Mass Spectrometric Analysis of 21 Phosphorylation Sites in the Internal Repeat of Rat Profilaggrin, Precursor of an Intermediate Filament Associated Protein†

K. A. Resing,*,‡ R. S. Johnson,§ and K. A. Walsh§

Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80303, and Department of Biochemistry, University of Washington, Seattle, Washington 98195

Received January 9, 1995; Revised Manuscript Received May 4, 1995⊗

ABSTRACT: Profilaggrin, a highly phosphorylated protein synthesized in mammalian cornified epithelia, is the precursor of filaggrin, a protein that is involved in aggregation of keratin during terminal differentiation. Possible functions for the phosphorylation include preventing premature aggregation of keratin, packing profilaggrin into a storage granule, association of other proteins with the granule, and/or regulating proteolytic processing of profilaggrin. As a first step in characterizing the phosphorylation of rat profilaggrin, tryptic peptides of filaggrin and profilaggrin were fractionated by reverse-phase HPLC and analyzed by ionspray mass spectrometry. Nine putative phosphopeptides were identified as those with masses 80 Da (or multiples of 80 Da) greater than the predicted unphosphorylated masses. The six that were phosphorylated to a high stoichiometry were analyzed further. Several multiply phosphorylated peptides underwent neutral loss of H_3PO_4 during collisional activation, complicating interpretation of the MS/MS spectra. In order to circumvent this problem, an alternative strategy was applied in which peptide mixtures were treated with $Ba(OH)_2$, resulting in β-elimination of H_3PO_4 and generation of dehydrated serine or threonine at the site of phosphorylation. Peptides containing dehydrated serine or threonine fragmented well, providing unequivocal identification of multiple phosphorylation sites in peptides as long as 39 amino acids. The phosphopeptides (with phosphorylated residues underlined) were GQQHSGHPQVYYYGVEETEDESDAQQGHHQQQQQQR, GGQAGSHSESEASGGQAGR, HTSRPEQSPDTAGR, GESPAGQQSPDR, EASASQSSDSEGHSGAHAGIGQGQTSTTHR, and GSSESQASDSEGHSDYSEAHTQGAHGGIQTSSQR.

Aggregation of keratin intermediate filaments in cornifying mammalian epithelia involves filaggrin [for a review, see Dale et al. (1993)]. The expression of this protein in the cell must be closely regulated in order to prevent premature aggregation. Filaggrin is synthesized as a precursor, profilaggrin, which has multiple internal repeats (PIR,[1] profilaggrin internal repeat) whose size and number varies between species or even strains (Gan et al., 1990; Haydock & Dale, 1990; McKinley-Grant et al., 1989; Resing et al., 1985; Rothnagel & Steinert, 1990), with approximately 21 PIRs in the Sprague-Dawley rat profilaggrin used in this study (Haydock & Dale, 1990). Proteolytic processing to release profilaggrin occurs at a linker site that is near the middle of each PIR, so that the boundary of each PIR does not correspond to the termini of the filaggrin (Gan et al., 1990; Resing et al., 1989, 1993b). The number of repeats varies between species, as well as between individuals or strains (Gan et al., 1990), with approximately 20 repeats in the Sprague-Dawley rat profilaggrin used in this study (Haydock & Dale, 1990). Two independently regulated endoproteinases carry out the proteolytic processing in two stages (Resing et al., 1989, 1993a), apparently followed by exopeptidase processing of the amino and carboxyl termini of filaggrin (Resing et al., 1993b). Profilaggrin also has a 50 kDa calcium binding domain at the amino terminus (Markova et al., 1993; Presland et al., 1992) and a short carboxyl-terminal extension beyond the terminal PIR (Gan et al., 1990; Haydock & Dale, 1990; Rothnagel & Steinert, 1990); the function of these is unknown.

Synthesis of profilaggrin occurs just before cornification and involves extensive phosphorylation (approximately 400 mol/mol) by several kinases and deposition in a cytosolic nonmembrane bound granule (the keratohyalin granule). Phosphorylation of profilaggrin may prevent association with keratin, facilitate packing into the granule, and/or act as a docking site for other proteins in the granule. At terminal differentiation, the granules disperse as the profilaggrin is dephosphorylated by one or more phosphatases (Haugen-Scofield et al., 1988), including phosphatase 2A (Kam et al., 1993), and proteolyzed in a multistep process to yield filaggrin (Resing et al., 1989, 1993a). A role for profilaggrin phosphorylation in regulating the early proteolytic events has been proposed (Resing et al., 1993b); however, it may play other, unforeseen roles in the differentiation program leading to the cornification of epithelia.

The PIRs display a high degree of similarity to one another within an animal species (95% identity in mouse, 85% in human) but vary greatly between species (30−42% between

† Supported by NIH Grants AR39730 (K.A.R.) and RR05543 (K.A.W.) and by the University of Washington Royalty Research Fund (R.S.J.).

* To whom correspondence should be addressed at the Department of Chemistry and Biochemistry, Campus Box 215, University of Colorado, Boulder, CO 80309. Phone: 303-492-4604. Fax: 303-492-5394.

‡ University of Colorado.

§ University of Washington.

⊗ Abstract published in *Advance ACS Abstracts,* July 1, 1995.

[1] Abbreviations: ESI/MS, electrospray ionization mass spectrometry; Da/e, daltons per unit charge; HPLC, high-pressure liquid chromatography; LC/MS, liquid chromatography coupled to mass spectrometry; MS/MS, tandem mass spectrometry; *m/z*, mass to charge ratio; PIR, profilaggrin internal repeat.

0006-2960/95/0434-9477$09.00/0

mouse and human) (Gan et al., 1990; Haydock & Dale, 1990; Rothnagel & Steinert, 1990). Human, rat and mouse PIRs show limited areas of sequence identity which are most extensive around the site of proteolytic processing, extending 40 and 100 residues on the amino and carboxyl sides, respectively (McKinley-Grant, 1989; Haydock & Dale, 1990). Although the phosphorylated tryptic peptides of mouse profilaggrin have been identified (Resing et al., 1985), not all of those peptides are recognizable within the profilaggrin sequence from other species. To understand the roles of phosphorylation of this protein, it would be useful to identify the phosphorylation sites in several species; however, the large size, unusual composition, multidomain nature, and high extent of phosphorylation of these proteins has complicated the analysis of those containing the larger PIRs (mouse filaggrin is 26 kDa, whereas rat and human filaggrin are 42 and 37 kDa, respectively).

