J. Biochem. **120**, 1167-1175 (1996)

Cloning and Sequencing of cDNAs Encoding Plasma α-Macroglobulin and Murinoglobulin from Guinea Pig: Implications for Molecular Evolution of α-Macroglobulin Family[1]

Hiromitu Iwasaki, Yasuyuki Suzuki, and Hyogo Sinohara

Department of Biochemistry, Kinki University School of Medicine, Osaka-Sayama, Osaka 589

Received for publication, July 23, 1996

Several clones encoding plasma α-macroglobulin and murinoglobulin were isolated from guinea pig liver cDNA library and sequenced. The clones for α-macroglobulin contained overlapping sequences which together spanned a stretch of 4,546 nucleotides with one open reading frame coding for 1,476 amino acid residues. The clones for murinoglobulin contained overlapping sequences which together spanned a stretch of 4,578 nucleotides with one open reading frame coding for 1,464 amino acid residues. The phylogenetic analyses of 11 proteins of the α-macroglobulin family revealed that the mammalian tetrameric α-macroglobulins consist of two main branches: αM-1 subfamily (rat α_1- and mouse α-macroglobulins) and αM-2 subfamily (human α_2-, rat α_2-, and guinea pig α-macroglobulins). This dichotomy is in good accordance with their immunological, chemical, and physicochemical properties, and indicates that guinea pig α-macroglobulin is orthologous to human and rat α_2-macroglobulins but paralogous to rat α_1- and mouse α-macroglobulins. The divergence of the two subfamilies was a phylogenetically ancient event which occurred around the separation of metatherians and eutherians. The genes of the two subfamilies have been maintained in the rat, but either one became extinct in the mouse, guinea pig, or human. The tree also shows that guinea pig murinoglobulin forms one clade with mouse and rat murinoglobulins (α_1-inhibitor$_3$) prior to joining the αM-2 lineage, and suggests that murinoglobulin is not a primitive form of tetrameric α-macroglobulin, but rather has evolved under selective pressure which is different from that of the tetrameric paralogues.

Key words: guinea pig α-macroglobulin, molecular evolution, murinoglobulin, rodent phylogeny.

Rodents are the largest order of mammals, and are usually classified into four suborders; (i) sciuromorphs or squirrel-like rodents, (ii) myomorphs or mouse-like rodents, (iii) porcupines, and (iv) caviomorphs inclusive of the guinea pig (*1*). From the molecular evolutionary point of view, the guinea pig is unique among mammals, since a number of its proteins, such as insulin, lipoprotein lipase, and glucagon, have evolved severalfold faster than their respective orthologous proteins in other mammals (*2, 3*). The reason for this is not known at present. Based on the evolution of several genes, it was recently proposed that the caviomorphs do not belong to the order Rodentia, and that the caviomorphs should be elevated in taxonomical rank and regarded as a separate order distinct from Rodentia (*4, 5*). If this hypothesis is accepted, the average rate of molecular evolution in the guinea pig becomes closer to those in mouse and rat. Our results on the evolution of mammalian plasma α_1-antiproteinase (also called α_1-antitrypsin) support this hypothesis (*6*). However, there has been essentially unanimous agreement regarding the monophyly of rodents for more than a century, and virtually no contradictory data have so far been reported from paleontological, morphological, and ontogenetic points of view (*7*). Of special relevance to this issue is the evolution of murinoglobulin,[2] a member of the αM family. αMs are glycoproteins found in the blood and body fluid of not only vertebrates but also invertebrates (reviewed in Ref. *8*). Most αMs are tetramers assembled from 180-kDa identical subunits, but dimeric and monomeric αMs are also known. MUG, monomeric αM, was isolated in a homogeneous state from several rodent plasmas, such as mouse (*9*), rat (*10*), and hamster (*11*), and characterized chemically and physicochemically. Rat plasma protein, named α_1-inhibitor$_3$ (*12*) or α-X-protein (*13*), was thought to be the rat counterpart

[1] This work was supported in part by grants from the Japan Private School Promotion Foundation and the Environmental Science Research Institute of Kinki University. The nucleotide sequences of guinea pig α-macroglobulin and murinoglobulin reported in this paper have been submitted to the GenBank/EMBL/DDBJ Data Bank with accession numbers D84338 and D84339, respectively.

Abbreviations: αM, α-macroglobulin; α_1M, α_1-macroglobulin; α_2M, α_2-macroglobulin; MUG, murinoglobulin; NJ, neighbor-joining method; PZP, pregnancy zone protein; UPGMA, unweighted pair group method of arithmetic mean.

[2] See footnote 3 of Ref. *17* for nomenclature and identity of murinoglobulin-related proteins. Although "rat α_1-inhibitor 3" is widely used for the rat homologue of monomeric αM, "rat murinoglobulin" is used in this paper in order to facilitate comparison of different monomeric αMs.

of human inter-α-trypsin inhibitor, since they showed immunological cross-reactivity (*14, 15*). However, it was later identified as monomeric αM (*16*) that is a mixture of two isoforms, MUGs I and II (*10*). MUG has been detected in all rodents so far studied, *i.e.*, mouse, rat, hamster, and Mongolian gerbil, but not in other orders of mammal, such as rabbit, dog, sheep, bovine, goat, pig, or human (*18*). We purified MUG from guinea pig plasma, and found that its chemical and physicochemical properties are similar to those of MUG from plasma of other rodents (*19*). Thus, the distribution of MUG among mammals seems to favor the traditional view of rodent monophyly. Recently, Martignetti and Brosius showed that BC1 RNA, which is a neural-specific small cytoplasmic RNA and the product of a retropositionally generated gene, occurs in both myomorphs and caviomorphs, but not in other mammalian orders (*20*), supporting the traditional views. Since the guinea pig is widely used as an experimental animal, it is desirable to establish its phylogenetic position among mammals in order to get an overall picture of the problems studied. Furthermore, it is known that αMs show marked species differences in the acute phase response. For example, human α_2M, guinea pig MUG, and rat α_1M show little change in concentration, whereas rat α_2M increases several hundred-fold (*17, 21-23*). On the other hand, rat and mouse MUGs are typical negative acute phase proteins, their concentration decreasing to less then 1/3 of the normal levels (*17, 22-24*). We also showed that there are several characteristic differences in physicochemical properties between guinea pig αM and human α_2M (*25*). Elucidation of these species differences in chemical and biological properties will provide a new insight not only into the molecular properties of αM, but also into the molecular mechanism of acute phase response. As an initial approach to these problems, the present paper describes sequence analysis of cDNAs encoding guinea pig αM and MUG, and the implications of the results for the evolution and function of the αM family.

