Supporting Information

Buller and Townsend 10.1073/pnas.1221050110

SI Text

Structures Displaying Outlier χ_{cat} Values. ***Clan SE, Family S11: Penicillin-binding protein 5.*** The structure of penicillin-binding protein 5 (PBP5) with a boronic acid inhibitor (PDB ID: 1Z6F) shows the Lys base of its Ser-Lys-Asn catalytic triad approaches from an anomalously low $\chi_{cat} = -66°$ (1). From this angle, the ε-amino group is 4.5 Å from the leaving group and unable to affect proton transfer. This discrepancy between structure and function has been recognized within the literature and an *in silico* QM/MM analysis of PBP5 supports either water penetration or motion of the Lys base closer to the leaving group, which raises $|\chi_{cat}| > 90°$ (2).

Clan SE, Family S13: PBP4. Similar to PBP5, the *apo* structure of PBP4 shows $\chi_{cat} = -73°$ (PDB ID: 1TVF). Further inspection of the electron density within this 2.0-Å structure clearly shows an unidentified citrate ligand has cocrystallized within the active site, which likely distorts the geometry of the catalytic groups. A comparable motion of the Lys base that was uncovered for PBP5 may also apply to PBP4.

Clan SJ, Family S16: Lon protease. Lon proteases function with a Ser-Lys catalytic diad and in the absence of a mechanistically relevant cocrystal structure there is a great degree of structural ambiguity within the active site. The structure of a Ser→Ala inactivated *Escherichia coli* LonA (PDB ID: 1RR9) is similar to that of signal peptidase I (SPI), which functions through a *g*+ rotamer, and $\chi_{cat} = 148°$ (3, 4). Modeling this rotamer for *E. coli* LonA shows $\chi_{cat} = -176°$ with the Lys base 4.4 Å away from the catalytic Ser. Rotation of the Lys base closer also brings χ_{cat} closer to a value consistent with that of other *si*-face–attacking enzymes. This ambiguity was not neglected in the publication of the initial Lon structure and considerable analysis was performed on the stereochemical mechanism of the protease. The authors identified the structural similarity of Lon with other Ser-Lys diad systems (including SPI) and correctly identify the *g*+ rotamer as the reactive conformation. It was also recognized that these Ser-Lys proteases bind their substrates for a *si*-face attack, in contrast to the classic catalytic triad systems, such as subtilisin and sedolisin that use a *re*-face attack. However, the cause of this difference, the link between χ_{cat} and stereochemistry reported here, was not commented upon.

Clan SC, Family S28: Human dipeptidyl peptidase. The 2.45-Å cocrystal structure of human dipeptidyl peptidase (hDPP) with inhibitor GSK237826A was determined (PDB ID: 3N0T), and the His base is positioned with $\chi_{cat} = 80°$. This enzyme contains similar features to its homolog prolyl oligopeptidase, for which catalytically relevant complexes have been trapped. Superimposition of the nucleophile for these structures shows the position of the His base of hDPP is 3.8 Å away from the position of the amine leaving group. We propose rotation of the His base must occur to populate a catalytic orientation.

Clan PA, Family S39: Sesbania mosaic virus (SMV) protease. Clan PA contains seven structurally characterized families of Ser proteases and for one, Sesbania mosaic virus (SMV) protease of family S39 (PDB ID: 1ZYO), the structure shows $\chi_{cat} = 175°$ (5). Further inspection shows the His base is not hydrogen bonded to the Ser nucleophile, which was reported incorrectly in the original publication. We hypothesize that considerable protein motion must occur to generate a reactive state that, consistent with other members of clan PA, moves the base into a position where $\chi_{cat} < -90°$.

Clan SJ, Family S50: Birnavirus VP4 protease. An acyl enzyme structure has been determined for bVP4 protease (PDB ID: 1ZYO), which is a homolog of the Lon proteases (6). This structure clearly establishes *si*-face attack and $\chi_{cat} = 173°$. Although the sign of χ_{cat} is correct, the value itself is unusually large. The simple rotation of χ_4 side-chain angle on the Lys base to 60° decreases the hydrogen bonding distance to the Ser nucleophile and places it closer to a reasonable catalytic orientation.

