Intrinsic evolutionary constraints on protease structure, enzyme acylation, and the identity of the catalytic triad

Andrew R. Buller and Craig A. Townsend[1]

Departments of Biophysics and Chemistry, The Johns Hopkins University, Baltimore MD 21218

The study of proteolysis lies at the heart of our understanding of biocatalysis, enzyme evolution, and drug development. To understand the degree of natural variation in protease active sites, we systematically evaluated simple active site features from all serine, cysteine and threonine proteases of independent lineage. This convergent evolutionary analysis revealed several interrelated and previously unrecognized relationships. The reactive rotamer of the nucleophile determines which neighboring amide can be used in the local oxyanion hole. Each rotamer–oxyanion hole combination limits the location of the moiety facilitating proton transfer and, combined together, fixes the stereochemistry of catalysis. All proteases that use an acyl-enzyme mechanism naturally divide into two classes according to which face of the peptide substrate is attacked during catalysis. We show that each class is subject to unique structural constraints that have governed the convergent evolution of enzyme structure. Using this framework, we show that the γ-methyl of Thr causes an intrinsic steric clash that precludes its use as the nucleophile in the traditional catalytic triad. This constraint is released upon autoproteolysis and we propose a molecular basis for the increased enzymatic efficiency introduced by the γ-methyl of Thr. Finally, we identify several classes of natural products whose mode of action is sensitive to the division according to the face of attack identified here. This analysis of protease structure and function unifies 50 y of biocatalysis research, providing a framework for the continued study of enzyme evolution and the development of inhibitors with increased selectivity.

peptidase | antibiotic stereochemistry | N-terminal nucleophile

The acceleration of chemical reactions is essential to all biochemistry. The generation of reactive species by enzymes is tightly controlled and limited by the chemical makeup of the 20 proteinogenic amino acids. In isolation, no protein residue harbors a strong nucleophile at physiological pH. Rather, enzymatic activity arises after protein folding introduces cooperative interactions that selectively amplify the reactivity of their otherwise weakly active functional groups. The preeminent system through which this phenomenon has been studied is the Ser-His-Asp catalytic triad of the Ser proteases (1). Several distinct molecular strategies have been identified that contribute to rate enhancement. Substrate binding increases the local concentration of interacting species, general acids and bases facilitate proton transfer, and a high-energy step can be split into multiple lower-energy steps (2). These distinct molecular strategies may be unified through the formalism of transition-state theory (3) and the idea that the active site is electrostatically preorganized for transition-state stabilization (4, 5). Although the biophysics of rate acceleration are intricate and sensitive to even minor structural perturbations, evolution has converged on a catalytic triad (or diad) with a reactive Ser, Cys, or Thr nucleophile more than 25 separate times to facilitate central biochemical reactions such as hydrolysis (6), transacylation (7), and phosphorylation (8).

Our understanding of biochemistry has progressed into an increasingly complex picture but many basic questions about enzyme evolution remain unanswered. Because evolution operates through random forces, rationalizing why a particular outcome occurs is a difficult challenge. For example, the hydroxyl nucleophile of a Ser protease was swapped for the thiol of Cys at least twice in evolutionary history (9). However, there is not a single example of Thr naturally substituting for Ser in the protease catalytic triad, despite its greater chemical similarity (9). Instead, the Thr proteases generate their N-terminal nucleophile through a posttranslational modification: *cis*-autoproteolysis (10, 11). These facts constitute clear evidence that there is a strong selective pressure against Thr in the catalytic triad that is somehow relieved by *cis*-autoproteolysis. Because the catalytic triad is the principal system through which enzymatic rate acceleration has been studied, the inability to rationalize the evolutionary selection of the catalytic nucleophile is a fundamental weakness in our understanding of biocatalysis. Motivated by this simple, yet unexplained structure–function relationship, we sought to answer the question: How do the chemical and structural constraints of proteolysis combine to shape the evolution of a protease active site?

Two strategies can be undertaken to probe the properties of diverse catalytic architectures. The first is the contemporary experimental approach whereby all parameters of interest in an active site are systematically varied and analyzed *in silico* for catalytic viability. This technique was used by Smith et al. to show that the catalytic triad of esterases adopts a consensus geometry that minimizes reorganization during the multistep mechanism (12). Computational explorations are unsurpassed in their detail and broad applicability, but hindered by their need to explicitly anticipate the structural and mechanistic diversity generated through natural selection. We elected to pursue an alternative strategy to elucidate the constraints that guide active site evo-

Significance

The structure–function relationship of proteases is central to our understanding of biochemistry. Nature has evolved at least 23 independent solutions to this problem, using an acylation mechanism. We examined the structures of these proteases, using a new framework to characterize the geometric relationships within each active site. This analysis revealed the orientation of the base determines the stereochemistry of catalysis and elucidated why threonine does not substitute for serine in the catalytic triad. These observations explain how the absolute stereostructures of natural protease inhibitors prevent off-target inhibition and serve as boundary conditions to enzyme design.

lution that is not computational, but observational. Studying the large set of naturally occurring structures limits analyses to coarser features, but greatly increases the universality of the results. This technique requires only a diverse set of enzyme structures and a rudimentary understanding of organic and biochemistry. By recognizing the constraints imposed by bond geometries and steric clashes, it is easy to compare the potential catalytic architectures for a given reaction.

Results from the veritable explosion of gene sequencing and protein structure determination have identified at least 23 folds related through convergent evolution that support proteolysis through an acyl-enzyme intermediate (9). We cross-compared the active site architecture among these "analogous" enzymes (13), each of which has been optimized for catalysis through natural selection, to interrogate the entire known structural diversity that supports efficient proteolysis. We first identified key parameters that define active site architecture, such as the orientation of substrate binding, the φ, ψ angles of the nucleophile, and the stereochemistry of catalysis. Analysis of all Ser, Cys, and Thr proteases of independent lineage showed they were all highly interrelated. By rationalizing these correlations between analogous enzymes in terms of their underlying chemistry, we describe how the function of proteases has guided the evolution of their structure. These data show that proteases naturally divide into two mechanistically distinct classes that are inhibited by separate groups of naturally occurring antibiotics. By identifying prohibitive steric interactions of the Thr γ-methyl within each class, we explain why evolution selected Ser, and not Thr, as the nucleophile of the catalytic triad.

Results and Discussion

Chemical Constraints of Proteolysis. To understand how the chemical constraints of proteolysis affect the evolution of a protease active site, it is necessary to invoke the chemical principles that underlie its mechanism. As the present work is not a review of the protease literature, we limit this discussion to only the most central facts (Fig. 1). The Ser hydroxyl group is not itself strongly nucleophilic and enzymes must supply a general base to facilitate its deprotonation. The traditional residue for this role is His, although other residues are also used (6, 14). Nucleophilic attack into a peptide bond generates an oxyanion intermediate that is stabilized by the "oxyanion hole" composed of two hydrogen bond donors, one of which typically comes from a backbone NH flanking the nucleophilic residue (14, 15). The nitrogen of the scissile bond becomes basic upon formation of the tetrahedral intermediate and proton transfer to the amine leaving group occurs concomitantly with collapse of the oxyanion, yielding the acyl-enzyme intermediate. Subsequent hydrolysis of this enzyme-bound species also proceeds through a tetrahedral intermediate. Although
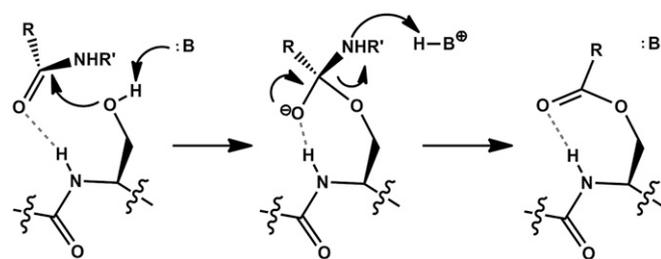

Fig. 1. The acylation mechanism of a Ser protease. After substrate binding, the Ser side chain attacks the scissile peptide bond to generate a tetrahedral oxyanion intermediate. Protonation of the amine leaving group allows collapse of the intermediate and formation of an acyl-enzyme species. Dashed bonds represent the hydrogen bond of the oxyanion hole. See text for details.

the energetics of hydrolysis are different from enzyme acylation, they share many structural constraints. As is shown in this investigation, it is sufficient to consider only the acylation step of the reaction, as displayed in Fig. 1, to rationalize much of the observable convergent evolution of proteases.

