α-Helix Dipoles and Catalysis: Absorption and Raman Spectroscopic Studies of Acyl Cysteine Proteases†

John D. Doran‡ and Paul R. Carey*,§

Department of Biochemistry, University of Ottawa, Ottawa, Ontario K1H 8M5, Canada, and Department of Biochemistry, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, Ohio 44106-4935

Received March 18, 1996; Revised Manuscript Received May 28, 1996⊗

ABSTRACT: Raman and absorption spectroscopic data are combined with the deacylation rate constants for a series of acyl cysteine proteases to provide insight into the role of α-helix dipoles in rate acceleration. The Raman spectra, obtained by Raman difference spectroscopy, of (5-methylthienyl)acryloyl adducts with papain, cathepsins B and L, and two oxyanion hole mutants of cathepsin B (Q23S and Q23A) show that extensive polarization throughout the π-electron chain occurs for the bound acyl group in the active sites. A similar result is obtained using the specific chromophoric substrate ethyl 2-[(*N*-acetyl-L-phenylalanyl)amino]-3-(5-methylthienyl)acrylate. By using $^{13}C{=}O$ substitution it is possible to detect the acyl C═O stretching frequency, $\nu_{C=O}$, for each acyl enzyme. A correlation between $\nu_{C=O}$ and log k_3, where k_3 is the deacylation rate constant, is found where $\nu_{C=O}$ increases with increasing reactivity. This is exactly the opposite sense to the relationship found for a series of acyl serine proteases [Carey & Tonge (1995) *Acc. Chem. Res. 28*, 8]. The opposite trend in the direction of the correlation for the acyl cysteine proteases is ascribed to the strong electron polarizing forces in the active site, due principally to the active-site α-helix dipole, giving rise to canonical (valence bond) forms of the acyl group which change the hybridization about the C═O carbon atom. A correlation is also observed between the absorption maximum, λ_{max}, and log k_3 for each acyl cysteine protease. As the deacylation rate increases, 214-fold across the series, λ_{max} red-shifts from 367 to 384 nm. It is proposed that differential interactions between the active site's α-helix dipole and the acyl chromophore give rise to the observed changes in λ_{max}, with the red shift being caused principally by interactions with the excited electronic state, which has a high degree of charge separation, and the dipole. Similar interactions between the dipole and the negatively charged tetrahedral intermediate, which resembles the transition state, are proposed as the source of differential rates in deacylation. It is interesting to note that similar energies are operating in both cases. A shift in λ_{max} from 367 to 384 nm corresponds to a change in electronic absorption transition energies of 3.2 kcal/mol and a change of deacylation rate constants of 214-fold also corresponds to a change of activation energies of 3.2 kcal/mol.

α-Helical segments in proteins possess a substantial dipole moment and it has been proposed that the α-helix dipole has important structural and catalytic roles, although the latter has not been experimentally characterized (Wada, 1976; Hol *et al*., 1978). Independently, Nicholson *et al*. (1988) and Sali *et al*. (1988) demonstrated that interactions between α-helix dipoles and charged side chains can increase protein stability. For example, in the small ribonuclease barnase the interaction between the protonated form of His-18 and the α-helix stabilizes the native structure by 2.1 kcal mol^{-1} at neutral pH (Sali *et al*., 1988). Aqvist *et al*. (1991) have pointed out the importance of using a correct value for the effective dielectric constant when calculating the interactions of an α-helix dipole, and these authors are in agreement with the findings of Nicholson *et al*. (1991) that the α-helix dipole is best considered not as a giant macrodipole but rather as a manifestation of the concentration of local unsatisfied charges on the amides and carbonyls within the first and last turns of the helix. These findings are also in accord with the Hartree−Fock self-consistent field (SCF) calculations on the active site of papain by Rullmann *et al*. (1989). They found that the protein matrix stabilizes the zwitterionic form of the active-site Cys(25)−His(159) pair, with the most important contribution to the stabilization coming from the α-helix to which Cys-25 is attached and more than half of this effect coming from the backbone atoms of Cys-25 itself. Although the structural and mechanistic importance of α-helix dipoles has been established by studies of the kind outlined above, there appears to be no experimental evidence relating changes in enzymatic reactivity to the effects of an α-helix dipole. The present work sets out to take the first steps in this direction.

Our original interest in establishing structure−reactivity relationships for acyl cysteine proteases stems from work with acyl serine proteases (Tonge & Carey, 1992, 1990, 1989a). Based on the resonance Raman spectra of a series of acyl chymotrypsins and acyl subtilisins, a strong correlation (see in Figure 5) was found between the carbonyl stretching frequency, $\nu_{C=O}$, and log k_3, where k_3 is the deacylation rate constant (Tonge & Carey, 1992; Carey & Tonge, 1995). Overall, it was possible to develop relationships between the rate, the carbonyl bond length, and the

† This work was supported by the Natural Sciences and Engineering Research Council of Canada and by NIH Grant GM-54072-01.

* To whom correspondence should be addressed.

‡ University of Ottawa.

§ Case Western Reserve University.

⊗ Abstract published in *Advance ACS Abstracts,* September 1, 1996.

S0006-2960(96)00649-6 CCC: $12.00

hydrogen bonds in the active site between the C═O group and the oxyanion hole (Tonge & Carey, 1992). There are several known differences between the properties of acyl serine and acyl cysteine proteases, including the differences in chemistry between esters and thiolesters (Grunwell *et al*., 1977; Smolders *et al*., 1988), the differences in efficacy of the oxyanion holes in the two classes of proteases (Ménard *et al*., 1995), and the existence of strong electrostatic polarizing forces in the active site of members of the papain superfamily which appear to be absent in the serine proteases (Carey *et al*., 1978). In order to investigate how these differences might affect spectra−structure−reactivity relationships for the papain superfamily, we have combined the kinetic characterization of a series of acyl cysteine proteases, presented in the accompanying paper (Doran *et al*., 1996), with a Raman and electronic absorption spectroscopic analysis.

