Ameloblastin binds to phospholipid bilayers via a helix-forming motif within the sequence encoded by exon 5

*Jingtan Su[1], Natalie C. Kegulian[1], Rucha Arun Bapat[1], and Janet Moradian-Oldak[1]**

[1]Center for Craniofacial Molecular Biology, Herman Ostrow School of Dentistry, University of Southern California, 2250 Alcazar St. Los Angeles, CA 90033, USA

***All correspondence may be addressed to:**

Dr. Janet Moradian-Oldak, Center for Craniofacial Molecular Biology, University of Southern California, Los Angeles, CA.

Phone: 323-442-7504

Fax: 323-442-2981

Email: joldak@usc.edu

Table S1. GenBank accession numbers of Ambn protein sequences from 47 species for bioinformatic analysis reported in Table 1 and figure 7.

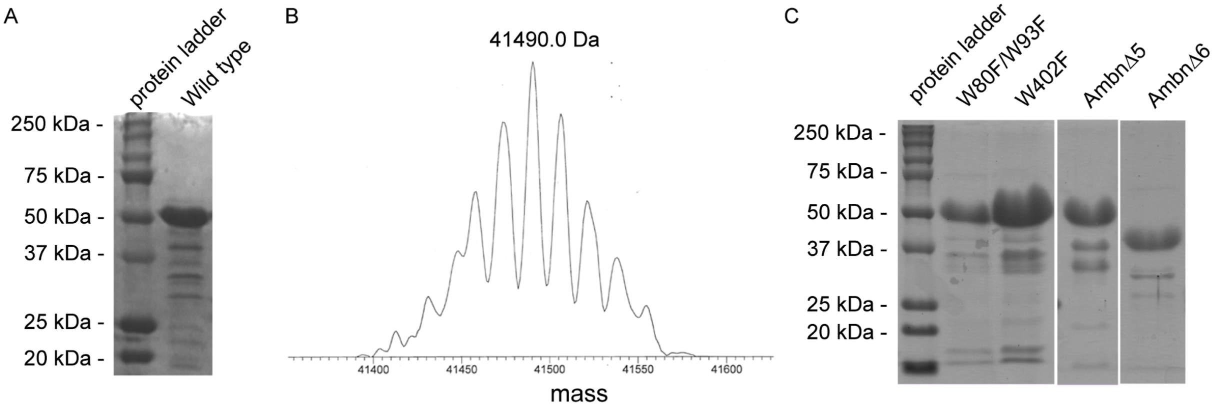
No.	Species	GenBank No.	Lineage
1	*Ailuropoda melanoleuca*	EFB23251.1	Mammalia
2	*Balaenoptera acutorostrata*	XP_007193625.1	Mammalia
3	*Bos mutus*	ELR55446.1	Mammalia
4	*Bos taurus*	AAD39833.2	Mammalia
5	*Camelus ferus*	XP_006188736.1	Mammalia
6	*Canis lupus familiaris*	XP_539304.3	Mammalia
7	*Castor canadensis*	APY20149.1	Mammalia
8	*Cavia porcellus*	XP_005002729.1	Mammalia
9	*Ceratotherium simum simum*	XP_004419200.1	Mammalia
10	*Chrysochloris asiatica*	XP_006873854.1	Mammalia
11	*Cricetulus griseus*	ERE87609.1	Mammalia
12	*Dasypus novemcinctus*	XP_012374726.1	Mammalia
13	*Elephantulus edwardii*	XP_006896906.1	Mammalia
14	*Equus przewalskii*	XP_008528890.1	Mammalia
15	*Fukomys damarensis*	KFO22787.1	Mammalia
16	*Galeopterus variegatus*	XP_008584653.1	Mammalia
17	*Heterocephalus glaber*	EHB10658.1	Mammalia
18	*Homo sapiens*	AAF37355.1	Mammalia
19	*Ictidomys tridecemlineatus*	XP_005333329.1	Mammalia
20	*Macaca fascicularis*	EHH53750.1	Mammalia
21	*Macaca mulatta*	XP_014994033.1	Mammalia
22	*Monodelphis domestica*	XP_007495581.1	Mammalia
23	*Mus musculus*	AAB93765.1	Mammalia
24	*Myotis brandtii*	XP_005881709.1	Mammalia
25	*Myotis davidii*	ELK25525.1	Mammalia
26	*Ornithorhynchus anatinus*	XP_016083225.1	Mammalia
27	*Orycteropus afer afer*	XP_007941009.1	Mammalia
28	*Oryctolagus cuniculus*	XP_008265993.1	Mammalia
29	*Pteropus alecto*	XP_006917342.1	Mammalia
30	*Pteropus vampyrus*	XP_011358803.1	Mammalia
31	*Rattus norvegicus*	CAA90414.1	Mammalia
32	*Sarcophilus harrisii*	XP_012407378.1	Mammalia
33	*Sorex araneus*	XP_004617937.1	Mammalia
34	*Sus scrofa*	AAA85588.1	Mammalia
35	*Trichechus manatus latirostris*	XP_004383413.1	Mammalia
36	*Tupaia chinensis*	XP_014441212.1	Mammalia

