Quantifying tetrahedral adduct formation and stabilization in the Cysteine and the Serine Proteases

Jennifer A. Cleary[a], William Doherty[b], Paul Evans[b] and J.Paul G. Malthouse[a]*

[a]School of Biomolecular and Biomedical Science, Centre for Synthesis and Chemical Biology, Conway Institute, University College Dublin, Dublin 4, Ireland.
[b]School of Chemistry, Centre for Synthesis and Chemical Biology, Conway Institute, University College Dublin, Dublin 4, Ireland.

*Corresponding author. UCD School of Biomolecular and Biomedical Science, UCD Centre for Synthesis and Chemical Biology, Conway Institute, University College Dublin, Dublin 4, Ireland. Tel.: 0035317166872; fax 0035317166893

E-mail address: J.Paul.G.Malthouse@ucd.ie (J.P.G. Malthouse)

ABSTRACT

Two new papain inhibitors have been synthesized where the terminal α-carboxyl groups of Z-Phe-Ala-COOH and Ac-Phe-Gly-COOH have been replaced by a proton to give Z-Phe-Ala-H and Ac-Phe-Gly-H. We show that for papain replacing the terminal carboxylate group of a peptide inhibitor with a hydrogen atom decreases binding 3-4 fold while replacing an aldehyde or glyoxal group with a hydrogen atom decreases binding by 300,000-1,000,000 fold. Thiohemiacetal formation by papain with aldehyde or glyoxal inhibitors is shown to be ~10,000 times more effective than hemiacetal or hemiketal formation with chymotrypsin. It is shown using effective molarities, that for papain, thiohemiacetal stabilisation is more effective with aldehyde inhibitors than with glyoxal inhibitors. The effective molarity obtained when papain is inhibited by an aldehyde inhibitor is similar to the effective molarity obtained when chymotrypsin is inhibited by glyoxal inhibitors showing that both enzymes can stabilize tetrahedral adducts by similar amounts. Therefore the greater potency of aldehyde and glyoxal inhibitors with papain is not due to greater thiohemiacetal stabilization by papain compared to the hemiketal and hemiacetal stabilization by chymotrypsin, instead it reflects the greater intrinsic reactivity of the catalytic thiol group of papain compared to the catalytic hydroxyl group of chymotrypsin. It is argued that while the hemiacetals and thiohemiacetals formed with the serine and cysteine proteases respectively can mimic the catalytic tetrahedral intermediate they are also analogues of the productive and non-productive acyl intermediates that can be formed with the cysteine and serine proteases.

Key words: aldehyde inhibitor; glyoxal inhibitor; tetrahedral intermediate; cysteine protease; serine protease.

Abbreviations: Z, benzyloxycarbonyl; Ac, Acetyl: Ac-Cys, N-Acetyl-L-Cysteine; -COCHO, α-keto-β-aldehyde; ZAAF -COOH, Z-Ala-Ala-NHCH(CH$_2$C$_6$H$_5$)-COOH; ZAAF-H, ZAla-Ala-NHCH$_2$CH$_2$C$_6$H$_5$; ZAAF-CHO, Z-Ala-Ala-NHCH(CH$_2$C$_6$H$_5$)-CHO; ZAAF-COCHO, Z-Ala-Ala-NHCH(CH$_2$C$_6$H$_5$)-COCHO; ZAPF -COOH, Z-Ala-Pro-NHCH(CH$_2$C$_6$H$_5$)-COOH; ZAPF-H, Z-Ala-Ala-NHCH$_2$CH$_2$C$_6$H$_5$; ZAPF-COCHO, Z-Ala-Pro-NHCH(CH$_2$C$_6$H$_5$)-COCHO; ZFA-COOH, Z-Phe-NHCH(CH$_3$)-COOH; ZFA-H, Z-Phe-NHCH$_2$CH$_3$; ZFA-COCHO, Z-Phe-NHCH(CH$_3$)-COCHO; Ac-FG-COOH, Ac-Phe-NHCH$_2$-COOH; Ac-FG-H, Ac-Phe-NHCH$_3$; Ac-FG-CHO, Ac-Phe-NHCH$_2$-CHO.

1. Introduction

Specific substrate derived aldehyde inhibitors are usually highly potent and specific inhibitors of the cysteine proteases. Therefore it is not surprising that they have been used to specifically inhibit cysteine protease such as the NOV protease[1], the SARS 3Cl protease[2] and the falcipains[3] which are potential targets for treating gastroenteritis, severe acute respiratory syndrome and malaria.

Recently we have developed a method for quantifying the contribution of a peptide warhead to the potency of an inhibitor [4]. This has allowed us to determine how the aldehyde-, carboxylate- and glyoxal groups contribute to inhibitor potency with the serine proteases [4, 5]. In this paper, we use the same approach to quantify how the carboxylate-, aldehyde- and glyoxal groups contribute to inhibitor potency with the cysteine protease papain.

For both the cysteine and serine proteases catalysis proceeds via an acyl intermediate whose formation and breakdown proceeds via a tetrahedral intermediate. The formation or breakdown of the tetrahedral intermediate is thought to be rate limiting. Westerik and Wolfenden [6] first suggested that the tight binding of aldehyde inhibitors to papain might be attributed to formation of a thiohemiacetal analogous to the tetrahedral intermediate formed during catalysis. Saturation transfer experiments suggested that papain formed a thiohemiacetal with aldehyde inhibitors [7]. This was confirmed by the direct observation of thiohemiacetal formation with papain [8, 9] and with the hepatitis A virus protease [10] using ^{13}C-NMR. Likewise ^{13}C-NMR has been used to show that aldehyde inhibitors can form two diastereomeric hemiacetals with the serine proteases [5, 11, 12]. However, with the cysteine protease papain thiohemiacetal formation was stereospecific with only one diastereoisomeric thiohemiacetal being formed with both aldehyde [9] and glyoxal inhibitors [13]. Likewise with glyoxal inhibitors hemiketal formation with the serine proteases has been shown to be stereospecific [14-16] and stabilised by up to a factor of ~200,000 [4] by enzyme interactions.

Early X-ray crystallographic studies have shown that both the serine proteases chymotrypsin and subtilisin [17, 18] and also the cysteine protease papain[19, 20] have an "oxyanion hole" which contain

groups which hydrogen bond to the oxyanion of the tetrahedral intermediate formed during catalysis. It has been argued [21] that with substrates or inhibitors with large groups extending into the S1' subsite the carbonyl oxygen must point into the oxyanion hole as the bulky leaving group is too large to occupy the oxyanion hole. However, with aldehyde inhibitors the aldehyde proton could fit in the oxyanion hole and the carbonyl oxygen could point towards the catalytic histidine residue [21].

More recent X ray crystallographic studies have shown that with the serine proteases and aldehyde inhibitors two hemiacetals can be formed, one with the oxyanion oxygen in the oxyanion hole and the other with it hydrogen bonding to $N\varepsilon 2$ of the catalytic histidine [22-26]. Likewise with the cysteine proteases two thiohemiacetals can be formed in the same way [1, 27-30].

There has been considerable controversy about whether tetrahedral intermediate stabilization is more effective in the serine proteases than in the cysteine proteases[31-35]. Therefore in this paper we have quantified thiohemiacetal formation with both aldehyde and glyoxal inhibitors of papain. The effective molarity of the catalytic thiol group of papain is the molarity of an analogous low molecular weight thiol group which would be required to be as effective as the catalytic thiol group of papain under the same experimental conditions. For chymotrypsin the effective molarity is the molarity of water required to be as effective as the catalytic serine hydroxyl group of chymotrypsin. By comparing the effective molarity of the catalytic thiol group of papain with that of the catalytic serine hydroxyl group of chymotrypsin we intend to quantify the formation and stabilization of tetrahedral adducts by these enzymes. This should allow us to determine whether or not papain and chymotrypsin stabilize the tetrahedral adducts of aldehyde and glyoxal peptide inhibitors by similar amounts.

2. Materials and methods

2.1. Materials

All materials were obtained from Sigma-Aldrich Chemical Co., Gillingham, Dorset, U.K. L-[1-^{13}C]Alanine (99 at.%) was obtained from Cambridge Isotope Laboratories, Inc (50 Frontage Road, Andover, MA 01810-5413 USA).

2.2. Synthesis of Z-Phe-Ala-CO^{13}CHO, Z-Phe-Ala-H and Ac-Phe-Gly-H

Z-Phe-Ala-CO^{13}CHO was synthesized as described previously [13]. Z-Phe-Ala-H and Ac-Phe-Gly-H were synthesised by coupling ethylamine and methylamine to Z-Phe and Ac-Phe- respectively, using 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride as a coupling reagent[36, 37]. Ac-Phe-Gly-^{13}CHO was synthesized by converting Ac-Phe-Gly-^{13}COOH to its Weinreb amide[38] which was then reduced using lithium aluminum hydride [39] to give the aldehyde.

2.3. NMR spectra of Z-Phe-Ala-H and Ac-Phe-Gly-H

^{13}C-NMR analysis of Z-Phe-Ala-H gave the folowing data (75.475 MHz, CDCl$_3$) δ: 14.4 (1C, CH$_3$CH$_2$), 34.3 (1C, CH$_3$$CH_2$), 38.9 (1C, CH$CH_2$Ph), 56.4 (1C, C$_6H_5CH_2$$C$H), 66.9 (1C, O-$CH_2$Ph), 126.9-129.3 (10C, CH=CH), 136.1 (1C, CH=C=), 136.6 (1C, CH=C=) 155.9 (1C, O-CO-NH), 170.5 (1C, CO-NH).

^{13}C-NMR analysis of Ac-Phe-Gly-H gave the folowing data (125.772 MHz, d$_6$-DMSO) δ: 22.5 (1C, CH$_3$$C$O), 25.6 (1C, CH$_3$NH), 37.8 (1C, CHCH$_2$Ph), 54.2 (1C, C$_6H_5CH_2$$C$H), 126.2-129.1 (5C, CH=CH), 138.2 (1C, CH=C=) 169.1 (1C, CH$_3$$C$O), 171.7 (1C, CO-NH).

