

Analysis of Bovine β -Casein Tryptic Digest by Continuous-flow Fast-atom Bombardment Mass Spectrometry

D. S. Jones,* W. Heerma,[†] P. D. van Wassenaar[‡] and J. Haverkamp[†]

Kratos Analytical, Barton Dock Road, Urmston, Manchester, M31 2LD, UK

Liquid chromatography/fast-atom bombardment mass spectrometry was used to partially confirm the amino-acid sequence of the protein, β -casein. The study demonstrates that the technique is capable of the rapid and accurate identification of peptide fragments from tryptic digests and that chromatographic integrity is maintained during the analysis. The power of the technique derives from the ability to determine both the retention time and the molecular weight of the eluting components. Each of the components yields a prominent pseudo-molecular ion (MH^+), the majority exhibiting sufficient fragmentation to confirm their structure unambiguously.

Caseins are phosphoproteins and are present in appreciable quantities in bovine milk. There are four principal types of casein, each of different amino acid composition. These are designated as α s1-, α s2-, β - and κ -casein. The relative proportions of these proteins, present in the micelles, are 38%, 11%, 38%, and 13%, respectively. The aim of this study is to investigate the possibility of determining the primary structure of β -casein by continuous-flow fast-atom bombardment (FAB) mass spectrometry.

A commonly employed method for elucidating the primary structure of large proteins is proteolysis. An enzyme, in this case trypsin, is used to break down the protein into smaller, more manageable fragments. Efficient separation of these complex protein digests has been possible for many years, by using high-performance liquid chromatography (HPLC). However, a major limitation has always been the inability to identify the eluting components unambiguously using conventional detectors. Fortunately, the advent of FAB¹ enabled the analyst to obtain molecular weight and structural information from underivatized peptides with reasonable sensitivity. However, not until recently was it possible to combine the two techniques,^{2,3} thereby allowing on-line separation and identification of individual components. This method has recently been used to analyse the tryptic digest of the protein Lys₇₈-plasminogen.⁴

EXPERIMENTAL

Tryptic digest of β -casein

β -casein was isolated from raw skimmed milk following the method of Fox.⁵ The material, obtained after urea precipitation and dialysis against distilled water, was purified by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). Solutions of casein (10 mg/mL in 0.1 M ammonium bicarbonate pH 8.3), were digested with L-1-tosylamide-2-phenylethyl-chloromethylketone (TPCK)-treated trypsin (Worth-

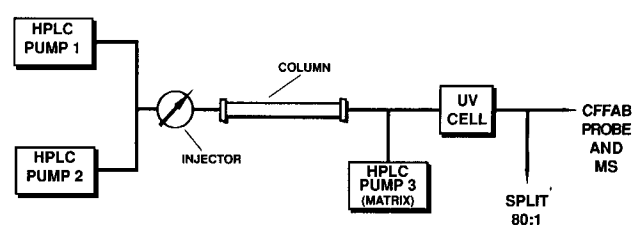


Figure 1. Liquid chromatography separation and inlet system.

ington Biochemical Company, Freehold, NJ, USA) at 37 °C for 16 h (ratio of enzyme to substrate was 1:100), followed by lyophilization.

Mass spectrometer

The study was performed using a Concept 1H mass spectrometer (Kratos Analytical, Manchester, UK), operating at 2000 resolution (10% valley definition). Accelerating voltage was maintained at 8 kV and the magnet was scanned from mass 3500 to 400, and in subsequent experiments, 2750 to 50, at 10 s per decade. A standard FAB gun was used to produce an 8 keV primary xenon atom beam. The most stable conditions were observed at a source temperature of 70 °C and source pressure of 2×10^{-4} Torr. Acquired centroided data were processed using a Kratos Mach 3 data system.

Table 1. Amino acid sequence of bovine β -casein A² variant

	Residue
RELEELNVPG EIVEZLZZZE ESITRINKKI	
EKFQZEEQQQ TEDELQDKIH	50
PFAQTQSLVY PFPGPINSL PQNIPPLTQT	
PVVVPPFLQP EVMGVSKVKE	100
AMAPKHKEMP FPKYPVQPF ESQSLTLTDV	
ENLHLPLPLL QSWMHOPHOP	150
LPPTVMFPPQ SVLSLSQSKV LPVPQKAVPY	
PQRDMPIQAF LLYEQPVLGP	200
VRGPFPIIV	209

Underlined sequences verified by this study.

