and if so the effect of this on kinetic measurements. Mitochondrial forms of the enzyme have been shown to increase in activity upon dissociation (Dickinson & Allanson, 1985; Weiner & Takahashi, 1985).

Registry No. NADH, 58-68-4; propionaldehyde, 123-38-6; aldehyde dehydrogenase, 9028-86-8.

REFERENCES

Bennett, A. F., Buckley, P. D., & Blackwell, L. F. (1983) *Biochemistry 22*, 776-784.

Blackwell, L. F., Bennett, A. F., & Buckley, P. D. (1983) *Biochemistry 22*, 3784-3791.

Blackwell, L. F., Motion, R. L., MacGibbon, A. K. H., Hardman, M. J., & Buckley, P. D. (1987) *Biochem. J. 242*, 803-808.

Buckley, P. D., & Dunn, M. F. (1982) *Prog. Clin. Biol. Res. 114*, 23-35.

Dawson, R. M. C., Elliot, D. C., Elliot, W. H., & Jones, K. M. (1969) *Data for Biochemical Research*, pp 483-500, Oxford University Press, New York.

Dickinson, F. M. (1985) *Biochem. J. 225*, 159-165.

Dickinson, F. M. (1986) *Biochem. J. 238*, 75-82.

Dickinson, F. M., & Hart, G. J. (1982) *Biochem. J. 205*, 443-448.

Dickinson, F. M., & Allanson, S. (1985) *Prog. Clin. Biol. Res. 174*, 71-82.

Dickinson, F. M., & Haywood, G. W. (1986) *Biochem. J. 233*, 877-883.

Dickinson, F. M., Hart, G. J., & Kitson, T. M. (1981) *Biochem. J. 199*, 573-579.

Dunn, M. F., & Buckley, P. D. (1985) *Prog. Clin. Biol. Res. 174*, 15-27.

Hart, G. J., & Dickinson, F. M. (1982) *Biochem. J. 203*, 617-627.

Hempel, J., & Jörnvall, H. (1987) *Prog. Clin. Biol. Res. 232*, 1-14.

Klotz, I. M., Darnall, D. W., & Langerman, N. R. (1975) *Proteins (3rd Ed.) 1*, 293-411.

MacGibbon, A. K. H., Buckley, P. D., & Blackwell, L. F. (1977a) *Biochem. J. 165*, 455-462.

MacGibbon, A. K. H., Blackwell, L. F., & Buckley, P. D. (1977b) *Biochem. J. 167*, 469-477.

MacGibbon, A. K. H., Blackwell, L. F., & Buckley, P. D. (1977c) *Eur. J. Biochem. 77*, 93-100.

MacGibbon, A. K. H., Motion, R. L., Crow, K. E., Buckley, P. D., & Blackwell, L. F. (1979) *Eur. J. Biochem. 96*, 585-595.

Segel, I. H. (1975) *Enzyme Kinetics. Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*, pp 506-523, Wiley, New York.

Weiner, H., & Takahashi, K. (1985) *Prog. Clin. Biol. Res. 174*, 83-90.

Cooperativity of Papain-Substrate Interaction Energies in the S_2 to S_2' Subsites[†]

Paul J. Berti,[‡,§] Carlos H. Faerman,[‡] and Andrew C. Storer*[,‡]

Protein Engineering Section, Biotechnology Research Institute, 6100 avenue Royalmount, Montréal, Québec, H4P 2R2, Canada, and Department of Biochemistry, McGill University, 3655 Drummond Street, Montréal, Québec, H3G 1Y6, Canada

Received August 21, 1990; Revised Manuscript Received October 17, 1990

ABSTRACT: Enzyme-substrate contacts in the hydrolysis of ester substrates by the cysteine protease papain were investigated by systematically altering backbone hydrogen-bonding and side-chain hydrophobic contacts in the substrate and determining each substrate's kinetic constants. The observed specificity energies [defined as $\Delta\Delta G_{obs} = -RT \ln [(k_{cat}/K_M)_{first}/(k_{cat}/K_M)_{second})]]$ of the substrate backbone hydrogen bonds were −2.7 kcal/mol for the P_2 NH and −2.6 kcal/mol for the P_1 NH when compared against substrates containing esters at those sites. The observed binding energies were −4.0 kcal/mol for the P_2 Phe side chain, −1.0 kcal/mol for the P_1' C=O, and −2.3 kcal/mol for the P_2' NH. The latter three values probably all significantly underestimate the incremental binding energies. The P_2 NH, P_2 Phe side-chain, and P_1 NH contacts display a strong interdependence, or cooperativity, of interaction energies that is characteristic of enzyme-substrate interactions. This interdependence arises largely from the entropic cost of forming the enzyme-substrate transition state. As favorable contacts are added successively to a substrate, the entropic penalty associated with each decreases and the free energy expressed approaches the incremental interaction energy. This is the first report of a graded cooperative effect. Elucidation of favorable enzyme-substrate contacts remote from the catalytic site will assist in the design of highly specific cysteine protease inhibitors.

Papain (EC 3.4.22.3) is the most studied and consequently best understood member of the papain superfamily of cysteine proteases. The hydrolytic mechanism is minimally represented kinetically by a three-step process:

$$E + S \underset{k_{-1}}{\overset{k_{+1}}{\rightleftharpoons}} E{\cdot}S \xrightarrow{k_{+2}} EA + \text{Prod1} \xrightarrow{k_{+3}} E + \text{Prod2}$$

The free enzyme and substrate (E + S) associate to form a Michaelis complex (E·S) with a dissociation constant K_S (= k_{-1}/k_{+1}). The carbonyl carbon of the scissile ester or amide bond undergoes nucleophilic attack from the active-site cysteine (Cys 25), forming the covalent acyl-enzyme intermediate (EA) and releasing the alcohol or amine product (Prod1) with

---

[†] NRCC Publication No. 32396.

[‡] Biotechnology Research Institute.

[§] McGill University.

0006-2960/91/0430-1394$02.50/0

a rate constant k_{+2}. The acyl–enzyme then deacylates with a rate constant k_{+3} by undergoing nucleophilic attack from a water molecule to yield the free enzyme and the carboxylic acid product (Prod2). The acylation and deacylation steps are believed to proceed through tetrahedral intermediates [e.g., see Brocklehurst et al. (1987)].

Enzymic Catalysis. Enzymes catalyze reactions by preferentially stabilizing the transition state of the reaction over the reactants' ground state. Transition-state stabilization by enzymes can be achieved by two means, by using enzyme/substrate binding energy to effect ground-state destabilization (Jencks, 1975), thereby lowering the free energy of transition-state formation, and by compensation for entropy loss. Entropy is lost on forming an enzyme–substrate transition state due to the loss of translational and overall rotational motion of the two molecules relative to each other. Entropy loss on formation of the transition state is only weakly dependent on the identity of the substrate and has been variously estimated to be on the order of -35 to -50 cal/(K·mol) for reactions with "tight" transition states, including acyl transfer reactions (Page & Jencks, 1971; Finklestein & Janin, 1989), corresponding to 10–15 kcal/mol at 300 K. If an enzyme does nothing more in catalyzing a bimolecular reaction than to bind the two reactants in the active site, then the free energy of activation will be reduced by $\approx$10–15 kcal/mol.[1]

Incremental Binding/Specificity Energies. In examining a particular functionality of a substrate, one seeks to determine its incremental binding energy. As described by Fersht (1988), several criteria must be met in order for an observed binding energy of a substituent to be considered an incremental binding energy. It is necessary to delete the substituent under investigation, since substitution with another group that can interact with the enzyme will lead to a strict specificity effect; the deletion must be such that solvent has free access to the binding site on the enzyme, otherwise the free energy cost of excluding solvent will be included in the observed energies; and finally, there must be no difference in the entropy of the uncomplexed substrates and in the enzyme–substrate complexes in the presence and absence of the substituent. The latter condition becomes only approximately true for very good substrates; so one may only determine, at best, the lower limit of a group's incremental binding energy.

The observed binding energies, $\Delta\Delta G_{\text{obs}}$, are determined from

$$\Delta\Delta G_{\text{obs}} = -RT \ln \left[(k_{\text{cat}}/K_{\text{M}})_{\text{A}}/(k_{\text{cat}}/K_{\text{M}})_{\text{B}}\right]$$

where $k_{\text{cat}}/K_{\text{M}}$ is the specificity constant for substrates A and B. $k_{\text{cat}}/K_{\text{M}}$, which equals k_{+2}/K_{s} when $k_{-1} \gg k_{+2}$, is inversely proportional to the dissociation constant for the quasi-equilibrium between the *free* enzyme plus *free* substrate to the enzyme–substrate transition state.

