ELSEVIER

Journal of Chromatography A, 864 (1999) 237–245

JOURNAL OF CHROMATOGRAPHY A

www.elsevier.com/locate/chroma

# Eliminating disulfide exchange during glutamyl endopeptidase digestion of native protein

Shelly J. Dormady, Jianming Lei, Fred E. Regnier*

*Department of Chemistry, Purdue University, Lafayette, IN 47907, USA*

Received 21 April 1999; received in revised form 15 September 1999; accepted 24 September 1999

**Abstract**

Numerous advantages of using immobilized enzymes over free-solution protein digests have been cited in the literature. This investigation examines both the rate of hydrolysis and the extent of disulfide bond exchange in disulfide bridged dipeptide fragments formed during proteolysis of native protein. Gutamyl endopeptidase as both an immobilized enzyme and in free solution was used in these studies. It was found that extensive hydrolysis of insulin was achieved in 2 min with immobilized enzyme cartridges operated in the stopped-flow mode orders. This is orders of magnitude faster than was seen in free solution. Other advantages ranging from ease of use and reduction in sample size to the potential for automation were also noted with the immobilized enzyme cartridge. Normal free-solution proteolysis generally requires 12–24 h, based on the lower enzyme-to-substrate ratio in solution. A disturbing feature noted in these lengthy free-solution reactions was the tendency to form disulfide bridged peptide artifacts. This could lead to the erroneous conclusion that disulfide bonding in a sample was not that of the native protein. It is concluded that the advantage of immobilized enzymes over free-solution reactions will be most important in the pharmaceutical industry where proteolytic fragment ''fingerprinting'' of recombinant proteins is being used to confirm structure. 

*Keywords:* Immobilized proteins; Glutamyl endopeptidase; Enzymes; Proteins; Insulin

## 1. Introduction

Human therapeutic proteins are now widely produced in microorganisms through recombinant DNA technology [1–3]. An essential element of this process is that there be some mechanism to guarantee the quality and efficacy of the product [4]. This means that production materials must be confirmed to be the same as the material demonstrated to be therapeutically efficacious in clinical trials. Because the efficacy of human therapeutic proteins is due to a single protein, international regulatory agencies now regulate these materials in terms of their chemical structure. Confirming quality and efficacy becomes a matter of verifying the various structural elements of the protein product. Ideally, in a process environment this could be done quickly and easily.

A particular three-dimensional structure is generally essential to the biological activity of a protein. Although primary structure is important in directing three-dimensional structure formation, disulfide bonds also frequently play a critical role in organizing and maintaining protein structure. A problem in producing proteins on a large scale is that disulfide

*Corresponding author. Department of Chemistry, Purdue University, 1393 Brown Building, West Lafayette, IN 47907, USA. Tel.: +1-765-4943-878; fax: +1-765-4940-359.

0021-9673/99/$ – see front matter 
PII: S0021-9673(99)01028-6

bonds may either not form properly in vivo or be scrambled during purification. Thus, biological activity can be compromised. Establishing that a protein has the correct, native disulfide bonds has been a lengthy process, frequently involving several days. A recent report indicates that immobilized trypsin columns coupled directly to a liquid chromatograph can shorten this process to 1–2 h in the case of human growth hormone [5,6]. Further coupling to an electrospray mass spectrometer [7] produces an even more powerful system without compromising speed [8].

The dramatic reduction in analysis time achieved with this integrated system results from use of an immobilized enzyme column where the concentration of enzyme is equal to or in excess of analyte concentration. Immobilized enzymes are now widely used in analytical chemistry [9–12]. Proteolysis can be achieved in a few minutes in many cases using this regime. The great advantage of immobilized proteolytic enzymes is that (i) the sample will not be contaminated with proteolytic fragments from autodigestion, (ii) the enzyme is reusable, (iii) very rapid proteolysis is possible because of the high enzyme-to-substrate ratio, (iv) micro-reactions may be achieved in μl volume columns, (v) contamination is diminished, and (vi) immobilized enzyme columns integrate easily into multidimensional analytical systems [13–17]. The established protocol for proteolysis in solution, in contrast, is to use an analyte-to-enzyme ratio of 50/1 in a 24 h digestion. One of the differences in these two approaches is that proteolysis may not be complete in the case of immobilized enzyme column because of the much shorter incubation time. The issue of reaction time in disulfide bond analysis will be addressed in this paper.