Recent developments in mass spectrometry have greatly facilitated analysis of posttranslational modifications, such as phosphorylation. Recently, this technology proved useful in understanding proteolytic processing of rat profilaggrin (Resing et al., 1993b). In this technique, nebulization-assisted electrospray ionization (ESI/MS) is used to volatilize peptides for mass analysis [recently reviewed in Siuzdak (1994)] with a triple quadrupole mass analyzer. This process allows on-line analysis of HPLC eluates, facilitating characterization of phosphorylation sites without the need for purification of the phosphopeptides. Sequence information from peptides in mixtures is obtained by selecting a precursor ion of particular mass to charge ratio (m/z) in the first of three quadrupole m/z-analyzers (Q1), collisionally activating the selected ion in the second quadrupole (Q2), and analyzing the resulting product ions in the third quadrupole (Q3). The overall process is known as tandem mass spectrometry or MS/MS (Biemann, 1990a,b; Hunt et al., 1986). In this study, these new techniques were employed to identify the phosphorylation sites on rat profilaggrin.

MATERIALS AND METHODS

Preparation of Proteins and Peptides. Epidermis was prepared from cervically dislocated, 2−3 day old, Sprague-Dawley rats which were heated in phosphate buffered saline with 10 mM EDTA for 3−4 min, so that the epidermis could be peeled from the dermis, eliminating contact of the skin with blood. Profilaggrin was extracted from the epidermis at room temperature with 9 M urea, 10 mM EDTA, 10 μg/mL phenylmethanesulfonyl fluoride, 10 mM glycerol phosphate, and 50 mM Tris, pH 7.8. The extracts were vortexed to incorporate air and allowed to oxidize at room temperature for 1 h; this precipitated the keratins, which were removed by centrifugation at 22 000g for 30 min at 20 °C. The supernatant was applied to DE52 equilibrated in 9 M urea and 50 mM Tris, pH 7.8, at room temperature; filaggrin did not bind and profilaggrin was eluted with a gradient from 0 to 0.4 M NaCl in the same buffer. Filaggrin was purified further as previously reported (Resing et al., 1993b); profilaggrin was precipitated with 75% MeOH or by 10-fold dilution with H_2O. Tryptic digests were prepared by suspending 50−100 μg of each protein in 200 mM Tris and 2 mM $CaCl_2$, pH 8, and digesting with 1−5% (w/w) trypsin at 38 °C for 2 h. Digests were acidified with trifluoroacetic acid and analyzed on Hewlett Packard C18, Synchropak C18, or Ultrasphere C18 reverse-phase columns. In some cases, digests were acidified with phosphoric acid, following a protocol previously developed to minimize nonspecific association of [^{32}P]ATP to peptides (Haugen-Scofield et al., 1988; Resing et al., 1993). To β-eliminate phosphate from peptides in a tryptic digest, solid $Ba(OH)_2$ was added to saturation, and the sample was incubated for 2 h at 30 °C under argon. To stop the reaction, the sample was acidified with HCl until the $Ba(OH)_2$ was solubilized; the acidified sample was applied immediately to an HPLC column for analysis. It should be noted that profilaggrin lacks cysteine, eliminating the need for reduction and alkylation. If alkylating agents are used, all remaining reagent must be removed before treatment with $Ba(OH)_2$ to prevent modification of primary amines and histidines, which would otherwise eliminate the positive charges that enable detection by ion spray MS.

Mass Spectrometry. For direct coupling to the mass spectrometer (LC/MS), an Applied Biosystems Model 140A liquid chromatograph was connected directly to the mass spectrometer, with flow rates between 200 and 300 μL/min. A postcolumn split directed approximately 10−15% of the HPLC effluent to the mass spectrometer and the remainder to an ABI Model 785A UV detector with fractions collected by hand. Fractions were lyophilized, resuspended in 0.1% formic, 25 or 50% MeOH for MS/MS. Chromatography on 500 micron capillary columns packed with POROS resin was carried out as previously described (Resing et al., 1995).

Mass and most of the sequence determinations were carried out as previously described (Resing et al., 1993b) on a triple quadrupole mass spectrometer (API-III, Sciex, Thornhill, ON, Canada) equipped with an ionspray ion source. MS/MS on the peptide P46 was obtained using a high pressure collision cell, as described in Resing et al. (1995). Interpretation of the MS/MS spectra was carried out with the assistance of computer programs written in-house, as previously described (Resing et al., 1993b). Other programs used were provided with the Sciex instrument, except that MS and MS/MS manuscript figures were prepared by importing data into Sigmaplot (Jandel) or Photoshop (Adobe).

Peptide and Ion Nomenclature. The nomenclature of the tryptic peptides follows that of the previous study of rat filaggrin peptides (Resing et al., 1993b); P refers to peptides from profilaggrin and F to peptides from filaggrin. Peptides are numbered sequentially from P1, which encompasses both the amino-terminal F1 and the carboxyl-terminal F49 peptides of filaggrin plus the residues removed during proteolytic processing. In numbering the specific residues of the filaggrin sequence, numbering initiates at the first residue of the P1 peptide. Ion nomenclature is that of Biemann (1990a,b) where the observed electrospray ionization masses are determined by the equation $(M+z)/z$, where M = mass of the peptide and z = charge of the ion. When ionized peptides are observed by MS, they are referred to as peptide or parent ions and their masses are represented by MH_n^{n+}, where n represents the number of protons that have been added to the peptide to achieve a charge of n.

Several of the profilaggrin peptides have multiple phosphorylation sites; the number of phosphates on each peptide is indicated by an asterisk followed by the number of phosphates, for example, $(P1^{*2})^{3+}$ indicates the triply charged ion of the diphosphorylated P1 peptide (for simplicity, the protons are not indicated when this nomenclature is used).

