Proc. Natl. Acad. Sci. USA
Vol. 82, pp. 2282–2286, April 1985
Biochemistry

Nucleotide sequence of cDNA encoding human α_2-macroglobulin and assignment of the chromosomal locus

(human serum proteins/proteinase inhibitors/human chromosome mapping/thiol ester bond)

CHEN-CHEN KAN*, ELLEN SOLOMON†, K. TERTIA BELT‡, AILEEN C. CHAIN*, LYNNE R. HIORNS†, AND GEORG FEY*

*Department of Immunology, Research Institute of Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, CA 92037; †Imperial Cancer Research Funds Laboratory, P.O. Box 123, Lincoln's Inn Fields, London WC2A 3PX, England; and ‡Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, England

Communicated by Frank J. Dixon, December 6, 1984

ABSTRACT **Six α_2-macroglobulin (α_2M) cDNA clones were isolated from a human liver cDNA library by using synthetic oligonucleotides as hybridization probes. One of these, pα_2M1, carries a 4.6-kilobase-pair insert, which was sequenced. The insert contains the coding sequences for the mature α_2M polypeptide (1451 amino acids) and for a 23-amino acid signal peptide at the NH_2 terminus of the precursor pro-α_2M. At the 3′ end of the insert a poly(A) addition signal A-A-T-A-A-A and part of the poly(A) tail of the messenger RNA were found. The protein sequence deduced from the nucleotide sequence agrees with the published α_2M amino acid sequence for all except three residues. The α_2M locus was assigned to human chromosome 12 by Southern blot analysis with DNA from a panel of mouse/human somatic cell hybrids, using α_2M cDNA as a hybridization probe.**

α_2-Macroglobulin (α_2M) is a serum glycoprotein and a major plasma proteinase inhibitor with a wide specificity. α_2M-related proteins are present in all vertebrate species (1–4). Human α_2M is a tetramer of four identical 185-kDa subunits, arranged as a pair of dimers each consisting of two disulfide-linked monomers (5, 6). The α_2M polypeptide has a so-called bait region and an internal thiol ester bond, which account for its properties as a proteinase inhibitor. The bait region, composed of a series of target peptide bonds for plasma proteinases (7, 8), is located close to the center of the monomeric subunit. The thiol ester bond is a covalent linkage between the γ-carboxyl function of glutamate residue 975 and the sulhydryl group of cysteine residue 972 (9–12). Cleavage of the bait region by proteinases leads to a conformational change of the α_2M molecule. This conformational change activates the thiol ester bond, leading to its hydrolysis and as a consequence to covalent binding between the proteinase and α_2M via the reactive glutamate 975. α_2M receptors are present on the surfaces of fibroblasts and macrophages. These receptor proteins preferentially bind α_2M–proteinase complexes or α_2M molecules with hydrolyzed thiol ester bonds (13, 14). This suggests that the conformational change, which accompanies the complex formation between α_2M and proteinases and the hydrolysis of the thiol ester bond, exposes regions of the α_2M molecule that are recognized by these receptors. Receptor-mediated endocytosis of proteinase–α_2M complexes by macrophages and liver cells leads to clearance of the complexes from the circulation.

Internal thiol ester bonds are also found in the complement proteins C3 and C4 (15), which are derived from precursor polypeptides of size similar to α_2M (180–200 kDa). Their thiol ester sites are found in positions comparable to those in α_2M, and the amino acid sequences of all three thiol ester sites are conserved. These observations led to the proposal that the C3, C4, and α_2M genes are derived from a common ancestral gene (16). The sequences of murine and human C3 and human C4 and partial cDNA sequences of murine C4 have recently been determined (17–21). Comparison of human α_2M with murine C3 revealed a 25% overall sequence homology (17, 18, 22), supporting the proposal of their common evolutionary origin. The human C3 gene is located on chromosome 19 (23), and the human C4 genes map to the histocompatibility complex on chromosome 6 (24).