EXPERIMENTAL PROCEDURES

Materials—Restriction enzymes, DNA ligase, reverse transcriptase, and other DNA-modifying enzymes were purchased from Takara Shuzo. cDNA Synthesis Kit and cDNA Sequencing Kit were obtained from Pharmacia and Stratagene, respectively. Antibodies against guinea pig αM (*25*) and MUG (*19*) were raised in rabbits as described. Other materials were essentially the same as those used for the previous experiments (*19, 25, 26*).

Construction of cDNA Library and Screening—The guinea pig liver cDNA library was constructed in λgt11 and screened with antisera against αM or MUG essentially as described previously for guinea pig α_1-antiproteinase (*26*).

DNA Sequencing—Positive clones were isolated and their restriction fragments were subcloned into M13mp18 and M13mp19. The products were used for the construction of a series of deletion mutants containing varying lengths of the inserts by application of the Cyclone system of Dale *et al.* (*27*). Both strands of several clones were sequenced by the dideoxynucleotide chain termination method (*28*). The 5′-end of the MUG transcript was amplified by the method of rapid amplification of cDNA ends (RACE) according to Frohman *et al.* (*29*), except that the oligo(dT) primer was replaced by dC 15-mer. The antisense primers used for this technique were 5′-ACTTGCACTTCAAATCTGGG-3′, complementary to positions 742-723 (see Fig. 2); 5′-TGTGACTGCCTTTGGACACTTTACTT-3′, complementary to positions 764-739; and 5′-CATCCAGAAATGTGACTGCC-3′, complementary to positions 774-755. The melting points of these primers and their cognate cDNA were higher than those of oligo(dA-dT) pairs, and attempts to amplify the 5′ end of cDNA by the original oligo(dA-dT) method were unsuccessful. This was overcome by the replacement of oligo(dT) by oligo(dC).

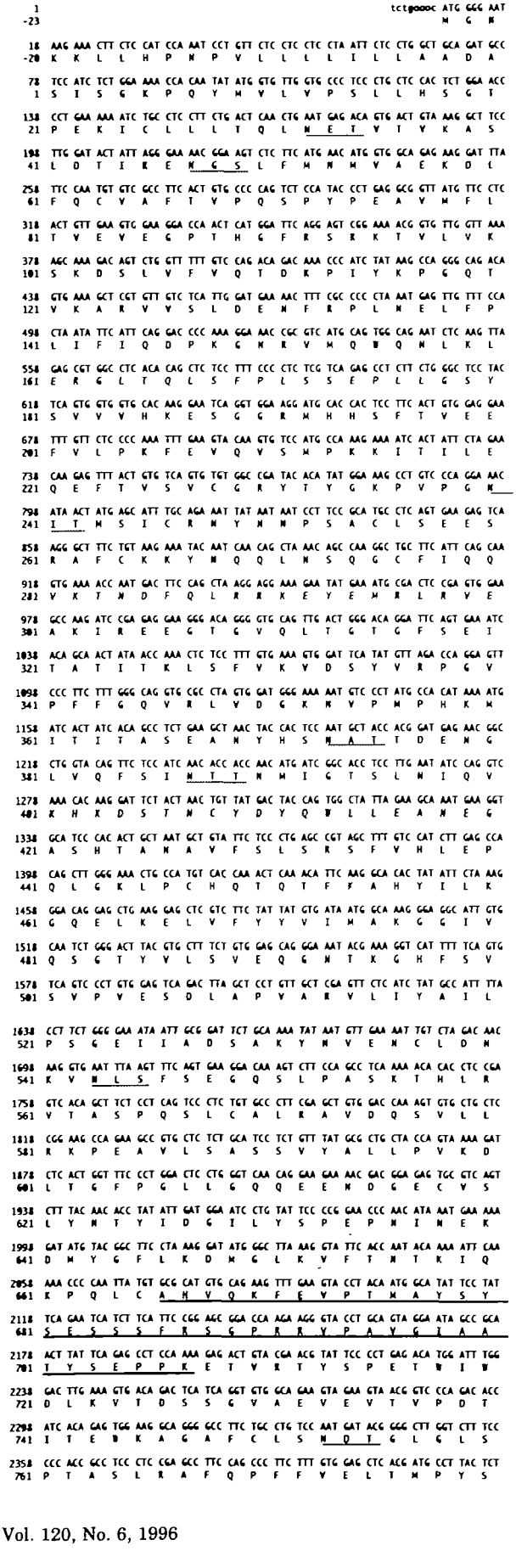
Phylogenetic Analysis—It is generally thought that synonymous nucleotide substitution, which does not affect amino acid sequences, is selectively neutral, and accumulates at a constant rate during evolution (*30*), and that the number of nucleotide substitutions per synonymous site (K_S) reflects the evolutionary relationships among the orthologues. However, all αM members contain a bait region, which varies greatly in length and shows much greater sequence diversity than the directly flanking introns of both sides (*31, 32*). This indicates that the bait region has evolved under a strong selective pressure to change, so it is impossible to estimate the evolutionary distance by the usual methods based on the neutral theory (*30*). Furthermore, αM members contain additional diversity regions which are located mainly in the N-terminal half of the molecule (*33, 34*). Accordingly, the C-terminal half of the molecule was used to calculate the number of nucleotide substitutions per synonymous site (K_S) and nonsynonymous site (K_A) between every two members of the αM family by the method of Li *et al.* (*35*). The evolutionary distances were also estimated by three other methods, *i.e.*, using a one-parameter model (*36*), a two-parameter model (*37*), and the proportion of differences (*38*), for comparison. Although the C-terminal half represents only half of the molecule, its length, about 760 amino acid residues, is longer than that of many proteins, and its evolutionary data could be reliably compared with those obtained for other genes (see "RESULTS"). Phylogenetic trees were constructed by the UPGMA (*38*) and NJ methods (*39*).

Estimation of Divergence Time—The divergence time (T) for each pair of species was calculated from the equation $T = K_S/(r_1 + r_2)$ where r_1 and r_2 are evolutionary rates (changes/site/year) for lineages leading to the two species (*38*). The evolutionary rates used for rodent and nonrodent lineages were 6.5×10^{-9} and 3.0×10^{-9}, respectively (*40*).