Z denotes phosphoserine residue.

* Author to whom correspondence should be addressed.

[†] Laboratory of Analytical Chemistry, University of Utrecht, Croesestraat 77A, 3522 AD Utrecht, The Netherlands.

[‡] Unilever Research Laboratory, PO Box 114, 3130 AC Vlaardingen, The Netherlands.

Table 2. Identified peptide fragments (ordered by elution time)

Peptide code	Residues	Identity	MH ⁺	Expected?	Origin
A	106-107	HK	284.2	Yes	t
B	98-99	VK	246.2	Yes	t
C	26-28	INK	374.2	Yes	t
D	26-29	INKK	502.3	Yes	t (p)
E	30-32	IEK	389.2	Yes	t
F	164-169	SLSQSK	649.3	Yes	c, t
G	29-32	KIEK	517.3	Yes	t (p)
H	100-105	EAMAPK	646.3	Yes	t
I	33-48	FQZEEQQQTEDELQDK	2061.8	Yes	t
J	177-183	AVPYPQR	830.4	Yes	t
K	170-176	VLPVPQK	780.5	Yes	t
O	191-193	LLY	408.2	Yes	c
P	194-202	EQPVLGPVR	994.5	Yes	c, t
Q	106-111	HKEMPF	788.4	Yes	t, c
R	108-113	EMPFPK	748.4	Yes	t
T	49-52	IHPF	513.3	Yes	t, c
U	184-190	DMPIQAF	821.4	Yes	t, c
V	33-52	FQZEEQQQTEDELQDKIHFP	2556.1	Yes	t, c (p)
W	191-202	LLYEOPVLGPVR	1383.8	Yes	c, t
X	203-209	GPFPIIV	742.4	Yes	t

t = Due to tryptic action.

c = Due to chymotryptic action.

(p) = Ragged Nor C-terminus peptide.

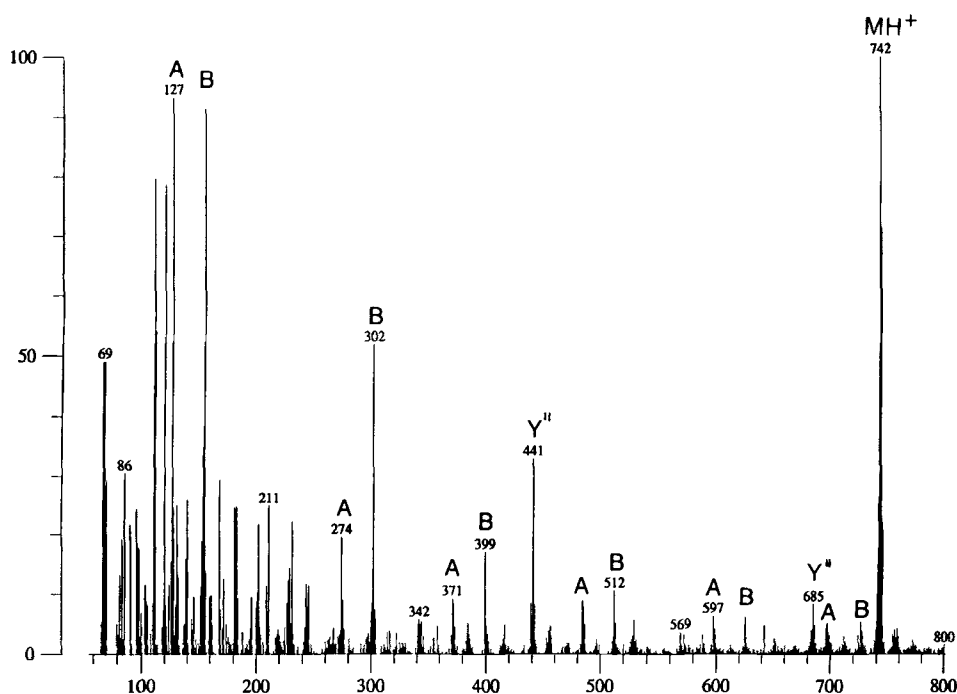
Table 3. Unexpected peptide fragments (ordered by elution time)

Peptide code	Residues	Identity	MH ⁺	Expected?	Origin
L	49-56	IHPFAQTQ	941.5	No	t, ?*
M	114-123	YPVQPFTESQ	1195.6	No	t, ?*
N	49-54	IHPFAQ	712.4	No	Unknown
S	86-91	PFLQPE	730.4	No	Unknown
Y	124-141	SLTLTDVENLHLPLLLQ	2016.1	No	?* ?*

t = Due to tryptic action.

