

Novel Polymorphisms Influencing Transcription of the Human CHRM2 Gene in Airway Smooth Muscle

Anthony G. Fenech*, Charlotte K. Billington*, Caroline Swan, Susan Richards, Therese Hunter, Martin J. Ebejer, Alex E. Felice, Roger Ellul-Micallef, and Ian P. Hall

Department of Clinical Pharmacology and Therapeutics, and Department of Physiology and Biochemistry, University of Malta, Msida, Malta; and Division of Therapeutics, University of Nottingham, Nottingham, United Kingdom

Muscarinic receptors are a functionally important family of G-protein-coupled receptors. Using a combination of rapid amplification of 5' cDNA ends and reporter gene assays, we characterized the 5' untranslated region of the CHRM2 gene as expressed in human airway smooth muscle (HASM) cells. A splice site is present 46 bp upstream from the ATG start codon. Five exons with alternative splicing patterns are present upstream of this splice site, separated by introns ranging from 87 bp to > 145 kb. There is evidence for the gene being under the control of a TATA-less promoter with Sp1, GATA, and activator protein-2 binding sites. Multiple transcription start sites (TSSs) were identified. We identified a novel 0.5-kb hyper-variable region located 648 bp upstream of the most 5' TSS, a multiallelic (CA) tandem repeat 96 bp downstream of the most 5' TSS, and a common C→A SNP located 136 bp upstream of the most 5' TSS. Functional studies in primary HASM cells and the BEAS-2B cell line demonstrated highest promoter activity to be upstream of the most 3' TSS, with potential repressor elements operating in a cell type-dependent manner, located upstream of the most 5' TSS. We present functional data to show that the CA repeat may influence the transcription of the gene in HASM and BEAS-2B cells.

The human muscarinic M₂ receptor is a functionally important membrane-bound protein that is negatively coupled to adenylyl cyclase and is widely expressed in a range of tissues, including cardiac myocytes, prejunctional cholinergic nerve endings, and smooth muscle. In the airways, M₂ receptors are constitutively expressed on airway smooth muscle (ASM) and, in contrast to muscarinic M₃ receptors, continue to be expressed during primary culture (1). Muscarinic M₂ receptors play an important role in the control of ASM cAMP regulation, being responsible for acetylcholine-mediated inhibition of adenylyl cyclase activity. This action ameliorates in part the relaxant effect of agents such as isoproterenol in the airways (2).

Despite the cloning of the human muscarinic M₂ receptor gene several years ago (3, 4) the elements important for tran-

scriptional regulation of the gene in ASM have not so far been identified. In recent years, knowledge of the regulatory regions of the different muscarinic genes has been enhanced due to studies in other species (rat M₁ [5], chicken M₂ [6], porcine M₂ [3, 4], rat M₄ [7]).

In the human genome, the promoter region has been identified for the M₁ receptor subtype (5) and the M₃ subtype (8). Zhou and coworkers recently described a muscarinic M₂ promoter arrangement based on work performed on mRNA transcripts obtained from human heart muscle and the human neuroblastoma cell line IMR-32 (ATCC No. CCL-127) (9). In the majority of studies on different muscarinic receptor subtypes, the promoter is spatially separated from the coding exon by at least one noncoding exon and more than 4.4 kb of intronic sequence. To date, the human muscarinic M₃ 5' untranslated region (UTR) is reported to be the most complex. Findings by Forsythe and colleagues show the existence of seven upstream exons ranging from 99–298 bp and introns of lengths ranging between 171 bp and > 65.2 kb (8).

All the muscarinic gene regulatory regions studied to date appear to be TATA-less promoters, containing consensus binding sites for a number of transcription factors reported to be important in promoter activity (10). Consistent with being a TATA-less promoter, transcription of the chick muscarinic M₂ gene is initiated at 5 or more sites within a 146-bp segment located 321 bp upstream from the 5'-end of the 5'UTR. These sites are preferentially used in a cell-specific manner (6).

In this article we describe a series of studies in which the upstream arrangement of the human muscarinic M₂ receptor gene was characterized and the putative promoter regions identified by studying mRNA transcripts obtained from primary cultures of human ASM (HASM) cells. In addition, we define novel polymorphic regions that have the potential to influence transcriptional activity of the M₂ receptor gene.

Materials and Methods

5'RACE Procedure

Preparation of RNA. Primary cultures of HASM cells were prepared from explants of trachealis muscle obtained from individuals without respiratory disease within 12 h of death as previously described (11). Following the establishment of a confluent monolayer of HASM cells in a 75 cm² flask, cells were harvested by trypsinisation and pelleted by centrifugation, as previously described (11). Total RNA was then extracted from the cells using the RNeasy Qiagen kit. RNA samples were then DNase treated using DNase I (Gibco-BRL, Paisley, UK) to eliminate potentially contaminating DNA.

5' Rapid amplification of cDNA ends. 5' rapid amplification of 5' cDNA ends (5'RACE) of HASM cell cDNA was performed using the GeneRacer system (Invitrogen Life Technologies, Paisley, UK) according to the manufacturer's recommended protocol. In contrast to conventional 5'RACE, the GeneRacer system ensures that only mature capped mRNA transcripts participate in the reaction. In brief, total RNA extracted from HASM cells was treated with calf intestinal phosphatase (CIP) to dephosphorylate the 5' ends of truncated mRNA and non-mRNA molecules and thereby inhibit their participation in downstream

(Received in original form January 12, 2003 and in revised form September 21, 2003)

No part of the research presented in this paper has been funded by tobacco industry sources.

*These authors contributed equally to the work described in this paper.

Address correspondence to: Prof. Ian P. Hall, Head, Division of Therapeutics, C Floor, South Block, University Hospital, Nottingham NG7 2UH, UK. E-mail: ian.hall@nottingham.ac.uk

Abbreviations: rapid amplification of 5' cDNA ends, 5'RACE; 5' untranslated region, 5'UTR; activating protein-2, AP-2; airway smooth muscle, ASM; calf intestinal phosphatase, CIP; chronic obstructive pulmonary disease, COPD; cyclic AMP-responsive element binding protein, CREB; G-protein-coupled receptor, GPCR; human airway smooth muscle, HASM; polymerase chain reaction, PCR; pGL3 control plasmid, pGL3C; pGL3 enhancer plasmid, pGL3E; single nucleotide polymorphism, SNP; tobacco acid pyrophosphatase, TAP; transcription start sites, TSS.

This article has an online data supplement, which is accessible from this issue's table of contents online at www.atsjournals.org

Am. J. Respir. Cell Mol. Biol. Vol. 30, pp. 678–686, 2004
Originally Published in Press as DOI: 10.1165/rcmb.2003-00110C on September 25, 2003
Internet address: www.atsjournals.org

reactions. The dephosphorylated RNA was treated with tobacco acid pyrophosphatase (TIP) to remove the 5' cap structure from mature full-length mRNA, leaving a 5' phosphate group available. A manufacturer-supplied RNA oligonucleotide (5'-CGA CUG GAG CAC GAG GAC ACU GAC AUG GAC UGA AGG AGU AGA AA-3') was ligated to the decapped mRNA using T₄ RNA ligase, thus providing a priming site for later amplification procedures. First-strand cDNA synthesis was performed using SuperScript II reverse transcriptase and a manufacturer-supplied oligo dT primer (5'-GCT GTC AAC GAT ACG CTA CGT AAC GGC ATG ACA GTG (T)₁₈-3'). Double-stranded cDNA was subsequently prepared by amplifying the first strand product using a reverse gene-specific primer (reverse GSP) designed to be within the M₂ receptor gene coding region, and the GeneRacer 5' primer (homologous to the previously ligated RNA oligonucleotide). A second nested polymerase chain reaction (PCR) was performed using a manufacturer-supplied 5' nested primer and a reverse nested gene-specific primer (reverse nested GSP) (see Table 1 for primer details). This step increased the product yield available for later cloning, and in addition eliminated any potential nonspecific products of the first PCR. The nested PCR products were gel-purified and cloned as described below. The whole 5'RACE procedure was performed twice using different populations of HASM cells.

