Volume 14 Number 2 1986 Nucleic Acids Research

The sequence and structure of a new serum amyloid A gene

Robert S.Stearman+, Clifford A.Lowell, Cynthia G.Peltzman and John F.Morrow

Howard Hughes Medical Institute and Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA

Received 4 October 1985; Accepted 2 December 1985

ABSTRACT

The acute phase response is characterized by changes in the serum concentrations of many proteins. A 1000-fold increase in the concentration of serum amyloid A (SAA) protein occurs within 24 hours of LPS injection in the mouse. We have isolated a cDNA clone and its corresponding genomic phage for a third, previously unreported SAA protein. The sequences of the cDNA, the gene's exons and neighboring DNA are presented along with the mapping evidence supporting the gene structure.

INTRODUCTION

Infection, inflammation, and neoplasia in mammals cause major changes in the concentrations of many serum proteins. This acute phase response produces changes which range from a 2-fold decrease in the level of serum albumin to a 1000-fold increase of serum amyloid A (SAA) protein concentration (1). In mice, the acute phase response can be induced experimentally by a single injection of bacterial lipopolysaccharide (LPS) (2). The serum concentration of murine SAA increases from its normal level of about 1 μg/ml to 1000 μg/ml 24 hours after LPS administration. During the acute phase response, SAA is found associated with high density lipoprotein (HDL), comprising up to 20% of its protein (3).

Mammalian liver is the site of synthesis for many serum proteins (4), and the rapid regulation of many acute phase proteins occurs by modulation of their mRNA levels (5, 6). Murine liver SAA mRNA increases at least 500-fold 12 hours after LPS injection while albumin mRNA decreases 2-fold by 24 hours (5). Baumann et al. have demonstrated an increase in the mRNA levels for other acute phase proteins in primary hepatocytes (6).

Two different molecular species of murine SAA protein, called isotypes SAA1 and SAA2, have been identified in several inbred strains including BALB/c (7). The amino terminal sequence was determined for the BALB/c isotypes and they are identical at 24 of the 26 residues that could be compared. SAA1 and SAA2 differ in their physiology since amyloid A fibrils contain fragments of

© IRL Press Limited, Oxford, England.

the SAA2 isotype (7). This suggests differential proteolytic processing to form AA fibrils (8). Recently, the cDNA sequences for these two isotypes were reported (9). SAA1 and SAA2 cDNAs encode proteins of 103 amino acids with 91% protein sequence homology. The amino acid and cDNA sequence results indicate there are at least two non-allelic murine SAA genes. The murine SAA genes have been shown to be closely linked on chromosome 7 (33). A comparison of SAA and AA proteins from a variety of species shows a pattern of amino acid substitutions (9, 10). In particular, there is a highly conserved region from residues 32 to 54 that is present in all mammals studied as well as in duck.

We have isolated a cDNA clone that encodes a third murine SAA isotype, SAA3, which was previously unreported. Although the liver mRNA used to produce the cDNA clone was from the random outbred Swiss mouse, the complete gene has been isolated from two BALB/c genomic phage libraries. We present the cDNA sequence, the genomic DNA sequence, and the evidence to support the exon-intron structure of the complete SAA3 gene.

MATERIALS AND METHODS

Animals, Enzymes, and Chemicals -- Male Swiss and BALB/c mice were obtained from Jackson Labs (20-30 g). LPS (Sigma L-2630) was *E. coli* type 0111:B4, phenol extracted, and was dissolved in 4% rat serum diluted in sterile water. Intraperitoneal injections of 100 μg per Swiss mouse or 50 μg per BALB/c mouse were made. Livers from LPS-treated mice were removed 9 to 24 hours post injection, as indicated, and were frozen in liquid nitrogen. Restriction enzymes, T4 polynucleotide kinase, and T4 DNA ligase were purchased from New England Biolabs or B.R.L., and used as recommended by the supplier. Restriction sites were mapped by single and multiple digests, and, in some cases by partial digestion of [^{32}P]end-labelled DNA fragments (11). S1 nuclease, *E. coli* DNA polymerase I and its Klenow fragment, and *M. luteus* DNA polymerase I were purchased from Boehringer-Mannheim. Mung bean nuclease was from P-L Biochemicals. Bacterial alkaline phosphatase was bought from Worthington (BAPF), heat treated (11) and used prior to kinase labelling according to Weiss *et al*. (12). AMV reverse transcriptase was supplied by J. Beard (Division of Cancer Cause and Prevention, N.C.I.), and R. Ratcliff (Los Alamos) supplied the purified calf thymus terminal deoxynucleotidyltransferase. Oligo dT_{12-18} and oligo dT cellulose were obtained from Collaborative Research. Radiolabelled nucleotides were from N.E.N. or Amersham.