(p) = Ragged Nor C-terminus peptide.

?* Fragments L, M and Y: Yet unknown aspecific cleavage at Q-S probably by chymotryptic activity.

**Figure 2.** Spectrum of a typical digest peptide GPFPIIV (Peptide X), indicating A, B and Y'' backbone cleavage ions.

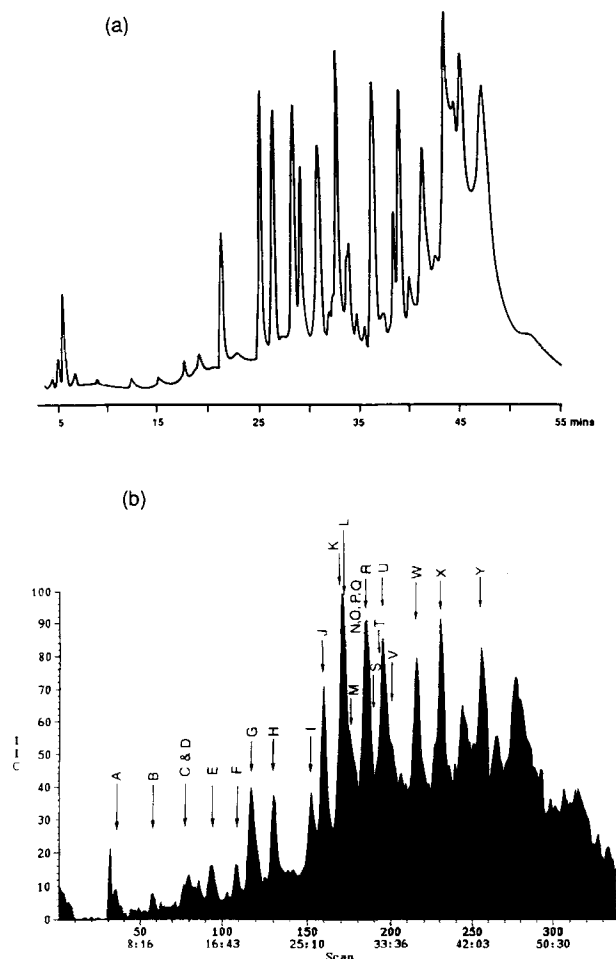


Figure 3. (a) UV and (b) total-ion chromatograms of the β -casein tryptic digest.

Liquid chromatography

Equipment. Binary gradient separation of the digest was carried out using two model LC 9A HPLC pumps (Shimadzu Corporation, Kyoto, Japan). A third LC 9A pump was used for post-column addition of FAB matrix via a zero dead volume tee piece. Separated components were monitored at 214 nm with a Shimadzu SPD-6AV UV detector fitted with a standard 8 μ L flow cell. The column was maintained at a temperature of 30 $^{\circ}$ C during the analysis.

LC separation. Solvent A: 0.06% trifluoroacetic acid (TFA) in water. Solvent B: 0.06% TFA in 80% acetonitrile 20% water. The gradient employed was 100% A, isocratic for 5 min, followed by a linear gradient to 100% B over the next 50 min. The column used for separation was ODS1 Spherisorb (2mm $\times$ 150mm) (Phase Separations, Deeside, UK), with flow rate of 0.2 mL per min. Injection volume was 20 μ L, corresponding to 20 μ g of the total digest. A further 0.2 mL per min of 2% glycerol in acetonitrile was added post-column to provide a FAB matrix. A standard splitting device maintained a constant flow-rate to the mass spectrometer of about 5 μ L per min, using effectively a split ratio of 80:1. A schematic diagram of the liquid chromatograph introduction system is given in Fig. 1.

