Investigation of the applicability of *cis*-urocanic acid as a model for the catalytic Asp–His dyad in the active site of serine proteases based on ^{1}H NMR hydrogen bonding studies and spectroscopic pK_a measurements

Helmi Neuvonen * and Kari Neuvonen

Department of Chemistry, University of Turku, FIN-20014 Turku, Finland

The intramolecular hydrogen bond of *cis*-urocanic acid, (*Z*)-3-(1*H*-imidazol-4-yl)prop-2-enoic acid (*cis*-UCA) in DMSO has been evaluated by the ^{1}H NMR chemical shift approach to assess the suitability of *cis*-UCA as a model for the Asp–His dyad in the active site of serine proteases. Both the acidic OH proton and the C2 proton of the imidazole ring of *cis*-UCA resonate at exceptionally low field, at *ca.* 17.4 and 8.2 ppm, respectively, indicating a strong N–H–O type intramolecular hydrogen bond. The nucleophilic reactivity of *cis*-UCA toward the acyl carbon of 4-nitrophenyl chloroacetate has been compared with that of the *trans*-isomer and other imidazole derivatives with the aid of kinetic studies. The non-linear kinetic behaviour observed for the reaction of *cis*-UCA is attributed to complex formation with the reacting acyl derivative and *cis*-UCA. pK_a determinations of imidazole derivatives in DMSO have also been performed and a value of 3.2 has been observed for *cis*-UCA.

Introduction

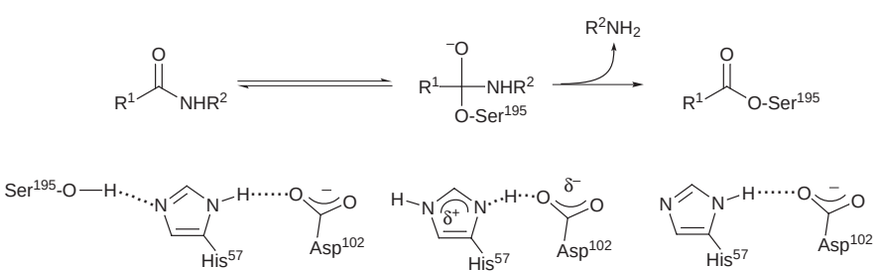
Frey *et al.*[1a] recently suggested that a low-barrier hydrogen bond (LBHB), *i.e.* a short, strong hydrogen bond, between the carboxyl group of Asp[102] and the imidazole ring of His[57] increases the reactivity of His[57] as a general base in the catalysis of the nucleophilic attack of the hydroxy group of Ser[195] of chymotrypsin on the acyl carbon (Scheme 1). Further, *cis*-urocanic acid, (*Z*)-3-(1*H*-imidazol-4-yl)prop-2-enoic acid (**1**, *cis*-UCA) was suggested as a first approximation to a model for the Asp–His part of the catalytic triad in serine proteases.[1a] In the enzymes the imidazole group of His[57] is hydrogen bonded both to the carboxyl group of Asp[102] and to the hydroxy group of Ser[195].[2] The occurrence of LBHBs in enzymes and their energetic advances have been recently discussed extensively.[1,3–9] The ΔpK_a value of about 4 between the pK_a of His[57] and that of Asp[102] in the native serine proteases[2,10–14] contradicts the suggestion of an LBHB in the Asp–His dyad in the active site of the serine proteases because the matched pK_a values are considered to be needed for the bases sharing the proton in an LBHB.[3] Further, the rise of the pK_a of His[57] during the catalytic process has been proposed.[11,15–18] However, as introduced recently by us,[19] there are some experimental data which indicate that the pK_a of Asp[102] can also increase during the formation of the tetrahedral intermediate, and this renders possible a strong mutual interaction through hydrogen bonding between Asp[102] and His[57].

The value of *cis*-UCA as a model for the Asp–His dyad of the catalytic system of serine proteases hinges on the following facts. First, although for 4-substituted imidazoles, such as histidine, the tautomeric equilibrium shown in Scheme 2 is

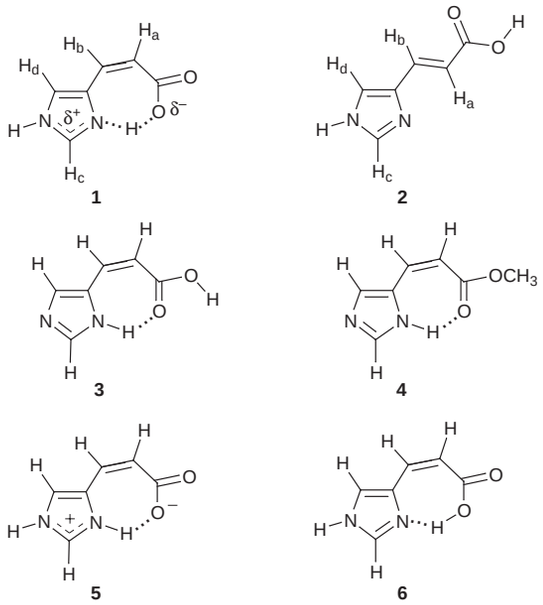
Scheme 2

usually far on the right, for the histidine residue in serine proteases the N3(H) tautomer predominates, and a similar predominance prevails for the anion of *cis*-UCA in aqueous solution.[20] Second, the aqueous pK_a values of both the carboxylic group (pK_a = 3.3) and the imidazole group (pK_a = 7.0) of *cis*-UCA[20] are very similar to those of the carboxyl group of Asp[102] and the imidazole group of His[57], respectively, in the active site of serine proteases.[2,10–14] Last, an exceptionally low-field ^{1}H NMR chemical shift (17–18 ppm) has been detected at low pH for several serine proteases in aqueous solution.[1,10,12,21] This shift has been confidently assigned to the proton engaged in the hydrogen bond between the dissociated Asp[102] and the protonated His[57].[2,10] A low-field chemical shift of 17.4 ppm was also detected by Frey *et al.*[1a] for *cis*-UCA in DMSO. The authors assigned this shift to the proton involved in the intramolecular hydrogen bond (*cf.* **1**) which they further assigned as an LBHB.

