^{13}C NMR Study of the Stereospecificity of the Thiohemiacetals Formed on Inhibition of Papain by Specific Enantiomeric Aldehydes†

Neil E. Mackenzie,* Stephan K. Grant, A. Ian Scott, and J. Paul G. Malthouse

Center for Biological NMR, Department of Chemistry, Texas A&M University, College Station, Texas 77843

Received October 17, 1985; Revised Manuscript Received December 6, 1985

ABSTRACT: The inhibition of papain by *N*-acetyl-D- and *N*-acetyl-L-phenylalanyl[1-^{13}C]glycinal was investigated by ^{13}C nuclear magnetic resonance (NMR) spectroscopy. Both the L- and D-aldehyde enantiomers formed thiohemiacetals with papain. The ^{13}C-enriched carbon of the thiohemiacetals formed with the L- and D-aldehydes has chemical shifts at 74.7 and 75.1 ppm, respectively. The difference in chemical shift for the two inhibitor complexes is attributed to each forming a different diastereomeric papain thiohemiacetal. Each enantiomeric inhibitor formed two diastereomeric thiohemiacetals with chiral thiols but produced a single diastereoisomer with papain. It is concluded that with papain thiohemiacetal formation is stereospecific. The D inhibitor is bound only 5-fold less tightly than the L inhibitor, which suggests that in both these inhibitor complexes the phenyl ring of the inhibitor phenylalanine is bound at the S_2 hydrophobic pocket of papain. This is supported by computer modeling studies that show that both the *N*-acetyl-D- and *N*-acetyl-L-phenylalanine moieties can be separately fitted into the S_2 subsite with the phenyl ring of phenylalanine in the S_2 hydrophobic pocket. It is concluded that thiohemiacetal formation at S_1 (S_1 and S_1' are the active center amino acid binding sites) is stereospecific with both D and L inhibitors. Computer modeling studies support this showing that, due to steric hindrance between the thiohemiacetal hydroxyl group and the backbone amide nitrogen of serine-24, only one of the two possible thiohemiacetal enantiomers can be formed at the S_1 subsite. The thiohemiacetals formed from both the D- and L-aldehyde inhibitors therefore have only one permitted conformation at S_1. In this conformation the thiohemiacetal hydroxyl *does not point toward the oxyanion hole* but instead points away from it into the bottom of the active site cleft of papain, where there are no significant hydrogen-bonded interactions with the thiohemiacetal hydroxyl. It is therefore inferred that the stabilization of the oxyanion of a tetrahedral intermediate by hydrogen-bonded interactions with groups in the oxyanion hole may not be possible in papain-catalyzed reactions.

The cysteine protease papain catalyzes the hydrolysis of amide and ester substrates by a process that can be minimally represented by three steps:

$$E + S \underset{}{\overset{K_s}{\rightleftharpoons}} ES \xrightarrow{k_2} ES' \xrightarrow{k_3} E + P_2$$
$$\qquad\qquad\qquad\quad + \qquad\quad P_1$$

The Michaelis complex ES gives rise to an acyl enzyme derivative (ES'), formed by reaction at the thiol group of cysteine-25, and the first product P_1 (either an amine or an alcohol). Deacylation with base-catalyzed addition of water gives the second product P_2, a carboxylic acid. Both acylation and deacylation are thought to proceed via tetrahedral intermediates (Lowe, 1970, 1976; Glazer & Smith, 1971), a mechanistic feature that is shared by the serine proteases (Kraut, 1977). These tetrahedral intermediates of catalysis, particularly in the case of the serine proteases, are thought to be stabilized by hydrogen bonding of the oxyanion to proximate amide functions (Polgar, 1982). Recently, Asboth et al. (1985) have cast doubt upon the requirement for stabilization of the tetrahedral intermediate by oxyanion binding, suggesting the latter nonessential in catalysis by cysteine proteases.

To assess the relevance and stability of tetrahedral intermediates in proteolytic hydrolyses, many studies have involved aldehyde derivatives of specific substrates (Thomson, 1973; Thomson & Bauer, 1979; Westerick & Wolfenden, 1972; Mattis et al., 1977; Breaux & Bender, 1975). These compounds are potent inhibitors of both serine and cysteine proteases and are thought to form covalent hemiacetals and thiohemiacetals by reaction with the active site serine hydroxyl and cysteine thiol group, respectively. These stable, tetravalent complexes can then be considered as analogues of tetrahedral transition states.

A convenient method for the detection of covalent enzyme substrate/inhibitor species is nuclear magnetic resonance (NMR)[1] spectroscopy (Jardetsky & Roberts, 1981; Steitz & Shulman, 1982). This technique employing ^{19}F (Gorenstein & Shah, 1982) and ^{1}H (Chen et al., 1979; Lowe & Nurse, 1977; Clark et al., 1977; Bender et al., 1977) as the reported nuclei has, by indirect means, characterized the covalently enzyme-bound species involved in the aldehydic inhibition of a variety of proteases. More recently, carbon-13 (^{13}C) has proven more useful as the observed nucleus. Thus, ^{13}C NMR spectroscopy has been used to directly observe and characterize several covalent adducts of the proteases [Mackenzie et al. (1984) and references cited therein; Shah et al., 1984; Rich, 1985; Malthouse et al., 1985].

In this work [preliminary data communicated by Gamcsick et al. (1983)], we have used *N*-acetyl-D- and *N*-acetyl-L-phenylalanyl[1-^{13}C]glycinal to determine whether these al-

† This work was supported by a grant from the National Institutes of Health (GM32596).

