Vol. 140, No. 1, 1986 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
October 15, 1986 Pages 28-37

ASSIGNMENT OF PHOSPHORYLATION SITES IN BUFFALO β-CASEIN BY FAST ATOM BOMBARDMENT MASS SPECTROMETRY

P.Petrilli*, P.Pucci†, H.R.Morris ‡ and F.Addeo*

* Istituto di Industrie Agrarie, Università di Napoli, 80138 Napoli, Italy
† Dipartimento di Chimica Organica e Biologica, Università di Napoli, 80138 Napoli, Italy
‡ Department of Biochemistry, Imperial College of Science and Technology, London

Received July 11, 1986

Fast atom bombardment mass spectrometry has been applied to the localization of phosphorylation sites in buffalo β-casein.

Two complementary strategies of identification are described. Phosphorylated residues in the tryptic peptide Tp 1 have been assigned by measuring the masses of peptide fragments obtained by enzymatic degradations.

The phosphoserine residue in peptide Tp 2 has been identified by determining the intact molecular weight and confirmed by partial sequence information.

This rapid and sensitive procedure appears of a great interest in structural studies of a wide range of post-translational modifications in proteins.

Following the aminoacid sequence determination of a protein, the final step in primary structure determination is the identification of post-translational modifications and the S-S bridge assignments.

Phosphorylation is one of the major post-translational events which occurs at specific sites along the protein sequence involving a recognition mechanism by protein kinase.

Recently we reported the determination of about 70% of the buffalo β-casein primary structure (1). When this sequence is compared with that of the bovine β-casein, it appears that the two proteins have a considerable degree of homology. In particular, those regions which are phosphorylated in bovine casein are conserved in the buffalo molecule.

0006-291X/86 $1.50

According to Mercier et al.(2) (see review (3) for details), the configuration SER/THR-X-A, where X represents any aminoacid and A an acidic residue, might be a recognition site for phosphorylation of caseins.

The identification of phosphorylation sites in caseins isolated from different species can be helpful to point out other factors involved in the recognition mechanism by protein kinase.

This investigation is a difficult task to achieve by classical sequencing methods because the phosphate esters are easily hydrolyzed under Edman degradation conditions (4).

The applicability of Fast Atom Bombardment mass spectrometry in protein and peptide analysis has already been demonstrated by a number of authors (5-11), including a method for S-S bridge assignment (9) and assignment of phosphorylated peptide sequences (11).

Structural information on proteins can be obtained by analyzing peptide mixtures generated by chemical and/or enzymatic cleavages, and this procedure has been christened "FAB-mapping" (6,7).

This paper reports the identification of phosphorylation sites in buffalo β-casein by FAB-mapping mass spectrometry.

MATERIALS AND METHODS

Buffalo β-casein was cleaved by trypsin in water at 37 ℃ according to (12). The digestion was stopped by lowering the pH to 4.6; the precipitate was collected by centrifugation . The supernatant was first fractionated on a Dowex 50 column; phosphorylated peptides, Tp 1 and Tp 2, were further purified by gel filtration on Sephadex G-25 eluted with 30% acetic acid.

Digestions with thermolysin and Staphilococcus Aureus V-8 protease were performed on purified peptides in 0.4% ammonium bicarbonate pH 8.0 at 37 ℃ for 4 hours. Alkaline phosphatase hydrolysis was carried out in 0.4% ammonium bicarbonate pH 8.5 for 17 hours at 37 ℃.

Positive Fast Atom Bombardment mass spectra were recorded on a VG ZAB HF (High Field) double focusing mass spectrometer equipped with an M-Scan FAB source and an M-Scan FAB gun, using Xenon as primary ionising beam (current 20 μA at 8KeV).

Samples were dissolved in 5% acetic acid and loaded onto a glycerol-coated probe tip; where necessary, thioglycerol and/or 0.1 N HCl were added directly onto the probe.

Sephadex G-25 fine grade was obtained from Pharmacia, Dowex 50 from Bio-Rad, trypsin, thermolysin, glycerol and α-monothioglycerol

from Sigma; alkaline phosphatase and Staphilococcus Aureus V-8 protease were from Miles Laboratories.

Fragment recognition within the peptide sequence was accomplished by a computer program based on molecular weight assignment (9).