Nitrocellulose Filter Hybridization -- Two lambda Charon 4A mouse

libraries were screened by plaque hybridization (13). The first library was of partial *Eco*RI digested BALB/c 3T3 fibroblast cell line DNA prepared by K. Peden and P. Mounts. About 600,000 phage were screened using [^{32}P]nick-translated pRS48 DNA (5) as a probe. One phage was isolated, lambda SAA1, which contained most of the SAA3 gene. The second library screened was prepared by M. Davis, R. Joho, I. Weissman, and L. Hood using partial *Alu*I and *Hae*III digestion of BALB/c mouse sperm DNA. Approximately 800,000 phage were screened using [^{32}P]nick-translated pRS48 and [^{32}P]extended primer probes (see below). Four independent phage which hybridized to both probes were isolated after two additional rounds of low density plaque screenings. Three of these were further characterized. Southern nitrocellulose filter analyses (14) of recombinant phage DNA, subcloned phage DNA, and mouse genomic DNA were prepared, hybridized, and washed as described (15).

Synthesis of Extended Primer Probe -- A gel purified DNA restriction fragment from within the cDNA sequence of pRS48 (5) was used as a primer for reverse transcriptase. Primers were first 5' end-labelled with [γ-^{32}P]ATP using T4 polynucleotide kinase (16), and then hybridized to poly A+ liver RNA from LPS-treated mice by the method of Rosbash *et al*. (17). The RNA/DNA hybrids were diluted 10-fold into binding buffer and the nucleic acid was rechromatographed on oligo dT cellulose. Eluted hybrids were ethanol precipitated several times to remove SDS and redissolved in reverse transcriptase reaction buffer (18) containing a final concentration of approximately 1 µM (about 600 Ci/mmole) [α-^{32}P]dCTP. The reaction was carried out at 42°C for 90 minutes with 6 units AMV reverse transcriptase per µg RNA. To insure complete cDNA synthesis, 1mM unlabelled dCTP (final concentration) was then added and the reaction continued at 42°C for 15 minutes. The reaction was stopped by adding EDTA and ammonium acetate, phenol:chloroform (1:1) extracted, and ethanol precipitated. The RNA was degraded by treating at 60°C in 0.1N NaOH for 60 min.

DNA Sequence Analysis -- DNA fragments were 5' or 3' end-labelled and sequenced by the Maxam and Gilbert method (16). Pst I sites were labelled with [α-^{32}P] cordycepin triphosphate using terminal deoxynucleotidyltransferase as described (19). Some regions were also sequenced by the Sanger method (20). All of the cDNA sequence was determined on both strands. Most of the SAA3 exons and 5' upstream region were sequenced on both strands and/or by separate methods. Thin (0.4 mm) denaturing urea-polyacrylamide gels (16) were used for electrophoresis of labelled

sequencing products and nuclease resistant fragments. All autoradiographs were made with preflashed Kodak XAR-5 film at -70°C. For some exposures, an Ilford ParSpeed screen was also used.

Nuclease Mapping of Exon Boundaries -- DNA fragments used as hybridization probes for nuclease mapping were either 5' end-labelled using T4 polynucleotide kinase or 3' end-labelled using *M. luteus* DNA polymerase I. In order to facilitate hybridization of DNA probes to short regions of RNA, single stranded DNA probes were prepared (21). The DNA probes were then annealed to RNA (17), and the RNA/DNA hybrids were adjusted to S1 nuclease reaction conditions (22) or mung bean nuclease conditions (as recommended by P-L Biochemicals). Nuclease concentrations and times of incubation were tested to ensure a large excess of enzymatic activity. The nuclease resistant fragments were analyzed on urea-polyacrylamide gels and autoradiographed. pBR322 *Hpa*II DNA fragments for standards were either 5' end-labelled with [γ-32P]ATP using T4 kinase, or 3' end-labelled with [α-^{32}P]dCTP using DNA polymerase.

RESULTS AND DISCUSSION

A collection of cDNA clones was made from Swiss mouse liver poly A+ RNA 24 hours after LPS administration (5). One of the 142 cDNA clones screened (pRS48) showed specific hybridization to [^{32}P]-cDNA from induced liver RNA and, by hybrid selected translation (23), hybridized to pre-SAA mRNA. This cDNA clone was used, in conjunction with extended primer probe, to screen two phage genomic libraries. Lambda SAA1, a genomic DNA clone containing part of the SAA3 gene, was isolated from a partial *Eco*RI library prepared from BALB/c 3T3 fibroblast DNA, using pRS48 as the probe. A partial *Hae*III-*Alu*I library from BALB/c mouse sperm DNA was also screened using both pRS48 and extended primer probes. Three independent phage were analyzed and they were found to contain three different SAA genes, for isotypes SAA1, SAA2, and SAA3, as well as one pseudogene (Lowell, Potter, Stearman, and Morrow, manuscript submitted). The phage containing the complete SAA3 gene was called Lambda SAA 25.