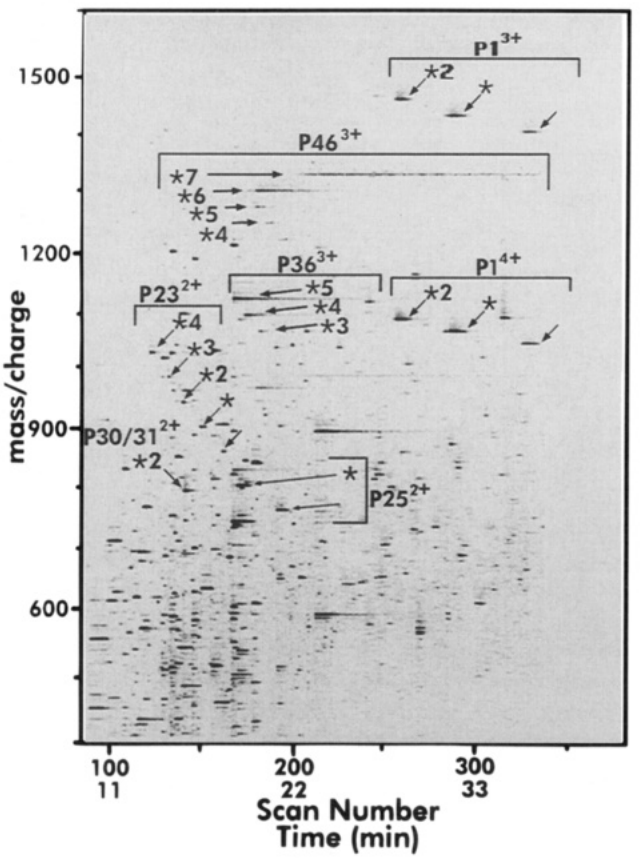

FIGURE 1: LC/MS analysis of a tryptic digest of profilaggrin. Data were collected on-line as the peptides eluted from a C18 HPLC column (*x* axis records scan number and elution time in minutes, *y* axis is *m/z*). Only the major ions of the phosphorylated peptides are labeled. For clarity, those peptides with more than two phosphates are labeled as a group, with an arrow pointing to each form of the peptide. The number following the asterisk indicates the number of phosphates (an unphosphorylated peptide in the series has no asterisk). In general, the other ions can be identified as filaggrin domain peptides [as in Resing et al. (1993b)].

Treatment with $Ba(OH)_2$ to β-eliminate H_3PO_4 from phosphorylated serine or threonine yields a residue mass that is 18 Da less than serine or threonine (β-elimination is equivalent to the loss of water); the presence of β-eliminated residues is indicated by the symbol Δ followed by the number of such residues, for example $(P1^{\Delta 2})^{3+}$ is the β-eliminated equivalent of $(P1^{*2})^{3+}$. The nomenclature of the fragment ions produced by MS/MS follows that of Biemann (1990a,b), except that b_n^{+2}, y_n^{+2}, and y_n^{+3} refer to $(b_n+H)^{2+}$, $(y_n+H)^{2+}$, and $(y_n+2H)^{3+}$, respectively.

RESULTS

Identification of Phosphopeptides. Initial studies indicated that profilaggrin phosphatase activity in epidermis was difficult to inhibit unless 9 M urea was used as an extractant. With these extraction conditions, less than 5% of the [^{32}P]-phosphate incorporated *in vivo* was released during homogenization and purification [measured as described in Haugen-Scofield et al. (1988)]. Oxidative cross-linking of cysteines allowed most of the keratin (as well as some other contaminants) to be selectively removed by precipitation, after which the filaggrin and profilaggrin were purified by DE52 chromatography.

Phosphorylated tryptic peptides were identified by analysis of a tryptic digest of profilaggrin on reverse-phase HPLC coupled on-line to an electrospray ionization mass spectrometer (LC/MS). The on-line data provided a display of the ion masses vs the HPLC elution time (Figure 1). Computer-assisted analysis of the LC/MS data identified virtually all of the peptides predicted by the cDNA of rat profilaggrin; the identity of the profilaggrin peptides were confirmed by MS/MS as previously described for filaggrin [data not shown; see Table 1 in Resing et al. (1993)].

In order to determine the potential phosphorylation sites, a list of possible phosphopeptides was generated by adding

Table 1: Putative Phosphopeptides Detected in LC/MS Analysis of Rat Profilaggrin

peptide no[a]	predicted mass[b] (Da)	observed mass (Da)	elution time (min)	ion current[c] ($\times 10^6$ cpm)
P1	4158.2	4157.8	35.5	0.96
P1*	4238.2	4237.6	32.0	2.59
P1*2	4318.2	4317.4	29.4	2.51
xP1[d]	3018.4	3017.9	29.9	0.14
xP1*	3098.4	3098.0	27.0	0.42
xP1*2	3178.4	3178.4	24.5	0.46
P9[e]	1087.5	1087.3	15.7	3.46
P9*	1167.5	1067.5	14.1	0.25
P16[e]	1015.4	1015.6	14.2	2.16
P16*	1095.4	1095.4	13.1	0.18
P23	1728.9	1729.0	18.6	0.77
P23*	1808.9	1809.0	17.0	0.87
P23*2	1888.9	1889.3	15.7	0.81
P23*3	1968.9	1969.0	14.6	0.21
P23*4	2048.9	2048.9	13.4	0.53
P25	1537.7	1538.1	22.6	2.15
P25*	1617.7	1617.9	19.8	5.85
P30/31	1440.7	not seen		
P30/31*	1520.7	1520.8	18.2	0.32
P30/31*2	1600.7	1600.8	15.9	2.73
P31	1227.7	not seen		
P31*	1307.7	1307.8	16.3	0.22
P31*2	1387.7	1387.5	14.4	1.14
P36	2937.3	not seen		
P36*	3017.3	3017.5	23.9	0.19
P36*2	3097.3	3097.3	22.5	0.26
P36*3	3177.3	3177.7	21.3	0.23
P36*4	3257.3	3257.6	~20[g]	~1.4[g]
P36*5	3337.3	3337.8	~19[g]	~3.6[g]
P37[e]	835.4	835.5	11.4	0.95
P37*	915.4	915.5	10.2	0.11
P46	3387.4	not seen		
P46*	3467.4	not seen		
P46*2	3547.4	not seen		
P46*3	3627.4	not seen		
P46*4	3707.4	3708.6[f]	21.8	0.14
P46*5	3787.4	3788.1[f]	21.0	0.32
P46*6	3867.4	3867.9[f]	21.3	1.87
P46*7	3947.4	3948.4[f]	~24[g]	~1.9[g]

