Volume 308, number 2, 149–153 FEBS 11416 August 1992

Mapping and sequence-specific identification of phosphopeptides in unfractionated protein digest mixtures by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry

Tai-Tung Yip and T. William Hutchens

Protein Structure Laboratory, USDA/ARS Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, 1100 Bates, Houston, TX 77030, USA

Received 20 May 1992; revised version received 23 June 1992

We have demonstrated a procedure for the rapid (minutes), sensitive (<pmol), and *sequence*-specific identification of phosphopeptides in *unfractionated* digests of phosphoproteins using matrix-assisted UV laser desorption/ionization (MALDI) time-of-flight (TOF) mass spectrometry. The mass-dependent identification of one specific 13-residue phosphopeptide (S105–K117), observed among the 153 possible trypsin digest fragments of human β-casein (211 residues), was confirmed by amino acid sequence analysis of the ^{32}P-labeled peptide after isolation by reverse-phase HPLC. MALDI-TOF was also used to monitor the rate and extent to which an 18-residue N-terminal β-casein peptide (R1–K18) was phosphorylated in vitro. These results demonstrate that MALDI-TOF may be used (i) to facilitate the identification of *sequence*-specific sites of protein phosphorylation and dephosphorylation, (ii) to monitor protein and peptide phosphorylation and dephosphorylation reaction rates, even in complex unfractionated mixtures, (iii) to determine the minimum primary structure necessary for the phosphorylation of specific protein surface domains, and (iv) to evaluate the effects of intact protein phosphorylation and dephosphorylation on susceptibility to subsequent proteolytic events.

Protein sequence; Phosphorylation; Phosphopeptide; Casein; Mass spectrometry; Laser desorption

1. INTRODUCTION

A variety of important biochemical reactions are regulated by the activities of specific kinases and phosphatases (e.g. [1]). The protein and peptide targets of one or more phosphorylation and/or dephosphorylation reactions are often much more readily identified than are the specific amino acid sequences and structural domains involved [1]. To facilitate our understanding of phosphorylation-dependent alterations in the structure and function of proteins and peptides, simplified and more sensitive procedures are required to monitor the kinetics and extent of *sequence*-specific phosphorylation and dephosphorylation reactions in complex unfractionated mixtures. Detailed investigations of sequence-specific phosphorylation reaction mechanisms and higher order structural alterations associated with either phosphorylation or dephosphorylation (e.g. [1–3]) frequently involve the preparation of suitable model peptides as substrates (e.g. [4]). A desirable strategy for the identification and/or design of such model substrates requires knowledge of the minimum primary structure necessary to maintain phosphorylation of the selected protein target residues and domains; these efforts presently necessitate the use of labor intensive biochemical isolation and screening procedures. Finally, there is increasing need to evaluate the effects of specific phosphorylation and dephosphorylation reactions on the susceptibility of the target proteins to subsequent proteolytic events, both in vitro and in vivo [5]. Each of these events involve reactions that result in stable alterations in mass.

We have recently reported the development of mass spectrometric techniques that enable the detailed evaluation of post-translational modifications in peptide structure, including intramolecular disulfide bond formation ([6,7], in preparation), coordinate covalent interactions with specific transition metal ions [6–12], and metal ion-induced conformational changes [7]. Although most of these developments involved the evaluation of purified synthetic peptides, one procedure employs the use of matrix-assisted UV laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF) to map and identify *sequence*-specific transition metal-binding protein surface domains and peptides in *unfractionated* protein digest mixtures [10]. We report here procedures developed for the rapid (minutes), sensitive (<pmol), and sequence-specific identification of phosphopeptides in unfractionated digests of phosphoproteins using MALDI-TOF. Human β-casein was chosen as a model for these investigations because phosphopeptides derived from the proteolytic digestion of this milk protein remain the subject of intense investigations [13], particularly with respect to the regulation

Correspondence address: T.W. Hutchens, Department of Pediatrics (CNRC), Baylor College of Medicine, 1100 Bates, Houston, TX 77030, USA. Fax: (1) (713) 798-7057.

of mineral and metal ion absorption (e.g. [14]) and cell growth (e.g. [15]).