Detailed restriction maps were generated for pRS48, Lambda SAA 1, and Lambda SAA 25. Figure 1 shows these maps with the overlap of the two phage inserts. The gene is composed of four exons of lengths 54, 95, 139, and 267 nucleotides (nt) in the 5' to 3' direction. They are separated by introns of about 550, 2300, and 250 nt. The asterisks denote the approximate location of members of the B1 family of interspersed repeats (24) that were identified by

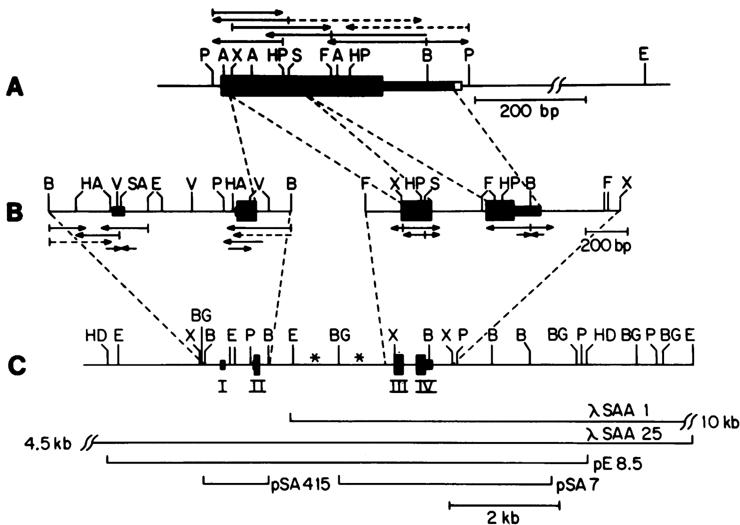

FIGURE 1. Structure and sequencing strategy of the murine SAA3 cDNA and genomic clones. For all the figures, the 5' to 3' transcription orientation is from left to right. In addition, the wide and narrow shaded boxes represent respectively the protein coding regions, and the 5' and 3' untranslated regions of the exons. 1A: The restriction map of the cDNA clone pRS48 is shown with its sequencing strategy. The unfilled box at the cDNA's 3' end shows the location of the poly dA sequence preserved from the 3' end of the SAA3 mRNA. The thin line represents the pBR322 plasmid vector DNA as well as homopolymer deoxynucleotide sequences which resulted from the cDNA cloning. The DNA sequencing is represented by arrows, with the arrowhead indicating the 3' end of the determined sequence. Solid arrows indicate sequences determined by the Maxam and Gilbert technique (16) while dashed arrows show sequencing done by the Sanger method (20). 1B: The exons of the SAA3 gene are shown with their sequencing strategy. Flanking mouse genomic DNA is represented by the thin line. 1C: The overall restriction map and structure of the SAA3 gene is shown. Lambda SAA1 and lambda SAA25 are independent genomic clones containing the SAA3 gene. The extent of mouse genomic DNA contained in the lambda Charon 4A phage and their plasmid subclones is indicated by the lower five lines. Asterisks indicate restriction fragments which contain interspersed B1 repeated elements. Restriction sites are as follows: A, AluI; B, BamHI; BG, BglII; E, EcoRI; F, HinfI; HA HaeIII; HD, HindIII; HP, HpaII; P, PstI; S, SmaI; SA, SacI; V, AvaII; X, XbaI.

Southern analysis (data not shown).

The DNA sequence of the cDNA clone is shown in Fig. 2. Also indicated is the amino acid sequence translated from its unique long open reading frame and the location of the splice sites. The homopolymer deoxyadenosine sequence at the 3' end of the cDNA indicates that pRS48 contains the 3' sequence of its mRNA. Also present is the conserved AATAAA sequence which has been implicated

```
        10          Splice   15                      20                      25                      30
A GAA   GCT GGT CAA G/GG TCT AGA GAC ATG TGG CGA GCC TAC TCT GAC ATG AAG AAA GCT AAC TGG AAA AAC TCA GAC
  Glu   Ala Gly Gln Gly  Ser Arg Asp Met Trp Arg Ala Tyr Ser Asp Met Lys Lys Ala Asn Trp Lys Asn Ser Asp

        35                   40                      45                      50                      55
  AAA   TAC TTC CAT GCT  CGG GGG AAC TAT GAT GCT GCC CGG AGG GGT CCC GGG GGA GCC TGG GCT GCT AAA GTC ATC
  Lys   Tyr Phe His Ala  Arg Gly Asn Tyr Asp Ala Ala Arg Arg Gly Pro Gly Gly Ala Trp Ala Ala Lys Val Ile

  Splice                     65                      70                      75                      80
  AG/C  GAT GCC AGA GAG  GCT GTT CAG AAG TTC ACG GGA CAT GGA GCA GAG GAC TCA AGA GCT GAC CAG TTT GCC AAT
  Ser   Asp Ala Arg Glu  Ala Val Gln Lys Phe Thr Gly His Gly Ala Glu Asp Ser Arg Ala Asp Gln Phe Ala Asn

        85                   90                      95                      100                 104
  GAG   TGG GGC CGG AGT  GGC AAA GAC CCC AAC CAC TTC CGA CCT GCT GGC CTG CCT AAA AGA TAC TGA GTTTTCTCTTC
  Glu   Trp Gly Arg Ser  Gly Lys Asp Pro Asn His Phe Arg Pro Ala Gly Leu Pro Lys Arg Tyr STOP

CTGTTGTTCCCAGTCATGCTGCCCCCCGAGAAGAGGAGCAACTACTGGGTTGAGATATTTTCTAAAATCTGGATCCCTAAACATCCCAATGTGCTGAATAAA

TACTTGTGAAATGCAAAAAAAAAAAAAA
```

