Real-Time Fluorescence Monitoring of Tryptic Digestion in Proteomics

ARTICLE *in* JOURNAL OF PROTEOME RESEARCH · FEBRUARY 2008

Impact Factor: 4.25 · DOI: 10.1021/pr0704480 · Source: PubMed

research articles Journal of proteome research

Real-Time Fluorescence Monitoring of Tryptic Digestion in Proteomics

Peter Karuso,*,†,‡,§ Angela S. Crawford,|| Duncan A. Veal,†,‡ Graham B. I. Scott,|| and Hung-Yoon Choi‡,§

Department of Chemistry & Biomolecular Sciences, Macquarie University, Sydney NSW 2109, Australia, FLUOROtechnics Pty. Ltd., Macquarie University, Sydney NSW 2109, Australia, and Sigma-Aldrich Biotechnology, 2909 Laclede Avenue, St. Louis, Missouri 63103

Received July 19, 2007

Ensuring that proteolytic digestions are complete before submitting samples for downstream proteomic analyses is important, as failure or partial digestion can waste valuable instrument time and make results difficult to interpret. Conversely, overdigestion can also be problematic, such as when removing affinity tags from recombinant proteins or using nonspecific proteases. The techniques of HPLC, circular dichroism, SDS-PAGE, and MS have each been used to assess protein digestion. These techniques are slow, may require expensive instrumentation, can be inaccurate, and/or are unsuitable for real-time monitoring. Epicocconone is a natural fluorophore that reacts reversibly with proteins to form a highly fluorescent adduct and has previously been used to quantify proteins in 1D and 2D gels and in solution. Here, we describe a new method for the real-time monitoring of protein digestion based on epicocconone. This unique *in situ* fluorescent assay can tracelessly follow proteolysis of samples, at low microgram levels, destined for proteomics analysis or purification.

Keywords: proteomics • epicocconone • tryptic digestion • mass spectrometry • PMF • LC−MS • real-time assay • high-throughput proteomics • expression tag removal • enzyme kinetics

Introduction

Protein proteolysis is a central procedure in many proteomics workflows and is routinely carried out upstream of various mass spectrometric (MS) analyses (viz. MALDI-TOF Peptide Mass Fingerprinting [PMF], MALDI TOF/TOF, or LC-ESI−MS/MS experiments). Most in-solution tryptic digestions are routinely incubated overnight. Ensuring complete digestion, while avoiding overdigestion (with proteases such as Lys-C or proteinase K), is critical before analyzing samples via mass spectrometric or HPLC techniques. Failure or partial digestion can waste valuable instrument time and makes results difficult to interpret, while overdigestion can result in unexpected products or produce peptides too small to positively identify, thereby hindering protein identification. Increasingly, MS is used quantitatively in bottom-up proteomics via techniques such as absolute quantification (AQUA) that rely on consistent and complete proteolysis for reliability.[1] Indeed one of the key assumptions inherent to the AQUA technique is that proteolysis of the biological sample or cell lysate is fully driven to completion, otherwise the veracity of the quantitative data obtained may be questionable. Historically, HPLC,[2] circular dichroism,[3,4] SDS-PAGE,[5] and MS[6] have been used to assess the extent of protein digestion. These techniques are slow, may require expensive instrumentation, can be imprecise, are generally unsuitable for real-time monitoring, and are not amenable for coupling with high-throughput workflows. There is thus a place for a simple and reliable assay that enables accurate measurement of protease-driven hydrolysis prior to identification via mass spectrometry (e.g., PMF or LC−MS/MS). A fluorescently labeled protein, such as α-casein,[7] can be used to simply measure protease activity, but this method does not follow the digestion of the protein(s) of interest, whose rates of hydrolysis can vary widely (*vida infra*) and is thus only an approximate method of measuring proteolytic activity. Similarly, fluorophores conjugated to peptides are only pseudosubstrates, and the rates of hydrolysis are merely indicative of enzyme activity. It may also be advantageous in proteomics to curtail a digestion at some point prior to complete digestion to increase the number of missed cleavages and obtain a larger number of big peptides. For some proteins, this can improve the coverage (e.g., improve the MASCOT score) and make MS−MS peptide sequencing easier. There is currently no method to follow *in situ* proteolytic digestion in real-time on actual substrates.

Expression tags and fusion proteins such as hexa- or penta-HIS, FLAG, and GST or GBP fusions are routinely used in protein expression systems. In the production of recombinant proteins, expression tags are often used to aid purification and identification of the sought after protein but should be removed

* To whom correspondence should be addressed. Professor Peter Karuso, Department of Chemistry & Biomolecular Sciences, Macquarie University, Sydney New South Wales 2109, Australia. Phone: +61298508290. Fax: +61298508313. E-mail: peter.karuso@mq.edu.au.

† Department of Chemistry & Biomolecular Sciences, Macquarie University.

