High-Temperature Protein Mass Mapping Using a Thermophilic Protease

Steven J. Bark,* Nemone Muster, John R. Yates, III, and Gary Siuzdak

The Scripps Research Institute
Center for Mass Spectrometry
Departments of Chemistry,
Cell Biology and Molecular Biology
La Jolla, California 92037

Received August 7, 2000

The revolution of genomics has recently generated massive sequence databases for entire genomes. Proteomics is the study of the proteins encoded by these DNA sequences. The primary techniques for proteomic analysis consist of 2-D gel electrophoresis followed by proteolytic digestion, mass spectrometry, and computer-facilitated data analysis.[1] Masses for proteolytic fragments of the target protein generated by enzymatic degradation can be measured directly with high accuracy (±0.005 Da for a 1000 Da peptide). Alternatively, liquid chromatography tandem mass spectrometry can be used to analyze the protein digest fragments. This information is then compared with the theoretical proteolytic fragments predicted for proteins within a database, and matches are statistically evaluated. The success of this strategy relies on the existence of the protein sequence within the database, but with the sequences of whole genomes being completed, the likelihood for matches is high.

Two limitations of this method are the time required for adequate proteolysis and the resistance of some proteins to digestion. Adequate digestion with trypsin typically requires 4−24 h. In addition, many compact globular proteins exhibit resistance to cleavage by proteases and, thus, cannot be effectively analyzed by available procedures. These proteins are often constituents in larger protein complexes with critical biological functions.

The primary factor affecting the proteolysis of proteins is the mobility of protein structure,[2] a process influenced by chemical denaturants or cosolvents. Alternatively, high temperature also has the effect of increasing flexibility and, therefore, the proteolytic susceptibility of a protein.[3] This thermal denaturation process has the advantage of requiring no extra sample preparation or purification. We have taken advantage of thermal denaturation to generate a facile proteolysis method for identifying proteins. The key to this method is the use of a thermophilic enzyme, thermolysin, that exhibits optimal activity at elevated temperatures.[4]

We utilized a substrate:enzyme ratio in the range of 100:1 to 10:1 in either 100 mM Tris-HCl or 100 mM ammonium bicarbonate buffer, pH ≈ 7.5. The cleavage reaction was incubated in a water bath at 65 °C and then analyzed by either a matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF), MALDI-TOF reflector, or a liquid chromatography tandem mass spectrometer (LC/MS/MS). Using these conditions, we observed rapid proteolytic fragmentation of myoglobin, bovine serum albumin (BSA), cytochrome C, ribonuclease A, and a highly interlinked capsid from Hong Kong 97 (HK97) virus. MALDI-MS experiments were performed on a PE Biosystems Voyager DE or DE-STR instrument. LC/MS/MS data was obtained on a Finnigan LCQ system equipped with an Agilent HP 1100 pump system. For time-course experiments, an aliquot of the digest was removed, mixed with α-cyano-4-hydroxycinnammic acid, placed on a MALDI plate and analyzed. Prior to MALDI-TOF analysis, all samples were subjected to an on-plate wash to remove salts by placing 1 μL of ice water on the crystallized sample for 10 s followed by removal of the water. LC/MS/MS data was acquired on the crude digests without any further processing.

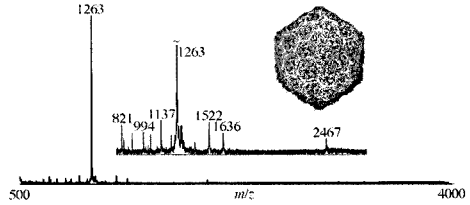

Figure 1. Thermolysin rapidly cleaves the highly interlinked capsid from HK97 virus at 65 °C. Cleavage reaction was performed for 2 h in 100 mM Tris-HCl, 10 mM $CaCl_2$. Trypsin cleavage showed no significant fragmentation within this time. (Inset) An image of HK97.

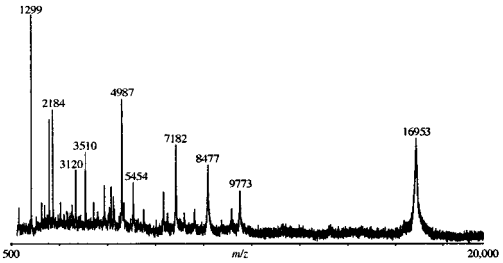

Figure 2. On-plate digest of apomyoglobin. A sample of apomyoglobin was mixed with thermolysin on-plate and heated to dryness (∼30 s). After digestion, MALDI analysis was performed using α-cyano-4-hydroxy-cinnammic acid matrix.

HK97 icosohedral virus capsid[5] presented an interesting system for these analyses because it has a highly stable interlinked "chain-mail" capsid that is relatively impervious to enzyme digestion. Our attempts at trypsin digestion of this capsid generated minimal peptide cleavage, even after 24 h, which is consistent with a previous report.[6] However, high-temperature digestion with thermolysin generated significant cleavage within 2 h (Figure 1).