* Address correspondence to this author at the Control Division, Spectroscopy Development, Upjohn Co., Kalamazoo, MI 49001.

[1] Abbreviations: Z-Phe-Ala-CMK, N^{α}-benzyloxycarbonyl-L-phenylalanyl-L-alanine chloromethyl ketone; L-BAPA, N^{α}-benzyloxycarbonyl-L-arginine-*p*-nitroanilide; NMR, nuclear magnetic resonance.

0006-2960/86/0425-2293$01.50/0

dehyde enantiomers form thiohemiacetals with papain. It was also hoped that binding studies and computer modeling would allow us to define the stereochemistry of inhibition and assess whether the thiohemiacetal hydroxyl is stabilized within the oxyanion hole of papain.

Materials and Methods

Synthesis of Aldehyde Derivatives. (A) N-Acetyl-L-phenylalanylglycinal Dimethyl Acetal. To a stirred solution of *N*-acetyl-L-phenylalanine (0.621 g, 3 mmol) and triethylamine (0.42 mL, 3 mmol) in dry tetrahydrofuran (30 mL) at −5 °C, isobutyl chloroformate (0.4 mL, 3.1 mmol) was added. After 0.5 h, a solution of aminoacetaldehyde dimethyl acetal (0.33 mL, 3 mmol) in tetrahydrofuran (5 mL) was added with continued stirring. The reaction mixture was then allowed to warm to room temperature over 15 h after which 1 N HCl was added. The solution was then extracted with diethyl ether, and the organic phase was washed with water, saturated aqueous sodium bicarbonate, and water and then dried ($MgSO_4$). Evaporation of the ethereal solution under reduced pressure gave a colorless solid that was recrystallized from ethyl acetate: yield 0.794 g, 90%; mp 135–136.5 °C; $[\alpha]^{25}_{365}$ +25.7 ± 1° (methanol); ^{1}H NMR (500 MHz, $C_2{}^2H_6SO$) δ 8.072 (2 H, br m, >NH), 7.260–7.165 (5 H, m, Ar–H), 4.481 [1 H, ddd, ^3J = 4.8, 5.5, and 9.6 Hz, >CH–C(O)–], 4.282 [1 H, t, ^3J = 5.5 Hz, $-CH-(OCH_3)_2$], 3.238 (3 H, s, $-O-CH_3$), 3.229 (3 H, s, $-O-CH_3$), 3.140, 3–1.34 [2 × 1 H, t, ^3J = 5.5 Hz, $CH_2-CH(OCH_3)_2$], 2.921, 2.708 (2 × 1 H, dd, ^2J = 13.5 Hz, ^3J = 5.5 Hz, Ar–CH_2–), 1.736 [3 H, s, –C(O)CH_3].

(B) N-Acetyl-D-phenylalanylglycinal Dimethyl Acetal. This was prepared as for the L isomer with substitution of *N*-acetyl-D-phenylalanine: yield 0.780 g, 88%; mp 134.5–136 °C; $[\alpha]^{25}_{365}$ −23.9 ± 0.9° (methanol). ^{1}H NMR was as for the L isomer.

Hydrolysis of both L- and D-dimethyl acetals by 1 mM HCl at room temperature for 20 h gave *N*-acetyl-L- and *N*-acetyl-D-phenylalanylglycinal, which were used without further purification in the inhibition studies. For ^{13}C-enriched acetals, hydrolysis was carried out at pH 1 (0.1 M HCl), 25 °C, under a N_2 atmosphere, and on completion, aldehyde concentrations were determined by their phenylalanine UV absorption (ϵ_{258} = 200 M^{-1} cm^{-1}).

The synthesis of ^{13}C-enriched materials proceeded in an identical fashion with [1-^{13}C]aminoacetaldehyde dimethyl acetal (W. U. Primrose and A. I. Scott, unpublished work).

Papain. Fully active papain was isolated and characterized as described earlier (Mackenzie et al., 1985).

Determination of K_I Values. *N*-Acetyl-L-phenylalanylglycinal is a tight binding competitive inhibitor of papain (Westerick & Wolfenden, 1972; Mattis et al., 1977). Inhibitor dissociation constants (K_I) were determined by the method of Henderson (1973), eq 1, where v_0 and v_i are the initial rates

$$\frac{I_t}{1 - v_i/v_0} = E_t + K_I\frac{S_0 + K_m}{K_m}(v_0/v_i) \qquad (1)$$

for substrate hydrolysis in the absence and presence of the inhibitor and I_t and E_t are total inhibitor and total enzyme concentration, respectively.

