A Single-Stranded DNA Binding Site in the Human A$_1$ Adenosine Receptor Gene Promoter

HONGZU REN and GARY L. STILES

Departments of Medicine and Pharmacology, Duke University Medical Center, Durham, North Carolina 27710

This paper is available online at http://www.molpharm.org

ABSTRACT

Human A$_1$ adenosine receptor gene expression is controlled by two independent promoters. The upstream promoter, promoter A, is subject to tissue specific regulation because not all cells express the mRNA associated with this promoter. One potential regulatory sequence located downstream of the TATA box is an AGG element appearing in a tandem repeat. In a previous study, transient transfection assays showed that mutations made in those AGG elements substantially reduced promoter activity. In the current study, DNase I footprinting indicated nuclear protein binding to this sequence between the TATA box and transcriptional start site. Electrophoretic mobility shift as-say confirmed further the presence of an AGG element binding protein (AGBP) in human brain nuclear protein extracts. This binding protein has much higher affinity for single-stranded than for double-stranded DNA, and the binding is sequence specific. A series of assays also showed that AGBP is not related to the nuclear factor SP1 and the binding does not require metal cofactors. Therefore, AGBP is likely to be a spe-cific single-stranded DNA binding protein that is required for the full expression of A$_1$ adenosine receptor gene and particu-larly abundant in brain tissue.

Adenosine is present in all mammalian cells and can in-duce a wide variety of physiological responses in the heart, brain, and vascular system. Extracellular adenosine pro-duces these actions via a group of membrane bound recep-tors; namely, the adenosine receptors, which belong to the super family of G protein-coupled receptors. These adenosine receptors share the common structural motif of other seven-transmembrane domain receptor molecules and are coupled to effector systems through the heterotrimeric G proteins.

At present, four adenosine receptors (A$_1$, A$_{2a}$, A$_{2b}$, and A$_3$) have been cloned from a variety of different species (Libert *et al.*, 1991, 1992; Olah *et al.*, 1992; Ren and Stiles, 1994a; Reppert *et al.*, 1991; Tucker *et al.*, 1992; Zhou *et al.*, 1992). Recently, the gene structures (or partial structures) of all four adenosine receptors have been reported (Aguilar *et al.*, 1995; Bhattacharya *et al.*, 1993; Chu *et al.*, 1996; Jacobson *et al.*, 1995; Le *et al.*, 1996; Murrison *et al.*, 1996; Ren and Stiles, 1995), and all have only one intron interrupting the coding sequence and forming two unequal parts. The last exon of all these genes contains part of the coding sequence and the entire 3′-untranslated sequence. There are, however, differences in the 5′-untranslated region. For the human A$_1$ receptor, the 5′-untranslated region of the gene contains two exons interrupted by an intron. These exons are expressed in a mutually exclusive fashion depending on which one of the two promoters is used for transcription (Ren and Stiles, 1995). In contrast, no intron has been reported in the 5′-untranslated region of human A$_3$ receptor gene (Murrison *et al.*, 1996) or rat A$_{2a}$ receptor gene (Chu *et al.*, 1996). Human A$_1$ receptor and rat A$_{2a}$ receptor gene expressions are con-trolled by two independent promoters (Chu *et al.*, 1996; Ren and Stiles, 1995). In the human A$_3$ receptor and rat A$_{2a}$ receptor promoters, no TATA box has been found (Chu *et al.*, 1996; Murrison *et al.*, 1996). Many of these gene motifs are shared with other G protein-coupled receptors (Lee *et al.*, 1996; Li *et al.*, 1996).

Both promoters A and B in the human A$_1$ adenosine recep-tor gene have nonclassic TATA boxes, although the TATA box in promoter B is more transcriptionally active than that in promoter A (Ren and Stiles, 1995). The A$_1$ receptor tran-scripts derived from each promoter are similar in size. The transcript derived from promoter B has multiple AUG codons in the 5′-untranslated region corresponding to exon 1B, whereas the transcript from promoter A does not have an upstream AUG codon in the corresponding exon 1A. The upstream AUG codons in exon 1B hinder the expression of receptor protein at the post-transcriptional level (Ren and Stiles, 1994b). Transient transfection of a plasmid containing

ABBREVIATIONS: AGBP, AGG element binding protein; EMSA, electrophoretic mobility shift assay; -F, coding sequence; -R, noncoding sequence; EGTA, ethylene glycol bis(β-aminoethyl ether)-*N,N,N′,N′*-tetraacetic acid; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; EGF-r, epidermal growth factor receptor; DTT, dithiothreitol; WT, wild-type.

promoter A linked to the luciferase gene as a reporter showed that a series of AGG elements immediately downstream of TATA box are responsible for a major part of the transcriptional activity of this promoter (Ren and Stiles, 1995).

To explore the mechanism of transcriptional regulation, we analyzed the binding activities of nuclear proteins derived from human tissues or various nonhuman mammalian cells to the AGG elements in promoter A. Interestingly, the protein or proteins in nuclear extracts were found to bind single-stranded DNA (AGG elements) much better than to double-stranded DNA of the same sequence. These binding proteins are particularly abundant in brain tissue.

Methods

Oligonucleotide DNA probes used in EMSA. The oligonucleotide DNA fragments used as probes in EMSA were synthesized by GIBCO BRL (Gaithersburg, MD). Their sequences are: AGG-F (coding strand), 5′-GGGAGGAGACGGAGGATGAGGAGGGAGGGG-3′; AGG-R (noncoding strand), 5′-CCCCTCCCTCCTCATCCTCCGTCTCCTCCC-3′; WT, double-stranded probe made by annealing AGG-F and AGG-R; MUT, 5′-GGGTTTACTCGGGTTA-TGTCAACCGCTTGG-3′ (the mutated bases are underlined); MUT-1, 5′-GGGTTTACT-CGGAG-GATGAGGAGGGAGGGG-3′; MUT-2, 5′-GGGAGGAGACGGGTTAT-GAGGAGGG-AGGGG-3′; MUT-3, 5′-GGGAGGAGACGGAGGATG-TCAACCGCTTGG-3′; MUT-4, 5′-GGGAGGAGACGGTACGATCTG-CCGACTCGGTCAGGATGAGGAGGGAGGGG-3′ (a 20-base randomly selected sequence was inserted between the first AGG element and the second AGG element); AGGE-1, 5′-GATCGGGAGGAGACGGACGT-3′; AGGE-3, 5′-GATCAG-GATGAGGAGGGAGGGGTAGC-3′; and SP1, 5′-GATCAGCTGGGGCGGGGCTAAC-3′. The mutant sequences used in this study were the same as those used in our previous study (Ren and Stiles, 1995).

Nuclear protein extract from cultured cells and frozen human tissues. The frozen human brain tissue was obtained from the Human Brain Bank of the Bryan Neurosciences Program (Duke University Medical Center, Durham, NC). Other frozen human tissues were obtained from the laboratory of Dr. Robert J. Lefkowitz (Duke University Medical Center). Nuclear proteins were extracted from various cultured mammalian cells according to Dignam *et al.* (1983).