Cloning. The pCR4-TOPO vector (Invitrogen Life Technologies) was used for TA cloning. The purified nested PCR product mixture was ligated to pCR4-TOPO by incubation at 22°C for 5 min in the presence of 200 mM NaCl and 10 mM MgCl₂. The ligation mixture was then used to transform One Shot TOP10 chemically competent *Escherichia coli* cells (Invitrogen Life Technologies) using heat shock for 30 s at 42°C. The cells were grown in SOC broth for 1 h at 37°C, after which they were plated onto Lennox Broth (LB) agar plates containing ampicillin for selection of positive transformants, and incubated overnight at 37°C. Several colonies were subsequently picked and grown in selective LB broth at 37°C with shaking (250 rpm) overnight. Each broth (5 μl) was streaked out on a selective LB agar plate, and the plates were incubated at 37°C overnight. One colony was picked from each plate, and grown in 5 ml selective LB broth at 37°C with shaking (250 rpm) overnight.

Plasmid DNA extraction. Plasmid DNA was extracted from each *E. coli* broth using Wizard Plus SV Minipreps DNA Purification System (Promega Corporation, Madison, WI), using sterile deionized water to elute the final product. An aliquot of each extract was restricted using EcoRI to excise the insert, and viewed on a 1% agarose gel. Successful extracts were submitted for sequencing.

Sequencing. DNA dideoxy sequencing was performed using an ABI 377 Genetic Analyzer and the ABI Big Dye Terminator Ready Reaction Mix (Applied Biosystems, Foster City, CA). Each clone was sequenced in both directions using vector primers M13-forward and M13-reverse, respectively.

Sequence data analysis. The data obtained from sequencing was analyzed by comparing it to the NCBI human genome database, using the Basic Local Alignment Search Tool available at <http://www.ncbi.nlm.nih.gov/BLAST/>.

TABLE 1. Oligonucleotide primers used for 5'RACE analysis

Primer	Sequence
M ₂ RACE GSP (reverse)	5'-GGT CTG GAG GTG GCG GTT GAC TT-3'
GeneRacer 5' primer (forward)	5'-CGA CTG GAG CAC GAG GAC ACT GA-3'
M ₂ RACE nested GSP (reverse)	5'-GCC ACC AGG ACA ATA AAC ACC ACT TC-3'
GeneRacer 5' nested primer (forward)	5'-GGA CAC TGA CAT GGA CTG AAG GAG TA-3'

The initial PCR was performed using the manufacturer-supplied GeneRacer 5' forward primer and our designed M₂ RACE reverse gene-specific primer (GSP). A second nested PCR was performed using the manufacturer-supplied GeneRacer 5' nested primer, and our designed M₂ RACE nested GSP.

Preparation of Promoter Deletion Constructs and Reporter Assay Analysis

PCR. Suitable regions for promoter activity analysis were selected, based on information obtained from the analysis of the 5'RACE results (Figure 1). Regions upstream of each of the three identified major transcription start sites were investigated. Each region was amplified from human genomic DNA (extracted from whole blood using the Qiagen DNA Blood Mini kit [Qiagen, West Sussex, UK]) using PCR primers with restriction site consensus sequences for *Mlu*I and *Xho*I built into the terminal regions of the oligonucleotides. This was necessary to enable subsequent directional cloning into the pGL3E firefly luciferase reporter vector (Promega Corporation). Tables 2 and 3 give the oligonucleotide primer sequences and PCR conditions used. All PCR reactions were performed using 0.5 μM of each primer; 0.2 mM each of dATP, dCTP, dGTP, and dTTP; ~ 0.3 μg of genomic DNA; and 2.6 U of High Fidelity polymerase (Roche Molecular Biochemicals, Indianapolis, IN). A 5-min hotstart at 95°C was followed by a 1-min denaturation, 1-min annealing, and 2-min extension at 72°C. Thirty PCR cycles were performed per reaction.

Ligation reactions. PCR products were gel purified using Strataprep minicolumns (Stratagene, La Jolla, CA), and were double-digested with *Mlu*I and *Xho*I at 37°C for 3 to 6 h to produce sticky-ended products. pGL3E plasmid was treated in the same way. All restriction products were gel-purified once more and their concentrations were estimated by electrophoresing on a 1% agarose gel, together with a quantitative DNA ladder. Ligation reactions were set up using T₄ DNA ligase (Promega Corporation), while maintaining an insert:vector molar ratio of 3:1. Ligations were performed overnight at 4°C.

Transformation and cell culturing. Chemically competent DH5α cells were removed from -80°C cold storage, and allowed to thaw on ice for ~ 5 min. Each ligation reaction (5 μl) was added to 50 μl of cells, gently mixed, and kept on ice for 20 min. The cell suspensions were then heat-shocked at 42°C for 45 s, after which they were returned to ice for 2 min. Ampicillin-free LB broth (950 μl) was added to each cell suspension, and the suspensions were incubated at 37°C and 150 rpm for 1.5 h. Each transformant suspension (150 μl) was then plated onto selective ampicillin-containing LB agar plates, and incubated overnight at 37°C. The next day, colonies were picked from each plate and grown overnight in 5 ml ampicillin-containing LB broths at 37°C and 250 rpm. Each broth (5 μl) was streaked to a single colony on ampicillin-containing LB agar plates, and the plates were incubated overnight at 37°C. One colony was picked from each plate, and grown in a 5-ml ampicillin-containing LB broth at 37°C and 250 rpm overnight.

Plasmid DNA extraction. Plasmid DNA was extracted from each *E. coli* broth using Wizard Plus SV Minipreps DNA Purification System (Promega Corporation), using sterile deionized water to elute the final product. An aliquot of each extract was restricted using *Mlu*I and *Xho*I to excise the insert, and viewed on a 1% agarose gel. Successful plasmids were submitted for confirmatory sequencing to ascertain that the correct insert was cloned.

Preparation of plasmids for transfection. Large-scale plasmid preparation was performed by inoculating 100 ml ampicillin-containing broths with 100–200 μl of broth from a 5-ml culture, and incubating overnight at 37°C with shaking at 250 rpm. Plasmid maxipreps were then prepared using the Plasmid Maxiprep (Qiagen) purification kit, according to manufacturer-recommended instructions. At the end of the procedure, an extra ethanol precipitation step was performed under sterile conditions, to ensure optimum purity and sterility of the plasmid DNA obtained.

Culturing of mammalian cells. HASM cells and BEAS-2B cells (ATCC number CRL-9609) were cultured in 75 cm² flasks until confluent, using high-glucose Dulbecco's modified Eagle's medium containing 10% heat-inactivated bovine calf serum and 4 mM L-glutamine. The cells were harvested by trypsinization, followed by centrifugation and resuspension in growth medium. The cell density was adjusted to 10⁵ cells/ml, and 0.5 ml of cell suspension was transferred to each well of a 24-well tissue culture plate, thus giving a cell count of 5 × 10⁴ cells/well. Sufficient tissue culture plates were prepared to allow for transfections with each different plasmid to be performed in triplicate on the same plate, together with negative and positive controls. The plates were incubated at 37°C and 5% CO₂, until ~ 70–80% confluent.

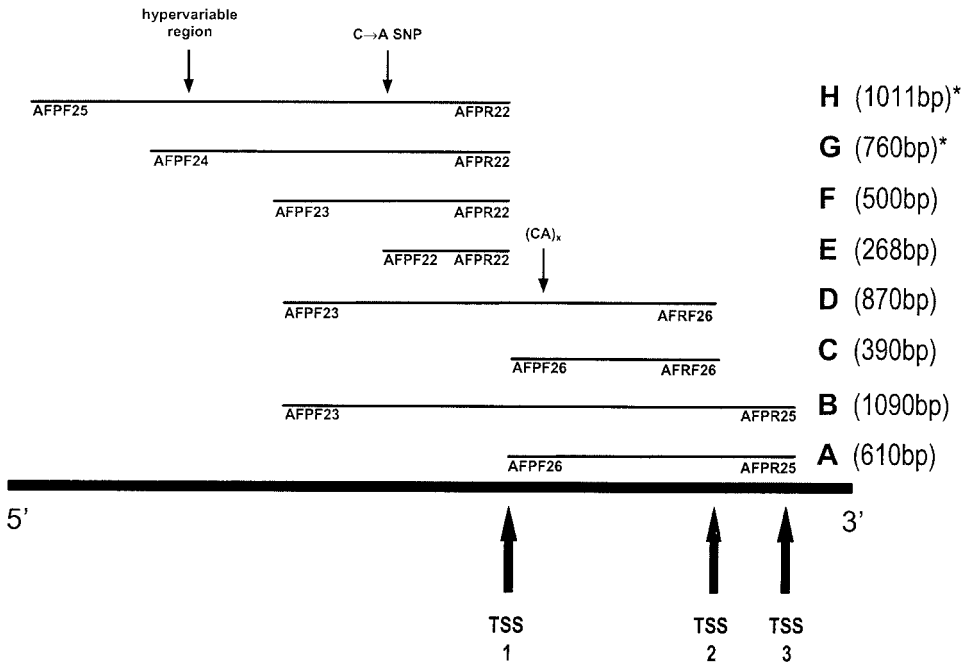


Figure 1. Regions of the muscarinic M₂ receptor 5'UTR amplified for pGL3E cloning. (*These fragment lengths are theoretical and are based on NCBI human genome data. The actual fragment lengths for these two constructs were actually 1,550 bp and 1,300 bp due to the presence of the novel 540-bp region.)

