Chang et al., 2007; Postle and D'Esposito, 1999; Postle and D'Esposito, 2003; Postle et al., 2000). Statistically significant group differences were reported as voxel-intensity z values. For anatomical localization, statistical maxima of activation were converted to conform to the standard space of Talairach and Tournoux (Talairach and Tournoux, 1998).
Statistics

Allelic expression imbalance (AEI) was determined by comparing normalized cDNA ratios (peak area ratios of cDNA/ mean of the peak area ratios of DNA) with a cut-off value of 1.2 (~20% difference of mRNA ratios from DNA ratios) as the reliably detectable minimal allelic expression ratio. Student t-test, were used to test for statistical significance for deviation from unity in the mRNA ratios (presence of AEI). Reproducibility of allelic ratios obtained with two indicator SNPs (rs6275 and rs6279; rs6277 and rs6279) was assessed with Pearson correlation analysis in heterozygous samples for both SNP pairs. Fisher’s Exact tests were conducted for association between genotype status (heterozygous or homozygous) with AEI. False Discovery Rates were controlled using the method of Benjamini and Hochberg (Benjamini, 1995). To test the possibility of more than one SNP contributing to the observed AEI (epistasis), 2-loci P values (adjusted p value considering the effects of multiple test) were calculated for the combination of any two SNPs using Helix Tree software. To test the association of SNPs with differences in allelic ratios between the splice variants DRD2S and DRD2L – indicative of a genetic factor in splicing – we used the ratios of allelic expression ratios for DRD2L (AEI (L)) versus the ratios for DRD2S (AEI (S)) as the phenotype for SNP scanning along the DRD2 locus in the same fashion as described for allelic ratios of the total mRNA. We first tested whether the allelic ratios differed significantly between
DRD2S and L in the same subject, and then used the presence or absence of any difference for SNP scanning. Multiple regressions were conducted. The interaction effects between rs12364283 and rs1799732, however, were not significant.

3.3 Results

**Allelic expression imbalance (AEI) in human autopsy brain samples**

To measure mRNA allele-specific expression, 68 heterozygous samples (54 from prefrontal cortex, and 14 from striatum) for at least one of the marker SNPs, rs6275, rs6277 and rs6279, were chosen and tested for AEI. Comparing allelic ratios measured with different indicator SNPs, Pearson correlation analysis revealed excellent correlations (Fig. 3.3b & c), rs6277 *versus* rs6279, Person r = 0.9626, p < 0.01; rs6275 *versus* rs6279, person r = 0.931, p < 0.0001. Shown in Fig. 3.3a, 8 tissues displayed clearly detectable AEI with ratios above unity (the main allele/over the minor allele: C/T for rs6277, A/G for rs6275, and C/G for rs6279), whereas 7 had ratios below unity. Marker SNP rs6277 had been reported to cause altered mRNA turnover rate (Duan et al., 2003) but is not associated with the observed allelic ratios, indicating the existence of other functional polymorphisms not linked with marker SNP (because allelic ratios were below and above unity).
Genotype analysis and SNP scanning of the DRD2 locus using allelic expression ratios

To search for polymorphisms accounting for AEI, we genotyped 19 polymorphisms in all samples (Table 3.1) to perform association studies with AEI serving as a phenotypic trait. Predicted haplotype frequency was calculated by Helix Tree (Table 3.3a), and D’ scores between each pair of SNPs are listed in table 3.3b. Consistent with previous studies, SNPs residing in the regions from intron 2 to 3’ downstream of DRD2 are highly linked, generating 4 major haplotypes with Freq > 4% (No. 1 – 4, table 3.3a). Haplotype blocks are more fragmented in intron 1 and 5’ upstream regions, producing several haplotypes with frequency from 1% - 2% (haplotypes 5-14, table 3.3a).

Association analysis between SNPs and with AEI ratios, using Fisher’s Exact tests, revealed that promoter region SNP (rs12364283) is highly associated with AEI (P = 0.001, multiple-comparison-adjusted P=0.019) (Table 3.4). To determine whether allelic ratios reflect the combined effects of multiple factors, we further used a two-loci linkage analysis with Helixtree (Golden Helix), which confirmed rs12364283 as the main contributor (adjusted P < 0.001, Fig. 3.4a). To test whether other SNPs can also contribute, samples heterozygous for rs12364283 were removed from the pool and the remainder reanalyzed. None of the other SNPs yielded a significant association with
allelic expression ratios, indicating that the results are consistent with only rs12364283 accounting significantly to the observed AEI (figure 3.4b). Marker SNP rs6277, although previously suggested to affect mRNA turnover, had no significant association.

**Relative expression of splice isoforms and splice isoform-specific allelic expression**

Three alternative splice isoforms of DRD2 have distinct distributions in brain regions and subcellular locations (Khan et al., 1998; Liu et al., 2000; Usiello et al., 2000). We used fluorescently labeled primers to measure relative expression of splice isoforms in prefrontal cortex and striatum. Comparable to previous studies (Khan et al., 1998), DRD2S mRNA expression higher in prefrontal cortex (DRD2S / (DRD2S + DRD2L) *100% = 22 ± 8%, n = 37) than in striatum (9.5 ± 6.1%, n = 25) (F = 42.50, p < 0.0001 (one way ANOVA, SPSS)) while D2L has the reverse trend with higher expression in striatum. No detectable D2 longer isoforms was found in our tissues from two regions. Differences in regional expression of the two forms are likely due to trans-acting regulation. To test the additional involvement of any cis-acting regulation of DRD2 splicing, we measured allele-specific expression separately for each splice form (DRD2L or DRD2S). A portion of the tissues displayed up to twofold differences in allelic ratios between short and long isoforms (figure 3.5a). Because we also knew the relative expression of each splice variant, we could calculate the expected allelic ratios for the previously determine ratios of combined mRNA (total AEI, figure 3.5a). The combined
allelic ratios calculated from each splice variant were indistinguishable from the measured combined ratios, providing independent validation for the accuracy of the assays.

Serving as a phenotype, the AEI differences between long and short were used to scan the \( DRD2 \) gene locus, showing association with several SNPs, rs7103679, rs1125394, rs2075654, rs1076560 and rs1800497 with rs1125394, rs2075654, and rs1076560 being the most significant associated SNPs (\( p < 0.0001, n = 37, \) Fig. 3.5b).

The location of rs1076560 in intron 6 suggests a role in alternative splicing of exon 6; however, full sequencing of the region spanning introns 5 and 6 identified another candidate SNP, rs2283265, in intron 5 completely linked with rs1076560 in 37 Stanley brain tissues tested (allele frequency 17%). To illustrate the impact of these two intronic SNPs on splicing, we divided the samples into two groups according to the genotype of rs1076560 and compared the AEI ratios between DRD2L and DRD2S for each sample (Fig. 3.6a). Each sample heterozygous for rs2283265/rs1076560 yielded significantly lower AEI ratios for DRD2S compared to DRD2L (Fig. 3.6a). Mean AEI ratios sorted by rs2283265/ rs1076560 genotype are shown in figure 3.6b. In subjects homozygous for these two intronic SNPs but heterozygous for marker SNP rs6277, we detect a small (10-20%) but insignificant increase of allelic expression for the main allele (C allele for rs6277), which is consistent with the notion that the marker SNP rs6277 affects mRNA turnover, but the effect is small and insignificant (Duan et al., 2003).
difference is observed between total \( DRD2 \) mRNA and the two splice variants. In contrast, heterozygous subjects for rs2283265 or rs1076560 display a significantly lower AEI ratio for total mRNA \((p < 0.05)\), indicating that the minor intronic alleles (linked to the main allele of the marker SNP) yield lower mRNA levels. This is almost entirely caused by a marked reduction in DRD2S formation when the minor intronic alleles are present. These results provide strong evidence that either of the two intronic SNPs, or both acting together, cause a shift of DRDR2 splicing from S to L, while in addition causing somewhat lower overall mRNA levels. Since rs2283265/rs1076560 are linked to the main allele of rs6277, this confounds the potential opposite effect of rs6277 on mRNA turnover (Duan et al., 2003).

Distinct allelic expression ratios between long and short splice forms could result from at least two mechanisms: one allele is less stable for one or both splice variants, or SNPs affect the splicing process. To test this further, we first compared the percentage of short and long between CC carries and CA + AA carriers of rs1076560 (one of the two implicated intronic SNPs) in both prefrontal cortex and striatum. Consistent with our AEI results, CA + AA carriers of rs1076560 has less relative mRNA expression for the short form than CC carriers, in both brain regions (prefrontal cortex, \( F = 18.70, p < 0.0001, n = 40 \); in striatum, \( F = 10.92, p = 0.003, n = 25 \) (One-way ANOVA)) (Fig. 3.6c). Real-time
RT-PCR did not reveal significant differences in total mRNA expression level (including both long and short forms) between different genotypes (data not shown), possibly a result of high variability in total levels, and/or equal turnover rates for DRD2S and L.

