Published in final edited form as:

J Mol Evol. 1994 February ; 38(2): 121–131.

Evolution of Murine α_1 -Proteinase Inhibitors: Gene Amplification and Reactive Center Divergence

Carol Rheaume¹, Richard L. Goodwin¹, Jean J. Latimer^{2,*}, Heinz Baumann², and Franklin G. Berger¹

¹Department of Biological Sciences, University of South Carolina, Columbia, SC 29208, USA

²Department of Molecular and Cellular Biology, Roswell Park Cancer Institute, Buffalo, NY 14263, USA

Abstract

The organization and sequence of genes encoding the α_1 -proteinase inhibitor (α_1 PI), a major serine proteinase inhibitor of the mammalian bloodstream, have been compared in several species, including murine rodents (genus *Mus*). Analysis of gene copy number indicates that amplification of α_1 PI genes occurred at some time during evolution of the *Mus* genus, leading to fixation of a family of about three to five genes in several existing species (e.g., *M. domesticus* and *M. saxicola*), and only a single gene in others (e.g., *M. caroli*). A phylogeny for the various mammalian α_1 PI mRNAs was constructed based upon synonymous substitutions within coding regions. The mRNAs in different murine species diverged from a common ancestor before the formation of the first species lineages of the *Mus* genus, i.e., about 10–13 million years ago. Thus, α_1 PI gene amplification must have occurred prior to *Mus* speciation; gene families were retained in some, but not all, murine species. The reactive center region of the α_1 PI polypeptide, which determines target protease specificity, has diverged rapidly during evolution of the *Mus* species, but not during evolution of other mammalian species included in the analysis. It is likely that this accelerated evolution of the reactive center, which has been noted previously for serine proteinase inhibitors, was driven by some sort of a positive Darwinian selection that was exerted in a taxon-specific manner. We suggest that evolution of α_1 PI genes of murine rodents has been characterized by both modification of gene copy number and rapid reactive center divergence. These processes may have resulted in a broadened repertoire of proteinase inhibitors that was evolutionarily advantageous during *Mus* speciation.

Keywords

Gene families; mRNA divergence; Alpha-1-proteinase inhibitor

Correspondence to: Franklin G. Berger.

*Present address: Magee Women's Research Institute, Department of Genetics, University of Pittsburgh, Pittsburgh, PA 15213, USA

Introduction

A primary feature of the evolutionary process is the modification of both the structure and regulation of protein-encoding genes. Structural changes may result in amino acid sequence alterations that could impact upon the function of the encoded polypeptide; regulatory modifications perturb expression patterns, including developmental timing, cell specificity, and responses to metabolic effectors. In total, these effects generate diversity in gene structure and function among existing species (Wilson et al. 1977; Ohta 1991). A fundamental question pertains to the nature and extent of selection pressures that drive the appearance and fixation of such diversity.

α_1 -Proteinase inhibitor (α_1 PI), also called α_1 -antitrypsin, is a mammalian serine proteinase inhibitor that specifically recognizes trypsinlike proteinases (Travis and Salvesen 1983; Carrell and Travis 1985; Carrell et al. 1987). It is produced almost exclusively in the liver, and is secreted into the bloodstream, where it functions primarily in the control of neutrophil elastase activity (Travis and Salvesen 1983). The molecular elements responsible for liver specificity of α_1 PI gene transcription have been studied in detail, and a number of *cis*-acting elements that interact with a variety of DNA binding factors have been identified (Grayson et al. 1988; Monaci et al. 1988; Costa et al. 1989).

The number of α_1 PI genes per haploid genome varies among mammalian species. Primates, including baboons and humans, contain a single gene (Rosenberg et al. 1984; Kurachi et al. 1981), as do sheep (Brown et al. 1989), rats (Chao et al. 1990), and the mouse species *Mus caroli* (Berger and Baumann 1985; Latimer et al. 1987; Rheaume et al. 1988). However, multiple genes are found in the mouse species *M. domesticus* (Hill et al. 1985; Krauter et al. 1986; Borriello and Krauter 1990, 1991), indicating that α_1 PI gene amplification occurred recently in mammalian evolution.

The tissue specificity of α_1 PI expression also differs among species. In *Mus caroli*, the gene is abundantly transcribed in renal tubular epithelial cells in addition to the liver (Berger and Baumann 1985; Latimer et al. 1987; Rheaume et al. 1988); this contrasts sharply with other species, where transcription is predominantly liver-specific (Grayson et al. 1988; Monaci et al. 1988; Costa et al. 1989). Renal α_1 PI expression in *M. caroli* is regulated by androgens during postnatal development, while hepatic expression is at the adult level at birth (Latimer et al. 1987).

Primary sequence comparisons among a variety of serine proteinase inhibitors, both within and between mammalian species, have revealed that the reactive center, which determines the inhibitors' specificity for proteinases, shows a high level of sequence divergence compared to other regions of the protein (Hill et al. 1984; Hill and Hastie 1987; Borriello and Krauter 1990). It has been suggested that this reflects an accelerated rate of evolution that may have occurred in response to some sort of positive Darwinian selection (Hill and Hastie 1987). This may be a general feature of serine proteinase inhibitor evolution.

It is clear, therefore, that the organization, sequence, and expression of α_1 PI genes have been modified during evolution. In the present study, we have made a detailed comparison of mammalian α_1 PI genes and their encoded mRNAs. We have placed particular focus on

the timing and species specificity of α_1 PI gene amplification and the extent of reactive center divergence. Our analysis indicates that amplification of the α_1 PI gene occurred prior to the onset of speciation within the *Mus* genus; evolution of the genus was characterized by modification of the size of this primordial gene family. In addition, we find that rapid divergence of the reactive center, a process driven by positive Darwinian selection, occurred among *Mus* species, but not among other mammals. Thus, selective forces driving evolution of the reactive center have been exerted in a taxon-specific manner.

Materials and Methods

Animals

Mice of species *M. domesticus* (strain C57BL/6J) were purchased from the Jackson Laboratory (Bar Harbor, ME). *Mus hortulanus*, *Mus spretus*, and *Mus caroli* were obtained from Dr. Verne Chapman (Roswell Park Cancer Institute, Buffalo, NY). *Mus cervicolor*, *Mus cookii*, *Mus saxicola*, and *Mus pahari* were provided by Dr. Michael Potter under terms of a contract between the National Cancer Institute and Hazleton Laboratories. All animals were used at 2–4 months of age.

Southern Blot Analysis of α_1 PI Gene Organization

DNA was extracted from the livers of mice representing each of the *Mus* species by the methods of Sampsell and Held (1985). Samples containing 10 μ g DNA were digested with *Eco*RI, and the resulting fragments were fractionated on 1.0% agarose gels, transferred to a nylon filter, and hybridized to an α_1 PI-specific probe overnight at 65°C in buffer containing 4 $\times$ SSC (1 $\times$ SSC = 0.015 M sodium citrate, 0.15 M sodium chloride, pH 7.0); the membrane was washed for 2 h at 65°C in buffer containing 2 $\times$ SSC, and for 30 min in buffer containing 0.5 $\times$ SSC. The probe was the cDNA insert of plasmid pCR12, which is a 503-bp *Pst*I fragment of an α_1 PI cDNA isolated from *M. domesticus* strain BALB/c (Rheume et al. 1988).