Transfection. Transient transfection was performed using Fugene 6 reagent (Roche Molecular Biochemicals). Each cloned pGL3E firefly luciferase reporter plasmid was cotransfected with pRL-SV40 (Promega Corporation), a plasmid expressing renilla luciferase under the influence of an SV40 promoter. The latter plasmid was used as a transfection efficiency control. Transfection solutions (100 μ l) containing 0.75 μ g cloned pGL3E DNA, 18.75ng pRL-SV40 DNA (DNA ratio of 40:1) and 2.31 μ l of Fugene 6 (DNA:Fugene ratio of 3:1), were freshly prepared in serum-free cell culture medium. Transfection was performed by the dropwise addition of each transfection solution to a well containing growing cells, gentle mixing by swirling, and further incubation at 37°C and 5% CO₂ for 48 h. Each plate contained five different plasmid transfections, a positive transfection control with pGL3C plasmid (Promega Corporation), a baseline expression control (transfection with empty pGL3 Enhancer vector) and an untreated control (no transfection), all performed in triplicate. Four replicates of each 24-well plate were prepared. The pGL3 Enhancer was used rather than pGL3 Basic

plasmid because of the low levels of luciferase activity seen in HASM cells following transfection with the latter construct.

Dual luciferase reporter assay. Luciferase assays were performed using the Dual Luciferase Reporter assay system (Promega Corporation) according to manufacturer's instructions. The 24-well plates were removed from the incubator, and the growth medium was aspirated. The wells were rinsed with phosphate-buffered saline and aspirated. Passive lysis buffer (100 μ l) was then added to each well, and the plates were placed on a rocking platform for 15 min. The lysates were assayed for firefly luciferase activity by the addition of 20 μ l lysate to 100 μ l of freshly prepared Luciferase Assay Reagent II, and luminescence was measured in a TD20e (Turner Designs, Sunnyvale, CA) luminometer programmed to effect a 3-s initial delay followed by a 10-s integration period. Renilla luciferase activity was assayed by the addition of 100 μ l Stop and Glo reagent, rapid vortexing, and luminometric measurement.

Results were normalized for variations in transfection efficiency, by using the ratio of firefly to renilla luciferase activity as an index of promoter activity. The promoter activity of each deletion construct was expressed as fold values over the baseline reporter expression activity (transfection with empty pGL3E vector).

Determination of potential transcription factor-binding sites. The online bioinformatics application, MatInspector (10), available at <http://www.genomatix.de/cgi-bin/matinspector/matinspector.pl> and the DNAssist software application (<http://www.dnassist.org>) were used to identify transcription factor consensus binding sites in the putative promoter sequence.

Genotyping. Genotyping of the identified multiallelic (CA)_n repeat was performed by Lark Technologies (Essex, UK) by analysis of PCR

TABLE 2. Oligonucleotide primer sequences used in order to amplify promoter fragments for subsequent cloning

Primer	Sequence
AFPF22	5'-CAA TGC GGA AAC GCG TAA TTG CAG ATG GAG AGC TGG-3'
AFPF23	5'-CAA TGC GGA AAC GCG TCA TTT GTC AAA GCT CCA AAG-3'
AFPF24	5'-CAA TGC GGA AAC GCG TGG AGA TCA CAA ATT CCC TGA-3'
AFPF25	5'-CAA TGC GGA AAC GCG TGG ATG AGA GAA AAT AAA GCC-3'
AFPF26	5'-CAA TGC GGA AAC GCG TTG CTG TAC TAA AGG CGC CAG-3'
AFPR22	5'-TGA ACT TGC ACT CGA GCT GGC GCC TTT AGT ACA GCA-3'
AFPR25	5'-TGA ACT TGC ACT CGA GAC CGC TTA GAG TCC GAG GCC-3'
AFPR26	5'-TGA ACT TGC ACT CGA GTG TGA CCT GTT CGT GGT TTA-3'

The boxed nucleotides contain the recognition sequence for *Mlu*I (forward primers) (A/CGCGT) or *Xho*I (reverse primers) (C/TCGAG). The region 5' of the recognition sequence consists of a 10-bp noncomplementary tail, designed to reduce restriction efficiency problems associated with digestion of terminal fragments from dsDNA.

TABLE 3. Primers and annealing temperatures used for amplification of 5'UTR regions for pGL3E cloning

PCR for Construct	Forward Primer	Reverse Primer	Annealing Temperature (°C)
A	AFPF26	AFPR25	62
B	AFPF23	AFPR25	62
C	AFPF26	AFPR26	63
D	AFPF23	AFPR26	63
E	AFPF22	AFPR22	54
F	AFPF23	AFPR22	54
G	AFPF24	AFPR22	51
H	AFPF25	AFPR22	51

product sizes using an ABI 3,100 genetic analyzer. We generated PCR products using fluorescently labeled forward primer 5'-FAM-5'-ATG GAG AGA AGC GAA AAA GAG C-3' and reverse primer 5'-TTT GGG AGG CAA CAC CTA TTC G-3'. Thermal cycling was performed using an Eppendorf Mastercycler as follows: hotstart 95°C for 5 min, denaturing 94°C for 1 min, annealing 64°C for 1 min, and extension 72°C for 2 min for a total of 30 cycles. A final extension at 72°C for 10 min was performed. A set of 14 control samples consisting of PCR products generated from plasmids of known genotype (two separate PCR products for each genotype) was also submitted to Lark together with the population samples.

Ethical approval. The genomic DNA used in this study was obtained from an anonymous random DNA bank of Maltese individuals, maintained at the Laboratory of Molecular Genetics, Department of Physiology and Biochemistry, University of Malta, and from the Outpatient Asthma Clinic, St. Luke's Hospital, Guardamangia, Malta. Prior written ethical approval was obtained from the Research Ethics Committee of the Faculty of Medicine and Surgery, University of Malta.

Results

5'RACE Experiments

At the time of experiment, our 5'RACE sequence data aligned to contig NT_007933.10 in the NCBI human genome database, as well as partially to AC009332.6 and partially to AC009329.20 found in the general NCBI DNA sequence database. NT_007933.10 has since been removed and is currently in revision 12 (NT_007933.12). We have noted a change in the size of our largest intron when calculated using the most recent NT_007933.12 data, then when calculated using data from the superceded contigs.

Analysis of the sequence data arising from both 5'RACE experiments identified the presence of six different mRNA transcripts (Figure 2) and three different transcription start sites (TSS) for the human muscarinic M₂ receptor gene. The second experiment confirmed the earlier results, and also identified two new transcripts. We have submitted the sequence data describing all our identified 5'UTR arrangements to GenBank (Accession nos. AY219704, AY034603, AY034604, AY219705, AY219703, and AY034605).

The M₂ coding sequence is preceded by a 46-bp exon that is invariably expressed in all mRNA transcripts we obtained. Upstream of this, we have identified five additional exons of which exons 5, 4, and 2 are alternatively spliced. The different

upstream arrangements, together with the exon and intron sizes are shown in Figure 2. Donor and acceptor splice site sequences were identified at the exon/intron boundaries. This arrangement suggests the presence of three different transcription start sites, all of which are active in HASM cells, with the most 5' TSS (TSS1) lying more than 146 kb upstream from the ATG start codon of the gene. We have identified a GenBank mRNA sequence (Accession no. AL832585.1) submitted by Ansorge and coworkers (2002) from a cDNA library, which is in perfect agreement with our first 5'UTR arrangement in Figure 2.

In the transcripts where Exon 3 was preceded by Exon 2 or Exon 1, Exon 3 splicing occurred between positions 122,903 and 122,904 in AC009332.6. In transcripts resulting from transcription initiation at TSS3, Exon 3 extended by a further upstream span of 24 nucleotides. We also obtained other transcripts which initiated within very close proximity of TSS3, and which suggest that a number of specific TSSs may actually be present around the region designated as TSS3 (Figure 3).