Another issue is the particular proteolytic enzyme chosen for the analysis. Although trypsin is the most widely used proteolytic enzyme for protein structure characterization, it is often inadequate for disulfide bond analysis. When the requisite basic amino acids for trypsin cleavage are not appropriately located to allow disulfide bond analysis, another enzyme must be chosen. Glutamyl endopeptidase (Glu-C) is frequently the trypsin alternative [18–20]. The specificity of Glu-C depends on the buffer. Peptide bond hydrolysis occurs on the C-terminal side of both aspartate and glutamate residues in phosphate buffer (pH 7.8), but in either ammonium hydrogencarbonate (pH 7.8) or ammonium acetate (pH 4.0) hydrolysis occurs exclusively at glutamate residues [21]. Altering the specificity of cleavage with Glu-C and a specific buffer frequently produces peptides more amenable to disulfide bond analysis. One of the disturbing features of Glu-C is that it can also catalyze disulfide exchange during proteolysis; often creating new disulfide bonds not found in the native protein [22]. This presents a serious analytical problem. During the course of disulfide bond analysis, artifacts can be generated that falsely imply a protein has un-natural disulfide linkages. The hypotheses to be examined in this paper is that decreasing incubation times through the use of immobilized Glu-C columns will diminish or eliminate the disulfide exchange problem in disulfide bond analysis of native proteins. The model protein chosen for study was insulin.

## 2. Experimental

### 2.1. Materials

Recombinant human insulin, Glu-C and α-cyano-4-hydroxycinnamic acid (α-cyano) were purchased from Sigma (St. Louis, MO, USA). HPLC-grade acetonitrile (ACN) was a generous gift from EM Science (Gibbstown, NJ, USA). HPLC-grade trifluoroacetic acid (TFA), was purchased from Pierce (Rockford, IL, USA). Ammonium hydrogencarbonate was obtained from Fisher Scientific. Peptide mixtures used in the calibration of MALDI-TOF (matrix-assisted laser desorption/ionization time-of-flight) spectra were obtained from PE Biosystems (Framingham, MA, USA).

### 2.2. Instrumentation

Chromatographic analyses were executed on a BioCad perfusion chromatography workstation (PE Biosystems). Analyte detection was achieved by monitoring absorbance at 215 nm. On-column digestions were performed using a tandem column con-

figuration with automated column switching. The first column was a 100×2.1 mm Porozyme immobilized endopeptidase Glu-C cartridge (PE Biosystems). During digestion, the Glu-C cartridge was maintained at a temperature of 37°C using a LC-22A column heater (Bioanalytical Systems, West Lafayette, IN, USA). Peptides generated via on-column and free-solution digestion were analyzed by reversed-phase chromatography using a 250×4.6 mm PepMap $C_{18}$ silica column (PE Biosystems).

MALDI-TOF mass spectra of peptides were obtained with a Voyager-Delayed Extraction Model (DE-RP) Biospectrometry Workstation from PE Biosystems. The instrument was equipped with a nitrogen laser operating at 337 nm and an Omega 48 channel plate detector operated at an overall voltage of 1.8 kV. The parameters of the spectrometer were set at the following settings: reflector and positive modes, the accelerating voltage was 20 kV, grid voltage 58%, 75 ns delayed extraction, and a low mass gate of $M_r$ 300.

### 2.3. Free-solution Glu-C digestion of recombinant human insulin

Digestion was performed in a reaction vessel containing 2 mg/ml of recombinant human insulin and 0.04 mg/ml of Glu-C endopeptidase in digestion buffer (0.05 *M* ammonium hydrogencarbonate, pH 7.8). This produced an enzyme-to-substrate ratio of 1:50 (w/w). Proteolytic digestions were performed in a water bath at 37°C for various lengths of time ranging up to 24 h. Upon completion of digestion, a 1-ml aliquot of the solution was removed and placed in a 1.5-ml Eppendorf tube. Immediately following, the sample was placed in liquid nitrogen in order to quench the enzyme reaction. The sample was then placed in a freezer for further use.

Preceding reversed-phase analysis of the peptide fragments, the sample was removed from the freezer and thawed. The sample was then filtered through a 0.25-μm PVDF (polyvinylidene fluoride) syringe filter (Alltech, Deerfield, IL, USA). The digested insulin, 500 μl, was flushed through the 100-μl injection loop. A 20-μl volume of the sample was injected on the PepMap $C_{18}$ reversed-phase column for analysis.