FIGURE 2. Nucleotide sequence of the SAA3 cDNA clone pRS48. The coding strand nucleotide sequence and predicted amino acid sequence of pRS48 are displayed in the 5' to 3' direction. The division of the cDNA sequence into exons is indicated as splice sites above the nucleotide sequence. The conserved AATAAA sequence near the 3' end of the cDNA is underlined and the poly dA sequence determined is shown.

as a signal for polyadenylation (25). The numbering of the amino acids is based on a comparison with other known amyloid protein sequences. Comparison of the nt coding sequence of pRS48 with the SAA1 and SAA2 cDNAs shows 57 differences out of 289 nt (9). This results in 24 amino acid substitutions (out of 95 residues), 18 of which are not of a chemically conservative nature. The highly conserved region between residues 32 to 54 (9, 10) is present in the SAA3 protein sequence. Within this region, residues 45 and 51 of the SAA3 protein are not identical to the murine SAA1 and SAA2 isotypes. The amino acids which are substituted at these two positions occur either in other mammals (residue 51) or in duck (residues 45 and 51). Even though SAA3 protein has only 70% overall homology to murine SAA1 and SAA2, it is a member of this gene family based on the conservation of this region. At least 200 million years of evolution, since the divergence of birds, reptiles, and mammals, is represented by this interspecies comparison.

A primer extension experiment (18) was carried out to estimate the length of mRNA sequence not contained in pRS48. The 51 base pair (bp) AluI fragment at the 5' end of the cDNA (Fig. 1) provided an SAA3 mRNA specific primer. The AMV reverse transcriptase extension products were analyzed on the urea-polyacrylamide gel shown in Fig. 3A. The major product is a doublet of about 192 nt in length. Subtraction of 56 nt (51 nt for the AluI primer plus the 5 nt of cDNA in pRS48 not included in this fragment) yields 136 nt of the 5' end of SAA3 mRNA not present in pRS48.

