1 Structural basis of ribosomal peptide macrocyclization in plants

2	Joel Haywood ^{1,2} , Jason W. Schmidberger ^{1,2} , Amy M. James ^{1,2} , Samuel G. Nonis ^{1,2} ,
3	Kirill V. Sukhoverkov ^{1,2} , Mikael Elias ³ , Charles S. Bond ¹ , Joshua S. Mylne ^{1,2}
4	The University of Western Australia, ¹ School of Molecular Sciences & ² The ARC Centre of Excellence in
5	Plant Energy Biology, 35 Stirling Highway, Crawley, Perth 6009, Australia; ³ University of Minnesota,
6	Department of Biochemistry, Molecular Biology and Biophysics & Biotechnology Institute, St. Paul, MN
7	55108, USA
8	Constrained, cyclic peptides encoded by plant genes represent a new generation of drug
9	leads. Evolution has repeatedly recruited the Cys-protease asparaginyl endopeptidase (AEP)
10	to perform their head-to-tail ligation. These macrocyclization reactions use the substrates
11	amino terminus instead of water to deacylate, so a peptide bond is formed. How solvent-
12	exposed plant AEPs macrocyclize is poorly understood. Here we present the crystal structure
13	of an active plant AEP from the common sunflower, <i>Helianthus annuus</i> . The active site
14	contained electron density for a tetrahedral intermediate with partial occupancy that
15	predicted a binding mode for peptide macrocyclization. By substituting catalytic residues we
16	could alter the ratio of cyclic to acyclic products. Moreover, we showed AEPs from other
17	species lacking cyclic peptides can perform macrocyclization under favorable pH conditions.
18	This structural characterization of AEP presents a logical framework for engineering superior

19 enzymes that generate macrocyclic peptide drug leads.

20 Introduction

21 Asparaginyl endopeptidases (AEPs) are a group of asparagine/aspartic acid (Asx) specific 22 proteases that have been classified as belonging to the C13 family of cysteine proteases based 23 on the presence of a His-Gly-spacer-Ala-Cys motif (Hara-Nishimura et al., 1993; Chen et al., 24 1997; Mathieu et al., 2002; Shafee et al., 2015). First described in plants as vacuolar processing 25 enzymes based on their propensity for processing seed proteins stored in vacuoles, AEPs have 26 since been described in a variety of organisms and shown to be involved in a wide range of 27 processes including, cell death, antigen processing and hemoglobin degradation (Hara-28 Nishimura et al., 1993; Manoury et al., 1998; Hatsugai et al., 2004; Kuroyanagi et al., 2005; 29 Yamada et al., 2005; Sojka et al., 2007). In addition to the proteolytic function observed in these 30 processes, AEP has become well known for its curious ligation reactions (Min and Jones, 1994; 31 Sheldon et al., 1996; Mylne et al., 2012; Nguyen et al., 2014; Zhao et al., 2014; Dall et al., 2015). 32 The ability of endoproteases to perform ligation reactions was first observed by Bergmann and 33 Fruton in 1938 with chymotrypsin (Bergmann and Fruton, 1938). Later, in vitro ligation 34 reactions were performed with AEP from jack bean (Canavalia ensiformis) seeds (Bowles et al., 35 1986; Min and Jones, 1994). The recent discovery that evolutionarily distinct plant families have 36 repeatedly recruited AEPs to catalyze the formation of ribosomally synthesized and post-37 translationally modified peptides (RiPPs), through the macrocyclization of linear precursor 38 sequences, has caught the attention of drug designers keen to overcome the current 39 inefficiencies in native chemical ligation that limit the therapeutic use of cyclic peptides 40 (Pattabiraman and Bode, 2011; Mylne et al., 2012; Arnison et al., 2013). Such therapeutic cyclic

41 peptides are viewed by many to have the potential to capitalize on a niche in the current 42 pharmaceutical market by virtue of their intermediate size between small molecule drugs and 43 large protein structures, and their unique capacity to combine favorable bioavailability and 44 stability characteristics with high target specificity facilitated by tolerance to site-directed 45 mutagenesis (Clark et al., 2005; Clark et al., 2010; Gould et al., 2011; Ji et al., 2013; Poth et al., 46 2013; Truman, 2016). Moreover, as computational techniques for the discovery of RiPPs 47 improve and the number of cyclic peptides described continues to expand, an ever-wider array 48 of scaffolds might be exploited to tailor molecules to specific drug targets (Bhardwaj et al., 49 2016; Truman, 2016; Hetrick and van der Donk, 2017) (Figure 1). 50 Sunflower trypsin inhibitor-1 (SFTI-1) is a 14-residue, bicyclic peptide with a cyclic backbone and 51 an internal disulfide bond (Luckett et al., 1999). Its biosynthesis is rather unusual as it sequence 52 is buried within a precursor that also encodes seed storage albumin. Seed storage albumins are 53 a major class of seed storage protein that constitute over 50% of total seed protein and become 54 a source of nitrogen and sulfur during seed germination (Youle and Huang, 1978; Shewry and 55 Halford, 2002). The common sunflower (Helianthus annuus) has many genes encoding 56 precursors for these napin-type or 2S seed storage albumins that are synthesized in the rough 57 endoplasmic reticulum before undergoing cleavage maturation by AEP and localizing to storage 58 vacuoles (Bollini and Chrispeels, 1979; Franke et al., 2016; Jayasena et al., 2016). Along with an 59 adjacent albumin, SFTI-1 is post-translationally processed by AEP from within a unique seed 60 storage albumin precursor called Preproalbumin with SFTI-1 (PawS1) (Mylne et al., 2011). SFTI-1 61 is a potent inhibitor of serine proteases, and its intrinsic stability and cellular penetration

capabilities have led to its application as a bioactive scaffold (White and Craik, 2016; Swedberg
et al., 2017).

64 AEP-catalyzed macrocyclization of SFTI-1 is hypothesized to proceed via a cleavage coupled 65 intramolecular transpeptidation reaction whereby catalysis begins with the deprotonation of 66 the active site Cys by a localized His of the catalytic center, facilitating the nucleophilic attack on 67 the carbonyl carbon of the Asp by the activated Cys thiol of AEP. This attack culminates in the 68 formation of a thioacyl intermediate between the substrate and AEP, and the removal of the C-69 terminus. The reaction is proposed to be subsequently concluded by the nucleophilic attack of 70 this intermediate by a Gly at the N-terminus of the substrate, resulting in the macrocyclization 71 of SFTI-1 in a head-to-tail manner (Mylne et al., 2011; Bernath-Levin et al., 2015). Notably, this 72 reaction proceeds in competition with nucleophilic attack upon the thioacyl intermediate by 73 any nearby water molecules which would produce hydrolyzed, acyclic-SFTI. In vitro studies 74 revealed the ratio of acyclic to cyclic SFTI-1 to be in the order of 5.8:1, with the less stable 75 acyclic products hypothesized to be quickly degraded *in vivo* (Bernath-Levin et al., 2015). 76 Evidence for AEP-mediated and hydrolysis-independent transpeptidation was demonstrated through the exclusion of a heavy atom O¹⁸ in the cyclic SFTI-1 product from an *in vitro* jack bean 77 78 AEP (CeAEP1) catalyzed reaction (Bernath-Levin et al., 2015). The formation of a much larger cyclic peptide by an AEP from Oldenlandia affinis (OaAEP1) was also shown to lack O¹⁸ 79 80 incorporation, suggesting a conserved mechanism of macrocyclization despite differences in 81 substrate sequences (Harris et al., 2015). However, the suggestion that an AEP from butterfly 82 pea (Clitoria ternatea) termed butelase 1 functions only as a ligase (Nguyen et al., 2014), 83 combined with the proposal of a succinimide-driven, cleavage-independent ligation event

84	based on the crystal structure of human AEP (hAEP) and its ability to ligate substrate in the
85	absence of the catalytic Cys, has cast uncertainty on the mechanism of AEP-catalyzed
86	macrocyclization (Dall et al., 2015).
87	In order to elucidate a structural explanation why plant AEPs have been recruited by distinct
88	plant lineages to perform macrocyclization and to understand the catalytic and structural
89	nuances that might allow preferences towards cleavage or ligation reactions, we sought the
90	crystal structure of a sunflower AEP.
91	Herein, we describe the first structure of an active plant AEP; one capable of performing
92	peptide macrocyclization. This AEP, the most abundant AEP of five AEPs in the common
93	sunflower (HaAEP1), displays structural similarity to previously published active AEPs from
94	mammals, with subtle differences at residues involved in substrate recognition. Our
95	characterization by site-directed mutagenesis of HaAEP1 residues integral to macrocyclization
96	will facilitate the bioengineering of plant AEPs for improved macrocyclization efficiency,
97	diversifying the scaffolds usable as cyclic therapeutic leads.

98 Results

99 Catalytic domain of HaAEP1 purified at pH 4

100 AEPs are synthesized as inactive precursors that have been shown to undergo irreversible auto-

- 101 activation into their mature form on exposure to a low pH environment that resembles the
- 102 acidic pH in the vesicles/vacuole where these proteins are active *in vivo* (Dall and Brandstetter,
- 103 2013; Shafee et al., 2015). In order to obtain an active form of a plant AEP, a ~51 kDa pro-

104 HaAEP1 (residues 28-491) lacking an endoplasmic reticulum signal sequence and including a N-105 terminal His-tag was expressed in *Escherichia coli* and purified by nickel affinity 106 chromatography before being activated at pH 4.0 overnight. The activated form of HaAEP1 was 107 then further purified by size exclusion chromatography, enabling separation of the core domain 108 from the 'cap' domain (Zhao et al., 2014), and crystal trials were undertaken (Figure 2-figure 109 supplement 1). SDS-PAGE analysis of the ~38 kDa core domain peak revealed the disassociation 110 of the core domain from the cap domain but also showed several bands of HaAEP1, suggesting the presence of several cleavage sites at the termini of the core domain, as seen previously for 111 112 several AEPs (Hara-Nishimura et al., 1998; Nguyen et al., 2014; Zhao et al., 2014; Harris et al., 113 2015). Indeed sequence comparison reveals the conservation of several of these predicted 114 cleavage sites but notably lacks a previously described C-terminal di-Asp motif (Hiraiwa et al., 115 1999) (Figure 2—figure supplement 2). This led us to hypothesize that Asp52, Asn57, Asn338, 116 Asp356, and Asp358 might represent the dominant autocatalytic cleavage sites in HaAEP1.

117 HaAEP1 structure reveals subtle differences to dictate substrate specificity

118 Crystallization trials of the ~38 kDa activated HaAEP1 yielded diffraction quality crystals that 119 diffracted to a resolution of 1.8 Å (Table 1). The crystal structure was solved by molecular 120 replacement yielding a single monomer in the asymmetric unit, and revealed an active 121 monomeric HaAEP1 (residues 58-338 with weak electron density for Asn338) that forms a 122 canonical C13 caspase structure, with a central six-stranded β -sheet region confined by five α -123 helices (Hara-Nishimura et al., 1993; Yamada et al., 2005) (Figure 2A, Figure 2—figure 124 supplement 3). The structure of HaAEP1 lacks the C-terminal cap domain and N-terminal His-Tag, displaying dimensions of approximately 44 Å x 42 Å x 39 Å. Sequence analysis of this core 125

domain suggests that the aforementioned pro-domains are likely to have been auto catalytically processed during maturation as the previously predicted Asn cleavage sites
 precede and follow the defined active structure.

129 HaAEP1 displays structural similarity to hAEP and a recently published inactive structure of 130 OaAEP1 (PDB ID: 4N6O and 5HOI) with an r.m.s.d. value of 1.0 and 0.7 Å over 262 and 267 131 carbon alpha residues, respectively (Holm and Rosenström, 2010; Yang et al., 2017) (Figure 2B). 132 Due to such topological conservation it is expected that subtle differences around the substrate 133 binding pocket will define substrate specificity and catalytic efficiency. Indeed, comparison of 134 these three structures reveals HaAEP1 exhibits a unique flexible extension, reflected by weak 135 electron density, in the α 5- β 6 loop and differences between the residues that are local to the 136 catalytic His and Cys and those that delineate the S3-S5 pockets (following the protease 137 nomenclature defined by Schechter and Berger where the cleavage site residue is termed P1 138 and residues prior to and following the cleavage site are labelled P5-P2 and P1'-P2', 139 respectively, and where the corresponding binding sites on the protease are described as S5-140 S2') (Schechter and Berger, 1967). Specifically, differences in the substrate pocket include 141 residues YGT 249-251 in HaAEP1 (hAEP: YAC 217-219, OaAEP1: WCY 246-248), a bulky Trp232 in 142 hAEP versus Leu271 in HaAEP1 and Leu268 in OaAEP1, and the presence of an additional 143 proline prior to the β_{IV} - β_V polyproline loop which orients E257 away from the S4 region in 144 HaAEP1 (Figure 2B). In OaAEP1, the C247 residue at the entrance to the S4 pocket was recently 145 proposed to function as a 'gate keeper for ligation' with large bulky side chains inhibiting 146 ligation (Yang et al., 2017). Differences local to the active site include residues P181, Q245, 147 N247 in the β_1 sheet and β_5 - β_{1V} loop (OaAEP1: A178, T242, S244, hAEP: T151, R213, S215) and

G185, E189, H191 in the β_{II}-β_{III} region (OaAEP1: G182, K186, Y188, hAEP: V155, N158, D160)
(Figure 2B).

150 A 6-residue insertion in the α 5- β 6 loop results in the disruption of a potential N-linked 151 glycosylation site that was hypothesized to affect substrate binding upstream to P5 in hAEP 152 (Fuentes-Prior and Salvesen, 2004). Interestingly, none of the four conserved potential N-linked 153 glycosylation sites in hAEP and mouse AEP (mAEP) are found in HaAEP1 (Figure 2—figure 154 supplement 2). Given that these glycosylation sites have been predicted to protect AEP from 155 non-specific protease activation it is intriguing to find that the only two potential N-linked 156 glycosylation sites in HaAEP1 (N138 and N143) are located on the opposite side of the protein 157 to the activation peptide in mammalian AEPs, at the beginning and center of the α 2 helix, 158 respectively (Dall and Brandstetter, 2016). Moreover, only N138 is prominently surface 159 exposed, suggesting that N-linked glycosylation in HaAEP1 is not utilized for mitigation against 160 non-specific premature activation.

161 HaAEP1 displays a tetrahedral intermediate in the active site

The 1.8 Å resolution of our HaAEP1 structure allowed us to observe a succinimide moiety (Snn) below the catalytic His in a location identical to human AEP (Figure 3A) that was hypothesized to play a role in peptide ligation (Dall et al., 2015). We distinguished dual conformations of the catalytic His178 and Cys220 residues which we hypothesize represent conformational changes that occur during catalysis and correspond to substrate free and reaction intermediate states (Figure 3A). In the intermediate state the catalytic Cys Sδ is oriented ~95° towards the Nδ of the catalytic His imidazole ring which is orientated ~3.8 Å closer to this residue in the corresponding intermediate state. Moreover, flexibility of His in a relatively open pocket free of steric
hindrance (Figure 3—figure supplement 1) suggests an additional role for conformational shifts
of the His in catalysis as seen with a range of proteases (McLuskey et al., 2012; Clark, 2016;
Chekan et al., 2017). In the proposed alternate resting state the catalytic Cys Sõ is oriented
towards the backbone amine of the highly conserved Gly179, reducing the distance between
them from 5.4 Å to 3.6 Å. In this orientation the backbone amine might function in stabilizing
proton abstraction from the Cys thiol.

176 Close examination of the electron density in the active site of HaAEP1 suggested a small 177 peptide chain is intermittently bound to the intermediate conformation of Cys220 (Figure 3B). 178 Given AEPs specificity for Asx and the nature of the electron density, we built and refined a 179 tetrahedral complex of a three residue peptide ligand (AAN) bound to the active site Cys of 180 HaAEP1 with partial occupancy (Figure 3B). Residues Ala-Ala were modelled upstream of the P1 181 Asn due to the weak electron density away from the peptide backbone at these residues. The 182 AAN peptide ligand allowed characterization of interactions likely to exist between HaAEP1 and 183 P1 Asn and main chain of residues P2-P3 during substrate recognition. Due to the high 184 sequence conservation of the HaAEP1 active site with mammalian AEPs and conserved 185 substrate orientation, many of the interactions with the substrate match those observed for 186 inhibitors of hAEPs (Dall et al., 2015) (Figure 3C). The presence of this unexpected substrate in 187 the HaAEP1 active site is likely to be a product of auto-activation. The continuous electron 188 density between Cys220 and the P1 Asn supports the interpretation of the formation of a 189 tetrahedral intermediate that is stabilized by the presence of an oxyanion hole formed by 190 His178, Gly179 and the backbone amine of Cys220 that is more congruous to the electron

191 density than an acyl intermediate or free peptide (Figure 3B — figure supplement 2).

192 Moreover, the observed short distance between the tetrahedral intermediate carbon atom and 193 catalytic cysteine sulfur atom is incompatible with the absence of a covalent interaction. Main 194 chain amino groups of Gly residues have previously been proposed to function in the creation 195 of an oxyanion hole stabilizing the formation of a tetrahedral intermediate with the substrate in 196 a wide range of cysteine proteases (Fuentes-Prior and Salvesen, 2004). Furthermore, previous 197 observations of tetrahedral intermediates and enzyme-product complexes with serine 198 proteases have been shown to exhibit a pH-dependent equilibrium (Wilmouth et al., 2001; 199 Radisky et al., 2006; Lee and James, 2008). Similarly, the trapping of this intermediate state may 200 have been fortuitously facilitated the activation of HaAEP1 at pH 4, followed by crystallization 201 at pH 7. Although alternative conformations of the active site Cys in hAEP have been described 202 previously, this tetrahedral intermediate state has not been described before (Dall et al., 2015).

203 Modulation of pH enables HaAEP1 to perform macrocyclization

204 The similarity of the HaAEP1 active site to OaAEP1 and the revelation it contained a reactive succinimide prompted us to test an enzyme preparation similarly taken to pH 4.0 against the 205 206 modified SFTI-1 precursor substrate SFTI(D14N)-GLDN substrate. HaAEP1 had previously been 207 unable to create a macrocyclic product from SFTI(D14N)-GLDN, but had been shown to efficiently cleave it at a rate, k_{cat}/K_m value of 610 M⁻¹ S⁻¹, similar to rates published for other 208 209 AEPs (Bernath-Levin et al., 2015). To our surprise, our new preparations of HaAEP1 taken to pH 210 4.0 produced cyclic SFTI(D14N) when the reactions were conducted at pH 6.5 (Figure 4B WT, 211 Figure 4—figure supplement 1A WT). Previously, HaAEP1 had been purified at pH 8, activated 212 at pH 5 and then used in reactions at pH 5. The HaAEP1 preparations that were able to

213 macrocyclize were similarly purified at pH 8, but activated at pH 4 then returned to pH 6.5.

214 Activation at lower pH has been shown to be more effective at removing the cap domain, which

in mammalian AEPs had a propensity to re-ligate (Zhao et al., 2014). Higher pH could also favor

216 macrocyclization by facilitating the deprotonation of the Gly N-terminus at the active site,

217 priming it to attack its C-terminus, which is held in the thioacyl intermediate at the active site.

218 Residues local to the catalytic dyad influence product formation

219 To investigate the mechanism of macrocyclization by HaAEP1 we identified several residues for 220 site directed mutagenesis. Firstly, we hoped to clarify the roles of Cys220 and Snn177 in 221 macrocyclization through Ser and Ala mutations, respectively. Furthermore, we also mutated 222 Snn177 to Gly, as the C13 protease GPI8 has also been proposed to carry out an intramolecular 223 transpeptidation reaction yet displays a Gly residue at a location equivalent to Snn177 (Zacks 224 and Garg, 2006) (Figure 2—figure supplement 2). Secondly, we hoped to alter the ability of 225 HaAEP1 to macrocyclize its native substrate SFTI-1 by altering the residues that fine tune this 226 catalysis. By modelling the binding of an NMR structure of PawS1 (Franke et al., 2017) and the N-terminally cleaved SFTI-1 precursor (SFTI-GLDN) to HaAEP1 we were able to hypothesize the 227 228 location of a hydrophobic S2' binding region encompassing strand β_{II} with G185 located at the 229 center of this region (Figure 4—figure supplement 2). In addition, previous studies have also 230 implicated that Asn, Glu and Asp residues proximal to the catalytic Cys and His function in 231 catalysis (Dall and Brandstetter, 2013; Zhao et al., 2014). Inspection of the HaAEP1 structure 232 revealed that several of these residues (including Asn73 and Glu221) are conserved in HaAEP1. 233 Although, Glu221 is oriented away from Cys220, in a direction similar to that seen in hAEP

bound to human cystatin E, it might assume alternate conformations due to its solventexposure and high B-factor.

The mutant HaAEP1 proteins were expressed in *E.coli* and analyzed by circular dichroism
(Figure 4—figure supplement 3), with wild-type (WT) HaAEP1 displaying a melting temperature
of ~52°C similar to previous reports for hAEP (Dall and Brandstetter, 2013), and mutants
displaying similar spectra to WT.

240 Incubation of WT HaAEP1 with a fluorophore labelled (BODIPY) activity-based probe (Lu et al., 241 2015) illustrated its heterogeneity in size following activation at pH 4 and incubation with 242 substrate at pH 6.5, which has previously been observed for AEP proteins both in vivo and in 243 vitro and speculated to be the result of processing of non-glycosylated forms (Zhao et al., 2014) 244 (Figure 4a). This probe also revealed a C220S mutation in HaAEP1 to result in the enzyme 245 becoming incapable of pH-dependent activation (**Figure 4a**). Further analysis of AEP mutants 246 activity using the BODIPY probe illustrates the substantial heterogeneity in size between active 247 WT and N73A, N73D, D177G, D177A or E221K mutant AEPs, likely due to autocatalytic processing, yet indicates that they remain active (Figure 4—figure supplement 1D). 248 249 The use of a seleno-modified synthetic SFTI(D14N)-GLDN substrate, which we have previously 250 shown to be processed by HaAEP1 and CeAEP1(Bernath-Levin et al., 2015), allowed for a 251 comparison of activity profiles of HaAEP1 mutants through the quantification of distinctive 252 isotopic cyclic, acyclic and unprocessed peak areas by MALDI-MS (Figure 4B-C, Figure 4—figure

253 **supplement 1A**). This comparison reveals subtle changes between mutants in the ratio of cyclic,

acyclic and unprocessed peptides and confirmed that the HaAEP1-C220S mutant is unable to

255 cleave as evidenced by the lack of a peak at mass 1608 or 1626 (Figure 4, Figure 4—figure 256 supplement 1). As expected, mutation of the second residue of the catalytic dyad (H178A) also 257 results in a drastic reduction in activity, based on SFTI(D14N)-GLDN processing and activity 258 based BODIPY probe results, confirming the significance of C220 and H178 in AEP activity 259 (Figure 4, Figure 4—figure supplement 1). Interestingly, the H178A mutation does not abolish 260 HaAEP1 activity, as previously seen with mAEP, and suggests a third residue could facilitate 261 proton transfer at the active site (Zhao et al., 2014). Similar results to the H178A mutation were 262 also observed for G185S that was directed at altering the hydrophobic S2' binding region. The 263 effect of the G185S mutation suggests G185 has a role in substrate recognition and could 264 sterically alter the conformation of H178 and Snn177 (Figure 4, Figure 3—figure supplement 1). 265 The mutation E221K, which has previously been shown to increase endopeptidase activity in 266 hAEP, resulted in a loss of cyclic product as shown by an absence of a peak of mass 1608 (Figure 267 **4B, Figure 4—figure supplement 1A)**. Mutation of N73A leads to a higher ratio of cyclic to 268 acyclic product as shown by an increased peak area relative to WT of mass 1608 and a reduced 269 acyclic product peak of mass 1626. N73D and D177G mutants appear to process SFTI(D14N)-270 GLDN in a manner similar to WT. Whereas the large fraction of unprocessed SFTI(D14N)-GLDN, 271 mass 2025, after incubation with D177A illustrates a reduction in processing efficiency, as 272 previously noted for the mAEP D149A mutant (Zhao et al., 2014). 273 The effect of the N73A, E221K and D177G mutations on AEP catalyzed ligation was probed 274 further by investigating each mutant's ability to revert to its inactive form via re-ligation of its 275 cap domain upon shift to neutral pH. As described for mAEP and equivalent mutants (Zhao et 276 al., 2014), incubation of HaAEP1 at pH 4 led to an irreversible dissociation of its cap domain

(Figure 4—figure supplement 4). However, following activation at pH 5.5 and 6.5, upon shifting
to pH 8 the WT, N73A, E221K and D177G mutants were able to re-ligate the cap onto the core
domain resulting in the formation of the inactive pro-enzyme with a Mw ~52 kDa as evidenced
by SDS-PAGE and activity based probes (Figure 4—figure supplement 4).

281 These results confirm the importance of the catalytic dyad in AEP function and show that

HaAEP1, like mAEP and the closely related C13 family member GPI8, is able to perform its

283 ligation reaction in the absence of a Snn residue. Moreover, subtle mutations affecting the

stability of the catalytic dyad might favor either hydrolysis or macrocyclization.

285 Active HaAEP1 exhibits an open surface amenable to macrocyclization

286 In order to discern the structural determinants that favored AEPs to be recruited independently

287 by evolution multiple times for macrocyclization we compared the structure of HaAEP1 and its

288 predicted binding mode with the crystal structures of closely related cysteine proteases:

- sortase A, papain and metacaspase MCA2 (Suree et al., 2009; Chu et al., 2011; McLuskey et al.,
- 290 **2012) (Figure 5)**.
- 291 Sortase A is a *Staphylococcus aureus* cysteine protease which catalyzes a similar

transpeptidation reaction to AEP, ligating proteins bearing a LPXTG motif to peptidoglycan

293 precursors in the bacterial cell wall (Mazmanian et al., 1999). NMR studies have shown that the

- resolution of the transpeptidation thioacyl intermediate reaction occurs through nucleophilic
- attack of a lipid terminal amine in a steep valley between the β 7- β 8 and β 4- β 5 loops (Suree et
- al., 2009). In comparison to HaAEP1, the sortase A β7-β8 loop protrudes much further from the
- active site and the aromatic residues F122, Y128, W194 in these loops orient over the catalytic

Cys (Figure 5B). Together these loops, despite reported flexibility, would likely impart
considerable steric hindrance for macrocyclization by inhibiting both the resolution of the
intermediate by the N-terminal amine group and binding at the S2' substrate 'tail'.

Papain from the melon tree *Carica papaya* is the archetypal plant C1 family cysteine protease and has been found to bind to a cystatin homologue, tarocystatin, in a manner analogous to that of hAEP binding cystatin (Otto and Schirmeister, 1997; Chu et al., 2011; Dall et al., 2015). However, inspection of the papain active site reveals a topology that is not conducive for macrocyclization with the catalytic triad buried deep within the protein and steric hindrance for peptide N-terminal attack likely to be imparted by W177, D158 and the extended alpha helical loops of the α 3 and α 7 regions (**Figure 5C**).

308 Metacaspases are expressed in plant, fungi and protozoa and display a C14 caspase domain 309 that is structurally homologous to human caspases (Tsiatsiani et al., 2011). Metacaspases reside 310 within the same CD clan as AEP, but exhibit a strict specificity for a cleavage following Arg or Lys 311 (Vercammen et al., 2004). Currently no crystal structure is available for a plant metacaspase, 312 however the crystal structure of the protozoan metacaspase MCA2 reveals an architecture that 313 like papain would likely be unfavorable for macrocyclization due to steric hindrance around the 314 active site from several prominent loops; including the $\beta 1-\alpha 1$, $\beta A-\beta B$ and 280-loop (McLuskey et 315 al., 2012) (Figure 5D).

316 The crystal structure of active HaAEP1 suggests that the combination of a relatively open 317 reaction interface with space around the active site allowing for catalytic residue flexibility has

resulted in the convergence upon AEPs for macrocyclization ahead of the other 30 families ofcysteine proteases in plants (Rawlings et al., 2016).

320 AEPs have an intrinsic ability to perform peptide macrocyclization

321 Given the sequence similarity between AEPs and the conservation of residues involved in 322 catalysis we hypothesized that the ability to macrocyclize peptides might be inherent to AEPs. 323 To test this hypothesis we recombinantly expressed two AEPs from species which are currently 324 not thought to contain cyclic peptides of any kind; Arabidopsis thaliana (AtAEP2) and Ricinus 325 communis (RcAEP1), respectively. These AEPs were purified and activated as described for 326 HaAEP1 (pH 4) and then incubated with non-native substrates; SFTI-GLDN and SFTI(D14N)-327 GLDN, at a pH that favors ligation (pH 6.5). Under these conditions RcAEP1 was able to 328 macrocyclize both SFTI-GLDN and SFTI(D14N)-GLDN substrates whereas AtAEP2 was able to 329 macrocyclize only SFTI(D14N)-GLDN (Figure 6). These findings further supports our hypothesis 330 that the structural features of AEPs described above have allowed for the convergence upon 331 AEP for peptide macrocyclization reactions.