The enzymes studied are papain, cathepsins B and L, and oxyanion hole mutants of the cathepsins. Initially, we planned to undertake resonance Raman studies of these enzymes acylated with the small chromophoric group (5-methylthienyl)acrylate (5MTA) (MacClement *et al*., 1981) and specific substrate analogs labeled with this group (Doran *et al*., 1995). However, due to the highly photolabile nature of the thiolester intermediates we could not obtain reliable resonance Raman data. Fortunately, this drawback could be overcome by obtaining high-quality Raman difference spectra under nonresonance conditions (Callender & Deng, 1994; Kim *et al*., 1993; Kim & Carey, 1993). Difference spectra provide the spectrum of the bound ligand after the Raman spectrum of the enzyme is subtracted from that of the acyl enzyme. Since the acyl groups are strong scatterers, even under nonresonance conditions, the difference spectrum contains contributions only from the bound acyl group. In view of the correlations observed for the serine protease series of intermediates, the results of the Raman and kinetic studies for the papain superfamily were initially unexpected and surprising. However, when taken with the data from absorption spectroscopy, they point to the importance of α-helix dipoles in cysteine protease active sites and pave the way for an understanding of the role of these dipoles in modulating enzymatic reactivity.

MATERIALS AND METHODS

Enzymes and their protein-engineered variants were cloned and purified as described in the accompanying paper (Doran *et al*., 1996). The preparation and purification of the acyl enzymes is also described therein. The syntheses of the simple imidazolyl (5-methylthienyl)acrylate (5MTAIm) substrate and its ^{13}C-labeled analogs are given in Tonge *et al*. (1991). The synthesis of the specific chromophoric substrate ethyl 2-[(*N*-acetyl-L-phenylalanyl)amino]-3-(5-methylthienyl)acrylate is given in Doran *et al*. (1995). The model compound (5-methylthienyl)acrylate ethyl thiolester was synthesized as follows. 5MTAIm (50 mg) was added to 1 mL of ethanethiol and stirred overnight at room temperature. The next day 300 mL of 50 mM KP_i, pH 7.0, were added, and the product 5MTA thiolester was extracted using CH_2Cl_2. The crude product was then HPLC-purified on a silica gel supelcosil LC-SI column using a hexane−CH_2Cl_2 gradient. The purified product was authenticated by gas chromatography−mass spectroscopy (GC-MS).

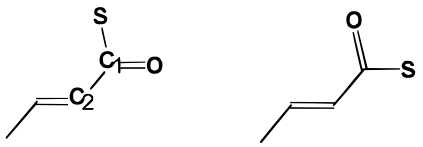
Absorption spectroscopic data were measured using a Cary 3 UV−visible spectrophotometer. FTIR data for the model compound were recorded using a FTS-40A Bio-Rad instrument, KBr windows and a path length of 0.05 mm. Raman difference spectra were recorded using the spectrometer described by Kim *et al*. (1993). Peak positions were calibrated using cyclohexanone as a standard and the program Spectracalc (Galactic Industries). Peak positions are believed to accurate to ±1 cm^{-1} for well-resolved features. Typical conditions for collecting Raman data were 488-nm laser excitation, 350 mW power, acquisition time 8 s × 40 accumulations, spectral slit c. 13 cm^{-1}, 90° irradiation−collection. The Raman spectra of the acyl enzymes were recorded at 200 μM in D_2O buffer. D_2O was used to remove contributions from water's H−O−H bending mode near 1645 cm^{-1}. The reported pD values were calculated by adding 0.4 to the observed pH meter reading (Glascoe & Long, 1960).

Molecular modeling was performed using the Insight II (95.0) and Discover (2.9.7 and 95.0/3.00) software packages from Biosym/Molecular Simulations. 5MTA was built manually and minimized independently before being bonded to Cys-25 of a relaxed crystal structure of papain. The ligand was placed initially in a position parallel to the axis of the α-helix and energy minimization was carried out within a 10-Å radius of the acyl carbonyl oxygen. Subsequently, all restraints were removed and the entire molecule was minimized till a maximum derivative of 0.1 kcal mol^{-1} Å^{-1} was achieved.

RESULTS AND DISCUSSION

Raman Spectra. The Raman data provide information on three aspects of the bound acyl group, *viz*., the conformation(s) of the 5-MTA moiety, the degree of electron polarization in the ground electronic state of the bound acyl moiety, and the chemistry of the acyl carbonyl.

(A) Conformational Heterogeneity. α,β-Unsaturated thiolesters usually exist as a mixture of *s-cis* and *s-trans* conformational isomers about the C═C−C(═O) single bond (Fausto *et al*., 1994; O'Connor *et al*., 1996).



s-trans *s-cis*

For the thiolester *S*-ethyl thiocrotonate, *ab initio* and vibrational calculations have shown that a marker band, having a contribution from the stretching motion of the same C1−C2 bond, occurs for the *trans* isomer in the 1150-cm^{-1} region while the *s-cis* isomer gives rise to a corresponding feature near 1040 cm^{-1} (Fausto *et al*., 1994). Our experience with 5-MTA bound to the viral cysteine protease HAV-3C (unpublished work, this laboratory) and the semisynthetic enzymes selenosubtilisin and thiolsubtilisin (O'Connor *et al*., 1996; unpublished work, this laboratory) indicates that these marker bands remain valid for the 5-MTA moiety. The Raman spectra of the "simple" acyl group 5-MTA bound to any of the cysteine proteases discussed in this paper suggest that the 5-MTA is present as a mixture of *s-cis* and *s-trans*

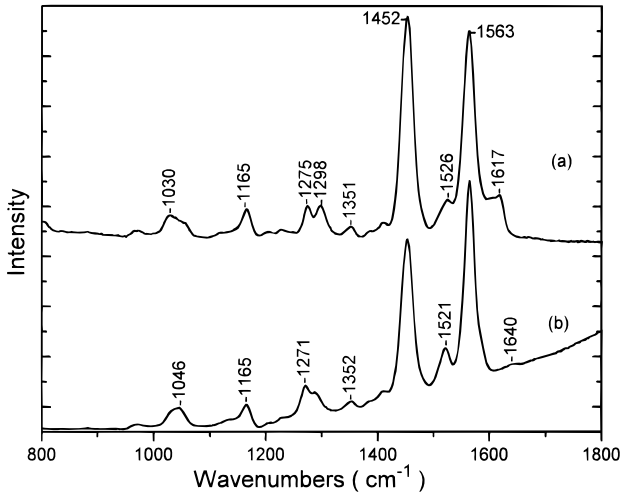


FIGURE 1: Raman difference spectra of 5MTA-cathepsin B, (a) labeled $^{13}C{=}O$ and (b) labeled $C{=}^{13}C\text{-}C{=}O$. Acyl enzyme 0.2 mM, pD 4.0; 488-nm, 300-mW excitation. The carbonyl feature in (a) at 1617 cm^{-1} shifts to 1620 cm^{-1} at active pD, i.e., pD 6 (Table 1).