Human tissue was frozen at $-70°$. The frozen solid human brain tissue ($\sim$10 g) was homogenized in 40 ml of homogenization buffer (Fei *et al.*, 1995) with the use of a motor-driven Teflon pestle and glass homogenizer. The homogenate was layered over a 3-ml cushion of the same buffer and centrifuged at 38,000 rpm in a Sorvall TH-641 rotor for 30 min. The pellet (nuclei) was resuspended in the nuclei extraction buffer (10 mM HEPES-KOH, pH 7.9, 1.5 mM $MgCl_2$, 0.1 mM EGTA, 0.5 mM DTT, 5% glycerol, 400 mM NaCl, 0.5 mM phenylmethylsulfonyl fluoride) and slowly swirled at 4° for 30 min. The final dialysis solution contains 20 mM HEPES-KOH, pH 7.9, 75 mM NaCl, 0.1 mM EDTA, 0.5 mM DTT, and 20% glycerol. The protein concentration of the nuclear extract was determined according to Bradford (1986) using bovine serum albumin as standard.

EMSA. The experimental method was that of Scott *et al.* (1994) with slight modification. The single-stranded DNA oligonucleotide (20 pmol) was labeled radioactively in a 25-μl reaction containing 50 μCi of γ-^{32}P-ATP, kinase buffer, and 10 units of T4 kinase. The reaction was incubated at 37° for 10 min and stopped by heating at 70° for 10 min. The labeled DNA probe was purified on a spin column (Boehringer-Mannheim Biochemicals, Indianapolis, IN). To make a double-stranded probe, a portion of the labeled single-stranded probe was mixed with 1.5 molar excess of unlabeled complementary oligonucleotide plus 1 μl of 20× oligo annealing buffer (200 mM Tris·HCl, pH 7.9, 40 mM $MgCl_2$, 1 M NaCl, 20 mM EDTA) and water to a final volume of 20 μl. The mixture was heated to 90° for 5 min in water bath and allowed to cool slowly to room temperature. The probe was

purified by 10% nondenaturing polyacrylamide gel electrophoresis, eluted into 200 μl of extraction buffer (2.5 M ammonium acetate and 1 mM EDTA) at 37° for 2 hr with shaking, and ethanol precipitated.

The EMSA reaction mixture contains 10 mM Tris·HCl, pH 7.6, 50 mM KCl, 1 mM EDTA, 5% glycerol, 1 mM DTT, 1 μg of poly(dI/dC)·poly(dI/dC), 1 μl of ^{32}P-labeled DNA probe (500,000 cpm), and 0.5 μg (or as otherwise indicated in the figure legends) of nuclear protein extract in a final volume of 15 μl. In each experiment, a master reaction mix containing all common components were prepared to ensure sample uniformity. The above components except the labeled DNA probe were incubated at room temperature for 15 min. Then, the DNA probe was added, and incubation was continued for an additional 30 min. For the experiments using single-stranded labeled DNA as a probe, 5 pmol of another unlabeled single-stranded DNA (Luc-7, 5′-ACGACGATTCTGTGATTTGTATTCAGC-3′ from luciferase gene coding sequence) was included as a nonspecific competitor. Samples were directly loaded onto an 8% acrylamide gel in 0.5× TBE buffer (44.5 mM Tris, 44.5 mM boric acid, 1 mM EDTA). Substitution of poly(dI/dC)·poly(dI/dC) with poly(dA/dT)·poly(dA/dT) made little difference in band shift patterns. All experiments were repeated at least three times to ensure reproducibility.

For the effects of EGTA and zinc on the binding complex formation with single-stranded probes, 4 mM EGTA alone or EGTA plus 100 μM $ZnCl_2$ was included in the reaction mixture and preincubated with the nuclear protein extract before the addition of radiolabeled probes.

DNase I footprinting. For labeling of a DNA fragment, the plasmid nif/PmtA was digested with restriction enzymes *Eco*RI and *Nhe*I. A 1-kb fragment was isolated after agarose gel electrophoresis separation. The DNA fragment was purified by Qiaex DNA purification kit (Qiagen, Studio City, CA). Approximately 3.5 μg of this DNA fragment was used in a 50-μl Klenow fill-in reaction, containing 50 μCi each of [α-^{32}P]dATP and [α-^{32}P]dTTP, reaction buffer, and 1 μl of Klenow fragment (GIBCO BRL, Gaithersburg, MD). The labeling reaction was carried out at room temperature for 25 min; 2 μl of 2.5 mM concentration of each dNTP mix was added. The incubation continued for an additional 5 min and the labeled fragment was purified through a spin column. The DNA fragment was then digested with the restriction enzyme *Bsr*BI and the larger fragment was isolated and purified as described above. This single-end labeled DNA fragment was used for DNase I footprinting experiments.

In a 100-μl volume, $\sim$10–15 μg of human brain nuclear protein extract was mixed with reaction buffer containing 10 mM HEPES, pH 7.9, 10% glycerol, 50 mM KCl, 0.1 mM EDTA, 1 mM DTT, and 3 μg of poly(dI/dC) and incubated at room temperature for 20 min. Approximately 10,000 cpm labeled DNA fragment was added, and the binding reaction was incubated on ice for 40 min. After the mixture was allowed to warm to room temperature for 2 min, 5 μl of a solution containing 100 mM $MgCl_2$ and 50 mM $CaCl_2$ was added. An empirically adjusted amount of freshly diluted DNase I was then added, and the digestion proceeded at room temperature for 1 min. The reaction was stopped by the addition of 100 μl of a solution containing 100 mM Tris·HCl, pH 8.0, 0.6 M sodium acetate, 1% sodium dodecyl sulfate, 10 mM EDTA, and 50 μg/ml tRNA. The mixture was phenol/chloroform extracted, and DNA was recovered by ethanol precipitation. The partially digested DNA was fractionated on an 6% acrylamide/7 M urea gel alongside DNA sequencing ladders as the molecular standard.
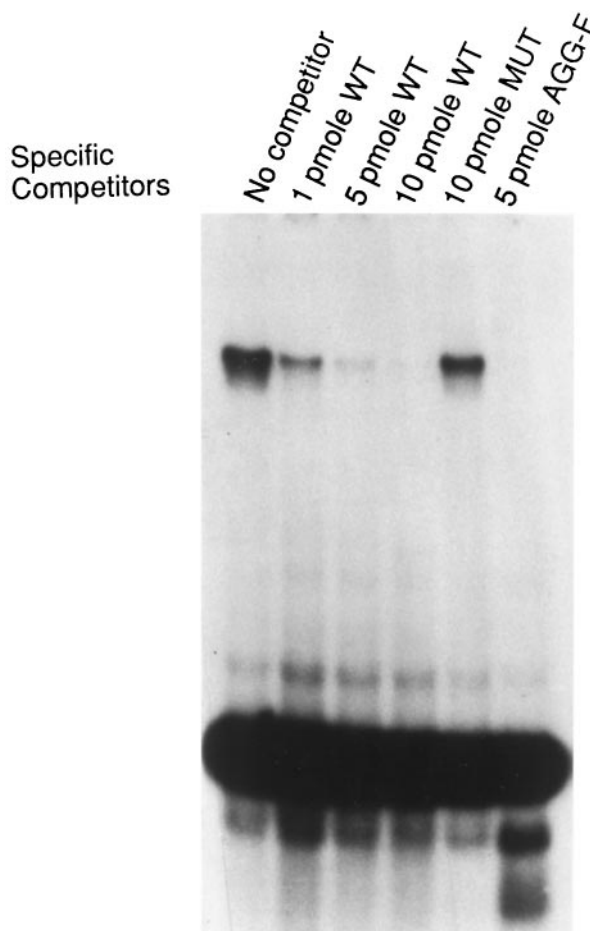
Results

