THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 281, NO. 4, pp. 1935–1942, January 27, 2006
Printed in the U.S.A.

Function and Evolution of a Mosquito Salivary Protein Family*[S]

Received for publication, September 21, 2005, and in revised form, November 14, 2005 Published, JBC Papers in Press, November 21, 2005, DOI 10.1074/jbc.M510359200

Eric Calvo[1], **Ben J. Mans**[1], **John F. Andersen, and José M. C. Ribeiro**[2]

From the Section of Vector Biology, Laboratory of Malaria and Vector Research, NIAID, National Institutes of Health, Bethesda, Maryland 20852

Saliva of blood-sucking arthropods contains a complex and diverse mixture of antihemostatic, antiinflammatory, and immunomodulatory compounds. The D7 salivary family of proteins is abundantly expressed in blood-feeding Diptera and is distantly related to the odorant-binding protein superfamily. In mosquitoes, two subfamilies exist, the long and short D7 proteins. Ticks and kissing bugs evolved salivary lipocalins that act as efficient scavengers of biogenic amines, and a similar function was postulated for the D7 proteins. Accordingly, we expressed the five members of the small D7 family of the African malaria vector *Anopheles gambiae* and a D7 long form from *Aedes aegypti* and showed by isothermal microcalorimetry, a modified and very sensitive non-equilibrium chromatography/spectrum distortion method, and by smooth muscle bioassay that four of these five short D7 proteins and the D7 long form bind serotonin with high affinity, as well as histamine and norepinephrine. The nonbinding D7 protein is poorly expressed in the salivary glands and appears to be on the path to becoming a pseudogene. Scavenging of host amines would antagonize their vasoconstrictor, platelet-aggregating, and pain-inducing properties. It appears that counteracting biogenic amines is of strong adaptive value in the convergent evolution of arthropods to hematophagy. This adaptation has been solved independently in ticks, bugs, and mosquitoes by co-option of either member of the lipocalin or, as shown here, by the odorant-binding protein families.

At least 14 orders or families of arthropods (containing over 400 different genera and more than 15,000 species) independently evolved to feed on vertebrate blood (1). To accomplish this task, these animals evolved sophisticated cocktails of salivary pharmacologic reagents that affect blood clotting, platelet aggregation, vascular contraction, host immunity, inflammation, and angiogenesis. With the development of transcriptome analysis, the salivary compositional diversity of several hematophagous arthropods is being revealed at a fast pace; however, the majority of these proteins have no known function (2).

Among many different families of proteins unique to hematophagous arthropods, the D7 family has been recognized to be specifically expressed in the salivary glands of adult Diptera. These proteins are distant relatives of the odorant-binding protein superfamily, of which they are a distinct branch. In mosquitoes, two D7 subfamilies exist, the short family having molecular mass of 15–20 kDa and the long with 27–30 kDa, whereas sand flies appear to have only the long forms (3–5). According to a recent sialotranscriptome analysis, *Anopheles gambiae*, the main African malaria vector, has three long and five short D7 proteins in chromosome arm 3R (6), arranged in an inverted tandem repeat. In the closely related mosquito *Anopheles stephensi*, one short D7 protein, named hamadarin, has been expressed and shown to have anticlotting activity (7). No other function has been described for the other protein members, which are so different as to have <40% amino acid identity to hamadarin.

It has previously been hypothesized that the D7 proteins could function as scavengers of biogenic amines or other hemostasis agonists (5), as happens with the salivary lipocalins of ticks and kissing bugs (8, 9). In this work, we expressed all five short D7 proteins of *An. gambiae* and one long D7 from *Aedes aegypt* and demonstrate that four of the short proteins, including the hamadarin homolog, avidly bind serotonin as well as histamine and norepinephrine (NE),[3] as does the Aedes protein. The nonbinding D7, the last gene in the short D7 gene cluster, which has two instead of three exons, is poorly expressed in the salivary glands and appears to be on the way to becoming a pseudogene. It thus appears that scavenging biogenic amines has been of strong adaptive value in the evolution to hematophagy and that this has been solved independently in ticks, mosquitoes, and bugs by co-option of members of either the lipocalin or odorant-binding protein families.

EXPERIMENTAL PROCEDURES

Mosquitoes and cDNA Preparation—Mosquitoes were reared in the National Institutes of Health Medical Entomology Section as described previously (10). Adult female *An. gambiae* (Giles strain) were dissected at day 2–3 post-emergence to remove the salivary glands. Salivary glands (20 pairs) were dissected under a stereomicroscope in 20 mM HEPES buffer, pH 7.4, plus 150 mM NaCl and used immediately for total RNA extraction (TRIzol; Invitrogen) and cDNA synthesis (SuperScript™ RT; Invitrogen) following standard procedures.

Expression and Purification of D7 Proteins in a Mammalian Expression System—PCR fragments coding for mature D7 proteins were amplified (Platinum Spermix; Invitrogen) from salivary gland cDNA using gene-specific primers as follows: D7rp1 forward 5′-aacacggttaagaagtgtgagaag-3′ and reverse 5′-ttagttgcaaatcttgtcatcgattt-3′; D7rp2 forward 5′-cgaaaggagtcaacggtggagg-3′ and reverse 5′-ctagcacaaaccatcatcgatttcc-3′; D7rp3 forward 5′-agacaagaggaaacggttgaagaatgc-3′ and reverse 5′-tcagttacacagcccatcatcaa-3′; D7rp4 forward 5′-gagactgtgcaagattgtgagaat-3′ and reverse 5′-tcagcagtttaatgccttatcataatc-3′; D7rp5 forward 5′-gtgagtgattgtgtgaggcatg-3′ and reverse 5′-ttagcaccgtaactgactgttcaac-3′. The PCR-amplified products were cloned into VR2001-TOPO vector (modified version of the VR1020 vector; Vical Inc., San Diego, CA) and their sequences and orientation verified by DNA

* This work was funded by the Division of Intramural Research, NIAID, National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S3 and supplemental Tables S1 and S2.

[1] Both authors contributed equally to this work.

[2] To whom correspondence should be addressed: NIAID/LMVR, 12735 Twinbrook Pkwy., Rm. 2E32D, Rockville, MD 20852. Tel.: 301-496-9389; Fax: 301-480-2571; E-mail: jribeiro@niaid.nih.gov.

[3] The abbreviations used are: NE, norepinephrine; 5-HT,serotonin.

sequencing (DTCS Quick Start kit; Beckman Coulter Inc., Fullerton, CA). Recombinant proteins were produced by transfecting FreeStyle™ 293-F cells (Invitrogen) with 30 μg of purified VR2001-D7rp plasmids following the manufacturer's recommendations (Invitrogen). After 72 h, transfected cell cultures were harvested and the supernatant (containing secreted D7 recombinant proteins) concentrated 60-fold in AmiconR Ultra-15 (10 kDa MWCO; Millipore Corp., Bedford, MA). Recombinant D7 protein separation was carried out in two-step high pressure liquid chromatography purification using a Spectra System P400 pump in conjunction with a dual-wavelength UV-visible detector model SM4100 (Thermo Separation Products, Riviera Beach, FL). Briefly, recombinant D7 recombinant proteins were loaded on a Superdex 75 column (Amersham Biosciences) and fractions eluted at 0.5 ml/min with 20 mM HEPES, 150 mM NaCl, pH 7.4. Fractions corresponding to 15–20-kDa range from the gel filtration chromatography were loaded on a C_{18} analytic column (Vydac™, 1 × 250 mm) equilibrated with 10% acetonitrile/0.1% trifluoroacetic acid. Elution of D7 recombinant proteins was performed with a linear gradient of 10–80% acetonitrile/trifluoroacetic acid over 60 min at a 0.1 ml/min flow rate. Fractions were vacuum dried and reconstituted in 20 mM HEPES, 100 mM NaCl, pH 7.4. See supplement for additional data.

