Enzyme:Substrate Hydrogen Bond Shortening during the Acylation Phase of Serine Protease Catalysis[†]

Krisztián Fodor,[‡] Veronika Harmat,[§] Richard Neutze,[||] László Szilágyi,[⊥] László Gráf,*,[‡] and Gergely Katona*,[#]

Biotechnology Research Group of the Hungarian Academy of Sciences, Pázmány Street 1/C, 1117 Budapest, Hungary, Protein Modeling Group of the Hungarian Academy of Sciences, Pázmány Street 1/C, 1117 Budapest, Hungary, Department of Chemistry and Bioscience, Chalmers University of Technology, Box 462, 40530 Gothenburg, Sweden, Department of Biochemistry, Eötvös Loránd University, Pázmány Street 1/C, H-1117 Budapest, Hungary, and LCCP, UMR 9015 IBS-CEA/CNRS/Université J. Fourier, 41 Avenue Jules Horowitz, 38027 Grenoble, Cedex 1, France

ABSTRACT: Atomic resolution (≤1.2 Å) serine protease intermediate structures revealed that the strength of the hydrogen bonds between the enzyme and the substrate changed during catalysis. The well-conserved hydrogen bonds of antiparallel β-sheet between the enzyme and the substrate become significantly shorter in the transition from a Michaelis complex analogue (*Pontastacus leptodactylus* (narrow-fingered crayfish) trypsin (CFT) in complex with *Schistocerca gregaria* (desert locust) trypsin inhibitor (SGTI) at 1.2 Å resolution) to an acyl−enzyme intermediate (*N*-acetyl-Asn-Pro-Ile acyl−enzyme intermediate of porcine pancreatic elastase at 0.95 Å resolution) presumably synchronously with the nucleophilic attack on the carbonyl carbon atom of the scissile peptide bond. This is interpreted as an active mechanism that utilizes the energy released from the stronger hydrogen bonds to overcome the energetic barrier of the nucleophilic attack by the hydroxyl group of the catalytic serine. In the CFT:SGTI complex this hydrogen bond shortening may be hindered by the 27I−32I disulfide bridge and Asn-15I of SGTI. The position of the catalytic histidine changes slightly as it adapts to the different nucleophilic attacker during the transition from the Michaelis complex to the acyl−enzyme state, and simultaneously its interaction with Asp-102 and Ser-214 becomes stronger. The oxyanion hole hydrogen bonds provide additional stabilization for acyl−ester bond in the acyl−enzyme than for scissile peptide bond of the Michaelis complex. Significant deviation from planarity is not observed in the reactive bonds of either the Michaelis complex or the acyl−enzyme. In the Michaelis complex the electron distribution of the carbonyl bond is distorted toward the oxygen atom compared to other peptide bonds in the structure, which indicates the polarization effect of the oxyanion hole.

Serine proteases are ubiquitous in prokaryotes and eukaryotes and serve important and diverse biological functions. These include central biological processes such as fibrinolysis, hemostasis, complement reaction, and the digestion of dietary proteins. Early studies of trypsin and other serine protease systems established the structural basis of serine protease catalysis (*2−6*). While serine proteases have a very broad range of specificities, the main catalytic steps and catalytic residues are presumed to be identical (*7*). The first step is the noncovalent binding of the polypeptide substrate to the active site cleft which leads to the formation of a Michaelis complex. In the acylation phase, the catalytic serine attacks the scissile amide bond and concomitantly a tetra-

hedral oxyanion intermediate is generated that breaks down to an acyl−enzyme complex with the following release of the C-terminal product fragment. In the deacylation phase of catalysis, attack of a water molecule onto the ester bond results in a second tetrahedral intermediate that collapses, releasing the N-terminal product fragment and regenerating the native enzyme.

Even though there is a consensus on the core steps, the relative importance of various factors influencing the catalytic efficiency remains unclear. Hydrogen bonds have always been in the focus of interest. In particular the precise nature of the interaction between the catalytic histidine and aspartate remained controversial (*8−14*). Kinetic and solvent isotope effect experiments indicated hydrogen transfer in the rate-limiting transition state (*15, 16*). A large scale conformational change of the catalytic histidine was also proposed which would facilitate the reprotonation of the leaving amino group (*8, 17*); this process is thought to be essential for making the amide a better leaving group. More subtle, but nevertheless important structural changes of catalytic residues during the reaction cycle were also emphasized based on calculational and theoretical grounds (*18, 19*). Earlier low-resolution observations indicated geometrically strained acyl−enzyme intermediates (*20*), and a strained scissile peptide bond was

* To whom correspondence should be addressed. G.K.: tel, +33 (0)4 38 78 59 22; fax, +33 (0)4 38 78 51 22; e-mail, gergely.katona@ibs.fr. L.G.: tel, +36 (0)1 38 12 171; fax, +36 (0)1 38 12 172; e-mail, graf@ludens.elte.hu.
[‡] Biotechnology Research Group of the Hungarian Academy of Sciences.
[§] Protein Modeling Group of the Hungarian Academy of Sciences.
[||] Chalmers University of Technology.
[⊥] Eötvös Loránd University.
[#] UMR 9015 IBS-CEA/CNRS/Université J. Fourier.

found in a subtilisin:elastase inhibitor complex (*21*). Later, based on an atomic resolution structure, the acyl−enzyme ester bond was found to be planar within experimental error (*22*). The key tetrahedral intermediate was observed in high quality data obtained from pH jumped flash frozen crystals of porcine pancreatic elastase acyl−enzymes (*23*).

Hamdaoui et al. (*24*) and Malik et al. (*25*) independently isolated trypsin-specific canonical inhibitors from a desert locust (*Schistocerca gregaria*). SGTI[1] (*Schistocerca gregaria* trypsin inhibitor) was found to be only a modest inhibitor of bovine pancreatic trypsin, while it inhibited very efficiently trypsin isolated from narrow-fingered crayfish (CFT) (*26*). The NMR structure of SGTI revealed an exposed, convex loop which enables SGTI to act as a canonical inhibitor of trypsin (*27*). The crystal structure of the CFT:SGTI complex revealed an extended interaction between the enzyme and the inhibitor (*28*). Canonical inhibitors inhibit serine proteases by forming a tight complex and thereby preventing the substrate molecules from gaining access to the substrate binding pocket. Surprisingly, the enzymes remain capable of cleaving peptide bonds of canonical inhibitors, which was demonstrated by the isolation of the cleaved form of the inhibitor in the presence of the enzyme (*29, 30*). This observation has led to the formulation of the standard mechanism of canonical inhibition: the prolonged association of the enzyme and the inhibitor leads to an equilibrium between the cleaved and uncleaved forms of the inhibitor.