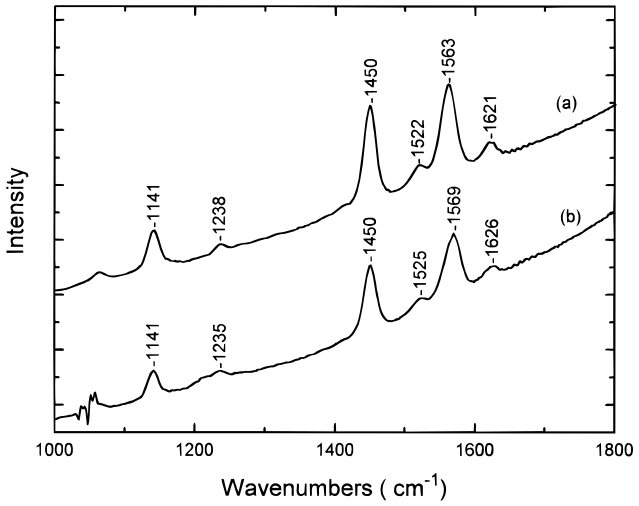


FIGURE 2: Raman difference spectra of 2-[(*N*-acetyl-L-phenylalanyl)amino]-3-(5-methylthienyl)acryloyl-cathepsin B, (a) WT and (b) Q23S. Acyl enzyme 0.2 mM, pD 4.0; 488-nm, 300-mW excitation.

conformers. Thus, in Figure 1, the medium intensity feature near 1030−1040 cm^{-1} may be associated with the *s-cis* population while the weak shoulder near 1140 cm^{-1} (seen most clearly in Figure 3) may originate from the conformer *s-trans* about the C1−C2 linkage. These assignments should be regarded as tentative since the *s-cis* and *s-trans* marker bands were developed for "simple" α,β-unsaturated thiol esters. The C(═O)−S−C fragment [or C(═O)−O−C fragment] in thiolesters, or esters, is almost invariably planar and *cis* about the C−S (or C−O) bond. However, this may not be the conformation in the corresponding fragment in the acyl enzymes. Rotation about the C−S bond is probably not expensive in terms of energy since this linkage in thiolesters has minimal double-bond character (Fausto, 1994), and it is worth noting that the X-ray crystallographic structure of indolylacryloyl-chymotrypsin reveals a nonplanar structure within the C(═O)−O−C fragment (Henderson, 1970). For the specific substrate Phe5-MTA [2-[(*N*-acetyl-L-phenylalanyl)amino]-3-(5-methylthienyl)acrylate; see Chart 1 in the accompanying paper], the conformation about the 5MTA's C1−C2 bond in the active site cannot be deduced. It is unlikely that the marker bands discussed above will be of value since the 5MTA is now carrying the phenylalanine-based substituent on C2. Additionally, close scrutiny of the Raman difference spectra for Phe5-MTA acyl enzymes does not reveal new putative candidates for marker bands.

Table 1: Deacylation, Absorption, and Raman Spectroscopic Data for 10 Acyl Cysteine Proteases[a]

acyl enzyme	k_3 deacylation ($s^{-1} \times 10^{-3}$)	λ_{max} (nm)	$\nu_{C=C}$ (cm^{-1})	$\nu^{13}{}_{C=O}$ (cm^{-1})
5MTA-papain (5)	1.1	378	1579	1619
Phe5MTA-papain (7)	3.3	384	1573	1624
5MTAcatB (3)	0.5	378	1579	1620
Phe5MTAcatB (6)	3.3	384	1576	1623
5MTAcatBQ23A (2)	0.18	374	no data	1618
5MTAcatBQ23S (1)	0.07	374	no data	1612
Phe5MTAcatBQ23S (4)	0.6	375	no data	1624
5MTAcatL (8)	7.7	381	1579	1635
5MTAcatLQ19S (9)	0.28	367	no data	no data
Phe5MTAcatL (10)	15	384	no data	no data
$5MTASC_2H_5$ (CCl_4)		338	1603	
$5MTASC_2H_5$ (CH_3CN)		340	1600	
$5MTASC_2H_5$ (H_2O)		350	1595	

[a] The spectroscopic data are recorded at "active pH", where deacylation is maximal. The values for $\nu_{C=C}$ are for the unlabeled intermediates. Also tabulated are some absorption and Raman data for the model compound, $5MTASC_2H_5$, in various solvents. The acyl enzymes are numbered in accordance with the numbers used in Figures 4−6.

(B) π-Electron Polarization. The first resonance Raman studies on substituted cinnamoyl-papains revealed the presence of strong electron polarization forces in papain's active site (Carey *et al*., 1976, 1978). The consequent π-electron reorganization in the acyl group is so large that the resonance Raman spectrum of the bound acyl group is totally distinct from that of the substrate or the product. The major changes in the Raman spectra are accompanied by perturbations to the electronic absorption spectrum, a large red shift (≈50 nm) in the acyl group's λ_{max} occurs upon binding. One advantage of carrying out Raman or resonance Raman measurements is that Raman peak positions are a property solely of the ground electronic state. Thus, unlike in absorption spectroscopic measurements, ground-state and electronic excited-state effects can be distinguished (Carey & Storer, 1984). The early resonance Raman studies focused mainly on the intense double-bond feature which could be ascribed to $\nu_{C=C}$ or a coupled mode involving $\nu_{C=C}$ (Carey *et al*., 1976, 1978). The double bond stretching feature moves to lower wavenumber with increasing polarization and there is a correlation between $\nu_{C=C}$ and λ_{max}, with a reduction in the frequency of $\nu_{C=C}$ being accompanied by a red shift (Carey & Storer, 1984; Storer *et al*., 1981). Thus, 4-(dimethylamino)-3-nitrocinnamic acid thiolester in CH_3CN has a $\nu_{C=C}$ of 1609 cm^{-1} and a λ_{max} of 363 nm, but in the strongly electron-polarizing environment of papain's active site these values change to 1570 cm^{-1} and 411 nm, respectively.

