# Proteolysis of Native Proteins

## TRAPPING OF A REACTION INTERMEDIATE*

### Chenyi Wu, Duncan H. L. Robertson, Simon J. Hubbard, Simon J. Gaskell‡, and Robert J. Beynon§

*From the Department of Biomolecular Sciences and the ‡Michael Barber Centre for Mass Spectrometry, University of Manchester Institute of Science and Technology, Manchester M60 1QD, United Kingdom*

When limited proteolysis of the mouse major urinary proteins by trypsin was stopped by rapid denaturation of the proteinase, a covalent adduct of the two proteins was observed. The formation of this complex required active trypsin, was favored at low pH, and could be reversed by the addition of covalent or non-covalent trypsin inhibitors. Electrospray mass spectrometry of the complex demonstrated that it was an acyl-enzyme complex, formed after an unusual exopeptidase attack on the C-terminal-Arg-Glu-OH sequence by trypsin. The complex could sequester over 50% of the trypsin in a digestion mixture, and as anticipated, the protein was an effective trypsin inhibitor.

Proteolysis of native proteins is influenced by higher order structural features of the interaction between substrate and proteinase, in addition to the restrictions imposed by the primary specificity of the proteinase. Proteolysis can be limited, and the rate of digestion of peptide bonds with equivalent primary specificity seems to depend on several parameters, the most important of which is the capability for local flexibility of the polypeptide chain in the region of the scissile bond (reviewed recently in Ref. 1). Proteolytic enzymes can therefore be used as probes of flexible or exposed regions of a polypeptide chain, to monitor changes in such parameters after conformational shifts such as those elicited by ligand binding, and to generate functional subdomains by excision of accessible connecting regions (2, 3).

We have used limited proteolysis as a structural probe of major urinary proteins (MUPs).[1] MUPs are a group of closely related proteins that are secreted into mouse urine and that are members of the lipocalin class of proteins (4–6). They bind semiochemicals (7, 8) to effect their slow release (9, 10) or to protect them from, for example, atmospheric oxidation. We have become interested in the role of these proteins in ligand binding and release and in the ability of proteolysis to affect the loss of ligands by disruption of the structure of the native holoprotein. Further, an N-terminal MUP-derived hexapeptide may be important in communicating additional chemosignals

(11), and we have explored the role of proteolysis in the generation, or destruction, of this motif. As part of this investigation we have exposed MUPs to different proteinases and analyzed the products. Digestion by trypsin yielded an unusual pattern of behavior, specifically, the generation of substantial amounts of a high mass covalent complex of trypsin and MUP, which we have characterized.

## MATERIALS AND METHODS

Pooled urine was collected from unanesthetized mice by bladder massage from mature BALB/c males, housed together in groups of 8–16 using a 12-h dark/12-h light cycle and given free access to food and water. Immediately after collection, the urine was desalted on 5-ml "spun columns" of Sephadex G-25, previously equilibrated with water, and further concentrated with Centricon centrifugal concentrators (Amicon, Stonehouse, Gloucester, United Kingdom) with an $M_r$ = 10,000 cut-off. The resultant stock solutions contained MUPs at 40–60 mg/ml ($E_{280}^{1\%}$ = 6.1). This pool comprises a mixture of very similar allelic variants (12), but when we have resolved these variants by ion exchange chromatography, we have observed no difference in behavior during limited proteolysis[2] and we have routinely used the mixed pool. The mixture of proteins in the preparation differ in conservative substitutions of only four amino acids (12). This pool of very similar allellomorphs is referred to in the singular, as MUP, in this paper.

Trypsin (EC 3.4.21.4) from porcine pancreas (Sigma-Aldrich Chemical Co, Poole, Dorset, UK) was incubated with MUPs in a volume of 120 $\mu$l with 1:1 (by weight) enzyme/substrate ratio. Incubation time, temperature, and pH were varied in different experiments. The pH was buffered as follows: chloroacetate, pH 3.0; formate, pH 4.0; benzoate, pH 5.0; maleate, pH 6.0; Hepes, pH 7.0–8.0; borate, pH 9.0–10.0; and CAPS, pH 11.0. All buffers were 20 mM, and the ionic strength was adjusted to 0.1 M with NaCl. Buffers were thermodynamically corrected for temperature and ionic strength effects (http://www.bi.umist.ac.uk/buffers.html). Where other reactants, such as soybean trypsin inhibitor and TLCK were included, they were dissolved in buffer (10 mg/ml) or methanol (5 mM), respectively. Carboxypeptidase B was added from a 10 mg/ml stock as supplied by the manufacturer (Boehringer Mannheim, Lewes, East Sussex, UK).

Enzyme reactions were stopped by addition of trichloroacetic acid to a final concentration of 10% (w/v). Precipitated proteins were pelleted by centrifugation for 5 min at 12,500 rpm. The pellets were washed 2–3 times with 200 $\mu$l of diethyl ether to remove excess trichloroacetic acid, resuspended in SDS-PAGE sample buffer, and heated at 100 °C for 5 min prior to separation by SDS-PAGE under reducing conditions in 17.5% acrylamide gels (13). For quantitative analysis, Coomassie Blue-stained gels were scanned at 300 pixels per inch in an 8-bit grayscale and analyzed using Quantifier software (Phoretix International, Newcastle, UK). The relationship between band volume and amount of protein was linear over the range used in the experiment 0–100 $\mu$M MUP ($r$ = 0.98).

In some experiments, proteins separated by SDS-PAGE were transferred to a 0.45-$\mu$m nitrocellulose membrane (Schleicher and Schuell, Kingston-upon-Thames, Surrey, UK) by electroblotting overnight at a constant 30 V in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol, pH 8.3). The membrane was then incubated for 1 h in a

[1] The abbreviations used are: MUP, major urinary protein; PAGE, polyacrylamide gel electrophoresis; ESI, electrospray; FPLC, fast protein liquid chromatography; TLCK, $N^\alpha$-*p*-tosyl-L-lysine chloromethyl ketone; CAPS, 3-(cyclohexylamino)propanesulfonic acid.

[2] C. Wu, D. H. L. Robertson, S. J. Hubbard, S. J. Gaskell, and R. J. Beynon, unpublished observations.

1:1000 rabbit anti-MUP serum solution in phosphate-buffered saline/Tween (10 mM phosphate, 150 mM NaCl, 0.05% Tween (v/v), pH 7.3) and visualized by a 1-h incubation with 1:3000 goat anti-rabbit IgG conjugated to alkaline phosphatase (Bio-Rad, Hemel Hempstead, Herts, UK) followed by a further incubation with a mixture of 5-bromo-4-chloro-3-indolyl phosphate and nitro blue tetrazolium (Sigma). Proteins from SDS-gels were also blotted on to polyvinylidene difluoride membrane in transfer buffer (10 mM CAPS, 10% methanol, pH 11.0). After a brief staining with 0.1% (w/v) Coomassie Blue in 40% (v/v) methanol, 1% (v/v) acetic acid, the desired protein bands were cut from the membrane and subjected to 5 cycles of Edman degradation on an Applied Biosystems 476A protein sequencer.