N^α-Benzoyl-L-arginine-*p*-nitroanilide (L-BAPA) has a K_m value of 3 mM at pH 7 (Mole & Horton, 1973) and, therefore, provided $S_0 \ll K_m$; K_I values can be obtained by using eq 2.

$$\frac{I_t}{1 - v_0/v_i} = E_t + K_I(v_0/v_i) \qquad (2)$$

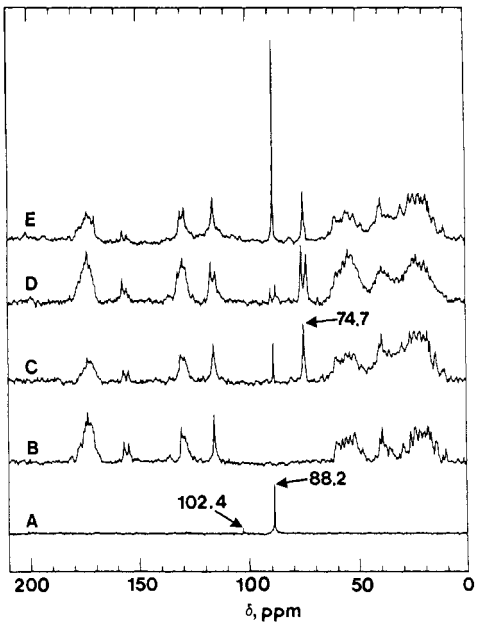


FIGURE 1: Addition of *N*-acetyl-L-phenylalanyl[1-^{13}C]glycinal to papain. Experimental conditions were as follows: pH 7.17; 10 mM sodium phosphate; 10% v/v D_2O; sample of volumes of 8.6–9.0 mL. In spectra A–E, the number of accumulations and concentrations of papain (fully active) and *N*-acetyl-L-phenylalanyl[1-^{13}C]glycinal were as follows: (A) 5000, 0.0 mM, 9 mM (2-mL sample volume); (B) 29 000, 0.8 mM, 0.0 mM; (C) 24 000, 0.73 mM, 0.87 mM; (D) 103 000, 0.73 mM, 0.87 mM; (E) 120 000, 0.38 mM, 1.08 mM. All spectra are proton-decoupled except for spectrum D, which is proton-coupled.

K_I values for *N*-acetyl-L- and *N*-acetyl-D-phenylalanylglycinal binding to papain were determined in this way at pH 7.0 in 20 mM sodium phosphate at 25 °C (see Table I).

NMR Spectra. Spectra (Figures 1–3) were recorded on a Bruker WM 300 wide-bore spectrometer (sample volume 8–10 mL) at 7.045 T. Chemical shifts are reported relative to tetramethylsilane at 0 ppm. Spectral conditions were 8000 time domain data points, 65-μs pulse width (112 μs = 90° pulse), ~0.25-s acquisition time and ~220 ppm spectral width, 20-Hz exponential weighting factor (line broadening), and broad band ^{1}H decoupling (1.0 W).

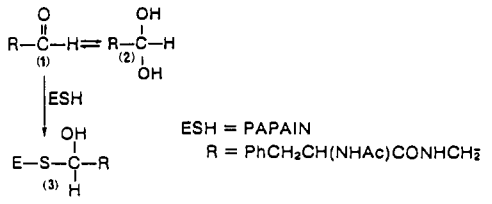
Spectra (Figure 4) were recorded as above except that a 2-Hz exponential weighting factor was used. Some model compound studies (Figure 5) were carried out at 11.74 T with a Bruker WM 500 spectrometer (sample volume ~2 mL). Spectral conditions were 32 000 time domain data points, 18-μs pulse width (21 μs = 90° pulse), 0.59-s acquisition time, 2-Hz exponential weighting factor (line broadening), and proton decoupling.

Computer Modeling. The Chem Graf modeling system was used: CHEM GRAF, created by E. K. Davies, Chemical Crystallography Laboratory, Oxford University; developed and distributed by Chemical Design Ltd., Oxford. The molecular coordinates of the papain/N^α-benzyloxycarbonyl-L-phenylalanyl-L-alanine chloromethyl ketone (Z-Phe-Ala-CMK) inhibitor complex (Drenth et al., 1976) were used as a starting point. The bound inhibitor was modified to give the bound N^α-acetyl-L-phenylalanylglycinal as described in the text.

Results and Discussion

N-Acetyl-L-phenylalanyl[1-^{13}C]glycinal (**1**, Scheme I) exists predominantly as its hydrate (**2**, Scheme I; δ = 88.2 ppm, Figure 1A) in aqueous solution; the carbonyl resonance of the aldehyde is barely discernible at 200.9 ppm (Figure 1A,E). The aldehyde was prepared (see Materials and Methods) by

Scheme I



hydrolysis of *N*-acetyl-L-phenylalanyl[1-^{13}C]glycinal dimethyl acetal (δ = 102.4 ppm, Figure 1A). Hydrolysis of the acetal was continued until no acetal resonance could be detected at 102.4 ppm. When papain was mixed with a slight excess of *N*-acetyl-L-phenylalanyl[1-^{13}C]glycinal, a new broad resonance ($\simeq$50 Hz) at 74.7 ppm was observed (Figure 1C). With a 3-fold excess of aldehyde, the intensity of this resonance (δ = 74.7 ppm) was unchanged though that of the hydrate was increased (Figure 1E), demonstrating that the new resonance is stoichiometric with respect to enzyme. The resonance at 74.7 ppm is shifted 14 ppm upfield relative to that of the hydrate (Figure 1C,E). This upfield shift relative to the hydrate is characteristic of thiohemiacetal formation due to the less efficient deshielding by sulfur (compared to oxygen) and is also observed with model compounds (see Model Studies). The proton-coupled spectrum (Figure 1D) confirms that both the hydrate and thiohemiacetal have a single directly bonded proton, both of which have ^{13}C–H coupling constants of 164.8 Hz. The resonance at 74.7 ppm can therefore be unambiguously assigned to the formation of a thiohemiacetal between the thiol group of cysteine-25 of papain and the aldehyde carbonyl of *N*-acetyl-L-phenylalanyl[1-^{13}C]glycinal (**3**, Scheme I).

