Structure and Expression of Mouse α_1 -Acid Glycoprotein Gene-3 (AGP-3)

CHING-JIN CHANG, MING-YANG LAI, DING-SHINN CHEN,
and SHENG-CHUNG LEE

ABSTRACT

The genome of *Mus domesticus* has multiple genes of the α_1 -acid glycoprotein (AGP). Two cDNA clones were identified corresponding to AGP-1 and AGP-2. Moreover, two alleles of AGP-1 exist in inbred mice. The genomic DNA of the AGP-2 gene has been cloned and studied. Here we report the genomic organization of three *M. domesticus* AGP genes, the sequence analysis of the AGP-3 genomic DNA, and the expression of the AGP-3 gene. The major structural differences between AGP-2 and AGP-3 genes are located in introns 1 and 5. The low level of AGP-3 mRNA can be detected by the polymerase chain reaction (PCR). The molecular basis of the low level expression of AGP-3 and the possible classification of AGP-3 as a pseudogene are discussed.

INTRODUCTION

α_1 -ACID GLYCOPROTEIN (AGP) is a single-chain polypeptide with a molecular weight of approximately 44,000, containing about 45% carbohydrate. This plasma protein is an acute-phase reactant synthesized in the liver of *Mus domesticus*. Plasma concentrations of AGP in the mouse increase several fold during infection, inflammation, or are induced following subcutaneous injection of turpentine or lipopolysaccharide. Expression of the AGP gene is regulated by interleukin-1 (IL-1), tumor necrosis factor- α (TNF- α), interleukin-6 (IL-6), and glucocorticoids (Baumann and Maquat, 1986; Klein *et al.*, 1987; Prowse and Baumann, 1988; Won and Baumann, 1990). The biological function of AGP is unknown, but there are indications that it may suppress the immune response (Bennett and Schmid, 1980). Some experiments suggest that AGP is a nonspecific antiinfection agent (Friedmann, 1983), and that it possesses nerve growth-promoting activity (Liu *et al.*, 1988).

Multiple forms of mouse AGP can be demonstrated by two-dimensional gel electrophoresis (Baumann *et al.*, 1984), and a high degree of heterogeneity in the amino acid composition of human AGP has been described (Schmid, 1975). Dente *et al.* (1987) cloned and sequenced genomic DNA segments coding for three human AGP genes. The one designated AGP-A, is transcribed to a greater extent

than AGP-B or AGP-C (Dente *et al.*, 1987). Previously, we reported the existence of at least two AGP genes, AGP-1 and AGP-2, in *M. domesticus* (Lee *et al.*, 1989). Both *cis*-elements and *trans*-acting factors reportedly involved in the regulation of AGP expression have been analyzed extensively (Lee *et al.*, unpublished observations). One of these, AGP/EBP (a liver-enriched transcription factor), has been analyzed and is key to regulating AGP expression (Chang *et al.*, 1990). To study the genomic organization, the molecular basis of the differential expression, and the regulation of AGP expression, we analyzed the *M. domesticus* AGP genes. We now show that there are three AGP genes in the *M. domesticus* genome; we determined the complete sequence of the newly identified gene, AGP-3, and show that AGP-3 mRNA is expressed at very low levels.

MATERIALS AND METHODS

Isolation of AGP genomic clones

A genomic library from T lymphocytes of C57BL/6 *M. domesticus* and *E. coli* K803 (r^- , m^- , gal^- , met^-) were gifts of Dr. L. Mori (University of Milan, Italy). The library was screened with an AGP-1 cDNA probe (Lee *et al.*, 1989). Of the approximately 4×10^5 plaques screened, four positive plaques were recovered and purified by secondary screening. Sequencing was carried out by the dideoxynucleotide termination method (Sanger *et al.*, 1977).