Southern hybridization analysis of Lambda SAA 1 and SAA 25 indicated

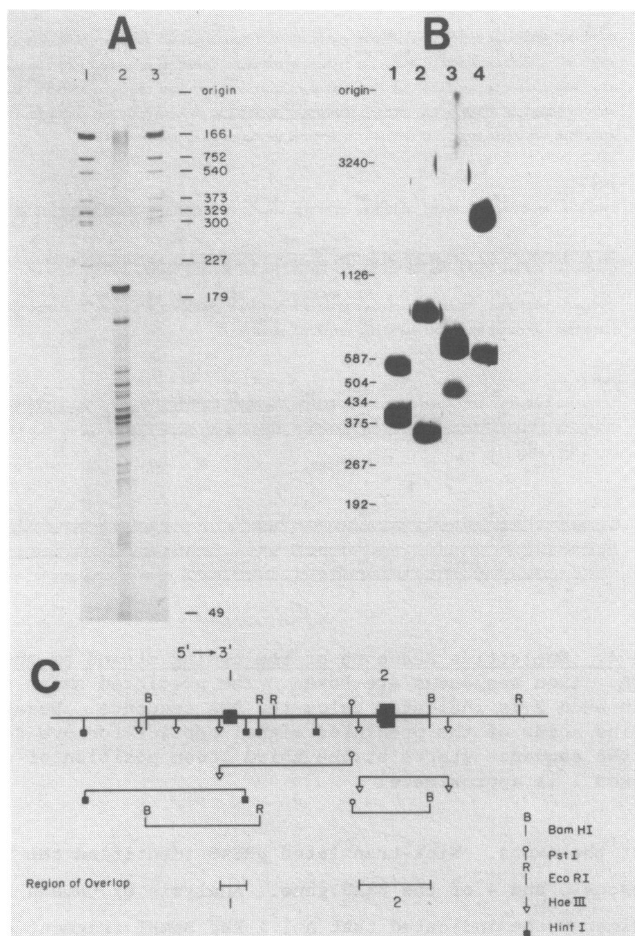

FIGURE 3. Southern analysis to locate exons 1 and 2. 3A: The products of the extended primer reaction (lane 2) were analyzed on a polyacrylamide-urea gel and autoradiographed. Lanes 1 and 3 show [^{32}P]SV40 virion DNA digested with HaeIII. 3B: Southern analysis of pSA415 DNA. The [^{32}P]extended primer probe was prepared by using the 355 bp XbaI-BamHI fragment from pRS48. A single 2% agarose gel was used to separate the pSA415 DNA fragments. The restriction enzyme digests were: lane 1, HaeIII and PstI; lane 2, PstI and BamHI; lane 3, EcoRI and BamHI; lane 4, HinfI. Molecular weight standards are from a variety of pBR322 DNA fragments which were visualized by ethidium bromide staining prior to transfer. 3C: Schematic interpretation of the pSA415 Southern analysis. The detailed restriction map of pSA415 was determined by standard multiple enzyme digestions, partial restriction digestions using end-labelled DNA fragments (11), and from the determined sequence. The restriction fragments which hybridized to [^{32}P]extended primer probe are shown relative to the restriction map. The positions and lengths of the two overlap regions, corresponding to exons 1 and 2, are indicated.

```
EXON #1
 BamhI          20                    40                    60                    80                   100
GGATCCCATG ATTTATCACA CATTCATATT TTTCTTTCTA ATATTCCTCC CCCCCCCCAA GAAACTCGAT ATAAAGCATT TGTCTTCATT ATTTCCAACC
               120          HaeIII 140                   160                   180                   200
AGGATGGCGA AGACTTCACA AAGTCCTAAC TGGCCACAAT TTCTGGAAAT GCCTAGATGG CAGCAATCTG GGGAAAGAAG ATGTACAAGA ACGAAAAGTT
               220                   240                   260                   280                HaeIII
ATCTTCTGAA AGAGAAATTA TGGGTAAGTG GCATTATGCA AGATCAAACA GGGATTGCTC CATTAGTGGA GTGACAGTAT ATATAGATAT CCTCCGCCAC
               320                  AvaII          SacI  360                   380                   400
ATCTCCTC AGCAAAGGACAG AAGCCACTCA CCTGGGGTCC CAGAAGGAGC TCGCAGCACGAG GTGAGTTG CCAGAACTTT CCAGCATTTC CTCTTTGCAG
               420                   440                   460              Eco RI
ATTCTTTTCC CCGCCAGCAT TTGCTTGTTG TGTTTAATGT TCTATCTCTA TCCTGAATCC ACTAGATGGT TAGAATTC

EXON #2
 PstI           20                    40               HaeIII           -18         -15                -10
CTGCAGAAGC ATCCTCCACC CCACGAAATG GCATGCTCAT CTTTGTTTTC CTGGCCCTAT AG CAGG ATG AAG CCT TCC ATT GCC ATC ATT CTT
                                                                         MET Lys Pro Ser Ile Ala Ile Ile Leu

              -5              +1          AvaII +5               +10                                         177
TGC ATC TTG ATC CTG GGA GTT GAC AGC CAA AGA TGG GTC CAG TTC ATG AAA GAA  GCT GGT CAA G  GTAGGCTTCA ACCATGCGGC
Cys Ile Leu Ile Leu Gly Val Asp Ser Gln Arg Trp Val Gln Phe Met Lys Glu  Ala Gly Gln

               197                   217  AvaII            237                   257                   277
CTGGACGTTC TATTTCTGAG TTTTAAGAAT TTACTTCAGC AGGGACCACA TAATCAAGCC CAAACATGCT TGTAAAATCC CCATAGAGTG GCGGTGTAGC
               297                   317              BamHI
TCTTCCTGAG AACTTTGTAG CTTTTTTGCT CAGTCTTGAC CTTTTTCTCT GAGGATCC

EXON #3
               20                    40         XbaI        60                    80                   100
TACCTCAACA CAACCACCGC CATCAACTCT CTTCCCTCTCCAG GGTCTAG AGACATGTGG CGAGCCTACT CTGACATGAA GAAAGCTAAC TGGAAAAACT
               120                   140            SmaI  160                   180                   200
CAGACAAATA CTTCCATGCT CGGGGGAACT ATGATGCTGC CCGGAGGGGT CCCGGGGGAG CCTGGGCTGC TAAAGTCATCAG GTAACACC GCTCTGGGTC
          216
AACAGAGATT GACTGC

EXON #4
               20                    40                    60                    80                   100
CGATGCCAGA GAGGCTGTTC AGAAGTTCAC GGGACATGGA GCAGAGGACT CAAGAGCTGA CCAGTTTGCC AATGAGTGGG GCCCGAGTGG CAAAGACCCC
               120                   140                   160                   180                   200
AACCACTTCC GACCTGCTGG CCTGCCTAAA AGATACTGAC TTTTCTCTTC CTGTTGTTCC CAGTCATGCT CCCCCCCCAC AAGAGCAGCA ACTACTCGGT
               220  BamHI            240                   260                   280                   300
TGAGATATTT TCTAAAATCT GGATCCCTAA ACATCCCAAT GTGCTGAATA AATACTTGTC AAATGCA CTTTGTCGCTGTC ACTGTTGTTT TGACAGCACA
     310
GGGCAGTCAC
```

Figure 4. Nucleotide sequence of the coding strand of the SAA3 gene and flanking DNA. Exon sequences are boxed. The predicted amino acid sequence contained in exon 2 is indicated below the DNA sequence. Negative numbers indicate amino acids of the predicted signal peptide for pre-SAA3 protein. The pRS48 cDNA sequence starts at the third codon position of residue +8. The 5' end of exon 1 is approximate.

locations of the exons. Nick-translated pRS48 identified the region that contained exons 3 and 4 of the SAA3 gene. Analysis of Lambda SAA 25 with extended primer probe indicated that a 1.1 kbp BamHI fragment contained the 5' end of the SAA3 gene. This DNA fragment was subcloned into pBR322 to make pSA415 (Fig. 1). Southern analysis of pSA415 with extended primer probe (Fig. 3B) defined two separate hybridizing regions. Exon 1 was contained within a 110 bp HaeIII-HinfI fragment while exon 2 was within a 295 bp HaeIII-BamHI fragment.