In addition to binding to proteinases, α_2M can also adhere to small molecules, including lectins, cations (zinc), basic proteins (myelin basic protein, histone H4), and platelet-derived growth factor (PDGF) (1, 25). The binding mechanisms involved are probably unrelated to the trapping of proteinases because adherence of small molecules does not prevent the proteinase binding (1). A physiological role for the complex formation between α_2M and these small molecules has not been established. The striking similarity of the amino acid sequences of PDGF and $p28^{sis}$, the transforming protein of the simian sarcoma virus (26, 27), suggests that viral and other cellular transformation mechanisms may be mediated by growth factor-like proteins. Thus, α_2M may serve as an important role in limiting the effects of PDGF-like mitogens.

The amino acid sequence of α_2M has recently been reported (28). For our future studies of the structural properties of α_2M that underlie its functions and of the regulation of expression of the α_2M gene, it is necessary to prepare cDNA clones and to derive the complete cDNA sequence. Here we report the isolation and sequence analysis of human α_2M cDNA clones and the assignment of the human α_2M locus to chromosome 12.

MATERIALS AND METHODS

Materials. Restriction enzymes were from Boehringer Mannheim, DNA polymerase I (Klenow fragment) was from Bethesda Research Laboratories, and the 17-nucleotide universal sequencing primer was from Collaborative Research. T4 DNA polymerase and dideoxy- and deoxynucleotide triphosphates were from P-L Biochemicals. The radionucleotides [α-^{32}P]dCTP (>400 Ci/mmol; 1 Ci = 37 GBq) and deoxyadenosine 5′-(α-[^{35}S]thio)triphosphate (>400 Ci/mmol) were from Amersham. T4 DNA ligase and the *Escherichia coli* strain JM101-TG1 were gifts from D. Bentley and T. Gibson. Oligonucleotide probes were synthesized by Creative Biomolecules (South San Francisco, CA).

Synthetic Oligonucleotide Probes and Isolation of cDNA Clones. The construction of the human liver cDNA library enriched for mRNA species over 2000 nucleotides in length

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: C3 and C4, third and fourth components of complement; α_2M, α_2-macroglobulin; kb, kilobase pair(s).

with the plasmid vector pAT 153 Pvu II-8 has been previously described (19). Two methionine residues at positions 1378 and 1385 of the human α_2M sequence (28) and their surrounding sequences were chosen as probe sites because they were encoded by nucleotide triplets with minimal degeneracy. These amino acid sequences are Asn-$\overset{1378}{\text{Met}}$-Ala-Ile-Val-Asp and Asp-Val-Lys-$\overset{1385}{\text{Met}}$-Val. The corresponding oligonucleotide mixtures were

$$5'\text{-T-C-N-A-C-}\begin{matrix}\text{A}\\\text{T}\end{matrix}\text{-G-A-T-N-G-C-C-A-T-}\begin{matrix}\text{A}\\\text{G}\end{matrix}\text{-T-T-3}'$$

(N represents all four nucleotides) and

$$5'\text{-A-C-C-A-T-}\begin{matrix}\text{T}\\\text{C}\end{matrix}\text{-T-T-N-A-C-}\begin{matrix}\text{A}\\\text{G}\end{matrix}\text{-T-C-3}'.$$

This cDNA library consisted of about 15,000 different recombinant clones; 30,000 colonies were plated on nitrocellulose discs. A set of replica filters were prepared from the master filters (29). Colonies on replica filters were grown, then transferred to chloramphenicol plates to amplify the plasmid. The colonies were processed (30, 31) and hybridized with kinase ^{32}P-labeled (32) 14-mer and 17-mer oligonucleotide mixtures. Positive colonies were rescreened until genetic homogeneity was reached. The length of cDNA inserts was determined by restriction analysis with the enzymes *Hin*dIII and *Sal* I.

DNA Sequence Analysis. Plasmid DNAs carrying the longest cDNA inserts were digested with *Hin*dIII and *Sal* I. The fragments were separated by electrophoresis in low-melting agarose gels and recovered from the agarose by extraction with phenol. Shotgun cloning of random 200- to 400-base-pair subfragments in the *Sma* I site of the phage M13 mp8 vector and preparation of single-stranded DNA from phage particles were performed as described (33). DNA sequences were determined with the dideoxynucleotide chain-termination method (34), using deoxyadenosine 5'-(α-[^{35}S]thio)triphosphate and buffer gradient gels (35). DNA sequence data were analyzed by using Staden's computer programs (36–38).