EMSA showing specific binding of nuclear proteins to AGG elements. In our previous report, we indicated that the AGG repeat sequence near the TATA box and the transcriptional start site is responsible for a major portion of the transcriptional activity of promoter A in the human A_1 aden-
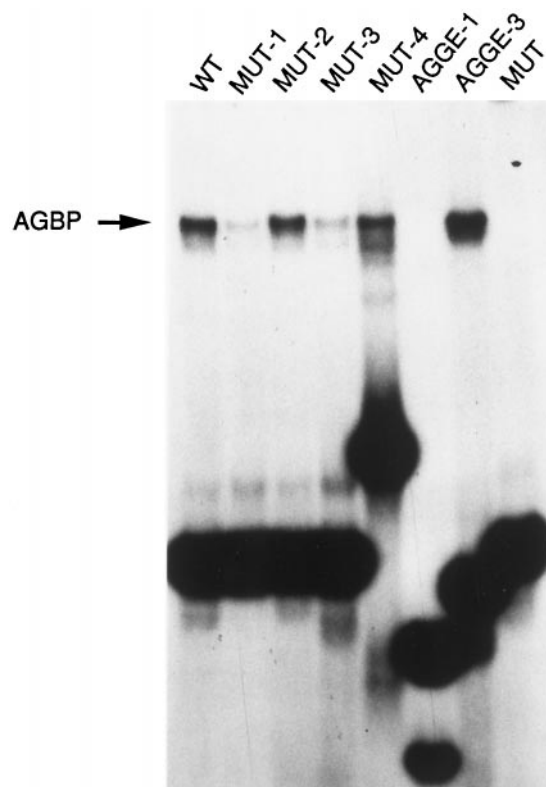
osine receptor gene when transfected transiently into Chinese hamster ovary cells (Ren and Stiles, 1995). To show whether any transcription factor or factors are bound to the sequence containing the AGG repeat, we conducted EMSAs. Fig. 1 shows that the nuclear protein extract from human brain tissue contains protein or proteins that bind to the DNA fragment containing the AGG elements and AGG repeat. This binding is sequence specific because the nonspecific competitor DNA, poly(dI/dC), which was included in all EMSA samples, did not compete for the binding. In contrast, the binding is effectively blocked by the addition of excess unlabeled oligonucleotide of the same sequence. When the AGG elements were mutated (probe MUT), the oligonucleotide lost its ability to block the binding (Fig. 1).

Specific mutations in the individual AGG elements and/or AGG repeats reduced or totally eliminated the ability of nuclear proteins to bind to their cognate sequences. Fig. 2 shows that when the first AGG element in the fragment was mutated (MUT-1), the binding was reduced, but mutation of the second AGG element alone in the fragment (MUT-2) did not result in any major change in binding.

The mutation of the AGG repeat (MUT-3) also reduced binding. The insertion of a 20-base random sequence between

Fig. 2. EMSA of WT and mutant double-stranded oligonucleotide probes. The DNA sequences of the probes are shown. *Underlined*, bases targeted for mutations in WT and mutated bases in mutants. The reaction conditions are described in Materials and Methods. The "free" unbound probes migrate (*bottom*), and their positions vary according to the size of the probe.

Fig. 1. EMSA of nuclear proteins from human brain binding to double-stranded WT probe. The ^{32}P-labeled oligonucleotide AGG-F was annealed with the complementary strand AGG-R and used as the WT probe for EMSA. The nonlabeled competitor is the unlabeled double-stranded WT, double-stranded MUT, or single-stranded oligonucleotide AGG-F. *Top*, quantity of specific competitors used in the reaction mixture. Assay conditions are described in Materials and Methods.

the first and second AGG elements (MUT-4) changed the natural sequence context of the fragment, but the binding complex was still observed along with some additional minor bands. The specificity of these minor bands remains unknown. A minor band also appears close to the bottom of the gel in WT, MUT-1, MUT-2, and MUT-3 (Figs. 1 and 2), which may represent nonspecific complex formation because it was not competed out by excess unlabeled probe (Fig. 1). When the sequence containing the first AGG element was separated from the remainder of the fragment (AGGE-1), there was no binding complex formation corresponding to the band in WT, but a band that migrates faster than the free probe is visible. The separation of the AGG repeat from the first AGG element (AGGE-3) resulted in somewhat enhanced binding.

Mutation of all of the AGG elements (MUT) totally abolished protein binding to this fragment.

These results indicate that the protein or proteins binding to this DNA fragment are sequence specific and related to individual AGG elements and their repeat sequence. Although the first AGG element by itself does not seem to form the binding complex seen with the WT probe, its presence is likely important to formation of the complex, as indicated in Fig. 2. However, the insertion of a 20-base unrelated sequence between the first and second AGG elements had only a minor effect on the formation of the major binding complex. The results in Fig. 2 suggest that the first AGG element and AGG repeats may play a major role in the binding complex formation.

The nuclear protein has higher affinity for single-stranded probe than for double-stranded probe. Although most of the reported transcription factors bind to specific sequence motifs in double-stranded DNA, a number of nuclear binding proteins have been described that prefer single-stranded DNA. These function as regulatory factors in the transcription process of protooncogenes such as c-*myc* (Duncan *et al*., 1994; Michelotti *et al*., 1995; Takai *et al*., 1994). In the c-*myc* gene promoter, there is an upstream regulatory sequence that binds a protein factor only on the noncoding strand. This protein factor is known to be required for c-*myc* expression (Duncan *et al*., 1994). However, this factor does not function as a traditional enhancer and is active only when working in concert with other *myc* promoter elements. Some single-stranded DNA binding proteins also bind to RNA, linking translation regulation with transcriptional regulation.