**Reporter gene assay to test the effects of promoter SNP rs12364283**

We cloned several DNA fragments of different lengths in the promoter region of *DRD2*, and inserted them into a luciferase reporter plasmid. Since a ‘GAA (AAA)’ repeat region located 3’ of rs12364283 (-806 ~ -629) contain multiple deletion/insertion polymorphisms, we designed PCR primers labeled with Fam fluorescent dye, to test variations in length of this region. While 11 repeat variants were found in 105 Stanley Samples (Table 3.5), these were not significantly correlated with allelic mRNA expression. To study further the effect of the repeat polymorphisms, we selected two homozygous samples for two variants, 361 and 364 (distances the PCR fragment migrate after electrophoresis, represent the length of each fragment), with 8 and 4 nucleotide deletions compared to reference haplotype 368 (UCSC genome browser, [http://www.genome.ucsc.edu](http://www.genome.ucsc.edu)). The promoter constructs were labeled PRO_S, PRO_M, and PRO_LC/T1 and 2, the latter two referring to repeat variants 360 and 364 (Fig. 3.7a)

As shown in Figure 3.7b, Pro_S and Pro_M have the highest promoter activities, while those four Pro_L fragments have less activities compared to Pro_S in both cell lines (P < 0.05 or 0.01, Dunnett’s multiple comparison tests), indicating a silencer
harbored in the region from –600 ~ -963. The C allele (minor) for rs12364283 apparently has significantly higher promoter activity than T allele in both cell lines tested (in HEK cells, Pro_LC1 (76% ± 10) vs. Pro_LT1 (41% ± 4), p < 0.001 (one-way ANOVA, Bonferroni’s Multiple comparison test), Pro_LC2 (72% ± 10) vs. Pro_LT2 (50% ± 9), p < 0.05; in SHSY5Y cells, Pro_LC1 (70% ± 12) vs. Pro_LT1 (52% ± 9), p < 0.05, Pro_LC2 (69% ± 10) vs. Pro_LT2 (49% ± 1), p < 0.05). This demonstrated a disinhibitory effect of the C allele on promoter activity. There were no significant differences between the repeat variants (PRO_L1 and 2) for both genotypes, suggesting that the repeat polymorphisms do not contribute to promoter activity.

**Minigene studies to determine the effects of intronic SNPs**

Two minigene constructs carrying major or minor allele for both intronic SNPs were transfected into HEK cells separately, and the ratios between two isoforms were calculated. Consistent with the findings in tissues, the construct carrying A allele for both intronic SNPs generated much less short isoforms than C allele constructs in transfected cells (Figure 3.8)(24 hrs: 0.021 ± 0.005 vs. 0.039 ± 0.004, p < 0.01; 48 hrs: 0.029 ± 0.009 vs. 0.057 ± 0.002, p < 0.01, student t-test). Furthermore, the expression of short isoforms in transfected HEK cells was lower than those in brain tissues, indicating tissue-specific regulation in splicing by *trans* or in the regions not included in our *DRD2* minigenes.
The association of promoter SNP rs12364283 with schizophrenia

In a first test to assess the clinical penetrance of rs12364283, we performed an association analysis using the 105 samples from the Stanley DNA collections (35 schizophrenic, 35 bipolar, and 35 control subjects). This revealed a significant association of rs12364283 with schizophrenia with minor allele carriers having much higher frequency in patient population ($\chi^2 = 6.89$, $P = 0.009$ (n = 95), odds ratio (AA/AG) = 0.442, 95% confidence interval, 0.274 – 0.712, table 3.6a & b). However, because of the small cohort size, independent clinical association studies should be conducted to test these results further.

The association of promoter and intronic SNPs with working memory in vivo

Demographics and behavioral data

ANOVAs showed no significant differences between genotype groups in any demographical variable (all $p>0.1$). ANOVAs on behavioral data did not show any statistically significant main effect of genotypes on accuracy and response time (all $p>0.1$).
fMRI data

Analysis of the working memory imaging data in the whole sample revealed significant BOLD responses in the working memory cortical and subcortical network, including dorsolateral prefrontal cortex (BA 9), anterior cingulate (BA 24 and BA 32), premotor area (BA 6), parietal cortex (BA 39/40), caudate, and putamen, consistent with earlier reports (Bertolino et al., 2006a; Bertolino et al., 2004; Callicott et al., 2000; Callicott et al., 1999; Callicott et al., 2003; Chang et al., 2007; Postle and D'Esposito, 1999; Postle and D'Esposito, 2003; Postle et al., 2000).

ANOVA of the fMRI data analysis showed a significant effect of rs1076560 genotype: the C/A genotype was associated with greater BOLD activity in several brain regions, including bilateral head of the caudate (Figure 3.9) (x,y,z: -19, 22, -4; Z= 4.10; k= 127; FWE p=0.001, and 19, 18, -7; Z= 3.95; k= 102; p=0.002; Figure 5, left middle frontal gyrus (-51, 34, 12; Brodmann’s Area, BA, 46, Z= 3.53; k= 6; p= 0.005, and BA 11; x,y,z: -30, 40, -2; Z= 3.53; k= 13; p= 0.007), left precentral gyrus (BA 6; -51, 2, 49; Z= 3.51; k= 12; p= 0.008), left anterior cingulate (BA 24; -11, 20, 22; Z= 3.49; k= 4; p= 0.008), left thalamus (-8, -11, 1; Z= 3.24; k= 3; p= 0.017), left superior frontal gyrus (BA 10; -30, 49, 22; Z= 3.21; k= 3; p= 0.019) and left caudate tail (-11, -20, 22; Z= 3.20; k= 7; p= 0.019). The opposite contrast (C/C > C/A) did not show any significant difference.
ANOVA also revealed a very similar effect of rs2283265 genotype: again the C/A genotype was associated with greater activity in left caudate body (x,y,z: -10, -17, 25; Z= 3.31; k= 8; FWE p=0.018), left claustrum (-22, 19, 9; Z= 3.71; k= 55; p= 0.005), left inferior frontal gyrus (BA 46; -51, 38, 12; Z= 3.42; k= 5; p= 0.011), left superior temporal gyrus (BA 22; -34, -50, 13; Z= 3.26; k= 3; p=0.018), right posterior cingulate (BA 23; 10 -24 22; Z= 3.36; k= 3 ; p= 0.013) and right insula (BA 13; 34 19 9 ; Z= 3.14; k= 3; p= 0.025 ) (data not shown). No significant difference was found on the inverse contrast (C/C > C/A).

ANOVA of the rs12364283 SNP did not indicate any statistically significant difference in any brain region.

3.4 Discussion

This study demonstrates the presence of frequent *cis*-acting regulatory polymorphisms in the gene encoding the dopamine receptor DRD2. Whereas earlier studies had implicated DRD2 variants in CNS disorders and suggested relevant polymorphisms, the functional variants and their impact in human brain remained uncertain. Our approach using allelic expression analyses of human brain autopsy tissue samples, followed by SNP scanning of the DRD2 locus, has revealed one upstream promoter polymorphism and a haplotype affecting DRD2 splicing. In contrast, previously suggested regulatory variants (-141Ins/Del), a SNP affecting mRNA turnover (rs6277),
and a marker SNP in an adjacent region (Taq1A) appear not to contribute significantly to DRD2 expression. In view of the central role of DRD2 in dopaminergic neurotransmission, these novel variants are likely to have clinical significance, which can now be assessed by genotyping in peripheral blood lymphocytes in large association studies.

The use of allelic expression imbalance has gained rapid acceptance as a powerful tool for detecting regulatory polymorphisms (Johnson et al., 2005). We have successfully identified frequent functional polymorphisms or haplotypes correlated to AEI for candidate genes implicated in CNS disorders, including OPRM1 (Zhang et al., 2005), MAOA (Pinsonneault et al., 2006), and TPH2 (Lim et al., 2007). In the present study, we have demonstrated the presence of significant and frequent AEI for DRD2 in 68 tissue samples from human prefrontal cortex and striatum. Use of three marker SNPs (rs6277, rs6275 and rs6279) validated the AEI results in independent assays. Detection limit for allelic mRNA expression ratios was 20-30% deviation from unity, normalized to genomic DNA.

The frequent marker SNPs rs6277 (C957T) had previously been shown to alter mRNA turnover rate in vitro (Duan et al., 2003); therefore, we had expected a bias of the allelic mRNA expression ratios in favor of the C allele, allegedly the more stable allele. However, no such bias was detectable (Fig.3.3a), and moreover, homozygous samples for rs6277 also displayed AEI. A small effect of SNP rs6277 on mRNA yield (10-20%) is
possible, but only in the context of the main haplotype lacking intronic SNPs rs2283265/rs1076560, the latter apparently overriding effects of rs6277. Taken together, these results indicate that any effect of rs6277 on mRNA expression level is marginal and may not have strong penetrance in vivo.

To search for functional polymorphism(s), we genotyped a total of 19 SNPs to scan the DRD2 locus. Haplotype analyses reveal a strong haplotype block ranging from intron 2 to 3’ UTR, and the 3’ downstream region, whereas linkage deteriorates in the 5’ upstream region (Table 3.3b). As a result, SNPs in the promoter region are not linked with the marker SNPs (Table 3.3b), representing putative candidate SNP that can account for the pattern of allelic expression ratios. Indeed, SNP rs12364283, located 844 bps upstream of the transcription start site, was significantly associated with AEI ratios, while all other variants failed to contribute to the phenotype. It is noted that rs12364283 is present in only 6 of 15 subjects (sample No. 2, 3, 5, 64, 65, and 68) showing significant AEI; therefore, additional factors not accounted for in this study - genetic and/or epigenetic - remain to be discovered.