[a] Tryptic peptides from profilaggrin; peptide nomenclature is described in Materials and Methods. Asterisks distinguish phosphopeptides from their unphosphorylated counterparts. The number following the asterisk denotes the number of phosphates attached. [b] MW was calculated using monoisotopic masses for peptides less than 1200 Da and average masses for larger peptides. [c] Determined by summing the area under the peaks in the selected ion chromatograms of each of the peptide ions; this represents the relative number of ions that reached the detector. [d] The xP1 forms are minor products of cleavage at Tyr_{12}; identity of these peptides was confirmed by MS/MS (not shown). [e] The putative phosphorylated form of this peptide was of low intensity relative to the unphosphorylated form. [f] The masses of the P46 peptides were approximately 1 Da too high, suggesting that some deamidation may have occurred in this experiment. In repeat experiments, the observed masses were closer to the predicted masses. [g] The elution time and the ion intensity are not well determined, because this peptide was spread over 15 min of elution (see Figure 1).

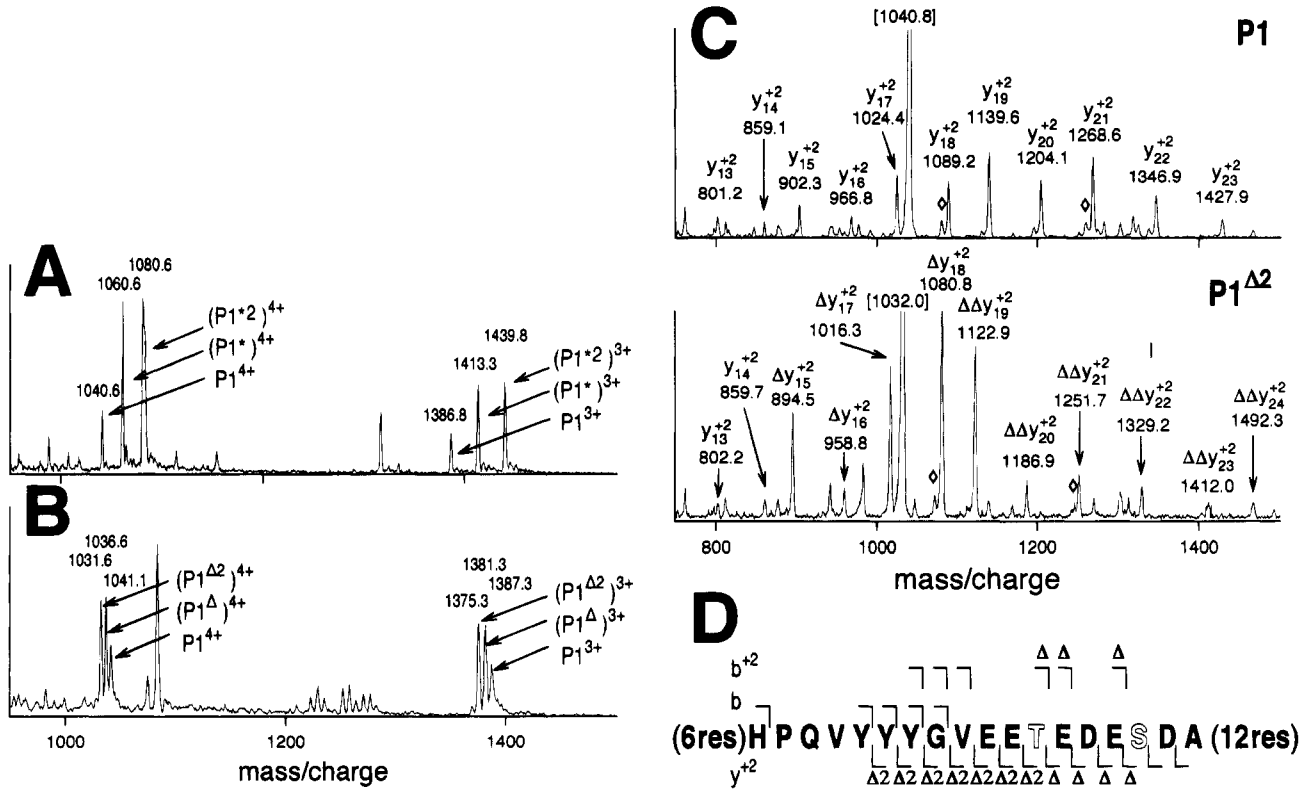

FIGURE 2: Analysis of P1 phosphopeptides. (A) MS spectrum of the three forms of the P1 peptide. Scans 268−330 from the LC/MS analysis of a tryptic digest of profilaggrin shown in Figure 1 were summed to obtain this spectrum. The peptide ions, labeled P1, P1*, and P1*^{2}, were each observed as two major ions, MH_4^{4+} and MH_3^{3+}. The observed molecular masses of the three peptides were calculated from the masses of the observed peptide ions as 4157.9, 4237.7, and 4317.4 Da; these observed masses differed by 80 Da, indicating the presence of one or two phosphates. (B) MS spectrum of the three P1 peptides after treatment of the tryptic digest with $Ba(OH)_2$ (summation of scans 288−335 of an LC/MS analyzed in parallel with that shown in Figure 1). The observed molecular masses of the three peptides, calculated from the observed peptide ions (labeled as the two charge forms of P1, P1$^\Delta$, and P1$^{\Delta 2}$) were 4159.7, 4141.7, and 4122.7 Da. Due to increased hydrophobicity of these peptides upon β-elimination, they eluted later than those seen in (A); hence the other ions observed in these spectra are different. (C) Comparison of portions of the MS/MS spectra of the nonphosphorylated $(P1)^{4+}$ and β-eliminated $(P1^{\Delta 2})^{4+}$ ions; for clarity, only those regions containing the doubly charged y ions that allowed localization of phosphate to Ser_{18} and Thr_{22} are shown. The presence of one or two dehydrated residues is revealed by a decrease of 9 or 18 Da/e in the m/z of the doubly-charged fragment ions and are denoted by the symbols Δ and $\Delta\Delta$, respectively. There is a small loss of water from some of these fragment ions (e.g., $\diamond$), even when not phosphorylated, but this appears to be nonspecific dehydration because it is similar in both $(P1)^{4+}$ and $(P1^{\Delta 2})^{4+}$. The parent ions are bracketed. (D) Summary of all ions observed in the MS/MS spectrum of the $(P1^{\Delta 2})^{4+}$ ion, displayed on part of the sequence denoted by the single-letter code (the first six and last 12 residues are not shown). The deduced phosphorylated residues are indicated by S or T. The b^+ and b^{2+} ions are displayed above the sequence (⌉) and the y^{2+} ions are below (⌊). A Δ symbol by the ⌊ or ⌉ indicates the observed ion corresponded to a fragment containing a dehydrated residue in place of a serine or threonine and that there was no detectable ion at the mass corresponding to the unmodified serine or threonine. A $\diamond$ indicates loss of H_2O from an ion, where there is also a high intensity unmodified ion also present. The number following the Δ symbol indicates that more than one residue was dehydrated.