Urocanic acid is a naturally occurring deamination metabolite of histidine which accumulates in the *stratum corneum* of mammalian skin (the uppermost layer of epidermis) as the *trans*-isomer (**2**) and undergoes a *trans* to *cis* photoisomerization upon UV-irradiation.[22] Some experimental results in solution,[23] and quantum chemical calculations[24,25] as well, sug-



Scheme 1



gest a planar conformation and intramolecular hydrogen bonding for *cis*-UCA and its esters. The *ab initio* calculations performed [24b] show that in the gas phase the most stable structure of *cis*-UCA is like **3.** An analogous structure has also been depicted by Mawlawi *et al.* [26a] and Lauth-de Viguerie *et al.* [26b] Further, for *cis*-methyl urocanate the hydrogen bonded structure **4** with a relatively weak hydrogen bond has been proposed in $CHCl_3$ solution.[23] On the other hand, ^{1}H NMR T_1 relaxation time measurements support the hydrogen bonded structure **1** for *cis*-UCA in DMSO.[27a]

In continuation of the work concerning the NMR studies and *ab initio* calculations of urocanic acid and its derivatives,[24,27] we have now studied by a ^{1}H NMR approach the structure and the intramolecular hydrogen bond of *cis*-UCA in a dipolar aprotic solvent DMSO to evaluate the applicability of *cis*-UCA as a model for the Asp–His dyad of serine proteases. ^{1}H NMR measurements were performed for both *cis*- and *trans*-UCA as well as for *trans*-methyl urocanate (*trans*-MeU). Furthermore, we determined the nucleophilic reactivity of *cis*-UCA toward the ester carbonyl group of 4-nitrophenyl chloroacetate in DMSO as compared to that of *trans*-UCA and some other imidazole bases. The solvation properties of the active sites of many enzymes are assumed to be mimicked better by dipolar aprotic solvents than by water.[28] The non-aqueous environment can enhance catalytic interactions which are not significant in aqueous solution. To evaluate the expected nucleophilicities, pK_a determinations in DMSO were also performed. Before this study, the pK_a values of *cis*-UCA were not known in any organic solvent, but those determined for *trans*-UCA in DMSO–water mixtures with low water content show that the organic solvent induces appreciable changes in the pK_a values of the ionizing groups. The difference between the pK_a values of the carboxyl and imidazolium groups of *trans*-UCA in DMSO ($\Delta pK_a = 5.9$) is even larger than in water ($\Delta pK_a = 2.1$) and further, the order of the successive acid constants is reversed if compared with that in aqueous solution.[29]

Experimental

Materials

trans-Urocanic acid (Sigma, anhydrous), imidazole (Aldrich), 1-methylimidazole (Fluka AG, *purum*) and benzimidazole (Fluka AG, *puriss*) were commercial products. Imidazole was recrystallized from benzene and benzimidazole from acetone. 1-Methylimidazole was distilled under reduced pressure (107 °C, 47 mmHg). *cis*-Urocanic acid was prepared according to the method of Morrison *et al.* [30] and *trans*-methyl urocanate according to the method described by Mawlawi *et al.* [26a] 4-Nitrophenyl acetate and chloroacetate were synthesized and purified as previously described.[31] Dimethyl sulfoxide (Merck, *pro analysi*) employed in kinetic studies and [$^2H_6$]dimethyl sulfoxide (Aldrich, 99.9% D) employed in NMR measurements were used as received. 2,4-Dinitrophenol (Merck, zur Synthese) was purified by recrystallization twice from ethanol and drying in a vacuum desiccator.

NMR measurements

The ^{1}H and ^{13}C NMR spectra were recorded at 25 °C on a JEOL JNM-A500 spectrometer working at 500.16 and 125.78 MHz, respectively. The samples with varying concentrations were prepared in 5 mm od tubes in 0.6 ml of [$^2H_6$]dimethyl sulfoxide. The deuterium of the solvent was used as a lock signal. The shift values δ_H are given in ppm from TMS and J values are given in Hz. ^{13}C NMR spectra were measured with ^{1}H broad-band decoupling and NOE ^{1}H non-decoupling techniques. The other conditions were as follows: ^{1}H: spectral width 10 000 Hz, 64 K data points, digital resolution 0.15 Hz per point, pulse width 11.2 μs (90°), acquisition time 6.55 s, number of transients 64–1024, pulse delay 3 s, pulse sequence SGNON; ^{13}C: spectral width 30 030 Hz, 32 K data points (^{1}H decoupled)/64 K data points (^{1}H coupled), digital resolution 0.92 Hz per point (^{1}H decoupled)/0.46 Hz per point (^{1}H coupled), pulse width 4.35 μs (45°), acquisition time 1.09 s (^{1}H decoupled)/2.18 s (^{1}H coupled), number of transients 1000–12 000, pulse delay 3 s (^{1}H decoupled)/5 s (^{1}H coupled), pulse sequence SGBCM (^{1}H decoupled)/SGNOE (^{1}H coupled). Exponential windowing with line-broadening terms 0.02 Hz (^{1}H) and 2 Hz (^{13}C, ^{1}H decoupled)/1 Hz (^{13}C, ^{1}H coupled) were used.

