Characterization of Two Forms of Mouse Salivary Androgen-Binding Protein (ABP): Implications for Evolutionary Relationships and Ligand-Binding Function † , ‡

ARTICLE *in* BIOCHEMISTRY · JULY 2003

Impact Factor: 3.02 · DOI: 10.1021/bi027424l · Source: PubMed

2 AUTHORS:

Robert Karn
The University of Arizona
86 PUBLICATIONS **1,560** CITATIONS
SEE PROFILE

Christina Laukaitis
The University of Arizona
32 PUBLICATIONS **784** CITATIONS
SEE PROFILE

Characterization of Two Forms of Mouse Salivary Androgen-Binding Protein (ABP): Implications for Evolutionary Relationships and Ligand-Binding Function[†,‡]

Robert C. Karn* and Christina M. Laukaitis

Department of Biological Sciences, Butler University, Indianapolis, Indiana 46208

Received December 26, 2002; Revised Manuscript Received April 15, 2003

ABSTRACT: Mouse salivary androgen-binding protein (ABP) is a member of the secretoglobin family produced in the submaxillary glands of house mice (*Mus musculus*). We report the cDNA sequences and amino acid sequences of the β and γ subunits of ABP from a mouse cDNA library, identifying the two subunits by their p*I*s and molecular weights. An anomalously high molecular weight of the α subunit is likely due to glycosylation at a single site. A phylogenetic comparison of the three subunits of ABP with the chains of other mammalian secretoglobins shows that ABP is most closely related to mouse lachrymal protein and to the major cat allergen Fel dI. An evaluation of the most conserved residues in ABP and the other secretoglobins, in light of structural data reported by others [Callebaut, I., Poupon, A., Bally, R., Demaret, J.-P., Housset, D., Delettre, J., Hossenlopp, P., and Mornon, J.-P. (2000) *Ann. N.Y. Acad. Sci. 923*, 90−112; Pattabiraman, N., Matthews, J., Ward, K., Mantile-Selvaggi, G., Miele, L., and Mukherjee, A. (2000) *Ann. N.Y. Acad. Sci. 923*, 113−127], allows us to draw conclusions about the critical residues important in ligand binding by the two different ABP dimers and to assess the importance of ligand binding in the function of the molecule. In addition to the cDNAs, which represent those of the *musculus* subspecies of *Mus musculus*, we also report the coding regions of the β and γ subunit cDNAs from two other mouse inbred strains which represent the other two subspecies: *M. musculus domesticus* and *M. musculus castaneus*. The high nonsynonymous/synonymous substitution rate ratios (K_a/K_s) for both the β and γ subunits suggest that these two proteins are evolving under strong directional selection, as has been reported for the α subunit [Hwang, J., Hofstetter, J., Bonhomme, F., and Karn, R. (1997) *J. Hered. 88*, 93−97; Karn, R., and Clements, M. (1999) *Biochem. Genet. 37*, 187−199].

The relationship of structure and function is a central theme of biology. Mammalian salivary proteins come in a variety of different forms with diverse functions. In some cases a salivary protein's function has yielded to biochemical studies. For example, salivary amylase digests dietary starches, glycoproteins lubricate hard and soft tissues, and some peptides retard the growth of bacteria and fungi in the oral cavity (see ref *5* for a review). For other salivary proteins, however, the function has remained elusive. Mouse salivary androgen-binding protein (ABP)[1] is one of those. This small globular protein is composed of two different disulfide-bridged subunits and, as its name implies, binds male sex steroid hormones (*6*, *7*).

Mouse salivary ABP is a member of the newly erected family of proteins called secretoglobins [see the report of the nomenclature committee (*8*)]. This interesting group has been described thus far only in mammals (*9*), where they can be found in a wide variety of tissues and secretions (reviewed in ref *10*). The earliest member of this family to be identified originally went by the names blastokinin (*11*) and uteroglobin (*12*). Subsequently, this protein was shown to be identical to another protein under intensive study, lung clara cell protein, and the molecule is now often referred to as uteroglobin/clara. A variety of functions have been ascribed to uteroglobin/clara; however, its true biological role remains elusive (*13*). The same is true of the other members of the secretoglobin family. Although a great deal of effort has been put into studying them, none has yet been unequivocally associated with a specific function (*14*). One important consequence is that there are few clues to shared functional aspects of these proteins; they are currently assigned to this family only on the basis of structural similarities (*1*). Nonetheless, it appears that secretoglobins may be grouped roughly by their mode of secretion. Some, such as uteroglobin/clara and rat prostatic steroid-binding protein (PBP; which also goes by several other names in the literature), are secreted in internal organs (*10*). Others, such as cat Fel dI, the hamster Harderian gland proteins, and mouse ABP, are secreted by exocrine glands and find their way almost immediately after secretion to the animal's exterior as the result of licking and grooming (*15*−*17*).

Over the past decade, this laboratory has focused its efforts on determining the function of mouse salivary ABP. We have

[†] This work was supported in part by the Holcomb Research Institute, Butler University.

[‡] The nucleotide sequences reported in this paper have been submitted to GenBank with accession numbers AY293278, AY293279, AY293280, AY293281, AY293282, and AY293283.

* To whom correspondence and reprint requests should be addressed. Tel: 317-940-8301. Fax: 317-940-9519. E-mail: rkarn@butler.edu.

[1] Abbreviations: ABP, mouse salivary androgen-binding protein; CC, lung clara cell protein; Fel dI, feline allergen; HGB, hamster heteroglobin; IEF, isoelectric focusing; LPP, lipophilin; MGB, mammaglobin; PBP, prostatic steroid-binding protein; RT-PCR, reverse transcription−polymerase chain reaction; Utero, uteroglobin; DHT, dihydrotestosterone.

10.1021/bi027424l CCC: $25.00
Published on Web 05/22/2003

pursued several lines of investigation, including biochemical comparisons (*4*, *6*, *7*, *18*−*20*), population genetic studies (*3*, *21*−*23*), and direct tests of the proposal that the molecule is a recognition signal involved in isolation of different gene pools (*17*, *24*), called subspecies by some (*25*) and species by others (*26*). It was our goal in the studies reported here to answer a number of fundamental questions about the genetics, biochemistry, and evolution of ABP. Our purpose is to better understand ABP in the context of the secretoglobin family and to draw direct comparisons to the hamster heteroglobins, molecules that also may be implicated in a behavioral role (P. Dominguez, personal communication; *27*).

One of the most important genetic questions we have asked is, how many genes are responsible for encoding ABP? The protein is secreted into saliva in two different dimeric forms: an α subunit disulfide-bridged to either a β (the A·B dimer) or a γ (the A·G dimer) subunit. On the basis of genetic evidence, the three subunits were assigned separate genes, named *Abpa*, *Abpb*, and *Abpg*, respectively, but that proposal was formulated and supported only by indirect evidence (*28*). Until this report, the only subunit with a known sequence was the α subunit, the one common to all forms of ABP observed to date (*7*). We wished to determine the amino acid sequences of the other two subunits in order to differentiate among several possibilities: (1) that the three subunits each have a separate gene or (2) that two or more of the subunits could share a transcriptional element with alternative splicing patterns for one or more exons or (3) that two or more of the subunits are differentiated only by posttranslational modifications. The results from this part of the study also allowed us to make an evolutionary comparison of the ABP subunits with each other and with the other secretoglobins, which provides additional insights about groups of proteins within this family that may have more closely related functions.