Cloning and Sequencing of an α_1 PI cDNA from *M. saxicola*

A cDNA library was prepared from liver mRNA of *M. saxicola* in the vector λ gt10 (Chaudhuri et al. 1991). Phage clones containing cDNAs corresponding to α_1 PI mRNAs of *M. saxicola* were isolated from the library using plasmid pCR12 as probe. The 1.4-kb cDNA of one phage was subcloned into pBluescript SK(+) (Stratagene, Inc.), generating plasmid pCR173; it was sequenced by the dideoxy chain termination method (Sanger et al. 1977). Sequencing was facilitated by construction of nested deletions generated with exonuclease III (Henikoff 1984), resulting in a number of overlapping sequences. About two-thirds of the sequence was verified on both strands.

To screen for reactive center heterogeneity among *M. saxicola* α_1 PIs, 87 additional cDNA clones, and a 1 μ l aliquot of each phage stock was grown overnight as a single plaque on a lawn of *E. coli* C600 cells; the plaques were transferred to nylon filters (Benton and Davis 1973) and hybridized to an end-labeled oligonucleotide corresponding to the reactive center region of pCR173 (nucleotides 1,147–1,166; see Fig. 2). The cDNA inserts in the nonhybridizing clones were amplified by asymmetric-PCR using primers corresponding to

the vector/ insert junctions of λ gt10; the single-stranded products were sequenced with a primer representing residues 1,093–1,112, which is just outside the reactive center. (See Fig. 2.) Oligonucleotides corresponding to newly identified reactive centers were hybridized to the nylon filter containing the collection of phage clones. This process was continued until all 87 clones were accounted for.

Computer analysis of mammalian α_1 PIs was done using the sequence analysis software package of the Genetics Computer Group (Devereux et al. 1984). Synonymous and nonsynonymous substitution rates (K_S and K_A , respectively) were calculated using a program developed by Li *et al.* (1985); this program was provided by Dr. W.-H. Li. Phylogenetic trees were constructed from distance matrices and were based upon K_S values; branch lengths were determined by the method of Fitch and Margoliash (1967), as described by Li (1981).

Results

α_1 PI Gene Copy Number in *Mus* Species

Previous studies indicated that the number of functional α_1 PI genes varies among mammalian species; inbred strains of *M. domesticus* contain five genes (Hill et al. 1985; Boriello and Krauter 1990, 1991) while *M. caroli*, rats, sheep, and primates contain only one (Kurachi et al. 1981; Rosenberg et al. 1984; Berger and Baumann 1985; Latimer et al. 1987; Rheume et al. 1988; Brown et al. 1989; Chao et al. 1990). To obtain a rough estimate of α_1 PI gene copy number in the species *M. saxicola*, which is derived from a lineage that separated early in *Mus* evolution (Tseng-Crank and Berger 1987; Jouvin-Marche et al. 1988), genomic DNAs were digested with *Eco*RI and analyzed by Southern blotting. The probe, a 503-bp *Pst*I cDNA fragment (see Materials and Methods), spans a single *Eco*RI site within the *M. domesticus* and *M. caroli* α_1 PI genes and should hybridize to two DNA fragments totaling a length of about 9 kb (Rheume et al. 1988; Latimer et al. 1990). As seen in Fig. 1, *M. domesticus* and *M. saxicola* exhibited four to six relatively intense bands that represent a total of about 22–35 kb of DNA. In contrast, *M. caroli* contains two light-intensity bands that represent a total of about 9 kb of DNA. Since it is known that *M. domesticus* contains five α_1 PI genes (Hill et al. 1985; Krauter et al. 1986; Boriello and Krauter 1990, 1991), and *M. caroli* contains only one (Latimer et al. 1990), we estimate that there are approximately three to five genes in *M. saxicola*. The presence of multiple α_1 PI genes in both *M. domesticus* and *M. saxicola*, which separated from each other 9–10 Mya (Tseng-Crank and Berger 1987; Jouvin-Marche et al. 1988), suggests that amplification of the gene may have occurred early in the evolution of the *Mus* genus, leading to the fixation of multigene families in separate species.

Sequence Analysis of α_1 PI cDNAs from *M. saxicola*

To verify and extend the conclusions drawn from Southern blot analysis, we conducted a more detailed analysis of the α_1 PI gene family in *M. saxicola*. A cDNA library was prepared from *M. saxicola* liver mRNA and screened for α_1 PI clones; plasmid pCR173, containing a 1.4-kb α_1 PI cDNA, was isolated and sequenced (Fig. 2). Alignment with the sequences from other *Mus* species (Sifers et al. 1990; Latimer et al. 1990) indicated 91% identity between

M. domesticus and *M. caroli*, 92% identity between *M. domesticus* and *M. saxicola*, and 88% identity between *M. caroli* and *M. saxicola*. The amino acid sequences for seven mammalian α_1 PIs are compared in Fig. 3. Among the *Mus* sequences, there is 85% identity between *M. domesticus* and *M. caroli*, 85% between *M. domesticus* and *M. saxicola*, and 81% between *M. caroli* and *M. saxicola*; similarities (Gribskov and Burgess 1986) are 92%, 93%, and 89%, respectively, indicating that about half of the amino acid substitutions are chemically conservative. It was rather surprising to find that the *Mus* α_1 PI mRNAs are equally divergent from one another. The *M. saxicola* lineage separated from the others about 8–10 Mya, while the *M. domesticus* and *M. caroli* lineages diverged about 4–6 Mya (Tseng-Crank and Berger 1987; Jouvin-Marche et al. 1988); thus, one would have expected the *M. domesticus* and *M. caroli* sequences to be more closely related to each other than either is to *M. saxicola*. The near-equal extents of divergence among the three species indicate that their α_1 PI genes radiated from a common ancestor at about the same time.

The reactive center regions of a number of α_1 PI clones from the cDNA library were sequenced. It was expected that this region would be hypervariable among the α_1 PI genes within *M. saxicola*, so the number of distinct reactive center sequences should be an indicator of the size of the family in this species. Indeed, this type of analysis was used to define α_1 PI gene copy number in *M. domesticus* (Boriello and Krauter 1990). An oligonucleotide corresponding to the reactive center section of pCR173 (nucleotides 1,147–1,166; see Fig. 2) was hybridized to a filter containing 87 α_1 PI cDNA phage clones. Only 31 of the 87 clones hybridized to this oligonucleotide; in contrast, all of the clones hybridized to an oligonucleotide corresponding to a region just 5' to the reactive center. This suggests that the *M. saxicola* mRNAs are particularly divergent within their reactive centers. In a control experiment, an oligonucleotide corresponding to the reactive center region of the *M. caroli* α_1 PI mRNA hybridized to all α_1 PI-specific cDNA phage clones isolated from a *M. caroli* library. Thus, as expected for a species with a single α_1 PI gene, *M. caroli* does not show reactive center heterogeneity.

One of the *M. saxicola* clones that did not hybridize to the reactive center oligonucleotide of pCR173 was chosen for further analysis. PCR amplification, followed by direct sequencing, led to identification of a distinct reactive center in this clone. An oligonucleotide corresponding to this region was found to hybridize to 13 clones, none of which hybridized to the first oligonucleotide. Through additional cycles of oligonucleotide hybridization and PCR amplification/sequencing of nonhybridizing clones, we identified two more reactive center sequences: the third corresponded to 38 clones while the fourth corresponded to five clones. Thus, based upon reactive center sequences, we could classify the 87 cDNAs of *M. saxicola* into four groups (Fig. 4), indicating the existence of at least four α_1 PI genes in this species. This is consistent with conclusions drawn from the Southern blot experiments. (See Fig. 1.) The four mRNAs are currently being sequenced in their entirety. Partial sequences obtained thus far (R. Goodwin, unpublished) indicate that the *M. saxicola* gene family, like that in *M. domesticus* (Boriello and Krauter 1991), exhibits a high degree of sequence divergence within the reactive center region.