Both rat and human DRD2 contain a promoter region ~300 bp upstream of the transcription start site (Minowa et al., 1992) (Arinami et al., 1997), but sequences further upstream can also regulate transcription, involving tissue specific expression or silencer domains (Kamakura et al., 1997; Minowa et al., 1992). To identify new regulatory elements of human DRD2, we used reporter gene assays revealing a repressor region
located -600 to -963 bp’s upstream, capable of inhibiting transcription from a previously tested promoter region (-283 to +292) (Pro_S) (Arinami et al., 1997). Importantly, the C allele (minor allele) of rs12364283 confers higher transcriptional activity compared to the T allele (major allele), indicating a disinhibition or gain function for the C allele. This effect was seen in two transfected cell lines, and since AEI ratios correlate with rs12364283 in two brain regions, the effects do not appear to be tissue selective. Sequences flanking rs12364283 (~ 20 bp) are highly conserved among different species (Figure 3.10), predicted to bind to several transcription factors, such as E47 (AT (C) CTGGC), ANF (GAATCTGGCAAA), NF-X3 (AGAATCTG), and HSF1 (long & short) (CACAGAAT) (TRANSFAC, version 8.3) (Farre et al., 2003; Messeguer et al., 2002). Importantly, the T to C change leads to a loss of predicted binding sites for ANF and HSF1 but generates a new putative site for AREB6 (AGAACCTG, dissimilarity, 7.59%). Detailed molecular studies are needed to resolve the nature of the regulatory events and the role of rs12364283, but a gain of function polymorphism is expected to have penetrant properties in vivo.

Another promoter SNP, rs1799732 (-141 In/Del), had been reported to cause altered mRNA transcription in vitro (Arinami et al., 1997); however, rs1799732 was not significantly associated with allelic expression in our study, suggesting low if any effect on transcription in vivo. A two-loci analysis (Helix Tree) and a multiple regression analysis to test any interaction of rs12364283 with rs1799732 also failed to yield
significant results. Possibly, the promoter region harboring -141 Ins/Del is masked by an upstream silencer, which may diminish the effects of this SNP at least in the tissues examined here. Moreover, SNPs rs1799732 and rs12364283 are highly linked (D’ = 0.972, Table 3.3b). Paradoxically, the -141 C Ins allele displayed higher promoter activity in vitro (Arinami et al., 1997) but it is almost always linked to the T allele of rs12364283, shown here to convey lower activity. This shows that the in vivo penetrance of any promoter SNP studied in vitro is difficult to predict without use of AEI results.

While it is possible that more complex interactions and further factors cooperate in regulating DRD2 transcription, we ascribe the main effect to rs12364283.

A second striking finding is the discovery of a novel haplotype of two intronic SNPs flanking exon 6, regulating DRD2 splicing. Alternative splicing mediated in trans can account for the observed region-specific expression of the two main splice variants (Khan et al., 1998), as observed here in prefrontal cortex and striatum. However, applying allelic expression analysis to each splice variant selectively, we demonstrate that cis-acting polymorphisms also affect the splicing process independent of region-specific regulation in trans. Exploiting the presence or absence of allelic expression differences between splice variants, we were able to scan the DRD2 locus, identifying two strongly associated SNPs in intronic regions of DRD2 flanking exon 6. Application of ESE finder (Cartegni et al., 2003) revealed that both rs1076560 and rs2283265 modulate distinct SRP protein binding sites, with the G allele generating a putative binding site for SRP55.
but none in rs1076560, and the T allele forming an SC35 (rs2283265) or SRP40 (rs1076560) binding site. To test the effects of these two SNPs, we generated two minigenes carrying major or minor alleles for both SNPs, and compared the relative expression of two splice isoforms in transfected cells. Comparable to our findings in tissues, minor allele produced much less short isoforms than major allele (2-folds). Moreover, our results indicate this kind of effects is sequence-specific rather than tissue-specific (cis-regulation) since HEK cells are not neuron-like cell type. On the other hand, the much less short forms produced from transfected HEK cells than from brain tissues indicate the existence of other regulations for DRD2 splicing harbored beyond minigene region. As rs1076560 and rs2283265 are tightly linked to each other, further studies need to be done to determine the effects of individual SNPs.

The finding of genetic variants affecting DRD2 splicing has potential clinical significance, considering the distinct physiological and pharmacological characters of DRD2S and L. In striatum with high expression level of DRD2, we have observed greater than twofold differences of DRD2S formation relative to DRD2L, as a function of genotype. As an autoreceptor in the presynaptic dopaminergic pathway, DRD2S controls the release of dopamine (Cragg and Greenfield, 1997), secretion of prolactin (Kelly et al., 1997), and dampens dopamine neurotransmission (Usiello et al., 2000). DRD2L on the other hand, mainly thought to be localized in the postsynaptic membrane, mediates haloperidol-induced catalepsy, and positively modulates or facilitates D1-dependent
reactions (Usiello et al., 2000), thereby exerting opposite effect on neurotransmission compared to DRD2S. Therefore, the relative expression levels between DRD2S and L are critical for normal neurophysiologic activities (Bozzi and Borrelli, 2006; Iaccarino et al., 2002).

To test the clinical relevance of functional SNPs rs12364283 (promoter), and rs1076560/rs2283265 (intronic), we performed an association study with cognitive processing in normal human subjects, using fMRI (by collaborating with Dr. Bertolino from the University of Bari, Italy). The results reveal robust associations between the two intronic SNPs with activity of the ventral striatum, thalamus, dorsolateral prefrontal cortex, and premotor cortex during working memory. Heterozygote subjects had greater activity in these brain regions despite similar accuracy and reaction time suggesting that these effects are not because of behavioral differences. On the other hand, the promoter SNP did not show statistically significant association with any of the above anatomical regions.

These results are consistent with the known role of dopamine in modulating the cortico-striato-thalamo-cortical network. An important part of this modulation is mediated via striatal DRD2 receptors; moreover, both splice forms of the DRD2 receptor are implicated in modulating GABA and glutamate release in the striatum (Centonze et al., 2003). Greater density of DRD2S provides greater GABA transmission in the striatum (Centonze et al., 2004). On the other hand, the ratio between DRD2S and
DRD2L is implicated in inhibition of glutamate release, the greater the ratio the greater the release (Centonze et al., 2004). Consistent with these electrophysiology experiments in rodents, our results demonstrate that the minor allele of the two intronic SNPs associated with reduced expression of DRD2S is also associated with greater activity in the human striatum during working memory. Our results also demonstrate greater activity in other regions of the working memory network, including the dorsolateral prefrontal cortex, parietal cortex, and thalamus. It is indeed possible that greater activity of the striatum commands greater activity in all other brain regions in the cortico-striato-thalamo-cortical network, as predicted by the net effect of increasing striatal activity in this loop (Tisch et al., 2004). In addition, because of the cis-acting splicing effects from those intronic SNPs, the trend of generating less DRD2S in heterozygote is likely also present in areas outside the striatum (Figure 3.6c).

Our findings suggesting *in vivo* penetrance of the two intronic SNPs affecting DRD2 splicing has potentially broad implications. DRD2 is one of the primary targets for typical and atypical antipsychotics. In particular, clozapine has higher affinity for DRD2S than DRD2L (Malmberg et al., 1993), implying that patients carrying different genotypes of rs1076560 or rs2283265 could potentially experience different treatment outcomes. The Taq1A A1 allele – being in linkage disequilibrium (D’ = 0.855) with the minor allele of the intronic SNPs - was found to be more frequent in non-responders than responders
under clozapine treatment (Hwang et al., 2005). Therefore, the A1 allele is partially linked to low DRD2S formation, possibly accounting for reduced sensitivity to clozapine treatment.

Representing the most intensely studied SNP, Taq1A (rs1800497) is located in an adjacent gene (Figure 3.1a) but was shown to be associated with reduced DRD2 density in vivo (Jonsson et al., 1999; Pohjalainen et al., 1998; Ritchie and Noble, 2003) and cocaine self-administration in both rhesus monkeys and humans (Nader et al., 2006; Volkow et al., 1993). Therefore, low DRD2 levels in striatum could be a predictor of high vulnerability to cocaine impulsivity (Dalley et al., 2007), and probably other addictive drugs. However, it is noted that Taq1A is in linkage equilibrium with the promoter SNP rs12364283 reported here, which likely confounds interpretation of the results. Several clinical studies also support the correlation of Taq1A with multiple drug abuse and other mental disorders (Dubertret et al., 2004; Hallikainen et al., 2003; Lawford et al., 2000). While the distant Taq1A is unlikely to affect DRD2 activity, our results support the view that these previous association results reflect strong linkage with rs1076560/rs2283265, suggesting that association studies with the latter could reveal stronger penetrance of DRD2 than previously recognized.

It is noted that some samples displayed AEI that was not linked to any specific polymorphisms; therefore, other factors including epigenetic regulation could also contribute to DRD2 expression. This could have resulted in lack of association between
the promoter SNP and striatal activation during working memory observed here. Alternatively, variation of overall expression mediated by a promoter variant could affect different phenotypes, such as drug addiction or mental disorders. While the Stanley Foundation collection is grouped into schizophrenics, bipolar and control subjects (35 schizophrenic, 35 bipolar, and 35 control subjects), and we did find significant associations of rs12364283 with schizophrenia ($\chi^2 = 6.89, P = 0.009 (n = 95)$, odds ratio (AA/AG) = 0.442, 95% confidence interval, 0.274 – 0.712, Table 3.6a & b). However, because of the small cohort size, I need to be cautious in the interpretation of this result, and independent clinical association studies should be conducted to test these results further.