2.4. Synthesis of Ac-Phe-Gly-^{13}CON(OMe)Me

To a stirred solution of Ac-Phe-Gly-^{13}COOCH$_3$ (220 mg, 0.79 mmol, 1 equiv.) in methanol (4 mL) caesium carbonate (515 mg, 1.58 mmol, 2 equiv.) in distilled water (4 mL) was added. The solution was stirred for 24 hours at room temperature and then the solvent was removed *in vacuo*. The residue was

acidified with 1 M HCl (20 mL) and extracted with ethyl acetate (3 × 20 mL). The combined organic layers were washed with brine (20 mL), dried over MgSO$_4$, filtered and then the solvent was removed *in vacuo* to give the carboxylic acid (209 mg, 0.79 mmol, 100%). This was dissolved in dichloromethane (8 mL) and treated with 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (165 mg, 0.86 mmol, 1.1 equiv.) followed by Hydroxybenzotriazole hydrate (133 mg, 0.86 mmol, 1.1 equiv.). After 10 minutes of stirring, *N,O*-dimethylhydroxylamine hydrochloride (85 mg, 0.87 mmol, 1.1 equiv.) was added followed by *N,N*-Diisopropylethylamine (0.20 mL, 1.15 mmol, 1.4 equiv.) and the reaction mixture was left to stir overnight at room temperature. The mixture was diluted with dichloromethane (35 mL), washed with 1 M HCl (10 mL) and the aqueous layer was extracted with dichloromethane (40 mL). The combined organic layers were washed with brine (20 mL), dried over MgSO$_4$, filtered and then the solvent was removed *in vacuo* to give the crude product. Purification by silica gel chromatography (EtOAc → EtOAc/MeOH; 4:1) gave the product as an off white crystalline solid (143 mg, 59%). ^{13}C NMR analysis gave the following data δ_C (100.577 MHz, CDCl$_3$): 23.2 (1C, $\underline{C}$H$_3$CO), 32.5 (1C, N$\underline{C}$H$_3$), 38.5 (1C, C$_6$H$_5\underline{C}$H$_2$), 40.8 (d, J_{CC} = 54 Hz, 1C, NH$\underline{C}$H$_2$), 54.3 (1C, $\underline{C}$HCO), 61.6 (1C, NO$\underline{C}$H$_3$), 127.0-129.3 (5C, $\underline{C}$H=CH), 136.7 (1C, CH=$\underline{C}$=), 169.3 (1C, CH$_2$ - $^{13}\underline{C}$O), 170.1 (1C, CH$_3\underline{C}$O), 171.3 (1C, CH$\underline{C}$ONH)

2.5. Synthesis of Ac-Phe-Gly-^{13}CHO

The Ac-Phe-Gly-^{13}CON(OMe)Me (85 mg, 0.28 mmol, 1 equiv.) was stirred in dry tetrahydrofuran (5 mL) at 0 °C under nitrogen. Lithium aluminium hydride (15 mg, 0.40 mmol, 1.4 equiv.) was added to the flask and the mixture was stirred for 1 hour. The reaction was quenched by dropwise addition of 1 M KHSO$_4$ (2 mL) followed by water (5 mL). The resulting mixture was extracted with ethyl acetate (3 × 10 mL). The combined organic layers were washed with saturated aqueous NaHCO$_3$ (10 mL), brine (10 mL), dried over MgSO$_4$, filtered and solvent was removed *in vacuo* to give the crude product. Purification by silica gel chromatography (CH$_2$Cl$_2$/MeOH; 10:1) gave the product as a white foam (27 mg, 39%). This material was immediately dissolved in 1 mM HCl containing 10% (v/v) deuterium oxide (~ 2 mL) in order to prevent polymerisation. ^{13}C NMR analysis gave the following data δ_C (125.758 MHz, 1 mM

HCl/D$_2$O; 9:1): δ 21.6 (1C, $\underline{C}$H$_3$CO), 37.1 (1C, C$_6$H$_5$$\underline{C}H_2$), 44.9 (d, J_{CC} = 47.3 Hz, 1C, NH$\underline{C}$H$_2$), 55.5 (1C, $\underline{C}$HCO), 88.2 (1C, - $^{13}\underline{C}$H(OH)$_2$), 127.0-129.0 (5C, $\underline{C}$H=CH), 136.4 (1C, CH=$\underline{C}$=), 173.4 (1C, CH$_3$$\underline{C}$O), 174.0 (1C, CH$\underline{C}$ONH), 200.8 (1C, - $^{13}\underline{C}$HO)

2.6. Papain solutions

Papain was obtained from the Sigma-Aldrich Chemical Company as a crystallized and lyophilized powder (Lot no. 098K70201). Commercial papain contains irreversibly denatured papain and papain reversibly inhibited by forming a mixed disulfide with cysteine. Consequently, it is nescessary to activate the reversibly inhibited papain by incubating it with a large excess of cysteine at alkaline pH. Therefore ~50 uM Papain was activated by incubation for 30 minutes in 0.1 M Tris / HCl buffer containing 30 mM EDTA and 50 mM N-Acetyl-L-cysteine at pH 8.3. The amount of papain was estimated using a (E$_{280}$ = 56,000 M^{-1}cm^{-1} [40]. A mean value (12 determinations) of 54.79 ± 6.69 M^{-1}s^{-1} for k$_{cat}$/K$_M$ was determined from the inhibition experiments at pH 7.0 with Ac-Phe-Gly-H and Ac-Phe-Gly-COOH (Table 1) as described in the legend to Figure 1. For fully active papain at pH 7.0 a k$_{cat}$/K$_M$ value of 231 M^{-1} s^{-1} is expected [41]. Therefore from these numbers we estimate that after activation the papain was 24 ± 3 percent fully active. Active papain concentrations were calculated (Figure 1) by assuming that the papain was 24 % fully active.

2.7. Inhibition of Papain by Ac-Phe-Gly-H, Ac-Phe-Gly-COOH, Z-Phe-Ala-H and Z-Phe-Ala-COOH

The inhibition of the papain catalysed hydrolysis of α-N-Benzoyl-L-Arginine-*p*-nitroanilide was studied at 25°C in 3 ml cuvettes containing 0.1M potassium phosphate buffer, 5 mM N-Acetyl-L-cysteine, 1 mM EDTA and 1.67% (v/v) dimethylsulphoxide. The initial rate of hydrolysis of α-N-Benzoyl-L-Arginine-*p*-nitroanilide was followed by measuring the release of p-nitroaniline (E$_{410}$ = 8800 M^{-1}cm^{-1} [42] over a 5-15 minute period. The pH in the reaction mixture was determined. Stock solutions of substrate and inhibitor were dissolved in dimethyl sulphoxide. These studies were carried out using a Cary-50 spectrophotometer.

The inhibition of papain by ZFA-COCHO [13] and Ac-FG-CHO [6] has been shown to be competitive and the K_i values were estimated when $[S_0] \ll K_M$. Therefore, the equation for competitive inhibition $d[P]/dt = k_{cat}[E][S]/([S]+K_M(1+[I]/K_i))$ reduces to $d[P]/dt = (k_{cat}/K_M)[E][S]K_i/([I]+K_i)$. K_i values were estimated by using a nonlinear least squares regression program [43].

2.8. Determination of hydration constants (K_{HYD}) and thiohemiacetal association constants (K_{THA})

Hydration constants (K_{HYD} =[Hydrate]/[Carbonyl]) and the thiohemiacetal association constants (K_{THA} =[THA]/([Aldehyde][Ac-Cys])) were determined by quantitive ^{13}C-NMR spectroscopy. All ^{13}C-NMR spectra were obtained without proton decoupling to ensure that no NOE was generated and an interpulse delay 5-10 times the longest T_1 value was used to ensure all signals were at their maximum intensities. T_1 values were determined using the inversion recovery pulse sequence [44]. The T_1 values of the ^{13}C-enriched carbon of Ac-Phe-Gly-^{13}CHO (Scheme 1C) and its hydrate (Ac-Phe-Gly-^{13}C(OH)$_2$) were 1.32 $\pm$ 0.03 s and 1.53 $\pm$ 0.04 s respectively. Therefore a pulse delay of 16 s was used and the hydration constant was determined from the ratio of the areas of the signals from the aldehyde carbonyl and its hydrate. The diastereoisomers formed at 76.8 ppm and 77.9 ppm when N-Ac-Cysteine reacted with Ac-Phe-Gly-^{13}CHO had T_1 values of 0.68 $\pm$ 0.02 s and 0.67 $\pm$ 0.03 s respectively.
The diastereoisomers formed at 75.5 ppm and 77.0 ppm when N-Ac-Cysteine reacted with Z-Phe-Ala-CO^{13}CHO (Scheme 1D) had T_1 values of 0.39 $\pm$ 0.06 s and 0.56 $\pm$ 0.06 s respectively. An interpulse delay of 8 s was used for measuring thiohemiacetal formation.

The concentrations of stock solutions of Ac-Phe-Gly-^{13}CHO (Scheme 1 C) in 1mM HCl were determined using E_{258} =200 M^{-1} cm^{-1} [9]. Z-Phe-Ala-CO^{13}CHO was in d$_6$-DMSO and its concentration in aqueous solutions was estimated using $E_{257.5}$ = 390 M^{-1} cm^{-1}. The amount of thiol present in stock solutions of N-Ac-Cysteine was determined by titrating with 2,2' dipyridyl disulphide at pH 8.1 using an E_{343} = 7920 M^{-1} cm^{-1} [45]. When determining K_{THA} using 1.5 mM Z-Phe-Ala-CO^{13}CHO (Scheme 1D) NMR spectra obtained over ~16 h were used and thiol concentrations were determined at the start and end of the NMR spectrum and the mean thiol concentration was used to calculate K_{THA}.