For electrospray mass spectrometry (ESI/MS), the pellets recovered after proteolysis were redissolved with gentle heating for a few seconds on a boiling water bath in a 100-$\mu$l ESI/MS mobile phase (50% (v/v) acetonitrile, 0.1% (v/v) formic acid). Spectra were acquired on a Quattro triple quadrupole mass spectrometer (Micromass, Manchester, UK) fitted with an electrospray ionization source and upgraded to Quattro II specifications. The instrument was tuned and calibrated with a 20 pmol/$\mu$l solution of horse heart myoglobin made up in 50% (v/v) aqueous acetonitrile, 0.1% (v/v) formic acid. Samples were introduced into the source at 10 $\mu$l/min. Scanning was from $m/z$ 950 to 1500 at 10 s/scan. Twenty to thirty scans were averaged to produce the final spectra. Data acquisition, instrument control, and post-run processing were by the VG MassLynx software (Micromass). Mass assignments were made from maximum entropy-processed data (14). The protein masses thus calculated were within 2 Da of the cDNA-inferred masses (12), and repeated measurements over a 6-month period of three of the proteins yielded mass assignments of 18644.2 ± 0.78 Da, 18694 ± 0.84 Da, and 18708.2 ± 1.75 Da (mean ± S.D., $n$ = 10 for each protein).

Anion-exchange chromatography was on an FPLC system (Amersham Pharmacia Biotech) fitted with a Resource-Q column ($V_t$ = 6 ml). The column was equilibrated with 20 ml of 50 mM formate buffer, pH 5.0, before application of 500 $\mu$l of the tryptic digest of MUP (1 mg/ml MUP, 7 mg/ml trypsin, pH 4.0, incubated for 30 min at room temperature in 50 mM formate, pH 5.0). The column was developed using a 100-ml linear salt gradient (0–1 M NaCl) at a flow rate of 6.0 ml/min. Protein fractions were collected and reconcentrated with Centricon (Amicon) centrifugal concentrators.

Anhydrotrypsin was prepared by a published method (15) with some modifications, and purified by affinity chromatography. Porcine pancreatic trypsin (200 mg) and phenylmethylsulfonyl fluoride (70 mg), dissolved in 100 ml of 50 mM Tris/HCl, pH 7.0, containing 20 mM CaCl$_2$, were incubated at room temperature for 30 min with occasional additions of 20 mM NaOH to maintain the pH at 7.0. Trypsin activity was monitored throughout the incubation period. The reaction was stopped by adjusting the pH to 3.0 with 1 M HCl. After a brief centrifugation, the supernatant was dialyzed overnight against 3 liters of 1 mM HCl at 4 °C. The solution of phenylmethylsulfonyl trypsin thus obtained was brought to 50 mM KOH at 0 °C and incubated for 12 min, then adjusted to pH 5.0 with formic acid to stop the reaction. The mixture was finally adjusted to pH 8.0 by the addition of 2 M Tris/HCl. Affinity chromatography of the anhydrotrypsin was performed on immobilized soybean trypsin inhibitor (Pierce & Warriner, Chester, UK) packed in a 5-ml spun column. Typically 2 ml of gel was used with a capacity of 20 mg of trypsin per ml of gel. The column was equilibrated with 20 ml of 50 mM Tris/HCl, pH 8.0, before application of 10 ml (20 mg) of anhydrotrypsin. The bound material was washed with 20 ml of the same buffer and eluted with 3 ml of 100 mM sodium chloroacetate buffer, pH 2.0. The anhydrotrypsin thus purified was concentrated/desalted with centrifugal concentrators. The final concentration was estimated spectrophotometrically ($E_{280}^{1\%}$ = 14.7). Tryptic activity at 25 °C toward Bz-Arg-4-nitroanilide was monitored by following the increase in absorbance at 405 nm. MUP, MUP$^{1/9-161}$, and the synthetic peptides, SLQARE and SLQAR, were tested for trypsin inhibitory properties at low pH in a total volume of 1 ml of assay solution consisting of 20 mM formate, pH 4.0, 0.2 mM $N^\alpha$-benzoyl-DL-arginine-$p$-nitroanilide, 2.1 $\mu$M trypsin. The concentrations of potential inhibitors ranged from 0 to 25 $\mu$M.

## RESULTS AND DISCUSSION

Purified MUPs comprise over 98% of the proteins in mouse urine and were used after desalting but without further purification. This preparation comprises a mixture of MUPs with a high level of sequence identity, differing in only 4 of 162 amino acids (9), each with a molecular mass of approximately 18–19 kDa. When the MUP mixture (referred to subsequently in the singular) was treated with trypsin, a product band (MUP′),
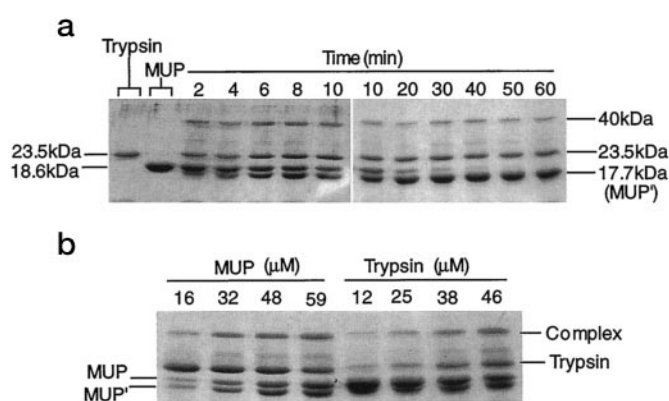


FIG. 1. **Formation of a high mass complex during proteolysis of MUP by trypsin.** *a*, trypsin (50 $\mu$M) was incubated with MUP (65 $\mu$M) in 20 mM Hepes buffer, pH 7.5, for between 2 and 60 min. At various times, samples were removed, and the reaction was terminated by addition of trichloroacetic acid to a final concentration of 10% (w/v). The proteins in the reaction mixture were then recovered by centrifugation, residual acid was washed away with diethyl ether, and the reaction mixture was analyzed on nonreducing SDS-PAGE. *b*, the analysis was conducted in essentially the same fashion as in *a* except that the time was fixed at 5 min and the concentrations of MUP or trypsin were varied. In the first experiment, trypsin was maintained at 50 $\mu$M, and MUP was varied from 16 to 60 $\mu$M. In the second experiment, MUP was maintained at 65 $\mu$M, and trypsin was varied from 12 to 45 $\mu$M. MUP′ refers, in this and subsequent figures, to the mixture of products of tryptic attack on MUP at Arg-8 and Arg-161.

approximately 1 kDa smaller than the MUP, was formed quite rapidly and was resistant to further proteolysis (Fig. 1*a*). Thus, proteolysis was limited and was consistent with removal of short peptide fragments from the N and/or C termini of the molecule. However, SDS-PAGE analysis also demonstrated an unexpected outcome of the proteolytic reaction. Specifically, additional bands, predominantly in the region of 40 kDa and to a lesser extent at 32 kDa, appeared on the gel. These bands were not present in the starting materials and were not formed when the two proteins were incubated singly under identical conditions.