a mass increment of 80 Da for each hypothetical phosphate (Table 1). The search was facilitated by the knowledge that a phosphopeptide normally eluted 1−5 min earlier than the corresponding unphosphorylated form. With the exception of P36 and P46, the ion intensity of the unphosphorylated peptides from filaggrin digests (which were completely dephosphorylated) was similar to the sum of the ion intensities of the phosphorylated forms of that peptide from profilaggrin (not shown), indicating that phosphorylation does not significantly reduce their recovery. In the case of P36 and P46, the highly phosphorylated peptides eluted anomalously during reverse-phase HPLC, exhibiting broad and trailing peaks (Figure 1); this behavior interfered with estimation of their yields. In preliminary experiments, several of the profilaggrin peptides, both phosphorylated and unphosphorylated, formed noncovalent adducts with H_3PO_4 when the tryptic digest was acidified with phosphoric acid before chromatography; these were detected as ions that were 98 Da larger than predicted and produced nearly complete loss of 98 Da during collisional activation, yielding a major ion with the mass of the predicted peptide. In contrast, a normal phosphorylated peptide produced only partial loss of 98 Da from a peptide that was 80 Da larger than the predicted mass.

Nine peptides were identified as possible phosphopeptides (Table 1). Three of these (P9, P16, and P37) had a signal intensity for the putative phosphorylated form of less than 15% of the unphosphosphorylated form, which was too low for further analysis. Total recovery of these three peptides was similar from both profilaggrin and filaggrin (not shown), indicating that the low stoichiometries were not due to preferential loss of the phosphorylated forms. The six remaining peptides (P1, P23, P25, P31, P36, and P46) contained high stoichiometry phosphorylation sites and were characterized further after fractionating the peptides, either from the split stream during LC/MS or from HPLC without the mass spectrometer in line. MS/MS confirmed the identity of the peptides, as well as the presence of phosphate, but

often did not provide sufficient information for assignment of the phosphate to a specific residue.

A problem encountered in sequencing these phosphorylated peptides by MS/MS was the preferential elimination of H_3PO_4 (neutral loss of 98 Da), which complicates interpretation of the spectra and is usually accompanied by reduced fragmentation at the peptide bonds during collision-induced fragmentation. To circumvent this problem, peptides were treated with $Ba(OH)_2$, yielding dehydroalanine or dehydrothreonine at the phosphorylated residues by base-catalyzed β-elimination of H_3PO_4 (Byford, 1991). There was little effect on peptide recovery when treatment was limited to 2 h or less and the samples were analyzed immediately. For example, the relative intensities of the MH_4^{4+} and MH_3^{3+} ions of the P1 peptide were nearly identical before and after treatment with $Ba(OH)_2$ (Figure 2A,B). A side reaction that was observed was the occasional conversion of glutamine to glutamic acid (e.g., see discussion of P1). Derivatization of the resulting dehydrated residues was unnecessary, eliminating variability in the peptides that could have been generated by side reactions and/or nonquantitative derivatization. The beta-eliminated peptides P36 and P46 exhibited improved chromatographic behavior (see above), eliminating one source of loss of the highly phosphorylated peptides. Interestingly, replacement of a phosphorylated residue with a dehydroamino acid sometimes appeared to produce a more efficient fragmentation at those sites (i.e., y_{18} and y_{19} in Figure 2C, b_8 in Figure 3, and b_5 in Figure 4). Furthermore, elimination of partially dephosphorylated ions (i.e., y_{10} in Figure 3A, panel 2, and y_7 and y_9 in Figure 4A, panel 2) simplified analysis of the MS/MS spectrum.

Analysis of Phosphopeptide P1: GQQGSGHPQVYYY-GVEET$_{18}$EDES$_{22}$DAQQGHHQQQQQQR. P1, which contains the protease processing sites, has been characterized as a phosphopeptide previously (Resing, et al., 1993b). In that study, approximately equal amounts of unphosphorylated P1 and monophosphorylated P1* were observed as MH_4^{4+} ions of *m/z* 1040.6 and 1060.6 Da/e and MH_3^{3+} ions of *m/z* 1386.8 and 1413.3 Da/e, corresponding to peptides with masses of 4157.9 and 4237.7 Da, respectively (also see Figure 2A). Neutral loss of phosphate was not observed during collisional activation of P1*, allowing the site of phosphorylation to be identified as Ser_{22}. In the present study, when profilaggrin was purified in the presence of 9 M urea, very little unphosphorylated P1 was observed, and 40% of the ion current in the P1 peptides was found as the diphosphorylated $P1^{*2}$ (observed in Figure 2A as ions of 1080.6 and 1439.8 Da/e, corresponding to a mass of 4317.4 Da). The $P1^{*2}$ ions readily lost phosphate during collision induced dissociation, with little subsequent fragmentation.

