Tetrahedron Vol. 42, No. 12, pp. 3269 to 3276, 1986
Printed in Great Britain.

0040-4020/86 $3.00 + .00
Pergamon Journals Ltd.

INHIBITION OF TRYPSIN BY CARBOBENZYLOXYLYSYL CHLOROMETHYL KETONE: ^{13}C NMR AND X-RAY DIFFRACTION ANALYSES OF THE ENZYME-INHIBITOR COMPLEX†

A. I. SCOTT,* N. E. MACKENZIE,[1] J. P. G. MALTHOUSE, W. U. PRIMROSE, P. E. FAGERNESS and A. BRISSON

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

LE ZONG QI and W. BODE

Max-Planck Institut fur Biochemie, D-8033 Martinsried, West Germany.

CARL M. CARTER and YI JIN JANG

Department of Chemistry, University of Utah, Salt Lake City, Utah 84112, U.S.A.

(*Received in UK 5 February* 1986)

ABSTRACT - Comparison of the X-ray diffraction data and both solution and solid (CPMAS) NMR spectra for the covalent adduct of trypsin and carbobenzyloxylysyl chloromethyl ketone provides unequivocal evidence that the serine-195 hydroxyl group of the enzyme forms a covalent bond with the carbonyl group of the bound inhibitor. This result provides important correlation of solution and solid-state geometry in such adducts.

Chloromethyl ketone derivatives of specific substrates are potent, irreversible inhibitors of serine proteases, alkylating N-3 of the imidazole of histidine-57.[2,3] Further reaction involves the nucleophilic addition of the side chain hydroxyl of serine-195 across the ketonic functionality.[4] In particular, inhibition of trypsin (E.C. 3.4.21.4) by [2-^{13}C]carbobenzyloxylysyl chloromethyl ketone (ZLMK) (1, Fig. 1) has been the subject of intensive ^{13}C nuclear magnetic resonance (NMR) studies,[5-7] summarized in Fig. 1.

In aqueous solution the inhibitor displays two resonances in the ^{13}C NMR spectrum (Fig. 1A) at 204.7 and 95.4 ppm which are assigned to the carbon-13 enriched ketone (1) and its hydrate (2), respectively. In 100% dimethyl sulphoxide a single resonance was detected at 200.6 ppm. Incubation of trypsin, whose ^{13}C NMR spectrum at pH 3.2 is shown in Fig. 1B, with ZLMK produced no loss of activity at pH 3.2 and no perturbation of the spectrum of the chloromethyl ketone (Fig. 1C). When the pH was adjusted to 6.9, enzyme activity was decreased and the inhibitor resonances at 95.4 and 204.7 ppm were replaced by a new signal at 98.0 ppm (Fig. 1). The enzyme-inhibitor complex characterized by the resonance at 98.0 ppm (and the analogous complex enriched from [1-^{13}C]ZLMK[7] (δ = 53.8 ppm)) was determined[5,6] from linewidth calculations and pH titration shifts on intact and denatured species, to be the covalent, tetrahedral adduct (3) (Fig. 1). This is in contrast to the findings of Rich *et al*[8-10] who have recently reported ^{13}C NMR studies involving the inhibition of aspartic proteases by ^{13}C-enriched ketonic inhibitors. They found that the inhibiting species was the ketone hydrate, with

† It is a pleasure for us to dedicate this paper towards the celebration of Professor Ralph A. Raphael's 65th birthday. One of us (A.I.S.) was privileged to be his first graduate student at Glasgow University (1949-52) and to have experienced the pleasure of learning organic chemistry from an inspiring teacher.

