Selective Alteration of Substrate Specificity by Replacement of Aspartic Acid-189 with Lysine in the Binding Pocket of Trypsin†

Laszlo Graf,‡,§ Charles S. Craik,‖,⊥ Andras Patthy,‖ Steven Rozcniak,‡ Robert J. Fletterick,‖ and William J. Rutter*,‡,‖

Hormone Research Institute, Department of Biochemistry and Biophysics, and Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California 94143

Received September 30, 1986; Revised Manuscript Received December 19, 1986

ABSTRACT: To test the role of Asp-189 which is located at the base of the substrate binding pocket in determining the specificity of trypsin toward basic substrates, this residue was replaced with a lysine residue by site-directed mutagenesis. Both rat trypsinogen and Lys-189 trypsinogen were expressed and secreted into the periplasmic space of *Escherichia coli*. The proteins were purified to homogeneity and activated by porcine enterokinase, and their catalytic activities were determined on natural and synthetic substrates. Lys-189 trypsin displayed no catalytic activity toward arginyl and lysyl substrates. Further, there was no compensatory change in specificity toward acidic substrates; no cleavage of aspartyl or glutamyl bonds was detected. Additional studies of substrate specificity involving gas-phase sequence analyses of digested natural substrates revealed an inherent but low chymotrypsin-like activity of trypsin. This activity was retained but modified by the Asp to Lys change at position 189. In addition to hydrolyzing phenylalanyl and tyrosyl peptide bonds, the mutant enzyme has the unique property of cleaving leucyl bonds. On the basis of computer graphic modeling studies of the Lys-189 side chain, it appears that the positively charged NH_2 group is directed outside the substrate binding pocket. The resulting hydrophobic cavity may explain the altered substrate specificity of the mutant enzyme. The relatively low chymotrypsin-like activity of both recombinant enzymes may be due to distorted positioning of the scissile bond with respect to the catalytic triad rather than to the lack of sufficient interaction between the hydrophobic side chains and the substrate binding pocket of the enzyme.

The serine proteases are members of a large family of homologous proteins which require the amino acid serine (as well as aspartic acid and histidine) at their active site and appear to use the same mechanism for catalysis. Each member of the family exhibits a unique substrate specificity that is primarily determined by the structural characteristics of the substrate binding pocket. The substrate is bound such that the scissile bond is properly positioned relative to the conserved active-site residues (Kraut, 1977; Steitz & Shulman, 1982).

Trypsin catalyzes the hydrolysis of lysyl or arginyl peptide and ester bonds. Our earlier studies showed that replacement of the glycine residues at positions 216 and 226 at the entrance of the substrate binding pocket with alanine altered the discrimination between arginyl and lysyl substrates (Craik et al., 1985). The three-dimensional structure of trypsin bound to various inhibitors or pseudosubstrates suggests that Asp-189, which resides at the base of the substrate binding pocket, is a primary determinant of the preference for basic amino acids. Asp-189 forms an electrostatic bond either directly or indirectly (mediated through water) with the Lys or Arg residue of the substrate (Ruhlmann et al., 1973; Stroud et al., 1974; Sweet et al., 1974; Bode & Schwager, 1975).

On the basis of steric considerations alone, replacement of Asp-189 with Lys might alter the substrate specificity; for example, by reversing the geometry of the Asp/Lys charge pair, this mutant enzyme might cleave peptide bonds after Asp or Glu residues. We have produced the mutant rat trypsin (Asp-189 → Lys) hereafter referred to as Lys-189 trypsin, purified the protein to homogeneity, and investigated the substrate specificity of the mutant enzyme. Consistent with the proposed crucial role of Asp-189, the selective activity toward basic substrates was lost; however, there was no corresponding increase in activity toward acidic substrates. Instead, the intrinsic low level of activity toward hydrophobic residues (chymotrypsin-like activity) was retained and modified.

Materials and Methods

Materials. (Diethylaminoethyl)cellulose (DEAE-cellulose, DE52)[1] and (carboxymethyl)cellulose (CM-cellulose CM52) were purchased from Whatman. A prepacked Superose 12 (cross-linked agarose) column (volume 25 mL) was obtained

† This work was supported by NSF Grant PCM830610 (W.J.R.), NIH Grant AM26081 (R.J.F.), NSF Grant DMB860806 (C.S.C.), a Bristol-Myers Co. grant of the Research Corp. (C.S.C.), and a CRCC Grant (C.S.C.).

* Address correspondence to this author at the Department of Biochemistry and Biophysics, University of California, San Francisco.

‡ Hormone Research Institute.

§ Present address: Biochemistry Department, L. Eotvos University of Budapest, Budapest, Hungary.

‖ Department of Biochemistry and Biophysics.

⊥ Department of Pharmaceutical Chemistry.

[1] Abbreviations: DEAE, diethylaminoethyl; CM, carboxymethyl; swMb, sperm whale apomyoglobin; β_p-LPH, porcine β-lipotropin; α_h-ACTH, human adrenocorticotropic hormone; hGH, human growth hormone; LH-RH, luteinizing hormone releasing hormone; ABI, Applied Biosystems, Inc.; HPLC, high-performance liquid chromatography; PAGE, polyacrylamide gel electrophoresis; DFP, diisopropyl fluorophosphate; TPCK, *N*-tosyl-L-phenylalanine chloromethyl ketone; AFC, 7-amino-4-(trifluoromethyl)coumarin; AMC, 7-amino-4-methylcoumarin; Suc, succinyl; β-NA, β-naphthylamide; SDS, sodium dodecyl sulfate; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid; MES, 2-(*N*-morpholino)ethanesulfonic acid; PTH, phenylthiohydantoin; TFA, trifluoroacetic acid.