This study reveals the presence of frequent functional polymorphisms in D2, affecting the dopamine neurotransmission in vivo. Different hypotheses were generated to account for the functions of genetic factors in complex diseases, including mental disorders. On the one hand, relatively rare variants in multiple genes may cooperate with many other variants to affect diseases; on the other hand, frequent variants in relatively fewer genes with high penetrance could also affect disease. Our results suggested that it is critical to study candidate genes in mental disorders for the presence of functional polymorphisms and resolve the underlying molecular mechanism. As shown in the example of D2, regulatory polymorphisms may remain hidden even to intense study over long times. The use of allelic expression analysis has proven a useful tool in the
discovery of regulatory polymorphisms. The implications of finding frequent functional polymorphisms go beyond the D2 receptor but may affect our understanding of additional key candidate genes that may also harbor hidden regulatory variants. Whether and to what extent frequent functional variants exist in these key genes, and whether they have been selected for in phenotypic evolution in the human lineage and could play a role in human disease, remains to be addressed.
<table>
<thead>
<tr>
<th>number</th>
<th>SNP</th>
<th>positions (genome)</th>
<th>regions (gene)</th>
<th>MAF Stanley F.</th>
<th>MAF Univ. Pari</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>rs10891556</td>
<td>G/T</td>
<td>5' upstream</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>rs12364283</td>
<td>A/G</td>
<td>5' upstream</td>
<td>0.07</td>
<td>0.09</td>
</tr>
<tr>
<td>3</td>
<td>rs1799978</td>
<td>A/G</td>
<td>promoter</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>rs1799732</td>
<td>C/-C</td>
<td>promoter</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>rs4938019</td>
<td>T/C</td>
<td>intron 1</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>rs4350392</td>
<td>C/A</td>
<td>intron 1</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>rs4648317</td>
<td>C/T</td>
<td>intron 1</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>rs4581480</td>
<td>T/C</td>
<td>intron 1</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>rs1257447</td>
<td>C/T</td>
<td>intron 1</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>rs4648318</td>
<td>A/G</td>
<td>intron 1</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>rs7125415</td>
<td>C/T</td>
<td>intron 1</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>rs7103679</td>
<td>C/T</td>
<td>intron 1</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>rs1125394</td>
<td>A/G</td>
<td>intron 1</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>rs2075654</td>
<td>G/A</td>
<td>intron 2</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td>15a</td>
<td>rs2283265</td>
<td>C/A</td>
<td>intron 5</td>
<td>0.17</td>
<td>0.10</td>
</tr>
<tr>
<td>15b</td>
<td>rs1076560</td>
<td>C/A</td>
<td>intron 6</td>
<td>0.17</td>
<td>0.12</td>
</tr>
<tr>
<td>16</td>
<td>rs6275</td>
<td>C/T</td>
<td>exon 7</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>rs6277</td>
<td>C/T</td>
<td>exon 7</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>rs6279</td>
<td>C/G</td>
<td>3' UTR</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>rs1800497</td>
<td>C/T</td>
<td>3' downstream</td>
<td>0.18</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1. Genotyped SNPs of DRD2. Allele frequencies were calculated from the 105 samples of the Stanley Foundation. Allele frequencies in the cohort of 100 subjects from the University of Bari are also provided where available.
Table 3.2. Oligonucleotides designed for experiments. a. Oligonucleotides (primers) for genotyping using GC_Clamp PCR and SNaPshot assay. b. PCR conditions and oligonucleotides sequence for splice variant amplification and testing, promoter region constructs, repeat region detection, and minigenes.

<table>
<thead>
<tr>
<th>SNPs ID #</th>
<th>Oligos</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs10891556 (GC_Clamp)</td>
<td>GCWild: 5’ CCTCTTTCCAGCCTCATTGTTG</td>
</tr>
<tr>
<td></td>
<td>GCSPN: 5’ GGCCGCAGCAGCACGTGTTG</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’ CACCACCTCCATTGTTG</td>
</tr>
<tr>
<td>rs12364283 (GC_Clamp)</td>
<td>GCWild: 5’CTGACTCTCAGTTTGCCGGA</td>
</tr>
<tr>
<td></td>
<td>GCSPN: 5’ GCCCCGCCGCCGCCGCCCTGTCTCATTTGCTGCTG</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’ CAGCACCTCCATTGTTG</td>
</tr>
<tr>
<td>rs1799978 (GC_Clamp)</td>
<td>GCWild: 5’ CAGCCCTCAGTTGAGCTTGC</td>
</tr>
<tr>
<td></td>
<td>GCSPN: 5’ GCCCCGCCGCCGCCGCCCTCATTTGCTGCTG</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’ CAGCACCTCCATTGTTG</td>
</tr>
<tr>
<td>rs1799732 (GC_Clamp)</td>
<td>GCWild: 5’ TCCACTCTCCAGTGTTGAGCTTGC</td>
</tr>
<tr>
<td></td>
<td>GCSPN: 5’ CAGCCCTCAGTTGAGCTTGC</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’ TCCACTCTCCAGTGTTGAGCTTGC</td>
</tr>
<tr>
<td>rs4938019</td>
<td>GCWild: 5’ GGGCCGGCGCCGCGGGAAGAGGCTGATTAAAAATATCAG</td>
</tr>
<tr>
<td></td>
<td>GCSPN: 5’ GAAAAGGCTGATTAAAAATATCAG</td>
</tr>
<tr>
<td></td>
<td>Forward: 5’ GCCCAAACCTACACTAAGCTGATACC</td>
</tr>
<tr>
<td>rs4350392</td>
<td>GCWild: 5’ CCGGGCCGGCGCGGCGGCGGCGGCTCAGCTCAGCTCAGCTG</td>
</tr>
<tr>
<td></td>
<td>GCSPN: 5’ CCGGGCCGGCGGCGGCGGCGGCTCAGCTCAGCTCAGCTG</td>
</tr>
<tr>
<td></td>
<td>Reverse: 5’ TGGTTCCAGGACAGCAATGCTG</td>
</tr>
<tr>
<td>rs</td>
<td>GC_SNP</td>
</tr>
<tr>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td>4648317</td>
<td>GCWild: 5' GCACAGGATGGCTGGAGCTTC</td>
</tr>
<tr>
<td>4581480</td>
<td>GCWild: 5' GACCCAGGCTTGCTTTTG</td>
</tr>
<tr>
<td>12574471</td>
<td>GCWild: 5' CAGGTGGGAGGAGGATGCA</td>
</tr>
<tr>
<td>4648318 (GC_Clamp)</td>
<td>GCWild: 5' CCGGCGGGGCCGCCGCCACAGCAACACAAATCTCGCT</td>
</tr>
<tr>
<td>(SNaPShot)</td>
<td>Forward: 5' CGTAGAAAGTCTAAAAGCAATGGAA</td>
</tr>
<tr>
<td>PEP: 5' GAGCACAGCAACAAATCTCCTG</td>
<td>Forward: 5' CGTAGAAAGTCTAAAAGCAATGGAA</td>
</tr>
<tr>
<td>7125415</td>
<td>GCWild: 5' AAAGCAGGGGACCTGTCTAAG</td>
</tr>
<tr>
<td>7103679</td>
<td>GCWild: 5' AGGTCTCCATTTTCTCTGTGCG</td>
</tr>
<tr>
<td>2283265</td>
<td>GCWild: 5' GAGAACAGGCTCATAGAAGGTATGC</td>
</tr>
<tr>
<td>PEP: 5' ACTGGAGAACAGGCTTAGGAAGGTAGG</td>
<td>GCWild: 5' GAGAACAGGCTCATAGAAGGTATGC</td>
</tr>
<tr>
<td>1076560</td>
<td>GCWild: 5' TTGCAGGAGATCTCTTGGAGTGG</td>
</tr>
<tr>
<td>6275 (SNaPshot)</td>
<td>Forward: 5' CCAGGTGACTCTCCCGAC</td>
</tr>
<tr>
<td>PEP: 5' GGACGTGCTTGAGTAGACC</td>
<td>Forward: 5' CCAGGTGACTCTCCCGAC</td>
</tr>
<tr>
<td>6277 (SNaPshot)</td>
<td>Forward: The same as the forward primer for rs6275</td>
</tr>
<tr>
<td>6279</td>
<td>Forward: 5' AGCCTGAGTCAGGGCCC</td>
</tr>
</tbody>
</table>

100
<table>
<thead>
<tr>
<th>SNP</th>
<th>Forward</th>
<th>Reverse</th>
<th>GCWild</th>
<th>GCSNP</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs1800497</td>
<td>5’ ACCGCTGCTCCACG</td>
<td>5’ CCCAGGGCTGAGTTTCT</td>
<td>5’ CATCCTCAAAGTGGTGACG</td>
<td>5’ GGCCGGGGCGGGGCAATCAAGCAAGTGCTGGCCA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5’ AGCTCAGCTCCATCTGGACG</td>
<td></td>
</tr>
<tr>
<td>SNaPshot</td>
<td>5’ AGCTCAGCTCCATCTGGACG</td>
<td>5’ CAACAGCCATCCTCAAGT</td>
<td>5’ AGCTCAGCTCCATCTGGACG</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5’ CCATCCTCAAAGTGGTACC</td>
<td></td>
</tr>
<tr>
<td>rs12363125</td>
<td>5’ AAGATCTCAAGCAATGACTACC</td>
<td>5’ AGAGGGAGGCAGGGGTCC</td>
<td>5’ AAGATCTCAAGCAATGACTACC</td>
<td>5’ CCATCCTCAAAGTGGTACC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5’ AGAGGGAGGCAGGGGTCC</td>
<td></td>
</tr>
<tr>
<td>rs2511521</td>
<td>5’ CCGGGCGGGCGGGCAGAAGGCGTCAGAGATCCTG</td>
<td>5’ TTCCAGCTCCCTCCTTGG</td>
<td>5’ CCGGGCGGGCGGGCAGAAGGCGTCAGAGATCCTG</td>
<td>5’ TTCCAGCTCCCTCCTTGG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5’ AAGGCCTTGAGGCTCATGACACTAAATAAACAAGG</td>
<td>5’ AAGGCCTTGAGGCTCATGACACTAAATAAACAAGG</td>
</tr>
</tbody>
</table>