There is a difficulty in determining the incremental binding energy of a functionality such as a substrate's P_1 NH in papain-catalyzed ester hydrolysis insofar as that the requirement of no entropy difference between different substrates requires the use of a substrate that extends into at least the S_2 subsite to achieve a sufficiently high $k_{\text{cat}}/K_{\text{M}}$. Obviously, deletion of the P_1 NH is then impossible. If that functionality is substituted with another, then potential interactions of the new substituent with the enzyme will interfere with determination of the incremental binding energy; that is, in this instance $\Delta\Delta G_{\text{obs}}$ is purely a specificity energy rather than a binding energy. When considering a substitution, one can only determine a quantity that will be referred to here as the *incremental specificity energy*, by analogy to the incremental binding energy. Use of the phrase incremental specificity energy makes the same assumptions as determination of incremental binding energy except that the effect of the interaction of the enzyme with the new substituent is taken into consideration. If the free energy of interaction of the new substituent can be determined independently, then the incremental binding energy can be estimated indirectly.

Specificity. Early studies on the specificity of papain included the work of Schechter and Berger [1967, 1968; see also Berger and Schechter (1970)] who demonstrated seven subsites in papain's active site, four on the acyl group side (S_4–S_1) and three on the leaving group side (S_1'–S_3').[2] They identified a specificity for Phe at P_2. The specificities of a number of other cysteine proteases have also been investigated [e.g., see Brocklehurst et al. (1987)]. The study of cleavage patterns of a number of polypeptides has failed to yield clear specificity requirements for any of the cysteine proteases examined. The inability to discern unambiguous specificites at any single subsite from peptide cleavage maps demonstrates that a number of enzyme–substrate contacts at different locations contribute to specificity and that no single contact is indispensable [cf. serine proteases that have very strong specificities at P_1, such as trypsin, which requires Lys or Arg, or *Staphylococcus aureus* protease V-8, which requires Asp or Glu at P_1 (Polgar, 1987)]. Because of the implication of cysteine proteases in a number of disease states, including myocardial infarction, cancer metastasis, muscular dystrophy [see Rich (1986)], and parasitic (Dresden et al., 1982; Coombs, 1982) and bacterial pathogenicity (Bjorck et al., 1989; Shah & Gharbia, 1989), recent specificity studies have focused on the inhibition of cysteine proteases, or more specifically, on selectivity inhibiting one cysteine protease in the presence of others [see Demuth (1990)]. Selective inhibition of particular cysteine proteases in vivo would also allow elucidation of the roles of those proteases. Studies on the specificity of inhibition have used side-chain variation of peptide-analogue inhibitors to effect specificity for a particular protease. The contribution of the peptide backbone to specificity has received less attention (Asboth et al., 1988).

In this study we have identified and characterized hydrogen-bonding interactions between peptide substrate backbone atoms in the S_2–S_2' subsites as well as side-chain interactions in the S_2 subsite. We discuss the importance of using these hydrogen bonds in maximizing the differentiation between the reactivities of different cysteine proteases. While the importance of compensation for entropy loss has been well recognized in theoretical discussions (Page & Jencks, 1971; Jencks, 1975, 1981), to our knowledge this is the first experimental report of a graded effect. The observed binding or specificity energies for a given substituent are dependent strongly on the presence of other enzyme–substrate interactions; that is, they show a strong cooperativity.

MATERIALS AND METHODS

Materials

Mercuripapain was prepared from papain (Sigma Chemical Co., Type III) by the method of Sluyterman and Wijdenes (1970) and activated as described previously (Storer et al., 1983). Freshly activated papain contained 1.0 mol of SH/mol

---

[1] We neglect here the entropy changes involved in releasing bound water from the enzyme and ligand and changes in the vibrational entropies on complex formation.

[2] Subsite numbering proceeds outward from the catalytic site. The corresponding amino acid residues in the peptide substrate are numbered: P_4-P_3-P_2-$P_1 \uparrow P_1'$-P_2'-P_3', where $\uparrow$ indicates the scissile amide or ester bond.

of protein and was stored on ice until used.

HPLC[3]-grade acetonitrile and high-purity H_2O (resistivity > 18 MΩ·cm) were used for HPLC. All other reagents and solvents were the highest available grade.

All compounds gave acceptable elemental analyses [±0.02(calculated percentage), except %N, which was (calculated percentage) ± 0.2%] and gave only one peak by HPLC. Analytical HPLC was performed on C-18, end-capped, reverse-phase columns, 4 mm × 15 cm with 5-μm irregular particles, with absorbance detection at 205 nm. Preparative HPLC used a C-18 reverse-phase column, 19 mm × 30 cm, with absorbance detection at 205 nm.

AcPheGlyOMe (**1**). AcPhe [synthesized as described previously (Carey et al., 1984)] and GlyOMe were coupled by the mixed anhydride method as described previously (Carey et al., 1984). Anal. Calcd for **1**: C, 60.42; H, 6.52; N, 10.07. Found: C, 60.59; H, 6.59; N, 9.89.

Glycolic Acid Methyl Esters and Glycolamides [MocPhe-O-CH$_2$-C(O)-OMe (**4**), *Ph-(CH$_2$)$_2$-C(O)-O-CH$_2$-C(O)-OMe* (**8**), *MocPhe-O-CH$_2$-C(O)-NH$_2$* (**13**), *and Ph-(CH$_2$)$_2$-C(O)-O-CH$_2$-C(O)-NH$_2$* (**17**)]. MocPhe (**6**; Carey et al., 1984) or Ph-$(CH_2)_2$-COOH (**10**) (1.1 equiv, approximately 10 mmol) was dissolved in ca. 25 mL of acetone with 1 equiv of methyl bromoacetate or iodoacetamide and 1 equiv of anhydrous K_2CO_3, and the mixture was refluxed for 5 h. The acetone was removed under vacuum. Remaining material was dissolved in ca. 30 mL of water and ca. 30 mL of CH_2Cl_2. The organic layer was washed with 1 N HCl, 1 N NaOH, and saturated NaCl solution, placed overnight over Na_2SO_4, and then filtered, and the solvent was removed under vacuum. The remaining material was dissolved in methanol, filtered, and dried under an N_2 stream, then extensively under vacuum. Anal. Calcd for **4**: C, 56.95; H, 5.80; N, 4.74. Found: C, 57.09; H, 5.79; N, 4.84. Anal. Calcd for **8**: C, 64.85; H, 6.35. Found: C, 64.57; H, 6.44. Anal. Calcd for **13**: C, 55.69; H, 5.75; N, 10.03. Found: C, 55.19; H, 5.81; N, 10.20. Anal. Calcd for **17**: C, 63.74; H, 6.32; N, 6.79. Found: C, 63.47; H, 6.34; N, 6.65.

MocPhe-O-CH$_2$-COOH (**5**) *and Ph-(CH$_2$)$_2$-C(O)-O-CH$_2$-COOH* (**9**). **13** and **17** were deamidated by the method of Nefkens and Nivard (1965). **5** was purified by flash chromatography on silica gel in ethyl acetate/hexane (7:3, v/v) and dried under an N_2 stream, then under vacuum. HPLC analysis indicated a ratio of [**5**]/[**6**] of 32:1. It was further purified by preparative HPLC in (0.1% TFA/H_2O)/acetonitrile (7:3, v/v) to give a ratio of 500:1. The pure resinous material was dried under vacuum and characterized by NMR: δ ([$^2H_4$]methanol) 2.91–2.96, 3.25–3.30 (2 H, 2 m, CH-CH_2-Ph), 3.53 (3 H, s, O-CH_3), 4.51–4.54 [1 H, m, NH-CH-C(O)], 4.62–4.72 (2 H, m, O-CH_2-COOH), 7.22 (5 H, m, C_6H_5). Anal. Calcd: C, 56.95; H, 5.80; N, 4.74. Found: C, 57.09; H, 5.79; N, 4.84. **9** was purified by HPLC in (0.1% TFA/H_2O)/acetonitrile (4:6 v/v), giving white crystals. Anal. Calcd: C, 63.45; H, 5.81. Found: C, 63.21; H, 5.86.

