Ionization characteristics and chemical influences of aspartic acid residue 158 of papain and caricain determined by structure-related kinetic and computational techniques: multiple electrostatic modulators of active-centre chemistry

Michael A. NOBLE*, Sheraz GUL*¹, Chandra S. VERMA† and Keith BROCKLEHURST*²

*Laboratory of Structural and Mechanistic Enzymology, Department of Molecular and Cellular Biology, Division of Biomedical Sciences, Queen Mary and Westfield College, University of London, Mile End Road, London E1 4NS, U.K., and †Structural Biology Laboratory, Department of Chemistry, University of York, Heslington, York YO1 5DD, U.K.

The pK_a of (Asp¹⁵⁸)-CO₂H of papain (EC 3.4.22.2) was determined as 2.8 by using 4-chloro-7-nitrobenzofurazan (Nbf-Cl) as a reactivity probe targeted on the thiolate anion component of the Cys²⁵/His¹⁵⁹ nucleophilic-acid/base motif of the catalytic site. The possibility of using Nbf-Cl for this purpose was established by modelling the papain-Nbf-Cl Meisenheimer intermediate by using QUANTA/CHARMM and performing molecular orbital calculations with MOPAC interfaced with Cerius 2. A pH-dependent stopped-flow kinetic study of the reaction of papain with Nbf-Cl established that the striking rate maximum at pH 3 results from reaction in a minor ionization state comprising (Cys²⁵)-S⁻/(His¹⁵⁹)-Im⁺H (in which Im represents imidazole) produced by protonic dissociation of (Cys²⁵)-SH/(His¹⁵⁹)-Im⁺H with pK_a 3.3 and (Asp¹⁵⁸)-CO₂H. Although the analogous intermediate in the reaction of caricain (EC 3.4.22.30) with Nbf-Cl has similar geometry, the pH- k profile (k being the second-order rate constant) lacks a rate maximum under acidic conditions. This precludes the experimental de-

termination of the pK_a value of (Asp¹⁵⁸)-CO₂H of caricain, which was calculated to be 2.0 by solving the linearized Poisson–Boltzmann equation with the program UHBD ('University of Houston Brownian dynamics'). A value lower than 2.8 had been predicted by consideration of the hydrogen-bonded networks involving Asp¹⁵⁸ and its microenvironments in both enzymes. The difference between these pK_a values (values not previously detected in reactions of either enzyme) accounts for the lack of the rate maximum in the caricain reaction and for the differences in the electronic absorption spectra of the two S-Nbf-enzymes under acidic conditions. The concept of control of cysteine proteinase activity by multiple electrostatic modulators, including (Asp¹⁵⁸)-CO₂H, which modifies traditional mechanistic views, is discussed.

Key words: 4-chloro-7-nitrobenzofurazan, cysteine proteinases, Poisson–Boltzmann equation, reactivity probe.

INTRODUCTION

The value of the papain family of cysteine proteinases in the study of enzyme active-centre chemistry and catalytic mechanism [1,2] derives from the natural variation within it of both structure and functional characteristics and from the existence of a highly reactive nucleophilic centre (S⁻) in the common Cys-His ion-pair system (Cys⁻His⁺) of the catalytic site, which has a nucleophilic role in the catalytic act. Papain and caricain are two such variants. Both are constituents of *Carica papaya*, with well-defined three-dimensional structures [the crystal structure of papain, from the Brookhaven Protein Data Bank, is 9pap.pdb [3]; the homology-modelled structure of caricain (papaya proteinase Ω) is described in [4] and its crystal structure is 1ppo.pdb [5]]. The major features of these structures are closely similar but they differ in detail in certain regions, including in the environments close to the catalytic sites (see the Results and discussion section). The existence of the Cys⁻His⁺ ion pair permits data from kinetics of catalysis to be usefully complemented by those from reactions with thiol-specific time-dependent inhibitors used as reactivity probes. Such data from reactions with thoughtfully

designed probes can provide mechanistic insights that are not always available from a kinetic study of catalysis alone.

One of the least well understood aspects of enzyme catalysis and active-centre chemistry is the range and nature of the kinetic consequences of electrostatic effects initiated by ionizations at various distances from and orientations to the catalytic site [6–9]; evidence continues to accumulate that electrostatic effects [10] are a major factor in determining the behaviour of cysteine proteinases [4,11–22]. The carboxy group of Asp¹⁵⁸ is the ionizable group nearest to the catalytic sites in both enzymes [being separated from the (Cys²⁵)-S⁻/(His¹⁵⁹)-Im⁺H (in which Im represents imidazole) ion pairs by approx. 7 Å [1]] and its role in active-centre chemistry has long been controversial [2,20]. Interest in this side chain was renewed by the demonstration that the ion-pair state [produced by protonic dissociation from (Cys²⁵)-SH/(His¹⁵⁹)-Im⁺H across pK_a 3.3 in papain and 2.9 in caricain] is not sufficient for catalytic competence [1]. This requires the protonic dissociation of another group with a pK_a of approx. 4 in both enzymes. In the case of caricain, k_{cat}/K_m for the catalysed hydrolysis of some substrates is further enhanced by protonic dissociation with a pK_a of approx. 5.5–5.8 [4,20]. The

Abbreviations used: L-BAPNA, α-N-benzoyl-L-arginine 4-nitroanilide; eu, electron unit; HOMO, highest occupied molecular orbital; Im, imidazole; LUMO, lowest unoccupied molecular orbital; Nbf-Cl, 4-chloro-7-nitrobenzofurazan; 2PDS, 2,2'-dipyridyl disulphide.

¹ Present address: Division of Protein Structure, MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, U.K.

² To whom correspondence should be addressed (e-mail kb1@qmw.ac.uk).

deletion of both electrostatic and hydrogen-bonding properties of Asp¹⁵⁸ by site-directed mutagenesis results in a substantial loss of catalytic activity [20,23] but definitive parameter values from pH-dependent kinetic experiments for the relevant mutants have not been reported. In some cases, such as for papain Asp¹⁵⁸ → Ala, the activity was too low to measure [23], and in others the analysis and interpretation of the complex pH- k_{cat}/K_m data were not soundly based, in part because the report of the pK_a values for (Cys²⁵)-S⁻/(His¹⁵⁹)-Im⁺H formation [1] had not yet appeared. Clarification required the direct determination of the ionization characteristics of (Asp¹⁵⁸)-CO₂H in each enzyme.

In the present study, as part of a programme of a more general investigation of electrostatic and hydrogen-bonding influences on catalytic-site characteristics of the cysteine proteinases, we have performed a combination of reactivity probe kinetics and spectroscopic and computer modelling studies targeted on the characteristics and influences of the side chain of Asp¹⁵⁸ in papain and caricain. Modelling suggested that a study of the pH-dependent kinetics of the reaction of the catalytic-site thiol group (Cys²⁵) of papain with 4-chloro-7-nitrobenzofurazan (Nbf-Cl) at low pH might provide the pK_a of (Asp¹⁵⁸)-CO₂H because of the predicted assistance to the reaction by the undissociated form of the carboxy group. This proved to be true and the pK_a thus determined (2.8) is lower than any of the kinetically influential pK_a values reported previously, the lowest being 3.3, the value for Cys-His⁺ formation [1]. The pK_a of (Asp¹⁵⁸)-CO₂H of caricain was not accessible kinetically but was calculated to be 2.0 by solving the linearized Poisson-Boltzmann equation with finite difference methods implemented in the program UHBD ('University of Houston Brownian dynamics'). These pK_a values eliminate Asp¹⁵⁸ as the source of either of the kinetically influential ionizations with pK_a values of approx. 4 and 5.7 respectively and demonstrate that the (Cys²⁵)-S⁻/(His¹⁵⁹)-Im⁺H ion pairs are always under the influence of (Asp¹⁵⁸)-CO₂⁻. Removal of this carboxylate by site-directed mutagenesis results in a substantial decrease in catalytic activity. (Asp¹⁵⁸)-CO₂⁻ can therefore be regarded as an electrostatic modulator permanently in the 'on' state, whose influence is augmented markedly by ionization of the unidentified modulator across pK_a approx. 4 in both enzyme variants and additionally, for some caricain-substrate systems, by the ionization of another unidentified modulator across pK_a 5.7. This necessity for electrostatic control modifies the traditional views of the cysteine proteinase mechanism, including previous assignments of the kinetically influential pK_a of 4 variously to Asp¹⁵⁸, His¹⁵⁹ and Cys²⁵. The structural basis of the substantial differences in the kinetic behaviour of papain and caricain and in the spectroscopic characteristics of their S-Nbf-derivatives, deduced by modelling, is discussed.