AmpF_In6 5’ ATGGAACTCAGACTGGGCA
AmpR_In6 5’ AGAGGGAGGGCAGGGGTCC
<table>
<thead>
<tr>
<th>Experiment</th>
<th>PCR Conditions Cycles, TM, time</th>
<th>Oligos</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Splice forms detection</td>
<td>1, 95°C, 5 min; 30(23), 95 °C 30 sec, 60 °C 1 min, 72 °C 1 min; 1, 72 °C, 5 min Taq polymerase</td>
<td><strong>Forward:</strong> 5’ACATTGTCCCTCCGCAGACG  &lt;br&gt; <strong>Reverse:</strong> 5’ GCATGCCCATTTCTTCTCTGG</td>
<td>Forward primer was Fam labeled</td>
</tr>
<tr>
<td>Splice-specific AEI testing</td>
<td>1, 95°C, 5 min; 30, 95 °C 30 sec, 60 °C 1 min, 72 °C 1 min; 1, 72 °C, 5 min Taq polymerase (L), Sigma PCR mix (S)</td>
<td><strong>Forward (L):</strong> 5’GCTCCACTAAAGGGCAACTGT A  &lt;br&gt; <strong>Forward (S):</strong> 5’TGAGGGCTCCACTAAAGGAGG C  &lt;br&gt; <strong>Reverse:</strong> 5’GCATGCCCATTTCTTCTCTGG</td>
<td></td>
</tr>
<tr>
<td>Genetic variants for repeat region</td>
<td>1, 95°C, 5 min; 30, 95 °C 30 sec, 62 °C 1 min, 72 °C 1 min; 1, 72 °C, 5 min Taq polymerase</td>
<td><strong>Forward:</strong> 5’ CGCTGCTACACCTTGGCCA  &lt;br&gt; <strong>Reverse:</strong> 5’CAGCACCCTGTTTAAAGCCTCAGT</td>
<td></td>
</tr>
<tr>
<td>Reporter gene assay</td>
<td>1, 95°C, 5 min; 30, 95 °C 15 sec, 62 °C 1 min, 72 °C 1.5 min; 1, 72 °C, 5 min Taq High fidelity polymerase + 0.5× Enhancer</td>
<td><strong>D2upstreamFS:</strong> 5’GGGGTACCCACCCCTGG &lt;br&gt; <strong>D2upstreamFM:</strong> 5’CCGGTACCCCTGG &lt;br&gt; <strong>D2upstreamFL:</strong> 5’GGGGTACCCCTTC &lt;br&gt; <strong>Reverse:</strong> 5’GAAGATCTTCCGGGACAG &lt;br&gt; <strong>AGACCGGCGCGGCGTCTT</strong></td>
<td>Forward primers have Kpn I site, reverse primer has Bgl II site</td>
</tr>
<tr>
<td>Minigene constructs</td>
<td>1, 95°C, 2 min; 29, 95 °C 15 sec, 62 °C 1 min, 72 °C 3.5 min; 1, 72 °C, 10 min Taq High fidelity polymerase</td>
<td><strong>Forward:</strong> 5’CCCAAGCTTACCAGAAC &lt;br&gt; <strong>GAGTGCATCATGTC</strong>  &lt;br&gt; <strong>Reverse:</strong> 5’CGGCTTCGAGCGAACA &lt;br&gt; <strong>ATGGCGAGCATCTGA</strong>  &lt;br&gt; <strong>SP6:</strong> 5’ CATTAGTGACACTATAG</td>
<td>Forward primer has HindIII site, reverse primer has XhoI site</td>
</tr>
</tbody>
</table>
Table 3.3. Linkage disequilibrium analysis for 19 SNPs of DRD2 in Stanley samples.

a. Predicted haplotype frequencies using 19 SNPs of DRD2 in Stanley samples \((n = 105)\). The SNPs are ordered from 1-19 as indicated in Table 1 of the main text.

b. Linkage disequilibrium analysis of 19 SNPs in DRD2 using Helix Tree\(^R\) \((n=105)\).

The SNPs are ordered from 1-19 as indicated in Table 1 of the main text.

<table>
<thead>
<tr>
<th>#</th>
<th>Haplotype</th>
<th>EM Freq.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNP</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>-----</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>1</td>
<td>0.294</td>
<td>0.789</td>
</tr>
<tr>
<td>2</td>
<td>0.958</td>
<td>1.000</td>
</tr>
<tr>
<td>3</td>
<td>0.958</td>
<td>1.000</td>
</tr>
<tr>
<td>4</td>
<td>0.962</td>
<td>0.976</td>
</tr>
<tr>
<td>5</td>
<td>0.998</td>
<td>1.000</td>
</tr>
<tr>
<td>6</td>
<td>0.998</td>
<td>1.000</td>
</tr>
<tr>
<td>7</td>
<td>0.958</td>
<td>0.737</td>
</tr>
<tr>
<td>8</td>
<td>0.958</td>
<td>0.737</td>
</tr>
<tr>
<td>9</td>
<td>0.958</td>
<td>0.737</td>
</tr>
<tr>
<td>10</td>
<td>1.000</td>
<td>0.997</td>
</tr>
<tr>
<td>11</td>
<td>0.931</td>
<td>0.726</td>
</tr>
<tr>
<td>12</td>
<td>0.959</td>
<td>0.959</td>
</tr>
<tr>
<td>13</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>14</td>
<td>0.999</td>
<td>0.996</td>
</tr>
<tr>
<td>15</td>
<td>0.998</td>
<td>0.726</td>
</tr>
<tr>
<td>16</td>
<td>0.961</td>
<td>0.959</td>
</tr>
<tr>
<td>17</td>
<td>0.880</td>
<td>0.380</td>
</tr>
<tr>
<td>18</td>
<td>0.980</td>
<td>0.979</td>
</tr>
<tr>
<td>19</td>
<td>0.880</td>
<td>0.979</td>
</tr>
</tbody>
</table>

p
Table 3.4. The association of rs12364283 with AEI.

Exact P = 0.001, adjusted p-value = 0.019 (n = 68), 0 = no AEI. 1 = with AEI.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>AEI</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>51</td>
<td>60</td>
</tr>
<tr>
<td>AG</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>53</td>
<td>68</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Haplotype</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>343</td>
</tr>
<tr>
<td>2</td>
<td>348</td>
</tr>
<tr>
<td>3</td>
<td>350</td>
</tr>
<tr>
<td>4</td>
<td>354</td>
</tr>
<tr>
<td>5</td>
<td>358</td>
</tr>
<tr>
<td>6</td>
<td>360</td>
</tr>
<tr>
<td>7</td>
<td>364</td>
</tr>
<tr>
<td>8</td>
<td>368</td>
</tr>
<tr>
<td>9</td>
<td>372</td>
</tr>
<tr>
<td>10</td>
<td>376</td>
</tr>
<tr>
<td>11</td>
<td>379</td>
</tr>
</tbody>
</table>

Table 3.5. Repeat variants in the upstream of DRD2. Altogether, 11 repeat variants detected in 105 Stanley DNA samples with different lengths.
<table>
<thead>
<tr>
<th>rs12364283</th>
<th>Case-control</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>57</td>
<td>84</td>
</tr>
<tr>
<td>AG</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>60</td>
<td>95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Value</th>
<th>95% Confidence Interval</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Odds Ratio for DRD2 12364283 (A_A / A_G)</td>
<td>5.630</td>
<td>1.383</td>
<td>22.914</td>
</tr>
<tr>
<td>For cohort Profile = 3</td>
<td>2.488</td>
<td>.937</td>
<td>6.604</td>
</tr>
<tr>
<td>For cohort Profile = 4</td>
<td>.442</td>
<td>.274</td>
<td>.712</td>
</tr>
<tr>
<td>N of Valid Cases</td>
<td>95</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.6. The association of rs12364293 with Schizophrenia. Table 3.6a. Cross-tabulation and Chi square analysis. \( \chi^2 = 6.89, P = 0.009 \ (n = 95); 0 = \text{control}; 1 = \text{Schizophrenia}. \)  

b. Risk estimation of rs12364283 in schizophrenia.
Figure 3.1. Gene map and linkage blocks for DRD2. Panel a. Gene Maps of DRD2, including both long and short isoforms. The locations of 20 SNPs genotyped were highlighted. As the gene is located in the reverse strand in the chromosome 11q 23, the promoter region begin from the right side. Panel b. Linkage block of DRD2 in Hapmap database.
Figure 3.2. Control experiments to test PCR efficiency and calculate standard curves for DRD2S and S expression. Panel a. Control experiments to test the PCR efficiency for DRD2S and L using upstream and downstream primers in regions common to both isoforms. Two plasmids carrying only long or short isoforms were used to test the amplification efficiencies for the two splice variants. Panel b. Standard curves for calculating relative mRNA expression of DRD2S and DRD2L using a fluorescently labeled forward primer. To exclude differential fluorescent yield between different lengths of PCR products, a series of mixtures of two cDNA plasmids (DRD2S and DRD2L) with differential ratios were used to obtain the standard curve, with a linear regression line showing a correlation coefficient of 0.998, p < 0.0001.
\[ y_L = -3.426x + 24.7 \]
\[ y_S = -3.219x + 24.5 \]

\[ y = 1.1462x + 0.133 \]
Figure 3.3. The measurement of Allele-specific expression of DRD2 using three indicator SNPs. Panel a. Allele-specific expression of DRD2 mRNA for human autopsy samples from two regions, prefrontal cortex (54) and striatum (14).