Our data suggest that the TSS3 region appears to be the most commonly used transcription start site (8 out of 14 clones), whereas TSS2 is the rarest (1 out of 14 clones).

Identification of Novel Polymorphic Regions in the Human Muscarinic M₂ Promoter Sequence

During the preparation of the PCR products for cloning into pGL3E for reporter expression studies, agarose gel analysis showed that the length of products G and H (Figure 1) was ~ 500 bp more than expected. Sequencing of this region revealed a novel 540-bp sequence containing the repeating motif TC(C)TGG(AC)[AT]_n (Figure 4) located between positions 121,646 and 121,647 of AC009332.6. We have submitted this novel sequence to GenBank (Accession no. AY221504). At the time of the experiment, this region did not align with any contig in the NCBI human genome database, although the sequences flanking it were in perfect agreement with human genome sequence data. However, a recently revised NCBI contig NT_007933.12, dated April 28, 2003 includes sequence data within this region that is very similar, though not identical, to ours. Therefore, the real length of regions G and H, used for cloning in our reporter experiments, is 540 bp more than indicated in Figure 1. Preliminary gel analysis of PCR products from 22 other human genomic DNA templates obtained from

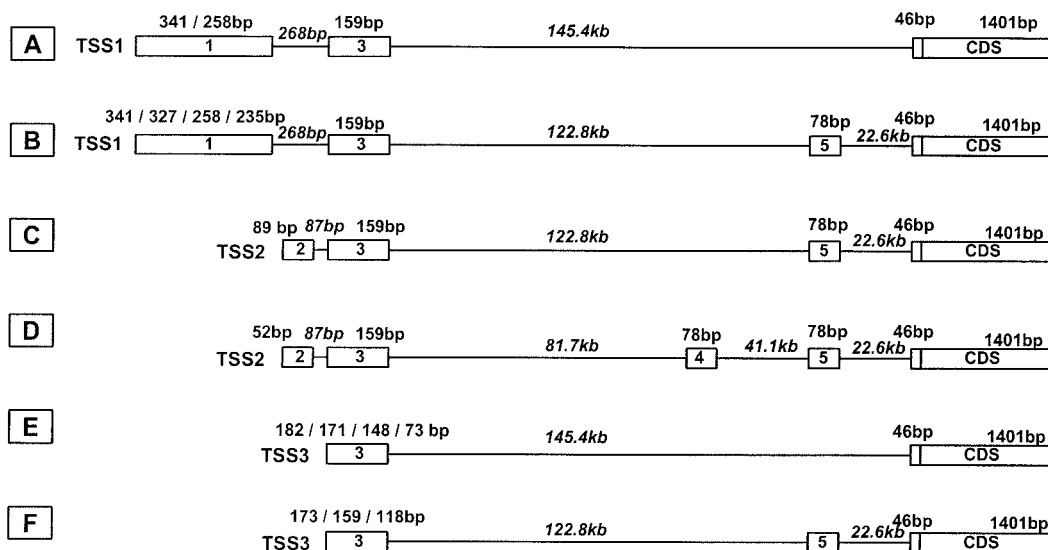


Figure 2. Identified arrangements of the human muscarinic receptor 5'UTR. CDS, coding sequence; TSS, transcription start site; *open boxes*, exonic sequences; *lines*, intronic sequences. Exonic sizes are shown in bold and intronic sizes in bold italic.

AC009332.6 (NCBI)

```

121321 cacttcttgc tctttatttg attgaagtgg gatgagagaa aataaagcct atgtaatata AFPF25
121381 taatatattg atgttatatt cacatacaca tatgctttta caaaagcaag ttttggccta
121441 agctacactt tagaatccca tcattgtact tttatcaaat gaggttagta tctaaaaata
121501 atagatacaa ttaattggca aaatttctca ttaaaataaa gaaataaaca tcaaaggcag
121561 gtctgaacac agaaaagttg catcctaata cttttttcat ggagatcaca aattccctga AFPF24
121621 agcatcaaac atatatatat tctggatata tataatctata tattctggat atatatatcc
121681 tggacatata tatatatatt ctagatata atattctgga tatacacata ttctgaatat
121741 agatatagat acagatatag atatatagat attctgtagt ctactgcccc tttagctgaag
121801 aaaacatggt cactggctca tatttaaagt aagagaaaaa aattacttct ccttcggata
121861 catttgtcaa agctccaaag aaataaacia tctgtgaata gttcagcaca attatatatt AFPF23
121921 taaaaagcat agtggattta gtctaacatt tccccacctg gaaaaactgg gcttaagtat
121981 ctccgacaaa caaaaagcaa tactgcgttc tcccaatat gccatgtcac cttttagca
122041 ctcacacact tgcagggaa ggaactgggg ttgggggtgg ttctgatgta gaaattgca AFPF22
122101 atggagagct gggctctggg gaggaaatg gagagaagcg aaaaagagcc tggagaaa gg
122161 gctgtcatat ttgctcagaa gaaccgcaa atcagagagc acagaagatt tcattggagg
122221 agcgccttca acagcaacta atcctgtcat actgaaagaa aattgtaaat gttctctcca
122281 gaaattgtcc gcccAttgca gctaategAa cacacacgca cgaatacaga cacgctctcc
122341 tgctgtacta aaggcgccag ggcgcaaaaga cctagggAgc gcgcgcgggg acacacacac AFPR22/AFPF26
122401 Acacacacac acacacagac acacacacac tcacacactc caggctcggg gttggccaga
122461 gcttggaaat cgcagatttc accaggaat cgcgcaggta ccogaatagg tgttgcctcc
122521 caaaactaat tctactctcc ccaagcactt tggagaagaa atgaatccag ccagctcgc
122581 egaggcatcc aggtctccat tctattttgg tctccaggct cttctgtaa gccaggtttg
122641 aaagtctttg ttgctctctc tttagtggtt tggcgtggag ggcgggggca gcgtcaattt AFPR26
122701 aactccagtc taaaccacga acaggtcAca gactagccaa gtggctgaga cgagtgtggg
122761 gtgcGtgact ctgctcgcc gcgcgccagc cccgcagctc tgcagagc cttggggtcg
122821 gttctctccc gcctctccc cctctcccgc cccaccctca cagtgtctcg gcatcaactA
122881 atcttttct Ttctttttgc aaGatcaagg gagAaagaga accggcagct ggcctcggac AFPR25
122941 tctAagcggg gcgcagctcc agcccagcgc gatcggccct gaaccacAa aggactcctc
123001 gctccttcaa gcctccacga cctcgcagcc ggggagcga cctggagcga accagcagca
123061 ggttaagggc ccgcgctgc cagctctccg cttaccact ccaagggccc tggcaccaga
123121 ggtctctccc ttcactctcc cccaagccgg aggtcaggg gtcggggctt . . . . .

```

AC009329.20 (NCBI)

```

17041 cagcctgggc tcaagtgat ctcctgcctc agcctctgga gtatctagaa ctacagggc
17101 atgccaccat gtccggctaa tttttttat ttttttgcat attgcctagg ttgtccaggc
17161 tgatcttata ctcctggcct caagtgatcc tctctgctca gctccaaaa gcaactgggaa
17221 tacataggtg agccagcccg acatggcctc cttctccctga gactcattct caaacagagg
17281 aggccccctg ggtgaagcta cttgcataag accatgtgga atatccagct cgaatggttc
17341 ttgttacttc agcttcattc ttttaataga aatgaggcaa aaggtgcccag gcacatggat

58141 catgtacttc agatgttggg gtctaaactg agtcacaaca gacttttctc tcaagaatte
58201 cagtttagga tctgtccttg taaagactaa tatactccat tttctgttt gttgtaatta
58261 tttttactgt ttcccacaca tggttttggt ctcttttcta ccaggttgtt agcgcatga
58321 gtacagggtc tggatcttac ccattcaaga gtgtattctc agcaccatg gcttggcata
58381 tagtaagtgc tcaataaatt tttaatgaat aaattaatagc atttatagac aaaacattaa
58441 gttgttggta agtcacaata atcctttgga aaggtaaata tgtcaagtaa . . . . .

80701 tatctggata aggacatatt taaaacatta ctggacttca agttcattgc cagcctgaaa
80761 ggctcttttg aatcctccat ttgaaggtct tgetttttac atggggaatt gaggcaggta
80821 gacacagtaa tcaatgcaggg tcaatggagat ttggggagaaa ataatgtggt ttaaaaaggag
80881 aaacaacatt atgtatttta aaccaatggt tatattatgt ttgttaattt tattctattt
80941 ccttgcaggt ttaaatgttt atttgctact tggctactga ttagagaacg caaaATGAAT
81001 AACTCAACAA ACTCCTCTAA CAATAGCCTG GCTCTTACAA GTCTTATAA GACATTGTAA
81061 GTGGTGTFTA TTGTCCTGGT GGCTGGATCC CTCAGTTTGG TGACCATTAT CCGGAACATC
81121 CTAGTCATGG TTTCCATTAA AGTCAACCGC CACCTCCAGA CCGTCAACAA TTACTTTTTA
81181 TTCAGCTGG CCTGTGCTGA CCTTATCATA GGTGTTTTCT CCATGAACTT GTACACCTC
81241 TACTACTGTA TTGGTTACTG GCCTTTGGGA CCTGTGGTGT GTGACCTTTG GCTAGCCCTG
81301 GACTATGTGG TCAGCAATGC CTCAGTTATG AATCTGTCTA TCATCAGCTT TGACAGGTAC
81361 TTCTGTGTCA CAAAACCTCT GACTACCCA GTCAAGCGGA CCACAAAAAT GGCAGGTATG
81421 ATGATTGCAG CTGCCTGGT CCTCTTTTC ATCCTCTGGG CTCAGCCCAT TCTTCTGTG
81481 CAGTTCATTG TAGGGGTGAG AACTGTGGAG GATGGGGAGT GCTACATTCA . . . . .