Earlier lower resolution X-ray diffraction studies on serine protease:inhibitor complexes did not reveal significant deviation from the native active site geometry that would presumably render the enzyme inactive (*21*). Herein we report the refinement of the structure of the canonical inhibitor complex formed between crayfish trypsin (CFT) and SGTI (PDB entry: 2f91) which is assumed to be a suitable model for studying the catalytic Michaelis (loose) complex (*28, 31*), in which the substrate peptide forms an antiparallel β-sheet with the enzyme in order to position the scissile peptide bond in the active site. It is important to note however, that the scissile peptide does not hydrolyze rapidly and the cleaved and uncleaved forms of the inhibitor reach equilibrium over an extended period. Substrate binding via a canonical antiparallel β-sheet was also called into question (*32*), but as yet alternative binding models fail to explain why the acyl−peptide also binds in a canonical conformation (*22*) and why the K_d for canonical inhibitors correlate with the k_{cat}/K_m of their analogous substrates (*33*). Nevertheless, the very high occupancy of SGTI combined with the accuracy of the atomic resolution data allowed the detailed description of the scissile peptide bond geometry in the CFT: SGTI complex relative to the acyl−ester bond (*22*) and the tetrahedral intermediate (*23*).

MATERIAL AND METHODS

Refinement Details. The CFT:SGTI complex and the elastase acyl−enzyme were refined with SHELX (*34*) as described previously (*22, 28*) with the following modifications: Identical geometric and *B*-factor weights were used during the two refinements. The same number of hydrogens

were modeled on the catalytic residues and enzyme:substrate β-sheet (main chain) in the two structures, and no hydrogens were included for the His-57 residue. In order to perform an unbiased comparison, the geometry of the P2−P2′ residues of SGTI was not restrained in the final rounds of the refinement. Similarly the ester linkage between the acyl−peptide and Ser-195 in the elastase acyl−enzyme was simulated unrestrained in the final SHELXL refinement steps. The updated model of the CFT:SGTI complex has been submitted to the PDB (*35*) (entry code: 2f91).

Analysis and Treatment of Errors. The radial positional errors of the atoms and errors of bond and plane angles were deduced from the inversion of the least squares minimization matrix in SHELX without geometric and *B*-factor restraints (DFIX, DANG, FLAT, DELU, SIMU, ISOR, and BUMP) and using the BLOC 1 and DAMP 0 0 parameters. Calculation of "radial positional errors" takes into account the variance of the coordinate along the crystallographic axes and the covariances (i.e. the off-diagonal elements of the inverse of the LS matrix) between them. The radial positional errors were propagated into the errors of distances and distance differences as described by Schneider (2000) (*36*). We also ignored the fact that for estimating distance errors the perpendicular component of the radial position errors is not necessary, and we neglected the covariance terms between the two atoms. As a consequence the distance errors reported here are slightly overestimated compared to the ones calculated with the RTAB command of SHELX. Bonding and plane angles and their associated estimated errors were calculated directly with the commands RTAB and MPLA, respectively.

Analysis of the Electron Density and Crystallographic Models. 2mFo-DFc electron density maps were calculated with Shelxpro (*34*). Absolute scaling of the reflections was done by SHELXL using diffraction intensities as input. The program MAPMAN (*37*) was used to calculate the electron density values at the atomic positions with the PEek VAlue interpolate option. The structural models of CFT:SGTI and elastase acyl−enzyme were superpositioned with iterative sieve fitting using the program LSQMAN (*38*) and compared with the program O (*39*).

RESULTS AND DISCUSSION

Mechanism for Conversion of Binding Energy into Catalytic Action. SGTI forms an antiparallel β-sheet with CFT (Figure 1) via six hydrogen bond connections. Two of those form the oxyanion hole at the scissile bond (not shown); three connect the N terminal part of the inhibitor to the enzyme, and only one holds the C terminal part which would normally constitute the first product of the catalytic cycle. These hydrogen bonds are structurally homologous in other serine protease:inhibitor complexes of the chymotrypsin family (Figure 1B,C). In the atomic resolution structure of the elastase acetyl-NPI acyl−enzyme (PDB code: 1gvk) (*22*) all three β-sheet forming non-oxyanion hole hydrogen bonds are shorter than their counterparts in the CFT:SGTI structure (P1 N to S1 O 3.56 Å vs 3.04 Å, P3 O to S2 N 3.14 Å vs 2.97 Å, and P3 N to S2 O 2.95 Å vs 2.72 Å).[2] The hydroxyl

---

[1] Abbreviations: CFT, crayfish trypsin; SGTI, *Schistocerca gregaria* trypsin inhibitor; rms, root mean square.