RNase protection assay

The RNase protection assay was performed as described by Melton *et al.* (1984). Total RNA (20 μ g) was dissolved in 20 μ l of hybridization buffer (60% formamide, 40 mM PIPES pH 6.7, 400 mM NaCl, 1 mM EDTA) containing the riboprobe ($2-5 \times 10^5$ cpm). The hybridization mixture was heated at 85°C for 5 min, followed by incubation at 45°C for a minimum of 4 hr. Then 300 μ l of RNase digestion buffer (10 mM Tris-HCl pH 7.5, 5 mM EDTA, 300 mM NaCl) containing 40 μ g/ml of RNase A and 50 U/ml of RNase T1. The reaction was continued at 30°C for 30 min, then stopped by addition of 20 μ l of 10% NaDodSO₄ and 50 μ g of proteinase K. After incubating at 37°C for 15 min, the mixture was extracted with phenol/chloroform and precipitated with ethanol (with 20 μ g carrier tRNA included). The sequences of the RNA samples were then determined and analyzed.

Polymerase chain reaction

Each polymerase chain reaction (PCR) included 10 ng of cloned DNA fragments or 100 ng of single-stranded cDNA, 1.25 units of *Taq* polymerase, 100 μ M deoxyribonucleotides, 2.5 nM each oligonucleotide primer, 1.5 mM MgCl₂, 50 mM KCl, and 10 mM Tris-HCl pH 8.3 in a final volume of 100 μ l. The reaction mixture was overlaid with mineral oil, and run through 30 cycles of 1 min each at 94°C, 1 min at 57°C, and 30 sec at 72°C. The product was analyzed by native 20% polyacrylamide gel.

RESULTS

Isolation and characterization of *M. domesticus* AGP genomic clones

Several genomic clones were obtained by screening the EMBL-3 library (C57BL/6) with an AGP cDNA probe. Four positive plaques of λ Agp-1 contained an insert of ap-

proximately 16 kb, while the λ Agp-25 plaques contained an insert of approximately 18 kb. Three distinct AGP genes were identified by restriction mapping and partial sequencing of these five clones (Fig. 1). cDNAs corresponding to two of them, AGP-1 and AGP-2, had been characterized previously (Cooper and Papaconstantinou, 1986; Lee *et al.*, 1989), and a third gene designated AGP-3 was identified. AGP-1 is contained in λ clone Agp-1 while AGP-2 and AGP-3 are in λ Agp-25 clones.

Only two types of cDNAs and the corresponding proteins of AGP have been identified in the *M. domesticus* (Baumann, 1984; Lee *et al.*, 1989). To address the issue of whether or not AGP-3 is expressed, the genomic DNA containing the coding and noncoding sequences of AGP-3 was sequenced. This gene spans 4 kb from 548 bp upstream from the putative transcription initiation site to 130 bp downstream of the polyadenylation signal, AATAAA (Fig. 2). This gene has 6 exons and 5 introns, a structure analogous to the rat (Liao *et al.*, 1985; Reinke and Feigelson, 1985), human (Dente *et al.*, 1985; Dente *et al.*, 1987), and mouse AGP-2 genes (Cooper *et al.*, 1987; the Papaconstantinou group designates it AGP-1, while our designation is AGP-2). The consensus sequence for splice junctions is preserved in both AGP-2 and AGP-3 genes.

The mouse AGP-3 and AGP-2 gene (Cooper *et al.*, 1987) sequences are very similar; however, there are two major differences between them: AGP-3 lacks 86 nucleotides in intron 1 (between +331 and +332), while AGP-2 has an additional (GT)₂₈ tract in intron 5 (+2815), not found in AGP-3.

Differential expression of AGP genes

It would be interesting to know the relative level of mRNA expression among AGP-1, -2, and -3 genes in *M. domesticus* liver under normal conditions and during the acute-phase reaction. Because of the high degree of similarity among these three genes, an RNase protection analysis was performed to discriminate among these mRNAs.

FIG. 1. Partial restriction map of α_1 -acid glycoprotein genomic clones, λ Agp1 and λ Agp25. Shaded boxes represent AGP-1, -2, and -3 genes. Exons are represented by solid boxes. Abbreviations: B, *Bam* HI; E, *Eco* RI; H, *Hind* III; S, *Sal* I.