37	*Tursiops truncatus*	XP_019788214.1	Mammalia
38	*Alligator mississippiensis*	KYO35091.1	Reptilia
39	*Anolis carolinensis*	ALA62321.1	Reptilia
40	*Caiman crocodilus*	AAK92227.1	Reptilia
41	*Chelonia mydas*	XP_007057987.1	Reptilia
42	*Ophiophagus hannah*	ETE66901.1	Reptilia
43	*Python bivittatus*	XP_007425899.1	Reptilia
44	*Xenopus laevis*	AAO65388.1	Amphibia
45	*Xenopus tropicalis*	XP_002938667.2	Amphibia
46	*Latimeria chalumnae*	XP_006011890.1	Sarcopterygii
47	*Lepisosteus oculatus*	AMD08894.1	Actinopterygii

Table S2. Helical regions of mouse Ambn as calculated by Heliquest. AH: Amphipathic Helix. The sequences from 239 to 260 were not assigned as amphipathic helix because these sequences had proline in the middle of the sequences.

Potential AH motifs	Net charge	Hydrophobic moment	Hydrophobic face	AH
^{58}LQGLNALSQYSRLGFGKA^{75}	2	0.233	Y G G L A	Y
^{59}QGLNALSQYSRLGFGKAL^{76}	2	0.233	Y G G L A	Y
^{60}GLNALSQYSRLGFGKALN^{77}	2	0.255	Y G G L A	Y
^{69}RLGFGKALNSLWLHGLLP^{86}	2	0.129	A P L F G L	Y
^{239}AYGTLFPRFGGFRQTLRR^{256}	4	0.406	A F L L F Y	?
^{240}YGTLFPRFGGFRQTLRRL^{257}	4	0.450	L F L L F Y	?
^{241}GTLFPRFGGFRQTLRRLN^{258}	4	0.399	L F L L F	?
^{242}TLFPRFGGFRQTLRRLNQ^{259}	4	0.411	L F L L F	?
^{243}LFPRFGGFRQTLRRLNQN^{260}	4	0.422	L F L L F	?

Figure S1. Characterization of WT mouse Ambn and its mutants. HPLC system (Varian) equipped with a Reversed Phase Phenomenex C4 column (10 × 250 mm, 5 μm) was used to purify Ambn. The purified protein was lyophilized, kept at -20°C, and dissolved in distilled water before use. The concentration of protein was determined using a Pierce BCA protein assay kit (Thermofisher). SDS-PAGE and mass spectrometry confirmed that the purified protein was Ambn. (A) SDS-PAGE showed that the purity of wild type Ambn was of sufficient quality for biochemical and biophysical experiments, while the apparent molecular weight was higher than the theoretical value. (B) Mass spectra of the band around 50 kDa in SDS-PAGE showed that the exact molecular weight of the purified protein was close to the theoretical value (41459.8 Da), suggesting that the purified protein was Ambn. (C) SDS-PAGE of AmbnW80F/W93F, AmbnW402F, AmbnΔ5 (36547.3 Da) and AmbnΔ6 (33991.4 Da) variants. (D) The sequence of mouse Ambn. The signal peptide is highlighted in yellow. The purified Ambn does not include this signal peptide. The sequence encoded by exon 5 is highlighted in green and is deleted in AmbnΔ5. The sequence encoded by exon 6 is highlighted in pink and is deleted in AmbnΔ6. W80 and W93 are in red and are replaced by F in W80F/W93F. W402 is in blue and replaced by F in W402F.