The evolutionary question of most interest is whether the strong selection that has apparently driven the microevolution of the coding region of *Abpa* (*3*, *22*, *23*) has also operated on the coding regions of *Abpb* and *Abpg*. To further study the extent to which selection may have been involved in the evolution of ABP, we examined the extent of nucleotide polymorphism between the β and γ subunit cDNAs and the predicted amino acid replacements in the subunit proteins of each of the three subspecies. The claim that selection has acted directly on *Abpa*, rather than a closely linked gene, finds its strongest support in the significant deviation of the nonsynonymous/synonymous nucleotide substitution rate ratios (K_a/K_s) from predictions of the neutral hypothesis (*3*, *22*), and it was intriguing to ask whether that extends to the other two subunit genes.

The biochemical question that we wished to answer involved the anomalously high molecular weight of the α subunit as determined by physical−chemical means (*6*) compared to the stoichiometric molecular weight determined from its amino acid sequence (*7*). We investigated the possibility that nonproteinaceous moieties might be contributing to the molecular weight of one or more of the three subunits. This also provided the opportunity for a more thorough comparison of the structures of the ABP subunits to those of other modified secretoglobins such as the glycosylated heteroglobins (*9*).

MATERIALS AND METHODS

Materials. Inbred strains of mice, C3H/HeJ, DBA/2J, and CAST/Ei, were purchased from Jackson Laboratory. A DBA/2J submaxillary cDNA library was provided by Ken Gross (*7*). Salivas (C3H/HeJ and DBA/2J) and submaxillary glands (C3H/HeJ and CAST/Ei) were obtained as previously described (*18*).

Protein Chemistry Techniques. Purification of ABP from mouse saliva was performed as described (*7*). Peptides were produced by digestion of the ABP dimer with trypsin at 37 °C (*7*), and the peptides were separated by HPLC in the Laboratory for Macromolecular Structure at Purdue University. Molecular weights were determined by analysis in a mass spectrometer in the Laboratory for Macromolecular Structure at Purdue University. Selected peptides were sequenced in the Laboratory for Macromolecular Structure at Purdue University (*7*).

Molecular Biology Techniques. Submaxillary glands were removed from mice as previously described (*3*). RNA was purified from mouse submaxillary glands with a Sigma total RNA purification kit (St. Louis, MO). cDNA libraries were constructed and RT-PCR products were obtained from mouse submaxillary gland RNA essentially as described in ref *3* using Sigma M-MLV reverse transcriptase for first strand cDNA synthesis and Sigma Taq polymerase for subsequent amplification. Fragments were manually sequenced as previously described using a Circumvent cycle sequencing kit [NEB, Beverly, MA (*3*)] or using the Amersham Thermo Sequenase cycle sequencing kit (Piscataway, NJ).

Data Analysis. Sequences were aligned with the DNASIS program (Hitachi, Yokohama, Japan) and, for the purpose of constructing a phylogenetic tree, with CLUSTAL (*29*, *30*). Predictions of protein secondary structures, stoichiometric molecular weights, and isoelectric points were made with the PROSIS program (Hitachi). A phylogenetic tree was calculated by neighbor-joining with 10000 bootstrap replicates, using the program PAUP* (*31*). The tree was printed using TreeView (*32*). The program MEGA, version 2.1 (*33*), was used to calculate numbers of synonymous and nonsynonymous substitutions, as well as their rates (K_a and K_s), using the modified method of Nei and Gojobori (*34*), with Jukes−Cantor correction for multiple hits and a transition/transversion rate of 2.

RESULTS

The cDNA Sequences and Amino Acid Sequences of the β and γ Subunits of ABP. Until now, only the amino acid sequence of the α subunit of ABP and partial sequences for the dimer had been determined (*7*). To obtain a cDNA for the β subunit, we designed a mixed oligomer probe from the partial sequence of the putative β subunit and used it to probe a mouse submaxillary gland cDNA library. A candidate cDNA was obtained and its nucleotide sequence determined (Figure 1). The putative amino acid sequence determined by translation of the cDNA and the presence of a polyadenylation signal near the 3′ end suggested that the cDNA was complete. Cleavage of the signal peptide at the third Ala would yield a secreted protein of 89 amino acids, beginning with a Cys, as noted above from the partial sequence of the HPLC-purified chain, and otherwise in agreement with the earlier sequences. When a sample of the

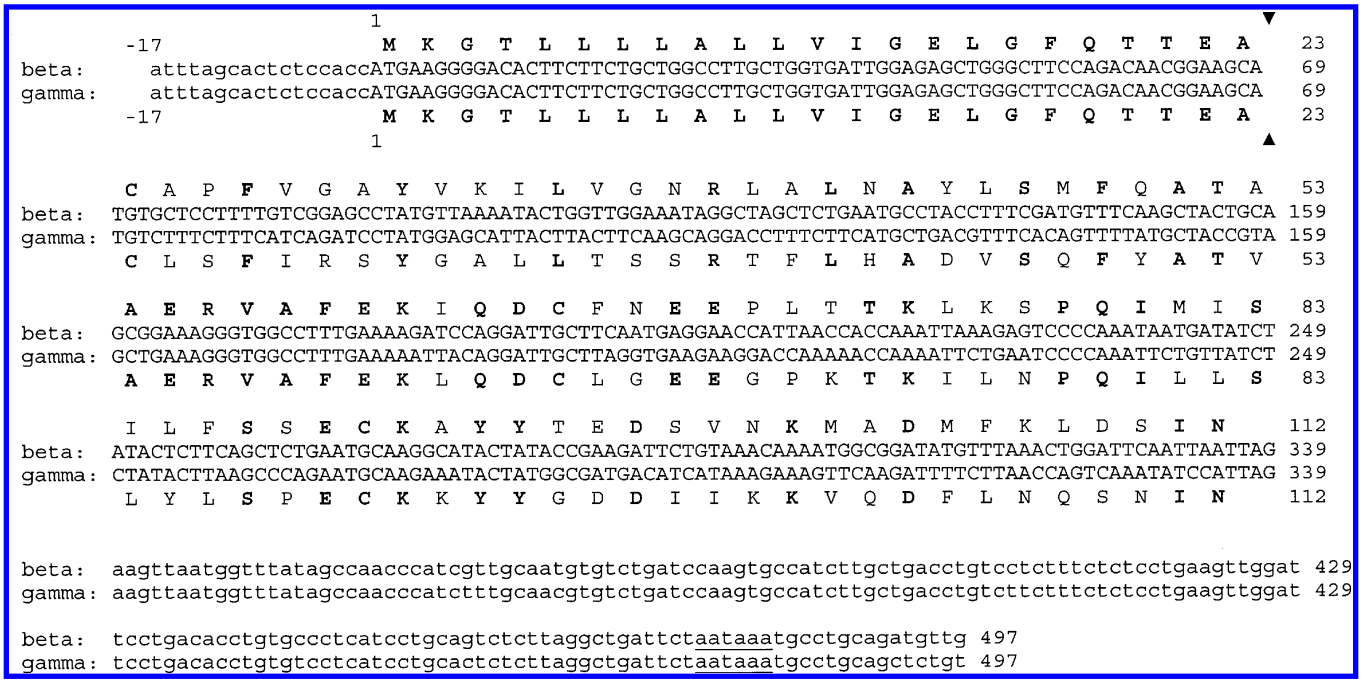

FIGURE 1: cDNAs of the β and γ subunits of mouse salivary androgen-binding protein (ABP). The amino acids in common between the two subunits are in boldface type. The cleavage points for the removal of the signal peptides are indicated by the black triangles above and below the sequences.

reduced, alkylated protein was chromatographed by high-performance liquid chromatography (HPLC), a peak was recovered with the N-terminal sequence of the β subunit but which had a carboxymethyl-Cys as the first residue (not shown), rather than the Asp reported earlier (*7*). It is likely that the first residue observed by Karn and Russell (*7*) was an oxidation product of the N-terminal Cys of this subunit.