The 1263 (1263.66) mass peptide fragment appears most rapidly during the digestion process. This peptide corresponds to residues 172−181 or 173−182 in the capsid sequence. These peptides are situated at a β-sheet interface between capsid proteins and demonstrate this to be the first region of the virus coat to be susceptible to denaturation.

Further experiments with BSA demonstrated that thermolysin cleavage at elevated temperatures was rapid (<15 min) and efficient (as determined by MALDI analysis). Subsequent tests demonstrated that a proteolytically susceptible protein, myoglobin, could be digested in a high-throughput format directly on a MALDI plate. A sample of myoglobin and thermolysin were mixed on-plate and heated to dryness (∼30 s). After the same sample preparation and wash described above, MALDI-TOF mass spectra were obtained. Analysis of this extremely rapid digest demonstrated incomplete but significant digestion of myoglobin (Figure 2). These experiments demonstrate the high efficiency of thermolysin for proteolysis of standard proteins.

To further examine the generality of this approach with other proteolytically resistant proteins, a digest of ribonuclease A was examined. Ribonuclease A maintains structural stability under

(1) Pandey, A.; Mann, M. *Nature* **2000**, *405*, 837−846.
(2) (a) Fontana, A.; Fassina, G.; Vita, C.; Dalzoppo, D.; Zamai, M.; Zambonin, M. *Biochemistry* **1986**, *25*, 1847−1851. (b) Hubbard, S. J.; Eisenmenger, F.; Thornton, J. M. *Protein Sci.* **1994**, *3*, 757−768.
(3) Park, Z.-Y.; Russell, D. H.; *Anal. Chem.* **2000**, *72*, 2667−2670.
(4) (a) Matsubara, H.; Feder, J. in *The Enzymes*; Academic Press: New York, 1970; Vol. 3, p 721. (b) Petra, P. H. In *Methods in Enzymology*; Academic Press: New York, 1970; Vol. 19, p 460.
(5) Wikoff, W.; Liljas, L.; Duda, R. L.; Tsuruta, H.; Hendrix, R. W.; Johnson, J. E. *Science* **2000**, *289*, 2129.
(6) (a) Duda, R. L. *Cell* **1998**, *94*, 55−60. (b) Popa, M. P.; McKelvey, T. A.; Hempel, J.; Hendrix, R. W. *J. Virol.* **1991**, *65*, 3227−3237

10.1021/ja002909n CCC: $20.00
Published on Web 01/31/2001

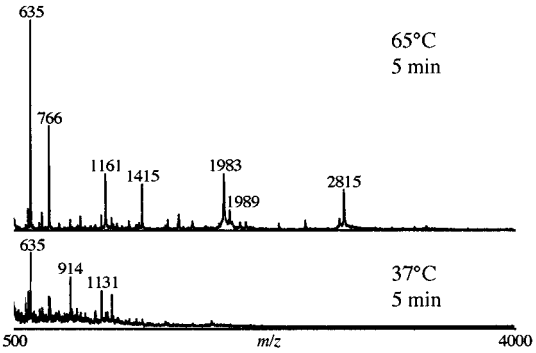

Figure 3. Comparison of thermolysin cleavage of ribonuclease A at 65 °C (Top) and 37 °C (Bottom) in 100 mM ammonium bicarbonate, pH ≈ 7.5. Cleavage reactions were performed for 5 min.

normal cleavage conditions and rapid digestion with trypsin requires denaturation. Thermolysin digestion at elevated temperatures rapidly fragmented this protein within 5 min without chemical denaturants. A low-temperature control digest demonstrates that elevated temperature is critical for good protein fragmentation (Figure 3).

We then attempted a digestion on a mixture of proteins to identify cleavage site preferences for thermolysin. A digestion of a mixture of BSA, cytochrome C, myoglobin, and ribonuclease A were analyzed by electrospray tandem mass spectrometry. After a 10 min digest at 65 °C, the sample was loaded onto a capillary reversed-phase column and eluted with a water/acetonitrile gradient. The MS/MS data was analyzed using SEQUEST based on a database of these proteins.[7] The cleavage sites were catalogued, and all four proteins were identified accurately. However, only one peptide from ribonuclease A was obtained. This is probably because of the resistance of this protein to proteolysis and the high concentration of more easily fragmented proteins.

The cleavage propensity for thermolysin was analyzed from all data (>130 cleavage sites identified). While the specificity of thermolysin centers on hydrophobic residues (W, Y, F, I, L, V, A, M),[8,9] we observed that cleavages were primarily at F, L, I, and V. We did not observe the large numbers of smaller peptides expected from a promiscuous protease. One interpretation for this observation is that there are likely to be structural constraints preventing access to all possible cleavage sites.[9] Ribonuclease A retains significant structural integrity at higher temperature,[10] and it is probable that a highly interlinked virus capsid could also retain significant structure under our cleavage conditions. This interpretation does not account for our observations with BSA, cytochrome C, or myoglobin. In these experiments, complete digestion would generate many more fragments than were observed, and BSA, cytochrome C, or myoglobin are not thermostable. Incomplete digestion at short times could account for some of the perceived specificity. However, even at longer time courses, larger polypeptide fragments persisted. On the basis of these experiments, it appears that thermolysin has greater specificity at higher temperatures.