Evolution has selected for many different residues to activate the nucleophile, but the cooperative interactions that engender reactivity are effectively "tuned" differently for each individual enzyme. This principle was first tested by the chemical swapping of the hydroxyl nucleophile of subtilisin for a thiol and observing loss of proteolytic activity (16, 17). More exhaustive studies using site-directed mutagenesis of both subtilisin and trypsin showed that alteration of any component of the catalytic triad results in a large decrease in enzyme efficiency (18, 19). Interestingly, the Ser→Thr variant of trypsin created a worse enzyme than simultaneously mutating all three residues of the catalytic triad. Conversely, when the N-terminal Thr nucleophile of the proteasome β-subunit was mutated to Ser, there was a reduction in its catalytic power (20). The juxtaposition of these results and the inability to rationalize the opposite effects of the Thr γ-methyl further stimulated our desire to understand how evolution follows the chemical constraints of proteolysis.

Reactive Rotamer Determines the Identity of the Local Oxyanion Hole. To our knowledge, the extent of active site variation among different proteases has not been systematically explored. Relying on the clan/family nomenclature developed in the MEROPS database (9) and several excellent reviews (6, 14, 21, 22), we used a simple strategy to analyze the existing collection of protease structures in the Protein Data Bank (PDB). Because our investigation is based upon structural analysis of actual proteases, the residues facilitating this chemistry are evolutionarily optimized for catalysis. Cross-comparison of these structures enables us to probe the amount of structural diversity present when the biophysical constraints for rate acceleration are met and ask how this diversity relates to the underlying chemistry.

There are three staggered rotamers of Ser and they are denoted according to whether the γOH is *gauche* or *trans* to its peptide NH (Fig. 2A). Table 1 shows that all three staggered rotamers have been exploited for enzyme acylation and we denote the state from which chemistry arises as the "reactive rotamer." Table 1 also shows that all but one clan of Ser proteases (23) use an amide NH flanking the nucleophile to stabilize the oxyanion hole. It has been established that utilization of a local oxyanion hole moiety forces a suboptimal geometry of oxyanion stabilization (24, 25). Although the geometry of the oxyanion holes is well understood, it was proposed that the identity of the oxyanion hole "... is a direct consequence of the handedness of the catalytic triad" (ref. 26, p. 165). "Handedness" refers to the direction from which the base approaches the nucleophile, but is a limited descriptor because it cannot be extended to the Thr proteases. We introduce a term that both encompasses the handedness of the active site and is generalizable: the dihedral angle between the side chain of the nucleophile and the basic atom, designated χ_{cat} (Fig. 2B). Contrary to the earlier hypothesis, the identity of the oxyanion hole is not determined by the handedness of the catalytic triad, but is correlated with the reactive rotamer. When either of the *gauche* (g+ or g−) rotamers is used, it is the amide NH of the nucleophilic residue that stabilizes the oxyanion intermediate (Fig. 2 D–H). When the *trans* (t) rotamer is used, it is the amide NH of the residue C-terminal to the nucleophile (denoted the N + 1 residue) that is accessible (Fig. 2 F and I). The molecular basis for this relationship is simple: Only the backbone NH that is adjacent to the nucleophile can be used to facilitate catalysis.

Natural Limits on the Location of the Proton Transfer Residue. We observed that the identity of the reactive rotamer is tightly cor-
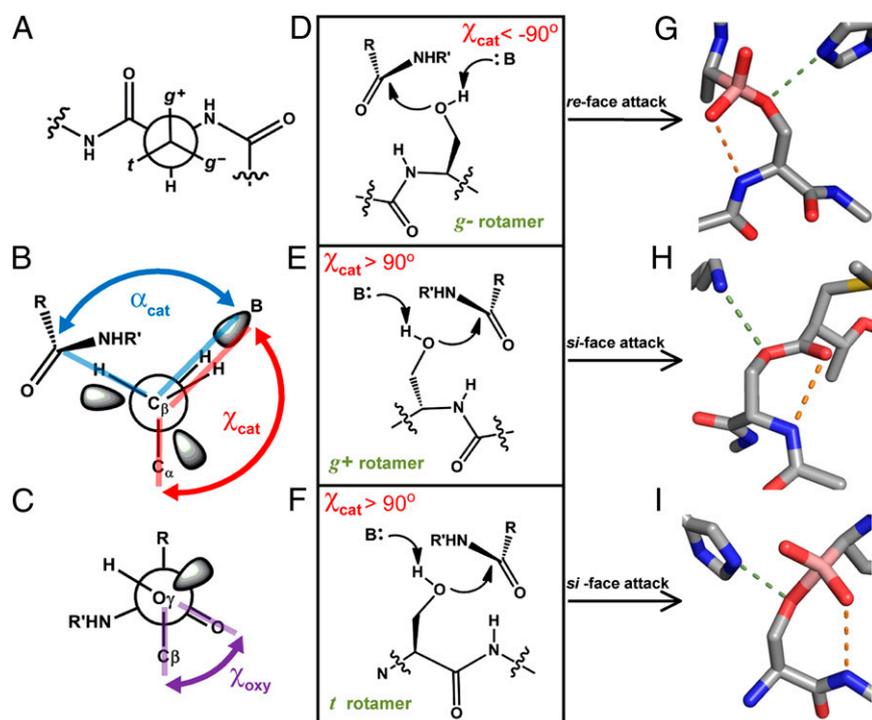
Fig. 2. Active site architectures of Ser Proteases. (*A*) Newman projection down the C_β-C_α bond of a side chain on a peptide. The rotamers of the γ-atom are named $g+$, $g-$, or t, as shown. (*B*) Newman projection down the C_β-O_γ bond of a Ser nucleophile. The dihedral angle χ_{cat} is measured between C_α, C_β, O_γ, and the base (B). α_{cat} is defined as the angle between the base, O_γ, and carbonyl carbon. (*C*) The dihedral angle between C_β, O_γ of the nucleophile, and the carbonyl of the scissile bond, χ_{oxy}, is used to measure the relative orientation of the substrate. (*D-F*) Schematic of the three possible active site arrangements when a local oxyanion hole is used. (*G–I*) Representative examples of a Ser protease with each reactive rotamer. Dashes represent hydrogen bonds between the nucleophile and the base (green) or the protein backbone and substrate (orange). PDB IDs: *G*, 3VGC, γ-chymotrypsin with a boronic acid inhibitor; *H*, 2EEP, signal peptidase I with a β-lactam inhibitor; *I*, 1B12, prolyl aminopeptidase with a boronic acid inhibitor.

related with the location of the base, as indicated by χ_{cat} (Table 1). The only exception to this correlation is clan SH, which does not use the local oxyanion hole. Therefore, we hypothesized that constraints introduced by the reactive rotamer and oxyanion hole combine to severely limit the catalytically viable location of the base. For proteases that share a common reactive rotamer, there is ~50° variability in χ_{cat}. Some of this variation may be due to the fact that conformational heterogeneity in the ground state can hide features that are used to stabilize the transition state(s) (12).