After treatment with $Ba(OH)_2$, the masses of the resulting peptides were 4159.6, 4141.7, and 4122.7 Da (Figure 2B), corresponding to P1 + 1 Da, $P1^{\Delta}$ +1 Da, and $P1^{\Delta 2}$ + 1 Da. As expected, these derivatives differ by 18 Da due to loss of phosphoric acid from the phosphorylated residues. The 1 Da increase in mass of each derivative is most likely due to the conversion of one of the carboxyl-terminal glutamines to glutamic acid. This is supported by the MS/MS spectrum of the MH_4^{4+} ion of the altered P1 (Figure 2C); however, no fragmentation within the sequence QQGHHQQQQQQR was observed so that it was not possible to determine if a single glutamine was converted or if the treated peptide is heterogeneously deamidated.

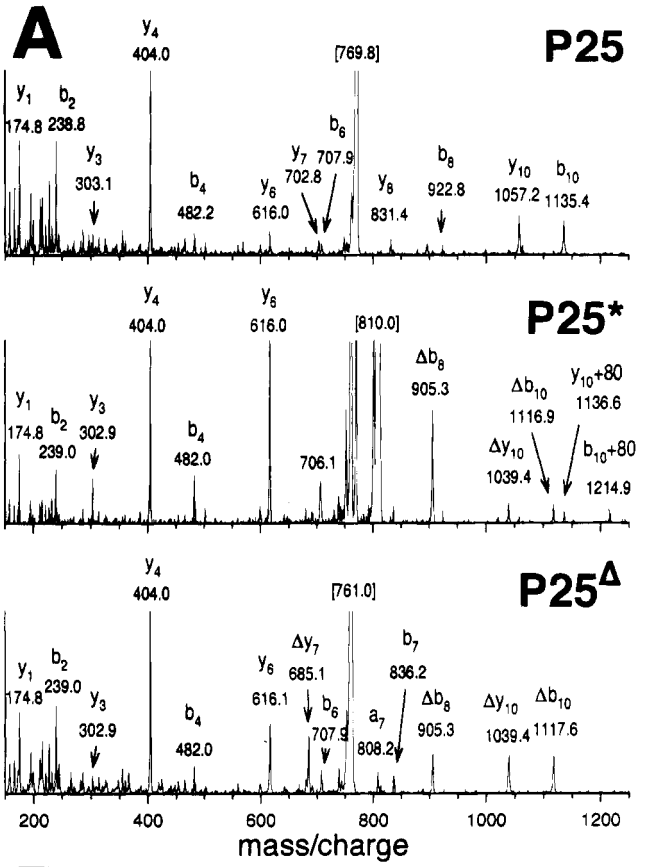

FIGURE 3: Analysis of the P25 phosphopeptide. (A) Comparison of MS/MS spectra of $P25^{2+}$, $(P25^{*})^{2+}$, and $(P25^{\Delta})^{2+}$. An asterisk indicates an ion that has phosphate, while Δ indicates an ion that has lost phosphoric acid, either by β-elimination with $Ba(OH)_2$ or in the collision cell (e.g., b_8 in P25*). The major unlabeled peaks in P25* are 800.8 Da/e (loss of 18/2 from the parent), 760.7 Da/e (loss of 98/2 Da/e from the parent), 751.7 Da/e (further loss of 18/2 Da from 760.7), and two unidientified ions at 706.1 Da/e (**), and 769.7 Da/e (**). (B) Summary of the observed b ions (⌈) and y ions (⌊) in the $P25^{\Delta}$ sequence. See Figure 2 legend for the descriptions of other symbols in this panel.

MS/MS of the $(P1^{\Delta 2})^{4+}$ ion located the two phosphates on Thr_{18} and Ser_{22} (Figure 2D). There was an increased relative abundance of ions resulting from fragmentation at peptide bonds near the dehydrated residues.

Analysis of Phosphopeptide P25: HTSRPEQS$_{199}$PDTAGR. Ions corresponding to the MH_2^{2+} (770.1 and 810.0 Da/e) and MH_3^{3+} (513.8 and 540.8 Da/e) ions of P25 and P25* were observed and their identities confirmed by MS/MS. The phosphate was partially stable during MS/MS, allowing localization of the phosphorylation site at Ser_{199} (Figure 3A), although analysis of the spectra was complicated by the neutral loss of H_3PO_4 from the sequence-specific fragment ions, in that some fragment ions containing Ser_{199} were present in two forms, one with phosphate and one with dehydroalanine. Furthermore, the presence of phosphate apparently altered the ion chemistry (note the large increase in y_3, y_6, and b_8), as previously observed in phosphorylation sites of MAPK kinase (Resing et al., 1995). Similar results have been obtained with peptides containing cysteic acid