MATERIALS AND METHODS

Enzymes

Procedures for the purification of papain [24] and caricain [4], including those for the production of the fully active enzymes by covalent chromatography [25,26], and evaluation by spectroscopic titration with 2,2'-dipyridyl disulphide (2PDS) [27] have been described. The source of both enzymes was dried papaya latex supplied by Sigma (e.g. catalogue no. P-3375, lot no. 127F0630).

Substrate and thiol-specific reagents

α -N-Benzoyl-L-arginine 4-nitroanilide (L-BAPNA) was purchased from Sigma, and 2PDS and Nbf-Cl from Aldrich. 2PDS

was purified to constant melting point (58 °C) as described by Pinitglang et al. [1]. Nbf-Cl was recrystallized from 50 % (v/v) ethanol as yellow needles, melting point 96 °C (Shipton et al. [28] give a melting point of 96–97 °C).

Assay of catalytic activity

The catalytic activity of papain or caricain during the purification procedure was measured by determining initial rates of release of 4-nitroaniline from the substrate L-BAPNA by spectral analysis at 410 nm (ϵ_{410} 8800 M⁻¹·cm⁻¹ [29]) by a procedure similar to that described for ficin [30]. When fractions containing the chymopapains needed to be collected, catalytic activity towards t-butyloxycarbonylglycine 4-nitrophenyl ester ('Boc-Gly-4NPE') was monitored in addition to that towards L-BAPNA, to identify fractions containing chymopapain M by using the procedure described in [31].

Stopped-flow kinetics of the reactions of papain and caricain with Nbf-Cl

An approx. 20 mM stock solution of Nbf-Cl (molecular mass 199.6 Da) was prepared by dissolving recrystallized Nbf-Cl (40 mg) in spectroscopic ethanol. Solutions were prepared freshly for use on the same day and stored in a stoppered bottle protected from light with aluminium foil. Accurate concentrations were determined by the measurement of A_{343} of a suitably diluted sample at pH 6.0 and quantified by using an ϵ_{343} of 1.17×10^4 M⁻¹·cm⁻¹ [28]. Reactions of papain and caricain with Nbf-Cl were monitored at 25 °C in aqueous buffers [I = 0.1 M, containing 6.7 % (v/v) ethanol] by stopped-flow analysis; for papain at A_{412} , the isosbestic point of the pH-dependent spectra of the product, S-Nbf-papain, and for caricain at A_{417} , the pH-independent value of λ_{max} of S-Nbf-caricain. Kinetic runs were performed only in the pH range 2–5 to ensure specific reaction with the catalytic-site thiol group (see the Results and discussion section). Having shown that the reactions obeyed overall second-order kinetics at pH 3.0 and 4.5, the reactions at various pH values were studied with $[E]_0 = 10 \mu\text{M}$ and $[\text{Nbf-Cl}]_0 = 800 \mu\text{M}$. Initial rates (v) were determined from the linear increase in A_{412} or A_{417} recorded over periods from 10 s to 4 min. Values of the second-order rate constant (k , M⁻¹·s⁻¹) were calculated by using the values of $[E]_0$, $[\text{Nbf-Cl}]_0$ and either ϵ_{412} 11360 M⁻¹·cm⁻¹ or ϵ_{417} 8400 M⁻¹·cm⁻¹. In some cases complete progress curves were recorded and the values of k , calculated from the observed (pseudo-)first-order rate constant, and $[\text{Nbf-Cl}]_0$ were shown to be in good agreement with the values calculated from initial rates.

Analysis of pH- k data

Kinetic models differing in the number and reactivity of the reactive ionization states (Scheme 1) [18,32] were evaluated by a combination of methods with SKETCHER [4,33], weighted non-linear regression and the statistical procedure discussed by Mannervik [34] as described by Pinitglang et al. [1].

Preparation and spectral analysis of S-Nbf-papain and S-Nbf-caricain

An activated, activator-free solution of the enzyme (15 ml of approx. 25 μM) in 0.2 M sodium acetate buffer, pH 4.0, containing 1 mM EDTA was mixed with an equal volume of 3 mM Nbf-Cl in the same buffer but containing 13 % (v/v) spectroscopic ethanol; the mixture was stirred overnight at room temperature. The volume of the solution was then decreased to approx. 3–4 ml

For models (a) and (b) the rate equations are equations (1) and (2) respectively.

$$k = \frac{\tilde{k}_1}{1 + \frac{[\text{H}^+]}{K_I} + \frac{K_{II}}{[\text{H}^+]} + \frac{K_{II}K_{III}}{[\text{H}^+]^2}} + \frac{\tilde{k}_2}{1 + \frac{[\text{H}^+]^2}{K_I K_{II}} + \frac{[\text{H}^+]}{K_{II}} + \frac{K_{III}}{[\text{H}^+]}} + \frac{\tilde{k}_3}{1 + \frac{[\text{H}^+]^3}{K_I K_{II} K_{III}} + \frac{[\text{H}^+]^2}{K_{II} K_{III}} + \frac{[\text{H}^+]}{K_{III}}} \quad (1)$$

$$k = \frac{\tilde{k}_1}{1 + \frac{[\text{H}^+]}{K_I} + \frac{K_{II}}{[\text{H}^+]}} + \frac{\tilde{k}_2}{1 + \frac{[\text{H}^+]^2}{K_I K_{II}} + \frac{[\text{H}^+]}{K_{II}}} \quad (2)$$

Scheme 1 Kinetic models and associated pH-dependent rate equations for reactions in three protonic states (a) and two protonic states (b)

XH₃ to X represent generalized reactant protonic states with charges omitted; K_I–K_{III} are macroscopic acid dissociation constants and $\tilde{k}_1$ – $\tilde{k}_3$ are pH-independent rate constants; for reactions of time-dependent inhibitors these relate to second-order reactions of Cys²⁵ with the electrophilic reagent.

by ultrafiltration through a PM10 membrane (Amicon, Danvers, MA, U.S.A.) and excess Nbf-Cl was removed by gel filtration with a Sephadex G-25 column (25 cm × 2.5 cm) and 0.01 M sodium acetate buffer, pH 4.5, containing 1 mM EDTA as eluent, to produce approx. 15 ml of a yellow solution of S-Nbf-enzyme. Absorption spectra were recorded over 250–500 nm in aqueous buffers, I = 0.1 M, in the pH range approx. 2.5–5.5. The set of pH-dependent spectra of S-Nbf-papain contains an isosbestic point at 412 nm with ϵ_{412} 11360 M⁻¹·cm⁻¹ based on ϵ_{402} 12100 M⁻¹·cm⁻¹ [35]. The pH-independent spectrum of S-Nbf-caricain has a λ_{max} of 417 nm with ϵ_{max} 8400 M⁻¹·cm⁻¹.

Computer modelling of the microenvironments of Asp¹⁵⁸ in papain and caricain and in their Meisenheimer intermediates with Nbf-Cl

The crystal structures of papain [3] and caricain [5], each with the blocking group bonded to S^γ of the catalytic-site Cys²⁵ removed, were used as the basis for modelling. Buried crystallographic water molecules (25 in papain and 24 in caricain) were retained. The CHARMM force field [36] was used to establish the modelled structures as described previously [9]. This included a 12 Å (1 Å = 0.1 nm) cut-off imposed on the non-bonded interactions and solvent screening by the use of a distance-dependent dielectric with a scaling coefficient of 1. When explicit solvent was modelled, a constant dielectric of 1 was used. The geometry of Cys²⁵ was allowed to relax while, initially, the rest of the protein was restrained with a harmonic potential of 10.0 kcal/mol per Å (1 kcal = 4.18 kJ). Then the restraint on the residues within 8 Å of the Cys²⁵ side chain was decreased to 1.0 kcal/mol per Å and minimizations were performed until the potential gradient was less than 0.001 kcal/mol per Å. The structures thus obtained were used for the pK_a calculations described below.