In this proteolytic reaction, the digestion was stopped by addition of trichloroacetic acid to a final concentration of 10% (w/v). When the reaction was stopped using SDS-PAGE sample buffer, without acid precipitation, no higher molecular weight bands were evinced. These bands were present in the presence or absence of reducing agents during SDS-PAGE, precluding the possibility of disulfide-linked complexes. The amount of the high mass complexes that were formed was dependent on the concentration of both trypsin and MUP (Fig. 1*b*).

The precise pattern of limited proteolytic attack on MUP was defined by electrospray mass spectrometry. We have demonstrated previously that ESI/MS combined with maximum entropy data analysis is a powerful method of MUP characterization and is able to identify several gene products that are consistent (within 2 Da) with cDNA-predicted molecular masses (12, 16). ESI/MS analysis of the undigested MUP mixture identified three masses, 18646, 18694, and 18708 Da, each of which is consistent with a different, but intact, MUP in urine (12). After tryptic digestion, the overall distribution of peaks remained the same but the pattern had shifted down the mass scale, to 17699, 17746, and 17760 Da, by measured decrements of 948, 947, and 948 Da, respectively (Fig. 2). The mass decrements and SDS-PAGE data are thus consistent and suggest the removal of an N-terminal or a C-terminal peptide. However, the loss of approximately 948 Da cannot be explained by a simple hydrolytic release of a C- or N-terminal peptide, generated by cleavage at the first or last tryptic sites in the sequence. Specifically, cleavage at the first tryptic site in the sequence:
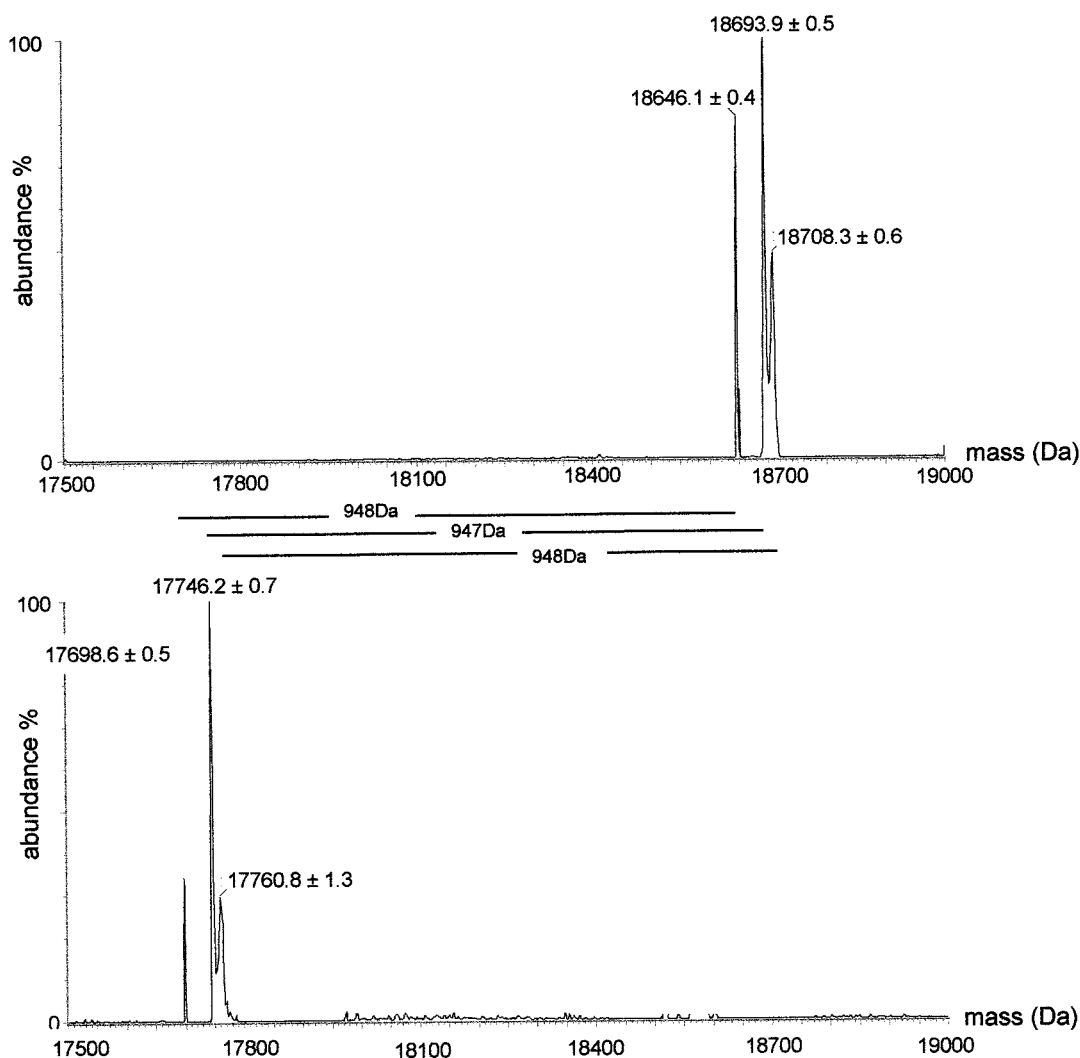
FIG. 2. **Limited proteolysis of MUPs by trypsin.** A mixture of MUPs derived from a single mouse strain (total MUP concentration = 65 $\mu$M) was incubated with trypsin (50 $\mu$M) for 120 min in 20 mM Hepes, pH 7.5. At the end of this time, the reaction was terminated by precipitation with trichloroacetic acid, as described under "Materials and Methods." The reaction mixture was then analyzed by electrospray mass spectrometry, using maximum entropy data processing (14). *Panel a* is the transformed spectrum of the undigested mixture. *Panel b* is the mass spectrum after digestion with trypsin. The estimated errors in the mass assignments are derived from maximum entropy processing of single accumulations of ESI/MS data.