Purified ABP was subjected to digestion with trypsin. The dimer proved to be resistant to a single treatment of 2% w/w trypsin at 37 °C overnight, even in the presence of 2 M urea. Digestion was only obtained when trypsin was added in increments of 0.5% w/w every hour for at least 4 h at 37 °C in the presence of 2 M urea, the last addition being followed by overnight incubation at 37 °C. This suggests that ABP secreted in mouse saliva may remain stable for long periods of time in the animal's environment, as suggested by behavioral studies (*17*).

Subsequent isolation and sequencing of peptides from an HPLC separation of trypsin-digested ABP verified all of the peptides predicted for the secreted protein from the β subunit cDNA sequence (not shown). In addition, a unique peptide with the sequence TF**L**H**A**DV**S**Q**F**Y**AT**V**AER** was obtained. This peptide corresponded to the peptide LA**L**N**A**YL**S**M-**F**Q**AT**A**AER** in the β subunit (the common residues appear in bold type; 10/17 = 59% identity). A mixed oligomer probe was designed from this peptide and used to probe the mouse submaxillary gland cDNA library. As before, a candidate cDNA was obtained and its nucleotide sequence determined, yielding a complete cDNA including the 5′ untranslated sequence (5′ UTR) and a polyadenylation signal near the 3′ end. Figure 1 shows the sequences of the β and γ subunits with their identical amino acids in bold type. These two ABP subunits are more similar to each other (39%) than either is to the α subunit (17% and 13%, respectively). They are not, however, nearly identical to each other in their first 18 residues as previously reported (*4*).

Identification of β and γ Subunits. The identity of the α subunit cDNA was easily established in previous studies because it is the only one of the three that is common to all forms of ABP observed and because it has a higher apparent molecular weight than either of the other two (*4*, *6*, *28*). Since the β and γ subunits are the variable subunits in the two forms of the dimer, another objective means was needed to identify one from the other. Isoelectric focusing (IEF) in two-dimensional gel electrophoresis (*28*) showed that the γ subunit has a much higher p*I* than does the β subunit, whereas the p*I* of the α subunit is intermediate between these. Estimates of p*I* clearly distinguish the putative proteins produced from translating the cDNAs shown in Figure 1. The β subunit has a predicted p*I* of 5.13, and the γ has a predicted p*I* of 7.06. The predicted p*I* of the α subunit is 6.48. As shown below, genetic polymorphism also provides an identification of these two subunits consistent with that based on p*I*.

Molecular Weights of the ABP Dimer and Its Subunits. The molecular weights of the ABP dimer and its subunits were obtained by mass spectrometric analysis of the proteins separated by HPLC. Table 1 shows the molecular weight values in comparison with the stoichiometric molecular weights determined from their amino acid sequences and with the estimates originally derived from physical−chemical analysis (*6*). The molecular weight of the γ subunit obtained from a mixture of reduced/alkylated dimers is in good agreement with that calculated from the amino acid sequence shown in Figure 1. Apparently β was not recovered in the HPLC separation since it should have been resolved from the γ in the mass spectrometric analysis. While the value obtained by mass spectrometric analysis of the α subunit is relatively close to the estimate from physical−chemical methods, it is higher than the stoichiometric value by approximately 3000 Da. This suggests that a nonproteinaceous moiety is attached to that subunit as a posttranslational modification.

Table 1: Molecular Weights of the ABP Dimer and Its Subunits Obtained by Mass Spectrometer Analysis and Stoichiometric Molecular Weights Determined from the Amino Acid Sequences of the Subunits

	ABP dimer	α subunit	variable subunits (β and/or γ)	sum of the subunits
SDS electrophoresis (*6*)	16700 (±300)	12000 (±425)[a]	8800 (±800)	20800[a]
mass spectrometer analysis	20923	11006[a]	10105	21111[a]
stoichiometric values				
α		10841[a]/7882[b]		
β			9997	
γ			10132	
α−β dimer				20838[a]
α−γ dimer				20973[a]

[a] Includes carbohydrate moiety. [b] Without carbohydrate moiety.

Since the amino acid sequence (*7*) of the α subunit contains a glycosylation consensus sequence (N15 G16 T17), the most likely explanation of the molecular weight discrepancy is that the protein contains one asparagine-linked oligosaccharide moiety with a mass of approximately 3000 Da (*35*). This is consistent with the inability to identify the PTH-amino acid residue in cycle 15 of the sequence of the alkylated α chain (*7*). To explore this possibility, a sample of the α subunit (C3H strain) was digested with trypsin, and the peptides were separated by HPLC. One of the peptides recovered from HPLC had the amino acid sequence KVDLFLXGTTEEYVEYLK (where X denotes a cycle whose PTH-amino acid could not be determined), which represents residues 9−26 of the α chain (*7*). This peptide, which unmodified would have a mass of 2161 Da, yielded a mass of 5094 Da from mass spectrometric analysis. The difference of nearly 3000 Da is likely due to glycosylation at N15, which probably interfered with determination of this residue in the peptide.

Relationship of the ABP Subunits to Other Secretoglobins. Previous studies showed that the α subunit of mouse salivary ABP shared significant identity with a number of secretoglobin subunits, especially with chain 1 of the major cat allergen Fel dI (*10*, *16*, *19*). Figure 2, panel A, shows a comparison of the three ABP subunits with each other and with representative members of other groups of secretoglobins. The alignment of the first 67 amino acids of both the β and γ subunits of ABP with the sequences of the other uteroglobin subunits strongly suggests that they are members of this family. As is the case with the secreted forms of mouse lachrymal protein, rat PBP, and cat Fel dI, the secreted forms of mouse ABP β and γ subunits begin with an N-terminal Cys that corresponds to the third amino acid in the sequence of their partner (α in the case of ABP) in the dimer. It is also interesting that the same six subunits are appreciably longer at the C-terminus than the others by 14−17 residues, and ABP β and γ are the longest of the six.

A phylogenetic tree constructed from the alignment shown in Figure 2 is illustrated in Figure 3. The uteroglobin/clara cluster was used as the outgroup in constructing the tree because of two important differences between them and the other secretoglobins: the lack of a middle Cys in this group in the uteroglobin/clara cluster and their formation of homo- rather than heterodimers. Bootstrap values for the nodes of the tree are generally high, especially at nodes that define five groups, labeled A−E. These five groups are consistent with the major groups described by Ni et al. (*36*) for a similar collection of secretoglobin subunits, except that the branches leading to D and E are swapped, probably because they included the signal peptides in their comparison. Our new sequences for ABP β and γ, along with the addition of the 11 kDa protein from mouse lachrymal gland, expand group E, which only contained chain 2 of the cat allergen Fel dI in the previous analysis (*36*). Since the α subunit groups with chain 1 of Fel dI and the other mouse lachrymal protein in cluster B, it is apparent that mouse salivary ABP, mouse lachrymal proteins, and feline allergen Fel dI are evolutionarily quite closely related.