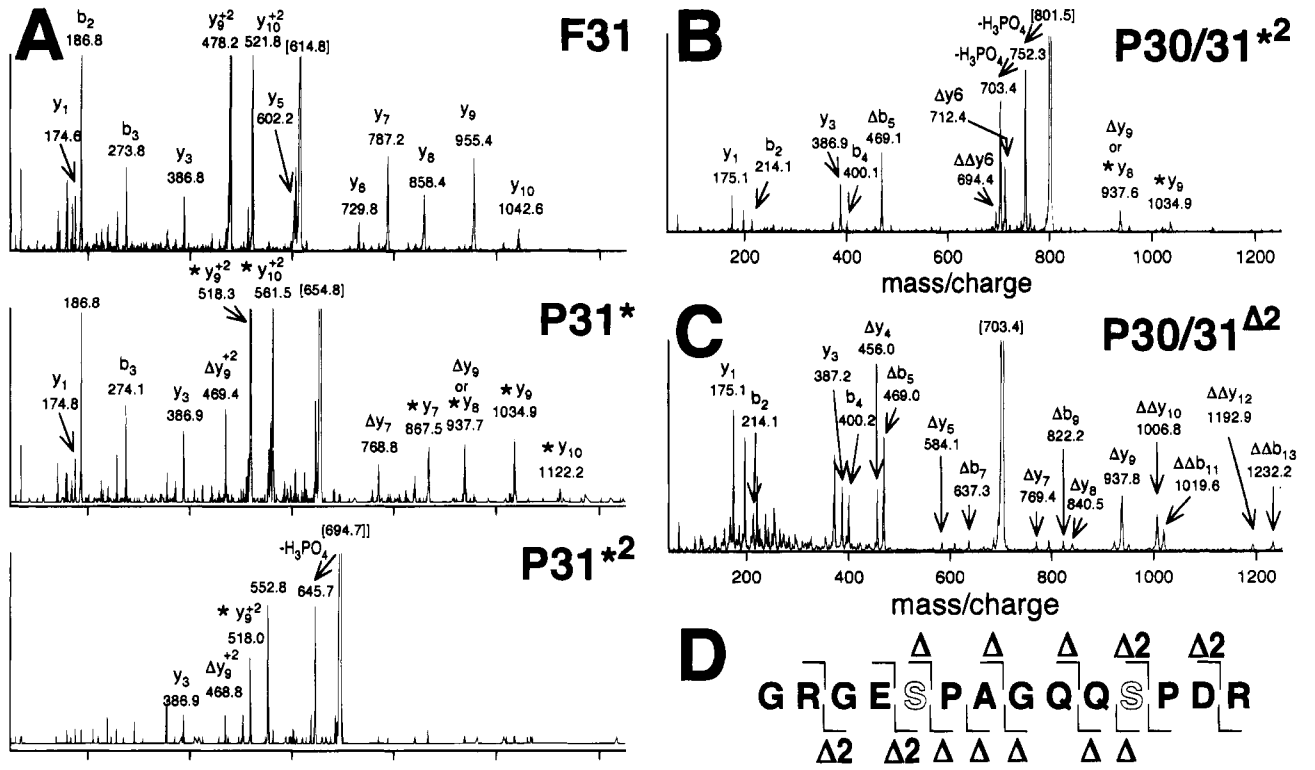

FIGURE 4: Analysis of P31 phosphopeptides. (A) MS/MS spectra of the MH_2^{2+} ions of F31, P31*, and P31*2. Comparison of MS/MS spectra of F31 and P31* localizes the single phosphate in P31* to S_{242} (the ninth residue in this peptide). The $(P31^{*2})^{2+}$ ion produced primarily neutral loss of phosphate; however, sufficient fragmentation was observed to identify it as a form of P31. The 552.8 Da/e ion in P31*2 spectrum is the doubly charged y_{10} with one phosphate and one dehydroserine from neutral loss of H_3PO_4 from the second phosphoryl group (98/2 Da/e). (B) MS/MS of the incomplete proteolytic product $(P30/31^{*2})^{2+}$. Two ions resulting from neutral loss of phosphoric acid are indicated as well as a few fragment ions. (C) MS/MS of $P30/31^{\Delta 2}$ demonstrates the improved cleavage of this peptide ion (compared to B). (D) Summary of the observed b ions (⌈) and y ions (⌊) in the $P30/31^{\Delta 2}$ sequence. See the legend to Figure 2 for a description of the symbols in this panel.

(Burlet et al., 1992). In contrast, after chemical β-elimination of P25*, the MS/MS spectrum of $(P25^{\Delta})^{2+}$ ion (Figure 3A) was similar in appearance to P25, except that there was an increase in peptide bond cleavage next to the dehydroalanine.

Analysis of Phosphopeptide P31: $GES_{236}PAGQQS_{242}PDR$. MH^+ and MH_2^{2+} ions of P31* (1308.8 and 654.9 Da/e) and P31*2 (1388.5 and 694.8 Da/e) were identified in the LC/MS data. The identity of P31* was confirmed by MS/MS (Figure 4A, panel 2), and the phosphate was localized to S_{242}. There was no neutral loss of H_3PO_4 from the P31* parent ion, although at least one fragment ion showed neutral loss of H_3PO_4 in addition to cleavage at the backbone. In contrast, few sequence specific ions were present in the MS/MS spectrum of P31*2, although they were sufficient to identify the 694.7 Da/e peptide ion as $(P31^{*2})^{2+}$ (Figure 4A, panel 3). The location of the two phosphates is clear, as there are only two serines and no threonine or tyrosine in this peptide. The total recovery of P31, P31*, and P31*2 was low compared to the recovery of F31 from filaggrin (approximately 30%). A search revealed high intensity peptide ions with observed *m/z* similar to those predicted for the MH_2^{2+} and MH_3^{3+} ions of the incomplete proteolytic product P30/31*2 (801.4 and 534.8 Da/e) as well as low intensity ions with *m/z* corresponding to the predicted MH_2^{2+} and MH_3^{3+} ions of P30/31* (761.4 and 507.9 Da/e). These incomplete proteolytic products were rarely observed in filaggrin digests, indicating that the presence of phosphate on P31 was interfering with typtic cleavage. The identification of P30/31*2 was confirmed by MS/MS (Figure 4B). Again, there are only two phosphorylatable residues in this peptide, at Ser_{236} and Ser_{242}. MS/MS spectra of $(P30/31^{\Delta 2})^{2+}$ showed improved fragmentation of the peptide bonds compared to $(P30/31^{*2})^{2+}$ (Figure 4C compared with Figure 4B).

