JOURNAL OF EXPERIMENTAL ZOOLOGY (MOL DEV EVOL) 288:165–174 (2000)

Evolutionary History of the Uterine Serpins

MORGAN R. PELTIER,[1] LEE C. RALEY,[2] DAVID A. LIBERLES,[2] STEVEN A. BENNER,[2,3] AND PETER J. HANSEN[1]*
[1]*Department of Dairy and Poultry Sciences, University of Florida, Gainesville, Florida 32611–0920*
[2]*Department of Chemistry, University of Florida, Gainesville, Florida 32611–0920*
[3]*Department of Anatomy and Cell Biology, University of Florida, Gainesville, Florida 32611–0920*

ABSTRACT A bioinformatics analysis was conducted on the four members of the uterine serpin (US) family of serpins. Evolutionary analysis of the protein sequences and 86 homologous serpins by maximum parsimony and distance methods indicated that the uterine serpins proteins form a clade distinct from other serpins. Ancestral sequences were reconstructed throughout the evolutionary tree by parsimony. These suggested that some branches suffered a high ratio of nonsynonymous to synonymous mutations, suggesting episodes of adaptive evolution within the serpin family. Analysis of the sequences by neutral evolutionary distance methods suggested that the uterine serpins diverged from other serpins prior to the divergence of the mammals from other vertebrates. The porcine uterine serpins are paralogs that diverged from a single common ancestor within the *Sus* genus after pigs separated from other artiodactyls. The uterine serpins contain several protein kinase C and tyrosine kinase phosphorylation sites. These sites may be important for the lymphocyte-inhibitory activity of OvUS if, like other basic proteins, OvUS can cross the cell membrane of an activated lymphocyte. Internalized OvUS could serve as an alternative target to protein kinases important for the mitogenic response to antigens. *J. Exp. Zool. (Mol. Dev. Evol.) 288:165–174, 2000.*

Uterine serpin (US) is a protein that is secreted under the influence of progesterone from the endometrial glands of sheep (Moffatt et al., ’87; Ing et al., ’89; Leslie and Hansen, ’91), cattle (Leslie and Hansen, ’91; Mathialagan and Hansen, ’96) and pigs (Malathy et al., ’90). Studies on OvUS have suggested that these proteins may mediate the immunosuppressive effects of progesterone to allow for tolerance of the fetal allograft during mid- to late pregnancy (reviewed in Hansen, ’98). The proteins have been classified as members of the serpin superfamily of proteinase inhibitors based on the alignment of their sequences (Ing and Roberts, ’89; Malathy et al., ’90; Mathialagan and Hansen, ’96). However, no target serine proteinase has been identified for this group of proteins although OvUS has some inhibitory activity to pepsin A and pepsin C (Mathialagan and Hansen, ’96). Additionally, OvUS bound to members of the pregnancy associated glycoproteins (Mathialagan and Hansen, ’96), an inactive group of aspartic proteinases produced by the binuclear cells of the syncytiotrophoblast (Xie et al., ’91, ’94, ’95, ’97) and to the growth factor, activin (McFarlane et al., ’99).

Modern methods of bioinformatics model the phylogeny of a protein family (reviewed in Benner et al., ’98). Other analyses can identify motifs within the sequence of a protein that may be of biological significance. The objective of this study was to perform bioinformatics analyses on the uterine serpin family of proteinase inhibitors as a means of determining how these proteins may have evolved from other serpins and to determine if the proteins contain peptide motifs that may explain their biochemical and immunobiological activities.

Grant sponsor: United States Department of Agriculture National Research Initiative Competitive Grants Program; Grant number: 96-35203; Grant sponsor: National Institutes of Health; Grant numbers: HG 01729 and GM 54075.
*Correspondence to: Dr. Peter J. Hansen, P.O. Box 110920, Department of Dairy and Poultry Sciences, University of Florida, Gainesville FL 32611-0920. E-mail: hansen@dps.ufl.edu
Received 29 November 1999; Accepted 29 March 2000

MATERIALS AND METHODS

Multiple sequence alignments

Potential homologs of OvUS (SwissProt: P21814) were selected from the SwissProt protein database using the BLAST search algorithm (Altshul et al., '97) run on the National Center for Bioinformatics web server (www.ncbi.org). The search utilized the BLOSUM62 protein matrix with a gap penalty of 0.85 and was limited to mammalian sequences. Protein sequences from animals with homology scores of greater or equal to 50 (E values 2×10^{-6}) were downloaded from the database and a multiple sequence alignment was generated by the program ClustalW (Thompson et al., '94). For all computer analyzes, the signal sequence (where known) was removed prior to entry into the programs.

For analyzes of DNA sequences, the coding sequences for the proteins identified in the BLAST search were downloaded from the European Institute of Bioinformatics webserver (www2.edi.org) and regions encoding the signal sequences, introns, and stop codons were removed and a multiple sequence alignment was generated by the program ClustalW (Thompson et al., '94). The nucleic acid sequence for rat corticosteroid binding globulin was not present in the databases; instead the published sequence was used (Smith and Hammond, '89).