Origin of the Murine α_1 PI Gene Family

We used the extent of synonymous substitution to more accurately determine the evolutionary relationships among the various α_1 PI mRNAs. Since synonymous base substitutions within the coding regions of mRNAs do not affect amino acid sequences, they are generally considered to be selectively neutral, and accumulate at a constant rate over time. Thus, the number of synonymous changes between two sequences is a measure of the time since they shared a common ancestor (Kimura 1977; Jukes and King 1979; Li et al. 1985). Nonsynonymous substitutions, on the other hand, change amino acid sequences and are subject to selective constraints upon protein function (Kimura 1977; Li et al. 1985).

We determined the numbers of both synonymous and nonsynonymous substitutions in pairwise comparisons of representative α_1 PI mRNA coding regions of eight mammalian species, including four within the genus *Mus*. Using a computer program developed by Li et al. (1985), we calculated K_S (the number of synonymous changes per synonymous site) and K_A (the number of nonsynonymous changes per nonsynonymous site) for each pairwise comparison among the α_1 PI sequences (Table 1). The mean K_A/K_S ratio for the 28 comparisons was 0.37 ± 0.14 . It was not surprising to find that $K_S > K_A$, since synonymous substitutions are well tolerated, while nonsynonymous changes are under purifying selection (Kimura 1977; Jukes and King 1979; Li et al. 1985). Indeed, the K_A/K_S ratios for 34 different mRNAs in a variety of mammalian species, as tabulated by Li et al. (1985), averaged 0.22 ± 0.12 , and ranged from 0.030 to 0.47; this wide range primarily reflects large variations in rates of evolution at nonsynonymous sites (Li et al. 1985).

The K_S values obtained from the six pairwise comparisons among the four murine α_1 PI mRNAs were very similar (Table 1). This is clearly distinct from what is predicted based upon the phylogeny of the species themselves: *M. saxicola* separated from *M. domesticus* and *M. caroli* about 8–10 Mya; *M. caroli* separated from *M. domesticus* about 4–6 Mya (Tseng-Crank and Berger 1987; Jouvin-Marche et al. 1988); the two *M. domesticus* inbred strains were established in laboratory colonies less than a century ago (Festing 1989). Indeed, other genes that have been analyzed in these species, including *RP2* and *Odc*, have diverged in concert with the species themselves (Chaudhuri et al. 1991; Johannes and Berger 1993).

To determine when the α_1 PI multigene family arose, we utilized the K_S values in Table 1 to calculate the separation times for the various mRNAs. Evolutionary rates of 6.5×10^{-9} changes/site/year for rodent lineages, and 3.0×10^{-9} changes/site/year for sheep and primate lineages were assumed in these calculations; these represent average rates derived from analysis of a number of mammalian genes (Li et al. 1987). The divergence times are presented in Table 2, and the corresponding phylogeny is shown in the lower panel of Fig. 5. Separation of the rat, sheep, and primate sequences parallels divergence of the species themselves, indicating that these mRNAs are products of a single allelic α_1 PI gene that is evolving in a conventional fashion. In contrast, the four *Mus* sequences radiated from a common ancestor about 10–13 Mya (Fig. 5), which is the approximate time of formation of the first lineages leading to existing *Mus* species (Tseng-Crank and Berger 1987; Jouvin-Marche et al. 1988); as mentioned above, this pattern of mRNA divergence does not parallel

Mus phylogeny. The results, in total, suggest that amplification of the α_1 PI gene occurred prior to the time of formation of the first species lineages of the *Mus* genus; existing murine species have fixed various members of this primordial family.

Rapid Evolution of the α_1 PI Reactive Center Specifically Within the *Mus* Genus

Several studies have shown that the reactive centers of serine proteinase inhibitors, including α_1 PI, have been subjected to a positive Darwinian selection that has accelerated the rate of evolution of this region of the protein (Hill et al. 1984; Hill and Hastie 1987; Laskowski et al. 1987; Boriello and Krauter 1990, 1991). However, analyses to date have not indicated whether this phenomenon has occurred in all mammalian lineages. To determine if rapid reactive center evolution is taxon-specific to any degree, we measured the extents of divergence in this region among eight species for which appropriate sequence data are available.

The coding regions of the α_1 PI mRNAs were divided into two domains: domain 1 spans the region between codons 367 and 387, and includes the reactive center (Boriello and Krauter 1990, 1991); domain 2 includes the 366 N-terminal codons. (See Fig. 2.) For each of the 28 pairwise comparisons among the eight mammalian species, the K_A s for domains 1 and 2 (termed K_{A1} and K_{A2} , respectively) were calculated and expressed relative to the K_S for domain 2 (termed K_{S2}). The ratio K_{A1}/K_{S2} measures the rate of amino acid replacement in the reactive center relative to the overall neutral rate of α_1 PI mRNA evolution. A ratio <1 would indicate that the amino acid replacement rate is less than the neutral rate, suggesting a selection against amino acid substitutions in the reactive center region. Conversely, a ratio >1 would indicate that the amino acid replacement rate exceeds the overall neutral rate; such would be consistent with the existence of a positive, or diversifying, selection. Table 3 shows that the K_{A1}/K_{S2} ratios for mammalian α_1 PIs averaged 0.79 ± 0.49 . Interestingly, the average ratio for murine species was much higher than for the other mammalian species. While the ratios for *Mus* sequences averaged 1.48 ± 0.40 , those for rat, sheep, and primate sequences averaged 0.27 ± 0.068 (Table 3). The reactive centers of human and baboon are, in fact, identical (Fig. 3, Table 3). Thus, within the genus *Mus*, the α_1 PI reactive center has accumulated amino acid replacements more rapidly than have other regions of the inhibitor. This has not occurred in other mammalian species, indicating that the phenomenon is taxon-specific. The fact that $K_{A1}/K_{S2} > 1$ for all comparisons among the murine species suggests the existence of a positive selection pressure driving the accumulation of amino acid replacements in the reactive center region.

The K_{A2}/K_{S2} ratios, reflecting the amino acid replacement rate throughout most of the α_1 PI molecule relative to the neutral rate, averaged 0.35 ± 0.095 among the eight mammalian species (Table 3). The mean ratios were 0.45 ± 0.14 among *Mus* species and 0.29 ± 0.063 among the others. Thus, most of the amino acid coding region is under negative selection and does not show an overall high rate of accumulation of nonsynonymous substitutions, as does the reactive center.

Discussion

Earlier studies had shown that inbred strains of *M. domesticus* contain five α_1 PI genes (Hill et al. 1985; Krauter et al. 1986; Borriello and Krauter 1990, 1991). In contrast, only a single gene exists in the murine species *M. caroli*, as well as in rats, sheep, and primates (Rosenberg et al. 1984; Kurachi et al. 1981; Brown et al. 1989; Chao et al. 1990; Berger and Baumann 1985; Latimer et al. 1987). Thus, at some time during evolution of the *Mus* genus, α_1 PI gene amplification occurred, and led to fixation of a gene family in *M. domesticus*. Southern blotting experiments (Fig. 1) indicate that *M. saxicola*, like *M. domesticus*, contains multiple α_1 PI genes. Sequence analysis of the reactive center regions of cDNA clones corresponding to α_1 PI mRNAs of *M. saxicola* led to the identification of four distinct genes in this species (Fig. 4). It is thus clear that a family of genes has been fixed in at least two distantly related *Mus* species.