[2] The naming of the substrate binding subsites follows the Schechter−Berger convention (*1*).
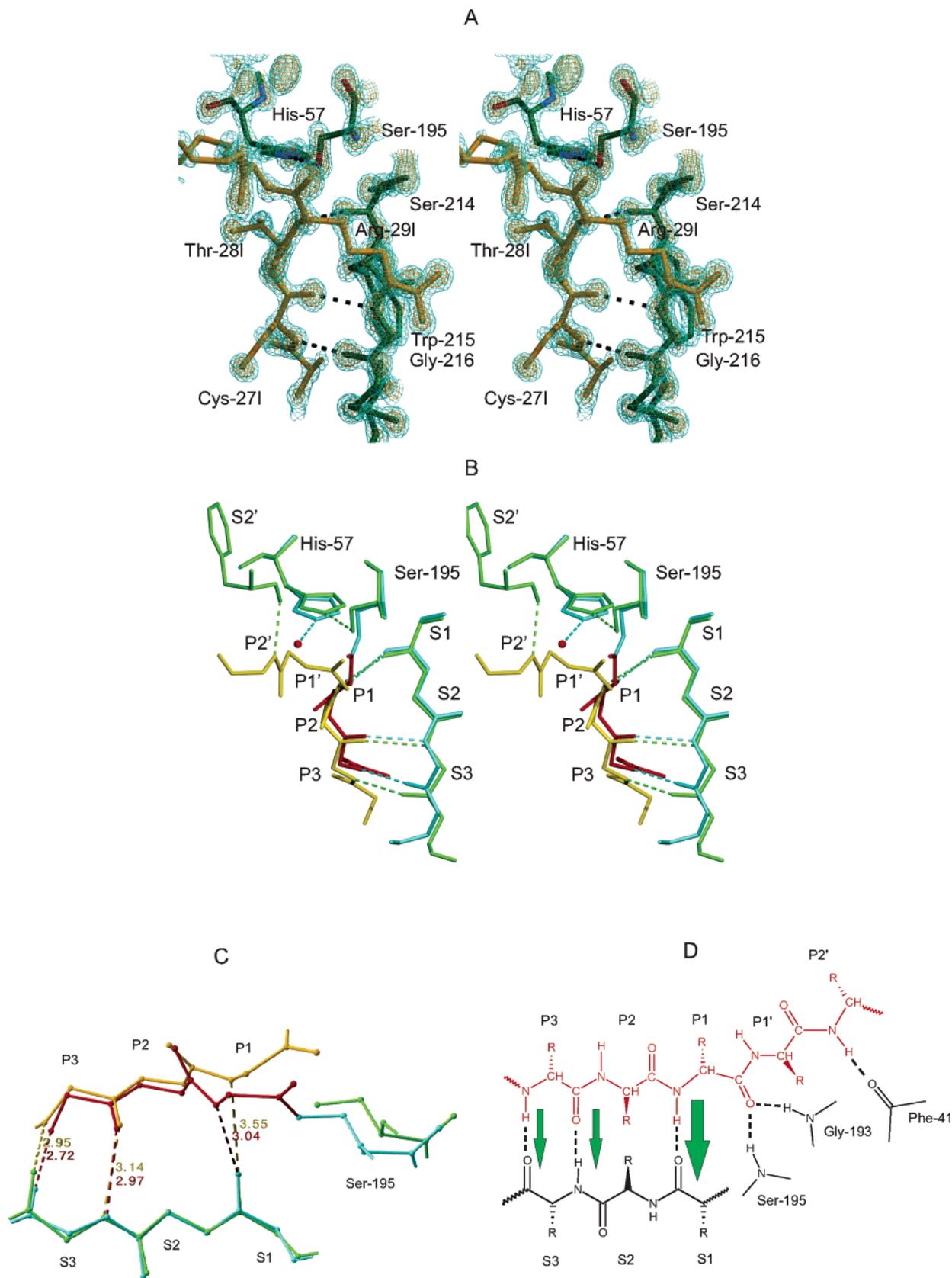
A

B

C

D

FIGURE 1: (A) Stereoview of the X-ray structure and electron density near the active site of the CFT:SGTI complex at 1.2 Å resolution. The bonds of the enzyme are colored in green while the inhibitor is yellow. The 2mFo-DFc electron density map is contoured at 2.5σ (cyan) and 4.0σ (gold). (B) Superposition of the elastase Ac−NPI acyl−enzyme complex onto the CFT:SGTI complex. CFT is green, SGTI is yellow, porcine pancreatic elastase is cyan, and the Ac−NPI acyl−peptide is red. The position of the attacking water molecule is also indicated as a red sphere. (Stereoview.) (C) Hydrogen bond lengths in the CFT:SGTI complex and the acyl−enzyme intermediate. The color code is identical to that in Figure 1B. (D) Schematic representation of the enzyme−substrate β-sheet. This figure was produced with the programs Povscript+ (56) and Chemsketch (http://www.adclabs.com).

group of the catalytic serine is wedged between the peptide and the enzyme, which probably accounts for the longer hydrogen bonds and prevents the formation of an ideal, cooperative interaction. It is also worth noting that the distance difference between the donor and acceptor is the largest in the hydrogen bond which is closest to the scissile peptide bond. This may implicate that, once the hydroxyl group of the serine is activated, part of the activation energy that is required for the tetrahedral intermediate formation could be derived from the stronger hydrogen bonds, which forces the carbonyl carbon to interact more closely with the hydroxyl of the catalytic serine. The cooperative strengthening of three hydrogen bonds may therefore compensate the energetic barrier of the nucleophilic attack (Figure 1C). This hypothesis complements the recent idea that, during the deacylation reaction, the transition from the planar to the tetrahedral geometry shifts the position of the carbonyl carbon atom which leads to the "unzipping" of the acyl−peptide and eventually the departure of the product (*23*). It may also provide an explanation to why only the breakdown of cognate substrates is catalyzed at an elevated rate (*40*).

If we assume that the enzyme:substrate interaction is finely tuned then any substrate which cannot make stronger hydrogen bonds with the enzyme due to steric or electrostatic hindrances will not be able to utilize the rate acceleration provided by the enzyme:substrate β-sheet. Negative cooperativity could play an important role, since even a single noncognate residue in the substrate may have the potential to prevent the shortening of multiple hydrogen bonds, in which case the strain imposed by the hydrogen bonds on the Ser-195 hydroxyl−carbonyl carbon interaction may not be sufficient for carrying out the nucleophilic attack and it may promote substrate dissociation instead (see similar arguments in ref *41*). It is well-known that longer synthetic peptides are better substrates than shorter ones, which also points to the importance of β-sheet hydrogen bond formation in catalytic action (*7*). Further away from the P1 residue hydrogen bond shortening is less pronounced and also noncognate residue at distant positions cannot disturb proximal hydrogen bonds as much, which may limit their effect on specificity. On the other hand, since the largest energy gain can be expected at the P1 residue, the noncognate binding at this position has the most severe effect (*42*). Early structural studies on boronate and phosphonate esters of serine proteases (*43*) also revealed a closer interaction between P1 and the main chain of Ser-214 compared to canonic inhibitor complexes, albeit the binding of an artificial intermediate and the limited crystallographic precision of these studies may have masked the shortening at more distant hydrogen bonds. Our studies on natural intermediates further emphasize the potential role of the "catalytic β-sheet", which can be extended to the acylation phase of the catalytic cycle.

Substrate vs Inhibitor. Although strong and specific binding seems to be paramount in efficient catalytic action, canonic inhibitors do not adhere to this rule. Canonic inhibitors typically bind to serine proteases with a K_d well below the nanomolar range, yet their scissile peptide bond hydrolyzes very slowly. According to the standard mechanism, the inhibitor complex can dissociate into both the uncleaved and the cleaved form of the inhibitor (*44*). The inhibitory loop is remarkably similar in unrelated families of polypeptide inhibitors, and it is assumed that it takes up

a similar conformation as a substrate for the target protease (*45*). Beyond similarity, we also investigated the potential structural differences between the inhibitory loop and real substrates. The transition from the planar peptide bond to the tetrahedral intermediate necessarily amounts to conformational changes in the polypeptide backbone of the inhibitor. An analogous conformational change is expected during the change from the planar ester bond to the tetrahedral intermediate in the deacylation phase (*23*). The CFT:SGTI complex contains numerous intermolecular hydrogen bonds, yet the homologous hydrogen bonds in the elastase acyl−enzyme are shorter and stronger. Are there structural features in the CFT:SGTI complex that prevent the hydrogen bond shortening and the necessary conformational change in the reactive bond? One likely candidate is the 27I−32I disulfide bridge, which connects the N- and C-terminal ends of the reactive loop, one of the three disulfide bridges in SGTI, which is conserved in the grasshopper family (*33*). The fully conserved Asn-15I residue in this inhibitor family provides an additional anchor to the C-terminal leaving group by coordinating it with a further two hydrogen bonds (Asn-15I $O_{\delta 1}$ to Lys-30I N and Asn-15I $N_{\delta 2}$ to Lys-30I O). These hydrogen bonds are on the diagonally opposite side of the catalytic serine, hence as the carbonyl carbon of the scissile peptide bond approaches the catalytic serine, they have to break (leading to negative cooperativity (*41*)) or the asparagine residue has to move in concert, which may cause strain elsewhere in the inhibitor structure. Additional interactions outside the primary binding region may also stabilize and tether the enzyme:inhibitor complex in an unproductive state (*28*).