MocPhe-O-CH$_2$-C(O)-CH$_3$ (**14**) *and MocPhe-O-CH$_2$-C(O)-CH$_2$CH$_3$* (**15**). **6** (5 mmol) was dissolved with stirring in tetrahydrofuran with 6 mmol of 1-hydroxypropanone or 1-hydroxy-2-butanone, plus 5.5 mmol of 1-hydroxybenzotriazole and 3 drops of *N*-methylmorpholine at −10 °C; then 5.5 mmol of dicyclohexylcarbodiimide was added. Stirring was continued overnight, with warming to room temperature. The solution was rotary-evaporated to dryness, then ca. 30 mL of CH_2Cl_2 was added. The precipitate was filtered off after sitting overnight, and the material was purified by silica gel flash chromatography in CH_2Cl_2/methanol (95:5, v/v). **14** was further purified by HPLC in (0.1% TFA/H_2O)/acetonitrile (55:45, v/v). **15** was purified by HPLC in (0.1% TFA/H_2O)/acetonitrile/H_2O [times are cumulative: 0 min, 0:50:50 (v/v/v); 4 min, 50:50:0; 28 min, 50:50:0; 38 min, 20:80:0; flow = 5.0 mL/min]. Both were white solids (yield = 240 mg for **14**, 150 mg for **15**). Anal. Calcd for **14**: C, 60.19; H, 6.13; N, 5.04. Found: C, 59.89; H, 6.11; N, 5.20. Anal. Calcd for **15**: C, 61.41; H, 6.53; N, 4.79. Found: C, 61.21; H, 6.28; N, 4.96.

O-Acetyl-β-phenyllactylGlyOMe (**2**). L-β-phenyllactic acid was synthesized from L-phenylalanine by treatment with nitrous acid (Cohen & Weinstein, 1964). *O*-acetyl-L-β-phenyllactic acid was synthesized by refluxing with acetic anhydride (Ingles & Knowles, 1968). **2** was then made by the mixed anhydride method as used in Carey et al. (1984) and purified by silica gel chromatography in CH_2Cl_2/methanol (9:1, v/v). The material was rotary-evaporated to give a yellow oil. The compound was characterized by NMR: δ ([$^2H_6$]-acetone) 2.1 [3 H, s, CH_3-C(O)], 3.1–3.4 (2 H, m, CH-CH_2-Ph), 3.8 (3 H, s, C(O)-O-CH_3), 4.0 [2 H, d, J = 7 Hz, NH-CH_2-C(O)], 5.4 [1 H, m, O-CH-C(O)], 7.4 (5 H, s, C_6H_5), 7.7 (1 H, br s, NH). Stoichiometry of the hydrolysis by papain showed 97.8% purity.

MocPheGly-OMe (**3**; Angus et al., 1986) and *N*-benzoyl-Gly-OMe (**18**; Storer & Carey, 1985) have been described previously. MocGlyGly-OMe (**11**), MocPhe-OMe (**16**), and benzoyl-O-CH$_2$-C(O)-OMe (**19**) were the generous gifts of Dr. Ron Angus.

Methods

pH-stat kinetics were performed as previously described (Storer & Carey, 1985). Reaction solution was 300 mM NaCl, 1 mM EDTA, and 20% (v/v) acetonitrile, pH 6.0, with enzyme concentrations of 0.14–6 μM. Kinetics constants were determined directly from initial velocities ($v_0/[E]_0$) by the algorithm of Cornish-Bowden (1976), which assumes that random errors in observed $v_0/[E]_0$'s are proportional to the true $v_0/[E]_0$; this is true for pH-stat kinetic determinations (Storer & Carey, 1985).

MocPhe-O-CH$_2$-C(O)-OMe (**4**) *and Ph-(CH$_2$)$_2$-C(O)-O-CH$_2$-C(O)-OMe* (**8**). These two substrates each contain two ester functionalities. In each case, hydrolysis may proceed at either or both esters; for example, at MocPhe↑O-CH$_2$-C(O)-OMe and MocPhe-O-CH$_2$-C(O)↑OMe (where ↑ indicates the scissile bond). If it is assumed that productive binding for hydrolysis of one ester functionality precludes hydrolysis at the other without the Michaelis complex first dissociating and reforming (i.e., effectively binding a different substrate), then hydrolysis at each ester functionality proceeds as it would for two chemically distinct competitive substrates (Berti, 1990). In this case

$$(k_{\mathrm{cat}}/K_{\mathrm{M}})_{\mathrm{T}} = (k_{\mathrm{cat}}/K_{\mathrm{M}})_{\mathrm{B}} + (k_{\mathrm{cat}}/K_{\mathrm{M}})_{\mathrm{C}}$$

and

$$v_{0,\mathrm{B}}/v_{0,\mathrm{C}} = (k_{\mathrm{cat}}/K_{\mathrm{M}})_{\mathrm{B}}/(k_{\mathrm{cat}}/K_{\mathrm{M}})_{\mathrm{C}}$$

where T refers to the total reaction and B and C refer to hydrolysis at the individual ester functionalities.

To determine the kinetic constants for hydrolysis at each site, it was necessary to separate and quantitate the reaction products by HPLC. Analytical HPLC was performed as

---

[3] Abbreviations: All amino acids and derivatives are abbreviated as recommended by the IUPAC–IUB Joint Commission on Biochemical Nomenclature (1985); HPLC, high-pressure liquid chromatography; Moc-, (methoxycarbonyl)-; TFA, trifluoroacetic acid; ↑, in substrate names, the scissile ester bond.

described above, with flow = 2.5 mL/min for all runs. For **4** and its products, elution was with (0.1% TFA/H_2O)/acetonitrile, from 9:1 (v/v) to 3:7 over 6 min. For **8**, the same solvents were used, from 75:25 to 35:65 over 6 min. Standard curves were determined for both substrates and each of the potential products: **5**, **6**, **9**, and **10**.

In quantitating substrate and products in the reaction mixtures, it was necessary to quickly and gently separate them from papain, which irreversibly bound to and destroyed the HPLC column. This was accomplished by a dialysis arrangement using two nested test tubes separated by dialysis tubing with a cylindrical diameter of 1.6 cm (2.0 mL/cm) and a molecular weight cutoff of 3500. The outer test tube was a silanized 30-mL Corex centrifuge tube and contained 20 mL of reaction mixture. The inner test tube was a 13 mm × 100 mm disposable culture tube. Use of the inner test tube considerably increased the surface area to volume ratio and therefore the rate of dialysis. A small volume (200 μL) of reaction buffer was placed inside the dialysis membrane and allowed to equilibrate for 5–10 min before being removed and injected directly onto the HPLC column. The reaction buffer was 100 mM sodium phosphate, 200 mM NaCl, pH 6.0 and 20% (v/v) acetonitrile. This is the same ionic strength, percentage acetonitrile, and pH as the pH-stat reaction solution. Phosphate ions have been shown not to have an effect on papain activity (Whitaker & Bender, 1965; Storer & Carey, 1985).

Samples taken at various times from the dialysis setup in the reaction of papain with substrate were directly injected onto the column. For the reactions with **4**, the resulting peak areas for each component were converted to nanomoles injected, and the initial substrate concentration was then used to calculate the concentration of each component present in the reaction mixture from the relative amounts present. The ratio of products [**5**]/[**6**] was also determined. Initial velocities for substrate depletion and product formation were determined by the direct linear plot method (Cornish-Bowden, 1975). Michaelis–Menten kinetic constants determined from HPLC reactions were identical within experimental error with those determined by pH-stat (not shown).

The products of this reaction that still contained an ester functionality were examined as substrates. Glycolic acid methyl ester does not produce a measurable rate by pH-stat when 2 mM is reacted with 2 μM papain. **5** was tested as substrate by reacting 3.3 mM with 15 μM papain for 60 min in the dialysis setup and then submitting the products to HPLC analysis with the solvent program as for **4**.

MocPhe-O-CH_2-C(O)-NH_2 (**13**). Compound **13** was reacted with papain by the dialysis method to confirm hydrolysis at the ester. **13** from the dialysis setup before addition of papain gave a single peak. After a 15-min reaction of 1.9 mM **13** with 2 μM papain, a gradient of (0.1% TFA/H_2O)/acetonitrile [100:0 (v/v) to 35:65 over 6 min at 2.5 mL/min) gave peaks for **13** and **6**.

Ph-$(CH_2)_2$-C(O)-O-CH_2-C(O)-NH_2 (**17**). As no reaction rate could be detected by pH-stat, **17** was tested as a substrate by HPLC. **17** in the dialysis setup before addition of papain gave a single peak with (0.1% TFA/H_2O)/acetonitrile [8:2 (v/v) to 35:65 over 6.5 min at 2.5 mL/min]. After reacting 5 mM **17** with 2 μM papain for 80 min, only **17** was found by HPLC. Spiking an aliquot of the 80-min sample with Ph-$(CH_2)_2$-COOH gave two peaks.

RESULTS

MocPhe-O-CH_2-C(O)-OMe (**4**). Initial velocities ($v_0/[E]_0$'s) were calculated for **4** depletion and for **5** and **6** formation. In all cases they appear to follow Michaelis–Menten kinetics, as would be expected. The $v_0/[E]_0$'s calculated for **4** depletion were identical within experimental error with those determined by pH-stat. The ratio of products formed, [**5**]/[**6**], was 8.2 ± 0.4. Within experimental error, the ratio [**5**]/[**6**] was independent of elapsed time or $[S]_0$ (not shown).