2. EXPERIMENTAL

2.1. *Purification of human β-casein from milk*

Human milk was collected by breast pump and stored at −20°C until use. Thawed milk was dialyzed first against 50 mM EDTA, 0.5 M NaCl (pH 6.5), and then 20 mM phosphoserine in 50 mM 2-(*N*-morpholino)ethanesulfonic acid (MES) and 1 M NaCl (pH 6.5). An aliquot (25 ml) of this sample was pumped into a 50-ml column (2.2 × 13 cm) of immobilized iminodiacetate (IDA)-Fe(III) (Chelating Sepharose Fast Flow, Pharmacia) equilibrated with 20 mM phosphoserine in 50 mM MES containing 1 M NaCl (pH 6.5). After extensive washing with the same buffer to remove all unadsorbed proteins, casein was eluted with 100 mM EDTA, pH 7.0. High-performance reverse phase (phenyl, Vydac) and size-exclusion (Superose 12, Pharmacia) chromatography were used to separate the intact β-casein from casein fragments. The final product was detected as a single silver-stained band after polyacrylamide gradient gel electrophoresis and appeared as a single peak by MALDI-TOF with a determined molecular mass of 24,430 Da; the chemical average molecular mass of non-phosphorylated human β-casein calculated from the cDNA sequence is 23,835 Da [16]. The identity of the purified β-casein was confirmed by amino acid sequence analysis.

2.2. *Enzymatic treatment*

Trypsin (Sigma) digestion of the purified human β-casein was initiated at an enzyme-to-substrate ratio of 1:50 at 37°C in 10 mM Tris-HCl at pH 7.8. Phosphorylation was initiated by incubating the sample in 20 mM Tris-HCl (pH 7.2) containing 9 mM magnesium acetate, 5 mM ATP with 1 μCi [^{32}P]ATP (New England Nuclear, DuPont) and 10 U of cAMP-dependent protein kinase, catalytic subunit (Promega) at 37°C. Dephosphorylation was performed by incubating the sample with 0.25 U of bovine intestine mucosal alkaline phosphatase (Sigma) in 5 mM Tris-HCl (pH 7.8) at 37°C.

2.3. *Separation of casein digest peptides by HPLC*

The casein digest peptides were separated on a C_8 column (4.6 mm i.d. × 250 mm; 10 μm particle size, 10.0 nm pore size, Amicon) with a 10–50% gradient of acetonitrile in 0.1% TFA at a flow rate of 1 ml/min; fractions of 0.2 min each were collected. Eluted peaks were detected by absorbance at 220 nm; each peak was evaluated by MALDI-TOF. An aliquot from alternate fractions was evaluated for ^{32}P by liquid scintillation counting.

2.4. *Synthesis of the model phosphopeptide*

The 18-residue human β-casein N-terminal peptide (R1-K18) was synthesized on an Applied Biosystems Model 430A automated peptide synthesizer using Fmoc/NMP(9-fluorenylmethyloxycarbonyl/*N*-methylpyrrolidone) chemistry (FastMoc, Applied Biosystems, Foster City, CA) and purified using methods described previously [9]. Mass and purity were evaluated by electrospray ionization mass spectrometry and by LDTOF, as described previously [9]; sequence was verified with an Applied Biosystems Model 473A automated amino acid sequence analyzer.

2.5. *Matrix-assisted UV laser desorption/ionization time-of-flight mass spectrometry*

Peptides in the unfractionated β-casein digest were mixed 1:1 (v/v) with matrix; either a saturated aqueous solution of 2,5-dihydroxybenzoic acid (154.12 Da) or a saturated solution of 3,5-dimethoxy-4-hydroxycinnamic acid (224.21 Da) in 30% acetonitrile. 1 μl of this mixture was analyzed on a Vestec Model 2000 mass spectrometer (Vestec Corp., Houston, TX) exactly as described previously [8,9]. All spectra shown were taken in the positive ion mode. Real-time signal averages of multiple (100) laser shots were used to generate each spectrum. Calibration was performed before and verified after each sample analysis using bovine insulin (5734.5 Da).