Phylogenetic analysis

Phylogenetic trees were built using the method of maximum parsimony with the PHYLIP analysis package (Felsenstein, '89). Parsimony-based algorithms are optimized for determining the topology (arrangement of branches) of an evolutionary tree and function by predicting the ancestral sequences for all possible arrangements of the branches. The topology that would require the fewest mutations in the evolutionary history is then selected as the most probable route of evolution (reviewed in Nei, '87).The reliability of the assignment of the branches on the phylogeny trees produced was estimated by bootstrapping (Felsenstein, '85). This method involves the resampling of the data with the introduction of random changes (some residues are deleted and other are duplicated with no change in the size of the original data set) so that the variability of the branch assignment can be tested. With a large data set as used for the generation of this serpin tree, bootstrap values less than 70% are considered nonsignificant, values of 70–94% are significant and percentages of 95–100% are highly significant (Li and Ford, '98). Trees were constructed by the programs and a consensus tree was produced with the program CONSENSE and plotted using the program TREEVIEW 1.5 (distributed by Rod Page, University of Glasgow, UK). Because of the large number of sequences analyzed, some clades were abbreviated for clarity of presentation.

Calculation of the K_a:K_s ratios

To determine if adaptive evolution was apparent amongst the uterine serpins, the K_a:K_s ratios were assigned to a smaller parsiomony-based tree consisting of the uterine serpins with rat corticosteroid binding globulin assigned a priori as the outgroup. The ratio of nonsilent to synonymous mutations was determined for each branch as previously described (Li et al., '85; Pamilo and Bianchi, '93; Ina, '95, '96; Endo et al., '96; Trabesinger-Ruef et al., '96; Messier and Stewart, '97). The K_a:K_s values represents the normalized ratio of nucleic acid substitutions that result in a change in the protein sequence (nonsynonymous substitutions) divided by nucleic acid substitutions that do not result in a change in the polypeptide sequence (synonymous substitutions). Low K_a:K_s ratios reflect purifying selection during episodes of divergent evolution where physiological function of the protein is likely to remain constant. High K_a:K_s ratios suggest adaptive evolution where certain mutations in the amino acid sequence are favored because the protein is developing a different biological function.

Estimation of time of divergence for uterine serpin sequences

Because nonsynonymous mutations are subject to selection pressures, they do not accumulate in a clock-like manner and can confuse simple models of molecular history, which assume a stochastic behavior. Therefore, we used pairwise neutral evolutionary distances (NEDs; D.A. Liberles et al., unpublished) which were calculated for the aligned uterine serpin genes and the rat corticosteroid binding globulin gene that was the nearest outgroup. The NEDs represent the proportion of conserved twofold degenerate codons (Asp, Cys, Glu, Phe, His, Lys, Asn, Gln, and Tyr) between pairs of nucleic acid sequences. The twofold degenerate amino acids were used because the differences between each of these codons are represented solely by transitions (substitutions of a pyrimidine for a pyrimidine or a purine for a purine) at the third position. All of the three-, four-, and sixfold degen-

erate amino acid codons involve transitions or transversions (substitution of a purine for a pyrimidine or vice versa) between codons that are not limited to the third position. These additional changes in the nucleic acid sequence within the different types of synonymous codons would occur with different probabilities than those due to transitions only (reviewed in Nei, '87). With time this twofold redundant codon system would approach equilibrium according to a first order rate law described by an exponential decay to an equilibrium model ($NED = 0.5e^{-kt} + 0.5$) at an assumed first order rate constant (k) (D.A. Liberles et al., unpublished observations). To convert these distances to number of years (t), k was determined to be 3×10^{-9} changes per base per year by calibration with the fossil record. This value was estimated under the assumption that sheep and cattle diverged 18 million years ago (Mya) and that pigs diverged from ruminants 60 Mya (Carroll, '88).

Motif analysis of the uterine serpins

Motif analysis was performed on uterine serpins and other members of the serpin superfamily to determine whether there were specific amino acid motifs present for uterine serpins that could provide some clues to their function. Motif analysis was performed using the program MOTIF (available at www.motif.genome.jp) and the PROSITE PATTERN library of sequence motifs (Bairoch et al., '97).

RESULTS

Sequence alignments and evolutionary trees for serpin

Table 1 lists the proteins identified as potential homologs of OvUS by the BLAST program (also listed are all the abbreviations used in this report). All of the proteins with significant similarity to OvUS were members of the serpin superfamily of proteinase inhibitors. The most significant scores were found with other members of the uterine serpin family, BoUS, and PoUS-1 and PoUS-2.

A ClustalW alignment of the uterine serpins is shown in Figure 1. The complete CLUSTALW alignment of all proteins with homology to OvUS is published elsewhere (Peltier, 2000). The ovine sequence has a unique insertion at 303–311 that was not present in any of the other serpins. The bovine sequence contains a unique insertion into the putative P_1–P_1' site that had previously been described (Mathialagan and Hansen, '96) for a more limited set of serpin sequences.

An unrooted consensus tree of OvUS and 86 other serpins was built using parsimony methods (Fig. 2). The tree places all uterine serpins within their own clade with high certainty (bootstrap 100%). However, the placement of the limb containing the uterine serpin family on the tree was much less certain, due to low bootstrap values. A tree built with a distance-based method using the PAM distance (point accepted mutation; the number of amino acid differences per 100 residues of polypeptide sequence between protein pairs) with the NEIGHBOR algorithm gave similar results (data not shown).