Bacterial Protein Expression and Purification—cDNA for D7r1, D7r2, D7r3, D7r4, and D7r5 were obtained from an *An. gambiae* cDNA library and were modified by PCR methods so that the mature proteins (signal peptide as predicted with SignalP removed) all obtained an N-terminal NdeI and a C-terminal XhoI restriction site. The NdeI site added a 5′-methionine codon to all sequences that acts as start codon in the bacterial expression system, whereas the XhoI site was incorporated after the stop codon to facilitate directional cloning into the pET17b expression vector. PCR products were first cloned into the TOPO 2.1 vector and sequences confirmed before moving into the pET 17b vector by subcloning. pET 17b constructs were again confirmed before transformation of *Escherichia coli* strain BL21(DE3) pLys-E.

For recombinant protein production, 50 ml of Luria Bertani broth (with added chloramphenicol and carbenicilin) was inoculated with a single colony and grown overnight (maximum of 16 h). Luria Bertani broth (1 liter, with added chloramphenicol and carbenicilin) was inoculated with 10 ml of the overnight culture and grown at 37 °C with shaking at 250 rpm until an optical density of 0.6–0.8 was reached before isopropyl-1-thio-b-D-galactopyranoside (1 mM final concentration) was added to induce expression. The flask was shaken for 3 h under the same conditions; cells were harvested by centrifugation and washed once in 20 mM Tris-HCl, pH 8.0, before the cell pellet was frozen and stored until use.

The frozen cell pellet was resuspended in 200 ml of 20 mM Tris-HCl, pH 8.0, and cells disrupted using a probe sonicator before collecting the inclusion bodies by centrifugation. Inclusion bodies were extracted with 20 mM Tris-HCl, pH 8.0, and 1% Triton X-100 for 3 h followed by centrifugation. The remaining pellet was washed three times with 20 mM Tris-HCl, pH 8.0, before solubilization in 20 ml of 20 mM Tris-HCl, pH 8.0, 6 M guanidinium hydrochloride, 10 mM dithiothreitol. The solubilized material was diluted into 1 liter of 20 mM Tris-HCl, pH 8.5, 0.4 M arginine monohydrochloride and incubated overnight. Protein was bulk precipitated using saturated ammonium sulfate, and the resultant precipitate was collected by filtration. The filtrate was solubilized in 100 ml of 20 mM Tris-HCl, pH 8.0, and further concentrated by ultrafiltration. Samples were then purified using size-exclusion chromatography on Sephacryl S-100 (16/60 column; Amersham Biosciences) using 20 mM Tris-HCl, pH 8.0, 0.15 mM NaCl. D7r2, D7r3, D7r4, and D7r5 were also subjected to another round of purification using anion exchange with a mono-Q column.

Most of the D7 short proteins gave appreciable levels of expression in an *E. coli* expression system, with final yields of purified protein being 1.1, 3.5, 2.7, and 3.2 mg for D7r2, D7r3, D7r4, and D7r5, respectively. D7r1, the ortholog of hamadarin, showed signs of cytotoxicity, as cell growth was halted after induction with isopropyl-1-thio-β-D-galactopyranoside. Expression of D7r1 was scaled up to 3-liter cultures, and the final yield obtained was 274 μg. Purity and the correct identity of the preparations was assessed by SDS-PAGE, mass spectrometry, and N-terminal sequencing (See supplement for additional data).

Spectral Distortion/Size-exclusion Chromatography Experiments—Size-exclusion chromatography was performed with a 4.6 × 300-mm Super SW2000 column obtained from Tosoh Bioscience (Montgomeryville, PA), eluted with 150 mM NaCl, 25 mM HEPES, pH 7.0, at either 0.05, 0.1, 0.2, or 0.4 ml/min. Recombinant proteins (10 μg) were injected in each run, either preincubated with 100 μM serotonin chloride or not. The eluate was monitored by a scanning light-absorbance detector (Model SPD-M10AV; Shimadzu Corp., Columbia, MD). The absorbance data were exported to an Excel spreadsheet and normalized by zeroing the absorbance value at 320 nm and setting to 1 the maximum value observed between 320 and 240 nm. An average of three-five experiments was used to compare spectra of native or serotonin-treated recombinant protein.

Isothermal Titration Calorimetry—Isothermal titration calorimetry was performed using a VP-ITC calorimeter (Microcal, Northhampton, MA). Proteins were dialyzed against 20 mM Tris-HCl, pH 7.4, 0.15 M NaCl, and dialysis buffer was used to prepare all ligands tested. All solutions were degassed under vacuum for 5 min before use. Cell concentrations between 0.75 and 3.0 μM protein were used, and ligand concentrations from 7.5–30 μM were used in the syringe. Aliquots (10 μl) were injected every 120 s, and the syringe was stirred at 290 rpm while measuring heat of binding at 30 °C. After subtraction of the heats of dilution, the net enthalpy data were analyzed with a single binding model using the Microcal Origin software package.

Anticlotting Measurements—These were obtained by measuring either the recalcification time as described before (11) or prolongation of the activated partial thromboplastin time as supplied by the ALEXIN LS kit (Sigma Diagnostics). Sample (10 μl) and normal (50 μl) plasma was incubated for 5 min at 37 °C before adding ALEXIN LS (50 μl) and incubating for a further 5 min. Clotting was induced with 50 μl of 20 mM $CaCl_2$ and measured at 650 nm every 11 s with shaking before and after each reading step. Clotting time was measured at ~50% of the control clot absorbance.

Smooth Muscle Bioassays—The guinea pig ileum contractions to histamine and the rat uterus contractions to serotonin were measured isotonically, and the rat aortic ring preparation contractures elicited by NE were measured isometrically. A modified Tyrode solution (with 10 mM HEPES buffer, pH 7.4) was used for the ileum and aortic assays (12), and low calcium HEPES-buffered solution was used for the uterus assay (13). All solutions were kept oxygenated by bubbling air.

RESULTS AND DISCUSSION

Uncovering of the *An. gambiae* genome (14) and random sequencing of 3,000 clones of a salivary gland cDNA library from adult female mosquitoes identified three long and five short D7 proteins on chromosome 3R, arranged as an inverted tandem repeat (Fig. 1; modified from Ref. 6). These genes were abundantly expressed in adult female salivary glands except for the last sequence on each cassette (in the transcription order), which had one less exon/intron. No expression of these genes was found

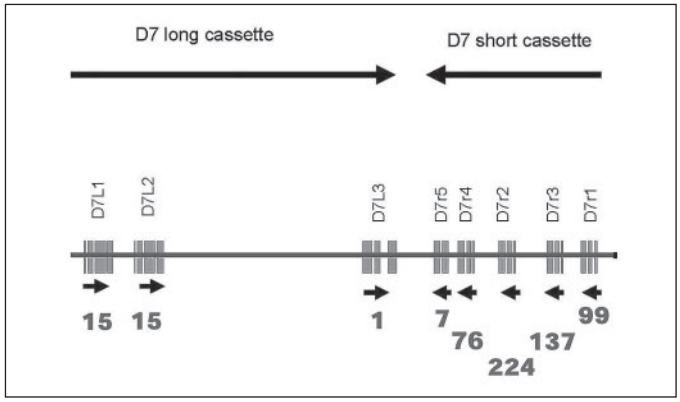

FIGURE 1. **Diagram of *An. gambiae* chromosome 3R showing the D7 salivary gene region.** *Arrows* indicate direction of transcription. *Numerals* indicate number of mRNA sequences deriving from matching genes from a total of 3,000 randomly sequenced clones from a cDNA library made from adult female salivary glands.