332 **Discussion**

333 Herein, we have described the structure of an active plant AEP containing a peptide ligand with

334 partial occupancy bound to the active site catalytic Cys as a tetrahedral intermediate,

335 illustrating conformational flexibility in the AEP catalytic dyad. This structure has enabled us to

336 predict a model for SFTI-1 macrocyclization by HaAEP1 where the GLDN tail of the SFTI-1

337 precursor orients in a manner analogous to cystatin binding to hAEP, over the catalytic His to a

338 hydrophobic patch on the β_{\parallel} region, and where the N-terminus attack occurs between

diminutive β_{I} sheet and $\beta_{5}-\beta_{IV}$ loop in a manner analogous to the attack of thioacyl intermediates by peptidoglycan precursors in sortase-catalyzed transpeptidation reactions (Suree et al., 2009; Clancy et al., 2010). This model complements the requirement for a small amino acid followed by a hydrophobic residue at P1' and P2' due to their orientation over the catalytic His and towards the hydrophobic $\beta_{II}-\beta_{III}$ region, respectively.

344 Several Cys proteases, including AEPs, have previously described a minor role for a third 345 Asx/Arg residue in catalysis (Vernet et al., 1995; Dodson and Wlodawer, 1998; Fuentes-Prior 346 and Salvesen, 2004; Clancy et al., 2010; Buller and Townsend, 2013; Dall and Brandstetter, 347 2013). This third residue has been suggested to function in stabilization and orientation of the 348 His imidazole ring. Interestingly, functional analysis of HaAEP1 catalytic triad residue N73 349 revealed that an N73A mutation resulted in the production of a higher ratio of cyclic SFTI-1. The 350 observation of potential conformational shifts in the crystal structure of HaAEP1 during 351 catalysis suggests that the loss of an orientating N73 side chain allows for further 352 conformational flexibility of H178. This increased flexibility of His in a relatively open pocket that might accommodate a range of rotamers could reduce steric hindrance of an N-terminus 353 354 entry towards the acyl intermediate and also facilitate deprotonation of the substrate N-355 terminus.

The requirement for space and flexibility between the catalytic dyad favors the convergence for peptide macrocyclization upon Cys proteases over Ser proteases due to the close proximity between Ser-His residues in these proteases, an inherent requirement based upon the reduced nucleophilic properties of Ser (Buller and Townsend, 2013). Indeed, a recent structure of a

360 macrocyclizing Ser protease, PCY1, suggested that a shift in the catalytic His away from Ser is 361 required for macrocyclization, based on a comparison to their hydrolytic relatives, and 362 suggested a role for the catalytic His in deprotonation of the attacking peptide N-terminus 363 (Chekan et al., 2017). Furthermore, comparison with other Cys proteases suggests that AEPs 364 have been converged upon for macrocyclization based on their relatively flat and open catalytic 365 site. In contrast to a recent hypothesis for efficient macrocyclization based on a comparison of 366 AEP structures focused on a region closer to the S4 pocket, here we suggest that differences in 367 AEP catalytic efficiency and substrate specificity will be defined by subtle amino acid differences 368 in the β_{II} - β_{III} region, β_{I} sheet and β_{5} - β_{IV} loop and the S1-S5 pocket (Yang et al., 2017). Moreover, 369 this orientation of the tail away from the β_1 sheet and β_5 - β_{1V} loop is hindered in other caspases 370 which exhibit a relatively straight substrate channel (Figure 5) and thus prevent simultaneous 371 attack of the thioacyl intermediate by the N-terminus loop upon scissile bond cleavage. 372 Following activation, AEP functionality has been illustrated to be dependent on a delicate 373 balance between pH and stability, where endopeptidase activity is favored at a low pH ($^{PH 4}$) 374 and ligase activity is favored at a higher pH (~pH 6) (Dall et al., 2015). A pH-dependent activity 375 has been well documented for cysteine proteases and is a function of the formation of a 376 thiolate and imidazolium ions on the catalytic dyad of Cys and His residues (Fuentes-Prior and 377 Salvesen, 2004; Frankel et al., 2005). Given the hypothesized role of His in deprotonating the 378 attacking N-terminal peptide chain in peptide macrocyclization it would therefore be expected 379 that at a low pH the catalytic His would be more readily protonated and hence this reaction 380 might be less frequent resulting in more hydrolysis at low pH. This hypothesis is supported by 381 the low level of endopeptidase activity observed with Ala mutations of the catalytic His in

382 HaAEP1 and other proteases, illustrating that thiolates might form and that hydrolytic 383 resolution of the thioacyl intermediate might occur in the absence of His (Frankel et al., 2007; 384 Zhao et al., 2014). Previous reports have also suggested that a local Gly amide backbone might 385 facilitate catalysis via a transfer of a proton to the leaving group of the tetrahedral intermediate 386 from a hydrogen bonded water molecule; analysis of the HaAEP1 structure reveals Gly179 387 could potentially perform this role in the absence of His (Brady et al., 1999). Moreover, 388 mutation of the acidic residue following the catalytic Cys to a basic residue in hAEP has 389 previously been shown to enhance catalytic efficiency by decreasing the local pKa of the Cys 390 residue (Dall and Brandstetter, 2013). Interestingly, we found an equivalent mutation in 391 HaAEP1 to result in the loss of peptide macrocyclization indicating that E221 could also aid in 392 deprotonation of the incoming N-terminus or that the residue's larger side chain and expected 393 orientation in activating the catalytic Cys could impart steric hindrance upon the N-terminus 394 attack of the thioacyl intermediate (Figure 4, Figure 4—figure supplement 1). However, given 395 the finding that E221K mutation is unable to prevent re-ligation of the cap domain to the active 396 core domain upon pH shift in HaAEP1, there is likely to be redundancy between the local 397 residues in creating a nucleophilic amine group to complete a transpeptidation reaction. 398 The finding that AEPs from species that lack cyclic peptides may be coaxed into performing 399 peptide macrocyclization of a linear peptide under favorable conditions significantly expands 400 the potential use of AEPs for the production of cyclic peptides (Figure 6). Moreover, further 401 investigation into the subtle nuances that define substrate specificity and catalytic activity

402 between AEPs is warranted, with differences in the β_{II} - β_{III} region, β_{I} sheet, β_{5} - β_{IV} loop and

 $403 \qquad \text{around the S1-5 pocket including the variable $\alpha5-\beta6$ and $\beta_{IV-}\beta_V$ loops likely to be key (James et$

404 al., 2017). Indeed, in comparison to HaAEP1 and OaAEP1, the efficient peptide macrocyclizing 405 AEP butelase 1 displays several different amino acids that could be responsible for this AEPs 406 efficacy. These differences include shorter sidechains in the β_1 sheet and β_5 - β_{1V} loops that may 407 reduce steric hindrance on a peptides N-terminus during attack on a thioacyl intermediate 408 (HaAEP1: P181, Q245, N247 OaAEP1: A178, T242, S244 Butelase-1: A168, G232, S234) and 409 differences in the β_{1V} - β_V region that could generate substrate specificity (**Figure 5—figure** 410 **supplement 1**).

411 Active HaAEP1 exhibits a succinimide at the same position as hAEP which has previously been 412 postulated to perform a Cys-independent ligation reaction through cyclic rearrangement with 413 the P1 Asx side chain (Dall et al., 2015). Functional analysis reveals that HaAEP1 is able to 414 perform peptide macrocyclization despite D177G/A mutations that cannot form a succinimide 415 group (Figure 4B-C). Moreover, these mutants continued to perform re-ligation of the cap 416 domain when activated at pH >5.5 (Figure 4—figure supplement 4), a result that has previously 417 been shown with mAEP (Zhao et al., 2014). In the absence of a critical requirement for succinimide formation in macrocyclization the question remains as to why this relatively 418 419 unstable aspartimide appears stable in AEP crystal structures and is largely conserved in the 420 C13 proteases. Succinimides have been shown to form more readily when adjacent to His 421 residues as the His N\delta abstracts a proton from the His backbone NH allowing the deprotonated 422 main chain amine to attack the Asx side chain (Brennan and Clarke, 1995; Takahashi et al., 423 2016). Once formed this succinimide could enhance the activity of the catalytic His by virtue of 424 reducing stabilizing hydrogen bonding interactions with the carboxyl terminus and be involved 425 in the proper positioning of the catalytic His. A further potential role for succinimides could

present upon their hydrolysis through racemization and favored formation of iso-Asp (Geiger
and Clarke, 1987; Reissner and Aswad, 2003). This iso-Asp could potentially represent a subtle
mode of AEP regulation as the orientation of this side chain towards the S1 pocket is likely to
disrupt substrate binding. Such iso-Asp residues have previously been proposed to regulate
protein activity by a time-dependent molecular switch (Geiger and Clarke, 1987; Ritz-Timme
and Collins, 2002).

432 In the absence of caspases, plants have evolved a wide range of cysteine proteases to ensure 433 functional redundancy in a myriad of highly regulated programmed cell death pathways in 434 response to environmental and developmental cues (Fagundes et al., 2015). Furthermore, 435 plants have developed a variety of strategies to control the destructive prowess of these 436 proteases including the expression of proteases as inactive zymogens with cofactor 437 dependency, compartmentalization and the production of protease inhibitors (Martinez et al., 438 2012). Of these cysteine proteases the AEPs have recently attracted considerable interest due 439 to their ability to carry out peptide macrocyclization and their potential application in the 440 synthesis of pharmacoactive cyclic peptides. Herein, the structural and functional analysis of 441 HaAEP1 has revealed residues that are able to favor the production of cyclic or acyclic products 442 from SFTI(D14N)-GLDN. Moreover, we have modelled a binding mode for productive 443 macrocyclization of the HaAEP1 natural ligand SFTI-1, based on a comparison with related 444 cysteine proteases that is likely to be conserved between AEPs, where substrate specificity is 445 defined by the amino acids around the binding site for respective AEPs. Furthermore, we have 446 shown that AEPs from diverse species lacking cyclic peptides are able to perform 447 macrocyclization under favorable pH conditions. These findings provide the foundation for

further optimization of AEPs, potentially widening the array of peptide substrates that could becyclized by AEPs.

450 Methods

451 **Protein expression and purification**

452 DNA sequence encoding residues 28-491 of HaAEP1 (accession code: KJ147147), Ricinus 453 communis AEP (RcAEP1) residues 58-497 (accession code: D17401) and Arabidopsis thaliana 454 AEP (AtAEP2) residues 47-486 (accession code: Q39044) were cloned into a pQE30 (QIAGEN) 455 expression vector with an N-terminal six-histidine tag and expressed in the SHuffle strain E. coli 456 (New England Biolabs) transformed with pREP4 (QIAGEN). Briefly, cultures were grown at 30°C 457 to an OD₆₀₀ of 0.8-1.0 in Luria Broth medium containing 100 μ g/mL ampicillin and 35 μ g/mL 458 kanamycin with expression induced at 16°C with 0.1 mM isopropyl β -D-1-thiogalactopyranoside 459 then cells cultured overnight. Cell pellets were collected by centrifugation and lysed by 460 ultrasonication in 50 mM Tris (pH 8.0), 100 mM sodium chloride, 0.1% Triton X-100. The soluble 461 fraction was then harvested by centrifugation and the supernatant was incubated (batch wise) 462 with Ni-NTA resin overnight at 4°C. The resin was then washed with 20 mL of 50 mM Tris (pH 463 8.0) and 20 mL of 50 mM Tris (pH 8.0) 20 mM imidazole and protein was eluted stepwise with 464 20 mL of 50 mM Tris (pH 8.0) 300 mM imidazole. For purification for crystal screens, six-465 histidine tagged inactive protein was purified, using an ÄKTA FPLC platform, by anion-exchange 466 chromatography (HiTrap Q HP 5 mL) with a gradient of 0 to 500 mM sodium chloride over 60 467 min and then either activated by dialysis in 100 mM citric acid - sodium citrate buffer (pH 4.0) 468 containing 50 mM sodium chloride and 5 mM DTT, overnight at 16°C, or directly concentrated

and further purified by size-exclusion chromatography (HiLoad 16/600 Superdex 200) in 50 mM
Tris (pH 8.0), 50 mM sodium chloride. Following dialysis active AEP was separated from
insoluble material by centrifugation and purified by size-exclusion chromatography (HiLoad
16/600 Superdex 200) in 100 mM citric acid – sodium citrate buffer (pH 4.0) containing 50 mM
sodium chloride. Protein was assessed for purity by SDS-PAGE.

474 HaAEP1 site-directed mutations were made following the Stratagene QuickChange protocol 475 with the following primers: The N73A mutation was made with forward primer 5'-GTA GCA AAG 476 GTT ATG GTG CTT ATC GTC ATC AGG CC-3' and reverse primer 5'-GGC CTG ATG ACG ATA AGC 477 ACC ATA ACC TTT GCT AC-3'; the N73D mutation with 5'-GTA GCA AAG GTT ATG GTG CTT ATC 478 GTC ATC AGG CC-3' and 5'-GCC TGA TGA CGA TAA TCA CCA TAA CCT TTG CTA C-3'; The D177A 479 mutation with 5'-CTG TTT TAT AGC GCT CAT GGT GGT CCG G-3' and 5'-CCG GAC CAC CAT GAG 480 CGC TAT AAA ACA G-3'; the D177G mutation with 5'-CTG TTT TAT AGC GGC CAT GGT GGT CCG 481 GG-3' and 5'-CCC GGA CCA CCA TGG CCG CTA TAA AAC AG-3'; the H178A mutation with 5'-CTG 482 TTT TAT AGC GAT GCT GGT GGT CCG GGT G-3' and 5'-CAC CCG GAC CAC CAG CAT CGC TAT AAA 483 ACA G-3'; the G185S mutation with 5'-GTC CGG GTG TTC TGA GTA TGC CGA ATG AAC-3' and 5'-484 GTT CAT TCG GCA TAC TCA GAA CAC CCG GAC-3'; the C220S mutation with 5'-TGA TTT ATC TGG 485 AAG CAT CTG AAA GCG GCA GCA T-3' and 5'-ATG CTG CCG CTT TCA GAT GCT TCC AGA TAA ATC 486 A-3'; and the E221K mutation with 5'-GAT TTA TCT GGA AGC ATG TAA GAG CGG CAG CAT TTT 487 TGA AGG-3' and 5'-CCT TCA AAA ATG CTG CCG CTC TTA CAT GCT TCC AGA TAA ATC-3'. 488 Mutations were verified by sequencing and expressed as described for WT.

489 Crystallization and data collection

490 Protein was concentrated to 10-15 mg/mL and crystal screening performed using the sitting-491 drop-vapor diffusion method with 80 µL of reservoir solution in 96-well Intelli-Plates at 16°C. 492 Protein to mother-liquor ratios for the sitting drops were varied in each condition: 0.1:0.1, 493 0.1:0.2, and 0.2:0.1 μL. Crystals of active HaAEP1 were obtained in 0.1 M HEPES (pH 7.5), 1.4 M 494 sodium citrate tribasic dihydrate. Single crystals were soaked in mother-liquor supplemented 495 with 20% glycerol as a cryoprotectant prior to being flash-frozen and stored in liquid nitrogen. 496 Data collection was performed at 100 K on the Australian MX2 (micro-focus) beamline using a 497 wavelength of 0.9537 Å and diffraction data for crystals were collected to a resolution of 1.8 Å.

498 Structural determination, refinement and model building.

499 Diffraction data were processed using iMOSFLM and scaled with AIMLESS from the CCP4 program suite (Battye et al., 2011; Winn et al., 2011) in space group P3₁21 with unit cell 500 501 dimensions a = b = 77.03, c = 108.17. A sequence alignment of HaAEP1 and human AEP1 was 502 generated using ClustalO and used to create a search model of HaAEP1 based on the last 503 common atom of human AEP (4FGU) using CHAINSAW. The structure of HaAEP1 was solved by 504 molecular replacement with PHASER using this search model, followed by automatic building 505 with ARP/WARP. Manual building and refinement was performed in iterative cycles with COOT 506 and REFMAC5 using the CCP4 program suite. Structural analysis and validation were carried out 507 with COOT and MolProbity (Emsley and Cowtan, 2004; Chen et al., 2010). Crystallographic data 508 and refinement statistics are summarized in **Table 1** with Ramachandran plot values calculated 509 from COOT. The peptide AAN modelled into the HaAEP1 active site was oriented into the active 510 site based on the similar mode of cystatin binding to human AEP (4N6O)(Dall et al., 2015). The

511 tetrahedral intermediate was evidenced from initial visualization in Fo-Fc electron density 512 difference maps using a polder OMIT map as implemented in phenix.polder (Liebschner et al., 513 2017). Coordinates and structure factors were deposited into the Protein Data Bank (PDB) 514 under accession code 6AZT. Tetrahedral complex of a three residue peptide ligand (AAN) bound 515 to the active site Cys220 defined as CX9 in PDB file. Figures illustrating structures were 516 generated using PyMol, electrostatic surface potentials were contoured at ±10 kT/e using an 517 APBS PyMol plugin (Dolinsky et al., 2007; Schrodinger, 2010). Models of C. ensiformis and C. 518 ternatea AEPs were generated using the I-TASSER server (Roy et al., 2010). 519 **Circular Dichroism** 520 Proteins purified in 50 mM Tris (pH 8.0) 300 mM imidazole were concentrated in an 30 kDa 521 Amicon centrifugal filter and buffer exchanged with an excess of 10 mM sodium phosphate 522 buffer (pH 8). Concentrations were checked by absorbance at 280 nm with a NanoDrop using 523 the extinction coefficient of Pro-AEP and samples diluted to 0.1 mg/ml or 0.2 mg/ml. CD 524 measurements were made in triplicate using a Jasco J-810 CD spectrometer with a 0.1-cm 525 quartz cuvette using a 1 nm bandwidth on a 0.1 mg/ml sample. CD wavelength spectra were 526 collected from between 200-260 nm a rate of 1 nm/sec at 20°C. Melt curves were collected

527 using the same bandwidth at 222 nm with temperature increasing at a rate of 1°C/min from 20

- 528 95°C on a 0.2 mg/ml sample. WT HaAEP1 melting temperature was interpolated from melt
- 529 curve using a sigmoidal four parameter logistic regression fit (GraphPad Prism).

530 **Protein activity analysis**

531 AEPs were activated by dialysis for four hours at room temperature in activation buffer (20 mM 532 sodium acetate pH 4.0, 5 mM DTT, 100 mM sodium chloride, 1 mM EDTA) followed by a second 533 dialysis into ligation buffer (20 mM MES pH 6.5, 0.5 mM DTT, 100 mM sodium chloride, 1 mM 534 EDTA acid). Protein concentrations were determined by measuring absorbance at A280 using a 535 NanoDrop (1 Abs = 1.0 mg/mL). For mass spectrometry analysis of the processing of 536 SFTI(D14N)-GLDN or native SFTI-GLDN by WT HaAEP1, RcAEP, AtAEP2 and mutant HaAEPs, AEPs 537 at a concentration of 40 µg/mL were incubated with 0.25 mM peptide with a diselenide bond 538 and 25 mM native (i.e. disulfide) SFTI-1 as an internal standard in activity buffer (20 mM MES 539 pH 6.5, 5 mM DTT, 1 mM EDTA). Reactions were carried out at 37°C for 16 hours. Three 540 independent reactions were performed for each protein tested. The reactions were stopped by 541 dilution 100-fold in 50% acetonitrile, 0.1% formic acid and spotted with an α -cyano-4-542 hydroxycinnamic acid matrix onto a plate for analysis by MALDI-MS. Quantification of peak area 543 by MALDI-MS was calculated using the internal standard to normalize for ionization efficiency 544 as described previously (Bernath-Levin et al., 2015). Briefly, for activity probe analysis 50 µL of 545 AEP at 10 µg/mL was incubated with 1 µM of the BODIPY probe JOPD1 (Lu et al., 2015) at room 546 temperature overnight and protected from light. The labelling reaction was stopped by the 547 addition of 10 μL of 6x SDS-PAGE loading buffer containing β-mercaptoethanol and proteins 548 separated using 4-12% Bis-Tris SDS-PAGE gels as previously described (Lu et al., 2015). Labeled 549 proteins were visualized in gel with excitation and emission wavelengths of 532 and 580 nm 550 using a Typhoon 9500 (GE Healthcare).

551 Analysis of reversal of AEP activation

552 WT, D177G, N73A and E221K proteins were purified by affinity chromatography as described 553 above and activated by overnight (4°C) dialysis at pH 4.0/5.5/6.5 in 100 mM citric acid - sodium 554 citrate buffer containing 50 mM sodium chloride and 5 mM DTT, with pH adjusted by addition 555 of 1.0 M Tris-HCl pH 8. Prior to pH-shift, sample aliquots were flash frozen and stored at -80°C 556 for subsequent activity probe analysis and SDS-PAGE analysis. Remaining activated protein was 557 then returned to the previous buffer but with pH adjusted to pH 8.0 by adding 1.0 M Tris-HCl 558 pH 8.0, as previously described, and incubated overnight at 4°C (Hara-Nishimura et al., 1998). 559 Following dialysis at pH 8.0 samples were flash frozen as described above. Proteins were then 560 analyzed for activity, as described above, using the BODIPY activity based probe. Following in 561 gel visualization of active protein gels were immediately fixed in Coomassie Blue stain for 562 comparison of active and inactive protein.

563 **Competing interests**

564 The authors declare no financial and non-financial competing interests.

565 Acknowledgements

566 The authors thank Dan Tawfik for critical insights during project design and Renier van der

- 567 Hoorn for the BODIPY probe JOPD1. J.S.M. was supported by an Australian Research Council
- 568 (ARC) Future Fellowship (FT120100013). S.G.N. and K.V.S. were supported by the Australian
- 569 Government's Research Training Program. J.H. and this work were supported by ARC grant
- 570 DP160100107 to J.S.M. and Dan Tawfik. HaAEP1 is deposited in PDB under accession 6AZT.

572 Figure Captions

573 Figure 1. Examples of enzyme catalyzed formation of cyclic RiPPs. Cyclic RiPPs that are enzyme 574 catalyzed from linear peptide precursors are commonly flanked by an N-terminus leader 575 sequence and a C-terminus follower sequence prior to cyclization. These flanking sequences 576 commonly aid in substrate recognition and catalysis. Well characterized examples of enzymes 577 that are able to perform peptide macrocyclization currently include cysteine proteases 578 (Helianthus annuus AEP – HaAEP, and Oldenlandia affinis AEP – OaAEP; following removal of 579 the leader peptide by an as yet undefined enzyme), serine proteases (patellamide G – PatG, and 580 peptide cyclase 1 – PCY1; following removal of the leader peptide by oligopeptidase 1 – OLP1) 581 and the ATP dependent ATP-grasp ligases (microviridin C and B – MdnC/B, and microcin J25 C -582 McjC; following removal of the leader peptide by microcin J25 B – McjB). These examples also 583 illustrate a range of cyclic peptide scaffolds that may have potential therapeutic applications.

584 Figure 2. Architecture of HaAEP1 versus other AEP structures. (A) Cartoon representation of 585 the overall topology of HaAEP1 with major α -helices and β -sheets annotated. Residues 304-309 586 that exhibited weak electron density and are missing from the model are shown as red dotted 587 loop. (B) Comparison of HaAEP1 core domain (green) with OaAEP1 (magenta) and hAEP (grey) 588 with bound chloromethylketone inhibitor (black) illustrating high overall structural similarity. 589 Expanded surface and cartoon representations of highlighted regions of β_{IV} - β_V substrate binding 590 region (blue backgrounds) and catalytic region (β_1 sheet, $\beta_5-\beta_{1V}$ loop and $\beta_1-\beta_{111}$ region) 591 orientated over the catalytic His residue (yellow backgrounds) are shown below illustrating the

residue differences that could alter substrate specificity (shown in stick format). Also see Figure
2–figure supplement 1-3.

594 Figure 3. Outstanding features of the HaAEP1 active site. (A) Catalytic residues with dual 595 conformations illustrated in the HaAEP1 active site with simulated annealing omit electron 596 density maps (2 F_{obs} - F_{calc}) contoured at 1σ level. (B) Cross-eyed stereo view of polder OMIT map (F_{obs} - F_{calc}) calculated in the absence of the shown overlaid AAN tetrahedral intermediate, 597 598 contoured at 3 σ level. (C) Schematic representation of interactions between AAN tetrahedral 599 intermediate (purple sticks) and labelled active site residues generated using LigPlot⁺ (Laskowski 600 and Swindells, 2011). Residues forming hydrogen bonds (green) are illustrated with black sticks 601 with distances shown (Å). Residues and atoms that provide hydrophobic interactions are 602 highlighted with orange eyelash symbols. Also see Figure 3–figure supplement 1.

603 Figure 4. Mutagenesis studies of HaAEP1. (A) Activity based probe analysis of HaAEP1 WT and 604 C220S mutant illustrates that WT is active at pH 6.5 after activation at pH 4 whereas C220S 605 mutant remains inactive. In-gel fluorescence of activity based probe (right) and post-606 fluorescence Coomassie stain of SDS-PAGE gel (left). (B) MALDI-TOF spectra of SFTI(D14N)-607 GLDN processing by WT and a range of HaAEP1 mutants directed at altering the ability of 608 HaAEP1 to cleave and macrocyclize the substrate SFTI(D14N)-GLDN. NEC – no enzyme control. 609 (C) Quantitation of peak areas from B. Peak areas of mass 1608 - cyclic SFTI(D14N), mass 1626 -610 acyclic-SFTI(D14N) and mass 2025 - unprocessed seleno-Cys modified SFTI(D14N)-GLDN 611 substrate, were normalized for ionization efficiency using an internal standard mass 1515 -612 native SFTI-1. Black - cyclic SFTI-1(D14N), grey - acyclic-SFTI(D14N) and white - unprocessed

seleno-Cys modified SFTI(D14N)-GLDN substrate. Peak areas with acyclic and cyclic forms have
previously been shown to exhibit similar ionization efficiencies (Bernath-Levin et al., 2015).
Error bars illustrate standard deviation n=3 (D177A n=2) technical replicates. Also see Figure 4–
figure supplement 1-4.

617 Figure 5. Comparison of Cys protease active site topologies. Substrates are oriented towards 618 the catalytic dyad of Cys and His, highlighted in yellow and blue sticks, respectively. Regions and 619 residues likely to impart steric hindrance on the proposed mode of macrocyclization by AEPs in 620 other Cys proteases are labelled. (A) Proposed binding mode for SFTI-1 macrocyclization (cyan 621 sticks and dots) based on the position of the tetrahedral intermediate in the HaAEP1 structure 622 with the N-terminus attacking the intermediate and the GLDN-tail (red-dots) oriented over the 623 catalytic His towards the β_{II} - β_{III} region. (B) Structure of sortase A (PDB 2KID) covalently bound to 624 an analog of the sorting signal (purple sticks). Residues implicated in transpeptidation are 625 highlighted with green sticks, in a region analogous to the $\beta 5-\beta_{\rm IV}$ loop and $\beta_{\rm I}-\beta_{\rm II}$ loop of HaAEP1. 626 (C) Structure of papain (PDB 3IMA) bound to residues 2-7 and 49-53 of tarocystatin (purple 627 sticks) illustrating restricted access to the catalytic dyad. (D) Structure of the metacaspase 628 MCA2_{C213A} (PDB 4AFV) with residues 30-33 of the N-terminal domain (purple sticks) bound in 629 the predicted direction for substrate binding. Flexible regions of low electron density are shown 630 with grey dots. Also see Figure 5–figure supplement 1.

Figure 6. The ability to macrocyclize is inherent to AEPs. MALDI-TOF spectra of SFTI-GLDN (A)
 and SFTI(D14N)-GLDN (B) with AEPs from species thought not to contain cyclic peptides, namely
 Arabidopsis thaliana (AtAEP2) and *Ricinus communis* (RcAEP1). (C) Quantitation of peak areas

from seleno-Cys modified substrates in **A** (mass 1609 - cyclic SFTI-1, mass 1627 - acyclic-SFTI and mass 2026 - unprocessed SFTI-GLDN substrate) and **B** (mass 1608 - cyclic SFTI(D14N), mass 1626 - acyclic-SFTI(D14N) and mass 2025 - unprocessed seleno-Cys modified SFTI(D14N)-GLDN substrate) normalized to mass 1515 - native SFTI-1. Black – cyclic peptide, grey – acyclic peptide and white – unprocessed substrate. Error bars illustrate standard deviation n=3 technical replicates. (**D**) SDS-PAGE analysis of inactive (pH 8) and active (pH 6.5) AtAEP2 and RcAEP1 proteins.

Table

Table 1. Crystallography data collection and refinement statistics. Numbers in parenthesis refer

644 to the highest resolution bin.