3. RESULTS AND DISCUSSION

Intact human β-casein, peptides resulting from the digestion of this protein with trypsin, and a synthetic 18-residue peptide from the N-terminus of casein (R1–K18) were used as model protein and peptide substrates for these investigations.

Peptides were generated by the digestion of native human β-casein with trypsin and evaluated by MALDI-TOF before and after phosphorylation by incubation with cAMP-dependent protein kinase (Fig. 1). Mass differences in the spectra corresponding to 80 Da (and multiples thereof) revealed that several peptides were phosphorylated by the treatment with this kinase. One goal of this investigation, however, was to identify phosphorylation sites on the intact protein that might represent stable domains and unaltered kinase substrates even after fragmentation by cleavage with trypsin. Therefore, the identities and relative quantities of specific phosphopeptides in the unfractionated protein digest maps were also evaluated by initiation of phosphorylation before trypsin digestion of the intact β-casein (not shown). The mass accuracy (0.01% up to 30 kDa) afforded by MALDI-TOF allowed the mass-dependent identification of specific peptides that were phosphorylated equally well both before and after trypsin digestion.

One specific phosphopeptide was identified by phosphorylation of the intact β-casein prior to digestion with trypsin and also after phosphorylation of the unfractionated trypsin digest mixture. This peptide, with a protonated molecular mass $[M+H]^+$ determined to be

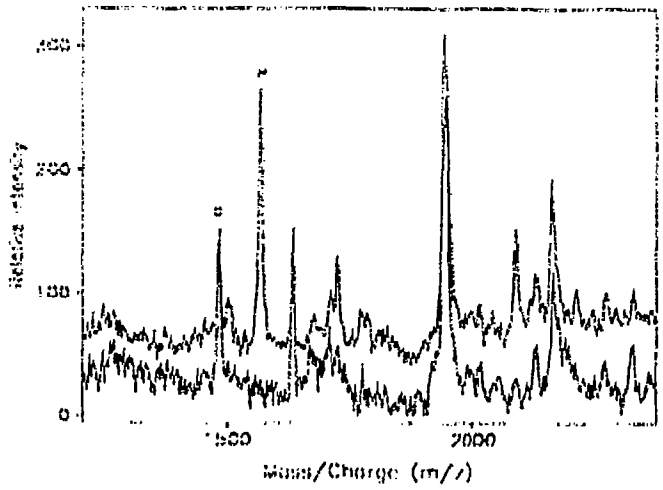

Fig. 1. Proteolytic digest map of unfractionated human β-casein fragments obtained by MALDI-TOF before (bottom) and after (top) phosphorylation of the peptide digest mixture by incubation with cAMP-dependent protein kinase. The specific increase in mass of 80 Da (peak P at m/z = 1567.1) observed for one particular peptide (peak 0 at m/z = 1487.1 Da) in the digest mixture is indicated.

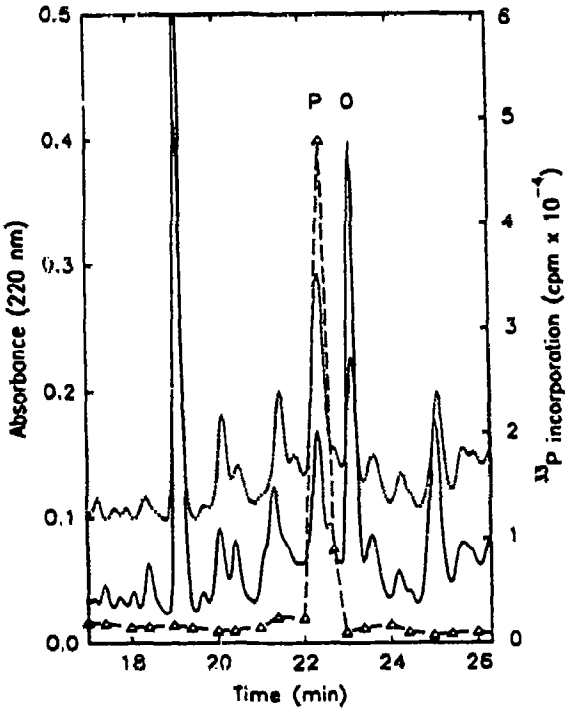

Fig. 2. Reverse-phase HPLC profile of the human β-casein trypsin digest mixture before (bottom) and after (top) phosphorylation. The elution position of ^{32}P-labeled phosphopeptides is also shown (triangles). The mass of each peptide was determined by MALDI-TOF; the mass spectrum of peaks labeled O (before phosphorylation) and P (after phosphorylation) are shown in Fig. 3.