The chemical shift of the papain thiohemiacetal resonance was pH-independent over the pH range examined (pH 3.0–8.6). Denaturation of papain at lower pHs and its decreased solubility at higher pHs precluded studies at higher or lower pHs. The pH independence of the papain thiohemiacetal resonance could result from an interaction between histidine-57 and the thiohemiacetal hydroxyl. Frankfater and Kuppy (1981) observed that the fluorescence intensity of tryptophan-177 was essentially pH independent from pH 3.5 to pH 8.0, suggesting that the pK_a of histidine-57 in the papain/*N*-acetyl-L-phenylalanylglycinal inhibitor complex was not in the pH range 3.5–9.0. Alternatively, the ^{13}C-enriched thiohemiacetal carbon may be insensitive to the ionization state of histidine-57 and of any other nearby groups, e.g., the carboxylate group of aspartate-158. Denaturation of papain and dissociation of the aldehyde at pHs <3.5 makes interpretation of decreases in fluorescence intensity at lower pHs ambiguous.

Using cross-saturation techniques, Bendall et al. (1977) obtained evidence for thiohemiacetal formation between papain and *N*-benzoylaminoacetaldehyde. Although the thiohemiacetal was not directly observed, they were able to estimate the chemical shift of the thiohemiacetal proton (δ = 6.51 ppm) and the coupling constant to the thiohemiacetal ^{13}C nucleus ($J_{^{13}C-H}$ = 182 Hz). In these studies, by direct observation of the residual splitting of the papain thiohemiacetal carbon as a function of the proton-decoupling frequency (Figure 2), the chemical shift of the papain thiohemiacetal proton resonance was estimated to be 6.56 ppm, by the method of Pachler (1972). An excess of aldehyde was present, which permitted the *N*-acetyl-L-phenylalanyl[1-^{13}C]glycinal hydrate proton chemical shift to be estimated (δ = 5.18 ppm). The chemical shift of the papain thiohemiacetal proton is in excellent agreement with the value obtained by Bendall et al. (1977).

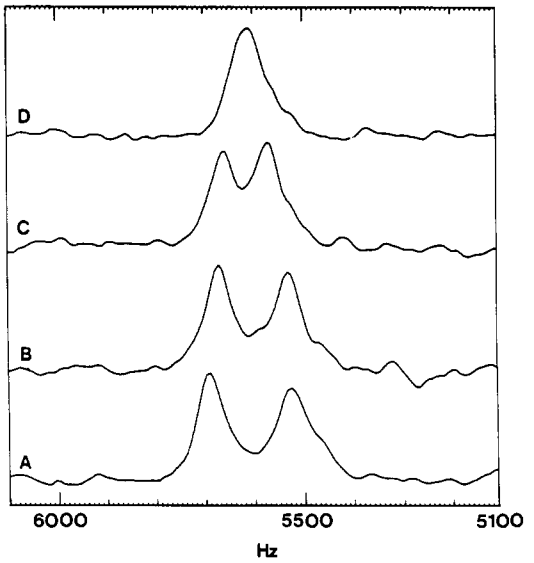


FIGURE 2: Residual coupling of the papain thiohemiacetal carbon as a function of the proton-decoupler frequency. Experiment conditions were as follows: pH 7.10; 10 mM sodium phosphate; 10% v/v D_2O; papain (fully active), 0.33 mM; *N*-acetyl-L-phenylalanyl[1-^{13}C]glycinal, 0.5 mM. Spectrum A is fully coupled (J_{C-H} = 164.8 Hz). The proton-decoupling frequencies for spectra B–D were 5000, 5600, and 5900 Hz, respectively. The residual splittings for spectra B and C were 141 and 89 Hz, respectively. A decoupler power of 0.4 W was used.

Table I: Binding Dissociation Constants for Papain/*N*-Acetyl-D- and *N*-Acetyl-L-phenylalanylglycinal Complex at 25 °C

isomer	K_I ($\times 10^8$ M)[a]	pH	ref
L	2.6	7.0	this work
L	1.0	6.5	Mattis et al., 1977
L	0.83	6.5	Frankfater & Kuppy, 1981
L	1.0	7.0	Gamcsik, 1983
L	4.6	5.5	Westerick & Wolfenden, 1972
D	13.0	7.0	this work

[a] Values are not corrected for hydration (Lewis & Wolfenden, 1977a,b).

It would therefore appear that the conformational, structural, and environmental effects on the chemical shift of the thiohemiacetal proton of the papain thiohemiacetals formed on reaction with *N*-benzoylaminoacetaldehyde and *N*-acetyl-L-phenylalanylglycinal are the same for these inhibitors. Direct measurement of $J_{^{13}C-H}$ for the directly bonded thiohemiacetal proton gave a value of 164.8 Hz. An identical value was also obtained for the aldehyde hydrate. These values are significantly less than the value of 182 Hz reported by Bendall et al. (1977).