Data collection	
Space group	P3 ₁ 2 1
Unit cell dimensions	
a, b, c (Å)	77.03, 77.03, 108.17
α, β, γ (°)	90.00, 90.00, 120.00
Wavelength	0.9537
Resolution (Å)	1.8
R _{merge} (%)	6.0 (43.3)
Ι/σΙ	14.7 (2.2)
Completeness (%)	93.8 (65.1)
Redundancy	4.1 (1.8)
CC 1/2	0.997 (0.727)
Refinement	
Resolution (Å)	66.71-1.80
No. reflections	31205
R _{work} /R _{free}	15.15/18.88
No. Atoms	2415
Protein	2168
Water	229
Ligand	18
Wilson <i>B</i> (Å ²)	15.7
Average refined B-factor (Å ²)	
Protein only (Ų)	22.2
Water (Å ²)	35.0
Ligand (Å ²)	41.1
r.m.s deviations:	
Bond lengths (Å)	0.018
Bond angles (°)	1.84
Ramachandran analysis	
Favored (%)	98.84
Allowed (%)	1.16
Outliers (%)	0

647 **Figure Supplements**

648 **Figure 2—figure supplement 1** | HaAEP1 auto-catalytically activates upon pH shift to pH 4. 649 HiLoad 16/600 Superdex 200 size exclusion chromatograph of HaAEP1 purified at pH 8 (black) 650 and HaAEP1 purified at pH 8 then dialyzed at pH 4 overnight at 16°C (purple). Right insert - SDS-651 PAGE analysis of fractions collected from labelled size exclusion peaks illustrating ~52 kDa 652 inactive HaAEP1 (Lane 1), heterogeneous self-cleavage products of HaAEP1 activation (Lane 2), 653 alternatively cleaved and removed C-terminal cap domains (Lane 3 and 4). 654 Figure 2—figure supplement 2 | Sequence alignment of C13 family of cysteine proteases. 655 Proteins labelled with UniProt codes, AOAOG2RI59 – HaAEP1. Plant AEPs grouped in green box. 656 Mammalian AEPs grouped in dark blue box. GPI8 proteins grouped in grey box. Identical 657 residues shown with white text and red box, similar residues shown with red text. Potential 658 cleavage sites of HaAEP1 shown with green stars. Catalytic triad highlighted with black stars. 659 Succinimide forming aspartic acid, not conserved in GPI8 proteins highlighted with gold stars. 660 Predicted N-linked glycosylation sites in HaAEP1 and human AEP shown with black and blue inverted triangles, respectively. Variable insertion sites between β_{IV} : β_V and α 5: β 6 strands in 661 662 HaAEP1 highlighted with cyan brackets. 663 Figure 2—figure supplement 3. Topology diagram of active HaAEP1. Topology diagram based 664 on previously published diagram of highly similar structure of hAEP (Dall and Brandstetter, 665 2013). Polyproline insertion site found in HaAEP1 between β_{IV} : β_V shown in red. Residues 304-666 309 which exhibit low electron density are illustrated with red dotted loop.

667 Figure 3—figure supplement 1 | HaAEP1 active site geometry. (A) Alternate conformations of 668 catalytic residues H178 and C220, distances illustrated with text (Å) and dots. Distance of 669 His178 to other local residues indicates that, with the exception of N73 and Snn177, H178 is 670 relatively free from steric hindrance. (B) Tetrahedral peptide conformation bound to C220 671 (purple sticks) suggests that H178 and the backbone amines of G179 and C220 contribute to 672 stabilizing this intermediate by forming an oxyanion hole, hydrogen bonds illustrated with text 673 (Å) and dots. Furthermore, H178 is ideally situated to deprotonate an attacking nucleophilic 674 group leading to formation of a tetrahedral intermediate and protonate the leaving group upon 675 resolution of the intermediate.

676 Figure 3—figure supplement 2 | Polder OMIT maps define the presence of a tetrahedral

677 intermediate in the HaAEP1 active site. Polder OMIT maps (F_{obs} - F_{calc}) calculated in the absence
 678 of an AAN peptide ligand, contoured at 3σ level, indicate that models of alternative

679 intermediate AAN peptide structures (thioacyl intermediate – purple) and cleavage product

680 (hydrolysis product – gray) appear not to fit the observed electron density as well as the

681 modelled tetrahedral intermediate (second tetrahedral intermediate – green). Model peptides

are shown from above (left) and side on (right) to aid visualization of electron density maps

683 with incongruous regions of models circled (red dashed line).

Figure 4—figure supplement 1 | Analysis of WT HaAEP1 and mutants activity. (A) Expanded
view of MALDI-TOF spectra of seleno-Cys SFTI(D14N)-GLDN processing by WT and a range of
HaAEP1 mutants. Peak of mass 1608 - cyclic SFTI-1(D14N) and mass 1626 - acyclic-SFTI(D14N)
indicated with blue and red dashed lines, respectively. (B) Structure of activity based probe with

fluorescent BODIPY structure (yellow), P1 Asp residue (blue) and acyloxymethyl ketone (AOMK) highlighted. (C) SDS-PAGE analysis of Pro-AEP (Mw ~52 kDa) HaAEP1 WT and mutants at 0.1 mg/ml (pH 8). (D) Activity based probe analysis of HaAEP1 WT and mutants at pH 6.5 following activation at pH 4 illustrates heterogeneity of active HaAEP1 mutants sizes and drastically reduced levels of activity of G185S, H178A and C220S mutants.

Figure 4—figure supplement 2 | Model of SFTI-1 processing from PawS1. Model aligned on
framework of tetrahedral intermediate in the active site and where the substrate should orient
through the S1-S5 pocket, between the β5-β6 and β_{IV}:β_V loops, using a PawS1 NMR structure
(PDB 5U87). HaAEP1 modelled PawS1 cleavage at Asn49 (A), Asn18 (B) and Asp14 (C) with
expanded views of active site shown below. HaAEP1 green, SFTI-1 precursor light cyan,
Preproalbumin cyan. Cleavage residues highlighted with dark blue sticks. N-terminal Gly, shown

as magenta sticks, acts as nucleophile to attack C-terminus of SFTI-GLDN (C).

700 Figure 4—figure supplement 3. Circular dichroism analysis of HaAEP1 secondary structure. (A)

701 Melt curve of pro-AEP (pH 8) HaAEP1 measured at 222 nm, [0.2 mg/mL], shows a melting

temperature of ~52°C. n=3 technical replicates data fitted with sigmoidal four parameter

703 logistic regression fit. (B) Circular dichroism spectroscopy of pro-AEP (pH 8) WT HaAEP1 and

704 active site mutants, [0.1 mg/mL], average of n=2 technical replicates.

Figure 4—figure supplement 4 | Re-ligation of cap domain at pH 8 favored after activation at
 pH 6.5. In-gel fluorescence of activity based probe (right) and post-fluorescence Coomassie
 stain of SDS-PAGE gel (left) of WT HaAEP1, N73A, E221K and D177G mutants. Activity based
 probe illustrates presence of active HaAEPs after dialysis at pH 4/5.5/6.5. Coomassie staining of

gels reveals that upon pH shift to pH 8 re-ligation of the cap domain is favored post-dialysis at
pH 6.5, with strong pro-AEP bands also seen post-dialysis at pH 5.5 with N73A and D177G
mutants.

712	Figure 5—figure supplement 1 Comparison of crystal structures of AEPs and a model of an
713	efficient macrocyclizing AEP. Crystal structures of AEPs with APBS generated electrostatic
714	potential maps contoured at ± 10 kT/e and catalytic Cys and His shown in yellow and blue sticks,
715	respectively. Overall topology of AEPs (left) appear similar with slight differences in surface
716	electrostatic potential. Expanded views of catalytic region (β_l sheet, β_l - β_{lv} loop and β_l - β_{lll}
717	region) orientated over the catalytic His residue (right) illustrates variation around the active
718	site, with residues that are proposed to alter substrate specificity shown with sticks. A model of
719	an efficient macrocyclizer, C. ensiformis AEP (accession code: AIB06797), built on the HaAEP1
720	structure which displays 76.3 % amino acid identity, is shown below the dotted line. C.
721	ensiformis AEP appears to exhibit a more accessible active site, in comparison to other AEP
722	structures, due to the presence of several short side-chain amino acids.

724 **References**

725	Arnison, P.G., Bibb, M.J., Bierbaum, G., Bowers, A.A., Bugni, T.S., Bulaj, G., Camarero, J.A.,
726	Campopiano, D.J., Challis, G.L., Clardy, J., Cotter, P.D., Craik, D.J., Dawson, M.,
727	Dittmann, E., Donadio, S., Dorrestein, P.C., Entian, K.D., Fischbach, M.A., Garavelli, J.S.,
728	Goransson, U., Gruber, C.W., Haft, D.H., Hemscheidt, T.K., Hertweck, C., Hill, C.,
729	Horswill, A.R., Jaspars, M., Kelly, W.L., Klinman, J.P., Kuipers, O.P., Link, A.J., Liu, W.,
/30	Marahiel, M.A., Mitchell, D.A., Moll, G.N., Moore, B.S., Muller, R., Nair, S.K., Nes, I.F.,
/31	Norris, G.E., Olivera, B.M., Onaka, H., Patchett, M.L., Piel, J., Reaney, M.J., Rebuttat, S.,
132	Ross, R.P., Sani, H.G., Schmidt, E.W., Seisted, IVI.E., Severinov, K., Shen, B., Sivonen, K.,
133	Smith, L., Stein, T., Sussmuth, R.D., Tagg, J.R., Tang, G.L., Truman, A.W., Vederas, J.C.,
734 725	Waish, C.I., Walton, J.D., Wenzel, S.C., Willey, J.W., and Van der Donk, W.A. (2013).
135	Ribosomally synthesized and post-translationally modified peptide natural products:
/36	overview and recommendations for a universal nomenciature. Nat. Prod. Rep. 30, 108-
737	
738	Battye, T.G.G., Kontogiannis, L., Johnson, O., Powell, H.R., and Leslie, A.G.W. (2011).
739	MOSFLM: a new graphical interface for diffraction-image processing with MOSFLM.
740	Acta Crystallogr D Biol Crystallogr 67 , 271-281.
741	Bergmann, M., and Fruton, J.S. (1938). Some synthetic and hydrolytic experiments with
742	chymotrypsin. J. Biol. Chem. 124, 321-329.
743	Bernath-Levin, K., Nelson, C., Elliott, A.G., Jayasena, A.S., Millar, A.H., Craik, D.J., and Mylne,
744	J.S. (2015). Peptide macrocyclization by a bifunctional endoprotease. Chem. Biol. 22,
745	571-582.
746	Bhardwaj, G., Mulligan, V.K., Bahl, C.D., Gilmore, J.M., Harvey, P.J., Cheneval, O., Buchko,
747	G.W., Pulavarti, S.V., Kaas, Q., Eletsky, A., Huang, P.S., Johnsen, W.A., Greisen, P.J.,
748	Rocklin, G.J., Song, Y., Linsky, T.W., Watkins, A., Rettie, S.A., Xu, X., Carter, L.P.,
749	Bonneau, R., Olson, J.M., Coutsias, E., Correnti, C.E., Szyperski, T., Craik, D.J., and
750	Baker, D. (2016). Accurate <i>de novo</i> design of hyperstable constrained peptides. Nature
751	538, 329-335.
752	Bollini, R., and Chrispeels, M.J. (1979). The rough endoplasmic reticulum is the site of reserve-
753	protein synthesis in developing <i>Phaseolus vulgaris</i> cotyledons. Planta 146, 487-501.
754	Bowles, D.J., Marcus, S.E., Pappin, D.J., Findlay, J.B., Eliopoulos, E., Maycox, P.R., and Burgess,
755	J. (1986). Posttranslational processing of concanavalin A precursors in jackbean
756	cotyledons. J. Cell Biol. 102, 1284-1297.
757	Brady, K.D., Giegel, D.A., Grinnell, C., Lunney, E., Talanian, R.V., Wong, W., and Walker, N.
758	(1999). A catalytic mechanism for caspase-1 and for bimodal inhibition of caspase-1 by
759	activated aspartic ketones. Bioorg. Med. Chem. 7, 621-631.
760	Brennan, T.V., and Clarke, S. (1995). Effect of adjacent histidine and cysteine residues on the
761	spontaneous degradation of asparaginyl- and aspartyl-containing peptides. Int. J. Pept.
762	Protein Res. 45, 547-553.
763	Buller, A.R., and Townsend, C.A. (2013). Intrinsic evolutionary constraints on protease
764	structure, enzyme acylation, and the identity of the catalytic triad. Proc. Natl. Acad. Sci.
765	U. S. A. 110, E653-661.

- Chekan, J.R., Estrada, P., Covello, P.S., and Nair, S.K. (2017). Characterization of the macrocyclase involved in the biosynthesis of RiPP cyclic peptides in plants. Proc. Natl. Acad. Sci. U. S. A. 114, 6551-6556.
 Chen, J.M., Dando, P.M., Rawlings, N.D., Brown, M.A., Young, N.E., Stevens, R.A., Hewitt, E., Watts, C., and Barrett, A.J. (1997). Cloning, isolation, and characterization of mammalian legumain, an asparaginyl endopeptidase. J. Biol. Chem. 272, 8090-8098.
 Chen, V.B., Arendall, W.B., III, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom
- Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom
 structure validation for macromolecular crystallography. Acta Crystallogr. Sect. D Biol.
 Crystallogr. 66, 12-21.
- Chu, M.H., Liu, K.L., Wu, H.Y., Yeh, K.W., and Cheng, Y.S. (2011). Crystal structure of
 tarocystatin-papain complex: implications for the inhibition property of group-2
 phytocystatins. Planta 234, 243-254.
- Clancy, K.W., Melvin, J.A., and McCafferty, D.G. (2010). Sortase transpeptidases: insights into
 mechanism, substrate specificity, and inhibition. Biopolymers 94, 385-396.
- 781 Clark, A.C. (2016). Caspase Allostery and Conformational Selection. Chem. Rev. **116**, 6666-6706.
- Clark, R.J., Jensen, J., Nevin, S.T., Callaghan, B.P., Adams, D.J., and Craik, D.J. (2010). The
 engineering of an orally active conotoxin for the treatment of neuropathic pain. Angew.
 Chem. Int. Ed. 49, 6545-6548.
- Clark, R.J., Fischer, H., Dempster, L., Daly, N.L., Rosengren, K.J., Nevin, S.T., Meunier, F.A.,
 Adams, D.J., and Craik, D.J. (2005). Engineering stable peptide toxins by means of
 backbone cyclization: stabilization of the alpha-conotoxin MII. Proc. Natl. Acad. Sci. U. S.
 A. 102, 13767-13772.
- Dall, E., and Brandstetter, H. (2013). Mechanistic and structural studies on legumain explain its
 zymogenicity, distinct activation pathways, and regulation. Proc. Natl. Acad. Sci. U. S. A.
 110, 10940-10945.
- Dall, E., and Brandstetter, H. (2016). Structure and function of legumain in health and disease.
 Biochimie 122, 126-150.
- Dall, E., Fegg, J.C., Briza, P., and Brandstetter, H. (2015). Structure and mechanism of an
 aspartimide-dependent peptide ligase in human legumain. Angew. Chem. Int. Ed. 54,
 2917-2921.
- 797 Dodson, G., and Wlodawer, A. (1998). Catalytic triads and their relatives. Trends Biochem. Sci.
 798 23, 347-352.
- Dolinsky, T.J., Czodrowski, P., Li, H., Nielsen, J.E., Jensen, J.H., Klebe, G., and Baker, N.A.
 (2007). PDB2PQR: expanding and upgrading automated preparation of biomolecular
 structures for molecular simulations. Nucleic Acids Res. 35, W522-525.
- 802 Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta
 803 Crystallographica Section D 60, 2126-2132.
- Fagundes, D., Bohn, B., Cabreira, C., Leipelt, F., Dias, N., Bodanese-Zanettini, M.H., and
 Cagliari, A. (2015). Caspases in plants: metacaspase gene family in plant stress
 responses. Funct. Integr. Genomics 15, 639-649.
- Franke, B., Colgrave, M.L., Mylne, J.S., and Rosengren, K.J. (2016). Mature forms of the major
 seed storage albumins in sunflower: A mass spectrometric approach. J. Proteomics 147,
 177-186.

810 Franke, B., James, A.M., Mobli, M., Colgrave, M.L., Mylne, J.S., and Rosengren, K.J. (2017). 811 Two proteins for the price of one: structural studies of the dual destiny Preproalbumin 812 with Sunflower Trypsin Inhibitor-1. J. Biol. Chem. 292, 12398-12411. 813 Frankel, B.A., Kruger, R.G., Robinson, D.E., Kelleher, N.L., and McCafferty, D.G. (2005). 814 Staphylococcus aureus sortase transpeptidase SrtA: insight into the kinetic mechanism 815 and evidence for a reverse protonation catalytic mechanism. Biochemistry 44, 11188-816 11200. 817 Frankel, B.A., Tong, Y., Bentley, M.L., Fitzgerald, M.C., and McCafferty, D.G. (2007). 818 Mutational analysis of active site residues in the Staphylococcus aureus transpeptidase 819 SrtA. Biochemistry 46, 7269-7278. 820 Fuentes-Prior, P., and Salvesen, G. (2004). The protein structures that shape caspase activity, 821 specificity, activation and inhibition. Biochem. J. **384**, 201-232. 822 Geiger, T., and Clarke, S. (1987). Deamidation, isomerization, and racemization at asparaginyl 823 and aspartyl residues in peptides. Succinimide-linked reactions that contribute to 824 protein degradation. J. Biol. Chem. 262, 785-794. 825 Gould, A., Ji, Y., Aboye, T.L., and Camarero, J.A. (2011). Cyclotides, a novel ultrastable 826 polypeptide scaffold for drug discovery. Curr. Pharm. Des. 17, 4294-4307. 827 Hara-Nishimura, I., Takeuchi, Y., and Nishimura, M. (1993). Molecular characterization of a 828 vacuolar processing enzyme related to a putative cysteine proteinase of Schistosoma 829 mansoni. Plant Cell 5, 1651-1659. 830 Hara-Nishimura, I., Kinoshita, T., Hiraiwa, N., and Nishimura, M. (1998). Vacuolar processing 831 enzymes in protein-storage vacuoles and lytic vacuoles. J. Plant Physiol. 152, 668-674. 832 Harris, K.S., Durek, T., Kaas, Q., Poth, A.G., Gilding, E.K., Conlan, B.F., Saska, I., Daly, N.L., van 833 der Weerden, N.L., Craik, D.J., and Anderson, M.A. (2015). Efficient backbone 834 cyclization of linear peptides by a recombinant asparaginyl endopeptidase. Nat. 835 Commun. **6,** 10199. 836 Hatsugai, N., Kuroyanagi, M., Yamada, K., Meshi, T., Tsuda, S., Kondo, M., Nishimura, M., and 837 Hara-Nishimura, I. (2004). A plant vacuolar protease, VPE, mediates virus-induced 838 hypersensitive cell death. Science **305**, 855-858. 839 Hetrick, K.J., and van der Donk, W.A. (2017). Ribosomally synthesized and post-translationally 840 modified peptide natural product discovery in the genomic era. Curr. Opin. Chem. Biol. 841 **38,** 36-44. 842 Hiraiwa, N., Nishimura, M., and Hara-Nishimura, I. (1999). Vacuolar processing enzyme is self-843 catalytically activated by sequential removal of the C-terminal and N-terminal 844 propeptides. FEBS Lett. 447, 213-216. 845 Holm, L., and Rosenström, P. (2010). Dali server: conservation mapping in 3D. Nucleic Acids 846 Res. 38, W545-W549. 847 James, A.M., Haywood, J., and Mylne, J.S. (2017). Macrocyclization by asparaginyl 848 endopeptidases. New Phytol. 10.1111/nph.14511. 849 Jayasena, A.S., Franke, B., Rosengren, K.J., and Mylne, J.S. (2016). A tripartite 'omic analysis 850 identifies the major sunflower seed albumins. Theor. Appl. Genet. 129, 613-629. 851 Ji, Y., Majumder, S., Millard, M., Borra, R., Bi, T., Elnagar, A.Y., Neamati, N., Shekhtman, A., 852 and Camarero, J.A. (2013). In vivo activation of the p53 tumor suppressor pathway by 853 an engineered cyclotide. J. Am. Chem. Soc. 135, 11623-11633.

854 Kuroyanagi, M., Yamada, K., Hatsugai, N., Kondo, M., Nishimura, M., and Hara-Nishimura, I. 855 (2005). Vacuolar processing enzyme is essential for mycotoxin-induced cell death in 856 Arabidopsis thaliana. J. Biol. Chem. 280, 32914-32920. 857 Laskowski, R.A., and Swindells, M.B. (2011). LigPlot+: multiple ligand-protein interaction 858 diagrams for drug discovery. J Chem Inf Model 51, 2778-2786. 859 Lee, T.-W., and James, M.N.G. (2008). 1.2Å-resolution crystal structures reveal the second 860 tetrahedral intermediates of streptogrisin B (SGPB). Biochim Biophys Acta 1784, 319-861 334. 862 Liebschner, D., Afonine, P.V., Moriarty, N.W., Poon, B.K., Sobolev, O.V., Terwilliger, T.C., and 863 Adams, P.D. (2017). Polder maps: improving OMIT maps by excluding bulk solvent. Acta 864 Crystallogr D Struct Biol 73, 148-157. 865 Lu, H., Chandrasekar, B., Oeljeklaus, J., Misas-Villamil, J.C., Wang, Z., Shindo, T., Bogyo, M., 866 Kaiser, M., and van der Hoorn, R.A. (2015). Subfamily-specific fluorescent probes for 867 cysteine proteases display dynamic protease activities during seed germination. Plant 868 Physiol. 168, 1462-1475. 869 Luckett, S., Garcia, R.S., Barker, J.J., Konarev, A.V., Shewry, P.R., Clarke, A.R., and Brady, R.L. 870 (1999). High-resolution structure of a potent, cyclic proteinase inhibitor from sunflower 871 seeds. J. Mol. Biol. 290, 525-533. 872 Manoury, B., Hewitt, E.W., Morrice, N., Dando, P.M., Barrett, A.J., and Watts, C. (1998). An 873 asparaginyl endopeptidase processes a microbial antigen for class II MHC presentation. 874 Nature **396,** 695-699. 875 Martinez, M., Cambra, I., Gonzalez-Melendi, P., Santamaria, M.E., and Diaz, I. (2012). C1A 876 cysteine-proteases and their inhibitors in plants. Physiol. Plant 145, 85-94. 877 Mathieu, M.A., Bogyo, M., Caffrey, C.R., Choe, Y., Lee, J., Chapman, H., Sajid, M., Craik, C.S., 878 and McKerrow, J.H. (2002). Substrate specificity of schistosome versus human legumain 879 determined by P1-P3 peptide libraries. Mol. Biochem. Parasitol. 121, 99-105. 880 Mazmanian, S.K., Liu, G., Ton-That, H., and Schneewind, O. (1999). Staphylococcus aureus 881 sortase, an enzyme that anchors surface proteins to the cell wall. Science **285**, 760-763. 882 McLuskey, K., Rudolf, J., Proto, W.R., Isaacs, N.W., Coombs, G.H., Moss, C.X., and Mottram, 883 J.C. (2012). Crystal structure of a *Trypanosoma brucei* metacaspase. Proc. Natl. Acad. 884 Sci. U. S. A. 109, 7469-7474. 885 Min, W., and Jones, D.H. (1994). In vitro splicing of concanavalin A is catalyzed by asparaginyl 886 endopeptidase. Nat. Struct. Mol. Biol. 1, 502-504. 887 Mylne, J.S., Colgrave, M.L., Daly, N.L., Chanson, A.H., Elliott, A.G., McCallum, E.J., Jones, A., 888 and Craik, D.J. (2011). Albumins and their processing machinery are hijacked for cyclic 889 peptides in sunflower. Nat. Chem. Biol. 7, 257-259. 890 Mylne, J.S., Chan, L.Y., Chanson, A.H., Daly, N.L., Schaefer, H., Bailey, T.L., Nguyencong, P., 891 Cascales, L., and Craik, D.J. (2012). Cyclic peptides arising by evolutionary parallelism via 892 asparaginyl-endopeptidase-mediated biosynthesis. Plant Cell 24, 2765-2778. 893 Nguyen, G.K.T., Wang, S., Qiu, Y., Hemu, X., Lian, Y., and Tam, J.P. (2014). Butelase 1 is an Asx-894 specific ligase enabling peptide macrocyclization and synthesis. Nat. Chem. Biol. 10, 732-895 738. 896 Otto, H.H., and Schirmeister, T. (1997). Cysteine proteases and their inhibitors. Chem. Rev. 97, 897 133-172.

- Pattabiraman, V.R., and Bode, J.W. (2011). Rethinking amide bond synthesis. Nature 480, 471479.
- Poth, A.G., Chan, L.Y., and Craik, D.J. (2013). Cyclotides as grafting frameworks for protein
 engineering and drug design applications. Biopolymers 100, 480-491.
- Radisky, E.S., Lee, J.M., Lu, C.J.K., and Koshland, D.E. (2006). Insights into the serine protease
 mechanism from atomic resolution structures of trypsin reaction intermediates. Proc.
 Natl. Acad. Sci. U. S. A. 103, 6835-6840.
- Rawlings, N.D., Barrett, A.J., and Finn, R. (2016). Twenty years of the MEROPS database of
 proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. 44, D343-D350.
- 907 Reissner, K.J., and Aswad, D.W. (2003). Deamidation and isoaspartate formation in proteins:
 908 unwanted alterations or surreptitious signals? Cell Mol Life Sci 60, 1281-1295.
- Ritz-Timme, S., and Collins, M.J. (2002). Racemization of aspartic acid in human proteins.
 Ageing Res. Rev. 1, 43-59.
- 911 Roy, A., Kucukural, A., and Zhang, Y. (2010). I-TASSER: a unified platform for automated
 912 protein structure and function prediction. Nat Protoc 5, 725-738.
- Schechter, I., and Berger, A. (1967). On the size of the active site in proteases. I. Papain.
 Biochem. Biophys. Res. Commun. 27, 157-162.
- 915 Schrodinger, L. (2010). The PyMOL Molecular Graphics System, Version 1.4.
- Shafee, T., Harris, K., and Anderson, M. (2015). Biosynthesis of cyclotides. In Advances in
 Botanical Research (Academic Press), pp. 227-269.
- Sheldon, P.S., Keen, J.N., and Bowles, D.J. (1996). Post-translational peptide bond formation
 during concanavalin A processing *in vitro*. Biochem. J. **320**, 865-870.
- Shewry, P.R., and Halford, N.G. (2002). Cereal seed storage proteins: structures, properties and
 role in grain utilization. J. Exp. Bot. 53, 947-958.
- Sojka, D., Hajdusek, O., Dvorak, J., Sajid, M., Franta, Z., Schneider, E.L., Craik, C.S., Vancova,
 M., Buresova, V., Bogyo, M., Sexton, K.B., McKerrow, J.H., Caffrey, C.R., and Kopacek,
 P. (2007). IrAE: an asparaginyl endopeptidase (legumain) in the gut of the hard tick
 Ixodes ricinus. Int. J. Parasitol. **37**, 713-724.
- Suree, N., Liew, C.K., Villareal, V.A., Thieu, W., Fadeev, E.A., Clemens, J.J., Jung, M.E., and
 Clubb, R.T. (2009). The structure of the *Staphylococcus aureus* sortase-substrate
 complex reveals how the universally conserved LPXTG sorting signal is recognized. J.
 Biol. Chem. 284, 24465-24477.
- Swedberg, J.E., Li, C.Y., de Veer, S.J., Wang, C.K., and Craik, D.J. (2017). Design of potent and
 selective cathepsin G inhibitors based on the sunflower trypsin inhibitor-1 scaffold. J.
 Med. Chem. 60, 658-667.
- Takahashi, O., Manabe, N., and Kirikoshi, R. (2016). A computational study of the mechanism
 of succinimide formation in the Asn-His sequence: intramolecular catalysis by the His
 side chain. Molecules 21, 327.
- 936 Truman, A.W. (2016). Cyclisation mechanisms in the biosynthesis of ribosomally synthesised
 937 and post-translationally modified peptides. Beilstein Journal of Organic Chemistry 12,
 938 1250-1268.
- Tsiatsiani, L., Van Breusegem, F., Gallois, P., Zavialov, A., Lam, E., and Bozhkov, P.V. (2011).
 Metacaspases. Cell Death Differ. 18, 1279-1288.

Vercammen, D., van de Cotte, B., De Jaeger, G., Eeckhout, D., Casteels, P., Vandepoele, K.,
 Vandenberghe, I., Van Beeumen, J., Inze, D., and Van Breusegem, F. (2004). Type II
 Metacaspases Atmc4 and Atmc9 of *Arabidopsis thaliana* cleave substrates after arginine
 and lysine. J. Biol. Chem. 279, 45329-45336.