Samples are heterozygous for at least one of the indicator SNPs, rs6275 (47 subjects), rs6277 (54) and rs6279 (49). 33 subjects were heterozygous for rs6275 and rs6277, 35 for both rs6277 and rs6279, and 47 for both rs6275 and rs6279. Allelic expression ratios of the two alleles (A/G for rs6275, C/T for rs6277, and C/G for rs6279) were normalized to genomic DNA ratios. Two cDNA syntheses and three PCR reactions for each cDNA preparation were done for each sample. Data are mean ± SD. Panel b & c. Correlations between two marker SNPs. rs6277 & rs6279, Pearson r = 0.9626, p < 0.01; rs6275 & rs6279, person r = 0.931, p < 0.0001. (Since rs6275 and rs6275 are belong to the same amplicon, no need for correlation analysis here)
Figure 3.4. Association analysis between single SNPs and AEI using HelixTree. Panel a. Association analysis between single SNPs and AEI. Adjusted P value was used to correct for multiple test effects. Panel b. Association analysis between single SNP and AEI using samples with heterozygotes of rs12364283 been excluded. To verify the effects of rs12364283, the same analysis was conducted for samples with the exclusion of heterozygous samples of rs12364283.
Figure 3.5. *cis*-Regulation of alternative splicing of DRD2. Panel a. Association studies to identify SNPs correlated with discrepant AEI ratio between DRD2S and L. Adjusted p value was used considering multiple testing. Panel b. AEI measurements for splice-specific isoforms (DRD2L and DRD2S) using indicator SNPs rs6277 and rs6275. All 37 RNA samples were from Stanley collection, including 30 heterozygous samples for rs6277 and 7 heterozygous for rs6275. Overall AEI ratios (AEI (T)) were the same as measured in Figure 3.3a.
Figure 3.6. The correlation of rs1076560 genotype status with altered splicing in brain tissues. Panel a. Comparison of AEI ratios between DRD2L and DRD2S divided by rs1076560 genotype. Data representing AEI ratios for DRD2L and DRD2S from the same sample are connected by solid lines. Panel b. Comparison of the mean AEI ratios of total mRNA and for each splice variant, divided by genotypes (CC vs. CA) of rs1076560. Data are mean ± SD, n = 3, * p < 0.05, ** p < 0.01, *** p < 0.001. Panel c. Relative expression of DRD2S mRNA for different genotypes (CC vs. CA + AA) of rs1076560 in two brain regions, prefrontal cortex and striatum. The mean of each genotype group was compared for both regions. Data are mean ± SD, n= 3, *** p < 0.001.
**AEI Ratios**

(a) CC and CA AEI Ratios

- CC AEI Ratios
- CA AEI Ratios

(b) Bar graph showing the mean of AEI (T), AEI (L), and AEI (S) for CC and CA.

(c) Scatter plot showing D2S (% of total) for CC and CA + AA in Prefrontal Cortex and Striatum.
Figure 3.7. Promoter region constructs and their transcription activities tested by reporter gene assay. Panel a. Gene map showing promoter regions and length of DNA fragments used for the reporter assays. +1 stand for the transcription start site, Pro_LT1 and Pro_LC1 have 8 nucleotides deletion in the repeat region compared to the reference sequence. Pro_LT2 and Pro_LC2 have a 4 nucleotides deletion. rs12364283 is located in position –847. Panel b. Luciferase report gene assay to test transcriptional activity of upstream regions of DRD2. Compared with Pro_S, the expression levels of Pro_L1 (2) T and Pro_L1 (2) C are less, and the C allele conferred higher expression than the T allele (Pro_L1 (2) T versus Pro_L1 (2) C, data are mean ± SD, *** p < 0.0001, * p < 0.05. In HEK cells, Pro_LC1 (76% ± 10) vs. Pro_LT1 (41% ± 4), p < 0.001 (one – way ANOVA, Bonferroni’s Multiple testing), Pro_LC2 (72% ± 10) vs. Pro_LT2 (50% ± 9), p < 0.05; in SHSY5Y cells, Pro_LC1 (70% ± 12) vs. Pro_LT1 (52% ± 9), p < 0.05, Pro_LC2 (69% ± 10) vs. Pro_LT2 (49% ± 1), p < 0.05.
a

Pro_S

Pro_M

Pro_LT1

Pro_LC1

Pro_LT2

Pro_LC2

Repeat region

+1 +292

-283

-600

-963

-963

-963

-963

-963


b

% of Pro_S Activities

0.00 40.00 80.00 120.00

Control vector Pro_S Pro_M Pro_LT1 Pro_LC1 Pro_LT2 Pro_LC2

HEK

SHSY5Y

*** * *

120
Figure 3.8. Alternative splicing from DRD2 minigene constructs in HEK cells. Two minigene constructs carrying two haplotypes of intronic SNPs were transfected into HEK cells, 24 hrs and 48 hrs after transfection, total RNA was isolated from the cells and the relative expression of two splice isoforms was measured using fluorescently labeled primers. Data are mean ± SD, ** p < 0.01.
Figure 3.9. rs1076560 genotype-based analysis of fMRI response during the working memory task. **Panel a:** Results of the ANOVA in SPM2 overlaid onto an average axial MRI at the level of the head of the caudate. The color bar indicates Z values of the difference in Blood Oxygen Level Dependent (BOLD) signal between the groups separated by CC and CA genotype. During the working memory task subjects with CA genotype had greater BOLD activity in bilateral head of the caudate compared with subjects with CC genotype. **Panel b:** Means ± .95 standard error plots reflecting percent signal change from the cluster in left caudate head showing a significant difference between CC and CA subjects. Subjects with CA genotype had greater engagement of caudate head, compared with the homozygous subjects (One way ANOVA: F (1, 42) =18.950, p = 0.00008). (Provided by Dr. Bertolino from the University of Bari, Italy)
Provided by Dr. Bertolino from the University of Bari, Italy
Figure 3.10. Alignment of DNA sequences flanking rs12364283 from several species.

The flanking sequence of rs12364283 (~ 40bp) is highly conserved between different species. In comparison with other species, the dominant allele ‘A’ of rs12364283 is unique to humans.
CHAPTER 4

Calmodulin interaction with peptides from G-protein coupled receptors measured with S-Tag labeling

Published in: Biochem Biophys Res Commun, 2005 Jul 29; 333(2): 390-5. Ying Zhang, Danxin Wang, and Wolfgang Sadée
Program in Pharmacogenomics, Department of Pharmacology, College of Medicine and Public Health, The Ohio State University, Columbus, OH, 43210

4.1 Introduction

Calmodulin (CaM), a universal Ca\(^{2+}\) sensor, interacts with numerous signaling proteins (Meyer et al., 1992). Recent studies have revealed a direct interaction of CaM with G protein coupled receptors (GPCRs). Metabotropic glutamate receptors subtype 5 and 7 (mGlu5, mGlu7) interact with CaM through binding motifs in the carboxyl terminus, affecting receptor phosphorylation (Minakami et al., 1997; Nakajima et al., 1999). Moreover, CaM binds to the third intracellular loop (i3 loop) of mu opioid receptor (MOR) at an overlapping region also required for G protein coupling (Wang et al., 1999), thereby regulating G protein coupling of MOR (Wang et al., 2000a).
Similarly, CaM binds to D2-dopamine (D2) receptors through a domain adjacent to the G protein coupling domain in the i3 N-terminus, suppressing G protein activation (Bofill-Cardona et al., 2000). Lastly, serotonin receptor 1A (5HT1A) harbors two distinct CaM binding domains in the C and N terminus of i3 loop, regulating protein kinase C-mediated receptor phosphorylation (Turner et al., 2004).

Structural motifs required for receptor interaction with CaM and G proteins appear to be similar (Wang et al., 1999), since amphipathic basic peptides, such as mellitin, bind with high affinity to both CaM and Goα proteins (Kataoka et al., 1989). Therefore, G protein-coupling regions in the i3 loop are likely candidates as CaM-binding domains. A sequence motif search for CaM binding domains (Wang et al., 1999) implicated the i3 loops of opioid receptors and other GPCRs, such as the muscarinic receptors, but specific sequence requirements for CaM binding are difficult to define. While a broad range of peptides can bind to CaM, single amino acid perturbation of the sequence may abrogate the interaction, indicating the interaction is selective. For example, substitution of Lys273 with Ala in MOR (K273A) leads to deficient CaM-receptor interaction, tested with intact receptors and peptides derived from the proposed binding domain (Wang et al., 1999).

A number of assay procedures have been developed to detect peptide-CaM interactions (Li and Villalobo, 2002; Minakami et al., 1997; Terpe, 2003; Turner et al., 2004). However, many of these techniques do not have sufficient sensitivity for kinetic analysis of the binding parameters, because required peptide and CaM concentrations
exceed Km values, enabling only stoichiometric analysis. Gel shift and overlay assays (Li and Villalobo, 2002; Minakami et al., 1997) including chemical cross-linking methods (Li and Villalobo, 2002; Martinez et al., 2003) provide largely qualitative information. Most previous assays require the use of tags for protein expression and purification, such as glutathione-S-transferase (GST) and maltose-binding protein (MBP). These relatively large tags (26kD for GST, 40kD for MBP) (Terpe, 2003) can affect solubility of the chimera and interfere with the interaction between proteins (peptides) and calmodulin. The S-Tag consists of only 15 amino acids, and the \textit{in vitro} translation system avoids the need for protein purification, avoiding possible degradation during the process. S-Tag detection is highly sensitive (f mole detection limit), and therefore permits kinetic analysis at low protein concentrations in small volumes. Other quantitative assays of calmodulin-protein interactions include surface plasmon resonance (SPR) and fluorescence measurement (Martinez et al., 2003; Tang et al., 2001; Turner et al., 2004) but require special instrumentation or lack sensitivity.