-548	GGATCCTTTCTGCTGTAGATACTGGGAGCTTTGCTGAACATAGATGTTCAAGTCA	2313	CCCCAGGATTCCTGGGAGAGGTGCCTGCACACAGACTGCCATTACAGCATGA
-493	GAATCAACCCCTTTTGGGCATTGGATGCCTCTAGGCTGGGAAGGGTCTCAGG	2368	CTTCCACCCACACCACTGGGCCAGTAGACTCATCTGCACCTGTGGACAGAAGTG
-438	AACATCACACTCCTTTGGAACTAATCCATCTTTGTCTTGGGCCCTTAACCTGAG	2423	TTAGATAATGCCTGCCCTTTGGGGATTCTGCTCAACAATCAATGGGTGAATAAGCC
-383	CCCCAAGTGTCTTCTATGTTCACTATGAACCTTGACCTGGGACCCCTTCTTATCA	2478	GGAGCTCAGAGATGAGGGACAACCTACCCCTAGACTAGTGGTTCTCTAGGATGGGT
-328	TGCTTGGGGGGGGTGTATGTATGTAGTGTTCACCTCTGCTAGGCAGCTTCAT	2533	CTCAACCTGTATGTCAGATGTCCTGCACGCTCAGGTATTTATTGATTATAATAGT
-273	GGGATAAGAGAGGGTGGGGACCACTGTCTGGGACCTAAGTATCATCAGGTACCC	2588	AGCATATAATTACAGTTATTGAAGTAGCAATGAAATCATGGTAGGTGATCACCAC
-218	TGTACCCACCTTGACCATGAATCAGCCACTCTGGTGTAGGGCAGGACGCTGTGC	2643	AACACAAGAACTGTGTTAAAGTTTTCATTATTAGGAAGGCTGAGAACCCTGT
-163	ACAGCCAGCTGGCTGAGAGAGCTGCACAAAGCTGGCTTGAGGGAACTTTTGCAC	2698	CCTGCCACTGCAGGGAGCCATGGCAGATCTAAGACACATCTGGTTGACACTACCG
-108	AAGACATTTATCAAGTCTGGTGTAGTGTGTGCTGCTCTACAGCCCTGGCTG	2753	GGCCATTTTGACCAACAACAGTACTCCCCCAACCCCTCACAATAGGTGTATT
-53	CAGTCCCATGCGCTCCCCACATCTGTATATAAGGCCACTGCACCCCTCCAGCCAC	2808	CATAGCTAGTGTGTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
3	CAGTTATCTCTCTCAAGCTCTGGTGCTCTGAGTATCTCTCAGCATGGAACACTAC	2863	CATGTGTGGATGTGAGGACACCTGTGGGCTATGTCTCTCTTACCATTCTCTC
58	ACGGTCTCATCATGTTGAGCCTCTGCCACTGTGGAAGCTCAGAACCCAGAGC	2918	CTGGGCTCTGGTTAAGGCTGGGTGGCTTCAAGCTGCCCTCAGGCTTACCTACC
113	ThrValLeuIleMetLeuSerLeuLeuProLeuLeuGluAlaGlnAsnProGluH	2973	TTGCCATTTTTTTTTTGTGTCTGTCTGTTTTTTCTGTTTTTGTGTTTTGT
168	ATGCCATCAACATAGGCGACCTTATCAACATGAGACCTGAGCTGGTGAAGTGT	3028	ATTAACTCTGCAGCCAGGCTACTCTACTGCAACTCATAGCAATCCTCTGCGCT
	isAlaIleAsnIleGlyAspProIleThrAsnGluThrLeuSerTrp	3083	CAGTATTCATCAACCTGGTGTGTGTCACGACCTGGCTTACTACTCTGCTCT
223	CTGCCGGGGCTGACCTGTTCTACTGTAGGATCTACTCTTCTCTGGGCTTT	3138	CCTCCCTGATATCTCCAGGACAGGTGCAGTGAACAGGAAAAGACATCTTGAG
278	CCCTTCCCTGGTGTCTGTGTGAGCTCTGGGCTCTGGTACTGCCCTTCCCACTT	3193	TTGGAGAAGGAGACCAAGAAAGATCCTGAGGAAGCCAGGCATGAACCTCAGCTCT
333	GTGTATACCTGGTGGCATCCCCCTGTCTCCAAAGGCAGAATCATCACTGTGAG	3243	LeuGluLysGluThrLysLysAspProGluGluSerGlnAla
388	CATAGCTTGCCTGCCCTCATCGTGGATGAATGCCAAGTCTCTACTACAAGGCC	3303	CTGGTCTCTTGGGCTGTCCCATGTGTACACACCTTACCCATCCTGGTCACT
443	TGCTCATCTGTGTGCTGCTTCTCCCGACTATCTGGGAAATGGTTTCTCATTCG	3358	TTGATCTGTCTCTGTAAATAAAGGTTGTGCTGACACTGTCAATATCATTTCTT
498	TGTGGTGACTCAGACCTGATTATAGCAGGAAATTCAAAAGGTACAGACTATA	3413	TGCTCCCTTCCTTTCTCCCTCCCTCCCTCCCTCTCGTGAGAGTCTTGAGTGA