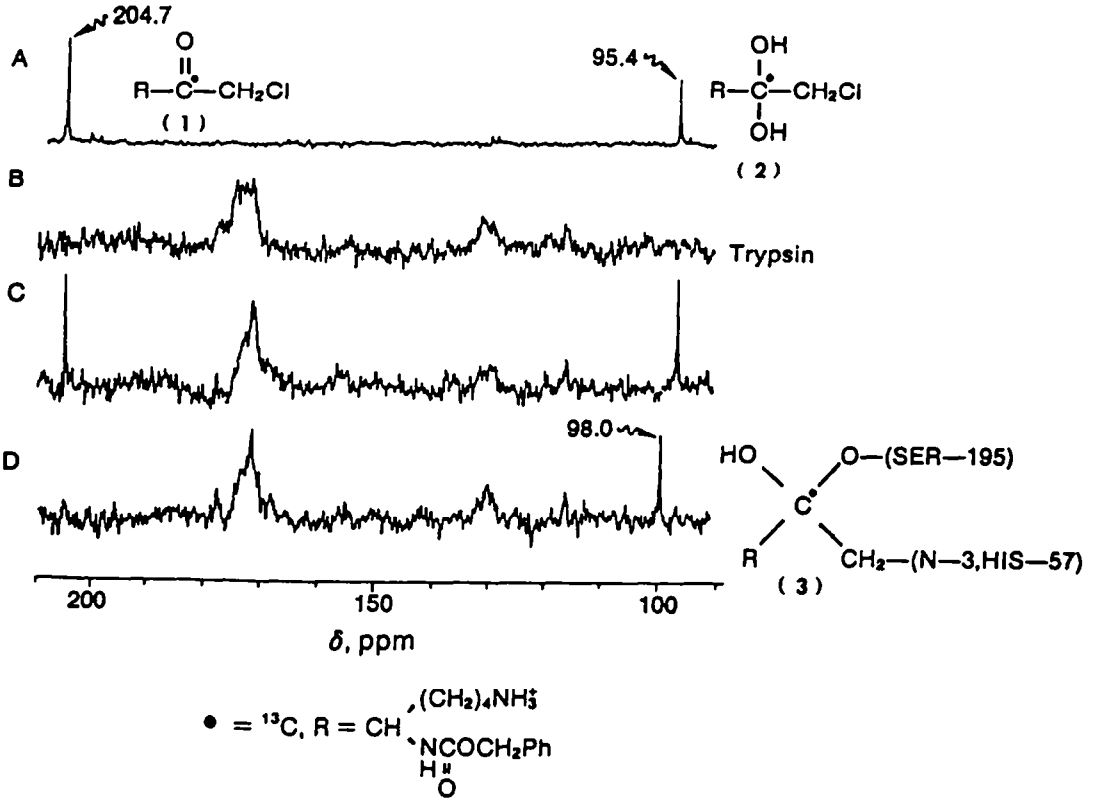

FIGURE 1 ^{13}C NMR spectra of (A) [2-^{13}C]ZLMK (47.6 mM) in 0.6 ml of 1 mM HCl containing 16.7% u/v D_2O. Number of accumulations = 1,660, pH 3.1. (B-D) 20 mM sodium phosphate; 12.5% u/v D_2O; 10,000 accumulations, trypsin 0.28 mM, volume 8 ml. [1-^{13}C]ZLMK concentrations, pH and enzyme activity were as follows: (B) 0.00 mM, pH 3.2, 100%; (C) 0.38 mM, pH 3.2, 100%; (D) 0.37 mM, pH 6.9, 0.06%.

a chemical shift of 99.0 ppm, i.e. there was no covalent interaction with the enzyme. For the above reason, viz. chemical shift comparison, the nature of the species giving rise to the resonance at 98.0 ppm uncovered in our previous work[5,6] has been reinvestigated in order to provide unequivocal evidence that the tetrahedral species (3) is indeed present.

It is apparent that a hydrate, e.g. (2, Fig. 1) can be distinguished from the hemiketal (3) since these species derive their two substituent oxygen atoms from different sources; both from water in the former, and one from Ser-195 and the other from water in the latter.

The effect of substitution of a ^{18}O atom for a ^{16}O atom directly bonded to a ^{13}C atom is to shift the ^{13}C NMR signal upfield.[11] This effect is additive and the magnitude of the shift can be determined by model experiments as shown in Fig. 2. Thus, Fig. 2A shows the ^{13}C NMR spectra of [2-^{13}C]ZLMK in a 1:1 mixture of $H_2{}^{16}O$ and $H_2{}^{18}O$. The ketone appears as a "doublet" with the heavy isotope resonance 0.05 ppm to higher field. The hydrate is an apparent "triplet" with resonances corresponding (from lower field) of ^{16}O, ^{16}O; ^{16}O, ^{18}O and ^{18}O, ^{16}O; ^{18}O, ^{18}O. The isotope effect in this case is 0.02 ppm per ^{18}O atom. This can be compared with Fig. 2B accumulated in 100% $H_2{}^{16}O$ in which both species give rise to singlets; the ketone resonance (3.0 Hz) is broader than the hydrate (1.1 Hz) due to its mode of relaxation (chemical shift anisotropy) at this field strength.