On the basis of the HPLC runs of **4** reacted with papain, particularly in light of the constant [**5**]/[**6**] with time, **5** would be expected to be an extremely poor substrate, if at all. After incubating 3.3 mM **5** with 15 μM papain for 60 min, 0.186 mM **6** had formed, giving $v_0/[E]_0 = 3.43 \times 10^{-3}\ s^{-1}$. By using the Michaelis–Menten equation and assuming $k_{cat} \geq 0.1\ s^{-1}$, 5-fold less than the lowest k_{cat} measured in this study, $K_M \geq$ 92.9 mM is obtained, yielding $k_{cat}/K_M \leq 1.1\ M^{-1}\cdot s^{-1}$. Thus **5** is a poor substrate, with an upper limit on the k_{cat}/K_M of 1.1 $M^{-1}\cdot s^{-1}$.

Ph-$(CH_2)_2$-C(O)-O-CH_2-C(O)-OMe (**8**). **8** was examined by HPLC, like **4**, to determine the products of the reaction with papain and their ratios of formation. **9** was detected and quantified, but **10** was not found. After a 76-min reaction time with 2 mM substrate and 6.2 μM papain, no **10** peak was detected. At 76 min, the **9** peak was 35.32 nmol. Peaks *at least* as low as 0.11 nmol **10** would have been detected. This gives a ratio [**9**]/[**10**] ≥ 321. Given $k_{cat}/K_M = 79\ M^{-1}\cdot s^{-1}$ for **8a** (see Table I), $k_{cat}/K_M \leq 0.24\ M^{-1}\cdot s^{-1}$ for **8b**.

Ph-$(CH_2)_2$-C(O)-O-CH_2-C(O)-NH_2 (**17**). As no reaction rate could be detected by pH-stat, **17** was tested as a substrate by HPLC. By using the same arguments as with **8b** and assuming the same absorptivity as **10**, the **17** peak was 80.2 nmol. Therefore [**10**] ≤ 0.0131 mM. This corresponds to $v_0/[E]_0 \leq 1.36 \times 10^{-3}\ s^{-1}$. Assuming again that $k_{cat} \geq 0.1\ s^{-1}$, then K_M = 31.8 mM and $k_{cat}/K_M \leq 0.31\ M^{-1}\cdot s^{-1}$.

The results of the pH-stat and HPLC kinetics are summarized in Table I.

DISCUSSION

By systematically varying substituents of substrates to fulfill specificity requirements for papain, we have been able to map out interactions in the S_2–S_2' subsites. The free energies of the hydrogen bonding interactions at the P_2 NH, P_1 NH, P_1' C=O, and P_2' NH, as well as of hydrophobic interactions of the P_2 Phe side chain, were characterized. Putative hydrogen-bonding partners were identified by reference to the X-ray crystallographic structures of α-chloroketone and stefin B inhibited papains (Drenth et al., 1976; Stubbs et al., 1990) (Figure 1).

P_1' and P_2' Subsites. The kinetic constants of **13** [↑O-CH_2-C(O)-NH_2][4] and related substrates demonstrate that there are hydrogen-bonding partners for the P_1'–P_2' amide group of peptide substrates. This is the first report of these interactions.

$\Delta\Delta G_{obs}$ = −2.2 kcal/mol[5] between **13** [↑O-CH_2-C(O)-NH_2] and **4b** [↑O-CH_2-C(O)-OMe] could be attributed to hydrogen bonding or to unfavorable steric interactions between the ester's methyl group of **4b** [↑O-CH_2-C(O)-OMe] and the enzyme. Isosteric analogues of each compound that are not hydrogen bonding in this region were synthesized. The difference between **14** [↑O-CH_2-C(O)-CH_3] and **15** [↑O-CH_2-C(O)-

---

[4] In this section, to facilitate the discussion of substrates involving the S_1' and S_2' subsites, the structure of the leaving group is indicated in brackets after the compound number. Thus, MocPhe↑OCH_2-C(O)-OMe is **4b** [↑OCH_2-C(O)-OMe].

[5] In all the schemes under Discussion, $\Delta\Delta G_{obs} = -RT \ln [(k_{cat}/K_M)_{first}/(k_{cat}/K_M)_{second}]$. A positive $\Delta\Delta G_{obs}$ means that the second substrate listed has a higher k_{cat}/K_M; a negative $\Delta\Delta G_{obs}$ means that the first substrate is better. Random errors in $\Delta\Delta G_{obs} \approx \pm 0.1$ kcal/mol.

Table I: Kinetic Constants As Determined by pH-Stat and HPLC

substrate	k_{cat}/K_M ($M^{-1}\cdot s^{-1}$)	k_{cat} (s^{-1})	K_M (mM)	no. of points
1 AcPheGly↑OMe	120000 ± 5000	9.4 ± 0.2	0.078 ± 0.004	37
2 *O*-acetyl-β-phenyllactyl-Gly↑OMe	1100 ± 100	2.1 ± 0.2	1.9 ± 0.3	32
3 MocPheGly↑OMe	39000 ± 4000	6.3 ± 0.2	0.16 ± 0.02	34
4 MocPhe-O-CH_2-C(O)-OMe (overall)	480 ± 30	12 ± 5	26 ± 11	44
4a MocPhe-O-CH_2-C(O)↑OMe[a]	430 ± 35			
4b MocPhe↑O-CH_2-C(O)-OMe[a]	52 ± 4			
5 MocPhe↑O-CH_2-COOH	≤1.1[b]			
7 Ph-$(CH_2)_2$-C(O)-Gly↑OMe[c]	710 ± 30	4.7 ± 0.2	5.9	
8a Ph-$(CH_2)_2$-C(O)-O-CH_2-C(O)↑OMe	79 ± 4	0.67 ± 0.07	8.4 ± 1.2	26
8b Ph-$(CH_2)_2$-C(O)↑O-CH_2-C(O)-OMe	≤0.24[b]			
11 MocGlyGly↑OMe	40 ± 2	1.8 ± 0.2	45 ± 6	22
12 AcGly↑OEt[d]	2.4 ± 0.2	1.4 ± 0.1	590 ± 100	16
13 MocPhe↑O-CH_2-C(O)-NH_2	2200 ± 100	4.7 ± 0.2	2.2 ± 0.2	48
14 MocPhe↑O-CH_2-C(O)-CH_3	39 ± 3	1.6 ± 1.0	40 ± 28	29
15 MocPhe↑O-CH_2-C(O)-CH_2CH_3	96 ± 5	15 ± 28	160 ± 300	27
16 MocPhe↑OMe	7.12 ± 0.04	*e*	*e*	9
17 Ph-$(CH_2)_2$-C(O)↑O-CH_2-C(O)-NH_2	≤0.31[b]			
18 *N*-benzoyl-Gly↑OMe	140 ± 6	2.8 ± 0.2	20 ± 2	33
19 benzoyl-O-CH_2-C(O)↑OMe	17 ± 2	0.23 ± 0.04	14 ± 4	49

[a] Calculated from the k_{cat}/K_M of the overall rate of hydrolysis of MocPhe-O-CH_2-C(O)-OMe and the ratio of product formation: [**5**]/[**6**] = 8.2 ± 0.4. [b] See text for explanation. [c] From Storer et al. (1988). [d] Williams et al. (1972) found k_{cat}/K_M = 2.25 $M^{-1}\cdot s^{-1}$ for AcGly↑OMe as opposed to 2.4 $M^{-1}\cdot s^{-1}$ for AcGly↑OEt reported here, showing a negligible contribution from the ethyl versus the methyl leaving group and indicating that the comparison with Ph-$(CH_2)_2$-C(O)-Gly↑OMe is valid. [e] $K_M \gg [S]_{0,max}$.

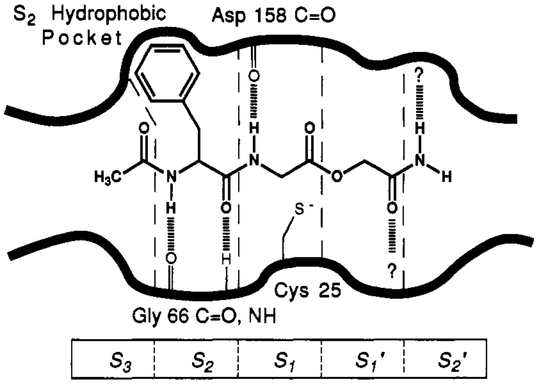

FIGURE 1: Schematic representation of binding of a hypothetical substrate, *N*-acetyl-PheGly↑O-CH_2-C(O)-NH_2, in papain's active site as deduced from the X-ray crystallographic structures of α-chloroketone-inhibited papains (Drenth et al., 1976) and the stefin B-papain complex (Stubbs et al., 1990). The active-site cysteine (Cys 25), the S_2 hydrophobic pocket (Tyr 67, Pro 68, Val 133, Val 157, and Ala 160), and those residues hydrogen-bonding to the substrate (Gly 66 and Asp 158) are shown. The hydrogen-bonding partners on the leaving-group side (subsites S_1' and S_2') are not known.