0006-2960/87/0426-2616$01.50/0

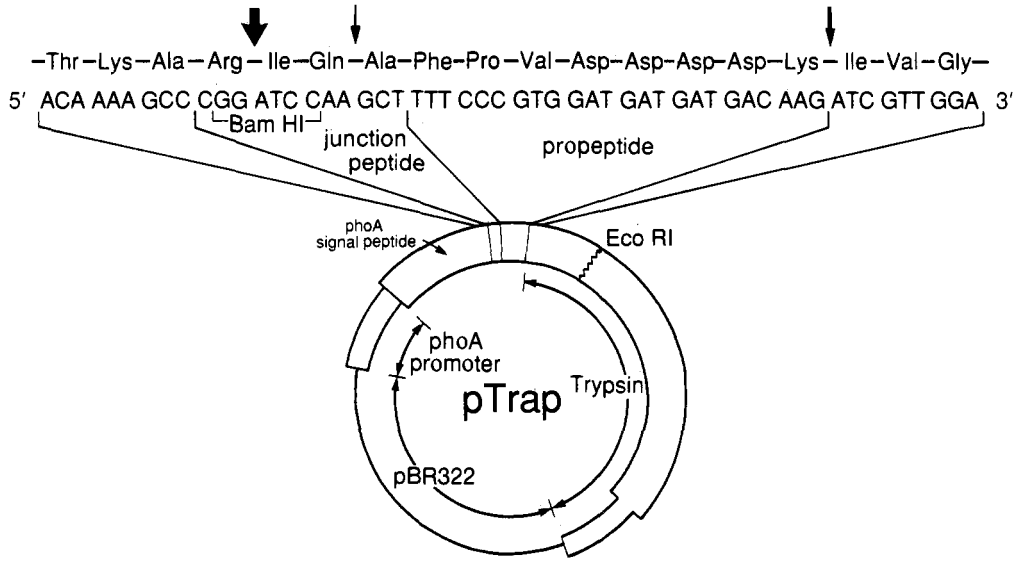


FIGURE 1: Bacterial expression plasmid for trypsinogen (pTrap). The regulatory sequences and signal peptide for bacterial alkaline phosphatase (phoA) were fused in-frame with the trypsinogen sequence. The amino acid and nucleotide sequences at the junction between the two proteins are shown in detail. The alkaline phosphatase/trypsinogen fusion is inserted into a *SalI*/*PvuII* DNA fragment of pBR322 to provide ampicillin resistance to bacteria and provide facile selection of the plasmid. The wide arrow between amino acids Arg and Ile and the narrow arrow between Gln and Ala represent primary and secondary cleavage sites, respectively, of the bacterial signal peptidase. The arrow between Lys and Ile is the cleavage by the natural activating enzyme enterokinase.

from Pharmacia, and gel filtration standards for calibrating the column were from Bio-Rad. Antibodies used for immunological analysis were raised in rabbits against rat trypsinogen and purified on DEAE-cellulose. Crude porcine enterokinase was purchased from Miles. Sperm whale apomyoglobin (swMb) was supplied by Applied Biosystems, Inc. (ABI), Foster City, CA. Porcine β-lipotropin (β_p-LPH) was prepared by the method of Graf and Cseh (1968) and further purified by high-performance liquid chromatography (HPLC). Natural human growth hormone (hGH) and synthetic human adrenocorticotropic hormone (α_h-ACTH) were generously supplied by Dr. C. H. Li (UCSF). Luteinizing hormone releasing hormone (LH-RH) was purchased from Peninsula Laboratories, Belmont, CA. Highly purified bovine pancreatic trypsin (type XIII, TPCK treated), α-chymotrypsin (3 times crystallized), diisopropyl fluorophosphate (DFP), *N*-tosyl-L-phenylalanine chloromethyl ketone (TPCK), and benzamidine hydrochloride were obtained from Sigma. [1,3-^{3}H]DFP (5.4 Ci/mmol) was purchased from Du Pont. Enzyme substrates with fluorogenic leaving groups, 7-amino-4-(trifluoromethyl)coumarin (AFC) and 7-amino-4-methylcoumarin (AMC), were obtained from Enzyme System Products, Livermore, CA. Gly-(Asp)$_4$-Lys-β-naphthylamide (β-NA) was supplied by Bachem.

Computer Graphics. The computer graphics program INSIGHT (Dayringer et al., 1986) was used with a color Evans and Sutherland PS-300 graphics computer to model the various trypsin structures.

Site-Directed Mutagenesis. A synthetic oligonucleotide of 30 bases was synthesized by solid-phase phosphoramidite chemistry on a 380B DNA synthesizer (Applied Biosystems, Inc., Foster City, CA). The oligonucleotide 5′GGCTTCCT*C*GAGGGAGGCAAG*AAA*TCCTGC3′ was used to prime the synthesis of a DNA strand encoding a trypsinogen template with a Lys codon (AAA) in place of the Asp codon (GAT) at position 189. The primer also introduced a silent mutation into the coding sequence incorporating a recognition sequence for the restriction endonuclease *Xho*I (CTCGAG). This site facilitated the characterization of the Lys-189 trypsinogen DNA. Mutagenesis, transformation of JM101, and screening and sequencing of the mutant template were carried out as described previously (Craik et al., 1985).

Heterologous Expression of Trypsinogen and Lys-189 Trypsinogen. Expression of trypsinogen in bacteria was accomplished by fusing the DNA coding sequences for trypsinogen (Craik et al., 1984) to the DNA sequences encoding the signal peptide and regulatory regions for bacterial alkaline phosphatase (Inouye et al., 1981). This vector, referred to as pTrap, is depicted in Figure 1. In this vector, the native signal peptide of pretrypsinogen is replaced with the signal peptide of alkaline phosphatase to ensure efficient secretion of the zymogen into the periplasmic space. The probability that the protein would fold with correct disulfide formation was maximized by avoiding the reducing environment of the cytoplasm of *Escherichia coli*. The alkaline phosphatase promoter–signal peptide/trypsinogen DNA fusion was ligated into pBR322 for selection by ampicillin resistance in *E. coli* (HB101). This system offers the regulated control of trypsinogen expression by varying the concentration of phosphate in the growth media. However, continuous production of trypsinogen during the growth cycle of *E. coli* is not deleterious to the host (C. S. Craik, unpublished results). Therefore, a strain that permits the constitutive production of trypsinogen was constructed by transforming *E. coli* SM138 [*E. coli* K12, F$^-$, araD139, Δ(lac)U169, relA, rpsL, phoR] with pTrap. The resultant strain was used to heterologously express trypsinogen and Lys-189 trypsinogen to high levels.

Protein Purification and Characterization. E. coli strain SM138 was transformed with the expression vector encoding trypsinogen or Lys-189 trypsinogen and grown in 12-L batches. Typically, two subsequent 12-L batches were combined for enzyme purification. On the basis of SDS–polyacrylamide gel electrophoresis (SDS–PAGE) (Laemmli, 1970) and immunological analysis by Western blotting (Burnette, 1981) of the total cell extract, the amounts of secreted rat trypsinogen and Lys-189 trypsinogen were about 0.7–1.2 mg/L. The experimental procedure for purification of the Lys-189 trypsinogen is as follows.

All purification procedures were carried out at 4 °C unless otherwise stated. Cells were collected by centrifugation at 10000*g* for 10 min and suspended in 480 mL of 25% sucrose and 10 mM Tris-HCl, pH 8. Fifty milliliters of a 0.5% ly-

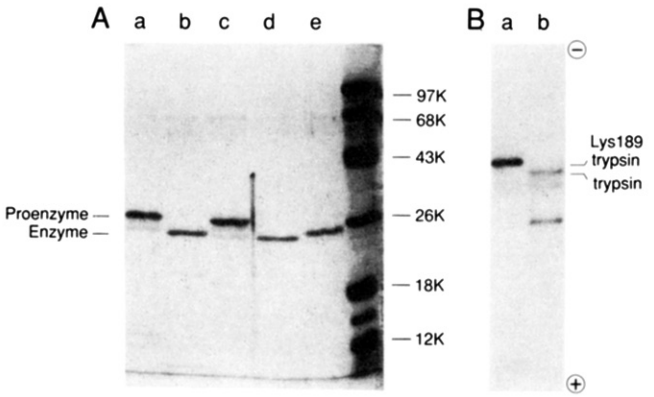


FIGURE 2: (A) SDS–PAGE in a 12.5% polyacrylamide gel of trypsinogen (lane a), trypsin (lane b), Lys-189 trypsinogen (lane c), Lys-189 trypsin (lane d), Asn-102 trypsin (lane e), and a mixture of proteins (phosphorylase *b*, bovine serum albumin, ovalbumin, α-chymotrypsin, β-lactoglobulin, and cytochrome *c*) as molecular weight standards (right lane); 0.8–1.2 μg of zymogens or activated enzymes was loaded onto the gel. (B) PAGE in a 12.5% polyacrylamide gel (in the absence of SDS and β-mercaptoethanol) of Lys-189 trypsin (lane a) and trypsin (lane b); 1 μg of protein was applied to the gel in lanes a–e.