It has been reported (Lowe, 1970; Lowe & Yuthavong, 1971a,b; Berger & Schechter, 1970) that papain has absolute stereospecificity for L-amino acids at the S_2 subsite. However, *N*-acetyl-D-phenylalanylglycinal (K_I = 0.13 μM) is only 5-fold less effective as an inhibitor of papain than its L isomer (Table I). In order to assess whether the decreased potency of the D isomer was due to the inability to form a papain thiohemiacetal, the D isomer was incubated with papain. A thiohemiacetal resonance at 75.1 ppm was observed (Figure 3A). A small difference in the chemical shift is expected for two diastereomeric thiohemiacetals (see Model Studies). The existence of two diastereomeric thiohemiacetals was confirmed by titration of the papain/*N*-acetyl-D-phenylalanylglycinal complex with *N*-acetyl-L-phenylalanyl[1-^{13}C]glycinal (Figure 3B–E).[2] Since the L-aldehyde binds 5 times more tightly than

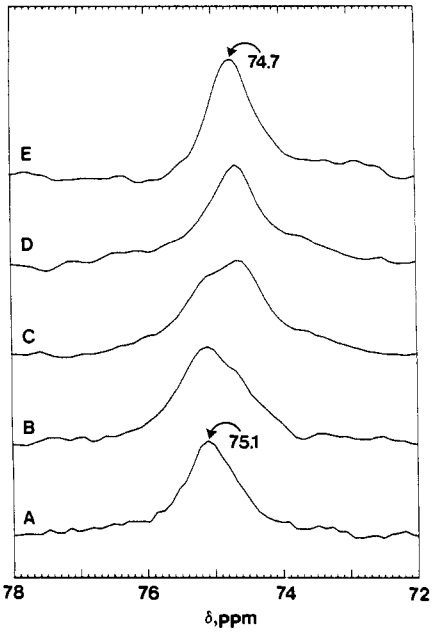
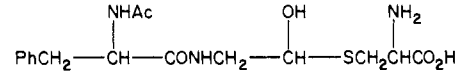


FIGURE 3: Titration of an equimolar mixture of papain and *N*-acetyl-D-phenylalanyl[1-^{13}C]glycinal with *N*-acetyl-L-phenylalanyl-[1-^{13}C]glycinal. Experimental conditions were as follows: pH 7.13; 10 mM sodium phosphate; 10% v/v D_2O; sample volumes of 8.2–9.4 mL. In Spectra A–E, the number of accumulations and concentrations of papain (fully active), *N*-acetyl-L-phenylalanyl[1-^{13}C]glycinal, and *N*-acetyl-D-phenylalanyl[1-^{13}C]glycinal were as follows: (A) 200000, 0.43 mM, 0.0 mM, 0.44 mM; (B) 90000, 0.36 mM, 0.13 mM, 0.43 mM; (C) 275000, 0.35 mM, 0.26 mM, 0.43 mM; (D) 23000, 0.33 mM, 0.63 mM, 0.40 mM; (E) 30000, 0.73 mM, 0.83 mM, 0.0 mM.

the D-aldehyde, then approximately equimolar concentrations of the D- and L-thiohemiacetals should be observed when the L-aldehyde concentration is approximately half the papain concentration. Under these conditions (Figure 3B,C), both resonances were detected while with an excess of the L-aldehyde gave only a single resonance at 74.7 ppm. Thus, the resonances at 74.7 and 75.1 ppm are due to the two diastereomeric papain thiohemiacetals formed from the L- and D-aldehyde, respectively.

Model Studies. Reaction of 2-mercaptoethanol with acetaldehyde gives rise to a thiohemiacetal of which the central sp^3 carbon has a single resonance of 73.27 ppm (Figure 4A). This value is in good agreement with that of approximately 75 ppm for the papain/inhibitor complexes. Introduction of a chiral center, as in the reaction between *N*-acetyl-L-cysteine and acetaldehyde, produces two diastereoisomeric thiohemiacetals. These arise by nucleophilic addition of the thiol group to the *si* or *re* face of the prochiral aldehydic carbonyl. The resonances at 74.19 and 73.02 ppm (Figure 4B) cannot be absolutely assigned to either diastereoisomer. A similar situation arises when the aldehydic moiety also contains a chiral center. Reaction of *N*-acetyl-L-phenylalanyl[1-^{13}C]glycinal with *N*-acetyl-L-cysteine produces again only two diastereoisomeric thiohemiacetals (not shown). The position (76.82 and 75.96 ppm) and separation (0.86 ppm = 65 Hz) of the chemical shifts in the ^{13}C NMR spectrum of these diastereoisomers is in accord with the resonances observed in a mixture of papain/inhibitor complexes (75.1 and 74.7 ppm, Figure 3).

Another mixture of diastereoisomers is displayed in the ^{13}C NMR spectrum (Figure 5) of the thiohemiacetals produced upon reaction *N*-acetyl-D- *and* *N*-acetyl-L-[1-^{13}C]phenylalanylglycinal with L-cysteine. Four diastereomers are formed (**4**–**7**) with chemical shifts of 65.2, 65.0, and 63.9 (and 63.9) ppm, which again cannot be absolutely assigned:



4,	D	D	L
5,	D	L	L
6,	L	D	L
7,	L	L	L

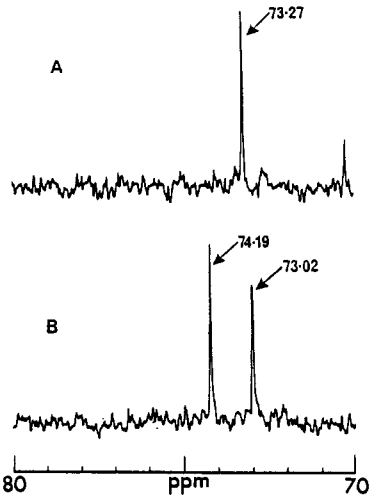


FIGURE 4: ^{13}C resonance(s) of thiohemiacetal(s) formed on reaction of acetaldehyde with 2-mercaptoethanol and *N*-acetyl-L-cysteine. Experimental conditions were as follows: (A) pH 3.0, 10% v/v D_2O, 200 mM acetaldehyde, 200 mM 2-mercaptoethanol, and 10000 accumulations with proton decoupling; (B) pH 3.0, 10% v/v D_2O, 200 mM acetaldehyde, 400 nM *N*-acetyl-L-cysteine, and 10000 accumulation with proton decoupling.

