Cloning and characterisation of the human adenosine A3 receptor gene

E.M. Murrison, S.J. Goodson, M.R. Edbrooke, C.A. Harris*

Division of Cellular Sciences, Glaxo Wellcome Medicines Research Centre, Gunnels Wood Road, Stevenage, Hertfordshire, SG1 2NY, UK

Received 9 February 1996; revised version received 16 March 1996

Abstract We have cloned the gene for the human adenosine A3 receptor and report characterisation of its intron/exon structure and upstream untranslated region. The open reading frame is interrupted by a single intron of approximately 2.2 kb, within the coding sequence for the second cytoplasmic loop. Sequence analysis of the upstream region reveals no TATA box but the transcriptional start site has been mapped to a common nucleotide in three tissues by 5′-RACE and RT-PCR analysis. Northern blotting, 5′-RACE PCR and analysis of upstream sequences, have provided no evidence for the occurrence of further introns in the upstream untranslated sequence or of transcriptional regulation by alternative splicing in this region.

Key words: Adenosine A3 receptor; Upstream sequence; Intron

1. Introduction

The adenosine A3 receptor subtype is the most recent member of this family of G-protein-coupled receptors to have been cloned. Degenerate PCR strategies were used to isolate cDNA clones from rat testis [1] and striatal [2] cDNA libraries. Pharmacological analysis of the receptor, expressed in CHO cells, confirmed it to be a novel adenosine receptor subtype [2]. Identification and cloning of a human homologue soon followed, isolated from heart [3], striatal [4], and HL60 [5] cDNA libraries. Whereas the cDNA sequences of A1, A2a and A2b subtypes are highly conserved between rat and human (e.g. rat and human A1 receptors are 89% identical across the open reading frame), A3 receptors of these species share only 79% identity across the open reading frame. Their pharmacological agonist and antagonist profiles and tissue distribution of mRNA differ also (reviewed in [6]), raising the possibility that these two receptors are not direct species homologues. They may, however, simply be unusually highly divergent between these species, as published cloning strategies employed thus far have revealed no better related homologue. A homologous sheep receptor has also been cloned which appears to resemble the A3 receptor of the human more closely than that of the rat, by both molecular biological and pharmalogical criteria [7].

The human A3 receptor resembles the A1 subtype more closely than A2a or A2b, both on the basis of sequence homology (61%, 54% and 52% to human A1, A2a and A2b respectively, across the open reading frames) and function. Recently control of receptor expression by alternative splicing in the 5′ untranslated region of the human adenosine A1 receptor was shown to occur [8,9]. The human A1 gene is divided into six exons, of which 1, 2, 3, 4 and part of 5 encode 5′ untranslated regions. Differential use of exons 3 or 4 directly affects levels of receptor expression in different tissues, the levels being considerably reduced when exon 4 is present. This exon contains two ATG codons (and an in-frame termination codon) which when mutated out relieve this inhibition of expression.

We describe here isolation and structure of the human adenosine A3 receptor gene and show that it is unlikely to be subject to a similar transcriptional control mechanism to that which regulates expression of A1 receptor mRNA, i.e. through alternative splicing of untranslated 5′ exons.

2. Materials and methods

2.1. Isolation of a genomic A3R

Approximately 400 000 recombinant phage from a custom-made HepG2 genomic DNA library in Lambda FIX II (Stratagene) were screened using a ^{32}P-labelled full length rat A3R cDNA probe (obtained from Prof. O. Civelli, [2]).

2.2. DNA sequencing

Taq Dyedeoxy Terminator Cycle sequencing reactions were analysed on an Applied Biosystems 373 Automated DNA Sequencer. Manual sequencing was also carried out on ^{35}S-labelled sequencing reactions generated using a Sequenase version II DNA sequencing kit (USB). Reactions were analysed using Biorad Sequi-Gen II sequencing apparatus. Oligonucleotide primers were synthesised using an Applied Biosystems 394 DNA/RNA synthesizer.

2.3. Northern and Southern blot analysis

Human Multiple Tissue Northern (MTN) blots (Clontech) were hybridised to a full length ^{32}P-labelled A3R cDNA in Hybridisol I solution (Oncor). The blots were washed at 42°C in 0.1×SSC, 0.05% SDS. Autoradiography was performed for 3–7 days at −70°C. Probes were labelled with ^{32}P using a Rediprime kit (Amersham). Oligonucleotide primer pairs 5′-ATGCCCAACAACAGGACTGCTCT-3′ and 5′-CTGACGGTAAGCTTGACCCGC-3′, and 5′-CGGGAGTTCAAGACGGCTAA-3′ and 5′-CTACTCAGAATTCTTCTCAATG-3′ were used to generate human A3 receptor 5′ and 3′ fragments for use as probes. A human adenosine A3 receptor cDNA used to generate probes was obtained from Dr. S. Munro [5].