NH$_2$-EEASSTGR $\downarrow$ NFNVEF . . . would yield new masses lower by 817 Da. An additional loss of 129 Da, provided by a second hydrolysis at the C-terminal end of the molecule, . . . CLQAR$^{161}$ $\downarrow$ E$^{162}$ yields a loss of 946 Da. N-terminal sequencing would have only revealed the new N terminus, and the CLQAR$^{161}$ $\downarrow$ E$^{162}$-OH cleavage, an exopeptidase attack, would not normally be considered a cleavage site for trypsin. The product of the N- and C-terminal proteolysis on the same substrate molecule is designated as MUP$^{9-161}$. Both the N and C termini of MUP are flexible, and the C terminus cannot be visualized in the crystal structure (17). Proteolytic attack at the extrema is therefore explicable in terms of accessibility of the termini of the protein, although the exopeptidase attack was unexpected.

The 40-kDa band was also unexpected, as there were no species of that size, or larger, in the initial incubation mixtures. When probed by Western blotting with a MUP polyclonal antibody it was evident that the 40-kDa species contained MUP (results not shown). The 40-kDa band was also blotted onto polyvinylidene difluoride membranes and subjected to automated Edman degradation. Three sequences could be identified (*a*) IVGGY . . . , at 6–8 pmol; (*b*) EEA . . . , at 4 pmol; and (*c*) NFNVE . . . , at 4 pmol. The first is attributable to trypsin, the

second to the N-terminal sequence of unproteolyzed MUP, and the third to the N-terminal sequence of the MUP$^{9-\text{C terminus}}$. Thus, the complex contained both MUP and trypsin, and the amounts of each protein were consistent with a 1:1 stoichiometry, as predicted by the mass of the complex.

The 40-kDa complex was stable to the conditions of SDS-PAGE and was likely therefore to be a covalent complex of trypsin and MUP (referred to hereafter as T-MUP) which is trapped by the virtually instantaneous denaturation by trichloroacetic acid and concomitant inactivation of trypsin. If one or the other protein components was first denatured with trichloroacetic acid before addition of the second protein, the complex was not formed. Further, the appearance of two MUP-derived N-terminal sequences in the complex suggested that the complex could also be formed by MUP$^{9-\text{C terminus}}$. To test this directly, the tryptic product MUP$^{9-\text{C terminus}}$ was prepared by tryptic digestion and subsequently purified and separated from trypsin and unproteolyzed MUP on anion-exchange chromatography (MonoQ FPLC). This product (MUP$^{9-\text{C terminus}}$) was then reincubated with trypsin. Under these circumstances, the T-MUP complex was formed once again (results not shown).
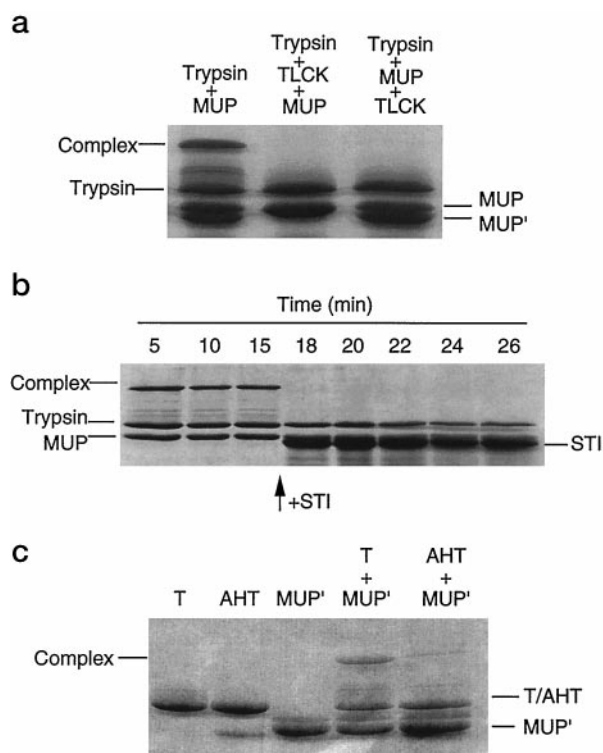
If trypsin was preincubated with the irreversible inhibitor TLCK (Fig. 3a), no complex could be generated, demonstrating

the requirement for active enzyme. TLCK alkylates the active site histidyl residue and reduces the nucleophilicity of the active site seryl residue. TLCK-modified trypsin is unable to bind to immobilized peptides with a C-terminal argininyl residue (18), consistent with some element of steric hindrance of the complex formation as well. In a further experiment, MUP was digested with trypsin for a fixed period, after which time a stoichiometric excess of soybean trypsin inhibitor was added. The complex could not be detected after addition of the soybean

trypsin inhibitor (Fig. 3*b*). Thus the strength of the interaction between the soybean inhibitor and trypsin is sufficient to sequester trypsin, at the expense of the T-MUP complex. This is good evidence for a reversible but covalent association of proteinase and substrate. The need for catalytically active trypsin and the role of the active site nucleophile were further confirmed by experiments in which MUP was incubated with anhydrotrypsin (Fig. 3*c*). Under such conditions, the complex was not formed.

The formation of the T-MUP complex was strongly dependent on pH (Fig. 4*a*). Trypsin and MUP were reacted in buffers from pH 4 and pH 11 at a constant ionic strength. Proteolysis of the MUP, leading to the formation of $MUP^{9-161}$, was evident at pH values of 6 and above, in accordance with the known pH optimum of trypsin. By contrast, complex formation was most evident at low pH values, and as the pH of the incubation increased, less of the complex was visible. At pH 4.0, approximately half of the MUP and trypsin was sequestered into the complex, as estimated by densitometry (Fig. 4*b*). The complex is formed preferentially at low pH, under conditions in which the hydrolytic activity of trypsin would be expected to be min-



FIG. 3. **Effect of inhibition of trypsin on complex formation.** *a*, trypsin (50 $\mu$M) was pretreated with the irreversible inhibitor TLCK for 30 min in 20 mM Hepes buffer, pH 7.6. The proteinase was fully inhibited, as confirmed by catalytic activity toward Bz-Arg-4-nitroanilide (see "Materials and Methods"). Subsequently, the inhibited trypsin was incubated with MUP (conditions as in Fig. 1), and the complex formation was assessed as described previously (see legend to Fig. 1). In a further experiment, uninhibited trypsin was incubated with MUP, and after a 5-min incubation, TLCK was added to a final concentration of 50 $\mu$M. After 30 min, the sample was precipitated by trichloroacetic acid and analyzed by SDS-PAGE. *b*, trypsin (65 $\mu$M) was incubated with MUP (40 $\mu$M) for up to 15 min, at which point soybean trypsin inhibitor (*STI*, 75 $\mu$M) was added to the reaction mixture. Samples were removed before and after addition of inhibitor and analyzed for complex formation as described previously (see legend to Fig. 1). *c*, in a separate experiment, MUP′ ($MUP^{9-161}$, 80 $\mu$M) was incubated with trypsin (*T*, 75 $\mu$M) or anhydrotrypsin (*AHT*, 75 $\mu$M) in 20 mM formate buffer, pH 4.0, $I$ = 0.1 M, for 5 min before the reaction was terminated by precipitation with trichloroacetic acid and analyzed by SDS-PAGE.
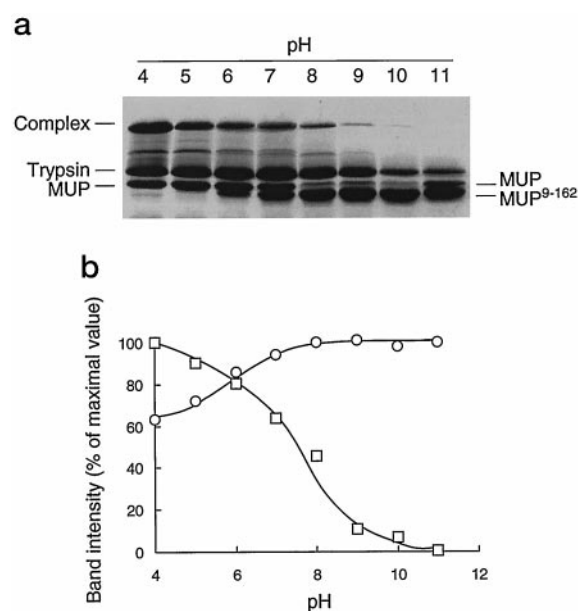