The absorption spectral data for the 10 5-MTA acyl cysteine proteases are compared to the model compound 5-MTA ethyl thiolester in Table 1. In each case, binding to the active site is accompanied by a red shift of approximately 30 nm. This large shift could be due to predominantly electronic ground-state or excited-state effects, or a mixture of both. The major changes in peak positions in the Raman spectra indicate that substantial electron rearrangement is occurring in the ground state. Figure 3 compares a typical

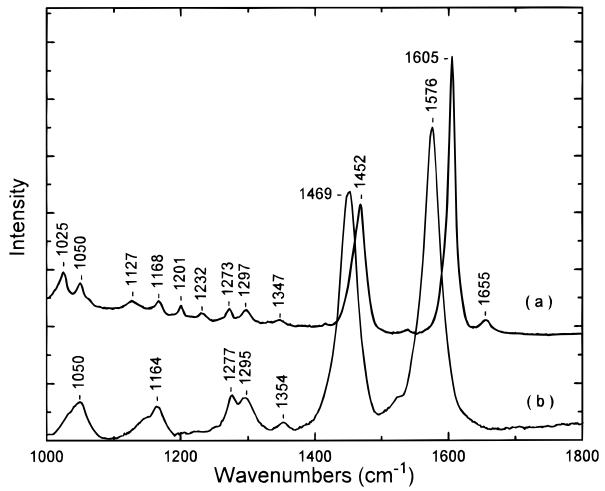


FIGURE 3: Raman difference spectra of 5-MTA ethyl thiolester and 5-MTA-papain. Thiolester 25 mM in CCl_4; 647.1-nm excitation. Conditions for the acyl enzyme spectrum were the same as in Figures 1 and 2.

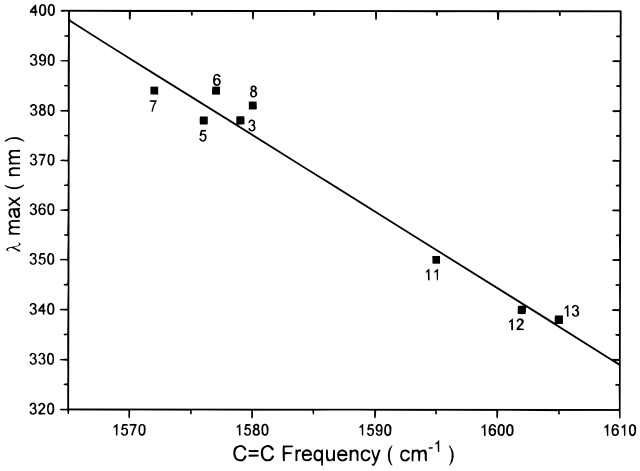


FIGURE 4: Plot of λ_{max} vs ethylenic stretching frequency, $\nu_{C=C}$, for acyl enzymes (at high or "active" pH) and ethyl esters of 5-MTA. Compound numbering: 3, 5MTAcatBwt; 5, 5MTA-papain; 6, Phe5MTAcatBwt; 7, Phe5MTA-papain; 8, 5MTAcatLwt; 11, $5MTASC_2H_5$ (in H_2O); 12, $5MTASC_2H_5$ (in CH_3CN); 13, $5MTASC_2H_5$ (in CCl_4).

acyl cysteine protease Raman spectrum, that of 5-MTA-papain, with the Raman spectrum of the model compound 5-MTA ethyl thiolester in CCl_4. Going from the model compound to papain's active site, polarization of the π electrons in the ground electronic state is evidenced by the drop in frequency of the intense C═C stretching feature from 1604 to 1573 cm^{-1}, which is accompanied by a drop of 16−18 cm^{-1} in the position of the thienyl ring modes at 1537 and 1468 cm^{-1}. The more intense bands in Figure 3 appear to be broader in the case of the acyl enzyme. This may be due to the fact that each band profile is made up of unresolved contributions from *s-cis* and *s-trans* conformers (about C1−C2) and, in the active-site differential interactions of the two forms with the polarization field, leads to increasingly large wavenumber differences for the equivalent bands from the *s-cis* and *s-trans* forms.

In Figure 4, the values of $\nu_{C=C}$ for five acyl cysteine proteases at "high" pH (where deacylation is maximal) and the three values for the model compound, 5-MTA ethyl thiolester, in different solvents are plotted against the corresponding values for λ_{max}. The general trend of a red shift in λ_{max} being accompanied by a drop in $\nu_{C=C}$ is observed. The straight line indicates that the change in the ground electronic state π-electron distribution, as probed by $\nu_{C=C}$, is accompanied by a proportional change in the difference between the ground- and excited-state energy levels as monitored by λ_{max}.

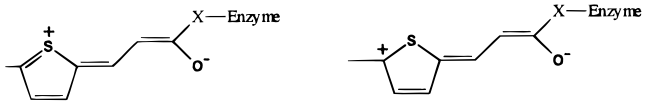
(C) Carbonyl Stretching Frequencies, $\nu_{C=O}$. Compared to the carbonyl stretch in 5-MTA esters, the Raman carbonyl stretching frequency, $\nu_{C=O}$, for 5-MTA thiolesters is approximately 7 times less intense, relative to the ethylenic stretch near 1600 cm^{-1}. This creates difficulties in observing the $\nu_{C=O}$ for the acyl cysteine proteases. However, as seen in Figure 1, isotopic labeling of the ethylenic linkage, as C═^{13}C−C(═O), shifts $\nu_{C=C}$ to lower wavenumber and makes $\nu_{C=O}$ somewhat easier to detect (it is often a poorly defined shoulder for the unlabeled complex). It still remains a very weak feature and cannot be seen for the majority of intermediates. In order to detect $\nu_{C=O}$ for each intermediate we had to label the carbonyl group with ^{13}C, whereupon $\nu_{C=O}$ could be seen in the Raman difference spectrum of every acyl enzyme—including intermediates formed from both 5-MTA itself (Figure 1) and from specific peptide substrates labeled with 5-MTA (Figure 2). Labeling the carbonyl introduces the complication of increased vibrational coupling to the ethylenic stretch and $\nu_{C=C}$ downshifts by about 10 cm^{-1} in the ^{13}C═O-labeled complexes. In a sense, the increased relative intensity of the $\nu_{C=O}$ may be due to intensity being borrowed from the ethylenic stretch. Carbonyl stretching frequencies are intrinsically lower for thiolesters compared to esters, due to the predominance of σ-orbital effects in weakening the C═O bond in thiolesters (Fausto, 1994). Thus, for α,β-unsaturated compounds a higher degree of vibrational coupling is expected for thiolesters than for esters; for thiolesters the coupling will be increased further upon ^{13}C═O substitution. The measured values for $\nu_{C=O}$ for each of eight acyl cysteine proteases are given in Table 1.