Clan PC, Family S51: Aspartyl dipeptidase. A single structure for aspartyl dipeptidase, the first structurally characterized member of clan PC, has been solved (PDB ID: 1FYE) and shows $\chi_{cat} = 84°$ (7). This enzyme functions with the *t*-rotamer of its Ser nucleophile, the N + 1 oxyanion hole, and a His base. Although no cocrystal structures have been determined that might suggest an inactive conformation for its His base, superimposition with prolyl oligopeptidase, which has a similar active site architecture and a cocrystal structure with a hemiketal present (PDB 1QFM), provides a good picture of the aspartyl dipeptidase tetrahedral intermediate (8). This model predicts a *si*-face attack and is consistent with the original proposal for the aspartyl dipeptidase mechanism (Fig. S2). As was predicted by the larger trends of our analysis, the His base of aspartyl dipeptidase is 3.6 Å from the position where the nitrogen of the tetrahedral intermediate will be positioned, and some small motion must raise $\chi_{cat} > 90°$ to bridge this gap.

Steric Clashes with a Thr Nucleophile. We identified a handful of *re*-face attacking Ser protease structures for which no steric interaction between the base and a hypothetical γ-methyl of a Thr could be detected. However, all of these structures were previously identified as those that crystallized with the acid/base in an inactive conformation. For example, for PBP5 there is no clash with the Thr γ-methyl but the ε-amino group is 4.5 Å from the leaving group. The same computational evidence that supports either water penetration or motion of the Lys base indicates that these events would be significantly impeded by the γ-methyl of Thr, as shown in Fig. S2 (2).

The bacterial asparaginase (BA) enzymes hydrolyze the terminal amide of asparagine with two separate catalytic triads. The first is a Thr-Tyr-Glu system that forms an acyl-enzyme species on the Thr nucleophile. The second is a Thr-Lys-Asp system that is used to activate a water molecule for the hydrolysis of the acyl-enzyme intermediate. Although the structure of the acyl-enzyme intermediate shows no clash between the γ-methyl and the Tyr base (9–13), it is not clear what role the Tyr base has in the catalytic mechanism. Mutation to Phe results in a 100-fold reduction in k_{cat}, which is small compared with mutation of other catalytic triads. We suggest that protonation of the leaving ammonia group may be performed by the second catalytic triad present in BA, which would leave the system in a state ready for water activation. Alternatively, the leaving group may be protonated through a water-mediated mechanism. In either case, the Tyr residue involved in acylation does not directly protonate the leaving group and is, therefore, able to occupy a position that does not clash with the γ-methyl. Although additional experiments are needed to determine the mechanism of BA, it is clear that its catalytic reaction is subject to weaker geometric constraints that enable use of a Thr nucleophile.

Finally, it was hypothesized on the basis of a series of X-ray structures that fluoroacetyl CoA thioesterase functions with a Thr nucleophile and a His base, but no Michaelis-like or covalently linked structures were determined (14). Suspiciously, this system showed $\chi_{cat} = -53°$, an unusually low value compared with that found for the proteases. Further reading of the literature identified a second mechanistic hypothesis involved anhydride formation with a Glu residue (15). Although both hypotheses

seem structurally reasonable, we await further experimentation that clarifies the mechanism of this unusual enzyme.

Stereospecific Antibiotics. In the course of this investigation we have placed great significance upon the stereochemical mechanism of proteolysis. Although our analysis itself shows the utility of this distinction in understanding the structure–function relationship, it is also pertinent to identify whether nature is sensitive to this distinction. We have extensively studied the biosynthesis of *5R*-carbapenem and it is well known that the bridgehead stereochemistry necessary for antibiotic activity forces the 5-membered ring to occlude the *si* face of the molecule (16). Because the penicillin-binding proteins function through a *re*-face attack, this does not impair their function. This configuration is common among all bioactive bicyclic β-lactam antibiotics (such as penicillin, thienamycin, and clavulanic acid), and the enzymes that generate this stereochemistry have been selected for independently at least four different times (17). The salinosporamides and cinnabaramides are a class of naturally occurring compounds that, similar to omuralide, inhibit the proteaseome (18, 19). These inhibitors share a β-lactone–γ-lactam core, and future studies are needed to determine the evolutionary relationship of these natural products.

These molecules represent the best-characterized, naturally occurring, stereospecific inhibitors. Other compounds that share these properties have also been described, such as vibralactone, a β-lactone pancreatic lipase inhibitor (20). Spongiolactone is a β-lactone marine natural product whose target is not known (21). If this compound functions through forming an acyl-enzyme intermediate, our analysis predicts that the target will be a *si*-face–attacking enzyme.