Chromosomal Assignment of the α_2M Locus. All human/rodent somatic cell hybrids grown for DNA preparations (Table 1), have been recharacterized as to human chromosome composition by karyotyping (39, 40), human isozyme analysis (41, 42), or both. High molecular weight DNA from human, rodent and hybrid clones was digested with the restriction enzymes *Eco*RI, *Pst* I (Boehringer Mannheim), or *Taq* I (P-L Biochemicals). Resulting fragments were electrophoretically separated in 0.7% agarose gels. DNA was denatured and transferred to nitrocellulose filters (Schleicher & Schuell) (43). Filters were hyridized with nick-translated (44) α_2M cDNA in 2.5× NaCl/Cit (1× NaCl/Cit is 150 mM sodium chloride/15 mM trisodium citrate), 10× Denhardt's solution (2% Ficoll/2% bovine serum albumin/2% polyvinylpyrrolidone), 0.1% sodium dodecyl sulfate, 6% polyethylene glycol 6000 (Fisons, Loughborough, England) or 10% dextran sulfate, and denatured salmon sperm DNA at 50 μg/ml. Hybridization was carried out at 65°C overnight. Filters were washed at 65°C successively with 2×, 1×, 0.3×, and 0.1× NaCl/Cit, three times for 20 min each.

RESULTS

Identification and Sequence Analysis of cDNA Clones. A human liver cDNA library was screened with two mixed synthetic oligodeoxynucleotide probes, a 17-mer and a 14-mer constructed on the basis of known amino acid sequences. Of 30,000 colonies screened, 34 clones hybridized with both the 17-mer and 14-mer probes.

After restriction enzyme analysis of the cDNA inserts, six recombinant plamids were selected for detailed analysis. One of them, pα_2M1, carries a 4.6-kilobase pair (kb) insert whose nucleotide sequence was determined. The nucleotide sequence from the 5' nontranslated region to the poly(A) tail of the mRNA and the derived amino acid sequence are given in Fig. 1. The complete coding sequence was determined on both DNA strands; each character was determined 3.75 times on average. A methionine encoding triplet (nucleotides 44 to 46) was identified as the translational initiation site. This methionine, together with the following 22 amino acid residues, forms the signal peptide of the α_2M protein, which has not previously been described. Thus, α_2M must be synthesized as a pro-α_2M precursor polypeptide. The α_2M signal peptide has a 10 out of 23 amino acid residue sequence homology with the mouse complement C3 signal peptide (Fig. 2 and ref. 17). This similarity further supports the idea that the C3 and α_2M genes must have a common evolutionary origin (17, 22). The amino acid sequence derived from the cDNA sequence was compared with the α_2M amino acid sequence determined on the protein level by Sottrup-Jensen and collaborators (28). Only three variant residues were found: cysteine vs. glutamic acid at position 563, glutamine vs. glutamic acid at 975, and valine vs. isoleucine at 1000. Residue 975 participates in the formation of the thiol ester bond (28). When the mature protein is sequenced, the bond is hydrolyzed and this residue manifests itself as glutamic acid. However, the cDNA sequence reveals the original residue at this position to be a glutamine. Similarly, cDNA sequence predicted glutamine residues in the thiol ester sequences of the related murine C3, C4, and sex-limited pro-

Table 1. Assignment of the α_2M locus to chromosome 12

		Human chromosomes																						
Hybrid	Human α_2M gene	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	X
1αA9.498	+	−	−	−	−	−	+	−	−	−	−	−	+	−	−	−	−	−	−	−	−	+	−	+
MOG2	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	−	+	+	+
DT2.1.2	+	−	−	−	−	−	−	−	−	−	+	+	+	+	−	+	−	+	−	−	−	+	+	+
SIR7.4ii	+	+	+	±		−	−			−		−	+		+	−	−		−	−	−	+	−	+
SCC16.5	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	+
HORL 9X	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+
MCP6	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+
MOG2.1ii	−	+	−	+	+	+	+	−	+	+	−	+	−	+	+	+	+	+	+	+	+	+	−	+
CTP41.2.A2	−	−	+		−	−	+	+		−	−	−	−		+	−			−	−	−			+