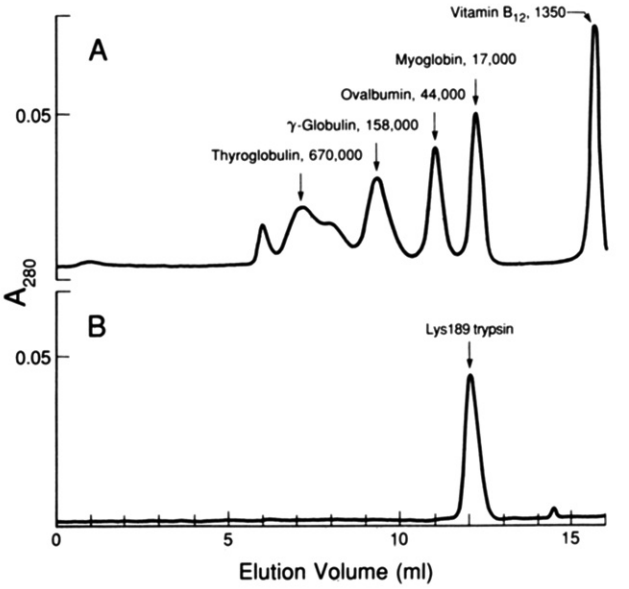


FIGURE 3: Chromatography of 1.5 mg of the Bio-Rad gel filtration standard (A) and 15 μg of activated Lys-189 trypsinogen on a Superose 12 column (B). Buffer, 0.5 M NaCl and 10 mM MES, pH 6; flow rate, 0.5 mL/min.

sozyme solution in 20 mM EDTA, pH 8, was added to the suspension, and the mixture was left for 1 h at room temperature. The periplasmic protein fraction released from the cells was obtained by centrifugation at 50000*g* for 20 min, and the supernatant was dialyzed extensively against 10 mM MES buffer, pH 6.0. The dialyzed protein fraction was loaded onto a DEAE-cellulose column (2.5 × 17 cm) preequilibrated with the same buffer. The column was eluted by a linear gradient from 0 to 0.5 M NaCl in 10 mM MES, pH 6.0. Fractions were monitored by SDS–PAGE and Western blotting. Lys-189 trypsinogen was eluted at 0.2 M NaCl. The fractions containing Lys-189 trypsinogen (~60 mL) were dialyzed against 10 mM sodium citrate buffer, pH 3.0. This dialysis step was used to precipitate the majority of the contaminating protein. Approximately 50% of Lys-189 trypsinogen remained in solution. Subsequent to clarification by centrifugation at 10000*g* for 10 min, this fraction was loaded onto a CM-cellulose column (1.4 × 40 cm) in 10 mM sodium citrate, pH 3.0. The column was eluted with a linear gradient from 0 to 0.5 M NaCl in 10 mM sodium citrate, pH 3.0. The fractions containing Lys-189 trypsinogen were greater than 90% homogeneous as determined by SDS gel electrophoresis. The total yield at this stage was about 5 mg, representing 20–30% of the zymogen present in the starting fermentation mixture (Figure 2A, lane c).

Lys-189 trypsinogen was concentrated by centrifugation in Centricon-10 tubes. The concentration of the final protein solution was determined by amino acid analysis to be 0.2 mM. For amino acid analysis, a Waters PICO-TAG system was used (Heinrikson & Meredith, 1984). The purification of trypsinogen was essentially identical with the procedure described above except that 95% of the trypsinogen precipitated during the 10 mM sodium citrate, pH 3.0, dialysis step. The resultant precipitate was extracted with 10 mM ammonium acetate, pH 5.0. CM-cellulose column chromatography was performed as described above except that 10 mM ammonium acetate, pH 5.0, substituted for sodium citrate, pH 3.0, to purify trypsinogen. This step resulted in about 3 mg of greater than 90% homogeneous trypsinogen (Figure 2A, lane a).

Purified trypsinogen was subjected to N-terminal sequence analysis by Edman degradation on a Model 470A protein sequencer and a Model 120A PTH amino acid analyzer (ABI). The primary site of cleavage by the bacterial signal peptide occurred between the Arg and Ile residues. Secondary cleavage (20%) occurred between the Gln and Ala residues at the signal peptide/trypsinogen junction (see Figure 1).

Activation of Trypsinogen to Trypsin. Crude porcine enterokinase was purified according to Liepnieks and Light (1974). This resulted in a preparation that was still quite heterogeneous as shown by chromatography on a Superose 12 column in 0.5 M NaCl and 10 mM MES, pH 6.0. The elution profile monitored at 280 nm revealed four major peaks. Each peak exhibited significant activity on the trypsin substrate D-Val-Leu-Arg-AFC, and the main peak of the enterokinase preparation coeluted with highly purified trypsin. The contaminating trypsin activity prevented the use of the major fraction of the enterokinase preparation for trypsinogen activation studies. The least retarded peak was about 100 times more active on the specific enterokinase substrate Gly-$(Asp)_4$-Lys-β-NA than on D-Val-Leu-Arg-AFC. The position of this peak on the Superose 12 column corresponded to that of γ-globulin with a molecular weight of 158 000 in agreement with the 145 000 molecular weight of bovine enterokinase (Anderson et al., 1977). The enterokinase peak was repurified 2 times on the same Superose 12 column to ensure that no contaminating trypsin was associated with the activated enzyme.

Three milligrams of Lys-189 trypsinogen was activated with 60 μL of highly purified enterokinase (enzyme to substrate ratio is 1:50, by weight) in 1 mL of 0.5 M NaCl, 2 mM $CaCl_2$, and 10 mM MES, pH 6.0, at 37 °C for 14 h. Due to the lack of trypsin-like activity of the mutant trypsin, autoproteolysis products which normally form during the "activation" step were not observed. Lys-189 trypsin was purified from the small propeptide fragment and the high molecular weight enterokinase by chromatography on the Superose 12 column using the conditions described above (Figure 3). The resulting sample was analyzed directly or aliquoted and kept frozen until use.

Trypsinogen was activated with the highly purified enterokinase in 0.25 M NaCl and 10 mM ammonium acetate, pH 5, for 1 h at room temperature (enzyme to substrate ratio was 1:10, w/w). After activation was complete, the pH of the solution was lowered to 3 by adding HCl, and the sample was

concentrated by Centricon-10 tubes. Protein concentration was determined by amino acid analysis and active-site titration with radiolabeled DFP.