Kinetics

Reaction rates for the reactions of imidazole bases with 4-nitrophenyl chloroacetate were determined by following the increase in the UV-absorption owing to the formation of 4-nitrophenol. The wavelength 310 nm at the absorbance maximum of 4-nitrophenol was used in the reactions of imidazole and benzimidazole. Due to the overlapping of the absorption of both *trans*- and *cis*-UCA and *trans*-MeU with that of 4-nitrophenol, somewhat longer wavelengths in the region 340–350 nm had to be used in their reactions. A Gilford 2600 spectrophotometer equipped with a Gilford Thermoset Temperature Controller was used. The temperature was accurate to ±0.1 °C. The reactions were carried out under pseudo-first-order conditions with the imidazole base in excess of the acyl derivative. The substrate concentration was $2 \times 10^{-5} - 2 \times 10^{-4}$ mol dm^{-3} depending on the imidazole base concentration. The rate coefficients were calculated by the method of Guggenheim.[32] The standard deviations of the individual rate coefficients were usually 0.2–0.8%. The k_{obs} values (*cf.* Fig. 1) are averages of three or more determinations.

pK_a determinations

pK_a values of the conjugate acids of the imidazole bases employed were determined in DMSO by a spectrophotometric indicator method using 2,4-dinitrophenol [pK_a(DMSO) = 5.1][33] as the indicator. Indicator solutions of fixed initial concentration [(2.5–7.5) × 10^{-5} mol dm^{-3}] were first reacted with increasing concentrations of triethylamine, until the absorbance of the reaction mixture reached its maximum value and remained constant on addition of an excess of the base. These maximum absorbance values were used to calculate the molar absorptivity of 2,4-dinitrophenolate ion at 432 nm in DMSO (ε/dm^3 mol^{-1} cm^{-1} 16 100 ± 200). By assuming the sentence of the activity coefficients to be one, eqn. (1) can be written

$$pK_a(BH^+) = pK_a(HIn) + \log \frac{[In^-][BH^+]}{[B][HIn]} \qquad (1)$$

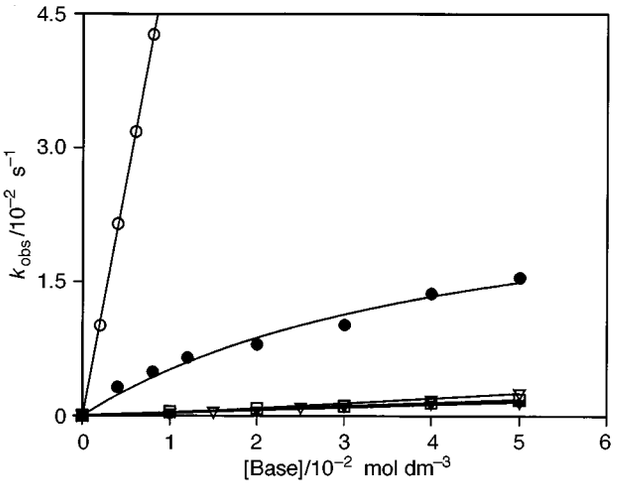


Fig. 1 A plot of k_{obs} *versus* the imidazole derivative concentration for the reactions of 4-nitrophenyl chloroacetate in DMSO at 298.2 K with imidazole (○), *cis*-UCA (●), *trans*-UCA (□), benzimidazole (▽) and *trans*- MeU (▼)

for the reaction between the indicator acid HIn and the imidazole base B investigated.

When the initial concentrations of the imidazole base and the indicator acid and also the molar absorptivity of the indicator anion are known, the equilibrium concentrations in eqn. (1), and also the $pK_a(BH^+)$ value, can be determined with the aid of absorbance measurements of the reaction solutions at 432 nm where the indicator anion is the only absorbing species. Several different (from six to nine) combinations of the initial base [(1–20) × 10^{-4} mol dm^{-3}] and indicator [(1.0–2.0) × 10^{-4} mol dm^{-3}] concentrations were used for each imidazole base. The concentrations of the indicator anion formed in the pK_a determinations were corrected by subtracting the fraction of the anion concentration due to the indicator–solvent interaction. For that purpose several aliquots of the indicator acid over a short concentration range of interest were added to the pure solvent and the concentration of the indicator anion was calculated for each aliquot by the $\varepsilon(In^-)$ value. With the aid of these results an empirical correction was obtained for the concentration of the anion form of the indicator as a function of the equilibrium concentration of the unreacted indicator.

Results and discussion

NMR indications

The ^1H NMR spectra were recorded in [$^2H_6$]dimethyl sulfoxide at several concentrations of *cis*-UCA (0.001–0.1 mol dm^{-3}), *trans*-UCA (0.005–0.1 mol dm^{-3}) and *trans*-MeU (0.01 and 0.1 mol dm^{-3}). The assignment of the different =CH protons was based, except for the consideration of the general trends of substituent effects, on NOE difference and HMQC (heteronuclear multiple quantum coherence) experiments. The assignment of the C2 protons of the imidazole rings for the hetero-correlation experiments was aided by the fact that the C2 carbons exhibit typical enhanced $^1J_{CH}$ couplings in comparison with C4 or C5.[34] The δ values of H_a, H_b, H_c and H_d protons (*cf.* **1** or **2**) were practically independent of the concentration: *cis*-UCA (H_a 5.67; H_b 6.89; $^3J_{ab}$ 12.8; H_c 8.14–8.16; H_d 7.65–7.66), *trans*-UCA (H_a 6.26–6.27; H_b 7.44; $^3J_{ab}$ 15.5–15.7; H_c 7.72–7.75; H_d 7.48–7.53), *trans*-MeU (H_a 6.33; H_b 7.53; $^3J_{ab}$ 15.6; H_c 7.74; H_d 7.57). A strong NOE was observed for the H_b proton of *cis*-UCA when irradiating the H_d proton (5.3%) suggesting the predominance of a planar conformation (structures **1**, **3**, **5** or **6**). For *trans*-UCA the shifts of H_b and H_d are too close to each other to allow reliable NOE measurements.