RESULTS

Isolation and Sequencing of cDNA Clones Coding for αM and MUG—Several positive clones of αM and MUG were isolated, and characterized by restriction enzyme analysis and partial sequencing. The αM clones overlapped each other, and spanned a stretch of 4,546 nucleotides with one open reading frame coding for 1,476 amino acid residues (Fig. 1). This molecule was concluded to consist of a 23-residue signal peptide and a 1,453-residue mature protein by comparison with other known sequences of the αM family members. On the other hand, MUG cDNA clones spanned a stretch of 4,578 nucleotides, and had one open reading frame coding for 1,464 amino acid residues (Fig. 2). This protein was considered to consist of a 23-

α-MACROGLOBULIN



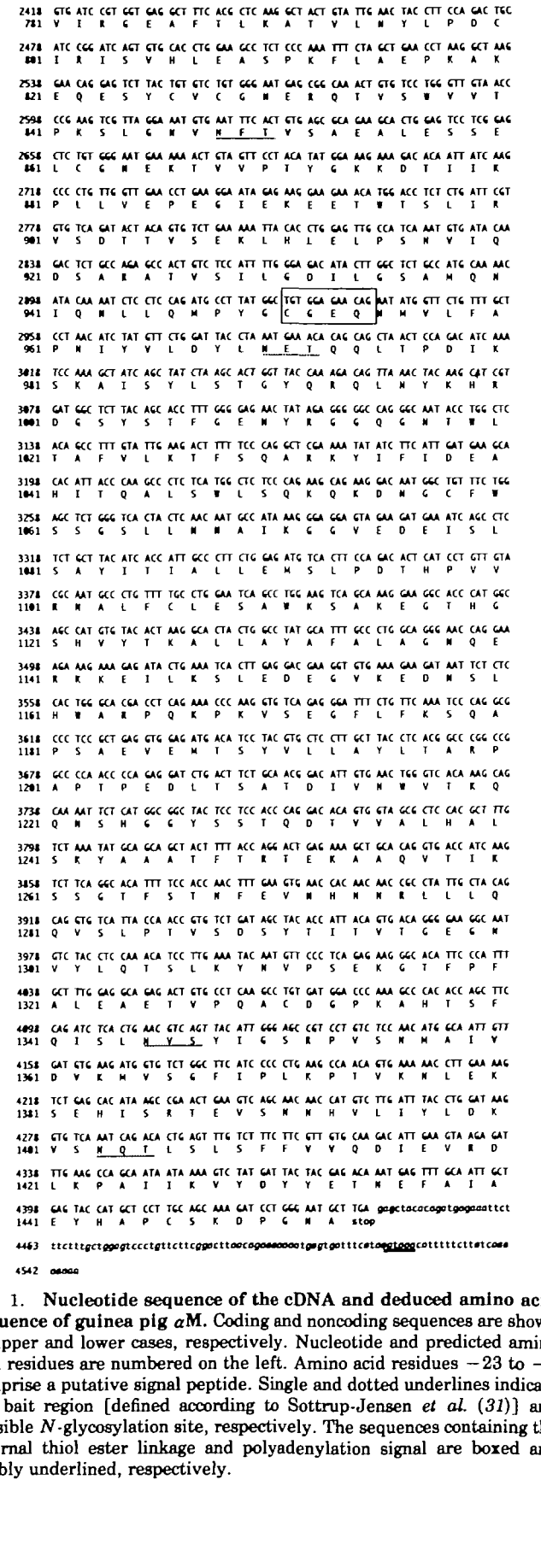
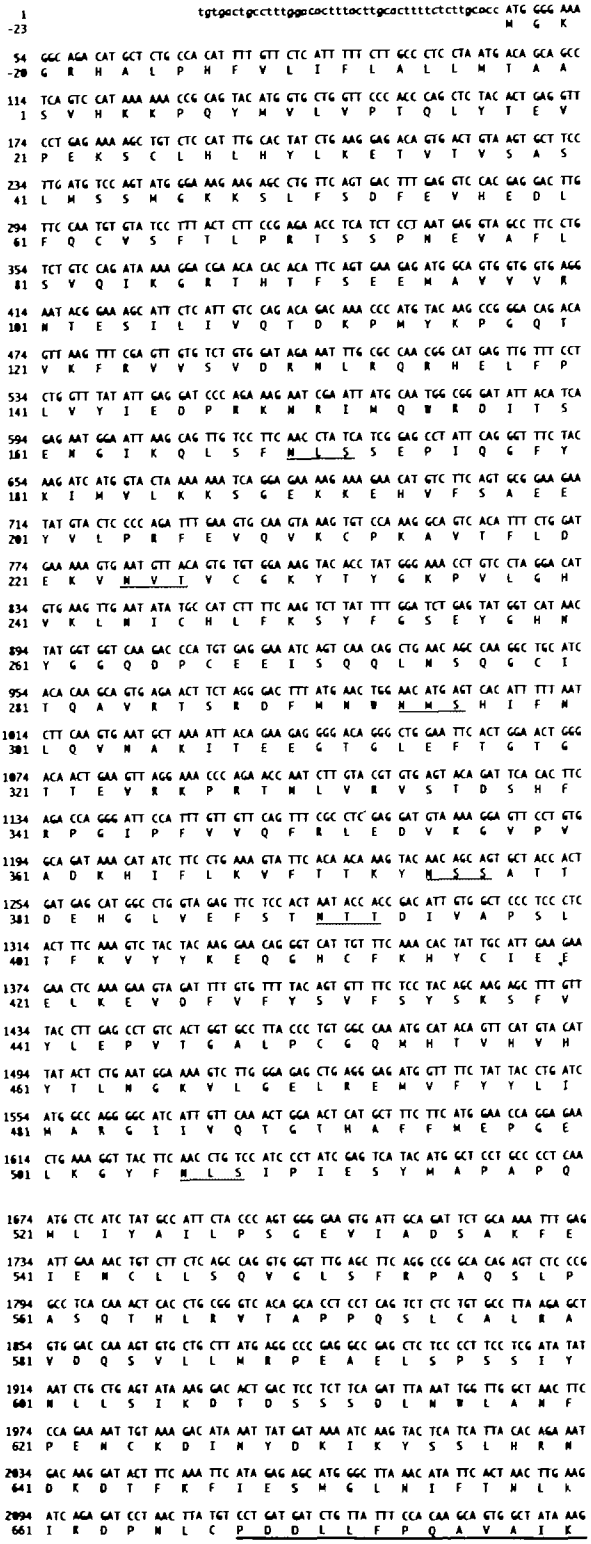


Fig. 1. **Nucleotide sequence of the cDNA and deduced amino acid sequence of guinea pig *α*M.** Coding and noncoding sequences are shown in upper and lower cases, respectively. Nucleotide and predicted amino acid residues are numbered on the left. Amino acid residues −23 to −1 comprise a putative signal peptide. Single and dotted underlines indicate the bait region [defined according to Sottrup-Jensen *et al.* (*31*)] and possible *N*-glycosylation site, respectively. The sequences containing the internal thiol ester linkage and polyadenylation signal are boxed and doubly underlined, respectively.