553	aValAlaAspSerAspProAspTyrArgGlnGluIleGlnLysValGlnThrIle		GCTAGCTAAGTCAATAACCTGCCAGGAATTGAAAGGCTCT
608	yrAsnSerAsnLruLeuGlyPheLeuArgGluAsnGlyThrLeuPheLysTyrG		
663	PhePheTyrLeuThrLeuAsnLeuIleAsnAspThrMetGluLeuArgGluTyrH		
718	ACACCAAGTGAGTCTTGTAAACAGCCGCCACCTGGCCCTGGCTTCCACTCCC		
773	isThrLy		
828	AGATTCTTAGAGACTGAGCAAACTGGCTCTGCCTGGCCCTCCCAACCACTTTCA		
883	GAAATGGGGACAGCTGTCTTGCTCTTGCCCTTCTACCTGGGCTAGTCAGAT		
938	CACCTCTCCATCAGTTGTCTCTCTTTGCTTTTATAGAGATGACCACTGTGTCT		
993	ATAACTCAACCTTCTGGGATCCAGAGAGAGAATGGGACCTCTTCAAGTATGG		
1048	sAspAspHisCysValT		
1103	TGAAGTGTGAACTCACTCTTCTTCTGGGAGGGTATTGCCAGTCTCTGAGGGACAG		
1158	CAGAACAGGGCAGTTTGGCTGTCAAGTCACTCTCTGGGCTTGTAAAGTGGACGT		
1213	AGATTTCAACTGGAGTCCAGCTGCCAGGCTGACTGTGTTTGTATCAGGCTTTAT		
1268	GACCTTCTGTGACCTCAGAGTGGAAAGCATGGGTGGGAAACAGTGACTTTAG		
1323	CCACACCCCAAGTACCAGGGACAGCATGGAGAGAACAACCCAATGTCTGGTG		
1378	GGCCAGATCAGACACTGGTTAGCTTTTAATTGCTTAAGCAGATGTTTGTATGTT		
1433	TAAGGGGAAGTGTAAACATTATTATGCCCCACCACACCAAGGCGCCTT		
1488	TCCAAGGCCCCAGCTCTTCCATACCTATGAAGAATGAGAAGTGAAGCTTGATCC		
1543	AGCATGAGGCTGAGCACATGGCAGCCTCAGGGAGCCCCAGGCACTTGCCATAGCT		
1598	ATGTGCTTCCCTTCCCTTGGGATGTGAACACATCATCTTCTAGTGACTCACA		
1653	AGACCTTCTCCACAGAGGAGAAGTAGAAAACCTTCTCACCTGAGAGTGCTA		
1708	lGlyGluValGluAsnProSerHisLeuArgValLeu		
1763	GAGAAACATGGGGCCATCATGCTTTTCTTTGACCTGAAGGATGAGAAGACGGG		
1818	ArgLysHisGlyAlaIleMetLeuPhePheAspLeuLysAspGluLysLysArgG		
1873	GACTGTCCCTGAGCGGTAGGGTCTCTCATCTCTGGTGGCCGACCTCACTGGC		
1928	lyLeuSerLeuSerA		
1983	CTTACTCTTGGTACCCACCCACATCCACATCCCTACCTGGCCCTCCCATCTACT		
2038	GGACCCATAGCCAGCATAACTTTGGATCCCTTCTCCCATGCAGCTAGAAGGCCAG		
2093	ATATCCCCCGGAGCTCGCGGAAGTATTTGAGAAGGCTGTACACACGTTGGGCAT		
2148	spliceProProGluLeuArgGluValPheGlnLysAlaValThrHisValGlyMe		
2203	GGATGAATCAGAAATCATATTGTGCTGACTGGAAGGTAAGTGAAGAGGCTGT		
2258	tAspGluSerGluIleIlePheValAspTrpLysLys		
	ACGATACCACCCAGCAGTGGCCAGTGTGAGTACCTTAGAGGCTCAGAGAGG		
	GCAAGCTTCTGGTTAAGGCAGCTCAGCGAGGCAGGTATCTGTAACTCTCTGC		
	CTCTCTCTCATCAGGATCAGACAGACCTTAGATGGCAGTGAGCCTCAGGGAG		
	GTGAAGTTAAGTAGGAGTCTGGAAGCTTGTGGAGGATAAGAGGAAGATCAGG		
	AGGGTCACTTAGGGAACAGCCAGTGCCAGGCTGCCAGGTTTCTCTGTCCTTCA		
	TATTACTACCTTTTCAAGCAGGAGTTTGTATTGACATCTTCCATGTCAACCAAC		
	TCCAGCAAGCCGATGGCTTTGATAGGAGGTTGACCACACTGAGACTCTTGAT		
	GTCCGGTCTACACATTGTGACAGGGAGAGGCAGCATCAGTTTGTGTTTTCACCC		
	ATGGCGAATGCATGGGATCAACAGTCACCTTGCATGTAGTTTAAGATACTCAAT		
	AGCTTTTGTAACTTCTTCTGTGTCACCTGAGCCTTCTCTGGCATCATCACCAG		