Adaptive evolution of the uterine serpins

As shown in Figure 3, the ratio of nonsynonymous to synonymous mutations (K_a:K_s ratio) was higher for all branches than would be typically seen for proteins divergently evolving under constant function. Thus, the uterine serpins may be undergoing episodes of adaptive evolution but similar to that observed in other proteinase inhibitors (Laskowski et al., '87). Although the K_a:K_s ratios were usually below 1, it has recently been shown for src homology 2 domains that K_a:K_s ratios less than 1 can still be correlated with adaptive evolution (D.A. Liberles et al., unpublished observations).

Estimation of the time of divergence of the uterine serpins

Analysis of the sequences using the NEDs revealed that the uterine serpin family diverged from the other known serpin sequences more than 60 Mya. The porcine uterine serpins appeared to have diverged from each other at 5 Mya, well after the divergence of pigs from the other artiodactyls.

Motif analysis of the uterine serpin family

Table 2 lists the results of motif analyses conducted on the artiodactyl serpins (EMBL accession numbers J04484, L22095, X62845, M30315, X15555, X63129, X88780, X73615, X69795, X78436, X16383, X68287, U48229, L23110). A complete listing of the results of motif analyses on the serpins listed in Table 1 is published in Peltier (2000). The ovine and bovine uterine serpins contained similar amounts of cAMP and casein kinase-2 phosphorylation sites but the bovine sequence also contained more sites for protein kinase C phosphorylation, tyrosine kinase phosphorylation, N-myristoylation, and N-linked glycosylation sites than OvUS (Table 2). The porcine sequences also contained sites for phosphorylation by casein kinase-2, protein kinase C, tyrosine kinase, tyrosine kinase as well as N-myristoylation and N-linked glycosylation sites.

TABLE 1. Protein sequences with homology to ovine uterine serpin as determined by a BLAST search of all mammalian proteins in the Swiss-Prot database[1]

Accession number	Swiss-Prot designation	Protein name and abbreviation used in paper	Score[2]	E value[3]
P21814	UTMP_SHEEP	Ovine Uterine Serpin (OvUS)	750	0
P46201	UTMP_BOVIN	Bovine Uterine Serpin (BoUS)	575	10^{-165}
P46202	UAB2_PIG	Porcine Uterine Serpin-2 (PoUS-2)	371	10^{-103}
P16708	UFBP_PIG	Porcine Uterine Serpin-1 (PoUS-1)	365	10^{-101}
P31211	CBG_RAT	Corticosteroid Binding Globulin (RatCBG)	164	6×10^{-41}
P01010	A1AT_PAPAN	Alpha 1 Antitrypsin (PapA1AT)	156	1×10^{-38}
P23035	A1AF_RABIT	Alpha 1 Antiproteinase F (LaA1AF)	155	3×10^{-38}
P17475	A1AT_RAT	Alpha 1 Antitrypsin (RatA1AT)	154	5×10^{-38}
P09006	CPI6_RAT	Contrapsin-like Proteinase Inhibitor 6 (RatCPI6)	153	1×10^{-37}
P23775	CBG_RABIT	Corticosteroid Binding Globulin (LaCBG)	152	3×10^{-37}
P05544	CPI3_RAT	Contrapsin-like Proteinase Inhibitor 3 (RatCPI3)	152	3×10^{-37}
P01011	AACT_HUMAN	Alpha 1 Antichymotrypsin (HuACT)	151	6×10^{-37}
P05545	CPI1_RAT	Contrapsin-like Proteinase Inhibitor 1 (RatCPI1)	150	1×10^{-36}
P01009	A1AT_HUMAN	Alpha 1 Antitrypsin (HuA1AT)	149	2×10^{-36}
P29621	KBP_MOUSE	Kallikrein-binding Protein (MuKBP)	148	3×10^{-36}
P50451	CBG_SAISC	Corticosteroid Binding Globulin (SaiCBG)	148	4×10^{-36}
P08185	CBG_HUMAN	Corticosteroid Binding Globulin (HuCBG)	146	1×10^{-35}
P12725	A1AT_SHEEP	Alpha 1 Antitrypsin (OvA1AT)	146	1×10^{-35}
Q06770	CBG_MOUSE	Corticosteroid Binding Globulin (MuCBG)	145	4×10^{-35}
Q00896	A1A3_MOUSE	Alpha 1 Antitrypsin 1-3 (MuA1A3)	144	7×10^{-35}
P34955	A1AT_BOVIN	Alpha 1 Antitrypsin (BoA1AT)	144	7×10^{-35}
P07758	A1A1_MOUSE	Alpha 1 Antitrypsin 1-1 (MuA1A1)	139	1×10^{-33}
Q00897	A1A4_MOUSE	Alpha 1 Antitrypsin 1-4 (MuA1A4)	139	1×10^{-33}
P50447	A1AT_PIG	Alpha 1 Antitrypsin (PoA1AT)	139	2×10^{-33}
P07759	COTR_MOUSE	Contrapsin (MuCOTR)	139	2×10^{-33}
P22599	A1A2_MOUSE	Alpha 1 Antitrypsin 1-2 (MuA1A2)	138	4×10^{-33}
P22324	A1AF_CAVPO	Alpha 1 Antiproteinase F (CaA1AF)	138	6×10^{-33}
P22325	A1AS_CAVPO	Alpha 1 Antiproteinase S (CaA1AS)	136	1×10^{-32}
P49920	CBG_SHEEP	Corticosteroid Binding Globulin (OvCBG)	136	1×10^{-32}
P05543	THBG_HUMAN	Thyroxine Binding Globulin (HuTHBG)	135	3×10^{-32}
Q00898	A1A5_MOUSE	Alpha 1 Antitrypsin 1-5 (MuA1A5)	135	3×10^{-32}
P22323	COTR_CAVPO	Contrapsin (CaCOTR)	131	6×10^{-31}
P50450	THBG_SHEEP	Thyroxine Binding Globulin (OvTHBG)	128	5×10^{-30}
P05154	IPSP_HUMAN	Plasma Serine Proteinase Inhibitor (HuIPSP)	124	5×10^{-29}
P29622	KAIN_HUMAN	Kallistatin (HuKAIN)	121	6×10^{-28}
P20848	A1AU_HUMAN	Alpha 1 Antitrypsin Related Protein (HuA1AU)	120	1×10^{-27}
Q60543	CBG_MESAU	Corticosteroid Binding Globulin (MesCBG)	120	1×10^{-27}
P26595	A1AT_MUSCR	Alpha-1-Antitrypsin (MusA1AT)	119	2×10^{-27}
Q03044	A1AT_DIDMA	Alpha-1-Antitrypsin (DiA1AT)	118	3×10^{-27}
P35577	THBG_RAT	Thyroxine Binding Globulin (RatTHBG)	116	1×10^{-26}
P28800	A2AP_BOVIN	Alpha 1 Antiplasmin (BoA2AP)	101	8×10^{-22}
P29508	SCC1_HUMAN	Squamous Cell Carcinoma Antigen 1 (HuSCC1)	100	2×10^{-21}
P08697	A2AP_HUMAN	Alpha 1 Antiplasmin (HuA2AP)	98	9×10^{-21}
P48594	SCC2_HUMAN	Squamous Cell Carcinoma Antigen 2 (HuSCC2)	96	3×10^{-20}
Q61247	A2AP_MOUSE	Alpha 2 Antiplasmin (MuA2AP)	95	4×10^{-20}