- Vernet, T., Tessier, D.C., Chatellier, J., Plouffe, C., Lee, T.S., Thomas, D.Y., Storer, A.C., and
 Menard, R. (1995). Structural and functional roles of asparagine 175 in the cysteine
 protease papain. J. Biol. Chem. 270, 16645-16652.
- White, A.M., and Craik, D.J. (2016). Discovery and optimization of peptide macrocycles. Expert
 Opin. Drug Discov. 11, 1151-1163.
- Wilmouth, R.C., Edman, K., Neutze, R., Wright, P.A., Clifton, I.J., Schneider, T.R., Schofield,
 C.J., and Hajdu, J. (2001). X-ray snapshots of serine protease catalysis reveal a
 tetrahedral intermediate. Nat. Struct. Biol. 8, 689-694.
- Winn, M. D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M.,
 Krissinel, E.B., Leslie, A.G.W., McCoy, A., McNicholas, S.J., Murshudov, G.N., S. Pannu,
 N., Potterton, E.A., Powell, H.R., Read, R.J., Vagin, A., and Wilson, K.S. (2011).
 Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol
 Crystallogr 67, 235-242.
- Yamada, K., Shimada, T., Nishimura, M., and Hara-Nishimura, I. (2005). A VPE family
 supporting various vacuolar functions in plants. Physiol. Plant 123, 369-375.
- Yang, R., Wong, Y.H., Nguyen, G.K.T., Tam, J.P., Lescar, J., and Wu, B. (2017). Engineering a
 catalytically efficient recombinant protein ligase. J. Am. Chem. Soc. 139, 5351-5358.
- Youle, R.J., and Huang, A.H.C. (1978). Albumin storage proteins in the protein bodies of castor
 bean. Plant Physiol. 61, 13.
- Zacks, M.A., and Garg, N. (2006). Recent developments in the molecular, biochemical and
 functional characterization of GPI8 and the GPI-anchoring mechanism Mol. Membr. Biol.
 23, 209-225.
- 2hao, L., Hua, T., Crowley, C., Ru, H., Ni, X., Shaw, N., Jiao, L., Ding, W., Qu, L., Hung, L.-W.,
 Huang, W., Liu, L., Ye, K., Ouyang, S., Cheng, G., and Liu, Z.-J. (2014). Structural analysis
 of asparaginyl endopeptidase reveals the activation mechanism and a reversible
 intermediate maturation stage. Cell Res. 24, 344-358.
- 971
- 972
- 973
- 974