The retention of the C-terminal half of the reactive loop in canonical inhibitors may have an additional role. The active site in the inhibitor complex is completely shielded from the solvent; the closest water molecule is 6.6 Å from the carbonyl carbon of the scissile peptide bond. The attacking water in the elastase Ac−NPI acyl−enzyme and the P1′ residue in the CFT:SGTI complex are in overlapping positions (Figure 1B), which means that, if the C-terminal product cannot depart, that also prevents the binding of second substrate, the water molecule. Therefore even if the acylation reaction is successful in the inhibitor complex, as it was demonstrated by Radisky et al. (*46*) for example, the catalytic cycle cannot proceed with the deacylation. Binding sites of the P1′ residue and the nucleophilic water most likely also overlap one another in real substrates; otherwise the departure of the first product would not be strictly required for the deacylation phase to proceed (ping-pong mechanism).

Active Site Residues. Important interatomic distances at the active site are compared in Table 1. The hydrogen bond between His-57 and Asp-102 has been the subject of intense debate. In the CFT:SGTI complex at pH 4.3 (when the histidine is protonated) the length of the His-57 $N_{\delta 1}$ to Asp-102 $O_{\delta 2}$ hydrogen bond is 2.74 Å, which is 0.12 Å longer than structure of subtilisin (pH 6) (*47*), 0.07 Å longer than the elastase Ac−NPI acyl−enzyme (pH 5) (*22*). Identical hydrogen bond distances (2.74−2.77 Å) were measured in *Fusarium oxysporum* trypsin structures at pH 4 and 5 (*48*) and in the crystal structure of α-lytic protease at pH 8 (2.77 Å) (*49*). At first glance the protonation state of the catalytic histidine does not seem to affect its hydrogen bond to the catalytic aspartate. However, if we compare the distances in

Table 1: List of Important Interatomic Distances and Their Estimated Errors for Each Pair of Atoms in the Active Site of the CFT:SGTI and Elastase Acyl−Enzyme (Å)

	CFT:SGTI	elastase acyl−enzyme	distance difference
P1 N to Ser-214 O	3.56 ± 0.04	3.04 ± 0.02	-0.52 ± 0.05
P3 O to residue-216 N	3.14 ± 0.04	2.97 ± 0.02	-0.17 ± 0.04
P3 N to residue-216 O	2.95 ± 0.05	2.71 ± 0.02	-0.24 ± 0.05
His-57 $N_{\epsilon 2}$ to nucleophile attacker	2.72 ± 0.04	2.65 ± 0.04	-0.07 ± 0.06
nucleophile attacker to carbonyl C	2.65 ± 0.05	2.78 ± 0.05	0.13 ± 0.07
Gly-193 N to P1 O	2.75 ± 0.04	2.85 ± 0.02	0.10 ± 0.05
Ser-195 N to P1 O	2.96 ± 0.03	2.72 ± 0.02	-0.24 ± 0.04
His-57 $N_{\delta 1}$ to Asp-102 $O_{\delta 2}$	2.74 ± 0.04	2.67 ± 0.02	-0.07 ± 0.04
His-57 $N_{\delta 1}$ to Asp-102 $O_{\delta 1}$	3.38 ± 0.04	3.47 ± 0.02	0.09 ± 0.04
Ser-214 O_{γ} to Asp-102 $O_{\delta 2}$	2.65 ± 0.03	2.68 ± 0.01	0.03 ± 0.03
Ser-214 O to His-57 $C_{\epsilon 1}$	3.21 ± 0.05	3.10 ± 0.02	-0.11 ± 0.05

the individual enzymes (trypsin and elastase) at higher pH, we find that the length of the His−Asp hydrogen bond also increases. Unfortunately the lack of sufficient data and the different refinement strategies used by different investigators make any reliable statistical analysis of these subtle differences impossible. The short donor−acceptor distance is important, but not the only characteristic of low barrier hydrogen bonds. The position of the hydrogen is also expected to shift toward the acceptor, so that the donor hydrogen and acceptor hydrogen distances become more equal. Such a position was reported in subtilisin (47), elastase acyl−enzyme (22), and proteinase K (50). While it was possible to detect hydrogens on many aromatic residues in the CFT:SGTI complex, on the imidazole ring of His-57 difference Fourier peaks corresponding to hydrogen atoms were absent.

Another type of unusual hydrogen bond has been proposed in serine proteases, namely, a CH−O hydrogen bond between the $C_{\epsilon 1}$ of His-57 and the backbone carbonyl oxygen of Ser-214 (17, 51). In the CFT:SGTI complex the C−O distance is 3.21 ± 0.05 Å, while the H−O distance is 2.66 Å (assuming ideal hydrogen atom position). These distances are among the longest observed in serine proteases (51), which suggests that this particular hydrogen bond is probably very weak in the Michaelis complex. Figure 1B illustrates that there is a shift of the imidazole ring of His-57 within its plane between the two structures. The conformational difference can be best described with a change in the χ_1 torsion angle (CFT:SGTI $70.2 \pm 2.3°$ vs elastase acyl−enzyme $82.9 \pm 1.1°$). This subtle movement is remarkably economic given the fact that during the catalytic cycle the histidine has to adapt to two nucleophilic attackers which are more than 2 Å apart. Bachovchin proposed the moving histidine hypothesis to explain the favorable protonation of the amino leaving group (8). Here we show an example of a conformational change that does not violate the principle of least motion and still prepares the histidine for its new role in activating the second nucleophilic attacker. It is also tempting to assume that the conformational change plays a role in shifting the pK_a of the histidine (52). For example it may change the character of the CH−O bond between His-57 and Ser-214, or the strength of the His-57 $N_{\delta 1}$ to Asp-102 $O_{\delta 2}$ hydrogen bond, which in turn alter the distribution the tautomeric forms of the histidine.