(continued)

TABLE 1. (continued)

Accession number	Swiss-Prot designation	Protein name and abbreviation used in paper	Score[2]	E value[3]
P09005	SI21_RAT	Serine Proteinase Inhibitor 2.1 (RatSPI21)	93	2×10^{-19}
P47776	HEP2_RABIT	Heparin Cofactor II (LaHep2)	92	4×10^{-19}
P41361	ANT3_BOVIN	Antithrombin III (BoANT3)	91	9×10^{-19}
P05121	PAI1_HUMAN	Plasminogen Activator Inhibitor 1 (HuPAI1)	89	3×10^{-18}
P01008	ANT3_HUMAN	Antithrombin III (HuANT3)	88	6×10^{-18}
P80229	ILEU_PIG	Leukocyte Elastase Inhibitor (PoILEU)	88	6×10^{-18}
Q99574	NEUS_HUMAN	Neuroserpin (HuNEUS)	87	1×10^{-17}
P32261	ANT3_MOUSE	Antithrombin III (MuANT3)	87	1×10^{-17}
P79335	PAI1_PIG	Plasminogen Activator Inhibitor 1 (PoPAI1)	87	2×10^{-17}
P05546	HEP2_HUMAN	Heparin Cofactor II (HuHep2)	86	2×10^{-17}
P22777	PAI1_MOUSE	Plasminogen Activator Inhibitor 1 (MuPAI1)	86	3×10^{-17}
P13909	PAI1_BOVIN	Plasminogen Activator Inhibitor 1 (BoPAI1)	85	4×10^{-17}
P50449	PAI1_MUSVI	Plasminogen Activator Inhibitor 1 (MuvPAI-1)	85	4×10^{-17}
P81105	A1A6_MOUSE	Alpha 1 Antitrypsin 1-6 (MuA1A6)	84	9×10^{-17}
P20961	PAI1_RAT	Plasminogen Activator Inhibitor 1 (RatPAI1)	83	1×10^{-16}
P50453	PTI9_HUMAN	Cytoplasmic Antiproteinase 3 (HuCAP3)	83	1×10^{-16}
P32262	ANT3_SHEEP	Antithrombin III (OvANT3)	82	4×10^{-16}
P49182	HEP2_MOUSE	Heparin Cofactor II (MuHep2)	82	4×10^{-16}
O35684	NEUS_MOUSE	Neuroserpin (MuNEUS)	82	6×10^{-16}
P35237	PTI6_HUMAN	Placental Thrombin Inhibitor 6 (HuPTI6)	81	7×10^{-16}
Q64268	HEP2_RAT	Heparin Cofactor II (RatHep2)	81	7×10^{-16}
P30740	ILEU_HUMAN	Leukocyte Elastase Inhibitor (HuILEU)	81	1×10^{-15}
Q07235	GDN_MOUSE	Glia Derived Nexin (MuGDN)	80	1×10^{-15}
Q95121	PEDF_BOVIN	Pigment Epithelium-Derived Factor (BoPEDF)	80	2×10^{-15}
P50452	PTI8_HUMAN	Placental Thrombin Inhibitor 8 (HuPTI8)	80	2×10^{-15}
P07092	GDN_RAT	Glia Derived Nexin (RatGDN)	79	3×10^{-15}
P29524	PAI2_RAT	Plasminogen Activator Inhibitor 2 (RatPAI2)	76	3×10^{-14}
P05619	ILEU_HORSE	Leukocyte Elastase Inhibitor (EqILEU)	75	6×10^{-14}
P36952	MASP_HUMAN	Maspin (HuMASP)	75	7×10^{-14}
P05120	PAI2_HUMAN	Plasminogen Activator Inhibitor 2 (HuPAI2)	75	7×10^{-14}
P07093	GDN_HUMAN	Glia Derived Nexin (HuGDN)	74	1×10^{-13}
P12388	PAI2_MOUSE	Plasminogen Activator Inhibitor 2 (MuPAI2)	74	1×10^{-13}
P48595	BOMA_HUMAN	Bomapin (HuBOMA)	73	2×10^{-13}
P36955	PEDF_HUMAN	Pigment Epithelium-Derived Factor (HuPEDF)	71	6×10^{-13}
P70564	MASP_RAT	Maspin (RatMASP)	71	8×10^{-13}
P70124	MASP_MOUSE	Maspin (MuMASP)	68	5×10^{-12}
P97298	PEDF_MOUSE	Pigment Epithelium-Derived Factor (MoPEDF)	67	2×10^{-11}
P05155	IC1_HUMAN	Plasma Protease C1 Inhibitor (HuIC1)	63	2×10^{-10}
P29457	HS47_RAT	Heat Shock Protein 47 (RatHSP47)	60	2×10^{-9}
P19324	HS47_MOUSE	Heat Shock Protein 47 (MuHSP47)	60	2×10^{-9}
P50454	CBP2_HUMAN	Collagen-Binding Protein 2 (HuCBP2)	60	2×10^{-9}
P29043	HS47_HUMAN	Heat Shock Protein 47 (MuHSP47)	59	4×10^{-9}
P50448	F12I_BOVIN	Factor XIIa Inhibitor (BoF12I)	50	2×10^{-6}