Figure 2, panel B, also shows an unexpected degree of identity between a portion of the sequences of the ABP subunits and residues 117−162 of the ATPase sequences of several representative vertebrates. β shows 30% identity to this region of the ATPase sequence shared by dog, sheep, pig, and horse, and γ is 17% identical to this region of ATPase. This region of ATPase forms a transmembrane helix (residues 130−150) and part of a loop (residues 117−129) connecting two such helices but has not been assigned a specific function (K. Axelsen, personal communication). Since ABP is a soluble protein secreted in mouse saliva, there is no reason to believe that these regions of the ABP subunits span membranes.

Genetic Polymorphism of the β and γ Subunits. We were interested in looking for polymorphism in the secreted forms of the β and γ subunits corresponding to the a, b, and c variants of the α subunit that are monomorphic in the subspecies of *Mus musculus* [*M. musculus domesticus*, *M. musculus musculus*, and *M. musculus castaneus*, respectively (*3*, *22*, *23*)]. Because the ABP phenotypes of these three wild subspecies are found in the three inbred mouse strains, C3H, DBA, and CAST, respectively (*3*), these three strains were investigated for β and γ variation.

Comparing the ABP subunit sequences among three subspecies of *M. musculus*, the β subunit is the most conserved of the three (Figure 4). The β subunits of the *M. musculus domesticus* and *M. musculus castaneus* subspecies have identical amino acid sequences, while the *M. musculus musculus* β sequence differs from them by three amino acids. The predicted p*I*s are identical for all three β sequences, consistent with the inability to identify electrophoretic variants that could be ascribed to the β subunit (*28*). By contrast, the γ subunit of ABP is the least conserved of the three subunits (Figure 4). It is similar in the *musculus* and *castaneus* subspecies (96% identical) but very different in the *domesticus* subspecies (81% identical between *musculus* and *domesticus*; 80% identical between *castaneus* and *domesticus*). The predicted p*I* of the *domesticus* γ (8.49) is

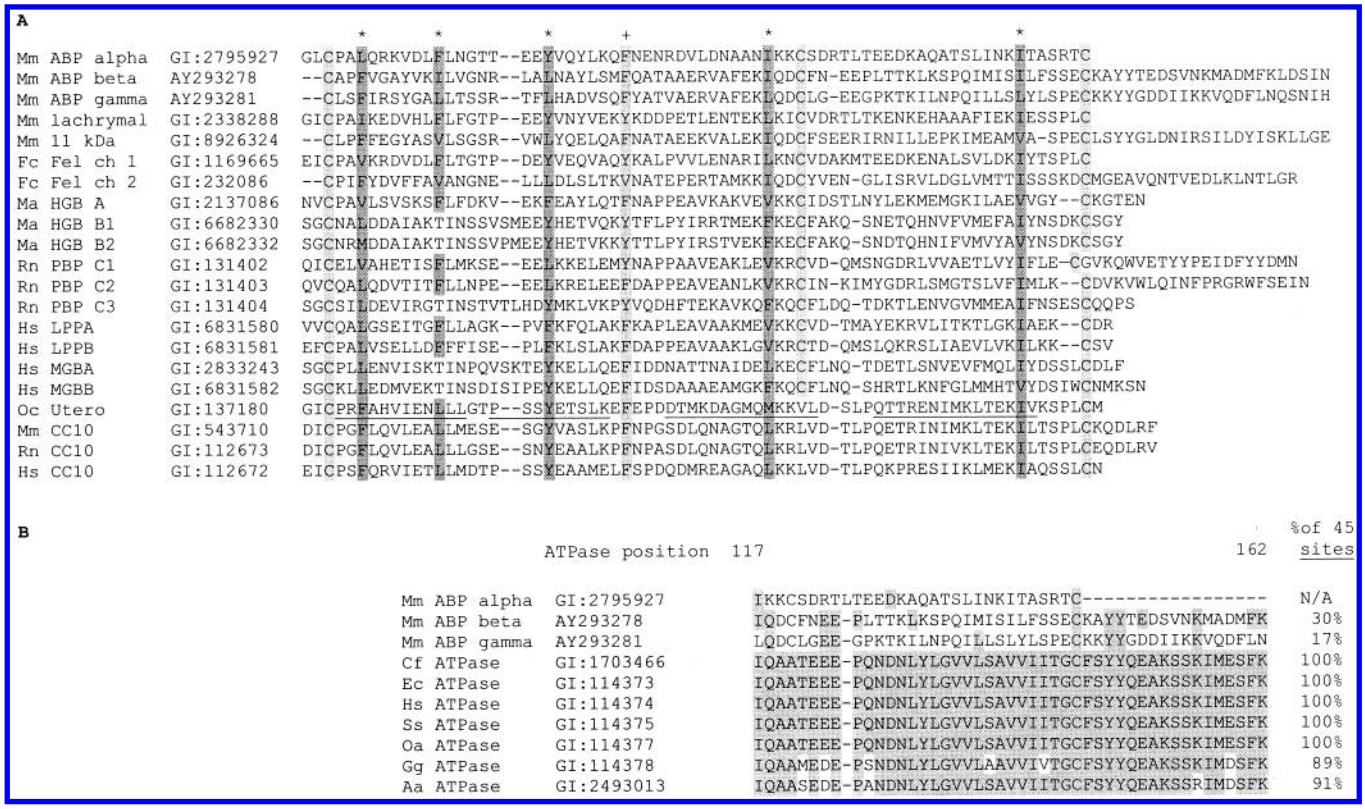

FIGURE 2: (A) Multiple alignment of the sequences of the secreted β and γ subunits of mouse salivary androgen-binding protein (ABP) with other secretoglobins. Signal peptides are not shown. The GenBank identifier or accession number (GI) follows the name of each protein. The helical regions experimentally determined for rabbit uteroglobin are underlined in that sequence. The most conserved positions among the sequences compared are indicated by asterisks, and the critical F28 (see text) is indicated by a plus. The cysteines and the other conserved residues are shaded. Names: ABP, mouse salivary androgen-binding protein; lachrymal, mouse allergen-like lachrymal protein; 11 kDa, mouse lachrymal 11 kDa secretory protein; Fel, feline allergen; HGB, hamster heteroglobin (chains A, B1, and B2 are represented); PBP, rat prostatic steroid-binding protein (chains C1, C2, and C3 are represented); LPPA, lipophilin A; LPPB, lipophilin B; MGBA, mammaglobin A; MGBB, mammaglobin B; Utero, uteroglobin; CC10, clara cell 10 kDa secretory protein. Species: Mm, *Mus musculus* (house mouse), Fc, *Felis catus* (domestic cat); Ma, *Mesocricetus auratus* (hamster); Rn, *Rattus norvegicus* (rat); Hs, *Homo sapiens*; Oc, *Oryctolagus cuniculus* (domestic rabbit). (B) Multiple alignment of the sequences of the ATPase-like regions of ABP α, β, and γ subunits with the sequences of various ATPases. Species: as above and Cf, *Canis familiaris* (dog); Ec, *Equus caballus* (horse); Ss, *Sus scrofa* (pig); Oa, *Ovis aries* (sheep); Gg, *Gallus gallus* (chicken); Aa, *Anguilla anguilla* (eel).

radically different from those of the other two (7.06 for *musculus* and 7.15 for *castaneus*) because many of the substitutions cause charge changes on the protein.