Human/rodent somatic cell hybrids tested for the presence of the human α_2M gene. References for the hybrids are 1αA9.498, SCC16.5 (gift of D. Cox, Department of Pediatrics, University of California, San Francisco); MOG2, MOG2.1ii (49); HORL9X (50); MCP6 (51); CTP41.2.A2 (52). A human chromosome is scored + if it is present in at least 10% of cells by karyotyping or the marker enzyme is clearly present by electrophoretic or immunological assay; ± indicates that the marker is present very weakly by enzyme assay. An empty space means not determined by either karyotyping or marker enzyme assay.

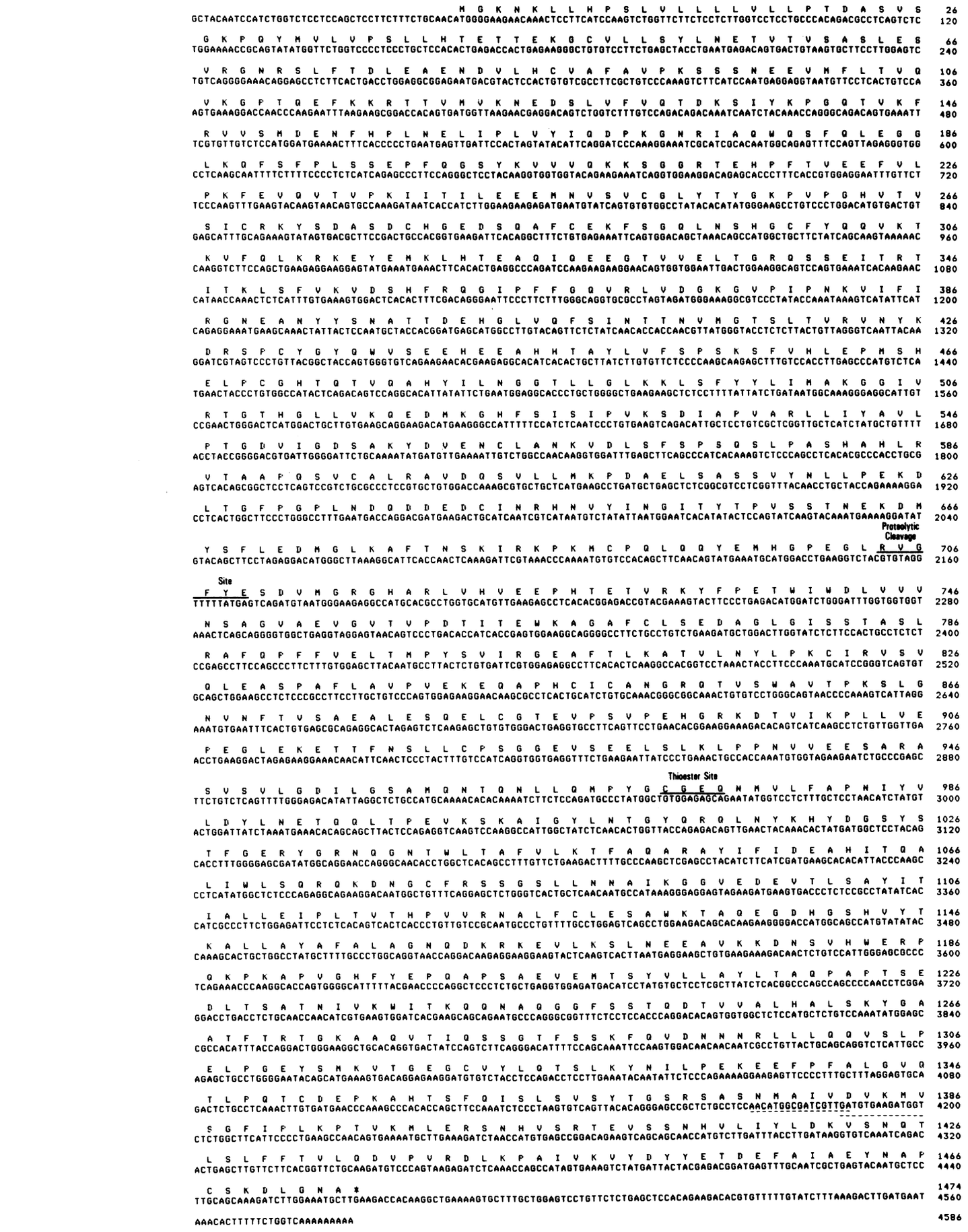

Fig. 1. (*Legend appears at the bottom of the next page.*)

Murine C3	[M G] P A S G S Q L - [L V L L L L] L A S S [P] - L [A] L G
Human A2M	[M G] K N K L L H P S [L V L L L L] - V L L [P] T D [A]

FIG. 2. Amino acid sequence homology between signal peptides of human pro-α_2M and murine prepro-C3. The sequences were aligned to give the best homology. The C3 sequence is from ref. 17.

tein and of human C3 and C4 (17–21, 46). At the 3′ end of the cDNA sequence (nucleotide positions 4558–4563) an A-A-T-A-A-A poly(A) addition signal was observed, 89 nucleotides downstream from the translational stop codon and 14 nucleotides upstream from a stretch of 9 adenines, which appears to be the beginning portion of the poly(A) tail.