Low-field N–H–O and N–H shifts of *cis*-UCA

An exceptionally low-field proton signal at 17.38 ± 0.01 ppm is observed for *cis*-UCA in [$^2H_6$]dimethyl sulfoxide in the concentration range 0.001–0.03 mol dm^{-3}. A weak upfield shift is observed with the increase in concentration (Δδ = 0.07 ppm at 0.05 mol dm^{-3} and Δδ = 0.24 ppm at 0.1 mol dm^{-3}). These values agree well with that, 17.4 ppm, reported previously by Frey *et al.*[1a] in the same solvent. The other low-field signal for *cis*-UCA was detected at 13.00 ± 0.01 ppm corresponding to the shift value of 12.9 ppm reported by Frey *et al.*[1a] The value of 12.2 ppm has been observed for the ^1H NMR chemical shift of the NH protons of imidazolium ion in [$^2H_6$]dimethyl sulfoxide.[1a] For the NH proton of the neutral imidazole we detected the value of 12.00 ppm (0.1 mol dm^{-3}) in the same solvent. For *trans*-UCA the OH proton signal appears at 12.1 ppm and the NH proton in the region 12.3–12.6 ppm in [$^2H_6$]-dimethyl sulfoxide (*cf.* discussion below). The OH groups of carboxylic acids usually resonate between 9.5–13 ppm.[35a] On the other hand, hydrogen bonded protons in complexes are known to resonate in the downfield region of 13–20 ppm.[36,37] For instance, hydrogen bonding of acetic acid and trifluoroacetic acid with 1-methylimidazole in chloroform causes downfield shifts of their OH resonances from 10.64 to 13.02 ppm and from 10.28 to 15.29 ppm, respectively.[37] The low-field shift observed for *cis*-UCA (17.38 ppm) indicates a high deshielding of the proton in question and suggests that it is equally or nearly equally shared between two heteroatoms.[36,37] Accordingly, the ^1H NMR shifts at 13.00 and 17.38 ppm for *cis*-UCA are assigned for the NH proton (possibly hydrogen bonded to solvent) and the hydrogen bonded N–H–O proton, respectively, and they suggest the structure **1** instead of the structure **3**, consistent with the proposal of Frey *et al.*[1a] An alternative solution is that the values are weighted averages for a fast equilibrium between structures **1**, **5** and **6,** the structure **1**, however, predominating. A fast equilibrium between **5** and **6** could not explain the low-field proton shift around 17–18 ppm because this value is larger than that expected for the acidic proton of structures **5** or **6**.

An anisotropic effect of the nearby C=O group could also contribute to the low-field shift of the N–H–O proton of *cis*-UCA.[35b] However, a careful comparison with low-field shifts observed for serine proteases around 18 ppm at low pH[1,10,12,21] suggests the predominance of the structure **1** for *cis*-UCA. The slight concentration dependence of the low-field shift of *cis*-UCA agrees with the involvement of the proton in question in a moderately strong hydrogen bond. The strong NOE observed for the H_b proton of *cis*-UCA when irradiating the H_d proton also suggests the predominance of a planar conformation which is optimal for a strong hydrogen bond.

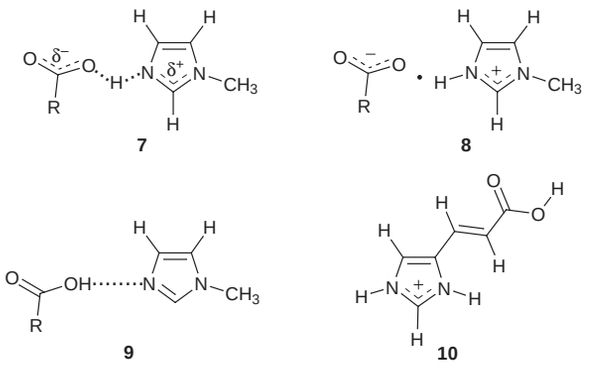
The scrutiny of the C2 proton shift

Further confirmation of the structure **1** for *cis*-UCA in DMSO is obtained by the C2 proton (H_c) ^1H NMR chemical shift, 8.15 ± 0.01 ppm. For the 1 : 1 complexes between 1-methylimidazole and a series of carboxylic acids with varying strength in $CDCl_3$, Tobin *et al.*[38] have shown the C2 proton shift of 1-methylimidazole to vary from 7.56 ppm for a neutral hydrogen bonded complex with a weak acid to 8.97 ppm for a hydrogen bonded 1-methylimidazolium ion in a complex with a strong acid. The chemical shift of the C2 proton was suggested as a probe of the degree of the positive charge of the imidazole ring.[38] In the complex with 2,2-dichloropropionic acid, the C2 proton of 1-methylimidazole resonates at 8.39 ppm and the acidic proton resonates at *ca.* 18 ppm. The coexistence of complex structures **7**, **8** and **9** has been suggested, the structure **7** predominating. The C2 proton chemical shifts of 7.54 ppm for 1-methylimidazole, 7.75 ppm for *trans*-UCA and 7.74 ppm for *trans*-MeU were recorded in the present work in [$^2H_6$]dimethyl sulfoxide (at the concentration of 0.1 mol dm^{-3}). The shift of 9.27 ppm has been previously detected for protonated *trans*-UCA (**10**) in the same solvent.[29] Accordingly, the value of 8.15 ppm found for the C2 proton of *cis*-UCA indicates a

Table 1 Second- and third-order rate coefficients (k_1 and k_2, respectively) for the reactions of 4-nitrophenyl chloroacetate with different imidazole bases in DMSO at 298.2 K and the pK_a values determined for the imidazole bases in DMSO

Base	k_1/dm^3 mol^{-1} s^{-1}	k_2/dm^6 mol^{-2} s^{-1}	p$K_a^{\text{this work}}$(DMSO)	p$K_a^{\text{lit.}}$(DMSO)	p$K_a^{\text{lit.}}$(water)
Imidazole	5.35 ± 0.05		5.1 ± 0.2	6.26;[a] 5.82[b]	7.1[c]
Benzimidazole	0.033 ± 0.001	0.40 ± 0.02	3.9 ± 0.1	4.36;[a] 4.04[b]	5.53[c]
cis-UCA	*d*		3.2 ± 0.1		7.0[e]
trans-UCA	0.035 ± 0.001		3.8 ± 0.1	3.84[b]	6.1[e]
trans-MeU	0.030 ± 0.003		3.8 ± 0.1		

[a] In DMSO, ref. 39. [b] In 91% (by weight) DMSO–water, ref. 29. [c] Ref. 40. [d] *cf.* Discussion. [e] Ref. 20.



partially protonated imidazolium ring suggesting the structure **1** or a fast equilibrium between structures **1**, **5** and **6**. A fast equilibrium between structures **5** and **6** can be discarded because of the low-field N–H–O signal observed, as discussed above.