Leader RiPP Follower















	÷			30	4 <u>0</u>	50	1 T	60	/ Ÿ	- ;
A0A0G2RI59	MVSRIIC	FTL.VLVTVVAL:	SYGAAGRES	SGGQKWRWGWD	PLIRSPVDAE	QE	VDEQMT	NGTKWAVLV	/ A G <mark>S</mark> K G Y G N	Y R H Q A D V C H A Y Q '
P49044	MGSSQLSTLLF	FT IVVTFLTV	VSSGRD.	L P G	DYLRLPSETS	RFFREP	KNDDDF	EGTR <mark>W</mark> AILI	AG <mark>S</mark> NGYW <mark>N</mark>	YRHQSDVCHAYQ
P49047	MTTVVS	FLALFLF	LVAA .	V S G	DVIKLPSLAS	KF F R P T	ENDDD.	. STKWAVLV	/ A G <mark>S</mark> S G Y W <mark>N</mark>	YRHQADVCHAYQ:
B8ASK4	MAARC	WVWGFVVALLAV	AAAADGE	E E E G K W E	PLIRMPTEEG	DDAEAAAPAPA	PAAADY	GGTRWAVLV	/ A G <mark>S</mark> S G Y G <mark>N</mark>	YRHQADVCHACQ:
P49042	METHKSLLFFTNYVL	FLVFTLS.FLPI	P.GLLAS	RLNPFE	PGILMPTEEA	E P V Q	VDDDDQ	LGTRWAVLV	/ A G <mark>S</mark> M G F G <mark>N</mark>	YRHQADVCHAYQ:
024326	MAVHRSLLNKPTWCRVAF	WWWMLVM.VMRI	Q. <mark>G</mark>	TNGKEQD	SVIKLPTQEV		DAESDE	VGTRWAVLV	/ A G <mark>S</mark> N G Y G <mark>N</mark>	YRHQADVCHAYQ
P49046	M	VMMLV.MLSL	H.GTAAR	LNRREWD	SVIQLPTEPV		DDE	VGTRWAVLV	A G <mark>S</mark> N G Y G <mark>N</mark>	YRHQADVCHAYQ:
P49045	MALDRSIISKTTWYSVVL	WMMVV.LVRV	H.GAAAR	PNRKEWD	SVIKLPTEPV		DADSDE	VGTRWAVLV	A G <mark>S</mark> N G Y G <mark>N</mark>	YRHQADVCHAYQ.
Q9R0J8		MTWRVAV.LLSL	VLGAGA			VHIG	VDDPED	GGKHWVVIV	AGSNGWYN	Y R H Q A D A C H A Y Q
089017		MTWRVAV.LLSL	VLGAGA			VPVG	VDDPED	GGKHWVVIV	AGSNGWYN	Y R H Q A D A C H A Y Q
Q99538		MVWKVAV.FLSV,	ALGIGA			VP	IDDPED	GGKHWVVIV	AGSNGWYN	YRHQADACHAYQ.
Q5K5D9		MVWKVAV.FLSA	ALVIGA			VP	IDDPED	GGKHWVVI	AGSNGWYN	YRHQADACHAYQ
Q95M12		MIWEFIV.LLSL	VLGIGA				LEDPED	GGKHWVVIV	AGSNGWYN	YRHQADACHAYQ.
P49048		MKHVLLI	FCALLATEA	L	LNIGLQL	KIDE	LFDIPG	HINNWAVLV	CISKEWEN	YRHVSNVLALYH:
QOLATI	MAADCEIT IDVA	T TAAT ATT	CTC C		NISVLFE	DONE	AAVKSI		UASKI WIN	V PHUANTI CUVD
092643	MAAFCELIBRVA	V IATV III	STC		VAAGHIE	DOAR	OFFRSG	HTNNWAVIN	CISKEWEN	V DHUANTI SUVD
058618	MAVIDOLO DANG	T TANY ITT			VAASHIB	DONE	OFFRSC	UTNNWAVIA	CTSPEWEN	V D H V A N T I S V I K
O3MHZ7	MUDICELS RGLT	T LAGI LIL	PFC C		LAASOT F	DOAE	OFFRSG	HTNNWAVLV	CISREWEN	V R H V A N T L S V Y R
O4KBV1	MVGTWFLC RGFT	T LAGL LLLI	PFG S		LAASOI E	DOAE	OFFRSG	HTNNWAVLV	CTSREWEN	Y B H V A N T L S V Y B
2111111		1			Linio y 1 D	<u>Dynu</u>	Q11R00			
	90 100	110_	120	130	140	150	160	17	' o 🔶	1.80 19
A0A0G2RI59	LKKGGLKDENIVVFMYDD	IAKSEMNPRPGI	IINSP.KGE	OVYA.GVPKDY	TGKNVTVDNL	SAVLLGDRSAV	K G .GSG	KVVDSKPED	RIFLFYSD	HGGPGVLGMPNE
P49044	LRKGGSKEENIIVFMYDD	IASNEE <mark>NP</mark> RPGV	IINKPD.GD	OVYA.G <mark>V</mark> PK DY	TGAEVHADNF	YAALLGNKSAL	TG.GSG	KVVDSGPNE	HIFVYYTD	H G G P G V L G M P V G I
P49047	LKKGGVKEENIVVFMYDD	IAKNEE <mark>NP</mark> RPGV	IINSPN.GE	OVYN.G <mark>V</mark> PK DY	TGDEVNVDNL	LAVIL <mark>G</mark> NKTAL	K <mark>G</mark> .GSG	K V V D S G P N E	HIFIYYSD	H G G P G V L G M P T S I
B8ASK4	LQKGGVKEENIVVFMYDD	IAHNILNPRPGT	IINHPK.GG	Ο <mark>νγ</mark> Α.G <mark>ν</mark> ρκ <mark>dγ</mark>	TGHOVTTENF	FAVLLGNKTAV	T <mark>G</mark> .GSG	KVIDSKPEI	HIFIYYSD	H G G P G V L G M P N L I
P49042	LRKGGLKEENIIVFMYDD	IAKNEL <mark>NP</mark> RPGV	IINHPQ.GE	DV <mark>Y</mark> A.G <mark>V</mark> PK <mark>DY</mark>	TGEH <mark>V</mark> TAKNL	YAVLLGDKSAV	Q <mark>G</mark> .GSG	K V V D S K P N I	RIFLYYSD	H G G P G V L G M P N L I
024326	LIKGGVKEENIVVFMYDD	IATHELNPRPGV	IINNPQ.GP	Ο <mark>Υ</mark> Α. G <mark>ν</mark> ρκ dy	TGESVTSHNF	FAVLLGDKSKV	K <mark>G</mark> .GSG	KVINSKPEI	RIFVYYSD	H G G P G V L G M P N M I
P49046	LIKGGVKEENIVVFMYDD	IAYNAM <mark>NP</mark> RPGV	IINHPQ.GP	DVYA.GVPKDY	TGEDVTPENL	YAVIL <mark>G</mark> DKSKV	K <mark>G</mark> .GSG	KVINSNPEI	RIFIFYSD	H G G P G V L G M P N A I
P49045	LIKG <mark>G</mark> LKEEN <mark>I</mark> VVFMYDD	IATNEL <mark>NP</mark> RHGV	IINHPE.GE	D L <mark>Y</mark> A . G <mark>V</mark> P K <mark>D Y</mark>	TGDNVTTENL	FAVIL <mark>G</mark> DKSKL	K <mark>G</mark> .GSG	K V I N S K P E E	RIFIYYSD	H G G P <mark>G</mark> I L G M P N M I
Q9R0J8	IHRNGIPDEQIIVMMYDD	IANNEE <mark>NP</mark> TPGV	VINRPN.GT	OVYK.GVPKDY	TGEDVTPENF	LAVLR <mark>G</mark> DEEAV	KGKGSG	KVLKSGPRI	HVFVYFTD	H <mark>G</mark> A T <mark>G</mark> I L V F P N E I
089017	IHRNGIPDEQIIVMMYDD	IANSEENPTPGV	VINRPN.GT	DVYK.GVLKDY	TGEDVTPENF	LAVLR <mark>G</mark> DAEAV	KGKGSG	K V L K S G P R I	HVFIYFTD	H G A T G I L V F P N D I
099538	I H R N G I P D E Q I V V M M Y D D	IAYSEDNPTPGI	VINRPN.GT	JVYQ.GVPKDY	TGEDVTPQNF	LAVLRGDAEAV	KGIGSG	KVLKSGPQI	HVFIYFTD	HGSTGILVFPNEI
Q5R5D9	THRNGIPDEQIVVMMYDD	IAYSEDNPTPGI	VINRPN.GT	NARO' CALKDA	TGEDVTPQNF	LAVLRGDAEAV	KGIGSG	K V L K S G P Q I	HVEVYSTD	HGSTGILVFPNEI
Q95M12	V H KINGIPDEQIIVMMYDD	IANSEDNPTPGI	VINRPN.GS	JVYQ.GVLKDY	IGEDVTPKNF	LAVLRGDAEAV	KGVGSG	KVLKSGPRI	HVEVYFTD	HGATGILVFPNEI
P49048	TARLGIPDSNIIMMLAED	V P C N S R N P R P G T	VIAAR.AGT	NILYGSDVEVDY	RGEEVIVESF	IKVLTGRHHPA	IP RS	KKLLTDHQS	NVLLYLG	HGGDSFMKFQDSI
Q8T4E1	VARESIPDSUIILMIADD	MACNARNPRPGQ	VINNANQHI		RGIEVIVENE	VKLLTGKTQNG	TAL. RS	KLLSDAGS	NVLLYLTG	HCONGELKFODSI
Q9CXY9	VKRLGIPDSHIVLMLADD	MACNARNPKPAT	VFSHKNMEL	NVYGDDVEVDY	RSYEVIVENE	LRVLTGRVPPS	TPRS	KRLLSDDRS	SNILIYMIG	HGGNGFLKFQDS.
092643	VKRLGIPDSHIVLMLADD	MACNPRNPKPAT	VFSHKNMEL	NVYGDDVEVDY	RSIEVIVENE	LRVLIGRIPPS	TP RS	KRLLSDDRS	NILIYMIG	HGGNGFLKFQDS.
QSK6L8	VKRLGIPDSHIVLMLADD	MACNPRNPRPAT	VESEKNMEL		RSIEVIVENE	TRVLIGRIPPS	TD DC	KKLLSDDKS	NILIIMIG	HGGNGFLKFQDS.
O4KRV1	VKRLGIPDSHIVIMLADD	MACNERNEREAT	V Y S H K N M F L	VV CDDVEVDI	RSILVIVENE	LRVLIGRIPSS	TP PS	KRILSDDKS	NILIIMIG	HCCNCELKEODS
2111111		Indon't Real of the					11			
	2 00 2	10 22	23	o, 240	250	260	-	270	280	290
A0A0G2RI59	HLVAKDLVDVLKKKHAMG	TYKEMVIYLEAC	ESG <mark>S</mark> IFEGI	LPEDLNIYATT	ASGAQENSYG	TYCPGTEPSPP	PEYITC	LGDLYS	VAWMEDSE	THNLKKESLEQQ
P49044	YLYASDLNEVLKKKHASG	TYKSLVFYLEAC	ESG <mark>S</mark> IFEGL	LPDDLNIYATT	A	YYCPGDKPPPP	PEYSTC	LGDLYS	JAWMEDSE	VHN <mark>LQ</mark> TE <mark>SL</mark> QQQ
P49047	NLYANDLNDVLKKKYASG	TYKSLVFYLEAC	ESG <mark>S</mark> IFEGL	LPEGLNIYATT	A	TYCPGEDPSPP	SEYETC	L G <mark>D</mark> L Y S	VAWIEDSE	K H N <mark>L Q</mark> T E T L H E Q '
B8ASK4	Y L Y A G D F I K V L Q K K H A S N	SYSKMVIYVEAC	ESG <mark>S</mark> IFEGL	MPENLNIYVTT	A	TYCPGEEPSPP	PEYITC	LG D M Y S	V A W M E D S E	THNLKKETIEDQ
P49042	Y L Y A M D F I E V L K K K H A A G	GYKKMVIYVEAC	ESG <mark>S</mark> IFEGI	MPKDVDIYVTT	A S N A Q 🖻 S <mark>S</mark> W G	TYCPGMEPSPP	PEFTTC	LGDLYS	S V A W M E D S E	SHNLKKETVKQQ
024326	YLYAMDFIDVLKKKHASG	GYKEMVIYVEAC	ESG <mark>S</mark> IFEGI	MPKDLNIYVTT	ASNAQENSWG	TYCPGMYPPPP	PEYITC	LGDLYS	SVAWMEDSE	SHNLKKESVEQQ
P49046	FVYAMDFIDVLKKKHASG	GYKEMVIYIEAC	ESGSIFEGI	MPKDLNIYVTT	ASNAQENSFG	TYCPGMNPPPP	EEYVTC	LGDLYS	SVSWMEDSE	THNLKRETVQQQ
P49045	YLYAMDFIDVLKKKHASG	SYKEMVIYVEAC	ESGSVFEGI	MPKDLNIYVTT	ASNAQENSWG	TYCPGMDPSPP	PEYITC		SVAWMEDSE	AHNLKRESVKQQ
Q9R018	. LHVKDLNKTIRYMYEHK	MYQKMVFYIEACI	ESGSMMN.H	LPDDIDVYATT	AANPNESSYA	CIIDE	. ERSTY		SVNWMEDSD	VEDLIKETLHKQ VEDLIKETLHKQ
089017	. LHVKDLNKIIKIMIEHK	MYDKMVEYIEACI	ESGSMMN.H	LPDDINVIAII	AANPKESSIA	CIIDE	. EKGII		VNWMEDSD	VEDLIKEILHKQ VEDLTKETLUKO
057570	. LAVEDINEITAIMIKAK	MYDYMVEYTEACI	ESGSMMN.H	LPDNINVIAII I DDNINVVATT	AANPRESSIA	CIIDE	. KRSII		VNWMEDSD	VEDLIKEILHKQ
0.05M12			SOSMMN.II	DE DUITUN TUTI	AUNT NEOOTU		• IV IV O I I	L G L W • • • L 🗸		VEDDINETHINV
	THARDINELIBAWAEHK	MMOKMVEYTEA []	FICICISMMN H		AANDRRSSVA	CYVDE	ORSTE	LCDW VS	VNWMEDSD	VEDITKETIHKO
P49048	LHVKDLNETIRYMYEHK	MYQKMVFYIEACI RYHEMLVIADSCI	ESG <mark>S</mark> MMN.H	LPPDINVYATT DSPN VISIS	AANPRESSYA	CYYDE	.QRSTF	LG <mark>D</mark> W <mark>Y</mark> S VIDRYTH Y I	VNELTKEV	VEDLTKETLHKQ KALNSSANMODY
P49048 08T4E1	LHVKDLNETIRYMYEHK ELTNVDLAYAIQTMFEDN ELTSOELADGIOOMWEKK	MYQKMVFYIEACI RYHEMLVIADSCI RYNELFFMVDTC	ESG <mark>S</mark> MMN.H RSA <mark>S</mark> MYEWI OAASLYEKF	LPPDINVYATT DSPNVLSL <mark>S</mark> TSPNVLAVA	AANPRESSYA SSLTHEESYS SSLVGEDSLS	CYYDE YDVD HHVD	.QRSTF TDIGVY PSIGVY	LGDWYS VIDRYTHYI MIDRYTYYZ	VNWMEDSD VNFLTKEV	VEDLTKETLHKQ KALNSSANMQDY P. FSKBTIGEF
Q95M12 P49048 Q8T4E1 O9CXY9	LHVKDLNETIRYMYEHK ELTNVDLAYAIQTMFEDN EITSQELADGIQQMWEKK ELTNIELADAFEOMWOKB	MYQKMVFYIEAC RYHEMLVIADSCI RYNELFFMVDTC RYNELLFTTDTC	ESG <mark>S</mark> MMN.H RSA <mark>S</mark> MYEWI QAA <mark>S</mark> LYEKF OGA S MYERF	LPPDINVYATT DSPNVLSLS ISPNVLAVA YSPNTMALA	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSOVGEDSLS	<u>CYYDE</u> D YDVD HHVD	.QRSTF TDIGVY PSIGVY PAIGVH	LGDWYS VIDRYTHYI MIDRYTYYA IMDRYTFYY	VNWMEDSD VNFLTKEV LEFLEKVQ	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF
Q95M12 P49048 Q8T4E1 Q9CXY9 O92643	L H V KD LNE T IR YMYEHK ELTN VD LA YA IQ TMFEDN EITS QE LAD G IQ QMWEKK EITN IE LAD AFE QMWQK R EITN IE LAD AFE OMWOK R	MYQKMVFYIEAC RYHEMLVIADSC RYNELFFMVDTC RYNELLFIIDTC RYNELLFIIDTC	ESG <mark>S</mark> MMN.H RSASMYEWI QAASLYEKF QGASMYERF OGASMYERF	LPPDINVYATT OSPNVLSLS ISPNVLAVA YSPNIMALA YSPNIMALA	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSQVGEDSLS SSOVGEDSLS	<u>CYYDE</u> YDVD HHVD HQPD HOPD	.QRSTF TDIGVY PSIGVY PAIGVH PAIGVH	LGDWYS VIDRYTHYI MIDRYTYY LMDRYTFYV LMDRYTFYV	VNWMEDSD VNFLTKEV LEFLEKVQ LEFLEEIN LEFLEEIN	VEDLTKETLHKQ KALNSSANMQDY PFSKRTIGEF PASQTNMNDL PASOTNMNDL
P49048 Q8T4E1 Q9CXY9 Q92643 Q5R6L8	LHVKDLNETIRYMYEHK ELITNVDLAYAIQTMFEDN EITSQELADGIQQMWEKK EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR	MYQKMVFYIEAC RYHEMLVIADSC RYNELLFIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC	E S G SMMN . H R S A SM YE W I Q A A S L YE K F Q G A SM YE R F Q G A SM YE R F Q G A SM YE R F	LPPDINVYATT DSPNVLSLS ISPNVLAVA YSPNIMALA YSPNIMALA YSPNIMALA	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS	CYYDE YDV D HHV D HQP HQP D HQP D HQP	QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH	LGDWYS VIDRYTHY MIDRYTYY LMDRYTFY LMDRYTFY LMDRYTFY LMDRYTFY	VNWMEDSD VNFLTKEV LEFLEKVQ LEFLEEIN LEFLEEIN LEFLEEIN	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL
Q95M12 Q8T4E1 Q9CXY9 Q92643 Q5R6L8 Q3MHZ7	LHVKDLNETIRYMYEHK ELTNVDLAYAIQTMFEDN ELTSQELADGIQQMWEKK EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR	MYQKMVFYIEAC RYHEMLVIADSC RYNELLFIDTC RYNELLFIDTC RYNELLFIDTC RYNELLFIDTC RYNELLFIDTC RYNELLFIDTC	ESG <mark>S</mark> MMN.H RSA <mark>S</mark> MYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF QGASMYERF	LPPDINVYATT OSPNVLSLS ISPNVLAVA YSPNIMALA YSPNIMALA YSPNIMALA YSPNIMALA	AANPRESSYA SSLTHEESYS SSLVGEDSLS SQVGEDSLS SQVGEDSLS SQVGEDSLS SQVGEDSLS SQVGEDSLS	CYYDE YDV HHV HQP HQP HQP HQP HQP D HQP D	.QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAIGVH	LGDWYS VIDRYTHY MIDRYTYYA LMDRYTFY LMDRYTFY LMDRYTFY LMDRYTFY LMDRYTFY	<u>S V N W MED SD</u> V N F L T K E V A LE F L E K V Q / LE F L E E I N / LE F L E E I N / LE F L E E I N / LE F L E E I N	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL
Q95M12 P49048 Q8T4E1 Q9CXY9 Q92643 Q5R6L8 Q3MHZ7 Q4KRV1	. LH VKD LNE T IR YMYEHK ELITNVD LAYA I OTMFEDN E ITS O ELAD G I O OMWEKK E ITN I ELAD AFE O MWOKR E ITN I E LAD AFE O MWOKR	MYQKMVFYIEA RYHEMLVIADSC RYNELFFMVDIC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC	ESGSMMN.H RSASLYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF	LPPDINVYATT DSPN.VLSLS ISPN.VLAVA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS	CYYDE YDV DHU HQP HQP HQP HQP HQP HQP DHQP DQP	.QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAVGVH	LGDWYS VIDRYTHYT MIDRYTYY LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV	VNWMEDSD VNFLTKEV LEFLEKVQ LEFLEEIN /LEFLEEN /LEFLEEN /LEFLEEN /LEFLEEN /LEFLEEN /LEFLENN	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL
095M12 P49048 Q8T4E1 Q9CXY9 Q92643 Q5R6L8 Q3MHZ7 Q4KRV1	LHVKDLAYAIOTMFEDN ELITNVDLAYAIOTMFEDN EITSQELADGIQQMWEKK EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR BITNIELADAFEQMWQKR 300 310	MYQKMVFYIEAC RYHEMIVIADSC RYNELFFMVDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF 330	LPPDINVYATT OSPN.VLSUS ISPN.VLSVA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA	$\begin{array}{c} \textbf{AANPRESSYA}\\ \textbf{SSLTHESSYS}\\ \textbf{SSLVGEDSLS}\\ \textbf{SSQVGEDSLS}\\ \textbf{SQVGEDSLS}\\ \textbf{SQVGEDSLS}\\ \textbf{SQVGEDSLS}\\ \textbf{SQVGEDSLS}\\ \textbf{SQVGEDSLS}\\ \textbf{SQVGEDSLS}\\ \textbf{SQVGEDSLS}\\ \end{array}$	CYYDE. YDVD HHVD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD	QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAIGVH	LGDWYS VIDRYTHYT MIDRYTYYY LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV	VNWMEDSD VNFLTKEV LEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL 380 39
Q95M12 P49048 Q8T4E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159	. LHVKD LNE TIRYMYEHK ELITNVDLAYAIOTMFEDN EITSQELADGIQOWWEKK EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR 300 310 NKVKKRTSNSNTYNTGSH	MYQKMVFYLEAC RYHEMLVIADSC RYNELFFMVDTC RYNELFFIDTC RYNELFFIDTC RYNELFFIDTC RYNELFFIDTC 320 VMEYGSKDIKPE	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF QGASMYERF 330 KVYYLYLGFD	LPPDINVYATT DSPN.VLSLS ISPN.VLAVA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA PSPN.IMALA	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS	CYYDE YDVD HHVD HQPD JSQ DKLDGVNIORD	QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAVGVH PAIGVH	LIGDWYS VIDRYTHYT MIDRYTFY LMDRYTFY LMDRYTFY LMDRYTFY LMDRYTFY LMDRYTFY S70 WORYKKSSE	VNWMEDSD VNFLTKEV VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL 380,39 JREITETLTHRG
Q95M12 Q8T4E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044	. LHVKDLNETIRYMYEHK ELITNVDLAYAIOTMFEDN EITSOELADGIQOMWEKK EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR 300 310 NKVKKRTSNSNTYNTGSH KLVKNRTI.SPYGN	MYQKMVFYLEAC RYHEMLVIADSC RYNELFFMVDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC 320 VMEYGSKDIKPE	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF QGASMYERF 330 KVYLYLGFD DLYOYLGTN	LPPDINVYATT DSPN.VLSIS ISPN.VLAVA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS NQIHF VDETENSLKL	CYYDE. YDVD HHVD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD JKLDGVNQRD CKLDGVNQRD	.QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAVGVH PAIGVH ADLIFFL ADLIFFL	LGDWY VIDRYTHYI MIDRYTYY LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV MDRYTFYV 370 WQRYKKSSE WEKFRKAPE	VNWMEDSD VNFLTKEV ALEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL 380 380 380 280 280 280 280 280 280 280 280 280 2
293M12 P49048 Q874E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49044	. LHVKD LNE TIRYMYEHK ELTNVD LAYAIOTMFEDN EITSQELADGIQQWWEKK EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQWWQKR EITNIELADAFEQWWQKR EITNIELADAFEQWWQKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEQKKR BITNIELADAFEX BITNI	MYQKMVFYLEAC RYHEMLVIADSC RYNELFFMVDIC RYNELLFIIDIC RYNELLFIIDIC RYNELLFIIDIC RYNELLFIIDIC RYNELFIIDIC RYNELFIIDIC WYNEYGSKDIKPE VMEYGDIGLSKNI	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF 330 KVYLYLGFD LYQYLGFN LVLYMGTN	LPPDINUYATT DSPN. VLSUS ISPN. VLAVA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA PATV. NLPA PATV. NLPA PATV. NSF PADE. NETF	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS NQIHF VDETENSLKL VNENSI	CYYDE YDV HHV HQP HQP HQP HQP HQP HQP DKLDGVNQRD S50 DKLDGVNQRD RTPSAAVNQRD RTPSAVNQRD	.QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIFL ADLIFH	LGDWYS VIDRYTHYP MIDRYTYP MIDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV MORYTFYV MORYTFYV MORYTFYV MORYTFYV MORYTFYV MORYTFYV WORYKKSSE WEKFRKAPP WHKYQKAPP	VNWMEDSO VNFLTKEV ALEFLEEVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN SS. TRPEI	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL E.SQUNMNDL B.ASQTNMNDL C.ASQTNMNDL B.ASQTNMNDL C.ASQ
Q92M12 Q874E1 Q92XY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4	. LHVKD LNE TIRYMYEHK ELTNVELAYAIOTMFEDN EITSQELADGIQQMWEKK EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR 300 310 NKVKRTSNSNTYNTGSH KLVKNRTSNSNTYNTGSH ELVKKRTSGSG.KSYGSH ELVKKRTSSNKLNEGSH	MYQKMVFYLEAC RYHEMLVIADSC RYNELFFMVDTC RYNELFFIDTC RYNELFIDTC RYNELFIDTC RYNELFIDTC VNELFIDTC VMEYGSKDIKPE VMEYGDIGLSKN VMEYGDKTFKDE	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF 330 KVYLIYLGFD DLYQYLGFN KLVLFMGFN KLVLFMGFN	LPPDINVYATT DSPN.VLSLS ISPN.VLAVA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA PATV.NLPA PATV.NLPA PAND.NSF PANG.NITN	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS NQIHF VDETENSLKL VNENSLL	CYYDE YDV HHV HQP HQP HQP HQP HQP HQP HQP BASO JSO DKLDGVNORD RPPSRVNORD PVPKATVNORD	. QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAVGVH PAVGVH ADLIFH ADLIFM	LGDWY3 VIDRYTHYF MIDRYTFY LMDRYTFY WQRYKKSSE WQRYKKSSE WEKFRKAPE WHKYQEQLNG	VNM MEDSO VNFLTKEV ILEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN CSSSQKNEA CSSARKVEA CSSARKVEA	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.EITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK LREIEDTIAHRK
Q30H12 Q8T4E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49042	. LHVKDLNE TIRYMYEHK ELITNVDLAYAIOTMFEDN ELITNVELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR BUTNIELADAFEQMWQKR 300 310 NKVKKRTSNSNTYNTGSH KLVKRRTI.SSPYGSH ELVKKRTSNSNTYNTGSH ELVKKRTSNSNKLNEGSH SSVKARTSNYNTYAGSH	MYQKMVFYLEAC RYHEMLVIADSC RYNELFFMVDTC RYNELFFIDTC RYNELFFIDTC RYNELFFIDTC RYNELFFIDTC VNEYSSKDIKPE VMEYGSKDIKPE VMEYGDIGLSKE VMEYGDKTFKDE	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF 330 KVYLYLGFD LLYQYLGTN KLYLYQGFN KLYLFQGFN	LPPDINVYATT DSPN.VLSLS ISPN.VLAVA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA PAND.NNSF PANG.NFTF PANG.NTN	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSQVGEDSLS SQVGEDSLS SQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS VGEDSLS VGEDSLS VGEDSLS VGEDSLS VGEDSLS	CYYDE YDV HHV HQP HQP HQP HQP HQP HQP HQP BSQ SSQ DKLDGVNQRD RPPSRVTNQRD PVPKATVNQRD NAPMEVVNQRD	. QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAVGVH PAVGVH PAUGVH ADLIFL ADLIFF ADLVHF AELHFM	IGDWY3 VIDRYTHYF MIDRYTFY IMDRYTFY	VNWMEDSO VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI GSSQKNEA GSSQKNEA GSSKKEA	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.C.SQTNMNDL B.C.SQTNMNDL LREITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK LQQIKDAIKHRS
Q30H12 Q474E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49042 Q24326	. LHVKD LNY AI OTMFEDN ELITNVELLAY AI OTMFEDN EITS QELADGI QQMWEKK EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR BITNIELADAFEQMWQKR BITNIELADAFEQMWQKR ELIVKKRTSNSNTYNTGSH KLVKKRTSNSNTYNTGSH ELVKKRTSNSNTYNTGSH ELVKKRTSNANKLNEGSH SSVKARTSNANKLNEGSH QSVKQRTSNFAYAMGSH	MYQKMVFYLEAC RYHEMLVIADSC RYNELFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC YNEYSSKDIKPE VMEYGSKDIKPE VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF 330 KVYLYLG CASMYERF CONT	LPPDINUYATT DSPN. VLSUS ISPN. VLAVA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA PATV. NLPA PATV. NLPA PATV. NFPP PATV. NFPP PATV. NFPP	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS NQIHF VDETENSLKL VNENSI NAHL NAHL	CYYDE YDV HHV HQP HQP HQP HQP HQP HQP S50 S50 DKLDGVNQRD RTPSAAVNQRD PPPSRVTNQRD NAPMEVVNQRD NAPMEVVNQRD NAPMEVVNQRD	.QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIFL ADLIFF ADLLFM AELFM	LGD WY3 VID RYTYYA MID RYTFYY LMD RYTFYY LMD RYTFYY LMD RYTFYY LMD RYTFYY LMD RYTFYY LMD RYTFYY WQ RYFYY WM RYTFYY WM RYTFYY WM RYTFYY WM RYTFYY WH RYCKSSE W EKFRKAPE W RYLYCASE W QLYKSSE W QV YQRSNE	<u>VNM MEDSD</u> <u>VNFLTKEV</u> <u>LEFLEKVQ</u> <u>LEFLEEVQ</u> <u>LEFLEEIN</u> <u>LEFLEEIN</u> <u>LEFLEEIN</u> <u>STRPEI</u> <u>GSSQKNEA</u> <u>SSSKKEA</u> <u>IGSEKKEI</u>	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL LREITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK LQQIKDAIKHRS LKQIEEIVKHRK
Q95M12 Q45048 Q874E1 Q92CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49047 B8ASK4 P49042 Q24326 P49046	. LHVKD LNE TIRYMYEHK ELTNVDLAYAIOTMFEDN ELTNVELADAFEQMWQKR ETN IELADAFEQMWQKR EITN IELADAFEQMWQKR EITN IELADAFEQMWQKR EITN IELADAFEQMWQKR ETN IELADAFEQMWQKR BOO NKVKKRTSNSNTYNTGSH ELVKKRTSNSNTYNTGSH ELVKKRTSNSNTYNTGSH ELVKKRTSNANKLNEGSH SSVKARTSNYNTYAAGSH QSVKRTSNYNTYAAGSH QSVKRTSNSNYRFGSH	MYQKMVFYLEAC RYHEM/VIADSC RYNELFFMVDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC VNEYGSKDIKPE VMEYGSKDIKPE VMEYGDIGLSKN VMEYGDIGLSKAD VMEYGDNKTFKDE VMQYGDQSIKAD	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF QGASMYERF 330 KVYLYLCGFD DLYQYLGTN KLYLFMOGFN KLYLFQOGFD KLYLFQGFD KLYLFQGFD KLYLFQGFD	LPPDINVYATT DSPN.VLSLS ISPN.VLAVA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA PADE.NFTP PATV.NFPP PATV.NFPP PATV.NFPP	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS NQIHF VDETENSLKL VNENSI E.LIW HNGRL HNGRL	CYYDE YDV HUV HQP HQP HQP DHQP HQP BAGONORIA S5Q DKLDGVNQRD RTPSAAVNQRD RPPSRVTNQRD PVPKATVNQRD NAPMEVVNQRD KSKMEVVNQRD EAKMEVVNQRD	.QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAUGVH PAUGVH PAUGVH ADLIFF ADLIFM ADLLFM AELLFM	IGDWY3 VIDRYTHYI IDRYTHYI IDRYTFY IMDRYTFY IMDRYTFY IMDRYTFY IMDRYTFY IMDRYTFY IMDRYTFY IMDRYTFY IMDRYTFY IMDRYTFY WQRYKSSE WEKFRKAPP WKRYEQLNC WQUYQRSNH WQWYQRSNH	<u>NNMMEDSO</u> <u>VNFLTKEV</u> <u>LEFLEKVQ</u> <u>LEFLEEIN</u> <u>LEFLEEIN</u> <u>LEFLEEIN</u> <u>LEFLEEIN</u> <u>CSSQKNEA</u> <u>CSSAKVEA</u> <u>CSSAKVEA</u> <u>CSSKKKEI</u> <u>LPEKKTDI</u> <u>LPEKKTDI</u>	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL 380 39 LREITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK LREIEDTIAHRK LQQIKDAIKHRS LKQIEEIVKHRK
295M12 P49048 Q8T4E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49042 O24326 P49045 P49045	. LHVVD LAY AI OTMFEDN ELITNVD LAY AI OTMFEDN ELITNVELAD AF EQMWQKR EITN IELAD AF EQMWQKR EITN IELAD AF EQMWQKR EITN IELAD AF EQMWQKR EITN IELAD AF EQMWQKR SOQ 310 NKVKKRTSNSNTYNTGSH KLVKNRTI.SSYNTGSH ELVKKRTSSNSNTYNTGSH SVKARTSNYNTYAAGSH SVKARTSNYNTYAAGSH CSVKRTSNSNFAYAMGSH CSVKRTSNSNYNTYAAGSH	MYQKMVFYLEAC RYHEMLVIADSC RYNELFFMVDTC RYNELFFIDTC RYNELFIDTC RYNELFIDTC QYNELFIDTC YNELFIDTC YNELFIDTC YNEYGSKDIKPE VMEYGSKDIKFE VMEYGDKTFKDE VMQYGDKANMTAE VMQYGDTNITAE VMQYGDTNITAE	ESGSMMN.H SASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF 330 KVYLYLGFD LYQYLGTN KLYLYQGFD KLYLYQGFD KLYLYQGFD KLYLYQGFD	LPPDINVYATT DSPN.VLSLS ISPN.VLAVA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA PAND.NNSF PADE.NFT PAND.NFT PANG.NITN PASV.NFPP PATV.NFPP PATV.NFPP	AANPRESSYA SSLVGEDSLS SSLVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS NQIHF VDETENSLKL VNENSLIW NAHL HNGRL 	CYYDE YDVD HHVD HQPD HQPD HQPD HQPD HQPD HQPD S30 DKLDGVNQRD RPPSRVTNQRD PVPKATVNQRD NAPMEVVNQRD KSKMEVVNQRD EAKMEVVNQRD ETKMEVVNQRD	QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAUGVH PAUGVH ADLIFL ADLLFM ADLLFM AELLFM AELLFM	LGDWY3 VIDRYTHYA MIDRYTHYA MIDRYTFYU LMDRYTFYU LMDRYTFYU <td< th=""><th>VNWMEDSO VNFLTKEV LEFLEKVQ LEFLEKVQ LEFLEEIN LEFLEEIN LEFLEEIN LEFLEEIN SSSQKNEA GSSQKNEA GSSQKNEA JGSEKKKEI LPEKKTDI QPEKKTHI QSENKTDI</th><th>VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.EITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK LQCIKDAIKHRS LKQIEEIVKHRK LQQIKDAIKHRS LKQIEETVKHRK</th></td<>	VNWMEDSO VNFLTKEV LEFLEKVQ LEFLEKVQ LEFLEEIN LEFLEEIN LEFLEEIN LEFLEEIN SSSQKNEA GSSQKNEA GSSQKNEA JGSEKKKEI LPEKKTDI QPEKKTHI QSENKTDI	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.EITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK LQCIKDAIKHRS LKQIEEIVKHRK LQQIKDAIKHRS LKQIEETVKHRK