‡ FLUOROtechnics Pty. Ltd., Macquarie University.

|| Sigma-Aldrich Biotechnology.

§ These authors contributed equally to this work.

10.1021/pr0704480 CCC: $40.75

Published on Web 12/01/2007

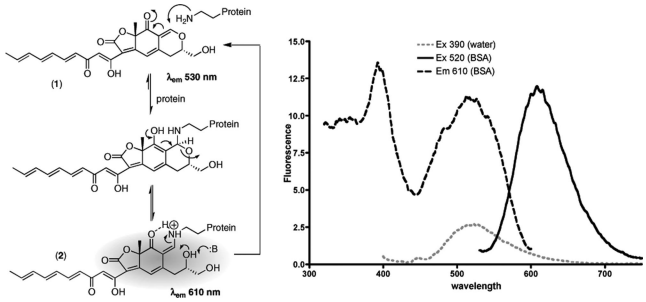

Figure 1. Mechanism of reversible staining of proteins by epicocconone. Lysine (and arginine/histidine) residues of a protein react with the enol-ether of epicocconone to produce a meta-stable enamine that is highly fluorescent (left panel). Emission profile of epicocconone (dotted grey line) and the [protein−epicocconone] enamine adduct (solid black line) in water (right panel). The excitation profile of the [protein−epicocconone] enamine (black dashed line) allows excitation at a range of wavelengths from 300–550 nm.

once no longer needed.[8] Current methodologies employed to determine complete removal of the tag include PAGE and Western blot analysis. Both techniques are slow and consume precious sample. With PAGE, there is often little difference in molecular weight between the tagged and untagged protein, and Western blots require the purchase of tag-specific antibodies, as well as detection reagents. Monitoring tag removal in real-time could ensure complete excision of the tag from the target protein, while avoiding overdigestion.

Epicocconone is a natural product[9] that reacts reversibly with proteins by a novel mechanism (Figure 1) to form an internal charge transfer (ICT) complex that is highly fluorescent in the relatively hydrophobic environment around proteins, but is in rapid equilibrium with unconjugated, non(red)-fluorescent dye at pH's above 5.[10] This mechanism has been exploited for accurate and sensitive protein quantification in solution.[11] We postulated that this mechanism could also be employed to follow the decrease in protein concentration during proteolysis. As the reaction of proteins with epicocconone is covalent, but reversible, the digested proteins could be used for downstream processes such as functional assays, HPLC, or MS. This technology would thus provide a useful and accurate method for assessing proteolytic digestion in real time for proteomics applications and for following the removal of expression tags from recombinant proteins.

Experimental Section

Chemicals and Reagents. Bovine serum albumin (A3059), α-casein (C6780), apo-transferrin (T2036), carbonic anhydrase (C7025), lactoferrin (L4765), lysozyme (L6876), *Escherichia coli* proteome (EC1), FluoroProfile (FP0010), and bicine (B3876) were acquired from Sigma-Aldrich, (St. Louis, MO). Human plasma was obtained from APAF (Macquarie University) and RapiGest SF from Waters. Epicocconone, in 20% acetonitrile/80% DMSO at a concentration of 0.5 mg/mL, was obtained from FLUOROtechnics Pty. Ltd., Sydney, Australia. Epicocconone solutions were freshly prepared by diluting 200-fold in bicine digestion buffer. Bicine is used to replace ammonium carbonate or ammonium acetate because epicocconone reacts with amines producing high backgrounds. A stock solution of bicine digestion buffer was made as a 10× concentrate (1 M, pH 8.5 or 7.7) and diluted to 1× as needed. Dithiothreitol (Bio-Rad, 161–0611) was prepared at a concentration of 200 mM in bicine buffer (100 mM, pH 8.5). Iodoacetamide (Sigma-Aldrich, I6125) was prepared at a concentration of 1 M in bicine buffer (100 mM, pH 8.5). Sodium dodecyl sulfate (SDS; BDH 442444H) was prepared at a concentration of 10% in water. Proteomics grade (TPCK treated) trypsin (Sigma-Aldrich, T6567) and chymotrypsin (Sigma-Aldrich, C4129) were prepared at a concentration of 1 μg/μL in 1 mM HCl. Lys-C (Sigma-Aldrich, P3428) was prepared at a concentration of 0.1 μg/μL in water. NuPAGE Novex 12% Bis-Tris Gels (Invitrogen, Carlsbad, CA; NP0341) were used for SDS-PAGE, and NuPAGE LDS sample buffer (Invitrogen, NP0007) was used for the preparation of SDS-PAGE protein samples. LMW markers (GE Healthcare/Amersham Biosciences, Piscataway, NJ; 17–0446–01) were used as a protein standard for SDS-PAGE runs. LavaPurple total protein gel kit (FLUOROtechnics, LP011005) was used for staining protein gels during SDS-PAGE validation. A Typhoon 9200 (Amersham Biosciences/GE Healthcare) scanner was used for obtaining images of the stained gels.