The intermediate 5-MTA-cathepsin L is unique in that, for reasons we do not fully understand, the carbonyl is more intense and it is possible to detect $\nu_{C=O}$ in the absence of isotopic labeling. It occurs at 1643 cm^{-1} at pD 6.0 (spectrum not shown) and enables us to gauge the strength of hydrogen bonding in the active site. For the model compound 5-MTA ethyl thiolester in CCl_4 solution, $\nu_{C=O}$ occurs in the normal Raman spectrum at 1655 cm^{-1} (Figure 3). The 12-cm^{-1} shift upon going from a totally non-hydrogen-bonding environment to the active site represents the formation of only weak hydrogen bond(s) to the C═O oxygen. Quantitative data on the relationship between the $\nu_{C=O}$ shift and the enthalpy of hydrogen bonding are not yet available for α,β-unsaturated thiolesters, but for α,β-unsaturated esters a shift in $\nu_{C=O}$ of 12 cm^{-1} corresponds to a change in hydrogen-bonding strength of approximately 10 kJ mol^{-1} (Dulce *et al*., 1991). On the basis of literature comparisons of simple esters and thiolesters (Grunwell *et al*., 1977; Smolders *et al*., 1988) we expect a similar value for the strength of the hydrogen bond(s) to the acyl C═O in the active-site thiolester. Undoubtedly, the carbonyl oxygen is hydrogen-bonded only weakly.

Correlation between $\nu_{C=O}$ and Reactivity (Log k_3). For acyl serine proteases, which included chymotrypsin, subtilisin, and oxyanion mutants of subtilisin BPN′, we were able to develop a striking correlation between $\nu_{C=O}$ and log k_3, over a range of 17 000 in k_3 (Tonge & Carey, 1992). This

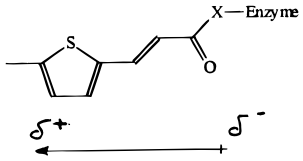
relationship is shown in Figure 5. As the acyl enzymes become more reactive, $\nu_{C=O}$ is found at lower values. The data shown in Figure 5 could also be used to evaluate quantitatively the changes in C═O bond length and hydrogen-bonding strengths to the C═O in the active site. The more reactive intermediates have a more polarized C═O group, with a longer C═O bond and stronger active site-to-C═O hydrogen bonds. Also plotted in Figure 5 are the data derived from the present work for eight 5-MTA cysteine proteases. Surprisingly, in Figure 5, exactly the opposite behavior to that for the serine proteases is found. As reactivity increases over the 214-fold range seen in Figure 5, there is an *increase* in $\nu_{C=O}$.

The explanation for the extraordinary reversal of slope seen in Figure 5 for the cysteine protease intermediates lies in ideas that were developed some time ago to explain the changes in the electronic absorption and resonance Raman spectra of furylacryloyl- and thienylacryloyl-chymotrypsins upon raising pH and activating the charge relay system (Phelps *et al*., 1981). As the pH is raised, these acyl serine proteases undergo a blue shift in λ_{max} and in the RR spectrum a modest upshift in $\nu_{C=C}$, both of which indicate a minor reduction in the polarization in the π-electron system. For example, for 5-MTA chymotrypsin, upon raising pH, λ_{max} blue-shifts by 7 nm and $\nu_{C=C}$ upshifts by 2.5 cm^{-1}. These changes can be explained in valence bond terms. Resonance (canonical) forms I and II, where X is oxygen, contribute to the true structure (Phelps *et al*., 1981)



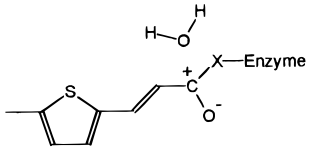
I II

and give rise to a small permanent dipole for the 5-MTA group in the active site (III):



III

If prior to hydrolysis the carbonyl is activated by, e.g., a water molecule, resonance forms of type IV become important.



IV

In these, the positive charge is localized on the carbonyl carbon, and forms I and II above, giving rise to the extended dipole, are reduced in importance. Thus, interactions of type IV associated with carbonyl activation give rise to a blue shift in λ_{max} and an increase in $\nu_{C=C}$. In these earlier studies (Phelps *et al*., 1981), the C═O stretching frequency was not observed clearly and thus carbonyl features were not

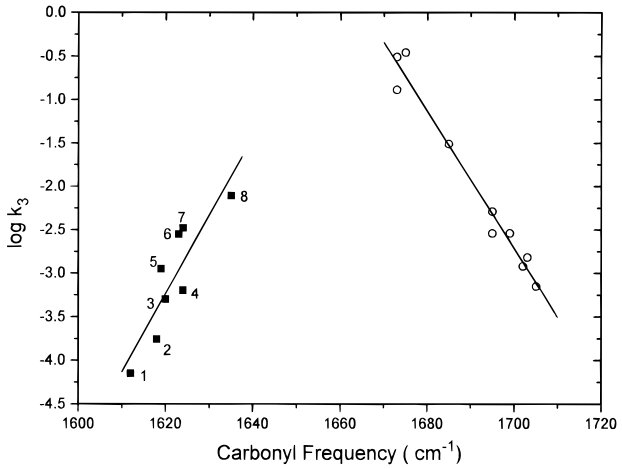
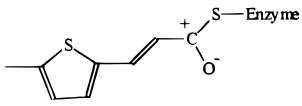


FIGURE 5: Plots of log k_3 vs carbonyl stretching frequency for acyl serine proteases (open circles) and acyl cysteine proteases (solid squares). The values for the latter are for the ^{13}C═O substituted isotopomers. Compound numbering: 1, 5MTAcatBQ23S; 2, 5MTAcatBQ23A; 3, 5MTAcatBwt; 4, Phe5MTAcatBQ23S; 5, 5MTA-papain; 6, Phe5MTAcatBwt; 7, Phe5MTA-papain; 8, 5MTA-catLwt.

discussed. However, the ideas just summarized are easily extended to take into consideration the C═O moiety and to the present data for acyl cysteine proteases.