We identified a common limit, $|\chi_{cat}| > 90°$, which is shared among the vast majority of protease structures. To explore the basis of this limit and whether divergent evolution may have altered active site architecture, we extended our analysis to every structurally characterized Ser protease family (Table S1). We observed that all outliers to the trend $|\chi_{cat}| > 90°$ also contained features that indicated an inactive state had been crystallized (*SI Text*). Hence, $|\chi_{cat}| > 90°$ is a naturally occurring constraint in protease architecture. A partial explanation for this limit may be

Table 1. Active site parameters of convergently related Ser proteases

Clan	Enzyme	PDB	Complex	Reactive rotamer	ϕ, ψ angles	Catalytic motif	Oxyanion hole	Face of attack	χ_{cat}	α_{cat}	χ_{oxy}	Reference
PA	Elastase	3HGN	Hemiketal	$g-$	(−43, 140)	S-H-D	N	*re*	−131	125	−32	(68)
SB	Sedolisin	1GA4	Ester	$g-$	(−62, −24)	S-E-D	N	*re*	−147	108	−18	(69)
SE	D-Ala-D-Ala peptidase	1MPL	Phophonate	$g-$	(−72, 0)	S-Y-K	N	*re*	−156	97	−48	(70)
SC	Prolyl aminopeptidase	1QFM	Hemiketal	t	(66, −115)	S-H-D	N + 1	*si*	117	106	32	(26)
SS	L,D-carboxypeptidase	1ZRS		t	(64, −130)	S-H-E	N + 1	*si*	109			(49)
PC	Aspartyl dipeptidase	1FYE		t	(57, −122)	S-H-E	N + 1	*si*	84			(71)
SH	CMV protease	1NKM	α-Ketoamide	t	(−177, 132)	S-H-H	Extrinsic	*re*	−125	111	−39	(23)
SK	ClpP	1FZS	HMK	t	(59, −122)	S-H-D	N + 1	*si*	130	74	50	(29)
SF	Signal peptidase	1B12	Ester	$g+$	(−68, −11)	S-K	N	*si*	148	109	1	(60)
SJ	VP4 protease	3R0B	Ester	$g+$	(−65, −13)	S-K	N	*si*	173	114	−9	(72)
ST	Rhomboid-1	2IC8		$g+$	(−63, −29)	S-H	N	*si*	113			(73)

Halomethyl ketone (HMK) inactivators give rise to unusually small α_{cat} values. See text for details.

attributed to the protein backbone flanking the nucleophile, which enforces an insurmountable barrier to an approach from more acute angles. Although additional work will be needed to ascertain the exact origin of this constraint on χ_{cat}, the fact that it occurs within all Ser proteases marks it as an intrinsic consequence of their activity.

A second constraint on the location of the base is the angle defined by the basic atom, nucleophile, and electrophile, designated α_{cat} (Fig. 2B). With respect to deprotonation, $\alpha_{cat} \sim 116°$ because the γO of the nucleophile is sp^3 hybridized during the concerted deprotonation and nucleophilic attack that initiates catalysis (27). Because this is inherent to proteolysis with Ser, all such proteases must converge on this geometry. Inspection of Table 1 shows this parameter is faithfully reported by most structures with a mechanism-based analog. An interesting exception is the family of halomethyl ketone (HMK) inactivators, which co-valently link to both the nucleophile and the base and show $\alpha_{cat} \sim 75°$, distorting the system that has moved away from its catalytic orientation (28–30).

These simple geometric constraints combine to define the catalytic architecture for all proteases that use the local oxyanion hole. For each reactive rotamer, only one candidate backbone amide can contribute to the oxyanion hole. Because the substrate and base must both be adjacent to the nucleophile, this limitation on the position of the substrate translates into a restriction on the location of the base. Consequently, χ_{cat} must be positive for the $g+$ and t rotamers and negative for the $g-$ rotamer, as was experimentally observed (Fig. 2 D–F). The location of the base is further defined by the requirements that $|\chi_{cat}| > 90°$ and $\alpha_{cat} \sim 116°$. Although additional constraints are imposed by the nuances of transition state stabilization to define the reaction trajectory, these fundamental constraints are sufficient to explain the convergent evolution of all proteases that use a local oxyanion hole.

Ser Proteases Divide into Two Mechanistic Classes. The most conspicuous feature of Table 1 is the interrelationship between χ_{cat} and the orientation of substrate binding. In principle, a peptide substrate can bind to present either the re or the si face of the scissile bond. For re-face attacking enzymes, our measurements show that χ_{cat} is always negative. Moreover, measurement of the dihedral angle between C_β and O_γ of the nucleophile and the carbonyl electrophile (χ_{oxy}, Fig. 2C) shows these enzymes bind their substrate in a similar, staggered, position. When the local backbone is used in oxyanion hole formation, only the $g-$ rotamer is observed to support re-face attack. Conversely, χ_{cat} is always positive for the si-face–attacking enzymes and only the $g+$ and t rotamers are used.

This complex convergence was wholly unexpected and we sought to rationalize this phenomenon in terms of the underlying structural and mechanistic constraints. For a protease to be catalytically productive, the leaving group must be positioned within hydrogen-bonding distance of the proton-transfer moiety. For $\chi_{cat} < -90°$ this constraint is satisfied only when a re-face attack occurs and, by symmetry, the opposite holds for $\chi_{cat} > 90°$ (Fig. 3). Because χ_{cat} is predetermined for each reactive rotamer/oxyanion hole pair, the convergence on a common stereochemical mechanism for each protease is the inevitable evolutionary consequence of using the local backbone to facilitate proteolysis.

As noted earlier, the proteases of clan SH do not use the local backbone in oxyanion hole formation but still attack into the re face of the scissile bond and have χ_{cat} and χ_{oxy} values similar to those of other re-face attackers (23). This prompted us to consider whether some additional mechanistic factor might preclude alternative geometries. For example, the oxyanion intermediate that is generated during enzyme acylation is highly basic and must be shielded from the acid moiety used in catalysis. In the framework described here, this chemical imperative translates
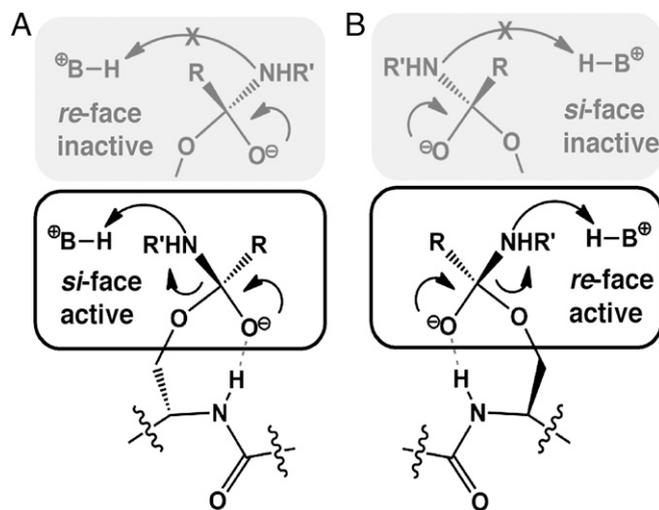

Fig. 3. Stereospecificity of proteolysis. (A) The substrate carbonyl points toward the local backbone for stabilization with a local oxyanion hole. When $\chi_{cat} > 90°$, the substrate must expose the si face of the peptide for attack, because a re-face attack is nonproductive for proton transfer. (B) When $\chi_{cat} < 90°$, the opposite holds. Only a re-face attack is viable for enzyme acylation and a si-face attack leads to an inactive complex.

into a requirement that χ_{oxy} and χ_{cat} must have the same sign (Fig. S1). We suggest additional constraints might be identified by *in silico* interrogation of nonobserved, hypothetical active site configurations. In sum, despite the great diversity in Ser protease structure and substrate specificity, they may be partitioned into precisely two distinct mechanistic classes according to which face of the peptide bond is attacked during catalysis.