Preparation of Protein Samples for In-Solution Digestion. Protein samples (substrates) were dissolved in 100 mM bicine (pH 8.5) at concentration of 10 mg/mL. Proteins (100 μL) were reduced, by adding 5 μL of 200 mM dithiothreitol (DTT) and 1 μL of 10% SDS, heating for 10 min at 70 °C, then alkylated, by adding 4 μL of 1 M iodoacetamide at room temperature for 45 min to 1 h. Residual iodoacetamide in the samples was neutralized by adding 20 μL of 200 mM DTT and incubated at room temperature for 45 min to 1 h. Plasma and *E. coli* proteomes were treated similarly except the protein concentration was determined using FluoroProfile.[11] For one sample of plasma, instead of 0.1% SDS, 0.1% RapiGest was used according to the manufacturer's protocol.

Real-Time Monitoring of Proteolytic Digestion Using Epicocconone with a 96-Well Fluorescence Reader. The reduced and alkylated protein samples were diluted 10-fold in 100 mM bicine buffer. Trypsin and Lys-C digestions were carried out in bicine buffer at pH 8.5, and chymotrypsin digestions were similarly conducted at pH 7.7 with the addition of 10 mM $CaCl_2$. The real-time monitoring assay for in-solution digestion of each protein required four wells in a black microtiter plate. Duplicates of the first or all of the wells were made as desired. Protein (substrate) in bicine (100 μL) was added to wells 1 and 2. Bicine buffer only (100 μL) was added to wells 3 and 4. A freshly prepared working solution of epicocconone (100 μL) was added to all four wells, and the plate was preincubated in the fluorescence plate reader at the optimal temperature (30 °C for chymotrypsin and 37 °C for Lys-C and trypsin) of each protease for 50 min. The final concentration of the proteins was approximately 1 nmol (i.e., 77 μg/well). After 50 min, the appropriate gain for the plate reader (FluoStar Optima, BMG) was obtained by reading the initial fluorescence of the samples (540 nm; 10 nm bandwidth excitation and 630 nm; 10 nm bandwidth emission filters) with 10 flashes/well. The FluoStar instrument required approximately 90 s to obtain the appropriate settings. Next, the protease (typically employed at a 1:20 ratio for Lys-C and a 1:40 ratio for chymotrypsin and trypsin) was added (2 μL) to wells 1 (substrate) and 3 (bicine buffer only) with mixing. Similarly, an equivalent amount of water or 1 mM HCl was added to wells 2 and 4, and fluorescence monitoring began immediately.

Data Analysis Using Exponential Decay. Fluorescence progress curves were manipulated in Microsoft Excel to subtract controls. The progress curve from well 3, containing the enzyme and bicine, was subtracted from well 1, containing the substrate

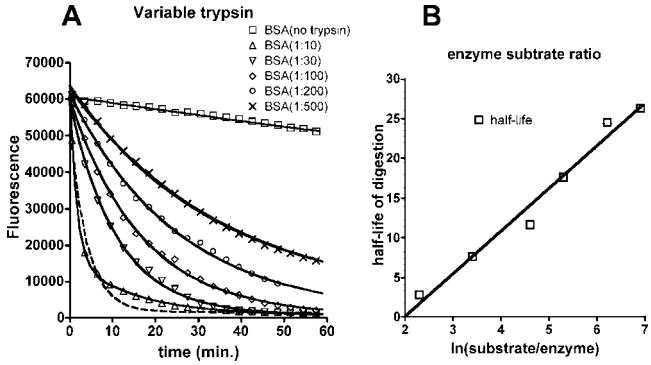

Figure 2. Fluorescence decay of BSA protein with variable amounts of trypsin, showing the breakdown of pseudo-first-order kinetics above a protein/enzyme ratio of 1:10 (A). Each plot is fitted to a single exponential decay except for the 1:10 condition, which is fitted to a double exponential. The one-phase exponential best fit is shown as a dotted line. (B) When the half-life of digestion is plotted against the natural logarithm of the enzyme/substrate ratio, a straight line is formed, indicating that the rate of fluorescence decay is directly related to the concentration of enzyme and is thus a true reflection of the hydrolysis rate.