Comparison of *cis*-UCA with *trans*-UCA and the effect of other imidazole bases or 4-nitrophenyl acetate

In the ^{1}H NMR spectra of *cis*-UCA the N–H–O and NH protons appear as two signals at 17.14–17.38 and at 13.00 ± 0.01 ppm, respectively, in the whole concentration range studied. In the ^{1}H NMR spectra of *trans*-UCA, there occur two low-field signals at 12.08 and at 12.41 ppm when the *trans*-UCA concentration is 0.1 mol dm^{-3}. However, at lower concentrations the lowest field shift is divided into two signals: *ca.* 12.3 and 12.6 ppm. Interestingly, a similar peak separation is observed in the ^{1}H NMR spectra of *trans*-MeU at concentrations of both 0.1 mol dm^{-3} and 0.01 mol dm^{-3}. This suggests that also in the spectra of *trans*-UCA the divided signal corresponds to the NH proton and suggests the coexistence of associated and non-associated species of *trans*-UCA in DMSO. The fact that we do not see two signals at high concentrations of *trans*-UCA can be explained by the increase of the exchange rate. We will study this topic further in the near future.

The ^{1}H NMR spectra of *cis*-UCA (0.05 mol dm^{-3}) were also recorded in the presence of imidazole bases with [*cis*-UCA]:[1-methylimidazole] ratios of 10:1, 5:1, 3.3:1, 2.5:1, 2:1, 1:1 and 1:2 and with [*cis*-UCA]:[imidazole] ratios of 10:1, 2.5:1 and 1:1. The N–H–O signal moved slightly upfield with increasing 1-methylimidazole concentration (*ca*. 0.5 ppm) and the N–H signal of the imidazole ring of *cis*-UCA disappeared at [*cis*-UCA]:[1-methylimidazole] ≤ 1:1, while other ^{1}H NMR signals of *cis*-UCA were practically unaffected. The ^{1}H NMR signals of 1-methylimidazole in DMSO were similar in the presence and absence of *cis*-UCA. The disappearance of the N–H signal is attributed to the increased rate of the proton exchange with the increase of the 1-methylimidazole concentration. The presence of imidazole had no more effect than that of 1-methylimidazole on the ^{1}H NMR shifts of *cis*-UCA. We can conclude that although imidazole and 1-methylimidazole are in DMSO considerably stronger bases than the imidazole group of *cis*-UCA[39] (*cf*. Table 1) they are not able essentially to affect the N–H–O hydrogen bond of *cis*-UCA. The effect of 1-methylimidazole on the ^{1}H NMR shifts of *trans*-UCA was also investigated at a *trans*-UCA concentration of 0.05 mol dm^{-3} with [*trans*-UCA]:[1-methylimidazole] ratios of 10:1, 2:1, 1:1 and 1:2. The OH signal of *trans*-UCA at 12.1 ppm disappeared in the presence of 1-methylimidazole, obviously because of the increased rate of exchange. Only one signal (at 12.3 ppm) was observed, also in the NH region. This supports the above depicted concept of the increase of the exchange rate which prevents the detection of two NH signals at high concentrations of *trans*-UCA. The behaviour of the ^{1}H NMR chemical shifts of *trans*-UCA in contrast to the constancy of the ^{1}H NMR chemical shifts of *cis*-UCA further supports the concept of a strong intramolecular hydrogen bond in the case of *cis*-UCA. We also tried to trace the kinetically exposed possibility of the association between *cis*-UCA and 4-nitrophenyl chloroacetate (*cf.* discussion below) by ^{1}H NMR measurements for solutions containing varying concentrations of *cis*-UCA and 4-nitrophenyl acetate ([*cis*-UCA]:[4-NPA] = 1:1, 1:4, 1:10 or 6:1). Because *cis*-UCA is much less reactive toward the latter ester than toward 4-nitrophenyl chloroacetate, the concentrations of the compounds stayed constant during the NMR measurement. No changes of NMR signals were, however, observed. After a prolonged reaction time, the signals of the *N*-acylated *cis*-UCA could be detected.

Kinetic results

The kinetic rate-laws. The rate coefficients were determined for the reactions between 4-nitrophenyl chloroacetate and imidazole, benzimidazole, *cis*-UCA, *trans*-UCA and *trans*-MeU. The spontaneous degradation of the ester in DMSO was not found. In the concentration range (0.002–0.01 mol dm^{-3}) of imidazole used, only the first-order reaction in amine could be detected [eqn. (2)] while the reaction with

$$k_{obs} = k_1[\text{Im}] \quad (2)$$

benzimidazole was of mixed first- and second-order in amine [eqn. (3)]. Linear dependence of k_{obs} on the concentration of

$$k_{obs} = k_1[\text{BzIm}] + k_2[\text{BzIm}]^2 \quad (3)$$

the imidazole derivative was also observed in the case of both *trans*-UCA and its methyl ester (*trans*-MeU). In the case of *cis*-UCA, the plot of k_{obs} *vs*. *cis*-UCA concentration showed a downward curvature (*cf.* Fig. 1). These saturation kinetics are discussed more below. Table 1 gives the second- and third-order rate coefficients determined. The pK_a values determined in this study for the different imidazole bases are also given.