3. Results

3.1. The binding of the α-carboxylate group in peptide binding by papain and chymotrypsin.

K_i values for the inhibition of the papain catalysed hydrolysis of α-N-Benzoyl-L-Arginine-*p*-nitroanilide by Ac-Phe-Gly-H (Scheme 1A) and Ac-Phe-Gly-COOH (Scheme 1B) were determined when the substrate was not saturating (Figure 1). The K_i of Z-Phe-Ala-H (Scheme 1A) was determined but due to the low solubility of Z-Phe-Ala-COOH (Scheme 1B) its K_i value with papain was not determined (Table 1). With the serine protease chymotrypsin replacing the α-carboxylate group of Z-Ala-Ala-Phe-COOH and Z-Ala-Pro-Phe-COOH with a hydrogen atom to give Z-Ala-Ala-H and Z-Ala-Pro-Phe-H respectively led to 3-4 fold tighter binding (Table 1). This was expected because we would expect that for optimal catalytic efficiency enzymes would bind products less tightly than substrates [4]. However, replacing the α-carboxylate group of Ac-Phe-Gly-COOH (Scheme 1B) with a hydrogen atom to give Ac-Phe-Gly-H (Scheme 1A) resulted in a 3.4 fold decrease in the tightness of binding (Table 1) showing that the α-carboxylate group promotes binding to papain. This demonstrates that the expulsion of the product in the papain catalysed reaction is less efficient than in chymotrpsin catalysed reactions.

3.2. Binding of aldehyde and glyoxal inhibitors to papain and chymotrypsin

For specific substrate derived glyoxal inhibitors of the serine and cysteine proteases the keto-carbon of the glyoxal group is expected to be in an equivalent position to the carbonyl carbon of the peptide bond which is hydrolyzed during substrate catalysis[15]. Therefore it is not surprising that specific peptide derived glyoxals (ZAPF-COCHO & ZAAF-COCHO) are potent inhibitors of the serine proteases (Table 1) with the catalytic serine hydroxyl group reacting with the glyoxal keto-carbon to form a hemiketal (Scheme 5B) which mimics the tetrahedral intermediate formed (Scheme 5A) during the catalysis of peptide substrates[14-16, 46, 47]. The specific peptide derived glyoxal Z-Phe-Ala-COCHO (Scheme 1D) is also a potent inhibitor of the cysteine protease papain[13] and its glyoxal keto-carbon is also expected to to be in a position equivalent to the carbonyl carbon of a substrate peptide bond . However, using ^{13}C-NMR it has been shown that the catalytic thiol group of papain does not react with

the keto-carbon of the glyoxal, instead it reacts with the aldehyde carbon of the glyoxal group [13]. Therefore the thiohemiacetal formed with papain and the glyoxal inhibitor is not a good mimic the tetrahedral intermediate formed during the catalysis of peptide substrates by papain. Despite this the dipeptide glyoxal inhibitor Z-Phe-Ala-COCHO (Scheme 1D) is a more potent papain inhibitor than the tripeptide glyoxal inhibitors of chymotrypsin (Table 1). Using ^{13}C-NMR it has been shown that the thiolate ion of the catalytic cysteine residue of papain forms a thiohemiacetal with the aldehyde carbon of the aldehyde inhibitor Ac-Phe-Gly-CHO (Scheme 1C) [8, 9]. The aldehyde carbon of Ac-Phe-Gly-CHO is equivalent to the carbonyl carbon of the peptide bond hydrolyzed during catalysis and so the thiohemiacetal formed by papain with the aldehyde inhibitor is expected to be a good mimic of the tetrahedral intermediate formed during substrate catalysis.

As the peptide sequences of the papain inhibitors we have studied are different (Table 1) it is not possible to determine the relative potency of the aldehyde or glyoxal warheads on the basis of just their K_i values. However, replacing the aldehyde and glyoxal groups with a hydrogen atom allows us to determine the contribution of the peptide moiety to inhibitor potency[4] and shows that the Z-Phe-Ala-H peptide (Scheme 1A) is bound ~10 x more tightly (Table 1) than the Ac-Phe-Gly-H peptide (Scheme 1A). By dividing the K_i values for the aldehyde and glyoxal inhibitors by the K_i values for Ac-Phe-H (Scheme 1A) and Z-Phe-Ala-H (Scheme 1A) respectively we can quantify how much the aldehyde and glyoxal group contribute to inhibitor binding. Doing this shows that with papain, the aldehyde (Scheme 1C) and glyoxal groups (Scheme 1D) increase inhibitor binding ~1,000,000 and ~300,000 fold respectively (Table 1). In contrast with serine protease chymotrypsin, the aldehyde (ZAAF-CHO) and glyoxal groups (ZAAF-COCHO and Z-APF-COCHO) increase inhibitor binding ~100 and 460-2300 fold respectively (Table 1). Therefore, the aldehyde and glyoxal groups were 100-10,000 fold more effective at increasing inhibitor binding with papain compared to their effect with chymotrypsin (Table 1). In order to understand how the aldehyde and glyoxal groups are able to increase the binding of inhibitors to the serine and cysteine proteases we have examined the mechanism of inhibition.

3.3. Mechanism of the inhibition of Papain by specific peptide derived aldehyde and glyoxal inhibitors

Using ^{13}C-NMR it has been shown that while Ac-Phe-Gly-CHO forms two diasteriomeric thiohemiacetals with Ac-Cys, it only forms one diastereomer when it reacts with papain [8, 9]. The aldehyde group of Acetyl-Phe-Gly-CHO (A in Scheme 2) undergoes hydration (AH in Scheme 2) in water and Scheme 2 is the minimal scheme we have used to analyze thiohemiacetal (THA) formation with papain. The hydration constant $K_{H1(obs)}$ = [AH]/[A], K_s is the dissociation constant (K_s =[E][A]/[EA]) of the non-covalent enzyme inhibitor complex (EA) and K_{THA} is the equilibrium constant for thiohemiacetal formation (K_{THA} = [THA]/ [EA]). The observed binding constant $K_{i(obs)}$ = [E]([A] + [AH]/([EA] + [THA]) and so after substituting for Ks, $K_{H1(obs)}$ and K_{THA} we get eqn. 1 which can be re-arranged to give eqn. 2. If it is assumed that the $K_{i(obs)}$ value for Acetyl-Phe-Gly-H is a good approximation for K_s then the equilibrium constant (K_{THA}) for the formation of the papain thiohemiacetal with Acetyl-Phe-Gly-CHO could be calculated (Table 2) using eqn. 2 and the experimentally determined values of $K_{H1(obs)}$ and $K_{i(obs)}$.

$$K_{iobs} = K_s[1+K_{H1(obs)}]/[1+K_{THA}] \tag{1}$$

$$K_{THA} = \{K_s[1+K_{H1(obs)}]-K_{i(obs)}\}/K_{i(obs)}) \tag{2}$$

Using ^{13}C-NMR it has been shown that the catalytic thiol group of papain reacts with the aldehyde carbon of the glyoxal inhibitor Z-Phe-Ala-COCHO (Scheme 1D) to form a thiohemiacetal [13]. Therefore we could also calculate the equilibrium constant (K_{THA}) for the formation of the papain thiohemiacetal with Z-Phe-Ala-COCHO (Table 2). It is clear that thiohemicetal formation with papain is ~ 4000-20,000 times more effective than either hemiacetal or hemiketal formation with chymotrypsin (Table 2). Therefore both aldehyde and glyoxal warheads would appear to have a much higher specificity for the cysteine protease papain than the serine protease chymotrypsin suggesting that these warheads should be used to preferentially target the cysteine proteases. For the aldehyde inhibitors the aldehyde carbon is expected to be in the same position as the peptide carbonyl carbon of a substrate. However, in glyoxal inhibitors, the aldehyde carbon is one carbon bond away from this position. Despite this thiohemiacetal formation (K_{THA}) is 1.6 times more effective with the glyoxal group compared to the aldehyde group (Table 2). This suggests that it is the greater reactivity of the more electrophilic glyoxal aldehyde group compared to the aldehyde carbonyl group that makes glyoxals more effective at thiohemiacetal formation.

11

3.4. The effective molarity of the active site thiol group of papain

For the serine proteases the effective molarity of the catalytic serine hydroxyl group is calculated by dividing the equilibrium constant for forming a hemiacetal (K_{HA}) or hemiketal (K_{HK}) by the association constant (K_{H1}) of the hydrated inhibitor (K_{H1} = K_{H1obs} /[H_2O]) (Table 3). This gives the molarity of water that would have the same reactivity as the hydroxyl group of catalytic serine. If we do this calculation (K_{THA}/K_{H1}) for papain we find that the molarity of water required to have the same reactivity as the thiol group of papain towards Ac-Phe-Gly-CHO is 58 million. This is, at least, 2 orders of magnitude greater than is observed with chymotrypsin (Table 3).

However, to quantify how papain increases the reactivity of the active site thiol group we need an appropriate reference compound. We have used Ac-Cys as a reference compound and we have used quantitive ^{13}C-NMR spectroscopy to determine the equilibrium constant ($K_{THA(obs)}$) for thiohemiacetal formation from Ac-Cys and aldehyde inhibitors ($K_{THA(obs)}$ = [THA]/(([Aldehyde]+[Hydrate])([Acetyl-cys])) (Table 4). However, Ac-Cys only reacts with the non-hydrated form of the aldehyde inhibitor (Scheme 3) and so the true equilibrium constant for thiohemiacetal formation (K_{THA}) by Ac-Cys is given by K_{THA}= [THA]/(([Ac-Cys]([Aldehyde])) = $K_{THA(obs)}$ (1 + 55.55K_{H1}). Therefore the effective molarity is obtained by dividing K_{THA} for papain (K_{THA}= [[THA]/[EA]) by K_{THA} for Ac-Cys (K_{THA}= [THA]/(([Ac-Cys][Aldehyde]) to give the molarity of Ac-Cys required to have the same reactivity as the catalytic thiol group of papain (Table 4).