FIG. 2. Nucleotide sequence of mouse AGP3 gene. The putative TATA box and poly(A) addition signal sequence (AATAAA) are boxed. Six exons and encoded amino acid sequences are shown. The site of initiation of transcription is depicted by +1.

The specific riboprobes of AGP-1 and AGP-2 (Fig. 3a, upper panel) was designed from their cDNA sequences (Lee *et al.*, 1989). The riboprobe of AGP-3 was derived from the genomic segments that contains the putative exon 4 (Fig. 3a). The result of RNase protection analysis is shown in Fig. 3a. When using RNA from normal liver and from lipopolysaccharide-stimulated liver, protected bands were detected corresponding to 247 bp of AGP-1 and 318 bp of AGP-2 (Fig. 3a, lower panel, lanes 1, 2 or N-L and LpS-L). However, there were no signals for the probe derived from AGP-3. The expression level of AGP-1 is about fivefold higher than that of AGP-2 and both genes respond to the acute-phase reaction, thus AGP appears as a liver-specific gene, because no signal was detected in RNA from the spleen.

Because the level of expression of AGP-3 is much less than AGP-1 and AGP-2, a more sensitive method (*e.g.*, PCR) was used to detect it. Primers for AGP-2 and AGP-3 were synthesized for PCR experiments (Fig. 3b, upper panel). Using genomic DNA of AGP-2 or AGP-3 genes as templates, we showed that the primers derived from AGP-2 and AGP-3 were specific for the corresponding templates (Fig. 3b, lower panel, lanes 1, 2, 4, 5). However, when these primers were used for reverse transcriptase (RT)-PCR using RNA derived from *M. domesticus* liver, an AGP-2, but not an AGP-3 fragment can be seen in the agarose gel by ethidium bromide. This does not exclude the possibility that the level of expression of AGP-3 might be below the detection limit of ethidium bromide staining.

