Asynchronous Evolutionary Origins of $A\beta$ and BACE1

D. Blaine Moore,¹ Madelyn A. Gillentine,¹ Nathalie M. Botezatu,¹ Kyle A. Wilson,¹ Ashley E. Benson,¹ and James A. Langeland^{*,1}

¹Department of Biology, Kalamazoo College, Kalamazoo, MI ***Corresponding author:** E-mail: Jim.Langeland@kzoo.edu.

Associate editor: Gregory Wray

Abstract

Neurodegenerative plaques characteristic of Alzheimer's disease (AD) are composed of amyloid beta (A β) peptide, which is proteolyzed from amyloid precursor protein (APP) by β -secretase (beta-site APP cleaving enzyme [BACE1]) and γ -secretase. Although γ -secretase has essential functions across metazoans, no essential roles have been identified for BACE1 or A β . Because their only known function results in a disease phenotype, we sought to understand these components from an evolutionary perspective. We show that APP-like proteins are found throughout most animal taxa, but sequences homologous to A β are not found outside gnathostomes and the β cut site is only conserved within sarcopterygians. BACE1 enzymes, however, extend through basal chordates and as far as cnidaria. We then sought to determine whether BACE1 from a species that never evolved A β could proteolyze APP substrates that include A β . We demonstrate that BACE1 from a basal chordate is a functional ortholog that can liberate A β from full-length human APP, indicating BACE1 activity evolved at least 360 My before A β .

Key words: Amyloid beta, APP, BACE, evolution.

Alzheimer's disease (AD) is characterized by progressive cognitive decline and the presence of extracellular plaques composed of amyloid beta (A β). The A β peptide is a 4-kDa fragment derived from sequential proteolysis of the amyloid precursor protein (APP) by β - and γ -secretase (Glenner and Wong 1984). Production of A β is precluded by activation of an alternative proteolytic pathway involving α -secretase, which is known to compete with β -secretase for cleavage of APP (Skovronsky et al. 2000).

Despite significant understanding of proteolytic pathways related to APP, little is known about its functional role. Indeed, some have characterized APP as an "all-purpose protein" (Shariati and De Strooper 2013) reflecting a vast array of proposed functions. These include dendritic spine formation (Tyan et al. 2012), neuronal cell migration (Young-Pearse et al. 2007), cell–cell adhesion (Soba et al. 2005), synaptogenesis (Wang et al. 2009), and long-term potentiation (Dawson et al. 1999). APP-like proteins (APPLPs), which lack the A β motif, are proteolyzed similar to APP but appear to have distinct, and quite different, roles (Shariati and De Strooper 2013).

 β -Secretase activity is thought to be the rate-limiting step in A β production (Citron et al. 1995). The beta-site APP cleaving enzyme (BACE1) responsible for β -secretase activity contains two consensus motifs [D(T/S,G(T/S)] characteristic of an aspartyl protease (Vassar et al. 1999). BACE1 is highly expressed in neurons in humans (Irizarry et al. 2001), is known to account for the majority of A β produced in the human brain (Ahmed et al. 2010), and BACE1 knockout mice are viable but do not produce A β (Luo et al. 2001).

BACE2 is a paralog containing 45% identity with BACE1 (Bennett et al. 2000). BACE2 expression in the brain is low or

undetectable (Bennett et al. 2000), and the enzyme is not a significant source of A β in humans (Ahmed et al. 2010). Hussain et al. (2000) have documented BACE2 cleavage of APP at the β -secretase cut site, with paradoxically no increase in A β . Indeed, Basi et al. (2003) demonstrated a role for BACE2 in lowering levels of secreted A β , and others have shown that the enzyme functions as an alternative α -secretase that precludes A β production by cleaving APP within the A β domain at a novel "theta" site just downstream from the alpha position (Sun et al. 2005).

Although much is known about the function of γ -secretase components across metazoans (reviewed in Zhang et al. [2013]), very little is understood regarding the normal role of two key components of Alzheimer's pathogenesis, the monomeric β -secretase and the A β peptide. Because their only known function results in a disease phenotype, we sought to understand these components from a functional evolutionary perspective.

Phylogenetic Analysis of APP, $A\beta$, and BACE1

We searched across animal taxa for APP, A β , and BACE1 sequences. Iterative searching of protein, transcript, and genomic databases (see Materials and Methods) indicates that APPLPs are found in all major animal taxa except sponges and ctenophores. Sequence alignments reveal extensive conservation across APP and APPLPs from all these taxa, including several previously characterized domains (Li and Südhof 2004; data not shown). Alignments also reveal that sequences homologous to A β are only found in APPs of gnathostomes (fig. 1A). Among the partially sequenced chondrichthyan genomes, *Raja erinacea* (little skate), has a clear APP/A β