The lack of p*I* variation in the β subunit and the marked p*I* difference between the *musculus* and *domesticus* versions of γ are another means by which the two subunits, β and γ, can be identified (see above). The relative differences are consistent with predictions from the IEF dimension of two-dimensional separations (*28*). Three nucleotide sites were polymporhic between the β subunits of C3H and DBA, resulting in three neutral amino acid replacements. The predicted p*I* for the C3H β subunit is 5.13, identical to that of the DBA β. This is consistent with the observations of Dlouhy et al. (*28*) that the β subunits did not differ in charge between these two strains. Thus the β subunit contains polymorphism hidden from procedures that separate molecular species on the basis of their charges.

In the region coding for the secreted protein, 28 nucleotide sites were polymorphic between the γ subunits of C3H and DBA, resulting in 17 amino acid replacements, 4 of which changed the charge on the protein. The predicted p*I* for the C3H γ subunit is 8.49, compared to 7.06 for the DBA γ. This is also consistent with the observations of Dlouhy et al. (*28*) that the γ subunits differed in charge between these two strains, conferring different isoelectric points on the dimers with any given α subunit. In this same region, there were 27 sites that differed between the CAST γ and the C3H γ and 5 sites that differed between the CAST γ and the DBA γ.

Nonsynonymous/synonymous substitution rate ratios (K_a/K_s) were calculated from the data in Figure 4 and are shown in comparison with those calculated from the data of Hwang et al. (*3*) in Table 2. The lowest K_a/K_s we obtained for the β and γ subunit genes (*Abpb* and *Abpg*) was that for the *domesticus*−*castaneus* comparison of *Abpg* (0.82) and was comparable to the value for the *musculus*−*castaneus* comparison of *Abpa*. As observed for *Abpa* by Karn and Nachman (*22*), these values stand in contrast to an average K_a/K_s value of 0.143 for 363 homologous genes (*37*). In the comparison made by Wolfe and Sharp (*37*) only one revealed $K_a/K_s > 1$ (for interleukin-3, $K_a/K_s = 1.12$). Thus, as is the case for *Abpa*, both *Abpb* and *Abpg* have K_a/K_s well above average.

We were interested in the impact of polymorphism on the conserved residues (F6, L13, Y21, F28, M41, and I63; numbers from the uteroglobin sequence) shown in Figure 2 and the predicted secondary structure of the secretoglobins, which is based on that of uteroglobin/clara. Figure 5 compares the secreted forms of all three ABP subunits, showing the amino acid substitutions, the likely positions

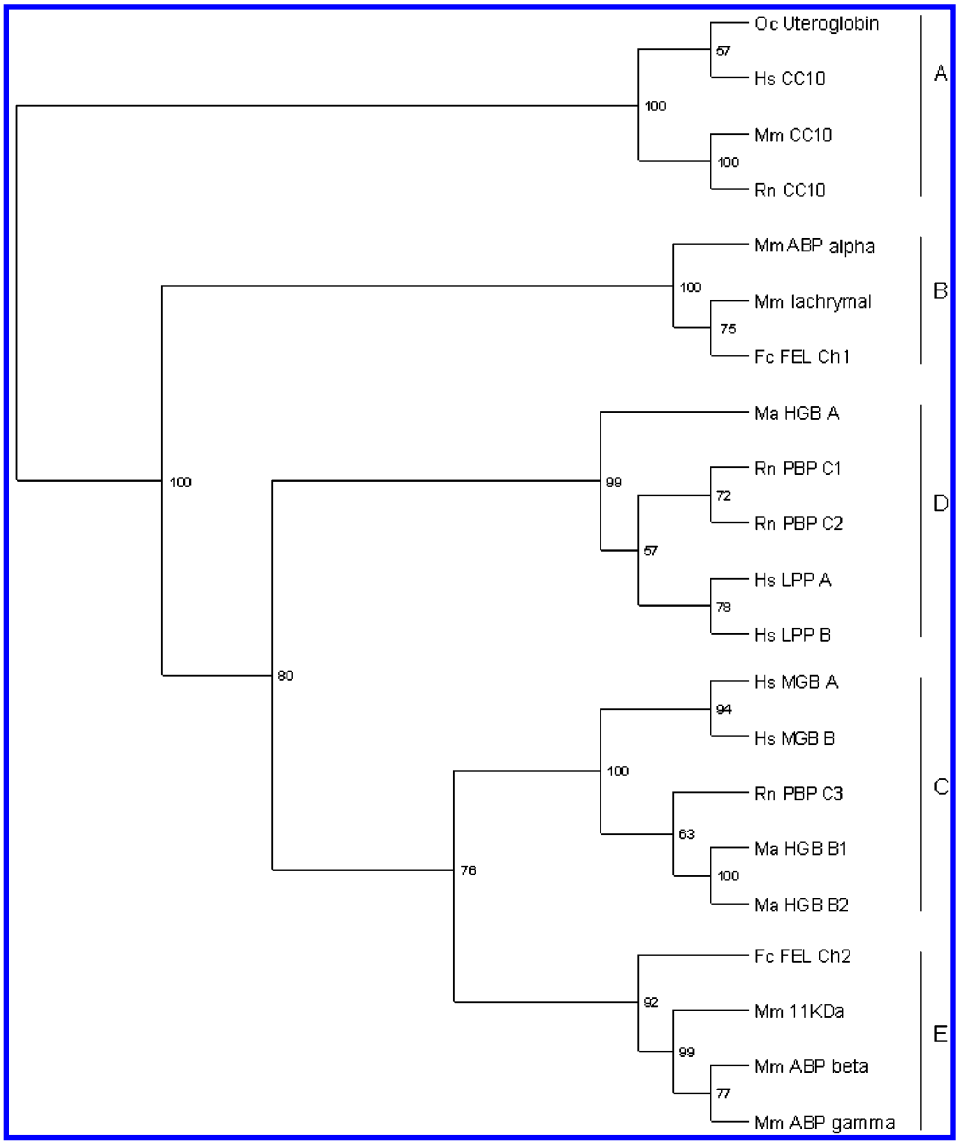

FIGURE 3: Phylogenetic tree of the secretoglobins shown in Figure 2. The tree is rooted by the four uteroglobin/CC10 sequences shown and bootstrap values are indicated at the nodes. Groups of secretoglobin proteins corresponding to those named A−E (*36*) are indicated by vertical lines and capital letters on the right-hand side of the tree. The branch lengths on the tree are not proportional to base substitutions. The abbreviations are explained in the legend to Figure 2.

Table 2: Values of K_a and K_s in Comparisons among *Abpa*, *Abpb*, and *Abpg* Alleles

	Abpa			*Abpb*			*Abpg*		
comparison	K_a (SE)	K_s (SE)	K_a/K_s	K_a (SE)	K_s (SE)	K_a/K_s	K_a (SE)	K_s (SE)	K_a/K_s
domesticus−*musculus*	0.041 (0.017)	0.000 (0.000)		0.016 (0.009)	0.000 (0.000)		0.108 (0.025)	0.126 (0.043)	0.86
domesticus−*castaneus*	0.041 (0.017)	0.000 (0.000)		0.000 (0.000)	0.014 (0.014)		0.102 (0.024)	0.125 (0.043)	0.82
musculus−*castaneus*	0.014 (0.010)	0.017 (0.017)	0.82	0.016 (0.009)	0.014 (0.014)	1.14	0.027 (0.012)	0.000 (0.000)	

of the helices, and the carbohydrate moiety on the α subunit. There are no amino acid substitutions at any of the six conserved sites in the α and the β subunits. There is an I/L13 variation and an I/L41 variation between the *domesticus* and *musculus* variants of the γ subunit, indicating flexibility in these two residues which are variable between the β and γ subunits (see Discussion).