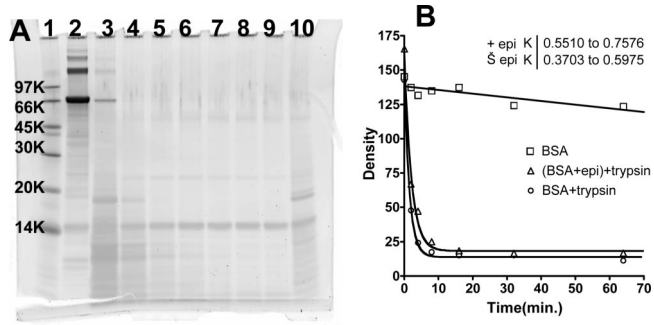

Figure 3. Independent validation of protein digestion with trypsin using gel electrophoresis. SDS-PAGE of subsampled tryptic digestion of BSA with epicocconone (A). Lane 1 is a molecular weight ladder, lanes 2–10 depict 0, 2, 4, 8, 16, 32, 64, 128 min, and overnight (18 h) digestion periods. The band at ~24 kDa is trypsin. Gels were stained with LavaPurple total protein stain and scanned on a Typhoon scanner (λ_{ex} 532, λ_{em} 560 LP). Kinetics of subsampled tryptic digestion of BSA with (open triangle) and without (open circles) epicocconone (B), analyzed by nonlinear least-squares regression of total fluorescence intensity of the stained gel. Pseudo-first-order rate constants (95% confidence interval) show no significant difference between samples with, or without, the addition of epicocconone.

and enzyme to make curve A. The progress curve from well 4, containing only buffer, was subtracted from well 2, containing the substrate–protein but no enzyme to make curve B. As duplicates were made, these were averaged to increase precision. Progress curves (A and B) were then fitted to a single or double exponential decay (Prism, version 4, GraphPad Software, San Diego, CA) to obtain kinetic constants (e.g., Figure 2). A more accurate analysis was achieved by first fitting a single phase exponential decay to curve B to obtain a rate constant (K_1). A two-phase exponential decay was then fitted to the digestion (curve A), while fixing the rate constant of the first exponential (K_1) to that of the control (curve B). This effectively subtracted the natural decay of fluorescent signal, unassociated with proteolysis. From this analysis, the half-life of digestion ($0.69/K_2$) and pseudo-first-order rate constant for digestion (K_2) were obtained (e.g., Figures 6 and 7).

Results and Discussion

Protein samples treated with epicocconone show strong orange/red fluorescence, unlike the unconjugated dye, which exhibits only a weak green fluorescence (Figure 1). The red fluorescence slowly decreases with time due to photobleaching and/or decomposition of the fluorescent adduct. However, with the addition of a protease (e.g., trypsin), a concentration-dependent decrease in fluorescence is observed (Figure 2). This decrease in fluorescence follows pseudo-first-order kinetics at low enzyme concentrations.

Independent validation of tryptic digestion was achieved by SDS-PAGE of subsampled tryptic digests, both with (Figure 3A), and without, epicocconone present (Supporting Information Figure 1). In each case, digestion appears to be complete after ~15 min, and the kinetics (pseudo-first-order rate constant; Figure 3B) were the same (within error) whether epicocconone was present or not. Further validation was achieved through HPLC analysis of a tryptic digestion of BSA (Supporting Information Figure 2) to observe the loss of the protein peak and concomitant increase of peptide peaks. The traces with, and without, epicocconone were qualitatively similar and showed major loss of the parent protein within the first 5 min of digestion.

The rate constants for hydrolysis of different proteins with trypsin varied, as, for example, carbonic anhydrase was hydrolyzed far more slowly than α-casein, though all proteins, regardless of p*I* and other physical characteristics (Supporting Information Table 3) behaved similarly (Figure 4). Notably, lysozyme, with an extremely high p*I* (11.35), was quite insoluble under the reaction conditions and was used at 10× lower concentration than the other proteins. With the use of nonlinear least-squares regression analysis, it was possible to quantify the pseudo-first-order rates constants for proteolysis using this method (Figure 4 insets).

Complex proteomes could also be monitored. Figure 5 showed real-time data of tryptic digestion of human plasma and *E. coli* proteomes. The close parallel between panels A and B in Figure 5 show that the epicocconone assay is compatible with RapiGest SF (Waters). RapiGest is claimed to increase the rate of proteolysis, but we found that the rates were quite similar compared to a similar concentration of SDS. The SDS-PAGE (Figure 5D) of subsamples taken after 2 h also show a similar degree of digestion. According to the calculated half-life, complete digestion would have been achieved after ~3 h. Similarly, *E. coli* proteins were relatively quickly digested ($t_{1/2} = 4$ min), but fluorescence remains at a high level due to the persistence of relatively large peptides (Figure 5D; lane 7). Of note was that the proteome alone (without trypsin) also dropped in fluorescence, which was interpreted as the action of endogenous proteases in the *E. coli* proteome.

Other proteolytic enzymes, such as Lys-C and chymotrypsin, are also used in proteomics due to their propensity for less frequent cleavage or different selectivity, respectively. Following chymotrypsin (Figure 6A,B) and Lys-C (Figure 6C,D) digestion of BSA and carbonic anhydrase (CA), in the presence of epicocconone, a qualitatively similar result to that of trypsin was revealed. In both cases, CA was cleaved more slowly than BSA. A limit of detection of between 1 and 10 μg/well was established through serial dilution of BSA from 100 to 0.79 μg/well (Supporting Information).