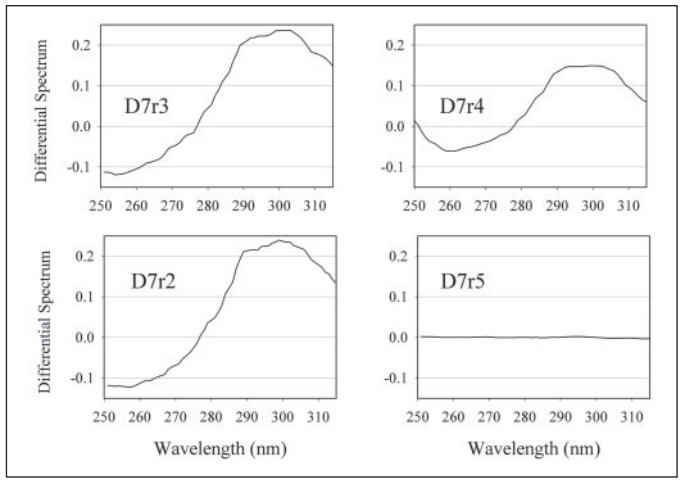

FIGURE 2. **Difference-normalized UV spectra of short D7 proteins when chromatographed with and without preincubation with 100 μM serotonin (5-HT).** The *graphs* represent the average of three-four individual chromatograms for control and 5-HT-treated proteins.

in adult male salivary glands when 1,500 randomly picked clones from a cDNA library were sequenced,[4] confirming the sex-specific nature of this gene family as previously described (5). Because male mosquitoes do not feed on blood, these salivary proteins must be associated with a blood-feeding function. The *An. gambiae* homolog of hamadarin (7), the salivary anticlotting and bradykinin production inhibitory protein from *An. stephensi*, is the first sequence on the short D7 cassette, named D7r1, to which it shares 58% identity.

To study the function of the short D7 proteins, these were expressed in a eukaryotic cell expression system. To investigate whether the short D7 proteins could bind serotonin (5-HT), we incubated 10 μg of protein in 10 μl with 0.1 mM 5-HT and injected the mixture into a molecular sieving column. Because UV absorption of proteins in the 250–300-nm region is dominated by tyrosine and tryptophane residues and because Tyr has an absorption spectrum skewed to the UV region when compared with Trp, we expected that tight binding of one serotonin to the protein would cause an increase in the absorbance in the longer region of the UV spectrum and a relative decrease in the shorter region of the spectrum. D7r1 was not assayed because it did not elute from the column under the conditions used. This is the only basic short D7 protein (pI, 9.2), and basic proteins are known to bind to silica-based columns (15). Preincubation of proteins D7r2, D7r3, and D7r4 with 5-HT led to a significant distortion of the spectrum in the mode predicted ($p < 0.01$ when the average change of spectrum at 300 nm was compared with that at 250 nm) (Fig. 2), whereas the D7r5 spectrum was not affected by 5-HT. The smaller spectral distortion seen in D7r4 compared with the other two proteins can be attributed to the presence of seven Tyr residues in the mature molecule, compared with three on the other proteins. The spectral distortion did not change with flow rates of 0.1–0.4 ml/min leading to retention times (*e.g.* separation of free 5-HT from protein) of 7–28 min, indicating that the K_{off} of the complexes were well above these values. Such values are seen with dissociation constants of the order of nM to pM range, indicating a very tight association of these D7 proteins with 5-HT.

To further characterize the biology of the short D7 proteins, these were expressed in a bacterial system that produces more protein necessary for microcalorimetric binding studies (on the mg protein range). We also succeeded in expressing *Ae. aegypti* D7 protein (NCBI accession number P18153) in high yields by the eukaryotic expression system (see "Experimental Procedures"). Isothermal titration calorimetry indicated that all short D7 proteins except D7r5 bind serotonin and histamine, and some also bind NE and epinephrine (Fig. 3; Table 1), whereas the Aedes long D7 protein bound all amines. In all cases, stoichiometries of 0.6–1 were obtained, suggesting a single binding site for all ligands tested. Results of competition experiments using serotonin at saturated concentrations of histamine suggested that all ligands bind to the same site (results not shown). The binding affinities calculated for a one-binding site model indicated that serotonin is bound with highest affinity (K_d values in the low nM range) by all the D7-related proteins. NE is bound with second highest affinity by D7r2 and D7r3, with K_d values also in the low nM range. Significant differences between the D7-related proteins could, however, be distinguished, with D7r1 not binding NE at all and D7r4 binding with relatively low affinity (K_d ~650 nM) that suggested that NE binding would not be physiologically significant. Histamine was bound by all four D7-related proteins and the Aedes long D7 with comparable affinities (K_d ~40–140 nM), suggesting that this would be physiologically significant. Significant differences appeared again with epinephrine binding, where D7r1 and D7r4 did not bind this ligand at detectable levels, whereas D7r2 and D7r3 showed K_d values comparable with those found for histamine.

Microcalorimetric measurements of the binding affinities of short D7 proteins to biogenic amines indicated that they would be able to remove the amines from their vertebrate receptors if supplied in sufficient amounts, because receptor affinities for these biogenic amines are on the order of 2–3 orders of magnitude lower (10^{-7}-10^{-6} M). Accordingly, we tested the ability of short D7 proteins to antagonize the contractions of the rat uterus to serotonin, of the guinea pig ileum to histamine, and of the rat aorta to NE (D7r1 was only tested in the rat uterus assay due to lack of enough recombinant protein). Results (Fig. 4) indicated that D7r2, D7r3, and D7r4 were effective inhibitors of histamine and serotonin in smooth muscle preparations, whereas only D7r2 and D7r3 were strong inhibitors of NE (Fig. 4). D7r4 gave a partial inhibition of the NE contraction when added at a concentration of 1.5 μM (when NE was added to the aortic ring at 0.5 μM), and D7r5 was inactive at 1.5 μM (results not shown). The Aedes long D7 protein also efficiently counteracted contraction of NE in the rat aortic ring (Fig. 4*C*). Smooth muscle bioassays thus correlated as expected with the binding affinities measured by microcalorimetric measurements.

When the recombinant short D7 proteins of *Anopheles* were tested for their anticlotting effect, only D7r1, the ortholog of hamadarin, showed significant inhibition of the activated partial thromboplastin

[4] E. Calvo and J. Riberro, unpublished data.