All three subunits of ABP are likely to consist of four α-helices in the first 70 amino acid residues. β and γ may also have a fifth helix in the extra C-terminal amino acids comprising the tail that appears to be unique to them and four other secretoglobin subunits (see above), although one analysis predicted sheet structure for this region in the b and c forms of γ. All of the variable residues in the α subunit are in helices while variable residues in β and γ fall in helices and in loops between them. A fifth helix is proposed for the long tails found on β and γ but not other secretoglobins (Figure 2). Amino acid substitutions between the *domesticus* γ and the *musculus* γ fall in that helix and the loop before it.

DISCUSSION

Characterization of Two Forms of ABP. In this paper, we describe the characterization of two forms of mouse salivary androgen-binding protein (ABP). ABP is a family of proteins composed of three subunits, α, β, and γ, at least two of which

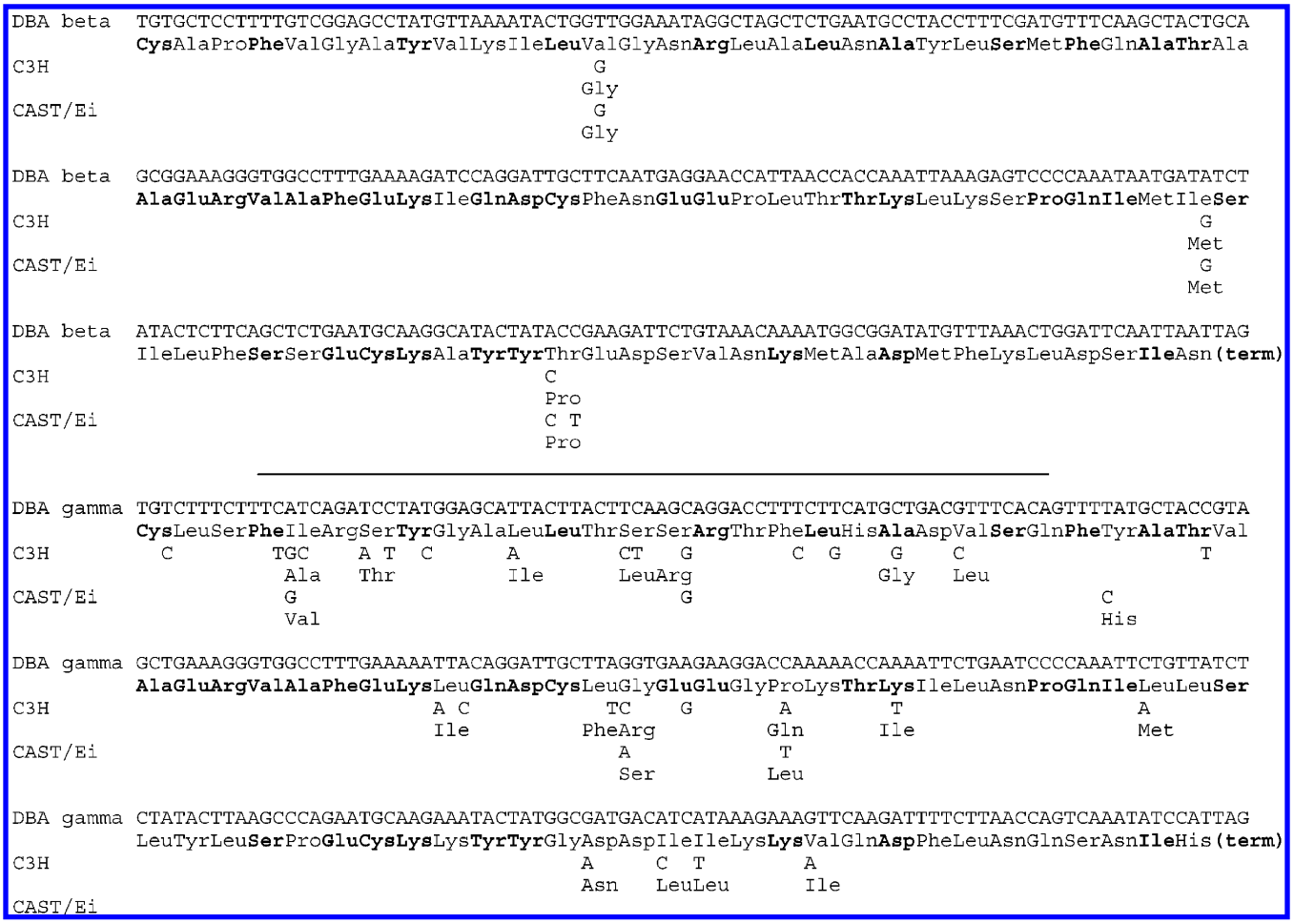

FIGURE 4: Nucleotide and amino acid substitutions in the sequences of the secreted forms of the β and γ sequences among three strains of mice that represent the three subspecies of *Mus musculus*. The ABP phenotypes of DBA, C3H, and CAST/Ei are those of *M. musculus musculus*, *M. musculus domesticus*, and *M. musculus castaneus*, respectively. The amino acid residues in common between the β and γ subunits are in bold type.

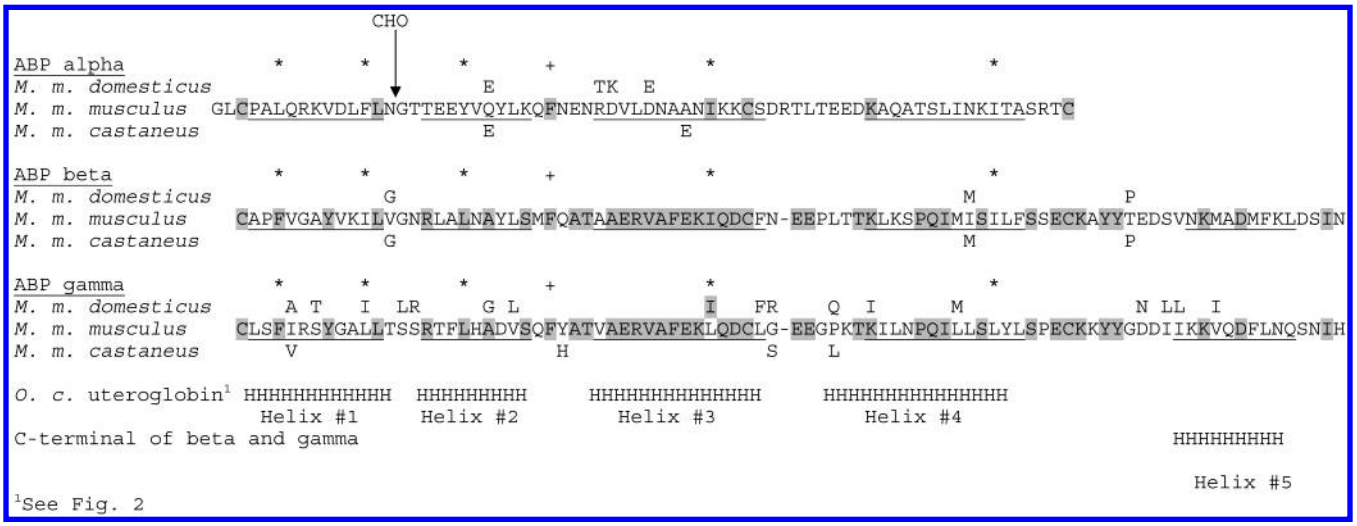

FIGURE 5: Comparison of the amino acid sequences of the secreted forms of the three subunits of mouse salivary ABP. The helices found in rabbit uteroglobin are shown by the series of H's at the bottom and are the underlined residues in the ABP sequences. A fifth helix is proposed for the tails of the β and γ sequences which are longer than uteroglobin and most other secretoglobins. Consensus residues are highlighted in gray, and the six most conserved residues in the secretoglobin family are shown by asterisks and a plus sign as they are in Figure 2. The carbohydrate moiety found on the α subunit is indicated by CHO with an arrow to N15.