1. Nicola G, et al. (2005) Crystal structure of Escherichia coli penicillin-binding protein 5 bound to a tripeptide boronic acid inhibitor: A role for Ser-110 in deacylation. *Biochemistry* 44(23):8207–8217.
2. Shi Q, Meroueh SO, Fisher JF, Mobashery S (2008) Investigation of the mechanism of the cell wall DD-carboxypeptidase reaction of penicillin-binding protein 5 of Escherichia coli by quantum mechanics/molecular mechanics calculations. *J Am Chem Soc* 130(29):9293–9303.
3. Botos I, et al. (2004) The catalytic domain of Escherichia coli Lon protease has a unique fold and a Ser-Lys dyad in the active site. *J Biol Chem* 279(9):8140–8148.
4. Paetzel M, Dalbey RE, Strynadka NCJ (1998) Crystal structure of a bacterial signal peptidase in complex with a beta-lactam inhibitor. *Nature* 396(6707):186–190.
5. Gayathri P, et al. (2006) Crystal structure of the serine protease domain of Sesbania mosaic virus polyprotein and mutational analysis of residues forming the S1-binding pocket. *Virology* 346(2):440–451.
6. Chung IYW, Paetzel M (2011) Crystal structure of a viral protease intramolecular acyl-enzyme complex: Insights into cis-cleavage at the VP4/VP3 junction of Tellina birnavirus. *J Biol Chem* 286(14):12475–12482.
7. Håkansson K, Wang AHJ, Miller CG (2000) The structure of aspartyl dipeptidase reveals a unique fold with a Ser-His-Glu catalytic triad. *Proc Natl Acad Sci USA* 97(26): 14097–14102.
8. Fülöp V, Böcskei Z, Polgár L (1998) Prolyl oligopeptidase: An unusual beta-propeller domain regulates proteolysis. *Cell* 94(2):161–170.
9. Derst C, Henseling J, Röhm KH (1992) Probing the role of threonine and serine residues of E. coli asparaginase II by site-specific mutagenesis. *Protein Eng* 5(8):785–789.
10. Derst C, Wehner A, Specht V, Röhm KH (1994) States and functions of tyrosine residues in Escherichia coli asparaginase II. *Eur J Biochem* 224(2):533–540.
11. Palm GJ, et al. (1996) A covalently bound catalytic intermediate in Escherichia coli asparaginase: Crystal structure of a Thr-89-Val mutant. *FEBS Lett* 390(2):211–216.
12. Aung HP, Bocola M, Schleper S, Röhm KH (2000) Dynamics of a mobile loop at the active site of Escherichia coli asparaginase. *Biochim Biophys Acta* 1481(2):349–359.
13. Yun M-K, Nourse A, White SW, Rock CO, Heath RJ (2007) Crystal structure and allosteric regulation of the cytoplasmic Escherichia coli L-asparaginase I. *J Mol Biol* 369(3):794–811.
14. Dias MVB, et al. (2010) Structural basis for the activity and substrate specificity of fluoroacetyl-CoA thioesterase FlK. *J Biol Chem* 285(29):22495–22504.
15. Weeks AM, Coyle SM, Jinek M, Doudna JA, Chang MCY (2010) Structural and biochemical studies of a fluoroacetyl-CoA-specific thioesterase reveal a molecular basis for fluorine selectivity. *Biochemistry* 49(43):9269–9279.
16. Stapon A, Li RF, Townsend CA (2003) Synthesis of (3S,5R)-carbapenam-3-carboxylic acid and its role in carbapenem biosynthesis and the stereoinversion problem. *J Am Chem Soc* 125(51):15746–15747.
17. Bodner MJ, et al. (2011) Definition of the common and divergent steps in carbapenem β-lactam antibiotic biosynthesis. *ChemBioChem* 12(14):2159–2165.
18. Feling RH, et al. (2003) Salinosporamide A: A highly cytotoxic proteasome inhibitor from a novel microbial source, a marine bacterium of the new genus salinospora. *Angew Chem Int Ed Engl* 42(3):355–357.
19. Rachid S, et al. (2011) Mining the cinnabaramide biosynthetic pathway to generate novel proteasome inhibitors. *ChemBioChem* 12(6):922–931.
20. Liu D-Z, et al. (2006) Vibralactone: a lipase inhibitor with an unusual fused beta-lactone produced by cultures of the basidiomycete Boreostereum vibrans. *Org Lett* 8(25):5749–5752.
21. Mayol L, Piccialli V, Sica D (1987) Spongiolactone, an unusual beta-lactone diterpene isovalerate based on a new rearranged spongiane skeleton from Spongionella-gracilis. *Tetrahedron Lett* 28(31):3601–3604.