4. Discussion

4.1. Comparing the inhibitor potency of aldehyde and glyoxal inhibitors with the serine and cysteine proteases

Aldehyde and glyoxal warheads are potent inhibitors of the serine protease chymotrypsin, increasing binding 100-2000 fold (Table 1). However, they are ~1000 times more potent with the cysteine protease papain increasing binding by as much as ~1000,000 fold (Table 1). The fact that Z-Phe-

Ala-H was bound ~10 times more tightly than Ac-Phe-Gly-H (Table 1) shows that the benzyloxycarbonyl group and/or the sidechain of the alanine residue must make a significant contribution to inhibitor binding to papain.

With papain the aldehyde inhibitor Ac-Phe-Gly-CHO is bound ~4 times less tightly than the glyoxal inhibitor Z-Phe-Ala-COCHO (Table 1). However, after allowing for the fact that Z-Phe-Ala-H is bound ~10x more tightly than Ac-Phe-Gly-H we find that the aldehyde group of the aldehyde inhibitor is bound ~3 times more tightly (958,000/321,000) to papain than the glyoxal group of the glyoxal inhibitor (Table 1). The opposite conclusion would be reached on the basis of the $K_{i(obs)}$ values of the inhibitors demonstrating the importance of determining how the peptide portion of the inhibitor contributes to binding. In the serine proteases the opposite result is obtained with the glyoxal group being bound 4.5 to 23 times more tightly than the aldehyde group of an aldehyde inhibitor (Table 1). A similar 10-fold difference has also been observed [48] for glyoxal and aldehyde inhibitors binding to chymotrypsin. However, with the cysteine protease cathepsin B, the glyoxal inhibitor was ~3 times more potent than the equivalent aldehyde inhibitor when both had the same peptide component [48]. Therefore showing that with cathepsin B, unlike papain, inhibition by the glyoxal group is three times more effective than with an aldehyde group.

4.2. Thiohemiacetal stabilization by papain

Thiohemiacetal formation with papain is almost equally favored with both aldehyde and glyoxal inhibitors (see the K_{THA} values in Table 2). Likewise with chymotrypsin both hemiacetal (K_{HA}) and hemiketal (K_{HK}) formation are stabilized by similar amounts. However, with papain thiohemiacetal formation is ~10,000 times more effective than either hemiacetal or hemiketal formation with chymotrypsin (Table 2). Therefore, we need to determine whether papain is 10,000 times more effective than chymotrypsin at stabilizing tetrahedral adducts or are thiohemiacetals 10,000 times more easily formed than hemiacetals or hemiketals, or is the factor of 10,000 a combination of both these factors? To answer this question we examined the effective molarity of the catalytic thiol group and catalytic serine hydroxyl of papain and chymotrypsin respectively. For chymotrypsin the effective molarity was

determined by the ratio of the reactivity of the catalytic serine hydroxyl relative to that of the hydroxyl groups of water (Table 3) while for papain the effective molarity was determined by the reactivity of the catalytic thiol group of papain relative to the thiol group of Ac-Cys (Table 4). Increases in effective molarity are usually due to the entropic advantage of having both reactants in the same molecule. However in a protein complex other factors such as specific hydrogen bonds, general acid-base catalysis and ion pair interactions may also make significant contributions to the effective molarity. Therefore the effective molarity reflects the reactivity of the nucleophile (sulphur or oxygen) as well as the total stabilization achieved within the enzyme-inhibitor complex.

For papain with the glyoxal inhibitor Z-Phe-Ala-COCHO the effective molarity of 607 was ~40 times less than was observed with the aldehyde inhibitor Ac-Phe-Gly-CHO (Table 4). With the glyoxal inhibitor Z-Phe-Ala-COCHO the glyoxal keto-carbon is expected to be in a position analogous to that of the peptide carbonyl carbon of the corresponding substrate [13]. However, it has been shown using ^{13}C-NMR that the catalytic thiol group of papain does not react with the keto carbon equivalent to the peptide carbonyl carbon, instead it reacts with the aldehyde carbon of the glyoxal inhibitor Z-Phe-Ala-COCHO to form a thiohemiacetal [13]. Therefore this thiohemiacetal will not be in a position equivalent to that of the catalytic tetrahedral intermediate and so it is not expected to be optimally stabilized by papain. This explains the low effective molarity observed when papain is inhibited by the glyoxal inhibitor Z-Phe-Ala-COCHO (Table 4). However, the aldehyde carbonyl carbon of the aldehyde inhibitor Ac-Phe-Gly-CHO reacts stereospecifically with the catalytic thiol group of papain[8, 9] and so it might be expected to be in a position analogous to that of peptide carbonyl during catalysis. Therefore the thiohemiacetal formed could have a similar structure to the catalytic tetrahedral intermediate formed during catalysis and so it might be expected to be optimally stabilized by papain. This could explain why its effective molarity is ~ 40 times greater than that observed with the glyoxal inhibitor Z-Phe-Ala-COCHO (Table 4).

However, X-ray studies of papain show that the side chain NH_2 group of Gln-19 and the amide proton of Cys-25 are located in the papain oxyanion hole in positions that should allow them to hydrogen bond to the oxyanion of the catalytic tetrahedral intermediate [19, 20]. Removing one of these hydrogen bonds by changing Gln-19 to Ala-19 resulted in a 60 fold reduction in k_{cat}/K_M [33], suggesting that

hydrogen bonding by Gln-19 to the oxyanion of the catalytic tetrahedral intermediate enhances catalysis by papain. Asn-155 is in an analogous position in the serine protease subtilisin and mutating it to an alanine produces a larger 180-1200 fold decrease in k_{cat}/K_M [49, 50] suggesting that hydrogen bonding to the oxyanion is 20-30 times more effective in the serine proteases than in the thiol proteases. The Asn155Ala mutation in a subtilisin-chloromethylketone inhibitor adduct raised the oxyanion pKa by 1.09 pKa units [50] even though oxyanion hydrogen bonding is not optimal in chloromethylketone inhibitor adducts [51]. In contrast, the analogous Gln19Ala mutation in papain decreased the K_i value for Ac-Phe-Gly-CHO from 36 to 21 nM instead of raising it, showing that Gln-19 does not stabilize thiohemiacetal formation in the papain / Ac-Phe-Gly-CHO complex [33]. X-ray crystallographic studies with a range of aldehyde inhibitor complexes formed with cysteine proteases have shown that the thiohemiacetal oxygen can lie in the oxyanion hole (Scheme 4A) [27, 28, 30] or it can point away from the oxyanion hole hydrogen bonding with Nε2 of the catalytic histidine (Scheme 4B or C) [1, 29, 30]. Therefore the fact that mutating Gln-19 to Ala-19 does not increase K_i for the binding of Ac-Phe-Gly-CHO to papain shows that the thiohemiacetal oxygen is not located in the oxyanion hole (Scheme 4A) and that it is expected to be hydrogen bonded to Nε2 of the catalytic histidine (Scheme 4B or Scheme 4C).

X-ray crystallography of a glyoxal inhibitor bound to cathepsin L (a member of the papain superfamily) showed that as in papain [8, 9] the catalytic thiol group of cathepsin L formed a thiohemiacetal with the glyoxal aldehyde group [52]. The thiohemiacetal oxygen is in the oxyanion hole 2.83 Å from the side chain amide nitrogen of Gln-19 and 3.39 Å from the backbone amide nitrogen of Cys-25 (Scheme 4A). The oxyanion is ideally ~3 Å from the NH of Cys-25 and so the thiohemiacetal oxyanion is slightly displaced from its ideal location in the oxyanion hole. If the thiohemiacetal formed between papain and the Z-Phe-Ala-COCHO glyoxal inhibitor has a similar structure (Scheme 4A) then this suggests that the more effective thiohemiacetal stabilization with Ac-Phe-Gly-CHO aldehyde inhibitor (Table 4) must be due to its interaction with the catalytic histidine (Scheme 4B and 4C). It has been suggested[21] that it is a neutral thiohemiacetal that interacts with the histidine (Scheme 4C). However, this neutral hydrogen bond (Scheme 4C) would be expected to be weaker than the charged

oxyanion hydrogen bonds formed in the oxyanion hole (Scheme 4A). However, the interaction between the oxyanion and a positively charged histidine (Scheme 4B) should be much stronger, which could explain why there is more effective thiohemiacetal formation with Ac-Phe-Gly-CHO (Table 4). It would also explain why for Ac-Phe-Gly-CHO the thiohemiacetal oxygen is not in the oxyanion hole but instead interacts with histidine-159.

Thiohemiacetal stabilization by papain with the glyoxal inhibitor is ~40 times less effective than that observed with the aldehyde inhibitor (compare effective molarities in Table 4). However, the glyoxal warhead is ~3 times more effective as an inhibitor than the aldehyde warhead (Table 1). This reflects the greater reactivity of the glyoxal aldehyde group relative to the aldehyde group of the aldehyde inhibitor (compare K_{THA} values for Ac-Cys in Table 1).

4.3. Hemiacetal and hemiketal formation by chymotrypsin

In contrast to the results with papain [13] it has been shown that in the serine proteases the catalytic serine hydroxyl group reacts stereospecifically with the glyoxal keto carbon to form hemiketals which mimic the tetrahedral intermediate formed during catalysis [14-16, 46, 47]. However, with the aldehyde inhibitor Z-Ala-Ala-Phe-CHO, hemiacetal formation with the serine protease chymotrypsin has been shown to be non-stereospecific with two diastereomeric hemiacetals being formed [5]. But only the hemiacetal which has its hemiacetal oxygen in the oxyanion hole is expected to be a good analogue of the catalytic tetrahedral intermediate. However, as both hemiacetals are observed then they must be stabilized by similar amounts. The fact that for the serine proteases the effective molarity observed with aldehyde inhibitors is lower than that observed with glyoxal inhibitors (Table 3) shows that hemiketal stabilization is more effective than hemiacetal stabilization in the serine proteases (Table 3). It has been suggested that this is because the larger glyoxal group can fix the catalytic histidine in position (Scheme 5B) as with substrates (Scheme 5A), helping raise the histidine pKa >11 which lowers the oxyanion pKa and ensures that the oxyanion is in the oxyanion hole [5].