2.4. Rapid amplification of 5′ cDNA ends (5′-RACE) and PCR

PCR was carried out on 5′-RACE Ready lung and liver cDNA (Clontech). An A3 receptor specific antisense primer (5′ GGTAGGAATGAGCAAG 3′) was used in conjunction with a sense anchored primer. PCR was also carried out on a human eosinophil cDNA library constructed in λZAP (kindly provided by Dr. J.T. Armstrong — Glaxo Wellcome). PCR products were cloned into pCRII (Invitrogen) and sequenced. All oligonucleotides used for PCR were synthesised on an ABI 394 DNA/RNA synthesiser. PCR products which hybridised to A3 receptor specific probes, were cloned into PCRII (TA Cloning kit — Invitrogen) for sequencing.

2.5. Computer analysis of nucleotide sequence

DNA sequences were analysed using the Genetics Computer Group Package (Wisconsin).

*Corresponding author. Fax: (44) (1438) 764887.
E-mail: CAH9466@GGR.CO.UK

SSDI S0014-5793(96)00324-9

3. Results and discussion

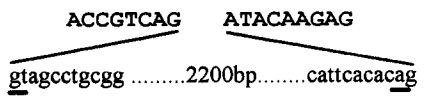
3.1. Isolation and structure of the human A3 receptor gene

A HepG2 genomic DNA library was screened using a ^{32}P-labelled, full length, rat adenosine A3 receptor cDNA probe and a single hybridising clone was isolated with an insert of greater than 12 kb. The A3 coding sequence was restriction mapped in its entirety to a 4.5 kb SstI-BamHI fragment, by hybridisation to ^{32}P-labelled probes made from 5′ and 3′ terminal human A3 cDNA PCR fragments (349 and 287 bp respectively). This 4.5 kb fragment was subcloned into pBluescript for subsequent analysis.

The A3 receptor coding region was found to be divided into two exons, separated by a single intron of approximately 2.2 kb (Fig. 1a). The intron occurs within the putative coding region for the second cytoplasmic loop, immediately after that for the third transmembrane domain. Its position within the open reading frame is precisely conserved when compared to that of introns described in the rabbit A1 [10], human A1 [8] and the rat A3 [2] receptor genes (Fig. 1b). The human A2a and A2b receptors have also been reported to contain an intron within this region [11,12]. Exon 1 of the A3 receptor consists of 5′ untranslated sequence and 323 bp of coding sequence and exon 2 consists of the remaining 606 bp of coding sequence and 3′ untranslated sequence. The nucleotides AG-GT [13] are conserved at the intron/exon acceptor and donor splice sites. The open reading frame codes for a protein of 319 amino acids, identical in sequence to that of the A3 receptor described by Salvatore et al. (1993) [4] but differing by a single nucleotide from that described by Sajjadi and Firestein (1993). The nucleotide at position 310 of the Sajjadi and Firestein sequence is an A in place of a G [3].

The 4.5 kb SstI-BamHI λ fragment contains 427 bp of 5′ sequence upstream to the initiating ATG codon. A further 899 bp upstream was determined by sequencing directly from the original λ clone, and from PCR products generated from this with a view to locating the promoter sequence (Fig. 2).

1a



1b

	111			↓					
Human A1	ATC	CCT	CTC	CG	G	TAC	AAG	ATG	GTG
	I	P	L	R		Y	K	M	V
Rabbit A1	ATT	CCT	CTC	CG	G	TAC	AAG	GCA	GTG
	I	P	L	R		Y	K	A	V
Human A3	CTT	ACC	GTC	AG	A	TAC	AAG	AGG	GTC
	L	T	V	R		Y	K	R	V
Rat A3	CTG	ACA	GTC	AG	A	TAT	AGA	ACG	GTT
	L	T	V	R		Y	R	T	V