### 2.4. On-column Glu-C digestion of recombinant human insulin

Prior to insulin digestion on the Porozyme column, the 100-μl injection loop was flushed with 500 μl of insulin solution (2 mg/ml). Following the flushing of the injection loop, the Porozyme column was equilibrated with digestion buffer (0.05 *M* ammonium hydrogencarbonate, pH 7.8) for 10 min at a flow-rate of 1 ml/min. Following equilibration, the Glu-C column was placed in tandem with the reversed-phase column. A 50-μl volume of the insulin solution was then injected into the Glu-C column and buffer flow halted 0.22 min after injection. After a fixed incubation time, ranging from 1–15 min, flow through the Glu-C column was resumed at a rate of 1 ml/min and continued 10 min thereafter. After transfer of the peptide fragments from the Porozyme column onto the reversed-phase column, switching valves were used to take the Glu-C column out of series with the reversed-phase column.

### 2.5. Reversed-phase analysis of peptide fragments

Prior to injection of the digested insulin on the reversed-phase column, the column was equilibrated with mobile phase A [0.1% TFA/1% ACN in double deionized (ddI) water] for 5 min at a flow-rate of 1 ml/min. Analyte elution was achieved isocratically using a linear solvent gradient ranging from 100% mobile phase A to 50% mobile phase B (0.1% TFA/5% ddI water in ACN) over 50 min. The mobile phase was then held at this composition for an additional 5 min. Throughout the analysis, the peptide fragments were observed by an on-line UV detector set at 215 nm. As the fragments were eluted from the reversed-phase column, they were manually collected and placed upon a Speed-Vac until the peptide solution was completely evaporated. Thereafter, MALDI-TOF-MS analysis was performed.

### 2.6. Preparation of peptide solutions for MALDI-TOF-MS

Peptide solutions were prepared by dissolving the individual peptide fractions in the matrix (α-cyano-4-hydroxycinnamic acid) employed in these experi-

ments. A 1-μl volume of this sample solution was applied to the MALDI plate and allowed to air dry. Spectra were calibrated externally using two different peptides – angiostensin I, monoisotopic $m/z$ 1296.69 and insulin (bovine), average $m/z$ 5730.61.

### *2.7. Capillary electrophoresis (CE)*

CE experiments were conducted using a laboratory-made instrument equipped with a fused-silica capillary (Polymicro Technologies, Phoenix, AZ, USA) of 45 cm (30 cm effective length)×75 μm I.D.×360 μm O.D.

A Spellman CZE 1000 R (Spellman High Voltage Electronics, Plainview, NY, USA) power supply was used to apply the electric field across the capillary. The power supply was connected to a 22 gauge platinum electrodes immersed in 1.5-ml buffer reservoirs along with the capillary ends.

High-voltage components of the system were contained in a Lucite cabinet equipped with a safety interlock that would interrupt electric power to the instrument when the cabinet opened.

Detection was performed at 220 nm with a variable-wavelength UV absorbance detector (Model $CV^4$, Isco, Lincoln, NE, USA). A Linear 1200 (Linear, Reno, NV, USA) was used to record the signal from the detector. Separations were performed by applying 16 kV over the capillary.

## 3. Results and discussion

### *3.1. Rate of proteolysis on-column and in free solution*

A comparison of insulin digestion with an immobilized Glu-C cartridge in the stopped-flow mode and in free solution reveals dramatic differences in the rates of hydrolysis (Fig. 1a and b, respectively). The immobilized enzyme system produced 60% hydrolysis in 2 min while less than 40% cleavage was achieved in 60 min in free solution. This large difference is attributed mainly to the fact that the enzyme to substrate ratio is generally 10–100-times larger in the immobilized enzyme system. The immobilized enzyme system obviously has major advantages in terms of speed of analysis, enzyme reuse, the ability to use small samples, ease of use, and the possibility that proteolytic digests may be directly transferred to the reversed-phase column.