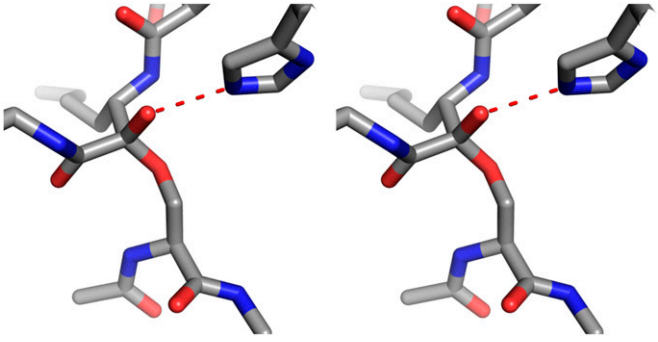

Fig. S1. Narlaprevir binding to cytomegaloviral protease. This structure shows the noncatalytically relevant interaction between the acid moiety and oxyanion when χ_{cat} and χ_{oxy} have the same sign upon substrate binding. The oxyanion must be shielded from the acid for efficient catalysis, which is observed in all hydrolases, regardless of whether the local backbone is used to facilitate catalysis.

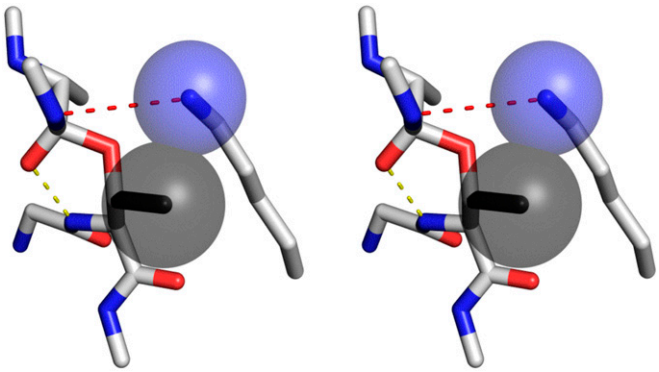

Fig. S2. Steric clash with Thr and the Lys base of PBP1b. A computational model of the first tetrahedral intermediate shows a small steric clash between the Lys base and γ-methyl of Thr. This repulsive interaction would increase as the acid moves from the current 3.8 Å to within hydrogen-bonding distance of the amine leaving group. Spheres are drawn with corresponding van der Waals radii. Yellow dashes indicated the oxyanion hole interaction. Red dashes are drawn between the proton donor and the acceptor.

Table S1. Active site parameters of representative structures from each Ser protease family