```

Exon 1

Exon 2

Exon 3

Exon 4

Exon 5

5' RACE nested GSP

5' RACE GSP

CDS

Figure 3. NCBI database sequence showing the position of the identified exonic (shaded) and intronic regions in the 5'UTR region of the human muscarinic M₂ receptor gene. The coding region (CDS) is shown in uppercase, primer regions are indicated with a single underline, and putative transcription start sites are indicated with a bold uppercase letter. **at** indicates the bases between which the hypervariable 0.5-kb region is inserted, c indicates the position of the C→A SNP, double underline indicates the position of the multiallelic (CA) tandem repeat. The primer and exon names are denoted by labels on the right.

different individuals all indicated the presence of this novel region. Cloning and sequencing of an additional 40 genomic templates in 8 pools, identified this to be a hypervariable region of ~ 0.4–0.6 kb, containing a repeating pattern that usually, but not unequivocally, fits the one described above. The significance of the polymorphic variation in this novel region is unknown, although given the functional data presented below, this region may not contain strong positive elements promoting transcription.

Sequencing of our plasmids also revealed a novel (CA) tandem repeat region located at 96 bp downstream of TSS1. The NCBI contig AC009332.6 contains 14 (CA) dinucleotide repeats in this region. Besides the wild-type alleles, pooled cloning followed by sequencing of an additional 40 samples, as described above, identified alleles with the following deletions $\Delta(\text{CA})_8$,

$\Delta(\text{CA})_4$, $\Delta(\text{CA})_3$, $\Delta(\text{CA})_2$, $\Delta(\text{CA})_1$, as well as an allele having an insertion of one (CA) repeat. Genotyping of samples obtained from an anonymous random DNA bank of Maltese individuals, maintained at the Laboratory of Molecular Genetics, Department of Physiology and Biochemistry, University of Malta, showed the most common alleles in the white population to be $\Delta(\text{CA})_1$ (44.0%) and wild type (22.3%) ($n = 186$) (Table 4). Repeat PCRs on the same genomic templates yielded a consistent number of (CA) repeats, suggesting that these are real alleles and not the result of PCR artifacts such as slippage.

We have also identified a common C→A SNP located at position 122,159 on NCBI AC009332.6 (136 bp upstream from the most 5' TSS). Our preliminary data suggest that the allelic frequency of the A allele is 52% ($n = 80$) in the white population.

AAAATTTTAA	CCNGCTTAAC	AATNTCCCN	TTCCCCCNT	TCAGGGGGG	NNCAANTTTT	60
TGGAANGGGG	GAATCGNGC	GGGCCCTTT	TTNGTTTTT	TCNNCCNCC	CCNAAGCCT	120
CCCNAGGNTA	AGGTAAGGTA	ATNNTTAAG	GTCCGGGGG	GGTCTTGGN	AGCCGCCNC	180
CAAAAAATA	TCCTTTATT	TCCANCNA	NCCGGGGNG	GGGNTTTTT	GGGGGAATT	240
CGAGAGTACT	AACCANCCG	TCPCNCCNA	AACAAANCG	AACAAAACA	AACTNGCAA	300
AATAGGCTGT	CCCCAGTGCN	AGTGCAGGTG	CCAGAACATT	TCTCTATCGA	TAGGGTACCG	360
AGCTTCTTAC	GCGTGGAGAT	CACAAATTC	CTGAAGCATC	ANACATATAT	ATATTCTNGGA	420*
GATATATATA	TATATATATT	TTTTTTTTGA	GAGATATATA	TATATATATA	TATTTTNGAG	480
ATATATATAT	ATATTTTCTG	GATATATATA	TATTTTCTGG	ATATATATAT	ATCTGGACA	540
TATATATATT	CTGGATATAT	ATATCCTGGA	CATATATATA	TATATTTCTGG	ATATATATAT	600
CCTGGACATA	TATATATATA	TATATATATT	CTGGATATAT	ATATATATAT	TCTGGATATA	660
TATATCCTGG	ACATATATAT	ATATATATAT	ATATATATAT	ATATTTCTGGA	TATATATATC	720
CTGGACATAT	ATATATATAT	ATATATATAT	TCTGGATATA	TATATCCTGG	ACATATATAT	780
ATATATATAT	ATATATATTC	TGGATATATA	TATCCTGGAC	ATATATATAT	ATATATCTG	840
GATATATATA	TATTTCTGGAT	ATATATATAT	ATTTCTGGATA	TATATATATA	TATTTCTGGAT	900
ATATATATAT	ATTTCTGGATA	TATATATATA	TTCTGGATAT	ATATATATAT	ATATTTCTGGA	960
TATATATATC	TATATATTTCT	GGATATATAT	ATCCTGGACA	TATATATATA	TATATTTCTAG	1020**
ATATATATAT	TCTGGATATA	CACATATTTCT	GAATATAGAT	ATAGATACAG	ATATAGATAT	1080

Figure 4. Novel sequence located between positions 121,646* and 121,647** of AC009332.6. Bases in italics denote the novel sequence, whereas bases in normal text denote regions which agree with the NCBI human genome sequence database information. Bold bases indicate the starts of the repeating motif TC(C)TGG(AC)[AT]_n (curved brackets denote bases that may or may not be present).

Dual Luciferase Reporter Gene Experiments

To investigate the potential promoter activity contained in regions upstream of the identified transcription start sites 1 to 3, a reporter gene approach was used. We transfected primary HASM cells and the human bronchial epithelial cell line BEAS-2B to look for airway tissue selective expression. Luciferase gene-based reporter assays were performed as previously described (12).

Reporter assay data showed that the highest promoter activity is present in the DNA regions upstream from TSS3 for both HASM and BEAS-2B cell transfectants, whereas low activities were obtained for regions upstream of TSS1 (Figures 5 and 6). HASM cell transfectants expressed lower luciferase activities than BEAS-2B, and also showed higher SEM values than BEAS-2B cells. This trend has been observed with several other transfection experiments from our group (12), and is likely to be due to the inherent variability in transfection of human primary cell culture systems.

The results obtained from HASM transfectants for activities upstream of TSS3 showed construct A to have higher activity than B (32.2 ± 3.3 versus 18.9 ± 3.5 fold over empty vector [mean \pm SEM], $P < 0.05$, $n = 4$). With respect to the region upstream of TSS2, construct C showed a trend toward higher activity than D (5.0 ± 1.6 versus 3.4 ± 0.3 fold over empty vector [mean \pm SEM], $n = 4$), although the differences did not attain statistical significance (Figure 5). These data suggest that the sequence upstream of TSS1 might contain elements which act as repressors in HASM cells. This might also explain the low activities observed for regions E, F, G, and H, upstream of TSS1.

This pattern was reversed in BEAS-2B cell transfectants (Figure 6). Construct B showed higher activity than A (57.4 ± 4.1 versus 34.2 ± 1.3 fold over empty vector, $n = 4$, $P < 0.01$) and D showed higher activity than C (31.9 ± 2.0 versus 14.3 ± 0.9 fold over empty vector, $n = 4$, $P < 0.001$), suggesting that such

postulated repressor elements may be cell type-specific, and may not operate in BEAS-2B cells.