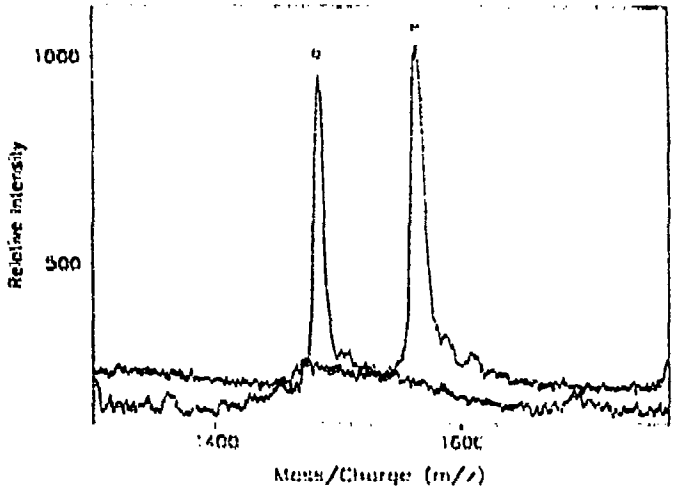

Fig. 3. Mass spectra of the peptides isolated by HPLC before (peak O) and after (peak P) phosphorylation of the unfractionated casein digest in vitro with ^{32}P-labeled ATP. The protonated molecular mass $[M+H]^+$ values determined for peptides O and P were 1487.9 and 1567.9 Da, respectively.

1487.1 Da before phosphorylation and 1567.1 Da after phosphorylation, was tentatively identified as the 13-residue sequence S105–K117 using the mass and sequence analysis program PROCOMP as described previously [10]. Although human β-casein (211 residues) contains 153 possible trypsin digest products (theoretical fragments from the partial and complete digestion at 11 Lys residues and 3 Arg residues), there are no other peptides that can be generated by tryptic digestion alone with mass values near 1486.74 Da. Nevertheless, the chemical identity of this peptide was confirmed by phosphorylation of β-casein with ^{32}P followed by trypsin digestion and isolation of the ^{32}P-labeled peptide fragments by reverse-phase HPLC (Fig. 2).

The ^{32}P-labeled phosphocasein peptides illustrated in Fig. 2, and peptides generated by trypsin digestion of the casein before its phosphorylation (negative control), were isolated and analyzed individually by MALDI-TOF. The peak labeled P in Fig. 2 was eluted with a retention time that was coincident with the elution of a major peak of ^{32}P activity; this peak was found to have a mass of 1566.9 Da. The later eluting peak labeled O in Fig. 2 was determined to have a mass of 1486.9 Da. Fig. 3 shows overlapping mass spectra for the two isolated casein peptides (O and P). The mass values determined for the peptides actually isolated before and after phosphorylation were within 0.05% of the mass values observed by MALDI-TOF for the unfractionated digest mixtures before and after phosphorylation; these values were essentially the same as mass values calculated for the casein peptide defined by sequence S105-K117 before (1486.74 Da) and after (1566.72 Da) one phosphorylation event. Finally, amino acid sequence analyses of the two isolated peptides (i.e. peaks O and P) confirmed the sequence identity (S105–K117) suggested from the determination of mass alone (Table I).

The kinase-treated casein peptide digest mixtures were also analyzed by MALDI-TOF before and after subsequent dephosphorylation with acid and alkaline phosphatase (not shown). Phosphorylation and

Table I

Determined mass values for the human β-casein peptides R1–K18 and S105–K117 observed by MALDI-TOF before and after phosphorylation

	Protonated molecular mass $[M+H]^+$ (Da)	
	Calculated[a]	Observed[b]
RETIESLSSSEESITEYK		
Before phosphorylation	2,089.23	2,089.4 ± 0.6
After phosphorylation	2,169.21	2,169.3 ± 0.9
SPTIPFFDPQIPK		
Before phosphorylation	1,487.74	1,487.9 ± 0.4
After phosphorylation	1,567.72	1,567.9 ± 0.4

[a] Calculated chemical average mass values for the 18-residue (R1–K18) and the 13-residue (S105–K117) human β-casein peptides were derived from the published cDNA sequence [16].