Two scenarios might explain the presence of α_1 PI multigene families in both *M. domesticus* and *M. saxicola*. It is possible that amplification of the α_1 PI gene occurred several times during evolution of the *Mus* genus, and thus the gene families present in existing species arose independently. Alternatively, α_1 PI gene amplification may have predated *Mus* speciation, occurring prior to formation of lineages leading to existing *Mus* species; subsequent evolution of the genus could have resulted in a multigene family in some species (e.g., *M. domesticus* and *M. saxicola*), but not in others (e.g., *M. caroli*). This second possibility is consistent with the phylogeny of the α_1 PI mRNAs (Fig. 5), which indicates that the mRNAs in different murine species diverged from a common ancestor before separation of the first lineages leading to current *Mus* species. This suggests that α_1 PI gene amplification did occur prior to *Mus* speciation, so existing species contain what are essentially evolutionary remnants of this primordial gene family.

If, in fact, an α_1 PI gene family was already present in the pre-*Mus* population, then how did interspecies variations in gene copy number originate? Perhaps the progenitor *Mus* population was polymorphic with regard to the number of α_1 PI genes; during speciation different haplotypes from this population were fixed. Gene copy number polymorphisms do exist within species (Takenaka et al. 1991; Zhu et al. 1991), presumably resulting from unequal crossover events during replication. According to this model, species lineages leading to *M. domesticus* and *M. saxicola* derived from ancestors with multiple genes, while the lineage leading to *M. caroli* derived from ancestors having a single gene. However, it is also possible that all existing *Mus* species are derived from progenitors containing multiple α_1 PI genes. Species-specific modifications of gene copy number may have occurred within particular lineages following their isolation. Thus, *M. domesticus* and *M. saxicola* retained the gene family inherited from this population, while *M. caroli* lost all but one gene.

Several studies have demonstrated that the reactive centers of serine proteinase inhibitors have evolved at an accelerated rate, driven by a positive Darwinian selection (Hill et al. 1984; Hill and Hastie 1987; Laskowski et al. 1987; Boriello and Krauter 1990). Such a mode of evolution was initially demonstrated for α_1 PI and *Spi-2* genes; the latter encode α_1 -antichymotrypsin in the human, contrapsin in the mouse, and a serine proteinase inhibitor of unknown specificity in the rat (Hill et al., 1984; Hill and Hastie 1987). More recently, Yoon

et al. (1987) have identified two members of a novel growth-hormone-regulated serine proteinase inhibitor family that show a high number of amino acid replacements within their reactive centers. Borriello and Krauter (1990, 1991) have shown that the reactive centers of the five α_1 PIs in C57BL/6 mice are extensively diverged; this also appears to be the case for the α_1 PI gene family members of *M. saxicola* (Fig. 4). The present study indicates that rapid evolution of the reactive centers of mammalian α_1 PIs is taxon-specific, having occurred in mice, but not in rats, sheep, or primates. However, the phenomenon may not be unique to the *Mus* genus. Suzuki et al. (1991) have pointed out that rabbit and guinea pig α_1 PIs, which are also encoded by small multigene families, show extensive divergence within their reactive centers. Thus, it is likely that both α_1 PI gene amplification, as well as acceleration of the amino acid substitution rate within reactive centers, occurred in several distinct mammalian lineages. An interesting correlation exists between the presence of multiple α_1 PI genes and the occurrence of rapid reactive center divergence (Suzuki et al. 1991); the only exception to this is *M. caroli* (Latimer et al. 1987; Table 3). It may be that gene amplification and rapid reactive center divergence are coupled—i.e., formation of gene families is a prerequisite to increasing the reactive center's amino acid substitution rate.

A role for positive selection in driving rapid reactive center divergence is implied by the observation that the number of nonsynonymous substitutions within the α_1 PI reactive center exceeds the number of synonymous substitutions in other regions of the mRNA (Table 3). Thus, the reactive center may have been under a diversifying selection pressure that favored the accumulation of amino acid replacements. The effects of such replacements upon α_1 PI function are not known. Some of the changes occur at the P₁ amino acid (see Fig. 3), which is a critical residue in the determination of target proteinase specificity (Travis and Salvesen 1983); thus, a possible consequence of reactive center divergence might have been alterations in the nature and/or number of proteinases recognized by the inhibitor. Another possibility relates to the observation that the reactive center forms a protruding loop that is highly susceptible to proteolytic attack (Kress and Catanese 1981; Huber and Carrell 1989; Stein et al. 1990). Perhaps divergence within the reactive center offers broadened protection against such attack. It is interesting to note that most of the amino acid replacements within the reactive center have occurred in the region that is N-terminal to the P₁ residue (Fig. 4; Boriello and Krauter 1990); this subregion forms the major portion of the protruding loop (Huber and Carrell 1989; Stein et al. 1990).

Hill and Hastie (1987) have suggested that infectious parasites may have been important sources of selection upon the reactive center. These organisms secrete a wide variety of proteolytic enzymes, including serine proteinases, that are involved in invasion of the host and digestion of extracellular proteins within the bloodstream (Banyal et al. 1981; McKerrow et al. 1985; Rosenthal et al. 1989). Amplification of genes encoding proteinase inhibitors, along with acceleration of divergence within their reactive centers, may have provided an effective strategy for combating parasitic invasion. Thus, there may be an evolutionary “interplay” between proteolytic enzymes and proteinase inhibitors, resulting in rapid coevolution of both classes of proteins (Neurath 1984; Creighton and Darby 1989). Regardless of the exact nature of the selective forces involved, the present study makes clear

that such forces were exerted in a taxon-specific manner, driving the evolution of α_1 PI reactive centers in some, but not all, phylogenetic lineages.

Acknowledgments

We thank Dr. W.-H. Li for providing the computer program for determining synonymous and nonsynonymous substitutions. This work was supported by a grant (DK33886) from the National Institutes of Health.