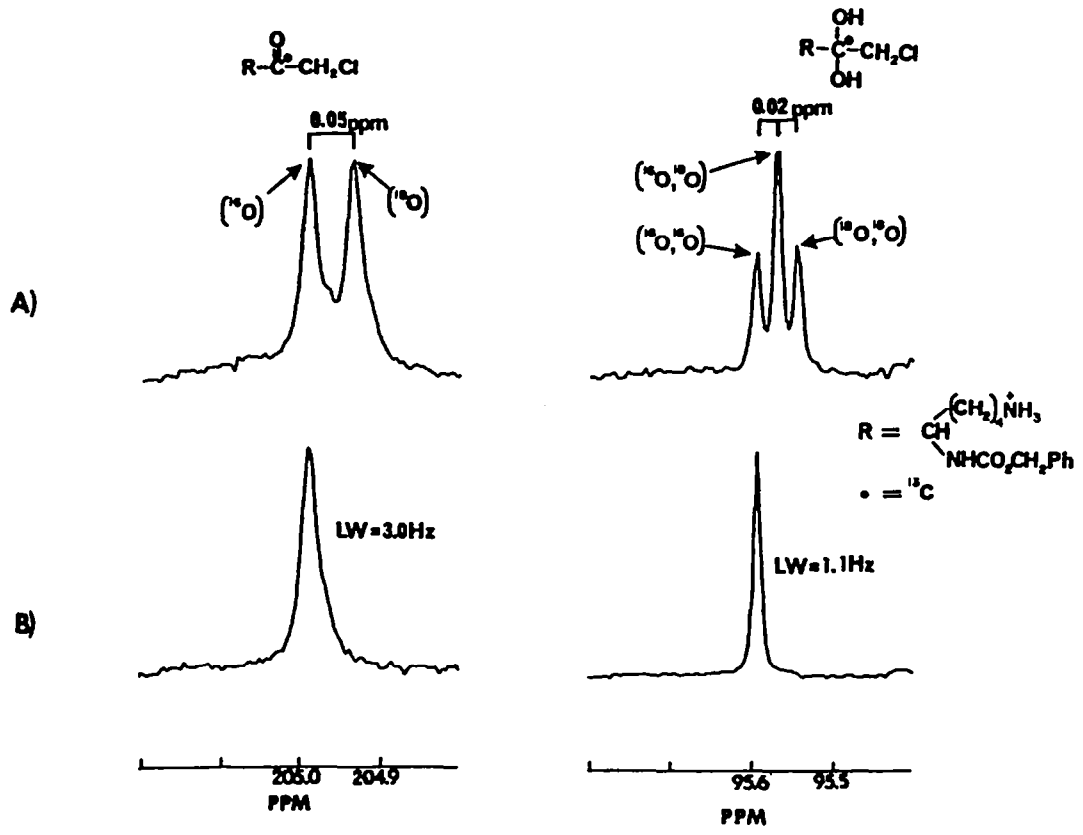

FIGURE 2 ^{13}C NMR spectra of [2-^{13}C]ZLMK (30 mM) in 2.0 ml of 1 m HCl containing 12.5% u/v D_2O. Number of accumulations = 400, pH 3.0. (A) In 50% $H_2{}^{18}O$/50% $H_2{}^{18}O$ and (B) in 100% $H_2{}^{16}O$.