To construct the models used to examine the orientation of Asp¹⁵⁸, with its side chain in both the undissociated and dissociated states in the free enzyme molecules and in their Meisenheimer intermediates with Nbf-Cl, the structures were subjected to further minimizations with the protocol outlined by Bruccoleri and Karplus [37]. The two oxygen atoms of carboxylate side chains were each assigned equal negative charge [–0.57 electron unit (eu)] and the central C atom a charge of +0.14 eu to give a net charge on the carboxylate anion of –1.0 eu. In the corresponding undissociated states a proton was added to one or other of the oxygen atoms and the net charge on the side chain set to zero. The structures were subjected to cycles of steepest-descent and adopted-basis-Newton–Raphson minimizations. In the final cycle, the whole structure was unrestrained and minimization was continued until the change in potential energy was less than 0.00001 kcal/mol per Å. This was necessary for computation of the values of the energy barriers to inter-conversion of conformers, which was performed with the conjugate peak refinement ('CPR') method [38] implemented in the TRAVel module of CHARMM. The CPR algorithm is very sensitive to the character of the potential-energy surface. This procedure has been found to give results in close agreement with experiment in other systems [39,40].

To construct models of the Meisenheimer intermediates of papain and caricain with Nbf-Cl, Asp¹⁵⁸ of the enzyme was first modelled un-ionized with the proton retained on atom OD2 of the carboxy side chain. This was then orientated to provide maximum interaction with its microenvironment, which resulted in OD2 pointing towards the active centre. The Nbf-Cl ligand was constructed with the Molecular Editor module of QUANTA. It was then docked manually into the active centre and S^γ of Cys²⁵ was bonded covalently to C4 of Nbf-Cl to create the

Meisenheimer intermediate as an approximation to the postulated transition state for the reaction. Initially the part of the system comprising the side chain of Cys²⁵ bonded to C4 of Nbf-Cl was subjected to energy minimization to relieve strain resulting from any bad contacts. An extensive search was then made of the minimum energy conformations resulting from variation of the dihedral angles across the C^α–C^β and C^β–C^γ bonds of Cys²⁵ and the S^γ–C4 bond of the intermediate. The lowest-energy structures were subjected to molecular dynamics simulations at 1000 K for 100 ps; structures were extracted from the trajectory at 5 ps intervals and minimized. The lowest-energy minimized structures were each solvated with a sphere of water molecules that extended to 15 Å from the S^γ of Cys²⁵. The added water molecules were subjected to minimizations followed by 10 ps of dynamics at 300 K and again minimized to allow them to relax to their local minimum energy in the field of the Meisenheimer intermediates. Then, with the Nbf-Cl ligand unconstrained, all atoms within 8 Å of it were restrained by a harmonic potential of 0.01 kcal/mol per Å, those between 8 and 15 Å from the ligand were restrained by a harmonic potential of 5.0 kcal/mol per Å and those beyond 15 Å were maintained in their original fixed positions. The whole system was subjected to minimization until the change in the potential energy gradient was less than 0.001 kcal/mol per Å.

Calculation of the difference between the pK_a values of (Asp¹⁵⁸)-CO₂H in papain and caricain

The procedure used to model the structures of the enzymes used for the ΔpK_a calculations was described above. The value of ΔG, the free energy change for the conversion of the side chain of Asp¹⁵⁸ from CO₂H (charge 0.0) to CO₂[−] (charge −1.0 eu) in each enzyme was calculated [41,42] by solving the linearized Poisson–Boltzmann equation by using finite-difference methods [43] as implemented in the program UHBD (version 5.1 [44], supplied by Professor J. Briggs, University of Houston, Houston, TX, U.S.A.) on an SGI R10000 computer. The following fractional charge distribution was used: each Lys and the N-termini, +1; each Arg, +0.5 on each terminal N; each His ring N, +0.5; the C-termini and each Asp and Glu, −1 except for Asp¹⁵⁸ initially. The dielectric constants of protein and solvent were set to 2 and 78 respectively [1]. Other experimental conditions were simulated with an ionic concentration of 100 mM, a 2 Å Stern layer around a Richards surface of accessibility [45], CHARMM atomic radii [36] and a temperature of 300 K.

The Richards accessible surface was generated for a probe of radius 1.4 Å; this delineated the regions of the two dielectrics [45]. Finite difference focusing and dielectric boundary smoothing [46] were used to compute electrostatic potentials in a three-step protocol. Initially a 400 Å × 400 Å × 400 Å cubic grid partitioned into 4 Å × 4 Å × 4 Å blocks was used. This was followed by two focusing runs with first a 100 Å × 100 Å × 100 Å cubic grid partitioned into 1 Å × 1 Å × 1 Å blocks and then a 10 Å × 10 Å × 10 Å cubic grid partitioned into 0.1 Å × 0.1 Å × 0.1 Å blocks. The initial grid accommodated the protein (longest length 50 Å) and a surrounding region of length 175 Å on each side of the protein. The energy required to change the charge on the side chain of Asp¹⁵⁸ from 0.0 to −1.0 eu was calculated by multiplying the potentials generated by the surrounding charges at the constituent atoms (C^γ, OD1 and OD2 of the carboxy group of Asp¹⁵⁸) by the formal charges on the atoms in the charged state (0.14, −0.57 and −0.57 eu respectively). For the calculation involving the uncharged side chain of Asp¹⁵⁸, the charges on the side chains of all carboxy groups were set to 0.0. This assumption was made because the kinetic data for the reaction of caricain

with Nbf-Cl suggested an abnormally low pK_a for (Asp¹⁵⁸)-CO₂H.

This pK_a was calculated with the analogous pK_a for papain, determined kinetically from the reaction with Nbf-Cl (2.8), and the calculated value of:

$$\Delta\Delta G = \Delta G_{\text{caricain}} - \Delta G_{\text{papain}} = RT \ln [K_{\text{a(papain)}} / K_{\text{a(caricain)}}]$$

where $R = 1.9872$ cal/K per mol, $T = 298.15$ K and $K_{\text{a(papain)}}$ and $K_{\text{a(caricain)}}$ are the acid dissociation constants for (Asp¹⁵⁸)-CO₂[−] formation in the two enzymes. Thus $\Delta pK_{\text{a}} = \Delta\Delta G / 1.3645$. To estimate the standard errors on the ΔpK_a values, the enzymes (with all ionizable groups ionized) were subjected to 100 ps of constrained molecular dynamics simulation with weak coupling to a heat bath at 300 K. The Poisson–Boltzmann calculations were repeated for eight structures taken at 5 ps intervals from the molecular dynamics trajectory during the final 45 ps with and without additional minimization as described above.

Calculation of the effect of a neighbouring carboxy group on the reactivity of Nbf-Cl towards nucleophilic attack at C4 by a thiolate anion

To assess quantitatively the effects of the presence of a carboxy group within hydrogen-bonding distance of the nitro group of Nbf-Cl on its reactivity towards nucleophilic attack at C4 by a thiolate anion, the coefficients of the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) for a model system were examined. The following procedure was performed with the program MOPAC, version 6.0 [47], interfaced with Cerius 2 (MSI, San Diego, CA, U.S.A.) for the calculations. Nbf-Cl (net charge 0.0) and CH₃CH₂CO₂H were brought together such that hydrogen-bonding of CO₂H to the oxygen atom of the nitro group remote from the furazan ring was optimized. The proton was then removed from the CO₂H group and the net effect of the opposing electrostatic interactions within the Nbf-Cl/CO₂[−] pair was optimized. The Nbf-Cl/CH₃CH₂-CO₂H and Nbf-Cl/CH₃CH₂CO₂[−] pairs were subjected to semi-empirical molecular orbital calculations employing the AM1 Hamiltonian operator and using Slater-type atomic orbitals ('STO')-3G basis sets. These use three Gaussian functions to form each Slater-type atomic orbital. In a separate set of calculations, the energy of the HOMO of CH₃CH₂S[−] (net charge −1 eu) and the value of the largest coefficient (which lies in the p_z direction, i.e. perpendicular to the plane of the benzofurazan ring system when the mercaptide is positioned for nucleophilic attack at C4 of Nbf-Cl) were determined.