MURINOGLOBULIN



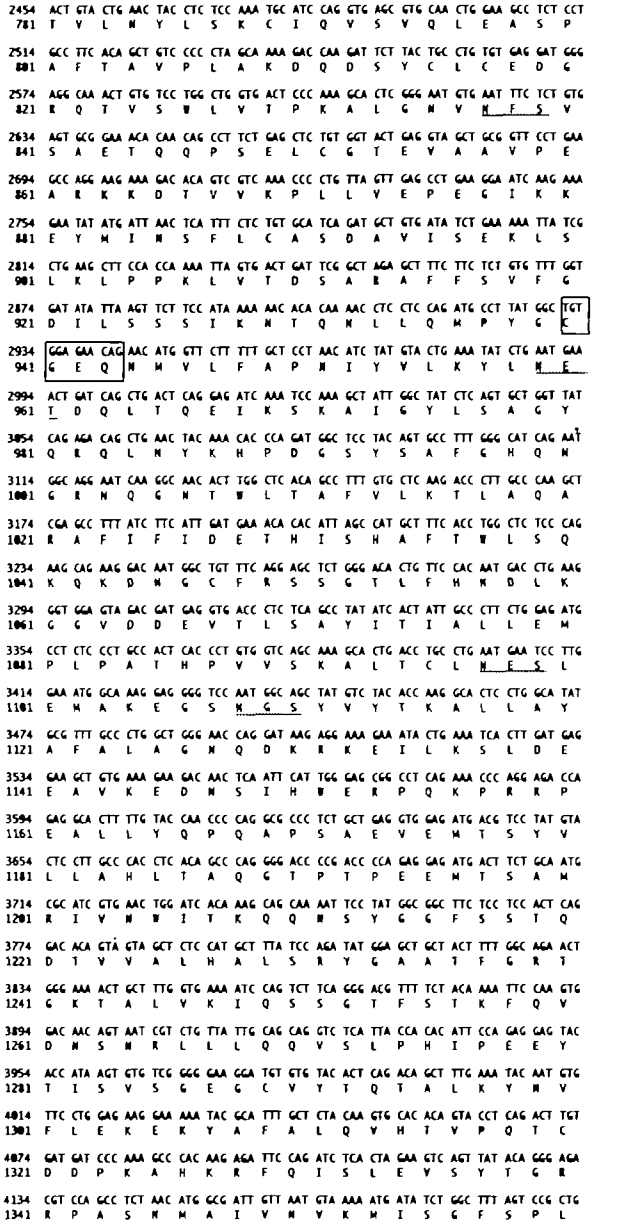


Fig. 2. **Nucleotide sequence of the cDNA and deduced amino acid sequence of guinea pig MUG.** Coding and noncoding sequences are shown in upper and lower cases, respectively. Nucleotide and predicted amino acid residues are numbered on the left. Amino acid residues -23 to -1 comprise a putative signal peptide. Single and dotted underlines indicate the bait region [defined according to Sottrup-Jensen *et al.* (*31*)] and possible *N*-glycosylation site, respectively. The sequences containing the internal thiol ester linkage and polyadenylation signal are boxed and doubly underlined, respectively.

residue signal peptide and a 1,441-residue mature protein by comparison with known sequences of other MUGs. No other clones having sequence similarity with either αM or MUG were encountered during the experiments. This suggests that the guinea pig contains a single mRNA each for αM and MUG, though more work is needed to rule out the possibility that the guinea pig contains additional active genes belonging to the αM family.

Disulfide Bridges—Guinea pig αM contains 23 cysteine residues, all of which are found in conserved positions[3] in human α_2M, rat α_2M, and human PZP. Polymeric αMs so far sequenced, however, contain two additional cysteines, Cys^{h898} and Cys^{h1298}, which form an intramonomer disulfide bridge. In guinea pig αM, these two residues are replaced by Arg and Asn, respectively. In contrast, the same disulfide bridge is conserved in guinea pig MUG. Another disulfide bridge connecting Cys^{h798} to Cys^{h826} is present in all polymeric αMs and guinea pig MUG but not in rat or mouse MUGs. Thus, guinea pig MUG bears a greater resemblance to the polymeric αMs than rat and mouse MUGs in this respect. All known mammalian polymeric αMs contain Cys^{h255} and Cys^{h408}, both of which are involved in intermonomer disulfide bridge formation. In contrast, all MUGs so far sequenced, including the guinea pig protein, lack Cys^{h255}, but contain one extra cysteine residue in the vicinity of Cys^{h408}. The remaining cysteine residues of MUGs are at the conserved positions of tetrameric α_2Ms, suggesting that the intramonomer disulfide structure is similar to that of tetrameric α_2Ms.

K_S and K_A of the C-Terminal Halves—The K_S and K_A values for each pairwise comparison among 11 members of mammalian α_2Ms are shown in Table I. The K_S values of rat α_1M *versus* mouse αM and rat MUG-1 *versus* mouse MUG-1 were 0.267 and 0.249, respectively. As will be discussed later, these two pairs of genes are orthologous. The K_S values between mouse and rat for nine other genes, tabulated by Li *et al.* (*40*), ranged from 0.115 to 0.281. This indicates that the C-terminal half of αM family members has accumulated synonymous substitutions at a rate similar to that of the other genes. The mean K_A/K_S ratio for all pairs between αM members was 0.211± 0.0053 (SE). The corresponding values for 34 different genes in mammalian species averaged 0.22±0.12 (SE), ranging from 0.030 to 0.47 (*35*). The wide range of this ratio reflects large variations in the rates at nonsynonymous sites (*35*). The ratios of the C-terminal halves of α_2Ms are close to that of growth hormone (ratio=0.217) and slightly greater than the average (ratio=0.189) of 42 genes examined (*35*). This indicates that the functional and structural constraints imposed on the C-terminal half of the

[3] In order to avoid confusion when comparing the different αM members, amino acid positions were numbered relative to the human mature α_2M, and expressed with the prefix "h" such as Cys^{h898}, "h" standing for "human."