Assignment of the α_2M Locus to Chromosome 12. The chromosomal locus of the α_2M gene was assigned by using somatic cell genetics. Fig. 3 shows human, mouse, and hybrid DNA digested with *Pst* I and probed with a nick-translated 1600-base-pair cDNA fragment representing the 3′ proximal portion of human α_2M mRNA. The human DNA in track 1 shows restriction fragments of 16, 5.8, 5.5, 4.3, and 2.5 kb, whereas the mouse DNA in track 10 shows major restriction fragments of 9.5, 6, and 5.5 kb. Three low-abundance mouse fragments are seen in some of the hybrids at 19, 14, and 3.5 kb. Hybrids in tracks 2, 4, and 9 are positive for the diagnostic human fragments. Track 4 contains DNA from the hybrid cell line 1αAq.498, which contains human chromosomes 6, 12, 21, and X. Hybrid cell line SCC16.5 (track 5) is negative and contains human chromosomes 21 and X. Hybrids MCP6 (track 8) and CTP41.2.A2 (track 6) are also negative and contain chromosome 6, among others. These combined results indicate that the human α_2M gene is located on chromosome 12. Table 1 presents all of the hybrids tested and their chromosome compositions. The only chromosome showing complete concordance with the α_2M gene is 12. The other chromosomes show at least two discordancies. Identical results were obtained when the same hybrids were digested with the restriction nuclease *Taq* I. We therefore conclude that the α_2M gene is located on chromosome 12.