FIG. 4. **pH dependence of complex formation.** *a,* trypsin (70 $\mu$M) and MUP (85 $\mu$M) were incubated for 5 min in a series of buffers at different pH values, all adjusted with NaCl to a constant ionic strength of 0.1 M (see "Materials and Methods"). The reaction was terminated with trichloroacetic acid, and the products were analyzed for complex formation by SDS-PAGE as described previously. *b*, the gel was scanned and analyzed by densitometry as described under "Materials and Methods." The band volumes for the complex (*open squares*) and the total of MUP and MUP digestion products (*open circles*) were plotted as a function of pH, expressed as a percentage of the maximal volume.

TABLE I
*Effect of pH on tryptic proteolysis of MUPs*

Trypsin (70 $\mu$M) and MUP (85 $\mu$M) were incubated for 5 min in formate (pH 4.0), Hepes (pH 7.0), or CAPS (pH 11.0) buffers, adjusted with NaCl to a constant ionic strength of 0.1 M (see "Materials and Methods"). The reaction was terminated by precipitation with trichloroacetic acid, and the MUP digestion products were analyzed by electrospray mass spectrometry.

| | Trypsin digestion | | |
| --- | --- | --- | --- |
| | pH 4.0 | pH 7.0 | pH 11.0 |
| Control | | 18690.7± 0.7Da | |
| + trypsin | 18563.5 ± 0.4 Da | 18562.9± 0.4Da | 17745.0± 0.3Da |
| | | 17745.9± 0.6Da | |
| ΔMass | −127.2 Da | −127.8Da | −945.7Da |
| | | −944.8Da | |
| Interpretation | EEASSTGRNFN. . . . . . QAR ↓ E | EEASSTGRNFN . . . . . . QAR ↓ E | EEASSTGR ↓ NFN. . . . . . QAR ↓ E |
| | | EEASSTGR ↓ NFN . . . . QAR ↓ E | |

imal. Indeed, the lack of digestion of MUP at low pH (below pH 6.0) is evident from the SDS-PAGE analysis. However, electrospray mass spectrometry of the digestion reactions at three pH values indicated that although the N-terminal endopeptidase cleavage was suppressed at low pH values, the C-terminal exopeptidase reaction still took place (Table I). Formation of the complex between MUP or MUP[9-161] and trypsin seems to preclude an intermediate derived directly from, and as an obligatory outcome of, the hydrolysis of the N-terminal octapeptide. Another possible explanation for complex formation would be a transpeptidation reaction, in which MUP[9-161] was
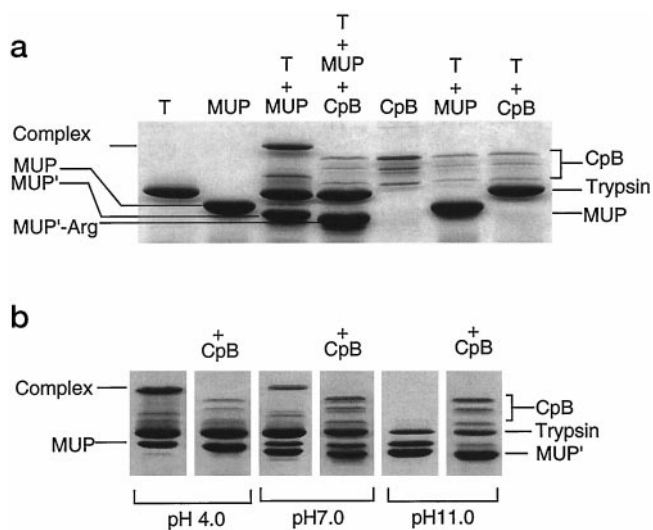