[©] The Author 2013. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Α	571	β	a <u>myloid co</u> re
Homo APP	ADRGLTTRPGSGLTNI KTE	EISEVKMDAEFRHDSGY	EVHHQKL VFFAEBVGSNKGAIIGL
Mus APP	ADRGLTTRPGSGLTNI KTE	ELSEVKMDAEFGHDSGF-	EVRHQKL VFFAEDVGSNKGAIIGL
Gallus APP	ADRGLTTRPGSGLTNVKTE	EVSEVKMDAEFRHDSGY	EVHHQKL VFFAEDVGSNKGAIIGL
Xenopus APP	ADRGLTTRPGSGLTNI KTE	EISEVKMDSEYRHDTAY	EVHHQKL VFFAEEVGSNKGAIIGL
Latimeria APP	ADRGLTTRPGSGLTGI KTE	EISEVKMDAEFRHDSGY	EVHHQKL VFFAEDVGSNKGAIIGL
Danio APP1	POLDLATRPVSGLKPD	DIFELRMEAEERHS	EVYHOKL VFFAEDVSSNKGAIIGL
Danio APP2	PERGLPTRP	EIFKVRLDIEERHNAGY	DVRDKRLWFLAEDMGSNKGAIIGL
RajaApp	PDRGLTTRPGSGLTGIKME	ELAEVKMETEFRODSGY	EVHHQKL VFF PEDVGSNKGAIIGL
Branchiostoma APPLP	DDDDKPAKMDYFGDD	EDEDAEGEGQAAKKVLVT	PQHLQKFAASKADRPPQAASIVKHITNLTTLSL
Homo APPLP1		YERKVNASVPRGFPFH	SSEI QRDELAPAGTGVSREAVSG L
Homo APPLP2	I DETL	DVKEMIFNAERVGGL-	EEERESVGPLREDFSLSSSALIGL
Drosophila APPLP	PDAALVTAANPNLETTKSEKDLSD	TEYGEATVSSTKVQTVLP-	TVDDDAVQRAVEDVAAKEGRNVYFTL
Caenorhabditis APPLP	VETSVQTDDEDDDEDSSSS	TSESDEDEDKNIKELRV	DIEPIIDEPASPEVERSASSVFQPYV
Nematostella APPLP		RHGNSKKHRKHDEED-	EPENERDFFGEPRSPATFAAVIGL
	γ40 γ42		695
Homo APP	MVGGVVIATVIVITLVMLKKKQ	VTCI UUCVVEVBAAVTD	L FERHISKMOONGVENPTVKFFFOMON
		- ITSTANGVVEVDAAVIP	
Mus APP	MVGGVVIATVIVITLVMLKKKQ	- YTSI HHGVVEVDAAVTP	EERHLSKMQQNGYENPTYKFFEQMQN
Mus APP Gallus APP	MVGGVVI ATVIVITLVMLKKKQ MVGGVVI ATVIVITLVMLKKKQ	- YTSI HHGVVEVDAAVTP - YTSI HHGVVEVDAAVTP	E E RHL SKMQQNGY ENP TYKFF E QMQN
Mus APP Gallus APP Xenopus APP	MVGGVVI ATVIVITLVMLKKKQ MVGGVVI ATVIVITLVMLKKKQ MVGGVVI ATVIVITLVMLKKKQ	- YTSI HHGVVEVDAAVTP - YTSI HHGVVEVDAAVTP - YTSI HHGVVEVDAAVTP - YTTI HHGVVEVDAAVTP	E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL TKMQQNGY ENP TYKFF E QMQN
Mus APP Gallus APP Xenopus APP Latimeria APP	MVGGVVIATVIVITLVMLKKKO MVGGVVIATVIVITLVMLKKKO MVGGVIATVIVITLVMLKKKO MVGGIIIATVIVITLVMLKKKO	- YTSI HHGVVEVDAAVTP - YTSI HHGVVEVDAAVTP - YTTI HHGVVEVDAAVTP - YASI HHGVVEVDAAVTP	E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL TKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN
Mus APP Gallus APP Xenopus APP Latimeria APP Danio APP1	MVGGVVIATVIVITLVMLKKKQ- MVGGVVIATVIVITLVMLKKKQ- MVGGVIATVIVITLVMLKKKQ- MVGGVIATVIVITLVMLKKKQ- MVGGVVIATVIVITLVMLKKQ-	YTSI HHGVVEVDAAVTP YTSI HHGVVEVDAAVTP YTTI HHGVVEVDAAVTP YASI HHGVVEVDAAVTP YTSI HHGLI EVDAAVTP	E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL TKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMHN
Mus APP Gallus APP Xenopus APP Latimeria APP Danio APP1 Danio APP2	MVGGVVIATVIVITLVMLKKKQ- MVGGVVIATVIVITLVMLKKKQ- MVGGVVIATVIVITLVMLKKKQ- MVGGVIATVIVITLVMLKKKQ- MVGGVVIATIVITLVMLRKKQ- MVGGVVIATVIVITLVMLRKKQ-	- YTSI HHGVVEVDAAVTP - YTSI HHGVVEVDAAVTP - YTTI HHGVVEVDAAVTP - YASI HHGVIEVDAAVTP - YTSI HHGI EVDAAVTP - YTSI HHGV	E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL TKMQQNGY ENP TYKFF E QMQN E ERHL TKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN
Mus APP Gallus APP Xenopus APP Latimeria APP Danio APP1 Danio APP2 RajaApp	MVGGVVI ATVIVITLVMLKKKQ MVGGVVI ATVIVITLVMLKKKQ MVGGVVI ATVIVITLVMLKKKQ MVGGVVI ATVIVITLVMLKKKQ MVGGVVI ATVIVITLVMLKKKQ MVGGVVI ATVIVITLVMLKKKQ MVGGVVI ATVIVITLVMLKKKQ	 YTSI HHGVVEVDAAVTP YTSI HHGVVEVDAAVTP YTTI HHGVVEVDAAVTP YTTI HHGVEVDAAVTP YTSI HHGVEVDAAVTP YTSI HHGVEVDAAVTP YTSI HHGVEVDAAVTP 	E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN
Mus APP Gallus APP Xenopus APP Latimeria APP Danio APP1 Danio APP2 RajaApp Branchiostoma APPLP	MVGGVVIATVIVITLVMLKKKQ MVGGVVIATVIVITLVMLKKKQ MVGGVIATVIVITLVMLKKKQ MVGGVIATVIVITLVMLKKKQ MVGGVVIATVIVITLVMLKKKQ MVGGVVIATVIVITLVMLKKKQ TVAAAILAVIVITLVMLKKKQ TVAAAILAVI	- YTSI HHGVVEVDAAVTP - YTSI HHGVVEVDAAVTP - YTTI HHGVVEVDAAVTP - YASI HHGVIEVDAAVTP - YTSI HHGVIEVDAAVTP - YTSI HHGVIEVDAAVTP - YTSI HHGVIEVDAAVTP MRQPGVNHAVVEVDPSMTP	E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL TKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMHN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QM
Mus APP Gallus APP Xenopus APP Latimeria APP Danio APP1 Danio APP2 RajaApp Branchiostoma APPLP Homo APPLP1	MVGGVVIATVIVITLVMLKKKQ MVGGVVIATVIVITLVMLKKKQ MVGGVIATVIVITLVMLKKKQ MVGGVIATVIVITLVMLKKKQ MVGGVVIATVIVITLVMLKKKQ MVGGVVIATVIVITLVMLKKKQ MVGGVVIATVIVITLVMLKKKQ IMGACCCSLVVSMLLIRAKKP	YTSI HHGVVEVDAAVTP - YTSI HHGVVEVDAAVTP - YTTI HHGVVEVDAAVTP - YASI HHGVIEVDAAVTP - YTSI HHGI EVDAAVTP - YTSI HHGVIEVDAAVTP MROPGVNHAVVEVDAAVTP - YGA SHGVVEVDPMLT	E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMMN E ERHL SKMQQNGY ENP TYKFF E QMMN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QM E ERHL SKMQQNGY ENP TYKFF E QM
Mus APP Gallus APP Xenopus APP Latimeria APP Danio APP1 Danio APP2 RajaApp Branchiostoma APPLP1 Homo APPLP1 Homo APPLP2	MVGGVVIATVIVITLVMLKKKQ MVGGVVIATVIVITLVMLKKKQ MVGGVIATVIVITLVMLKKKQ MVGGVIATIVITLVMLKKKQ MVGGVVIATIVITLVMLKKKQ MVGGVVIATVIVITLVMLKKKQ TVAAAILAVIIVITLVMLKKKQ TVAAAILAVIIVITLVMLKKKQ LIMGACCELIVISIVIRKKA LIMGACELIVISIVASIVA	YTSI HHGVVEVDAAVTP - YTSI HHGVVEVDAAVTP - YTSI HHGVVEVDAAVTP - YTSI HHGVIEVDAAVTP - YTSI HHGVIEVDAAVTP - YTSI HHGVIEVDAAVTP - YTSI HHGVIEVDAAVTP MRQP GVNHAVEVDAAVTP - YGSI SHGVEVDPMLTP - YGSI SHGVEVDPMLTP	E ERHL SKMQQ NGY ENP TYKFF E QMQN E ERHL SKMQ E ENP TYKFF E QMQN E ERHL SKMQ NGY ENP TYKFF E QMQN E E ERHL SKMQ NGY ENP TYKFF E QMQN E E E E E E E E E E E E E E E E E E E
Mus APP Gallus APP Xenopus APP Latimeria APP Danio APP1 Danio APP1 Branchiostoma APPLP Homo APPLP2 Homo APPLP2 Dosophila APPLP	MVGGVVI ATVIVITLVMLKKQ MVGGVVI ATVIVITLVMLKKKQ MVGGVVIATVIVITLVMLKKKQ MVGGVIATVIVITLVMLRKKQ MVGGVVIATVIVITLVMLRKKQ MVGGVVIATVIVITLVMLRKKQ TVBAALIAVIVIVITLVMLKKKQ LIMEAGGGSLVVSNLLRRKRP LVBAQAIATVIVISLVMLKKQ SFAEIALMAAVVSQAVAKWRTS	YTSI HHGVVEV DAAVTP - YTSI HHGVVEV DAAVTP - YTSI HHGVVEV DAAVTP - YTSI HHGVI EV DAAVTP - YTSI HHGVI EV DAAVTP - YTSI HHGVI EV DAAVTP MROPGVNHAVVEV DPSMTP - YTSI HHGVVEV DPSMTP - YTSI HHGVVEV DPSMTP - YGAI SHGVVEV DPMLTE - SPHAOGT I EV DONVTT	E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QS E ERHVSKMQQNGY ENP TYKFF E QS E ERHVSKMQN HSY ENP TYKFF E QS E ERHVSKMN HSY E E ENP TYKFF E QS E ERHVSKMN HSY E E ENP TYKFF E QS E E ERHVSKMN HSY E E ENP TYKFF E QS E E E E HSYN HSYN E E E E E E E E E E E E E E E E E E E
Mus APP Gallus APP Xenopus APP Latimeria APP Danio APP1 Danio APP2 RajaApp Branchiostoma APPLP Homo APPLP1 Homo APPLP2 Drosophila APPLP Caenorhabditis APPLP	MVGGVVI ATVIVITLVMLKKKQ MVGGVVI ATVIVITLVMLKKKQ MVGGVIATVIVITLVMLKKKQ MVGGVIATVIVITLVMLKKKQ MVGGVVIATVIVITLVMLKKKQ TVAAAIAVIVITLVMLKKKQ TVAAAIAVIVITLVMLKKKQ TVAAAIAVIVITLVMLKKKQ LIMGACGCELIVEMLLKKKP LVAAAIAVIVISVMLKKRD SFASJALMAAVVVVAVMKKS SFASJALMAAVVVVAVMKKS SFASJALMAAVVVVAVMKKS SFASJALMAAVVVVAVMKKS SFASJALMAAVVVVAVMKKS SFASJALMAAVVVVAVMKKS SFASJALMAAVVVVAVMKKS	YTSI HHGVVEVDAAVTP YTSI HHGVVEVDAAVTP YTSI HHGVVEVDAAVTP YTSI HHGVIEVDAAVTP YTSI HHGVIEVDAAVTP YTSI HHGVIEVDAAVTP YTSI HHGVVEVDAAVTP MRQPGVNHAVVEVDPMLTP - YGAI SHGVVEVDPMLTP - YGAI SHGVVEVDPMLTP - RSPHAQGFIEVDQNVTT - RRAMRGFIEVDVYTP-	E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL TKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E QMHN E ERHL SKMQQNGY ENP TYKFF E QMQN E ERHL SKMQQNGY ENP TYKFF E VYK E ERHL SKMQQNGY ENP TYKFF E VYK E ERHL SKMQQNGY ENP TYKFF E SKA



