Available online at www.sciencedirect.com
ScienceDirect
Analytical Biochemistry 356 (2006) 282–290
ANALYTICAL BIOCHEMISTRY
www.elsevier.com/locate/yabio

Rearrangement of terminal amino acid residues in peptides by protease-catalyzed intramolecular transpeptidation

Szilan Fodor, Zhongqi Zhang *

Amgen, Thousand Oaks, CA 91320, USA

Received 15 April 2006
Available online 5 July 2006

Abstract

Protease-catalyzed rearrangements of amino acid residues in peptides are observed during enzymatic digestion of proteins. When two enzyme-specific cleavage sites are within one or two residues of each other in the protein sequence, only one of the two sites usually is hydrolyzed by the protease, resulting in a peptide that contains an extra cleavage site near one of its termini. It is observed that in this type of peptide, the residues between the two cleavage sites often rearrange from one terminus of the peptide to the other terminus, catalyzed by the protease that created the peptide. It is proposed that the rearrangement is caused by protease-catalyzed intramolecular transpeptidation through a cyclic peptide intermediate. Several cases of this type of rearrangement were observed for different peptides generated by different proteases, indicating that this type of rearrangement is a general phenomenon occurring during enzymatic digestion of proteins.

Keywords: Rearrangement; Amino acid; Peptide; Transpeptidation; Protease; Protease-catalyzed; Cyclization; Digestion; Artifact

Enzymatic digestion followed by peptide mapping is an important analytical tool for protein characterization. However, artifacts often occur during the digestion process. Common artifacts include deamidation and isoaspartate formation, succinimide formation, pyroglutamic acid formation, oxidation, disulfide formation, disulfide interchange, carbamylation, and transpeptidation. To characterize a protein properly, it is important to understand these artifacts so as to distinguish them from impurities/modifications that are present in the sample.

Transpeptidation has been observed many times during proteolytic digestion of proteins [1–3]. In this article, we report protease-catalyzed rearrangements of amino acid residues that occur in the digestion process. When a proteolytic peptide has an extra protease-specific cleavage site close to one of its termini (by one or two residues), it was observed that these residues often rearrange from one terminus of the peptide to the other. The rearrangement is believed to be a protease-catalyzed intramolecular transpeptidation process through a cyclic peptide intermediate. A similar observation was reported previously on a tryptic peptide of hemoglobin using immobilized trypsin [2]. In the current work, we found that the artifact occurs much more frequently than was known previously.

Materials and methods

Materials

Purified recombinant human granulocyte colony-stimulating factor (G-CSF),[1] darbepoetin alfa (NESP), anti-NGF–Fc fusion protein, parathyroid hormone–Fc fusion protein (PTH–Fc), and two recombinant human monoclonal antibodies used in this study were manufactured at

* Corresponding author. Fax: +1 805 376 2354.
E-mail address: zzhang@amgen.com (Z. Zhang).

[1] *Abbreviations used:* G-CSF, granulocyte colony-stimulating factor; NESP, darbepoetin alfa; PTH–Fc, parathyroid hormone–Fc fusion protein; DTT, dithiothreitol; TFA, trifluoroacetic acid; MS/MS, tandem mass spectrometry; SIC, selected ion chromatogram.

0003-2697/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ab.2006.06.023

Amgen. Rabbit apo-transferrin, human hemoglobin, and horse cytochrome *c* were purchased from Sigma–Aldrich. Endoproteinase Glu-C, Lys-C, and trypsin were purchased from Roche Diagnostics. Urea was obtained from ICN Biomedicals. Sodium phosphate monobasic and dibasic (ultrapure reagent grade) were purchased from J.T. Baker. Hydroxylamine hydrochloride and methylamine hydrochloride were obtained from Sigma–Aldrich. Dithiothreitol (DTT) was obtained from Calbiochem. Trifluoroacetic acid (TFA) was obtained from Pierce. HPLC-grade acetonitrile was obtained from Burdick & Jackson.

LC/MS system

The LC/MS system used in this study consisted of an Agilent 1100 HPLC system directly connected to a ThermoFinnigan LCQ or LCQ DECA electrospray–ion trap mass spectrometer. Mobile phase A was 0.1% TFA in water, and mobile phase B was 0.1% TFA and 90% acetonitrile in water. All LC/MS analyses were performed on a reversed-phase column with gradient elution at a flow rate of 0.2 ml/min. The chromatogram was monitored by both UV (set at a wavelength of 214 nm) and MS. MS detection included full scans in positive mode, as well as data-dependent zoom scans and tandem mass spectrometry (MS/MS) scans of the most intense ions, with dynamic exclusion. Proteolytic peptides were identified automatically using a software program developed in-house that correlated the experimental tandem mass spectra against theoretical tandem mass spectra generated from known peptide sequences [4,5].