Enzyme Assays. Trypsin obtained from zymogen expressed in *E. coli* was indistinguishable from trypsin prepared from either the rat pancreas or a mammalian expression system (Craik et al., 1985) when assayed on D-Val-Leu-Arg-AFC. Highly purified Lys-189 trypsin, however, showed no activity on either D-Val-Leu-Arg-AFC or D-Val-Leu-Asp-AFC. In a search for proteolytic activity of this mutant protease, several natural protein and polypeptide substrates were used. swMb, hGH, α_h-ACTH, and β_p-LPH were incubated with the enzyme in 50 mM Tris-HCl, pH 8, and 10 mM $CaCl_2$ with a 1:(5–10) (w/w) ratio of enzyme to substrate at 37 °C for different periods of time. SDS–PAGE and high-performance liquid chromatography (HPLC) were used to analyze the proteolytic fragments. HPLC analysis was done by using a Vydac C-18 column, 0.46 × 25 cm, and a linear gradient from 0 to 80% acetonitrile in 0.1% TFA. Column fractions were monitored at 215 nm. The resulting fragments isolated by HPLC were subjected to gas-phase sequencing to identify the peptide bond(s) cleaved.

Twenty-microgram aliquots (about 20 nmol) of LH-RH in 100 μL of 50 mM ammonium acetate, pH 7.5, were digested with 2 μg of bovine trypsin, rat trypsin, and Lys-189 trypsin (1:10 enzyme:substrate ratios, w/w) and with 10 ng of bovine chymotrypsin (1:2000 enzyme:substrate ratio, w/w) at 37 °C for 10 h; 200-pmol aliquots were applied directly to the gas-phase sequenator to identify the cleavage sites and determine the relative rates of cleavage.

On the basis of information from the studies above, the following synthetic substrates were used for detailed kinetic analyses of wild-type and Lys-189 rat trypsins: Suc-Ala-Ala-Pro-Phe-AMC, Suc-Leu-Leu-Val-Tyr-AMC, and Suc-Phe-Gly-Ala-Leu-AMC. Enzyme assays were carried out in 10 mM $CaCl_4$ and 50 mM Tris-HCl, pH 8, using a Perkin-Elmer LS5 spectrofluorometer. Fluorescence values were converted to micromoles of the product using a standard solution of AMC or AFC. The concentrations of trypsin and mutant trypsin ranged from 10 to 120 nM, and the substrate concentrations were chosen so that the observed K_m values could be monitored over a 5–20-fold range. The concentrations of Suc-Ala-Ala-Pro-Phe-AMC were 0.025, 0.05, 0.075, 0.1, 0.5, 1.0, and 1.5 mM. Suc-Leu-Leu-Val-Tyr-AMC was hydrolyzed at concentrations of 0.025, 0.033, 0.05, 0.075, 0.1, and 0.15 mM. Poor solubility of this latter substrate did not allow its use at higher concentrations. The concentrations of Suc-Phe-Gly-Ala-Leu-AMC substrate were 0.25, 0.33, 0.5, 0.75, 1.0, 1.5, and 3.0 mM. The steady-state parameters k_{cat}, K_m, and k_{cat}/K_m were calculated by a program that carries out a nonlinear least-squares regression analysis (Cleland, 1979).

Active-site titrations were carried out by using radiolabeled DFP. Trypsin and Lys-189 trypsin (0.55 nmol on the basis of amino acid analysis) were treated with 0.1 μmol of tritiated DFP (3 μCi) in 1 mL of 50 mM phosphate buffer, pH 7.5, for 1 h at room temperature. The solutions were then adjusted to pH 3.0 with 1 N HCl and extensively dialyzed against 0.01 N HCl at 4 °C overnight to remove unincorporated radioactivity. Aliquots of the dialyzed samples were analyzed by liquid scintillation counting to determine the concentration of active enzyme. To establish that Ser-195 was essential for catalytic activity, trypsin and Lys-189 trypsin were assayed under standard conditions with Suc-Ala-Ala-Pro-Phe-AMC (0.1 mM) in the presence of 10^{-4} M DFP.

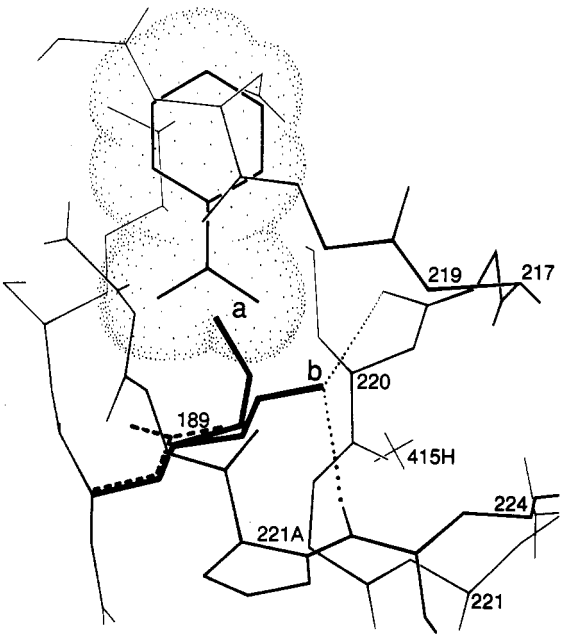


FIGURE 4: Computer graphics representation of part of the substrate binding pockets of trypsin. The position of the inhibitor, benzamidine, is taken from the X-ray analysis of bovine trypsin (Bode & Schwager, 1975). The van der Waals surface of the inhibitor defines the interior of the specificity pocket. The recognition residue Asp-189 is shown in dashed representation. Two positions, a and b, are shown for the modeled Lys-189. In position b, the side chain is extended into a hydrophilic environment where it would form two H bonds (shown as dashed lines) with the carbonyl oxygens of residues 219 and 224. In position a, the side chain extends into the pocket. Position 415H is a water molecule, 2.0 Å from the modeled position of Nζ of Lys-189b.

Attempts were made to inactivate trypsin and Lys-189 trypsins with an alkylating affinity reagent, TPCK (Schoellmann & Shaw, 1963). The highest amount of TPCK that could be dissolved in the aqueous assay solution gave a final concentration of 0.6 mM.

The kinetic analysis of the inhibitory effect of benzamidine (an arginine substrate analogue of trypsin) on the hydrolysis of D-Val-Leu-Arg-AFC and Suc-Ala-Ala-Pro-Phe-AMC by trypsin was also performed. Conditions of the enzyme assay were the same as described before. The inhibitor concentration was 10 μM. Kinetic data were analyzed as described above.