Because the proteolytic efficiency of the immobilized enzyme system is so high in the first 2 min of the reaction, it is surprising that the reaction appears to plateau at 70%. Why does the reaction not go to completion? The explanation is that the 50 μl sample volume is larger than that of the immobilized enzyme bed. Although the cartridge has a total volume of 100 μl, estimates by the cartridge manufacturer are that 65% of this volume is contributed by the connecting tubing and fittings. The remaining 35%, i.e., 35 μl, is the actual volume occupied by particles bearing the immobilized enzyme. It is believed that at least 30% of the insulin solution was not exposed to the immobilized enzyme during stopped flow incubation and that no hydrolysis occurred in this volume. In contrast, all the hydrolysis achieved was in 35 μl volume in the bed. Even when a 35 μl sample is used it is difficult to stop the flow at precisely the right time for the complete sample to be in the enzyme bed. This issue is most easily addressed by using sample volumes larger than the working volume of the enzyme cartridge and stopping the flow before all the sample has entered the immobilized enzyme bed.

### *3.2. Disulfide bond exchange in free-solution proteolysis*

Efforts to achieved complete proteolysis in solution involved the use of incubation times ranging up to 24 h. A comparison of the peptide maps from a 2 and 12 h incubation (Fig. 2a and b, respectively) show that longer incubation times reduce the amount of insulin in a sample to zero but that the chromatogram becomes more complex. It will be shown below that this increase in complexity is due to peptide artifacts not predicted from the theoretical cleavage pattern of insulin by Glu-C (Fig. 3). MALDI-MS showed (Table 1) that all the peptide fragments generated by Glu-C in 2 h were as predicted. Less than 0.2% deviation from the theoretical mass was required to assign a specific structure. Three new peaks (peaks 9, 10 and 11) not present in native insulin peptide maps were observed in the 12 h incubation (Fig. 2b). Mass spectrometry indicated

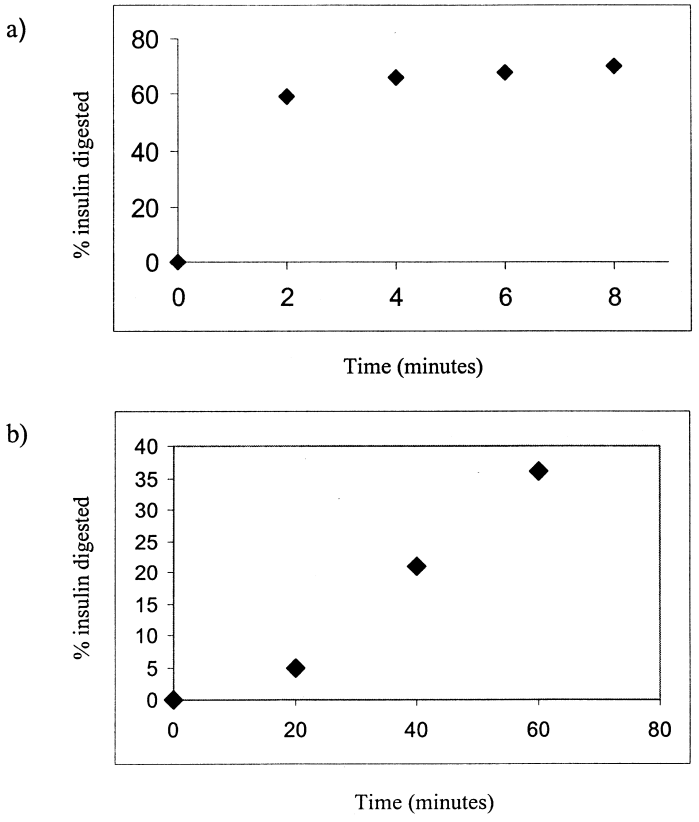


Fig. 1. Digestion of recombinant human insulin via Glu-C. (a) On-column. (b) Free-solution. Percent digestion was access by reversed-phase chromatography based on the amount of insulin remaining after digestion. Chromatographic resolution of sample components was achieved according to the procedure described in Experimental. The initial concentration of insulin was 2 mg/ml in both cases.

that these peptide fragments were A3–B1, A2–B2 and A2–B1. Being that intermolecular disulfide bonded peptides fragment at the disulfide bridge during MALDI-MS [23], the structure assignment of these peptide was made by identifying both the mass of the dipeptide and one of the peptides of which the dipeptide was composed. The counterpeptide was then deduced by subtracting the mass of the di-peptide from the single peptide. None of these disulfide bridged peptides occur in native insulin, even though peaks 5 and 11 are of the same mass as discussed below. It is significant that all three of these new peaks are only observed on extended incubation, after which they become prominent features in the chromatogram. Obviously, they arose by disulfide exchange during extended incubation. It is of critical importance that these peptides would have caused an analyst to conclude that the starting protein is of non-native structure if they were not know to be artifacts.