С



Fig. 1. Widely separate origins of A β and BACE 1. (A) Alignment of C-terminus ends of APP and APPLP sequences from selected animal species that represent major phylogenetic branches (some unique unaligned sequences trimmed for compactness). APPLPs are found throughout most animal taxa, extending basally as far as cnidaria (e.g., *Nematostella*). However, sequences homologous to A β (gray shaded box), including the central amyloidogenic core (A β 16-22, overlined) are not found outside of gnathostomes (e.g., *Raja*). Canonical gamma secretase cut sites are conserved throughout all

ortholog. Although several exons of APP-like sequences were found in the incomplete genomes of sea lamprey (*Petromyzon marinus*; Smith et al. 2013) and Japanese lamprey (*Lethenteron japnonicum*; Mehta et al. 2013; data not shown), no matches to $A\beta$ sequences were found in these or other non-gnathostome genomes, even under relaxed search constraints (see Materials and Methods). Our APP phylogenetic tree (fig. 1*B*) confirms and extends the analysis of Tharp and Sakar (2013), who used a different subset of species and identified three major clades of vertebrate APP and APPLPs. All $A\beta$ -containing proteins form a very well-supported clade and should be considered true APP orthologs.

Among APPs that share the A β motif, all also share the two gamma secretase sites (γ 40 and γ 42) and most also share the seven amino acid motifs (KLVFFAED) that has been demonstrated to be the central core of amyloid aggregates (Tjernberg et al. 1999). APP2 from Danio has a more divergent sequence in place of the amyloid core, indicating that it has been less constrained since teleost-specific genome duplication (Taylor et al. 2001). We find that the coelacanth, Latimeria chalumnae, is the most basal species containing the β -secretase recognition and proteolysis sequence (EVKMDAE; Li and Südhof 2004). This cut site is highly conserved within sarcopterygians (one amino acid change in Xenopus) but was not found in teleost (Danio, shown; Fugu, Oryzias, Gadus; not shown), other actinopterygians (Polypterus, Acipenser, Amia, and Lepisosteus; not shown), or chondrichthyan genomes. Some of these nonteleost genomes remain incomplete, and it is thus possible that they may harbor β -secretase recognition sequences. Such a finding would alter our proposed timing of the origin of the cut site and would indicate that modern teleosts secondarily lost this sequence. We were unable to detect signals of positive selection across A β and flanking sequences or anywhere in APP (data not shown). Non-A β -containing APPLPs form two separate, gnathostome-specific clades, APPLP1 and APPLP2. Invertebrate APPLPs and lamprey APPLP exons (not shown) group outside the cluster of gnathostome APP and APPLPs. Together, these analyses indicate that gnathostome vertebrates underwent duplications of ancestral APPLP genes. One APPLP paralog acquired the A β motif in at least the gnathostome ancestor.

Our analysis of BACE1 (fig. 1C) indicates that orthologs are widely present among animal taxa. Bona fide BACE1 orthologs can be found throughout the chordates, including all taxa that have $A\beta$ in their APPs, and also in taxa such as *Branchiostoma* that do not have $A\beta$ and only have APPLPs. Although there are some invertebrate taxa that are

conspicuously lacking BACE1 orthologs (e.g., *Drosophila* and *Caenorhabditis*), multiple edysozoans and lophotrochozoans have BACE1 orthologs (data not shown). Among basal animals, we found a bona fide BACE1 ortholog in the genome of hydra (*Hydra magnipapillata*). BACE2 orthologs appear to be limited to sarcopterygians, while proteins equally orthologous to BACE1 and BACE 2 can be found in ctenophores (*Mnemiopsis leidyi*) and choanoflagellates (*Monosiga bevicollis*). Cathepsin D is the closest family of aspartyl proteases to BACE1 and BACE2; our tree topology suggests that BACE1 properly arose by gene duplication within this family just after the origin of Eumetazoa (fig. 1C).

BACE1 Functional Evolutionary Analysis

Our phylogenetic analysis indicates that $A\beta$ peptide arose in early gnathostomes and the β cut site arose in early sarcopterygians. BACE1, in stark contrast, arose near the base of the animal phylum, in the common ancestor of bilaterians and cnidaria. Fossil evidence places this ancestor at least 570 Ma (Xiao et al. 2000) and sarcopterygians at 390 Ma (Benton 1990). The corresponding molecular clock estimates are 1,100 Ma (Cartwright and Collins 2007) and 430 Ma (Blair and Hedges 2005), respectively. Given this enormous temporal difference in the appearance of enzyme and substrate, we sought to test the degree of functional conservation of BACE1 during animal evolution. Specifically, we wished to determine whether BACE1 enzymes from basal species that never evolved A β would proteolyze human APP/A β substrates. BACE1 from Branchiostoma floridae was our test of this functional conservation. Branchiostoma are the sole surviving genus of the cephalochordate subphylum, the most basal chordate taxon (Putnam et al. 2008), and separated from humans by 520–890 My (Shu et al. 1996; Blair and Hedges 2005).

Although BACE1 from cnidarians (e.g., *Hydra*) are marginally more divergent from humans (570–1,100 Ma), we reasoned that functional conservation across this additional divergence time would not result in a significantly more informative result. Given the more well-developed and readily available cDNA resources (Langeland et al. 1998), we reasoned that cloning and expressing the *Branchiostoma* BACE1 cDNA would prove more tractable, yet still provide a robust test of deep functional conservation.