We also investigated the potential role of the novel (CA) tandem repeat on reporter gene expression in these systems. Construct A containing the wild-type allele, produced a significantly higher reporter expression than the $\Delta(CA)_8$ construct A, in both HASM cell transfectants (32.2 ± 3.3 versus 13.8 ± 5.4 , $n = 4$, $P < 0.05$) (Figure 5B) and BEAS-2B cell transfectants (34.2 ± 1.3 versus 20.6 ± 3.3 , $n = 4$, $P < 0.01$) (Figure 6B).

Determination of Potential Transcription Factor Binding Sites

Using MatInspector (13) and DNAssist, the potential promoter sequence was searched for transcription factor consensus binding sites. The selected results of this are shown in the online supplement.

Discussion

Muscarinic M₂ receptors play an important role in the regulation of adenylyl cyclase activity in HASM. In this article we describe the genomic organization of the human muscarinic M₂ receptor gene 5'UTR, and suggest regions of major transcriptional regulatory activity relevant for control of M₂ receptor expression in ASM.

The results of this work are based on two 5'RACE experiments performed on two separate occasions using different populations of HASM cells. The 5'RACE protocol adopted here includes some modifications over standard 5'RACE procedures. In brief, it eliminates the participation of uncapped or truncated mRNA transcripts in the 5'RACE reaction, by prior CIP dephosphorylation of the 5' ends, thus reducing the possibility of 5' ends of incomplete transcripts being falsely identified as TSSs. This approach has also been successfully used by other research groups (14–17).

In initial experiments, we also used a primer extension analysis approach, but found this to result in unreliable amplification for the 5'UTR of this gene. In this respect, other researchers have reported problems with primer extension analysis for a number of GPCR genes, and this has been ascribed to the presence of GC-rich sequences within the 5' noncoding regions (7, 18).

We identified three regions of transcription initiation in the human muscarinic M₂ receptor gene 5'UTR, with each region containing a cluster of specific transcription start sites (TSSs) in close proximity to each other. The region also contains five exons of which Exon 2, Exon 4, and Exon 5 are alternatively spliced (Figure 2). The most 5' TSS lies more than 146 kb upstream from the ATG start codon of the gene. An NCBI GenBank

TABLE 4. Allelic frequencies of muscarinic M₂ promoter (CA)_n variants in the random white population ($n = 186$)

Allele	Number of (CA) repeats present	Allelic frequency (%)
Insertion (CA) ₁	15	0.6
Wild-type	14	22.3
$\Delta(CA)_1$	13	44.0
$\Delta(CA)_2$	12	0.6
$\Delta(CA)_3$	11	14.9
$\Delta(CA)_4$	10	0.6
$\Delta(CA)_8$	6	17.1

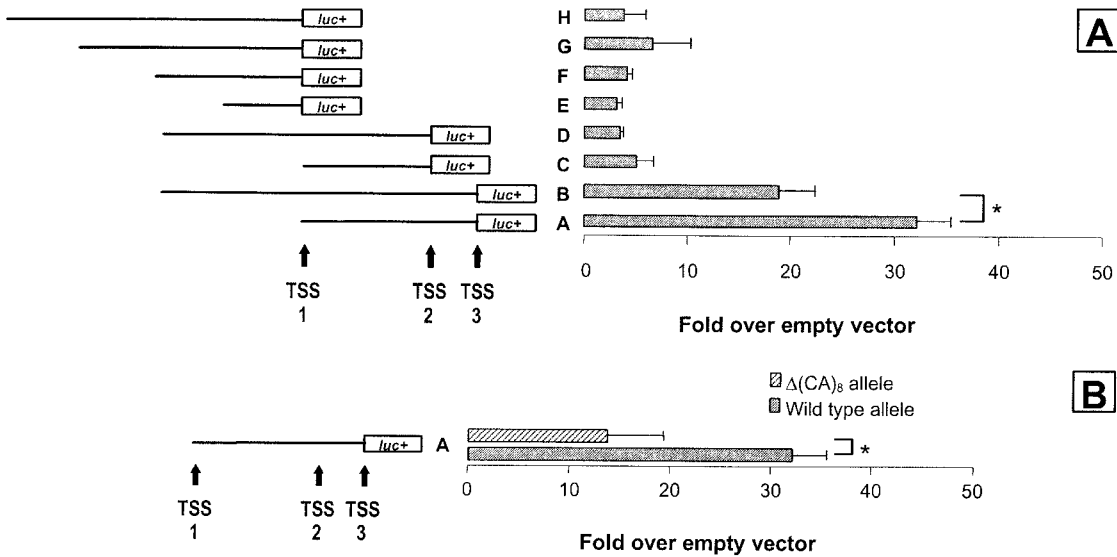


Figure 5. Luciferase reporter activity for HASM cell transfectants. Data are expressed as fold values over empty vector baseline control. (A) Luciferase activity results of all deletion constructs. (B) Luciferase activity results of the $\Delta(CA)_8$ and wild-type alleles of construct A. The error bars denote the SEM (* $P < 0.05$).

mRNA sequence (Accession no. AL832585.1), submitted by Ansonge and coworkers (2002) from a cDNA library, is in perfect agreement with the first 5'UTR arrangement shown in Figure 2. Further evidence of the existence of these arrangements arises from the presence of donor/acceptor splice sites at each exon/intron boundary.

Reporter gene expression analysis performed on primary cultures of HASM cells, transiently transfected with pGL3 Enhancer constructs, provided data which strongly suggest that the major regulatory region lies immediately upstream of TSS3. In view of this, it is interesting to note that transcription initiation region 3 also contains the majority of clustered TSSs and the highest frequency of mRNA transcripts appears to originate at TSS3 in primary HASM cell cultures. Moreover, from the differences in expression between construct A and B (Figure 5), it appears that repressor elements may operate upstream of TSS1. The decreased expression of construct D compared with C adds weight to this argument. All constructs containing regions upstream of TSS1 showed low levels of expression, and there may potentially be some repressor elements operating at the 5'

end region of H, considering that this construct showed lower expression levels than G.

The reporter assay results obtained for the BEAS-2B assay data also indicate the highest levels of positive transcriptional regulatory control to be present directly upstream of TSS3. However, in this cell type, the postulated repressor elements described above do not appear to exert an effect. Indeed construct B exerted higher positive transcriptional regulation than A and construct D exerted higher positive transcriptional regulatory control than C. In addition, construct H also showed higher promoter activity than G, as opposed to HASM cells, suggesting that any transcriptional repressor elements present in the 5' terminal region of H may operate in HASM cells but not in BEAS-2B.

These observations may be partially explained by the different transcription factor pool compositions expected to be present in different cell types. Although HASM cells express the M₂ receptor on the cell membrane, both in cell culture and in whole animal models, there is no current evidence to support the expression of M₂ receptors in BEAS-2B human airway epithelial

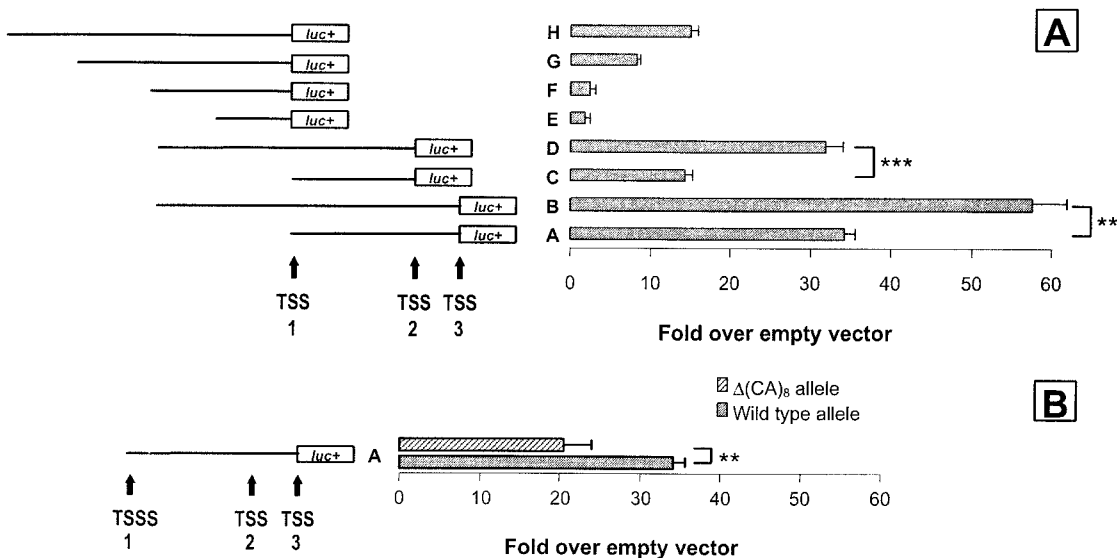


Figure 6. Luciferase reporter activity for BEAS-2B cell transfectants. Data are expressed as fold values over empty vector baseline control. (A) Luciferase activity results of all deletion constructs. (B) Luciferase activity results of the $\Delta(CA)_8$ and wild-type alleles of construct A. The error bars denote the SEM (** $P < 0.01$, *** $P < 0.001$).

cells. Different reporter expression profiles may be expected to be observed in further cell types.