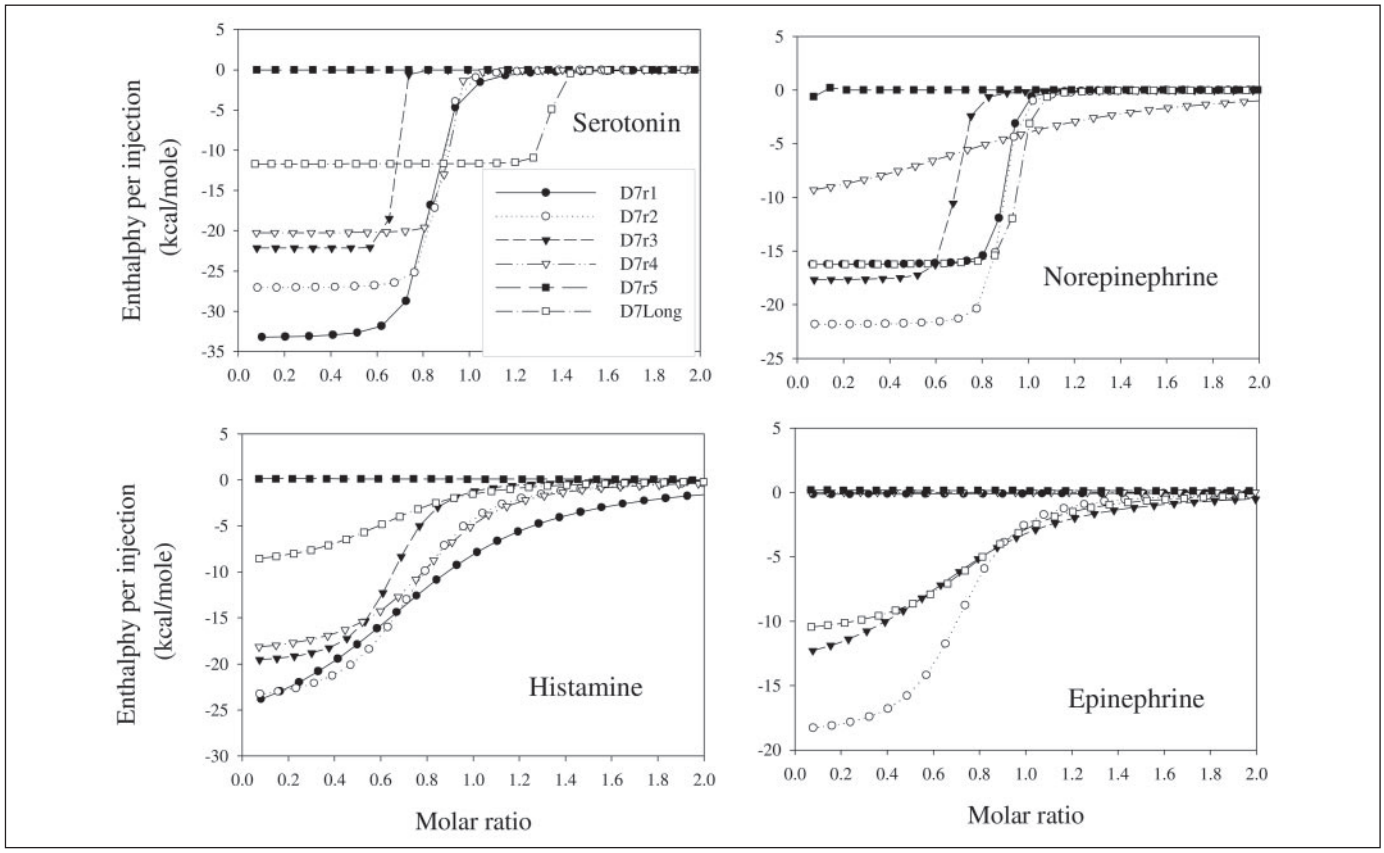

FIGURE 3. **Binding of biogenic amines by the D7-related proteins as measured by isothermal titration calorimetery.** Indicated are fits of the experimental data to a one-binding site model. In the case of D7r2, D7r3, D7r4, and D7r5, ~30 μM of ligand was titrated into ~3 μM of protein. In the case of D7r1, ~8 μM ligand was titrated into ~0.75 μM protein.

TABLE 1

Thermodynamic parameters for binding of biogenic amines to mosquito D7 proteins obtained from isothermal calorimetery analysis

	Stoichiometry	K_d	ΔH	TΔS
		nM	*Kcal/mol*	*Kcal/mol*
Serotonin				
D7r1	0.783	1.99 ± 0.398	−33.3 ± 0.36	−21.31
D7r2	0.828	2.93 ± 0.586	−27.1 ± 0.20	−15.27
D7r3	0.642	0.16 ± 0.074	−22.1 ± 0.11	−8.57
D7r4	0.864	0.93 ± 0.316	−20.2 ± 0.14	−7.73
AeD7L	1.310	0.39 ± 0.565	−11.7 ± 0.21	1.33
Norepinephrine				
D7r2	0.849	2.84 ± 0.79	−21.8 ± 0.23	−10.00
D7r3	0.650	3.19 ± 0.70	−17.7 ± 0.15	−5.91
D7r4	0.907	645 ± 75	−11.6 ± 0.45	−3.03
AeD7L	0.925	0.119 ± 0.05	−16.2 ± 0.21	−3.85
Histamine				
D7r1	0.793	103 ± 15	−28.5 ± 1.24	−18.88
D7r2	0.729	90 ± 7	−24.4 ± 0.26	−14.64
D7r3	0.631	41 ± 6	−20.0 ± 0.29	−9.76
D7r4	0.806	111 ± 9	−19.0 ± 0.21	−9.39
AeD7L	0.647	140 ± 47	−9.5 ± 0.68	0.003
Epinephrine				
D7r2	0.702	64 ± 6	−18.9 ± 0.23	−8.97
D7r3	0.713	312 ± 36	−14.3 ± 0.40	−5.27
AeD7L	0.792	102 ± 13	−11.1 ± 0.23	−1.45

time test, with an activity comparable with that observed for the *An. stephensi* protein (~40% prolongation of clotting time compared with the control at concentrations of ~145 nM) (7). The other D7-related proteins showed no inhibition at concentrations 10–30 times higher than that used for D7r1 (Fig. 5).

D7 proteins are among the most abundantly expressed salivary proteins recovered from SDS-PAGE gels (10, 16–20), consistent with their agonist binding function. Interestingly, the culicines *Culex quinquefasciatus* and *Ae. aegypti* have high expression of the long forms, whereas the anophelines *An. gambiae* and *An. stephensi* have poor expression of the long form and abundant protein expression of the short forms, as revealed by Edman degradation of one-dimension SDS-PAGE gel bands (10, 16–20). As proposed before (3), this nonenzymatic, nonreceptor-mediated way of antagonizing hemostasis needs large amounts of the binding protein, which would presumably have to achieve a concentration of 0.2–2 μM (the normal receptor saturating concentration of endogenous physiologic mediators such as histamine, serotonin, or ADP) or between 0.003 and 0.03 μg/μl of protein at the feeding wound