The fact that peaks 5 and 11 are of the same mass requires further discussion. Theory indicates that a disulfide bridged dipeptide of mass 2967.37 (A2–B1 in Fig. 3) should be produced by Glu-C. This peptide is unique in insulin in that it has an internal disulfide bridge within one of the peptides, A2. It is assumed that the prominent peak (peak 5) of this mass found early in both immobilized enzyme and free-solution digests is, in fact this peptide. Twelve-hour free-solution digests also show this peptide, in addition to a completely different peptide of the same mass (peak 11, in Fig. 2b). The only way this could occur is by a disulfide exchange within the initial dipeptide (peak 5). The fact that the new dipeptide is of the same mass means that B1 is bound to a different cysteine in A2 than in peak 5 and that a new

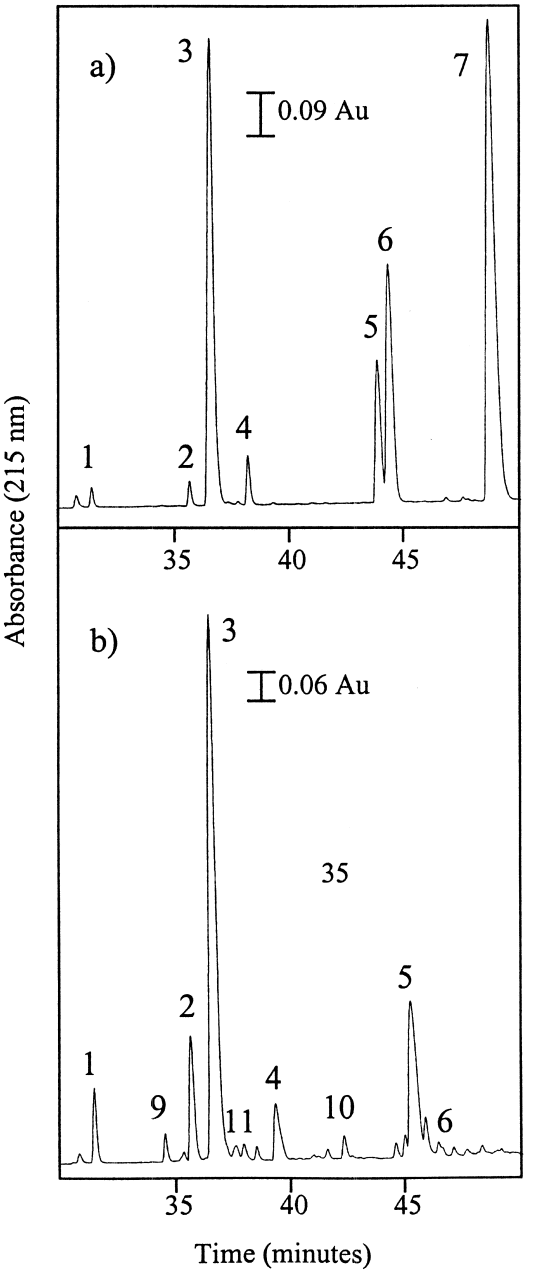


Fig. 2. Free-solution digestions: (a) 2 h and (b) 12 h.

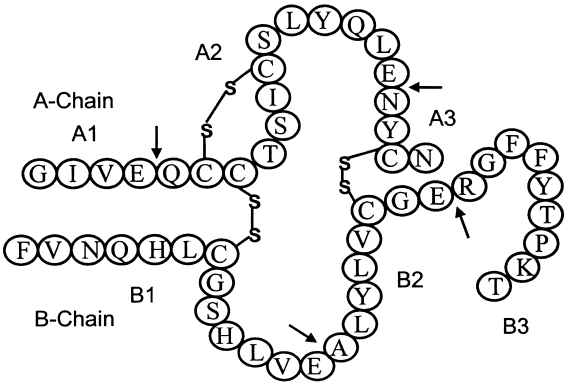


Fig. 3. The sequence and theoretical Glu-C cleavage sites of recombinant human insulin.

intra-peptide disulfide bridge is formed in A2. The most likely candidate is the dipeptide shown in Fig. 4. Although the dipeptides in peaks 5 and 11 are both made up of the peptides A2 and B1, they will differ in conformation. It is not unusual that conformationally different polypeptides are separated in reversed-phase chromatography [24]. In order to further confirm that peak 5 and 11 are indeed identical species, the collected peaks were individually analyzed by CE. The two species have identical migration times using running buffers of various pH (e.g., pH 7–10), indicating that the overall charge-to-mass ratio is identical (data not shown).