RESULTS

According to the phosphorylation code proposed for caseins (2), the buffalo β-casein contains five possible modification sites, all located in two very acidic tryptic peptides corresponding to the N-terminal region of the protein.

The sequences of these two peptides, named Tp 1 and Tp 2, are reported in fig 1 (13); four serine residues in peptide Tp 1 and the single one in peptide Tp 2 are therefore candidates for modification by phosphate groups.

In order to confirm this hypothesis, the two peptides were analyzed by FAB mass spectrometry.

PEPTIDE Tp 1

Fig. 2 shows the positive FAB mass spectrum in the 300-2300 atomic mass units (daltons) range of the thermolytic digest of peptide

Peptide Tp 1

```
 2          5                   10                  15                  20
GLU-LEU-GLU-GLU-LEU-ASN-VAL-PRO-GLY-GLU-ILE-VAL-GLU-SER-LEU-SER-SER-SER-GLU

21             25          28
GLU-SER-ILE-THR-HIS-ILE-ASN-LYS
```

Peptide Tp 2

```
33                          40                  45          48
PHE-GLN-SER-GLU-GLU-GLN-GLN-GLN-MET-GLU-ASP-GLU-LEU-GLN-ASP-LYS
```

Fig. 1: Aminoacid sequences of peptides Tp 1 and Tp 2 from the tryptic digest of buffalo β-casein.

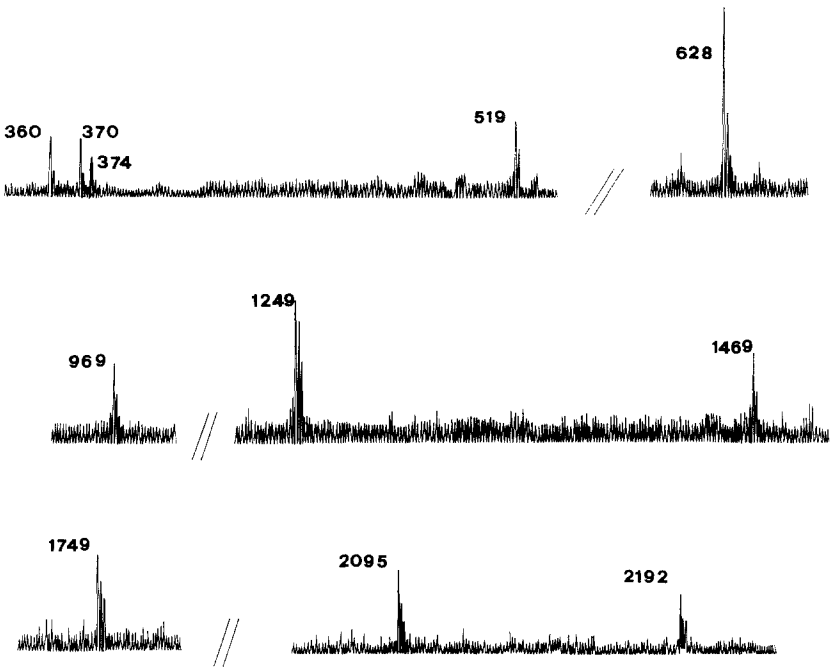

Fig. 2: Positive FAB mass spectrum of unfractionated thermolytic digest of peptide Tp 1.

Tp 1. The recognition of fragments along the peptide Tp 1 sequence has been made according to their molecular weights and the results are listed in Table 1.

The peaks were assigned to the sequences listed in the table, and it should be noted that m/z 1249, 1749, 2095 and 2149 can only be assigned to the sequence by incorporating the addition of phosphate into the molecular weight calculation. Signals at m/z 1249 and m/z 1749 can be assigned to sequences 6-16 and 2-16 respectively only if the presence of <u>one</u> phosphate ester (80 mass units) is incorporated. The molecular ions at m/z 2095 and 2192 can be interpreted as arising from sequences 6-22 and 12-28 respectively, only if <u>four</u> phosphate esters are assumed to be present in each peptide.