To demonstrate that this method is also potentially useful in following removal of expression tags, we monitored cleavage

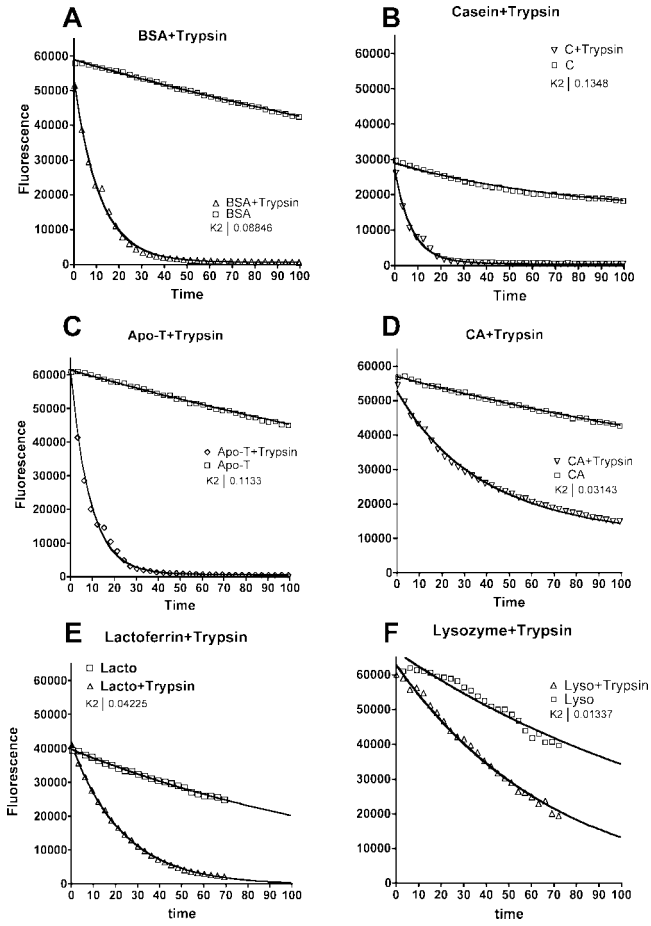

Figure 4. Kinetic analysis of proteolytic digestion of proteins using epicocconone. Trypsin hydrolysis of BSA (A), α-casein (B), Apotransferrin (C), carbonic anhydrase (D), lactoferrin (E), and lysozyme (F) under identical conditions (protein/trypsin ratio = 1:40, 100 mM bicine buffer, pH 8.5). Upper traces (open squares) are protein samples with no protease added.

of a small (octapeptide) FLAG tag from recombinant FLAG-BAP protein by use of the protease, enterokinase (Figure 7A). Unimpeded removal of the tag was independently validated by Western blot analysis using an anti-FLAG antibody (Figure 7B), which showed that the FLAG tag was no longer detectable after overnight treatment with enterokinase enzyme and epicocconone.

Tryptic digestions of four proteins, both with and without epicocconone, were analyzed via PMF, and the data were evaluated via the database search program, MASCOT (Matrix Science Ltd., London, U.K.). The peptides generated from tryptic digestion (with and without epicocconone) were virtually identical and unambiguously identified each protein (Supporting Information Table 1). There were no significant differences in sequence coverage obtained from the subsamples containing or deficient in epicocconone, as an optional additive to the tryptic digest. Similarly, there was little difference in the peptides identified by LC−MS/MS or the percentage coverage (Supporting Information Table 4) indicating that the addition of epicocconone does not interfere with the ability of LC−MS to identify proteins.

The success of proteolytic digestion is difficult to monitor in modern high-throughput proteomics experiments, but is a critical step in proteomics workflow(s) for protein identification. The reversible mechanism of fluorescent staining of proteins