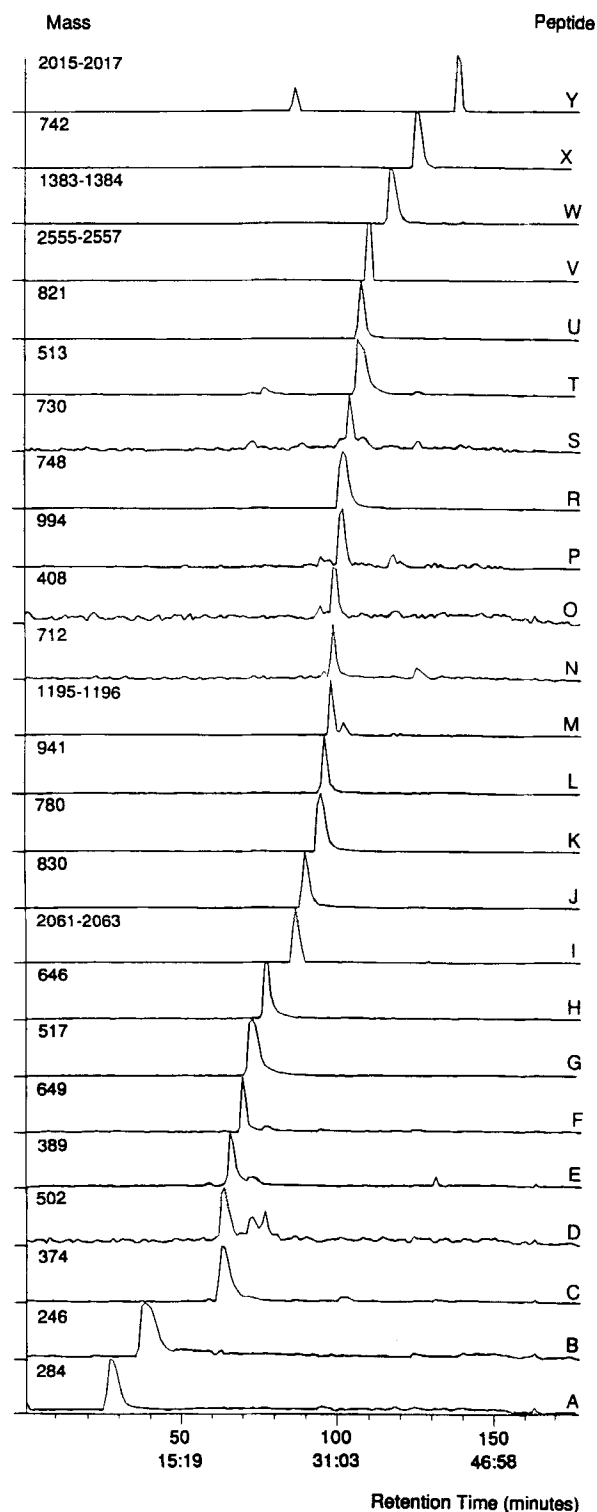


Figure 4. Protonated molecule ion chromatograms of identified peptides.

RESULTS AND DISCUSSION

The amino acid sequence of bovine β -casein (209 amino acids) has been previously documented⁶ and is reproduced as Table 1. This protein (molecular weight approximately 24 000 Da) is readily susceptible to hydrolysis. A study of the theoretically expected tryptic fragments has previously been carried out using liquid chromatography.⁷

A total of 25 peptides were observed in this investigation. Table 2 lists 20 identified components, proto-

nated molecule ions and chemical origin in elution-time order. These can all be explained relatively easily as arising from either tryptic and/or chymotryptic action, indicating that the trypsin used was not completely free of chymotryptic activity. Identified expected fragments were found to elute from the column in the predicted order.⁷ This is further confirmation of correct characterization. However, the 5 remaining peptides (L, M, N, S and Y), shown in Table 3, are rather more puzzling. Peptides L, M and Y each require Q-S cleavages of the casein molecule. These unexpected Q-S cleavages appear to arise from some unusual and unexplained enzyme activity. The assigned structures are, however, supported by observed ion fragmentation. Peptides N and S present even more of a problem and are regarded as highly unlikely from a chemical point of view. The structures shown represent the only possible sequence of amino acids, present in the previously characterized β -casein molecule, capable of yielding the observed protonated molecule ions. These structural assignments should therefore be regarded as rather speculative.