Extension of Convergent Analysis to the Cys Proteases. The mechanistic constraints discussed above are not unique to the Ser proteases. The availability of the local backbone to contribute to oxyanion hole formation, the relevance of the reactive rotamer, and the importance of χ_{cat} are all applicable to the Cys proteases. To validate these principles, we extended our analysis to each of the 11 structurally characterized clans of Cys proteases (Table 2). Like the Ser proteases, the Cys proteases evolved to use all three staggered rotamers of the nucleophile and we observed the same correlation between the reactive rotamer and the identity of the local oxyanion hole residue. We found that χ_{cat} did not correlate well with the face of attack because most Cys proteases use a thiolate/acid pair that forms before substrate binding and the larger radius of sulfur means the thiol is deprotonated from farther away (31). This important distinction from the Ser proteases allows the acid to be preorganized for protonation of the leaving group and not activation of the nucleophile. More generally, the thiol of Cys is intrinsically more reactive than the hydroxyl of Ser and the structural constraints for efficient catalysis are correspondingly less stringent. These positional differences are made clearest by analysis of clans PA and PC, which contain both Ser and Cys proteases that are related through divergent evolution (9). In each case, the Cys proteases show geometries that would be ill-suited for deprotonation of Ser, but the acidic moiety is well positioned for protonation of the amine leaving group. Although thiols and hydroxyls require different geometries of proton transfer, the orientation of substrate binding is conserved between these enzymes. Table 2 shows that all Cys proteases use the local oxyanion hole and $|\chi_{oxy}| < 90°$. Moreover, the relationship between the reactive rotamer and the face of attack is identical to that of the Ser proteases (with the exception of clan SH). Although divergent evolution may change the substrate specificity or the identity of the catalytic triad of

Table 2. Active site parameters of convergently related Cys proteases

Clan	Enzyme	PDB	Complex	Reactive rotamer	ϕ, ψ angles	Catalytic motif	Oxyanion hole	Face of attack	χ_{cat}	α_{cat}	χ_{oxy}	Reference
PA	SARS protease	3SND	Thioacetal	g−	(−56, 142)	C-H	N	re	−168	82	−41	(74)
CD	Caspase 8	1QTN	Thioacetal	g−	(−73, 159)	C-H	N	re	NA	NA	−2	(75)
CF	Pyrrolidone carboxylate peptidase	1IU8		g−	(−64, -52)	C-H-E	N	re	−169			(76)
CL	L,D-transpeptidase	1ZAT		g−	(−65, 164)	C-H	N	re	179			(77)
CM	HCV NS2 protease	2HD0		g−	(−66, 141)	C-H-E	N	re	153			(78)
PC	PH1704	1G2I		t	(60, −115)	C-E	N + 1	si	180			(79)
CA	Papain	1POP	Thioacetal	g+	(−56, −32)	C-H	N	si	−177	76	30	(80)
CE	Ulp1	1EUV	Thioacetal	g+	(−52, −46)	C-H-D	N	si	−169	82	36	(81)
CN	Equine encephalitis alphavirus nsP2	2HWK		g+	(−54, −26)	C-H	N	si	ND			(82)
CO	Diamino endopeptidase	2HBQ		g+	(−64, −47)	C-H-H	N	si	−160			(83)
CP	deSUMOylase DeSI-1	2WP7	Unmodeled	g+	(−59, −32)	C-H	N	si	179			(84)

Members of clan CD do not use the same general base for activation of the nucleophile and protonation of the leaving group. The structure 2HWK crystallized with the base in an inactive state and 2WP7 has a ligand covalently bound to the nucleophile that was not modeled.

a protease, we report here that the stereochemistry of catalysis is absolutely conserved.

Mechanistic Constraints Imposed by the Thr γ-Methyl. With this comprehensive framework for dissecting proteolysis in place, we can address the question that first motivated this investigation: Why do the Ser proteases not use Thr? For every system we examined, the Thr γ-methyl was modeled onto the Ser nucleophile and its interactions with the environment were recorded. Surprisingly, only two types of interference were necessary to account for the inability of Thr to function in the catalytic triad. For Ser proteases with $\chi_{cat} > 90°$, which attack into the *si* face of the scissile bond, the γ-methyl of Thr would form a preclusive steric clash with the oxyanion generated during catalysis (Fig. 4 *C* and *D*). Alternatively, when $\chi_{cat} < −90°$, the γ-methyl clashes with the residue facilitating proton transfer (Fig. 4 *A* and *B*). Because this residue is the active site component subject to the most variation, we expanded our test of the γ-methyl to every structurally characterized Ser protease family, which accounts for divergent evolutionary pressures. For the few structures in which no clash was observed there was strong evidence that a catalytically incompetent conformation had been crystallized (Fig. S2 and SI Text). Hence, a clash with a hypothetical γ-methyl was observed for all Ser proteases irrespective of the identity of the catalytic triad, the reactive rotamer, or the components of the oxyanion hole. Because this clash is intrinsic to protease structure, evolution always selects against a Thr nucleophile in all clans of Ser proteases.

In light of this conclusion, the report that trypsin may be converted to a weak, but functional, Thr protease with as few as three mutations is particularly striking (32). This variant was never tested against a true protein substrate and we hypothesize that additional experiments will show it avoids enzyme acylation or does not feature direct proton transfer to the leaving group. There do exist natural Thr nucleophiles in biochemistry but without posttranslational modification, none of them catalyze proteolysis. The most similar reaction is that of the bacterial asparaginase enzymes, which hydrolyze the terminal amide of asparagine to yield aspartate and ammonia. Evidence clearly supports the use of a Thr nucleophile in a Thr-Tyr-Glu catalytic triad. However, removal of the phenolic OH of the Tyr base has a relatively small 100-fold reduction on k_{cat}, compared with the ~1,000,000-fold reduction upon removal of the His residue of the Ser protease catalytic triad (18, 33). This and other circumstantial evidence suggest more than one residue or a bridging water may be facilitating proton transfer during acylation (SI Text).

When the chemistry in question is fundamentally altered from hydrolysis to phosphoryl group transfer, so are the mechanistic imperatives. Protonation of the leaving group is no longer necessary, and the structure of phosphothreonine lyase shows the longer P-O bonds may accommodate the Thr γ-methyl without a steric clash (8). Although some reactions have sufficiently weak chemical constraints to allow a Thr nucleophile, the vast majority of enzymes that function with an acyl-enzyme intermediate do not use Thr. Covalent catalysis is a common theme in biochemistry and the structural constraints identified here for the proteases also apply to many other enzymatic reactions including amidohydrolases (34), esterases (35), β-lactamases (36), thio-
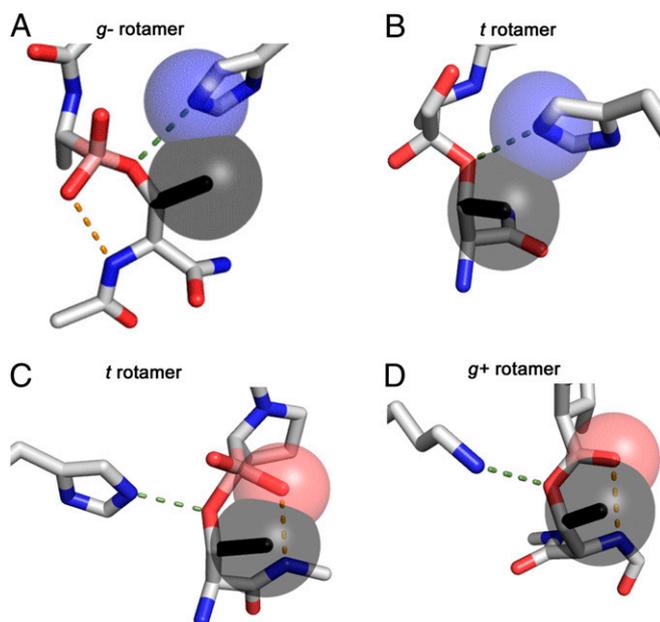