[b] Average mass observed (± S.D.) from a minimum of 6–8 separate determinations.

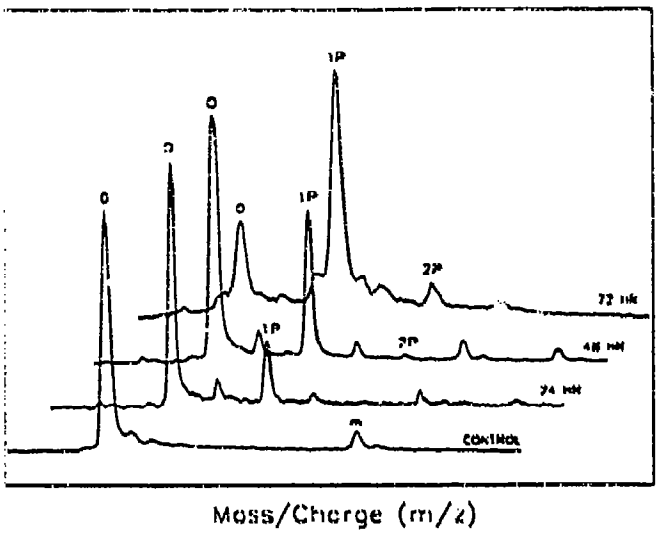

Fig. 4. Mass spectra showing the progressive phosphorylation of a synthetic 18-residue β-casein peptide (R1–K18). The original peptide (0), the *mono*phosphorylated peptide (1P), the *di*phosphorylated peptide (2P), and a small amount of the peptide observed with a sinapinic acid matrix adduct (M), are as indicated.

dephosphorylation was indicated by a sequential increase then decrease in mass corresponding to 80 Da. In addition, dephosphorylation of fragments resulting from the trypsin digestion of native (untreated) casein allowed us to identify one peptide with 4 phosphate groups as R1–K18 (2088.4 Da after dephosphorylation); the mass-dependent identification of the R1–K18 phosphopeptide was also confirmed by isolation and amino acid sequence analysis. In general, because there are no amino acid residues with a mass of 80 Da, identification of phosphopeptides by this procedure can be unequivocal. Our isolation and sequence analysis of the ^{32}P-labeled casein peptide is definitive evidence of the validity of this approach.

Although a specific increase in mass of 80 Da is associated with phosphorylation, other factors may complicate the identification of phosphopeptides by these procedures. A difference in mass of 102 Da is often observed for phosphopeptides in the presence of Na^+ because relatively stable Na adducts can be formed during MALDI-TOF. As we have demonstrated previously with other peptides [9,12], however, these Na adducts may be eliminated by washing the air-dried sample in distilled water. A mass increase of 102 Da can also be generated by the formation of phosphate (H_2PO_4) or sulfate (H_2SO_4) adducts. Again, these can also be eliminated or greatly reduced by removal of the salts that formed after the sample has air-dried on the probe tip. A specific phosphorylation event can be distinguished from these other possibilities by evaluation of reaction completion over time and by the enzymatic reversal of the process (i.e. incubation with phosphatase).

MALDI-TOF was also effective as a rapid (minutes) and efficient (<pmol) means to monitor both the rate and extent of synthetic peptide phosphorylation. The R1–K18 portion of the N-terminus of human β-casein (Table I) was chemically synthesized and purified as described in section 2. Phosphorylation of this bioactive peptide alters its iron-binding capacity and also its trophic activity in vitro (manuscripts in preparation). Fig. 4 shows the collection of MALDI-TOF mass spectra obtained at several different times after incubation of the R1–K18 casein peptide with kinase to achieve phosphorylation. A total of less than 10 pmol of this peptide was consumed to generate these spectra. Furthermore, using the same sample (i.e. on the same probe tip), after analyses of the metal-free apopeptide, the Fe(II)-, Fe(III)- and Ca(II)-binding properties were subsequently probed by MALDI-TOF (data not shown) using the in situ probe tip reaction strategy recently described elsewhere [11].