4.4. Comparison of tetrahedral adduct formation in papain and chymotrypsin

Thiohemiacetal formation for papain (K_{THA}, Table 2) is ~10,000 times more effective than hemiacetal formation for chymotrypsin (K_{HA}, Table 2) and also for Ac-Cys versus water (K_{THA}/K_{H1} for Ac-Cys in Table 4). However, we have found that thiohemiacetal stabilization (as measured by effective molarity) with aldehyde inhibitors and papain (Table 4) is comparable to the hemiketal stabilization (effective molarity) observed with chymotrypsin and glyoxal inhibitors (Table 3). Therefore we conclude that both papain and chymotrypsin can stabilize tetrahedral adducts by similar amounts. This suggests that earlier suggestions[31, 32] that oxyanion stabilization is much less effective in papain than in chymotrypsin are incorrect and that the later suggestions [33-35] that oxyanion stabilization is important in papain catalysis are correct. With specific peptide derived glyoxal inhibitors the glyoxal keto carbon is expected to be in the same position as the peptide carbonyl of a substrate. It is most probably the larger size of the sulfur nucleophile in papain that allows it to react with the more reactive glyoxal aldehyde group whereas the smaller oxygen nucleophile in chymotrypsin cannot do this.

4.5. What do hemiacetals and thiohemiacetals mimic best?

X-ray crystallographic studies of acyl intermediates in the serine proteases have shown that the carbonyl oxygen of the acyl intermediate can lie in the oxyanion hole (Scheme 6A) [53-55] or it can hydrogen bond directly to Nε2 of the catalytic histidine (Scheme 6B) [56] or it can hydrogen bond with a water molecule which hydrogen bonds with Nε2 of the catalytic histidine [17, 57] . It has also been observed in a position between these two states which has led to the suggestion that it was exchanging between its position in the oxyanion hole (Scheme 6A) and a position corresponding to that of the leaving peptide nitrogen of a peptide substrate (Scheme 5A, 5B) [57]. These are analogous positions to the two alternative hemiacetal conformations observed when aldehyde inhibitors form hemiacetals with chymotrypsin. In both cases the absence of a large leaving group allows the hemiacetal and the acyl intermediate to adopt more than one conformation. Therefore hemiacetals can adopt similar conformations to the catalytic acyl intermediate in the serine proteases. Likewise with thiohemiacetals formed with the cysteine proteases and aldehyde inhibitors the thiohemiacetal oxygen can lie in the oxyanion hole (Scheme 4A) [27, 28, 30] or it can point

away from the oxyanion hole hydrogen bonding with $N\varepsilon 2$ of the catalytic histidine (Scheme 4B or C) [1, 29, 30]. Suggesting these thiohemiacetals can adopt similar conformations (Scheme 4) to the catalytic thioacyl intermediate (Scheme 6C,D) formed during catalysis by the cysteine proteases.

Therefore it can be argued that the sp^3 hybridized hemiacetals and thiohemiacetals formed when aldehyde inhibitors react with the serine and cysteine proteases can mimic the conformations of the sp^2 hybridized catalytic acyl intermediate. This is because the lack of a large leaving group allows them to adopt similar conformations to acyl intermediates. Also the lack of a large leaving group with hemiacetals and thiohemiacetals means that the catalytic histidine is not fixed in position and its pKa is not raised to a value >10. Therefore the sp^3 hybridized thiohemiacetals and hemiacetals formed with aldehyde inhibitors are not good analogues of the catalytic tetrahedral intermediate but may well be better analogues of the sp^2 hybridized catalytic acyl or thioacyl intermediates which lack a group equivalent to a large substrate leaving group.

With glyoxal inhibitors the larger glyoxal group fixes the catalytic histidine in position and its pK_a is raised while the hemiketal oxyanion is stabilized in the oxyanion hole mimicking the catalytic process. It is therefore not surprising that the effective molarities observed with chymotrypsin and glyoxal inhibitors (Table 3) are greater than those observed with the aldehyde inhibitors and chymotrypsin (Table 3). This shows that with chymotrypsin the tetrahedral adducts formed with glyoxal inhibitors are much better analogues of the catalytic tetrahedral intermediate than the tetrahedral adducts formed with aldehyde inhibitors.

4.6. Conclusions

With papain a peptide aldehyde or glyoxal group increases inhibitor binding 300,000-1,000,000 fold. The effective molarity of the catalytic thiol group of papain is ~40 times greater with aldehyde inhibitors than with glyoxal inhibitors (Table 4). This shows that the inhibitor aldehyde carbonyl carbon must be in a position equivalent to a substrate peptide carbonyl for optimal thiohemiacetal formation. Thiohemiacetal formation when papain reacts with peptide aldehyde or glyoxal inhibitors is ~10,000 times more effective

than when chymotrypsin forms hemiacetals and hemiketals with aldehyde and glyoxal inhibitors respectively (Table 2). As the effective molarity of the catalytic thiol group of papain with the aldehyde inhibitor (Table 4) is similar to the effective molarity observed when the catalytic hydroxyl group of chymotrypsin reacts with glyoxal inhibitors (Table 3) then this shows that stabilization of the thiohemiacetal oxyanion (Scheme 4B) or its conjugate acid (Scheme 4C) by hydrogen bonding to Nϵ2 of the catalytic histidine is as effective as stabilization of hemiketal oxyanion in the oxyanion hole of chymotrypsin (Scheme 5B). Therefore we conclude that our results show that the cysteine protease papain can stabilize tetrahedral adducts by similar amounts to the serine protease chymotrypsin.

Acknowledgements

References

[1] Z. Muhaxhiri, L. Deng, S. Shanker, B. Sankaran, M.K. Estes, T. Palzkill, Y. Song, B.V. Prasad, Structural basis of substrate specificity and protease inhibition in Norwalk virus, Journal of virology, 87 (2013) 4281-4292.

[2] K. Akaji, H. Konno, H. Mitsui, K. Teruya, Y. Shimamoto, Y. Hattori, T. Ozaki, M. Kusunoki, A. Sanjoh, Structure-based design, synthesis, and evaluation of peptide-mimetic SARS 3CL protease inhibitors, J. Med. Chem., 54 (2011) 7962-7973.

[3] R. Prasad, Atul, V.K. Kolla, J. Legac, N. Singhal, R. Navale, P.J. Rosenthal, P.S. Sijwali, Blocking Plasmodium falciparum development via dual inhibition of hemoglobin degradation and the ubiquitin proteasome system by MG132, PLoS One, 8 (2013) e73530.

[4] T. Petrillo, C.A. O'Donohoe, N. Howe, J.P.G. Malthouse, Importance of tetrahedral Intermediate Formation in the catalytic Mechanism of the Serine Proteases Chymotrypsin and Subtilisin, Biochemistry, 51 (2012) 6164-6170.

[5] J.A. Cleary, W. Doherty, P. Evans, J.P. Malthouse, Hemiacetal stabilization in a chymotrypsin inhibitor complex and the reactivity of the hydroxyl group of the catalytic serine residue of chymotrypsin, Biochim. Biophys. Acta, 1844 (2014) 1119-1127.

[6] J.O. Westerik, R. Wolfenden, Aldehydes as inhibitors of papain, J. Biol. Chem., 247 (1972) 8195-8197.

[7] M.R. Bendall, I.L. Cartwright, P.I. Clark, G. Lowe, D. Nurse, Inhibition of papain by N-acyl-aminoacetaldehydes and N-Acyl-aminopropanones. Evidence for Hemithioacetal Formation by a Cross-saturation Technique in Nuclear-Magnetic-resonance Spectroscopy, Eur. J. Biochem., 79 (1977) 201-209.

[8] M.P. Gamcsik, J.P.G. Malthouse, W.U. Primrose, N.E. Mackenzie, A.S.F. Boyd, R.A. Russell, A.I. Scott, Structure and stereochemistry of tetrahedral inhibitor complexes of papain by direct NMR observation., J. Am. Chem. Soc., 105 (1983) 6324-6325.

[9] N.E. Mackenzie, S.K. Grant, A.I. Scott, J.P.G. Malthouse, ^{13}C NMR study of the stereospecificity of the thiohemiacetals formed on inhibition of papain by specific enantiomeric aldehydes, Biochemistry, 25 (1986) 2293-2298.

[10] B.A. Malcolm, C. Lowe, S. Shechosky, R.T. McKay, C.C. Yang, V.J. Shah, R.J. Simon, J.C. Vederas, D.V. Santi, Peptide aldehyde inhibitors of hepatitis A virus 3C proteinase, Biochemistry, 34 (1995) 8172-8179.

[11] C. Ortiz, C. Tellier, H. Williams, N.J. Stolowich, A.I. Scott, Diastereotopic Covalent Binding of the Natural Inhibitor Leupeptin to Trypsin: Detection of Two Interconverting Hemiacetals by Solution and Solid-State NMR Spectroscopy, Biochemistry, 30 (1991) 10026-10034.

[12] D.O. Shah, K. Lai, D.G. Gorenstein, ^{13}C NMR Spectroscopy of "Transition-State Analogue" Complexes of N-Acetyl-L-phenylalaninal and alpha-chymotrypsin, J. Am. Chem. Soc., 106 (1984) 4272-4273.

[13] J. Lowther, A. Djurdjevic-Pahl, C. Hewage, J.P.G. Malthouse, A ^{13}C-NMR study of the inhibition of papain by a dipeptide-glyoxal inhibitor, Biochem. J., 366 (2002) 983-987.