RESULTS

Modeling the Lys-189 in Rat Trypsin. The structures of bovine and rat trypsin were modeled with benzamidine bound (Bode & Schwager, 1975; Sprang et al., 1987) in the native form and with a lysyl side chain substituted for aspartyl in position 189. The positions of the Cβ, Cγ, and Cδ of lysine were taken approximately from their counterparts in Asp-189 (Cβ, Cγ, and OCδ). The Cε and Nζ atoms of lysyl were adjusted by systematic rotations of the side chain torsional angles χ_2 and χ_3. There are two modes of χ_2 and χ_3 angles that produce sterically feasible conformations of the side chain. In the first mode, the side chain is L shaped, and the positive charge of lysine points into the pocket (Figure 4, position a). This orientation would allow a stable interaction between Lys-189 and the negatively charged carboxylate of an Asp and Glu. Thus, acidic residues might be recognized as P_1 residues[2] of the substrate. In the second mode, χ_2 and χ_3 adopt the favored trans position, and the side chain is more extended

[2] In designating the position of an amino acid in the substrate relative to the scissile bond, the terminology of Schechter and Berger (1967) was used.

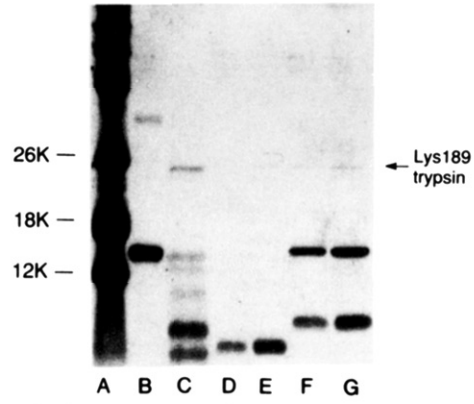


FIGURE 5: SDS–PAGE in a 15% polyacrylamide gel of molecular weight standards (see the legend for Figure 3) (lane A), 2 μg of swMb (lane B), 2 μg of swMb digested with Lys-189 trypsin for 5 h (lane C), and two aliquots (lanes D, E and F, G) of the two HPLC fractions of the swMb digest.

and reaches out of the base of the pocket into an internal hydrophilic environment (Figure 4, position b). The Nζ atom with its positive charge is positioned adjacent to three oxygen atoms—the main chain carbonyl oxygens of residues 217, 219, and 224. In this position, hydrogen bonds of 2.9 Å to residue 219 and 3.0 Å to residue 224 are feasible. The carbonyl oxygen of residue 217 is 3.5 Å distant. Solvent-accessible surface calculations (Richmond, 1984) on this model structure show the positive charge to be buried in this hydrophilic environment and inaccessible to solvent. The atoms Cγ, Cδ, and Cε present a hydrophobic area of 10 Å at the base of the pocket. The charged atom NZ and Cε are solvent accessible in the alternative structure with the L-shaped lysyl side chain (Figure 4, position a).

Production and Characterization of Recombinant Trypsin and Lys-189 Trypsin. We modified the rat trypsin cDNA/gene construct by site-directed mutagenesis such that the Asp codon at position 189 was replaced by a Lys codon (see Materials and Methods). The sequences encoding trypsinogen and Lys-189 trypsinogen were expressed in *E. coli* under the control of the alkaline phosphatase promoter. The proteins were purified and characterized. The purity of Lys-189 trypsin as analyzed by gel filtration chromatography is shown in Figure 3. Patterns of SDS–PAGE of trypsinogen, trypsin, Lys-189 trypsinogen, Lys-189 trypsin, and Asn-102 trypsin are shown in Figure 2A. As expected, SDS–PAGE distinguishes between the zymogen and activated forms of the enzymes but does not distinguish between the mutants and naturally occurring enzymes. On the other hand, PAGE under identical conditions (pH 8.8) but in the absence of SDS and β-mercaptoethanol clearly reveals the expected charge difference between Lys-189 trypsin and trypsin (Figure 2B). The breakdown products seen in lane b of Figure 2B (but not in lane b of Figure 2A) are clearly due to the autoproteolysis of trypsin that may have occurred under the nondenaturing conditions of the PAGE.

To confirm that enterokinase cleaved the proper peptide bond in Lys-189 trypsinogen, 150 pmol of the corresponding mutant trypsin was subjected to gas-phase sequencing, and the N-terminal sequence was as follows: Ile-Val-Gly-Gly-Tyr-. This N-terminal sequence is identical with the amino acid sequence deduced for trypsin from the nucleic acid sequence of the gene for rat trypsinogen II (Craik et al., 1984).

Action of Mutant and Wild-Type Trypsins on Natural Proteins and Polypeptides. Lys-189 trypsin exhibited no activity on the synthetic substrates D-Val-Leu-Arg-AFC and D-Val-Leu-Asp-AFC. However, the enzyme exhibited limited

	P_3 P_2 P_1	P_1' P_2' P_3'
	89	90
swMb	-Lys-Pro-Leu	Ala-Gln-Ser
	50	51
β_p-LPH	-Glu-His-Phe	Arg-Trp-Gly-
	74	75
	-Thr-Pro-Leu	Val-Thr-Leu
	78	79
	-Thr-Leu-Phe	Lys-Asn-Ala
	2	3
α_h-ACTH	NH_2-Ser-Tyr	Ser-Met-Glu
	7	8
	-Glu-His-Phe	Arg-Trp-Gly-
	37	38
	-Phe-Pro-Leu	Glu-Phe-COOH

FIGURE 6: Amino acid sequences around the peptide bonds cleaved by Lys-189 trypsin in sperm whale myoglobin (Edmundson, 1965), porcine β-LPH (Graf et al., 1971a), and human ACTH (Graf et al., 1971b).

BOVINE CHYMOTRYPSIN	7% ↓	10% ↓ pGlu-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2
BOVINE TRYPSIN		30% ↓ pGlu-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2
RECOMBINANT TRYPSIN		12% ↓ 24% ↓ pGlu-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2
LYS189 RAT TRYPSIN		7% ↓ 10% ↓ pGlu-His-Trp-Ser-Tyr-Gly-Leu-Arg-Pro-Gly-NH_2

FIGURE 7: Cleavage sites and percentages of the bonds cleaved in LH-RH (Matsuo et al., 1971) by four different enzymes (for assay conditions, see Materials and Methods).

proteolytic activity in selected natural proteins and peptides: well-defined fragments were generated from both swMb and hGH as determined by SDS–PAGE. The smaller fragment of the two major fragments formed from swMb (Figure 5) was isolated by HPLC, and its N-terminal sequence was determined to be Ala-Gln-Ser- (see Figure 6); this indicates that the Leu–Ala bond at sequence positions 89–90 in swMb has been cleaved by the enzyme. The relatively selective cleavage of swMb by Lys-189 trypsin may be due to the high conformational flexibility and accessibility of this particular peptide bond in the heme-free myoglobin as compared to the other 12 leucyl peptide bonds in swMb.

Unlike Mb, β_p-LPH and α_h-ACTH underwent a more random degradation with a 1:10 (w/w) enzyme:substrate ratio at 37 °C for 5 h. The fragments were resolved by HPLC and identified by sequencing. Results of this analysis are shown in Figure 6. It is clear from these data that the only bonds hydrolyzed followed Leu, Phe, and Tyr residues, indicating a chymotryptic-like activity of the mutant enzyme.