D

```
MSASKIPLFK MKGLILFLSL VKMSLAVPAF PQQPGAQGMA PPGMASLSLE TMRQLGSLQG LNALSQYSRL GFGKALNSLW 80
LHGLLPPHNS FPWIGPREHE TQQPSLQPHQ PGLKPFLQPT AATGVQVTPQ KPGPQPPMHP GQLPLQEGEL IAPDEPQVAP 160
SENPPTPEVP IMDFADPQFP TVFQIARSIS RGPMAHNKAS AFYPGMFYMS YGANQLNAPA RIGFMSSEEM PGERGSPMAY 240
GTLFPRFGGF RQTLRRLNQN SPKGGDFTVE VDSPVSVTKG PEKGEGPEGS PLQEANPGKR ENPALLSQMA PGAHAGLLAF 320
PNDHIPSMAR GPAGQRLLGV TPAAADPLIT PELAEVYETY GADVTTPLGD GEATMDITMS PDTQQPLLPG NKVHQPQVHN 400
AWRFQEP
```

Figure S2. DLS regularization curves showing particle size distribution (A, C, E, G, I &K) and correlation function (B, D, F, H, J & L) of WT Ambn, AmbnΔ6, and AmbnΔ5 at concentrations 1 and 3 µM in PBS buffer, pH 7.4. Lyophilized proteins were dissolved in pH 7.4, phosphate-buffered saline (PBS) composed of 10 mM sodium phosphate and 50 mM NaCl, and were diluted into 1 or 3 µM in pH 7.4 10.0 mM PBS buffer. The diluted protein solutions were kept on ice. The protein samples were taken out from the ice and were kept at room temperature for 1 hour before measurement. DLS was performed at room temperature using a DynaPro Nanostar light scattering instrument (Wyatt) with 10 µl cell. Each regularization curve was an average of ten acquisitions (10 scans for each acquisition). The data were analyzed using Dynamics 7.0, and the reported values were the hydrodynamic radii based on a Rayleigh sphere model.