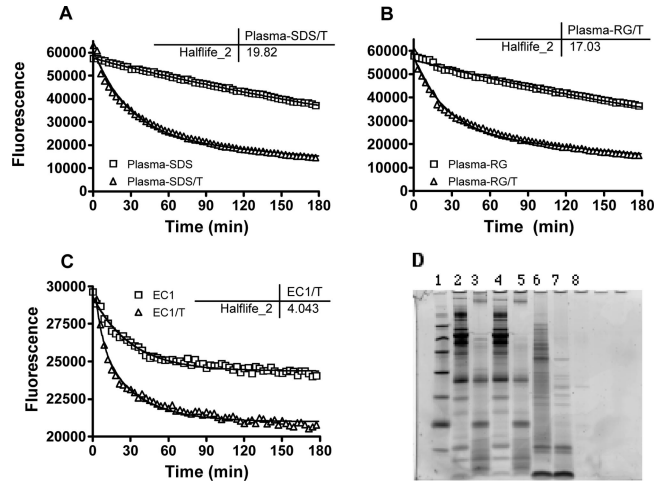

Figure 5. Real-time tryptic digestion of complex proteomes. Human plasma with SDS (A) or RapiGest (B) and *E. coli* proteome (C). Inserted are the calculated half-lives of tryptic digestion. All digestions were validated by SDS-PAGE (D): lane 1, LWM markers (97, 66, 45, 30, 20.1, and 14.4 kDa; lane 2, undigested plasma proteome; lane 3, trypsin-digested plasma protein containing SDS (plasma-SDS/T); lane 4, undigested plasma with 0.1% RapiGest; lane 5, digested plasma containing 0.1% RapiGest (plasma-RG/T); lane 6, undigested *E. coli* proteome (EC1), lane 7, *E. coli* proteome with trypsin (EC1/T); lane 8, trypsin (23.3 kDa) only. Human plasma samples were digested with trypsin at a ratio of 80:1 (w/w) and *E. coli* proteome at a ratio of 40:1.

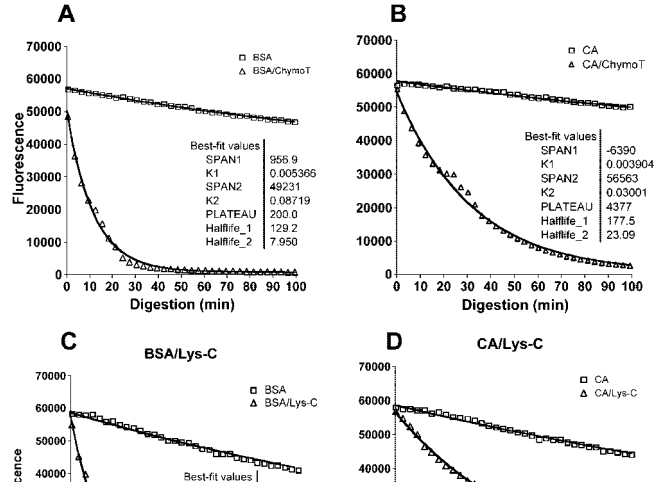

Figure 6. Digestion of BSA (open squares) and carbonic anhydrase (open triangles) with chymotrypsin (A and B) and Lys-C (C and D). The upper trace (open squares) shows the protein without enzyme, and the lower traces (open triangles) depict the protein with a 1:20 enzyme/substrate ratio for Lys-C and a 1:40 enzyme/substrate ratio for chymotrypsin.

with epiccconone (Figure 1) has allowed development of a method for real-time monitoring of proteolytic digestion (Figure 2). The linear correlation [of reaction half-life to substrate/enzyme ratio] shown in Figure 2B indicates that the progress curves (Figure 2A) are exponentially related to the enzyme concentration and, thus, evidence that the observed progress

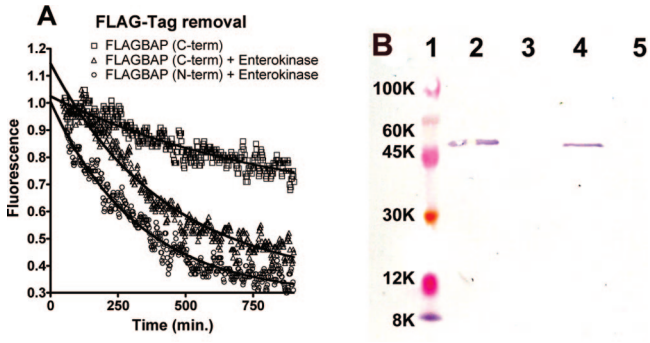

Figure 7. Removal of the protein expression tag (FLAG) from both C- and N-linked BAP (bacterial alkaline phosphatase) protein was monitored on a fluorescence spectrophotometer (A) and validated by Western blotting (B). Lanes 2 and 4 show the intact FLAG-BAP proteins, tagged at the C-terminus and N-terminus, respectively. Lanes 3 and 5 show the proteins after removal of the tag using enterokinase following overnight incubation. Lane 1 contains molecular weight markers.

curves follow the true hydrolysis rate of the added protein. Further validation was obtained through SDS-PAGE of subsampled incubations of BSA. Reconstruction of a progress curve (Figure 3B) indicated that there was no significant difference between the rates of digestion between samples that contained epicocconone at the working concentration to those that did not, as indicated by the 95% confidence intervals for the two calculated rate constants. Qualitative evidence that epicocconone does not interfere with protease activity was obtained from subsampled tryptic digestions of BSA followed by HPLC analysis of the resulting mixture (Supporting Information Figures 2A,B). The technique is applicable to all proteins tested (Figure 4) and even complex proteomes such as human plasma and *E. coli* extracts (Figure 5).