A comparative study of the activity of bovine chymotrypsin, TPCK-treated bovine trypsin, recombinant rat trypsin, and Lys-189 trypsin was carried out on LH-RH (Figure 7). LH-RH has a blocked N-terminus and three potential chymotrypsin cleavage sites: Trp–Ser, Tyr–Gly, and Leu–Arg. Since these bonds are separated by two amino acid residues, endopeptidase cleavage at one site would not prevent the action of the enzyme on the neighboring site. Since the amino acid sequences C-terminal to the scissile bonds are all different, direct sequencing of the enzymic digests determines the sites and the relative extents of all three possible cleavages simultaneously. The percentages of the bonds cleaved (relative

Table I: Kinetic Constants of Recombinant Trypsin and Lys-189 Trypsin on Different Substrates

substrate	enzyme	K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m (min^{-1} mM^{-1})	Lys-189 trypsin/ trypsin[b]
Suc-Ala-Ala-Pro-Phe-AMC	trypsin	1.73 ± 0.18[a]	6.77 ± 0.51	3.9 ± 0.1	0.9
	Lys-189 trypsin	1.13 ± 0.18	4.05 ± 0.37	3.6 ± 0.3	
Suc-Leu-Leu-Val-Tyr-AMC	trypsin	0.09 ± 0.01	0.61 ± 0.05	6.5 ± 0.4	1.2
	Lys-189 trypsin	0.07 ± 0.02	0.52 ± 0.10	7.6 ± 1.4	
Suc-Phe-Gly-Ala-Leu-AMC	trypsin	>10	<0.005	<0.0005	>6000
	Lys-189 trypsin	1.62 ± 0.25	0.44 ± 0.14	0.3 ± 0.1	

[a] Standard errors. [b] Ratio of k_{cat}/K_m values.

to the intact molecule) were calculated from the yields of the PTH–amino acids. The results were calculated from the second Edman cycle since the first cycle frequently contains contaminating amino acids.

Bovine chymotrypsin (used in a 200 times smaller amount than the other enzymes) split peptide bonds typical of chymotrypsin (Green & Neurath, 1954; Hill, 1965). The relatively higher extent of the tyrosyl than the tryptophanyl bond cleavage might be due to the presence of Arg at the P_3' position relative to Tyr (M. Laskowski, personal communication). TPCK-treated bovine trypsin selectively cleaved the Tyr–Gly bond. The relatively significant hydrolysis by recombinant rat trypsin of the Arg–Pro peptide bond is quite unexpected since this bond is resistant to bovine trypsin. Further investigations are required to establish the degree of difference in the substrate specificity of rat and bovine trypsin. The specificity difference might be due to differential recognition of subsite residues. Lys-189 trypsin, in good agreement with the specificity data in Figure 7, cleaved both the tyrosyl and leucyl peptide bonds in LH-RH. Thus, recombinant trypsin and Lys-189 trypsin also have different substrate specificities on chymotrypsin substrates.

Kinetic Studies. Table I summarizes the K_m and k_{cat} values for trypsin and Lys-189 trypsin with three different substrates. In addition, neither enzyme hydrolyzed Suc-Ala-Ala-Pro-Val-AMC or Suc-Ala-Ala-Pro-Ala-AMC. Because of the different P_2–P_4 residue structures of the substrates, in Table I, differences in the kinetic constants for the same enzyme on various substrates are not necessarily solely due to the variations of the P_1 (Phe, Tyr, and Leu) residue. Kinetic constants for the enzymes on the same substrate, however, clearly reflect differences in the substrate binding and catalytic events. Interestingly, the replacement of Asp-189 with Lys did not change either the binding affinity as measured by K_m or the hydrolysis rates as measured by k_{cat} of the substrates with Phe or Tyr at position P_1 (Table I). k_{cat}/K_m values are practically the same for both enzymes on both substrates (Table I). In contrast, the leucyl substrate strongly differentiates between the chymotrypsin-like activities of the enzymes, since Lys-189 cleaves Suc-Phe-Gly-Ala-Leu-AMC with a comparable efficiency to those with which the phenylalanyl and tyrosyl substrates are hydrolyzed, while recombinant rat trypsin is apparently inactive on the leucyl substrate. This agrees with the results of the natural substrate studies: the Leu–Arg bond in LH-RH was not attacked by either bovine or rat trypsin. The chymotrypsin-like activities (k_{cat}/K_m) of both trypsin and Lys-189 trypsin on the phenylalanyl and tyrosyl substrates are about 3–4 orders of magnitude lower than the activities of bovine and rate trypsins on their most reactive synthetic substrates (Craik et al., 1985).

DFP at 10^{-4} M assay mixture completely abolished the chymotrypsin-like activities of recombinant and Lys-189 trypsin within 15 min. Thus, Ser-195 appears to be essential for the catalytic activity of both proteases. The active enzyme concentration determined with radiolabeled DFP agreed, within experimental error, with the protein concentration of the enzymes determined by amino acid analysis.

Curiously, TPCK, an alkylating agent for His-57, did not affect enzyme activity. The highest concentration of TPCK used was 0.6 mM, about one-third of the K_m value determined for the structurally homologous phenylalanyl substrate used in this experiment (see Table I). The phenylalanyl substrate (perhaps due to the involvement of residues P_2–P_4 in its binding) may fit the binding pocket more appropriately than TPCK. In contrast to TPCK, benzamide competitively inhibited the hydrolysis of the trypsin substrate D-Val-Leu-Arg-AFC and on the chymotrypsin substrate (Suc-Ala-Ala-Pro-Phe-AMC). The K_i values determined for the arginyl and phenylalanyl substrates were 5.2 ± 1.0 and 7.9 ± 1.4 μM, respectively. These values are practically identical with and close to the K_i (16.6 ± 2.0 μM) reported by Mares-Guia (1968) for the inhibition by benzamidine of bovine trypsin on benzoyl-DL-arginine *p*-nitroaniline. These data together with our finding that the Asp → Lys change at the base of the trypsin binding pocket affects the substrate specificity toward chymotrypsin substrates strongly suggest that the trypsin and chymotrypsin substrates use the same binding pocket in trypsin.