RESULTS AND DISCUSSION

Nbf-Cl as a potential reactivity and spectroscopic probe for the investigation of the ionization characteristics of Asp¹⁵⁸ in papain and caricain

Nbf-Cl (1) was developed originally as a fluorescent labelling reagent for lysine, tyrosine and cysteine side chains in proteins to characterize their local environments and conformational states [48–54]. Reaction occurs at C4 of Nbf-Cl via a Meisenheimer intermediate (2) to yield an Nbf derivative (3) with the elimination of Cl[−], as exemplified in Scheme 2. The site of labelling and its environment can be characterized by the λ_{max} value of the long-wavelength band in the electronic absorption spectrum by relating it to those of Nbf derivatives of low-molecular-mass model compounds (amines, 475 nm [48]; mercaptans, 425 nm [50]; phenols, 390 nm [54]). The reaction of Nbf-Cl with cysteine proteinases such as papain and caricain at pH values of 5.0 or lower is known to be restricted to their catalytic-site thiol groups.

Scheme 2 Reaction of (Cys²⁵)-S⁻ of papain and caricain with Nbf-Cl (1) to form the S-Nbf-enzyme (3) via a Meisenheimer intermediate (2)

Scheme 3 Schematic representation of hydrogen-bonding interactions within the minimum energy conformations involving the ionized side chains (a, b) and the un-ionized side chains (c, d) of Asp¹⁵⁸ and their microenvironments in papain (a, c) and caricain (b, d)

For further details see the text.

This selectivity is plausibly attributed to the low p*K*_a values related to the generation of nucleophilic character in the (Cys²⁵)-S⁻/(His¹⁵⁹)-Im⁺H ion pairs (3.3 in papain and 2.9 in caricain [1]). At pH values greater than approx. 5 the reaction is not thiol-specific and reaction with amino groups occurs.

The potential of Nbf-Cl as a thiol-specific reactivity probe for this enzyme family was suggested by the results of early work [28,35,55]. This was necessarily restricted, however, by the then lack of appropriate kinetic technology required to obviate problems of protein unfolding at low pH. A key observation was the increase in *k* with decrease in pH for the reaction of Nbf-Cl with papain that was first demonstrated by Allen and Lowe [35].

In the present study the possibility that this increase in reactivity could be attributed to hydrogen-bonding of an acidic group in papain, such as (Asp¹⁵⁸)-CO₂H to the nitro group of Nbf-Cl, was explored by modelling prior to kinetic and spectroscopic studies. Semi-empirical molecular orbital calculations were performed on the reactions of CH₃CH₂S⁻ with Nbf-Cl in the presence and absence of a hydrogen bond between CH₃CH₂CO₂H and the nitro group of Nbf-Cl, as a model system. The largest LUMO coefficient for the model reaction is substantially negative (−0.53) when the CH₃CH₂CO₂H-nitro group hydrogen bond exists. This is not true in the presence of CH₃CH₂CO₂⁻, when the value of the corresponding coefficient is +0.003. These values predict an

increase in the reactivity of Nbf-Cl towards nucleophilic attack at C4 by a thiolate anion (largest HOMO coefficient for the model reaction, +0.97) consequent on hydrogen-bonding of an undissociated carboxy group with the nitro group of the probe. Similarly, the replacement of an undissociated carboxy group within hydrogen-bonding distance of the nitro group at the end of the π -system of Nbf-Cl by a carboxylate anion would be expected to result in a change in the $\lambda_{\max}$ of the electronic absorption spectrum of the Nbf moiety. It remained to be established that the dispositions of Cys²⁵ and Asp¹⁵⁸ of the enzymes with respect to C4 and the nitro group of Nbf-Cl respectively were appropriate, potentially, both to permit the Asp¹⁵⁸-assisted alkylation of Cys²⁵ and to provide a pH-dependent Nbf spectrum. The results of the investigation of the conformations of the enzyme–probe combinations by modelling are described below.

Interactions within the minimum energy conformations of the ionized and un-ionized side chains of Asp¹⁵⁸ in papain and caricain and their microenvironments

Investigation of these interactions in the free enzymes (summarized schematically in Scheme 3) was performed as a prelude to the analogous investigation of the interactions in the Meisenheimer intermediates formed by reaction with Nbf-Cl. When the side chain of Asp¹⁵⁸ of papain is ionized there is only one low-energy conformation (Scheme 3a), in which OD1 of Asp¹⁵⁸ accepts a hydrogen bond from the backbone –NH of Ala¹³⁷ and another hydrogen bond from the backbone –NH of Ala¹³⁶. This donor also supplies a hydrogen bond to OD2 of Asp¹⁵⁸, which points towards the region in which the Nbf-Cl ligand resides in the Meisenheimer intermediate (Figures 1a and 1c). When the side chain of Asp¹⁵⁸ is un-ionized there are two low-energy conformations that differ by only 0.1 kcal/mol and interconvert rapidly across a low-energy barrier (2.3 kcal/mol relative to the lower energy state). In one of these conformations (Scheme 3c) the hydrogen atom of the carboxy group of Asp¹⁵⁸ resides on OD2 and the hydroxy group donates a hydrogen bond to the oxygen atom of its own backbone carbonyl group; OD1 accepts two hydrogen bonds, one from each backbone –NH of Ala¹³⁶ and Ala¹³⁷. In this conformation the hydroxy group of (Asp¹⁵⁸)-CO₂H should be able to transfer its hydrogen bond readily from the backbone carbonyl oxygen atom to the nitro group of the Nbf-Cl ligand in the Meisenheimer intermediate (Figure 1c).

When the side chain of Asp¹⁵⁸ of caricain is ionized, as in papain there is one unique minimum-energy conformation (Scheme 3b). In this the carboxylate anion is stabilized by five hydrogen bonds as against three for papain. OD1 accepts hydrogen bonds from both the backbone –NH and the side-chain –N⁺H₃ of Lys¹³⁷ and from the backbone –NH of Ser¹³⁶. OD2 accepts hydrogen bonds from both the backbone –NH and the side-chain –OH of Ser¹³⁶. When the side chain of Asp¹⁵⁸ is un-ionized, as in papain there are two low-energy conformations. These conformations also differ in energy by only 0.1 kcal/mol and interconvert rapidly across a low-energy barrier (1.3 kcal/mol relative to the lower energy state). In both of these conformations the hydrogen atom of the carboxy group of Asp¹⁵⁸ resides on OD2. In the conformation shown in Scheme 3(d) the hydroxy group of (Asp¹⁵⁸)-CO₂H donates a hydrogen bond to the oxygen atom of the hydroxy group of Ser¹³⁶ and does not itself accept a hydrogen bond. The three hydrogen bonds accepted by OD1 are donated by the backbone –NH and the side chain –N⁺H₃ of Lys¹³⁷ and the backbone –NH of Ser¹³⁶.

The differences between the hydrogen-bonding patterns in the two enzymes arise from the fact that residues 136 and 137 are Ser

and Lys respectively in caricain, whereas they are both Ala in papain. These differences were suggested to be noteworthy by Topham et al. [4] when the homology-modelled structure of papaya proteinase Ω (i.e. caricain) was constructed. The hydrogen-bonding patterns shown in Scheme 3 are of interest in connection with attempts to predict or account for the relative pK_a values of (Asp¹⁵⁸)-CO₂H in the two enzymes. pK_a values derive from the differences in free energies of conjugate acid and base forms and it is well established for low-molecular-mass systems that acid dissociation is markedly affected by solvation, particularly of the charged species. The five hydrogen bonds donated to (Asp¹⁵⁸)-CO₂[–] in caricain, as against three in papain, might be expected to produce greater relative stabilization of the carboxylate anion in the former and thus both a lower pK_a value and a decrease in electrostatic influence. Both of these conclusions are in accord with the results of the experiments involving the Nbf label as a spectroscopic reporter group and Nbf-Cl (1, Scheme 2) as a reactivity probe, and with the result of the electrostatic potential calculations discussed below.

Interactions within the minimum energy conformations of the ionized and un-ionized side chains of Asp¹⁵⁸ of the Meisenheimer intermediates of papain and caricain with Nbf-Cl

The thiolate anion component of the (Cys²⁵)-S[–]/(His¹⁵⁹)-Im⁺H ion pairs of papain and caricain reacts with Nbf-Cl (1, Scheme 2) to form the *S*-Nbf-enzymes (3) via a Meisenheimer intermediate (2), which is considered to be similar in structure to the postulated transition state for the reaction [56]. Figure 1 shows views of the catalytic-site regions of papain and caricain with the side chain of Cys²⁵ bonded to C4 of Nbf-Cl in a Meisenheimer intermediate analogous to those of the free enzymes. When the carboxy group of Asp¹⁵⁸ is undissociated, the hydrogen atom resides on OD2 in the minimum-energy conformations for both enzymes shown in Figures 1(c) and 1(d). The arrangements of hydrogen bonds donated to the carbonyl oxygen atom (OD1) are similar to those in the analogous low-energy conformations of the free enzymes [two hydrogen bonds in papain (Scheme 3c and Figure 1c) and three hydrogen bonds in caricain (Scheme 3d and Figure 1d)].