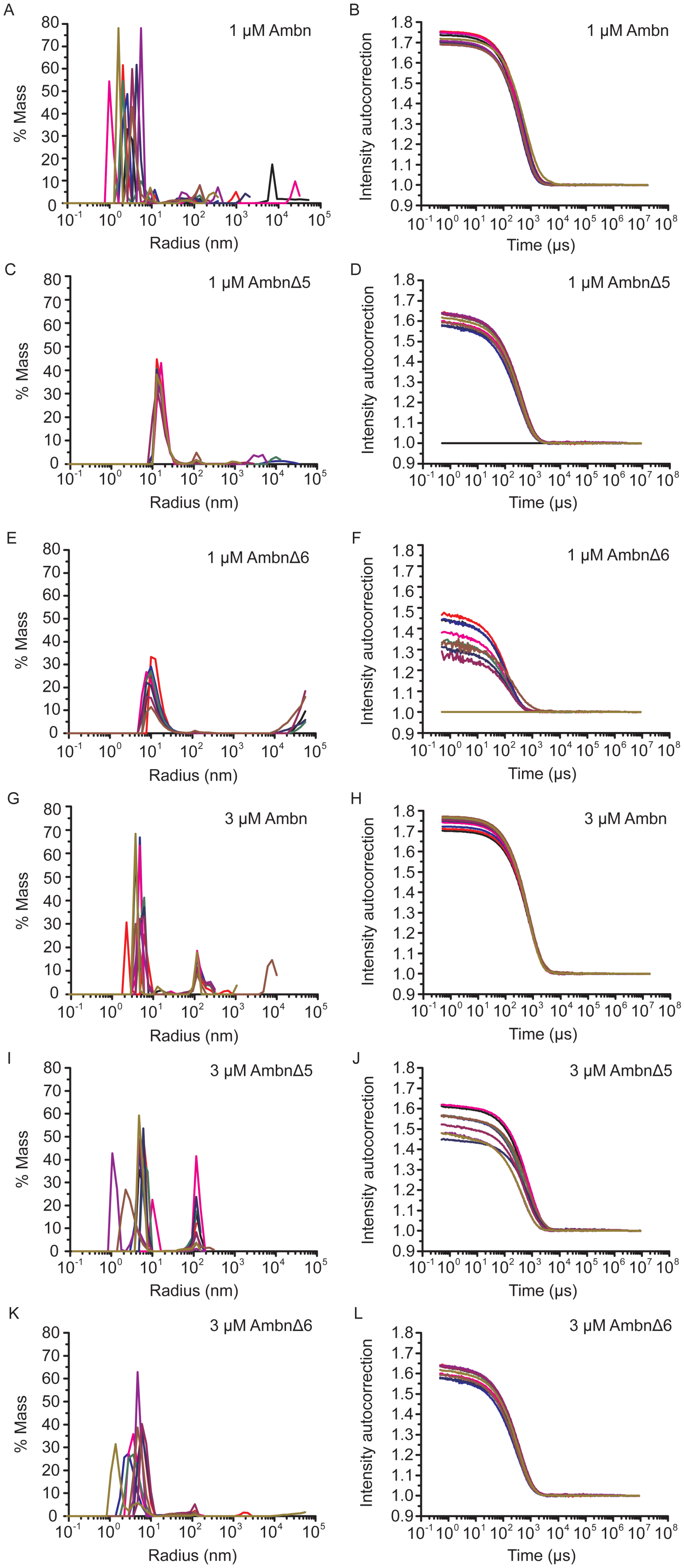

A
1 μM Ambn
% Mass
Radius (nm)
B
1 μM Ambn
Intensity autocorrection
Time (μs)
C
1 μM AmbnΔ5
% Mass
Radius (nm)
D
1 μM AmbnΔ5
Intensity autocorrection
Time (μs)
E
1 μM AmbnΔ6
% Mass
Radius (nm)
F
1 μM AmbnΔ6
Intensity autocorrection
Time (μs)
G
3 μM Ambn
% Mass
Radius (nm)
H
3 μM Ambn
Intensity autocorrection
Time (μs)
I
3 μM AmbnΔ5
% Mass
Radius (nm)
J
3 μM AmbnΔ5
Intensity autocorrection
Time (μs)
K
3 μM AmbnΔ6
% Mass
Radius (nm)
L
3 μM AmbnΔ6
Intensity autocorrection
Time (μs)
0
10
20
30
40
50
60
70
80
0.9
1.0
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
10^{-1} 10^{0} 10^{1} 10^{2} 10^{3} 10^{4} 10^{5}
10^{-1} 10^{0} 10^{1} 10^{2} 10^{3} 10^{4} 10^{5} 10^{6} 10^{7} 10^{8}

Figure S3. CD of Trp variants, AmbnW80F/W93F and AmbnW402F, compared to WT Ambn showing that the secondary structures were not affected by variations of Trp to Phe.

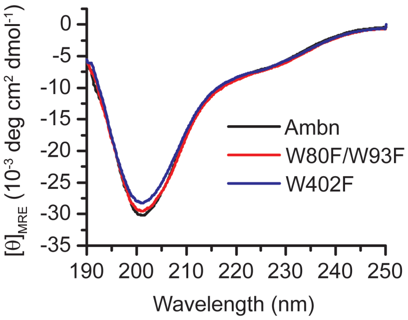


Figure S4. CD spectra of WT Ambn, AmbnΔ5 and AmbnΔ6 samples used for the study of LUV-protein interactions.