Fig. 1. (a) Position of the intron within the open reading frame of the human A3 adenosine receptor. Upper and lower case letters indicate coding and intervening sequence respectively. Conserved nucleotides at donor and acceptor sites are underlined. (b) Conserved position of the intron within the open reading frame of adenosine receptors. An alignment of the nucleotide and corresponding amino acid sequences is shown for adenosine receptor subtypes where the precise position of the intron is known. The intron position is indicated by an arrow.

```
-1000 CTGCTGAATT TTATTTTGGA CTGTACATAT TTAGATGCTT AAGGTAAAAA
                GATA
-950  TGATAAAGCC CTCAAGCCAC TGTGTGGGTT GGGTCCAAGT GTTCCTTGCT
                                                  SP1
-900  GCTGCCTCTC TAACACGCCT GGTTAAAATA ATCCCTTTGG ATGGTGCTGA
                                                  SP1      NF-IL6
-850  GAAGCACCTG AACCAAGTGG GTCCCCAAAT AACTATGGCG TGCAAGTGTC
                                                  SP1      NF-IL6
-800  TGGTTCCCAG AAGTTGGTGA CTAGGTAAGC GACTCAGGGA GAGGGGCTGA
                                                  SP1      NF-IL6
-750  TTCCCAGACA GTCGCCTGTT CCTGCTGGGA TGGGGCTGAG GCTTGGGGAA
                                                  SP1      NF-IL6
-700  TGTGGGCAGG AGGATATGCC ATTTGATTCT GTTGCACACG TTCTTTTCCC
-650  TTCTTTCTGT ATGTCTGGTC ATTCTGCTAT TCTGTCGTTC CTCACATAGG
-600  TTGGACATTG GCCGGCTGCC AGCATAAGTG CCAGTGTGAT TTTGCTAGGG
-550  TGTGAGCTGA GAAAGAGAGG TGGAGGCTAA GCAGGTGTGA TGCTTCTCAG
-500  AGGTGCTGAG TTTTTGCCCT TCTGAGCAGG GAATCTTTGC TTATCCCTTT
-450  GACCAAGGAT CTTTGCTCCA AAGGCTGGGT ATCGGCTGTG CTCAGCAAAG
                CRE
-400  CGTCAACTCG TGCAAGAACT TAGCAGGAAT AGTTCTGGCT AAGGTTAGGA
-350  GGCTGCCACC AAAGTCTCTT TTTTGTTCCT CTGCTTCTCC CGTTTGCCTC
-300  CTTATCATGA GATCTTTTTG CTAAGCTGGC AGAAAGATTG CATAATCAGT
-250  GCTTCCAGCT CCGCTCCCAC CTGATCCTGC ACTGTCCTCT GGTCCCTGAA
                CAAT box homology
-200  TGAATGAACT CTGATACCCA ATCTTGTCTC GAGCCTTCTC TATGCCACTC
                                                  SstI
-150  ATGGCTCCTC TTCTGCTCTT TCCATCTTTT TGCTGAGAGT TACTGAGCTC
-100  TGTACTTCCT CTTGGCCCAT CTCACTTCCT GAAACACCCC TGAAGAGGGT
-50   TGCTTATCTT GATGGAACTC AAAAAGCCAA AAAGCTGCAG GCAGAGGCGT
                ↓1                    ↓2 ↓3      Inr  Inr
0    TGaggacatc tgtttgggga actaagagca gcagcacttt cagattcagt
                ↓4
50   ccatatagag ctgtcctaca gcattctgga aacttgagga tgtgcggtgc
100  ataaaggggc tggaagtgac ccacctgtga tgagcccttt ctaaggagaa
150  gggtttccaa gagatcaccc caccagaaaa gggtaggaat gagcaagttg
200  ggaattttag actgtcactg cacatggacc tctgggaaga cgtctggcga
250  gagctaggcc cactggccct acagacggat cttgctggct cacctgtccc
300  tgtggaggtt cccctgggaa ggcaagATGC CCAACAACAG CACTGCTCTG
```

Fig. 2. Sequence of the region upstream to the initiating ATG of the human adenosine A3 receptor. Potential transcription factor recognition motifs are boxed. Coding sequence is shown in italics. The transcription initiation site demonstrated in lung and liver is marked by a vertical arrow and 1. Arrows followed by 2, 3 and 4 denote the 5′ most nucleotide of published A3 cDNAs cloned from HL60 cells ([5] and personal communication), heart [3] and brain [4] respectively. Small variations at the 5′ termini of the cDNA sequences from heart and brain with the genomic sequence occur (TCCTTCTG and CGTCCTAG respectively at the 5′ most ends — differences are highlighted in bold).

This upstream sequence contains no TATA-like motifs. The absence of a TATA box from the promoter region of G-protein-coupled receptors is not uncommon, e.g. the PAF receptor [14]. The sequence 'CCAAT' is present at −179 to −183 which may form part of a functional promoter, although it seems uncharacteristically far upstream to constitute a functional 'CAAT' box. Two perfect matches to the consensus Inr sequence YYANT/AYY ([14] and references therein), a motif occurring at the transcriptional start site in many TATA-less genes, occur close to the 5′ end of one of the published A3 receptor cDNA clones although the A position does not coincide precisely with the first transcribed nucleotide (Fig. 2).