The key difference between the free enzyme structures and those of the Meisenheimer intermediates is that in the latter, the hydroxy group of (Asp¹⁵⁸)-CO₂H (OD2-H) donates a hydrogen bond to the nitro group of the Nbf-Cl moiety rather than to a site on the protein. This enzyme–ligand interaction suggests the possibility for transition-state stabilization in the reactions with Nbf-Cl. This would be predicted to produce the rate enhancement required to provide the opportunity for kinetic determination of the pK_a value of (Asp¹⁵⁸)-CO₂H. In the caricain intermediate the nitro group accepts a hydrogen bond also from the side chain of Ser¹³⁶ (which is Ala¹³⁶ in papain) whether or not the carboxy group of Asp¹⁵⁸ is ionized. The carboxylate anion of Asp¹⁵⁸ accepts hydrogen bonds from the same donors in the Meisenheimer intermediates [three in the papain intermediate (Figure 1a) and five in the caricain intermediate (Figure 1b)] as in the free enzyme molecules (Schemes 3a and 3b). The interaction between Asp¹⁵⁸ and the Nbf-Cl ligand changes in both enzymes from repulsion when the Asp¹⁵⁸ side chain is ionized to attraction when it is undissociated. This leads to a greater movement of the ligand in the papain intermediate than in the caricain intermediate. The additional hydrogen bond from the side chain of Ser¹³⁶ to the nitro group in the caricain intermediate contributes to the different orientation of the ligand in this intermediate (Figure 1d) from that in the papain intermediate (Figure 1c). This difference is augmented when the carboxy group of Asp¹⁵⁸ is ionized by the greater electrostatic repulsion between the

Figure 1 Minimum energy conformations of the ionized side chain (a, b) and the un-ionized side chain (c, d) of Asp¹⁵⁸ of the Meisenheimer intermediates of Cys²⁵ of papain (a, c) and caricain (b, d) and Nbf-Cl and their respective microenvironments

Colour codes: for the protein, C (slate green), polar H (white), O (red), N (blue), S (yellow); for Nbf-Cl, C (white), O (magenta), N (orange-brown), Cl (pink); hydrogen bonds are shown as white broken lines. Geometrical characteristics are given for those involving Asp¹⁵⁸ and those involving the Nbf-Cl ligand. The sulphur atom of the catalytic-site Cys²⁵ is bonded to C4 of the Nbf-Cl moiety. OD2 of Asp¹⁵⁸ points towards the nitro group of Nbf-Cl. Hydrogen-bond characteristics [distance between heavy atoms X and Y (in Å) and bond angle X–H–Y] are as follows: (a) for (Ala¹³⁷)-NH... (Asp¹⁵⁸)-OD1, 2.1 and 159°; for (Ala¹³⁶)-NH... (Asp¹⁵⁸)-OD1, 2.0 and 146°; for (Ala¹³⁶)-NH... (Asp¹⁵⁸)-OD2, 2.5 and 138°; (the distance between (Asp¹⁵⁸)-OD1 and O1 of the nitro group of Nbf-Cl is 3.5 Å); (c) for (Ala¹³⁷)-NH... (Asp¹⁵⁸)-OD1, 2.1 and 163°; for (Ala¹³⁶)-NH... (Asp¹⁵⁸)-OD1, 1.8 and 144°; for (Asp¹⁵⁸)-OD2-H... (Nbf-Cl)-NO₂ (O1), 1.9 and 160°; (b) for (Lys¹³⁷)-N⁺H₃... (Asp¹⁵⁸)-OD1, 1.7 and 163°; for (Lys¹³⁷)-NH... (Asp¹⁵⁸)-OD1, 2.1 and 158°; for (Ser¹³⁶)-NH... (Asp¹⁵⁸)-OD1, 2.1 and 132°; for (Ser¹³⁶)-NH... (Asp¹⁵⁸)-OD2, 2.0 and 143°; for (Ser¹³⁶)-OH... (Asp¹⁵⁸)-OD2, 1.9 and 135°; for (Ser¹³⁶)-OH... (Nbf-Cl)-NO₂ (O1), 2.1 and 126°; (d) for (Lys¹³⁷)-N⁺H₃... (Asp¹⁵⁸)-OD1, 1.6 and 160°; for (Lys¹³⁷)-NH... (Asp¹⁵⁸)-OD1, 2.3 and 136°; for (Ser¹³⁶)-NH... (Asp¹⁵⁸)-OD1, 2.3 and 133°; for (Ser¹³⁶)-NH... (Asp¹⁵⁸)-OD2, 2.3 and 153°; for (Asp¹⁵⁸)-OD2-H... (Ser¹³⁶)-OH, 2.0 and 131°; for (Asp¹⁵⁸)-OD2-H... (Nbf-Cl)-NO₂ (O1), 2.0 and 128°; for (Ser¹³⁶)-OH... (Nbf-Cl)-NO₂ (O1), 1.9 and 132°.

carboxylate anion and the nitro group in the papain intermediate (Figure 1a) than in the caricain intermediate, in which the carboxylate anion accepts more hydrogen bonds (Figure 1b).

In summary, the results of the modelling studies demonstrate the possibility of using Nbf-Cl as a reactivity probe to determine the pK_a values of (Asp¹⁵⁸)-CO₂H in both free papain and free

Figure 2 pH dependence of the second-order rate constant (k) for the reaction of Nbf-Cl at 25 °C in aqueous buffers ($I = 0.1 \text{ M}$) containing 6.7% (v/v) ethanol with the catalytic-site thiol group (Cys²⁵) of papain (filled circles) (a) and caricain (open circles) (b)

The points are mean values for three determinations; S. D. values were at most $\pm 10\%$ of the means. For the papain reaction (a) the continuous line is theoretical for eqn (1) in Scheme 1 and the following values of the parameters, with those of the pH-independent rate constants ($\bar{k}_1$ – $\bar{k}_3$) in $\text{M}^{-1} \text{s}^{-1}$: $pK_1 = 2.70$, $pK_{II} = 3.32$, $pK_{III} = 4.18$, $\bar{k}_1 = 95.0$, $\bar{k}_2 = 4.0$ and $\bar{k}_3 = 1.5$. The broken line is theoretical for eqn (2) in Scheme 1 and the following values of the parameters with those of the pH-independent rate constants ($\bar{k}_1$ and $\bar{k}_2$) in $\text{M}^{-1} \text{s}^{-1}$: $pK_1 = 2.70$, $pK_{II} = 3.32$, $\bar{k}_1 = 95.0$ and $\bar{k}_2 = 1.5$. Standard errors on the pK_a values are in the range 0.04–0.06.

caricain molecules. The conformations predicted for the enzyme–probe complexes are consistent with simultaneous nucleophilic attack of (Cys²⁵)-S[−] at C4 of Nbf-Cl and hydrogen-bonding of (Asp¹⁵⁸)-CO₂H with its nitro group; this hydrogen-bonding interaction is calculated to increase the rate of the nucleophilic attack.

Kinetics of the reactions of Nbf-Cl with the catalytic-site thiol groups (Cys²⁵) of papain and caricain

Both reactions were shown to obey overall second-order kinetics, first-order in each reactant, at pH values 3.0 and 4.5. The pH-dependences of k were then determined in the pH ranges approx. 2.4–5.0 for the papain reaction and approx. 2.0–5.0 for the caricain reaction. The upper limit to the pH range was dictated by the need to retain thiol specificity, i.e. to avoid reaction of amino groups, and the lower limit by conformational instability of the enzyme resulting in the sudden loss of the high reactivity as the pH is decreased beyond the limits stated above. The pH-dependence of k for the papain reaction is shown in Figure 2(a). The pH– k profile contains a striking rate maximum at pH 3. By using modern stopped-flow kinetic technology, which obviates complications from protein unfolding at low pH, it was possible to define experimentally observed (macroscopic) pK_a values of the major kinetically influential ionizations as 2.7 and 3.3 and that of the minor influence as 4.2.