Fig. 5. **Effect of Arg-161 on complex formation.** *a*, trypsin (*T*, 70 $\mu$M) and MUP (85 $\mu$M) were incubated in 50 mM Hepes buffer, pH 7.5, for 5 min, in the presence or absence of 0.13 $\mu$g/ml carboxypeptidase B (*CpB*). Control incubations include carboxypeptidase B incubated alone and MUP incubated with carboxypeptidase. The reaction was terminated and analyzed for complex formation as described previously. *b*, in separate experiments, incubations of MUP and trypsin, in the presence of carboxypeptidase B, were conducted at pH 4.0, pH 7.0, and pH 11.0, all at a constant ionic strength of 0.1 M (see "Materials and Methods").
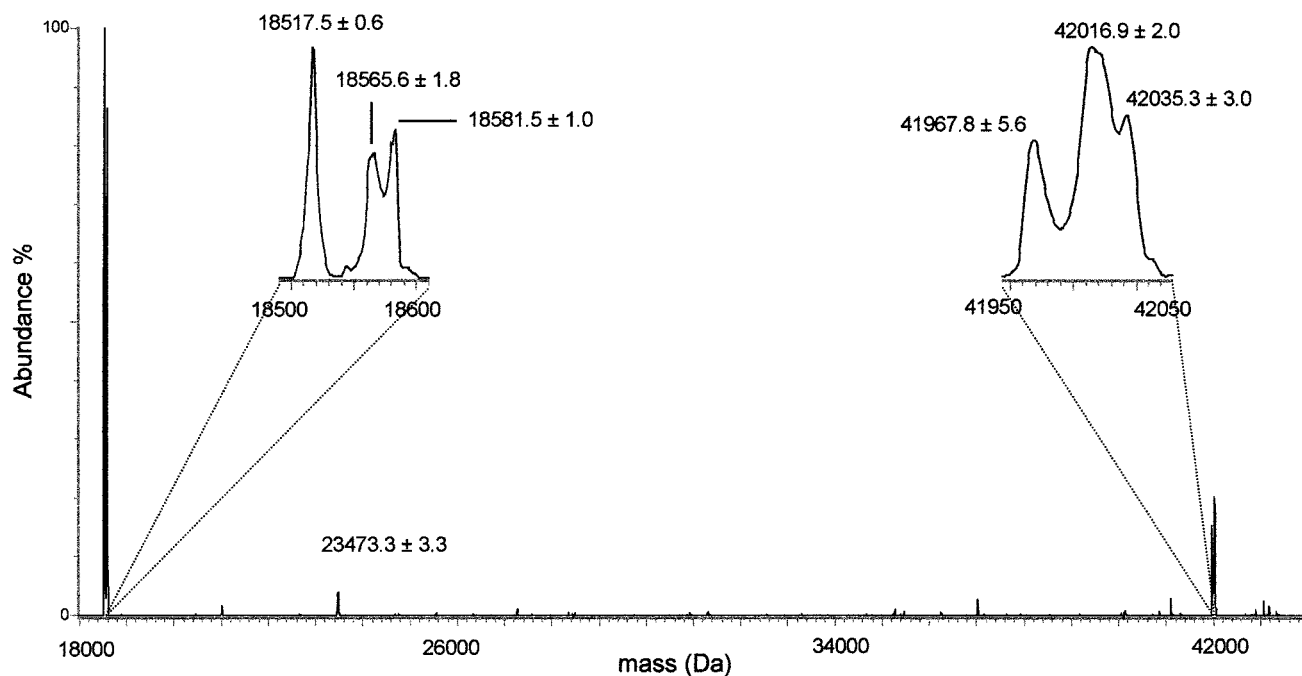
transferred to the N terminus of a second molecule of trypsin or MUP, which acted as an acceptor. This however is disproved by recovery of both the intact N-terminal trypsin sequence and the MUP sequence from the complex. Taken together, the data imply that the 40-kDa product is a stoichiometric, probably acyl-enzyme, complex between trypsin and MUP.

The formation of the complex seemed therefore to be associated with the unusual cleavage of the C-terminal glutamate residue. The role of the C-terminal, exposed Arg-161 was elaborated by an additional experiment in which MUP[9-161] was treated with carboxypeptidase B to remove the C-terminal Arg-161 before incubation with trypsin (Fig. 5). In this instance, formation of the complex was completely abolished, irrespective of the pH at which the reaction was conducted. This experiment also suggested that the complex was formed reversibly, since a C-terminal argininyl residue would not be susceptible to exopeptidase attack if the MUP was covalently and irreversibly linked via its C terminus to the catalytic seryl residue of trypsin. Since the complex is completely destroyed by incubation with the carboxypeptidase, it is reasonable to assume that the binding between proteolyzed MUP (with C-terminal arginine intact) and trypsin is rapid, favoring significant quantities of the complex.

We have made repeated attempts to observe the intact complex by electrospray mass spectrometry. Previous electrospray mass spectrometric analyses of trypsin (19, 20), however, have indicated significant heterogeneity, and it has proved impossible to purify a trypsin variant that yields a single peak. Trypsin was incubated with a mixture of three MUPs at low pH, and the complex was trapped by precipitation before redissolution and analysis by ESI/MS. The three MUPs (starting masses 18646, 18694, and 18708 Da) were all completely converted into products (18518, 18565, and 18581 Da) that indicated trypsin-mediated exopeptidase cleavage of the C-terminal Glu-162. In the 42-kDa region of the mass spectrum, three peaks were also identified (Fig. 6). Each of these was within 4 Da of the predicted mass of the (MUP isoform[1-161]) + trypsin − 17 Da. The masses of the three peaks are therefore consistent with formation of an acyl-enzyme complex between MUP[1-161] and



FIG. 6. **Electrospray mass spectrometry of the T-MUP complex.** Trypsin (320 $\mu$M) was incubated with MUP (80 $\mu$M) in 20 mM formate buffer, pH 4.0, for 5 min. The reaction was terminated with trichloroacetic acid, and the reaction products were analyzed by electrospray mass spectrometry (see "Materials and Methods").

trypsin.

Quantitative estimates of the ratio of complex to unbound MUPs could not be derived from the ESI/MS data, but SDS-PAGE analysis indicated that as much as 40% of a stoichiometric mixture of trypsin and MUP is sequestered as the complex (Fig. 4*b*). Why such a large proportion of trypsin should be sequestered into this T-MUP complex, particularly at low pH values, is unclear. It might be argued that the low pH conditions act in two ways. First, suppression of the C-terminal carboxylate charge might favor enhanced binding of the C-

terminal arginine-exposed peptide to the substrate-binding site of trypsin. Second, the lowered concentration of the nucleophilic hydroxyl ion would favor accumulation of a covalent catalytic intermediate that would normally be attacked by this species. This hypothesis has also been invoked to explain the effect of lowered pH on the enhanced peptide synthesis activity of trypsin (18). Indeed, the ability of trypsin to catalyze peptide bond synthesis is strong support of the existence of such acyl or tetrahedral complexes. However, these reaction intermediates are generally thought to be transient, and the accumulation that is seen in the present example is surprising.

The sequestration of large amounts of trypsin into the T-MUP complex implies that the MUPs should be capable of acting as trypsin inhibitors. This was tested by monitoring the ability of MUPs to alter the rate of hydrolysis of the chromogenic substrate Bz-Arg-4-nitroanilide at low pH values, where the complex is particularly favored (Fig. 7). As anticipated, MUPs are weak trypsin inhibitors, but we do not ascribe any physiological significance to this observation.

The unexpected accumulation of large quantities of a proteinase·substrate complex raised the possibility that the structure of MUP was an important contributory factor, and that the three-dimensional structure of the protein favors the accumulation or stabilization of the covalent intermediate. The structures of the C and N termini of MUPs are not visible by x-ray crystallography (17) and are generally held to be mobile segments which would be expected to be accessible to proteolytic attack (1, 2). Indeed, simple molecular modeling experiments have shown that a plausible covalent complex can be formed between trypsin and the MUP C terminus without introducing any contentious intra- or intermolecular stereochemistry. Furthermore, this can be achieved without breaking the Cys-64–Cys-156 disulfide bond. To test the ability of trypsin to bind unstructured C-terminal peptides, we synthesized two peptides SLQARE and SLQAR. The first is the full C-terminal hexapeptide sequence of MUP, with the substitution
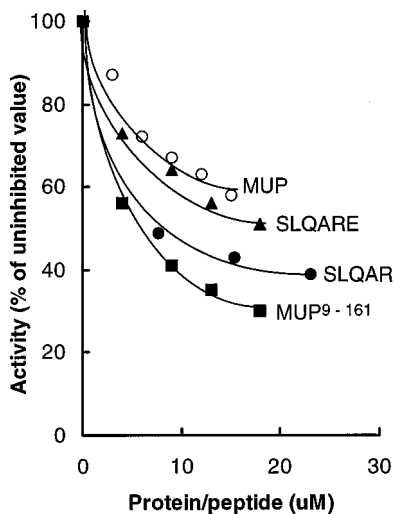