Comparison of the upstream sequence with a transcription factor recognition motif database revealed two SP1 consensus motifs or GC boxes, and consensus sites for recognition by NF-IL6 and GATA-1 and 3. Involvement of the latter factors in transcriptional control of this gene would be consistent with a role of the receptor in immune function. NF-IL6 is involved in the acute phase response and inflammation [15]

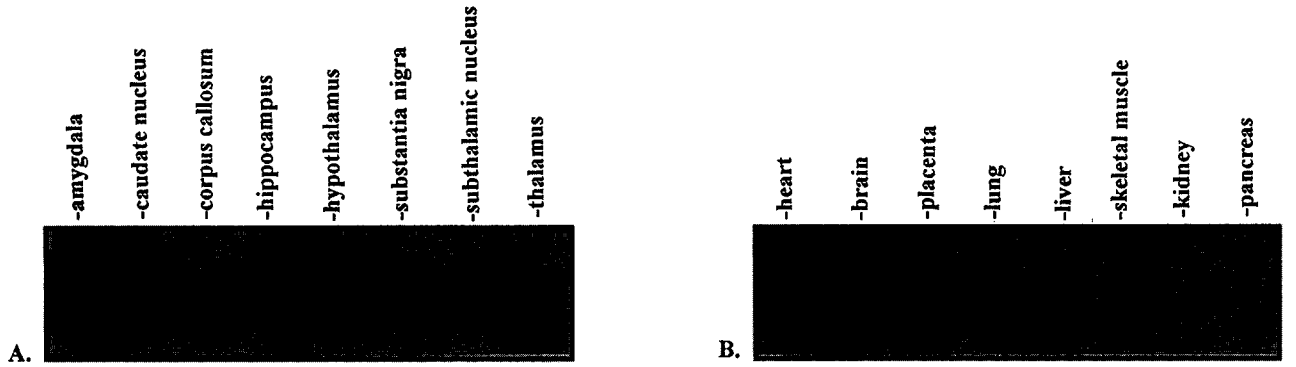


Fig. 3. Detection of a single A3 receptor transcript in a variety of tissues by Northern Blotting Hybridisation of a ^{32}-P labelled full length human A3 receptor cDNA probe to poly (A)+ RNA from various regions of the brain (A) and to peripheral tissues, including lung and liver, (B) is shown. In all cases hybridisation was seen to a single size of transcript of 2.1 kb.

and GATA binding factors have been demonstrated to regulate transcription in mast cells [16]. The sequence 'CGTCA' is also present and whilst this does not conform to standard cAMP responsive consensus elements, it has been shown to function as a CRE in the vasoactive intestinal peptide gene enhancer ([17] and references therein). It does not appear in the A3 gene as an inverted repeat, however, as is the case in the VIP CRE. Involvement of any of these factors in regulation of A3 receptor transcription is speculative at present as the functional significance of these motifs is not yet known. No matches to glucocorticoid responsive elements were found within this stretch of upstream sequence. It has been reported that rat A3 receptor mRNA expression is upregulated in response to dexamethasone in RBL cells [18]. If this also occurs in the human situation it must therefore be mediated by sequences elsewhere within the A3 gene, perhaps by control of mRNA stability.

The 265 bp [4], 293 bp [3] and 291 bp (Sean Munro, pers. comm.) of untranslated 5′ sequence present in published A3 adenosine receptor cDNA clones is contiguous with sequence determined from the genomic clone. This suggests that there are no introns and untranslated exons within this short region. No human A3 cDNA sequences have so far been reported in the literature that differ in their 5′ untranslated regions, as was the case for the human A1 receptor [8,9].

There is no significant homology between the immediate upstream sequences of the human A3 and A1 adenosine receptor genes. Comparison of the human A3 receptor 5′ flanking sequence to that available for the rat and sheep reveals homologies of 63% and 70% respectively. This suggests that this region is functionally conserved rather than intervening sequence.

3.2. 5′-RACE

To determine the site(s) of transcription initiation of A3 receptor mRNA, cDNA from lung and liver was PCR amplified using an A3 receptor specific antisense primer and a 5′ anchored primer. PCR was also performed on phage from a human eosinophil cDNA library using a sense vector specific primer and the same A3 specific antisense primer as used in the RACE analysis. Sequence determined from cloned PCR products (of approximately 200 bp) from all three RNA sources, i.e. lung, liver and granulocytes, began at a guanine nucleotide, 325 bp upstream of the initiating ATG (Fig. 2). No products initiating at nucleotides further upstream of this position or with alternative 5′ sequences were found using this protocol. This therefore appears to represent the major transcriptional start site in the tissues tested. This position is 33, 60 and 34 bp upstream from the start of cDNAs cloned from striatum [4], heart [3] and HL60 cells [5]. The differences could be tissue specific, cDNA cloning artefacts or indicate that other start sites are used rarely but were not detected in this experiment.