Peptides 6-16 and 2-16 contain a single hydroxylated residue, Serine in position 15 allowing unambigous assignment of the first modification site. However, signals at m/z 2095 and 2192 correspond to

TABLE 1
Observed mass values, sequence and position in the sequence of peptides generated by thermolytic digest of peptide Tp 1

Quasi-molecular ion	Sequence	Position in the sequence
360	I-V-E	12-14
370	I-T-H	23-25
374	I-N-K	26-28
519	E-L-E-E	2-5
628	L-N-V-P-G-E	6-11
969	L-N-V-P-G-E-I-V-E	6-14
1249	L-N-V-P-G-E-I-V-E-S-L + 1(P)	6-16
1469	E-L-E-E-L-N-V-P-G-E-I-V-E	2-14
1749	E-L-E-E-L-N-V-P-G-E-I-V-E-S-L + 1(P)	2-16
2095	L-N-V-P-G-E-I-V-E-S-L-S-S-S-E-E-S + 4(P)	6-22
2192	I-V-E-S-L-S-S-S-E-E-S-I-T-H-I-N-K + 4(P)	12-28

sequences containing five and six hydroxylated aminoacids respectively, including the phosphoserine just assigned; therefore, the other three modified residues cannot be identified on the basis of the recorded mass values alone. However, the presence of four phosphate groups in peptide 6-22 clearly indicates that the Threonine residue in position 24 is not modified.

In order to assign the remaining phosphorylated residues, peptide Tp 1 was digested by Staphilococcus Aureus V-8 protease. Fig. 3 shows the FAB mass spectrum of the resulting peptide mixture; the sequences corresponding to the recorded quasi-molecular ions are listed in Table 2. Extending the logic used in interpreting the thermolytic digest data, three signals could only be assigned to the Tp 1 sequence by assuming the presence of phosphate groups.

The signal at m/z 1249 confirms the already assigned phosphoserine residue in position 15. The related signals at m/z 1879

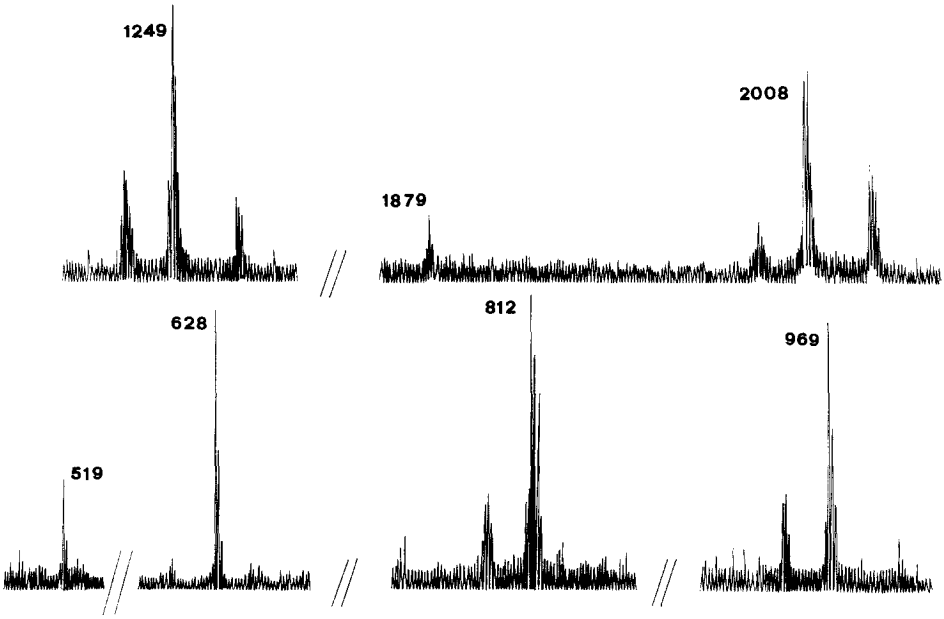

Fig. 3: FAB mass spectrum of unfractionated V-8 protease digest of peptide Tp 1.

and 2008 can be assigned to the peptides 6-20 and 6-21 each containing four phosphate groups. The 6-21 sequence contains only four hydroxylated aminoacids, the serine residues in position 15,17,18 and 19. Thus, each of these must be modified by a phosphate group to fit the molecular weight observed.