Fig. 4. Steric clashes of the Thr γ-methyl. (*A*) γ-Chymotrypsin with a boronic acid inhibitor (PDB ID: 3VGC). The γ-methyl of Thr was modeled (black) and its steric clash with the base is indicated by overlapping spheres drawn with the van der Waals radii. (*B*) CMV protease with an α-ketoamide inhibitor (PDB ID: 1NKM) shows a steric clash still occurs when the local oxyanion hole is not used. (*C*) Prolyl aminopeptidase with a boronic acid inhibitor (PDB ID: 2EEP). The "nucleophilic elbow" motif is incompatible with the Thr γ-methyl, which clashes with the oxyanion substrate (red sphere). (*D*) A similar clash is observed for signal peptidase I with a β-lactam inhibitor (PDB ID: 1B12).
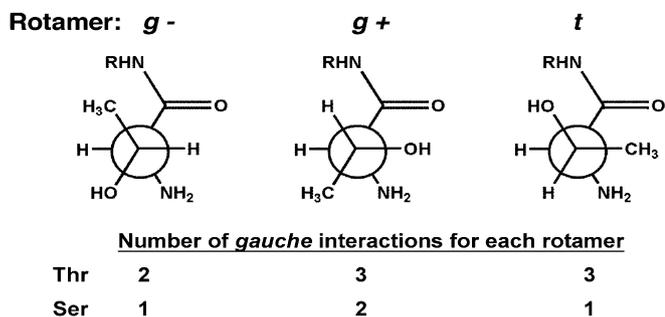
Rotamer: *g* -   *g* +   *t*

Number of *gauche* interactions for each rotamer

Thr	2	3	3
Ser	1	2	1

Fig. 5. *Gauche* interactions of an N-terminal nucleophile (Ntn). The *g*– rotamer of an Ntn Thr has the fewest *gauche* interactions. When the γ-methyl is removed by mutation to Ser, the *t* rotamer becomes favored. This effect is exacerbated by repulsive interactions with the neighboring carbonyl. Consequently, the Thr γ-methyl shifts the rotamer distribution toward a reactive state, which increases k_{cat}. These effects are intrinsic to Ntn structure and may contribute to the evolutionary selection of a Thr nucleophile.

esterases (37), and nitrilases (38), as well as more sophisticated enzymes such as glyceraldehye-3-phosphate dehydrogenase (39).

Positive Selection for an N-Terminal Thr Nucleophile. Of course, the structural barrier to the Thr nucleophile is not insurmountable. Seminal research into the mechanism of the proteasome β-sub-unit (PβS) established it as the first member of the "Thr pro-teases" (11, 40, 41). In a departure from the canonical Ser proteases, PβS and related enzymes require an intramolecular posttranslational modification, *cis*-autoproteolysis, to generate the active site (10, 42, 43). The resultant N-terminal nucleophile (Ntn) harbors an α-amine that acts as the proton transfer cata-lyst, analogous to the role played by His in the canonical catalytic triad. These enzymes are found to use all three possible nucle-ophiles (Cys, Ser, Thr). As indicated by the results discussed above, catalysis with an Ntn must be geometrically distinct from that of the Ser and Cys proteases to accommodate a Thr nu-cleophile. Furthermore, a thorough explanation of Ntn function should also account for the fact that mutational removal of the γ-methyl decreases the activity of PβS (20) and other Thr-using enzymes (44). Therefore, we extended our analysis to the Ntn enzymes in an effort to resolve the unusual evolutionary pref-erence for a Thr nucleophile.

There are at least two distinct protein folds that support autoproteolysis to generate an Ntn enzyme active site: the original Ntn-fold (10) and the DOM-fold (45). As described above, autoproteolysis generates an α-amine base with $\chi_{cat} = 30°$, which tightly constrains the potential catalytic geometries. Conse-quently, all Ntn-using enzymes bind their substrates with a com-mon orientation and attack into the *si* face of their substrates through the *g*– rotamer. More importantly, because the base approaches from a previously disallowed trajectory, it no longer has a steric clash with the γ-methyl, thereby allowing Thr to function as a catalytic nucleophile.

We recently reported that the Thr γ-methyl may accelerate the autoactivation of Ntn enzymes through the destabilization of inactive ground state conformations (46) through a phenomenon known as the reactive rotamer effect (RRE). Although the RRE is defined for intramolecular cyclizations (47), in principle it will apply to any reaction for which equilibrium between active and inactive conformational states is fast relative to catalysis. The apparent rate of the reaction (k_{cat}) can be increased if the addition of steric bulk shifts the ground state ensemble to favor the reactive state. This is precisely the situation presented by Thr→Ser mutation of PβS and related enzymes. As described by Fig. 5, the γ-methyl accelerates catalysis by intrinsically favoring the reactive *g*– rotamer. These rotamer distributions may be

further shifted by interactions with residues proximal to the active site (48). This rationale explains the reduction in k_{cat} observed upon Thr→Ser mutation of PβS (20) and glycosylasparaginase (44). Hence, the close association of multiple functional groups forces a steric clash with the Thr γ-methyl that precludes its use as a nucleophile for most enzymes. These constraints are relieved upon posttranslational autoproteolysis, whereupon the γ-methyl is not only tolerated but may intrinsically accelerate catalysis, explaining why it is the evolutionarily favored nucleophile in PβS and other Ntn enzymes.

Structural Constraints for Catalysis with a Local Oxyanion Hole. In addition to affecting the selection of the catalytic residue, the mechanistic constraints of enzyme acylation with a local oxy-anion hole limit the structural variation surrounding the nucle-ophile. For example, a Ramachandran plot showing the φ, ψ angles of the nucleophile is clustered by the reactive rotamer (Fig. 6). For the two *gauche* rotamers, φ ~ −55° is necessary to use the local backbone in the formation of an oxyanion hole. This in turn limits the possible ψ angles. When ψ ~ 150°, the carbonyl of the nucleophilic residue has a repulsive interaction with the *g*+ rotamer and only the *g*– rotamer is catalytically vi-able. Hence, all enzymes that use a local oxyanion hole and ac-ylate the *g*+ rotamer of the nucleophile have φ, ψ ~ −55°, −20°, where the *g*– rotamer is also allowed. Far more limiting, though, is use of the backbone of the N + 1 residue in the oxyanion hole, which requires ψ ~ −120°. This angle is allowed only in a small region of the φ, ψ space. When φ, ψ ~ 60°, −120°, a *g*+ rotamer has a strong steric repulsion with the N − 1 carbonyl, explaining why this rotamer is never associated with an enzyme that uses an N + 1 oxyanion hole.