For our analysis of BACE1 functional conservation, we chose a well-described system wherein the shortest, primarily neuronal isoform of human APP (APP695) is stably expressed in a Chinese Hamster ovary cell line (CHO 695 cells; Skovronsky et al. 2000). CHO 695 cells were transiently

FIG. 1. Continued

A β -containing APPs, while the canonical β cut site (overlined and arrow) are only conserved within sarcopterygians (e.g., *Latimeria*). (B) Maximum likelihood tree of APP and APPLPs. Nodes collapsed below 60% bootstrap support. A β -Containing APPs form a very well-supported clade (bootstrap support 99.4%), indicating that they are true orthologs. Several tetrapod APPLPs and all non-gnathostome APPLPs fall outside of the A β -containing clade. (C) Maximum likelihood tree of BACE1 and BACE2 sequences. Nodes collapsed below 60% bootstrap support. Bona fide BACE1 orthologs extend beyond sarcopterygians, through basal chordates (e.g., *Branchiostoma*), and extend basally as far as cnidaria (e.g., *Hydra*). BACE2 proteins have not been found outside of sarcopterygians. Sequences equally orthologous to BACE1 and BACE2 are found in ctenophores (e.g., *Mnemiopsis*) and choano-flagellates (e.g., *Monosiga*). The tree was rooted on cathepsin D as this is the closest family of aspartyl proteases to BACE1 and BACE2.

transfected with Branchiostoma BACE1 for 16 h. Conditioned media was collected and subjected to ELISA for human $A\beta$ 1-40, a major A β species produced by sequential β -secretase and γ -secretase proteolysis. For comparative purposes, Homo BACE1 was used as a positive control and GFP as a negative control (fig. 2). One-way analysis of variance (ANOVA) revealed a significant construct effect [F(2,18) = 4.686; P < 0.05]. Bonferroni post hoc analysis determined A β secretion was elevated compared with GFP controls after transfection of either Branchiostoma BACE1 or Homo BACE1 (P < 0.05) Indeed, transfection with Branchiostoma BACE1 elevated AB secretion by nearly 2-fold over GFP control, and was indistinguishable from Homo BACE1 in terms of its ability to proteolyze APP. As such, BACE1 from this basal chordate cuts human APP at the β cleavage site and represents a true functional ortholog.

Conclusion

We have demonstrated that BACE1 predates the origin of its A β substrate by at least 360 My based on molecular clock estimates. Furthermore, our observation that BACE1 from Branchiostoma, a basal chordate that never evolved A β , can proteolyze human APP and liberate $A\beta$ indicates a very high level of functional conservation since the origin of BACE1. These two results indicate that BACE1 has deeply conserved and essential functions that have nothing to do with APP processing. Nonamyloidogenic roles for BACE1 are beginning to emerge, such as peripheral myelination (Hu et al. 2006) and maintenance of synaptic function (Sheng et al. 2003). Moreover, BACE1 is known to cleave substrates other than APP, including APPLPs that lack the A β domain (Li and Südhof 2004), NRG-1 (Hu et al. 2006), the lipoprotein receptor-related protein LRP-1 (von Arnim et al. 2005), the cell adhesion protein PSGL-1 (Lichtenthaler et al. 2003), and the immune regulator ST6Gal1 (Kitazume et al. 2005). We found orthologs of each of these genes in both Branchiostoma and cnidaria (data not shown), indicating that these proteins are likely members of the ancestral set of essential BACE1 substrates. Although not as deeply conserved as BACE1, the A β peptide has been conserved for at least 430 My. This indicates that A β , too, has essential functions that have thus far escaped discovery.

Our results are consistent with established trends documenting unexpected roles for familiar proteins (Perona 2009), such as modulation of Wnt signaling by the telomerase component hTERT (Park et al. 2009) and activation of angiogenesis by the histone H2AX (Economopoulou et al. 2009). Components of the APP processing machinery are also known to have essential roles independent of APP proteolysis. For example, γ -secretase is known to cleave non-APP substrates including Notch (Song et al. 1999) and ErbB4 (Ni et al. 2001); these activities may represent its earliest biological function in metazoans. Furthermore, presenilin (PS, the catalytic component of γ -secretase) performs nonproteolytic functions in plants. Remarkably, this nonmetazoan PS rescues the growth deficiency phenotype in PS-deficient mouse embryonic fibroblasts (Khandelwal et al. 2007), indicating unknown nonproteolytic roles for PS in animals.



FIG. 2. Branchiostoma BACE1 proteolyzes human APP at the β -secretase site. CHO 695 cells stably transfected with human APP were transiently transfected with GFP control, *Homo* BACE1 or *Branchiostoma* (*Branch.*) BACE1. Conditioned media were harvested after 16 h and secretion of human A β 1–40 determined via ELISA. (*P < 0.05).

On the basis of our data, we suggest that the evolutionary origin of A β accumulation in neurodegenerative plaques necessarily included the following steps (fig. 3): BACE1 and APPLPs had and maintain essential ancestral functions dating to near the origin of animals. In the gnathostome ancestor, one APPLP paralog evolved the still-essential A β motif. The amyloidogenic core and the γ cut site seem to appear in one step, indicating that the intermediates have been lost. The definitive β cut site appears somewhat later in the sarcopterygian ancestor. Either concurrently or later, the two proteins came to be coexpressed and into physical contact in neurons. Because >95% of Alzheimer's cases manifest after age 60 (Holmes 2002), A β accumulation should be selectively neutral and would not affect the proposed essential functions of BACE1 and A β . Unresolved questions stemming from our analysis include when and how BACE1 began acting on ancestral A β containing APP substrates.