FIG. 3. a. Upper panel. RNase protection assay. Specific riboprobes and their protected length for AGP-1, -2, and -3 are represented. The riboprobes of AGP-1 and -2 are derived from their cDNAs in pGEM3 vector. The fragments are numbered relative to the mRNA transcription initiation site indicated by +1. The riboprobe of AGP3 contains exon 4 and intron segments in pGEM 4 vector. Lower panel. RNase protection experiment on RNA prepared from mouse liver and spleen. 1, 2, and 3 indicate the riboprobes derived from AGP1, 2, and 3. N-L, LPS-L, and LPS-S indicate the sources of RNA from normal liver, LPS-stimulated liver, and LPS-stimulated spleen, respectively. Specific protection bands are represented by arrows. b. PCR amplification of AGP-2 and AGP-3. The right (R) and left (L) primers of AGP-2 and AGP-3 are shown above, and the PCR products analyzed on 20% polyacrylamide gel are shown below. Lanes 1-6, Ethidium bromide staining patterns; lanes 7 and 8, autoradiographic patterns. Lanes 1-3, 7, AGP2 primer; lanes 4-6, 8, AGP3 primer. Templates used: lanes 1 and 4, AGP2 genomic DNA; lanes 2 and 5, AGP3 genomic DNA; lanes 3, 6, 7, 8, single-stranded cDNA derived from mouse liver.

To overcome this, 5'-end labeled primers were employed for RT-PCR. As demonstrated by this analysis, the signal for AGP-3 is at least two orders of magnitude weaker than AGP2 (Fig. 3b, compare lower panel lanes 7 and 8).

DISCUSSION

We have isolated the genomic clones for the entire sequence of three of the *M. domesticus* AGP genes, AGP-1,

-2, and -3. The newly identified gene, AGP-3, is located approximately 6 kb upstream from AGP-2. By the restriction map of the λ clones of Agp-1 and Agp-25 and the evidence that AGP-1 is proximal to the centromere (Baumann *et al.*, 1984), we predict the genomic organization of these three genes is: AGP-1, AGP-3, and AGP-2, arrayed in tandem. AGP-3 and AGP-2 are closely linked while there is some distance (not determined) between AGP-2 and AGP-3.

We have analyzed the sequence of AGP-3 from 548 bp

FIG. 4. Comparison of sequences in the promoter region of mouse AGP genes. The TATA box is boxed. The numbers are indicated according to the AGP3 sequence. The AGP/EBP binding motifs are indicated by solid underlining, GREs are indicated by overlining, and the consensus acute-phase sequence is indicated by dashed underlining.

upstream from the putative transcription initiation site to 130 bp 3' of the polyadenylation signal. This gene contains 6 exons and 56 introns. Its coding sequences are normal and contain no frameshift or nonsense mutations. The exon/intron splicing sites, and the 5'-flanking and 3'-flanking sequences are conserved, as compared to the AGP-2 gene. The 5'-flanking regions of three AGP genes are similar in the first 180 bp upstream from the cap site, as shown in Fig. 4. AGP-3 as well as AGP-1 and AGP-2 contain the potential glucocorticoid-responsive element (−125 to −113) and three AGP/EBP binding sites located at −119 to −110, −107 to −98, and −87 to −78 (Chang *et al.*, 1990). A sequence of 38 bp (−15 to +23) is similar to sequences observed in the three human acute-phase proteins (Dente *et al.*, 1985) located in AGP-3 gene. Part of the regulatory region (−180 to +60) of these three genes were separately fused to the chloramphenicol acetyl transferase (CAT) (Gorman *et al.*, 1982) reporter gene and then transfected into a BHK cell line. The expression was monitored by assaying the activity of the CAT enzyme. The activity of AGP-3 promoter is similar to those of AGP-1 and AGP-2 (data not shown).