[14] A. Djurdjevic-Pahl, C. Hewage, J.P.G. Malthouse, A ^{13}C-NMR study of the inhibition of delta-chymotrypsin by a tripeptide-glyoxal inhibitor, Biochem. J., 362 (2002) 339-347.

[15] A. Djurdjevic-Pahl, C. Hewage, J.P.G. Malthouse, Ionisations within a subtilisin-glyoxal inhibitor complex, Biochim. Biophys. Acta, 1749 (2005) 33-41.

[16] E. Spink, S. Cosgrove, L. Rogers, C. Hewage, J.P.G. Malthouse, ^{13}C and ^{1}H NMR studies of ionizations and hydrogen bonding in chymotrypsin-glyoxal inhibitor complexes, J. Biol. Chem., 282 (2007) 7852-7861.

[17] R. Henderson, Structure of Crystalline α-Chymotrypsin, IV. The Structure of Indoleacryloyl-α-Chymotrypsin and its Relevance to the Hydrolytic Mechanism of the Enzyme, J. Mol. Biol., 54 (1970) 341-354.

[18] J.D. Robertus, J. Kraut, R.A. Alden, J.J. Birktoft, Subtilisin; a Stereochemical Mechanism Involving Transition-State Stabilization, Biochemistry, 11 (1972) 4293-4303.

[19] J. Drenth, K.H. Kalk, H.M. Swen, Binding of Chloromethyl Ketone Substrate Analogues to Crystalline Papain, Biochemistry, 15 (1976) 3731-3738.

[20] J. Drenth, H.M. Swen, W. Hoogenstraaten, L.A. Sluyterman, Mechanism of papain action, Proc. K. Ned. Akad. Wet., Ser. C 78, (1975) 104-110.

[21] A. Frankfater, T. Kuppy, Mechanism of association of N-acetyl-L-phenylalanylglycinal to papain, Biochemistry, 20 (1981) 5517-5524.

[22] L.T. Delbaere, G.D. Brayer, Structure of the complex formed between the bacterial-produced inhibitor chymostatin and the serine enzyme Streptomyces griseus protease A, J. Mol. Biol, 139 (1980) 45-51.

[23] M.N. James, A.R. Sielecki, G.D. Brayer, L.T. Delbaere, C.A. Bauer, Structures of product and inhibitor complexes of Streptomyces griseus protease A at 1.8 A resolution. A model for serine protease catalysis, J. Mol. Biol, 144 (1980) 43-88.

[24] L.T.J. Delbaere, G.D. Brayer, The 1.8 A Structure of the Complex between Chymostatin and Streptomyces griseus Protease A. A model for Serine Protease Catalytic Tetrahedral Intermediates, J.Mol.Biol., 183 (1985) 89-103.

[25] I.V. Kurinov, R.W. Harrison, Two crystal structures of the leupeptin-trypsin complex, Protein Sci, 5 (1996) 752-758.

[26] G. Robin, K. Chappell, M.J. Stoermer, S.H. Hu, P.R. Young, D.P. Fairlie, J.L. Martin, Structure of West Nile virus NS3 protease: ligand stabilization of the catalytic conformation, J. Mol. Biol, 385 (2009) 1568-1577.

[27] J.M. LaLonde, B. Zhao, W.W. Smith, C.A. Janson, R.L. DesJarlais, T.A. Tomaszek, T.J. Carr, S.K. Thompson, H.J. Oh, D.S. Yamashita, D.F. Veber, S.S. Abdel-Meguid, Use of papain as a model for the structure-based design of cathepsin K inhibitors: crystal structures of two papain-inhibitor complexes demonstrate binding to S'-subsites, J. Med. Chem., 41 (1998) 4567-4576.

[28] E. Schroder, C. Phillips, E. Garman, K. Harlos, C. Crawford, X-ray crystallographic structure of a papain-leupeptin complex, FEBS Lett, 315 (1993) 38-42.

[29] S.E. Webber, K. Okano, T.L. Little, S.H. Reich, Y. Xin, S.A. Fuhrman, D.A. Matthews, R.A. Love, T.F. Hendrickson, A.K. Patick, J.W. Meador, 3rd, R.A. Ferre, E.L. Brown, C.E. Ford, S.L. Binford, S.T. Worland, Tripeptide aldehyde inhibitors of human rhinovirus 3C protease: design, synthesis, biological evaluation, and cocrystal structure solution of P1 glutamine isosteric replacements, J. Med. Chem., 41 (1998) 2786-2805.

[30] L. Zhu, S. George, M.F. Schmidt, S.I. Al-Gharabli, J. Rademann, R. Hilgenfeld, Peptide aldehyde inhibitors challenge the substrate specificity of the SARS-coronavirus main protease, Antiviral research, 92 (2011) 204-212.

[31] B. Asboth, L. Polgar, Transition-State Stabilization at the Oxyanion Binding Sites of Serine and Thiol Proteinases: Hydrolyses of Thiono and Oxygen Esters, Biochemistry, 22 (1983) 117-122.

[32] B. Asboth, E. Stokum, I.U. Khan, L. Polgar, Mechanism of action of cysteine proteinases: oxyanion binding site is not essential in the hydrolysis of specific substrates, Biochemistry, 24 (1985) 606-609.

[33] R. Menard, J. Carriere, P. Laflamme, C. Plouffe, H.E. Khouri, T. Vernet, D.C. Tessier, D.Y. Thomas, A.C. Storer, Contribution of the glutamine 19 side chain to transition-state stabilization in the oxyanion hole of papain, Biochemistry, 30 (1991) 8924-8928.

[34] R. Menard, C. Plouffe, P. Laflamme, T. Vernet, D.C. Tessier, D.Y. Thomas, A.C. Storer, Modification of the electrostatic environment is tolerated in the oxyanion hole of the cysteine protease papain, Biochemistry, 34 (1995) 464-471.

[35] K.L. Foje, R.P. Hanzlik, Peptidyl thioamides as substrates and inhibitors of papain, and as probes of the kinetic significance of the oxyanion hole, Biochimica et biophysica acta, 1201 (1994) 447-453.

[36] L.A. Carpino, 1-Hydroxy-7-azabenzotriazole. An efficient peptide coupling additive, J. Am. Chem. Soc., 115 (1993) 4397-4398.

[37] J.C. Sheehan, S.L. Ledis, Total Synthesis of a Moncyclic Peptide Lactone Antibiotic, Etamycin, J. Am. Chem. Soc., 95 (1973) 875-879.

[38] S. Nahm, S.M. Weinreb, N-methoxy-n-methylamides as effective acylating agents, Tetrahedron Lett., 22 (1981) 3815-3818.

[39] G. Wang, U. Mahesh, G.Y.J. Chen, S.Q. Yao, Solid-Phase Synthesis of Peptide Vinyl Sulfones as Potential Inhibitors and Activity-Based Probes of Cysteine Proteases, Org. Lett., 5 (2003) 737-740.

[40] K. Brocklehurst, J. Carlsson, M.P.J. Kierstan, E.M. Crook, Covalent Chromatography, Preparation of fully active papain from dried papaya latex, Biochem. J., 133 (1973) 573-584.

[41] J.E. Mole, H.R. Horton, Kinetics of Papain-Catalysed Hydrolysis of α-N-Benzoyl-L-arginine-p-nitroanilide, Biochemistry, 12 (1973) 816-823.

[42] B.F. Erlanger, N. Kokowsky, W. Cohen, The Preparation and Properties of Two New Chromogenic Substrates of Trypsin, Arch. Biochem. Biophys., 95 (1961) 271-278.

[43] J.P.G. Malthouse, W.U. Primrose, N.E. Mackenzie, A.I. Scott, ^{13}C NMR Study of the Ionizations within a Trypsin-Chloromethyl Ketone Inhibitor Complex, Biochemistry, 24 (1985) 3478-3487.

[44] R.L. Vold, J.S. Waugh, M.P. Klein, D.E. Phelps, Measurement of Spin Relaxation in Complex Systems, J. Chem. Phys., 48 (1968) 3831-3832.

[45] B.S. Baines, K. Brocklehurst, A Spectrophotometric Method for the detection of Contaminant Chymopapains in Preparations of Papain, Biochem. J., 173 (1978) 345-347.

[46] N. Howe, L. Rogers, C. Hewage, J.P.G. Malthouse, Oxyanion and Tetrahedral Intermediate Stabilization by subtilisin: detection of a new tetrahedral adduct, Biochim. Biophys. Acta, Proteins & Proteomics 1794 (2009) 1251-1258.

[47] E. Spink, C. Hewage, J.P.G. Malthouse, Determination of the structure of tetrahedral transition state analogues bound at the active site of chymotrypsin using ^{18}O and ^{2}H isotope shifts in the ^{13}C NMR spectra of glyoxal inhibitors, Biochemistry, 46 (2007) 12868-12874.

[48] B. Walker, N. McCarthy, A. Healy, T. Ye, M.A. McKervey, Peptide glyoxals: a novel class of inhibitor for serine and cysteine proteinases, Biochem. J., 293 (1993) 321-323.

[49] S. Braxton, J.A. Wells, The Importance of a Distal Hydrogen Bonding Group in Stabilizing the Transition State in Subtilisin BPN, J. Biol. Chem., 266 (1991) 11797-11800.

[50] T.P. O'Connell, R.M. Day, E.V. Torchilin, W.W. Bachovchin, J.P.G. Malthouse, A ^{13}C-NMR study of the role of Asn-155 in stabilizing the oxyanion of a subtilisin tetrahedral adduct, Biochem. J., 326 (1997) 861-866.

[51] A. MacSweeney, G. Birrane, M.A. Walsh, T.P. O'Connell, J.P.G. Malthouse, T.M. Higgins, Crystal structure of δ-chymotrypsin bound to a peptidyl chloromethyl ketone inhibitor, Acta Cryst., D56 (2000) 280-286.