### *3.3. Disulfide bond exchange during on-column digestion*

The fact that disulfide bond exchange is greater in free-solution proteolysis than with an immobilized enzyme cartridge must be due to some difference between the two systems. There are two easily identifiable differences. One is that the protein concentration was higher in the case of free-solution hydrolysis. The other is that reaction time was orders of magnitude shorter in the immobilized enzyme cartridge.

Concentration effects were studies by varying protein concentration in the immobilized Glu-C cartridge 10-fold. There was no apparent difference in the chromatographic pattern (Fig. 5). None of the artifact peaks noted in Fig. 2b were observed. In order to ensure that minor fragments were not hidden under major peaks in the chromatogram, segments spanning the entire chromatogram were collected and

Table 1
Mass spectral data for chromatographic peaks[a]

| Peak | Insulin fragment | Theoretical $M_r$ | MALDI-TOF-MS, deviation (%) | | | |
|---|---|---|---|---|---|---|
| | | | a | b | c | d |
| 1 | A3 | 416.23 | 0.18 | | – | – |
| 2 | B1–B2 | 2328.10 | 0.22 | 0.29 | – | – |
| 3 | B3 | 1115.28 | 0.07 | 0.15 | 0.08 | 0.14 |
| 4 | B1 | 1481.67 | 0.07 | 0 | – | – |
| 5 | A2–B1 | 2965.37 | 0.19 | 0.20 | 0.14 | 0.20 |
| 6 | A1–A2–B1 | 3363.67 | 0.20 | 0.18 | 0.11 | – |
| 7 | Insulin | 5801.60 | 0.08 | – | 0.11 | 0.17 |
| 8 | Insulin minus B3 | 4686.23 | – | – | 0.02 | – |
| 9 | A3–B1 | 1989.86 | – | 0.17 | – | 0.16 |
| | B1 | 1480.70 | – | 0.15 | – | 0.14 |
| 10 | A2–B2 | 2351.72 | – | 0.17 | – | 0.15 |
| | A2 | 1483.21 | – | 0.22 | – | 0.23 |
| 11 | A2–B1 | 2967.37 | – | 0.07 | – | 0.08 |
| | A2 | 1483.21 | – | 0.22 | – | 0.23 |

[a] a=2 h free solution digestion, b=12 h free solution digestion, c=15 min on-column digest, d=12 h on-column digest.

analyzed via MALDI-MS. No evidence of disulfide bond scrambling was found.

The impact of incubation time was examined by incubating a sample in the immobilized enzyme cartridge for 12 h. A comparison of the chromatograms obtained from a 15 min and 12 h incubations (Fig. 6) again show that product complexity increased in the long term incubation and artifact peaks 9, 10 and 11 appear as confirmed by MALDI-MS (Table 1). Mass spectral analyses also confirmed that peptide fragments generated during the 15 min incubation period were those predicted by the theoretical Glu-C cleavage pattern of insulin. Hence, the source of disulfide bond scrambling in Glu-C proteolysis is unquestionably due to extended incubation. Eliminating artifact formation requires reaction times not be longer than 1–2 h.

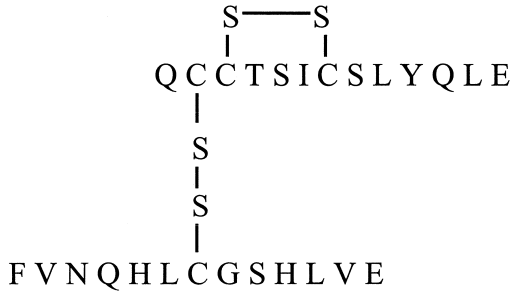


Fig. 4. Possible structure of scrambled A2–B1 dipeptide.

## 4. Conclusions

It may be concluded that disulfide bond scrambling during the digestion of a protein by Glu-C is the result of long incubation periods. This is most likely to occur where enzyme concentration is low and incubation time is necessarily long. Scrambling not only occurs, but the artifacts thus generated become major features of the chromatograms.