References

- Banyal HS, Misra GC, Gupta CM, Dutta GP. Involvement of malarial proteases in the interaction between the parasite and host erythrocyte in *Plasmodium knowlesi* infections. *J Parasitol.* 1981; 67:623–626. [PubMed: 7028937]
- Benton WD, Davis RW. Screening λ gt recombinant clones by hybridization to single plaques in situ. *Science.* 1973; 196:180–182. [PubMed: 322279]
- Berger FG, Baumann H. An evolutionary switch in tissue-specific gene expression: abundant expression of α_1 -antitrypsin in the kidney of a wild mouse species. *J Biol Chem.* 1985; 260:1160–1165. [PubMed: 2981835]
- Borriello F, Krauter KK. Reactive site polymorphism in the murine protease inhibitor gene family is delineated using a modification of the PCR reaction (PCR + 1). *Nucleic Acids Res.* 1990; 18:5481–5487. [PubMed: 2216721]
- Borriello F, Krauter KK. Multiple murine α_1 -protease inhibitor genes show unusual evolutionary divergence. *Proc Natl Acad Sci USA.* 1991; 88:9417–9421. [PubMed: 1946354]
- Britten RJ. Rates of DNA sequence evolution differ between taxonomic groups. *Science.* 1986; 231:1393–1398. [PubMed: 3082006]
- Brown WM, Dziegielewska KD, Foreman RC, Saunders NR, Wu Y. Nucleotide and deduced amino acid sequence of sheep α_1 -antitrypsin. *Nucleic Acids Res.* 1989; 17:6398. [PubMed: 2788872]
- Carrell RW, Pemberton PA, Boswell DR. The serpins: evolution and adaptation in a family of protease inhibitors. *Cold Spring Harbor Symp Quant Biol.* 1987; 52:527–535. [PubMed: 3502621]
- Carrell R, Travis J. α_1 -antitrypsin and the serpins: variation and countervariation. *Trends Biochem Sci.* 1985; 10:20–24.
- Chao S, Chai KX, Chao L, Chao J. Molecular cloning and primary structure of rat α_1 -antitrypsin. *Biochemistry.* 1990; 29:323–329. [PubMed: 2302382]
- Chaudhuri AC, Barbour K, Berger FG. Evolution of messenger RNA structure and regulation in the genus *Mus*: the androgen-inducible *RP2* mRNAs. *Mol Biol Evol.* 1991; 8:641–653. [PubMed: 1766361]
- Costa RH, Grayson DR, Darnell JE. Multiple hepatocyte-enriched nuclear factors function in the regulation of transthyretin and α_1 -antitrypsin genes. *Mol Cell Biol.* 1989; 9:1415–1425. [PubMed: 2786140]
- Creighton TE, Darby NJ. Functional evolutionary divergence of proteolytic enzymes and their inhibitors. *Trends Biochem Sci.* 1989; 14:319–324. [PubMed: 2678630]
- Devereux J, Haeberli P, Smithies O. A comprehensive set of sequence programs for the VAX. *Nucleic Acids Res.* 1984; 12:387–395. [PubMed: 6546423]
- Festing, MFW. Inbred strains of mice. In: Lyon, MF.; Searle, AG., editors. *Genetic variants and strains of the laboratory mouse.* Oxford: Oxford University Press; 1989. p. 636–648.
- Fitch WM, Margoliash E. Construction of phylogenetic trees. A method based on mutational distances as estimated from cytochrome c sequences is of general applicability. *Science.* 1967; 155:279–284. [PubMed: 5334057]
- Grayson DR, Costa RH, Xanthopoulos KG, Darnell JE. A cell-specific enhancer of the mouse α_1 -antitrypsin gene has multiple functional regions and corresponding protein-binding sites. *Mol Cell Biol.* 1988; 8:1055–1066. [PubMed: 2835657]
- Gribskov M, Burgess RR. Sigma factors for *E. coli*, *B. subtilis*, phage SPO1, and phage T4 are homologous proteins. *Nucleic Acids Res.* 1986; 14:6745–6763. [PubMed: 3092189]

- Henikoff S. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene*. 1984; 28:351–359. [PubMed: 6235151]
- Hill RE, Hastie ND. Accelerated evolution in the reactive centre regions of serine protease inhibitors. *Nature*. 1987; 326:96–99. [PubMed: 3493437]
- Hill RE, Shaw PH, Barth RK, Hastie ND. A genetic locus closely linked to a protease inhibitor gene complex controls the level of multiple RNA transcripts. *Mol Cell Biol*. 1985; 5:2114–2122. [PubMed: 2427931]
- Hill RE, Shaw PH, Boyd PA, Baumann H, Hastie ND. Plasma protease inhibitors in mouse and man: divergence within the reactive centre regions. *Nature*. 1984; 311:175–177. [PubMed: 6547997]
- Huber R, Carrell RW. Implications of the three-dimensional structure of α_1 -antitrypsin for structure and function of serpins. *Biochemistry*. 1989; 28:8951–8966. [PubMed: 2690952]
- Johannes GJ, Berger FG. Domains within the mammalian ornithine decarboxylase messenger RNA have evolved independently and episodically. *J Mol Evol*. 1993; 36:555–567. [PubMed: 8350350]
- Jouvin-Marche E, Cuddihy A, Butler S, Hansen J, Fitch WM, Rudikoff S. Molecular evolution of a single-copy gene: the immunoglobulin C_K locus in wild mice. *Mol Biol Evol*. 1988; 5:500–511. [PubMed: 3143045]
- Jukes TH, King JL. Evolutionary nucleotide replacements in DNA. *Nature*. 1979; 281:605–606. [PubMed: 492328]
- Kimura M. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature*. 1977; 267:275–276. [PubMed: 865622]
- Krauter KK, Citron BA, Hsu M-T, Powell D, Darnell JE. Isolation and characterization of the α_1 -antitrypsin gene of mice. *DNA*. 1986; 5:29–36. [PubMed: 3007061]
- Kress LF, Catanese JJ. Identification of the cleavage sites resulting from enzymatic inactivation of human antithrombin III by *Crotalus adamanteus* proteinase II in the presence and absence of heparin. *Biochemistry*. 1981; 20:7432–7438. [PubMed: 7034776]
- Kurachi K, Chandra T, Friezner-Degen SJ, White TT, Marchioro TL, Woo SL, Davie EW. Cloning and sequence of cDNA coding for α_1 -antitrypsin. *Proc Natl Acad Sci USA*. 1981; 78:6826–6830. [PubMed: 7031661]
- Laskowski M, Kato I, Kohr WJ, Park SJ, Tashiro M, Whatley HE. Positive Darwinian selection in evolution of protein inhibitors of serine proteases. *Cold Spring Harbor Symp Quant Biol*. 1987; 52:545–553. [PubMed: 3502622]
- Latimer JJ, Berger FG, Baumann H. Developmental expression, cellular localization, and testosterone regulation of α_1 antitrypsin in *Mus caroli* kidney. *J Biol Chem*. 1987; 262:12641–12646. [PubMed: 3497926]
- Latimer JJ, Berger FG, Baumann H. Highly conserved upstream regions of the α_1 -antitrypsin gene in two mouse species govern liver-specific expression by different mechanisms. *Mol Cell Biol*. 1990; 10:760–769. [PubMed: 1689000]
- Li W-H. Simple method for constructing phylogenetic trees from distance matrices. *Proc Natl Acad Sci USA*. 1981; 78:1085–1089. [PubMed: 6940127]
- Li W-H, Wu C-I, Luo C-C. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol Biol Evol*. 1985; 2:150–174. [PubMed: 3916709]
- Li W-H, Tanimura M, Sharp PM. An evaluation of the molecular clock hypothesis using mammalian DNA sequences. *J Mol Evol*. 1987; 25:330–342. [PubMed: 3118047]
- McKerrow JH, Pino-Heiss S, Lindquist R, Werb Z. Purification and characterization of an elastolytic proteinase secreted by cercariae of *Schistosoma mansoni*. *J Biol Chem*. 1985; 260:3703–3707. [PubMed: 3882713]
- Monaci P, Nicosia A, Cortese R. Two different liver-specific factors stimulate *in vitro* transcription from the human α_1 -antitrypsin promoter. *EMBO J*. 1988; 7:2075–2087. [PubMed: 3262058]
- Neurath H. Evolution of proteolytic enzymes. *Science*. 1984; 224:350–357. [PubMed: 6369538]
- Ohta T. Multigene families and the evolution of complexity. *J Mol Evol*. 1991; 33:34–41. [PubMed: 1909373]