We next examined whether protein binding occurred to the single-stranded oligonucleotides containing AGG element and AGG repeats. Fig. 3A demonstrates that both the coding strand (AGG-F) and noncoding strand (AGG-R) of the WT probe bind to proteins in human brain nuclear extract. The binding complex containing AGG-F migrates at the same position as the binding complex containing double-stranded WT (Fig. 3B). Single-stranded oligonucleotide AGG-F competes effectively with the binding of proteins to the WT probe (Fig. 1), indicating that the same or very similar proteins bind both single- and double-stranded probes. There is a minor band migrating faster than the free probe of AGG-F (Fig. 3A), and it seems to be strengthened by the presence of excess unlabeled probe, indicating the possibility of nonspecific binding. Two distinct complexes form with the AGG-R probe and display distinct mobilities. The complex formed that has slower mobility and lower abundance may contain the same protein as that which is present in the major band but with a different conformation, or it could represent the binding of a different binding protein. The single-stranded binding is also specific because an unrelated single-stranded oligonucleotide (Luc-7) did not compete with the binding, whereas the same amount of unlabeled AGG-F completely blocked the binding (Fig. 3A). When samples of single- and double-stranded binding are run on the same gel, the double-stranded binding is almost negligible (Fig. 3B), indicating that under the exact reaction conditions and with the same amount of labeled probe, nuclear proteins bind the single-stranded probe much more effectively than the double-stranded probe.

The interaction of nuclear proteins with single-stranded
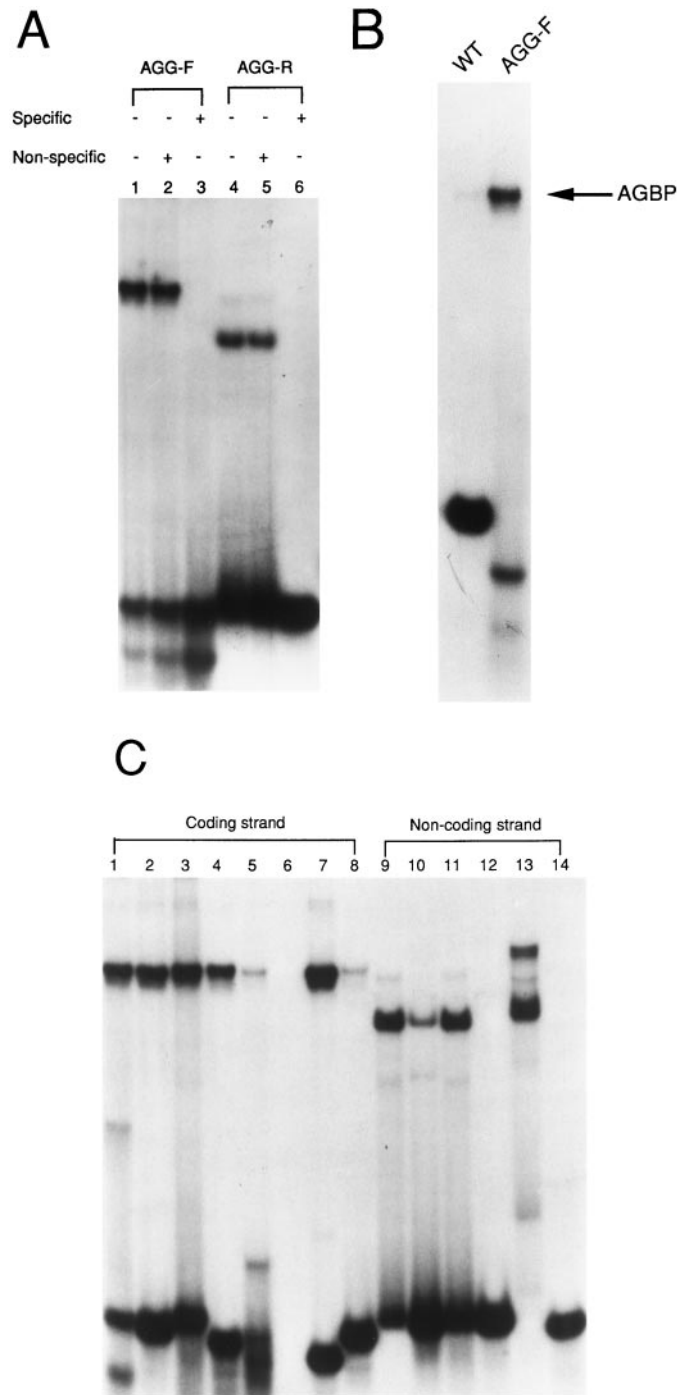
Fig. 3. EMSA of coding (AGG-F) and noncoding (AGG-R) single-stranded oligonucleotide probes. A, *Lanes 1–3*, coding (AGG-F). *Lanes 4–6*, noncoding (AGG-R). *Lanes 1 and 4*, no competitor. *Lanes 2 and 5*, an unrelated single-stranded oligonucleotide (Luc-7, 10 pmol) was preincubated with the human brain nuclear extract. *Lanes 3 and 6*, the unlabeled AGG-F or AGG-R (10 pmol) was preincubated with the nuclear proteins. B, Comparison of single- and double-stranded oligonucleotides: *left*, double-stranded WT; *right*, single-stranded AGG-F. Specific conditions are described in Materials and Methods. C, Single-stranded (WT and mutant) oligonucleotides. *Lanes 1–8*, coding strand (F). *Lanes 9–14*, noncoding strand (R). *Lanes 1 and 9*, WT. *Lanes 2 and 10*, MUT-1. *Lanes 3 and 11*, MUT-2. *Lanes 4 and 12*, MUT-3. *Lanes 5 and 13*, MUT-4. *Lane 6*, AGGE-1. *Lane 7*, AGGE-3. *Lanes 8 and 14*, MUT.

probes and its mutations produced quite different results from those seen with double-stranded probes and its mutations (Fig. 3C). Unlike the results seen with mutations in the

double-stranded experiments, protein binding to single-stranded sense mutants, such as MUT-1F, MUT-2F, and MUT-3F, did not seem to lead to any substantial loss of binding complex, but the amount of free probe remaining at the bottom of the gel indicates the change of binding affinity to the probes. For MUT-4F, the binding complex corresponding to that of WT was greatly reduced, and additional bands appear at the bottom of the gel. The probe AGGE-1F (Fig 3, *lane 6*) did not form the AGBP binding complex, and the small free probe migrated off the bottom of the gel. The probe AGGE-3F formed an enhanced binding complex (*lane 7*), whereas the total mutant probe MUT showed a barely visible band on coding strand (*lane 8*) and no band on noncoding strand (*lane 14*). Protein binding to the noncoding strand mutants MUT-1R, MUT-2R, and MUT-3R was similar to that was seen with the double-stranded binding in Fig. 2 (i.e., a reduction in MUT-1R binding and loss of binding in MUT-3R).