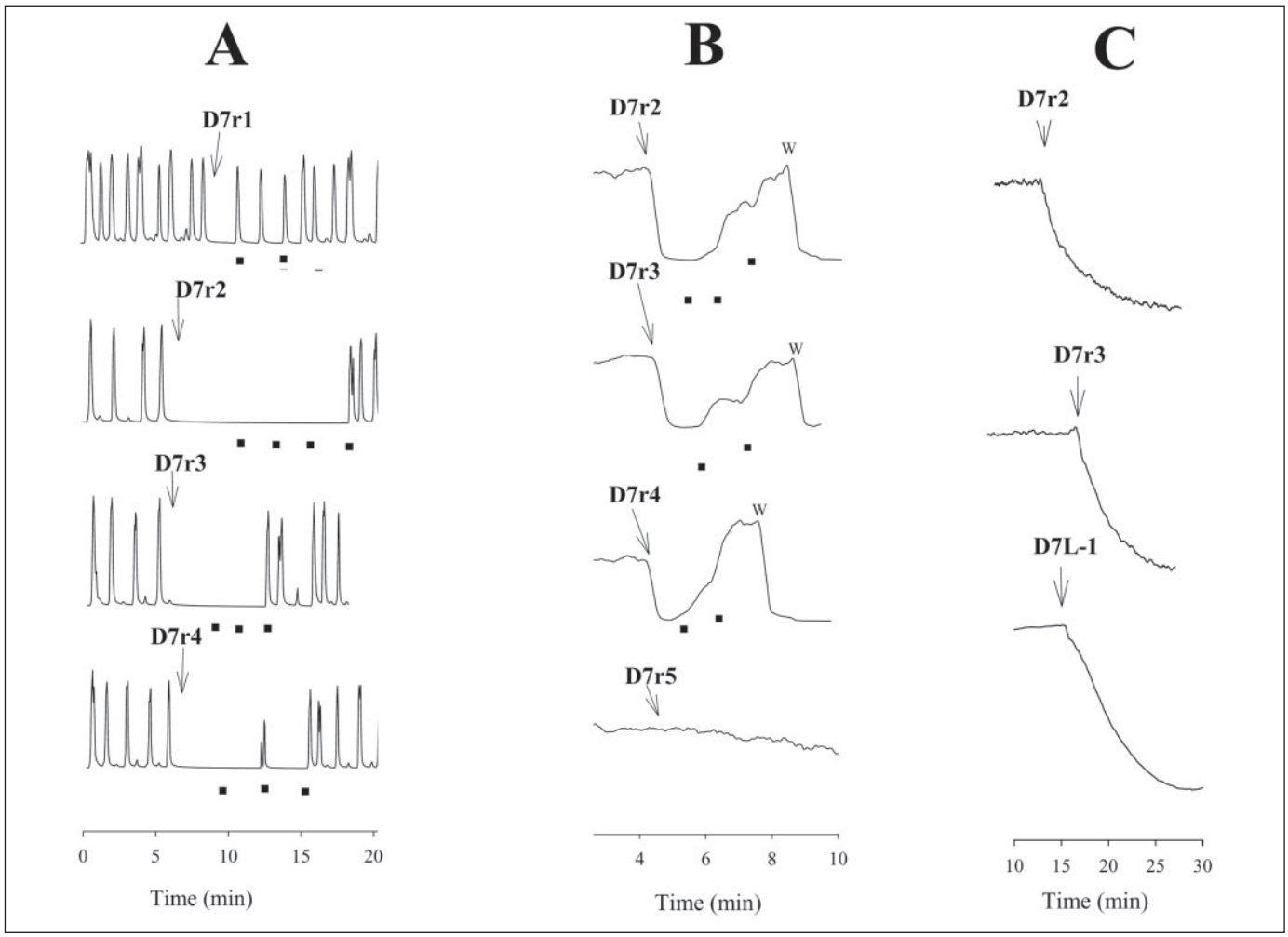

FIGURE 4. **Antagonism of short D7 proteins to smooth muscle-contracting activities of serotonin (*A*), histamine (*B*), and norepinephrine (*C*).** *A*, at time 0, 0.1 μM serotonin was added to the guinea pig ileum to produce rhythmic contractions. The *arrows* indicate addition of 0.5 μM of the indicated proteins. ■ indicates additions of 5-HT to increase the concentration by 0.1 μM. *B*, at time 0, 0.2 μM histamine was added to produce a tonic contraction of the guinea pig ileum. The *arrows* indicate addition of 1 μM of the indicated proteins. ■ indicates addition of 0.2 μM steps of histamine. *W* indicates washing of the preparation and return to baseline. *C*, at time 0, 0.5 μM norepinephrine was added to isolated rat aortic rings. The *arrows* indicate addition of 1 μM of the indicated proteins.

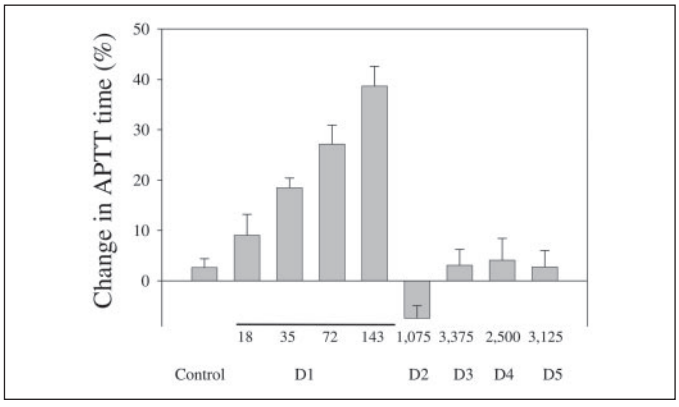

FIGURE 5. **Effect of recombinant short D7 proteins from *An. gambiae* on the activated partial prothrombin time (*APTT*) (mean ± S. E., $n = 3$).** The *numbers below the bars* indicate the final concentration in nM for each protein in the assay.

site for the short D7 proteins, assuming stoichiometric binding and an average molecular mass of 15 kDa. This contrasts with the activity of compounds that bind to cellular receptors such as the salivary vasodilatory maxadilan, which is active at nanomolar concentrations (21), or with enzymatic inhibitors such as apyrase that are active at subnanomolar quantities (22). Taking into consideration that between 0.003 and 0.03 μg/μl of short D7 proteins are needed for their effective function and that the feeding lesion would have an effective volume of 10 μl, then 0.03–0.3 μg of the protein is required to mop local biogenic amines. Considering that mosquito salivary glands have a protein content of 1–3 μg and that half of the protein is discharged during a blood meal (23–25), we concluded that D7 proteins should be at least between 5 and 20% of the salivary protein, a number compatible with the proposed major function of this protein family as proposed here.

The biogenic amines that are targeted by the D7 proteins all have important physiologic functions related to host defense against blood loss and parasite invasion. Serotonin is released from dense granules during platelet activation and leads to increased vascular permeability, acts as a weak platelet agonist, and is a vasoconstrictor (26). Histamine is released from tissue mast cells and platelets during inflammation and tissue damage and also leads to vascular permeability and infiltration of the injury site by monocytes (27). Both serotonin and histamine also induce host sensations of burning and itching that might disturb feeding (28). NE and epinephrine stimulate vasoconstriction, and epinephrine also acts as a platelet agonist (29). Sequestration of these biogenic amines during mosquito feeding would thus serve the important functions of inhibiting platelet aggregation, vasoconstriction, inflammation,