Using RNase protection to assay the endogenous expression of the three AGP genes, we showed that the amount of AGP-1 was about fivefold higher than AGP-2 in a normal *M. domesticus* liver and during acute-phase reaction. However, the expression of AGP-3 could not be detected by the RNase protection assay. To determine if the AGP-3 gene is a *bona fide* pseudogene or just expressed at very low level, we synthesized specific primers to amplify

AGP-3 mRNA by RT-PCR. AGP-3 mRNA was detected in this experiment, but the signal was weaker than that of AGP-2. Therefore, the AGP-3 gene is transcribed at a very low level. The expression of AGP-3 protein is yet to be determined.

We do not know why AGP-3 is so poorly expressed. One possibility is that structural aberrations are present in the introns. Comparing AGP-3 and AGP-2, the major differences are the lack of 86 bp of intron 1 and no (GT)₂₈ tract in intron 5. The 86 bp of intron 1 may contain the branch site of splicing; the deletion of this sequence may cause AGP-3 RNA to be spliced inefficiently. Another possible explanation is that alternating purine-pyrimidine sequences can form Z-DNA, which may introduce a potential site for gene regulation. Alternatively, the low-level expression may be due to the positional effect. Proudfoot (1986) has shown that transcription of the first gene in a gene cluster interferes with the transcription of the following gene. It is possible that the transcription of AGP-1 gene interferes with the transcription of AGP-3 gene.

Two AGP proteins have been identified corresponding to the products of AGP-1 and AGP-2; a protein product of AGP-3 has not been identified. By comparing the deduced amino acid sequence of AGP-3 and those of AGP-1 and AGP-2 (Fig. 5), AGP-3 would contain 206 amino

FIG. 5. The amino acid sequence of mouse AGP-1, -2, and the putative AGP-3. Dots above the single-letter amino acid mark the position of putative glycosylation sites (sequence Asn-X-Ser/Thr).

acids, including the 18-residue putative signal peptide. There are 45 amino acid substitutions between the encoded AGP-3 protein and that of AGP-1 or AGP-2. AGP is a highly glycosylated protein; therefore, it was of interest to localize potential carbohydrate attachment sites indicated by the sequence of Asn-X-Thr/Ser. Five potential sites can be found in the AGP-1 and six potential sites in AGP-2; however, only three potential glycosylation sites existed in the putative AGP-3 polypeptide. If AGP-3 expresses a functional protein, the amino acid substitutions would affect its function. The existence and the potential functional implications of AGP-3 protein remained to be investigated.

ACKNOWLEDGMENTS

We thank Dr. George Bolton for editing and Ms. Joanne Kahrmann and Ru-Ju Chen for typing this manuscript. This research was supported by Grant NSC80-0412-B002-09 from the National Science Council.