The Influence of the Oxyanion Hole on the Scissile Peptide Bond. The two best characterized enzyme:substrate hydrogen

bonds are formed between the main chain amide nitrogens of Ser-195 and Gly-193 (the so-called oxyanion hole) and the carbonyl oxygen of the scissile peptide bond. In the CFT:SGTI complex the scissile peptide bond is present in the uncleaved form, and there is no detectable residual electron density corresponding to a free carboxyl group or an acyl−ester bond.

Figure 2 illustrates the distinct environment of the scissile peptide bond and the acyl−ester bond during the catalytic cycle. For this purpose the equivalent atoms of the two bonds were superpositioned in the two intermediate complexes. It is immediately obvious that the two nucleophilic attacks (green arrow, acylation; gray arrow, deacylation) occur from exactly opposite directions, perpendicular to the plane of the corresponding bonds. The hydrogen bonds of the oxyanion hole in the Michaelis complex (green plane) are almost perpendicular to the plane of the scissile peptide bond ($74.3 \pm 0.7°$), which is not ideal for the interaction with the sp^2 orbitals of the carbonyl oxygen. On the other hand the oxyanion hydrogen bonds in the acyl−enzyme (gray plane) are closer ($34.0 \pm 0.9°$), but still not perfectly parallel with the plane of the acyl ester bond which offers better, but not too strong stabilization for the sp^2 orbitals. Therefore, the transition from the Michaelis complex to the acyl−enzyme intermediate eventually leads to a rotation of the carbonyl bond within the oxyanion hole. This twisting effect on the carbonyl bond and the relative stabilization of the acyl−ester carbonyl bond may help to push the equilibrium toward the acyl−enzyme.

The polarizing effect of the oxyanion hole on the carbonyl bond in the acyl−enzyme has been known for some time based on Raman and absorption spectroscopy (53, 54). The strength of the hydrogen bonds and hence the polarizing effect was also found to correlate with the rate of deacylation. In the CFT:SGTI structure the Gly-193 amide hydrogen bond is the shorter or stronger (Table 1) while in the elastase Ac−NPI acyl−enzyme the Ser-195 amide hydrogen bond is stronger, and in both cases the orientation of the stronger hydrogen bond approximately coincides with the direction of the nucleophilic attack (Figure 2). Our results suggest that the directionality of the polarization could also be a factor for the efficient nucleophilic attack.

There are indications of polarization of the C=O bond in the X-ray structure. Figure 3 shows the normalized *B*-factor difference ($2(B_C - B_O)/(B_C + B_O)$) between the carbonyl carbon and the oxygen in the P3−P2′ region of the inhibitor. This value also deviates significantly from the observed mean value of other peptide bonds in the structure (Figure 3A). The higher than average *B*-factor difference shows that the carbonyl oxygen has a lower *B*-factor relative to the carbonyl carbon compared to other peptide carbonyl bonds, which could mean that the bonding electrons are shifted to the carbonyl oxygen. Therefore the 2mFo-DFc electron density difference between the carbonyl oxygen and carbon was also investigated.

This value is 1.3 e$^-$/Å^3 in contrast to an average of 0.9 e$^-$/Å^3 with a standard deviation of 0.4 e$^-$/Å^3 in all other peptide bonds of the structure (the rms electron density (σ) of the map is 0.52 e$^-$/Å^3). Therefore the carbonyl oxygen atom at the scissile peptide bond has on average 0.4 e$^-$/Å^3 higher electron density than a peptide bond in other regions of the structure. To validate this method further, we statisti-
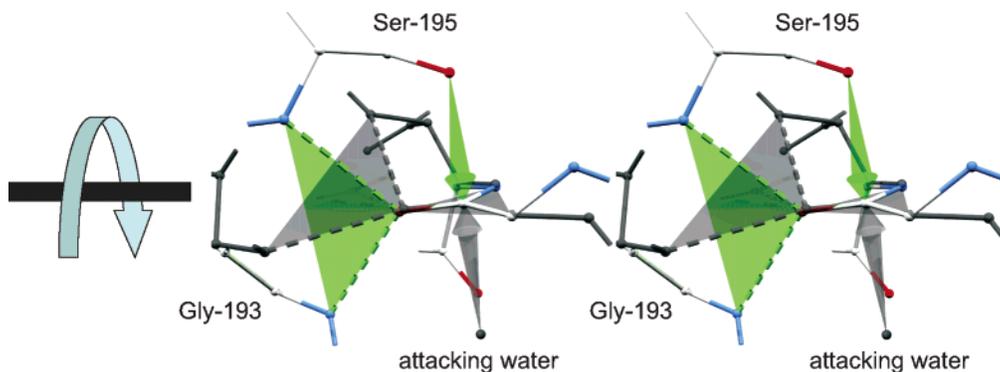
FIGURE 2: Superposition of the scissile peptide bond and the acyl−ester bond (stereoview). The green arrow shows the direction of the nucleophilic attack in the acylation phase, while the gray arrow indicates the direction of approach of the attacking water molecule in the deacylation phase. The CFT:SGTI model is colored according to atom types, and the elastase Ac−NPI acyl−enzyme is dark gray. Planes highlight the planarity of the peptide bond and the oxyanion hydrogen bonds (green acylation, gray deacylation). The curved blue arrow illustrates the torsional difference of the C_α−C−O−Ser-195 N angle between the Michaelis complex and the acyl−enzyme. The figure was produced with Deep Viewer (*57*) and Povray (http://www.povray.org).
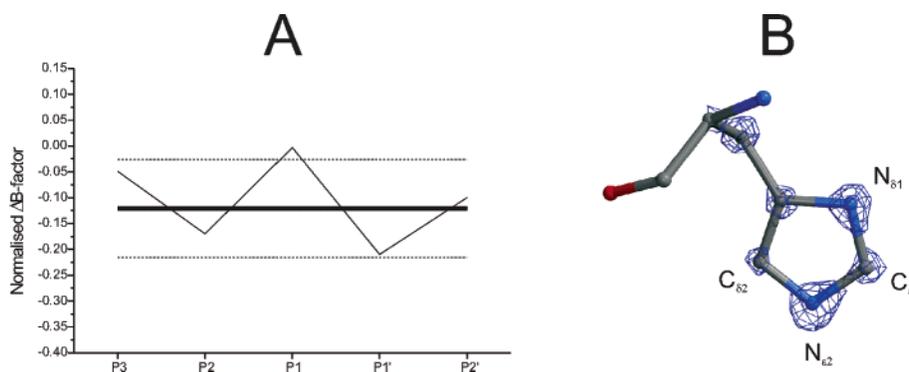