To characterize peptide-CaM interactions more accurately and expand the search for CaM binding motifs in GPCRs, we have developed a rapid quantitative peptide-CaM interaction assay using an \textit{in vitro} translation system with S-Tag labeled peptides, coupled to a sensitive S-Tag detection method (Kim and Raines, 1993). This assay served to measure CaM-binding of peptides derived from the i3loops of MOR, muscarinic, melanocortin, and serotonin receptors.
4.2 Materials and Methods

**S-Tag fusion peptide expression plasmids**

Sequences of human MOR (bases 772-858) and rat CaM kinase II (CaMKII) (bases 868~927) were amplified by PCR from cDNA plasmids. Melanocortin receptor 1 (MC1R) (bases 634~723), 3 (MC3R) (bases 742~852), serotonin receptor 1A (5HT1A) (bases 640~711) and serotonin receptor 2A (5HT2A) (bases 877~978) were amplified from human genomic DNA without intronic interruption. Primers were designed to carry BglII and KpnI restriction sites, and a stop codon (TAG). The amplified fragments were sub-cloned into PET-32a vector (Novagen Inc., WI), with T7 promoter and S-Tag 5’ of the insertion. All constructs were sequenced (Genetics/Microarray Core, OSU). The peptide sequences are shown in Table 4.1.

**In vitro translation and peptide detection**

We used the Ecopro coupled transcription and translation system from Novagen Inc. (Madison, WI). Generated S-Tag fusion peptides were measured by S-Tag rapid assay kit (protocol from Novagen Inc., WI). This assay is based on reconstitution of enzyme activity, using the binding of a small peptide S-Tag (15aa) with ribonuclease S-protein (150aa) to form active ribonuclease which break down poly(C) and cause the increasing of light absorbance at 280nM. Amounts of S-Tag fusion peptides were calculated in comparison to S-Tag peptide standard (Novagen Inc., WI).
S-Tag-peptide binding to CaM-agarose beads

After *in vitro* translation, a suspension of CaM-agarose beads (0.6 µM in 10 µl) was incubated with *in vitro* translation products in 20 mM Tris-HCl, 150 mM NaCl, pH 7.5, and 100 µM Ca\(^{2+}\) (binding buffer) for 30 minutes at room temperature. Beads were washed with binding buffer and eluted with elution buffer (20 mM Tris-HCl, 150 mM NaCl, 0.1% TrionX-100, 2 mM EGTA). Retention of S-Tag peptide on CaM beads was quantified with the S-Tag assay, after elution form the beads. For competition assays, CaM-agarose beads were pre-incubated with synthetic peptides (Table 4.1) for 20 minutes before adding the S-Tag peptides.

**Free Ca\(^{2+}\) concentration**

This was calculated with the Maxchelator program (C. Patton, Stanford University, CA), and calibrated using different concentrations of EGTA and CaCl\(_2\).

**Western blot of S-Tag peptides**

The translation products (5 µl) were treated with acetone and separated on 15% SDS-polyacrylamide gels. S-Tag peptides were detected with 1/5000 AP conjugated S-Protein (Novagen Inc., WI), followed by the ECL western blotting detection reagents (Amersham, Piscataway, NJ).
**Peptide synthesis**

Peptides corresponding to sections of i3 loop of MOR (Table 1) were prepared as described previously (Wang et al., 1999). Peptides derived from muscarinic receptors were synthesized by Biomedical Service Center at University of California San Francisco. CaM binding domain (CBD) peptide derived from CaM kinase II (CaMKII aa290-309) was from EMD Biosciences (San Diego, CA). Other reagents were from Sigma or Fisher Scientific.

**4.3 Results and Discussion**

For each S-Tag peptide, we performed Western blots after *in vitro* translation, showing bands of the expected molecular weight. This indicates full-length synthesis was achieved (Figure 4.1a shows the results of S-Tag CaMKII CBD and S-Tag MORi3 peptides). We next determined the dependence of peptide-CaM interactions on Ca$^{2+}$ ($10^{-7}$ to $10^{-3}$ M) (Jurado et al., 1999). Maximum binding of S-Tag MORi3 peptide occurred at 100 µM Ca$^{2+}$. No detectable binding was observed in the absence of Ca$^{2+}$ (data not shown).

Sensitivity of the S-Tag assay permits use of S-Tag peptides at a concentration below their Kd, enabling the estimation of binding affinities. With a constant amount (0.6 µM) of CaM-agarose beads, the binding of S-Tag CaMKII CBD (Table 1) to CaM was dose-dependent and has an estimated Bmax of 10 pmol (Figure 4.1b), while S-Tag MORi3
gave Bmax 3.0 pmol (Figure 4.1c). However, higher concentration of peptide (>1200 nM) reduced CaM binding, possibly due to peptide aggregation. Therefore, accurate calculation of binding affinity and Bmax from saturation experiments is confounded if peptide concentrations required for saturating CaM-binding exceed solubility. Relative affinities are then obtained from competition analysis (see below). Empty PET vector resulted in no detectable binding (data not shown).

Different regions of MORi3 (aa258-286) have been shown to display different binding potency with CaM, but a quantitative analysis was not feasible using a gel-shift assay (Wang et al., 1999). In a competitive binding experiment, synthetic peptides (60 µM) were pre-incubated with CaM-agarose beads (0.6 µM) before adding S-Tag MORi3 as a tracer (~180 nM). CaMKII CBD and MORi3 completely suppressed S-Tag MORi3 binding to CaM (Figure 2a). MORi3-C1, representing the i3 C-terminus (aa273-286) was >85% effective, consistent with previous results (Wang et al., 1999), while MORi3-C2 (aa267-286) and the N-terminal MORi3-N (aa258-272) reduced S-Tag MORi3-CaM binding by 50% and 60%, respectively (Figure 2a). Hence, as reported for other GPCRs (Bofill-Cardona et al., 2000; Turner et al., 2004), N-terminal portions of MORi3 appear to contribute to CaM-binding, as shown by MORi3-N induced inhibition (Figure 4.2a), and the fact that the entire MORi3 loop peptide was most effective. The reduced interaction of MORi3-C2 with 5 additional amino acids compared to MORi3-C1 indicates that the conformation of the i3 loop may affect CaM binding.
To quantify the binding of CaMKII CBD and MORi3 with CaM, we applied competitive binding assay with increasing concentration of synthetic peptides. The calculated IC$_{50}$’s were 0.7 and 0.8 µM, respectively, showing similar binding affinity for MORi3 and CaMKII CBD (Figure 4.2b and 4.2c). We also measured the binding capacity of a mutant MOR peptide (K273A-MORi3-C1, aa273-286), with a Lys$^{273}$-Ala substitution, a critical residue in the predicted CaM binding motif (Wang et al., 1999). The Ala substitution had been shown to reduce CaM binding, but the relative affinity had not been determined (Wang et al., 1999). K273A-MORi3-C1 binds to CaM less strongly than MORi3-C1 (P < 0.05, t-test) (Figure 4.2a). Competitive binding assays revealed a 5-fold difference in binding affinity between the wild type and the mutant peptides (IC$_{50}$ values: wild-type, 36 µM; mutant, 170 µM) (Figure 4.2d). A 5-fold difference is likely to affect the functional interaction between CaM and MOR in vivo, since CaM binding is titrated among numerous competing proteins. Functional assays in MOR-transfected cells showed that G-protein signaling of K273A-MOR is not regulated by CaM (Wang et al., 1999; Wang et al., 2000b). These results with known calmodulin binding peptides (CaMKII CBD, MORi3) validate the S-Tag assay as a rapid and sensitive method. Moreover, competition binding experiments using S-Tag MORi3 as the tracer at low concentrations (180 nM) provide quantitative results on relative peptide-CaM affinities.
We next screened peptides derived from the i3 loops of other GPCRs (Table 4.1) for CaM binding using S-Tag MORi3 as a tracer. Peptides derived from muscarinic receptors M1-3 receptors, representing 27 or 28 amino acid fragments in the C terminal of i3 loop, contain a CaM-binding motif (Wang et al., 1999). M3i3-C inhibited 90% of the S-Tag MORi3-CaM binding, while, M1i3-C was less effective, and M2i3-C inhibited binding by only 10% (Figure 4.3a). In competitive binding experiments, IC$_{50}$ for M3 peptide was 1.4 µM (Figure 4.3b). Comparing i3 loop sequences of M3i3-C with the less potent M1 and M2 derived peptides shows distinct residues adjacent to the conserved CaM-binding motif Lys/Arg-Glu-Lys-Lys (K/R-E-K-K) close to the i3 C terminus (Figure 4.4). Residues Ser-Leu-Val (S-L-V) in M1 and M3 are replaced with Pro-Pro-Ser (P-P-S) in M2 (Figure 4.4), possibly accounting for low binding affinity of M2 peptide. The greater binding potency of M3i3-C over M1i3-C may be due to a Gln$^{491}$ (Q491) in the C terminus of i3 loop, instead of Arg in M1 and M2. Point mutation Q490L in rat M3 (corresponding to human M3 Q491) was shown to result in robust-agonist independent signaling (Schmidt et al., 2003). Our previous studies indicated that a MOR mutant deficient in CaM binding (K273A) displays increased basal and agonist-stimulated GTP$_{\gamma}$S binding activity (Wang et al., 2000b). Apparently, CaM binding to MOR suppresses basal G protein coupling. If Gln$^{491}$ is crucial for CaM binding of M3, it is possible that the mutation reduced CaM-binding of the M3 receptor, thereby, increasing basal activity. The potential role of Gln$^{491}$ in CaM-binding requires further study.
To screen more GPCRs for CaM binding, we constructed additional S-Tag labeled peptides (Table 4.1) derived from the i3 loop of several additional GPCRs. After \textit{in vitro} translation, binding of the S-Tag peptides (~310 nM) to CaM beads was determined and compared to S-Tag CaMKII CBD. Shown in Figure 5a, S-Tag 5HT1Ai3-N binds to CaM with similar efficacy as CaMKII CBD, consistent with previous results (Turner et al., 2004). MC1Ri3 also showed moderate binding potency compared to MORi3 (Figure 4.5a), indicating that MC1R is another possible CaM binding protein. To test this further, we obtained a dose-response curve for S-Tag MC1Ri3 peptide. With increasing amounts of S-Tag MC1Ri3, the maximum binding increased to 4.5 pmol (Figure 4.5b), which is comparable to S-Tag MORi3 and S-Tag CaMKII CBD binding (Figure 1c and b). MC1R contains a predicted CaM binding motif with a pattern of hydrophobic residues in positions 1, 8, and 14 (Rhoads and Friedberg, 1997). The G protein-binding domain in the i3 loop of MC1R was reported to contain 4 amino acids, K226-R227-Q228-R229 (Frandberg et al., 1998). Notably, these 4 amino acids are also components of a predicted CaM-binding motif (Wang et al., 1999), further supporting the hypothesis that CaM and G proteins share a common binding motif. Therefore, CaM might regulate G protein activity of MC1R. A critical role of CaM in melanocortin receptor binding and activity has been previous reported (Gerst and Salomon, 1987), but the molecular mechanism remains to be clarified. S-Tag 5HT2Ai3-C and S-Tag MC3Ri3 bound only marginally to
CaM (Figure 4.5a), indicating that these two peptides did not contain a CaM binding motif, also serving as negative controls showing that S-Tag peptide-CaM binding varies with peptide primary structure.