These restrictions on the conformation of the local backbone have larger structural consequences. When the nucleophile has φ, ψ ~ −55°, 150°, there is no correlation in the local secondary structure between analogous enzymes (Fig. 7A). For φ, ψ ~ −55°,
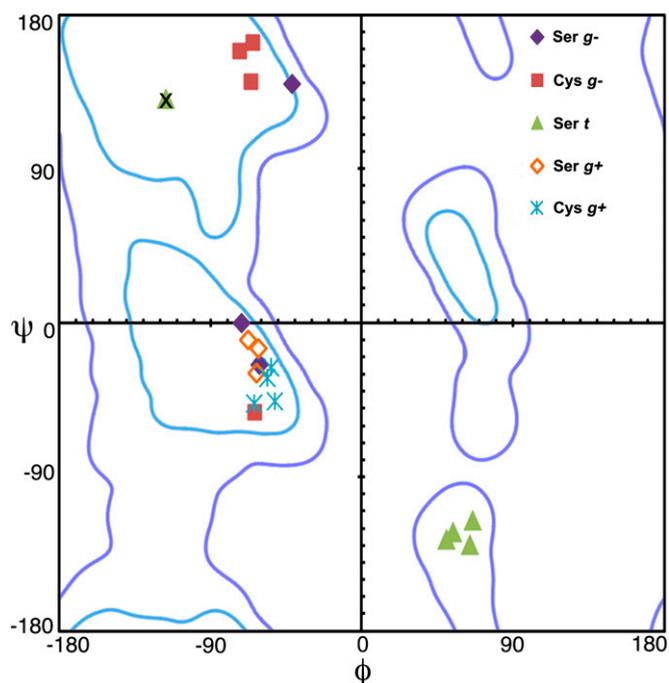

Fig. 6. Plot of the φ, ψ angles of the nucleophilic residue for Ser and Cys overlaying the Ramachandran plot. Proteases cluster according to the re-active rotamer in catalysis. The outlier, marked X, does not use the local backbone for oxyanion stabilization. Ramachandran plot was generated in MolProbity: light blue, strongly favored; dark blue, allowed (85).
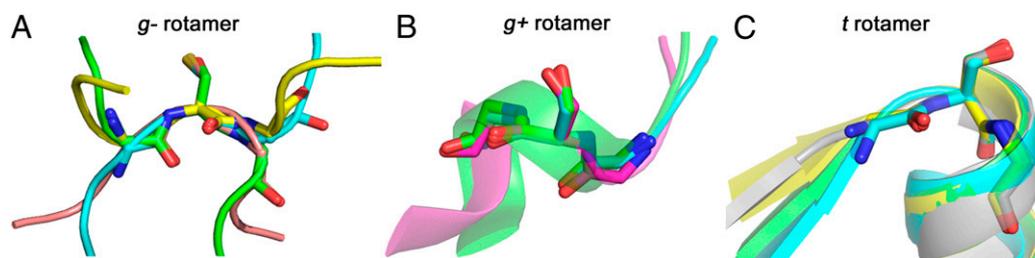
Fig. 7. Structural overlay of proteases according to their reactive rotamer. (*A*) Enzymes that use the *g*− rotamer show no structural similarity apart from their φ, ψ angles. (*B*) A reactive *g*+ rotamer has a loop at its N-terminal direction and a helix at it C-terminal end. (*C*) Use of the *t* rotamer and the N + 1 oxyanion hole imparts stringent structural constraints that are satisfied by a β-strand turning into an α-helix, the "nucleophilic elbow" motif. Structures shown are listed in Tables 1 and 2.

−20° there is a clear trend for the nucleophile to be positioned at the N terminus of a helix (Fig. 7*B*). When this circumstance occurs, the oxyanion is further stabilized by the helix dipole. Catalysis through the *t* rotamer with the N + 1 oxyanion hole shows evidence of the strongest structural constraints because ψ ∼ −120° is allowed when φ ∼ 60°. For these enzymes, the nucleophile is always preceded by a β-strand that leads into a helix (Fig. 7*C*). This motif has been recognized previously as the "nucleophilic elbow" and has strong sequence requirements for its formation (37, 49). This is the only viable conformation when a single residue separates a strand–helix transition in protein structure (50), which we suggest could be the first feature formed in a *t* rotamer proto-enzyme active site. Previously, there was no explanation for the convergent evolution of this elaborate motif for the proteases and more sophisticated nitrilases (37, 49, 51). When this question is posed in the framework described here, the nucleophilic elbow is simply a consequence of using the *t* rotamer and the N + 1 oxyanion hole for catalysis.

Recently, there has been great interest in developing artificial enzymes (52) and enzymes with novel function (53, 54). Landmark progress have been made with the design of Kemp Eliminases (55) and Diels-Alderases (56). For comparable success with enzymes that use an acylation mechanism, it would be beneficial to understand the interactions between the minimal active site components: nucleophile, oxyanion hole, and proton transfer catalyst. These were recently explored by the rational design of novel enzymes with esterase activity (57); but these enzymes had significantly less catalytic power than natural esterases. We believe that active site designs featuring the angles, distances, and structural motifs defined here should lead to catalysts with improved activity.

Stereospecific Mechanism-Based Inhibitors. Enzymes that use an acylation mechanism have a long and successful history as therapeutic targets for drug development. Since the discovery of penicillin, there has been an unceasing effort in the scientific community to both understand these enzymes and develop new molecules that inhibit them (58). Traditionally, these enzymes were classified according to the residue or metal ion used to generate the nucleophile (6). Unfortunately, this organization affords no perspective for the design of inhibitors that possess a mechanism of action that is sensitive to the stereochemistry of catalysis. Therefore, we sought to understand how nature exploits the distinction between *re*- and *si*-face–attacking enzymes by analyzing naturally occurring, stereospecific inhibitors in the context of our mechanism-based classification.

The β-lactam antibiotics inhibit penicillin-binding proteins responsible for the final cross-linking step in bacterial cell wall biosynthesis (59), a reaction that occurs through a *re*-face attack. Conspicuously, all classes of the naturally occurring bicyclic β-lactam antibiotics feature stereochemistry that occludes the *si* face of the β-lactam moiety (Fig. 8*A*). The bacterial signal pep-

tidase (SPase) is also present in the periplasm, but functions through a *si*-face attack. Consistent with our analysis, it has been shown that incubation of traditional β-lactam antibiotics fails to inhibit these enzymes (60). However, a β-lactam–based SPase inhibitor was successfully generated when the bridgehead stereochemistry of a penicillin-like molecule was inverted, to expose the *si* face of the synthetic β-lactam.

Omuralide is a naturally occurring β-lactone–γ-lactam (Fig. 8*B*) antibiotic that functions through inhibition of PβS, a *si*-face–attacking Ntn enzyme (61). Despite its high binding affinity for PβS, it has been shown that incubation of omuralide with the *Chlamydia* protease CPAF, which uses the nucleophilic elbow motif, successfully inactivates the enzyme (62). This is consistent with the hypotheses that the γ-lactam blocks the *re* face of the molecule, thereby restricting the inhibition profile to *si*-face–attacking enzymes. Similar properties may be attributed to the salinosporamides (63), vibralactone (64), and avibactam (65). In total, we identified seven independent cases where nature selected for the biosynthesis of an inhibitor that is sensitive to the stereochemistry of the active site of the target enzyme (*SI Text*).

Currently, efforts to develop new therapeutics are often focused on optimizing interactions within the binding pocket that are distinct from the catalytic center itself. This is an extremely challenging pursuit because there are numerous enzymes with overlapping substrate specificities. We suggest that incorporation of moieties that preclude attack into the nontarget face of an inhibitor may be a useful negative-design parameter in the development of new drugs with improved specificity.