[52] R.T. Shenoy, J. Sivaraman, Structural basis for reversible and irreversible inhibition of human cathepsin L by their respective dipeptidyl glyoxal and diazomethylketone inhibitors, Journal of structural biology, 173 (2011) 14-19.

[53] H. Blanchard, M.N. James, A crystallographic re-investigation into the structure of Streptomyces griseus proteinase A reveals an acyl-enzyme intermediate, J. Mol. Biol., 241 (1994) 574-587.

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

[55] N. Singh, T. Jabeen, S. Sharma, I. Roy, M.N. Gupta, S. Bilgrami, R.K. Somvanshi, S. Dey, M. Perbandt, C. Betzel, A. Srinivasan, T.P. Singh, Detection of native peptides as potent inhibitors of enzymes. Crystal structure of the complex formed between treated bovine alpha-chymotrypsin and an autocatalytically produced fragment, IIe-Val-Asn-Gly-Glu-Glu-Ala-Val-Pro-Gly-Ser-Trp-Pro-Trp, at 2.2 angstroms resolution, The FEBS journal, 272 (2005) 562-572.

[56] B.L. Stoddard, J. Bruhnke, N. Porter, D. Ringe, G.A. Petsko, Structure and activity of two photoreversible cinnamates bound to chymotrypsin, Biochemistry, 29 (1990) 4871-4879.

[57] M.M. Dixon, B.W. Matthews, Is gamma-chymotrypsin a tetrapeptide acyl-enzyme adduct of alpha-chymotrypsin?, Biochemistry, 28 (1989) 7033-7038.

[58] R.P. Hanzlik, S.P. Jacober, J. Zygmunt, Reversible binding of peptide aldehydes to papain. Structure-activity relationships, Biochimica et biophysica acta, 1073 (1991) 33-42.

[59] J.A. Mattis, J.B. Henes, J.S. Fruton, Interaction of papain with derivatives of phenylalanylglycinal, J. Biol. Chem., 252 (1977) 6776-6782.

[60] F. Willenbrock, K. Brocklehurst, A general framework of cysteine-proteinase mechanism deduced from studies on enzymes with structurally different analogous catalytic-site residues Asp-158 and -161 (papain and actinidin), Gly-196 (cathepsin B) and Asn-165 (cathepsin H). Kinetic studies up to pH 8 of

the hydrolysis of N-alpha-benzyloxycarbonyl-L-arginyl-L-arginine 2-naphthylamide catalysed by cathepsin B and of L-arginine 2-naphthylamide catalysed by cathepsin H, Biochem. J., 227 (1985) 521-528.

Figure legends

Figure 1: Inhibition of the papain catalysed hydrolysis of α-N-Benzoyl-L-Arginine-*p*-nitroanilide by Ac-Phe-Gly-H and Ac-Phe-Gly-COOH. All samples contained 1.67% (v/v) dimethyl sulfoxide and 0.1M potassium phosphate buffer at pH 7.0. (A) For inhibition by Ac-Phe-Gly-COOH the concentrations of α-N-Benzoyl-L-Arginine-*p*-nitroanilide and active papain were 0.25mM and 0.6 μM respectively. The solid line was calculated using the equation $d[P]/dt = ((k_{cat}/K_M)[E][S] K_i)/([I] + K_i)$ and the fitted values of 260 ± 10 M^{-1} s^{-1} and 3.30 ± 0.42 mM for k_{cat}/K_M and K_i respectively. (B) Inhibition by Ac-Phe-Gly-H was measured in the same way with substrate and active enzyme concentrations of 0.23 mM and 0.48 μM respectively. The solid line was calculated as in (A) using the fitted values of 237 ± 5 M^{-1} s^{-1} and 12.3 ± 1.0 mM for k_{cat}/K_M and K_i respectively. All the errors quoted in this figure legend are standard errors obtained on fitting the experimental data. Experimental data were obtained at 25˚C.

Scheme legends

Scheme 1: Structures of inhibitors[a]

[a] R and X are either Z and CH_3- for ZFA (Z-Phe-Ala-) or Ac and H for AcFG (Ac-Phe-Gly-)

Scheme 2: Minimal Scheme for Thiohemiacetal Formation by the Cysteine Proteases

Scheme 3: Minimal Scheme for Thiohemiacetal Formation with N-Acetyl-L-Cysteine

Scheme 4: Thiohemiacetals formed when aldehyde inhibitors inhibit papain and enzymes of the papain family

Scheme 5: Comparison of the structures of the catalytic tetrahedral intermediate formed with peptide substrates and the hemiketal formed when glyoxal inhibitors react with chymotrypsin

Scheme 6: Structures of the acyl intermediates formed with the serine and cysteine protease

Table 1
Binding of inhibitors to chymotrypsin and papain

Enzyme	pH	$K_{i(obs)}$ (µM)[a]	$K_{i(obs)}$ (µM)[a]	$K_{i(obs)}$ (µM)[a]	K_{iIH}/K_{iICOOH}	K_{iIH}/K_{iICHO} or $K_{iIH}/K_{iICOCHO}$
		ZAPF-H	ZAPF-COOH	ZAPF-COCHO	$Ki_{ZAPF-H}/$ $Ki_{ZAPF-COOH}$	$Ki_{ZAPF-H}/$ $Ki_{ZAPF-COCHO}$
Chymotrypsin [b]	7.2	77.4 ± 6.4	241 ± 7	0.0335	0.32	2310
		ZAAF-H	ZAAF-COOH	ZAAF-COCHO	$Ki_{ZAAF-H}/$ $Ki_{ZAAF-COOH}$	$Ki_{ZAAF-H}/$ $Ki_{ZAAF-COCHO}$
Chymotrypsin [b]	7.2	166 ± 9	532 ± 50	0.365	0.31	455
Chymotrypsin [c]				ZAAF-CHO 1.7 ± 0.4		$Ki_{ZAAF-H}/$ $Ki_{ZAAF-CHO}$ 100
		ZFA-H	ZFA-COOH	ZFA-COCHO	$Ki_{ZFA-H}/$ $Ki_{ZFA-COOH}$	$Ki_{ZFA-H}/$ $Ki_{FA-COCHO}$
Papain	7.2	1,060 ± 490(6)[d]	ND	0.0033 ± 0.0003[e]	ND	321,000
Papain		Ac-FG-H	Ac-FG-COOH	Ac-FG-CHO	$Ki_{AcFG-H}/$ $Ki_{AcFG-COOH}$	Ki_{AcFG-H}/ Ki_{AcFG-C}
	6.5 & 7.0	11,500 ± 5,100(9)[d]	3,360 ± 420 (3)[d]	0.012 ± 0.007(6)[f]	3.42	958,000

[a] Errors are the standard deviations of 3 or more determinations, calculated using n-1 degrees of freedom. The number of determinations is in parentheses. Experimental data were obtained at 25°C.
[b] $K_{i(obs)}$ data source[4]
[c] $K_{i(obs)}$ data source [5]
[d] Present work, data obtained at pH 7.0
[e] $K_{i(obs)}$ data source[13]
[f] Mean of 6 values obtained at pHs 6.5 and 7.0[9, 21, 58, 59]. K_i values for Ac-FG-CHO with papain are essentially the same from pH 6.0 to 7.5 [21] and so the mean K_i value obtained should be a good estimate of the K_i at pH 7.2.

Table 2
A comparison of thiohemiacetal, hemiacetal and hemiketal formation with papain and chymotrypsin

Enzyme	Inhibitor	K_s, μM	$K_{i(obs)}$, μM	$K_{H1(obs)}$[b]	K_{HK}[b]	K_{HA}[b]	K_{THA}[b]
α-Chymotrypsin	ZAPF-COCHO	77.4	0.0335	1.28[c]	5270		
α-Chymotrypsin	ZAAF-COCHO	166	0.365	1.58[d]	1170		
α-Chymotrypsin	ZAAF-CHO	166	1.7	13.9[e]		1500[e]	
Papain	ZFA-COCHO	1060	0.0033	63.3 ± 5.4(4)[f]			20,700,000
Papain	Ac-FG-CHO	11,500	0.012	12.2 ± 1.0(6)[g]			12,700,000

[a] Errors are the standard deviations of 3 or more determinations, calculated using n-1 degrees of freedom. The number of determinations is in parentheses. Experimental data were obtained at 25˚C.
[b] $K_{H1(obs)}$, K_{HK}, K_{HA} and K_{THA} are the equilibrium constants for formation of keto carbonyl or aldehyde carbonyl hydrates, hemiketals, hemiacetals and thiohemiacetals respectively.
[c] Glyoxal keto group[47]
[d] Glyoxal keto group[4]
[e] Aldehyde group [5]
[f] Glyoxal aldehyde group[60]
[g] Aldehyde group (Present work)

Table 3.
Effective molarity of the catalytic hydroxyl group of chymotrypsin when it forms hemiketals and hemiacetals with glyoxal and aldehyde inhibitors respectively

Enzyme	Inhibitor	K_{H1}[a]	K_{HK}	K_{HA}	Effective Molarity (M)[b]
α-Chymotrypsin [c]	ZAPF-COCHO	0.023	5270		229,000
α-Chymotrypsin [c]	ZAAF-COCHO	0.0284	1170		41,200
α-Chymotrypsin [d]	ZAAF-CHO	0.25		1500	6,000

[a] K_{H1} = [AH]/([A][H$_2$O]) = K_{H1obs}/[H$_2$O]. Experimental data were obtained at 25˚C.
[b] Effective molarity, K_{HK} / K_{H1} for glyoxal inhibitors or K_{HA} / K_{H1} for aldehyde inhibitors
[c] Data source[4]
[d] Data source[5]

Table 4
Effective molarity of the catalytic thiol group of papain when it forms thiohemiacetals with glyoxal and aldehyde inhibitors