Mass spectrometry is an important downstream technique that is used to identify proteins based on peptides liberated by proteolysis. It is thus critical that any method used for *in situ* monitoring of digestion does not interfere with mass spectrometric analysis or peptide mapping. To assess this, we employed two methods commonly used in proteomics, tryptic digestion followed by PMF and LC−MS/MS. Samples of four proteins digested with and without epicocconone present were submitted for peptide mass fingerprinting (PMF). All proteins were successfully identified whether epiccconone was present in the tryptic digestion or not. Furthermore, epicocconone did not affect the number of peptides identified (Supporting Information Table 1) nor the types of peptides identified (Supporting Information Table 2) even if acid was used to stop the reaction. Similarly, LC−MS/MS analysis of BSA digested with trypsin, with and without epicocconone present, showed little difference between the numbers of peptides identified (∼50) or the percentage coverage (∼80%) - irrespective of which time point was analyzed (Supporting Information Table 4).

Even though protease activity is a relatively complex phenomenon, it is well-known that the hydrolysis rates of model substrates at low enzyme concentrations follow pseudo-first-order kinetics.[12] However, simple kinetics breaks down when the enzyme concentration exceeds the substrate concentration. Here, we find that even at a 1:10 enzyme-to-substrate ratio the rate is clearly two-phase (Figure 2), but fits better to a single phase exponential for enzyme to substrate ratios ≥1:30. This greater sensitivity to enzyme concentration may be a result of this assay following the progress of hydrolysis on real (viz. complex protein) substrate(s) rather than a fluorescently tagged model substrate. This has previously not been possible to observe so directly, and hence, this method may also find utility in better understanding of enzyme catalyzed hydrolysis.

For the three proteases studied, all showed different rates and preferences for the four proteins tested. Carbonic anhydrase (CA) was universally the slowest protein to be hydrolyzed, and α-casein was the fastest by all enzymes tested (Figures 5 and 6). This may be related to CA's structure in solution, but the phenomenon is curious, as CA's sequence and size are similar to that of α-casein (Supporting Information Table 3). Trypsin and LysC cleaved apotransferrin (ApoT) faster than BSA, while chymotrypsin cleaved BSA faster than ApoT (data not shown). Again, the differences observed are curious because both BSA and ApoT are similar in sizes and p*I*'s and have approximately the same number of tryptic and chymotryptic cleavage sites (∼70 each). The differences observed are thus not related to sequence, size, composition, or p*I*. They are most likely associated with the secondary structure of each protein in solution (CA and ApoT contain the largest amount of β-sheet), which in turn affects the speed at which an endoprotease can process the substrate. However, in all cases, complete hydrolysis was achieved in minutes, not hours, suggesting that improvements in protein coverage could be achieved by optimizing digestion times to save time and resources in high-throughput proteomics or to stop a digestion at a predetermined point. Indeed, analysis by LC−MS/MS (Supporting Information Table 4) indicated that the largest number of peptides and the best coverage was achieved after a 15 min digestion of BSA. Overdigestion with trypsin is known to lead to artifacts because trypsin also has some chymotryptic activity (cleaves after phenylalanine, tyrosine, tryptophan, and leucine) and transpeptidase activity (i.e., N-terminal addition of lysine or arginine).[13] If these artifacts are avoided through optimization of the individual tryptic digestions, better proteomics results can be expected.

The method also provides an approach for the *in situ* monitoring of epitope tag removal from recombinant proteins. In this case, only a small change in hydrophobicity is observed in the protein, but epicocconone was found to respond adequately to the change (Figure 7A) and not interfere with enterokinase as independently validated by Western blotting (Figure 7B). Western blotting is extremely sensitive to the presence of specific epitopes, and the fact that lanes 3 and 5 contain no band is strong evidence that the FLAG tag has been completely removed. The half-life of digestion for both the C- and N-linked FLAG-tagged proteins was found to be 4–5 h, indicating that, under the conditions used here, the FLAG-tagged proteins require overnight digestion or a higher concentration of enterokinase. Again, this highlights the utility of monitoring cleavage by a protease.

Conclusions

We have developed a simple assay based on the chemistry of epicocconone that is suitable for real-time monitoring of in-solution proteolytic digestions. This method can be used to follow the hydrolysis of proteins in a fluorescence spectrometer, fluorescence scanner, or 96/384 well fluorescence plate reader, where the assay can be easily automated for high-throughput applications in proteomics, protein biotechnology, or other protein-focused fields, such as the food industry. In proteomics, the method may be automated through the use of robotics or

converted to an in-line assay by employing a syringe pump (to introduce trypsin and epicocconone) and a fluorescence detector between the HPLC and MS.