[1]Only mammalian proteins with a score of 50 or better were used for the phylogeny studies.
[2]The normalized sum of products of the background probabilities and the BLOSUM62 amino acid substitution matrix.
[3]The probability that the amino acid sequences would align by chance.

Fig. 1. Alignment and predicted secondary structure of the uterine serpin family of proteinase inhibitors and α_1 antichymotrypsin and α_1 antitrypsin. The origins of these sequences are given in Table 1. Regions corresponding to the nine α helices and strands of the three β sheets of HuA1AT and HuACT that form the consensus structure of the serpin superfamily are also indicated over their corresponding part of the sequence alignment.

Neither PoUS-1 nor PoUS-2 contained cAMP phosphorylation sites. The presence of these motifs was not unique to the uterine serpins; other serpins examined had at least some of these motifs.

DISCUSSION

These results support an earlier analysis of serpin evolution that included only OvUS and PoUS-1 and indicated that the uterine serpins were very different from other members of the serpin superfamily forming a clade (Marshall, '93). Striking about the uterine serpin family in pigs is the presence of two recently diverging paralogs. Analysis of silent substitutions using a rate constant for silent substitution calibrated from other events in the divergence of uterine serpins suggest that these

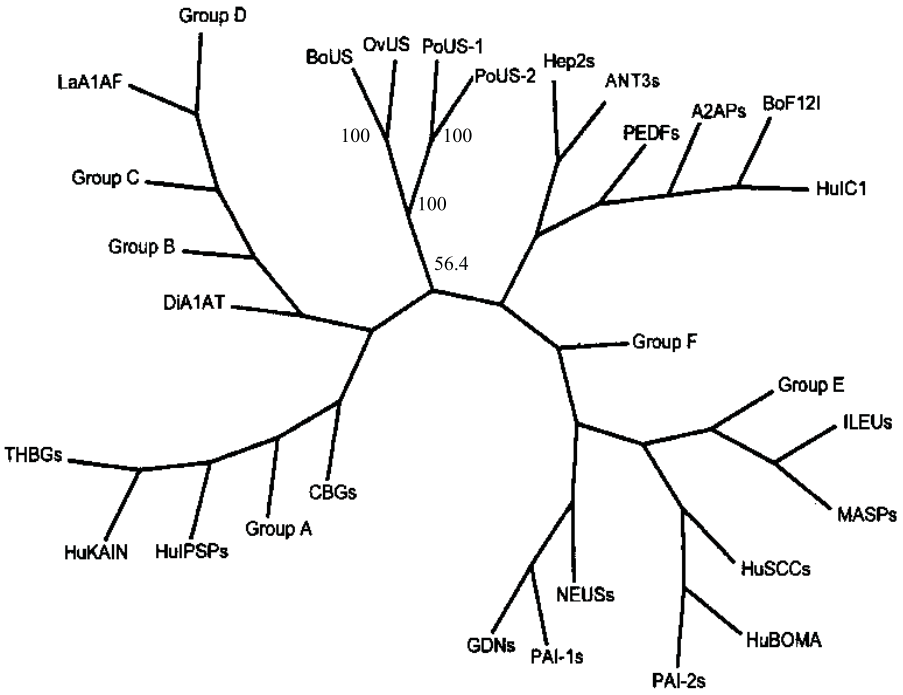

Fig. 2. Parsimony-based phylogeny tree of proteins with homology to the uterine serpins. Some of the clades of the tree were collapsed for clarity of presentation. The contents of these branches were: Group A: HuACT, MuKBP, MuCOTR, RatCPI6, Rat CPI1, Rat CPI6, MuA1A6, RatSI21; Group B: HuA1AU, HuA1AT, PapA1AT, Po A1AT, Bo A1AT, OvA1AT; Group C: RatA1AT, MusA1AT, MuA1A5, MuA1A4, MuA1A2, MuA1A1, MuA1A3: Group D: CaCOTR, CaA1AS, CaA1AF; Group E: HuPTI9. HuPTI8; Group F: RatHsp47; MuHSP47; HuHSP47; HuCBR-2. Numbers at the nodes on the uterine serpin limb refer to the number of bootstrap replicates where the branches were placed at the location.

paralogs diverged very recently, perhaps within the past 5 million years and well after the divergence of pigs from other *Artiodactyla*. To understand the function of these proteins, it would be useful to learn more about the reproductive physiology of *Suid* genera that diverged from the *Sus* genus about 5 Mya. These animals include the bush pig (*Potamochoerus*), the giant forest pig (*Hylochoerus*), the warthog (*Phacochoerus*), and perhaps the babirusa (*Babyrousa*).