Figure 4 summarizes the nt sequences determined for the four SAA3 gene exons. Exons 3 and 4 show 100% nt sequence homology to the SAA3 cDNA found in pRS48. As has been found with other eukaryotic genes, these exons have the conserved RNA splice sequences at their boundaries (26). The 3' end of exon 2 was also defined by homology to pRS48, where a donor RNA splice site is located.

Nuclease protection studies (27) were performed in order to determine

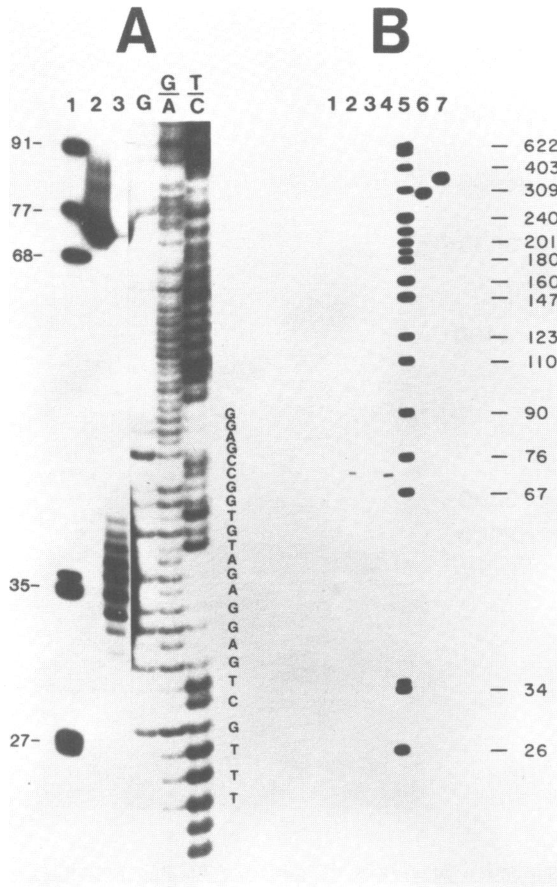

Figure 5. Nuclease mapping of the SAA3 gene. 5A: Mapping the 5' end of exon 1. The 5' end-labelled, non-coding strand of the 339 bp AvaII-BamHI fragment from pSA415 was used as the probe. The RNA/DNA hybrids were digested with S1 nuclease. The [^{32}P]probe was also used for Maxam and Gilbert sequencing reactions. Lane 1, [^{32}P]HpaII pBR322 DNA fragments; Lane 2, mapping the 5' end of exon 2 (see figure 5B) using the 290 bp [^{32}P]AvaII DNA fragment; Lane 3, mapping the 5' end of exon 1 using the 339 nt [^{32}P]AvaII-BamHI non-coding strand probe. The Maxam and Gilbert sequence ladder is shown from a longer exposure. 5B: Mapping the 5' end of exon 2. Two [^{32}P]end-labelled DNA fragments were used. The 290 bp AvaII fragment from pSA415 was 5' end-labelled with T4 kinase. The AvaII site within exon 2 used for this mapping experiment is shown in Fig. 4. The 303 bp AvaII fragment from pBR322 was labelled and served as a negative control probe. Both DNA fragments were gel purified. RNA/DNA hybrids were digested with mung bean nuclease. Lanes 1, 3, and 7 used the negative control probe while lanes 2, 4, and 6 had the exon 2 specific probe. Lanes 1 and 2, probes annealed to 200 μg of total liver RNA 9 hours after LPS injection; Lanes 3 and 4, probes annealed to 8 μg poly A+ liver RNA 24 hours after LPS injection; Lane 5, [^{32}P]HpaII pBR322 fragments; Lanes 6 and 7, untreated probes.