Evidence that the pK_a of (Asp¹⁵⁸)-CO₂H of papain is 2.8

Two of the kinetically influential pK_a values of Figure 2(a), 3.3 and 4.2, correspond closely to similar values observed in reactions with some other reactivity probes and in catalysis [1]. It is now established that the former characterizes the formation of the catalytic-site (Cys²⁵)-S[−]/(His¹⁵⁹)-Im⁺H ion pair and the latter the electrostatic modulator required to generate substantial catalytic

competence. The third pK_a , revealed by the Nbf-Cl reaction in Figure 2(a) as 2.7, has not been detected previously in reactions of papain. The evidence from the modelling studies strongly suggests that this pK_a is associated with (Asp¹⁵⁸)-CO₂H. The observed value would be expected to have been shifted downwards by 0.1 from the intrinsic value as a result of mixing with the known pK_a of (Cys²⁵)-SH/(His¹⁵⁹)-Im⁺H (3.3); the relationship between observed (macroscopic) and intrinsic (group) pK_a values and of overlapping kinetically influential ionizations is discussed in [18]. The corresponding upward perturbation of pK_a 3.3 would be suppressed by mixing with the pK_a of the modulator (4.2), which is seen to affect the pH– k profile of Figure 2(a) in the region around pH 4.

The single pK_a deduced by Allen and Lowe [35] as 3.7, in what is only a partial pH– k profile under acidic conditions, might be a composite of the pK_a values 3.3 and 4.2 of Figure 2(a). Their assignment of the pK_a of 3.7 to (Asp¹⁵⁸)-CO₂H takes no account of the additional requirement for the thiolate anion component of the ion pair (Cys²⁵)-S[−]/(His¹⁵⁹)-Im⁺H for nucleophilic attack at C4 of Nbf-Cl and must be incorrect. It is now clear that the reactivity in the rate maximum of Figure 2(a) must be due to the minor component of the central ionization state of Scheme 4 in which (Cys²⁵)-S[−]/(His¹⁵⁹)-Im⁺H and (Asp¹⁵⁸)-CO₂H coexist to a significant extent. The major component comprising (Cys²⁵)-SH/(His¹⁵⁹)-Im⁺H and (Asp¹⁵⁸)-CO₂H lacks thiolate anion and consequent nucleophilic reactivity towards Nbf-Cl.

Kinetic evidence that the pK_a of (Asp¹⁵⁸)-CO₂H of caricain is significantly lower than that of papain

In marked contrast with the pH– k profile for the reaction of papain with Nbf-Cl (Figure 2a), that for the corresponding reaction of caricain (Figure 2b) does not contain the striking rate maximum under acidic conditions. Rather, k increases with increasing pH across the pH region within which the (Cys²⁵)-S[−]/(His¹⁵⁹)-Im⁺H ion pair is produced (pK_a 2.9), and other kinetically influential ionizations (pK_a values 3.8 and 5.7) have been detected [1,4,20]. The absence of a rate maximum, and thus of a pK_a that can be convincingly assigned to (Asp¹⁵⁸)-CO₂H of caricain in Figure 2(b), is most obviously accounted for by the hypothesis that this side chain is essentially fully dissociated even at low pH. A pK_a for this group significantly lower than that in papain (2.8) might well not supply a sufficient concentration of the hydrogen-bond donor (Asp¹⁵⁸)-CO₂H to produce significant activation of Nbf-Cl by hydrogen-bonding. The hypothesis that the pK_a of (Asp¹⁵⁸)-CO₂H in caricain might be lower than that in papain is supported by the different hydrogen-bonding arrays involving the (Asp¹⁵⁸)-CO₂H/(Asp¹⁵⁸)-CO₂[−] conjugate pair (Scheme 3), the electrostatic potential calculations and the pH-independent nature of the electronic absorption spectrum of S-Nbf-caricain and its λ_{max} value, as discussed below.

Determination of the pK_a of (Asp¹⁵⁸)-CO₂H of caricain as 2.0 by electrostatic potential calculations

Poisson–Boltzmann calculations, performed as described in the Materials and methods section, provided a ΔpK_a , being [pK_a of (Asp¹⁵⁸)-CO₂H in caricain] – [pK_a of (Asp¹⁵⁸)-CO₂H in papain], as -0.8 ± 0.04 . This, together with the pK_a of (Asp¹⁵⁸)-CO₂H in papain (2.8) determined kinetically by using Nbf-Cl as a reactivity probe and correcting the observed value (2.7) for the perturbation resulting from overlap with pK_a 3.3, gives the calculated pK_a of (Asp¹⁵⁸)-CO₂H in caricain as 2.0.

Spectral characteristics of S-Nbf-papain and S-Nbf-caricain: spectroscopic evidence that the pK_a of (Asp¹⁵⁸)-CO₂H is lower in caricain than in papain

Electronic absorption spectra of S-Nbf-papain and S-Nbf-caricain at several pH values under acidic conditions are shown in Figure 3. For S-Nbf-papain, the long-wavelength band (Figure 3a) has a λ_{max} of 424 nm at pH 2.6 and 408 nm at pH 4.9; the set of spectra has an isosbestic point at 412 nm. By contrast, the spectrum of S-Nbf-caricain (Figure 3b), is pH-independent in the pH range 2.2–4.8, with a λ_{max} of 417 nm and ϵ_{417} 8400 M⁻¹·cm⁻¹. This value of λ_{max} , like that of S-Nbf-papain at pH 4.9 (408 nm), is blue-shifted relative to that of unperturbed low-molecular-mass models (λ_{max} 425 nm [35]) but to a smaller extent. The blue-shifted spectra are interpreted as those of the Nbf-chromophore perturbed by (Asp¹⁵⁸)-CO₂⁻ in the two enzymes. The differences between the two sets of spectra are consistent with greater transition energies when the Nbf chromophore is in the environment of the (Asp¹⁵⁸)-CO₂⁻ anion and the different pK_a values of (Asp¹⁵⁸)-CO₂H in papain (2.8) and caricain (2.0) discussed above. The smaller perturbation of λ_{max} for S-Nbf-caricain is consistent with the greater shielding of (Asp¹⁵⁸)-CO₂⁻ by hydrogen-bonding with Lys¹³⁷ and Ser¹³⁶ relative to that provided by Ala¹³⁶ and Ala¹³⁷ in S-Nbf-papain (Figure 1).

Conclusion

The pK_a values of (Asp¹⁵⁸)-CO₂H in papain and caricain have been determined for the first time as 2.8 and 2.0 respectively. Thus the catalytic-site (Cys²⁵)-S⁻/(His¹⁵⁹)-Im⁺H ion pairs produced by protonic dissociation across pK_a value 3.3 in papain and 2.9 in caricain [1] are always under the influence of the (Asp¹⁵⁸)-CO₂⁻ anion. Removal of the electrostatic and hydrogen-bonding capacity of this charge by site-directed mutagenesis results in a substantial decrease in catalytic activity [20,23], although an Asp → Asn mutation does not produce a significant change in the ionization characteristics of the Cys²⁵/His¹⁵⁹ system [57]. (Asp¹⁵⁸)-CO₂⁻ can therefore be regarded as an electrostatic modulator essentially in the 'on' state over the pH range in which catalytic activity develops. However, this modulator alone is not sufficient for efficient catalytic competence: another, with a pK_a value of approx. 4 in papain and approx. 3.7 in caricain, is required to be deprotonated. Both we (A. Khan, S. Hussain, I. F. Connerton, N. J. Cummings, M. A. J. Taylor, E. W. Thomas and K. Brocklehurst, unpublished work) and Ikeuchi et al. [22] have shown that the pK_a of approx. 3.7 in caricain cannot

Figure 3 Electronic absorption spectra of S-Nbf-papain in the pH range 2.6–4.9 (a) and S-Nbf-caricain in the pH range 2.2–4.8 (b)

For S-Nbf-papain, $\lambda_{\text{max}} = 424$ nm at pH 2.6 and 408 nm at pH 4.9; the set of pH-dependent spectra has an isosbestic point at 412 nm with ϵ_{412} 11360 M⁻¹·cm⁻¹. The spectrum of S-Nbf-caricain is pH-independent, with a λ_{max} of 417 nm and ϵ_{417} 8400 M⁻¹·cm⁻¹.

be assigned to (Glu⁵⁰)-CO₂H, the anion of which and the (Cys²⁵)-S⁻/(His¹⁵⁹)-Im⁺H ion pair have one of the highest electrostatic interaction energies in the enzyme [1]. Another modulator (pK_a

Scheme 4 Four ionization states of the Cys²⁵/Asp¹⁵⁸/His¹⁵⁹ systems of papain and caricain under acidic conditions; the other central ionization state containing (Cys²⁵)-SH, (Asp¹⁵⁸)-CO₂H and (His¹⁵⁹)-Im is not shown

K_I and K_{II} are the macroscopic acid dissociation constants characteristic of the two stages of ionization determined experimentally.

value approx. 5.7) is operative in caricain in its interaction with some substrates but not in papain [4,20]. The identity of the two unassigned electrostatic modulators in caricain should be revealed by a comprehensive protein engineering programme involving all of the potential candidates, which is currently in progress.