An injection of 20 μ g of digest, approximates to about 1 nmol of each tryptic peptide on-column. After making allowance for the 80:1 split at the interface, only 10 pmol of each component is delivered to the ion source. However, spectra were, in general, adequate to provide amino acid sequence information. A typical spectrum, for the peptide GPFPIIV, with fragment assignments, is shown in Fig. 2. The spectrum contains all six A and B amino terminus fragment ions and two Y'' ions from the carboxy terminus, agreeing well with predicted sequences.⁸

Good correlation was observed between the UV and the total-ion-current chromatogram (Fig. 3(a) and (b)) and is a promising basis for a future method of quantitative study. Upon close scrutiny, there is apparent evidence of slight degradation of chromatographic performance. However, this is dependent upon the selectivity of the detector. Single-ion chromatograms

demonstrate that it is possible to differentiate between spectra of closely eluting components to a far greater degree than is indicated by the total-ion chromatogram. This point is emphasised by the chromatograms of the protonated molecules for each component shown in Fig. 4. The complexity of the digest is demonstrated by the isometric plot of these data (Fig. 5) illustrating the mass, time and intensity distribution of ions from separated components during the course of the analysis.

The identified peptide fragments have been matched to the relevant portion of the β -casein molecule (shown as underlined text in Table 1). Unfortunately, there is a considerable amount of redundancy in these data as many of the peptides originate from overlapping sequences. Consequently, it has only been possible to map approximately 60% of the total protein.

CONCLUSION

The combination of HPLC and continuous-flow FAB mass spectrometry is capable of providing a wealth of useful information to the peptide chemist. Complex mixtures are analysed speedily, and with moderate ease, without resorting to off-line separation of components or derivatization procedures. Results indicate that chromatographic integrity is adequately maintained during analysis. Complete characterization of the entire protein was not possible in the present case, but nevertheless the potential of the general technique is demonstrated. Improved HPLC analysis, using different pH solvent systems, will assist in characterization of possibly missing peptides, that are at present masked by, or unresolved from, predominant hydrophobic species.

Acknowledgement

One of the authors (D.S.J.) would like to express his deep gratitude to Dr. Stephen Naylor of the Medical Research Council, Carshalton, Surrey for his invaluable comments and suggestions regarding interpretative aspects of peptide chemistry.

REFERENCES

1. M. Barber, R. S. Bordoli and R. D. Sedgwick, A. N. Tyler, *J. Chem. Soc. Chem. Comm.* 325 (1981).
2. Y. Ito, T. Takeuchi, D. A. Ishii and M. Goto, *J. Chrom.* **346**, 161 (1985).
3. R. M. Caprioli, T. Fan and J. D. Cottrell, *Anal. Chem.* **58**, 2949 (1986).
4. D. J. Bell, M. D. Brightwell, W. A. Neville and A. West, *Rapid Commun. Mass Spectrom.* **4**, 88 (1990).
5. P. F. Fox, *J. Dairy Sci.* **39**, 49 (1972).
6. A. F. Stewart *et al.*, *Nucleic Acids Res.* **12**, 3895 (1984).
7. L. Leadbeater and F. B. Ward, *J. Chrom.* **397**, 435 (1987).
8. P. Roepstorff and J. Fohlman, *Biomed. Mass Spectrom.* **11**, 601 (1984).

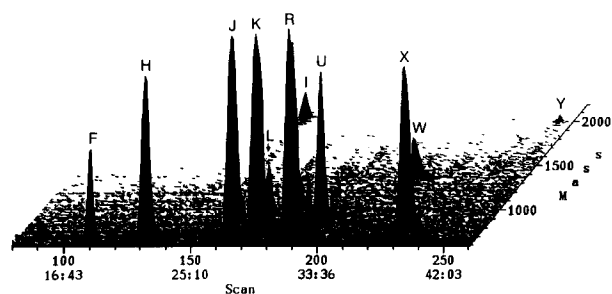


Figure 5. Limited mass range isometric plot of β -casein tryptic digest components.