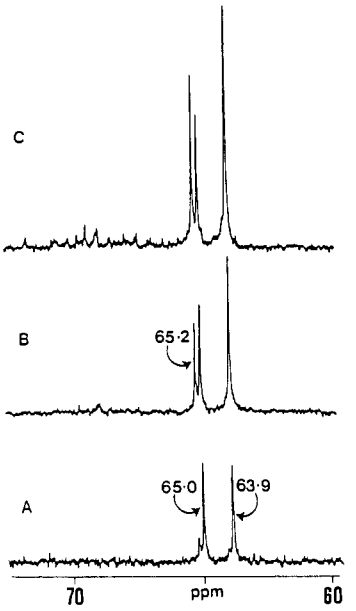


FIGURE 5: ^{13}C resonances of thiohemiacetals formed on reaction of L-cysteine and *N*-acetyl-D- and *N*-acetyl-L-phenylalanylglycinal. Experimental conditions were as follows: pH 3.7; 10 mM formic acid; 10% v/v D_2O. In spectra A–C, concentrations of L-cysteine, *N*-acetyl-D-phenylalanyl[1-^{13}C]glycinal, and *N*-acetyl-L-phenylalanyl-[1-^{13}C]glycinal were as follows: (A) 0.18 mM, 0.01 mM, 21.1 mM; (B) 0.17 mM, 0.15 mM, 20.2 mM; (C) 0.16 mM, 0.29 mM, 19.3 mM. For spectra A–C, 15000, 50000, and 15000 transients were accumulated, respectively.

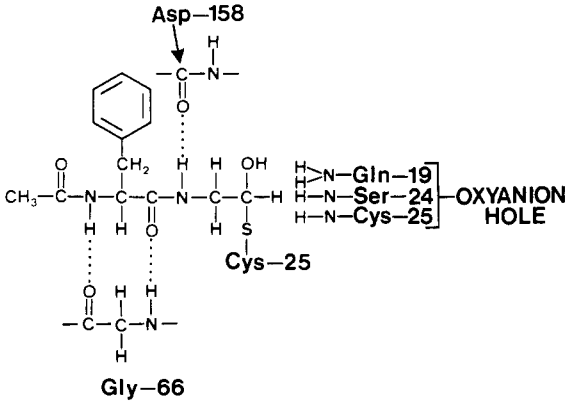
[2] Preliminary data, measured by progressive saturation, demonstrated that the enriched carbon of both D and L inhibitor/papain complexes has a T_1 of approximately 0.3 s. Hence, the relatively large pulse width (65 μs) and short acquisition time (0.25 s) were used throughout.

Although the possibility that this diastereoisomeric combination is present in the papain/inhibitor complex cannot be ruled out, it is unlikely, as both the D- and the L-aldehyde derivatives produced a thiohemiacetal resonance of identical line width (50 Hz).

The conclusion to be drawn from the model studies is that in the papain/inhibitor complexes the single resonances obtained from each thiohemiacetal implies that thiohemiacetal formation is stereospecific with the D- or L-aldehyde forming only one diastereomer with papain. Since the D-aldehyde is able to bind in S_2, we conclude that the stereochemistry of thiohemiacetal formation at S_1 must be the same for both the L- and L-aldehydes.

Computer Modeling. It is well established that the tight binding of phenylalanine-containing inhibitors and the high kinetic specificity of phenylalanine-containing substrates are due to the efficient binding of the phenylalanine phenyl ring at S_2. From the kinetic studies of Lowe and Yuthavong (1971a,b), it can be estimated that the binding of L-phenylalanine amino acid residues at the papain S_2 subsite results in *at least* a 1000-fold increase in binding. Therefore, if *N*-acetyl-D-phenylalanylglycinal cannot bind in S_2, a large increase in its dissociation constant (K_I) is expected (at least a 1000-fold). The comparatively small 5-fold increase in K_I observed for the D-aldehyde suggests that D-phenylalanine is bound at the S_2 subsite. Lowe (1976), however, stated that the S_2 subsite would be specific for L-amino acids since with D-amino acids the *N*-acylamino moiety of the D-amino acid bound at S_2 would point toward the protein and be sterically excluded. It was therefore decided to examine the binding of *N*-acetyl-D- and *N*-acetyl-L-phenylalanylglycinal to papain to determine if we could explain the apparently anomalous tight binding of *N*-acetyl-D-phenylalanylglycinal to papain.