As discussed above, the values of $\nu_{C=C}$ and λ_{max} for each acyl enzyme (Table 1 and Figure 4) demonstrate that the π-electron system for the 5-MTA group is highly polarized. That means that canonical forms of the kind I and II, where X is sulfur, make a major contribution to the true structure. This will perturb $\nu_{C=C}$ and $\nu_{C=O}$ strongly and have the effect of increasing vibrational coupling between the two groups. However, the reactivity of the carbonyl group will not be increased because that depends on canonical forms of the kind V which may be suppressed in the highly polarized 5-MTA bound in the active sites.



V

This is in contrast to the acyl serine proteases, where there is little polarization throughout the acyl group but where increasing hydrogen bonding at the C═O group does give rise to higher intrinsic reactivity by increasing the importance of forms equivalent to V (Carey & Tonge, 1995).

α-Helix Dipole-Induced Electron Polarization. There are three possible causes giving rise to the observed polarization, which lead also to changes in $\nu_{C=O}$. These are hydrogen bonding between the oxyanion hole and the C═O group, the presence of the positively charged imidazole side chain (from His-159 in papain) in the vicinity of the acyl C═O and which is present below pH 4.9 (the p*K* for deacylation), and the presence of the α-helix dipole due to the α-helix which terminates at Cys-25 in papain (Kamphius *et al*., 1984). The fact that strong polarization exists for 5-MTA-cathepsin L, where there is weak hydrogen bonding, suggests that hydrogen bonding at the acyl carbonyl is not a major contributor to polarization. This is confirmed by comparing the absorption maximum of the 5-MTA thiolester in water, 350 nm, with those for the acyl enzymes, which occur near

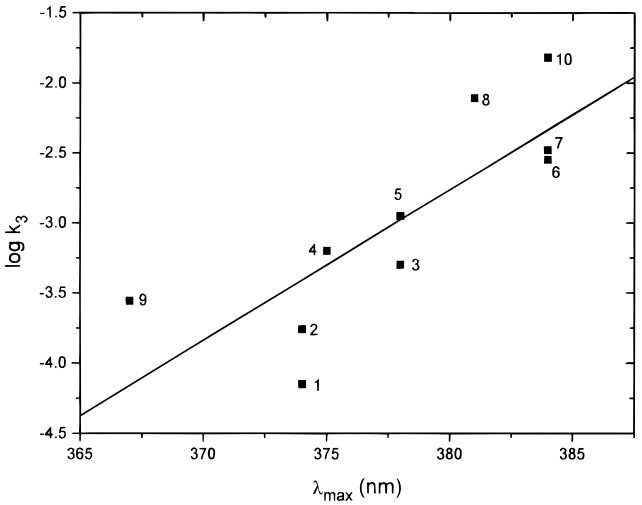


FIGURE 6: Plot of log k_3 vs λ_{max} for 5MTA cysteine proteases. Compound numbering: 1, 5MTAcatBQ23S; 2, 5MTAcatBQ23A; 3, 5MTAcatBwt; 4, Phe5MTAcatBQ23S; 5, 5MTA-papain; 6, Phe5MTAcatBwt; 7, Phe5MTA-papain; 8, 5MTAcatLwt; 9, 5MTAcatLQ19S; 10, Phe5MTAcatLwt.

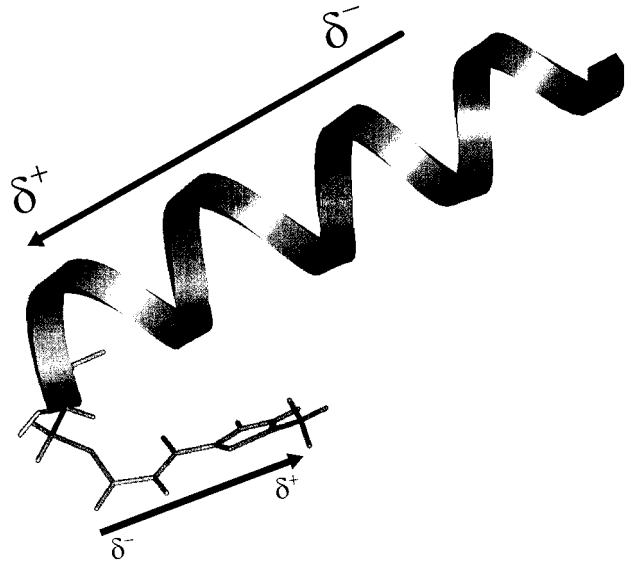


FIGURE 7: Representation of the acyl group 5MTA bound to cysteine-25 at the end of the Gly-43−Cys-25 α-helix (only the portion of the helix from Ile-40 to Cys-25 is shown).

380 nm (Table 1). For the thiolester, upon going from a non-hydrogen-bonding solvent such as CCl_4 to water, where we expect on average two medium-strength hydrogen bonds from water molecules to the acyl C═O, there is a shift in λ_{max} from 338 to 350 nm (Table 1 and Figure 4). Thus, it is likely that hydrogen bonding at the C═O cannot account for the full red shift of the acyl enzymes' λ_{max} values to 380 nm. Further confirmation comes from the consideration of the oxyanion hole mutants. For the intermediates involving 5-MTA-cathepsin B, removal of a putative oxyanion hole hydrogen-bonding donor (in the 5MTAcatBQ23A and 5MTAcatBQ23S forms) blue-shifts λ_{max} by only 4 nm (Table 1 and Figure 6) and $\nu_{C=C}$ increases by 2−4 cm^{-1} (the Raman data are for the ^{13}C═O-substituted acyl enzymes and are not shown). This strongly suggests the hydrogen bonding at the oxyanion hole is not a major contributor to polarization of the 5MTA group. Similarly, the modest changes seen in $\nu_{C=C}$ and λ_{max} upon taking the pH above 5 where the charge effects of the protonated imidazole are removed shows that the protonated imidazole is not a major contributor. Raising pH so that the intermediate side chain of His-159 is in its neutral state causes an approximately 5-nm blue shift in λ_{max} and a 3-cm^{-1} rise in $\nu_{C=C}$, as well as a small increase in the frequencies of other bands (data not shown). From these considerations we can see that hydrogen bonding at the acyl carbonyl and the electric field due to His 159's positively charged side chain make a small contribution to the observed polarization of the 5-MTA π-electrons. The major factor giving rise to polarization appears to be the α-helix dipole. The helix dipole has long been put forward as a key part of the mechanistic machinery of active sites where it is found (Hol *et al*., 1978; Wada, 1976), but experimental evidence for its effects has been lacking.