Figure 3 shows the results of a similar series of experiments in which an excess of [2-^{13}C]ZLMK was added to a solution of trypsin. It can be seen that addition of the enzyme caused considerable increase in the ^{13}C NMR linewidths. The excess, unbound inhibitor hydrate at 95.4 ppm, although no longer a resolved "triplet" (Fig. 3D), is discernible as such (Fig. 3B,C,E,F). The crucial series of spectra is shown in Fig. 3A-C. The intrinsically larger linewidth of the resonance of the enzyme inhibitor complex (Fig. 2B) at 98.0 ppm is isotopically increased (Fig. 3A). The magnitude of the increase (0.02 ppm, Fig. 2C) indicates that only one solvent-derived oxygen atom is involved and this confirms the assigned enzyme adduct as (3). However, the quality of the ^{13}C NMR spectra obtainable, although optimal, cast doubt upon the rigour of this technique when high molecular weight species (~ 24,000 M.W.) are involved. Thus, further evidence that the enzyme inhibitor complex was indeed (3) was adduced by X-ray analysis. Binding of various peptide chloromethyl ketone inhibitors to other serine proteases[4,12-14] has been previously investigated by crystallographic methods.

In our present studies it was found that covalent bond formation of trypsin with ZLMK replaced some solvent molecules found in free trypsin[15] and induced only very small changes as indicated by the positional and thermal parameters observed, even in the binding site of trypsin. Thus the enzyme represents a rigid scaffold

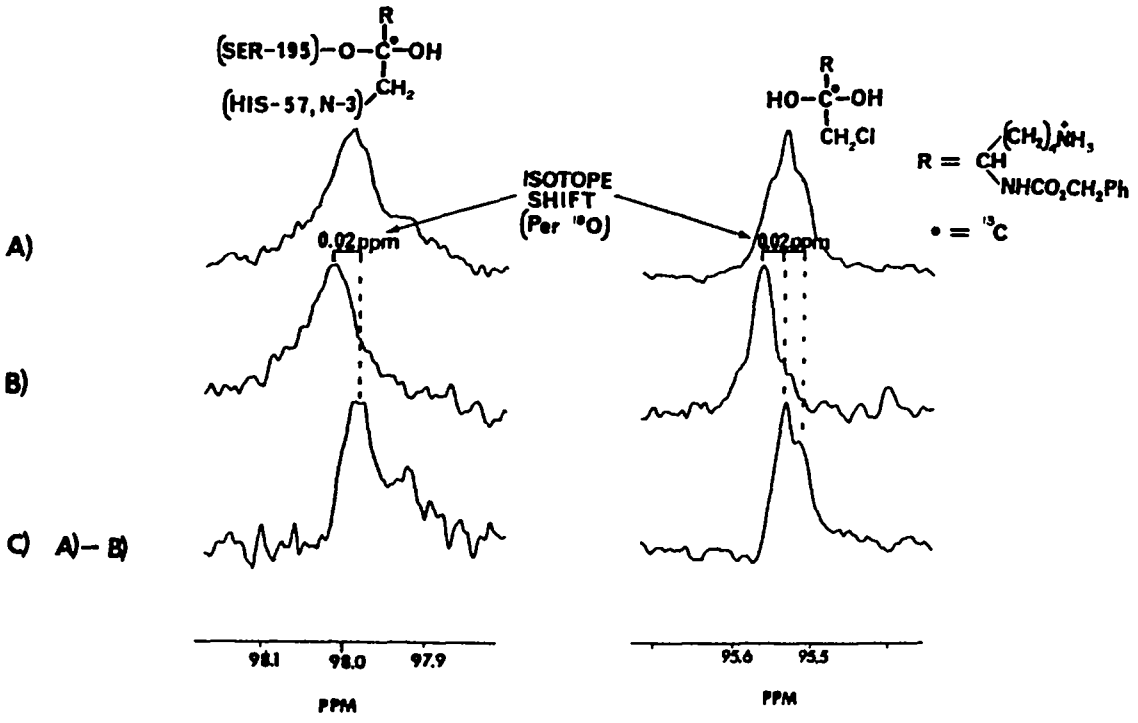

FIGURE 3 ^{13}C NMR spectra of [2-^{13}C]ZLMK (3.5 mM), trypsin (1 mM) in 20 mM sodium phosphate. pH was raised to 7.0 for 30 min, spectra recorded at pH 3.2. (A) In 50% $H_2{}^{16}O$/50% $H_2{}^{18}O$ with 9,000 accumulations. (B) In 100% $H_2{}^{16}O$ with 27,000 accumulations. (C) (A)-(B).

which forces ZLMK into a special conformation. Only the side chains of Gln-192, Ser-195 (OG movement of 0.5 Å) and His-57 (shifts of its imidazole ring up to 0.4 Å) respond significantly to the binding of the ZLMK group.