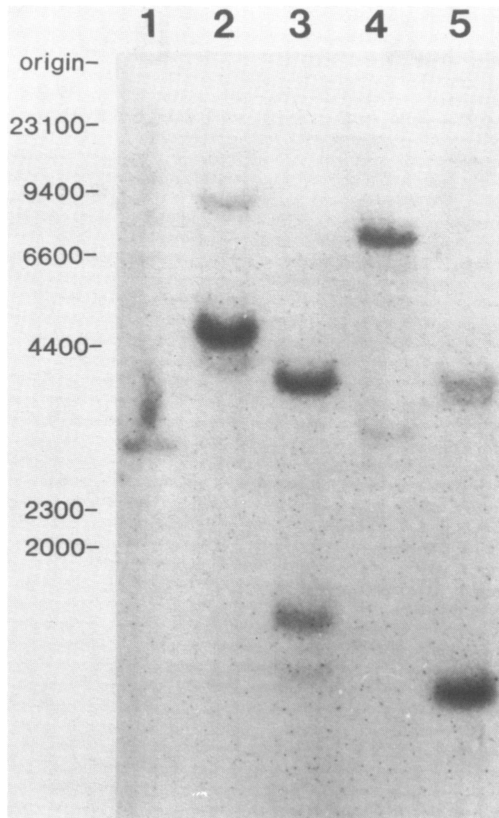

Figure 6. Southern analysis of genomic BALB/c mouse liver DNA. High molecular weight liver DNA was isolated from untreated BALB/c mice and digested with various restriction enzymes. The nitrocellulose filter was hybridized to [^{32}P]nick-translated pRS48 and washed as described. Lane 1, BamHI; Lane 2, BglII; Lane 3, PstI; Lane 4, EcoRI; Lane 5, XbaI. Molecular weight standards are from ethidium bromide staining of lambda phage DNA digested with HindIII. The restriction maps of the genomic phage (lambda SAA1 and lambda SAA 25) containing the SAA3 gene predicts the following size fragments should be the major hybridizing bands in murine DNA using pRS48 as the probe: BamHI 3.0 kb; BglII 4.4 kb; PstI 3.8 kb; EcoRI 7.4 kb; XbaI 1.1 kb.

the borders of the 5' ends of exon 1 and exon 2. Using a single stranded AvaII-BamHI fragment of the non-coding strand from exon 1, the 5' end of the SAA3 mRNA was mapped (Fig. 5A). Seven to nine protected fragments are observed each differing by one nt in length. Homogeneous mRNA has been shown to yield several fragments resistant to nuclease degradation when mapping the mRNA's 5' end (28, 29). We propose the extensive heterogeneity observed in

Fig. 5A is due to multiple 5' ends of the SAA3 mRNA. The doublet bands seen for extended primer (Fig. 3A) also supports this conclusion. Similar results were observed for the mRNAs derived from the SAA1 and SAA2 genes (Lowell, Potter, Stearman, and Morrow, manuscript submitted). Location of the 5' end of the SAA3 exon 1 has also taken into account the slower electrophoretic mobility of nuclease protected fragments compared to chemical degradation fragments (28). Located 31 bp upstream from the 5' end of exon 1 is the sequence TATATATAGATAT (Fig. 4) which has sequence and positional homology to the TATA-elements in other eukaryotic genes. The 308 bp of upstream DNA sequence contained in pSA415 is sufficient for accurate in vitro transcription in the HeLa cell extract system (30, data not shown). From the lengths of the transcripts produced in vitro, the location of the 5' end of exon 1 was also confirmed.

The 5' end of exon 2 was mapped using the 290 bp AvaII fragment from pSA415. Figure 5B shows a single nuclease resistant fragment 70 nt long. This maps the 5' end of exon 2 at a conserved acceptor RNA splice site (Fig. 4). The 3' end of exon 1 was not mapped directly, but its location was inferred from a donor RNA splice site. Having defined the exon boundaries, 138 nt are missing from pRS48 cDNA, which is within experimental error of the 136 nt calculated from the extended primer (Fig. 3A).

The SAA3 gene and cDNA encode a third SAA isotype which is 70% homologous to the SAA1 and SAA2 isotypes. Although the gene was isolated from a BALB/c mouse sperm library, it is important to demonstrate the existence of an unrearranged copy of the SAA3 gene in mouse liver DNA. A genomic DNA Southern analysis is shown in Fig. 6 using nick-translated pRS48 as the probe. The length of the major band in each lane corresponds exactly to the length predicted from the restriction map shown in Fig. 1. The minor bands can be accounted for by the maps of the SAA1 and SAA2 genes, and the SAA pseudogene (Lowell, Potter, Stearman, and Morrow, manuscript submitted).

The exons roughly divide the SAA3 gene into domains (31): exon 1 is 5' untranslated sequence, exon 2 contains the signal peptide and the first 12 residues of the mature protein, exon 3 encompasses the highly conserved region (residues 32 to 54), and exon 4 has the proteolytic cleavage site for amyloid A formation. It should also be noted that a high degree of amino acid sequence heterogeneity is found within the first fifteen amino terminal residues. These residues are mainly within exon 2. The significance of sequence heterogeneity at the amino terminus of SAA proteins is not understood. Though SAA3 protein is only 70% homologous to the other two

murine isotypes, the conservation of residues 32 to 54, as well as the conserved gene structure (Lowell, Potter, Stearman, and Morrow, manuscript submitted), clearly puts this gene into the SAA family. In addition, murine SAA3 shows slightly higher homology (76%) to human SAA1 (32) than to the other murine SAA proteins even though these species diverged more than 80 million years ago at the time of the mammalian radiation. The interspecies protein homology of known SAA and AA proteins (about 70% including duck, 10) is higher than the homology of chicken α-globin to human or mouse α-globin (about 50%).

The SAA3 protein was not observed by Hoffman et al. (7) in their analysis of SAA polypeptides. SAA3 protein may have different biochemical properties as it shares only 70% homology with SAA1 and SAA2 proteins. Possible differences include SAA3 protein being blocked at its amino terminus, not complexing with HDL, or having altered behavior in vivo.