(α and γ) have genetic variants (*28*). The sequence of the α subunit has been determined (*7*) and has been the subject of numerous evolutionary studies (*3*, *22*, *23*). We report here the sequences of the β and γ subunits, having identified them on the basis of a comparison of their isoelectric points (p*I*) with previous observations of their relative positions on IEF gels (*28*). We confirmed their identification by sequencing the β and γ cDNAs from the C3H strain where genetic variation of the γ, but not the β, can be detected on IEF gels (*28*). The p*I*s predicted for these two subunits in C3H

account for the significant increase in the positive charge of the dimer containing the γ variant. The β variant, while differing by three amino acid substitutions, does not change the charge of its dimer, consistent with IEF observations (*28*).

The Genes Encoding ABP Subunits. On the basis of several lines of indirect evidence, Dlouhy et al. (*28*) argued that three distinct genes, *Abpa*, *Abpb*, and *Abpg*, encoded the α, β, and γ subunits, respectively. The β and γ subunit sequences presented in this report strongly suggest that three separate genes encode the three subunits of mouse salivary androgen-binding protein (ABP). The many amino acid differences between the two subunits rule out posttranslational modification as the explanation of two different chains that bind alternatively to the α subunit. The longest sequence of amino acids that is identical between β and γ is residues 54−61, which makes it unlikely that the two subunits are the product of alternative splicing of exons in a single transcriptional element. The striking number of amino acid substitutions between the γ's from the *domesticus* and *castaneus* subspecies compared to an identical β sequence shared by the two supports this argument against alternative splicing strategies within a single gene.

Evolution of ABP and Other Secretoglobins. The secretoglobin gene family appears to be restricted to mammals, suggesting that it evolved and radiated relatively recently by gene duplication. Both our tree and that of Ni et al. (*36*) cluster the uteroglobin/clara proteins, which are homodimers, separately from the heterodimeric secretoglobins. The ABP β and γ subunits and mouse lachrymal 11 kDa protein cluster with cat Fel dI, chain 2, the only protein making up cluster E on their tree (*36*). Considering that the ABP α subunit groups with the mouse allergen-like lachrymal protein and chain 1 of cat Fel dI, it seems likely that clusters B and E represent expansions of recent duplication events, the first of which created the ancestral B and E genes in a common mammalian ancestor, followed by later duplications that created additional genes within the two groups, e.g., ABP β and γ in the mouse.

Comparing groups B and E with the other groups in Figure 3, it is interesting to note the striking differences between secretoglobins expressed in the same tissues in various rodents. Alvarez et al. (*9*) described two heteroglobin B (HGB.B) subunits (B1 and B2) expressed in hamster Harderian glands, one of which (B1) is also expressed in hamster submaxillary glands. The ABP and HGB subunit sequences are different enough that HGBB1 and HGBB2 fall into a different secretoglobin subfamily than ABP β and γ (groups C and E, respectively, in Figure 3). The other subunits of the HGB and ABP dimers, HGBA and ABP α, also fall into different families (groups D and B, respectively). In fact, the hamster HGB subunits appear to be more similar to rat PBP subunits than to mouse ABP subunits. HGBA shares group D with rat PBP C1 and C2, while HGBB1 and HGBB2 share group C with rat PBP C1. Another important difference between HGB and ABP is the location of the carbohydrate moieties on these proteins. HGBB1 and HGBB2 have multiple glycosyl moieties while ABP β and γ are not glycosylated. On the other hand, our results suggest that there is a single carbohydrate moiety on the α subunit of ABP but there does not appear to be a similar glycosylation of HGBA (*9*). Thus, while they share the characteristic of being secreted by rodent submaxillary glands, hamster HGB and mouse ABP are otherwise quite different secretoglobins.

Ligand Binding by Secretoglobins. Several secretoglobins, including rabbit uteroglobin (*38*), mouse salivary ABP (*6*), and rat PBP (*39*, *40*), have been shown to bind sex steroid hormones. However, the type of ligand normally bound, as well as the role of ligand binding in the function of the protein, has been the subject of much discussion. Karn (*20*) showed that ABP preferentially binds steroids with the more saturated A ring of progesterone and testosterone. Karn and Clements (*4*) demonstrated that both of the A·B and A·G dimers of ABP bind dihydrotestosterone (DHT) to about the same extent, while the A·B dimer binds substantially more testosterone that does the A·G dimer. Considering the results of these two studies, it appears that (1) the natural ligand of ABP should have a ring similar to that of testosterone and progesterone and (2) the two different dimers of ABP exist to bind structural variations of the ligand with different affinities. The evolution of two alternative subunits that change the binding affinity of the ABP dimer for variations of a ligand strongly suggests that ligand binding is an important aspect of ABP function. A role has been proposed for ABP in mate selection behavior (see, for example, ref *17*), making it likely that sex steroid hormones are natural ligands for the protein.

Residues Which May Be Involved in the Differential Ligand Binding by the Two Dimers of ABP. Because A·B and A·G bind testosterone and DHT with different affinities (*4*), it was of particular interest to examine amino acid residues at the positions shown by others to be critical for ligand binding by uteroglobin [see Callebaut (*1*) for details of the uteroglobin structure summarized here]. The uteroglobin dimer forms an internal hydrophobic cavity, located at the interface between the two subunits, the location of the binding of hydrophobic ligands. The subunits consist of four α-helices which do not form a canonical four-helix bundle motif but rather a boomerang-shaped structure. The subunits are connected in an antiparallel fashion to form a dimer where helices 3 and 4 are involved in the dimer interface. Six residues, F6, L13, Y21, F28, M41, and I63, in each subunit have been identified as being particularly important in this aspect of uteroglobin structure. Of these, only F28 (marked with a plus in Figures 2 and 5) is not accessible to the ligand and probably functions instead in maintaining the dimer interface. The other five (marked with asterisks in Figures 2 and 5) are involved in ligand binding. The aromatic residues F6 and Y21 are critical to this and cannot be replaced by aliphatic amino acids (*1*). Conversely, L13 is accessible to solvent in the hydrophobic pocket and is commonly substituted by aromatic amino acids. This suggests that L13 may be involved in determining ligand specificity.

All three mouse salivary ABP subunits have F at the position corresponding to uteroglobin F28, the residue responsible for maintaining the dimer interface (*1*). Both the β and γ subunits of ABP have F at the position corresponding to uteroglobin F6, but the α subunit has a Leu at that position. Conversely, the α has Y corresponding to uteroglobin Y21 while both the β and γ have Leu at that position. This suggests that, in a heterodimeric configuration, ligand binding can occur as long as at least one of the subunits in the dimer has an aromatic residue at each of those positions (utero-

globin 6 and 21), since all forms of ABP observed to date are capable of binding steroids (*4*, *20*, *28*). β and γ differ at residues corresponding to uteroglobin residues L13, M42, and I63, where in each case they have an Ile and a Leu, respectively. These differences in the internal hydrophobic cavities of the A·B and A·G dimers could account for their different binding affinities for testosterone and DHT (*4*).