Materials and Methods

Phylogenetic Analysis

For APP/A β searching, we used both full-length human APP695 isoform (NP_958817) and the 42 residue A β peptide from this isoform. For BACE1 searching, we used full-length human BACE1 (AAH36084). These sequences were used to exhaustively Blast the genomes of several species spanning progressively more basal animal taxa and also including choanoflagellates, yeast, and plants. Specific databases searched include GenBank, Ensemble, National Human Genome Research Institute, Joint Genome Institute, Skatebase, jlampreygenome.imcb.a-star.edu.sg, and the Broad Institute Origins of Mulicellularity database. Searches were current as of 9 November, 2013. To ensure maximum coverage, in addition to BlastP and PSI-Blast (where available) against protein sequences, we used TBlastN against transcript and genomic



Fig. 3. Timeline of BACE1 and A β evolution. Simplified animal phylogenetic tree annotated with key events demonstrated in this article. Numbers below nodes indicate divergence times in millions of years relative to the present. Top number is consensus estimate based on molecular clock data and bottom number is minimum number based on fossil evidence (see text for references). APPLP and BACE 1 proteins are shared with all animals inclusive of bilateria and cnidaria and thus arose at least 570 Ma and perhaps as long as 1,100 Ma. Our analysis of *Branchiostoma* BACE1 demonstrates that BACE1 function was present in the chordate ancestor at least 520 Ma and perhaps as long as 890 Ma. A β peptide arose in the common ancestor of jawed vertebrates and the canonical β cut site, on which BACE1 acts to release A β , arose in the common ancestor of tetrapods and lobe-finned fish, up to 600 My after the origin of BACE 1.

databases. We generally used default Blast settings (BLOSUM 62 matrix, 11/1 Gap penalties, a conditional compositional score matrix adjustment, and no filtering). We employed an iterative approach wherein e values, beginning at 10, were progressively raised until hits were generated. Regardless of whether or not they were previously annotated, the top hits from each search were return blasted against the human genome to determine their identity. Results were deemed negative if they did not return, respectively, A β , APP, or BACE. In taxa where we could not recover A β , we recovered a range of proteins including angiotensinogen, aquaporin, and esterase. For APP, negative candidates were typically cytoskeletal proteins such as myosin, tubulin, or ankyrin. In taxa where we could not recover BACE, the results were invariably cathepsins. Multiple sequence alignments were produced using T-coffee (Di Tommaso et al. 2011) with default settings. Maximum likelihood trees were produced using Mega5.2.2 (Tamura et al. 2007) with JTT substitution model and five gamma distributed rate categories. Trees were bootstrapped with 1,000 replicates, and the resulting consensus trees were visualized using FigTree v1.4 collapsing nodes below 60% bootstrap support. The following sequences were used for APP alignments, in order shown in figure 1B (GenBank Accession number, Ensemble sequence number, SkateBase Contig number, or jgi scaffold): NP 958817, NP 031497, AAG00593, NP_001082098, ENSLACG0000001469, NP_571639, NP_690842, SkateBase Contig91495, AAH00373, AAH52396,CAE75662, AAB96331, NP_031493, XP_002613121.1, NP_508871, NP_476626, XM_001637304. The following sequences were used for BACE alignments, in order shown in figure 1C: AAH36084, NP_035922, NP_001072485, ENSLACT0000020958, AAS64565, SkateBase Contig 11234, NP 991267, ENSPMAT 0000006274, Branchiostoma KJ018079, XP_002166168.2,

AAQ89286, NP_062390, AAS64566, NP_001080615, Skate-Base Contig 20593, ENSLACT00000017156, *Mnemiopsis* transcript ML154145a, jgi|Monbr1|28443|fgenesh2_pg.scaffold_26000054, NP_034113. Tests for positive selection were carried out using the codon-aligned APP nucleotide files in the Datamonkey implementation of HyPhy (Delport et al. 2010).

Cloning of Branchiostoma BACE1

Branchiostoma floridae genomic traces were blasted with human BACE1 cDNA sequence to obtain partial exon sequences. Primers were designed and used to amplify a 600-bp BACE1 fragment from our *B. floridae* cDNA library (Langeland et al. 1998). This fragment was used as a probe to screen for the full-length cDNA by hybridization. The resulting cDNA (2,790 bp; 500 amino acids) was subcloned into pcDNA3.1 for expression in mammalian cells.

Cell Culture and Transfection

CHO cells stably transfected with the 695 amino acid variety of human APP (CHO 695; gift of Dr Virginia M-Y. Lee, University of Pennsylvania) were maintained in minimum essential alpha media (αMEM; Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum, 1% glutamine, and 1% penicillin/streptomycin (Invitrogen). Cells were passaged as specified in Balan et al. (2010). Cells were seeded into sixwell plates at a density of 1.25×10^4 before transient transfection. CHO 695 cells at 70% confluence were transfected with GenePorter (Genlantis, San Diego, CA) 24 h after plating following manufacturer's instructions. Two micrograms of DNA (GFP, Homo BACE1, or Branchiostoma BACE1) and 10 µl transfection reagent was delivered in 0.5 ml serum-free α MEM for 5 h. Transfections were stopped by addition of 0.5 ml serum-containing α MEM. Visual inspection of GFP controls verified successful transfection.

Collection of Conditioned Media and ELISA

Conditioned media was harvested 16 h following transfection. A complete-mini protease inhibitor cocktail (Roche, Indianapolis, IN) was used to prevent degradation of secreted proteins. Media (1.0 mL) was transferred to microfuge tubes containing protease inhibitors and incubated on ice. Tubes were centrifuged at 13,000 rpm for 20 min at 4°C. Supernatant was transferred to new microfuge tubes and stored at -80°C until analysis. Secreted A β 1-40 was measured via a human-specific ELISA kit (Invitrogen) following manufacturer's instructions. Absorbance was measured at 450 nm on a µquant microplate spectrophotometer (Biotek, Winooski, VT).