The K_a:K_s ratios suggests that the evolution of the uterine serpins may be undergoing changes in biological function during the evolution of the species studied. Similar K_a:K_s values are known for other proteinase inhibitors. A characteristic feature of serpins is their propensity to interact with other proteins. The uterine serpins also bind other proteins—OvUS to the pregnancy-associated glycoproteins and pepsin (Mathialagan and Hansen, '96), IgA, and IgM (Hansen and Newton, '88), and activin (McFarlane et al., '99) and the PoUSs to the iron-containing uteroferrin (Baumbach et al., '86). Given that both the porcine (Baumbach et al., '86) and ovine (Newton et al., '89) uterine serpins cross the placenta, the uterine serpins may have evolved for transplacental transport of proteins or other molecules. The PoUS may accompany uteroferrin during placental transport so that iron can be transferred to transferrin in the fetal compartments (Buhi et al., '82). Also OvUS is present in colostrum (Hansen and Foti, '86) and may function to protect proteins from pepsin digestion in the neonatal gut. The highest K_a:K_s ratio was observed during the evolution of the OvUS sequence. Perhaps the inferred adaptive evolution of OvUS is related to its immunosuppressive activity or the relatively higher secretion rates in the uterus for this protein as compared to BoUS and the PoUSs. Frequently during divergent evolution of proteins, an ancestral sequence is duplicated and the two descendent sequences undergo adaptive evolution whereby each sequence is selected for a different function (Hughes, '94). It is not known whether the porcine paralogs have distinct functions; given the lower K_a:K_s ratios, however, the adaptive evolution of these proteins has been somewhat constrained.

Calibration of the rate of silent substitutions in the uterine serpins was based on the paleontologi-

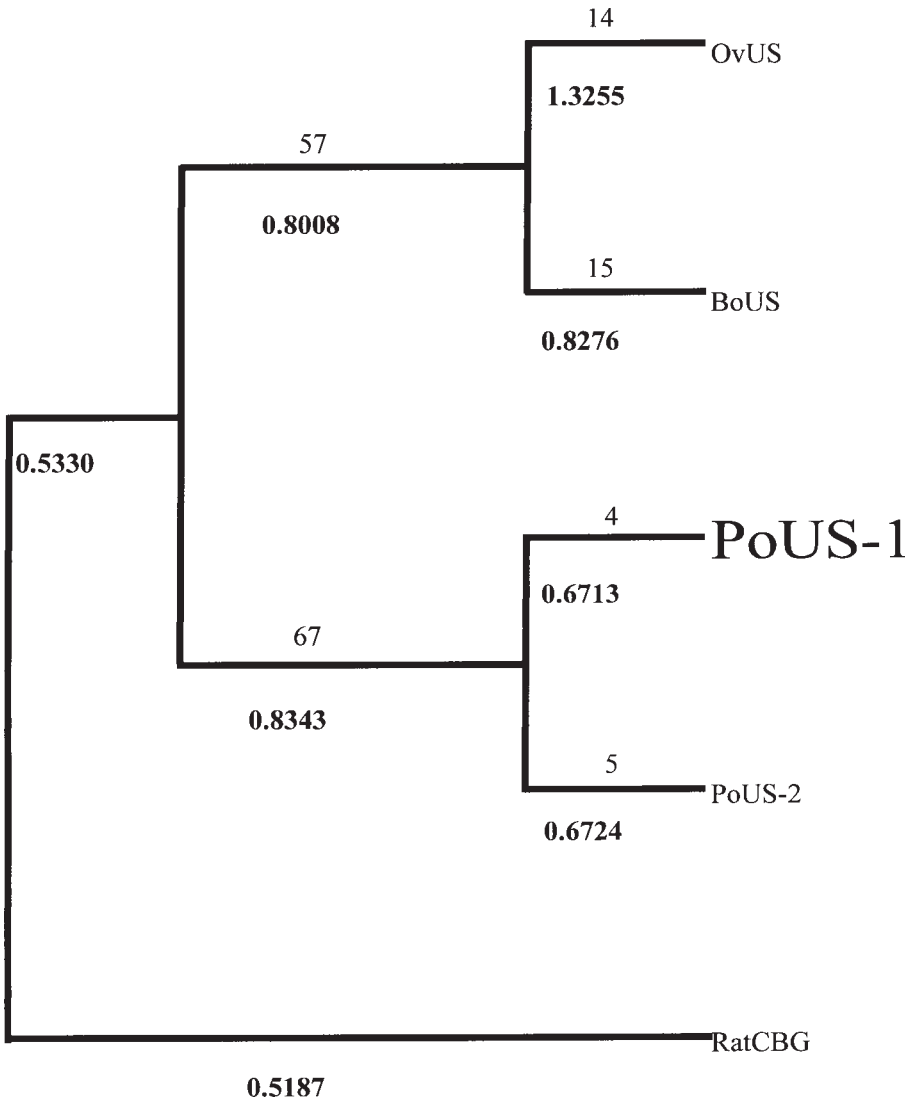

Fig. 3. A distance-based phylogeny of the uterine serpins using NEDs conducted on the uterine serpins with RatCBG defined a priori as the outgroup. Numbers in normal type along the branches refer to the millions of years for evolution to occur along that branch. Shown in bold type on the branches are the K_a:K_s ratios.