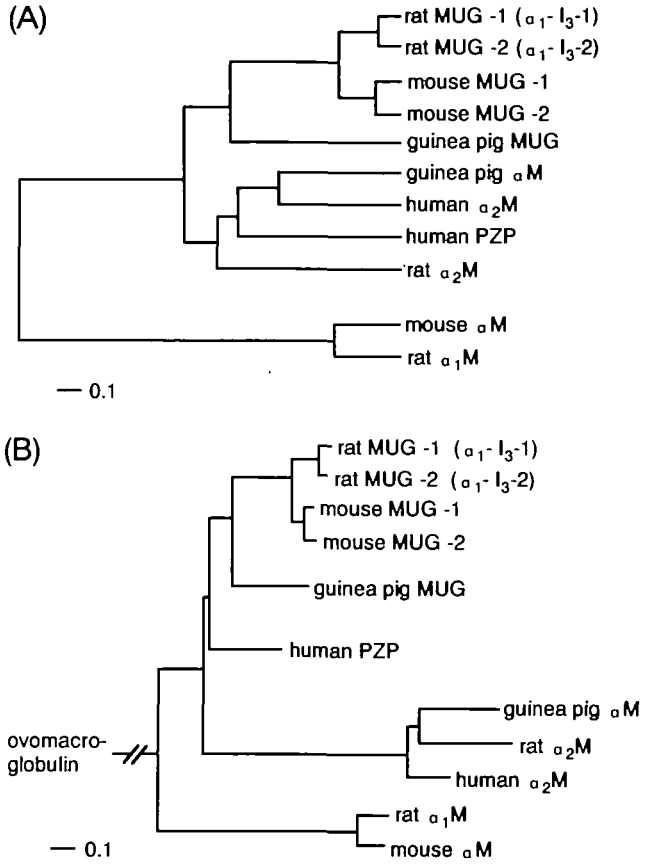


Fig. 3. **Phylogenetic tree of the mammalian αM family.** (A) UPGMA tree. (B) NJ tree. The branch lengths are drawn in proportion to the evolutionary distance and the scale indicates a genetic distance of 0.1. The sources of DNA sequences are given below: human α_2M (*52*), human PZP (*53*), rat α_1M (*54*), rat α_2M (*55*), mouse αM (*47*), rat MUG-1 (α_1I_3-1) (*56*), rat MUG-2 (α_1I_3-2) (*57*), mouse MUG-1 and -2 (*58*).

TABLE I. **Number of nucleotide substitutions per synonymous site (K_S: above diagonal) and per nonsynonymous site (K_A: below diagonal)×1,000 among αM members.** Standard error of each value is less than 10%.

	GA	HA	PZP	RA2	RA1	MA	GM	RM1	RM2	MM1	MM2
GA	–	503	819	695	1,531	1,706	729	1,053	1,129	956	983
HA	120	–	510	643	1,347	1,352	700	808	861	735	749
PZP	153	99	–	860	1,401	1,451	665	815	913	813	793
RA2	146	114	169	–	1,558	1,566	973	928	1,011	883	896
RA1	251	221	256	256	–	267	1,619	1,588	1,675	1,601	1,624
MA	249	225	254	254	68	–	1,578	1,550	1,552	1,445	1,506
GM	183	144	167	192	274	272	–	689	743	660	668
RM1	215	187	210	208	275	282	157	–	134	249	258
RM2	233	205	228	228	297	298	186	50	–	306	299
MM1	228	193	217	216	293	301	160	69	97	–	97
MM2	229	195	222	223	297	304	162	69	93	28	–

Abbreviations: GA, guinea pig αM; HA, human α_2M; PZP, human PZP; RA2, rat α_2M; RA1, α_1M; MA, mouse αM; GM, guinea pig MUG; RM1, rat MUG-1; RM2, rat MUG-2; MM1, mouse MUG-1; MM2, mouse MUG-2.

αM members have been similar to those of growth hormone.

Phylogenetic Trees—The phylogenetic trees constructed by the UPGMA and NJ methods are shown in Fig. 3, A and B, respectively. The former depends on the assumption of a constant rate of molecular evolution but gives a rooted tree (*38*). In this tree MUGs are clustered to form one clade, whereas the tetrameric αMs are separated into two clades; the subfamily αM-1 consists of rat α_1M and mouse αM, and the αM-2 subfamily consists of human and rat α_2Ms and guinea pig αM. The NJ method does not depend on the assumption of a constant evolutionary rate, but has been shown to find a correct unrooted tree efficiently (*39*). The tree shown in Fig. 3B is rooted by using chicken ovomacroglobulin (*41*) as an outgroup. Although the branch lengths connecting the molecular species are different, the NJ tree also has three major branches, αM-1, αM-2, and MUG subfamilies. The two trees, however, differ in the branching order within the αM-2 subfamily and the position of human PZP. The closest relative of guinea pig αM is human α_2M in the UPGMA tree, but it is rat α_2M in the NJ tree. PZP belongs to the αM-2 subfamily in the UPGMA tree, whereas it is closer to the MUG subfamily than the αM-2 subfamily in the NJ tree. The NJ trees based on the one-parameter and two-parameter models show the same branching pattern as the UPGMA tree although the branch lengths are different (data not shown). On the other hand, the NJ tree based on the proportion of difference method shows that the closest relative of PZP is human α_2M (data not shown). These results indicate that the mammalian αM family consists of three subfamilies, although the branching order within the αM-2 subfamily and the position of dimeric PZP differ depending upon the methods used.

DISCUSSION

Intramonomer Disulfide Bridges of Guinea Pig αM and MUG—Guinea pig αM is unique among the family members in that it lacks an intrachain disulfide bridge interconnecting Cys^{h898} and Cys^{h1298}, which holds the thiol ester-containing region in the intervening region. Thus, it is likely that the absence of this bridge is responsible for some physicochemical differences between guinea pig αM and human α_2M (*25*): (i) after treatment with methylamine, human α_2M was converted to a "fast" form which migrated faster in nondenaturing polyacrylamide gel electrophoresis, while no such conversion was observed with guinea pig αM under the same conditions; (ii) upon treatment with trypsin, tetrameric αMs underwent thiol ester cleavage, which proceeded with at least two different rates (*25, 42*); the rate constants of these reactions of guinea pig αM were severalfold greater than the respective constants of human α_2M; (iii) the amidase activity of trypsin bound to guinea pig αM was impaired by soybean trypsin inhibitor to a much greater degree than that of trypsin bound to human α_2M.