In the final Fourier map there is no adequate electron density for the benzyl moiety of CBZ 141 (Fig. 4), which is probably disordered. Some isolated density indicates its preferred location parallel to the imidazole ring of His-57, i.e. in the S2-subsite. The residual of ZLMK is well defined by electron density. The interaction of the Lym 151 (Fig. 4) side chain with the specificity pocket of trypsin is identical to that observed for Lys 151 in the trypsin -pancreatic trypsin inhibitor (trypsin-PTI) complex.[16] The hydrogen bond interactions of its terminal ammonium group (NZ, Fig. 4), especially its linkage to the carboxylate group of Asp-189 through the buried solvent molecule OH 414, are shown in Fig. 4 together with their distances.

Compared with the corresponding atom of Lys 151 in the trypsin-PTI complex, Lym 151, C-2, C°, N^{α} and the carbonyl oxygen (O) of ZLMK are considerably displaced to form covalent bonds with Ser-195 OG and the imidazole ring of His-57 and an improved hydrogen bond with Ser-214 O, respectively. Figure 5 shows part of the model around the carbonyl group of Lym 151, overlaid by electron density. There is continuous density between the ZLMK carbonyl carbon (C°) and Ser-195 OG and, similarly, between the ZLMK methylene group (1, Fig. 4) and His-57 NE2 as expected for a covalent bond. As shown by the bond and angle parameters around

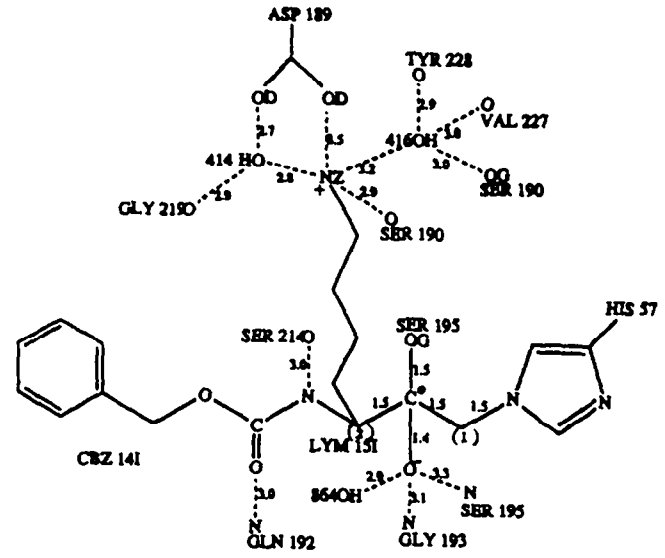

FIGURE 4 Schematic drawing of ZLMK and its specific covalent and hydrogen bond interactions with environmental groups of trypsin. Bond lengths are given in Angstrom unit.

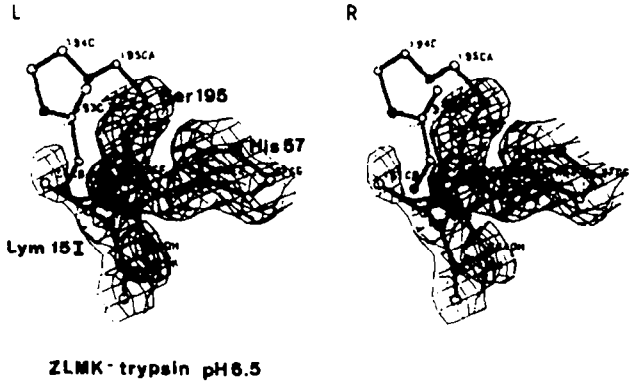

FIGURE 5 Part of the trypsin-ZLMK model around the carbonyl group of Lym 151 with a small section of the final 2Fobs-Fcalc electron density map overlaid. Contour level at 1 σ.

TABLE 1 Initial and refined bond lengths and angles around the carbonyl carbon (C°) and methylene group (1) (see Fig. 4).