In the competition analysis using single-stranded DNA, the nuclear protein or proteins binding to the AGG-F motif were blocked by 3 pmol of the unlabeled AGG-F (Fig. 4) and was slightly reduced by unlabeled MUT-F, suggesting that
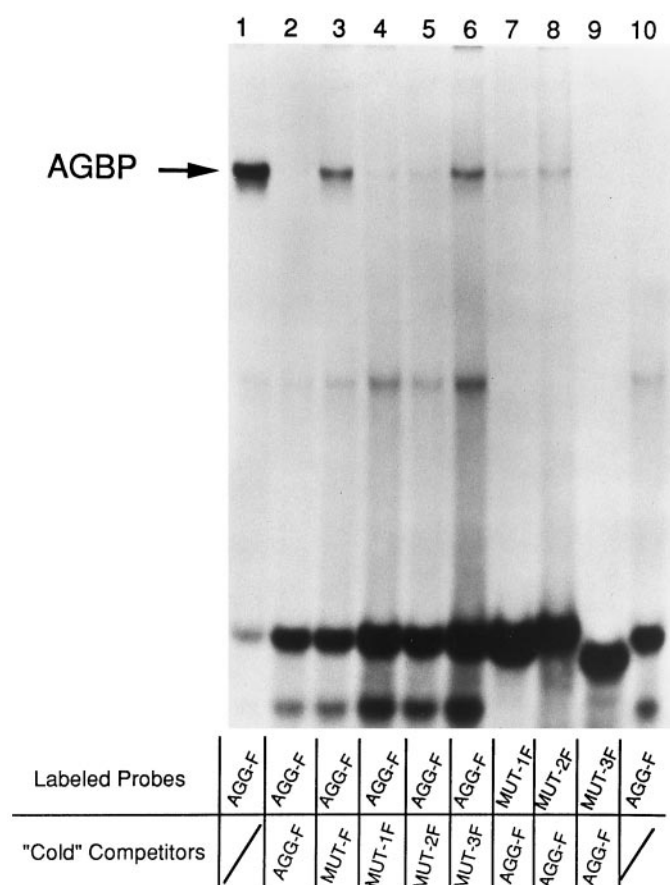
Fig. 4. EMSA of single-stranded oligonucleotides with various competitors. For the samples with competitors, 3 pmol of unlabeled oligonucleotide was mixed with 0.5 μg of human brain nuclear protein extract before the addition of the labeled probes: *lanes 1–6*, AGG-F. *Lane 7*, MUT-1F. *Lane 8*, MUT-2F. *Lane 9*, MUT-3F. Nonlabeled ("Cold") competitors: *lane 1*, no competitor. *Lane 2*, AGG-F. *Lane 3*, MUT-F. *Lane 4*, MUT-1F. *Lane 5*, MUT-2F. *Lane 6*, MUT-3F. *Lanes 7–9*, AGG-F. *Lane 10*, labeled AGG-F only without nuclear protein extract. The minor band below the AGBP resulted from an artifact of the labeled probe as seen in *lane 10* (no competitor; no protein).

the sequence around AGG elements may also play a role in binding. The same protein or proteins bound to AGG-F and mutant probes in that competition occurred when cold probes were included (Fig. 4). As one would expect, the mutation in AGG repeat makes the cold mutant probe MUT-3F a less effective competitor (Fig. 4, *lane 6*).

Tissue-specific expression of nuclear binding proteins. Because the nuclear protein extract from human brain exhibited strong binding to the single-stranded AGG elements in promoter A, we tested nuclear proteins from other human tissues, such as testis, prostate, and tongue muscle. The EMSA of nuclear proteins from those tissues are shown in Fig. 5A. Compared with the binding with proteins from brain, the nuclear proteins from those other tissues showed much less binding activity even though 3 μg of nuclear proteins from other tissues was used in the reaction compared with 0.5 μg of brain nuclear proteins. The use of 0.5 μg nuclear proteins from other tissues did not produce any binding complex (data not shown). The proteins from testis demonstrated detectable binding activity to the noncoding strand (AGG-R), although mobility differs compared with human brain proteins. In contrast, nuclear proteins from tongue muscle showed little binding to the coding strand and none to the noncoding strand. The nuclear proteins of prostate did not show any specific binding to the noncoding strand, and a faint, diffuse band appeared slightly below the AGBP band for the coding strand, which may or may not be the same binding complex as AGBP.

Nuclear proteins from various mammalian cell lines also were tested in EMSAs with single-stranded probes (Fig. 5B). When 0.5 μg of nuclear proteins was used in the reactions, no visible complexes were formed with the AGG-F probe, whereas complexes are detectable with the noncoding strand (AGG-R). The above results indicate that among all the tissues and cells tested, the nuclear proteins from human brain tissue are the most abundant source of binding proteins for the coding strand.

DNase I footprinting documents nuclear protein binding to the 5′ flanking sequence of human A$_1$ receptor promoter region. To demonstrate more directly the location of nuclear protein binding, a DNA fragment containing the sequence of $-513/+107$ relative to the transcriptional start site ($+1$) was radioactively labeled at the 3′ end of the coding strand. This probe was incubated with human brain nuclear protein extract, and after partial DNase I digestion, the DNA fragments were separated on the sequencing gel. The footprinting gel (Fig. 6) clearly shows that the sequence between TATA box (TTAAGA) and the AGG elements are protected by the DNA binding protein or proteins from the human brain nuclear extract (*lane 2*). When an excess amount of the competitor containing AGG elements (WT) was included in the reaction mixture, the protection was removed (*lane 3*). This result indicates the presence of binding protein or proteins in the human brain nuclear extract that specifically bind to the sequence immediately following the TATA box.