Immobilized enzymes were shown to provide a superior alternative to free solution enzymatic digestion. First, complete digestion takes place in a matter of minutes compared to hours. Second, incubation periods are short enough that competing disulfide bond scrambling reactions do not occur to a significant extent. Hence, it can be concluded that the use of immobilized enzymes is a superior alternative as an analytical tool to be used in the structural confirmation of recombinant proteins.

## Acknowledgements

The authors gratefully acknowledge support from

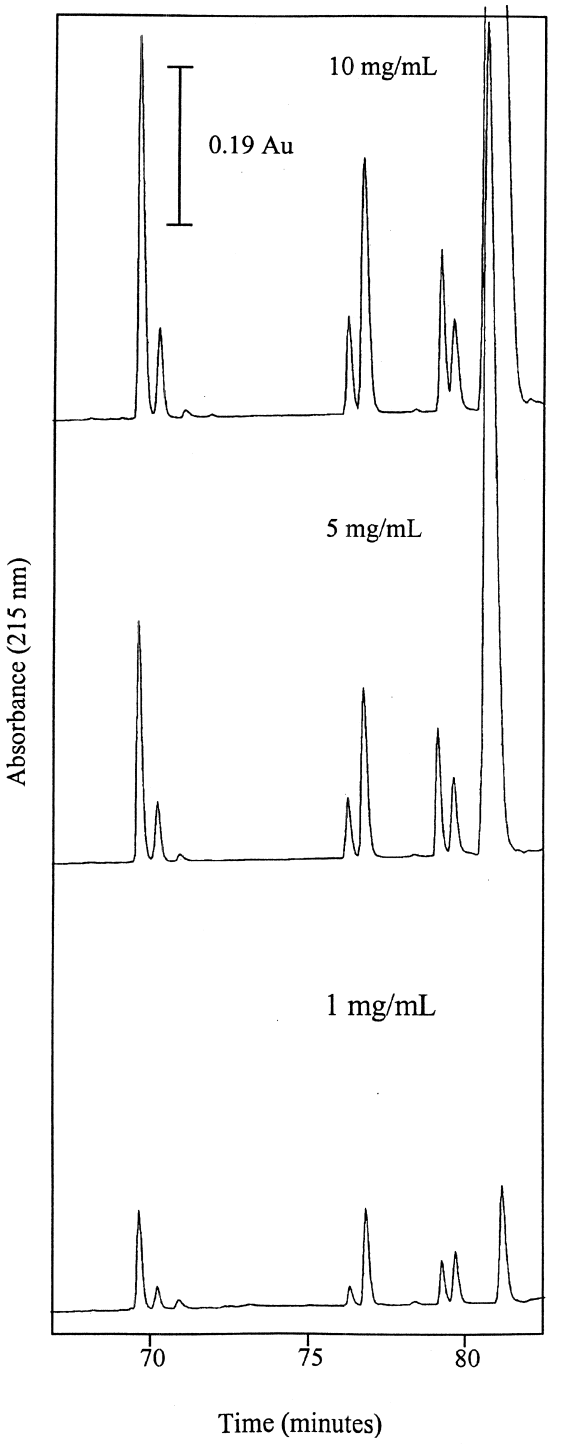


Fig. 5. On-column digestion of recombinant human insulin as a function of increasing concentration.