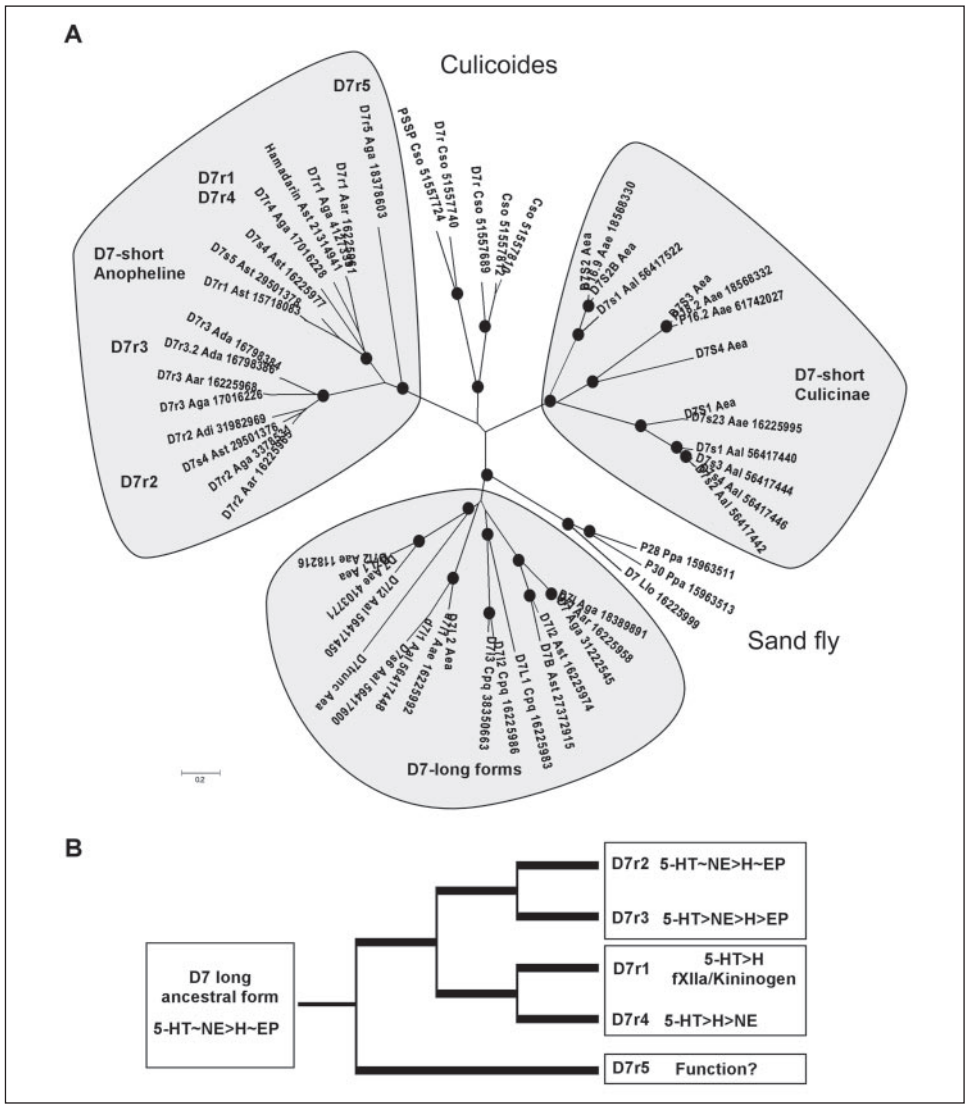

FIGURE 6. **Phylogenetic analysis of the D7 family of salivary proteins from blood-sucking Diptera.** *A*, an unrooted neighbor-joining tree of the D family. Confidence values >85% are indicated at branch points with *black dots* (10,000 bootstraps). *B*, a model for the functional evolution of the *Anopheles* D7-related short proteins. A D7 long form with bio-amine binding properties existed in the Culicidae ancestor. Gene duplications of the C-terminal D7 long domain within the Anophelines yielded the D7-related short forms. Conservation of function yielded branches similar to the ancestral form (D7r2 and D7r3), whereas functional specialization yielded branches more specific for serotonin and histamine (D7r1 and D7r4) with novel evolution of anti-coagulant activity in d7r1. D7r5 lost all bio-amine binding capability and either evolved an as yet unknown function or is becoming a pseudogene. *5-HT*, serotonin; *EP*, epinephrine; *NE*, norepinephrine.

and host behavioral responses that might disturb feeding. The fact that D7r1-D7r4, as well as the Aedes long D7 protein, all exhibit targeting of the biogenic amines underscores the importance of this function for this protein family.

Gene duplication accounts for a large proportion of genes in eukaryotic genomes (30). While most duplication events may not confer an advantage to their hosts and eventually become pseudogenes, a common immediate effect is increased dosage of the transcript and protein (31). In the case of the D7 proteins, this increased tissue dosage may have been of advantage in supplying the large amounts of protein needed for effective scavenging of the biogenic amines at the site of the blood meal. Further evolution of duplicated genes may lead to divergence of function, apparent in the *An. gambiae* D7 short proteins by their differential amine binding and anticlotting properties.

The differences observed between the different D7 proteins have interesting implications for the evolution of this protein family. Phylogenetic analysis of the D7-related proteins shows that three main mosquito branches can be distinguished, as well as one sand fly (family Psychodidae) and one Culicoides (family Ceratopogonidae) branch (Fig. 6*A*): (i) The D7 long proteins that form a monophyletic clade related to salivary proteins from sand flies; (ii) D7 short proteins from the Culicinae (Culex and Aedes families); and (iii) D7 short proteins from Anopheline mosquitoes. There is also a distinct clade formed by Culicoides D7 proteins (32). The anopheline D7 short proteins group into several paralogous subclades, with D7r2 and D7r3 forming a clade that also contains orthologs from *Anopheles arabiensis*, *Anopheles darlingi*, and *Anopheles dirus* B; D7r1 and D7r4 group in the hamadarin clade, for which D7r1 orthologs are found in *An. stephensi*, *An. gambiae*, and *An. arabiensis*. In the case of D7r4, no specific ortholog is found, suggesting that this is a gene duplication that occurred recently in *An. gambiae* or that because of its relatively low expression (Fig. 1) it has not yet been found in other species. Similarly, D7r5 groups basal to the rest of the D7 short proteins with no orthologous genes found, suggesting that this gene duplication was limited to *An. gambiae* or, more probably, remains to be discovered in other anophelines. It is interesting that D7 transcripts are also found in the salivary gland of other blood-sucking Diptera that evolved independently to blood feeding.

The fact that D7r1-D7r4 and the long Aedes D7 protein all bind serotonin suggests that this was the ancestral function associated with the D7 proteins. Even so, the significant differences observed for the binding of NE and epinephrine indicate that this function probably evolved within the ancestor of the D7 proteins but was then optimized within the D7r2/D7r3 clade and mostly lost in the D7r1/D7r4 clade, as noted by the absence of binding of epinephrine by D7r1 and D7r4 and

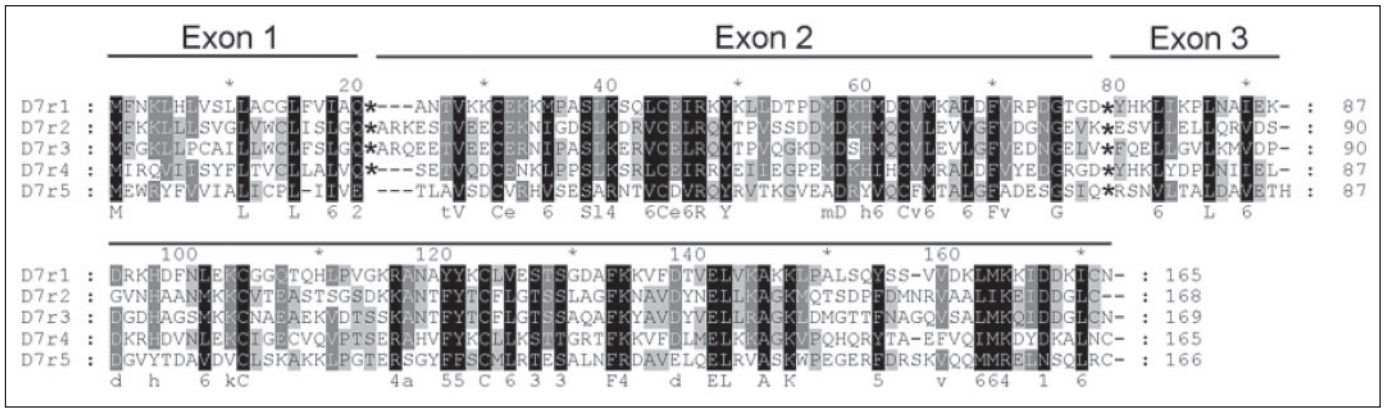

FIGURE 7. **Alignment of the D7 short proteins from *An. gambiae***. Indicated are the exons and the presence of introns (*, phase 0) within the alignment.

the low affinity of D7r4 for NE (Fig. 6*B*). This is supported by the presence of bioactive amine binding capability of the D7 long form from Aedes. The fact that the D7 long forms group into a monophyletic clade or a cluster-of-orthologs group that includes proteins from both Culicinae and Anophelinae suggests that the D7 long form was indeed the ancestral protein with bioactive amine binding functions. If this is the case, most probably the C-terminal odorant-binding protein domain of the D7 long form has the amine binding function, as this domain shows highest sequence similarity to the D7 short forms from the Anophelines (3).