Statistics

Significance of ELISA data was determined by one-way ANOVA followed by Bonferroni post hoc testing.

Acknowledgments

We gratefully thank Dr Virginia M-Y Lee, University of Pennsylvania, for the gifts of the CHO 695 cell line and the

Homo BACE1 construct; Carol McPherson and Kathryn Lightcap for technical assistance, and the Howard Hughes Medical Institute for summer research stipends to K.A.W. and A.E.B. Support for this project was provided by the Great Lakes Colleges Association as part of its New Directions Initiative, made possible by a grant from the Andrew W. Mellon Foundation.

References

- Ahmed RR, Holler CJ, Webb RL, Li F, Beckett TL, Murphy MP. 2010. BACE1 and BACE2 enzymatic activities in Alzheimer's disease. J Neurochem. 112:1045–1053.
- Balan AG, Myers BJ, Maganti JL, Moore DB. 2010. ER-targeted Bcl-2 and inhibition of ER-associated caspase-12 rescue cultured immortalized cells from ethanol toxicity. *Alcohol* 44:553–563.
- Basi G, Frigon N, Barbour R, Doan T, Gordon G, McConlogue L, Sinha S, Zeller M. 2003. Antagonistic effects of beta-site amyloid precursor protein-cleaving enzymes 1 and 2 on beta-amyloid peptide production in cells. J Biol Chem. 278:31512–31520.
- Bennett BD, Babu-Khan S, Loeloff R, Louis JC, Curran E, Citron M, Vassar R. 2000. Expression analysis of BACE2 in brain and peripheral tissues. *J Biol Chem.* 275:20647–20651.
- Benton MJ. 1990. Phylogeny of the major tetrapod groups: morphological data and divergence dates. J Mol Evol. 30:409-424.
- Blair JE, Hedges SB. 2005. Molecular phylogeny and divergence times of deuterostome animals. *Mol Biol Evol*. 22:2275–2284.
- Cartwright P, Collins A. 2007. Fossils and phylogenies: integrating multiple lines of evidence to investigate the origin of early major metazoan lineages. *Integr Comp Biol.* 47:744–751.
- Citron M, Teplow DB, Selkoe DJ. 1995. Generation of amyloid beta protein from its precursor is sequence specific. *Neuron* 14:661–670.
- Dawson GR, Seabrook GR, Zheng H, Smith DW, Graham S, O'Dowd G, Bowery BJ, Boyce S, Trumbauer ME, Chen HY, et al. 1999. Agerelated cognitive deficits, impaired long-term potentiation and reduction in synaptic marker density in mice lacking the beta-amyloid precursor protein. *Neuroscience* 90:1–13.
- Delport W, Poon AFY, Frost SDW, Kosakovsky Pond SL. 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* 26:2455–2457.
- Di Tommaso P, Moretti S, Xenarios I, Orobitg M, Montanyola A, Chang J-M, Taly J-F, Notredame C. 2011. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res.* 39:W13–17.
- Economopoulou M, Langer HF, Celeste A, Orlova VV, Choi EY, Ma M, Vassilopoulos A, Callen E, Deng C, Bassing CH, et al. 2009. Histone H2AX is integral to hypoxia-driven neovascularization. *Nat Med.* 15: 553–558.
- Glenner GG, Wong CW. 1984. Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem Biophys Res Commun.* 120:885–890.
- Holmes C. 2002. Genotype and phenotype in Alzheimer's disease. Br J Psychiatry. 180:131–134.
- Hu X, Hicks CW, He W, Wong P, Macklin WB, Trapp BD, Yan R. 2006. Bace1 modulates myelination in the central and peripheral nervous system. *Nat Neurosci.* 9:1520–1525.
- Hussain I, Powell DJ, Howlett DR, Chapman GA, Gilmour L, Murdock PR, Tew DG, Meek TD, Chapman C, Schneider K, et al. 2000. ASP1 (BACE2) cleaves the amyloid precursor protein at the beta-secretase site. *Mol Cell Neurosci.* 16:609–619.
- Irizarry MC, Locascio JJ, Hyman BT. 2001. Beta-site APP cleaving enzyme mRNA expression in APP transgenic mice: anatomical overlap with transgene expression and static levels with aging. Am J Pathol. 158: 173–177.
- Khandelwal A, Chandu D, Roe CM, Kopan R, Quatrano RS. 2007. Moonlighting activity of presenilin in plants is independent of

gamma-secretase and evolutionarily conserved. *Proc Natl Acad Sci U S A*. 104:13337–13342.