On the reaction mechanism. The nucleophilicity of imidazole and its derivatives toward esters with good leaving groups is well known, and the nucleophilic reactivity of imidazole toward 4-nitrophenyl chloroacetate particularly has been previously observed both in aqueous solution and in a dipolar non-hydroxylic solvent, acetonitrile, as well as in their mixtures.[41–43] In aqueous solution *cis*-UCA has been shown to react as a nucleophile toward 4-nitrophenyl acetate and it fits the Brønsted plot determined by other imidazole bases.[44] The mechanism generally accepted for aminolysis of esters and acid

chlorides in non-hydroxylic solvents involves the formation and breakdown of a zwitterionic addition intermediate.[28d,41,45] For all of the amines studied in the present work, the attacking nucleophile is much less basic than the leaving group of the ester [pK_a(4-nitrophenol)$_{DMSO}$ = 10.8].[33] Therefore, a mechanism involving a fast addition of the nucleophile followed by a rate-determining breakdown of the addition intermediate and consequently a high value of the Brønsted β value can be expected. This mechanism is supported by the overall third-order reaction detected for benzimidazole. The k_2 reaction indicates a contribution of the general base catalysis by a second molecule of the amine in the breakdown of the addition intermediate.[41,46] The β value of 1.24 has been determined by Nagy *et al.*[45b] for the uncatalysed nucleophilic reaction of amines toward 4-nitrophenyl 3,5-dinitrobenzoate in acetonitrile.

On the determined pK_a values and the reactivity of imidazole derivatives. The pK_a values obtained in this study (Table 1) for benzimidazole or *trans*-UCA are quite close to those measured by potentiometry in 91% DMSO–water, but they deviate somewhat from those given in DMSO. However, with the most basic amine, imidazole, the value 5.1 differs more from those measured in 91% DMSO–water (5.82)[29] or DMSO (6.26).[39] The reproducibility in the pK_a determination when using different base or indicator concentrations was also somewhat poorer with imidazole than with other bases. The pK_a values and the rate coefficients collected in Table 1, however, show that the three bases, benzimidazole, *trans*-UCA and *trans*-MeU, possess closely similar pK_a values and also closely similar uncatalysed nucleophilic reactivity (k_1) toward 4-nitrophenyl chloroacetate. The similar behaviour of *trans*-UCA and *trans*-MeU especially shows that the side-chain group of *trans*-UCA does not operate through its COOH functionality. Steric effects obviously prevent the overall third-order reaction (k_2) in the case of *trans*-UCA and *trans*-MeU, although it is observed for benzimidazole, because this rate term can be attributed to a general base-catalysed nucleophilic reaction of the imidazole derivative.[41,46] The third-order reaction is not detected in the case of imidazole either, but this is most obviously caused by the low amine concentrations used due to the high reactivity of imidazole. Although only an estimate is obtained, the k_1 terms in Table 1 indicate a β value higher than one for the reaction first-order in amine.

Association of *cis*-UCA. The kinetic behaviour of *cis*-UCA differs from the other imidazole bases studied (Fig. 1). One explanation for the non-linear saturation kinetics observed could be the self-association of *cis*-UCA at higher concentrations. However, a linear dependence of its UV-absorption was observed at least up to the concentration of 0.05 mol dm^{-3} [λ(DMSO)/nm 350, ε/dm^3 mol^{-1} cm^{-1} 197 ± 3] indicating that any concentration dependent association does not occur. ^{1}H NMR results agree with that interpretation. On the other hand, if a weak complexation between *cis*-UCA and the reacting ester is assumed, the observed inhibition of the reaction at higher *cis*-UCA concentrations can be interpreted. In Fig. 1 the observed rate coefficients are fitted to eqn. (4), and a good harmony is

$$k_{obs} = \frac{k_{max}[cis\text{-UCA}]}{K + [cis\text{-UCA}]} \quad (4)$$

detected. In eqn. (4), K represents the dissociation constant of the assumed *cis*-UCA–4-nitrophenyl chloroacetate complex. A mechanism where the complex is non-productive is kinetically indistinguishable from that where the complex is on the reaction path. A non-productive non-covalent complexation between 4-nitrophenyl trimethylacetate and various methylimidazoles has been previously proposed in aqueous solution to cause a non-linear kinetic behaviour at high amine concentrations.[47] The values k_{max} = (0.028 ± 0.005) s^{-1} and K = (0.045 ± 0.013) mol dm^{-3} are obtained by eqn. (4) for the reaction of *cis*-UCA and 4-nitrophenyl chloroacetate. If the complex would be non-productive ($k_{max} = k_1K$), the value of 0.62 dm^3 mol^{-1} s^{-1} is obtained for the second-order rate coefficient (k_1) of the reaction between the ester and *cis*-UCA. This value is *ca*. 20 times higher than the k_1 values for *trans*-UCA or *trans*-MeU. Because of the lower pK_a of *cis*-UCA as compared with that of *trans*-UCA or *trans*-MeU and because of the detected predominance of structure **1** for *cis*-UCA, one would, however, expect the reaction between *cis*-UCA to be slower than that with *trans*-UCA or *trans*-MeU. Accordingly, the *cis*-UCA–ester complex is assumed to lie on the reaction path ($k_{max} = k_1$) and the value of 0.028 s^{-1} is obtained for the rate coefficient of the breakdown of the complex to products.

It is possible that the charged character of *cis*-UCA (*cf.* **1**) results in its weak association with the ester possessing the polarized carbonyl group. Alternatively, it can be thought that the intramolecular hydrogen bonding of *cis*-UCA may increase the delocalization of its π-electrons and the system can therefore successfully interact with the π-electrons of the aromatic ester. Any ^{1}H NMR indications of an association between *cis*-UCA and 4-nitrophenyl acetate could not, however, be detected (*cf*. discussion above). The formation of a 1 : 1 complex between the substrate ester and caffeine or theophylline-7-acetic acid was quite recently verified to depress the rate of alkaline hydrolysis of substituted phenyl benzoates in 31% acetonitrile–water.[48] Also, aminolyses of some aromatic esters carrying carboxyl groups with several different tetra-, tri- and diamines in DMSO have also shown saturation kinetics due to complex formation.[49] The future study of this kind of stacking phenomena is of high interest because of the π–π interactions known to prevail in biomolecules.