- Rheume C, Latimer JJ, Baumann H, Berger FG. Tissue-and species-specific regulation of murine α_1 - antitrypsin gene transcription. *J Biol Chem*. 1988; 263:15118–15121. [PubMed: 3262614]
- Rosenberg S, Barr PJ, Najarian RC, Hallewell RA. Synthesis in yeast of a functional oxidation-resistant mutant of human α_1 -antitrypsin. *Nature*. 1984; 312:77–80. [PubMed: 6387509]
- Rosenthal PJ, McKerrow JH, Rasnick D, Leech JH. *Plasmodium falciparum*: inhibitors of lysosomal cysteine proteinases inhibit a trophozoite proteinase and block parasitic development. *Mol Biochem Parasitol*. 1989; 35:177–184. [PubMed: 2671722]
- Sampsel BM, Held WA. Variation in the major urinary protein multigene family in wild-derived mice. *Genetics*. 1985; 109:549–568. [PubMed: 3979814]
- Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA*. 1977; 74:5463–5467. [PubMed: 271968]
- Sifers RN, Carlson JA, Clift SM, DeMayo FJ, Bullock DW, Woo SLC. Tissue specific expression of the human alpha-1-antitrypsin gene in transgenic mice. *Nuc Acids Res*. 1987; 15:1459–1475.
- Sifers RN, Ledley FD, Reed-Fourquet L, Ledbetter DH, Ledbetter SA, Woo SLC. Complete cDNA sequence and chromosomal localization of mouse α_1 -antitrypsin. *Genomics*. 1990; 6:100–104. [PubMed: 2303252]
- Stein PE, Tewkesbury DA, Carrell RW. Ovalbumin and angiotensinogen lack serpin S-R conformational change. *Biochem J*. 1989; 262:103–107. [PubMed: 2818556]
- Stein PE, Leslie AGW, Finch JT, Turnell WG, McLaughlin PJ, Carrell RW. Crystal structure of ovalbumin as a model for the reactive centre of serpins. *Nature*. 1990; 347:99–102. [PubMed: 2395463]
- Suzuki Y, Yoshida K, Sinohara H. Molecular cloning and sequence analysis of cDNAs coding for guinea pig α_1 -antiproteinases S and F and contrapsin. *J Biol Chem*. 1991; 266:928–932. [PubMed: 1985973]
- Takenaka A, Ueda S, Terao K, Takenaka O. Multiple α -globin genes in crab-eating macaques (*Macaca fascicularis*). *Mol Biol Evol*. 1991; 8:320–326. [PubMed: 2072861]
- Travis J, Salvesen JS. Human plasma protease inhibitors. *Ann Rev Biochem*. 1983; 52:655–709. [PubMed: 6193754]
- Tseng-Crank J, Berger FG. Evolution of steroid-inducible *RP2* mRNA expression in the mouse kidney. *Genetics*. 1987; 116:593–599. [PubMed: 3623081]
- Wilson AC, Carlson SS, White TJ. Biochemical evolution. *Ann Rev Biochem*. 1977; 46:573–639. [PubMed: 409339]
- Yoon J-B, Towle HC, Seelig S. Growth hormone induces two mRNA species of the serine protease inhibitor gene family in rat liver. *J Biol Chem*. 1987; 262:4284–4289. [PubMed: 3494016]
- Zhu Z, Vincek V, Figueroa F, Schönbach C, Klein J. *Mhc-DRB* genes of the pigtail macaque (*Macaca nemestrina*): implications for the evolution of human DRB genes. *Mol Biol Evol*. 1991; 8:563–578. [PubMed: 1766359]

Fig. 1.

Southern blot analysis of α_1 PI genes in *Mus* species. Liver DNA (10 μ g) from each of several species was digested with *Eco*RI, and resulting fragments were fractionated on agarose gels, transferred to a nylon membrane, and hybridized to 32 P-labeled α_1 -PI-specific probe. (See Materials and Methods.) The lanes are *M. domesticus* strain C57BL/6J (DOM); *M. caroli* (CAR); *M. saxicola* (SAX). Each lane contains DNA from a single animal; all were run on the same blot. Size markers are *Hind*III fragments of λ DNA, and are indicated to the left.

SAX	1	CGGGTCACCAACCCAGCTCTGGGACAGCAAGCTGAAAAAGACTCCCTCCATCTCATGGCGTCTACTGCTTCTGGCAGGCTGTGTTGCCTGGTCCCAAGC	***
CAR		TCC.....T.....G.....T.....	
DOM		G.....	
SAX	101	ATCTGGCAGAGGATGTCCAAGAGACAGACACCTCCAGAAGGATCAGTCCCCAGCCTCCCATGAGATGGCTACAAACCTGGGAGACTTTGCCTTCAGCCT	
CAR		T...A..T.....T..G.....AGA.G....TCAGT.....T..CTC..CTAC.....--C..GC..T..A.....	
DOM		T.....T..G.....C.....AG.....C.....AA.....	
SAX	201	ATACCGGGAGCTGGTCCATCAGTCCAACACTTCCAACATCTTCTTCTCCCCAGTGAGCATTGCCACAGCCTTTGCTTTGCTCTCCCTAGGGAGCAAGGCT	
CAR		T..GA..CA.....C.....AG.....A.....GAG.....	
DOM		A.....	
SAX	301	GACACTCAGACACAGATTCTAGAGGGCCTGCAGTTCAACCTCACACAAACATCGGAGGCTGACATCCACAAGGTCTTCCAACACCTCCTCCAACCCCTCA	
CAR		C..G.....C.....C.....TC.....	
DOM		C..G.....C.....	
SAX	401	ACAGGCCAGACAGTGAAGCTGAGCTGAGCACAGGCAATGGCCTCTTGTCAACAATGACCTGAAGCTGGTGGAGAAGTTTCTGGAAGAGGCCAAGAACCA	
CAR		TCT..GC.....A.....	
DOM		A.....T.....	
SAX	501	TTACCAATCAGAAGTCTTCTCTGTCAACTTTGCAAAGTCAGAGGAGGCCAGGAAAATGATTAATGATTTCTGGGAGAAGGGAACCCAAGGAAAGATAGTT	
CAR		...T..C.....C.....G.....A..G...C.....T.....C.....	
DOM		...T..GG.....G.....A..G...T.....	
SAX	601	GATGCGGTGAAAGACTTGGATGAAGACACAGTTTTTGCCTGGCGAATTACATTTTCTTCAAGGCAAGTGAAGACGCCATTGATCCTGAGCACACTA	
CAR		..G.....CCT..C.....C.....C.....A.....A.....C.....A.....G	
DOM		..G..A.....AC...CC.....C.....GC.....C.T..A.....A.....C.....A.....G	
SAX	701	CGGAAGCTGACTTCCACGTGAACGAGTCGACCACGGTGAGGGTGCCCATGATGAACCTCATGAGAATGCTTGACGTGCACTATTGCAGCAGCTGTCCAG	
CAR		A.....A.....G..ACAGTC.....A.....C.....C.G.C.....C.....	
DOM		AA.....G.....G..A...C.....A.....C...TC.G.C.....T.....C.....A..C.T...	
SAX	801	CTGGGTGCTGATGATGATTACCTGGGCACGCCACTGCTGTCTTCTCTTGGCCGATGATGGCAAGATGCAGCATCTGGAGCAAACCTCTCAACAAGGAG	
CAR		C.....TCG.....C.....	
DOM		C.....GC.....G.....C.....A..G.....	
SAX	901	CTCATCTCTAAGTTCTCTGCTAAACAGGCACAGAAGTTTAGCCGAGATCCACTTCCCCAGACTGTCCATCTCTGGAAGCTATAACTTGAAGGCACTCATGG	
CAR		G.....C..G.....A.....T.....C.....A.....CT...CA.....A	
DOM		A.....G.....G...T.C...TA.....GAA.....A.....A	
SAX	1001	CTCCACTGGGCATCACCCGGGTCTTCAACAATGGGGCTGACCTCTCCGGAATCACAGAGGAGAATGCTCCCTGAGGCTCAGCAAGGCTGTGCATAAGGC	
CAR		G..AC.....A.....A.....CAG.....	
DOM		G.....A.....A.....	
SAX	1101	CGTGCTGACCATCGATGAGAGAGGAACAGAAGCTGCAGCAACCACAATCGTTGAAGCGTTTTTCATGTCTTTGCCCCCTATCTTGCACTTCAACCCACCT	
CAR		G.....C.....G.T...G..T.AC.....CCT.....A.....C..A.....A.A...	
DOM		T.....C.....G.T...G..T.....CCT.....A.....C...G...G.....	
SAX	1201	TTCTGTTTTCACAATAGTTGAAACACACACTCAGACCCCTCTTTGTGGGAAAAGTGGTGGATCCACACGTAAATGACCACCCCTAAGAAGTCATCCTTC	
CAR		...A.....T.....GA.....G.....A.....A.....A.....	
DOM		...C...T.T...T...GA.....G...A.....A.....A.....C..T.....	
SAX	1301	CTTCTGCATTGGGTTTCGCTTCCATTAAACACAGGCTGGCCTGGCACGTGCCTGATGTACAGCAAAAAAAAAA	
CAR		...TCTG..A.T..G..C...T.A..CTCTGACGGCTG..	
DOM		...A.....C.....T.....	