This work has identified polymorphic variation within the promoter region of the muscarinic M₂ receptor gene. A (CA)_n tandem repeat polymorphism present downstream of TSS1 has been shown to significantly alter reporter gene expression in HASM and BEAS-2B cell transfectants, and may therefore potentially affect muscarinic M₂ receptor expression in HASM cells *in vivo*. The 0.5-kb hypervariable region is present in a region of low promoter activity, and is therefore not expected to offer a significant contribution to the overall level of M₂ receptor expression in HASM cells, whereas the identified C→A SNP is of unknown relevance.

A map of transcription factor consensus sequences within the postulated regulatory regions is presented in the online supplement. In line with all muscarinic receptor promoters currently identified, the human muscarinic M₂ receptor promoter is TATA-less. Sp1, AP, and GATA transcription factors have previously been cited as relevant for TATA-less promoters (5, 7, 10, 19, 20). In view of this it is interesting to note that the highest incidence of Sp1, GATA, and AP sites lies within the region of maximum transcriptional regulatory activity, immediately upstream of TSS3. It is also interesting to note a cAMP receptor element-binding protein (CREB) consensus binding sequence present upstream of TSS2. cAMP activates the kinase PKA, which in turn phosphorylates CREB's activating region, increasing the affinity of CREB for CREB-binding protein (CBP). CBP is believed to interact with one or more parts of the transcriptional machinery, exerting a positive transcriptional regulatory effect. Therefore, cAMP may potentially influence M₂ transcriptional regulatory control.

One particularly striking aspect of the work reported here is the high homology between the human and porcine 5'UTR sequences: there is 77–86% homology between the porcine Exon 1B and significant sections of the human Exons 1 and 3. In addition to the exon homology, there is also a small but significantly homologous portion in the putative promoter where 32 bp match the published chick promoter region (GenBank Accession no: U61850) by 91%. Indeed, this significant homology has also been noted in the 5' arrangements of other muscarinic receptor subtypes: one of the five exons upstream of the M₃ receptor coding sequence is homologous to the reported bovine M₃ receptor cDNA (8), and the 5' exon in the rat M₁ receptor 5'UTR shares significant homology with the 5' end of porcine M₁ receptor cDNA (5, 21). Despite the observation that muscarinic receptor subtypes sharing common pharmacologic properties also share common structural features due to sections of coding sequence homology, there was no significant homology between the M₂ receptor upstream arrangement described in this paper and the rat M₄ receptor promoter region or exon 1 (GenBank Accession no: D78484.1) as analyzed using BLAST (<http://www.ncbi.nlm.nih.gov/blast>).

In addition to the nucleotide similarity between species, there appears to be a conserved pattern of exon organization between species and also between the muscarinic subtypes. A consensus splice acceptor point is observed 47 bp upstream of the ATG start codon in the human M₂ receptor gene, as confirmed in this paper and recognized by Bonner and colleagues (3), and this is largely replicated in the porcine M₂ (–46 bp) (4), the chick M₂ (–41 bp) (6), the rat M₁ (–69 bp) (5), the rat M₃ (–20 bp) (3), the rat M₄ (–32 bp) (19), and the human and rat M₅ (–78 bp) (22).

Alternative splicing was observed to occur for the human muscarinic M₂ receptor, as shown in Figure 2, as well as for the porcine M₂ receptor. The porcine M₂ receptor gene is reported to have two upstream exons, 1A and 1B, with exon 1B being

alternatively spliced (4). The 5' exons of human muscarinic M₂ receptor subtype have also been reported to be alternatively spliced (8), as have the rat M₁ and M₄ receptor genes. Not only do all these genes exhibit alternative splicing, but they also all include relatively large introns separating their single coding regions from upstream noncoding exons. We were particularly surprised by the size of the intron between exons 2 and 3 of the M₂ 5'UTR. It seems probable that it is the size of these introns and the complex upstream arrangement of these genes which has confounded their identification since the coding sequence was published in 1987 (3, 22).

All muscarinic receptor gene promoters found to date are TATA-less and generally have numerous transcription factor binding sites with Sp1, AP, and GATA being cited as relevant (5, 7, 10, 19, 20). These characteristics appear to be typical of promoters for housekeeping genes, which are constitutively expressed. However, recently a number of investigations have reported some highly regulated genes to have TATA-less promoters (23). GPCRs observed to have GC-rich, TATA-less promoters include D₁, D₂, and D₅ dopaminergic and the β₁ adrenergic receptor (18, 24–27). All of these promoters have also been observed to contain Sp1 sites, although the transcriptional requirement of Sp1 has not been demonstrated. However, Sp1 binding has been shown to be a critical requirement for transcription initiation in the TATA-less promoters of some genes (28, 29). The only investigation to date into the relative control by transcription factors over M₂ receptor regulation has shown the GATA family to be of importance: the expression of a construct containing the chick M₂ promoter was observed to be critically dependent on the GATA-responsive element in cardiac primary cultures (20).

Although in this study we have defined the transcription start sites important for transcriptional control of the M₂ receptor in ASM, this receptor is also important in other cell types, such as cholinergic nerves. Transcripts derived from different cells may exhibit different upstream arrangements. Rosoff and coworkers (6) observed the chick M₂ promoter to have five transcription start sites, some of which were used equally in both chick heart and brain but some were expressed preferentially in one of the tissues. Zhou and coworkers (9) recently described a muscarinic M₂ promoter arrangement based on work performed on mRNA transcripts obtained from human heart muscle and the human neuroblastoma cell line IMR-32 (ATCC no. CCL-127). Their work identified a single TSS that was present 55 bp upstream of the 22.6-kb intron, specifically at position 61,715,148 of NT_007933.12 or position 58,329 of AC009329.20. This corresponds to a location located within the Exon 5 reported in this article. These data, together with the data presented in this paper, suggest that tissue-specific regulation of M₂ receptor transcription may be an important mechanism controlling expression.

Analysis of promoter activity in our cell culture systems suggests that the most important 5' regulatory area for the M₂ receptor lies upstream of TSS3, with the presence of a multiallelic tandem CA repeat in this region potentially contributing to variations in activity. Although it is currently unclear how this tandem repeat actually exerts its influence, it may be postulated to affect the expression of muscarinic M₂ receptors in airway tissue. Such an alteration in expression might be expected to contribute to the interpatient variability in airway hyperresponsiveness observed in patients with asthma and in patients with COPD. In inflammatory conditions of the airways, there is an elevated degree of airway hyperresponsiveness that is due to both inflammatory and neurogenic factors, and that also potentially includes a degree of M₂ receptor dysfunction induced by allosteric modulation by EPO and MBP (30, 31) as well as in-

creased downregulation induced by tumor necrosis factor- α and interleukin-1 β (32). The CA repeat we report here has the potential to alter levels of M₂ receptor expression, which provides an additional mechanism controlling functional effects driven by M₂ receptor stimulation. Such effects include short-term inhibition of adenylate cyclase (1) as well as long-term sensitization of adenylate cyclase following chronic stimulation or modulation by interleukin-1 β or rhinovirus (33).

In summary, we have characterized the 5'UTR of the human muscarinic M₂ receptor gene and have identified the major transcriptional regulatory regions for this gene in HASM. In common with other muscarinic receptors, the promoter contains multiple exons, some of which are alternately spliced, and contains multiple transcriptional start sites. The region of major transcriptional activity is located more than 146 kb upstream from the gene coding region. We have also identified a novel hypervariable region (which is, however, located in a region of low promoter activity), a multiallelic CA tandem repeat region which alters transcriptional activity in reporter assays, and a biallelic C→A SNP, the relevance of which remains to be determined. Taken together, these data suggest that genetic variability may produce inter individual differences in the control of airway M₂ receptor expression.