Bonds/Angles	(2)-C°	(1)-C°	O-C°	OG-C°	(1)-NE2	(2)-C°-(1) [deg]	(2)-C°-O [deg]	(1)-C°-O [deg]
Starting values	1.52	1.52	1.26	1.43	1.46	91	121	130
"weak constraint"	1.52	1.51	1.40	1.41	1.54	106	104	118
"no constraint"	1.62	1.57	1.35	1.50	1.49	102	101	125

the carbonyl carbon (C°) (Table 1), the geometry of the refined model is virtually identical with that expected for a purely tetrahedral state: the sum of the three bond angles meeting at C° is 328° for "weak" constraint as well as for "no constraint" (i.e. close to 330°; the deviation of C° from the plane formed by its covalent neighbours (2), (1) and O is 0.47 Å (i.e. close to 0.5 Å expected for a tetrahedral state); the distances between the carbonyl carbon (C°) and OG are slightly below and above 1.4 Å, respectively (i.e. at values expected for a single C-O bond of standard length (1.43 Å). The carbonyl oxygen

points into the oxyanion hole (Fig. 4) to form two relatively long hydrogen bonds with the amide protons of Gly-193 and Ser-195. Solvent molecule 864 OH (Fig. 4) is in a favourable position and 2.9 Å away to fully exploit its hydrogen bonding capability.

The initial (Fph9-Fph6.5) difference Fourier map calculated with the intensities obtained at pH 9 and 6.5, respectively, showed only some stronger features at the side chain of His-57, which seemed mainly to indicate a considerable movement of its imidazole ring upon change of pH. A comparison of the refined structures showed, however, only a small rotation of the hemiketal group about the carbonyl carbon (C°) of Lym 151 and a very slight shift of the His-57 imidazole ring away from it. In the course of refinement, the covalent bond between the methylene group (1) and His-57 NE2 increased to above 1.7 Å. This intolerably large value is clearly the consequence of the simultaneous usage of normal constraints (as for the imidazole ring) and no constraint (for the (2)-NE2 interaction) and indicates the limits of this refinement procedure. The geometry at the carbonyl carbon, its arrangement relative to the oxyanion hole, and the position of solvent molecule 864 OH remain essentially constant upon change of pH.

In summary, from these X-ray crystallographic studies we conclude that the enzyme-inhibitor complex between ZLMK and trypsin is the covalent species (3) and that the considerable chemical shift changes induced by pH titration[5-7] are not due to one or more reversible conformational changes. Despite suggestions[17] that there exists an uneasy alliance between NMR spectroscopy and X-ray diffraction, a concerted use of these techniques to characterize protein complexes has proven useful.[18] It could still be argued that there is insufficient correspondence between the solution (NMR) and solid (X-ray) characteristics of the ZLMK-trypsin complex. To provide the final diagnosis, Figure 6 shows the ^{13}C cross-polarization magic angle spinning (CPMAS) NMR spectrum of the ZLMK-trypsin complex (^{13}C=O enriched) in the solid state. The resonance at 98.0 ppm (not present in the ^{13}C CPMAS spectrum of trypsin) can be assigned to the enriched carbon of (3 (Fig. 1). The broad overlap of resonance at 170-180 ppm and 10-60 ppm arises from natural abundance trypsin.

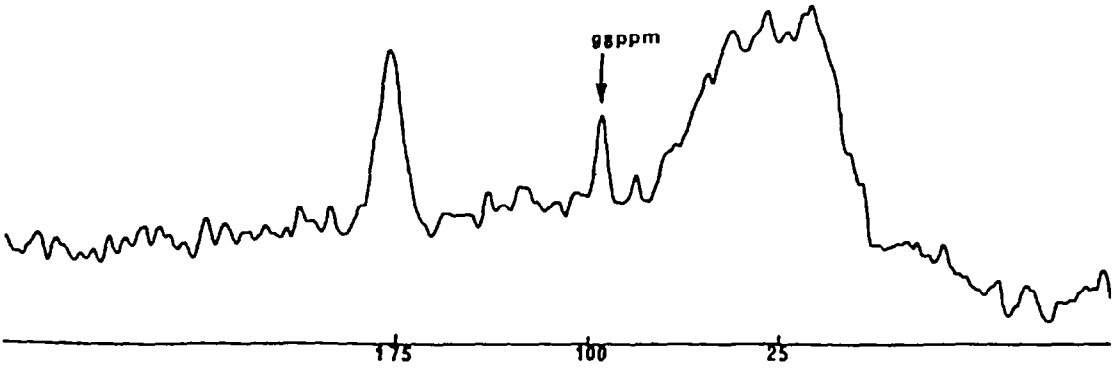

FIGURE 6 ^{13}C CPMAS NMR spectrum of 100 mgs of lyophilisate of a solution described in legend to Fig. 1D. The acquisition parameters were: 23,840 scans, 26 msec acquisition time, 20 KHz spectral width, 3.0 msec contact time, and rf power levels such that a 90° ^{1}H pulse was 7.0 µsec. The free induction decays were zero-filled four-fold and transformed with no line broadening.