Conclusions

In the present investigation, we defined geometric enzyme–substrate parameters that enabled systematic examination of enzyme function and applied them to every structurally characterized
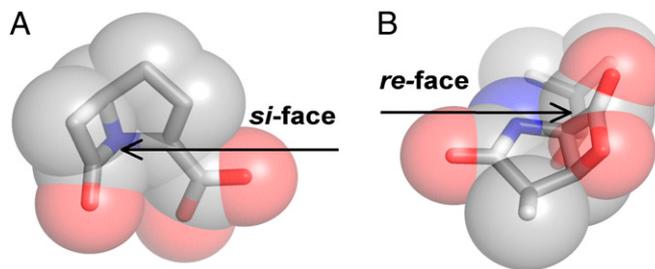

Fig. 8. Stereospecific antibiotics. (*A*) The β-lactam molecule carbapenem contains stereochemistry that blocks the *si* face of the elecrophilic carbonyl and this molecule may be used only to inhibit *re*-face attacking enzymes. (*B*) Conversely, omuralide features a β-lactone–γ-lactam core that occludes the *re* face of the eletrophilic carbonyl. Consequently, this inhibitor scaffold is effective only against *si*-face–attacking enzymes.

family of the Ser, Cys, and Thr proteases. This study revealed surprising correlations between seemingly independent active site components. We explained that utilization of the local backbone in oxyanion stabilization is related to the reactive rotamer for catalysis and that the location of the base determines the orientation of substrate binding. We advocate division of the proteases into two classes according to the stereochemistry of their nucleophilic attack, as is common in the study of NAD(P)H-dependent enzymes (66). By understanding the active site constraints imparted by proteolysis through an acyl-enzyme mechanism, we accounted for the selection against a Thr nucleophile, established why the nucleophilic elbow forms, and provided a basis for understanding the evolution of stereospecific inhibitors. These relationships constitute incontrovertible links between structure and function and limit the fitness landscape upon which the evolution of biocatalysis occurs. This knowledge is distinct from, but no less valid than, that accessible through detailed quantum mechanical calculations, time, and computational expertise. Our

method of convergent evolutionary analysis is broadly applicable and should aid in the continued study of enzymes and their elegant structure–function relationship.

Methods

Proteases of unrelated evolutionary lines were identified according to their classification in the MEROPS database (9). From this list, structures were selected according to their resolution and the presence of a mechanistically relevant ligand and uploaded into Coot for visual inspection of model quality within each active site (67). All of the measured values in Tables 1 and 2 contain error that is intrinsic to crystallography. These deviations are resolution dependent and, when possible, we have chosen the highest-resolution structure available. The face of attack for each enzyme was established by analysis of an inhibitor complex or inference from the location of the P1 site on the protein. Superimpositions and visualization of group van der Waals radii were performed in PyMol (Schrödinger, LLC).

1. Warshel A, Naray-Szabo G, Sussman F, Hwang JK (1989) How do serine proteases really work? *Biochemistry* 28(9):3629–3637.
2. Fersht A (1999) *Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding* (Freeman, New York).
3. Kraut J (1988) How do enzymes work? *Science* 242(4878):533–540.
4. Warshel A (1998) Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites. *J Biol Chem* 273(42):27035–27038.
5. Warshel A (1978) Energetics of enzyme catalysis. *Proc Natl Acad Sci USA* 75(11):5250–5254.
6. Ekici OD, Paetzel M, Dalbey RE (2008) Unconventional serine proteases: Variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci* 17(12):2023–2037.
7. Milkowski C, Strack D (2004) Serine carboxypeptidase-like acyltransferases. *Phytochemistry* 65(5):517–524.
8. Chen L, et al. (2008) Structural basis for the catalytic mechanism of phosphothreonine lyase. *Nat Struct Mol Biol* 15(1):101–102.
9. Rawlings ND, Barrett AJ, Bateman A (2012) MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* 40(Database issue, D1):D343–D350.
10. Brannigan JA, et al. (1995) A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. *Nature* 378(6555):416–419.
11. Seemüller E, et al. (1995) Proteasome from Thermoplasma acidophilum: A threonine protease. *Science* 268(5210):579–582.
12. Smith AJT, et al. (2008) Structural reorganization and preorganization in enzyme active sites: Comparisons of experimental and theoretically ideal active site geometries in the multistep serine esterase reaction cycle. *J Am Chem Soc* 130(46):15361–15373.
13. Galperin MY, Walker DR, Koonin EV (1998) Analogous enzymes: Independent inventions in enzyme evolution. *Genome Res* 8(8):779–790.
14. Polgár L (2005) The catalytic triad of serine peptidases. *Cell Mol Life Sci* 62(19–20):2161–2172.
15. Robertus JD, Kraut J, Alden RA, Birktoft JJ (1972) Subtilisin; a stereochemical mechanism involving transition-state stabilization. *Biochemistry* 11(23):4293–4303.
16. Neet KE, Koshland DE, Jr. (1966) The conversion of serine at the active site of subtilisin to cysteine: A "chemical mutation". *Proc Natl Acad Sci USA* 56(5):1606–1611.
17. Polgar L, Bender ML (1966) A new enzyme containing a synthetically formed active site. Thiol-subtilisin. *J Am Chem Soc* 88(13):3153–3154.
18. Carter P, Wells JA (1988) Dissecting the catalytic triad of a serine protease. *Nature* 332(6164):564–568.
19. Corey DR, Craik CS (1992) An investigation into the minimum requirements for peptide hydrolysis by mutation of the catalytic triad of trypsin. *J Am Chem Soc* 114(5):1784–1790.
20. Kisselev AF, Songyang Z, Goldberg AL (2000) Why does threonine, and not serine, function as the active site nucleophile in proteasomes? *J Biol Chem* 275(20):14831–14837.
21. Rawlings ND, Barrett AJ (1994) Families of serine peptidases. *Methods Enzymol* 244:19–61.
22. Paetzel M, Dalbey RE (1997) Catalytic hydroxyl/amine dyads within serine proteases. *Trends Biochem Sci* 22(1):28–31.
23. Khayat R, et al. (2003) Structural and biochemical studies of inhibitor binding to human cytomegalovirus protease. *Biochemistry* 42(4):885–891.
24. Simón L, Goodman JM (2010) Enzyme catalysis by hydrogen bonds: The balance between transition state binding and substrate binding in oxyanion holes. *J Org Chem* 75(6):1831–1840.
25. Simón L, Goodman JM (2012) Hydrogen-bond stabilization in oxyanion holes: Grand jeté to three dimensions. *Org Biomol Chem* 10(9):1905–1913.
26. Fülöp V, Böcskei Z, Polgár L (1998) Prolyl oligopeptidase: An unusual beta-propeller domain regulates proteolysis. *Cell* 94(2):161–170.
27. Wahlgren WY, et al. (2011) The catalytic aspartate is protonated in the Michaelis complex formed between trypsin and an in vitro evolved substrate-like inhibitor: A refined mechanism of serine protease action. *J Biol Chem* 286(5):3587–3596.
28. 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.
29. 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.
30. Mac Sweeney A, et al. (2000) Crystal structure of delta-chymotrypsin bound to a peptidyl chloromethyl ketone inhibitor. *Acta Crystallogr D Biol Crystallogr* 56(Pt 3):280–286.
31. Storer AC, Ménard R (1994) Catalytic mechanism in papain family of cysteine peptidases. *Methods Enzymol* 244:486–500.
32. Baird TT, Jr., Wright WD, Craik CS (2006) Conversion of trypsin to a functional threonine protease. *Protein Sci* 15(6):1229–1238.
33. 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.
34. Mileni M, et al. (2009) Binding and inactivation mechanism of a humanized fatty acid amide hydrolase by alpha-ketoheterocycle inhibitors revealed from cocrystal structures. *J Am Chem Soc* 131(30):10497–10506.
35. Cygler M, et al. (1993) Relationship between sequence conservation and three-dimensional structure in a large family of esterases, lipases, and related proteins. *Protein Sci* 2(3):366–382.
36. Strynadka NCJ, et al. (1992) Molecular structure of the acyl-enzyme intermediate in beta-lactam hydrolysis at 1.7 A resolution. *Nature* 359(6397):700–705.
37. Ollis DL, et al. (1992) The alpha/beta hydrolase fold. *Protein Eng* 5(3):197–211.
38. Brenner C (2002) Catalysis in the nitrilase superfamily. *Curr Opin Struct Biol* 12(6):775–782.
39. Garavito RM, Rossmann MG, Argos P, Eventoff W (1977) Convergence of active center geometries. *Biochemistry* 16(23):5065–5071.
40. Löwe J, et al. (1995) Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 A resolution. *Science* 268(5210):533–539.
41. Groll M, et al. (1997) Structure of 20S proteasome from yeast at 2.4 A resolution. *Nature* 386(6624):463–471.
42. Seemuller E, Lupas A, Baumeister W (1996) Autocatalytic processing of the 20S proteasome. *Nature* 382(6590):468–471.
43. Perler FB, Xu MQ, Paulus H (1997) Protein splicing and autoproteolysis mechanisms. *Curr Opin Chem Biol* 1(3):292–299.
44. Liu Y, Guan C, Aronson NN, Jr. (1998) Site-directed mutagenesis of essential residues involved in the mechanism of bacterial glycosylasparaginase. *J Biol Chem* 273(16):9688–9694.
45. Cheng H, Grishin NV (2005) DOM-fold: A structure with crossing loops found in DmpA, ornithine acetyltransferase, and molybdenum cofactor-binding domain. *Protein Sci* 14(7):1902–1910.
46. Buller AR, Freeman MF, Wright NT, Schildbach JF, Townsend CA (2012) Insights into cis-autoproteolysis reveal a reactive state formed through conformational rearrangement. *Proc Natl Acad Sci USA* 109(7):2308–2313.
47. Jung ME, Piizzi G (2005) gem-disubstituent effect: Theoretical basis and synthetic applications. *Chem Rev* 105(5):1735–1766.
48. Buller AR, et al. (2012) Autoproteolytic activation of ThnT results in structural reorganization necessary for substrate binding and catalysis. *J Mol Biol* 422(4):508–518.
49. 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.
50. Street TO, Fitzkee NC, Perskie LL, Rose GD (2007) Physical-chemical determinants of turn conformations in globular proteins. *Protein Sci* 16(8):1720–1727.
51. Pace HC, et al. (2000) Crystal structure of the worm NitFhit Rosetta Stone protein reveals a Nit tetramer binding two Fhit dimers. *Curr Biol* 10(15):907–917.
52. Motherwell WB, Bingham MJ, Six Y (2001) Recent progress in the design and synthesis of artificial enzymes. *Tetrahedron* 57(22):4663–4686.
53. Kipnis Y, Baker D (2012) Comparison of designed and randomly generated catalysts for simple chemical reactions. *Protein Sci* 21(9):1388–1395.