FIGURE 3: (A) Normalized *B*-factor difference between the carbonyl carbon and carbonyl oxygen of the P3−P2′ residues of SGTI. The thick line denotes the average normalized *B*-factor difference based on the whole structure, and the dashed lines are the standard deviations from the mean value. (B) 2mFo-DFc electron density map of His-57 contoured at 3.5σ. The positions of the nitrogen and carbon atoms can be directly identified by the size of the electron density peak. The figure was created with Bobscript (*58*).

cally analyzed the visually obvious electron density difference of carbon and nitrogen atoms of the imidazole ring of histidine residues (Figure 3B). The comparison of all $N_{\delta1}$−$C_{\epsilon2}$ and $N_{\epsilon2}$−$C_{\delta1}$ electron density differences yielded an average of 0.5 $e^-/Å^3$ with a standard deviation of 0.3 $e^-/Å^3$, which is due to the approximately one electron difference between the carbon and nitrogen atoms. While the polarization effect is detected by the electron density analysis, the extent of the polarization is not unique in the structure. Presumably structural hydrogen bonds may be also capable of polarizing other ordinary peptide bonds. From an evolutionary perspective this could mean that it is relatively easy to "convert" ordinary hydrogen bonds to a primordial oxyanion hole. While resonance Raman spectroscopy also indicated an increase of the carbonyl bond length (0.015−0.025 Å) (*53*), the change of this magnitude unfortunately falls outside the range of confidence of our crystallographic analysis.

It is interesting to note that the comparison of bond lengths and angles with ideal values shows that the acyl−ester bond is considerably more strained than the scissile peptide bond in the CFT:SGTI structure. For example in the acyl−enzyme the distance from the carbonyl carbon to the hydroxyl oxygen of the catalytic serine is 1.25 ± 0.03 Å vs 1.34 Å observed in small molecule esters (*22*). Both the acyl−ester and scissile peptide bonds are planar within coordinate error. In the CFT:SGTI complex the carbonyl carbon is only 0.01 ± 0.02 Å

Table 2: Comparison of Bond Lengths and Angles of the Scissile Peptide Bond in the CFT:SGTI Complex with the Ideal Values

| | Bond Lengths (Å) | | | | |
	inhibitor	ideal		inhibitor	ideal
N−C	1.35 ± 0.05	1.33	C_α−C	1.52 ± 0.05	1.52
C−O	1.26 ± 0.05	1.23	N−C_α'	1.50 ± 0.05	1.45

| | Bond Angles (deg) | | | | |
	inhibitor	ideal		inhibitor	ideal
O−C−N	120.3 ± 2.2	123	C_α−C−N	116.8 ± 2.1	116
C_α−C−O	123.0 ± 1.9	121	C−N−C_α'	118.1 ± 1.9	122

away from the plane defined by the carbonyl oxygen, the C_α of Arg-29, and N of Lys-30 (Table 2). The C_α−N−C angle of the scissile peptide bond is sharper by about 5° than the ideal value probably due to the orbital reorganization caused by the oxyanion hole and vicinity of the serine hydroxyl group. The distance of the carbonyl carbon to the hydroxyl group is 2.65 Å, which is 0.13 Å shorter than to the attacking water in the elastase Ac−NPI acyl−enzyme, signifying a stronger interaction.

CONCLUSION

Due to decades of intense research, serine protease catalysis became one of the best characterized enzymatic reactions. Recently a number of new serine protease structures have been solved at atomic resolution and have provided fresh mechanistic insight for studying serine

Fodor et al.

protease catalysis. The atomic resolution structure of the CFT:SGTI complex falls within this category, and detailed comparison to an elastase acyl—enzyme (*22*) revealed that the hydrogen bonds of the enzyme:substrate β-sheet become significantly shorter (stronger) in the acyl—enzyme. This observation leads to the hypothesis that the simultaneous strengthening of these hydrogen bonds helps to overcome the activation energy barrier in the acylation phase. Serine proteases may continuously scan polypeptides for potential cleavage sites and form transient β-sheets with them, but only those peptides, which can make tighter, stronger interactions with the enzyme, including their β-sheet, are actually hydrolyzed. On the other hand, SGTI, a canonic inhibitor, is able to form many good interactions with its target enzyme, except where it is crucial: in the proximity of the scissile bond. The β-sheet appears to be locked in a persistent loose position, which may be analogous to the catalytic Michaelis complex. The comparison of the two structures also revealed that in the acyl—enzyme the oxyanion hole stabilizes better the sp^2 orbitals at the carbonyl oxygen than in the scissile peptide bond. Together these interactions make the acyl—enzyme more stable, which favors the population of the acyl—enzyme state. The geometry of the scissile peptide bond does not deviate significantly from the consensus peptide bond geometry in the model of the Michaelis complex, and except for the shorter than van der Waals distance between the nucleophilic attacker and the reactive bond, there is no sign of geometric strain. By contrast, we found indications for the electrostatic polarization of the carbonyl moiety of the scissile peptide bond. This observation lends support to Warshel's original suggestion (*55*) that enzymes function by providing a preorganized electrostatic environment.

The hydrogen bond between the catalytic histidine and aspartate did not show any special character, although the position of the hydrogen atom could not be directly observed in the CFT:SGTI structure. The side chain of His-57 appears to be shifted in the acyl—enzyme intermediate compared to the native enzyme and enzyme:inhibitor complex as it adapts its conformation to the different nucleophilic attackers. While the present atomic resolution structures do not tell anything about the position of the histidine in the tetrahedral intermediate of the acylation phase, they reveal that the active site histidine enjoys a certain degree of freedom and responds to the binding of the second substrate, the water molecule.

The activity of serine proteases must be tightly controlled in the cell or the extracellular matrix. Nonspecific cleavage in proteins may lead to disruption of signaling processes or precipitation, or they may threaten the integrity of cellular or extracellular matrix. Therefore understanding the full range of molecular mechanisms that govern substrate recognition and hydrolysis is essential. Since all serine proteases share highly homologous enzyme:substrate interactions, their potential impact on catalysis and specificity can be easily generalized and may lead to the design of more efficient inhibitors or substrates for therapeutic purposes.