TABLE 2

Observed mass values, sequence and position in the sequence of peptides generated by S.A.V-8 protease digest of peptide Tp 1

Quasi-molecular ion	Sequence	Position in the sequence
519	E-L-E-E	2-5
628	L-N-V-P-G-E	6-11
812	S-I-T-H-I-N-K	22-28
969	L-N-V-P-G-E-I-V-E	6-14
1249	L-N-V-P-G-E-I-V-E-S-L + 1(P)	6-16
1879	L-N-V-P-G-E-I-V-E-S-L-S-S-S-E + 4(P)	6-20
2008	L-N-V-P-G-E-I-V-E-S-L-S-S-S-E-E + 4(P)	6-21

These data suggest that neither Ser-22 nor Thr-24 nor His-25 are phosphorylated and this is confirmed by the presence of the free peptide Ser-Ile-Thr-His-Ile-Asn-Lys at m/z 812.

The sequence of tryptic peptide Tp 1 can therefore be written as follows:

```
2                                   10                                  19
GLU-LEU-GLU-GLU-LEU-ASN-VAL-PRO-GLY-GLU-ILE-VAL-GLU-SER-LEU-SER-SER-SER
                                                        |       |   |   |
                                                        P       P   P   P

20                                  28
GLU-GLU-SER-ILE-THR-HIS-ILE-ASN-LYS
```

These assignments were accomplished on the basis of the unique mass values of peptides using the logic described above. However, since ambiguities can occasionally arise, the results were confirmed by incubating the unfractionated peptide mixtures with alkaline phosphatase; the shifts in mass of the respective signals accounted for the loss of the released phosphate moieties in each of the phosphopeptides assigned.

PEPTIDE Tp 2

Figs. 4A and 4B shows the quasi-molecular ion of the phosphorylated peptide Tp 2 of buffalo β-casein, recorded before (4A)

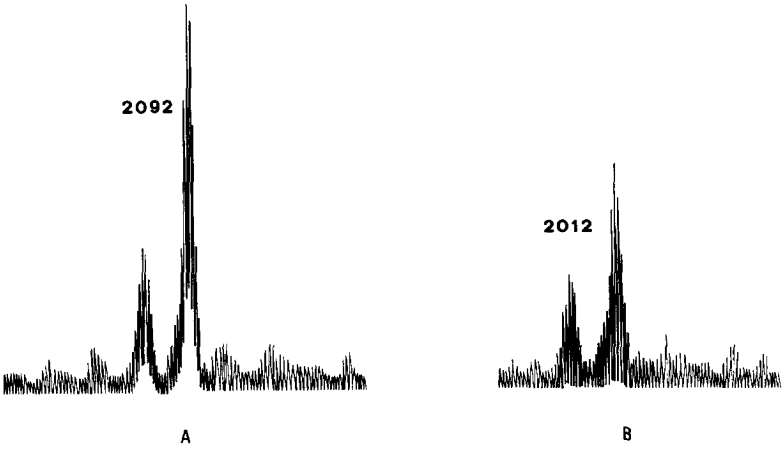

Fig. 4: Quasi-molecular ion region of peptide Tp 2 (A) before and (B) after alkaline phosphatase hydrolysis.

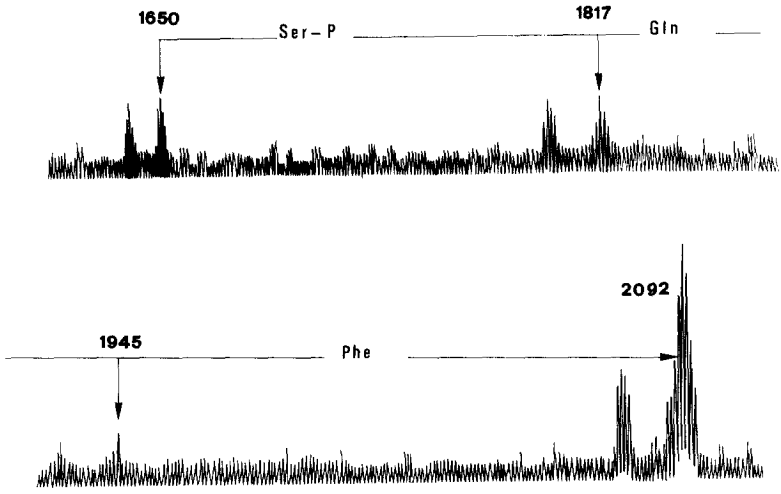

Fig. 5: Partial FAB mass spectrum of peptide Tp 2.