We thank Professor J. Briggs (University of Houston, Houston, TX, U.S.A.) for supplying the program UHBD. We also thank the Biotechnology and Biological Sciences Research Council for a postdoctoral appointment for C. S. V., a research studentship for M. A. N. and support for the infrastructure of the York Structural Biology Laboratory, and the Engineering and Physical Sciences Research Council for a research studentship for S. G.

REFERENCES

- Pinitglang, S., Watts, A. B., Patel, M., Reid, J. D., Noble, M. A., Gul, S., Bokth, A., Naeem, A., Patel, H., Thomas, E. W. et al. (1997) A classical enzyme active centre motif lacks catalytic competence until modulated electrostatically. *Biochemistry* **36**, 9968–9982.
- Brocklehurst, K., Watts, A. B., Patel, M., Verma, C. and Thomas, E. W. (1998) Cysteine proteinases. in *Comprehensive Biological Catalysis*. (Sinnott, M. L., ed.), vol. 1, pp. 381–423, Academic Press, London.
- Kamphuis, I. G., Kalk, K. H., Swarte, M. B. A. and Drenth, J. (1984) Structure of papain refined at 1.65 Å resolution. *J. Mol. Biol.* **179**, 233–256.
- Topham, C. M., Sali, E., Frazao, C., Kowlessur, D., Overington, J. P., Thomas, M., Brocklehurst, S. M., Patel, M., Thomas, E. W. and Brocklehurst, K. (1991) Structure–function relationships in the cysteine proteinases actinidin, papain and papaya proteinase Ω deduced by knowledge-based modelling and active centre characteristics determined by two-hydronic-state reactivity probe kinetics and kinetics of catalysis. *Biochem. J.* **280**, 79–92.
- Pickersgill, R. W., Rizkallah, P., Harris, G. W. and Goodenough, P. W. (1991) Determination of the structure of papaya proteinase Ω . *Acta Cryst.* **B47**, 766–761.
- Allewell, N. M. and Oberoi, H. (1991) Electrostatic effects in protein folding, stability, and function. *Methods Enzymol.* **202**, 3–19.
- Loewenthal, R., Sancho, J., Reinikainen, T. and Fersht, A. R. (1993) Long-range surface charge–charge interactions in proteins. Comparison of experimental results with calculations from a theoretical method. *J. Mol. Biol.* **232**, 574–583.
- Jackson, S. E. and Fersht, A. R. (1993) Contribution of long-range electrostatic interactions to the stabilization of the catalytic transition state of the serine protease subtilisin BPN'. *Biochemistry* **32**, 13909–13916.
- Plou, F. J., Kowlessur, D., Malthouse, J. P. G., Mellor, G. W., Hartshorn, M. J., Pinitglang, S., Patel, H., Topham, C. M., Thomas, E. W., Verma, C. and Brocklehurst, K. (1996) Characterization of the electrostatic perturbation of a catalytic site (Cys-S⁻/(His)-Im⁺H ion-pair in one type of serine proteinase architecture by kinetic and computational studies on chemically-mutated subtilisin variants. *J. Mol. Biol.* **257**, 1088–1111.
- Duncan, G. D., Huber, C. P. and Welsh, W. J. (1992) Molecular orbital studies of the structure and reactivities of model substrate intermediates in the deacylation of the cysteine proteinase papain. *J. Am. Chem. Soc.* **114**, 5784–5794.
- Willenbrock, F. and Brocklehurst, K. (1984) Natural structural variation in enzymes as a tool in the study of mechanism exemplified by a comparison of the catalytic-site structure and characteristics of cathepsin B and papain. pH-dependent kinetics of the reactions of cathepsin B from bovine spleen and from rat liver with a thiol-specific two-protonic-state reactivity probe (2,2'-dipyridyl disulphide) and with a specific synthetic substrate (*N*-benzyloxy-carbonyl-L-arginyl-L-arginine-2-naphthyl- amide). *Biochem. J.* **222**, 805–814.
- Willenbrock, F. and Brocklehurst, K. (1985) Preparation of cathepsins B and H by covalent chromatography and characterization of their catalytic sites by reaction with a thiol-specific two-protonic-state reactivity probe. Kinetic study of cathepsins B and H extending into alkaline media and a rapid spectroscopic titration of cathepsin H at pH 3–4. *Biochem. J.* **227**, 521–528.
- Kowlessur, D., Topham, C. M., Thomas, E. W., O'Driscoll, M., Templeton, W. and Brocklehurst, K. (1989) Identification of signalling and non-signalling binding contributions to enzyme reactivity. Alternative combinations of binding interactions provide for change in transition-state geometry in reactions of papain. *Biochem. J.* **258**, 755–764.
- Pickersgill, R. W., Sumner, I. G., Collins, M. E. and Goodenough, P. W. (1989) Structural and electrostatic differences between actinidin and papain account for differences in activity. *Biochem. J.* **257**, 310–312.
- Patel, M., Thomas, M. P., Kayani, I. S., Mellor, G. W., Thomas, E. W. and Brocklehurst, K. (1993) Ficin: a cysteine proteinase with binding site-catalytic site signalling characteristics intermediate between those of papain and actinidin. *Biochem. Soc. Trans.* **21**, 216S.
- Mellor, G. W., Patel, M., Thomas, M. P., Kowlessur, D., Sreedharan, S. K., Thomas, E. W. and Brocklehurst, K. (1993) New substrate-derived thiol-specific time-dependent inhibitors for the characterisation of the coupling of binding site interactions with catalytic site chemistry in the cysteine proteinases. *Biochem. Soc. Trans.* **21**, 215S.
- Mellor, G. W., Thomas, E. W., Topham, C. M. and Brocklehurst, K. (1993) Catalytic site characteristics of the porcine calpain II 80 kDa/18 kDa heterodimer revealed by selective reaction of its essential thiol group with two-hydronic-state time-dependent inhibitors: evidence for a catalytic site Cys/His interactive system and an ionizing modulatory group. *Biochem. J.* **290**, 289–296.
- Brocklehurst, K. (1994) A sound basis for pH-dependent kinetic studies on enzymes. *Protein Eng.* **7**, 291–299.
- Thomas, M., Verma, C., Boyd, S. M. and Brocklehurst, K. (1995) The structural origins of the unusual specificities observed in the isolation of chymopapain M and actinidin by covalent chromatography and the lack of inhibition of chymopapain M by cystatin. *Biochem. J.* **306**, 39–46.
- Taylor, M. A. J., Baker, K. C., Connerton, I. R., Cummings, N. J., Harris, G. W., Henderson, I. M. J., Jones, S. T., Pickersgill, R. W., Sumner, I. G., Warwicker, J. and Goodenough, P. W. (1994) An unequivocal example of cysteine proteinase activity affected by multiple electrostatic interactions. *Protein Eng.* **7**, 1267–1276.
- Katerelos, N. A. and Goodenough, P. W. (1996) Rapid reaction kinetics and structural determination of a cysteine proteinase mutant imply that residue 158 in caricain has a major effect upon the ability of the active site histidine to protonate a dipyriddy probe. *Biochemistry* **35**, 14763–14772.
- Ikeuchi, Y., Katerelos, N. and Goodenough, P. W. (1998) The enhancing of a cysteine proteinase activity at acidic pH by protein engineering, the role of glutamic 50 in the enzyme mechanism of caricain. *FEBS Lett.* **437**, 91–96.
- Ménard, R., Khouri, H. E., Plouffe, C., Laflamme, P., Dupras, R., Vernet, T., Tessier, D. C., Thomas, D. Y. and Storer, A. C. (1991) Importance of hydrogen-bonding interactions involving the side-chain of Asp 158 in the catalytic mechanism of papain. *Biochemistry* **30**, 5531–5538.
- Baines, B. S. and Brocklehurst, K. (1979) A necessary modification to the preparation of papain from any high quality latex of *Carica papaya* and evidence for the structural integrity of the enzyme produced by traditional methods. *Biochem. J.* **177**, 541–548.
- Brocklehurst, K., Carlsson, J., Kierstan, M. P. J. and Crook, E. M. (1973) Covalent chromatography: preparation of fully active papain from dried papaya latex. *Biochem. J.* **133**, 573–584.
- Brocklehurst, K. (1996) Covalent chromatography by thiol–disulfide interchange using solid-phase alkyl 2-pyridyl disulfides. In *Proteins Labfax* (Price, N. C., ed.), pp. 65–71, Bios Scientific Publishers Ltd., Oxford, and Academic Press, San Diego.
- Baines, B. S. and Brocklehurst, K. (1978) A spectrophotometric method for the detection of contaminant chymopapains in preparations of papain. Selective modification of one type of thiol group in the chymopapains by a two-protonic-state reagent. *Biochem. J.* **173**, 345–347.
- Shipton, M., Stuchbury, T. and Brocklehurst, K. (1976) 4-Chloro-7-nitrobenzo-2-oxa-1,3-diazole as a reactivity probe for the investigation of the thiol proteinases. Evidence that ficin and bromelain may lack carboxy groups conformationally equivalent to that of aspartic acid-158 of papain. *Biochem. J.* **159**, 235–244.
- Erlanger, B. P., Kokowsky, N. and Cohen, W. (1961) The preparation and properties of two new chromogenic substrates of trypsin. *Arch. Biochem. Biophys.* **95**, 271–278.
- Malthouse, J. P. G. and Brocklehurst, K. (1976) Preparation of fully active ficin by covalent chromatography and characterization of its active centre using 2,2'-dipyridyl disulphide as a reactivity probe. *Biochem. J.* **159**, 221–234.
- Thomas, M., Topham, C. M., Kowlessur, D., Mellor, G. W., Thomas, E. W., Whitford, D. and Brocklehurst, K. (1994) Structure of chymopapain M the late-eluted chymopapain deduced by comparative modelling techniques and active centre characteristics determined by pH-dependent kinetics of catalysis and of reactions with time-dependent inhibitors: the Cys-25/His-159 ion pair is insufficient for catalytic competence in both chymopapain M and papain. *Biochem. J.* **300**, 805–820.
- Brocklehurst, K. (1996) pH-dependent enzyme kinetics. In *Enzymology Labfax* (Engel, P. C., ed.), pp. 175–198, Bios Scientific Publishers Ltd., Oxford, and Academic Press, San Diego.
- Brocklehurst, S. M., Topham, C. M. and Brocklehurst, K. (1990) A general kinetic equation for multihydronic state reactions and rapid procedures for parameter evaluation. *Biochem. Soc. Trans.* **18**, 598–600.
- Mannervik, B. (1982) Regression analysis, experimental error, and statistical criteria in the design and analysis of experiments for discrimination between rival kinetic models. *Methods Enzymol.* **87C**, 370–390.
- Allen, G. and Lowe, G. (1973) Investigation of the active site of papain with fluorescent probes. *Biochem. J.* **133**, 679–686.