On the pK_a values of *cis*-UCA. The aqueous pK_a value of the carboxylic group of *cis*-UCA is 3.3[20] but the value in DMSO is unknown. For *trans*-UCA the pK_a(COOH) has been shown to change from 3.6 to 9.7 on going from water to 91% DMSO–water.[29] On the other hand, the intramolecular 'hydrogen bond solvation' can effectively stabilize the carboxylate anion in the case of *cis*-UCA leading to an only small change of pK_a-(COOH) for solvent change from water to DMSO. For maleic acid the change in pK_a of the first ionization step is only 2.2 for the solvent change from water to *N,N*-dimethylformamide, while for fumaric acid the respective change in pK_a is 6.2.[39,50] For *trans*-UCA the pK_a of the imidazole group changes from 6.1 to 3.8 on going from water to 91% DMSO–water.[20,29] In the present work the pK_a of the imidazole group of *cis*-UCA was determined in DMSO for the first time and the value of 3.2 was found. In aqueous solution the respective value is 7.0.[20] So, it is possible that for *cis*-UCA ΔpK_a between the imidazole group and the carboxylic group is smaller in DMSO than in aqueous solution [ΔpK_a(aq) = 3.7]. This can render possible the strengthening of the intramolecular hydrogen bond of *cis*-UCA on going from water to DMSO.[7c]

Conclusions

In previous studies[1a,24b,26] two different hydrogen bonded structures have been proposed for *cis*-UCA (**1** and **3**). The present ^{1}H NMR results, *i.e*. low-field N–H–O and C2 proton shifts and the observed NOE effect for *cis*-UCA as well as the constancy of the ^{1}H NMR chemical shifts of *cis*-UCA in contrast to the behaviour of the ^{1}H NMR chemical shifts of *trans*-UCA in the presence of 1-methylimidazole or imidazole, support the predominance of a strong intramolecular hydrogen bond (*cf*. **1**) for *cis*-UCA in DMSO. The pK_a determinations agree with this concept. Therefore, although we cannot assess the strength of this hydrogen bond, *cis*-UCA seems to be a relatively good model for the Asp–His dyad of the active site in serine protease for which the formation of an LBHB during the enzyme-catalysed reaction has been suggested. Furthermore, the reaction between aryl esters and *cis*-UCA seems to provide

an interesting new model to study the stacking phenomena usually encountered in biomolecules.