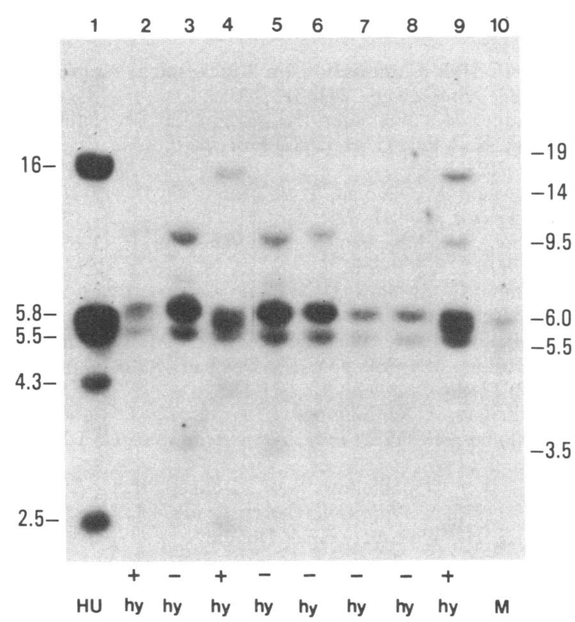

FIG. 3. Detection of human α_2M gene sequences in human/mouse somatic cell hybrids. Track 1, human (HU) line Molt-4 (47); track 2, hybrid (hy) MOG2; track 3, hybrid MOG2.1.ii; track 4, hybrid 1αA9.498; track 5, hybrid SCC16.5; track 6, hybrid CTP41.2.A2; track 7, hybrid Hor19X; track 8, hybrid MCP6; track 9, hybrid DT2.1.2; track 10, mouse (M) line IR (48). DNA was digested with *Pst* I and 25 μg of DNA was loaded per track. The + and − refer to the presence and absence of diagnostic human fragments.

DISCUSSION

α_2M cDNA clones were isolated in this study by using two synthetic oligonucleotide probes representing neighboring regions in the protein, located in its carboxyl-terminal portion. Successive screening with both probes eliminated the problem of false positives. One cDNA clone was found to contain an insert with the entire coding sequence for the α_2M protein, including its signal peptide. To date, no amino acid sequence data are available to confirm the signal peptide sequence predicted from cDNA; however, the sequence shows substantial similarity with the corresponding human and murine C3 signal peptides. The amino acid sequence derived from cDNA confirms the protein sequence reported by Sottrup-Jensen and co-workers (28) in all but three residues. The glutamine/glutamic acid difference in the thiol ester site (residue 975) has been explained in *Results*. The finding of a glutamine further supports our proposal that the formation of the thiol ester is initiated with a glutamine rather than a glutamic acid (46, 53). The difference between a cysteine and a glutamic acid residue at position 563 raises a question about the reported disulfide-bridge structure of the molecule (28). Further sequence analysis of this region of the protein would be desirable. The third difference, valine vs. isoleucine at position 1000, might be due to a DNA polymorphism, because the corresponding coding triplets can be converted from one to the other by a single base mutation resulting in a conservative replacement between two functionally similar amino acids.

The α_2M gene belongs in a gene family together with the complement C3, C4, and C5 genes. In mice, the C3 and C4 genes reside on the same chromosome, chromosome 17 (54); in humans they reside on different chromosomes: C3 on chromosome 19 (23) and C4 on chromosome 6 (24). Thus, no selective pressure is apparent to maintain these loci on the same chromosome after duplication and divergence. The assignment of the human α_2M locus to yet a different chromosome suggests that no selective pressure keeps the C3, C4, and α_2M loci together on the same chromosome.

From the data presented here we cannot definitively evaluate the number of α_2M loci on chromosome 12. The possibility exists that more than one α_2M locus or cross-hybridizing loci might be present on chromosome 12, in analogy to the duplicated C4 loci on chromosome 6. Rats and rabbits have two α-macroglobulins in their serum: α_1-macroglobulin (α_1M) and α_2M (4). Both are similar in size, are proteinase inhibitors, contain thiol ester bonds, share antigenic determinants, and show similar amino acid composition and glycosylation patterns (4). Thus, most likely they are products of duplicated genes. For humans, an equivalent of the α_1M protein has not been found. We had anticipated that an α_1M-related gene might be detected by cross-hybridization in the

FIG. 1 (*on preceding page*). cDNA and deduced amino acid sequences of pα_2M1, showing the 5′ nontranslated sequence, the signal peptide, the entire α_2M protein sequence, the poly(A) addition signal, and a portion of the poly(A) tail. The standard one-letter code for amino acids (45) is used. The proteolytic cleavage site and the thiol ester site are underlined. Also shown are regions of nucleotide sequences from which the 14-mer and 17-mer oligonucleotide probes were derived (broken underlines).

human genome, but until now we have not found its trace.

The cDNA sequences reported here can be used for isolation of the corresponding rat and human α_2M genomic DNA clones and for further studies of the control of expression of these genes during acute inflammations (55).