FIG. 7. **Inhibition of amidolytic activity of trypsin by MUPs and C-terminal peptides.** MUP and its tryptic product were tested for the ability to inhibit the trypsin-catalyzed hydrolysis of Bz-Arg-4-nitroanilide at pH 4.0, in 20 mM formate buffer. In separate experiments, peptides derived from the C-terminal sequence of MUP were also tested for inhibitory activity. MUP, *open circles*, MUP$^{9-161}$, *filled squares*; SLQAR, *closed circles*; SLQARE, *filled triangles*.
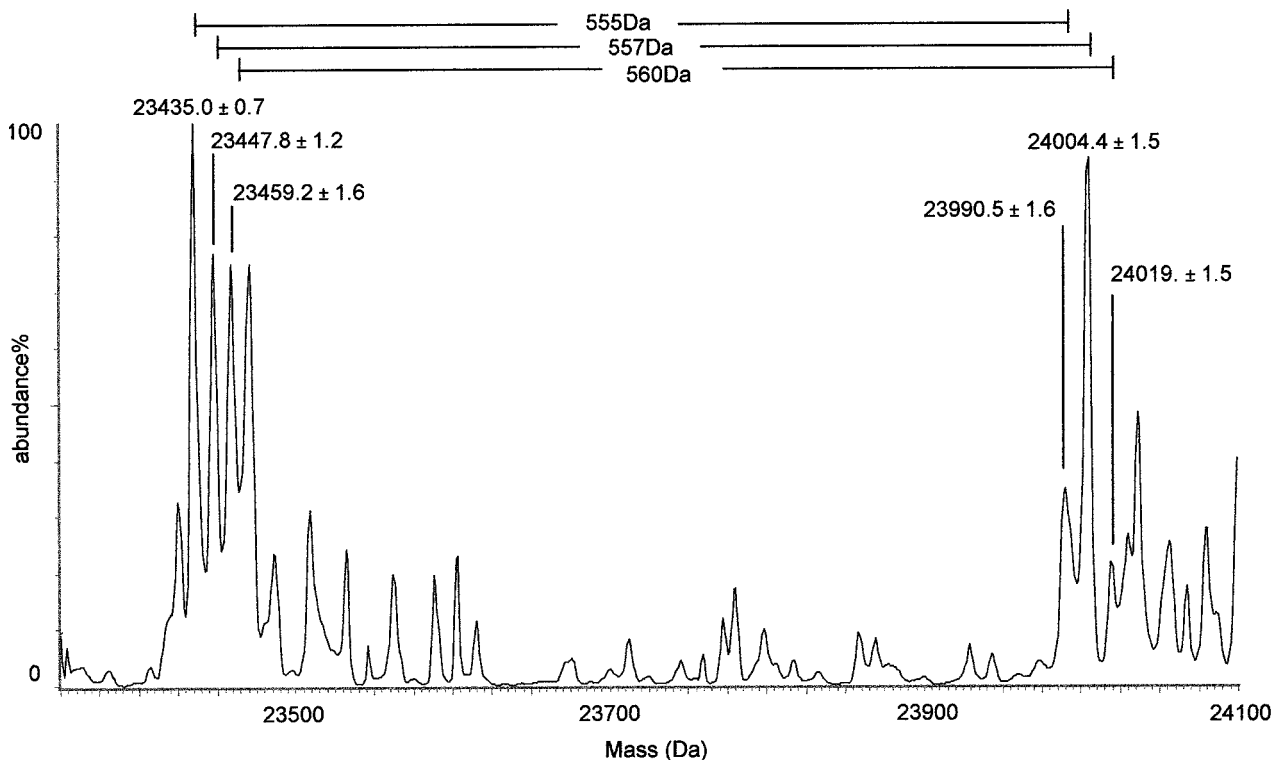


FIG. 8. **Electrospray mass spectrometry of the peptide-MUP complex.** Trypsin (35 μM) and the peptide SLQARE (600 μM) were incubated at room temperature in a solution of acetic acid titrated to pH 4.0 for 10 min, prior to analysis by electrospray mass spectrometry. The incubation mixture was introduced directly into the source, without additional organic solvent.
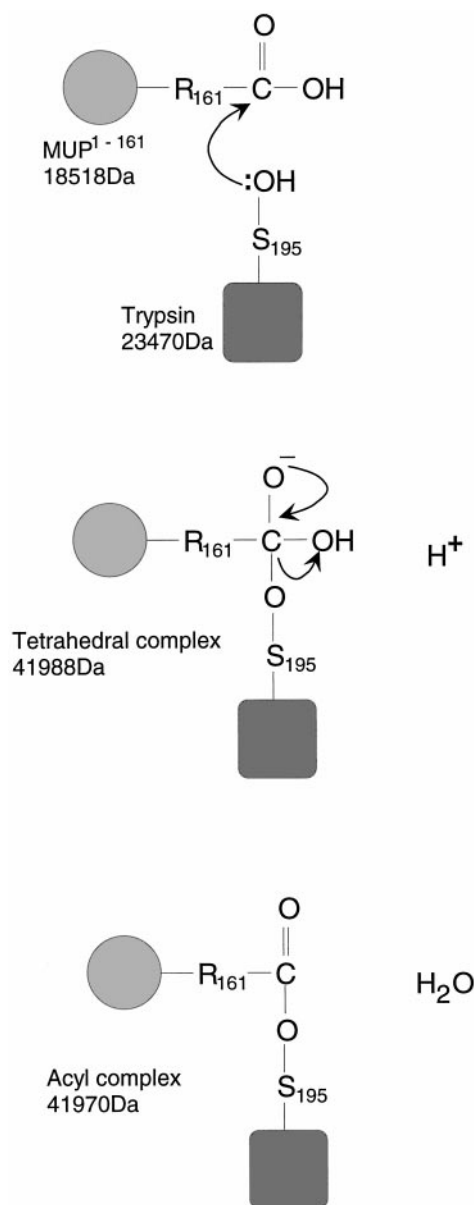
FIG. 9. **Schematic of the route of formation of the T-MUP complex.**

tides. Complex formation does not appear to be critically dependent on the three-dimensional structure of MUP, although the protein/protein interaction might enhance the amount of complex that is formed (Fig. 9).