Discussion

We have replaced aspartic acid with lysine at position 189 in the substrate binding pocket of trypsin to test its role in determining substrate specificity. The crystal structure of bovine trypsin complexed with pancreatic trypsin inhibitor (Ruhlmann et al., 1973) reveals an electrostatic bond between Lys-15 of the inhibitor and Asp-189 of the enzyme. It has been assumed that a similar interaction helps to orient the lysyl or arginyl bonds of peptide substrates with the active-site residues. Computer graphic modeling of Lys-189 in trypsin suggested that a similar charge pair could be formed with the Asp or Glu side chains of the substrate and therefore Lys-189 trypsin might hydrolyze peptide bonds after Asp or Glu residues. Consistent with the presumed crucial role of Asp-189 in determining the specificity toward basic substrates, Lys-189 trypsin displayed no catalytic activity toward arginyl and lysyl substrates. However, Lys-189 trypsin exhibited no compensatory enzymic activity for Asp or Glu residues on either natural or synthetic substrates. An alternative structure also suggested by computer graphic modeling experiments provides a possible explanation of these results: the NH_3^+ group of Lys-189 may interact with two or three neighboring oxygen atoms of main-chain carbonyls near the binding pocket. Buried lysines have been found hydrogen bonded to oxygen atoms that bear no formal charge (Rashin & Honig, 1986). In this proposed configuration, the positively charged NH_3^+ group would then be inaccessible to the COO^- group of the substrate (see position b of the Lys side chain in Figure 4), and the base of the pocket would be nonpolar. This would make the binding and desolvation of any charged side chain thermodynamically

unfavorable. The stabilization energy associated with the binding of a charged ligand that results in the formation of a buried charge pair is estimated to be 4–5 kcal/mol (Fersht et al., 1985). If the charge pair is not formed, this energy could not be compensated for by the loss of bound solvent molecules in the specificity pocket nor by other favorable contacts between the protein and substrate. Thus, the specificity pocket of this structure (Figure 4, position b) would not be expected to recognize charged side chains (basic or acidic).

Our studies on a variety of substrates revealed that, like chymotrypsin, trypsin as well as Lys-189 trypsin hydrolyzes peptide bonds after certain hydrophobic residues, but at a dramatically lower rate. An inherent chymotrypsin-like activity of highly purified trypsin had been suggested by early careful kinetic studies (McFadden & Laskowski, 1956; Inagami & Sturtevant, 1960; Vajda & Szabo, 1976). Our work provides the first direct evidence for this chymotryptic activity, since recombinant trypsin expressed in bacteria cannot be contaminated by chymotrypsin. Furthermore, this substrate specificity is characteristically different from that of chymotrypsin. Tryptophan is the favored residue in the P_1 substrate position for chymotrypsin (Green & Neurath, 1954; Hill, 1965), but in our studies, tryptophanyl peptide bonds are not cleaved by trypsin. This is consistent with earlier observations on soybean trypsin inhibitor in which replacement by enzymic semisynthesis of the critical Arg-63 residue with Trp destroyed activity while replacement with Phe retained potent trypsin inhibitor activity (Kowalski et al., 1974).

The activity of trypsin toward phenylalanyl and tyrosyl substrates is not influenced by the Asp to Lys change at position 189. However, Lys-189 trypsin displays the additional unique property of cleaving leucyl bonds. This is consistent with the proposed b configuration of Lys-189 (in Figure 4): the increased hydrophobicity of the pocket caused by the extended side chain carbons of Lys presumably allows Leu side chains to bind. Since Lys-189 trypsin exhibits a different substrate specificity from trypsin or chymotrypsin, it can be considered a novel enzyme. Although the activity of Lys-189 trypsin is low when compared to trypsin, it is well within the range of the highly specific serine proteases such as urokinase (Lottenberg et al., 1981) and the blood coagulation serine proteases (Cho et al., 1984).

In spite of the high degree of structural homology of the substrate binding pockets of trypsin and chymotrypsin, the two enzymes exhibit dramatic differences in substrate specificity. This suggests that the specificity pockets are delicately tailored for their natural substrates. In contrast, subtilisin tolerates great changes at the base of the substrate binding pocket (Estell et al., 1986) and exhibits a correspondingly relaxed substrate specificity. The very low catalytic efficiency of trypsin and Lys-189 trypsin on hydrophobic (chymotrypsin-like) substrates is not due to the lack of interaction of the P_1 residues with the binding pocket as evidenced by the reasonable K_m values for these substrates. Instead, the low activity may be a result of a distorted positioning of the scissile bond with respect to Ser-195 of the catalytic triad. Thus, to modify trypsin to produce high catalytic activity toward chymotrypsin substrates may require a number of subtle changes in the structure of the substrate binding pocket.

ACKNOWLEDGMENTS

We thank Dr. S. Michaelis for providing the *E. coli* cell line SM 138, Dr. H. Inouye for providing the DNA encoding the alkaline phosphatase promoter and signal peptide, Dr. E. Goldsmith for helpful discussions on computer modeling of protein structures, and L. Spector for preparing the manuscript.

REFERENCES

Anderson, L. E., Walsh, K. A., & Neurath, H. (1977) *Biochemistry 16*, 3354–3360.

Bode, W., & Schwager, P. (1975) *J. Mol. Biol. 98*, 693–717.

Burnette, W. N. (1981) *Anal. Biochem. 112*, 195–203.

Cho, K., Takumi, T., Cook, R., Kisiel, W., Fujikawa, K., Kurachi, K., & Powers, J. (1984) *Biochemistry 23*, 644–650.

Cleland, W. W. (1979) *Methods Enzymol. 63*, 103–138.

Craik, C. S., Choo, Q.-L., Swift, G. H., Quinto, C., MacDonald, R. J., & Rutter, W. J. (1984) *J. Biol. Chem. 259*, 14255–14264.

Craik, C. S., Largman, C., Fletcher, T., Roczniak, S., Barr, P. J., Fletterick, R., & Rutter, W. J. (1985) *Science (Washington, D.C.) 228*, 291–297.

Dayringer, H., Tramontano, A., Sprang, S., & Fletterick, R. J. (1986) *J. Mol. Graphics 4*, 82–87.

Edmundson, A. B. (1965) *Nature (London) 205*, 883–887.

Estell, D. A., Graycar, T. P., Miller, J. V., Powers, D. B., Burnier, J. P., Ng, P. G., & Wells, J. A. (1986) *Science (Washington, D.C.) 233*, 659–663.

Fersht, A. R., Shi, J.-P., Knill-Jones, J., Lowe, D. M., Wilkinson, A. J., Blow, D. M., Brick, P., Carter, P., Waye, M. M. Y., & Winter, G. (1985) *Nature (London) 314*, 235–238.

Graf, L., & Cseh, G. (1968) *Acta Biochim. Biophys. Acad. Sci. Hung. 3*, 175–177.

Graf, L., Barat, E., Cseh, G., & Sajgo, M. (1971a) *Biochim. Biophys. Acta 229*, 276–278.

Graf, L., Bajusz, S., Patthy, A., Barat, E., & Cseh, G. (1971b) *Acta Biochim. Biophys. Acad. Sci. Hung. 6*, 415–418.

Green, N. M., & Neurath, H. (1954) in *The Proteins* (Neurath, H., & Bailey, K., Eds.) Vol. 2, pp 1512–1587, Wiley, New York.

Heinrikson, R. L., & Meredith, S. C. (1984) *Anal. Biochem. 136*, 65–74.

Hill, R. L. (1965) *Adv. Protein Chem. 20*, 37–107.

Inagami, T. & Sturtevant, J. M. (1960) *J. Biol. Chem. 235*, 1019–1023.

Inouye, H., Michaelis, S., Wright, A., & Beckwith, J. (1981) *J. Bacteriol. 146*, 668–675.