Thus, comparison of the inhibitor complex X-ray data with the solid-state (CPMAS) spectrum taken together with the ^{18}O isotopic shift data in solution has provided unequivocal evidence that the serine-195 hydroxyl forms a covalent bond with the carbonyl carbon of the ZLMK moiety. These results provide a firm basis for the more detailed studies of the ionization of such complexes and adds to our overall understanding of the active site dynamics and mechanism of catalysis exhibited by the serine proteases, where nucleophilic enhancement of the Ser-195 hydroxyl is sufficient to develop covalent interaction with carbonyl groups inserted in site-specific inhibitors which serve as transition-state analogues of the scissile peptide bond of the natural substrates.

EXPERIMENTAL

[2-^{13}C]ZLMK was synthesized following literature procedures.[6] Oxygen-18 (99%) was purchased from Cambridge Isotope Laboratories. NMR spectra shown in Figures 1 and 3 were recorded on a Bruker WM300 wide-bore spectrometer operating at 75.5 MHz for carbon-13 nuclei and those in Figure 2 on a Bruker AM500 operating at 125.8 MHz for carbon-13 nuclei. The solid-state spectrum shown in Figure 6 was recorded on a Bruker CXP100 operating at 25.16 MHz for ^{13}C nuclei.

For X-ray diffraction, large orthorhombic crystals of approximate size 0.3 x 0.3 x 0.6 mm^3 and of space group $P2_12_12_1$ were grown from bovine β-trypsin at 20°C in 1.6 M ammonium sulfate, pH 5.0, in the presence of 10 mM benzamidine and 0.1 mM in $CaCl_2$ by seeding with crushed crystals and using the vapour diffusion technique.[19] After washing several times with benzamidine-free ammonium sulfate, these crystals were soaked for three weeks in a 2.5 M ammonium sulfate solution, buffered with 0.1 M phosphate to pH 6.5, which contained 1 mM [1-^{13}C]ZLMK. For measurements at higher pH values, these crystals were eventually transferred to 2.5 M ammonium sulfate, buffered to pH 9.

The cell constants of these crystals were according to precession and rotation photographs a = 55.0, b = 58.4, and c = 67.8 Å, i.e. virtually identical with those of the original crystals containing benzamidine inhibited trypsin.[20] Data to 1.6 Å (1.7 Å) (data at pH 9 are given in parentheses) resolution were collected on a rotation camera using a flat film cassette. Within four weeks, about 51,600 (53,600) reflections above the 1.5 σ level were collected and merged yielding 21,900 (19,100) independent reflections with an Rmerge (defined as $\sum(I-\{I\})/\sum\{I\}$, summation over all measurements, {I} averaged intensity, I intensity of individual measurements) of 0.072 (0.085).

An initial difference Fourier map calculated between ZLMK-trypsin and free trypsin,[15] using the phases of free trypsin, allowed modelling of the inhibitor moiety on a Vector General 3400 graphic display system using the programme FRODO.[21] The ZLMK model (as given in Fig. 4) was constructed by linking a carbobenzoxy group as found in carbobenzoxy-L-leucyl-p-nitrophenol ester[22] to the Lys 151 residue of the pancreatic trypsin inhibitor as observed in its complex with trypsin.[16] Crystallographic refinement of this ZLMK-trypsin complex was performed using the reciprocal space refinement with energy restraints EREF.[23] Except for the carbonyl carbon of Lym 151 (C° in Fig. 4), the energy parameters recommended by Levitt[24] were used. In addition, individual temperature factors were refined. In order not to restrict the OG atom of Ser-195 and the methylene group ((2) of Lym 151, Fig. 4) to freely approach the carbonyl group of the inhibitor and the imidazole ring of His-57, respectively, all non-bonded interactions were set to zero for Ser-195 OG and Lym 151 carbonyl carbon and methylene group.