Clan	Family	Enzyme	PDB	Complex	Reactive rotamer	ϕ, ψ angles	Catalytic motif	Oxyanion hole	Face of attack	χ_{cat}	α_{cat}	χ_{oxy}	Reference
PA	S1	Elastase	3HGN	Ketal	*g*−	(−42, 140)	S-H-D	N	*re*	−131	125	−32	(1)
PA	S3	Sindbis viral protease	2SNV	C terminus	*g*−	(−54, 119)	S-H-D	N	*re*	−155	85*		(2)
PA	S6	TSH protease	3AK5		*g*−	(−46, 124)	S-H-D	N	*re*	−150			(3)
PA	S7	Dengue virus protease	3U1I	Aldehyde	*g*−	(−64,139)	S-H-D	N	*re*	−177	90*		(4)
SB	S8	Subtilisin Carlsberg	1SCN	Carbamate	*g*−	(−53, −33)	S-H-D	N	*re*	−161	117	0	(5)
SE	S11	Penicillin-binding protein 5	1Z6F	Boronic acid	*g*−	(−66, −5)	S-K-N	N	*re*	−66*	127	−24	(6)
SE	S12	D-Ala-D-Ala carboxypeptidase B	1MPL	Phosphonate	*g*−	(−72, 0)	S-Y-K	N	*re*	−156	97	−49	(7)
SE	S13	Penicillin-binding protein 4	1TVF		*g*−	(−66, −4)	S-S-K	N	*re*	−73			(8)
PA	S29	Hepacivirin (hepatitis C virus)	3LON	α-ketoamide	*g*−	(−57, 131)	S-H-D	N	*re*	−150			(9)
PA	S32	Arterivirus protease	1MBM		*g*−	(−57, 135)	S-H-D	N	*re*	−153			(10)
PA	S39	Sesbania mosaic virus protease	1ZYO		*g*−	(−71, 134)	S-H-D	N	*re*	175*			(11)
SB	S53	Sedolisin	1GA4	Acyl	*g*−	(−62, −24)	S-E-D	N	*re*	−147	108	−18	(12)
SC	S9	Prolyl oligopeptidase	1QFM	Hemiketal	*t*	(66, −115)	S-H-D	N + 1	*si*	117	106	32	(13)
SC	S10	Carboxypeptidase Y	1WPX		*t*	(61, −120)	S-H-D	N + 1	*si*	104			(14)
SK	S14	*E. coli* ClpP	2FZS	CMK	*t*	(50, −126)	S-H-D	N + 1	*si*	137	74*	50	(15)
SC	S15	X-Prolyl dipeptidyl aminopeptidase	1LNS		*t*	(64, −123)	S-H-D	N + 1	*si*	106			(16)
SH	S21	CMV protease	1NKM	α-ketoamide	*t*	(−117,131)	S-H-H	Extrinsic	*re*	−125	111	−39	(17)
SC	S28	Human dipeptidil peptidase	3N0T		*t*	(62, −26)	S-H-D	N + 1	*si*	80			(18)
SC	S33	Prolyl aminopeptidase	2EEP	Boronic acid	*t*	(78, −120)	S-H-D	N + 1	*si*	99	107	39	(19)
SK	S41	Tricorn protease	1N6E	CMK	*t*	(60, −117)	S-H-S-D	N + 1	*si*	130	84*		(20)
SK	S49	Bacterial signal peptidase A	3BF0		*t**	(52, −109)	S-K	N + 1	*si*	116			(21)
PC	S51	Aspartyl dipeptidase	1FYE		*t*	(57, −122)	S-H-E	N + 1	*si*	84			(22)
SS	S66	LD-carboxypeptidase	1ZRS		*t*	(64, −129)	S-H-E	N + 1	*si*	99			(23)
SJ	S16	Lon-A protease *Methanocaldococcus jannaschii*	1XHK		*g*+	(-68, −7)	S-K-D	N	*si*	171			(24)
SF	S24	UmuD	1UMU		*g*+	(−67, −8)	S-K	N	*si*	155			(25)
SF	S26	Signal peptidase I	1B12	Ester	*g*+	(−68, −10)	S-K	N	*si*	148	109	1	(26)
SJ	S50	VP4 protease	3R0B	Ester	*g*+	(−62, −15)	S-K	N	*si*	173	110	−9	(27)
ST	S54	Rhomboid-1	2IC8		*g*+	(−63, −29)	S-H	N	*si*	113			(28)

*The structure is clearly perturbed from a catalytic state through either inactivation or crystallization. See *SI Text* for more details.