Apart from binding of the phenyl ring of phenylalanine at the papain S_2 subsite, three enzyme–substrate hydrogen bonds are also thought to be involved in substrate or inhibitor binding (Lowe, 1976; Drenth et al., 1976; Lowe & Yuthavong, 1971a). For Z-Phe-Ala-CMK, these hydrogen bonds are formed between (i) the α-amide nitrogen of the inhibitor phenylalanine residue and the enzyme backbone glycine-66 carbonyl oxygen, (ii) the α-amide carbonyl oxygen of the inhibitor phenylalanine carbonyl oxygen and the enzyme backbone glycine-66 amide nitrogen, and (iii) the α-amide nitrogen of the inhibitor alanine residue and the enzyme backbone amide carbonyl of aspartate-158. These H bonds are represented schematically as



By use of the molecular coordinates for the binding of Z-Phe-Ala-CMK to papain (Drenth et al., 1976) (see Materials and Methods), these hydrogen bond distances were measured. Z-Phe-Ala-CMK was converted into *N*-acetylphenylalanylglycinal, and this then was manipulated in order to (a) maintain or improve enzyme–inhibitor peptide hydrogen-bond interactions and (b) achieve tetrahedral adduct formation (i.e., S···C=O, distance 0.18 nm and all bond angles 109 ± 1°) between the Cys-25 thiol group and the inhibitor aldehyde carbonyl. These manipulations caused minimal movement of both enzyme and inhibitor. It was found that due to steric hindrance the conformation of the thiohemiacetal thus obtained was such that the hydroxyl function does not point toward the oxyanion hole but instead points into a cavity at the bottom of the active site cleft. This cavity is between the indole ring of tryptophan-26 and the backbone polypeptide chain residues of glycine-65 and glycine-66. In this conformation, the thiohemiacetal hydroxyl oxygen is too far from N^{ϵ_1} of histidine-159 for hydrogen bonding (distance 0.568 nm). When the L-phenylalanine of the inhibitor was converted into its D isomer, the *N*-acetyl moiety could readily adopt a nonsterically hindered position by rotation about the phenylalanine C_α–N bond. The *N*-acetyl moiety now projected into the solvent, which may account for the small decrease in K_I for the *N*-acetyl-D-phenylalanylglycinal.

These computer modeling studies therefore give some credence to our observation that both the D- and L-aldehydes bind tightly to papain. Apparently, the S_2 subsite does not have absolute stereospecificity for L-amino acids, at least when the α-amino group of the P_2 is acetylated. The fact that both the *N*-acetyl-D- and *N*-acetyl-L-phenylalanylglycinals only form one diastereoisomeric thiohemiacetal with papain is thus accounted for if, as indicated, both form thiohemiacetals at S_1 with the same stereochemistry.

These preliminary modeling experiments may suggest that the same S_1 stereochemistry would be adopted during substrate hydrolysis if it proceeds via a tetrahedral intermediate. In this case, stabilization of the oxyanion by hydrogen bonding to the oxyanion hole would not be possible. This lends further support to the proposal (Asboth et al., 1985; Asboth & Polgar, 1983) that tetrahedral intermediate stabilization by papain is negligible and implies that catalysis does not proceed via such intermediates with papain.

ACKNOWLEDGMENTS

We thank Dr. Paul E. Fagerness for help with the computer modeling studies.

Registry No. $HO(CH_2)_2SCH(OH)CH_3$, 100928-66-3; AcNHCH(CO_2H)$CH_2SCH(OH)CH_3$ (isomer 1), 100928-67-4; AcNHCH(CO_2H)$CH_2SCH(OH)CH_3$ (isomer 2), 100928-68-5; $AcPheNHCH_2CH(OH)SCH_2CH(NH_2)CO_2H$ (isomer 1), 100928-69-6; $AcPheNHCH_2CH(OH)SCH_2CH(NH_2)CO_2H$ (isomer 2), 100992-49-2; $AcPheNHCH_2CH(OH)SCH_2CH(NH_2)CO_2H$ (isomer 3), 100992-50-5; $AcPheNHCH_2CH(OH)SCH_2CH(NH_2)CO_2H$ (isomer 4), 100992-51-6; *N*-acetyl-L-phenylalanylglycinal dimethyl acetal, 100928-64-1; *N*-acetyl-L-phenylalanine, 2018-61-3; aminoacetaldehyde dimethyl acetal, 22483-09-6; *N*-acetyl-D-phenylalanylglycinal dimethyl acetal, 100928-65-2; *N*-acetyl-D-phenylalanine, 10172-89-1; *N*-acetyl-L-phenylalanylglycinal, 41036-40-2; *N*-acetyl-D-phenylalanylglycinal, 100992-48-1.

REFERENCES

Asboth, B., & Polgar, L. (1983) *Biochemistry 22*, 117–122.

Asboth, B., Stokum, E., Khan, I. U., & Polgar, L. (1985) *Biochemistry 24*, 606–609.

Bendall, R. M., Cartwright, I. L., Clark, P. I., Lowe, G., & Nurse, D. (1977) *Eur. J. Biochem. 79*, 201–209.

Berger, A., & Schechter, I. (1970) *Philos. Trans. R. Soc. London, B 257*, 249–264.

Breaux, E. J., & Bender, M. L. (1975) *FEBS Lett. 56*, 81–84.

Chen, R., Gorenstein, D. G., Kennedy, W. P., Lowe, G., Nurse, D., & Schultz, R. M. (1979) *Biochemistry 18*, 921–926.