Acknowledgments: This work was funded by the National Asthma Campaign, UK, and the University of Malta, Malta. All DNA dideoxy sequencing was performed by Ms Ingrid Davies Gibbs at the University of Nottingham Medical School.

References

- Widdop, S., K. Daykin, and I. P. Hall. 1993. Expression of muscarinic M₂ receptors in cultured human airway smooth muscle cells. *Am. J. Respir. Cell Mol. Biol.* 9:541–546.
- Barnes, P. J. 1990. Muscarinic receptors in airways: recent developments. *J. Appl. Physiol.* 68:1777–1785.
- Bonner, T. I., N. J. Buckley, A. C. Young, and M. R. Brann. 1987. Identification of a family of muscarinic acetylcholine receptor genes. *Science* 237:527–532.
- Peralta, E. G., J. W. Winslow, G. L. Peterson, D. H. Smith, A. Ashkenazi, J. Ramachandran, M. I. Schimerlik, and D. J. Capon. 1987. Primary structure and biochemical properties of an M₂ muscarinic receptor. *Science* 236:600–605.
- Klett, C. P., and T. I. Bonner. 1999. Identification and characterization of the rat M₁ muscarinic receptor promoter. *J. Neurochem.* 72:900–909.
- Rosoff, M. L., J. Wei, and N. M. Nathanson. 1996. Isolation and characterization of the chicken m2 acetylcholine receptor promoter region: induction of gene transcription by leukemia inhibitory factor and ciliary neurotrophic factor. *Proc. Natl. Acad. Sci. USA* 93:14889–14894.
- Wood, I. C., A. Roopra, C. Harrington, and N. J. Buckley. 1995. Structure of the m4 cholinergic muscarinic receptor gene and its promoter. *J. Biol. Chem.* 270:30933–30940.
- Forsythe, S. M., P. C. Kogut, J. F. McConville, Y. Fu, J. A. McCauley, A. J. Halayko, H. W. Liu, A. Kao, D. J. Fernandes, S. Bellam, E. Fuchs, S. Sinha, G. I. Bell, B. Camoretti-Mercado, and J. Solway. 2002. Structure and transcription of the human m3 muscarinic receptor gene. *Am. J. Respir. Cell Mol. Biol.* 26:298–305.
- Zhou, C., A. D. Fryer, and D. B. Jacoby. 2001. Structure of the human M₂ muscarinic acetylcholine receptor gene and its promoter. *Gene* 271:87–92.
- Saffen, D., M. Mieda, M. Okamura, and T. Haga. 1999. Control elements of muscarinic receptor gene expression. *Life Sci.* 64:479–486.
- Daykin, K., S. Widdop, and I. P. Hall. 1993. Control of histamine induced inositol phospholipid hydrolysis in cultured human tracheal smooth muscle cells. *Eur. J. Pharmacol.* 246:135–140.
- Le Jeune, I. R., M. Shepherd, G. Van Heeke, M. D. Houslay, and I. P. Hall. 2002. Cyclic AMP-dependent transcriptional up-regulation of phosphodiesterase 4D5 in human airway smooth muscle cells. Identification and characterization of a novel PDE4D5 promoter. *J. Biol. Chem.* 277:35980–35989.
- Quandt, K., K. Frech, H. Karas, E. Wingender, and T. Werner. 1995. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res.* 23:4878–4884.
- Chen, J., M. Sun, S. Lee, G. Zhou, J. D. Rowley, and S. M. Wang. 2002. Identifying novel transcripts and novel genes in the human genome by using novel SAGE tags. *Proc. Natl. Acad. Sci. USA* 99:12257–12262.
- Hubert, N., and M. W. Hentze. 2002. Previously uncharacterized isoforms of divalent metal transporter (DMT)-1: implications for regulation and cellular function. *Proc. Natl. Acad. Sci. USA* 99:12345–12350.
- Kressler, D., S. N. Schreiber, D. Knutti, and A. Kralli. 2002. The PGC-1-related protein PERC is a selective coactivator of estrogen receptor alpha. *J. Biol. Chem.* 277:13918–13925.
- Lin, S., Q. Shi, F. B. Nix, M. Styblo, M. A. Beck, K. M. Herbin-Davis, L. L. Hall, J. B. Simeonson, and D. J. Thomas. 2002. A novel S-adenosyl-L-methionine:arsenic(III) methyltransferase from rat liver cytosol. *J. Biol. Chem.* 277:10795–10803.
- Minowa, M. T., T. Minowa, F. J. Monsma, Jr., D. R. Sibley, and M. M. Mouradian. 1992. Characterization of the 5' flanking region of the human D_{1A} dopamine receptor gene. *Proc. Natl. Acad. Sci. USA* 89:3045–3049.
- Mieda, M., T. Haga, and D. W. Saffen. 1997. Expression of the rat m4 muscarinic acetylcholine receptor gene is regulated by the neuron-restrictive silencer element/repressor element 1. *J. Biol. Chem.* 272:5854–5860.
- Rosoff, M. L., and N. M. Nathanson. 1998. GATA factor-dependent regulation of cardiac m2 muscarinic acetylcholine gene transcription. *J. Biol. Chem.* 273:9124–9129.
- Kubo, T., A. Maeda, K. Sugimoto, I. Akiba, A. Mikami, H. Takahashi, T. Haga, K. Haga, A. Ichiyama, and K. Kangawa. 1986. Primary structure of porcine cardiac muscarinic acetylcholine receptor deduced from the cDNA sequence. *FEBS Lett.* 209:367–372.
- Bonner, T. I., A. C. Young, M. R. Brann, and N. J. Buckley. 1988. Cloning and expression of the human and rat m5 muscarinic acetylcholine receptor genes. *Neuron* 1:403–410.
- Azizkhan, J. C., D. E. Jensen, A. J. Pierce, and M. Wade. 1993. Transcription from TATA-less promoters: dihydrofolate reductase as a model. *Crit. Rev. Eukaryot. Gene Expr.* 3:229–254.
- Minowa, T., M. T. Minowa, and M. M. Mouradian. 1992. Analysis of the promoter region of the rat D₂ dopamine receptor gene. *Biochemistry* 31:8389–8396.
- Zhou, Q., P. M. Lieberman, T. G. Boyer, and A. J. Berk. 1992. Holo-TFIID supports transcriptional stimulation by diverse activators and from a TATA-less promoter. *Genes Dev.* 6:1964–1974.
- Beischlag, T. V., A. Marchese, J. H. Meador-Woodruff, S. P. Damask, B. F. O'Dowd, R. F. Tyndale, H. H. van Tol, P. Seeman, and H. B. Niznik. 1995. The human dopamine D5 receptor gene: cloning and characterization of the 5'-flanking and promoter region. *Biochemistry* 34:5960–5970.
- Collins, S., J. Ostrowski, and R. J. Lefkowitz. 1993. Cloning and sequence analysis of the human β_1 -adrenergic receptor 5'-flanking promoter region. *Biochim. Biophys. Acta* 1172:171–174.
- Pugh, B. F., and R. Tjian. 1990. Mechanism of transcriptional activation by Sp1: evidence for coactivators. *Cell* 61:1187–1197.
- Pugh, B. F., and R. Tjian. 1991. Transcription from a TATA-less promoter requires a multisubunit TFIID complex. *Genes Dev.* 5:1935–1945.
- Jacoby, D. B., G. J. Gleich, and A. D. Fryer. 1993. Human eosinophil major basic protein is an endogenous allosteric antagonist at the inhibitory muscarinic M₂ receptor. *J. Clin. Invest.* 91:1314–1318.
- Evans, C. M., A. D. Fryer, D. B. Jacoby, G. J. Gleich, and R. W. Costello. 1997. Pretreatment with antibody to eosinophil major basic protein prevents hyperresponsiveness by protecting neuronal M₂ muscarinic receptors in antigen-challenged guinea pigs. *J. Clin. Invest.* 100:2254–2262.
- Barnes, P. J., E. B. Haddad, and J. Rousell. 1997. Regulation of muscarinic M₂ receptors. *Life Sci.* 60:1015–1021.
- Billington, C. K., R. M. Pascual, M. L. Hawkins, R. B. Penn, and I. P. Hall. 2001. Interleukin-1beta and rhinovirus sensitize adenylyl cyclase in human airway smooth-muscle cells. *Am. J. Respir. Cell Mol. Biol.* 24:633–639.