REFERENCES

1. Schechter, I., and Berger, A. (1967) On the size of the active site in proteases. I. Papain, *Biochem. Biophys. Res. Commun. 27*, 157–162.
2. Wright, C. S., Alden, R. A., and Kraut, J. (1969) Structure of subtilisin BPN' at 2.5 angstrom resolution, *Nature 221*, 235–242.
3. Henderson, R. (1970) Structure of crystalline alpha-chymotrypsin. IV. The structure of indoleacryloyl-alpha-chyotrypsin and its relevance to the hydrolytic mechanism of the enzyme, *J. Mol. Biol. 54*, 341–354.
4. Kraut, J., Robertus, J. D., Birktoft, J. J., Alden, R. A., Wilcox, P. E., and Powers, J. C. (1972) The aromatic substrate binding site in subtilisin BPN' and its resemblance to chymotrypsin, *Cold Spring Harb. Symp. Quant. Biol. 36*, 117–123.
5. Stroud, R. M., Kay, L. M., and Dickerson, R. E. (1972) The crystal and molecular structure of DIP-inhibited bovine trypsin at 2.7 Angstrom resolution, *Cold Spring Harb. Symp. Quant. Biol. 36*, 125–140.
6. Huber, R., Kukla, D., Bode, W., Schwager, P., Bartels, K., Deisenhofer, J., and Steigemann, W. (1974) Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor: II. Crystallographic refinement at 1.9 Å resolution, *J. Mol. Biol. 89*, 73.
7. Hedstrom, L. (2002) Serine protease mechanism and specificity, *Chem. Rev. 102*, 4501–4524.
8. Bachovchin, W. W. (1986) 15N NMR spectroscopy of hydrogen-bonding interactions in the active site of serine proteases: evidence for a moving histidine mechanism, *Biochemistry 25*, 7751–7759.
9. Blow, D. M., Birktoft, J. J., and Hartley, B. S. (1969) Role of a buried acid group in the mechanism of action of chymotrypsin, *Nature 22*, 337–340.
10. Craik, C. S., Roczniak, S., Largman, C., and Rutter, W. J. (1987) The catalytic role of the active site aspartic acid in serine proteases, *Science 237*, 909–913.
11. Sprang, S., Standing, T., Fletterick, R. J., Stroud, R. M., Finer-Moore, J., Xuong, N.-H., Hamlin, R., Rutter, W. J., and Craik, C. S. (1987) The three-dimensional structure of Asn102 mutant of trypsin: role of Asp102 in serine protease catalysis, *Science 237*, 905–909.
12. Frey, P. A. (2004) Strong hydrogen bonding in chymotrypsin and other serine proteases, *J. Phys. Org. Chem. 17*, 511–520.
13. Frey, P. A., Whitt, S. A., and Tobin, J. B. (1994) A low-barrier hydrogen bond in the catalytic triad of serine proteases, *Science 264*, 1927–1930.
14. Ash, E. L., Sudmeier, J. L., DeFabo, E. C., and Bachovchin, W. W. (1997) A low-barrier hydrogen bond in the catalytic triad of serine proteases? Theory versus experiment, *Science 278*, 1128–1132.
15. Stein, R. L., Strimpler, A. M., Hori, H., and Powers, J. C. (1987) Catalysis by human leukocyte elastase: proton inventory as a mechanistic probe, *Biochemistry 26*, 1305–1314.
16. Polgar, L. (1971) On the mechanism of proton transfer in the catalysis by serine proteases, *J. Theor. Biol. 31*, 165–169.
17. Ash, E. L., Sudmeier, J. L., Day, R. M., Vincent, M., Torchilin, E. V., Haddad, K. C., Bradshaw, E. M., Sanford, D. G., and Bachovchin, W. W. (2000) Unusual H-1 NMR chemical shifts support (His) C-epsilon 1-H center dot center dot center dot O = C H-bond: Proposal for reaction-driven ring flip mechanism in serine protease catalysis, *Proc. Natl. Acad. Sci. U.S.A. 97*, 10371–10376.
18. Graf, L. (1995) Structural basis of serine protease action: the fourth dimension, in *Natural Sciences and Human Thought* (Zwilling, R., Ed.) pp 139–148, Heidelberg, Springer-Verlag, Berlin.
19. Topf, M., Varnai, P., and Richards, W. G. (2002) Ab initio QM/MM dynamics simulation of the tetrahedral intermediate of serine proteases: Insights into the active site hydrogen-bonding network, *J. Am. Chem. Soc. 124*, 14780–14788.
20. Wilmouth, R. C., Clifton, I. J., Robinson, C. V., Roach, P. L., Aplin, R. T., Westwood, N. J., Hajdu, J., and Schofield, C. J. (1997) Structure of a specific acyl-enzyme complex formed between beta-casomorphin-7 and porcine pancreatic elastase, *Nat. Struct. Biol. 4*, 456–462.

21. Marquart, M., Walter, J., Diesenhofer, J., Bode, W., and Huber, R. (1983) The geometry of the reactive site and peptide groups in trypsin, trypsinogen and its complexes with inhibitors, *Acta Crystallogr. B 39*, 480−490.

22. Katona, G., Wilmouth, R. C., Wright, P. A., Berglund, G. I., Hajdu, J., Neutze, R., and Schofield, C. J. (2002) X-ray structure of a serine protease acyl-enzyme complex at 0.95-Å resolution, *J. Biol. Chem. 277*, 21962−21970.

23. 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.

24. Hamdaoui, A., Wataleb, S., Devreese, B., Chiou, S. J., Vanden Broeck, J., Van Beeumen, J., De Loof, A., and Schoofs, L. (1998) Purification and characterization of a group of five novel peptide serine protease inhibitors from ovaries of the desert locust, Schistocerca gregaria, *FEBS Lett. 422*, 74−78.

25. Malik, Z., Amir, S., Pal, G., Buzas, Z., Varallyay, E., Antal, J., Szilagyi, Z., Vekey, K., Asboth, B., Patthy, A., and Graf, L. (1999) Proteinase inhibitors from desert locust, Schistocerca gregaria: engineering of both P(1) and P(1)′ residues converts a potent chymotrypsin inhibitor to a potent trypsin inhibitor, *Biochim. Biophys. Acta 1434*, 143−150.

26. Patthy, A., Amir, S., Malik, Z., Bodi, A., Kardos, J., Asboth, B., and Graf, L. (2002) Remarkable phylum selectivity of a Schistocerca gregaria trypsin inhibitor: the possible role of enzyme-inhibitor flexibility, *Arch. Biochem. Biophys. 398*, 179−187.

27. Gaspari, Z., Patthy, A., Graf, L., and Perczel, A. (2002) Comparative structure analysis of proteinase inhibitors from the desert locust, Schistocerca gregaria, *Eur. J. Biochem. 269*, 527−537.

28. Fodor, K., Harmat, V., Hetényi, C., Kardos, J., Antal, J., Perczel, A., Patthy, A., Katona, G., and Gráf, L. (2005) Extended interactions in a serine protease-canonical inhibitor complex account for high specificity and strong binding, *J. Mol. Biol. 350*, 156−169.