Clark, P. I., Lowe, G., & Nurse, D. (1977) *J. Chem. Soc., Chem. Commun.*, 451-453.
Drenth, J., Kalk, K. H., & Seven, H. M. (1976) *Biochemistry 15*, 3731-3738.
Frankfater, A., & Kuppy, T. (1981) *Biochemistry 20*, 5517-5524.
Gamcsik, M. P. (1983) Ph.D. Thesis, University of Edinburgh.
Gamcsik, M. P., Malthouse, J. P. G., Primrose, W. U., Mackenzie, N. E., Boyd, A. S. F., Russell, R. A., & Scott, A. I. (1983) *J. Am. Chem. Soc. 105*, 6324-6325.
Glazer, A. N., & Smith, E. L. (1971) *Enzymes (3rd Ed.) 3*, 501-546.
Gorenstein, D. G., & Shah, D. O. (1982) *Biochemistry 21*, 4679-4686.
Henderson, P. J. F. (1972) *Biochem. J. 127*, 323-333.
Jardetzky, O., & Roberts, G. C. K. (1981) in *NMR in Molecular Biology*, pp 417-447, Academic Press, New York.
Kraut, J. (1977) *Annu. Rev. Biochem. 46*, 331-358.
Lewis, C. A., Jr., & Wolfenden, R. (1977a) *Biochemistry 16*, 4886-4890.
Lewis, C. A., Jr., & Wolfenden, R. (1977b) *Biochemistry 16*, 4890-4895.
Lowe, G. (1970) *Philos. Trans. R. Soc. London, B 257*, 237-248.
Lowe, G. (1976) *Tetrahedron 32*, 291-302.
Lowe, G., & Yuthavong, Y. (1971a) *Biochem. J. 124*, 117-122.
Lowe, G., & Yuthavong, Y. (1971b) *Biochem. J. 124*, 107-115.
Lowe, G., & Nurse, D. (1977) *J. Chem. Soc., Chem. Commun.*, 815-817.
Mackenzie, N. E., Malthouse, J. P. G., & Scott, A. I. (1984) *Science (Washington, D.C.) 225*, 883-889.
Mackenzie, N. E., Malthouse, J. P. G., & Scott, A. I. (1985) *Biochem. J. 226*, 601-606.
Malthouse, J. P. G., Primrose, W. U., Mackenzie, N. E., & Scott, A. I. (1985) *Biochemistry 24*, 3478-3487.
Mattise, J. A., Henes, J. B., & Fruton, J. S. (1977) *J. Biol. Chem. 252*, 6776-6782.
Mole, J. E., & Horton, M. R. (1973) *Biochemistry 12*, 816-822.
Pachler, K. G. R. (1972) *J. Magn. Reson. 7*, 442-443.
Polgar, L. (1982) *Biochem. J. 207*, 1-10.
Rich, D. H. (1985) *J. Med. Chem. 28*, 263-273.
Shah, D. O., Lai, K., & Gorenstein, D. G. (1984) *J. Am. Chem. Soc. 106*, 4272-4273.
Steitz, T. A., & Shulman, R. G. (1982) *Annu. Rev. Biophys. Bioeng. 11*, 419-444.
Thomson, R. C. (1973) *Biochemistry 12*, 47-51.
Thomson, R. C., & Bauer, C.-A. (1979) *Biochemistry 18*, 1552-1558.
Westerick, J. O., & Wolfenden, R. (1972) *J. Biol. Chem. 247*, 8195-8197.

Evidence for Hemiketals as Intermediates in the Inactivation of Serine Proteinases with Halomethyl Ketones†

John S. McMurray and Douglas F. Dyckes*
Department of Chemistry, University of Houston, Houston, Texas 77004
Received September 10, 1985; Revised Manuscript Received December 13, 1985

ABSTRACT: The mechanism of inactivation of serine proteinases by peptide halomethyl ketone inhibitors was studied through the inhibition of trypsin with a series of model peptide ketones (Lys-Ala-LysCH_2X). In this series, X is a poor leaving group with increasing electron-withdrawing capacity (X = H, $CH_2CO_2CH_3$, $COCH_3$, $OCOCH_3$, and F), and as expected, the peptide ketones are reversible, competitive inhibitors of trypsin. The strength of binding of these inhibitors to trypsin increases with the electron-withdrawing ability of X, indicating that the inhibition constant K_i obtained is a measure of reversible hemiketal formation between the inhibitor ketone carbonyl group and the hydroxyl group of the active site serine. A Hammett plot of $-\log K_i$ vs. σ_I, the inductive substituent constant of X, reveals a linear relationship between the free energy of binding and the electron-withdrawing power of X. The reversible binding constant obtained for the corresponding chloromethyl ketone Lys-Ala-LysCH_2Cl falls on this line, indicating that the reversible binding involves hemiketal formation, which is followed by alkylation of the enzyme.

Halomethyl ketones have found wide use as both in vivo and in vitro probes of serine proteinases. These molecules inhibit the proteinases by N-alkylation of the imidazole ring of His-57 (chymotrypsin numbering system), which forms part of the enzyme catalytic triad (Ser-195, His-57, and Asp-102). Three possible mechanisms of inactivation of the enzyme have been reviewed by Powers (1977): (1) (a) formation of a reversible Michaelis-type enzyme–inhibitor complex, followed by (b) irreversible alkylation of His-57, followed by (c) hemiketal formation between the hydroxyl group of Ser-195 and the ketone carbonyl carbon; (2) (a) formation of a reversible Michaelis-type enzyme–inhibitor complex, followed by (b) hemiketal formation, followed by (c) irreversible alkylation of His-57; (3) (a) formation of a reversible Michaelis-type enzyme–inhibitor complex, followed by (b) hemiketal formation, followed by (c) displacement of the halogen by the original ketone carbonyl oxygen forming an epoxide, followed

† We are grateful to the Robert A. Welch Foundation for financial support of this work (Grant E-927).

0006-2960/86/0425-2298$01.50/0