54. Gerlt JA, Babbitt PC (2009) Enzyme (re)design: Lessons from natural evolution and computation. *Curr Opin Chem Biol* 13(1):10–18.

55. Röthlisberger D, et al. (2008) Kemp elimination catalysts by computational enzyme design. *Nature* 453(7192):190–195.

56. Siegel JB, et al. (2010) Computational design of an enzyme catalyst for a stereoselective bimolecular Diels-Alder reaction. *Science* 329(5989):309–313.

57. Richter F, et al. (2012) Computational design of catalytic dyads and oxyanion holes for ester hydrolysis. *J Am Chem Soc* 134(39):16197–16206.

58. Page MGP, Heim J (2009) New molecules from old classes: Revisiting the development of beta-lactams. *IDrugs* 12(9):561–565.

59. Fisher JF, Meroueh SO, Mobashery S (2005) Bacterial resistance to beta-lactam antibiotics: Compelling opportunism, compelling opportunity. *Chem Rev* 105(2):395–424.

60. 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.

61. Fenteany G, et al. (1995) Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin. *Science* 268(5211):726–731.

62. Huang ZW, et al. (2008) Structural basis for activation and inhibition of the secreted chlamydia protease CPAF. *Cell Host Microbe* 4(6):529–542.

63. Groll M, Huber R, Potts BCM (2006) Crystal structures of Salinosporamide A (NPI-0052) and B (NPI-0047) in complex with the 20S proteasome reveal important consequences of beta-lactone ring opening and a mechanism for irreversible binding. *J Am Chem Soc* 128(15):5136–5141.

64. Zeiler E, et al. (2011) Vibralactone as a tool to study the activity and structure of the ClpP1P2 complex from Listeria monocytogenes. *Angew Chem Int Ed Engl* 50(46):11001–11004.

65. Ehmann DE, et al. (2012) Avibactam is a covalent, reversible, non-β-lactam β-lactamase inhibitor. *Proc Natl Acad Sci USA* 109(29):11663–11668.

66. Fisher HF, Conn EE, Vennesland B, Westheimer FH (1953) The enzymatic transfer of hydrogen. I. The reaction catalyzed by alcohol dehydrogenase. *J Biol Chem* 202(2):687–697.

67. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.

68. 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.

69. Wlodawer A, et al. (2001) Carboxyl proteinase from Pseudomonas defines a novel family of subtilisin-like enzymes. *Nat Struct Biol* 8(5):442–446.

70. 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.

71. 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.

72. 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.

73. Wang Y, Zhang Y, Ha Y (2006) Crystal structure of a rhomboid family intramembrane protease. *Nature* 444(7116):179–180.

74. Zhu L, et al. (2011) Peptide aldehyde inhibitors challenge the substrate specificity of the SARS-coronavirus main protease. *Antiviral Res* 92(2):204–212.

75. Watt W, et al. (1999) The atomic-resolution structure of human caspase-8, a key activator of apoptosis. *Structure* 7(9):1135–1143.

76. Sokabe M, et al. (2002) The X-ray crystal structure of pyrrolidone-carboxylate peptidase from hyperthermophilic archaea Pyrococcus horikoshii. *J Struct Funct Genomics* 2(3):145–154.

77. Biarrotte-Sorin S, et al. (2006) Crystal structure of a novel beta-lactam-insensitive peptidoglycan transpeptidase. *J Mol Biol* 359(3):533–538.

78. Lorenz IC, Marcotrigiano J, Dentzer TG, Rice CM (2006) Structure of the catalytic domain of the hepatitis C virus NS2-3 protease. *Nature* 442(7104):831–835.

79. Du XL, et al. (2000) Crystal structure of an intracellular protease from Pyrococcus horikoshii at 2-A resolution. *Proc Natl Acad Sci USA* 97(26):14079–14084.

80. Schröder E, Phillips C, Garman E, Harlos K, Crawford C (1993) X-ray crystallographic structure of a papain-leupeptin complex. *FEBS Lett* 315(1):38–42.

81. Mossessova E, Lima CD (2000) Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol Cell* 5(5):865–876.

82. Russo AT, White MA, Watowich SJ (2006) The crystal structure of the Venezuelan equine encephalitis alphavirus nsP2 protease. *Structure* 14(9):1449–1458.

83. Xu Q, et al. (2009) Structural basis of murein peptide specificity of a gamma-D-glutamyl-l-diamino acid endopeptidase. *Structure* 17(2):303–313.

84. Suh H-Y, et al. (2012) Crystal structure of DeSI-1, a novel deSUMOylase belonging to a putative isopeptidase superfamily. *Proteins* 80(8):2099–2104.

85. Chen VB, et al. (2010) MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12–21.