We have also expressed representative members from the Culicinae D7 short forms (two proteins from Culex, two proteins from *Ae. aegypti*, and one protein from *Aedes albopictus*) but could not find any binding of bioactive amine using microcalorimetry. If we exclude the possibility that we did not obtain actively folded proteins, their inability to bind bioactive amines could be explained by a separate gene duplication and expansion event that occurred in the Culicinae to produce D7 short forms and might explain the particular topology obtained for the phylogenetic analysis. It is interesting in this context that short D7 proteins are poorly expressed in culicines but rich in anopheline salivary glands. Because of their low salivary quantities, the short D7 proteins of culicines may not be sufficient to bind mediators of inflammation and hemostasis that reach μM concentrations, and they may be adapted to mop mediators of higher receptor affinities, such as the inflammatory lipids, or they may have some other function that does not need the large amounts found associated with the large D7 of culicines and the short D7 of anophelines.

The absence of any anticlotting activity in the D7-related proteins that are paralogs to D7r1 and hamadarin indicate that inhibition of the intrinsic blood coagulation pathway evolved exclusively within the D7r1 hamadarin branch. The terminal position of this branch in the phylogenetic tree (Fig. 6) suggests that this anticlotting activity evolved exclusively within the anopheline lineage. The uniqueness of the hamadarin homolog as the only anticlotting protein among the other short D7 proteins is paralleled in the blood-sucking bug *Rhodnius prolixus*. In this insect, five salivary heme lipocalins are known carriers of nitric oxide and binders of histamine. An additional nonheme lipocalin binds serotonin, epinephrine, and NE (33), but only two of these proteins function as an anticlotting lipocalin, each one by a completely different mechanism (11, 34–36).

The fact that D7r5 does not show any biogenic amine binding properties or anticlotting activity could have several implications. D7r5 might have an unknown and as yet undetermined function or might be in the process of being eliminated. Although an unknown function could not be ruled out for D7r5, several lines of evidence support that it is becoming a pseudogene. Analysis of expression levels as estimated by the numbers of expressed sequence tags found for the different D7-related proteins in GenBank^{TM} indicated that D7r1-D7r4 have 70–200 expressed sequence tags each, whereas D7r5 only has 7 expressed sequence tags (Fig. 1). On the phylogenetic tree, D7r5 group basal to D7r1-D7r4 with a long branch, which suggests faster evolutionary rates compared with those of other D7 short proteins. Analysis of the exon-intron patterns of the D7 short proteins showed that D7r1-D7r4 all possess two introns with conserved position and phase 0 (Fig. 7). The first intron is located at the signal peptide splicing site, and the second intron is located toward the center of the mature protein. In D7r5, the second intron position is conserved with regard to position and phase, indicating a genomic origin for D7r5 and not retrotransposition of mRNA. The lack of the first intron could explain the low expressed sequence tag numbers, because the presence of introns and their processing is known to be important for successful nuclear export (37). Together, these properties suggest that D7r5 is in the process of gene death.

The common occurrence of D7 long and short proteins in both culicines and anophelines indicates that these proteins played a role in the early evolution of blood feeding in mosquitoes. It has been suggested that mosquitoes evolved in the Jurassic ~210 million years ago, that the subfamilies of Culicinae and Anophelinae diverged from each other 120 million years ago, and that by the end of the Cretaceous, before mammal irradiation, the genetic composition of the family Culicidae was well established (38). If we assume that the primeval function of D7 proteins was to bind serotonin, it is pertinent to ask whether vertebrates at that time had this amine in their thrombocytes. Interestingly, it has been recently found that 5-HT is associated with thrombocytes of birds and endothermic reptilian species but not with thrombocytes of cold-blooded vertebrates (39). Phylogenetic comparison of the presence of circulating 5-HT indicated an evolutionary divergence within reptilian species that might coincide with the emergence of endothermy ~320 million years ago (39). It is also interesting in this context that the integument, or specialized regions of the integument, is a major thermoregulatory organ in many vertebrates, both mammalian and nonmammalian. Such integument regions are characterized by a blood vessel density that far surpasses the metabolic needs of the organ, the increased blood flow serving as a medium to heat exchange between the organism and the environment (40–42). Accordingly, it is possible that evolution of skin vasculature specializations serving body thermal regulation increased the availability of blood at the animal surface, thus creating an opportunity for arthropods to exploit this protein-rich food resource. A contemporary similar scenario may be found in the tabanid fly *Chrysops callidus*, which feeds in the venous plexus of the carapace of

sun-basking turtles from which they acquire-and to which they transmit-the malaria-like protozoan *Hemoproteus metchnikovi* (40, 43). The time mosquitoes initially evolved (210 million years ago) was thus after the appearance of endothermic animals (320 million years ago) with their zones of skin with high blood flow and all possibly having serotonin-rich thrombocytes that could be counteracted by the D7 proteins.

Histamine binding by D7 proteins may have evolved concomitantly or later than the adaptive role of binding serotonin as vertebrate hosts "learned" to detect bloodsuckers by immune degranulation of mast cells and release of irritating histamine within 1 min of the bite. Some mammalian platelets contain as much histamine as they have serotonin (44), but the amounts of histamine in non-mammalian thrombocytes is unknown.

The evolution of NE binding by D7 proteins may be associated with the vascular control of heat exchange in homeothermic animals. Human skin, for example, has >10 times more blood vessels than needed for the organ metabolic supply. Under normal conditions, most of these vessels are closed due to active sympathetic noradrenergic vasoconstriction, but in hyperthermia, reduction of the noradrenergic tonus allows heat loss to the environment (45). Accordingly, further adaptation to hematophagy may have occurred when the bloodsucker evolved a mechanism to obtain a meal, even when the skin of the cool animal was vasoconstricted, by sequestering NE. Accordingly, D7 proteins continue to be functional after helping mosquitoes to feed on blood for the past 120 million years or longer.

Acknowledgments—We thank Van My Pham for expert technical assistance and Brenda Rae Marshall for editorial assistance. We thank Drs. Robert Gwadz and Thomas Wellems for support and encouragement.