N-terminal sequencing

To identify a peptide by N-terminal sequencing, the peak of interest was collected manually and loaded onto the Applied Biosystems ProSorb (polyvinylidine difluoride) membranes. Sequencing was performed on an Applied Biosystems Procise protein sequencer equipped with a Perkin-Elmer series 200 detector and an Applied Biosystems model 140 C syringe pump.

Glu-C digestion of anti-NGF–Fc fusion protein

The digestion was performed in a solution containing 0.6 mg/ml of the anti-NGF–Fc fusion protein, 0.03 mg/ml of Glu-C (1:20 enzyme/substrate ratio), 2.25 M urea, 7.5 mM DTT, and 75 mM sodium phosphate buffer at pH 7.0. The digestion was carried out for 18 h at 25 °C, followed by the addition of an appropriate amount of 5% TFA for a final pH of 2.0 to 3.0. The digest was analyzed with the LC/MS system on a Phenomenex Jupiter column (C-5, 300 Å pore, 5 μm particle size, 2.0 mm i.d. × 250 mm length) with a gradient of 2 to 30% mobile phase B in 30 min and then to 45% mobile phase B in 30 min.

To confirm that the rearranged peptide was generated from the original peptide, the peptide of interest was manually collected, dried in a SpeedVac (Savant), reconstituted in the same digestion buffer, and incubated with and without 0.03 mg/ml of Glu-C at 25 °C for 18 h. The two samples were analyzed by the same LC/MS procedures as described above.

Lys-C digestion of PTH–Fc fusion protein

Digestion was performed in a solution containing 0.8 mg/ml of PTH–Fc, 0.04 mg/ml of Lys-C (1:20 enzyme/substrate ratio), 2.5 M urea, and 0.1 M sodium phosphate buffer at pH 7.0. The digestion was carried out for 18 h at 37 °C, followed by the addition of an appropriate amount of 5% TFA for a final pH of 2.0 to 3.0. The digest was analyzed on the same LC/MS system as for anti-NGF–Fc fusion protein.

Trypsin digestion of G-CSF

G-CSF was digested in a solution containing 1.0 mg/ml of G-CSF, 20 μg/ml of trypsin (1:50 enzyme/substrate ratio), 2 M urea, and 0.1 M sodium phosphate buffer at pH 7.0. The digestion was carried out for 18 h at 37 °C.

Trypsin digestion of NESP

NESP was digested in a solution containing 0.5 mg/ml of NESP, 12.5 μg/ml of trypsin (1:40 enzyme/substrate ratio), and 0.1 M sodium phosphate buffer at pH 7.2. The digestion was carried out at 37 °C for 16.5 h.

Results

Endoproteinase Glu-C hydrolyzes a protein primarily at the C-terminal sides of glutamic acid residues (E). Anti-NGF–Fc fusion protein contains 15 glutamic acid residues. If the C-terminal sides of all glutamic acid residues were hydrolyzed, anti-NGF–Fc fusion protein would generate 16 peptides, here named E1 to E16. Fig. 1 shows the UV chromatogram of Glu-C-digested anti-NGF–Fc, with identified proteolytic peptides as well as their determined monoisotopic masses labeled on top of each peak. All peptides were confirmed by their tandem mass spectra. Most peaks in the chromatogram can be attributed to proteolytic peptides from anti-NGF–Fc fusion protein.

Characterization of Glu-C-digested anti-NGF–Fc fusion protein with LC/MS drew attention to two closely eluted peaks A and B at approximately 51 min (indicated by two arrows in Fig. 1). Mass spectrometric analysis showed that both peptide A and peptide B have identical monoisotopic masses of 2839.3 Da (Fig. 2). N-terminal Edman sequencing as well as MS/MS analysis demonstrated that peak A corresponds to peptide E12–13 with a sequence of LTKNQVSLTCLVKGFYPSDIAVEWE (theoretical monoisotopic mass of 2839.44 Da) in the Fc region of the protein. Fig. 3 (top panel) shows the tandem mass spectrum of peptide A and assignments of fragment ions