293M12 P49048 Q8T4E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49047 B8ASK4 P49042 O24326 P49046 P49045 O9R0J8	LHVYD LNY AI OTMFEDN ELITNYD LAY AI OTMFEDN ELITNYD LAY AI OTMFEDN ELITNYD LAY AI OTMFEDN ELITN IE LADAFE OMWOKR ELITN IE LADAFE OMWOKR ELITN IE LADAFE OMWOKR 300 300 310 NKVKKRTSNSNTYNTGSH KLVKKRTSNSNTYNTGSH KLVKKRTSNNAKLNEGSH SSVKARTSNYNTYAAGSH OSVKQRTSNFAYAMGSH OSVKRTSNSNSYFGSH KSVKQRTSNFAYAMGSH HLVKSHTN	MYQKMVFYLEAC RYHEMLVIADSC RYNELFFMVDIC RYNELLFIIDIC RYNELLFIIDIC RYNELLFIIDIC YNELLFIIDIC YNEYGSKDIKPE VMEYGDIGLSKNI VMEYGDIGLSKNI VMEYGDIGLSKNI VMEYGDIGLSKNI VMEYGDIGLSKSNI VMEYGDIGLSKSNI VMEYGDIGLSKSNI VMEYGDITLSE VMEYGDITLSE VMEYGDITNITAE VMQYGDNNITAE VMQYGDNSISTMI	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF QGASMYERF S30 KVYLYLG GFD LYQYLG KUYLYMG KUYLYNG KUYLYNG KUYLYHG KUYLYHG KUYLYHG KUYLYHG KUYLYHG CFD KUYLYHG CFD CFD CFD CFD CFD CFD CFD CFD	LPPDINUYATT DSPN. VLSLS ISPN. VLAVA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA PAND. NLPA PAND. NSF PAND. NFPP PAND. NFPP PATV. NFPP PATV. NFPP PATV. NFPP PATV. NFPP PATV. NFPP PATV. NFPP	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS 	CYYDE YDV HUV HUV HQP HQP HQP HQP HQP HQP B JSQ SSQ NAPMEVNQRD NAPMEVVNQRD KSKMEVVNQRD E TKMEVVNQRD LDITPSPD	. QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAIGVH ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI AELIFM	LGD WY3 VID RYTYYA NID RYTFYY LMD RYTFYY LMD RYTFYY LMD RYTFYY LMD RYTFYY LMD RYTFYY LMD RYTFYY WQ RYKKSSE WQRYKKSSE W QLYKRSSE W QUYQRSNF W QUYQRSNF W QUYQRSNF	VNWMEDSU VNFLTKEV LEFLEKVQ LEFLEKVQ LEFLEEIN LEFLEEIN LEFLEEIN CSSQKNEA CSSQKNEA CSSQKNEA CSSQKNEA CSSQKKEA USEDKLEA USEDKLEA USEDKLEA USEDKLEA USEDKLEA USENKTDI USENKTDI	VEDLTKETLHKQ KALNSSANMQOY PFSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL LREITETLTHRG EKQVLEAMSHRK LREIEDTIAHRK LQQIKDAIKHRS LQQIKDAIKHRS LQQIKDAIKHRS LKQIEEIVKHRK LEQITETVKHRK VGQIQHLLDARH
Q926H12 Q874E1 Q92CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49042 Q24326 P49045 Q9R0J8 O89017 O89522	LHVKDLNE TIRYMYEHK ELTNVDLAYAIOTMFEDN ELTNVELAYAIOTMFEDN EITSQELADGIQQWWEKK ETNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR SVKRTSNSNTYNTGSH KLVKKRTAGSG, KSYGSH ELVKKRTAGSG, KSYGSH ELVKKRTSNSNYNTYAGSH ELVKKRTSNSNSYRFGSH KSVKQRTSNFAYAMGSH HLVKSHTNTSH	MYQKMVFYLEAC RYHEMLVIADSC RYNELFFMVDTC RYNELFFIDTC RYNELLFIDTC RYNELLFIDTC RYNELLFIDTC VMEYGSKDIKPE VMEYGSKDIKPE VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDKTFKDE VMQYGDXNMTAE VMQYGDTNITAE VMQYGDKSISTM VMQYGNKSISTM	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF S30 KVYLYLGFD KUYLFQGFD KLYLFQGFD KLYLFQGFD KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYHG KLYLYH KLYLYHG KLYLYH KLYLYHG KLYLYH KLY	LPPDINVYATT DSPN.VLSLS ISPN.VLSV YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA	AANPRESSYA SSLTHESYS SSLVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS NQIHF VDETENSLKL VNENSI ELIW HNGRL HNGRL QNGRL	CYYDE YDV HUV HQP HQP HQP HQP HQP HQP HQP B JSQ JKLDGVNQRD RTPSAAVNQRD PPSRVTNQRD NAPMEVVNQRD KSKMEVVNQRD EAKMEVVNQRD LDLTPSPD LDLTPSPD LDLTPSPD	QRSTF TDIGVY PAIGVY PAIGVH PAIGVH PAIGVH PAIGVH ADLIFL ADLIFL ADLIFL ADLIFL ADLIFL ADLIFL VFLTIL VFLTIL	LGDWY3 VIDRYTHY1 WIDRYTFYV MIDRYTFYV LMDRYTFYV WQRYKSSE WQLYKRSEN WQVYQRSNH WQMYQRSNH WQMYQRSNH WQNYQRSNH WQNY WQNY	VNW MEDSO VNFLTKEV LEFLEKVQ LEFLEKVQ LEFLEEVQ LEFLEEIN LEFLEEIN LEFLEEIN CSSQKNEA CSSQKNEA CSSAKVEA VSEDKLRA GSEKKKEI LOPEKKTHI QPEKKTHI QPEKKTHI VK.ESQVL	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMDL B.A
295M12 P49048 Q8T4E1 Q92XY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49047 B8ASK4 P49046 P49046 P49045 Q9R0J8 Q889017 Q99538 O585D9	. LHVYD LAY AI OTMFEDN ELITNYD LAY AI OTMFEDN E LITNYD LAY AI OTMFEDN E ITN ELADAFE QMWQKR EITN IELADAFE QMWQKR EITN IELADAFE QMWQKR EITN IELADAFE QMWQKR EITN IELADAFE QMWQKR SVKARTSNSNTYNTGSH KLVKNRTSNSNTYNTGSH KLVKNRTSNSNTYNTGSH SVKARTSNSNTYNTGSH SVKARTSNNNTYAGSH QSVKARTSNNNTYAGSH QSVKARTSNNNTYAGSH CSVKARTSNNNTYAGSH HLVKSHTNTSH HLVKSHTNTSH	MYQKMVFYLEAC RYHEMLVIADSC RYNELFFMVDTC RYNELFFIDTC RYNELFIDTC RYNELFIDTC RYNELFIDTC VNELFIDTC VNEYGSKDIKFE VMEYGSKDIKFE VMEYGDKTFKDE VMQYGDKTFKDE VMQYGDTNITAE VMQYGDKSISTM VMQYGNKSISTM VMQYGNKSISTM	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF QGASMYERF 330 KVYLIYLGFD LYQYLGFD KLYLYGGFD KLYLYHGFO KLYLYHGFD KLYLYHGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYMOFO GMKK KVMOFO GMK	LPPDINVYATT DSPN.VLSLS ISPN.VLSV YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA PATV.NFP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS NQIHF VDETENSLKL VNENSLIW ELIW HNGNL QNGRL VTHVTH	CYYDE YDV HUV HQP HQP HQP HQP HQP HQP HQP HQP BASO SSQ DKLDGVNQRD RPPSRVTNQRD PVPKATVNQRD KSKMEVVNQRD EAKMEVVNQRD LDLTPSPD LDLTPSPD LDLTPSPD	QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIFL ADLIFL ADLIFL ADLIFH ADLLFM AELFFM AELFFU VPLTIL VPLTIV	LGDWY3 VIDRYTHY7 NIDRYTFY LMDRYTFY WQRYKKSSE WEKFRKAPE WHKYQKAPE WQLYKRSEN WQMYQRSNF WQMYQRSNF KRKLLRTNI KRKLLRTNI KRKLLRTNI KRKLLNTNI KRKLLNTNI	VNWMEDSO VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN SSSQKNEA GSSQKNEA GSSQKNEA GSSQKNEA IGSEKKKEI LPEKKTDI MK.ESQVL VK.ESQNL LLE SSQL	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.EITETLTHRG EKQVLEAMSHRK LKQIETUKHRK LQQUKDAIKHRS LKQIEEIVKHRK LQQIKDAIKHRS LKQIEEIVKHRK LQQIKDAIKHRS LKQIAETVKHRK UGQIQHLDARH TEEIQRHLDARH
293M12 P49048 Q8T4E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49042 O24326 P49046 P49045 O9R0J8 O89017 Q99538 Q585D9 O95M12	LHVVDLAYAIOTMFEDN ELITNVDLAYAIOTMFEDN ELITNVELAYAIOTMFEDN ELITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR ELITNIELADAFEQMWQKR SOQ SOV SOUCHARTSNSNTYNTGSH KLVKNRTSNSNTYNTGSH KLVKNRTSNSNTYNTGSH SSVKARTSNSNSYRFGSH QSVKQRTSNFNYAAGSH QSVKQRTSNFNYAAGSH USVKQRTSNFNYAAGSH HLVKSHTNTSH HLVKSHTNTSH HLVKSHTNTSH	MYQKMVFYLEAG RYHEMLVIADSC RYHELFFMVDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC VMEYGSKDIKPE VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDITITAE VMQYGDKTNITAE VMQYGNKSISTM VMQYGNKTISTM VMQYGNKTISTM	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF CONT CO	LPPDINMYATT DSPN. VLSLS ISPN. VLAVA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA PAUD. NSP PAUD. NFPP PAUD. NFPP PATV. NFPP	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS 	CYYDE YDV HUV HUV HQP HQP HQP HQP HQP HQP HQP B YDV HQP B YDV YDV YDV YDV NAPMEVVNORD NAPMEVVNORD XSKMEVVNORD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD	.QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAIGVH ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI VHF ALLFM AELIFM VPLTIL VPLTIM VPLTIM	LGD WY3 VID RYTYYA VID RYTYYA LMD RYTYYA LMD RYTFYV LMD RYTFYV LMD RYTFYV LMD RYTFYV LMD RYTFYV LMD RYTFYV WQ RYKKSSE WQRYKKSSE WQVYQRSNF WQVYQRSNF KRKLLRTNI KRKLLRTNI KRKLLMNTNI KRKLMNTNI KRKLMNTNI KRKLMNTNI KRKLMNTNI	VNWMEDSO VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI GSSQKNEA GSSQKNEA GSSQKNEA GSSKKEI LLPEKKTDI 4QPEKKTDI 4QPEKKTDI MK.ESQVL VK.ESQNL VLE.ESQL LLE.ESRQL	VEDLTKETLHKQ KALNSSANMQOY PFSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL LREITETLTHRG EKQVLEAMSHRK LREITETLKHRS LKQIELAMSHRK LQCIKADIKHRS LKQIEEIVKHRK LEQITETVKHRK VGQUQHLDARH TEEIQRHLDARH TEEIQHLDARH
Q92643 Q9748 Q92643 Q92643 Q92643 Q92643 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49047 B8ASK4 P49042 Q24326 P49046 Q9R0J8 089017 Q99538 Q9SR5D9 Q9555D9 Q955512 P49048	. LHVKD LNE TIRYMYEHK ELTNVD LAYAI OTMFEDN ELTNVELAYAI OTMFEDN EITS QELADGI QQMWEKK ETTN IELADAFEQMWQKR EITN IELADAFEQMWQKR EITN IELADAFEQMWQKR EITN IELADAFEQMWQKR EITN IELADAFEQMWQKR BITN IELADAFEQMWQKR SVKRTSNSNTYNTGSH ELVKKRTSNSNTYNTGSH ELVKKRTSNSNSYRGSH KSVKRTSNSNSYRGSH KSVKRTSNSNSYRGSH KSVKRTSNSNSYRGSH KSVKRTSNSNSYRGSH HLVKSHTNTSH HLVKSHTNTSH HLVKSHTNTSH HLVKSHTNTSH	MYQKMVFYLEAC RYHEMLVIADSC RYNELFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC VMEYGSKDIKPE VMEYGSKDIKPE VMEYGDIGLSKKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDANMTAE VMQYGDANMTAE VMQYGDANMTAE VMQYGDKSISTM VMQYGNKSISTM VMQYGNKTISTM VMQYGNKTISTM	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF	LPPDINMYATT DSPN. VLSLS ISPN. VLAVA YSPN. IMALA YSPN.	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS NQIHF VDETENSLKL VNENSI NAHL NAHL NAHL NAHL NAHL VTH	CYYDE YDV YDV HHV HQP HQP HQP HQP HQP HQP HQP BYDV ASP Stop BYDY Stop NAPMEVVNQRD KSKMEVVNQRD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD	QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIFL ADLIFL ADLIFL ADLIFL ADLLFM AELLFM AELLFM AELLFL VPLTIL VPLTIL VPLTIM VPLSIM	LGDWY3 VIDRYTYYA VIDRYTYYA MIDRYTYYA MIDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV LMDRYTFYV WQRYKSSE WQRYKSSE WQVYQRSNH WQNYQRSNH WQRNYQRSNH	VNNWMEDSO VNFLTKEV LEFLEKVQ LEFLEKVQ LEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN STRPEI CSSQKNEA CGSARKVEA VSEDKLRA GSEKKKEI LPEKKTDI LQPEKKTHI QQEKKTDI QDEKKTDI UPEKKTDI DLE.ESQLL LE.ESRQL	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL J.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C
Q95M12 Q874E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49047 B8ASK4 P49042 Q24326 P49046 P49045 Q9R0J8 Q9SM12 P49048 Q9SM12 P49048	. LHVKDLNE TIRYMYEHK ELTNVDLAYAIOTMFEDN ELTNVELAYAIOTMFEDN EITSQELADGIQQMWEKK EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR SVKARTSNSNTYNTGSH KLVKRTSNSNTYNTGSH KLVKRTSNSNTYNTGSH SVKARTSNSNTYNTGSH SVKARTSNNYNTYAGSH QSVKARTSNNYNTYAGSH QSVKARTSNNYNTYAGSH QSVKARTSNNYNTYAGSH CSVKARTSNNYNTYAGSH USVKARTSNSNYRFGSH HLVKSHTNTSH HLVKSHTNTSH HLVKSHTNTSH HLVKSHTNTSH HLVKSHTNTSH DSCPARKCLSNTGV	MYQKMVFYLEAC RYHEMLVIADSC RYNELFFMVDTC RYNELFFIDTC RYNELFIDTC RYNELFIDTC S20 VMEYGSKDIKPEI VMEYGDKIFKDEI VMEYGDKIFKDEI VMQYGDANMTAEI VMQYGDINITAEI VMQYGDINITAEI VMQYGDKSISTMI VMQYGNKSISTMI VMQYGNKSISTMI VMQYGNKSISTMI VMQYGNKSISTMI VMQYGNKSISTMI VMQYGNKSISTMI	ESGSMMN.H RSASMYEWI QAASLYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF 330 KVYLIYLGFD LYQYLGFD KLYLYGGFD KLYLYHGFD KLYLYHGFD KLYLYHGFD KLYLYQGFM KLYLYQGFD KLYLYQGFD KLYLYQGFD KLYLYQG CMK KLYLYQG CMK KLYLYQG CMK KLYLYQG CMK KLYLYQG CMK KLYLYQG CMK KLYLY CGMK KLYLYQ CGMK KLYLYQ CGMK KLYL CGMK KLYDFF CGMK CMC CMK CMC CMK CMC CMK CMC CMK CMC CMK CMC CMK CMC CMK CMC CMK CMC CMK CMC CMK CMC CMK CMC CMK CMC CMK CMC CMK CMC CMK CMC CMK CMC CMC	LPPDINVYATT DSPN.VLSLS ISPN.VLSV YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA PADU.NFP PATV.NFPP PATV.NFPP PATV.NFPP PATV.NFPP HRASSPISLPP RKASSPVSLPP HQASSPISLPA RIFQ.	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS NQIHF VDETENSLKL VNENSLIW HNGRL HNGRL QNGRL VTHVTH VTHVTH VTHVTH VTHVTH VTHVTH VTHVTH VTHVTH VTHVTH VTHVTH VTHVTH VTHVTH VTHVTH	CYYDE YDVD HHVD HQPD BYDKLDGVNQRD RPPSRVTNQRD NAPMEVVNQRD KSKMEVVNQRD EXKMEVVNQRD LDLTPSPD	QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIFI ADLIFI ADLIFI ADLIFI ADLLFM AELHFM AELHFM AELLFM VPLTIL VPLTIV VPLTIM VPLSIM FWA	LGDWY3 VIDRYTHY7 NIDRYTY7 LMDRYTFY0 WQRYKKSSE WQRYKKSSE WCLYKRSED WQMYQRSNF WQMYQRSNF KRKLLRTNI KRKLLRTNI KRKLLRTNI KRKLMNTNI KRKLMSTNI KRKLMSTNI KRKLMSTNI	VNNWMEDSO VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN SGSSQKNEA GSSQKNEA GSSQKNEA GSSQKNEA GSSKKKEI LOPEKKTDI UQPEKKTDI VK.ESQVL VK.ESQVL LLE.ESRQL LLE.ESRQL LQ.ESRL	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL REITETLTHRG EKQVLEAMSHRK LREITETLTHRG LKQIEEIVKHRK LQQIKDAIKHRS LKQIEEIVKHRK LQQIKDAIKHRS LKQIEEIVKHRK UGQIQHLDARH TEEIQQHLDARH TEEIQHLDARH TEEIQHLDARH VQKIDRHLEARN
295M12 P49048 Q8T4E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49047 B8ASK4 P49047 B8ASK4 P49042 O24326 P49046 P49045 O9S0J8 O89017 Q99538 Q5R5D9 Q95M12 P49048 Q8T4E1 Q9CXY9	LHVKDLNETIRYMYEHK ELTNVDLAYAIOTMFEDN ELTSOLLAYAIOTMFEDN EITSOLLAYAIOTMFEDN EITSOLLADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR SUTNIELADAFEQMWQKR UNKRTSNSNTYNTGSH KLVKNRTI.S.S.EPYGSH ELVKKRTSNNAKLNEGSH QSVKQRTSNFAYAMGSH QSVKQRTSNFAYAMGSH QSVKQRTSNFAYAMGSH USVKQRTSNFAYAMGSH HLVKSHTN.TSH HLVKSHTN.TSH HLVKSHTN.TSH HLVKSHTN.TSH USVKQRTSNFNYAGSH QSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH USVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH USVKQRTSNFNYAGSH HLVKSHTN.TSH HLVKSHTN.TSH USVKQRTSNFNYAGSH USVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH CSVKQRTSNFNYAGSH	MYQKMVFYLEAG RYHEMLVIADSC RYNELFFMVDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC YNEYGSKDIKPE VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDITITAE VMQYGNKSISTM VMQYGNKSISTM VMQYGNKTISTM XMQYGNKTISTM RKDHYPKDVKRV RKDLYPRDPHKV RKDLYPRDPHKV	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF CONT CO	LPPDINMYATT DSPN. VLSLS ISPN. VLAVA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA PADU. NLPA PAND. NSF PAND. NSF PANG. NITN PANG. NFPP PATV. SPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPA RIFQ	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS 	CYYDE YDV HUV HQP HQP HQP HQP HQP HQP B JSQ	.QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIF ADLIF ADLIF ADLIF ADLIF ADLIF ADLIF VPITIL VPLTIL VPLTIL VPLTIL VPLTIL VPLTIM FWA NEDDFIS	LGD WY3 VID RYTFYV VID RYTFYV MID RYTFYV LMD RYTFYV WQ RYKKSSE WQ RYKKSSE WQLYKRSEN WQVYQRSNH KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLMNTNE KRKLMNTNE KRKLMNTNE SK	VNM MEDSD VNFLTKEV UEFLEKVQ LEFLEKVQ LEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI GSSGKKEA GSSKKKEI LEPEKKTDI QPEKKTDI QPEKKTDI USSENKTDI USSENKTDI USSENKTDI USSENKTDI USSENKTDI USSENKTDI USSENKTDI USSENKTDI MK.ESQVL VK.ESQNL LE.ESRQL LE.ESRQL LLE.ESRQL LL.E.SRQL DLQ.ESRT	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL J.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C
295M12 P49048 Q8T4E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49047 B8ASK4 P49042 Q24326 P49046 Q9R0J8 089017 Q99538 Q9SN12 Q9SN12 P49048 Q8T4E1 Q9CXY9 Q92643	LHVKDLNETIRYMYEHK ELTIN VELAYAIOTMFEDN ELTIN ELADAFEQMWQKR ETIN IELADAFEQMWQKR ETIN IELADAFEQMWQKR EITIN IELADAFEQMWQKR EITIN IELADAFEQMWQKR EITIN IELADAFEQMWQKR EITIN IELADAFEQMWQKR ELTIN IELADAFEQMWQKR ELTIN IELADAFEQMWQKR SVKRTSNSNTYNTGSH ELVKKRTSNSNTYNTGSH ELVKKRTSNSNSYRGSH KSVKQRTSNFAXAMGSH HLVKSHTN	MYQKMVFYLEAC RYHEM/VIADSC RYNELFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC VMEYGSKDIKPEI VMEYGSKDIKPEI VMEYGDIGLSKKNI VMEYGDIGLSKNI VMEYGDIGLSKNI VMEYGDANMTAE VMQYGDANMTAE VMQYGDANMTAE VMQYGDKSISTMI VMQYGNKSI VMQYGNKSI VMX VMX VMX VMX VMX VMX VMX VMX	ESGSMMN.H ESGSMMN.H ESASMYEWI QAASLYEWI QAASLYEWF QGASMYERF	LPPDINVYATT DSPN.VLSLS ISPN.VLSV YSPN.IMALA	AANPRESSYA SSLTHESYS SSLVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS SSQVGEDSLS NQIHF VDETENSLKL VNENSI ELIW NAHL HNGRL QNGRL VTHVTH VTHVTH VTHVTH VTHVTH VTHVTH VTHVTH VTHVTH VTHVTH VTHVTH ILLEEIVL 	CYYDE YDV HUV HQP HQP HQP HQP HQP HQP HQP BYDV ASP JSQ DKLDGVNQRD PPSRVNQRD NAPMEVVNQRD KSKMEVVNQRD LDLTPSPD	QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIFL ADLIFL ADLIFL ADLUFM AELLFM AELLFM AELLFM AELLFL VPLTIL VPLTIM VPLTIM VPLTIM VPLSIM NEDDFI SQVVDS	LGDWY3 VIDRYTFYV VIDRYTFYV MIDRYTFYV LMDRYTFYV WQRYKSSE WQLYKRSEN WQVYQRSNH WQNYQRSNH SSK SSK	NMMMEDSD VNFLTKEV VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI CGSARKVEA VSEDKLRA GSEKKKEI IQPEKKTDI IQPEKKTDI QESNKTDI DLE.ESRQL DLE.ESRQL LQ.ESRL NWTKKP TAEERMEP	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMDL B
295M12 P49048 Q8T4E1 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49047 B8ASK4 P49042 O24326 P49046 P49046 P49046 P49045 Q9SMJ2 Q95M12 P49048 Q85R5D9 Q95M12 P49048 Q874E1 Q9CXY9 Q92643 Q55R6L8	. LHVKD LNE TIRYMYEHK ELTNVDLAYAIOTMFEDN ELTSOELADGIOOMWEKK EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR UNKKRTSNSNTYNTGSH KLVKNRTI.S.SPYGSH ELVKKRTSNSNTYNTGSH CVKRTSNSNSYRGSH CVKRTSNSNSYRGSH QSVRKRTSNSNSYRGSH QSVRKRTSNSNSYRGSH QSVRKRTSNSNSYRGSH UVKSHTNTSH HLVKSHTNTSH HLVKSHTNTSH HLVKSHTNTSH UVKSHTNTSH UVKSHTNTSH UVKSHTNTSH UVKSHTNTSH UVKSHTNTSH UVKSHTNTSH UVCSKLVSTPGH UVCPKSLCVSTPGH	MYQKMVFYLLAG RYHEM/VIADSC RYHELFFMVDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC YMEYGSKDIKPE VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKE VMEYGDIGLSKE VMEYGDINITAE VMQYGDINITAE VMQYGDINITAE VMQYGDINITAE VMQYGDINITAE VMQYGNKSISTM VMQYGNK ND ND ND ND ND ND ND ND ND ND	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF QGASMYERF 330 KVYLYLCGFD KLYLYLCGFD KLYLYLGGFD KLYLYLGGFD KLYLYLGGFD KLYLYLGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGG CMK KVMOFOGMK KVMOFOGMK KVMOFOGMK KVMOFOGMK KVMOFOGMK KLYLYFG CMK CF CG CMK CF CG CG CG CG CG CG CG CG CG CG	LPPDINVYATT DSPN.VLSLS ISPN.VLSV YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA YSPN.IMALA PADU.NFP PATV.NFPP PATV.NFPP PATV.NFPP HRASSPISLPP RKASSPISLPP HRASSPISLPP HRASSPISLPP HRASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASPISLPP RKASPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASSPISLPP RKASPISLPP	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS NQIHF VDETENSLKL VNENSLIW HNGRL HNGRL QNGRL VTHVTHVTH VTHVTHVTH VTHVTHVTH VSR VTRINV VSR	CYYDE YDVD HUPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD BYDV HQPD BYDY BYDY SKEVNQRD RDPSRVTNQRD VSKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD LDLTPSPD LDUTPSPD LDUTPSPD <t< th=""><th>QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAIGVH ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI ADLLFM AELFM AELFM VPLTIL VPLTIL VPLTIM VPLSIM FWA</th><th>LGDWY3 VIDRYTHY7 VIDRYTHY7 MIDRYTFYV MDRYTFYV LMDRYTFYV WQRYKSSE WQRYKSSE WQLYKRSEN WQMYQRSNF WQMYQRSNF KRKLLRTNI KRKLLRTNI KRKLLRTNI KRKLMNTNI KRKLMSTNI SSK DC SYK EDC</th><th>VNNWMEDSO VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI SS.S.TRPEI SS.S.TRPEI SS.S.KTRE SS.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S</th><th>VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL D.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNML B.ASQTNML B.AS</th></t<>	QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAIGVH ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI ADLLFM AELFM AELFM VPLTIL VPLTIL VPLTIM VPLSIM FWA	LGDWY3 VIDRYTHY7 VIDRYTHY7 MIDRYTFYV MDRYTFYV LMDRYTFYV WQRYKSSE WQRYKSSE WQLYKRSEN WQMYQRSNF WQMYQRSNF KRKLLRTNI KRKLLRTNI KRKLLRTNI KRKLMNTNI KRKLMSTNI SSK DC SYK EDC	VNNWMEDSO VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI SS.S.TRPEI SS.S.TRPEI SS.S.KTRE SS.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S.S	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL D.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNML B.ASQTNML B.AS
Q30H12 Q474E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49047 B8ASK4 P49042 Q24326 P49046 P49045 Q980J8 Q980J8 Q99538 Q5R0J8 Q955M12 P49048 Q874E1 Q92C43 Q92643 Q586L8 Q3MH27	L HVKD LNE TIRYMYEHK ELTINVELAYAIOTMFEDN ELTIN GLAADAFEQMWQKR EITIN IE LADAFEQMWQKR EITIN IE LADAFEQMWQKR EITIN IE LADAFEQMWQKR EITIN IE LADAFEQMWQKR EITIN IE LADAFEQMWQKR SURTIADAFEQMWQKR ELTIN IE LADAFEQMWQKR ELTIN IE LADAFEQMWQKR SUKRTSNSNTYNTGSH KLVKNRTI.S.S.EPYGSH ELVKRTAGSG.KSYGSH GSVKQRTSNFAYAMGSH QSVKQRTSNFAYAMGSH QSVKQRTSNFAYAMGSH QSVKQRTSNFNYAGSH KSVKQRTSNFNYAGSH HLVKSHTN.TSH HLVKSHTN.TSH HLVKSHTN.TSH UVKSHTN.TSH QLVKSHTN.TSH QVCPKSLCVSTP.GH	MYQKMVFYLEAG RYHEMLVIADSC RYHELFFMVDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC VMEYGSKDIKSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDINITAE VMQYGNKSISTM VMQYGNKSISTM VMQYGNKSISTM RKDHYPKDVKRV RKDLYPRDPHKV RTDLFQRDPKNV RTDLFQRDPKNV RTDLFQRDPKNV	ESGSMMN.H ESGSMMN.H ESAS MYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF QGASMYERF CONT C	LPPDIN/YATT DSPN.VLSLS DSPN.VLAVA YSPN.IMALA Y	AANPRESSYA SSLTHEESYS SSLVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS NQIHF VDETENSLKL 	CYYDE YDVD HHVD HQPD NAPMEVVNQRD NAPMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD VOQD QQD QQD QQD	.QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIF ADLI	LGD WY3 VID RYTFYV VID RYTFYV MID RYTFYV LMD RYTFYV LMD RYTFYV LMD RYTFYV LMD RYTFYV LMD RYTFYV LMD RYTFYV WQ RYKKSSE WQRYKKSSE WQLYKRSEN WQVYQRSNF KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLMNTNI KRKLMNTNI KRKLMNTNI SYK SYK	VNM MEDSD VNFLTKEV VLEFLEKVQ LEFLEKVQ LEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN STRPEI GSSQKNEA GSEKKKEI LOSENKKEN UQPEKKTHI UQSEKKTHI UQSEKKTHI VK.ESQVL VK.ESQVL VK.ESQVL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLQ.ESRRI DLQ ESRL DLQ ESRL DLQ ESRL DLQ ESRL DLQ ESRL DLQ ESRL DLQ ESRL DLQ ESRL DLQ ESRL	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL J.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C
Q92643 Q874E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49047 B8ASK4 P49042 Q24326 P49045 Q9R0J8 Q9R0J8 Q98017 Q99538 Q5R5D9 Q95M12 P49048 Q874E1 Q99CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1	L HVKD LNE TIRYMYEHK E LTNVD LAYAI OTMFEDN E LTNVD LAYAI OTMFEDN E ITS QELADGI QQMWEKK E ITNIE LADAFEQMWQKR E ITNIE LADAFEQMWQKR E ITNIE LADAFEQMWQKR E ITNIELADAFEQMWQKR E ITNIELADAFEQMWQKR E ITNIELADAFEQMWQKR E LTNIELADAFEQMWQKR E LTNKRTSNSNTYNTGSH KLVKRRTSNSNTYNTGSH E LVKRRTSNSNSYRFGSH KSVKQRTSNFNNYAMGSH HLVKSHTN	MYQKMVFYLEAC RYHEMLVIADSC RYHELFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC WELGSKDIKPE VMEYGSKDIKPE VMEYGDIGLSKKI VMEYGDIGLSKNI VMEYGDANMTAE VMQYGDANMTAE VMQYGDANMTAE VMQYGDANMTAE VMQYGDANMTAE VMQYGDANMTAE VMQYGDANTIAE VMQYGNKSISTMI MI MI MI MI MI MI MI MI MI	ESGSMMN.H ESGSMMN.H ESASMYEWI QAASLYEWI QAASLYEWF QAASMYERF QGASMYERF	LPPDIN/YATT LPPDIN/YATT DSPN.VLSLS ISPN.VLSV ISPN.VLSV YSPN.IMALA YSPN.INALA YS	AANPRESSYA SSLTHEESYS SSLTHEESYS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS NQIHF VDETENSLKL VNENSI ELIW HNGRL HNGRL VTHVNAHL VTHVNAHL VTHVNAHL VTHVNB VTHVI HLSEEIVL VTHVTHVI ITTETISL ITTETISL 	CYYDE YDV YDV HUP HQP HQP HQP HQP HQP HQP BYDV HQP BYDY SQ DKLDGVNQRD RTPSAAVNQRD PVPKATVNQRD KSKMEVVNQRD ETKMEVVNQRD LDLTPSPD VQD VQD VQD VQD VQD VQD VQD VQD	QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIF ADLIF ADLIF ADLIF ADLIF ADLIF ADLIF ADLIF MAELLFM AELLFM AELLFM VPLTIL VPLTIM VPLTIM VPLTIM VPLTIM VPLTIM SQVVDS SSEIMES SSEIMES SSGVVSS	IGDWY3 VIDRYTFYV VIDRYTFYV IDRYTFYV IDRYTFYV IMDRYTFYV WQRYKSSE WQRYQRSNH WQWYQRSNH WRKLLIRTNNI KRKLLRTNNI KRKLLRTNNI KRKLLNTNI KRKLNNTNI KRKLNNTNI KRKLNNTNI KRKLNNTNI SSK EDC SYK EDC SYK EDC SYK EDC SYK KRA	VNN W MEDSD VNFLTKEV LEFLEKVQ LEFLEKVQ LEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN QPEKKTDI QPEKKTDI QPEKKTDI QPEKKTDI QPEKKTDI QDELESRQL LLEESRQL LLEESRQL LLEESRQL LLEESRQL LLOESRL VK. ESQNL VK. ESQNL VK. ESQNL VK. ESQNL VK. ESQNL LEESRQL LLOESRL VK. TKKP TAEERMEP MDELMEP	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTTGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMDL B.
Q92643 Q874E1 Q92C43 Q92C43 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49047 B8ASK4 P49042 Q24326 P49046 P49046 P49046 P49046 Q95N38 Q95N38 Q95M12 P49048 Q874E1 Q92643 Q92643 Q576L8 Q3MH27 Q4KRV1	I. HYKD LNE TIRYMYEHK ELTNYD LAYAI OTMFEDN ELTS QELADGI QOMWEKK ETNIELADAFEQMWQKR ETNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR ELTVKRTSNSNTYNTGSH KLVKRTSNSNTYNTGSH ELVKKRTSSSKSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS	MYQKMVFYLEAC RYHEMLVIADSC RYHELFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC WEYGSKDIKPE VMEYGSKDIKPE VMEYGDIGLSKK VMEYGDIGLSKK VMEYGDKTFKDE VMEYGDANMTAE VMQYGNKSISTM VMQYGNK NTDF VMQYGNK NTDF VMQYGNK NTDF VMQYGN NTDF VMQYGN NTDF VMQYGN NTDF VMQYGN NTDF VMQYGN NTDF VMQYGN NTDF VMQYGN NTDF VMQYGN NTDF VMQYGN NTDF VMQYGN NTDF VMQ VMQYGN NTDF VMQ VMQ VMQ VMQ VMQ VMQ VMQ VM VMQ VMQ	ESGSMMN.H RSASMYEWI QAASLYEWI QAASLYEKF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF S30 KVYLYLG GFD DLYQYLGTNN KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD CMKK KVMOFOGMKK KVMOFOGMKK KVMOFOGMKK KVMOFOGMKK KVMOFOGMKK KVMOFOGMKK KVMOFOGSS LITDFFGSV LITDFFGSV LITDFFGSV LITDFFGSV	LPPDINVYATT DSPN.VLSLS ISPN.VLSV YSPN.IMALA	AANPRESSYS SSLTHESYS SSLVGEDSLS SSVGEDSSS SSVGEDSSS SSVGEDSSS SSVGEDSSS SSVG	CYYDE YDV HQV HQP HQP HQP DKLDGVNQRD RTPSAAVNQRD RTPSAAVNQRD PPSRVTNQRD VAAVNQRD NAPMEVVNQRD KSKMEVVNQRD LDLTPSPD	.QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAVGVH PAVGVH PAVGVH ADLIFF ADLUFF ADLUFF ADLUFF ADLUFF ADLUFF ADLUFF VPLTIL VPLTIL VPLTIL VPLTIL VPLTIL VPLTIL SQVVPS SEIMES SGTMKS 460	IGDWY3 VIDRYTHYT VIDRYTHYT IDRYTFYV IDRYTFYV IMDRYTFYV WQRYKSSE WQRYKSSE WQWYQRSNH WQWYQRSNH WQWYQRSNH WQWYQRSNH WQWYQRSNH SSKKLLRTNI KRKLLRTNI KRKLLRTNI KRKLLRTNI KRKLMNTNI KRKLMNTNI SYKEDC SYKEDC SYKKAC	VNN WMEDSU VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEIV VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VSEDKKTDI UQPEKKTDI UQPEKKTDI UQPEKKTDI UQPEKKTDI UQPEKKTDI ULE ESRQL LLE ESRQL DLQ ESRRL DLQ ESRRL DLQ ESRL DLQ E	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL 380 39 LREITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK LQUIKDAIKHRS LKQIEETVKHRK LQQIKDAIKHRS LKQIEETVKHRK LQQIEDAIKHRS LKQIEETVKHRK LQQIEDALDARH TEEIQRHLDARH TEEIQRHLDARH TEEIQRHLDARH TEEIQRHLDARH TEEIQRHLARN FKI LKY LKY LKY LKY LKY LKY LKY LKY