Molecular Dichotomy of Mammalian Tetrameric αMs—Overbergh *et al.* (*32*) revealed that mouse contains only one copy of αM gene, indicating that mouse αM and rat α_1M are orthologous (Fig. 3, A and B), *i.e.*, the difference in their DNA sequences reflects the time passed since the separation of the two species. The K_S value between them is 0.267 (Table I), indicating that they diverged 20×10^6 years ago. This is in agreement with the paleontological data of the rat-mouse divergence time ($15\text{-}30\times10^6$ years ago) (*40*). On the other hand, the K_S value between rat α_1M and rat α_2M, a paralogous pair, is 1.558 (Table I), which gives an estimate of divergence time of 111×10^6 years ago on the assumption that both genes evolved at the rodent rate throughout the whole period after gene duplication. The same calculation based on human α_2M and rat α_1M gave somewhat earlier date, 142×10^6 years ago, since the evolutionary rate in the nonrodent lineage was smaller than that of the rodent lineage (*40*). The actual divergence time of the two rat genes may be further antedated, since the rates in the rodent and nonrodent lineages should have been similar at the early stage of mammalian divergence, and the difference observed must have occurred mainly in more recent times. Thus, the two paralogous gene lines in the rat trace back to an event that occurred as early as in the Early Cretaceous Age prior to the separation of eutherians and metatherians (*43*). The eutherians are assumed to have diverged from each other within a short period of time, 75-80 million years ago, relative to the total length of time over which they have been evolving independently of each other. Thus it is likely that the divergence of the subfamilies αM-1 and αM-2 occurred before the eutherian radiation. Both genes have been maintained in the lineage leading to the rat, but the αM-2 gene became extinct in the mouse lineage after separation from the rat. The results obtained with the NJ tree, which does not depend on a constant evolutionary rate, also demonstrate that the divergence time of the two tetrameric αM subfamilies was earlier than the emergence of MUG and PZP (Fig. 3B).

Humans have three αM genes, one each for tetrameric and dimeric αM and one pseudogene (*44*). Although it is not clear whether this pseudogene is of the αM-1 type or not, only the αM-2 subfamily gene is active. It is also possible that the guinea pig may have lost the αM-1 type gene after separation from the rat, although further work is needed to substantiate this hypothesis. The loss of one copy within a multigene family is not uncommon in evolutionary history; for example, the η-globin locus became a pseudogene in the stem of primates and was deleted in the rodent and lagomorph lineages, while the γ-globin locus was lost in the artiodactyls (*45*).

We showed previously (*25*) that antibody raised against guinea pig αM cross-reacted partially with human and rat α_2Ms but not with rat α_1M, and that on treatment with trypsin, guinea pig αM was cleaved into 88- and 77-kDa fragments in essentially the same manner as human and rat α_2Ms. On the other hand, Schaeufele and Koo (*46*) showed that rat α_1M and mouse αM were cleaved by trypsin into fragments which were different from those of human and rat α_2Ms and that rat α_1M was serologically more closely related to the mouse homologue, as rat α_2M was to human α_2M. It has also been shown that some structural features are uniquely shared by rat α_1M and mouse αM, but not by other members of mammalian αMs (*9, 46, 47*): (i) rat α_1M and mouse αM are composed of two subunits of 165- and 35-kDa with only trace amounts of 180 kDa subunit, whereas the other members are composed of 180 kDa subunit; (ii) rat α_1M and mouse αM lack an intrachain disulfide bridge between Cys^{h1329} and Cys^{h1444}; and (iii) rat α_1M and mouse αM have a pentapeptide insert (SASSR in

rat α_1M or SESSR in mouse αM) between Thr^{h1195} and Ala^{h1196}. These features reflect the long history of the two mammalian tetrameric αMs after separation from the common ancestor (Fig. 3).

Evolution of Mammalian Monomeric and Dimeric αMs—It was recently reported that monomeric αMs are widely distributed among avian, reptilian, and amphibian classes (*48*). This raises the question of whether the mammalian MUG and nonmammalian monomeric αMs evolved directly from a common ancestral gene or whether they evolved in parallel from more distantly related ancestors. Starkey and Barrett (*49*) suggested the former possibility based on the general features of molecular evolution: the primitive gene emerges as a monomer, and then evolves into polymer with concomitant development of additional function. For instance, the primitive hemoglobin of metazoans was monomeric, but evolved into a homotetramer, then a heterotetramer. During this process the hemoglobin genes acquired potency to transport oxygen from mother to embryo and to efficiently deliver oxygen to respiring tissues by interacting cooperatively. The present finding (Fig. 3), however, is not compatible with this general feature of molecular evolution, since (i) the last common ancestor of the two mammalian tetrameric αM subfamilies was likely to have been a tetramer rather than a monomer; (ii) the two tetrameric subfamilies diverged prior to the emergence of MUGs; (iii) more ancient αM members, such as horseshoe crab (*50*) and plaice αMs (*49*), are dimers; and (iv) chicken ovomacroglobulin (*41*), a monomeric member, is a distant outgroup of the MUGs. Thus, the phylogenetic analysis suggests that rodent MUG emerged from the polymeric ancestor. As regards the capacity to inhibit proteinases, the tetrameric αMs are much more efficient than the monomeric αMs. Nonetheless, mice and rats have two functional MUG genes, both of which duplicated in each lineage after their separation (Fig. 3). Furthermore, both species have additional MUG pseudogenes (*32*), indicating that there is a strong driving force to duplicate the MUG gene in the mouse and rat lineages. Although the nature of this force is not known, these results indicate that the emergence of MUG had selective advantages. Covalent proteinase-binding by MUG is nearly quantitative, *i.e.*, as efficient as by tetrameric αMs (*51*), but its proteinase-inhibitory capacity is very weak; for example, the trypsin-inhibitory activity of hamster MUG levels off at about 20% inhibition (*11*). It is unlikely that in the presence of already abundant levels of αM, which was evolving under strong stabilizing pressure, new gene(s) with the same function but much weaker potency would have evolved. This suggests that the main function of MUG is not proteinase inhibition.

The phylogenetic position of human PZP is ambiguous (Fig. 3, A and B). However, it is likely that this protein also emerged from the tetrameric αM lineage. The invertebrate dimeric αMs, such as horseshoe crab αM (*50*) have functions similar to complement, in addition to proteinase-binding capacity. This suggests that they are functionally and phylogenetically very distantly related to human PZP, which has no complement-like activity, and that they have evolved in parallel.

Phylogeny of the Rodents—The NJ and UPGMA trees of the MUG subfamily, as well as the NJ tree of the αM-2 subfamily (Fig. 3, A and B), demonstrate that mouse, rat, and guinea pig genes form a clade prior to joining the human lineage, in agreement with rodent monophyly. On the other hand, the closest relative of the guinea pig αM is human α_2M in the UPGMA tree, which is inconsistent with rodent monophyly. However, it has been pointed out that the UPGMA tree tends to deviate from the real topology when the evolutionary rates among lineages are not constant (*38*). Thus, it may be concluded that the evolutionary history of the αM families favors rodent monophyly.

