

Lineage name	Family name	Higher-order	PDBID	N-term stacking hairpin	Core strand order	C-term $\alpha/\beta$ unit	S3 loop catalytic residue <sup>^</sup>	VR1 (forest green)	VR2 (sand)	VR3 (light blue)
Germination protease	HybD-like	HybD-like	1C8B_A	descending helix	21364	absent, possible displacement by N-term insert	--	2 helical turns, extended loop	ascending loop	small loop extension
HybD	HybD-like	HybD-like	1CFZ_A	none	21364	absent	--	none	ascending helix	small loop extension
peptidyl hydrolase	pth	pth-like	2PTH_A	present, first strand is poorly-formed	21364	possible compensatory $\beta$ -helix from Cterm ext	--	none	ascending minimal loop	extended loop
5'-methylthio-adenosine nucleosidase	MTAN	PNP-like	6AYM_A	present	213645	present	H (absent in MTAN family)	none	poorly-structured extended loop	none
YgiD-like	PCAD	PCAD-Memo	2PW6A	present, both strands broken	213645 162534cp	present	H	extended loop	extended loop	extended loop, helix 5 and 6a-b
Memo	Memo	PCAD-Memo	3BCZ_A	present, elongated at turn	213645 162534cp	present	H	none	extended loop housing 2 strands which stack in a barrel-like structure w/ third strand from Nterm	extended loop, helix 5 and 6
tRNA deacylase	tRNA deacylase	tRNA deacylase	1YQE_A	present, somewhat smaller than typical	213645	present	--	2 $\alpha$ -helical turns	extended loop with poorly-defined structure	none

bacPAC2	bacPAC2	PAC2-like	2P90_A	present	213645	present, not prominent: both strand and helix are primitive	--	slight bulge-like extension	none	slight bulge-like extension
endolysin Ply	amidase-like	Zn-dep. exo-peptidase assemblage	1XOV_A	absent	213645	present	--	none	none	slightly elongated loop
carboxy-peptidaseA	carboxy-peptidaseA	Zn-dep. exo-peptidase assemblage	1M4L_A	Nterm stacking	213645	present	N	large poorly-structure loop with 1-2 helices	somewhat elongated loop	slight bulge in loop
leucine amino-peptidase	leucine amino-peptidase	Zn-dep. exo-peptidase assemblage	1LAM_A	Nterm stacking	213645	present	N	extended loop with $\beta$ -hairpin	none	extended loop, helical segment
amino-peptidase	dinuclear Zn exo-peptidase	Zn-dep. exo-peptidase assemblage	1RTQ_A	Nterm stacking	213645	present	EE	$\alpha$ -helical turn	none	extended loop
transferrin receptor protease	FoIH-like	Zn-dep. exo-peptidase assemblage	1DE4_C	Nterm stacking, elongated strands	213645	present, strand considerably degraded	D	slight loop extension	none	extended loop
Succinyl-glutamate desuccinylase AstE	AstE/AspA-like	Zn-dep. exo-peptidase assemblage	2G9D_A	Nterm stacking, additional N-term strand	213645	present	N	single helix leading into extended loop	slightly extended loop	slightly extended loop
Glutaminy-peptide cyclo-transferase	Glutaminy-peptide cyclo-transferase	Zn-dep. exo-peptidase assemblage	2AFW_A	Nterm stacking, slightly	213645 (S4 and S5 degraded)	present, strand considerably degraded	EE	extended loop with bulge-like structure	none	striking extended loop

	like			elongated strands						pointing upwards
N-formyl-glutamate amidohydrolase	FGase-like	Zn-dep. exopeptidase assemblage	2Q7S_A	single stacking strand at N-terminus	213645	present	N	highly elaborate extended loop with little structure	small extended bulge loop	mostly extended loop with 1-2 helical elements
pyro-glutamate aminopeptidase	pyro-glutamate aminopeptidase	pyro-glutamate aminopeptidase	1A2Z_A	none	213645	present	--	none	very long extended loop with flap-like $\beta$ -element. reminiscent of Memo	extended, upwards-pointing loop