REFERENCES

- BAUMANN, H., and MAQUAT, L.E. (1986). Localization of DNA sequence involved in dexamethasone-dependent expression of the rat α_1 -acid glycoprotein gene. *Mol. Cell. Biol.* **6**, 2551–2561.
- BAUMANN, H., HELD, W.A., and BERGER, F.G. (1984). The acute phase response of mouse liver. Genetic analysis of the major acute phase reactants. *J. Biol. Chem.* **259**, 566–573.
- BENNETT, M., and SCHMID, K. (1980). Immunosuppression by human plasma α_1 -acid glycoprotein: Importance of the carbohydrate moiety. *Proc. Natl. Acad. Sci. USA* **77**, 6109–6113.
- CHANG, C.J., CHEN, T.T., LEI, H.Y., CHEN, D.S., and LEE, S.C. (1990). Molecular cloning of a transcription factor, AGP/EBP, that belongs to members of the C/EBP family. *Mol. Cell. Biol.* **10**, 6642–6653.
- COOPER, R., and PAPACONSTANTINOU, J. (1986). Evidence for the existence of multiple α_1 -acid glycoprotein genes in the mouse. *J. Biol. Chem.* **261**, 1849–1853.
- COOPER, R., ECKLEY, D.M., and PAPACONSTANTINOU, J. (1987). Nucleotide sequence of the mouse α_1 -acid glycoprotein gene 1. *Biochemistry* **26**, 5244–5250.
- DENTE, L., GILBERTO, G., and CORTESE, R. (1985). Structure of the human α_1 -acid glycoprotein gene: Sequence homology with other human acute phase protein genes. *Nucleic Acids Res.* **13**, 3941–3952.
- DENTE, L., PIZZA, M.G., METSPALU, A., and CORTESE, R. (1987). Structure and expression of the genes coding for human α_1 -acid glycoprotein. *EMBO J.* **6**, 2289–2296.
- FRIEDMAN, M.L. (1983). Control of malaria virulence by α_1 -acid glycoprotein (orosomucoid), an acute phase (inflammatory) reactant. *Proc. Natl. Acad. Sci. USA* **80**, 5421–5424.
- GORMAN, C.M., MOFFAT, L.F., and HOWARD, B.H. (1982). Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *Mol. Cell. Biol.* **2**, 1044–1051.
- KLEIN, E.S., REINKE, R., FEIGELSON, P., and RINGOLD, G.M. (1987). Glucocorticoid-regulated expression from the 5'-flanking region of the rat α_1 -acid glycoprotein gene. *J. Biol. Chem.* **262**, 520–523.
- LEE, S.C., CHANG, C.J., LEE, Y.M., LEI, H.Y., LAI, M.Y., and CHEN, D.S. (1989). Molecular cloning of cDNAs corresponding to two genes of α_1 -acid glycoprotein and characterization of two alleles of AGP-1 in the mouse. *DNA* **8**, 245–251.
- LIAO, Y.C., TAYLOR, J.M., VANNICE, J.L., CLAWSON, G.A., and SMUCKLER, E.A. (1985). Structure of the rat α_1 -acid glycoprotein gene. *Mol. Cell. Biol.* **5**, 3634–3639.
- LIU, H.M., TAKAGAKI, K., and SCHMID, K. (1988). *In vitro* nerve-growth-promoting activity of human plasma α_1 -acid glycoprotein. *J. Neurosci. Res.* **20**, 64–72.
- MELTON, D.A., KRIEG, P.A., REBAGLIATI, M.R., MANIATIS, T., ZINN, K., and GREEN, M.R. (1984). Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Res.* **12**, 7035–7056.
- PROUDFOOT, N.J. (1986). Transcriptional interference and termination between duplicated α -globin gene constructs suggests a novel mechanism for gene regulation. *Nature* **322**, 562–566.
- PROWSE, K.R., and BAUMANN, H. (1988). Hepatocyte-stimulating factor, α_2 -interferon, and interleukin-1 enhance expression of the rat α_1 -acid glycoprotein gene via a distal upstream regulatory region. *Mol. Cell. Biol.* **8**, 42–51.
- REINKE, R., and FEIGELSON, P. (1985). Rat α_1 -acid glycoprotein. Gene sequence and regulation by glucocorticoids in transfected L-cell. *J. Biol. Chem.* **260**, 4397–4403.
- SANGER, F., NICKLEN, S., and COULSON, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
- SCHMID, K. (1975). In *The Plasma Proteins*, vol. 1. F. Putnam, ed. (Academic Press, New York) pp. 184–228.
- WON, K.-A., and BAUMANN, H. (1990). The cytokine response element of the rat α_1 -acid glycoprotein gene is a complex of several interacting regulatory sequence. *Mol. Cell. Biol.* **10**, 3965–3978.

Address reprint requests to:

Dr. Sheng-Chung Lee
Institute of Biological Chemistry
Academia Sinica
Taipei, Taiwan

Received for publication October 21, 1991, and in revised form December 3, 1991.