CH_2-CH_3] should reflect steric effects only. $\Delta\Delta G_{obs}$ = 0.5 kcal/mol indicates that there is not a problem sterically with the size of the **4b** [↑O-CH_2-C(O)-OMe] leaving group and therefore that the difference between **13** [↑O-CH_2-C(O)-NH_2] and **4b** [↑O-CH_2-C(O)-OMe] is a hydrogen-bonding effect.

The k_{cat}/K_M of **16** [↑OMe] versus **13** [↑O-CH_2-C(O)-NH_2] gives $\Delta\Delta G_{obs}$ = 3.3 kcal/mol. It is possible to separate the contributions from the C=O and $-NH_2$. The hydrogen-bonding contribution of the carbonyl portion may be estimated thus: $\Delta\Delta G_{obs}$ = −1.0 kcal/mol between **4b** [↑O-CH_2-C(O)-OMe] and **16** [↑OMe]. Lowe and Williams (1965) found that the k_{cat}/K_M's of *N*-benzoyl-Gly↑OMe (133 $M^{-1}\cdot s^{-1}$), ↑OEt (k_{cat}/K_M = 133 $M^{-1}\cdot s^{-1}$), and ↑O-*n*Pr (118 $M^{-1}\cdot s^{-1}$, $\Delta\Delta G_{obs}$ > −0.1 kcal/mol) were the same. Thus, the favorable $\Delta\Delta G_{obs}$ between **16** [↑OMe] and **14** [↑O-CH_2-C(O)-CH_3] may be attributed to the binding of the carbonyl oxygen, with no apparent contribution from the methyl group or carbonyl carbon. For this comparison of *N*-benzoyl-Gly↑O-*n*Pr with **14** [↑O-CH_2-C(O)-CH_3] to be valid, it is necessary that the leaving-group carbons interact with the enzyme in the same conformation in each substrate. This assumption is supported by noting that, on adding a methylene to either leaving group, the k_{cat}/K_M is improved by the same amount; i.e., **14** [↑O-CH_2-C(O)-CH_3] versus **15** [↑O-CH_2-C(O)-CH_2-CH_3], $\Delta\Delta G_{obs}$ = 0.5 kcal/mol and *N*-benzoyl-Gly↑O-*n*Pr versus *N*-benzoyl-Gly↑O-*n*Bu, $\Delta\Delta G_{obs}$ = 0.4 kcal/mol. The $\Delta\Delta G_{obs}$ = −1.0 kcal/mol for the P_1' carbonyl is therefore due primarily to the carbonyl oxygen forming a hydrogen bond with a donor on the enzyme.

Given the contribution of −1.0 kcal/mol by the carbonyl group, the remaining −2.3 kcal/mol of the $\Delta\Delta G_{obs}$ between **16** [↑OMe] and **13** [↑O-CH_2-C(O)-NH_2] is contributed by hydrogen bonding of the $-NH_2$. This is the same as the free energy observed when comparing **13** [↑O-CH_2-C(O)-NH_2] with **14** [↑O-CH_2-C(O)-CH_3]. Given the unusual situation here that the terminal methyl of **14** [↑O-CH_2-C(O)-CH_3] does not affect specificity, the incremental binding energy of the P_2' NH may be estimated at −2.3 kcal/mol.

4b [↑O-CH_2-C(O)-OMe] versus **15** [↑O-CH_2-C(O)-CH_2-CH_3] has an energetic penalty of $\Delta\Delta G_{obs}$ = 0.4 kcal/mol for placing an oxygen at the P_2' NH position compared with a $-CH_2-$. The difference then in free energy between oxygen [**4b** [↑O-CH_2-C(O)-OMe]] and NH [**13** [↑O-CH_2-C(O)-NH_2]] at that position would be $\Delta\Delta G_{obs}$ = 2.7 kcal/mol. Thus, at the P_2' NH position, the specificity is NH $\gg$ CH_2 $\approx$ O. This observation is in contrast to the conclusions of Merz and Kollman (1989). In free-energy perturbation simulations of transition-state analogue inhibitors of thermolysin, they concluded that the hydrogen-bonded NH of a particular inhibitor could be substituted with a CH_2 with only a small increase in K_i, while an oxygen in the same position would decrease binding by 3.6 kcal/mol; that is, specificity is NH $\approx$ CH_2 $\gg$ O. These conclusions were supported by the K_i's determined for the inhibitors under question (Merz & Kollman, 1989). Clearly, different conditions obtain here. The energetics of interactions at hydrogen-bonding sites on the enzyme are too complex to admit of simple qualitative explanations and must be probed empirically or computationally.

Further evidence for a hydrogen-bond acceptor for the P_2' NH is the very poor k_{cat}/K_M ≤ 1.1 $M^{-1}\cdot s^{-1}$ for **5** [↑O-CH_2-COOH] where $\Delta\Delta G_{obs}$ ≥ 4.2 kcal/mol versus **13** [↑O-CH_2-C(O)-NH_2]. Removal of the $-NH_2$ accounts for 2.3 kcal/mol of this, with the other ≥1.9 kcal/mol being accounted for by the unfavorable interaction of the carboxylate with the hydrogen-bond acceptor. Finally, a carboxylate shifted by one methylene further away from the catalytic site interacts favorably with the enzyme, indicating the position dependence of the binding energy penalty of the carboxylate [N-benzoyl-Gly↑S-CH$_2$-CH$_3$ versus N-benzoyl-Gly↑S-(CH$_2$)$_2$-COOH, $\Delta\Delta G_{\text{obs}} = 0.5$ kcal/mol (Lowe & Williams, 1965)].

The evidence for the presence of hydrogen bonding partners for the P_1' C═O and the P_2' NH is in contrast with the observations of Brubacher and Bender (1966), who found that the presence of this amide functionality did not increase the second-order rate constants for added nucleophiles in the deacylation of *trans*-cinnamoyl–papain. The work of Alecio et al. (1974), however, tends to support the role of the P_1'–P_2' amide in the hydrolysis of amides. Brubacher and Bender (1966) found that there was no decrease in the second-order rate constants between L-tryptophanamide and L-tryptamine or methyl L-tryptophanate. Additionally, the similarity of specificity of glycinamide and aminoacetonitrile implies that the –NH$_2$ of glycinamide does not form a hydrogen bond. There are two different principles in effect with the tryptophan-based and the glycine-based nucleophiles. The large difference between the second-order rate constants for L-tryptophanamide and glycinamide in spite of their very similar inherent nucleophilicites indicates that there is binding of the tryptophan to the enzyme in a quasi-equilibrium step before nucleophilic attack on the acyl–enzyme by the amine group. The lack of effect of the P_1'–P_2' amide indicates that the binding of the tryptophan side chain dominates the association of nucleophile with enzyme and that the amide (or methyl ester) is in a different orientation than the amide of **13** [↑O-CH$_2$-C(O)-NH$_2$]. The difference in position need not be large given the strong length dependence and directionality of hydrogen bonds [particularly hydrogen bond donors (Baker & Hubbard, 1984)] and given the fact that a weak hydrogen bond is worse than none at all (Wilkinson et al., 1983).

In the case of glycinamide and aminoacetonitrile, it is unlikely, due to entropic effects, that an adsorptive complex is formed between the enzyme and the nucleophiles in the S_1' subsite. The loss of entropy entailed in binding a small molecule to an enzyme can be estimated variously to be worth 10–15 kcal/mol. Clearly, the two hydrogen bonds through –C(O)NH$_2$, even in the presence of strong dispersion energies acting on the C_α of glycinamide (for which there is no evidence for or against) will not lead to appreciable association of the nucleophile to the enzyme in a quasi-equilibrium step before nucleophilic attack. The observed second-order rate constants represent, then, the attack of the unbound nucleophile onto the acyl–enzyme.

P_2 and P_1 Subsites. (*A*) *P_2 NH Hydrogen Bonding.* The improvement to $k_{\text{cat}}/K_{\text{M}}$ by −3.0 kcal/mol between **1** (NH, Phe, NH)[6] and **7** (H, Phe, NH) indicates that the substrate's P_3–P_2 amide hydrogen bonds with the enzyme. Exchanging an ester for the amide group decreases specificity by $\Delta\Delta G_{\text{obs}} = 2.7$ kcal/mol [**2** (O, Phe, NH) versus **1** (NH, Phe, NH)], suggesting that it is the amide NH that is primarily responsible for the observed binding energy.