The AGBP is not related to SP1, and binding is not affected by EGTA and zinc. Recently, Handy and Gavras (1996) reported that in the rat α_{2A}-adrenergic receptor gene promoter, there are two variant GC boxes that bind the SP1 nuclear factor. The sequence of the variant GC box is very similar to the AGG repeat in the human A$_1$ receptor promoter; therefore, we tested the possibility that the AGG
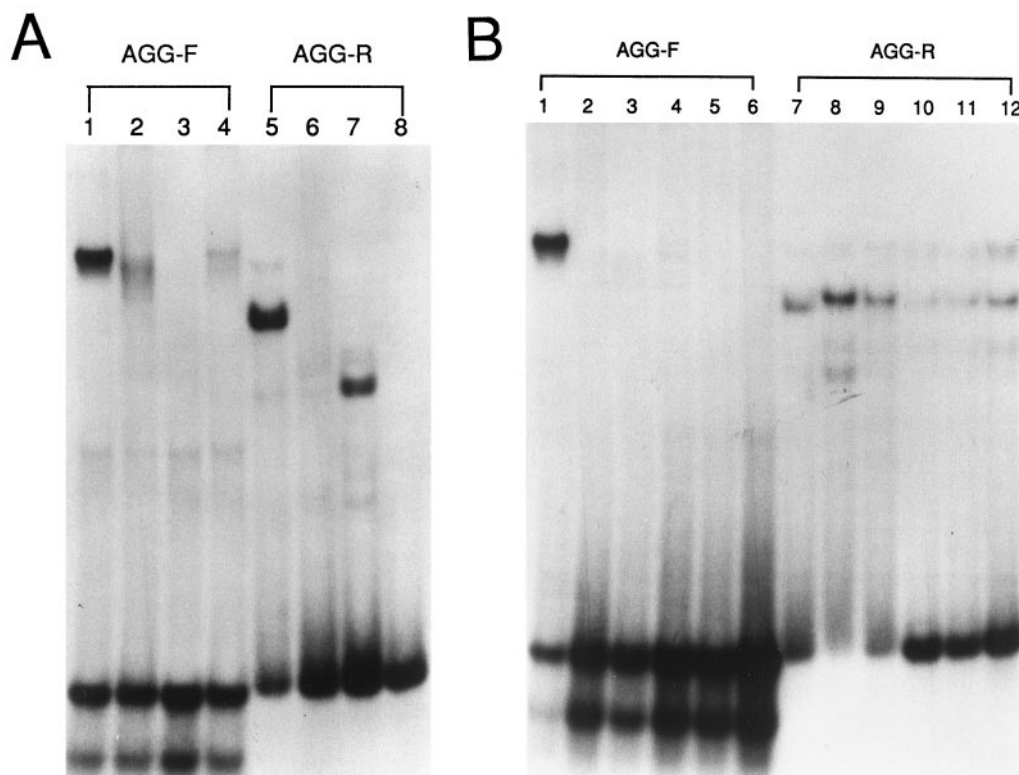
Fig. 5. EMSA of single-stranded oligonucleotides incubated with nuclear extracts from a variety of tissues and cell lines. A, Nuclear extracts from human tissues. *Lanes 1–4*, AGG-F. *Lanes 5–8*, AGG-R. *Lanes 1 and 5*, brain (0.5 μg). *Lanes 2 and 6*, prostate (3.0 μg). *Lanes 3 and 7*, testis (3.0 μg). *Lanes 4 and 8*, tongue muscle (3.0 μg). B, Nuclear extracts from human brain and cultured mammalian cells (0.5 μg of protein for each sample). *Lanes 1–6*, AGG-F. *Lanes 7–12*, AGG-R. *Lanes 1 and 7*, human brain. *Lanes 2 and 8*, SK-N-MC cells. *Lanes 3 and 9*, HeLa cells. *Lanes 4 and 10*, 293 cells. *Lanes 5 and 11*, CHO cells. *Lanes 6 and 12*, COS-7 cells.

binding protein from human brain nuclear extract is related to SP1. Fig. 7 shows that the oligonucleotide with SP1 consensus sequence did not compete with the binding to WT sequence, which indicates it is unlikely that the proteins that bind to the SP1 site are related to those binding to the AGG elements.

In another report, Michelotti *et al.* (1995) described a single-stranded binding protein present in HeLa cell extract that binds to the noncoding strand of the CT element of the human c-*myc* gene promoter. After purification and sequencing, this protein was identified as the cellular nucleic acid binding protein, a zinc finger protein previously shown to bind the purine-rich single-stranded of the sterol response element. Although the sequence of the purine-rich strand of CT element is somewhat different from the AGG repeat in A$_1$ receptor promoter, the effects were analyzed of EGTA and zinc on the binding of human brain nuclear proteins to the coding (AGG-F) and noncoding (AGG-R) strands of AGG-rich fragment. EGTA blocked the binding of the cellular nucleic acid binding protein to the purine-rich strand of the CT element and zinc restored the binding (Michelotti *et al.*, 1995), but both EGTA and zinc had no effect on the AGG binding protein (Fig. 8).

Discussion

Human A$_1$ adenosine receptor gene expression is under the control of two separate promoters, and specific AGG elements are important in transcriptional activity. We have identified specific regulatory sequences flanking the promoter and nuclear protein factors that are responsible for the control of transcription.

DNase I footprinting is a commonly used method to locate where specific nuclear proteins bind to DNA sequences. The

footprinting of a DNA fragment containing promoter A of the A$_1$AR and its flanking sequence showed that the sequence between TATA box and the transcriptional start was partially protected from DNase digestion by the inclusion of nuclear binding proteins from human brain. Protection was eliminated by the inclusion of an oligonucleotide containing these AGG elements in the reaction. This result indicates the presence of a specific AGG binding protein in the human brain nuclear extract. The results obtained from EMSA help explain why there was only partial protection; this is likely the case because there is only low affinity binding or incomplete binding to the double-stranded DNA by the AGBP.

EMSA with human brain nuclear extracts clearly show that the binding is specific because there is no competition from nonspecific DNA but binding is blocked by unlabeled DNA of the same sequence. Furthermore, specific mutations made in the AGG elements reduced or abolished the binding. A 20-base random sequence inserted between the first and second AGG elements and the AGG repeat (probe MUT-4) changed the natural sequence context, but the major binding complex is still present, although at a slightly reduced level with a few new minor bands appearing. The EMSA results are consistent with our previous findings showing that those AGG elements that bind to AGBP account for >80% of the promoter activity (Ren and Stiles, 1995).

Investigation of the AGBP produced some surprising results. The AGBP actually had much higher binding affinity for the single-stranded than the double-stranded DNA probes, as shown in Fig. 3B. The AGBP bound much more effectively to AGG-F than to WT under the same reaction conditions. Effective mutual competition in binding between AGG-F and mutant probes shown in Fig. 4 suggests that the same protein or similar proteins bind to these probes. The proteins that bind to coding and noncoding strands seem to
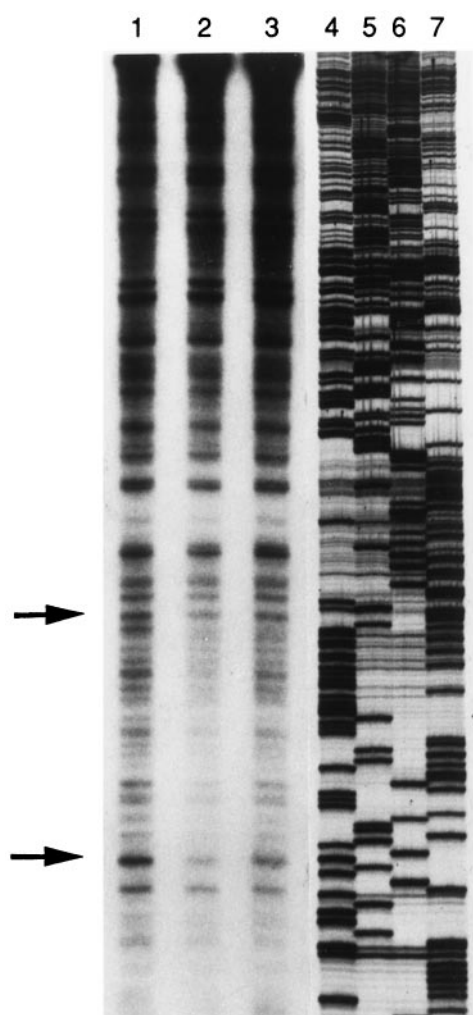
Protected DNA sequence between arrows:

5'-TTGTTAAGAGGAGGGGGTGGGGGAGG

AGACGGAGGATGAGGAGGGAGGGGCT-3'

Fig. 6. DNase I footprinting. The DNA fragment labeling and DNase I footprint procedures are described in Materials and Methods. *Lane 1*, labeled DNA fragment without added human brain nuclear extract. *Lane 2*, DNA with human brain nuclear extract. *Lane 3*, the same as *lane 2* except the nuclear extract was preincubated with competitor oligonucleotide. *Lanes 4–7*, DNA sequencing ladder used as molecular standard to count the bases. *Sequence between arrows*, protected by the human brain nuclear protein extract; this protection included the sequence between the TATA box and the transcriptional start site.


Fig. 7. Competition of SP1 consensus sequence oligonucleotide probe. The ³²P-labeled double-stranded WT and human brain nuclear extracts were used in all samples. *Lane 1*, no competitor. *Lane 2*, nonlabeled WT (15 pmol). *Lane 3*, nonlabeled SP1 consensus sequence oligonucleotide (15 pmol).

be different, or the protein/DNA complex has a different conformation. Proteins binding to the double-stranded DNA probes (especially MUT-1 and MUT-3) may be influenced by the presence of noncoding strand (Figs. 2 and 3C).

Compared with the nuclear protein extract from human brain, the nuclear proteins from other human tissues and mammalian cell lines have much less, if any, AGBP, indicating the possible tissue-specific expression of AGBP. However, we cannot rule out that AGBP inhibitors are present in or that certain cofactors required for AGBP binding are missing from the other tissues and cell lines. Because the noncoding

strand is CT rich, the protein binding to this strand should be named CT element binding protein. The binding to CT elements is also specific because it is sensitive to mutations in the sequence and not competed for by the unrelated single-stranded DNA.

In our previous study, we showed that among a number of tissues or organs, human brain has the highest level of transcript A and its production is under the control of AGG elements (Ren and Stiles, 1994a) and that brain tissue is a good source of A₁ adenosine receptors for ligand binding studies. In this report, we showed that human brain is an especially good source of AGBP, suggesting that it may function to increase transcription.

A search of the literature for similar DNA sequences and binding proteins revealed several interesting reports. The rat c-*neu* oncogene promoter, described by Suen and Hung (1991), has two closely connected AGG repeats downstream of a CCAAT box and a GC box. This sequence resembles a reversely oriented nuclear factor binding domain in the EGF-r promoter. SP1 and another specific TC factor can bind
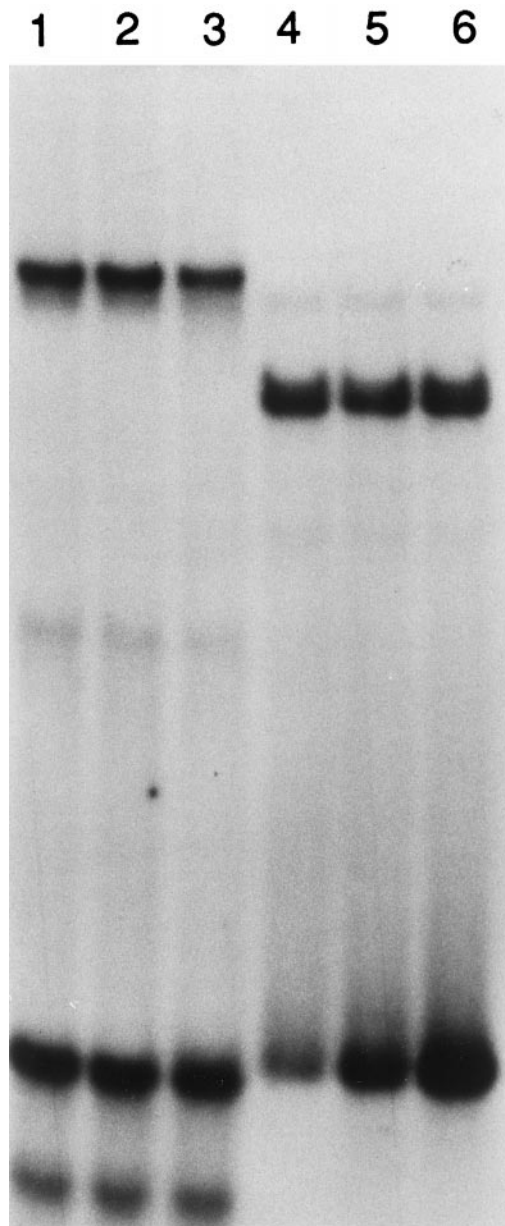
Fig. 8. Effects of EGTA and zinc on the binding complex formation. Radiolabeled single-stranded probes AGG-F (*lanes 1–3*) and AGG-R (*lanes 4–6*) are used. The reaction conditions are described in Materials and Methods. *Lanes 1 and 4*, no EGTA or zinc. *Lanes 2 and 5*, 4 mM EGTA. *Lanes 3 and 6*, 4 mM EGTA and 0.1 mM ZnCl₂.

to this domain and stimulate the transcription of EGF-r. However, both a DNA fragment from EGF-r promoter and SV40 promoter containing AGG repeat and six GC boxes failed to compete for protein binding on the c-*neu* promoter. In a related study, Johnson *et al.* (1988) showed that the human EGF-r promoter has four TCCTCCTCC repeats, a reversal of the AGG elements, and that those regions are sensitive to S1 nuclease, indicating they are single-stranded under certain conditions. They also showed that this part of promoter sequence is required for optimal transcriptional activity and a nuclear factor TCF was bound to it. However, they did not use single-stranded DNA probes for EMSA. A single-stranded DNA binding protein has been identified to bind to the thyrotropin receptor gene promoter (Shimura *et*

al., 1995) and is required for full expression of the receptor gene. Recently, Li *et al.* (1996) reported the gene structure and promoter activity of human thrombin receptor, which is another G protein-coupled receptor, and the most active promoter sequence in their report includes a string of AGG repeat. In addition, there are more recent reports about specific single-stranded DNA binding proteins involving in gene expression regulations of c-*myc* (Duncan *et al.*, 1994; Takai *et al.*, 1994), human granulocyte-macrophage colony-stimulating factor (Coles *et al.*, 1996), and catalase (Ito *et al.*, 1994).

We tested the possibility that AGBP is related to SP1 or requires a metal cofactor. The results in Figs. 7 and 8 show that the DNA probe containing the SP1 consensus sequence failed to compete for the binding to AGG elements in A_1 receptor promoter and that EGTA and zinc did not affect any binding activity. Therefore, those possibilities seem to be ruled out.