In conclusion, the use of S-Tag labeled peptides permits rapid and sensitive screening of GPCR-derived peptides as CaM binding candidates. Muscarinic receptors, especially M3, and melanocortin receptor MC1R may join GPCRs shown to bind CaM, with potential implications for receptor regulation and signaling.
<table>
<thead>
<tr>
<th>Designation</th>
<th>Description</th>
<th>Amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC1R i3</td>
<td>melanocortin 1 receptor (alpha melanocyte stimulating hormone receptor) i3 loop (aa212-241)</td>
<td>ARACQHAQGARHKRQRPVHQGFGLKGV</td>
</tr>
<tr>
<td>MC3R i3</td>
<td>melanocortin 3 receptor i3 loop (aa248-284)</td>
<td>FLFARLHVHKIAALPPADGVAPQHSMKGAVTITIL</td>
</tr>
<tr>
<td>5HT1A i3-N</td>
<td>serotonin receptor 1A i3 loop N-terminal peptides (aa214-237)</td>
<td>YGRIFRAARFRIRKTVKVEKTG</td>
</tr>
<tr>
<td>5HT2A i3-C</td>
<td>serotonin receptor 2A i3 loop C-terminal peptides (aa293-326)</td>
<td>KLFQRSIHREPGRGRTMQSISNEQKACKVLG</td>
</tr>
<tr>
<td>MORi3</td>
<td>Synthetic MOR i3 loop peptide (aa 258-286)</td>
<td>ILRLKSYRMLSGSKEKDRNRIRTRMVY</td>
</tr>
<tr>
<td>CaM KII CBD</td>
<td>Synthetic calmodulin binding domain (calmodulin kinase II (aa 290-309))</td>
<td>KKKFNARRKLKGAILTTLMA</td>
</tr>
<tr>
<td>MORi3-C1</td>
<td>Synthetic MOR i3 loop C-terminal short peptide (aa 273-286)</td>
<td>KDRNRIRTRMVY</td>
</tr>
<tr>
<td>MORi3-C2</td>
<td>Synthetic MOR i3 loop C-terminal long peptide (aa 267-286)</td>
<td>LSGSKEKDRNRIRTRMVY</td>
</tr>
<tr>
<td>MORi3-N</td>
<td>Synthetic MOR i3 loop N-terminal peptide (aa 258-272)</td>
<td>ILRLKSYRMLSGSKE</td>
</tr>
<tr>
<td>K273A-MOR i3-C1</td>
<td>Synthetic MOR i3 loop C-terminal short peptide with K273A mutation</td>
<td>ADRNRIRTRMVY</td>
</tr>
<tr>
<td>M1i3-C</td>
<td>Synthetic Masicrine receptor 1 i3 loop C-terminal (aa 341-368)</td>
<td>QKPRGKEQLAKRKTFLVKEKKAARTLS</td>
</tr>
<tr>
<td>M2i3-C</td>
<td>Synthetic Masicrine receptor 2 i3 loop C-terminal (aa 364-390)</td>
<td>KIVKMTQPAKKPSPREKVKTRTL</td>
</tr>
<tr>
<td>M3i3-C</td>
<td>Synthetic Masicrine receptor 3 i3 loop C-terminal (aa 467-494)</td>
<td>FALKTRSQITKRKRMLVKEKKAATLQ</td>
</tr>
</tbody>
</table>

Table 4.1. Peptide sequences. All S-Tag labeled peptides carry a 15aa S-Tag sequence (KQTAALKFERQHMDS) at their N-termini.
Figure 4.1. Dose-dependent binding of S-Tag CaMKII CBD and S-Tag MORi3 with CaM. Panel a. Western-blots of S-Tag CaMKII CBD and S-Tag MORi3 peptides using AP conjugated S-Protein. Panel b & c. S-Tag peptides were incubated with 0.6 µM CaM-agarose, and peptide binding was determined with the S-Tag assay. Bmax: S-Tag CaMKII CBD (panel b), 10.0 pmol; S-Tag MORi3 (Panel c), 3.0 pmol. Control PET-32a vector expressing S-Tag without peptide insertion did not show detectable binding. Data are mean ± SEM, n=4.
Figure 4.2. Titration of S-Tag MORi3 binding to CaM-agarose beads with unlabeled MOR peptides (Table 1). Panel a. Peptides (60 µM) were pre-incubated with CaM-agarose beads (0.6 µM) for 20 min at room temperature, followed by incubation with S-Tag MORi3 (~180 nM). S-Tag MORi3 binding to CaM was measured with S-Tag rapid assay. Data are % of control, in the absence of unlabeled peptide. Mean ± SEM, n=4, * P< 0.05. Panel b & c. Competition binding curves for CaMKII CBD (panel b) and MORi3 (panel c). IC$_{50}$ values: CaMKII CBD, 0.7 µM; MORi3, 0.8 µM. Data are mean ± SEM, n=4. Panel d. Competition binding curves for MORi3-C1 and K273A-MORi3-C1. IC$_{50}$ values: wild type, 36 µM; K273A, 170 µM. Data are mean ± SEM, n=4.
Figure a shows the S-Tag MOR3 bound to CaM (as a percentage of control) for different conditions: CaMKII CBD, MOR3, MOR3-C1, MOR3-C2, K273A-MOR3-C1. The bars indicate the range of values with error bars.

Figure b illustrates the S-Tag MOR3 bound to CaM (as a percentage of control) in relation to CaMKII CBD concentration (in log M). The graph includes data points with error bars.

Figure c presents the S-Tag MOR3 bound to CaM (as a percentage of control) in relation to MOR3 concentration (in log M). The graph includes a curve and data points with error bars.

Figure d shows the S-Tag MOR3 bound to CaM (as a percentage of control) for Peptides concentration (in log M). The graph includes two curves with data points for K273A-MOR3-C1 and MOR3-C1.
Figure 4.3. Titration of S-Tag MORi3 binding to CaM-agarose beads with unlabeled peptides derived from muscarinic receptors M1-3. Panel a. Synthetic peptides (60 µM) were pre-incubated with CaM beads, followed by incubation with S-Tag MORi3 (~180 nM). Panel b. Competition binding curves for M3i3-C. IC$_{50}$ value: 1.4 µM. Data are mean ± SEM, n=4.
M3 467 FALKTRSQITKRRMSLVEKKEKAAQLTLS
M1 341 QKPRGKEQLAKKFRSLSLVEKKEKKAARLTL
M2 364 KIVKMTKQPAKKK-PPSREKKVTRTI

Figure 4.4. Sequence alignment of M1, M2 and M3. Muscarinic receptors share the CaM-binding motif Lys/Arg-Glu-Lys-Lys (K/R-E-K-K, underlined).
Figure 4.5. CaM binding of S-Tag peptides derived from various GPCRs and

dose-dependent binding of S-Tag MC1Ri3 to CaM beads. Panel a. CaM binding of

S-Tag peptides derived from various GPCRs. S-Tag peptides (~310 nM) were

incubated with 0.6 μM CaM beads for 30 min. Data are expressed as % of S-Tag

CaMKII CBD binding. Panel b. Dose-dependent binding of S-Tag MC1Ri3 to CaM

beads. Bmax is 4.5 pmol for MC1Ri3 peptide. Data are mean ± SEM, n=3.
BIBLIOGRAPHY