29. Niekamp, C. W., Hixson, H. F., Jr., and Laskowski, M., Jr. (1969) Peptide-bond hydrolysis equilibria in native proteins. Conversion of virgin into modified soybean trypsin inhibitor, *Biochemistry 8*, 16−22.

30. Mattis, J. A., and Laskowski, M., Jr. (1973) pH dependence of the equilibrium constant for the hydrolysis of the Arg 63 -Ile reactive-site peptide bond in soybean trypsin inhibitor (Kunitz), *Biochemistry 12*, 2239−2245.

31. Helland, R., Leiros, I., Berglund, G. I., Willassen, N. P., and Smalas, A. O. (1998) The crystal structure of anionic salmon trypsin in complex with bovine pancreatic trypsin inhibitor, *Eur. J. Biochem. 256*, 317−324.

32. Coombs, G. S., Rao, M. S., Olson, A. J., Dawson, P. E., and Madison, E. L. (1999) Revisiting catalysis by chymotrypsin family serine proteases using peptide substrates and inhibitors with unnatural main chains, *J. Biol. Chem. 274*, 24074−24079.

33. Laskowski, M., and Qasim, M. A. (2000) What can the structures of enzyme-inhibitor complexes tell us about the structures of enzyme substrate complexes?, *Biochim. Biophys. Acta 1477*, 324−337.

34. Scheldrick, G., and Schneider, T. (1997) SHELXL: High-Resolution Refinement, *Methods Enzymol. 277*, 319−343.

35. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Res. 28*, 235−242.

36. Schneider, T. R. (2000) Objective comparison of protein structures: error-scaled difference distance matrices, *Acta Crystallogr. D 56* (Part 6), 714−721.

37. Kleywegt, G. J. (1996) xdlMAPMAN and xdlDATAMAN−Programs for Reformatting, Analysis and Manipulation of Biomacromolecular Electron-Density Maps and Reflection Data Sets, *Acta Crystallogr. D 52*, 826−828.

38. Kleywegt, G. J., and Jones, T. A. (1994) A super position, *CCP4/ESF-EACBM Newsl. Protein Crystallogr. 31*, 9−14.

39. Jones, T. A., and Kjeldgaard, M. (1997) Electron-density map interpretation, *Methods Enzymol. 277*, 173−208.

40. Perona, J. J., Hedstrom, L., Rutter, W. J., and Fletterick, R. J. (1995) Structural origins of substrate discrimination in trypsin and chymotrypsin, *Biochemistry 34*, 1489−1499.

41. Williams, D. H., Stephens, E., O'Brien, D. P., and Zhou, M. (2004) Understanding noncovalent interactions: ligand binding energy and catalytic efficiency from ligand-induced reductions in motion within receptors and enzymes, *Angew Chem., Int. Ed. Engl. 43*, 6596−6616.

42. Stein, R. L., and Strimpler, A. M. (1987) P1 Residue Determines the Operation of the Catalytic Triad of Serine Proteases during hydrolyses of Acyl-Enzymes, *J. Am. Chem. Soc. 109*, 4387−4390.

43. Read, R. J., Fujinaga, M., Sielecki, A. R., and James, M. N. (1983) Structure of the complex of Streptomyces griseus protease B and the third domain of the turkey ovomucoid inhibitor at 1.8-Å resolution, *Biochemistry 22*, 4420−4433.

44. Luthy, J. A., Praissman, M., Finkenstadt, W. R., and Laskowski, M., Jr. (1973) Detailed mechanism of interaction of bovine -trypsin with soybean trypsin inhibitor (Kunitz). I. Stopped flow measurements, *J. Biol. Chem. 248*, 1760−1771.

45. Krowarsch, D., Cierpicki, T., Jelen, F., and Otlewski, J. (2003) Canonical protein inhibitors of serine proteases, *Cell. Mol. Life Sci. 60*, 2427−2444.

46. Radisky, E. S., and Koshland, D. E., Jr. (2002) A clogged gutter mechanism for protease inhibitors, *Proc. Natl. Acad. Sci. U.S.A. 99*, 10316−10321.

47. Kuhn, P., Knapp, M., Soltis, S. M., Ganshaw, G., Thoene, M., and Bott, R. (1998) The 0.78 Å structure of a serine protease: Bacillus lentus subtilisin, *Biochemistry 37*, 13446−13452.

48. Schmidt, A., Jelsch, C., Ostergaard, P., Rypniewski, W., and Lamzin, V. S. (2003) Trypsin revisited: crystallography AT (SUB) atomic resolution and quantum chemistry revealing details of catalysis, *J. Biol. Chem. 278*, 43357−43362.

49. Fuhrmann, C. N., Kelch, B. A., Ota, N., and Agard, D. A. (2004) The 0.83 Å resolution crystal structure of alpha-lytic protease reveals the detailed structure of the active site and identifies a source of conformational strain, *J. Mol. Biol. 338*, 999−1013.

50. Betzel, C., Gourinath, S., Kumar, P., Kaur, P., Perbandt, M., Eschenburg, S., and Singh, T. P. (2001) Structure of a serine proteinase proteinase K from Tritirachium album limber at 0.98 Å resolution, *Biochemistry 40*, 3080−3088.

51. Derewenda, Z. S., Derewenda, U., and Kobos, P. M. (1994) (His)C-Epsilon-H.O=C Hydrogen-Bond in the Active-Sites of Serine Hydrolases, *J. Mol. Biol. 241*, 83−93.

52. Lin, J., Cassidy, C. S., and Frey, P. A. (1998) Correlations of the basicity of His 57 with transition state analogue binding, substrate reactivity, and the strength of the low-barrier hydrogen bond in chymotrypsin, *Biochemistry 37*, 11940−11948.

53. Tonge, P. J., and Carey, P. R. (1992) Forces, bond lengths, and reactivity: fundamental insight into the mechanism of enzyme catalysis, *Biochemistry 31*, 9122−9125.

54. Whiting, A. K., and Peticolas, W. L. (1994) Details of the acyl-enzyme intermediate and the oxyanion hole in serine protease catalysis, *Biochemistry 33*, 552−561.

55. Warshel, A. (1998) Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites, *J. Biol. Chem. 273*, 27035−27038.

56. Fenn, T. D., Ringe, D., and Petsko, G. A. (2003) POVScript+: a program for model and data visualization using persistence of vision ray-tracing, *J. Appl. Crystallogr. 36*, 944−947.

57. Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling, *Electrophoresis 18*, 2714−2723.

58. Esnouf, R. M. (1997) An extensively modified version of MolScript that includes greatly enhanced coloring capabilities, *J. Mol. Graphics Modell. 15*, 132−134, 112−113.