- Kitazume S, Nakagawa K, Oka R, Tachida Y, Ogawa K, Luo Y, Citron M, Shitara H, Taya C, Yonekawa H, et al. 2005. In vivo cleavage of alpha2,6-sialyltransferase by Alzheimer beta-secretase. J Biol Chem. 280:8589–8595.
- Langeland JA, Tomsa JM, Jackman WR Jr, Kimmel CB. 1998. An amphioxus snail gene: expression in paraxial mesoderm and neural plate suggests a conserved role in patterning the chordate embryo. *Dev Genes Evol.* 208:569–577.
- Li Q, Südhof TC. 2004. Cleavage of amyloid-beta precursor protein and amyloid-beta precursor-like protein by BACE 1. *J Biol Chem.* 279: 10542–10550.
- Lichtenthaler SF, Dominguez D-I, Westmeyer GG, Reiss K, Haass C, Saftig P, De Strooper B, Seed B. 2003. The cell adhesion protein P-selectin glycoprotein ligand-1 is a substrate for the aspartyl protease BACE1. *J Biol Chem.* 278:48713–48719.
- Luo Y, Bolon B, Kahn S, Bennett BD, Babu-Khan S, Denis P, Fan W, Kha H, Zhang J, Gong Y, et al. 2001. Mice deficient in BACE1, the Alzheimer's beta-secretase, have normal phenotype and abolished beta-amyloid generation. *Nat Neurosci.* 4:231–232.
- Mehta TK, Ravi V, Yamasaki S, Lee AP, Lian MM, Tay BH, Tohari S, Yanai S, Tay A, Brenner S, et al. 2013. Evidence for at least six Hox clusters in the Japanese lamprey (*Lethenteron japonicum*). Proc Natl Acad Sci U S A. 110:16044–16049.
- Ni CY, Murphy MP, Golde TE, Carpenter G. 2001. Gamma-secretase cleavage and nuclear localization of ErbB-4 receptor tyrosine kinase. *Science* 294:2179–2181.
- Park J-I, Venteicher AS, Hong JY, Choi J, Jun S, Shkreli M, Chang W, Meng Z, Cheung P, Ji H, et al. 2009. Telomerase modulates Wnt signalling by association with target gene chromatin. *Nature* 460: 66–72.
- Perona R. 2009. Old proteins harbour unexpected functions. *Clin Transl Oncol.* 11:489–490.
- Putnam NH, Butts T, Ferrier DEK, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu JK, et al. 2008. The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453:1064–1071.
- Shariati SAM, De Strooper B. 2013. Redundancy and divergence in the amyloid precursor protein family. *FEBS Lett.* 587:2036–2045.
- Sheng JG, Price DL, Koliatsos VE. 2003. The beta-amyloid-related proteins presenilin 1 and BACE1 are axonally transported to nerve terminals in the brain. *Exp Neurol.* 184:1053–1057.
- Shu D-G, Conway Morris S, Zhang X-L. 1996. A pikaia-like chordate from the Lower Cambrian of China. *Nature* 384:157–158.
- Skovronsky DM, Moore DB, Milla ME, Doms RW, Lee VM. 2000. Protein kinase C-dependent alpha-secretase competes with beta-secretase for cleavage of amyloid-beta precursor protein in the trans-golgi network. J Biol Chem. 275:2568–2575.
- Smith JJ, Kuraku S, Holt C, Sauka-Spengler T, Jiang N, Campbell MS, Yandell MD, Manousaki T, Meyer A, Bloom OE, et al. 2013. Sequencing of the sea lamprey (*Petromyzon marinus*) genome provides insights into vertebrate evolution. *Nat Genet.* 45:415–421, 421e1–2.
- Soba P, Eggert S, Wagner K, Zentgraf H, Siehl K, Kreger S, Löwer A, Langer A, Merdes G, Paro R, et al. 2005. Homo- and heterodimerization of APP family members promotes intercellular adhesion. *EMBO J.* 24:3624–3634.
- Song W, Nadeau P, Yuan M, Yang X, Shen J, Yankner BA. 1999. Proteolytic release and nuclear translocation of Notch-1 are induced by presenilin-1 and impaired by pathogenic presenilin-1 mutations. *Proc Natl Acad Sci U S A.* 96:6959–6963.
- Sun X, Wang Y, Qing H, Christensen MA, Liu Y, Zhou W, Tong Y, Xiao C, Huang Y, Zhang S, et al. 2005. Distinct transcriptional regulation and function of the human BACE2 and BACE1 genes. FASEB J. 19: 739–749.
- Taylor JS, Van de Peer Y, Braasch I, Meyer A. 2001. Comparative genomics provides evidence for an ancient genome duplication event in fish. *Philos Trans R Soc Lond B Biol Sci.* 356:1661–1679.

- Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24:1596–1599.
- Tjernberg LO, Callaway DJ, Tjernberg A, Hahne S, Lilliehöök C, Terenius L, Thyberg J, Nordstedt C. 1999. A molecular model of Alzheimer amyloid beta-peptide fibril formation. *J Biol Chem.* 274:12619–12625.
- Tyan S-H, Shih AY-J, Walsh JJ, Maruyama H, Sarsoza F, Ku L, Eggert S, Hof PR, Koo EH, Dickstein DL. 2012. Amyloid precursor protein (APP) regulates synaptic structure and function. *Mol Cell Neurosci.* 51: 43–52.
- Vassar R, Bennett BD, Babu-Khan S, Kahn S, Mendiaz EA, Denis P, Teplow DB, Ross S, Amarante P, Loeloff R, et al. 1999. Betasecretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* 286:735–741.
- von Arnim CAF, Kinoshita A, Peltan ID, Tangredi MM, Herl L, Lee BM, Spoelgen R, Hshieh TT, Ranganathan S, Battey FD, et al. 2005.

The low density lipoprotein receptor-related protein (LRP) is a novel beta-secretase (BACE1) substrate. *J Biol Chem.* 280: 17777–17785.

- Wang Z, Wang B, Yang L, Guo Q, Aithmitti N, Songyang Z, Zheng H. 2009. Presynaptic and postsynaptic interaction of the amyloid precursor protein promotes peripheral and central synaptogenesis. *J Neurosci.* 29:10788–10801.
- Xiao S, Yuan X, Knoll AH. 2000. Eumetazoan fossils in terminal proterozoic phosphorites? *Proc Natl Acad Sci U S A.* 97: 13684–13689.
- Young-Pearse TL, Bai J, Chang R, Zheng JB, LoTurco JJ, Selkoe DJ. 2007. A critical function for beta-amyloid precursor protein in neuronal migration revealed by in utero RNA interference. *J Neurosci.* 27: 14459–14469.
- Zhang S, Zhang M, Cai F, Song W. 2013. Biological function of presenilin and its role in AD pathogenesis. *Transl Neurodegener*. 2:15.