Reference thiol	Inhibitor	$K_{H1(obs)}$	K_{H1}	$K_{THA(obs)}$ Ac-Cys (M)[a]	K_{THA} Ac-Cys (M)[c]	K_{THA} Papain	Effective Molarity (M)[b]
Ac-Cys	ZFA-COCHO	63.3[d]	1.14	530	34,100	20,700,000	607
Ac-Cys	Ac-FG-CHO	12.2[e]	0.22	38.6 ±3.6 (11)	510	12,700,000	24,900

[a] Errors are the standard deviations of 3 or more determinations, calculated using n-1 degrees of freedom. The number of determinations is in parentheses. Experimental data were obtained at 25˚C.
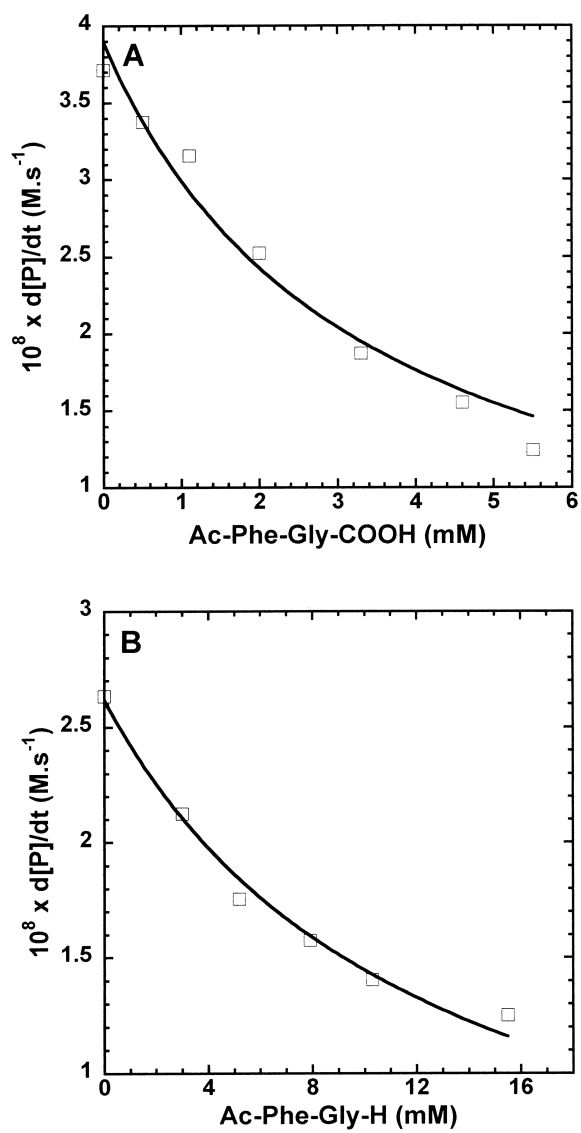[b] Effective molarity, $K_{THA(Papain)}$ / $K_{THA(Ac-Cys)}$
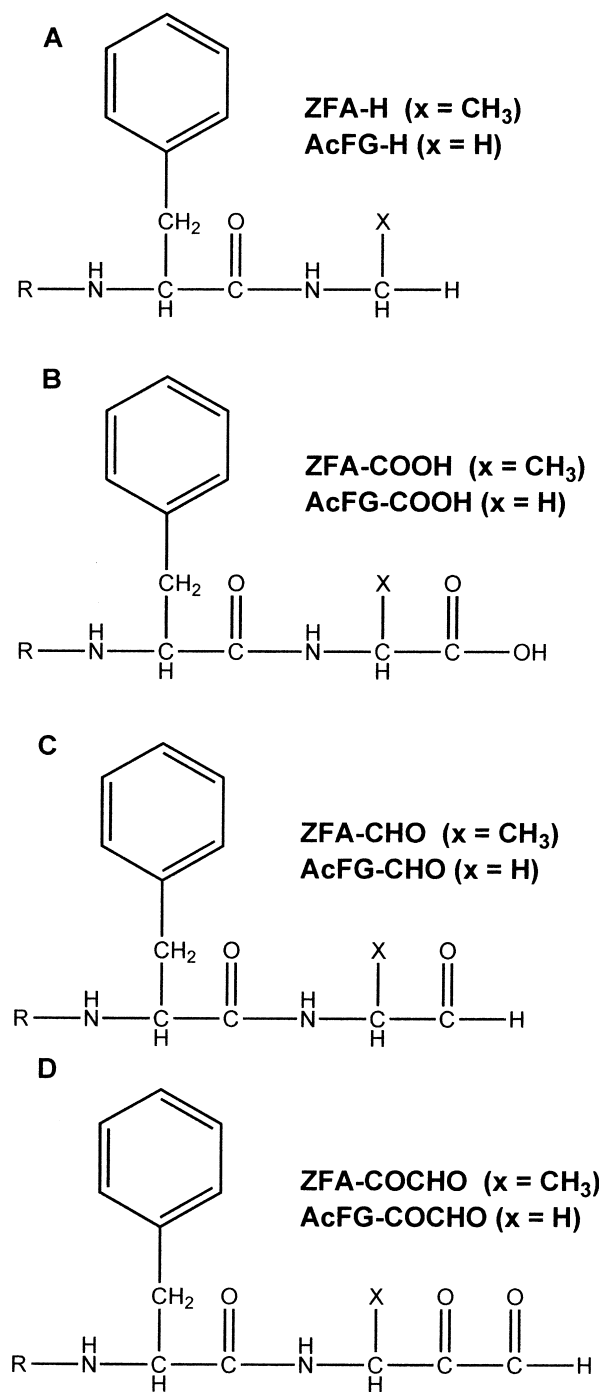[c] K_{THA} = $K_{THA(obs)}(1 + 55.55K_{H1})$
[d] Glyoxal aldehyde group and water (Present work)
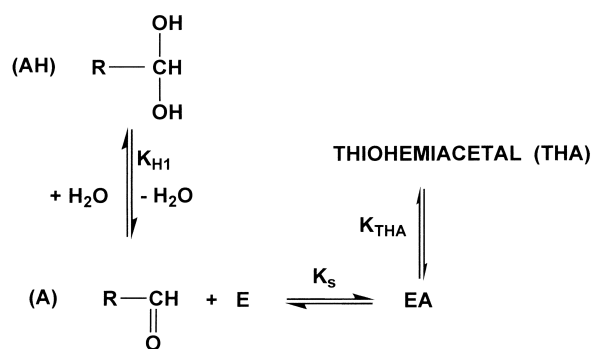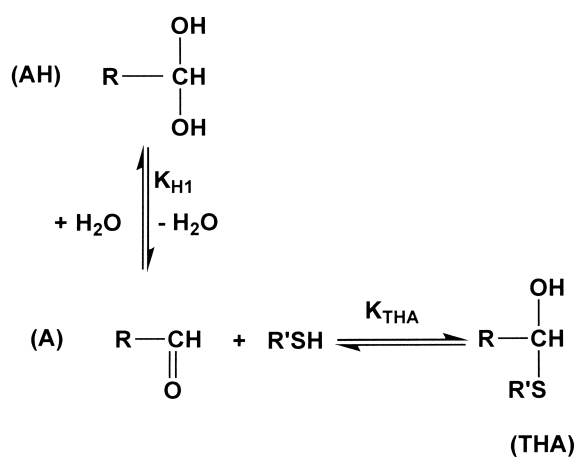[e] Aldehyde carbonyl group and water (Present work)

Fig. 1

Scheme 1

A

ZFA-H (x = CH₃)
AcFG-H (x = H)

B

ZFA-COOH (x = CH₃)
AcFG-COOH (x = H)

C

ZFA-CHO (x = CH₃)
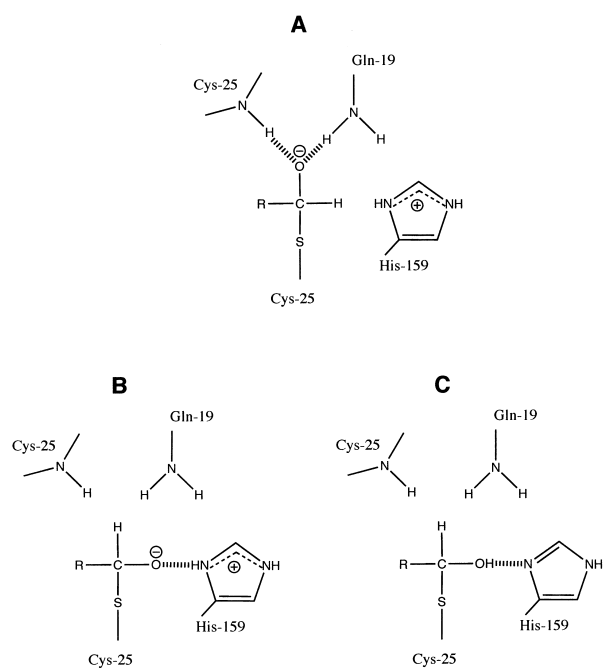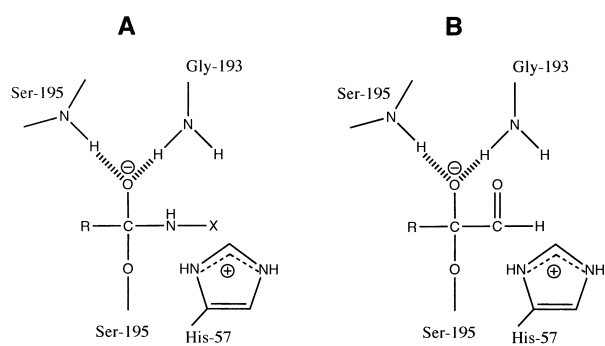AcFG-CHO (x = H)

D

ZFA-COCHO (x = CH₃)
AcFG-COCHO (x = H)

Scheme 2

Scheme 3

Scheme 4

Scheme 5

Scheme 6