1. Tamada T, et al. (2009) Combined high-resolution neutron and X-ray analysis of inhibited elastase confirms the active-site oxyanion hole but rules against a low-barrier hydrogen bond. *J Am Chem Soc* 131(31):11033–11040.
2. Tong L, Wengler G, Rossmann MG (1993) Refined structure of Sindbis virus core protein and comparison with other chymotrypsin-like serine proteinase structures. *J Mol Biol* 230(1):228–247.
3. Nishimura K, et al. (2010) Role of domains within the autotransporter Hbp/Tsh. *Acta Crystallogr D Biol Crystallogr* 66(Pt 12):1295–1300.
4. Schiering N, et al. (2011) A macrocyclic HCV NS3/4A protease inhibitor interacts with protease and helicase residues in the complex with its full-length target. *Proc Natl Acad Sci USA* 108(52):21052–21056.
5. Steinmetz ACU, Demuth HU, Ringe D (1994) Inactivation of subtilisin Carlsberg by N-((tert-butoxycarbonyl)alanylprolylphenylalanyl)-O-benzoylhydroxyl- amine: formation of a covalent enzyme-inhibitor linkage in the form of a carbamate derivative. *Biochemistry* 33(34):10535–10544.
6. Nicola G, et al. (2005) Crystal structure of Escherichia coli penicillin-binding protein 5 bound to a tripeptide boronic acid inhibitor: A role for Ser-110 in deacylation. *Biochemistry* 44(23):8207–8217.
7. Silvaggi NR, Anderson JW, Brinsmade SR, Pratt RF, Kelly JA (2003) The crystal structure of phosphonate-inhibited D-Ala-D-Ala peptidase reveals an analogue of a tetrahedral transition state. *Biochemistry* 42(5):1199–1208.
8. Rajashankar K, et al. (2009) Crystal structure of penicillin-binding protein 4 (PBP4) from Staphylococcus aureus. Available at Protein Data Bank (www.pdb.org). Accessed September 28, 2012.
9. Arasappan A, et al. (2010) Discovery of Narlaprevir (SCH 900518): A Potent, Second Generation HCV NS3 Serine Protease Inhibitor. *Acs Medicinal Chemistry Letters* 1(2):64–69.
10. Barrette-Ng IH, et al. (2002) Structure of arterivirus nsp4. The smallest chymotrypsin-like proteinase with an alpha/beta C-terminal extension and alternate conformations of the oxyanion hole. *J Biol Chem* 277(42):39960–39966.
11. Gayathri P, et al. (2006) Crystal structure of the serine protease domain of Sesbania mosaic virus polyprotein and mutational analysis of residues forming the S1-binding pocket. *Virology* 346(2):440–451.
12. Wlodawer A, et al. (2001) Carboxyl proteinase from Pseudomonas defines a novel family of subtilisin-like enzymes. *Nat Struct Biol* 8(5):442–446.
13. Fülöp V, Böcskei Z, Polgár L (1998) Prolyl oligopeptidase: An unusual beta-propeller domain regulates proteolysis. *Cell* 94(2):161–170.
14. Mima J, et al. (2005) Structure of the carboxypeptidase Y inhibitor IC in complex with the cognate proteinase reveals a novel mode of the proteinase-protein inhibitor interaction. *J Mol Biol* 346(5):1323–1334.
15. Szyk A, Maurizi MR (2006) Crystal structure at 1.9A of E. coli ClpP with a peptide covalently bound at the active site. *J Struct Biol* 156(1):165–174.
16. Rigolet P, Mechin I, Delage MM, Chich JF (2002) The structural basis for catalysis and specificity of the X-prolyl dipeptidyl aminopeptidase from Lactococcus lactis. *Structure* 10(10):1383–1394.

17. Khayat R, et al. (2003) Structural and biochemical studies of inhibitor binding to human cytomegalovirus protease. *Biochemistry* 42(4):885–891.
18. Dobrovetsky E, et al. (2010) Human dipeptidyl peptidase DPP7. Available at Protein Data Bank (www.pdb.org). Accessed September 28, 2012.
19. Xu Y, et al. (2008) Novel inhibitor for prolyl tripeptidyl aminopeptidase from Porphyromonas gingivalis and details of substrate-recognition mechanism. *J Mol Biol* 375(3):708–719.
20. Kim JS, et al. (2002) Navigation inside a protease: substrate selection and product exit in the tricorn protease from Thermoplasma acidophilum. *J Mol Biol* 324(5):1041–1050.
21. Paetzel M, Dalbey RE, Strynadka NCJ (1998) Crystal structure of a bacterial signal peptidase in complex with a beta-lactam inhibitor. *Nature* 396(6707):186–190.
22. Håkansson K, Wang AHJ, Miller CG (2000) The structure of aspartyl dipeptidase reveals a unique fold with a Ser-His-Glu catalytic triad. *Proc Natl Acad Sci USA* 97(26):14097–14102.
23. Korza HJ, Bochtler M (2005) Pseudomonas aeruginosa LD-carboxypeptidase, a serine peptidase with a Ser-His-Glu triad and a nucleophilic elbow. *J Biol Chem* 280(49):40802–40812.
24. Im YJ, et al. (2004) The active site of a lon protease from Methanococcus jannaschii distinctly differs from the canonical catalytic Dyad of Lon proteases. *J Biol Chem* 279(51):53451–53457.
25. Peat TS, et al. (1996) Structure of the UmuD′ protein and its regulation in response to DNA damage. *Nature* 380(6576):727–730.
26. Kim AC, Oliver DC, Paetzel M (2008) Crystal structure of a bacterial signal Peptide peptidase. *J Mol Biol* 376(2):352–366.
27. Chung IYW, Paetzel M (2011) Crystal structure of a viral protease intramolecular acyl-enzyme complex: Insights into cis-cleavage at the VP4/VP3 junction of Tellina birnavirus. *J Biol Chem* 286(14):12475–12482.
28. Wang Y, Zhang Y, Ha Y (2006) Crystal structure of a rhomboid family intramembrane protease. *Nature* 444(7116):179–180.