We are grateful to Dr. N. Saitou, National Institute of Genetics, Mishima, Japan, and Dr. W.-H. Li, Center for Demographic and Population Genetics, University of Texas, Houston, TX, USA, for their generous gifts of the computer programs used in this experiment.

REFERENCES

1. Freye, H. (1975) The rodents in *Grzimek's Animal Life Encyclopedia* (Grzimek, B., ed.) Vol. 11, pp. 191-198, Van Nostrand Reinhold, New York
2. Wriston, J.C., Jr. (1984) Comparative biochemistry of the guinea-pig: A partial checklist. *Comp. Biochem. Physiol.* **77B**, 253-278
3. Li, W.-H., Hide, W.A., Zharkikh, A., Ma, D.-P., and Graur, D. (1992) The molecular taxonomy and evolution of the guinea pig. *J. Hered.* **83**, 174-181
4. Graur, D., Hide, W.A., and Li, W.-H. (1991) Is the guinea-pig a rodent? *Nature* **351**, 649-652
5. D'Erchia, A.M., Gissi, C., Pesole, G., Saccone, C., and Arnason, U. (1996) The guinea-pig is not a rodent. *Nature* **381**, 597-600
6. Nakatani, T., Suzuki, Y., Yoshida, K., and Sinohara, H. (1995) Molecular cloning and sequence analysis of cDNA encoding plasma α-1-antiproteinase from Syrian hamster: Implications for the evolution of Rodentia. *Biochim. Biophys. Acta* **1263**, 245-248
7. Luckett, W.P. and Hartenberger, J.-L. (1993) Monophyly or polyphyly of the order Rodentia: Possible conflict between morphological and molecular interpretations. *J. Mammal. Evol.* **1**, 127-147
8. Sottrup-Jensen, L. (1989) α-Macroglobulins: Structure, shape, and mechanism of proteinase complex formation. *J. Biol. Chem.* **264**, 11539-11542
9. Saito, A. and Sinohara, H. (1985) Murinoglobulin, a novel protease inhibitor from murine plasma. Isolation, characterization, and comparison with murine α-macroglobulin and human α-2-macroglobulin. *J. Biol. Chem.* **260**, 775-781
10. Saito, A. and Sinohara, H. (1985) Rat plasma murinoglobulin: Isolation, characterization, and comparison with rat α-1- and α-2-macroglobulin. *J. Biochem.* **98**, 501-516
11. Miyake, Y., Shimomura, M., Ito, T., Yamamoto, K., Abe, K., Amemiya, S., and Sinohara, H. (1993) Hamster α-macroglobulin and murinoglobulin: Comparison of chemical and biological properties with homologs from other mammals. *J. Biochem.* **114**, 513-521
12. Gauthier, F. and Ohlsson, K. (1978) Isolation and some properties of a new enzyme-binding protein in rat plasma. *Hoppe-Seyler's Z. Physiol. Chem.* **359**, 987-992
13. Taga, H., Funakawa, K., and Nagase, S. (1979) Purification and characterization of α-X protein from rat serum (in Japanese with English summary). *Gan to Kagakuryoho* **6** (Suppl.), 321-329
14. Esnard, F. and Gauthier, F. (1980) Purification and physicochemical characterization of a new rat plasma proteinase inhibitor, α_1-inhibitor III. *Biochim. Biophys. Acta* **614**, 553-563
15. Esnard, F., Gauthier, F., and Maurizot, J.-C. (1981) The interaction between rat plasma α_1 inhibitor$_3$ and chymotrypsin. A protease-protease inhibitor system which gives partially active complexes. *Biochimie* **63**, 767-774
16. Esnard, F., Gutman, N., El Moujehed, A., and Gauthier, F. (1985) Rat plasma α_1-inhibitor$_3$: A member of the α-macroglobulin family. *FEBS Lett.* **182**, 125-129