Q30H12 Q474E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49047 B8ASK4 P49042 Q24326 P49046 P49046 Q9P538 Q3F0J8 Q95S12 P49048 Q85K12 Q95X12 Q95X12 Q95X12 P49048 Q85K12 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1	LINVED LAY AI OTMFEDN ELITNVELAY AI OTMFEDN ELITNVELAY AI OTMFEDN ELITN ELADAFEQMWQKR EITN IE LADAFEQMWQKR EITN IE LADAFEQMWQKR EITN IE LADAFEQMWQKR EITN IE LADAFEQMWQKR EITN IE LADAFEQMWQKR ELITN IE LADAFEQMWQKR SUVKRTSNSNTYNTGSH KLVKNRTSNSNTYNTGSH KLVKNRTSNSNTYNTGSH KLVKRTSNSNTYNTGSH SVKARTSNSNTYNTGSH SVKARTSNSNTYNTGSH QSVKQTSNFEAYAMGSH QSVKQTSNFEAYAMGSH QSVKQTSNFEAYAMGSH QSVKQTSNFFAYAMGSH HLVKSHTN	MYQKMVFYLLAG RYHEM/VIADSC RYHELFFMVDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC WYNEYGSKDIKPE VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDKTTAE VMEYGDKTISTM VMQYGNKSISTM VMQYGNKSISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM RKDLYPRDPHKV RTDLFQRDPKNV RTDLFQRDPKNV RTDLFQRDPKNV RTDLFQRDPKNV RTDLFQRDPKNV RTDLFQRDPKNV RTDLFQRDPKNV	ESGSMMN.H RSASMYEWI QAASLYEKF QGASMYERF	LPPDIN/YATT DSPN.VLSLS DSPN.VLAVA YSPN.IMALA YSPN.INFPP YSSY YSPN.INFPP YSY YSPN.INFPP YSY YSPN.INFPP YSY YSPN.INFPP YSY YSPN.INFPP YSPN.INFPP	AANPRESYA SSLTHESYS SSLTHESYS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS VDETENSLKL VDETENSLKL VNEN E.LIW VNEN VNEN VNEN VNEN VNEN VNEN VNEN VNEN VNEN VNEN VNEN VNEN VNEN VNEN VTH VTH VTH VTH VTH VTH VTH VTH VTEISL VTTEISL 	CYYDE YDV YDV HQP HQP HQP HQP HQP HQP HQP B HQP B P SQ NAPMEVNORD PVPKATVNQRD KSKMEVVNQRD EAKMEVVNQRD ETKMEVVNQRD LDLTPSPD	.QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAIGVH ADLIFI ADLIFI ADLIFI ADLIFI ADLIFI ADLLFM AELFFU VPLTIL VPLTIL VPLTIM VPLSIM FWA	IGD WY3 VID RYTFYV VID RYTFYV ID RYTFYV ID RYTFYV IMD RYTFYV IMD RYTFYV IMD RYTFYV IMD RYTFYV IMD RYTFYV WQ RYKKSSE WQRYKKSSE WQUYQRSNF WQUYQRSNF KRKLLRTNI KRKLLRTNI KRKLMNTNI KRKLMNTNI SYKEDC SYKEDC SYKKDC SHEKAC	VNNWMEDSU VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN SSSQKNEA GSSQKNEA GSSQKNEA GSSQKNEA GSSKKKEI LOPEKKTDI UQPEKKTDI UQPEKKTDI UQPEKKTDI UDE SSQVL VK ESQNL LLE SRQL LLE SRQL LLE SRQL LLE SRQL DLE SRQL DLE SRQL MDEELMEP MDEELMEP MDEELMEP MDEELMEP	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL LP.LTTLK LQ.LAST LS.LAST
Q30H12 P49048 Q8T4E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49047 B8ASK4 P49046 Q4326 Q9R0J8 Q9R0J8 Q9R0J8 Q9R0J8 Q9S017 Q99538 Q5R5D9 Q95M12 P49048 Q8T4E1 Q9CXY9 Q92643 Q5R6L8 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1	LINYKDLNETIRYMYEHK ELTNYDLNETIRYMYEHK ELTNYDLNYAIOTMFEDN ELTSOELADGIOQMWEKK ETNIELADAFEQMWQKR ETNIELADAFEQMWQKR ETNIELADAFEQMWQKR ETNIELADAFEQMWQKR ETNIELADAFEQMWQKR ETNIELADAFEQMWQKR ETNIELADAFEQMWQKR STRAFTSNSNTYNTGSH KLVKKRTSNSNTYNTGSH KLVKRTSNSNTYNTGSH KLVKRTSNSNTYNTGSH KLVKRTSNSNTYNTGSH KLVKRTSNSNTYNTGSH KLVKRTSNSNTYNTGSH CSVKRTSNSNSYRGSH CSVKRTSNS CSVKRTSNSNSYRGSH CSVKRTSNS CSVKRTSNSNSYRGSH CSVKRTSNS CSVKRTS	MYQKMVFYLLAG RYHEMLVIADSC RYNELFFMVDTC RYNELFFMVDTC RYNELFIIDTC RYNELLFIIDTC RYNELLFIIDTC WELLFIIDTC WELLFIIDTC VMEYGSKDIKSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMQYGNKSISTM VMQYGNK NTDLFQ	ESGSMMN.H ESGSMMN.H ESAS MYEWI QAASLYEWI QAASLYEWI QAASMYERF QGASMYERF	LPPDIN/YATT LPPDIN/YATT DSPN.VLSLS DSPN.VLSLS ISPN.IMALA YSPN.IMALA	AANPRESYA SATHESYS SSUTHESYS SSUTEDSUS SSUTEDSUS SSUTEDSUS SSUTEDSUS SSUTEDSUS SSUTEDSUS SSUTEDSUS SSUTEDSUS SSUTEDSUS VERNOUS	CYYDE YDVD HUPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD PYDKATVNQRD RTPSAAVNQRD PVPKATVNQRD KSKMEVVNQRD EXKMEVVNQRD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD VQD QQD QQD QPP 450	QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIFU VPLTIU VPLTIU VPLTIM VPL	IGDWY3 VIDRYTFY VIDRYTFY IDRYTFY IDRYTFY IMDRYTFY WEKFRKAPE WEKFRKAPE WQUYQRSNF WRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN SYKEDC SYKEDC SYKEDC SYKEDC SHE KKVEEVFI KRK	VNW MEDSD VNFLTKEV UEFLEKVQ UEFLEKVQ UEFLEEIN UEFLEEIN UEFLEEIN UEFLEEIN UEFLEEIN SS.TRPEI GSSQKNEA GSEKKKEI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UEESRQL UEESR	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMDL B.ASQTNMDL
Q30H12 Q49048 Q874E1 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49047 P49046 P49046 P49046 P49046 P49046 Q5R5D9 Q95M12 P49048 Q874E1 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044	I. HVKD LNE TIRYMYEHK ELTINVELAYAIOTMFEDN ELTIN ELADAFEQMWQKR ETINIELADAFEQMWQKR ETINIELADAFEQMWQKR ETINIELADAFEQMWQKR ETINIELADAFEQMWQKR ETINIELADAFEQMWQKR ETINIELADAFEQMWQKR ETINIELADAFEQMWQKR ETINIELADAFEQMWQKR ELTINIELADAFEQMWQKR ELVKRTSNSNSYRFGSH ELVKRTSNSNSYRFGSH ELVKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSNYRFGSH ELVKKRTSNSNSYRFGSH ELVKKRTSNSNSNYRFGSH ELVKKRTSNNYTYAAK ELVKKRTSNNYTYAA	MYQKMVFYLLAG RYHEMLVIADSC RYNELFINDTC RYNELLFINDTC RYNELLFINDTC RYNELLFINDTC RYNELLFINDTC RYNELLFINDTC WYNELG VMEYGSKDIKPE VMEYGSKDIKPE VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDKSISTN VMQYGDKSISTN VMQYGNKSISTM VMQYGNK SIL SAM SAM SAM SAM SAM SAM SAM SAM	ESGSMMN.H ESGSMMN.H ESASMYEWI QAASLYEWI QAASLYEKF QGASMYERF	LPPDIN/YATT LPPDIN/YATT DSPN.VLSLS DSPN.VLSLS ISPN.VLSLS YSPN.IMALA YSPN.INFPP YSYN.INFPP YSYN.INFPP YSYN.INFPP YSYN.INFPP YSYN.INFPP YSYN.INFPP YSYN.INFY Y	AANPRESSY SSLTHESYS SSLTHESYS SSVGEDSLS SSVGEDS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSS	CYYDE YDV YDV HQP HQP HQP HQP HQP HQP B YDV HQP B YDV S YDV S YDV B YDV LDLTPSPD	QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAVGVH PAVGVH ADLIFL ADLIFF ADLUFF ADLUFF ADLUFF ADLUFF ADLUFF ADLUFF ADLUFF SQVVS SSEIMES SSEIMES SSEIMES SSGIMKS 460 CNSSVE	IGDWY3 VIDRYTHY7 VIDRYTHY7 IDRYTFYV IDRYTFYV IMDRYTFYV WQRYKSSE WQRYQRSNH WQWYQRSNH WQWYQRSNH WQWYQRSNH WQWYQRSNH WQWYQRSNH SSK SSK SSK SSK SSK SYK SYK <	VNN WMEDSO VNFLTKEV LEFLEKVQ LEFLEKVQ LEFLEENVQ LEFLEEIN LEFLEEIN LEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI CSSQKNEA GSSRKVEA MCSSR SSRVEA SSRV	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTTGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.SQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.SQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNML P.ASQTNML P.ASQTNML P.ASQTNL P.ASQTNML P.ASQTNML P.ASQTNML P.ASQTNML P.ASQTNML P.ASQTNL P.AS
Q30H12 Q45048 Q874E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49046 P49046 Q9F0J8 Q9F0J8 Q9F0J8 Q95M12 P49048 Q874E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4	L HVYD LNY AI OTMFEDN EL ITNYD LAY AI OTMFEDN EL ITNYD LAY AI OTMFEDN EL ITN EL AD AFE QMWQKR EI TN I EL AD AFE QMWQKR EL TN I EL AD AFE QMWQKR EI TN I EL AD AFE QMWQKR SVKRT SN SN SYN F SYN SVKART SN SN SYN SYN SYN SYN SVKART SN SN SYN SYN SYN SYN SYN SVKART SN SYN SYN SYN SYN SYN SYN SYN SYN SYN	MYQKMVFYLLAG RYHEMLVIADSC RYHELFFMVDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC WHEYGSKDIKPE VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDITTAE VMQYGQSKTSISTM VMQYGDTNITAE VMQYGDTNITAE VMQYGNKSISTM VMQYGNKSISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKTISTM VMQYGNKSISTM VMQYGNKTISTM VMQYGNKTISTM TDLFQRDPKNV RTDLFQRDF RTDLFQRDF RTDLFQRDF RTDL	ESGSMMN.H ESGSMMN.H ESASMYEWI QAASLYEWI QAASLYEWF QGASMYERF	LPPDIN/YATT DSPN.VLSLS DSPN.VLSLS DSPN.VLAVA YSPN.IMALA Y	AANPRESYA SSLTHESYS SSLTHESYS SSVGEDSLS SSVGEDSS SSVGEDS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS	CYYDE YDV YDV HQP HQP HQP HQP HQP HQP HQP B HQP B YDV HQP B HQP B P SQ CMKHMRSIANI GMKHMRSFANI GMKHMRSFANI GMKKMRSFANI GMKKMRSFANI GMKKMRSFANI	- QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIF AD	IGD WY3 VID RYTFYV VID RYTFYV ID RYTFYV ID RYTFYV IMD RYTFYV WQ RYKKSSE WQRYKKSSE WQUYQRSNF WQUYQRSNF KRKLLRTNI KRKLLRTNI KRKLMNTNI KRKLMNTNI KRKLMNTNI SYKEDC SYKEDC <th>VNW MEDSD VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI GSSGKVEA GSSKKVEA GSSKKKEI UPEKKTHI UPEKKTHI UPEKKTHI USSENKTDI USSENKTDI USSENKTDI USSENKTDI USSENKTHI NK.ESQVL VK.ESQNL LE.ESRQL LE.ESRQL LE.ESRQL LE.ESRQL LE.ESRQL MV.TKKP MDELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP</th> <th>VEDLTKETLHKQ KALNSSANMQOY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL L.ELTETLTHRG EKQVLEAMSHL LREITETLKHRK LQUKULEAMSHRK LQUKAIKHSS LQUKAIKHSS LQUKAIKHSK LQUELVKHKK LEQITETVKHRK LEQITETVKHRK VGQIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH LKY. LKY. LKY. LKY. MANG GPYGTFGA.YSV NPWSSLQGGFSA SPWSSLDRGFSA</th>	VNW MEDSD VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI GSSGKVEA GSSKKVEA GSSKKKEI UPEKKTHI UPEKKTHI UPEKKTHI USSENKTDI USSENKTDI USSENKTDI USSENKTDI USSENKTHI NK.ESQVL VK.ESQNL LE.ESRQL LE.ESRQL LE.ESRQL LE.ESRQL LE.ESRQL MV.TKKP MDELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP MDEELMEP	VEDLTKETLHKQ KALNSSANMQOY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL L.ELTETLTHRG EKQVLEAMSHL LREITETLKHRK LQUKULEAMSHRK LQUKAIKHSS LQUKAIKHSS LQUKAIKHSK LQUELVKHKK LEQITETVKHRK LEQITETVKHRK VGQIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH LKY. LKY. LKY. LKY. MANG GPYGTFGA.YSV NPWSSLQGGFSA SPWSSLDRGFSA
Q30H12 Q45048 Q874E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49046 Q4326 Q9R0J8 Q9R0J8 Q9R0J8 Q9S017 Q99538 Q5R5D9 Q95M12 P49045 Q95M12 P49048 Q874E1 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49047 B8ASK4 P49047 B8ASK4 P49047 B8ASK4 P49042	LINVKDLNETIRYMYEHK ELTNVDLNYAIOTMFEDN ELTNVELAYAIOTMFEDN EITSQELADGIQQWEKK ETTNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR ELTNIELADAFEQMWQKR STANAKLASS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSSS SVKRTSSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSS SVKRTSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSS SVKRTSSSSSS SVKRTSSSSSS SVKRTSSSSSSS SVKRTSSSSSSSSS SVKRTSSSSSSSSSSSS SVKRTSSSSSSS S	MYQKMVFYLLAG RYHEMLVIADSC RYHELFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC WEYGSKDIKSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMQYGNKSISTM VMQYGNK NTDLFQ NT	ESGSMMN.H ESGSMMN.H ESAS MYEWI QAASLYEWI QAASLYEWI QAASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF GFD LYDCGFD GFD KUYLYH GFD CFD GFD KUYLYH GFD CFD GFD KUYLYH GFD CFD GFD KUYLYH GFD CFD GFD KUYLYH GFD CFD GFD KUYLYH GFD CFD GFD KUYLYH GFD CFD GFD KUYLYH GFD CFD GFD KUYLYH GFD CFD GFD CFD GFD CFD GFD CFD GFD CFD GFD CFD GFD CFD GFD CFD GFD CFD GFD CFD GFD CFD CFD GFD CFD CFD CFD CFD CFD CFD CFD C	LPPDIN/YATT LPPDIN/YATT DSPN.VLSLS DSPN.VLSLS ISPN.IMALA YSPN.IMALA YSPN.INFERDIC YSPN.INFERDIC XKASSPISLPP YSPN.INFERDIC XKVE XKVE XKVE XKVE XKVE	AANPRESYA SATHESYS SSUTHESYS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS NQIHF VDETENSLKL VNEN.E.LIW NNAHL HNGRL HNGRL NQHF VDETENSLKI VNEN.E.LIW NAHL HNGRL NAHL HNGRL NAHL HSELVL VTH VTH VTH VTH VTH USTDRINU HLSEEIVL ITTEKISL ITTETISL ITTETISL ITTETISL GLL GSL SQY GSL VO	CYYDE YDVD HUPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD PVKATVNQRD RTPSAAVNQRD PVPKATVNQRD KSKMEVVNQRD EXKMEVVNQRD LDLTPSPD HQTTAR QQD QMKHMRAFANI QMKHMRAFANI </th <th>QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIFU VPLTIU VPLTIU VPLTIM VPLTIM VPLTIM VPLTIM VPLTU VPLTIM VPLTU</th> <th>IGDWY3 VIDRYTFYV VIDRYTFYV IDRYTFYV IDRYTFYV IMDRYTFYV WQRYKSSE WEKFRKAPE WQLYKRSEN WQVYQRSNF WRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLNTNN KRKLLRTNN KRKLLRTNN</th> <th>VNN WMEDSD VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI GSSKKKEI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UEESRQL VK.ESQVL V</th> <th>VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL J.EEITETLTHRG EKQVLEAMSHRL LREITETLTHRG EKQVLEAMSHRL LREIEDTIAHRK LQQIKDAIKHRS LKQIEIVKHRK LQQIKDAIKHRS LKQIEIVKHRK LKQIZUVKANSH IKQIQGFLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH VQKIDRHLEARN FKI. LKY. LKY. LKY. LKY. QPYGTFGA.YSV NPWSSLDRGFSA SPWSSLDRGFSA ARWSPMTEGGHS</th>	QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIFU VPLTIU VPLTIU VPLTIM VPLTIM VPLTIM VPLTIM VPLTU VPLTIM VPLTU	IGDWY3 VIDRYTFYV VIDRYTFYV IDRYTFYV IDRYTFYV IMDRYTFYV WQRYKSSE WEKFRKAPE WQLYKRSEN WQVYQRSNF WRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLNTNN KRKLLRTNN	VNN WMEDSD VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI GSSKKKEI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UEESRQL VK.ESQVL V	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL J.EEITETLTHRG EKQVLEAMSHRL LREITETLTHRG EKQVLEAMSHRL LREIEDTIAHRK LQQIKDAIKHRS LKQIEIVKHRK LQQIKDAIKHRS LKQIEIVKHRK LKQIZUVKANSH IKQIQGFLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH VQKIDRHLEARN FKI. LKY. LKY. LKY. LKY. QPYGTFGA.YSV NPWSSLDRGFSA SPWSSLDRGFSA ARWSPMTEGGHS
Q30H12 Q450H12 Q92048 Q92048 Q92048 Q92048 Q92048 Q92041 Q4KRV1 A0A002R159 P49044 P49047 B8ASK4 P49048 Q95N12 P49048 Q874E1 Q92643 Q55R5D9 Q95M12 P49048 Q874E1 Q92643 Q56R618 Q3MH27 Q4KRV1 A0A002R159 P49044 P49047 B8ASK4 P49042 D49042 D49047 B8ASK4 P49042 D49042 D49042 D49047 B8ASK4 P49042 D49047 B8ASK4 P49042 D49042 D49042 D49042 D49042 D49042 D49042 D49044 D49047 D49042 D49042 D49042 D49047 D49042 D49042 D49047 D49042 D49042 D49042 D49047 D49042 D49047 D49042 D49047 D49042 D49042 D49047 D49042 D49047 D49042 D49047 D49042 D49047 D49042 D49047 D49042 D49047 D49042 D49047 D49042 D49047 D49042 D49047 D49042 D49047 D49042 D49047 D49042 D49047 D49042 D49042 D49047 D49042 D49042 D49047 D49042 D49042 D49047 D49042 D49042 D49047 D49042 D49047 D49042 D49042 D49047 D49042 D49042 D49047 D49042 D49042 D49047 D49042 D49042 D49047 D49042 D49042 D49047 D49042 D49042 D49047 D49042 D49042 D49047 D49042 D49042 D49047 D49042 D49042 D49047 D49042 D49042 D49047 D49042 D49042 D49042 D49047 D49042 D49042 D49042 D49047 D49042 D49042 D49042 D49047 D49042 D49042 D49042 D49047 D49042 D49	I. HYKD LINE TIRYMYEHK ELTIN VELAYA ATO TMFEDN ELTIN IELADAFEQMWQKR ETIN IELADAFEQMWQKR ETIN IELADAFEQMWQKR EITIN IELADAFEQMWQKR EITIN IELADAFEQMWQKR EITIN IELADAFEQMWQKR EITIN IELADAFEQMWQKR EITIN IELADAFEQMWQKR EITIN IELADAFEQMWQKR ELTIN KRITAGSG, KSYGSH ELVKKRTAGSG, KSYGSH ELVKKRTAGSG, KSYGSH ELVKKRTAGSG, KSYGSH ELVKKRTAGSG, KSYGSH ELVKKRTAGSG, KSYGSH KSVKQRTSNFNNYAMGSH HLVKSHTN	MYQKMVFYLLAC RYHEMLVIADSC RYNELFFMVDTC RYNELFFIDTC RYNELLFIDTC RYNELLFIDTC RYNELLFIDTC RYNELLFIDTC WELGSKDIKPE VMEYGSKDIKPE VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDKSISKN VMEYGDANMTAE VMQYGDKSISTM VMQYGNK SISTM REC REC REC REC REC REC REC REC	ESGSMMN.H ESGSMMN.H ESASMYEWI QAASLYEWI QAASLYEWI QAASLYEWF QGASMYERF		AANPRESSY SSLTHESYS SSLTHESYS SSVGEDSLS SSVGEDS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSS	CYYDE YDV YDV HUP HQP HQP HQP HQP HQP B YDV HQP B YDV B YDV B YDV B YDV SO SSO SO DKLDGVNQRD PYPSRVTNQRD NAPMEVVNQRD KSKMEVVNQRD LDLTPSPD GMKHMRAFANI GMKHMRAFANI GMKHMRAFANI GMKHMRAFANI </th <th>.QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAVGVH PAVGVH ADLIFL ADLIFL ADLIFL ADLIFL ADLIFL VPLTIL VPLTIL VPLTIL VPLTIL VPLTIL VPLTIL VPLTIL SQVVDS SSEIMES SSGIMKS 460 CNSSVE CNAGIQ CNSGVS</th> <th>IGDWY3 VIDRYTHY7 VIDRYTFYV IDRYTFYV IDRYTFYV IMDRYTFYV WQRYKSSE WQRYQRSNF WQWYQRSNF WQWYQRSNF WQWYQRSNF WQWYQRSNF WQWYQRSNF SSKEDC SYKEDC SYKEDC</th> <th>VNN WMEDSO VNFLTKEV UNFLTKEV UNFLTKEV UEFLEEKVQ UEFLEEIN UEFLEEIN UEFLEEIN UEFLEEIN UEFLEEIN UEFLEEIN UEFLEEIN UEFLEEIN UEFKTDI QPEKKTDI QPEKKTDI QPEKKTDI QPEKKTDI UEESRQL ULEESRQL ULEESRQL ULEESRQL ULEESRQL ULEESRQL ULEESRQL ULEESRQL QACESRL MODELMEP</th> <th>VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL 380 380 LREITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK QKQVLEAMSHRK LQUIKDAIKHRS LKQIEETVKHRK LQUIKDAIKHRS LKQIEETVKHRK LQUIKDAIKHRS LKQIETVKHRK LQUIKDAIKHS LKQIEUVKHRK LKQIETVKHRK LKQIEUVKHRK LKQIEDVKHRK LKY EKI VGQIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH KY LKY LKY LKY LKY LKY SPWSSLQGFSA SPWSSLQGFSA ARWSPMTEGGHS GQUHPTNQGYSA</th>	.QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAVGVH PAVGVH ADLIFL ADLIFL ADLIFL ADLIFL ADLIFL VPLTIL VPLTIL VPLTIL VPLTIL VPLTIL VPLTIL VPLTIL SQVVDS SSEIMES SSGIMKS 460 CNSSVE CNAGIQ CNSGVS	IGDWY3 VIDRYTHY7 VIDRYTFYV IDRYTFYV IDRYTFYV IMDRYTFYV WQRYKSSE WQRYQRSNF WQWYQRSNF WQWYQRSNF WQWYQRSNF WQWYQRSNF WQWYQRSNF SSKEDC SYKEDC	VNN WMEDSO VNFLTKEV UNFLTKEV UNFLTKEV UEFLEEKVQ UEFLEEIN UEFLEEIN UEFLEEIN UEFLEEIN UEFLEEIN UEFLEEIN UEFLEEIN UEFLEEIN UEFKTDI QPEKKTDI QPEKKTDI QPEKKTDI QPEKKTDI UEESRQL ULEESRQL ULEESRQL ULEESRQL ULEESRQL ULEESRQL ULEESRQL ULEESRQL QACESRL MODELMEP	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL 380 380 LREITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK QKQVLEAMSHRK LQUIKDAIKHRS LKQIEETVKHRK LQUIKDAIKHRS LKQIEETVKHRK LQUIKDAIKHRS LKQIETVKHRK LQUIKDAIKHS LKQIEUVKHRK LKQIETVKHRK LKQIEUVKHRK LKQIEDVKHRK LKY EKI VGQIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH KY LKY LKY LKY LKY LKY SPWSSLQGFSA SPWSSLQGFSA ARWSPMTEGGHS GQUHPTNQGYSA
Q30H12 Q45048 Q874E1 Q92643 Q5R6L8 Q30H127 Q4KRV1 A0A0G2R159 P49044 P49047 B8A5K4 P49042 Q4326 Q99538 Q5R5D9 Q95M12 P49048 Q874E1 Q95X12 P49048 Q874E1 Q95X12 P49048 Q874E1 Q925M12 P49048 Q874E1 Q925M2 P49046 P49044 P49047 B8A5K4 P49044 P49047 B8A5K4 P49046	L HVYD LNY AI DYMYE HK ELTINYD LNY AI OTMFEDN ELTINYD LAY AI OTMFEDN ELTIN ELADAFEQMWQKR EITIN IE LADAFEQMWQKR EITIN IE LADAFEQMWQKR EITIN IE LADAFEQMWQKR EITIN IE LADAFEQMWQKR EITIN IE LADAFEQMWQKR ELTIN IE LADAFEQMWQKR ELTIN IE LADAFEQMWQKR SUVKRTSNSNTYNTGSH KLVKNRTI. S. EPYGSH ELVKRTSNSNTYNTGSH KLVKNRTSNSNTYNTGSH KLVKRTSNSNTYNTGSH KLVKRTSNSNYFFAYAMGSH QSVKQTSNFAYAMGSH QSVKQTSNFAYAMGSH QSVKQTSNFAYAMGSH QSVKQTSNFAYAMGSH QSVKQTSNFAYAMGSH DSVKQRTSNFNYTAAGSH HLVKSHTN. TSH HLVKSHTN. TSH HLVKSHTN. TSH DSCPARKCLSNT. GV QVCPKSLCVSTP. GH QVCPKSLCVSTP. GH QVCPKSLCVSTP. GH QVCPKSLCVSTP. GH QVCPKSLCVSTP. GH QVCPKSLCVSTP. GH QVCPKSLCVSTP. GH QVCPKSLCVSTP. GH QVCPKSLCVSTP. GH DSCSIDMIGVLFGPQNG IDNSVKLIGQLLFGIEKG LDSSIDTIGKLYGFENG LDSSIDTIGKLYGFENG	MYQKMVFYLLAG RYHEMLVIADSC RYHELFFMVDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC WYNEYGSKDIKPE VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDKTISTM VMEYGDKSISTM VMQYGNKSISTM VMQYGNKSISTM VMQYGNKSISTM VMQYGNKSISTM VMQYGNKTISTM RKDHYPKDVKRV RKDLYPRDPHKVN RKDLYPRDPHKVN RKDLYPRDPKNV RTDLFQRDV RTDLFQRDV	ESGSMMN.H ESGSMMN.H ESAS MYEWI QAASLYEWI QAASLYEWI QAASLYEWF QGASMYERF	LPPDIN/YATT DSPN.VLSLS DSPN.VLSLS DSPN.VLSLS YSPN.IMALA Y	AANPRESSYS SSLTHESYS SSLTHESYS SSVGEDSLS SSVGEDSS SSVGEDSLS SSVGEDSLS SSVGEDS SSVGEDSLS SSVGEDSLS SSVGEDSS	CYYDE YDVD HHVD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD DKLDGVNQRD KTPSAAVNQRD VPKATVNQRD VPKATVNQRD VPKATVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EATKMEVVNQRD LDLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DQD QPD QPD QPD QPD QPD QPD QPD QPD QPD QPD 	.QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIF AD	IGD W Y IGD W Y VID RYTFY Y MID RYTFY Y IMD RYTFY Y WQ RYKSSE Y WEKYCASE W Y WU YQRSNF Y KRKLLRTNN K K KRKLLRTNN K K KRKLLNNTNI K K K.KKLLNTNI K K K.K.LMNTNI K K SYK EDC S SYK EDC S SYK EDC S K K K K K K <th>VNW MEDSD VNFLTKEV VLEFLEKVQ VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI GSSGKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKEA GSSKKEA GSSKKEA GSSKKEA GSSKKEA GSSKKEA UPEKEA UPEKEA DEESRQL DESRQL DES</th> <th>VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMDL B.A</th>	VNW MEDSD VNFLTKEV VLEFLEKVQ VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI GSSGKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKKEA GSSKKEA GSSKKEA GSSKKEA GSSKKEA GSSKKEA GSSKKEA UPEKEA UPEKEA DEESRQL DESRQL DES	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMDL B.A
Q30H12 Q45048 Q874E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49046 Q9R0J8 Q9R0J8 Q9R0J8 Q9S0J7 Q9S538 Q5R5D9 Q95M12 P49045 Q95M12 P49048 Q874E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49047 B8ASK4 P49047 B8ASK4 P49047 B8ASK4 P49042 Q24326 P49045	I LIVKD LIAY AI OTMFEDN ELITNVELAY AI OTMFEDN ELITN IE LADAFEQMWQKR EITN IE LADAFEQMUTT IS H EIVKSHTN	MYQKMVFYLLAGC RYHEMLVIADSC RYHELFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC WELLFILDTC WEYGSKDIKSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDANMTAE VMEYGDANTAE VMEYGDA	ESGSMMN.H ESGSMMN.H ESAS MYEWI QAASLYEWI QAASLYEWI QAASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF CONSTRUCTION SOUTHOUS	LPPDIN/YATT LPPDIN/YATT DSPN.VLSLS DSPN.VLSV YSPN.IMALA YSPN.INFPP YSPN.INFPP YSPN.INFPP YSPN.INFPP YSPN.INFPP YSPN.INFPP YSPN.INFPP YSPN.INFPP Y	AANPRESYA SATHES YA SSLTHES YS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS VDETENSLKL VEENSLKL VEENSLKL VEENSLKL VEENSLKL VEENSLKL VEENSLKL VEENSLKL VEENSLKL VEENSLKL VEENSLK VE	CYYDE YDVD HUVD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD PVKATVNQRD RTPSAAVNQRD PVKATVNQRD VVPKATVNQRD EXKMEVVNQRD EXKMEVVNQRD LDLTPSPD HQTTPSPD HQTTPSPD HQTTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDUTPSPD LDUTPSPD LDUTPSPD HQTANNIG GMKHMRAFANIG GMKHMRAFANIG GMKHMRAFANIG GMKHMRAFANIG GMKHMRAFANIG GMKHMRAFANIG GMKHMRAFANIG	QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIFU VPLTIU VPLTIU VPLTIM VPLTIM VPLTIM VPLTU V	IGDWY3 VIDRYTFYV VIDRYTFYV IDRYTFYV IDRYTFYV IMDRYTFYV WQRYKSSE WQRYQRSNF WRKRYEQLNC WQVYQRSNF WRKKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLNNTNI KRKLLRTNN KRKLARASA KSYK KSK KSK KSK KSK <	VIN WHEDSD VINFLTKEV VINFLTKEV VINFLTKEV VIEFLEEKVQ VIEFLEEIN VIEFLEEIN VIEFLEEIN VIEFLEEIN VIEFLEEIN VIEFLEEIN VIEFLEEIN VIEFLEEIN SS.TRPEI GSSKKKEI UPEKKTDI GOPEKKTHI UOPEKKTDI UOPEKKTHI UOPEKKTDI UDEESRQL VK.ESQVL VK	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL 380 39 LREITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK LQUEAMSHRL LREIEDTIAHRK LQUIKDAIKHRS LKQIEEIVKHRK LGQUKDAIKHRS LKQIEIVKHRK VGQIQHLLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLARH KY KY LKY LKY LKY LKY LKY SPWSSLQGGFSA ARWSPMTEGGHS. GQWHPTGGYSA GLLHPSNTGYSA GLLHPSNTGYSA
Q92643 Q874E1 Q92643 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49047 B8ASK4 P49046 P49046 Q95N32 Q95N12 P49048 Q874E1 Q92643 Q5R5D9 Q95M12 P49048 Q874E1 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49044 P49044 P49047 B8ASK4 P49042 Q24326 P49045 Q9R0J8	I. HYVD LIAY AI OTMFEDN ELTNVELLAY AI OTMFEDN ELTN IE LAD AFLOMWQKR EITN IE LAD AFLOMWQKR ELTN KRTAGSG, KSYGSH ELVKKRTAGSG, KSYGSH ELVKKRTAGG, KSYGSH ELVKKRTAG, KSYGSH ELUKKRTAG, KSYGSH ELUKKRTAGG, ELUGK, KSYGN, KSYGH ELUKKRTAGG, KSYGH ELUKKRTAG, KSYGN, KSYN, KSYGN, KSYN,	MYQKMVFYLEAG RYHEMUVIADSC RYHELFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC WAEYGSKDIKPEI VMEYGDIGLSKNI V	ESGSMMN.H ESGSMMN.H ESASMYEWI QAASLYEWI QAASLYEWI QAASLYEWF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF GFD KUYLYGGFD KUYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLYGGFD KLYLGFG GSV LITDFFGSV LIT		AANPRESSYS SSLTHESEYS SSLTHESEYS SSVGEDSLS SSVGEDS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGE	CYYDE YDV YDV HQP HQP HQP HQP HQP HQP HQP BASO JSO JSO DKLDGVNQRD RTPSAAVNQRD PYPKATVNQRD KSKMEVVNQRD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD SGMKHMRAFANI GMKHMRAFANI	CNSSVE CN	IGD WY3 VIDRYTHYI RYTHYI IDRYTFYV X IDRYTFYV X IMDRYTFYV X WQRYKSSE X WQRYQRSNF WQNYQRSNF WQNYQRSNF WQNYQRSNF WQNYQRSNF X KRKLLRTNI K KRKLMNTNI K KRKLMNTNI K KRKLMNTNI K KSK EDC SYK EDC KSKVEEVFI NEPMAEASSI HTSMEEACN	NMMMEDSU VNFLTKEV VNFLTKEV VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEN SS.TRPEI GSSAKVEA GSSEKKEI ILPEKKTDI IQPEKKTDI IQESNKTDI ULE.ESRQL NV.TKKP TAECGGKNI QACPELMEP MDELMEP MDELMEP MOELMEP ACGGRYNA AACCGRYNS AACCGYNA AACCGYNA	VEDLTKETLHKQ XALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNML P.ASQTNML P.ASQ