In the course of refinement the energy constraints imposed on the covalent bonds formed with CF were gradually removed. In a first series of refine cycles, weak force constants (one-third of those given by Levitt) were applied ("weak constraint"), whereas in a subsequent series all bond angles and bond lengths involving CF were completely unconstrained ("contraint").

The final structure contains, besides the ZLMK-trypsin model, the unique calcium ion[16] and 91 stereochemically reasonable solvent molecules with temperature factors below 42 Å. The final R-value defined as $(\sum(|Fobs|-|Fcalc|)/\sum|Fobs|)$ is 0.182 (0.184) for all reflections in the resolution range between 7.0 and 1.6 Å (1.7 Å). The r.m.s. deviation of all main chain atoms is 0.11 Å between the ZLMK-trypsin structures at pH 6.5 and 9, and 0.14 Å compared with free trypsin.

ACKNOWLEDGEMENTS

We thank the National Institutes of Health and the Robert A. Welch Foundation for financial support.

REFERENCES

1. Present address: The Upjohn Company, Kalamazoo, Michigan 69001, U.S.A.
2. J.R. Coggins, W. Kray and E. Shaw, Biochem. J., 138, 579 (1974).
3. E. Shaw and S. Springhorn, Biochem. Biophys. Res. Commun., 27, 391 (1967).
4. T.L. Poulos, R.A. Alden, S.T. Freer, J.J. Birktoft and J.J. Kraut, J. Biol. Chem. 252, 1097 (1976).
5. J.P.G. Malthouse, N.E. Mackenzie, A.S.F. Boyd and A.I. Scott, J. Am. Chem. Soc., 105, 1685 (1983).
6. J.P.G. Malthouse, W.U. Primrose, N.E. Mackenzie and A.I. Scott, Biochemistry, 24, 3478 (1985).
7. W.U. Primrose, A.I. Scott, N.E. Mackenzie and J.P.G. Malthouse, Biochem. J., 231, 677 (1985).
8. D.H. Rich, M.S. Bernatowicz and P.G. Schmidt, J. Am. Chem. Soc., 104, 3535 (1982).
9. D.H. Rich, F.G. Salituro, M.W. Holladay and P.G. Schmidt, in "Conformationally Directed Drug Design", (J.A. Vida and M. Gordon, eds.), American Chemical Society, Washington, D.C.; ACS Symp. Ser. No. 251, pp. 211-237.
10. D.H. Rich, J. Med. Chem., 28, 263 (1985).
11. J.C. Vederas and T.T. Nakashima, J. Chem. Soc., Chem. Commun., 183 (1980).
12. D.M. Segal, J.C. Powers, G.H. Cohen, D.R. Divies and P.E. Wilcox, Biochemistry, 10, 3728 (1971).
13. J.D. Robertus, J. Kraut, R.A. Alden and J.J. Birktoft, Biochemistry, 11, 4293 (1972).
14. M.N.G. James, G.D. Brayer, L.T.J. Delbaerre and A.R. Sielecki, J. Mol. Biol., 139, 423 (1980).
15. W. Bode and P. Schwager, J. Mol. Biol., 98, 693 (1975).
16. R. Huber, D. Kukla, W. Bode, P. Schwager, K. Bartels, J. Deisenhofer and W. Steigemann, J. Mol. Biol., 89, 73 (1974).
17. J.S. Cohen and A.W. Codawer, T.I.B.S., 389 (1982).
18. D. Ringe, B.A. Seaton, M.H. Gelb and R.H. Abeles, Biochemistry, 24, 64 (1985).
19. J. Walter and W. Bode, Hoppe-Seyler's Z. Physiol. Chem., 364, 949 (1983).
20. H. Fehlhammer and W. Bode, J. Mol. Biol., 98, 683 (1975).
21. T.A. Jones, J. Appl. Crystal., A11, 268 (1978).
22. V.M. Coiro, F. Mazza and G. Mignucci, Acta Crystallogr., 30, 2607 (1974).
23. A.T. Jack and M. Levitt, Acta Crystallogr., A34, 931 (1978). 24. M. Levitt, J. Mol. Biol., 82, 393 (1974).