Lineage name	Family name	Higher-order grouping	'VR4'	core strand 5 insert	Hairpin insert	Comments	HHPRED search	DALI search
Germination protease	HybD-like	HybD-like	helical cap with flap (dimerizing domain?)	n/a	n/a	elaborate version of other HybD	returns pth family profiles	hits to pth structures recovered first, often PNP structures are observed next
HybD	HybD-like	HybD-like	extended loop	n/a	n/a	minimal HybD	returns pth family profiles	pth is strongest, maybe pyro-glutamyl aminopeptidase is next? Then PNP hits
peptidyl hydrolase	pth	pth-like	helix cap	n/a	n/a	reverse reciprocity with HybD	returns HybD family profiles	PNP and minimal HybD structures recovered with about the same scores
5'-methylthioadenosine nucleosidase	MTAN	PNP-like	extended loop	helix with strand stacking	none	barrelizing structure formed at S5 insert with C-	returns exclusively PNP family profiles	returns several families with similar scores including PAC2, Zn-dep.

				between S3 and S5		terminal strands		exo-peptidases, and deacylases
YgiD-like	PCAD	PCAD-Memo	none	none	elongation at turn	--	PCAD tends to return low-scoring hits with tRNA deacylase profiles and some Zn-dep. exo-peptidases	difficult to discern, PCAD members return several other families with roughly equal scores
Memo	Memo	PCAD-Memo	none	lateral shelf has strongly helicized	elongation at turn	VR2 insert forms something like a barrel... distinct from PNP barrel	low-scoring hits to tRNA deacylase family profiles	difficult to discern, returns several families including PNP and the Zn-dep. exopeptidases with roughly equal scores
tRNA deacylase	tRNA deacylase	tRNA deacylase	small loop has helicized	some extension, interactions with VR1 and VR2	none	C-term domain helix stacks at the Cterm $\alpha/\beta$ unit	returns PCAD family profiles with strongest scores, Zn-dep. exo-peptidases at borderline significance	returns PNP and several other families at roughly equal scores
bacPAC2	bacPAC2	PAC2-like	1-2 $\alpha$ -helical turns	standard lateral shelf	poorly-formed helix gives way to hairpin insert	Nterm ext. stacks small strand against Nterm hairpin	none	PAC2-like family members and PNPs tend to be the best-scoring hits, followed by scattered Zn-dep. exo-peptidase assemblage hits
endolysin Ply	amidase-like	Zn-dep. exo-peptidase assemblage	helix cap	none	n/a	--	detects tRNA deacylase family profiles and other exopeptidases, notably the AstE_AspA family	after detecting amidases, appears to largely detect other families of the Zn-dep. exopeptidase assemblage

carboxy-peptidaseA	carboxy-peptidaseA	Zn-dep. exo-peptidase assemblage	slightly extended loop	lateral shelf with 1-2 $\alpha$ -helical turns	none	--	detects other peptidase families, with AstE_AspA, SpoIIP, and amidases as outgroups	sees other peptidase families with some PNP and polyglutamyl aminopeptidase structures also scoring well
leucine amino-peptidase	leucine amino-peptidase	Zn-dep. exo-peptidase assemblage	extended loop that folds back on itself	helicized lateral shelf with extended loop tracing backwards	2-helix insert	4 <sup>th</sup> sandwich layer formed by core S5, VR1, Nterm hairpin	returns "classical" Zn-dep. exo-peptidase profiles (top assemblage in Fig. 4)	mostly returns hits with other Zn-dep. exopeptidase family members
amino-peptidase	dinuclear Zn exo-peptidase	Zn-dep. exo-peptidase assemblage	extended cap-like loop with some helicity	lateral shelf with extended bulge	none	S6 is 'broken' by strange kind of spreading of core sheet	returns "classical" Zn-dep. exo-peptidase profiles (top assemblage in Fig. 4)	mostly returns hits with other Zn-dep. exopeptidase family members
transferrin receptor protease	FolH-like	Zn-dep. exo-peptidase assemblage	extended cap-like structure with at least 1 helix	extended lateral shelf with helical segment	PA (protease-associated) domain insertion	--	returns "classical" Zn-dep. exo-peptidase profiles (top assemblage in Fig. 4)	mostly returns hits with other Zn-dep. exopeptidase family members
Succinyl-glutamate desuccinylase AstE	AstE/AspA-like	Zn-dep. exo-peptidase assemblage	small loop extension	standard lateral shelf	none	Cterm fusion to barrel sandwich hybrid domain	returns profiles from Fig. 4 bottom assemblage of the zn-dep. exo-peptidases	mostly returns hits with other Zn-dep. exopeptidase family members
Glutaminyl-peptide cyclo-transferase	Glutaminyl-peptide cyclo-transferase like	Zn-dep. exo-peptidase assemblage	extended loop folding back over active site	standard lateral shelf	none, slightly extended strands and loop	unusual extended loop after H1	returns "classical" Zn-dep. exo-peptidase profiles (top assemblage in Fig. 4)	sees other peptidase families before detecting some PNP and other family structures with Z-scores ~5-6