We thank Dr. Lars Sottrup-Jensen for making the α_2M sequence available prior to publication, Dr. Maarten de Bruijn for making C3 data available and for helpful discussions, Dr. Hans Müller-Eberhard for support, Dr. Michael C. Carroll and Prof. Rodney R. Porter for their cooperation, and Keith Dunn for preparation of the manuscript. This work was supported by Research Grant AI19651 from the National Institute of Allergy and Infectious Diseases. This is publication no. 3597IMM from the Research Institute of Scripps Clinic.

1. Barrett, A. J. (1981) *Methods Enzymol.* **80,** 737–753.
2. Harpel, P. C. & Brower, M. S. (1983) *Ann. N.Y. Acad. Sci.* **241,** 1–9.
3. Starkey, P. M. & Barrett, A. J. (1982) *Biochem. J.* **205,** 91–95.
4. Gordon, A. H. (1976) *Biochem. J.* **159,** 643–650.
5. Harpel, P. C. (1973) *J. Exp. Med.* **138,** 508–521.
6. Wang, D., Yuan, A. I. & Feinman, R. (1983) *Ann. N.Y. Acad. Sci.* **241,** 90–97.
7. Barrett, A. J. & Starkey, P. M. (1973) *Biochem. J.* **133,** 709–724.
8. Sottrup-Jensen, L., Lønblad, P. B., Stepanik, T. M., Petersen, T. E., Magnusson, S. & Jornvall, H. (1981) *FEBS Lett.* **127,** 167–173.
9. Howard, J. B. (1983) *Ann. N.Y. Acad. Sci.* **241,** 69–80.
10. Swenson, R. P. & Howard, J. B. (1980) *J. Biol. Chem.* **255,** 8087–8091.
11. Sottrup-Jensen, L., Petersen, T. E. & Magnusson, S. (1980) *FEBS Lett.* **121,** 275–279.
12. Sottrup-Jensen, L., Hansen, H. F., Mortensen, S. B., Petersen, T. E. & Magnusson, S. (1981) *FEBS Lett.* **123,** 145–148.
13. Hanover, J. A., Willingham, M. C. & Pastan, I. (1983) *Ann. N.Y. Acad. Sci.* **241,** 410–423.
14. Kaplan, J. & Keogh, E. A. (1983) *Ann. N.Y. Acad. Sci.* **241,** 442–456.
15. Tack, B. F. (1983) *Springer Semin. Immunopathol.* **6,** 259–282.
16. Sim, R. B. & Sim, E. (1981) *Biochem. J.* **193,** 129–141.
17. Fey, G. H., Lundwall, Å., Wetsel, R. A., Tack, B. F., de Bruijn, M. H. L. & Domdey, H. (1984) *Philos. Trans. R. Soc. London Ser. B* **306,** 333–344.
18. de Bruijn, M. H. L. & Fey, G. H. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 708–712.
19. Belt, K. T., Carroll, M. C. & Porter, R. R. (1984) *Cell* **36,** 907–914.
20. Ogata, R. T., Schreffler, D. C., Sepich, D. S. & Lilly, S. P. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 5061–5065.
21. Nonaka, M., Takahashi, M., Natsuume-Sakai, S., Nonaka, M., Tanaka, S., Shimizu, A. & Honjo, T. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 6822–6826.
22. Sottrup-Jensen, L., Stepanik, T. M., Kristensen, T., Lønblad, P. B., Jones, C. M., Wierzbicki, D. M., Magnusson, S., Domdey, H., Wetsel, R. A., Lundwall, Å., Tack, B. F. & Fey, G. H. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 9–13.
23. Whitehead, A. S., Solomon, E., Chambers, S., Bodmer, W., Povey, S. & Fey, G. H. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 5021–5025.
24. Carroll, M. C., Campbell, R. D., Bentley, D. R. & Porter, R. R. (1984) *Nature (London)* **307,** 237–241.
25. Huang, J. S., Huang, S. S. & Deuel, T. F. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 342–346.