Small modulations of the α-helix dipole 5-MTA interaction, which we shall explore below, will result in variations in the contributions of forms I and II to the overall structure. Since the hybridization at the C═O carbon must be altered by the overall polarization due to the α-helix dipole, any perturbations to this effect on $\nu_{C=O}$, due to a change in the dipole−chromophore interaction, are difficult to predict on the basis of a simple model.

In Figure 7, we show a representation of the 5-MTA acyl group bound to papain's cysteine-25. The only part of the active shown is a portion of the α-helix, residues Ile-40−Cys-25, pointing toward the active site. The orientation and conformation of the 5-MTA group and the cysteine linkages were obtained via energy minimization. Due to the open cleft-like nature of papain's active site, there may be other conformers with similar energies to that shown in Figure 7. However, the energy minimization demonstrates clearly that the axis of the 5-MTA group (along the ring methyl group to the C═O carbon) can lie along that of the α-helix, such that the direction of the α-helix dipole and the dipole induced in the 5-MTA group are antiparallel. Confirmation of the molecular details of the helix-induced electron polarization will require extensive X-ray structural and/or modeling analysis. For the latter, papain may not be the best system since papain's open groove-like active site offers too many possible orientations to acyl groups such as 5-MTA. However, the semisynthetic enzymes thiolsubtilisin (Philipp & Bender, 1983) or selenosubtilisin (Syed *et al*., 1993) may provide better opportunities since their active sites appear to restrict 5-MTA's conformational space. For both these systems, the 5-MTA acyl group binds in the active site in a single *s-trans* conformer about C1−C2 (O'Connor *et al*., 1995; unpublished work, this laboratory). Moreover, no electron polarization is detected in the bound 5-MTA group. Since the active-site serine (which is changed to cysteine or selenocysteine in the semisynthetic homologs) is at the end of an α-helix in subtilisin (Abrahmsen *et al*., 1991), the lack of polarization demonstrates that this effect must be dependent on acyl group conformation or orientation. For thiolsubtilisin, however, it was shown that a polarized conformer could be generated photochemically, and that this conformer exists along with the original nonpolarized conformer (Tonge & Carey, 1989b). Thus, acyl thiolsubtilisins and selenosubtilisins, perhaps examined by a combined spectroscopic and crystallographic strategy, hold promise for delineating the stereochemical determinants of substrate electron polarization by an α-helix electric field.

Correlation between λ_{max} and Log k_3. For the acyl serine proteases there is little or no variation in λ_{max} as reactivity

across the series varies 17 000 fold (Tonge & Carey, 1992). However, across the cysteine protease series, with its 200-fold reactivity change, λ_{max} varies by 17 nm, and as Figure 6 shows, there is a correlation between λ_{max} and log k_3 (correlation coefficient 0.76), with the most red-shifted intermediates having the fastest deacylation rates. The implications of this relationship are important. The electronic absorption maximum represents a $\pi \rightarrow \pi^*$ transition between the ground and first excited states, and a red shift in λ_{max} represents a narrowing of this energy gap. In organic-based chromophores such as 5-MTA the wave function for the excited electronic state contains a major contribution from forms, such as I and II above, where there is charge separation. That is, the excited state has a larger dipole moment than the ground state. That is why there is a red shift for 5-MTA ethyl thiolester upon going from a solvent with a low dielectric constant to a solvent with a high one (Suzuki, 1967; Barltrop & Coyle, 1975). Polarization does increase also in the ground state as evidenced by the drop in $\nu_{C=C}$ (Figure 4), but the fact that a red shift is observed indicates that interactions in the excited state predominate by lowering the energy of that state to a greater extent than that of the ground state. Thus, the major part of the progressive red shift seen in Figure 6 is due to increasingly favorable stabilization of the excited electronic state and the prime candidate for bringing this about is the active-site α-helix. The changes in the red shift undoubtedly involve minor variations in the α-helix−chromophore alignment and interactions. Thus, the data in Figure 6 can be interpreted as the most reactive acyl enzymes having the most effective α-helix dipole−chromophore interactions in terms of stabilizing the dipolar nature of the excited electronic state.

The connection to reactivity is made by remembering that the hydrolysis of the acyl enzyme likely goes through a transition state which resembles a tetrahedral intermediate (Ménard *et al*., 1995). In the latter species the acyl C═O is converted to a negatively charged C−O$^-$ group. Thus, we propose that the α-helix dipole−acyl group interactions that stabilize the charge separation in the excited electronic state and bring about a red shift also stabilize negative charge buildup in the transition state and bring about enhanced reactivity. Of course, the enzyme reaction proceeds solely through the electronic ground state, but the commonality involves the ability of the α-helix dipole to stabilize charge buildup, be it in the transition state on the reaction pathway or in the excited electronic state which determines electronic spectral properties.

Toward Quantitating the Catalytic Effect of an α-Helix Dipole. The deacylation rate constants for the present series of acyl cysteine proteases vary by a factor of 200. On the basis of a simple Arrhenius equation calculation, this means that the activation energies $\Delta\Delta E^{catalysis}$ vary by 3.2 kcal mol^{-1} throughout the series. In the previous section we postulated that the interaction between the α-helix dipole and the negative charge of the tetrahedral intermediate C−O$^-$ oxygen helps stabilize the charge buildup and thus contributes to rate acceleration. We postulated further that there are differences in the helix−acyl group interactions throughout the series, due to subtle changes in chromophore−helix alignment, that lead to the observed changes in λ_{max}. The origin of these changes lies in the differential contibutions of canonical structures such as I and II seen above, with the more dipolar, more polarized acyl groups having the largest red shifts. It is of interest to note that throughout the series of acyl cysteine proteases λ_{max} varies from 367 to 384 nm, and if we calculate the differences in the electronic energy level transitions for these extremes λ_{max} using Planck's relationship we find that $\Delta\Delta E^{electronic}$ is 3.2 kcal mol^{-1} across the series. The identity of the values for $\Delta\Delta E^{catalysis}$ and $\Delta\Delta E^{electronic}$ must be fortuitous. However, it may be an indication that our underlying postulates have some validity and, crucially, that the observed variation in rate of deacylation is caused by differential helix−acyl group interactions. If this is the case, then we have the first experimental outline of the effects of α-helices on catalytic rate.