If the P_3 C═O were to hydrogen-bond to the enzyme, then a more negative $\Delta\Delta G_{\text{obs}}$ between **2** (O, Phe, NH) and **7** (H, Phe, NH) would be expected than the weakly favorable −0.3 kcal/mol. At least part of the 0.3 kcal/mol is accounted for by differences in conformational entropy. The presence of the acetyl group of **2** hinders internal rotations of those bonds analogous to χ_1 and Ψ of a Phe residue, meaning that less entropy must be lost on forming the enzyme–substrate complex. Also, in the X-ray crystallographic structures of α-chloroketone-inhibited papains where the inhibitor extends as far as the P_5 subsite, the P_3 C═O is solvent-exposed and not in contact with the enzyme (Drenth et al., 1976).

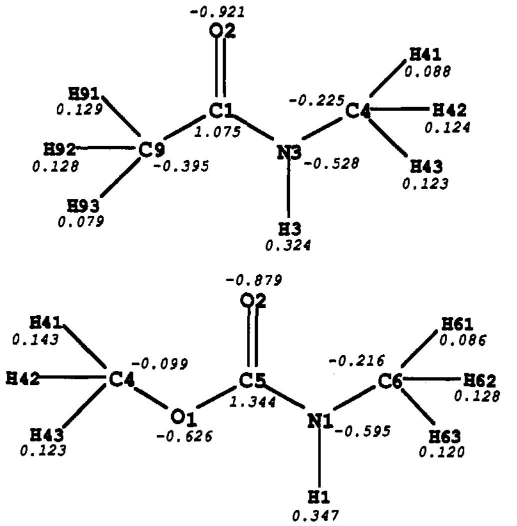

FIGURE 2: Atomic labeling for (top) N-methylacetamide and (bottom) methyl N-methylcarbamate, including atomic charges.

The use of a methyl carbamate group in **3** (NH, Phe, NH) in the place of an acetamide in **1** (NH, Phe, NH) leads to a loss of specificity by $\Delta\Delta G_{\text{obs}} = 0.7$ kcal/mol. Ab initio self-consistent field (SCF) calculations of the charge distributions of carbamate- and amide-containing model compounds were performed by using distributed multipoles analysis (Stone, 1981) (see Appendix). While the partial positive charges on the hydrogen-bonding protons of the carbamate and amide are nearly equal in vacuo (0.347e and 0.324e, respectively; Figure 2), hydrogen bonds from the carbamate would be less favorable due to the partial negative charge of −0.626e on the nearby sp^3 oxygen. Over 7630 grid points surrounding these NH groups, the average difference in electrostatic potential in vacuo is −9.4 kcal/mol, favoring the amide as a hydrogen-bond donor. The electrostatic potential in vacuo is not related to hydrogen-bonding free energy in solution in any simple manner, but it does clearly indicate that amides are better hydrogen-bond donors than carbamates.

The $\Delta\Delta G_{\text{obs}}$ on addition of a P_3–P_2 carbamate, that is, the addition of a hydrogen bond at P_2 NH, depends on the presence of other enzyme–substrate interactions. Specifically, the presence of a P_1 NH makes the added P_2 NH more favorable than when there is an ester at P_2–P_1; i.e., **3** (NH, Phe, NH) versus **7** (H, Phe, NH), $\Delta\Delta G_{\text{obs}} = -2.3$ kcal/mol, and **4a** (NH, Phe, O) versus **8a** (H, Phe, O), $\Delta\Delta G_{\text{obs}} = -1.0$ kcal/mol. The same effect is observed in the presence or absence of the P_2 Phe side chain; **3** (NH, Phe, NH) versus **7** (H, Phe, NH), $\Delta\Delta G_{\text{obs}} = -2.3$ kcal/mol, and **11** (NH, Gly, NH) versus **12** (H, Gly, NH), $\Delta\Delta G_{\text{obs}} = -1.6$ kcal/mol.

(*B*) *P_2 Side Chain.* $k_{\text{cat}}/K_{\text{M}}$ increases as the S_2 binding subsite is progressively occupied with hydrophobic interactions from a number of different side chains (unpublished results). On addition of a benzyl group to **11** (NH, Gly, NH) to give **3** (NH, Phe, NH), $\Delta\Delta G_{\text{obs}} = 4.0$ kcal/mol. The higher conformational entropy of Gly over Phe may account for part of this difference. The calculations of Brant et al. (1967) in-

---

[6] In this section, to facilitate the discussion of substrates involving the S_2 and S_1 subsites, the presence or absence of various groups will be denoted in parentheses after the compound number as follows: the atoms or groups occupying the positions for the P_2 NH, P_2 side chain, and P_1 NH are shown. Thus, MocPhe-O-CH$_2$-C(O)↑OMe is **4a** (NH, Phe, O). A proton is considered the equivalent of a Gly "side chain".

dicated that the conformational entropy of a Gly residue about the Φ and Ψ angles is only slightly higher than that of an amino acid with a non-β-branched side chain, with an upper limit on the order of 1.5 cal·K^{-1}·mol^{-1}. If Phe's lower conformational entropy is directly converted into binding energy, specificity would increase by a maximum of 0.4 kcal/mol with respect to a substrate with a P_2 Gly.

As is the case with the hydrogen-bonding interactions, the effect of side-chain binding in S_2 is interdependent with other substrate–enzyme interactions. The added side chain has the greatest effect when the P_2 and P_1 NH's are present, as witnessed by the difference for **12** (H, Gly, NH) versus **7** (H, Phe, NH) ($\Delta\Delta G_{obs}$ = 3.3 kcal/mol) and for **11** (NH, Gly, NH) versus **3** (NH, Phe, NH) ($\Delta\Delta G_{obs}$ = 4.0 kcal/mol).

(C) P_1 Hydrogen Bonding. The presence of the hydrogen bond to the P_1 NH was demonstrated and its strength evaluated as for the P_2 NH. Substitution of a P_2–P_1 ester for the amide demonstrates that the P_1 NH is hydrogen-bonded to the enzyme with $\Delta\Delta G_{obs}$ = −1.2 to −2.6 kcal/mol: **3** (NH, Phe, NH) versus **4a** (NH, Phe, O), $\Delta\Delta G_{obs}$ = −2.6 kcal/mol, **7** (H, Phe, NH) versus **8a** (H, Phe, O), $\Delta\Delta G_{obs}$ = −1.3 kcal/mol, and **18** (–, benzoyl, NH) versus **19** (–, benzoyl, O), $\Delta\Delta G_{obs}$ = −1.2 kcal/mol. The X-ray structures of α-chloroketone and stefin B inhibited papains (Drenth et al., 1976, Stubbs et al. 1990) indicate that the Asp 158 backbone carbonyl is the probable hydrogen-bond acceptor. Once again, the $\Delta\Delta G_{obs}$ of the P_1 NH depends on other enzyme–substrate interactions, with the presence of the P_2 NH and the P_2 Phe side chain cooperating to increase the effect of the P_1 NH.

Ester for Amide Substitution. The $\Delta\Delta G_{obs}$'s for the P_2 (−2.7 kcal/mol) and P_1 NH's (−2.6 kcal/mol) are based, at least in part, on the substitution of amide functionalities with esters at the site of the NH in question and therefore represent specificity energies. Initially, one might argue that these free energies represent a specificity *against* the sp^3 oxygen rather than *for* the amide NH. However, because an sp^3 oxygen has a smaller van der Waals radius than an NH in the hydrogen-bonding direction, selection against an oxygen would have to be electrostatic rather than steric. Any putative negative charge on the enzyme that could interact unfavorably with an ester's sp^3 oxygen would necessarily interact favorably with an amide NH; therefore, any discrimination by the enzyme between esters and amides will include a considerable favorable free-energy contribution from the amide NH.

For the P_2' NH, in contrast to the P_1 and P_2 NH's, the binding energy of −2.3 kcal/mol was determined independently of the ester and represents a true binding energy. The penalty to k_{cat}/K_M on placing an ester's alcoholic oxygen adjacent to the enzyme's hydrogen-bond acceptor in the S_2' subsite is approximately 0.4 kcal/mol as determined by the comparison of **4b** with **15**.

It is not possible to directly predict the energetic penalty for an oxygen in the position of the P_1 NH and P_2 NH; however, the $\Delta\Delta G_{obs} \geq$ 5.2 kcal/mol for the addition of a methyl carbamate to **17** to give **4b** argues for a strong hydrogen bonding of the P_1 NH. There is no hydrogen-bond donor in the position that the sp^3 oxygen would be expected to occupy, and the terminal methyl group is positioned adjacent to Gly 66 NH. A binding energy of better than −5.2 kcal/mol is too large to have been contributed by the C=O alone. Even assuming 1 kcal/mol in favorable dispersion energy being contributed by the H_3C-O–, a contribution on the order of 2 kcal/mol from the NH is required. This compares favorably with the specificity energy of −2.6 kcal/mol between ester and amide functionalities at this site and therefore represents the favorable binding energy of the NH.