26. Waterfield, M. D., Scrace, G. T., Stroobant, P., Johnsson, A., Wasteson, A., Westermark, B., Heldin, C.-H., Huang, J. S. & Deuel, T. F. (1983) *Nature (London)* **304,** 35–39.
27. Doolittle, R. F., Hunkapiller, M. W., Hood, L. E., Devare, S. G., Robbins, K. L., Aaronson, S. A. & Antoniades, H. N. (1983) *Science* **221,** 275–277.
28. Sottrup-Jensen, L., Stepanik, T. M., Wierzbicki, D. M., Jones, C. M., Lønblad, P. B., Kristensen, T., Mortensen, S. B., Petersen, T. E. & Magnusson, S. (1983) *Ann. N.Y. Acad. Sci.* **241,** 41–60.
29. Hanahan, D. & Meselson, M. (1980) *Gene* **10,** 63–67.
30. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
31. Grunstein, M. & Hogness, D. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 3961–3965.
32. Studencki, A. B. & Wallace, R. B. (1984) *DNA* **3,** 7–15.
33. Bankier, A. T. & Barrell, B. G. (1983) *Techniques in Nucleic Acid Biochemistry*, ed. Flavell, R. A. (Elsevier Scientific, Limerick, Ireland), Vol. B5, pp. 1–34.
34. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
35. Biggin, M. D., Gibson, T. J. & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3963–3965.
36. Staden, R. (1982) *Nucleic Acids Res.* **10,** 4731–4751.
37. Staden, R. (1984) *Nucleic Acids Res.* **12,** 521–538.
38. Staden, R. (1982) *Nucleic Acids Res.* **10,** 2951–2961.
39. Bobrow, M. & Cross, J. (1974) *Nature (London)* **251,** 77–79.
40. Solomon, E., Bobrow, M., Goodfellow, P. N., Bodmer, W. F., Swallow, D. M., Povey, S. & Noel, B. (1976) *Somatic Cell Genet.* **2,** 125–140.
41. Harris, H. & Hopkinson, D. A. (1976) *Handbook of Enzyme Electrophoresis in Human Genetics* (North-Holland, Amsterdam); supplements 1977 and 1978.
42. Swallow, D. M., Solomon, E. & Pajunen, L. (1977) *Cytogenet. Cell Genet.* **18,** 136–148.
43. Southern, E. (1980) *Methods Enzymol.* **68,** 152–174.
44. Bernards, R. & Flavell, R. A. (1980) *Nucleic Acids Res.* **8,** 1521–1534.
45. IUPAC-IUB Commission on Biochemical Nomenclature (1968) *J. Biol. Chem.* **243,** 3557–3559.
46. Domdey, H., Wiebauer, K., Kazmaier, M., Muller, M., Odink, K. & Fey, G. H. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 7619–7623.
47. Minowada, J., Ohnoma, T. & Moore, G. E. (1972) *J. Natl. Cancer Inst.* **49,** 891–895.
48. Nabholz, M., Miggiano, V. & Bodmer, W. F. (1969) *Nature (London)* **223,** 358–363.
49. Solomon, E., Swallow, D. M., Burgess, S. & Evans, L. (1979) *Ann. Human Genet.* **42,** 273–281.
50. Goodfellow, P. N., Banting, G., Wiles, M. V., Tunnacliffe, A., Parkar, M., Solomon, E., Dalchau, R. & Fabre, J. W. (1982) *Eur. J. Immunol.* **12,** 659–663.
51. Goodfellow, P. N., Banting, G., Trowsdale, J., Chambers, S. & Solomon, E. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 1190–1194.
52. Jones, E. A., Goodfellow, P. N., Kennett, R. H. & Bodmer, W. F. (1976) *Somatic Cell. Genet.* **2,** 483–496.
53. Fey, G. H., Wiebauer, K. & Domdey, H. (1983) *Ann. N.Y. Acad. Sci.* **421,** 307–312.
54. Fey, G. H., Domdey, H., Wiebauer, K., Whitehead, A. S. & Odink, K. (1983) *Springer Semin. Immunpathol.* **6,** 119–147.
55. Kushner, I. (1982) *Ann. N.Y. Acad. Sci.* **389,** 39–48.