The reason Yamamoto and Migita (9) did not isolate a SAA3 cDNA clone may be due to the lower level of this mRNA at 24 hours after LPS injection, the time they prepared RNA. We have followed the timecourse of SAA1, SAA2, and SAA3 mRNAs in BALB/c mouse liver using gene-specific probes (Lowell, Stearman, and Morrow, manuscript submitted). These results indicated the total SAA mRNA reached a maximum level 12 hours after LPS administration, with each gene contributing about one-third of the SAA mRNA synthesized. However, after this time the SAA3 mRNA was degraded, while SAA1 and SAA2 mRNAs maintained their maximum level. At 24 hours, SAA3 mRNA comprised less than 20% of the total SAA mRNA.

Acknowledgments -- We wish to thank Drs. Thomas Kelly, Barbara Sollner-Webb, and David Potter for critical review of this work, and Holly Porter for manuscript preparation. This work was supported in part by grant 5P01CA16519-09 from the National Cancer Institute, and C.A.L. was a recipient of the Medical Scientist Training Program Grant, N.I.H. 2T32GM07309.

+ Present address: Department of Biophysics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205.

REFERENCES

1. Koj, A. (1974) in Structure and Function of Plasma Proteins, Allison, A.C. Ed., vol. 1, pp.73-125, Plenum Press, New York.
2. McAdam, K.P.W.J., and Sipe, J.D. (1976) J. Exp. Med. 144, 1121-1127.
3. Hoffman, J.S., and Benditt, E.P. (1982) J. Biol. Chem. 257, 10510-10517.

4. Kushner, I. (1982) Ann. N.Y. Acad. Sci. 389, 39-48.
5. Morrow, J.F., Stearman, R.S., Peltzman, C.G., and Potter, D.A. (1981) Proc. Natl. Acad. Sci. USA 78, 4718-4722.
6. Baumann, H., Jahreis, G.P., and Gaines, K.C. (1983) J. Cell Biol. 97, 866-876.
7. Hoffman, J.S., Ericsson, L.H., Eriksen, N., Walsh, K.A., and Benditt, E.P. (1984) J. Exp. Med. 159, 641-646.
8. Glenner, G.G. (1980) N. Engl. J. Med. 302, 1283-1292.
9. Yamamoto, K., and Migita, S. (1985) Proc. Natl. Acad. Sci. USA 82, 2915-2919.
10. Gorevic, P.D., Greenwald, M., Frangione, B., Pras, M., and Franklin, E.C. (1977) J. Immunol. 118, 1113-1118.
11. Smith, H.O., and Birnsteil, M.L. (1976) Nucleic Acids Res. 3, 2387-2398.
12. Weiss, B., Live, T.R., and Richardson, C.C. (1968) J. Biol. Chem. 243, 4530-4542.
13. Benton, W.D., and Davis, R.W. (1977) Science 196, 180-182.
14. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
15. Pearson, W.R., Mukai, T., and Morrow, J.F. (1981) J. Biol. Chem. 256, 4033-4041.
16. Maxam, A., and Gilbert, W. (1980) Meth. Enzymol. 65, 499-560.
17. Rosbash, M., Blank, D., Fahrner, K., Hereford, L., Ricciardi, R., Roberts, B., Ruby, S., and Woolford, J. (1979) Meth. Enzymol. 68, 454-469.
18. Ghosh, P.K., Reddy, V.B., Piatak, M., Lebowitz, P., and Weissman, S.M. (1980) Meth. Enzymol. 65, 580-595.
19. Tu, C.-P.D., and Cohen, S.N. (1980) Gene 10, 177-183.
20. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
21. Durnam, D.M., and Palmiter, R.D. (1983) Analytical Biochem. 131, 385-393.
22. Vogt, V.M. (1973) Eur. J. Biochem. 33, 192-200.
23. Ricciardi, R.P., Miller, J.S., and Roberts, B.E. (1979) Proc. Natl. Acad. Sci. USA 76, 4927-4931.
24. Bennett, K.L., Hill, R.E., Pietras, D.F., Woodworth-Gutai, M., Kane-Haas, C., Houston, J.M., Heath, J.K., and Hastie, N.D. (1984) Mol. Cell. Biol. 4, 1561-1571.
25. Proudfoot, N.J., and Brownlee, G.G. (1976) Nature 263, 211-214.
26. Mount, S.M. (1982) Nucleic Acids Res. 10, 459-472.
27. Berk, A.J., and Sharp, P.A. (1977) Cell 12, 721-732.
28. Sollner-Webb, B., and Reeder, R.H. (1979) Cell 18, 485-499.
29. Hovemann, B., and Galler, R. (1982) Nucleic Acids Res. 10, 2261-2274.
30. Manley, J.L., Fire, A., Cano, A., Sharp. P.A., and Gefter, M.L. (1980) Proc. Natl. Acad. Sci. USA 77, 3855-3859.
31. Gilbert, W. (1985) Science 228, 823-824.
32. Parmelee, D.C., Titani, K., Ericsson, L.H., Eriksen, N., Benditt, E.P., and Walsh, K.A. (1982) Biochemistry 11, 2934-2938.
33. Taylor, B.A., and Rowe, L. (1984) Mol. Gen. Genet. 195, 491-499.