References

1 (*a*) P. A. Frey, S. A. Whitt and J. B. Tobin, *Science*, 1994, **264**, 1927; (*b*) C. D. Cassidy, J. Lin and P. A. Frey, *Biochemistry*, 1997, **36**, 4576.
2 W. W. Bachovchin and J. D. Roberts, *J. Am. Chem. Soc.*, 1978, **20**, 8041 and references therein.
3 F. Hibbert and J. Emsley, *Adv. Phys. Org. Chem.*, 1990, **26**, 255; C. L. Perrin and J. B. Thoburn, *J. Am. Chem. Soc.*, 1992, **114**, 8559; J. A. Gerlt and P. G. Gassman, *J. Am. Chem. Soc.*, 1993, **15**, 11 552; P. Gilli, V. Bertolasi, V. Ferretti and G. Gilli, *J. Am. Chem. Soc.*, 1994, **116**, 909.
4 W. W. Cleland and M. M. Kreevoy, *Science*, 1994, **264**, 1887; J. A. Gerlt, M. M. Kreevoy, W. W. Cleland and P. A. Frey, *Chem. Biol.*, 1997, **4**, 259.
5 Y. Pan and M. A. McAllister, *J. Am. Chem. Soc.*, 1997, **119**, 7561.
6 B. Schwartz and D. G. Drueckhammer, *J. Am. Chem. Soc.*, 1995, **117,** 11 902.
7 (*a*) S. Shan, S. Loh and D. Herschlag, *Science*, 1996, **272,** 97; (*b*) S. Shan and D. Herschlag, *J. Am. Chem. Soc.*, 1996, **118**, 5515; (*c*) S. Shan and D. Herschlag, *Proc. Natl. Acad. Sci. USA*, 1996, **93**, 14 474.
8 Y. Kato, L. M. Toledo and J. Rebek, Jr., *J. Am. Chem. Soc.*, 1996, **118**, 8575.
9 J. P. Guthrie, *Chem. Biol.*, 1996, **3**, 163.
10 W. W. Bachovchin, *Biochemistry,* 1986, **25**, 7751.
11 A. A. Kossiakoff and S. A. Spencer, *Biochemistry*, 1981, **20**, 6462.
12 G. Robillard and R. G. Shulman, *J. Mol. Biol.*, 1974, **86**, 519; G. Robillard and R. G. Shulman, *J. Mol. Biol.*, 1974, **86**, 541.
13 J. L. Markley and J. B. Ibanez, *Biochemistry*, 1978, **17**, 4627.
14 A. Warshel and S. Russell, *J. Am. Chem. Soc.*, 1986, **108**, 6569.
15 J. P. G. Malthouse, W. U. Primrose, N. E. Mackenzie and A. I. Scott, *Biochemistry,* 1985, **24**, 3478.
16 M. D. Finucane and J. P. G. Malthouse, *Biochem. J.*, 1992, **286**, 889.
17 T. P. O'Connell and J. P. G. Malthouse, *Biochem. J.*, 1996, **317**, 35.
18 T.-C. Liang and R. H. Abeles, *Biochemistry*, 1987, **26**, 7603.
19 H. Neuvonen, *Biochem. J.*, 1997, **322**, 351.
20 J. D. Roberts, C. Yu, C. Flanagan and T. R. Birdseye, *J. Am. Chem. Soc.*, 1982, **104**, 3945.
21 J. L. Markley, *Biochemistry*, 1978, **17**, 4648.
22 E. DeFabo and F. Noonan, *J. Exp. Med.*, 1983, **157**, 84; H. Morrison, *Photodermatology*, 1985, **2**, 158.
23 R. Quinn, J. Mercer-Smith, J. N. Burstyn and J. S. Valentine, *J. Am. Chem. Soc.*, 1984, **106**, 4136; F. D. Lewis, D. K. Howard, J. D. Oxman, J. D. Upthgrove and S. L. Quillen, *J. Am. Chem. Soc.*, 1986, **108**, 5964.
24 (*a*) A. Lahti, M. Hotokka, K. Neuvonen and P. Äyräs, *J. Mol. Struct.* (*THEOCHEM*), 1995, **331**, 169; (*b*) A. Lahti, M. Hotokka, K. Neuvonen and P. Äyräs, *Struct. Chem.*, 1997, **8**, 331.
25 A. H. de Vries and P. Th. van Duijnen, *Biophys. Chem.*, 1992, **43**, 139.
26 (*a*) H. Mawlawi, M. C. Monje, A. Lattes and M. Rivière, *J. Heterocycl. Chem.*, 1992, **29**, 1621; (*b*) N. Lauth-de Viguerie, N. Sergueeva, M. Damiot, H. Mawlawi, M. Rivière and A. Lattes, *Heterocycles*, 1994, **37**, 1561.
27 (*a*) K. Neuvonen, C.-P. Askolin, T. Kaila, P. Pasanen, J. Mattinen and J. Jansén, unpublished results; (*b*) A. Lahti, M. Hotokka, K. Neuvonen and G. Karlström, *J. Mol. Struct.* (*THEOCHEM*), in the press.
28 (*a*) G. Wallerberg, J. Boger and P. Haake, *J. Am. Chem. Soc.*, 1971, **93**, 4938; (*b*) J. W. Henderson and P. Haake, *J. Org. Chem.*, 1977, **42**, 3989; (*c*) F. M. Menger, *J. Am. Chem. Soc.*, 1966, **88**, 3081; (*d*) F. M. Menger and A. C. Vitale, *J. Am. Chem. Soc.,* 1973, **95**, 4931; (*e*) R. L. Snell, W.-K. Kwok and Y. Kim, *J. Am. Chem. Soc.*, 1967, **89**, 6728; (*f*) E. Ciuffarin, M. Isola and P. Leoni, *J. Org. Chem.*, 1981, **46**, 3064; (*g*) A. B. Maude and A. Williams, *J. Chem. Soc.*, *Perkin Trans. 2*, 1997, 179.
29 J.-C. Halle, C. Pichon and F. Terrier, *J. Biol. Chem.*, 1984, **259**, 4142.
30 H. Morrison, D. Avnir and T. Zarrella, *J. Chromatogr.*, 1980, **183**, 83.
31 H. Neuvonen, *J. Chem. Soc.*, *Perkin Trans. 2*, 1990, 669.
32 A. A. Frost and R. G. Pearson, *Kinetics and Mechanisms*, Wiley, New York, 1969, p. 27.
33 F. G. Bordwell, *Acc. Chem. Res.*, 1988, **21**, 456.
34 R. E. Wasylishen and G. Tomlinson, *Biochem. J.*, 1975, **147**, 605; M. Schumacher and H. Günther, *J. Am. Chem. Soc.*, 1982, **104**, 4167.
35 (*a*) P. Sohár, *Nuclear Magnetic Resonance Spectroscopy*, CRC Press, Boca Raton, 1983, vol. 2, ch. 3, p. 1; (*b*) P. Sohár, *Nuclear Magnetic Resonance Spectroscopy*, CRC Press, Boca Raton, 1983, vol. 1, ch. 1, p. 1.
36 Z. Dega-Szafran and M. Szafran, *Heterocycles*, 1994, **37**, 627; N. S. Golubev, S. N. Smirnov, V. A. Gindin, G. S. Denisov, H. Benedict and H.-H. Limbach, *J. Am. Chem. Soc.*, 1994, **116**, 12 055.
37 I. I. Schuster and J. D. Roberts, *J. Am. Chem. Soc.*, 1979, **44**, 3864.
38 J. B. Tobin, S. A. Whitt, C. S. Cassidy and P. A. Frey, *Biochemistry*, 1995, **34**, 6919.
39 *Acid-Base Dissociation Constants in Dipolar Aprotic Solvents*, ed. K. Izutsu, Blackwell, Oxford, 1990.
40 D. D. Perrin, *Dissociation Constants of Organic Bases in Aqueous Solution*, Butterworths, London, 1965.
41 H. Neuvonen, *J. Chem. Soc.*, *Perkin Trans. 2*, 1995, 951.
42 K. Koehler, R. Skora and E. H. Cordes, *J. Am. Chem. Soc.*, 1966, **88**, 3577; R. C. Knowlton and L. D. Byers, *J. Org. Chem.*, 1988, **53**, 3862.
43 H. Neuvonen, *Ber. Bunsenges. Phys. Chem.,* 1996, **100**, 1231.
44 S. C. Zimmerman, J. S. Korthals and D. Cramer, *Tetrahedron*, 1991, **47**, 2649.
45 (*a*) F. M. Menger and J. H. Smith, *J. Am. Chem. Soc.*, 1972, **94**, 3824; (*b*) O. B. Nagy, V. Reuliaux, N. Bertrand, A. Van der Mensbrugghe, J. Leseul and J. B. Nagy, *Bull. Soc. Chim. Belg.*, 1985, **94**, 1055; (*c*) M. Jedrzejczak, R. E. Motie, D. P. N. Satchell and W. N. Wassef, *J. Chem. Soc.*, *Perkin Trans. 2*, 1994, 1471.
46 H. Neuvonen, *J. Chem. Soc.*, *Perkin Trans. 2*, 1987, 159; H. Neuvonen, *J. Chem. Soc.*, *Perkin Trans. 2*, 1988, 2051.
47 R. A. Gagliano, Jr., R. C. Knowlton and L. D. Byers, *J. Org. Chem.*, 1989, **54**, 5247.
48 N. Pirinccioglu and A. Williams, *J. Chem. Soc.*, *Perkin Trans. 2*, 1998, 37.
49 J. Suh, Y. Kim, E. Lee and S. H. Chang, *Bioorg. Chem.*, 1986, **14**, 33.
50 *Dissociation Constants of Organic Acids in Aqueous Solution*, eds. G. Kortüm, W. Vogel and K. Andrussow, Butterworths, London, 1961.

Paper 7/08832I
Received 8*th December* 1997
Accepted 15*th May* 1998