- 36 Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. and Karplus, M. (1983) CHARMM — a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217
- 37 Bruccoleri, R. E. and Karplus, M. (1986) Spatially constrained minimization of macromolecules. *J. Comput. Chem.* **7**, 165–175
- 38 Fischer, S. and Karplus, M. (1992) Conjugate peak refinement — an algorithm for finding reaction paths and accurate transition states in systems with many degrees of freedom. *Chem. Phys. Lett.* **194**, 252–261
- 39 Verma, C. S., Fischer, S., Caves, L. S. D., Roberts, G. C. K. and Hubbard, R. E. (1996) Calculation of the reaction pathway for the aromatic ring flip in methotrexate complexed to dihydrofolate reductase. *J. Phys. Chem.* **100**, 2510–2518
- 40 Fischer, S., Verma, C. S. and Hubbard, R. E. (1998) Rotation of structural water inside a protein: calculation of the rate and vibrational entropy of activation. *J. Phys. Chem. B* **102**, 1797–1805
- 41 Gilson, M. K., Rashin, A., Fine, R. and Honig, B. (1985) On the calculation of electrostatic interactions in proteins. *J. Mol. Biol.* **184**, 503–516
- 42 Gilson, M. K. and Honig, B. (1988) Calculation of the total electrostatic energy of a macromolecular system — solvation energies, binding-energies, and conformational-analysis. *Proteins Struct. Funct. Genet.* **4**, 7–18
- 43 Sharp, K. and Honig, B. (1989) Lattice models of electrostatic interactions — the finite-difference Poisson–Boltzmann method. *Chem. Scr.* **29A**, 71–74
- 44 Madura, J. D., Briggs, J. M., Wade, R. C., Davis, M. E., Luty, B. A., Ilin, A., Antosiewicz, J., Gilson, M. K., Bagheri, B., Scott, L. R. and McCammon, J. A. (1995) Electrostatics and diffusion of molecules in solution: simulations with the University of Houston Brownian Dynamics Program. *Comput. Phys. Commun.* **91**, 57–95
- 45 Antosiewicz, J., McCammon, J. A. and Gilson, M. K. (1994) Prediction of pH-dependent properties of proteins. *J. Mol. Biol.* **238**, 415–436
- 46 Gilson, M. K., Davis, M. E., Luty, B. A. and McCammon, J. A. (1993) Computation of electrostatic forces on solvated molecules using the Poisson–Boltzmann equation. *J. Phys. Chem.* **97**, 3591–3600
- 47 Stewart, J. J. P. (1990) Special Issue — MOPAC — a semi-empirical molecular-orbital program. *J. Comp.-Aid. Mol. Design* **4**, 1–45
- 48 Ghosh, P. B. and Whitehouse, M. W. (1968) 7-Chloro-4-nitrobenzo-2-oxa-1,3-diazole: a new fluorogenic reagent for amino acids and other amines. *Biochem. J.* **108**, 155–156
- 49 Ghosh, P. B. and Whitehouse, M. W. (1968) Potential antileukemic and immunosuppressive drugs. Preparation and in vitro pharmacological activity of some 2,1,3-benzoxadiazoles (benzofurazans) and their N-oxides (benzofuroxans). *J. Med. Chem.* **11**, 305–311
- 50 Birkett, D. J., Price, N. C., Radda, G. K. and Salmon, A. G. (1970) The reactivity of SH groups with a fluorogenic reagent. *FEBS Lett.* **6**, 346–348
- 51 Price, N. C., Cohen, M. and Schirmer, R. H. (1975) Fluorescent and spin label probes of the environments of the sulfhydryl groups of porcine muscle adenylate kinase. *J. Biol. Chem.* **250**, 644–652
- 52 Stappen, R., Dierks, T., Broer, A. and Kramer, R. (1992) Probing the active site of the reconstituted aspartate/glutamate carrier from mitochondria. *Eur. J. Biochem.* **210**, 269–277
- 53 Ben-Efraim, I., Strahilevitz, J., Bach, D. and Schai, Y. (1994) Secondary structure and membrane localization of synthetic segments and a truncated form of the IsK (minK) protein. *Biochemistry* **33**, 6966–6973
- 54 Fayer, R. S., Katina, C. B. and Abrahamson, E. W. (1973) The use of NBD chloride (7-chloro-4-nitrobenzo-2-oxa-1,3-diazole) in detecting amino acids and as an N-terminal reagent. *Anal. Biochem.* **53**, 290–294
- 55 Baines, B. S., Allen, G. and Brocklehurst, K. (1977) The highly electrophilic character of 4-chloro-7-nitrobenzofurazan and possible consequences for its application as a protein-labelling reagent. *Biochem. J.* **163**, 189–192
- 56 Bunnett, J. F. (1958) Mechanism and reactivity in aromatic nucleophilic substitution reactions. *Quart. Rev. Chem. Soc.* **12**, 1–16
- 57 Watts, A., Hafeez, A., Gul, S., Verma, C., Thomas, E. W. and Brocklehurst, K. (1997) Effects of site-specific mutations on the kinetically influential ionisations of papain. *Biochem. Soc. Trans.* **25**, 84S

Received 19 May 2000/20 July 2000; accepted 9 August 2000