17. Goto, K., Saito, A., Nagase, S., and Sinohara, H. (1988) Acute phase response of plasma proteins in analbumineminic rats. *J. Biochem.* **104**, 952-955
18. Saito, A., Suzuki, Y., Yamamoto, K., and Sinohara, H. (1985) Plasma murinoglobulin: A novel family of proteins resembling α-2-macroglobulin in *Abstracts of the 13th International Congress of Biochemistry, Amsterdam*, p. 704
19. Suzuki, Y. and Sinohara, H. (1986) Guinea pig plasma murinoglobulin. Purification and some properties. *Biol. Chem. Hoppe-Seyler* **367**, 579-589
20. Martignetti, J.A. and Brosius, J. (1993) Neural BC1 RNA as evolutionary marker: Guinea pig remains a rodent. *Proc. Natl. Acad. Sci. USA* **90**, 9698-9702
21. Gauthier, F. and Mouray, H. (1976) Rat α_2 acute-phase macroglobulin. Isolation and physicochemical properties. *Biochem. J.* **159**, 661-665
22. Yamamoto, K., Tsujino, Y., Saito, A., and Sinohara, H. (1985) Concentrations of murinoglobulin and α-macroglobulin in the mouse serum: Variations with age, sex, strain, and experimental inflammation. *Biochem. Int.* **10**, 463-469
23. Suzuki, Y., Yoshida, K., Ichimiya, T., Yamamoto, T., and Sinohara, H. (1990) Trypsin inhibitors in guinea pig plasma: Isolation and characterization of contrapsin and two isoforms of α-1-antiproteinase and acute phase response of four major trypsin inhibitors. *J. Biochem.* **107**, 173-179
24. Lonberg-Holm, K., Reed, D.L., Roberts, R.C., Herbert, R.R., Hillman, M.C., and Kutney, R.M. (1987) Three high molecular weight protease inhibitors of rat plasma: Isolation, characterization, and acute phase changes. *J. Biol. Chem.* **262**, 438-445
25. Suzuki, Y. and Sinohara, H. (1986) Isolation and characterization of α-macroglobulin from guinea pig plasma. *J. Biochem.* **99**, 1655-1665
26. Suzuki, Y., Yoshida, K., Honda, E., and Sinohara, H. (1991) Molecular cloning and sequence analysis of cDNAs coding for guinea pig α_1-antiproteinases S and F and contrapsin. *J. Biol. Chem.* **266**, 928-932
27. Dale, R.M.K., McClure, B.A., and Houchins, J.P. (1985) A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: Application to sequencing the corn mitochondrial 18S rDNA. *Plasmid* **13**, 31-41
28. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467
29. Frohman, M.A., Dush, M.K., and Martin, G.R. (1988) Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**, 8998-9002
30. Kimura, M. (1989) A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111-120
31. Sottrup-Jensen, L., Sand, O., Kristensen, L., and Fey, G.H. (1989) The α-macroglobulin bait region. Sequence diversity and localization of cleavage sites for proteinases in five mammalian α-macroglobulins. *J. Biol. Chem.* **264**, 15781-15789
32. Overbergh, L., Hilliker, C., Lorent, K., Van Leuven, F., and Van den Berghe, H. (1994) Identification of four genes coding for isoforms of murinoglobulin, the monomeric mouse α_2-macroglobulin: Characterization of the exons coding for the bait region. *Genomics* **22**, 530-539
33. Rubenstein, D.S., Enghild, J.J., and Pizzo, S.V. (1991) Limited proteolysis of the α-macroglobulin rat α_1-inhibitor-3. Implications for a domain structure. *J. Biol. Chem.* **266**, 11252-11261
34. Braciak, T.A., Northemann, W., Hudson, G.O., Shiels, B.R., Gehring, M.R., and Fey, G.H. (1988) Sequence and acute phase regulation of rat α_1-inhibitor III messenger RNA. *J. Biol. Chem.* **263**, 3999-4012
35. Li, W.-H., Wu, C.-I., and Luo, C.-C. (1985) A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon change. *Mol. Biol. Evol.* **2**, 150-174
36. Jukes, T.H. and Cantor, C.R. (1969) Evolution of protein molecules in *Mammalian Protein Metabolism* (Munro, H.N., ed.) pp. 21-132, Academic Press, New York
37. Kimura, M. (1980) A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111-120
38. Li, W.-H. and Graur, D. (1991) *Fundamentals of Molecular Evolution*, Sinauer Associates, Sunderland, MA
39. Saitou, N. and Nei, M. (1987)The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406-425
40. Li, W.-H., Tanimura, M., and Sharp, P. (1987) An evaluation of the molecular clock hypothesis using mammalian DNA sequences. *J. Mol. Evol.* **25**, 330-342
41. Nielsen, K.L., Sottrup-Jensen, L., Nagase, H., Thogersen, H.C., and Etzerodt, M. (1994) Amino acid sequence of hen ovomacroglobulin (ovostatin) deduced from cloned cDNA. *DNA Seq.* **5**, 111-119
42. Christensen, U. and Sottrup-Jensen, L. (1984) Mechanism of α_2-macroglobulin-proteinase interactions. Studies with trypsin and plasmin. *Biochemistry* **23**, 6619-6626
43. Air, G.M., Thompson, E.O.P., Richardson, B.J., and Sharman, G.B. (1971) Amino-acid sequences of kangaroo myoglobin and haemoglobin and the date of marsupial-eutherian divergence. *Nature* **229**, 391-394
44. Devriendt, K., Zhang, J., Van Leuven, F., Van den Berghe, H., Cassiman, J.J., and Marynen, P. (1989) A cluster of α_2-macroglobulin genes on human chromosome 12p: Cloning of the pregnancy-zone protein gene and α2M pseudogene. *Gene* **81**, 325-334
45. Goodman, M., Czelusniak, J., Koop, B.E., Tagle, D.A., and Slightom, J.L. (1987) Globins: A case study in molecular phylogeny. *Cold Spring Harbor Symp. Quart. Biol.* **52**, 875-890
46. Schaeufele, J.T. and Koo, P.H. (1982) Structural comparison of rat alpha-1 and alpha-2-macroglobulins. *Biochem. Biophys. Res. Commun.* **108**, 1-7
47. Van Leuven, F., Torrekens, S., Overbergh, L., Lorent, K., De Strooper, B., and Van den Berghe, H. (1992) The primary sequence and subunit structure of mouse α_2-macroglobulin, deduced from protein sequencing of the isolated subunits and from molecular cloning of the cDNA. *Eur. J. Biochem.* **210**, 319-327
48. Rubenstein, D.S., Thogersen, I.B., Pizzo, S.V., and Enghild, J.J. (1993) Identification of monomeric α-macroglobulin proteinase inhibitors in birds, reptiles, amphibians and mammals, and purification and characterization of a monomeric α-macroglobulin proteinase inhibitor from the American bullfrog *Rana catesbiana*. *Biochem. J.* **290**, 85-95
49. Starkey, P.M. and Barrett, A.J. (1982) Evolution of α_1-macroglobulin: The structure of a protein homologous with human α-2-macroglobulin from plaice plasma. *Biochem. J.* **205**, 105-115
50. Enghild, J.J., Thogersen, I.B., Salvesen, G., Fey, G.H., Figler, N.L., Gonias, S.L., and Pizzo, S.V. (1990) α-Macroglobulin from *Limulus polyphemus* exhibits proteinase inhibitory activity and participates in a hemolytic system. *Biochemistry* **29**, 10070-10080
51. Sottrup-Jensen, L. (1994) Role of internal thiol esters in the α-macroglobulin-proteinase binding mechanism. *Ann. N.Y. Acad. Sci.* **737**, 172-187
52. Kan, C.C., Solomon, E., Belt, K.T., Chain, A.C., Hiorns, L.R., and Fey, G. (1985) Nucleotide sequence of cDNA encoding human α_2-macroglobulin and assignment of the chromosomal locus. *Proc. Natl. Acad. Sci. USA* **82**, 2282-2286
53. Devriendt, K., Van den Berghe, H., Cassiman, J.-J., and Marynen, P. (1991) Primary structure of pregnancy zone protein. Molecular cloning of full-length PZP cDNA clone by the polymerase chain reaction. *Biochim. Biophys. Acta* **1088**, 95-103
54. Warmegard, B., Martin, N., and Johansson, S.E. (1992) cDNA cloning and sequencing of rat α_1-macroglobulin. *Biochemistry* **31**, 2346-2352
55. Gehring, M.R., Shiels, B.R., Northemann, W., Bruijn, M.H.L., Kan, C.C., Chain, A.C., Noonan, D.J., and Fey, G.H. (1987) Sequence of rat liver α_2-macroglobulin and acute phase control of its messenger RNA. *J. Biol. Chem.* **262**, 446-454
56. Regler, R., Sickinger, S., and Schweizer, M. (1991) Differential

regulation of the two mRNA species of the rodent negative acute phase protein, α_1-inhibitor 3. *FEBS Lett.* **282**, 368-372
57. Northemann, W., Shiels, B.R., Braciak, T.A., and Fey, G.H. (1989) Structure and negative transcriptional regulation by glucocorticoids of the acute-phase rat α_1-inhibitor III gene. *Biochemistry* **28**, 84-95
58. Overbergh, L., Torrekens, S., Van Leuven, F., and Van den Berghe, H. (1991) Molecular characterization of the murinoglobulins. *J. Biol. Chem.* **266**, 16903-16910