Specificity has been an important consideration for proteolysis in protein identification in the past. Trypsin has been the most commonly used protease for proteomic digestion because it generates peptides with useful ionization and fragmentation characteristics. Because of advances in mass spectrometry technology and the data presented above, this may not be as relevant a consideration today. Modern mass spectrometers are capable of accurately measuring masses to 5 ppm or less which, in some cases, allows identification of a protein using only one or a few peptides.[11,12] Thermolysin possesses enough specificity under our conditions to generate peptides with cleavage sites at F, L, I, or V on both N- and C-termini and allow accurate identification. And sequence information can be obtained directly from tandem mass spectrometry, which allows identification of virtually any protein using proteases with any cleavage site preference.

To establish whether this methodology was capable of identification based on mass mapping, we digested both BSA and HK97 virus capsid and obtained internally calibrated spectra for each. The molecular masses of the fragments were obtained, and searches were undertaken using Profound.[13] Using only F, L, I, and V as amino acid cleavage sites, we obtained an unequivocal identification of both BSA and HK97 virus capsid proteins. These identifications were made using the molecular mass range of 0−100 kDa, and the database searches were constrained to mammals for BSA and viruses for HK97. Expanding the searches to include all taxa in the mass range of 0−500 kDa did not alter the identification significantly. These data indicate that thermolysin allows for identification of proteins that would be difficult or impossible to identify otherwise.

This facile digestion technique has been used for structural analysis, as exemplified by the limited proteolysis experiments.[9,14,15] The use of elevated temperatures for these studies have also been reported, albeit rarely. One such method was reported for the structural identification of a thermal unfolding event in ribonuclease A.[10] These studies demonstrate that, not only can we obtain identification information using this rapid proteolysis method, we can obtain a degree of structural information as well.

The use of thermolysin in protein identification, as exemplified in this paper, is a complementary technique to the available methodologies currently in use. This work also suggests that other thermophilic proteases, obtained naturally or through protein engineering, may also offer accelerated rates for reaction and could exhibit different and enhanced selectivity. In addition, thermophilic enzymes other than proteases could extend rapid biochemical and analytical analysis at elevated temperature to other important biochemical processes.

In conclusion, we have demonstrated that high-temperature proteolytic digestion using thermolysin is a more efficient (and general) method for protein identification of normal proteins and, more importantly, proteins that are resistant to normal proteolysis methods.

Acknowledgment. We gratefully acknowlege Dr. Xia Giao, Dr. Jane Wu, and David Tabb for useful discussions. We also gratefully acknowlege Dr. Jack Johnson for providing a sample of HK97 bacteriophage for these studies and Dr. Robert Duda and Dr. Roger Hendrix, University of Pittsburgh, for providing HK97 viral capsid heads. This work was supported by NIH grants RR11823-04 (JRY) and GM55775 (G.S.) from the National Institutes of Health.

JA002909N

(7) Eng, J. K.; McCormack, A. L.; Yates, J. R., III. *J. Am. Soc. Mass. Spectrom.* **1994**, *5*, 976−989.
(8) Voet, D.; Voet, J. G. *Biochemistry*, 2nd ed.; John Wiley & Sons: 1995; pp 112 (T).
(9) Arnold, U.; Rucknagel, K. P.; Schierhorn, A.; Ulbrich-Hoffmann, R.; *Eur. J. Biochem.* **1996**, *237*, 862−869.
(10) Hermans, J., Jr.; Scheraga, H. A. *J. Am. Chem. Soc.* **1961**, *83*, 3283.
(11) Conrads, T. P.; Anderson, G. A.; Veenstra, T. D.; Pasa-Tolic, L.; Smith, R. D. *Anal. Chem.* **2000**, *72*, 3349−3354.
(12) Goodlett, D. R.; Bruce, J. E.; Anderson, G. A.; Rist, B.; Pasa-Tolic, L.; Fiehn, O.; Smith, R. D.; Aebersold, R. *Anal. Chem.* **2000**, *72*, 1112−1118.
(13) Zhang, W.; Chait, B. T. *Anal. Chem.* **2000**, *72*, 248−249
(14) Gervasoni, P.; Staudenmann, W.; James, P.; Pluckthun, A.; *Biochemistry* **1998**, *37*, 11660−11669.
(15) Mantscheff, M.; Weiss, V.; Glocker, M. O. *Biochemistry* **1999**, *38*, 11012−11020.