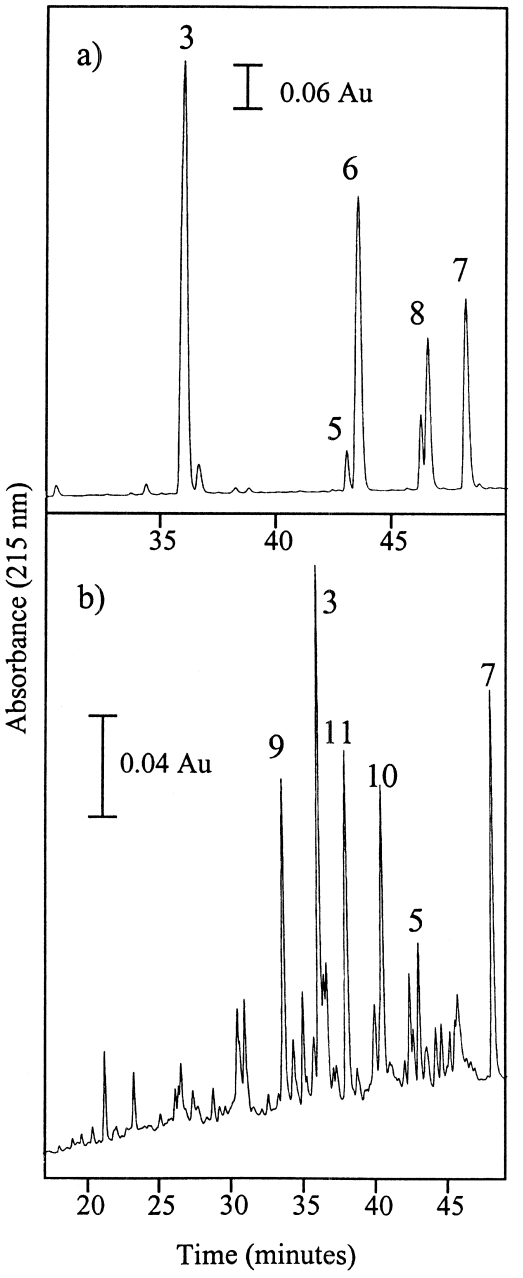


Fig. 6. On-column digestions: (a) 15 min and (b) 12 h.

NIH grant 35421. The authors would also like to thank PE Biosystems for the loan of the Voyager DE-RP Mass Spectrometer and BioCad Workstation and gratefully acknowledge Sara Poehlein for her assistance in the mass spectral analysis and Dr.

Ahmad Amini for his assistance with the capillary electrophoresis experiments.

## References

[1] R.L. Garnick, N.J. Solli, P.A. Papa, Anal. Chem. 60 (1988) 2546–2557.
[2] D. Blohm, C. Bollschweiler, H. Hillen, Angew. Chem., Int. Ed. Engl. 27 (1988) 207–225.
[3] S.K. Sharma, Adv. Drug Deliv. Rev. 4 (1990) 87–111.
[4] T. Nadler, C. Blackburn, J. Mark, N. Gordon, F.E. Regnier, G. Vella, J. Chromatogr. A 743 (1996) 91–98.
[5] C.S. Fullerman, R.H. Wasserman, J. Biol. Chem. 254 (1979) 7208–7212.
[6] L.J. Janis, P.M. Kovach, R.M. Riggins, J.K. Towns, Methods Enzymol. 271 (1996) 86–113.
[7] S.A. Carr, M.E. Hemling, G.D. Bean, Anal. Chem. 63 (1991) 2802–2824.
[8] W.J. Kohr, R. Keck, R.N. Harkins, Anal. Biochem. 122 (1992) 348–359.
[9] L.D. Bowers, P.W. Carr, Anal. Chem. 48 (1976) 544A–558A.
[10] E. Canova-Davis, T.J. Kessler, V.T. Ling, Anal. Biochem. 196 (1991) 39–45.
[11] G.G. Guilbault, Anal. Chem. 38 (1966) 527R–536R.
[12] G.G. Guilbault, Anal. Chem. 42 (1970) 334R–349R.
[13] W.-K. Chui, W. Wainer, Anal. Biochem. 210 (1992) 237–245.
[14] D.N. Gray, M.H. Keyes, B. Watson, Anal. Chem. 49 (1977) 1067A–1078A.
[15] D.N. Nguyen, G.W. Becker, R.M. Riggin, J. Chromatogr. A 705 (1995) 21–45.
[16] J. Ronnenberg, B. Preitz, G. Wostemeier, S. Diekmann, J. Chromatogr. B 656 (1994) 169–177.
[17] R.D. Voyksner, D. Chen, H.E. Swaisgood, Anal. Biochem. 188 (1990) 72–81.
[18] J.J. Birktoft, K. Breddam, Methods Enzymol. 244 (1994) 114–127.
[19] G.R. Drapeau, Can. J. Biochem. 56 (1978) 534–544.
[20] A.C. Ryden, L. Ryden, L. Philipson, Eur. J. Biochem. 44 (1974) 105–114.
[21] J. Houmard, G.R. Drapeau, Proc. Nat. Acad. Sci. USA 69 (1972) 3506–3509.
[22] S. Hara, R. Rosenfeld, H.S. Lu, Anal. Biochem. 243 (1996) 74–79.
[23] S.D. Patterson, V. Katta, Anal. Chem. 66 (1994) 3727–3732.
[24] F.E. Regnier, Science 238 (1987) 319.