In the case of the P_2 NH, it is not possible formally to estimate the incremental binding energy. However, given the known proximity of a hydrogen-bond acceptor (Gly 66 C=O) in the X-ray structures of α-chloroketone-inhibited papains (Drenth et al., 1976), it is likely that, of the $\Delta\Delta G_{obs}$ = 2.7 kcal/mol for an ester versus an amide at P_3–P_2 (**2** versus **1**), the incremental binding energy of the P_2 NH will be on the order of 2 kcal/mol.

Interdependence of Interaction Energies. The observed interdependence, or cooperativity, of effects at the P_2 NH, the P_2 side chain, and the P_1 NH are typical of enzyme-substrate interactions. While certain processes are recognized as being highly cooperative, such as protein folding (Creighton, 1983) and the association of myosin subunits (Erickson, 1989), and while the importance of cooperativity has been noted in theoretical discussions of catalysis (Jencks, 1975, 1981; Page & Jencks, 1971), there is a dearth of empirical evidence, and clear manifestations of cooperativity have not been recognized as such (Kowlessur et al., 1989). Our observation of increasingly favorable binding energies with better substrates is, to our knowledge, the first report of a gradation of cooperative effects.

When a favorable interaction, for example a hydrogen bond, is added to an enzyme–substrate complex, there is a loss of internal entropy of the complex. This includes contributions from the restriction of vibrational and overall rotational motions of the substrate and enzyme relative to each other, as well as restriction of dihedral bond rotations. This loss of entropy causes, then, an increase in the free energy of the enzyme–substrate complex (E·S) and the transition state (ES‡). The magnitude of the free energy increase depends on the tightness of binding of the substrate *before* the addition of the hydrogen bond. Thus the incremental specificity energy of the hydrogen bond is offset to a certain extent by this higher free energy of the enzyme–substrate complexes (i.e., E·S and ES‡). A "loose" complex where there is considerable freedom of motion of the substrate in the active site will undergo a greater entropy loss than a "tight" complex. The observable effect of this condition is that, on the addition of a hydrogen bond of a given incremental specificity energy to a substrate, the observed change in free energy ($\Delta\Delta G_{obs}$) depends on the presence of other interactions and will be greater when it is added to a good versus a poor substrate.

This effect is illustrated by comparing the addition of a P_1 NH to **4a** (NH, Phe, O) versus adding the same P_1 NH to **8a** (H, Phe, O) or **19** (–, benzoyl, O). The $\Delta\Delta G_{obs}$'s are −2.6, −1.3, and −1.2 kcal/mol, respectively. The presence in **4a** (NH, Phe, O) of the P_2 NH [and P_2 Phe side chain compared with **19** (–, benzoyl, O)] reduces the motions of complexed substrate and enzyme molecules relative to each other, making the entropic penalty upon addition of the P_1 NH less than with the other two substrates.

It is not necessary to specifically consider ground-state destabilization to account for the interdependence of interaction energies. Changes in the second-order rate constant k_{cat}/K_M reflect changes in the overall binding energy that is available, which may be expressed as changes in either k_{cat} or K_M. The concept of ground-state destabilization is concerned only with the partitioning of binding energy between formation of the noncovalent enzyme–substrate complex and decreasing the activation energy of the reaction, not the total amount of binding energy available.

Alternately phrased, because the entropy of the enzyme–substrate complex is higher for poor substrates, the restriction of motion necessary to form the hydrogen bond with the P_1

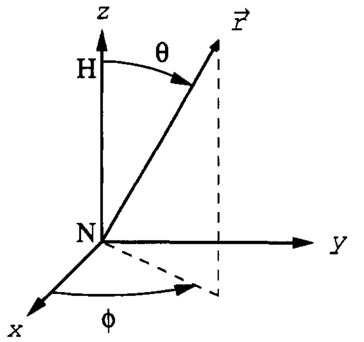

FIGURE 3: Coordinate system used for investigating region around hydrogen-bonding NH's.

NH raises the free energy of the enzyme–substrate complex (E·S) and the transition state ($ES^{\ddagger}$) *more* for poor substrates than for good. In each case, the NH group interacts equally strongly with the hydrogen-bond acceptor, but for poor substrates a larger portion of the interaction energy compensates for the higher free energy of the complex and therefore appears as a poorer $\Delta\Delta G_{obs}$. Precisely the same effect is observed with the P_2 NH and the P_2 Phe side chain as with the P_1 NH.

This principle is further illustrated by the results of Asboth et al. (1988), as compared with those of Williams et al. (1972). Substitution of the P_2 C=O of Z-PheGly↑OEt (k_{cat}/K_M = 120 000 $M^{-1}\cdot s^{-1}$ for papain) with C=S decreased specificity by 4.4 kcal/mol for papain (Asboth et al., 1988). These $\Delta\Delta G_{obs}$'s represent the loss of the C=O hydrogen bond plus the unfavorable free energy of fitting the C=S into a site that normally accommodates the smaller C=O. For *N*-benzoyl-Gly↑OMe (k_{cat}/K_M = 246 $M^{-1}\cdot s^{-1}$), $\Delta\Delta G_{obs}$ = 1.6 kcal/mol on substituting the P_2 carbonyl oxygen with a sulfur atom (Williams et al., 1972). As with the substrates reported in this study, the effect of oxygen to sulfur substitution at P_2 C=O depends on other enzyme–substrate interactions. The better the substrates, the more unfavorable the resulting effect.

The $\Delta\Delta G_{obs}$'s for the interactions discussed above represent the lower limits of incremental specificity (P_2 NH, P_1 NH) or binding (P_2 Phe side chain) energies. In particular, the incremental binding energy for a P_2 Phe side chain is more than the −4.0 kcal/mol observed here between **3** and **11**. The poorer of the two substrates in the comparison, **11**, is quite poor indeed and may therefore have a much higher entropy in the E·S and $ES^{\ddagger}$ complexes than the corresponding species for **3**. With the leaving-group interactions, the $\Delta\Delta G_{obs}$'s are based on relatively poor substrates and therefore may also significantly underestimate the incremental binding energies. The strengths of the hydrogen bonds relative to each other (P_1' C=O, $\Delta\Delta G_{obs}$ = −1.0 kcal/mol; P_2' NH, −2.3 kcal/mol) may represent differences in the entropy of hydrogen-bond formation. Upon formation of the hydrogen bond to the P_1' C=O, the movement of the P_2' NH becomes very restricted due to the planar geometry of the amide, making the entropic cost of forming the P_2' NH hydrogen bond relatively smaller.

Application of this observed cooperativity of energies of interaction is in the design of synthetic inhibitors of cysteine proteases, or indeed, in the manipulation of intermolecular binding in general. The highest possible specificity of an inhibitor for one protease over another will be achieved by utilizing all the bonding interactions common to the two enzymes, maximizing the effectiveness of those interactions that are different. For example, considering two hypothetical cysteine proteases that share the same peptide backbone hydrogen bonding as observed here but have different specificities for the P_2 side chain, the use of all the common backbone interactions in an inhibitor will allow the incremental binding energy of the P_2 side-chain interactions to be expressed as specificity between the two proteases and not be "lost" in entropy compensation.

The hydrogen bonds to the substrate's peptide backbone at P_2 NH, P_2 C=O, and P_1 NH all involve backbone atoms of the papain molecule: Gly 66 C=O, Gly 66 NH, and Asp 158 C=O, respectively. As such, these hydrogen-bonding interactions are expected to be conserved through most or all cysteine proteases. It would therefore be useful to include these interactions in any synthetic inhibitors of cysteine proteases in order to maximize selectivity.

In conclusion, we have shown that the active-site cleft of papain has hydrogen-bonding partners for the peptide backbone of substrates extending at least from the S_2 to the S_2' subsites. The lower limits of the incremental specificity energies for the P_2 NH and P_1 NH's are, respectively, −2.7 and −2.6 kcal/mol versus esters at these sites. The incremental binding energies are estimated to be on the order of 2 kcal/mol. The $\Delta\Delta G_{obs}$'s for P_2 Phe side chain (−4.0 kcal/mol), P_1' C=O (−1.0 kcal/mol), and P_2' NH (−2.3 kcal/mol) probably all underestimate the incremental binding energy at these points.

The observed interdependence of binding energies at each site on the acyl side of the catalytic site is characteristic of enzyme–substrate interactions and is largely due to a decreased entropic penalty with successive additions of favorable interactions.

The use of all the backbone hydrogen bonding interactions observed here will aid in producing synthetic inhibitors of cysteine proteases with the greatest possible selectivity.

ACKNOWLEDGMENTS

We gratefully acknowledge Dr. Ron Angus for the gift of several substrates.