Fig. 2.

Sequences of α_1 PI cDNAs in *Mus* species. The sequence of the *M. saxicola* (SAX) α_1 PI cDNA is shown aligned with those of *M. caroli* (CAR; Latimer et al. 1990) and *M. domesticus* (DOM; Sifers et al. 1990). Numbering is based on the *M. saxicola* sequence. Differences relative to *M. saxicola* are indicated; identities are depicted by dots (.) and deletions by dashes (-). Asterisks (*) mark the translational initiation codon at nucleotides 37–39 and the termination codon at nucleotides 1276–1278.


```

1                               44
SAX MTPSISWRLL LLAGLCCLVP SYLAEDVQE- -----TDTSQ KDQSPASHEM
DOM .....G.....F.....- .....I
CAR .....G.....F.....- .....R R.SV....DT
R .A...RG...A...A...TS...GVLGH AVQETD...AH QE--A.C.KI
SH .AL...TRG...A...A...TS...GVLGH AVQETD...AH QE--A.C.KI
BA .....L.....GS...P.GD AAQKTDTPPH DQNH.TLNKI
H .PS.V...GI.....VS...P.GD AAQKTDTSHH DQDH.TFNKI

45                               94
SAX ATNLGDFAFS LYRELVHGSN TSNIFFSPVS IATAFALLSL GSKGDTQTQI
DOM .....I.....- .....M.....H...
CAR PY...LELSI...G.K.T.....Q.....M.....E...H...
R SS..A.....- .....M.....T.....M.....RK..
SH .P..AN...I.HK.A.....S.....M.....A..N.H.E.
BA TPS.AE...Q.A...ST.....M.....T.A..HSE.
H TP..AE...Q.A...ST.....M.....T.A..HDE.

95                               144
SAX LEGLQPNLTQ TSEADIHKVF QHLLQTLNRP DSELQLSTGN GLFVNNDLKL
DOM .....S.....- .....- .....- .....-
CAR .....A.....- .....SLL.....
R .....E.....IP.....A...H.....N.....KN...
SH .....G.....E LA..E...G...H...Q...NHQ...T...I.ESA..
BA .....N.....E IP..QV.EG..E...R...K...Q...T...L.KS..V
H .....N.....E IP..QV.EG..E...R...Q...Q...T.DG...LSEG..

145                               194
SAX VEKFLFEAKN HYQSEVFSVN FAKSEEARMK INDFVEKGTQ GKIVDAVKDL
DOM .....A.....E...K.V.....E...E..
CAR .....N.H.....E...K.V.....AE...P
R .....V...N.H.A.....D...K.V.....LM.Q
SH .DT...DV...L.H.KA..I...RDA..K.K...SH .....L...
BA .D...DV...L.H.A...EDT...K.Q...NY.....V...L...E.
H .D...DV.K L.H.A.T...GDT...K.Q.....D...E.

195                               244
SAX DEDTVFALAN YIFPQCKWKT PFDPEHTTEA DFHVNESTTV RVPMMNLIM
DOM .....G...L.K...K...N.E...E...DK...K...T.SG.
CAR .....L.K...K...K...E...DTV...K...T.TG.
R .....V...K...R...N...RD...DK...K...RLG.
SH .Q...V...S.K...EK...EV...R...Q...K...RLG.
BA .R...V...K...ER...EV.A.E.E...DQA...K...TLG.
H .R...V...K...ER...EVKD.E.E...DQV...K...KRLG.

245                               294
SAX LDVHYCSTLS SWVLMMDYLG NATAVFLLPD DGKMQHLEQT LNKELISKLE
DOM ...H...F...L...A...S...E...I...
CAR ...H...L...R...I...I...
R F.M...I...I...T.D...R...
SH F.L...DK.A...LL.V...V.C.I...L.L.Q...DK...N..LA...
BA FNIYH.EK...L.K...I.F...E...L...NE...THDI.T...
H FNIQH.KK...L.K...N.I.F...E...L...NE...THDI.T...

295                               344
SAX LNRHRSIAEI HFPRLSISGS YNLKALMAPL GITRVFNNGA DLSGITEENA
DOM ...R.R.VQ...I...E...T...S...I...
CAR ...R.QV...L...L...N...T.NT...SH...I...
R ...QTRS.IL Y.K...T...T.LSS...D...D...
SH EKKFA.S.NL L.K...ET...D...TVLGE...N...S...Q
BA E.EK.RS.L L.K.A.T.T D...TVLGH...K...S...V...D...
H E.ED.RS.SL L.K...T.T D...SVLGG...K...S...V...D...

345                               394
SAX PLRLSKAVHK AVLTIDERGT EAAATTIVEA VFMSLPPIILH FNHPFVFETIV
DOM ..K...T...T...A.VF...P...M...R...D...L.I.F
CAR ..K...AD...M...T...A.VLQ...P...M...N...K...I...
R ..K...Q...L...GA.V...P...QVK...D...I...M...
SH ..MV...L...A...K...GA.FL...IP...DVE...R.LCILY
BA ..K...K...K...GAMFL...IP...I...EVK...K...LMI
H ..K...K...K...GAMFL...IP...I...EVK...K...LMI

395                               413
SAX ETHHTQTPLFV GKVVDPTRK
DOM ...S.I...H...
CAR ...S...I...
R ...S...I...
SH DRN.KS...N...QA
BA .QN.KS...I...N...Q
H .QN.KS...M...N...Q

```

Fig. 3.