cal record, which suggests ruminants and pigs diverged 60 Mya and sheep and cattle diverged 18 Mya (Carroll, '88). If this calibration is accepted then the uterine serpins diverged from other serpin families well before the divergence of artiodactyls and possibly as early as the divergence of mammals from other vertebrates. If so, the members of the uterine serpin clade may be present in members of other mammal orders, including primates.

Sequence analysis revealed that the uterine serpins contain several PKC, tyrosine kinase, and cyclic AMP phosphorylation sites. The OvUS inhibits PKC-induced lymphocyte proliferation (Peltier et al., 2000) and it is conceivable that this action involves competitive inhibition of PKC-regulated proteins. If OvUS can cross the plasma membrane, it could conceivably serve as an alternative target for protein kinase C. Another lymphocyte-inhibitory protein with basic pI, bovine seminal RNAse, does enter the lymphocyte (Mancheno et al., '94; Mastronicola et al., '95). Whether OvUS can cross the plasma membrane and serve as an alternative phosphorylation target for PKC in vivo requires further study.

ACKNOWLEDGMENTS

This manuscript is no. R-07180 of the Florida Agricultural Experiment Station. This work was

TABLE 2. Number of motifs on HuA1AT and artiodactyl serpins with homology to OvUS

Serpin	cAMP[1]	CK-2[2]	PKC[3]	TYR[4]	MYS[5]	ASN[6]	Other
OvUS	2	4	5	1	1	2	0
BoUS	2	4	9	2	2	1	0
PoUS-1	0	5	10	1	2	4	0
PoUS-2	0	5	10	2	3	3	0
PoA1AT	0	5	3	4	1	2	0
OvA1AT	0	6	4	1	8	4	0
BoA1AT	0	7	5	1	6	4	0
OvTHBG	1	2	1	0	6	4	0
OvANT3	1	7	7	0	3	4	0
PoPAI1	1	8	3	0	4	3	0
BoPAI1	1	9	3	0	4	3	0
BoF12I	1	9	10	0	3	4	0
BoPEDF	1	11	5	1	4	1	1[7]
BoA2AP	0	7	6	0	2	5	2[8]

[1]Cyclic AMP phosphorylation sites.
[2]Casein kinase-2 phosphorylation sites.
[3]Protein kinase C phosphorylation sites.
[4]Tyrosine kinase phosphorylation sites.
[5]N-myristoylation sites.
[6]N-linked glycosylation sites.
[7]Gram-positive cocci surface proteins anchoring hexapeptide site.
[8]2 Leucine zipper sites and 1 RGD recognition sequence.

funded in part by the United States Department of Agriculture NRICGP (96-35203) to P.J.H. and the National Institutes of Health (HG 01729 and GM 54075) to S.A.B.

LITERATURE CITED

Altshul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W Lipman DJ. 1997. Gapped BLAST and PSI-Blast: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402.

Bairoch A, Bucher P, Hofman K. 1997. The PROSITE database: its status in 1997. Nucleic Acids Res 25:217–221.

Baumbach GA, Ketcham CM, Richardson DE, Bazer FW, Roberts RM. 1986. Isolation and characterization of a high molecular weight stable pink form of uteroferrin from uterine secretions and allantoic fluid of pigs. J Biol Chem 261: 12869–12878.

Benner SA, Trabesinger N, Schreiber D. 1998. Post-genomic science: converting primary structure into physiological function. Adv Enzyme Reg 38:155–180.

Buhi WC, Ducsay CA, Bazer FW, Roberts RM. 1982. Iron transfer between the purple phosphatase uteroferrin and transferrin and its possible role in iron metabolism of the fetal pig. J Biol Chem 257:1712–1723.

Carroll RL. 1988. Vertebrate paleontology and evolution. New York: W.H. Freeman.

Endo T, Ikeo K, Gojobori T. 1996. Large-scale search for genes on which positive selection may operate. Mol Biol Evol 13:685–690.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791.

Felsenstein J. 1989. PHYLIP-Phylogeny interface package (version 3.2). Cladistics 5:164–166.

Hansen PJ. 1998. Regulation of uterine immune function by progesterone-lessons from the sheep. J Reprod Immunol 40:63–79.

Hansen PJ, Foti SA. 1986. Proteins similar to the major proteins in uterine secretions of pregnant ewes are also found in colostrum. J Anim Sci 63(Suppl 1):53–54 (abstr).

Hansen PJ, Newton GR. 1988. Binding of immunoglobulins to the major progesterone-induced proteins secreted by the sheep uterus. Arch Biochem Biophys 260:208–217.

Hughes A. 1994. The evolution of functionally novel proteins after gene duplication. Philos Trans Roy Soc Lond B 256:119–124.