From all the data we have collected thus far concerning the AGBP, it is clear that AGBP is a specific single-stranded DNA binding protein that is particularly abundant in human brain tissue and that its cognate AGG sequence in the A_1 receptor promoter is necessary for the full expression of the gene. The regulation of human A_1 adenosine receptor gene is probably more complex, and multiple regulatory binding proteins are likely involved, as our previous study of promoter activity suggested. We will continue to explore the mechanism of human A_1 adenosine receptor gene expression.

References

Aguilar JS, Tan F, Durand I, and Green RD (1995) Isolation and characterization of an avian A_1 adenosine receptor gene and a related cDNA clone. *Biochem J* **307**:729–734.

Bhattacharya S, Dewitt DL, Burnatowska-Hledin M, Smith WL, and Spielman WS (1993) Cloning of an adenosine A_1 receptor-encoding gene from rabbit. *Gene* **128**:285–288.

Bradford MM (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**:248–254.

Chu Y, Tu K, Lee Y, Kuo Z, Lai H, and Chern Y (1996) Characterization of the rat A_{2a} adenosine receptor gene. *DNA Cell Biol* **15**:329–337.

Coles LS, Diamond P, Occhiodoro F, Vadas MA, and Shannon MF (1996) Cold shock domain proteins repress transcription from the GM-CSF promoter. *Nucleic Acids Res* **24**:2311–2317.

Dignam JD, Lebovitz RM, and Roeder RB (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* **11**:1475–1489.

Duncan R, Bazar L, Michelotti G, Tomonaga T, Krutzsch H, Avigan M, and Levens D (1994) A sequence-specific, single-strand binding protein activates the far upstream element of c-*myc* and defines a new DNA-binding motif. *Genes Dev* **8**:465–480.

Fei Y, Matragoon S, and Liou GI (1995) Simple and efficient method for the preparation of nuclear extracts. *Biotechniques* **18**:984–987.

Handy DE and Gavras H (1996) Evidence for cell-specific regulation of transcription of the rat α_{2A}-adrenergic receptor gene. *Hypertension* **27**:1018–1024.

Ito K, Sato K, and Endo H (1994) Cloning and characterization of a single-stranded DNA binding protein that specifically recognizes deoxycytidine stretch. *Nucleic Acids Res* **22**:53–58.

Jacobson MA, Johnson RG, Luneau CJ, and Salvatore CA (1995) Cloning and chromosomal localization of the human A_{2b} adenosine receptor gene (ADORA2B) and its pseudogene. *Genomics* **27**:374–376.

Johnson AC, Jinno Y, and Merlino GT (1988) Modulation of epidermal growth factor receptor proto-oncogene transcription by a promoter site sensitive to S1 nuclease. *Mol Cell Biol* **8**:4174–4184.

Le F, Townsend-Nicholson A, Baker E, Sutherland GR, and Schofield PR (1996) Characterization and chromosomal localization of the human A_{2a} adenosine receptor gene: ADORA2A. *Biochem Biophys Res Commun* **223**:461–467.

Lee S, Minowa MT, and Mouradian MM (1996) Two distinct promoters drive transcription of the human D_{1A} dopamine receptor gene. *J Biol Chem* **271**:25292–25299.

Li F, Baykal D, Horaist C, Yan C, Carr BN, Rao GN, and Runge MS (1996) Cloning and identification of regulatory sequences of the human thrombin receptor gene. *J Biol Chem* **271**:26320–26328.

Libert F, Schiffmann SN, Lefort A, Parmentier M, Gerard C, Dumont JE, Vanderhaeghen J, and Vassart G (1991) The orphan receptor cDNA RDC7 encodes an A_1 adenosine receptor. *EMBO J* **10**:1677–1682.

Libert F, Van Sande J, Lefort A, Czernilofsky A, Dumont JE, Vassart G, Ensinger

HA, and Mendla KD (1992) Cloning and functional characterization of a human A$_1$ adenosine receptor. *Biochem Biophys Res Commun* **187**:919–926.

Michelotti EF, Tomonaga T, Krutzsch H, and Levens D (1995) Cellular nucleic acid binding protein regulates the CT element of the human c-*myc* protooncogene. *J Biol Chem* **270**:9494–9499.

Murrison EM, Goodson SJ, Edbrooke MR, and Harris CA (1996) Cloning and characterization of the human adenosine A$_3$ receptor gene. *FEBS Lett* **384**:243–246.

Olah ME, Ren H, Ostrowski J, Jacobson KA, and Stiles GL (1992) Cloning, expression, and characterization of the unique bovine A$_1$ adenosine receptor. *J Biol Chem* **267**:10764–10770.

Ren H and Stiles GL (1995) Separate promoters in the human A$_1$ adenosine receptor gene direct the synthesis of distinct messenger RNAs that regulate receptor abundance. *Mol Pharmacol* **48**:975–980.

Ren H and Stiles GL (1994a) Characterization of the human A$_1$ adenosine receptor gene: evidence for alternative splicing. *J Biol Chem* **269**:3104–3110.

Ren H and Stiles GL (1994b) Posttranscriptional mRNA processing as a mechanism for regulation of human A$_1$ adenosine receptor expression. *Proc Natl Acad Sci USA* **91**:4864–4866.

Reppert SM, Weaver DR, Stehle JH, and Rivkees SA (1991) Molecular cloning and characterization of a rat A$_1$ adenosine receptor that is widely expressed in brain and spinal cord. *Mol Endocr* **5**:1037–1048.

Scott V, Clark AR, and Docherty K (1994) The gel retardation assay, in *Methods in Molecular Biology* (Harwood AJ, ed) pp 339–347, Humana Press, Totowa, NJ.

Shimura H, Shimura Y, Ohmori M, Ikuyama S, and Kohn LD (1995) Single strand DNA-binding proteins and thyroid transcription factor-1 cojointly regulate thyrotropin receptor gene expression. *Mol Endocr* **9**:527–539.

Suen T and Hung M (1991) c-*myc* reverses neu-induced transformed morphology by transcriptional repression. *Mol Cell Biol* **11**:354–362.

Takai T, Nishita Y, Iguchi-Ariga SM, and Ariga H (1994) Molecular cloning of MSSP-2, a c-*myc* gene single-strand binding protein: characterization of binding specificity and DNA replication activity. *Nucleic Acids Res* **22**:5576–5581.

Tucker AL, Linden J, Robeva AS, D'Angelo DD, and Lynch KR (1992) Cloning and expression of a bovine adenosine A$_1$ receptor cDNA. *Fed Eur Biochem Soc* **297**:107–111.

Zhou Q-Y, Li C, Olah ME, Johnson RA, Stiles GL, and Civelli O (1992) Molecular cloning and characterization of an adenosine receptor: the A$_3$ adenosine receptor. *Proc Natl Acad Sci USA* **89**:7432–7436.