REFERENCES

1. Ribeiro, J. M. C. (1995) *Infect. Agents Dis.* **4,** 143–152
2. Ribeiro, J. M., and Francischetti, I. M. (2003) *Annu. Rev. Entomol.* **48,** 73–88
3. Valenzuela, J. G., Charlab, R., Gonzalez, E. C., Miranda-Santos, I. K. F., Marinotti, O., Francischetti, I. M., and Ribeiro, J. M. C. (2002) *Insect Mol. Biol.* **11,** 149–155
4. Hekmat-Scafe, D. S., Dorit, R. L., and Carlson, J. R. (2000) *Genetics* **155,** 117–127
5. Arca, B., Lombardo, F., Lanfrancotti, A., Spanos, L., Veneri, M., Louis, C., and Coluzzi, M. (2002) *Insect Mol. Biol.* **11,** 47–55
6. Arcà, B., Lombardo, F., Valenzuela, J. G., Francischetti, I. M. B., Marinotti, O., Coluzzi, M., and Ribeiro, J. M. C. (2005) *J. Exp. Biol.* **208,** 3971–3986
7. Isawa, H., Yuda, M., Orito, Y., and Chinzei, Y. (2002) *J. Biol. Chem.* **13,** 13
8. Andersen, J. F., Gudderra, N. P., Francischetti, I. M., and Ribeiro, J. M. (2005) *Arch. Insect Biochem. Physiol.* **58,** 97–105
9. Sangamnatdej, S., Paesen, G. C., Slovak, M., and Nuttall, P. A. (2002) *Insect Mol. Biol.* **11,** 79–86
10. Francischetti, I. M., Valenzuela, J. G., Pham, V. M., Garfield, M. K., and Ribeiro, J. M. (2002) *J. Exp. Biol.* **205,** 2429–2451
11. Ribeiro, J. M. C., Schneider, M., and Guimaraes, J. A. (1995) *Biochem. J.* **308,** 243–249
12. Webster, M., and Prado, E. S. (1970) *Methods Enzymol.* **19,** 681–699
13. Gaddum, J. H., and Hameed, K. A. (1949) *J. Physiol.*(*Lond.*) **108,** 240–248
14. Holt, R. A., Subramanian, G. M., Halpern, A., Sutton, G. G., Charlab, R., Nusskern, D. R., Wincker, P., Clark, A. G., Ribeiro, J. M., Wides, R., Salzberg, S. L., Loftus, B., Yandell, M., Majoros, W. H., Rusch, D. B., Lai, Z., Kraft, C. L., Abril, J. F., Anthouard, V., Arensburger, P., Atkinson, P. W., Baden, H., De Berardinis, V., Baldwin, D., Benes, V., Biedler, J., Blass, C., Bolanos, R., Boscus, D., Barnstead, M., Cai, S., Center, A., Chatuverdi, K., Christophides, G. K., Chrystal, M. A., Clamp, M., Cravchik, A., Curwen, V., Dana, A., Delcher, A., Dew, I., Evans, C. A., Flanigan, M., Grundschober-Freimoser, A., Friedli, L., Gu, Z., Guan, P., Guigo, R., Hillenmeyer, M. E., Hladun, S. L., Hogan, J. R., Hong, Y. S., Hoover, J., Jaillon, O., Ke, Z., Kodira, C., Kokoza, E., Koutsos, A., Letunic, I., Levitsky, A., Liang, Y., Lin, J. J., Lobo, N. F., Lopez, J. R., Malek, J. A., McIntosh, T. C., Meister, S., Miller, J., Mobarry, C., Mongin, E., Murphy, S. D., O'Brochta, D. A., Pfannkoch, C., Qi, R., Regier, M. A., Remington, K., Shao, H., Sharakhova, M. V., Sitter, C. D., Shetty, J., Smith, T. J., Strong, R., Sun, J., Thomasova, D., Ton, L. Q., Topalis, P., Tu, Z., Unger, M. F., Walenz, B., Wang, A., Wang, J., Wang, M., Wang, X., Woodford, K. J., Wortman, J. R., Wu, M., Yao, A., Zdobnov, E. M., Zhang, H., Zhao, Q., Zhao, S., Zhu, S. C., Zhimulev, I., Coluzzi, M., Della Torre, A., Roth, C. W., Louis, C., Kalush, F., Mural, R. J., Myers, E. W., Adams, M. D., Smith, H. O., Broder, S., Gardner, M. J., Fraser, C. M., Birney, E., Bork, P., Brey, P. T., Venter, J. C., Weissenbach, J., Kafatos, F. C., Collins, F. H., and Hoffman, S. L. (2002) *Science* **298,** 129–149
15. Schluter, H., and Zidek, W. (1993) *J. Chromatogr.* **639,** 17–22
16. Valenzuela, J. G., Pham, V. M., Garfield, M. K., Francischetti, I. M., and Ribeiro, J. M. C. (2002) *Insect Biochem. Mol. Biol.* **32,** 1101–1122
17. Valenzuela, J. G., Francischetti, I. M., Pham, V. M., Garfield, M. K., and Ribeiro, J. M. (2003) *Insect Biochem. Mol. Biol.* **33,** 717–732 Del Capurro, M.
18. Calvo, E., Andersen, J., Francischetti, I. M., Del Caputto, M., deBianchi, A. G., James, A. A., Ribeiro, J. M., and Marinotti, O. (2004) *Insect Mol. Biol.* **13,** 73–88
19. Ribeiro, J. M., Charlab, R., Pham, V. M., Garfield, M., and Valenzuela, J. G. (2004) *Insect Biochem. Mol. Biol.* **34,** 543–563
20. James, A. A., Blackmer, K., Marinotti, O., Ghosn, C. R., and Racioppi, J. V. (1991) *Mol. Biochem. Parasitol.* **44,** 245–254
21. Moro, O., and Lerner, E. A. (1997) *J. Biol. Chem.* **272,** 966–970
22. Ribeiro, J. M. C. (1987) *Annu. Rev. Entomol.* **32,** 463–478
23. Marinotti, O., James, A., and Ribeiro, J. M. C. (1990) *J. Insect Physiol.* **36,** 545–548
24. Nascimento, E. P., dos Santos Malafronte, R., and Marinotti, O. (2000) *Arch. Insect Biochem. Physiol.* **43,** 9–15
25. Ribeiro, J. M. C., Rossignol, P. A., and Spielman, A. (1985) *J. Insect Physiol.* **9,** 551–560
26. Van Neuten, J. M., Janssens, W. J., and Vanhoutte, P. M. (1985) *Pharmacol. Res. Comm.* **17,** 585–608
27. Jutel, M., Blaser, K., and Akdis, C. A. (2005) *Int. Arch. Allergy Immunol.* **137,** 82–92
28. Millan, M. J. (1999) *Prog. Neurobiol.* **57,** 1–164
29. Barnett, D. B., Swart, S. S., Nahorski, S. R., and Cook, N. (1985) *Adv. Exp. Med. Biol.* **192,** 97–108
30. Sankoff, D. (2001) *Curr. Opin. Genet. Dev.* **11,** 681–684
31. Kondrashov, F. A., Rogozin, I. B., Wolf, Y. I., and Koonin, E. V. (2002) *Genome Biol.* **3,** Research 0008
32. Campbell, C. L., Vandyke, K. A., Letchworth, G. J., Drolet, B. S., Hanekamp, T., and Wilson, W. C. (2005) *Insect Mol. Biol.* **14,** 121–136
33. Andersen, J. F., Francischetti, I. M., Valenzuela, J. G., Schuck, P., and Ribeiro, J. M. (2003) *J. Biol. Chem.* **278,** 4611–4617
34. Andersen, J. F., Gudderra, N. P., Francischetti, I. M., Valenzuela, J. G., and Ribeiro, J. M. (2004) *Biochemistry* **43,** 6987–6994
35. Gudderra, N. P., Ribeiro, J. M., and Andersen, J. F. (2005) *J. Biol. Chem.* **280,** 25022–25028
36. Zhang, Y., Ribeiro, J. M., Guimaraes, J. A., and Walsh, P. N. (1998) *Biochemistry* **37,** 10681–10690
37. Aguilera, A. (2005) *Curr. Opin. Cell Biol.* **17,** 242–250
38. Rai, K. S., and Black, W. C., IV (1999) *Adv. Genet.* **41,** 1–33
39. Maurer-Spurej, E. (2005) *Cell. Mol. Life Sci.* **62,** 1881–1889
40. Weathers, W. W., and White, F. N. (1971) *Am. J. Physiol.* **221,** 704–710
41. Hochscheid, S., Bentivegna, F., and Speakman, J. R. (2002) *Physiol. Biochem. Zool.* **75,** 66–76
42. Smith, E. N. (1976) *Am. J. Physiol.* **230,** 1205–1210
43. DeGiusti, D. L., Sterling, C. R., and Dobrzechowski, D. (1973) *Nature* **242,** 50–51
44. Humprey, J. H., and Jaques, R. (1954) *J. Physiol. (Lond.)* **124,** 305–310
45. Abramson, D. I. (1989) *Handbook Exp. Pharmacol.* **87,** 89–116