Ina Y. 1995. New methods for estimating the numbers of synonymous and nonsynonymous substitutions. J Mol Evol 40:190–226.

Ina Y. 1996. Pattern of synonymous and nonsynonymous substitutions: an indicator of mechanisms of molecular evolution. J Genetics 75:91–115.

Ing NH, Roberts RM. 1989. The major progesterone-induced proteins secreted into the sheep uterus are members of the serpin superfamily of protease inhibitors. J Biol Chem 264:3372–3379.

Ing NH, Francis H, McDonnell JJ, Amann JF, Roberts RM. 1989. Progesterone induction of the uterine milk proteins: major secretory proteins of sheep endometrium. Biol Reprod 41:643–654.

Laskowski M, Kato I, Kohr WJ, Park SJ, Tashiro M, Whatley HE. 1987. Positive Darwinian selection in evolution of protein inhibitors of serine proteinases. Cold Spring Harbor Symp Q Biol 52:545–553.

Leslie MV, Hansen PJ. 1991. Progesterone-regulated secretion of the serpin-like proteins of the ovine and bovine uterus. Steroids 56:589–597.

Li MD, Ford JJ. 1998. A comprehensive evolutionary analysis based on nucleotide and amino acid sequences of the α and β-subunits of glycoprotein hormone gene family. J Endocrinol 156:529–542.

Li W-H, Wu C-I, Lou C-C. 1985. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. Mol Biol Evol 2:150–174.

Malathy PV, Imakawa K, Simmen RC, Roberts RM. 1990.

Molecular cloning of the uteroferrin-associated protein, a major progesterone-induced serpin secreted by the porcine uterus, and the expression of its mRNA during pregnancy. Mol Endocrinol 4:428–440.

Mancheno JM, Gasset M, Onaderra M, Gavilanes JG, D'Alessio G. 1994. Bovine seminal ribonuclease destabilizes negatively charged membranes. Biochem Biophys Res Commun 199:119–124.

Marshall CJ. 1993. Evolutionary relationships among the serpins. Philos Trans Roy Soc Lond B 342:101–119.

Mastronicola MR, Piccoli R, D'Alessio G. 1995. Key extracellular and intracellular steps in the antitumor action of seminal ribonuclease. Eur J Biochem 230:242–249.

Mathialagan N, Hansen TR. 1996. Pepsin-inhibitory activity of the uterine serpins. Proc Natl Acad Sci USA 93:13653–13658.

McFarlane JR, Foulds LM, O'Connor AE, Phillips DJ, Jenkin G, Hearn MT, de Kretser DM. 1999. Uterine milk protein, a novel activin-binding protein, is present in ovine allantoic fluid. Endocrinology 140: 4745–4752.

Messier W, Stewart CB. 1997. Episodic adaptive evolution of primate lysozyme. Nature 385:155–180.

Moffatt J, Bazer FW, Hansen PJ, Chun PW, Roberts RM. 1987. Purification, secretion, and immunocytochemical localization of the uterine milk proteins, major progesterone-induced proteins in uterine secretions of the sheep. Biol Reprod 36:419–430.

Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.

Newton GR, Hansen PJ, Bazer FW, Leslie MV, Stephenson DC, Low BG. 1989. Presence of the major progesterone-induced proteins of the sheep endometrium in fetal fluids. Biol Reprod 40:417–424.

Pamilo P, Bianchi NO. 1993. Evolution of the Zfx and Zfy genes: rates and interdependence between genes. Mol Biol Evol 10:271–281.

Peltier MR. 2000. Immunoregulatory activity, biochemistry, and phylogeny of the ovine uterine serpin. Gainesville, Florida: University of Florida. Thesis.

Peltier MR, Liu W-J, Hansen PJ. 2000. Regulation of lymphocyte proliferation by uterine serpin: interleukin-2 mRNA production, CD25 expression and responsiveness to interleukin-2. Proc Soc Exp Biol Med 223:75–81.

Smith CL, Hammond GL 1989. Rat cortisosteroid binding globulin: primary structure and messenger ribonucleic acid levels in the liver under different physiological conditions. Mol Endocrinol 3:420–426.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680.

Trabesinger-Ruef N, Jermann T, Zankel T, Durrant B, Frank G, Benner SA. 1996. Pseudogenes in ribonuclease evolution: a source of new biomacromolecular function? FEBS Lett 382:319–322.

Xie S, Low BG, Nagel RJ, Kramer KK, Anthony RV, Zoli AP, Beckers JF, Roberts RM. 1991. Identification of the major pregnancy-specific antigens of cattle and sheep as inactive members of the aspartic proteinase family. Proc Natl Acad Sci USA 88:10247–10251.

Xie S, Low BG, Nagel RJ, Beckers JF, Roberts RM. 1994. A novel glycoprotein of the aspartic proteinase gene family expressed in bovine placental trophectoderm. Biol Reprod 51:1145–1153.

Xie S, Green J, Beckers JF, Roberts RM. 1995. The gene encoding bovine pregnancy-associated glycoprotein-1, an inactive member of the aspartic proteinase family. Gene 159:193–197.

Xie S, Green J, Bixby JB, Szafranska B, DeMartini JC, Hecht S, Roberts RM. 1997. The diversity and evolutionary relationships of the pregnancy-associated glycoproteins, an aspartic proteinase subfamily consisting of many trophoblast-expressed genes. Proc Natl Acad Sci USA 94:12809–12816.