A comparison of the regions of the β and γ subunits corresponding to helices 3 and 4 of uteroglobin are additionally interesting in this assessment. Residues 30−43 in β and γ, corresponding to helix 3 in uteroglobin, are 79% conserved, over twice the average identity between these two subunits. By contrast, residues 48−62 in β and γ, corresponding to helix 4 in uteroglobin, are only 40% conserved, about the average for these two. If these are helices 3 and 4 in the ABP A·B and A·G dimers, the conservation of helix 3 between the β and γ subunits may be due to a critical role in the dimer interface where the ligand is bound.

Evolution of the Subunits of ABP. Molecular evolutionary analyses suggest that *Abpa* has evolved under strong directional selection (*3*, *22*, *23*), which is consistent with sexual selection (*17*, *24*) driving ABP evolution. The best evidence that selection is operating directly on ABP is the observation of a high nonsynonymous/synonymous (K_a/K_s) substitution rate ratio (*3*, *22*). In this investigation, we asked whether there was evidence for similarly elevated K_a/K_s ratios in comparisons of the β and γ cDNAs representing the three subspecies of *Mus musculus.* The ratios we obtained (Table 2) are higher than the average observed for proteins generally, in agreement with ratios for the α subunit (*3*, *22*) and suggesting that a similar kind of selection is operating on the β and γ subunit genes.

While most of the results support the sexual selection role, it is by no means certain, and there have been suggestions that it might not be the main and/or original function of this protein (*17*, *23*). A number of questions remain unanswered; for example, what is the physiological role, if any, of androgen binding? The data we report here suggest that androgen binding is an important part of the function of ABP, but more work will be required to show what its role is.

ACKNOWLEDGMENT

The authors thank M. Hermodson and M. Nachman for many helpful discussions and M. Bower for technical assistance. We also thank K. Axelsen for information about ATPase.

REFERENCES

1. Callebaut, I., Poupon, A., Bally, R., Demaret, J.-P., Housset, D., Delettre, J., Hossenlopp, P., and Mornon, J.-P. (2000) *Ann. N.Y. Acad. Sci. 923*, 90−112.
2. Pattabiraman, N., Matthews, J., Ward, K., Mantile-Selvaggi, G., Miele, L., and Mukherjee, A. (2000) *Ann. N.Y. Acad. Sci. 923*, 113−127.
3. Hwang, J., Hofstetter, J., Bonhomme, F., and Karn, R. (1997) *J. Hered. 88*, 93−97.
4. Karn, R., and Clements, M. (1999) *Biochem. Genet. 37*, 187−199.
5. Karn, R. (1991) in *Aspects of Oral Molecular Biology* (Ferguson, D., Ed.) pp 117−140, S. Karger, Basel.
6. Dlouhy, R., and Karn, R. (1983) *Biochem. Genet. 21*, 1057−1070.
7. Karn, R., and Russell, R. (1993) *Biochem. Genet. 31*, 307−319.
8. Klug, J., Beier, H., Bernard, A., Chilton, B., Fleming, T., Lehrer, R., Miele, L., Pattabiraman, N., and Singh, G. (2000) *Ann. N.Y. Acad. Sci. 923*, 348−354.
9. Alvarez, J., Vinas, J., Alonso, J., Albar, J., Ashman, K., and Dominguez, P. (2002) *J. Biol. Chem. 277*, 233−242.
10. *The Uteroglobin/Clara Cell Protein Family* (2000) Vol. 923, New York Academy of Sciences, New York.
11. Daniel, J. (2000) *Ann. N.Y. Acad. Sci. 923*, 1−8.
12. Beier, H. (2000) *Ann. N.Y. Acad. Sci. 923*, 9−24.
13. Singh, G., and Katyal, S. (2000) *Ann. N.Y. Acad. Sci. 923*, 43−58.
14. Gail, D. (2000) *Ann. N.Y. Acad. Sci. 923*, 355−356.
15. Morgenstern, J., Griffith, I., Brauer, A., Rogers, B., Bond, J., Chapman, M., and Kuo, M. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 9690−9694.
16. Dominguez, P. (1995) *FEBS Lett. 376*, 257−261.
17. Laukaitis, C., Critser, E., and Karn, R. (1997) *Evolution 51*, 2000−2005.
18. Dlouhy, R., Nichols, W., and Karn, R. (1986) *Biochem. Genet. 24*, 743−673.
19. Karn, R. (1994) *Biochem. Genet. 32*, 271−277.
20. Karn, R. (1998) *Biochem. Genet. 36*, 105−117.
21. Karn, R., and Dlouhy, R. (1991) *J. Hered. 82*, 453−458.
22. Karn, R., and Nachman, M. (1999) *Mol. Biol. Evol. 16*, 1192−1197.
23. Karn, R., Orth, A., Bonhomme, F., and Boursot, P. (2002) *Mol. Biol. Evol. 19*, 462−471.
24. Talley, H., Laukaitis, C., and Karn, R. (2001) *Evolution 55*, 631−634.
25. Boursot, P., Auffray, J.-C., Britton-Davidion, J., and Bonhomme, F. (1993) *Annu. Rev. Ecol. Syst. 24*, 119−152.
26. Sage, R., Atchley, W., and Capanna, E. (1993) *Syst. Biol. 42*, 523−561.
27. Thiessen, D. (1992) in *Harderian Glands* (Webb, S., Hoffman, R., Puig-Domingo, M., and Reiter, R., Eds.) pp 127−140, Springer-Verlag, Berlin.
28. Dlouhy, R., Taylor, B., and Karn, R. (1987) *Genetics 115*, 535−543.
29. Jeanmougin, F., Thompson, J., Gouy, M., Higgins, D., and Gibson, T. (1998) *Trends Biochem. Sci. 23*, 403−405.
30. Thompson, J., Gibson, T., Plewniak, F., Jeanmougin, F., and Higgins, D. (1997) *Nucleic Acids Res. 25*, 4876−4882.
31. Swofford, D. L. (1998) *PAUP*. Phylogenetic Analysis Using Parsimony* (**and Other Methods*) Version 4, Sinauer Associates, Sunderland, MA.
32. Page, R. (1996) *Comput. Appl. Biosci. 12*, 357−358.
33. Kumar, S., Tamura, K., Jakobsen, I., and Nei, M. (2001) *Bioinformatics 17*, 1244−1245.
34. Nei, M., and Gojobori, T. (1986) *Mol. Biol. Evol. 3*, 418−426.
35. Moore, K., Dunbar, B., Bousfield, G., and Ward, D. (1990) *Endocrinology 127*, 1477−1486.
36. Ni, J., Kalff-Suske, M., Gentz, R., Schageman, J., Beato, M., and Klug, J. (2000) *Ann. N.Y. Acad. Sci. 923*, 25−42.
37. Wolfe, K., and Sharp, P. (1993) *J. Mol. Evol. 37*, 441−456.
38. Beato, M., and Baier, R. (1975) *Biochim. Biophys. Acta 392*, 346−356.
39. Heyns, W., and De Moor, P. (1977) *Eur. J. Biochem. 78*, 221−230.
40. Parker, M., Needham, M., and White, R. (1982) *Nature 298*, 92−94.

BI027424L