In the present work, we have used the chromophoric acyl group 5-MTA for a variety of historical reasons. For this type of chromophore, interactions with enzyme dipoles give rise to a complementary dipole in the acyl group's π-electron system, which we term "polarization", and which extends along the long axis of the acyl group (see Figure 7 and canonical forms I and II). For a natural substrate, which lacks an extended π-electron chain, the polarization will be limited to the C═O group itself, and we predict that a serine protease-type relationship (as seen in Figure 5) between $\nu_{C=O}$ and log k_3 will assert itself. In other words, the reverse slope relationship seen in Figure 5 is a consequence of our choice of substrate. Experiments are planned to confirm or refute this hypothesis.

ACKNOWLEDGMENT

We are grateful to Mr. Deendayal Dinakarpandian and Dr. Frank Soennichsen for assistance with molecular modeling.

REFERENCES

Abrahmsen, L., Tom, J., Burnier, J., Butcher, K. A., Kossiakoff, A., & Wells, J. A. (1991) *Biochemistry 30*, 4151.
Aqvist, J., Luecke, H., Quiocho, F. A., & Warshel, A. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2026.
Barltrop, J. A., & Coyle, J. D. (1975) *Excited States in Organic Chemistry*, Wiley, Chichester, U.K.
Callender, R., & Deng, H. (1994) *Annu. Rev. Biophys. Biomol. Struct. 23*, 215.
Carey, P. R., & Storer, A. C. (1984) *Annu. Rev. Biophys. Bioeng. 13*, 25.
Carey, P. R., & Tonge, P. J. (1995) *Acc. Chem. Res. 28*, 8.
Carey, P. R., Carriere, R. G., Lynn, K. R., & Schneider, H. (1976) *Biochemistry 15*, 2387.
Carey, P. R., Carriere, R. G., Phelps, D. J., & Schneider, H. (1987) *Biochemistry 17*, 1081.
Doran, J. D., Tonge, P. J., Carey, P. R., & Arya, P. (1995) *Bioorg. Med. Chem. Lett. 5*, 2381.
Doran, J. D., Mort, J. S., Tonge, P. J., & Carey, P. R. (1996) *Biochemistry 35*, 12487−12494.
Dulce, G., Faria, M., Teixeira-Dias, J. J. C., & Fausto, R. (1991) *J. Mol. Struct. 263*, 87.
Fausto, R. (1994) *J. Mol. Struct. 315*, 123.
Fausto, R., Tonge, P. J., & Carey, P. R. (1994) *J. Chem. Soc., Faraday Trans. 90*, 3491.
Glascoe, P. K., & Long, F. A. (1960) *J. Phys. Chem. 64*, 188.
Grunwell, J. R., Foerst, D. L., Kaplan, F., & Siddigui, J. (1977) *Tetrahedron 33*, 2781.
Henderson, R. (1970) *J. Mol. Biol. 54*, 341.
Hol, W. G. J., van Duijen, P. T., & Berendsen, H. J. C. (1978) *Nature 273*, 443.
Kamphius, I. G., Kalk, K. H., Swarte, M. B. A., & Drenth, J. (1984) *J. Mol. Biol. 179*, 233.
Kim, M., & Carey, P. R. (1993) *J. Am. Chem. Soc. 115*, 7015.
Kim, M., Owen, H., & Carey, P. R. (1993) *Appl. Spectrosc. 47*, 1780.

MacClement, B. A., Carriere, R. G., Phelps, D. J., & Carey, P. R. (1981) *Biochemistry 20*, 3438.
Ménard, R., Plouffe, C., Laflamme, P., Vernet, T., Tessier, D. C., Thomas, D. Y., & Storer, A. C. (1995) *Biochemistry 34*, 464.
Nicholson, H., Becktel, W. J., & Matthews, B. W. (1988) *Nature 336*, 651.
Nicholson, H., Anderson, D. E., Dao-pin, S., & Matthews, B. W. (1991) *Biochemistry 30*, 9816.
O'Connor, M. J., Dunlap, R. B., Odom, J. D., Hilvert, D., Pusztai-Carey, M., Shenoy, B., & Carey, P. R. (1966) *J. Am. Chem. Soc. 118*, 239.
Phelps, D. J., Schneider, H., & Carey, P. R. (1981) *Biochemistry 20*, 3447.
Philip, M., & Bender, M. L. (1983) *Mol. Cell Biochem. 51*, 5.
Rullmann, J. A. C., Bellido, M. N., & van Duijnen, P. T. (1989) *J. Mol. Biol. 206*, 101.
Sali, D., Bycroft, M., & Fersht, A. R. (1988) *Nature 335*, 740.
Smolders, A., Maes, G., & Zeegers-Huyskens, T. (1988) *J. Mol. Struct. 172*, 23.
Storer, A. C., Phelps, D. J., & Carey, P. R. (1981) *Biochemistry 20*, 3454.
Suzuki, H. (1967) *Electronic Absorption Spectra and Geometry of Organic Molecules*, Chapter 6, Academic Press, New York.
Syed, R., Wu, Z.-P., Hogle, J. M., & Hilvert, D. (1993) *Biochemistry 32*, 6157.
Tonge, P. J., & Carey, P. R. (1989a) *Biochemistry 28*, 6701.
Tonge, P. J., & Carey, P. R. (1989b) *J. Mol. Liq. 42*, 195.
Tonge, P. J., & Carey, P. R. (1990) *Biochemistry 29*, 10723.
Tonge, P. J., & Carey, P. R. (1992) *Biochemistry 31*, 9122.
Tonge, P. J., Pusztai, M., White, A. J., Wharton, C. W., & Carey, P. R. (1991) *Biochemistry 30*, 4790.
Tonge, P. J., Anderson, V. E., Fausto, R., Kim, M., Pusztai-Carey, M., & Carey, P. R. (1995) *Biospectroscopy 1*, 387.
Wada, A. (1976) *Adv. Biophys. 9*, 1.

BI960649+