APPENDIX

It is well-known that the electrostatic potential can describe the hydrogen-bonded geometries in small van der Waals complexes (Buckingham & Fowler, 1985). We have compared the electrostatic potential around NH for both amide- and carbamate-containing molecules as a means to assess their tendency to form hydrogen bonds. The electrostatic potential was calculated by using a distributed multipoles analysis (Stone, 1981) of the SCF wave function, calculated with a 3-21G basis set by using the ab initio program suite CADPAC (Amos, 1984). The atomic dipoles, quadrupoles, and higher moments explicity represent the nonspherical features in the atomic charge distribution, such as the lone-pair density.

The molecules were superimposed in such a way that the NH groups had identical coordinates, thus facilitating the comparison of the electrostatic potential. Figure 2 illustrates both molecules with the atomic labeling, while Figure 3 illustrates the coordinate system chosen to describe the vicinity of the NH bond. Some geometrical parameters from the work of Taylor et al. (1983) on the geometry of the NH···O=C hydrogen bond were used. The electrostatic potential was studied in the region of space around the NH bond, namely $\phi < 27°$ and $\theta < 360°$ (both angles were stepped in 3° increments). In addition, the distance from the nitrogen to the observation point was varied from 2.6 to 3.2 Å in 0.1-Å steps. This gave 7630 grid points in the vicinity of NH. The electrostatic potential from the atomic multipole moments up to rank 4[7] was calculated by using ORIENT (Price & Stone, 1987).

---

[7] A complete list of these atomic moments is available from the authors upon request.

The difference between the electrostatic potential of amide and carbamate was calculated by using

$$\Delta V(\bar{r}) = V(\bar{r})_{\text{carb}} - V(\bar{r})_{\text{amide}}$$

where $\bar{r}$ is within the region described above.

The results obtained are $\Delta V(\bar{r})_{\text{avg}} = -9.4$ kcal/mol over the 7630 grid points. This clearly shows in a quantitative manner that the electrostatic potential of carbamate is systematically smaller than that of amide in the region around the potentially hydrogen bonding NH. This is in accord with the experimental observation that carbamates are worse hydrogen-bond donors than amides.

Registry No. **1**, 7625-55-0; **2**, 131041-20-8; **3**, 101653-92-3; **4**, 131010-72-5; **5**, 131010-73-6; **6**, 41844-91-1; **8**, 131010-74-7; **9**, 131010-75-8; **10**, 501-52-0; **11**, 131010-76-9; **12**, 1906-82-7; **13**, 131010-77-0; **14**, 131010-78-1; **15**, 131010-79-2; **16**, 41844-71-7; **17**, 60397-80-0; **18**, 1205-08-9; **19**, 29747-05-5; AcPhe, 2018-61-3; GlyOMe, 616-34-2; papain, 9001-73-4; bromoacetate, 68-10-0; iodoacetamide, 144-48-9; 1-hydroxypropanone, 116-09-6; 1-hydroxy-2-butanone, 5077-67-8.

REFERENCES

Alecio, M. R., Dann, M. L., & Lowe, G. (1974) *Biochem. J. 141*, 495–501.

Amos, R. D. (1984) CADPAC: The Cambridge Analytical Derivatives Package, Publication CCP1/84/4 91984, SERC Daresbury Laboratory, Daresbury, Warrington, U.K.

Angus, R. H., Carey, P. R., & Storer, A. C. (1986) *Biochemistry 25*, 3304–3310.

Asboth, B., Majer, Z., & Polgar, L. (1988) *FEBS Lett. 233*, 339–341.

Baker, E. N., & Hubbard, R. E. (1984) *Prog. Biophys. Mol. Biol. 44*, 97–179.

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

Berti, P. J. (1990) M.Sc. Thesis, University of Ottawa, Ottawa, Ontario, Canada.

Bjorck, L., Akesson, P., Bohus, M., Trojnar, J., Abrahamson, M., Olafsson, I., & Grubb, A. (1989) *Nature 337*, 385–386.

Brant, D. A., Miller, W. G., & Flory, P. J. (1967) *J. Mol. Biol. 23*, 47–65.

Brocklehurst, K., Willenbrock, F., & Salih, E. (1987) in *Hydrolytic enzymes* (Neuberger, A., & Brocklehurst, K., Eds.) pp 39–158, Elsevier Biomedical Press, Amsterdam.

Brubacher, L. J., & Bender, M. L. (1966) *J. Am. Chem. Soc. 88*, 5871–5879.

Buckingham, A. D., & Fowler, P. W. (1985) *Can. J. Chem. 63*, 2018–2025.

Carey, P. R., Angus, R. H., Lee, H., & Storer, A. C. (1984) *J. Biol. Chem. 259*, 14357–14360.

Cohen, S. G., & Weinstein, S. Y. (1964) *J. Am. Chem. Soc. 86*, 5326–5330.

Coombs, G. H. (1982) *Parasitology 84*, 149–155.

Cornish-Bowden, A. (1975) *Biochem. J. 149*, 305–312.

Cornish-Bowden, A. (1976) in *Principles of enzyme kinetics*, Butterworths, London.

Creighton, T. E. (1983) in *Proteins, structures and molecular properties*, W. H. Freeman and Company, New York.

Demuth, H.-U. (1990) *J. Enzyme Inhib. 3*, 249–278.

Drenth, J., Kalk, K. H., & Swen, H. M. (1976) *Biochemistry 15*, 3731–3738.

Dresden, M. H., Payne, D. C., & Basch, P. F. (1982) *Mol. Biochem. Parasitol. 6*, 203–208.

Erickson, H. P. (1989) *J. Mol. Biol. 206*, 465–474.

Fersht, A. R. (1988) *Biochemistry 27*, 1577–1580.

Finklestein, A. V., & Janin, J. (1989) *Protein Eng. 3*, 1–3.

Ingles, D. W., & Knowles, J. R. (1968) *Biochem. J. 108*, 561–569.

IUPAC–IUB Joint Commission on Biochemical Nomenclature (1985) *J. Biol. Chem. 260*, 14–42.

Jencks, W. P. (1975) *Adv. Enzymol. Relat. Areas Mol. Biol. 43*, 219–410.

Jencks, W. P. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4046–4050.

Kowlessur, D., Topham, C. M., Thomas, E. W., O'Driscoll, M., Templeton, W., & Brocklehurst, K. (1989) *Biochem. J. 258*, 755–764.

Lowe, G., & Williams, A. (1965) *Biochem. J. 96*, 199–204.

Merz, K. M., Jr., & Kollman, P. A. (1989) *J. Am. Chem. Soc. 111*, 5649–5658.

Nefkens, G. H. L., & Nivard, R. J. F. (1965) *Recl. Trav. Chim. Pays-Bas 84*, 1315–1318.

Page, M. I., & Jencks, W. P. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 1678–1683.

Polgar, L. (1987) in *Hydrolytic enzymes* (Neuberger, A., & Brocklehurst, K., Eds.) pp 159–200, Elsevier, Amsterdam.

Price, S. L., & Stone, A. J. (1987) *J. Chem. Phys. 86*, 2859–2868.

Rich, D. H. (1986) in *Proteinase inhibitors* (Barrett, A. J., & Salvesen, G., Eds.) pp 153–178, Elsevier, Amsterdam.

Schechter, I., & Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157–162.

Schechter, I., & Berger, A. (1968) *Biochem. Biophys. Res. Commun. 32*, 898–902.

Shah, H. N., & Gharbia, S. E. (1989) *FEMS Microbiol. Lett. 61*, 213–218.

Sluyterman, L. A. A. E., Wijdenes, J. (1970) *Biochim. Biophys. Acta 200*, 593–595.

Stone, A. J. (1981) *Chem. Phys. Lett. 83*, 233–239.

Storer, A. C., & Carey, P. R. (1985) *Biochemistry 24*, 6808–6818.

Storer, A. C., Lee, H., & Carey, P. R. (1983) *Biochemistry 22*, 4789–4796.

Storer, A. C., Angus, R. H., & Carey, P. R. (1988) *Biochemistry 27*, 264–268.

Stubbs, M. T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B., & Turk, V. (1990) *EMBO J. 9*, 1939–1947.

Taylor, R., Kennard, O., & Vershichel, J. (1983) *J. Am. Chem. Soc. 105*, 5761–5766.

Whitaker, J. R., & Bender, M. L. (1965) *J. Am. Chem. Soc. 87*, 2728–2737.

Wilkinson, A. J., Fersht, A. R., Blow, D. M., & Winter, G. (1983) *Biochemistry 22*, 3581–3586.

Williams, A., Lucas, E. C., Rimmer, A. R., & Hawkins, H. C. (1972) *J. Chem. Soc., Perkin Trans. 1*, 627–633.