Q30H12 P49048 Q8T4E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49046 Q9S0J8 Q9S0J8 Q9S0J8 Q95M12 Q95M12 P49048 Q8T4E1 Q9CXY9 Q925M12 Q95K12 P49048 Q8T4E1 Q9CXY9 Q925M12 P49046 P49047 B8ASK4 P49048 B7 B7 B8ASK4 P49048 B7 B7 B7 B7 B7 B7 B7 B7 B7 B7 B7 B7 B7	I. HVVD LNY AI OTMFEDN ELTNVD LNY AI OTMFEDN ELTN IE LADAFEQMWQKR EITN IE LADAFEQMWQKR SOV QNQ G G G G G G G G G G G G G G G G G G	MYQKMVFYLADSC RYHEMLVIADSC RYHELFFMVDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC WYNEYGSKDIKFE VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDKTISTM VMEYGDKTISTM VMEYGDKSISTM VMEYGDKSISTM VMEYGDKSISTM VMEYGDKSISTM VMEYGDKSISTM VMEYGDKSISTM VMEYGDKSISTM VMEYGDKSISTM VMEYGDKSISTM VMEYGDKSISTM VMEYGDKSISTM VMEYGDKSISTM VMEYGDKSISTM VMEYGDKSISTM VMEYGDKSSISTM VMEYGDKSISTM VMEYGDKSISTM VMEYGDKSSSISTM VMEYGDKSSISTM VMEYGDKSSSS VMEYGDKSSSS VMEYGDKSSSS VMEYGDKSSSS VMEYGDKSSSS VMEYGDKSSSS VMEYGDKSSSS VMEYGDKSSSS VMEYGDKSSSS VMEYGDKSSSS VMEYGDKSSSS VMEYGDKSSSS VMEYGDKSSSS VMEYGDKSSSS VMEYGDKSSSSS VMEYGDKSSS	ESGSMMN.H ESGSMMN.H ESAS MYEWI QAASLYEWI QAASLYEWI QAASLYEWF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF GFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD GMLK KUYLOFGGMLK KUYLYEGFD GMLK KUYLYEGFD GMLK KUYLFG GGNLK KUYLFG CULF	LPPDIN/YATT DSPN.VLSLS DSPN.VLSLS DSPN.VLAVA YSPN.IMALA Y	AANPRESSYS SSITHESYS SSITHESYS SSVVGEDSLS SSVVGEDSLS SSVVGEDSLS SSVVGEDSLS SSVVGEDSLS SSVGEDSS SSVGEDSLS S	CYYDE YDVD HHVD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD DKLDGVNQRD KTPSAAVNQRD VPKATVNQRD VPKATVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD LDLTPSPD LDLTPSPD LDLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD QQD QPD QPD QPD QPD QPD QPD QPD QPD QPD QPD QPD QPD QPD QPD 	.QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIF A	IGD W Y IGD W Y VID RYTFY Y MID RYTFY Y IMD RYTFY Y WEKS SYK S WEKYCASPE WEKYCASPE WQUYYCRSNF WQUYYCRSNF KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLNTNE KRKLLNTNE KRKLMNTNE KRKLMNTNE KRK SYK EDC SYK EDC SYK EDC SYK EDC SYK EDC SYK EDC KSKVEE ATO KSKVEE ATO KSK <th>VNW MEDSD VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI GSSGKVEA GSSKKKEI LEPEKKTHI UPEKKTHI UPEKKTHI UPEKKTHI USSENKTESQVL VK.ESQVL VK.ESQVL VK.ESQVL VK.ESQVL UE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL ALCOGKNI ACCGKNI NACGGYNA AACCGGYAA AAACGGYAA AAACGYAA AAACGYAA AAACGYAA AAACGYAA</th> <th>VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNML P.ASQTNML P.ASQT</th>	VNW MEDSD VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN SS.TRPEI GSSGKVEA GSSKKKEI LEPEKKTHI UPEKKTHI UPEKKTHI UPEKKTHI USSENKTESQVL VK.ESQVL VK.ESQVL VK.ESQVL VK.ESQVL UE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL DLE.ESRQL ALCOGKNI ACCGKNI NACGGYNA AACCGGYAA AAACGGYAA AAACGYAA AAACGYAA AAACGYAA AAACGYAA	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNML P.ASQTNML P.ASQT
Q30H12 Q45048 Q874E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49046 Q9R0J8 Q9R0J8 Q9S017 Q92643 Q5R5D9 Q92643 Q5R5D9 Q95M12 P49044 Q92643 Q5R6L8 Q9CXY9 Q92643 Q5R6L8 Q9CXY9 Q92643 Q5R6L8 Q9CXY9 Q4KRV1 A0A0G2R159 P49044 P49045 Q9CXY9 Q4KRV1 A0A0G2R159 P49046 P49045 Q9CMJ8 Q89017 Q99538	. LHVKDLNY TIRYMYEHK . LHVKDLNY AI OTMFEDN ELTNYELADAFEQMWQKR EITNIELADAFEQMWQKR SVKKRTSNSNTYNTGSH GVKRSTSNNNYTAAGSH GVKRSTSNNSNSYFGSH KSVKQRTSNFNNYAMGSH QVKKSTNNNNTNNTSSNSNSTSH BLVKSHTNNNNTSSNSTSH BLVKSHTNNNNNNNTSSNSTSH GVCPKSLCVSTP	MYQKMVFYLLAG RYHEMLVIADSC RYHELFIDTC RYNELLFIDTC RYNELLFIDTC RYNELLFIDTC RYNELLFIDTC RYNELLFIDTC WELLFIDTC WEYGSKDIKSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDITTTAE VMQYGNKSISTM VMQYGNK SSI MILSAN MQYGNK SSI MILSAN MQYGNK MULSAN MQYGNK MULSAN MQYGNK MULSAN M	ESGSMMN.H ESGSMMN.H ESAS MYEWI QAASLYEWI QAASLYEWI QAASLYEWF QAASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF GFD KVYLYEGFD KVYLYEGFD KVYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEG SUYDD VE SUYDD VE SUYDD SUYD SUYDD SUYDD SUYDD SUYDD SUYDD SUYDD SUYDD SUYD SUYDD SUYDD SUYDD SUY		AANPRESSYA SATHESYS SSUTHESSS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS SSUVGEDSLS VDEN.SLKL VDEN.SLK VDEN.SLKL VDEN.SLK VDEN	CYYDE YDVD HHVD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD PVPKATVNQRD PVPKATVNQRD PVPKATVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVNQRD EXKMEVNQRD EXKMEVNQRD EXKMEVNQRD EXKMEVNQRD EXKMEVNQRD EXKMEVNQRD EXKMEVNQRD EXKMEVNQRD EXKMEVNQRD EXKMEVNQRD EXKMEVNQRD EXKMEVNQRD EXKMENTANI GMKHMRAFANI GMKHMRAFANI GMKHMRAFANI GMKHMRAFANI GMKHMRAFANI GMKHMRAFANI GMKHMRAFANI GMKHMRAFANI CMKHMRAFANI	QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIF AD	IGD W Y IGD W.I. Y VID RYTHYT MID RYTFYV IMD RYTFYV WQ RYCRSP WKRVEQUNC WQVYQRSNF WRRVEQNSNF KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRK KRKLLRTNN KRK KRKLLRTNN KRK KRKLNNN RK KRKLLNNNN KRK KRK KRK KSYK EDC SYK EDC SYK EDC </th <th>VNN WMEDSU VNFLTKEV VNFLTKEV UEFLEKVQ VLEFLEEVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEN GSSKKKEI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UEESRQL UEESRQL ULEESRQL ULEESRQL ULEESRQL VVTKKP MDEELMEP MDE</th> <th>VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL 380 39 LREITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK LQUEAMSHRL LREIEDTIAHRK LQUIKDAIKHRS LKQIEUVKHKK LQUIKDAIKHRS LKQIEUVKHKK LGQUFLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLARH KY. LKY. LKY. LKY. LKY. LKY. LKY. LKY.</th>	VNN WMEDSU VNFLTKEV VNFLTKEV UEFLEKVQ VLEFLEEVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEN GSSKKKEI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UEESRQL UEESRQL ULEESRQL ULEESRQL ULEESRQL VVTKKP MDEELMEP MDE	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL 380 39 LREITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK LQUEAMSHRL LREIEDTIAHRK LQUIKDAIKHRS LKQIEUVKHKK LQUIKDAIKHRS LKQIEUVKHKK LGQUFLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLARH KY. LKY. LKY. LKY. LKY. LKY. LKY. LKY.
Q95M12 Q874E1 Q92643 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49042 Q49045 Q95M32 Q95M32 Q874E1 Q925X39 Q95M2 Q874E1 Q92X39 Q95M12 P49048 Q874E1 Q92X9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49044 P49044 P49044 P49044 P49044 P49044 P49044 P49044 P49044 P49044 P49044 P49045 Q24326 P49045 Q9R0J8 O89017 Q99538 Q95559 Q95538 Q95559 Q95538 Q95538 Q95538 Q95538 Q95538 Q95538 Q95538 Q95559 Q95555 Q95555 Q95555 Q95555 Q95555 Q95555 Q	. LHVKDLNE TIRYMYEHK ELTINVELAYAIOTMFEDN ELTINGELANGIOOMWEKK EITINIELADAFEQMWQKR EITINIELADAFEQMWQKR EITINIELADAFEQMWQKR EITINIELADAFEQMWQKR EITINIELADAFEQMWQKR EITINIELADAFEQMWQKR ELTINIELADAFEQMWQKR ELTINIELADAFEQMWQKR SUVKRTSNSNTYNTGSH KLVKNRTSNSNTYNTGSH KLVKNRTSNSNTYNTGSH KLVKRTSNSNTYNTGSH SVKARTSNSNTYNTGSH SVKARTSNSNTYNTGSH SVKARTSNSNTYNTGSH SVKARTSNSNTYNTGSH SVKARTSNSNTYNTGSH SVKARTSNSNTYNTGSH SVKARTSNSNTYNTGSH SVKARTSNSNTYNTGSH SVKARTSNSNTYNTGSH SVKARTSNNTYNTGSH SVKARTSNYNTYAAGSH SVKARTSNYNTGST SVKARTSNYNTYAAGSH SVKARTSNYNTYAAGSH SVKARTSNYNTGST SVKARTSNYNTYAAGSH SVKARTSNYNTYAGST SVKARTSNYNTYAAGSH SVKARTSNYNTYA	MYQKM VFYLEAG RYHEM/VIADSC RYHEM/VIADSC RYNELFFM/DTC RYNELFFM/DTC RYNELFFIDTC RYNELFIDTC RYNELFIDTC RYNELFIDTC RYNELFIDTC RYNELFIDTC RYNELFIDTC VMEYGSKDIGLSKN VMEYGSKDIGLSKE VMEYGQUSKS VMEYGQUSKS VMEYGQUSKS VMEYGDANMTAE VMQYGDANMTAE VMQYGDANMTAE VMQYGDINITAE VMQYGDTNITAE VMQYGDTNITAE VMQYGDKSISTM VMQYGNK SSVLRSV MNQYGNK SSVLRSV MND MND MND MND MND MND MND MND	ESGSMMN.H ESGSMMN.H SSASMYEWI QAASLYEWI QAASLYEWF QGASMYERF	LPPDIN/YATT DSPN.VLAVA VLSIS ISPN.VLAVA YSPN.IMALA YSTON<	AANPRESSYA SATHERSYA SSLTHESYS SSVGEDSLS SSVGE	CYYDE YDVD HHVD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD TSO SSO DKLDGVNQRD FYPSRVTNQRD PYPKATVNQRD FAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD LDLTPSPD LDLTPSPD LDLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD QVD QPD 	.QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAVGVH PAVGVH ADLIFI ADLIFF ADLUFM ADLIFF ADLUFM AELLFM AELLFM AELLFM AELLFM VPLTIL VPLTIL VPLTIL VPLTIM VPLTIL SQVVDS SSEIMES SSGVVS SSGMKS 460 CNSSVE CNAGIQ CNSSVE CNAGIQ CNSGVS CNSGVS CNSGVS	IGD W Y IGD W Y VID RYTHYT WID RYTFYV IMD RYTFYV WQRYCS RYK WQRYQRSNH WQWYQRSNH WQWYQRSNH WQWYQRSNH WQWYQRSNH RKKLLRTNNI KRKLLRTNNI KRKLLRTNNI KRKLLNTNI KRKLMNTNI KRKLMNTNI KRK FDK SYK EDC KKK EX<	VIN WHED SO VINFLTKEV VINFLTKEV VINFLTKEV VIEFLEEKVQ VIEFLEEIN VIEFLEEN VIEF	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL P.ASQTNMNDL B.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C.C
Q30H12 Q42048 Q874E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49046 Q950J8 Q950J8 Q950J8 Q95538 Q95538 Q95538 Q95538 Q95538 Q95538 Q95538 Q95538 Q95538 Q95538 Q95538 Q95542 Q9542 Q9542 Q9542 Q92443 Q92443 Q92443 Q327 Q4KRV1 A0A0G2R159 P49044 P49045 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49045 Q9242 Q24326 P49046 P49045 Q980J8 Q880J7 Q99538 Q585D9 Q99538 Q555D9 Q95535 Q95532 Q95552 Q95552 Q95552 Q95552 Q95552 Q95552 Q95552 Q95552	L HVVD LNY AI OTMFEDN ELTINVELAY AI OTMFEDN ELTIN ELADAFEQMWQKR EITIN IE LADAFEQMWQKR EITIN IE LADAFEQMWQKR SUVKRTSNSNTYNTGSH KLVKNRTI.ADAFEQMWQKR EUVKRTSNSNTYNTGSH KLVKNRTSNNNKLNEGSH SVKARTSNNNKLNEGSH SVKARTSNNNKLNEGSH SVKARTSNNNKLNEGSH QSVKQRTSNFAYAMGSH QSVKQRTSNFAYAMGSH QSVKQRTSNFAYAMGSH USVKQRTSNFAYAMGSH USVKQRTSNFAYAMGSH USVKQRTSNFAYAMGSH USVKQRTSNFAYAMGSH USVKQRTSNFAYAGSH USVKQRTSNFAYAMGSH USVKQRTSNFAYAMGSH USVKQRTSNFNYTAGSH HLVKSHTNTSH HLVKSHTNTSH USCPARKCLSNTGV QVCPKSLCVSTPGH QVCFKSLCVSTPGH QVCFKSLCVSTPGH QVCFKSLCVSTPGH QVCFKSLCVSTPGH QVCFKSLCVSTPGH QVCFKSLCVSTPGH QVCFKSLCVSTPGH QVCFKSLCVSTPGH QVCFKSLCVSTPGH QVCFKSLCVSTPGH QVCFKSLCVSTPGH QVCFKSLCVSTPGH QVCFKSLCVSTPGH QVCFKSLCVSTPGH QVCFKSLCVSTPGH QVCFKS	MYQKMVFYILAG RYHEM/VIADSC RYHEM/VIADSC RYNELFFM/DJC RYNELFFM/DJC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC VMEYGSKDIKSE VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDVGSKKN VMEYGDVGSKKN VMQYGQSKTISTM VMQYGQKKTISTM VMQYGNKSISTM VMQYGNKSISTM VMQYGNKSISTM VMQYGNKSISTM VMQYGNKSISTM VMQYGNKSISTM RKDHYPKDVKV RTDLFQRDPKNV RTDLFQRDV RTDV RTDV RTDV RTDV RTDV RTDV RTDV RTDV RTDV RTDV RTDV RTDV RTDV RTDV RTDV	ESGSMMN.H ESGSMMN.H ESGSMMN.H ESAS MYEWI QAASLYEWI QAASLYEWI QAASLYEWI QAASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF GFD EVUL EVUL EVUL EVUL EVUL EVUL EVUL EVUL EVUL EVUL EVUL EVUL EVUL EGSV EVUL EVUL EGSV EVUL EGSV EVUL		AANPRESSYS SSITHESYS SSITHESYS SSITHESYS SSVGEDSLS SSVGE	CYYDE YDV HUV HQP	.QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIF AD	IGD W Y IGD W Y VID RYTHYT WID RYTHYT MID RYTFYV IMD RYTFYV IMNO RYTT KRKLLA	VNW MEDSD VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEN VLEFLE VLES VLES VLES VLES VLES VLES VLES VL	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL 380 39 LREITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK LQQIKDAIKHRS LKQIEEIVKHRK LQQIKDAIKHRS LKQIEEIVKHRK LQQIKDAIKHRS LKQIETVKHRK LKQIQGFDARH TEEIQHLDARH TGQIQGFDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH CKY LKY LKY LKY LKY LKY LKY LKY LKY LKY L
Q30H12 Q42048 Q874E1 Q9CXY9 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49046 Q4326 Q9R0J8 Q9R0J8 Q9R0J8 Q9R0J8 Q9R0J8 Q9SR5D9 Q9SR5D9 Q9SR5D9 Q9SR5D9 Q92643 Q5R5D9 Q92643 Q5R5D9 Q92643 Q5R5D9 Q92643 Q5R5D9 Q92643 Q5R5D9 Q92643 Q5R5D9 Q92643 Q5R5D9 Q92643 Q5R5D9 Q92643 Q585D9 Q92643 Q585D9 Q92643 Q585D9 Q92643 Q585D9 Q92643 Q585D9 Q92643 Q585D9 Q92643 Q585D9 Q92643 Q585D9 Q92643 Q585D9 Q92643 Q585D9 Q92643 Q585D9 Q92643 Q585D9 Q92643 Q585D9 Q92643 Q585D9 Q92643 Q585D9 Q95017 Q92538 Q585D9 Q955012 P49048	. LHVKDLNETIRYMYEHK ELTNVDLNYAIOTMFEDN ELTNVELAYAIOTMFEDN ELTSQELADGIQQWEKK ETTNIELADAFEQMWQKR ETTNIELADAFEQMWQKR ETTNIELADAFEQMWQKR ETTNIELADAFEQMWQKR ETTNIELADAFEQMWQKR ETTNIELADAFEQMWQKR ETTNIELADAFEQMWQKR ELTNIELADAFEQMWQKR STANAKLASS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSSS SVKRTSSS SVKRTSSSS SVKRTSSS SVKTTSSSS SVKTTSSSS SVKTTSSSS SVKTTSSSS SVKTTSSSS SVKTTSSS SVKTTSSSS SVKTTSSSS SVKTTSSSS SVKTTSSSS SVKTTSSSS SVKTTSSSS SVKTTSSSS SVKTTSSSS SVKTTSSSS SVKTTSSSS SVKTTSSSS SVKTTSSSS SVKTTSSSSS SVKTTSSSS SVKTTSSSS SVKTTSSSSS SVKTTSSSSSS SVKTTSSSSSSS SVKTTSSSSSSS SVKTTSSSSSSSS SVKTTSSSSSSSSSS	MYQKMVFYLLAG RYHEMLVIADSC RYHELFILJTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC WELLFILDTC WELLFILDTC VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDANMTAE VMEYGDANTTAE VMEYGDA	ESGSMMN.H ESGSMMN.H ESAS MYEWI QAASLYEWI QAASLYEWI QAASLYEWF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF GFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEGFD KUYLYEG SUYDDFFGSV LITDFFGSV LVDDWTCL LITGHSCY LITGHSCY LITGHSCY		AANPRESSYA AANPRESYS SSLTHEESYS SSLTHEESYS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS SSVGEDSLS 	CYYDE YDVD HUVD HUVD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD PQPD HQPD HQPD PVKATVNQRD KSKEVVNQRD EXKMEVVNQRD EXKMEVVNQRD LDLTPSPD HQTPSPR HMKHMRAFANI GMKHMRAFANI ALRYLYVLANLALALALRHLYVL	QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIFU ADLIFU ADLIFU ADLIFU ADLIFU ADLIFU ADLIFU AELLFM AELLFM AELLFM AELLFM AELLFM VPLTIL VPLTIM VPLTM VPLTIM VPLTM V	IGD W Y IGD W Y VID RYTHYT MID RYTFYV MID RYTFYV IMD RYTFYV WQ RYCRSNF WQUYQRSNF WQUYQRSNF WQUYQRSNF RYKKILRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KSKULLRTNN KRKLLRTNN	VNN WMEDSU VNFLTKEV VNFLTKEV UEFLEKVQ UEFLEKVQ UEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEN VKESQVL	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL 380 39 LREITETLTHRG EKQVLEAMSHRL QKQVLEAMSHRL LREIEDTIAHRK LQUEDTIAHRK LQUEDTIAHRK LQUETVKHRK LQUTETVKHRK LQUTETVKHRK LQUTETVKHRK LQUTETVKHRK LQUTETVKHRK LKY TEEIQHLLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLARH TEEIQHLARH SPWSLQGGFSA ARWSPMTEGGS SPWSSLQRGFSA ARWSPMTEGS SPWSSLQRGFSA ARWSPMTEGS SPWSSLQRGFSA ARWSPMTEGS SPWSSLQRGFSA ARWSPMTEGS SPWSSLQRGFSA ARWSPMTEGS SPWSSLQRGFSA
Q95M12 Q874E1 Q92643 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49042 Q92643 Q95M12 P49046 P49046 P49045 Q95M12 P49048 Q874E1 Q92XY9 Q92643 Q3MH27 Q4KRV1 A0A0G2RI59 P49048 Q874E1 Q92XY9 Q92643 Q5R6L8 Q3MH27 Q45R6L8 Q3MH27 Q45R5D9 Q92643 Q5R6L8 Q3MH27 Q45R5D9 Q92643 Q5R6L8 Q3MH27 Q45R5D9 Q95M12 P49044 P49047 B8ASK4 P49045 Q24326 P49045 Q99017 Q99538 Q95012 P49048 Q8585D9 Q99017 Q99538 Q585D9 Q99017 Q99538 Q585D9 Q99048 Q874E1 Q99048 Q874E1 Q48047 Q5865 Q5875 Q5875 Q49047 Q5875 Q5875 Q49047 Q9505 Q5875 Q49048 Q5875 Q97038 Q99048 Q5875 Q97038 Q99048 Q5875 Q97038 Q95012 Q97038 Q97	I HAY KOLLAY A IO TMFEDN ELTNY IE LADAFEQMWQKR ELTN IE LADAFEQMWQKR ELVKRTSNNTYNTGSH KLYKNRTSNNTYNTGSH KLYKNRTSNNTYNTAGSH GVCPKSLCVSTPGH OVCPKSLCVSTPGH OVCKSLCVSTPGH OVCKSLCVSTPGH OVCPKSLCVSTP	MYQKMVFYLLAG RYHEM/VIADSC RYHELFFMVDTC RYNELFFMVDTC RYNELFFIDTC RYNELFIDTC RYNELFIDTC RYNELFIDTC RYNELFIDTC RYNELFIDTC RYNELFIDTC RYNELFIDTC RYNEFGDIGLSKN VMEYGSKDIGLSK VMEYGQVSK VMEYGQVSK VMEYGQVSK VMEYGQVSK VMEYGQVSK VMEYGQVSK TOLFQD RYNE R	ESGSMMN.H ESGSMMN.H SSASMYEWI QAASLYEKF QGASMYERF	LPPDINUYATT DSPN. VLSIS ISPN. VLAVA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA YSPN. IMALA PATU. NLPA PATU. NLPA PATU. NFPP PATU. NF	AANPRESSYA AANPRESYS SSLTHEESYS SSLTHEESYS SSVGEDSLS SSV	CYYDE YDVD HUPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD S50 350 DKLDGVNQRD RTPSAAVNQRD RTPSAAVNQRD PVPKATVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD DLTPSPD QQD QPD 	QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAVGVH PAVGVH PAUGVH ADLIFL ADLIFL ADLIFL ADLIFK ADLLFM AELLFM AELLFM AELLFM AELLFM VPLTIL VPLTIL VPLTIV VPLTIM VPLTIS SQVVS SSEIMES SSEIMES SSGVVS SSEIMES SCVVS SSGVS CNSGVS CNSGVS CNSGVS CNSGVS CNSGVS CNSGVS CNSGVS	IGD WY VID RYTHYT VID RYTHYT MID RYTFYV IMD RYTFYV WQ RYCRSNF WQMYQRSNF WQMYQRSNF WQMYQRSNF WQMYQRSNF WQMYQRSNF SSK SSK EDC SYK EDC	VNN WMEDSU VNFLTKEV VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VSEDKKTDI IQPEKKTHI IQPEKKTHI IQPEKKTHI IQPEKKTHI IQPEKKTDI VK.ESQVL	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL 380 380 LREITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK QKQVLEAMSHRK LQUKDAIKHRS LKQIEETVKHRK LQUKDAIKHRS LKQIEETVKHRK LQUKDAIKHRS LKQIEETVKHRK LQULABTVKHRK LQUIEDAIKHS LKQIEETVKHRK LKQIETVKHRK LKQIETVKHRK LKQIETVKHRK LKY EKI KY LKY LKY LKY LKY QFYGTFGA.YSV NPWSSLQGFSA ARWSPMTEGGHSA GQWHPTNQGYSA GLHPSNTGYSA GLFHPSNTGYSA GLFHPSNTGYSA
Q30112 Q87481 Q87481 Q92643 Q58618 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49044 Q905018 Q95017 Q99538 Q95K12 P49044 Q95K12 Q95K12 Q95K12 Q4KRV1 A0A0G2R159 P49044 Q92643 Q5R618 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49047 B8ASK4 P49047 B8ASK4 P49045 Q95R0J8 Q880J7 Q95S38 Q5FSD9 Q95M12 P49048 Q874E1 Q9CXY9 Q9CXY9	I. HYKD LNE TIRYMYEHK ELTINYD LAYA AI OTMFEDN ELTIN ELADAFEQMWQKR EITIN IE LADAFEQMWQKR SOQ 310 NKVKKRTSNSNTYNTGSH KLVKNRTI.S.S.EPYGSH ELVKKRTSNSNTYNTGSH KLVKNRTSNANKLNEGSH SSVKARTSNSNSNSYRFGSH KLVKRTSNANKLNEGSH SSVKARTSNNSNSYRFGSH KSVKQRTSNFAYAMGSH QSVKQRTSNFAYAMGSH QSVKQRTSNFAYMGSH LVKSHTNTSH UVKSHTNTSH QUVCPKSLCVSTPGH QVCPKSLCVSTPGH QVCPKSLCVSTP	MYQKMVFYILAG KYHEMLVIADSC RYNELFFMVDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELSFIIDTC RYNELSFIIDTC RYNELSFIIDTC RYNELSFIIDTC RYNERD RYNELSFIIDTC RYNERD RYNE	ESGSMMN.H ESGSMMN.H ESGSMMYEWI QAASLYEWI QAASLYEWI QAASLYEWF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF CONSTRUCTION CONSTRUCT		AANPRESSYA SSLTHESYS SSLTHESYS SSVGEDSLS SSVGE	CYYDE YDV YDV HQP HQP HQP HQP HQP HQP HQP DKLDGVNQRD RTPSAAVNQRD PVPKATVNQRD VPKATVNQRD VPKATVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD EAKMEVVNQRD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD QQD QDD 	QRSTF TDIGVY PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH PAIGVH ADLIF ADL	IGD W Y IGD W Y VID RYTFY Y ID RYTFY Y ID RYTFY Y IMD RYTTY Y IMD RYTY RY IMN RK L	VNN WEDSD VNFLTKEV VNFLTKEV VLEFLEKVQ VLEFLEEVQ VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEN VLEFLEEN VLEFLEEN VLEFLEEN VLEFLEEN VLEFLEEN VLEFLEEN VLEFLEEN VLEFLEEN VLEFLEEN VLEFLEEN VLEFLE VLEFLEEN VLE	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL S.ASQTNMNDL S.ASQTNMNDL S.ASQTNMNDL S.ASQTNMNDL S.ASQTNMNDL S.ASQTNMNDL S.ASQTNMNDL S.ASQTNMNDL S.ASQTNMNDL S.ASQTNMNDL S.ASQTNMNDL S.ASQTNMDL
Q30112 Q43048 Q874E1 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2RI59 P49044 P49047 B8ASK4 P49047 B8ASK4 P49047 B8ASK4 Q95S18 Q55R5D9 Q95M12 P49048 Q3MH27 Q4KRV1 A0A0G2RI59 P49045 Q95M12 P49047 B8ASK4 P49047 B8ASK4 P49047 B8ASK4 P49045 Q9R0J8 Q89S017 Q99538 Q5R5D9 Q980J8 Q898017 Q99538 Q5R5D9 Q95M12 P49045 Q98538 Q955729 Q9538 Q9572 P49048 Q8742	. LHVKDLNETIRYMYEHK . LHVKDLNYAIOTMFEDN ELTSQELADGIQQMWEKK EITSQELADGIQQMWEKK EITSQELADGIQQMWEKK EITSQELADGIQQMWEKK EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIELADAFEQMWQKR EITNIFLADAFEQMWQKR SVKKRTSNSNTYNTGSH KLVKRTSNSNSNSYFSGH KSVKQRTSNFNNYAAGSH QVKRSLTSNFNNYAAGSH BLVKSHTN. SVKRTSNSNSNSYFSGH HLVKSHTN. TSH DSCPARKCLSTNT. GV QVCPKSLCVSTP. GH QVCPKSLCVSTP. GH QVCPKSLCVSTP.	MYQKMVFYLLAG RYHEMUVIADSC RYHELFILJTIC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC RYNELLFILDTC WELLFILDTC WELLFILDTC WEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKN VMEYGDANMTAE VMEYGDANTTAE VMEYGDA	ESGSMMN.H ESGSMMN.H ESAS MYEWI QAASLYEWI QAASLYEWI QAASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF QGASMYERF GFD KUYLYIGFO GFD KUYLYIGFO GFD KUYLYIG GFD KUYLYIG GFD KUYLYIG GFD KUYLYIG GFD KUYLYIG GFD KUYLYIG GFD KUYLYIG GFD KUYLYIG GFD KUYLYIG GFD KUYLYIG GFD KUYLYIG GFD KUYLYIG GFD KUYLYIG GFD KUYL GFD GFD GFD KUYL GFD GFD KUYL GFD GFD GFD KUYL GFD GFD KUYL GFD GFD GFD GFD KUYL GFD GFD GFD GFD GFD GFD GFD GFD		AANPRESSYA AANPRESSYA SSUTHESSYS SSUTEESYS SSUTEESSS SSUTEESSS SSUTEESSS SSUTEESSS SSUTEESSS SSUTEESS SSSUTEESS SSSUTEES	CYYDE YDVD HHVD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD HQPD SSQ DKLDGVNQRD RTPSAAVNQRD PVPKATVNQRD PVPKATVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD EXKMEVVNQRD LDLTPSPD DLTPSPD	QRSTF TDIGVY PSIGVY PAIGVH PAIGVH PAIGVH PAVGVH PAVGVH PAUGVH ADLIFU ADL	IGD W Y IGD W.T.D RYTHYT ID RYTFYV ID RYTFYV IMD RYTFYV WQ RYCRSNF WQUYQRSNF WQUYQRSNF WQUYQRSNF RQUYQRSNF WQUYQRSNF RXKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRK KSKULRTNN KRK KSKULLRTNN	VNN WMEDSU VNFLTKEV VNFLTKEV UNFLTKEV UNFLTKEV UEFLEKVQ UEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEIN VLEFLEEN VSEDKLRA GSEKKKEI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI UPEKKTDI VK ESQVL VK ESQVL VK ESQVL UE ESRU ULE ESRU	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMNDL B.ASQTNMDL B
Q95M12 Q874E1 Q92643 Q92643 Q5R6L8 Q3MH27 Q4KRV1 A0A0G2R159 P49044 P49047 B8ASK4 P49042 Q92643 Q95M12 P49046 P49045 Q95M38 Q874E1 Q95X12 P49048 Q874E1 Q92XY9 Q92643 Q3MH27 Q49047 B8ASK4 P49047 B8ASK4 P49042 Q5R6L8 Q3MH27 Q49047 B8ASK4 P49047 B8ASK4 P49042 Q24326 P49046 P49047 B8ASK4 P49047 B8ASK4 P49047 Q97538 Q97538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q99538 Q575D9 Q92643 Q5	. LHVKDLNY AL OTMFEDN ELTIN VELLAY AL OTMFEDN ELTIN SELLAD GIO OMWEKK EITIN IE LAD AFE QMWQKR EITIN KRT SN YNTYN GSH KLVKRT SN YNTYN AG SH SVK ART SN YNTYN AG SH QSVR KRT SN YNTYA GSH HLVKSHTN	MYQKM VFYLEAG RYHEM/VIADSC RYHELFFM/DTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC RYNELLFIIDTC YMEYGSKDIGLSKN VMEYGDIGLSKN VMEYGDIGLSKE VMEYGDSIKSIST VMEYGDSIKSIST VMQYGDLNITAE VMQYGDLNITAE VMQYGDLNITAE VMQYGDLNITAE VMQYGDLNITAE VMQYGDKSIST VMQYGNK RTDLFQ RSTLLS R	ESGSMMN.H ESGSMMN.H ESGSMMYERF QGASMYERF		AANPRESY AANPRESY SSUTHESYS SSUTESYS SSUTESSS SS	CYYDE YDV YDV HQP BYDKATVNQRD RPPSRVNQRD KSKMEVVNQRD EXKMEVVNQRD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD LDLTPSPD CMKHMRAFANI GMKHMRAFANI GMKHMRAFANI GMKHMRAFANI GMKHMRAFANI GMKHMRAFANI GMKHMRAFANI GMKHMRAFANI GMKHMRAFANI ALRYLYVLANL ALRYLYVLANL ALRYLYVLANL ALRHLYVLVVLANL ALRHLYVLVNL CIKHMKFIF GIKHMKFIF GIKHMKFIF GIKHMKFIF GIKHMKFIF IKHMKFIF CIKHMKFIF	ORSTF TDIGVY PSIGVY PAIGVH PAIGVH PAVGVH PAVGVH ADLIFL ADLIFL ADLIFL ADLIFL ADLIFN AELLFM AELLFM AELLFM AELLFM AELLFM SQVVDS SSEVDS SEIMES SSGVVS SSGVVS SSGVVS SSGVS CNSGVS CNSGVS CNSGVS CNSGVS CNSGVS CNSGVS	IGD W Y IGD W.T.D RYTHYT WID RYTFYV IMD RYTFYV WQ RYK WQRYCSNF WQVYQRSNF WQMYQRSNF WQWYQRSNF WQMYQRSNF KRKLLRTNN KRKLLRTNN KRKLLRTNN KRKLLRTNN KRK KRKLLRTNN KRK KRKLMNTNE KRK KRKLMNTNE KRK KRKLMNTNE KRK KRK KE SYK EDC SYK EDC SYK EDC SYK EDC SYK EDC SYK EDC	VIN WHEDSO VINFLTKEV VINFLTKEV UEFLEKVQ UEFLEEIN VIEFLEEIN VIEFLEEIN VIEFLEEIN VIEFLEEIN VIEFLEEIN VIEFLEEIN VIEFLEEIN VIEFLEEIN VIEFLEEIN VIEFLEEIN VIEFLEEIN VESDKKTDI VESDKKT	VEDLTKETLHKQ KALNSSANMQDY P.FSKRTIGEF P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL P.ASQTNMNDL 380 380 LREITETLTHRG EKQVLEAMSHRK QKQVLEAMSHRK QKQVLEAMSHRK LQUKDAIKHRS LREIEDTIAHRK LQUKDAIKHRS LQUKDAIKHRS LQUKDAIKHRS LQUEDAIKHRS LQUEDAIKHRS LQUEDAIKHS LKY IEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH TEEIQHLDARH KY LKY LKY LKY LKY QFYGTFGA.YSV NPWSSLQGFSA ARWSPMTEGGSSA GQWHPTNQGYSA GLLHPSNTGYSA GLFHPSNRGYSA