3.3. Northern blot analysis

Northern blot analysis revealed a single 2.1 kb transcript hybridising to a ^{32}P-labelled full length human A3 AR cDNA probe, expressed primarily in the liver, lung, placenta and in all regions of the brain tested (Fig. 3). No signal was observed in heart, skeletal muscle, kidney or pancreas using this technique. This transcript size and distribution is consistent with published data for human tissues (with the exception of expression in the kidney) where available [3] and provides no evidence for the existence of alternatively spliced mRNA variants of different sizes in these tissues.

The contiguous nature of published cDNA sequences with the A3 genomic sequence presented here, the presence of a single size of transcript and the single transcriptional start site, as defined by 5′-RACE analysis, is strongly suggestive of the absence of alternative splicing occurring in the 5′ untranslated region of the adenosine A3 receptor gene. Of course, the possibility that transcripts differing significantly at their 5′ ends may occur in different tissues, not yet examined, cannot be ruled out. Levels of A3 receptor mRNA vary considerably in different tissues examined, as demonstrated by this and other studies, and transcriptional control of A3 mRNA must therefore operate by other mechanisms.

Acknowledgements: We are grateful to Nikki Crocker for oligonucleotide synthesis and to Victoria Perreau and Steve Foord for help with the library screening.

References

[1] Meyerhof, W., Muller-Brechlin, R. and Richter, D. (1991) FEBS Lett. 284, 155–160.
[2] Zhou, Q.-Y., Li, C., Olah, M.E., Johnson, R.A., Stiles, G.L. and Civelli, O. (1992) Proc. Natl. Acad. Sci. USA 89, 7432–7436.
[3] Sajjadi, F.G. and Firestein, G.S. (1993) Biochim. Biophys. Acta 1179, 105–107.

[4] Salvatore, C.A., Jacobson, M.A., Taylor, H.E., Linden, J. and Johnson, R.G. (1993) Proc. Natl. Acad. Sci. USA 90, 10365–10369.
[5] Munro, S., Thomas, K.L. and Abu-Shaar, M. (1993) Nature 365, 61–65.
[6] Linden, J. (1994) TIPS 15, 298–306.
[7] Linden, J., Taylor, H.E., Robeva, A.S., Tucker, A.L., Stehle, J.H., Rivkees, S.A., Fink, J.S. and Reppert, S.M. (1993) Mol. Pharmacol. 44, 524–532.
[8] Ren, H. and Stiles, G.L. (1994) J. Biokl. Chem. 269, 3104–3110.
[9] Ren, H. and Stiles, G.L. (1994) Proc. Natl. Acad. Sci. USA 91, 4864–4866.
[10] Bhattacharya, S., Dewitt, D.L., Burnatowska-Hledin, M., Smith, W.L. and Spielman, W.S. (1993) Gene 128, 285–288.
[11] Peterfreund, R.A., MacCollin, M., Gusella, J. and Fink, J.S. (1994) Drug Dev. Res. 31, 306.
[12] Jacobson, M.A., Johnson, R.G., Luneau, C.J. and Salvatore, C.A. (1995) Genomics 27, 374–376.
[13] Breathnach, R. and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349–383.
[14] Pang, J.-H., S., Hung, R.-Y., Wu, C.-J., Fang, Y.-Y. and Chau, L.-Y. (1995) J. Biol. Chem. 270, 14123–14129.
[15] Akira, S., Isshiki, H., Sugita, T., Tanabe, O., Kinoshita, S., Nishio, Y., Nakajima, T., Hirano, T. and Kishimoto, T. (1990) EMBO J. 9, 1897–1906.
[16] Zon, L.I., Gurish, M.F., Stevens, R.L., Mather, C., Reynolds, D.S., Austen, K.F. and Orkin, S.H. (1991) J. Biol. Chem. 266, 22948–22953.
[17] Fink, J.S., Verhave, M., Kasper, S., Tsukada, T., Mandel, G. and Goodman, R.H. (1988) Proc. Natl. Acad. Sci. USA 85, 6662–6666.
[18] Ramukar, V., Wilson, M., Dhanraj, D.N., Gettys, T.W. and Ali, H. (1995) J. Immunol. 154, 5436–5443.