It has long been recognized that trypsin is able to bind peptides that have an exposed C-terminal lysine or arginine residue. Indeed, this is the basis of the methods of recovery of such peptides using anhydrotrypsin as a trapping matrix (20, 21). It is generally held that the association reflects non-covalent binding of the C-terminal residue to the specificity pocket. There has not, to our knowledge, been any indication previously that the interaction with intact trypsin might be covalent and reflect a partial reversal of the hydrolytic reaction. Obviously, the acyl or tetrahedral complex must be formed with the carbonyl group of one of the substrates when trypsin functions as a peptide synthetase; the amino group of the N-terminal donor then attacks this complex (reviewed in Ref. 22). However, this reaction is relatively weak, and significant accumulation of acylated enzyme would be surprising. It remains to be seen whether such complexes involving C-terminal basic residues are a generic feature of trypsin catalysis or even whether the phenomenon can be extended to other serine proteases. In this respect, the recent structural elucidation of an acyl-enzyme complex between porcine pancreatic elastase and a heptapeptide, trapped at low pH values (23), is particularly relevant, as is the formation of SDS-stable complexes between serine proteases and serpins (24). However, in the latter example, the evidence pointed to covalent links additional to that generated in the acyl-enzyme complex. In this respect, the data presented here are simpler, as they clearly identify the role of the active site seryl residue.

There may be new applications that can make use of the covalent trapping, should this association manifest itself as a common occurrence. For example, recovery of peptides that have a C-terminal argininyl residue would be enhanced at low pH values where adventitious tryptic endoproteolysis would be avoided. Such a tool may lend itself to methods of peptide analysis by mass spectrometry where, for example, a combined trypsin-peptide complex could be recovered. An endoproteolytic treatment at neutral pH might be followed by an acidification, in which case the only free peptide would be the C-terminal peptide. This additional identification tag might enhance the identification of the parent protein, which, together with new methods of N-terminal analysis (25), could reduce the complexity of the identification problem in proteome analysis.

REFERENCES

1. Hubbard, S. J. (1998) *Biochim. Biophys. Acta* **1382,** 191–206
2. Fontana, A., Polverino de Laureto, P., De Fillipis, V., Scaramella, V., and Zambonin, M. (1997) *Fold. Design* **2,** 17–26
3. Ellison, D., Hinton, J., Hubbard, S. J., and Beynon, R. J. (1995) *Protein Sci.* **4,** 1337–1345
4. Finlayson, J. S., Potter, M., and Runner, C. R. (1963) *J. Natl. Cancer. Inst.* **31,** 91–107
5. Flower, D. R., North, A. C. T., and Attwood, T. K. (1993) *Protein Sci.* **2,** 753–761
6. Flower, D. R. (1994) *FEBS Lett.* **354,** 7–11
7. Robertson, D. H., Beynon, R. J., and Evershed, R. P. (1992) *J. Chem. Ecol.* **19,** 1405–1416
8. Bacchini, A., Gaetani, E., and Cavaggioni, A. (1992) *Experientia* **48,** 419–421
9. Robertson, D., Hurst, J. L., Hubbard, S. J., Gaskell, S. J., and Beynon, R. J. (1998) *J. Chem. Ecol.* **24,** 1127–1140
10. Hurst, J. L., Robertson, D. H. L., Tolladay, U., and Beynon, R. J. (1998) *Anim. Behav.* **55,** 1285–1297
11. Mucignat-Caretta, C., Caretta, A., and Cavaggioni, A. (1995) *J. Physiol. (Lond.)* **486,** 517–522
12. Robertson, D. H. L., Cox, K. A., Gaskell, S. J., Evershed, R. P., and Beynon, R. J. (1996) *Biochem. J.* **316,** 265–272
13. Laemmli, U. K. (1970) *Nature* **227,** 680–685
14. Ferrige, A. G., Seddon, M. J., Green, B. N., Jarvis, S. A., and Skilling, J. (1992) *Rapid. Commun. Mass Spectrom.* **6,** 707–711
15. Ako, H., Foster, R. J., and Ryan, C. A. (1972) *Biochem. Biophys. Res. Commun.* **47,** 1402–1407
16. Evershed, R. P., Robertson, D. H. L., Beynon, R. J., and Green, B. N. (1993) *Rapid Commun. Mass Spectrom.* **7,** 882–886
17. Bocskei, Z., Groom, C. R., Flower, D. R., Wright, C. E., Phillips, S. E. V.,

of a seryl for a cysteinyl residue to obviate peptide dimerization through disulfide bonds. The second is the C-terminal sequence after removal of the glutamyl residue and is thus equivalent to the C-terminal sequence of MUP[9–161]. Both peptides were weak trypsin inhibitors (Fig. 7), and the shorter of the two was more potent. This may reflect the need for exopeptidase attack on the longer peptide.

Trypsin was incubated with the peptides SLQAR or SLQARE, and the reaction products were analyzed by electrospray mass spectrometry (Fig. 8). Trypsin was resolved as multiple peaks, and a second set of peaks, offset by approximately 557 Da, was also observed. These are consistent with addition of the peptide SLQAR to trypsin and loss of a water molecule as a consequence of formation of an acyl-enzyme intermediate. The mass difference between trypsin and the peptide-acyl-enzyme adduct is not sufficient to be visible on SDS-PAGE, and we have already noted that ESI/MS cannot yield accurate measurements of the relative abundance of the different peaks. It is therefore not possible to assess the amount of the complex that was formed with the short pep-

Cavaggioni, A., Findlay, J. B. C., and North, A. C. T. (1992) *Nature* **360,** 186–188

18. Chowdhury, S. K., Katta, V., and Chait, B. T. (1990) *Rapid Commun. Mass Spectrom.* **4,** 81–87
19. Ashton, D. E., Ashcroft, A. E., Beddell, C. R., Cooper, D. J., Green, B. N., and Oliver, R. W. A. (1994) *Biochem. Biophys. Res. Commun.* **199,** 694–698
20. Kasai, K. (1992) *J. Chromatogr.* **597,** 3–18
21. Chung, S. Y., Ullah, A. H. J., and Sanders, T. H. (1994) *J. Agric. Food Chem.* **42,** 623–628
22. Kasche, V. (1989) in *Proteolytic Enzymes. A Practical Approach* (Beynon, R. J., and Bond, J. S., eds) pp. 125–143, IRL Press, Oxford
23. Wilmouth, R. C., Clifton, I. J., Robinson, C. V., Roach, P. L., Aplin, R. T., Westwood, N. J., Hajdu, J., and Schofield, C. J. (1997) *Nat. Struct. Biol.* **4,** 456–462
24. Christensen, S., Valnickova, Z., Thogersen, I. B., Pizzo, S. V., Nielsen, H. R., Roepstorff, P., and Enghild, J. J. (1995) *J. Biol. Chem.* **270,** 14859–14862
25. Summerfield, S. G., Bolgar, M. S., and Gaskell, S. J. (1997) *J. Mass. Spectrom.* **32,** 225–231