Amino acid sequences for mammalian α_1 PIs. The sequence of the α_1 PI of *M. saxicola* (SAX) is shown aligned with those of *M. domesticus* (DOM; Sifers et al. 1990), *M. caroli* (CAR; Latimer et al. 1990), rat (R; Chao et al. 1990), sheep (SH; Brown et al. 1989), baboon (BA; Kurachi et al. 1981), and human (H; Rosenberg et al. 1984). Numbering is based on *M. saxicola*. Identities are indicated by dots (.) and deletions by dashes (–). The reactive center is overlined. An asterisk (*) marks the P₁ amino acid at residue 377.

SAX1: GCAACCACAATCGTTGAAGCCGTTTTTCATGTCTTTGCCCCCTATCTTGCACTTCAACCACCCTTTCGTTTTTC
 SAX2:G...A.....A.....
 SAX3:G..C..C..G.AG..A.....GC.....A.....TG.....C....A
 SAX4:G..A...G.AG.AC.C..TAT...GC....G.....T.....

*

SAX1: A T T I V E A V F M S L P P I L H F N H P F V F
 SAX2: G N
 SAX3: . . . V L A G S Y . . A N . . C . . L L
 SAX4: . . . V I G S T L Y . A . A

Fig. 4.

Sequences of the reactive centers of *M. saxicola* α_1 PIs. The four reactive center regions deduced from analysis of 87 cDNA clones (see text) are aligned in the upper panel, with *asterisks* (*) indicating the codon that corresponds to the P₁ amino acid. Amino acid sequences for the reactive centers are aligned in the lower panel, with the P₁ amino acid indicated by an *asterisk*. Sequence differences are highlighted; identities are indicated by *dots* (.).

Fig. 5.

Phylogeny of mammalian α_1 PI mRNAs. The evolutionary relationships among mammalian species are shown in the upper panel, along with estimated times of divergence as determined primarily from the fossil record (Li et al. 1985; Tseng-Crank and Berger 1987; Britten 1986; Jouvin-Marche et al. 1988). BL/6 and BALB/C represent inbred strains C57BL/6J and BALB/cJ, respectively, of *M. domesticus*. The phylogeny of α_1 PI mRNAs is shown in the lower panel, and is based upon the synonymous substitution rates in Table 1;

branch lengths (in changes/100 sites) are indicated. The thick line highlights the relationships among mRNAs of *Mus* species. (See text for details.)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Number of substitutions per synonymous site and per nonsynonymous site among mammalian α_1 PI mRNAs^a

	D _{B6}	D _C	Car	Sax	R	Sh	Ba	H
D _{B6}	—	0.041 (0.009)	0.082 (0.010)	0.075 (0.009)	0.12 (0.012)	0.25 (0.019)	0.22 (0.018)	0.22 (0.017)
D _C	0.13 (0.032)	—	0.086 (0.014)	0.12 (0.017)	0.16 (0.019)	0.28 (0.028)	0.26 (0.022)	0.26 (0.022)
Car	0.15 (0.026)	0.14 (0.035)	—	0.11 (0.011)	0.15 (0.014)	0.26 (0.020)	0.24 (0.018)	0.23 (0.018)
Sax	0.13 (0.02)	0.17 (0.038)	0.16 (0.027)	—	0.12 (0.012)	0.23 (0.012)	0.23 (0.012)	0.23 (0.012)
R	0.32 (0.041)	0.32 (0.058)	0.36 (0.045)	0.26 (0.036)	—	0.22 (0.017)	0.19 (0.016)	0.19 (0.016)
Sh	0.92 (0.10)	1.06 (0.28)	0.89 (0.099)	0.91 (0.10)	0.86 (0.092)	—	0.18 (0.015)	0.18 (0.015)
Ba	0.82 (0.093)	0.83 (0.13)	0.71 (0.081)	0.70 (0.079)	0.83 (0.093)	0.62 (0.073)	—	0.038 (0.006)
H	0.85 (0.096)	0.76 (0.12)	0.76 (0.089)	0.72 (0.082)	0.85 (0.095)	0.58 (0.067)	0.11 (0.021)	—

^aValues for K_S (below diagonal) and K_A (above diagonal), in substitutions per site, were determined using a computer program developed by Li et al. (1985); standard errors are given in parentheses. The following species were analyzed: *Mus domesticus* inbred strains C57BL/6 (DB6) and BALB/c (DC); *Mus caroli* (Car); *Mus saxicola* (Sax); rat (R); sheep (Sh); baboon (Ba); and human (h). All sequences were from the literature (Rosenberg et al. 1984; Kurachi et al. 1981; Brown et al. 1989; Chao et al. 1990; Hill et al. 1990; Sifers et al. 1990; Latimer et al. 1990), except that for *M. saxicola*, which was determined in the present work (Fig. 1). Comparisons were made for the entire coding region (428 codons, see Fig. 2); for strain BALB/c, they were made for the C-terminal 213 codons, since only that region of the mRNA has been sequenced (Hill et al. 1984).

Table 2Divergence times among mammalian α_1 PI mRNAs^a

Species pair	Divergence time
D _{B6} vs D _C	10.0
vs Car	11.5
vs Sax	10.0
vs R	24.6
vs Sh	96.8
vs Ba	86.3
vs H	89.5
D _C vs Car	10.8
vs Sax	13.1
vs R	24.6
vs Sh	112
vs Ba	87.4
vs H	80.0
Car vs Sax	12.3
vs R	27.7
vs Sh	93.7
vs Ba	74.7
vs H	80.0
Sax vs R	20.0
vs Sh	95.8
vs Ba	73.7
vs H	75.8
R vs Sh	90.5
vs Ba	87.4
vs H	89.5
Sh vs Ba	103
vs H	96.7
Ba vs H	18.3

^aThe species are listed in Table 1. For each species pair, the divergence time (in Myr) was calculated from the equation $t_d = K_S / (r_1 + r_2)$ where t_d is the divergence time, K_S is the fractional number of synonymous substitutions (in changes/site) as taken from Table 1, and r_1 and r_2 are evolutionary rates (in changes/ site/year) for lineages leading to the two species. The evolutionary rate used for rodent lineages was 6.5×10^{-9} while that used for sheep and primate lineages was 3.0×10^{-9} (Li et al. 1987)

Table 3Divergence of the reactive center region of α_1 PI mRNAs^a

	D _{B6}	D _C	Car	Sax	R	Sh	Ba	H
D _{B6}	—	0.21	0.46	0.49	0.35	0.29	0.28	0.27
D _C	1.20	—	0.41	0.51	0.39	0.30	0.33	0.35
Car	1.00	1.39	—	0.63	0.42	0.33	0.36	0.33
Sax	1.54	2.17	1.56	—	0.42	0.26	0.32	0.30
R	0.68	1.24	0.78	1.12	—	0.25	0.23	0.22
Sh	0.46	0.66	0.35	0.45	0.33	—	0.32	0.32
Ba	0.55	0.90	0.42	0.82	0.32	0.19	—	0.38
H	0.53	0.95	0.38	0.77	0.30	0.20	0.00	—

^aThe species are listed in Table 1. For each species pair, the following values were calculated: K_{A1}, the fractional number of nonsynonymous substitutions in domain 1 (the reactive center region, i.e., codons 367–387); K_{S2}, the fractional number of synonymous substitutions in domain 2 (codons 1–366); and K_{A2}, the fractional number of nonsynonymous substitutions in domain 2. The K_{A1}/K_{S2} ratios are shown below the diagonal, while the K_{A2}/K_{S2} ratios are shown above the diagonal