and after (4B) incubation with alkaline phosphatase. The shift in mass value (80 daltons) indicates the pre-existence of a single phosphate group. As peptide Tp 2 has a single serine residue in position 35 (see fig. 1) and because of the enzyme specificity, the modification site is easily assigned. The identification has been confirmed by the sequence information in the FAB mass spectrum of the native peptide (fig 5). The fragment ion at m/z 1650 corresponds to the loss of a phosphoserine residue from position 3 in the N-terminus of the peptide.

DISCUSSION

The assignment of phosphorylation sites in proteins by classical methods remains a difficult problem which is sometimes only resolved by a major committment of time and sample.

Two complementary strategies of identification have been discussed in this paper, both of them using FAB mass spectrometry. The first one is based upon mass spectrometric analysis of proteolytic digests of phosphorylated peptides. The modification sites have been easily assigned by measuring the masses of modified peptides following

the logic we described. This method allowed the assignment of all the phosphorylated serine residues in peptide Tp 1.

The localization of the phosphoserine residue in peptide Tp 2 has been accomplished by measuring the intact molecular weight and by the analysis of the fragmentation ions in the positive FAB spectrum of the phosphorylated peptide. A similar approach has already been described in the characterization of other caseins (11) and riboflavin-binding protein (14); however, it should be pointed out that fragmentation is often absent in FAB spectra.

In both methods, the use of alkaline phosphatase is recommended to confirm the assignment.

Knowledge of the dephosphorylated peptide sequence was required in this study but this is normally the case when work on post-translational modification is initiated. In these circumstances, this analysis of the phosphorylated peptides of buffalo β-casein demonstrates that the combined use of High Field (or either High Mass) mass spectrometry with Fast Atom Bombardment provides a principal method of assignment of post-translation modifications in proteins.

ACKNOWLEDGEMENT

This work was supported by the Medical Research Council ,Great Britain, Programme Grant to HRM; Ministero Pubblica Istruzione, Italy, Progetto di ricerca di interesse nazionale "Livelli di Organizzazione Strutturale di Proteine: metodi di indagine" to P. Petrilli, C.N.R., Italy, Progetto Strategico "Biotecnologie" and Ministero Pubblica Istruzione, Italy, to F.A.

REFERENCES

1. Petrilli, P., Addeo, F. and Chianese L. (1983) Ital. J. Biochem. 32, 336-344
2. Mercier, J.C., Grosclaude, F. and Ribadeau Dumas, B. (1971) Eur. J. Biochem. 23, 41-51
3. Mercier, J.C. (1981) Biochimie 63, 1-17
4. Hamazuma, Y., Mega, T. and Kenaka, T. J. Biochem. (Tokio) 95, 1633-1644

5. Barber, M., Bordoli, R.S., Sedgwick, R.D. and Tyler, A.N. (1981) J. Chem. Soc. Chem. Commun. 325-327
6. Morris, H.R., Panico, M., Barber, M., Bordoli, R.S., Sedgwick, R.D. and Tyler, A.N. (1981) Biochem. Biophys. Res. Commun. 101, 623-631
7. Morris, H.R., Panico, M. and Taylor, G.W. (1983) Biochem. Biophys. Res. Commun. 117, 299-305
8. Morris, H.R., Panico, M., Kamplus, A., Lloyd, P.E. and Riniker, B. (1982) Nature 300, 643
9. Morris, H.R. and Pucci, P. (1985) Biochem. Biophys. Res. Commun. 126, 1122-1128
10. Pucci, P., Borgia, R. and Marino, G. (1985) J. Chromat. 331, 425-431
11. Hunt, D.F., West, D., Panico, M. and Morris, H.R. (1986) Biochem. J. in press
12. Ribadeau Dumas, B., Brignon, G., Grosclaude, F. and Mercier, J.C. (1972) Eur. J. Biochem. 25, 505-514
13. Pucci, P. and Petrilli, P. unpublished results
14. Fenselau, C., Heller, D.N., Miller, M.S. and White III, H.B. (1985) Anal. Biochem. 150, 309-314