N-formyl-glutamate amido-hydrolase	FGase-like	Zn-dep. exo-peptidase assemblage	small extended loop folding back over active site	lateral shelf with notable bulge	n/a	--	mostly detects AstE/AspA and SpoIIIP before other peptidases	strongest hits are with Fig. 4 bottom assemblage of Zn-dep. exo-peptidase structures
pyro-glutamate amino-peptidase	pyro-glutamate amino-peptidase	pyro-glutamate amino-peptidase	extended, upwards-pointing loop w/ $\alpha$ -like elements	standard lateral shelf	n/a	--	nothing recovered outside of this family	strongest hits with PNPs

Lineage name	Family name	Higher-order grouping	phyletic distributions
Germination protease	HybD-like	HybD-like	mostly firmicutes, some scattered in tenericutes and bacteroidetes
HybD	HybD-like	HybD-like	largely pan-bacteria, pan-archaeal (including Asgardian)
peptidyl hydrolase	pth	pth-like	absolutely conserved in bacteria, eukaryotes with mitochondria
5'-methylthio-adenosine nucleosidase	MTAN	PNP-like	PNP-like has a branch clearly traceable to LUCA
YgiD-like	PCAD	PCAD-Memo	bacteria, some eukaryotes
Memo	Memo	PCAD-Memo	universal
tRNA deacylase	tRNA deacylase	tRNA deacylase	pan-archaeal, land plants, scattered bacteria
bacPAC2	bacPAC2	PAC2-like	pan-archaeo-eukaryotic, pan-actinobacteria and a scattering of other bacterial lineages
endolysin Ply	amidase-like	Zn-dep. exo-	largely pan-bacterial, quite a bit of caudovirus representation

		peptidase assemblage	
carboxy-peptidaseA	carboxy-peptidaseA	Zn-dep. exo-peptidase assemblage	pan-eukaryotic, largely pan-bacterial with notable underrepresentation in cyanos. Limited euryarchaeal presence
leucine amino-peptidase	leucine amino-peptidase	Zn-dep. exo-peptidase assemblage	pan-bacterial, largely pan-eukaryotic with basal lineages in question. Limited archaea look like horizontal transfers
amino-peptidase	dinuclear Zn exo-peptidase	Zn-dep. exo-peptidase assemblage	probably pan-eukaryote and pan-bacteria. Archaea found sporadically across many lineages but relatively strong presence in Asgard
transferrin receptor protease	FoH-like	Zn-dep. exo-peptidase assemblage	see above row for joint phylogenetic distribution, but this specific family appears pan-euk with some scattered bacteria
Succinyl-glutamate desuccinylase AstE	AstE/AspA-like	Zn-dep. exo-peptidase assemblage	pan-euryarchaea, a few other archaeal sequences (Bathyarchaeota), euk: vertebrates, scattered fungi, SAR representation. bac: pan-bacteria with likely underrepresentation in actinos
Glutaminyl-peptide cyclo-transferase	Glutaminyl-peptide cyclo-transferase like	Zn-dep. exo-peptidase assemblage	core family involved in peptide cyclotransferase reaction (e.g. neuropeptide processing in animals) is crown group euk, sporadic bacterial distribution, and Asgardian archaea. Considering the broader assemblage of related families, could probably be considered pan-bacteria and sporadically present across archaea
N-formyl-glutamate amido-hydrolase	FGase-like	Zn-dep. exo-peptidase assemblage	found in most bacterial lineages but often sporadic (concentrated in $\alpha$ -proteos, but underrepresented in firmicutes and cyanos)
pyro-glutamate	pyro-glutamate	pyro-glutamate	sporadic in bacterioidetes, otherwise pan-bacteria. pan-euk but absent in giardia and probably apicomplexa. Sporadic presence in eury- and cren-archaea, only a single Asgardian.

amino-peptidase	amino-peptidase	amino-peptidase	
-----------------	-----------------	-----------------	--

^: S3 here refers to core strand ordering across the PNP-ptf fold; from the vantage of the PCAD-Memo superfamily this is S1