In summary, MALDI-TOF has proven to be an extremely useful new technology for the identification and evaluation of phosphopeptides, particularly in the case of unfractionated mixtures. Although phosphopeptide detection can be favored in the negative ion mode, for the purpose of these investigations we found no significant or consistent advantages to the evaluation of spectra collected in this mode. In fact, a diminished signal intensity for the corresponding unphosphorylated (or dephosphorylated) peptides was sometimes observed in the negative ion mode, thus creating a disadvantage.

The utility of this approach is dependent primarily on mass resolving power (i.e. $m/\delta m$ at FWHM) and not necessarily on the complexity of the digest mixture; phosphopeptides arising from the enzymatic digestion of proteins in excess of 75 kDa have been identified using the approach outlined here (unpublished). The availability of reflectron MALDI-TOF instruments with mass resolving powers of up to 2000 should allow the evaluation of phosphorylation reactions even on intact proteins with molecular weights in excess of 80 kDa.

Acknowledgements: We thank Chee Ming Li and Gary L. Cook in our laboratory for peptide synthesis and amino acid sequencing. Funded, in part, with federal funds from the US Department of Agriculture, Agricultural Research Service under Cooperative Agreement number 58-6250-1-003. The contents of this publication do not necessarily reflect the views or policies of the US Department of Agriculture, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

REFERENCES

[1] Taylor, S.S., Beuchler, J.A. and Yonemoto, W. (1990) Annu. Rev. Biochem. 59, 971–1005.
[2] Sood, S.M., Chang, P. and Slattery, C.W. (1990) Arch. Biochem. Biophys. 277, 415–421.
[3] Hollosi, M., Urge, L., Perczel, A., Kajtar, J., Teplan, I., Otvos, L. and Fasman, G.D. (1992) J. Mol. Biol. 223, 673–682.

[4] Meggio, F., Perich, J.W., Meyer, H.E., Hoffman-Posorske, E., Lennon, D.P.W., Johns, R.B. and Pinna, L.A. (1989) Eur. J. Biochem. 186, 459–464.
[5] Meisel, H. and Frister, H. (1988) in: Milk Proteins: Nutritional, Clinical, Functional and Technological Aspects (Barth, C.A. and Schlimme, E. eds.) pp. 150–154, Springer-Verlag, Darmstadt.
[6] Allen, M.H. and Hutchens, T.W. (1992) Rapid Commun. Mass Spectrom. 6, 308–312.
[7] Hutchens, T.W. and Allen, M.H. (1992) Rapid Commun. Mass Spectrom. 6 (in press).
[8] Hutchens, T.W., Nelson, R.W. and Yip, T.-T. (1991) J. Mol. Recog. 4, 151–153.
[9] Hutchens, T.W., Nelson, R.W., Allen, M.H., Li, C.M. and Yip, T.-T. (1992) Biol. Mass Spectrom. 21, 151–159.
[10] Hutchens, T.W., Nelson, R.W. and Yip, T.-T. (1992) FEBS Lett. 296, 99–102.
[11] Nelson, R.W. and Hutchens, T.W. (1992) Rapid Commun. Mass Spectrom. 6, 4–8.
[12] Hutchens, T.W., Nelson, R.W., Li, C.M. and Yip, T.-T. (1992) J. Chromatogr. 604, 125–132.
[13] Fiat, A.-M. and Jolles, P. (1989) Mol. Cell. Biochem. 87, 5–30.
[14] Yuan, Y.V. and Kitts, D.D. (1991) Nutr. Res. 11, 1257–1272.
[15] Shimizu, M., Kato, M., Miwa, K. and Kaminogawa, S. (1991) Biochem. Int. 24, 1127–1133.
[16] Menon, R.S. and Ham, R.G. (1989) Nucleic Acids Res. 17, 2869–2871.