Kowalski, D., Leary, T. R., McKee, R. E., Sealock, R. W., Wang, D., & Laskowski, M., Jr. (1974) *Bayer-Symp. 5*, 311–324.

Kraut, J. (1977) *Annu. Rev. Biochem. 46*, 331–358.

Laemmli, U. K. (1970) *Nature (London) 227*, 680–685.

Liepnieks, J. J., & Light, A. (1974) *Anal. Biochem. 60*, 395–404.

Lottenberg, R., Christensen, U., Jackson, C. M., & Coleman, P. L. (1981) *Methods Enzymol. 80*, 341–361.

Mares-Guia, M. (1968) *Arch. Biochem. Biophys. 127*, 317–322.

Matsuo, H., Baba, Y., Nair, R. M. G., Arimura, A., & Schally, A. V. (1971) *Biochem. Biophys. Res. Commun. 43*, 1334–1339.

McFadden, M. L., & Laskowski, M., Jr. (1956) Abstracts of the 130th Meeting of the American Chemical Society, Washington, D.C., p 71C.

Rashin, A., & Honig, B. (1986) *J. Mol. Biol. 173*, 515–521.

Richmond, T. J. (1984) *J. Mol. Biol. 178*, 63–89.

Ruhlmann, A., Kukla, D., Schwager, P., Bartels, K., & Huber, R. (1973) *J. Mol. Biol. 77*, 417–436.

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

Schoellmann, G., & Shaw, E. (1963) *Biochemistry 2*, 252–255.
Sprang, S., Standing, T., Fletterick, R., Hamlin, R., Xuong, N. G., Rutter, W. J., & Craik, C. S. (1987) *Science (Washington, D.C.)* (in press).
Steitz, T. A., & Shulman, R. G. (1982) *Annu. Rev. Biophys. Bioeng. 11*, 419–444.
Stroud, R. M., Kay, L. M. & Dickerson, R. E. (1974) *J. Mol. Biol. 83*, 185–208.
Sweet, R. M., Wright, H. T., Janin, J., Chothia, C. H., & Blow, D. M. (1974) *Biochemistry 13*, 4212–4228.
Vajda, T., & Szabo, T. (1976) *Acta Biochim. Biophys. Acad. Sci. Hung. 11*, 287–294.

Evidence from Fourier Transform Infrared Spectroscopy for Polarization of the Carbonyl of Oxaloacetate in the Active Site of Citrate Synthase†

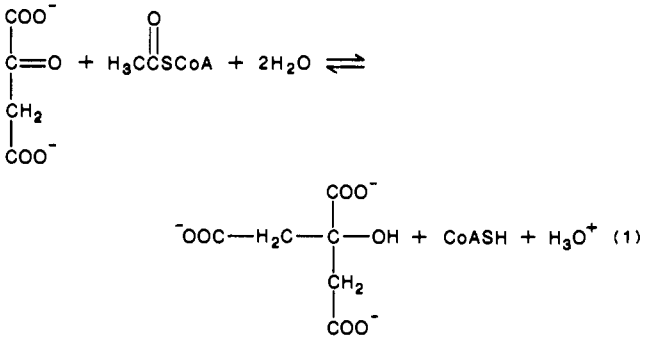
Linda C. Kurz* and George R. Drysdale

Department of Biological Chemistry, Division of Biology and Biomedical Sciences, Washington University School of Medicine, St. Louis, Missouri 63110

Received November 3, 1986; Revised Manuscript Received December 19, 1986

ABSTRACT: The infrared spectrum of oxaloacetate bound in the active site of citrate synthase has been measured in the binary complex and in the ternary complex with the acetyl coenzyme A (CoA) enolate analogue carboxymethyl-CoA. The carbonyl stretching frequency of oxaloacetate in binary and ternary complexes is found at 1697 cm^{-1}, a shift of 21 cm^{-1} to lower frequency relative to that of the free ligand. The line widths of the carbonyl absorption in enzyme complexes differ from that of the free ligand, decreasing from a value of 20 cm^{-1} for the free ligand to 10 cm^{-1} in the binary complex and 7 cm^{-1} in the ternary complex with carboxymethyl-CoA. The integrated absorbance of the carbonyl absorption in these enzyme complexes is significantly increased over that of the free ligand at the same concentration, increasing ~2-fold in the binary complex and ~3-fold in the ternary complex. These results indicate strong polarization of the carbonyl bond in the enzyme–substrate complexes and suggest that ground-state destabilization is a major catalytic strategy of citrate synthase.

Citrate synthase (EC 4.1.3.7) catalyzes the condensation of oxaloacetate (OAA) with acetyl coenzyme A (CoA) to form citrate (eq 1). The chemical mechanism is thought to involve



generation of the carbanion (enolate) of acetyl-CoA, which condenses with the carbonyl of OAA to form *S*-citryl-CoA as an intermediate (Eggerer, 1965; Weidman & Drysdale, 1979; Eggerer & Remberger, 1963; Bayer et al., 1981). It has been proposed that the carbonyl of OAA could interact with an electrophilic residue resulting in polarization of the C=O bond with substantial positive charge development at the carbonyl carbon (Srere, 1966). Polarization of the carbonyl would enhance the reactivity of the carbonyl carbon toward carbanion addition.

We have previously reported (Kurz et al., 1985) very large C-13 chemical shift changes (+6.8 ppm, downfield) for the carbonyl of OAA upon binding to the enzyme, which we have interpreted as evidence for polarization of the carbonyl of OAA in the active site. However, our interpretation of even these large changes in ^{13}C shifts as resulting from carbonyl bond polarization has been questioned (Malthouse, 1986). We now present independent evidence from Fourier transform infrared spectroscopy (FTIR) supporting the validity of our interpretation of the NMR results.

Materials and Methods

Crystalline citrate synthase was a product of Sigma Chemical Co., St. Louis, MO. Enzyme crystals were first dissolved in 50 mM tris(hydroxymethyl)aminomethane hydrochloride (Tris-HCl) buffer, pH 7.50. After extensive dialysis to remove $(NH_4)_2SO_4$, enzyme samples were concentrated to ~300 mg/mL (0.5 mL) by centrifugation in a CF25 Centriflo concentrator (Amicon Corp.). The sample was rediluted with 4.5 mL of 50 mM Tris-DCl, pD 7.90, in D_2O and reconcentrated to 0.5 mL, repeating this procedure 3 times to ensure complete exchange of solvent protons. The sample was removed from the membrane cone, which was rinsed with 0.5 mL of buffer, to obtain a final concentration of 1–3 mM active sites.

For experiments with the OAA–enzyme complex, sample preparation included a Sephadex G-25 purification step to remove small molecule contaminants (pyruvate, α-ketoglutarate, etc.). After concentration of the enzyme as described above, a 10% molar excess of OAA in D_2O was added to form the binary complex. The sample was applied to the top of a 2.8-mL column of Sephadex G-25 equilibrated in 50 mM Tris-DCl, pD 7.9 (in D_2O), and centrifuged briefly to

†Supported by National Institutes of Health Grant GM33851.

0006-2960/87/0426-2623$01.50/0