Epicocconone provides a simple, traceless technology for real-time monitoring of tryptic digestion that is suitable for all proteins tested and does not appear to interfere with proteolytic activity. The method is fully compatible with many downstream analyses, and potentially replaces expensive/time-consuming gel electrophoresis, HPLC, or CD for validation of proteolytic digestion. Moreover, this assay can be used to calculate pseudo-first-order rate constants for proteolytic digestion on actual substrate(s) of interest, rather than on a model substrate. The same sample can then be used for downstream analytical processes, as the fluorescent stain does not permanently derivatize substrates. The method can also be useful for the monitoring tag removal from recombinant proteins. In these cases, the removal of FLAG, or other fusion tags used for expression, purification or solublization, can be followed, and thus optimized. Monitoring the digestion "*in situ*" ensures a complete digestion and avoids overdigestion, or can be used to determine when to arrest a digestion at a specific time-point.

Acknowledgment. The authors thank Richard J. Mehigh, Justin Wildsmith, and Kevin B. Ray of Sigma-Aldrich Biotechnology and Narelle Jay of APAF for mass spectrometric data and Dian Er Chen (Sigma-Aldrich) and Jim Mackintosh (FLUOROtechnics) for assistance and helpful discussions.

Supporting Information Available: Supporting Information is available on the independent validation of tryptic digestions using HPLC, LC−MS, and PMF. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Gerber, S. A.; Rush, J.; Stemman, O.; Kirschner, M. W.; Gygi, S. P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Nat. Acad. Sci. U.S.A.* **2003**, *100* (12), 6940–5.

(2) Kohse, K. P.; Wisser, H. Quantitative monitoring by high-performance liquid chromatography of the dissociation of human serum cholinesterase by limited proteolysis. *J. Chromatogr.* **1987**, *420* (1), 25–33.

(3) Caballero, R.; Fernandez, B. A.; Moran, F.; Montero, F. Kinetics of conformational changes in the globular head of calf thymus H1 upon tryptic digestion. *Biochem. Int.* **1985**, *11* (5), 765–72.

(4) Noda, Y.; Fujiwara, K.; Yamamoto, K.; Fukuno, T.; Segawa, S. Specificity of trypsin digestion and conformational flexibility at different sites of unfolded lysozyme. *Biopolymers* **1994**, *34* (2), 217–26.

(5) Thomas, K.; Aalbers, M.; Bannon, G. A.; Bartels, M.; Dearman, R. J.; Esdaile, D. J.; Fu, T. J.; Glatt, C. M.; Hadfield, N.; Hatzos, C.; Hefle, S. L.; Heylings, J. R.; Goodman, R. E.; Henry, B.; Herouet, C.; Holsapple, M.; Ladics, G. S.; Landry, T. D.; MacIntosh, S. C.; Rice, E. A.; Privalle, L. S.; Steiner, H. Y.; Teshima, R.; Van Ree, R.; Woolhiser, M.; Zawodny, J. A multi-laboratory evaluation of a common in vitro pepsin digestion assay protocol used in assessing the safety of novel proteins. *Regul. Toxicol. Pharmacol.* **2004**, *39* (2), 87–98.

(6) Bothner, B.; Chavez, R.; Wei, J.; Strupp, C.; Phung, Q.; Schneemann, A.; Siuzdak, G. Monitoring enzyme catalysis with mass spectrometry. *J. Biol. Chem.* **2000**, *275* (18), 13455–9.

(7) Haugland, R. P. Detecting Peptidases and Proteases. In *The Handbook—A Guide to Fluorescent Probes and Labeling Technologies*, 9th ed.; Molecular Probes: Eugene, OR, 2002, Section 10.4, pp 430–9.

(8) Jarvik, J. W.; Telmer, C. A. Epitope tagging. *Annu. Rev. Genet.* **1998**, *32*, 601–18.

(9) Bell, P. J. L.; Karuso, P. Epicocconone, a novel fluorescent compound from the fungus Epicoccum nigrum. *J. Am. Chem. Soc.* **2003**, *125* (31), 9304–5.

(10) Coghlan, D. R.; Mackintosh, J. A.; Karuso, P. Mechanism of reversible fluorescent staining of protein with epicocconone. *Org. Lett.* **2005**, *7* (12), 2401–4.

(11) Mackintosh, J. A.; Veal, D. A.; Karuso, P. Fluoroprofile, a fluorescence-based assay for the rapid quantification of proteins in solution. *Proteomics* **2005**, *5* (18), 4673–7.

(12) Srividhya, J.; Schnell, S. Why substrate depletion has apparent first-order kinetics in enzymatic digestion. *Comp. Biol. Chem.* **2006**, *30* (3), 209–14.

(13) Schaefer, H.; Chamrad, D. C.; Marcus, K.; Reidegeld, K. A.; Blüggel, M.; Meyer, H. E. Tryptic transpeptidation products observed in proteome analysis by liquid chromatography-tandem mass spectrometry. *Proteomics* **2005**, *5* (4), 846–52.

PR0704480