Analysis of P23: $GGQAGS_{171}HS_{173}ES_{175}EAS_{178}GGQAGR$. This peptide produced a series of MH_2^{2+} ions, with observed *m/z* of 866.0, 906.0, 946.0, 985.7, and 1026.1 Da/e, and a series of MH_3^{3+} ions, 577.8, 604.3, 630.8, 657.4, and 683.9 Da/e (Figure 5B), corresponding to the presence of up to four phosphates. The reverse-phase elution behavior of these peptides was typical, in that the addition of each phosphate led to a 1.5 min decrease in elution time from reverse-phase HPLC (Figure 5A). The presence of phosphate on these ions was confirmed by neutral loss experiments and by MS/MS of the phosphorylated peptide ions (for example, Figure 5C), where losses of 49 Da/e from the doubly charged precursor ions were detected. No sequence specific ions were observed in the MS/MS spectra of the phosphorylated peptides. The relationship of this phosphorylated series to P23 was confirmed in the MS/MS spectrum of the 866.0 Da/e MH_2^{2+} ion of the putative unphosphorylated peptide (Figure 5D). The peptide ions were of low intensity, even when analyzed in digests of filaggrin (Resing et al., 1993), and there was a low yield of fragment ions; nevertheless, sufficient information was obtained to unambiguously identify this ion as $P23^{2+}$. As there are only four phosphorylatable residues (S_{171}, S_{173}, S_{175}, S_{178}) in P23, further characterization of this peptide was not required to identify the four phosphorylation sites.

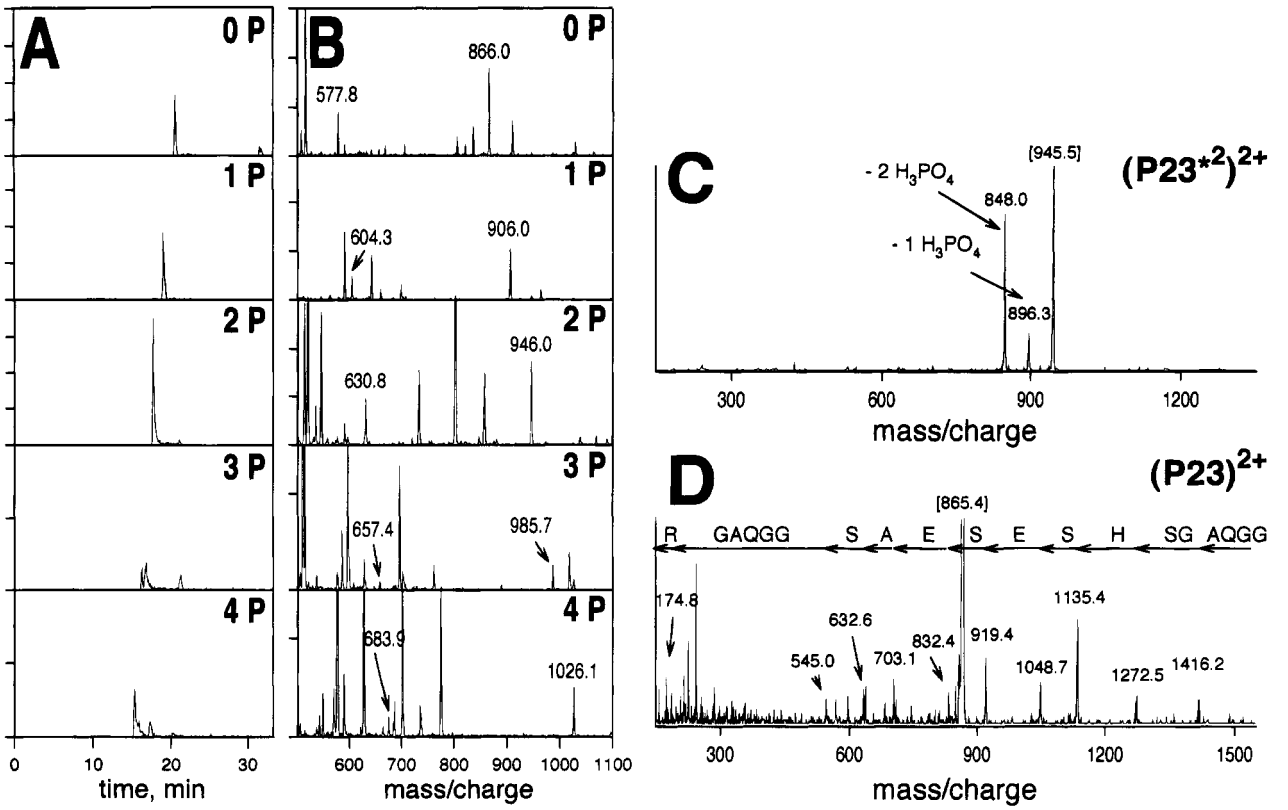

FIGURE 5: Analysis of P23 phosphopeptides. Five forms of this peptide were recovered, each in low yield. (A) Ion chromatograms corresponding to the predicted *m/z* for the MH_2^{2+} ions of P23 (0P), P23* (1P), P23*^{2} (2P), P23*^{3} (3P), and P23*^{4} (4P). Each panel represents a chromatogram for an individual peptide ion of predicted *m/z* 865.9, 905.9, 945.9, 985.9, and 1025.9 Da/e (using a 1.5 Da wide window). The two peaks in the 3P panel appear to represent two forms of the $(P23^{*3})^{2+}$ ion. The peaks at 21 min in 3P and 17.3 min in 4P are shoulders of intense ions close to the selected *m/z*. The *y* axes of the 3P and 4P panels are expanded 2-fold, relative to the 0, 1, and 2P panels. (B) MS spectra of each peak detected in the scans of panel A. Four scans across each peak were summed; for clarity, only a limited mass range is shown. The observed MH_2^{2+} and MH_3^{3+} ions for the different P23 peptides are labeled in each panel. The unlabeled ions represent other profilaggrin peptides. The *y* axis of the 3P and 4P panels is expanded 2-fold, relative to the 0, 1, and 2P panels. (C) MS/MS spectrum of the $(P23^{*3})^{2+}$ ion (parent ion bracketed). The major fragment ions represent the loss of one or two H_3PO_4 groups from the two phosphorylated residues. (D) MS/MS spectrum of the $P23^{2+}$ ion (label in brackets). Singly charged y ions are the major (labeled) product ions The 1272.5 Da/e ion is b_6 (GGQAGS). Also observed were 242.9 (b_3), 595.7 (b_7), 682.4 (b_8), 636.5 (y_{13}^{2+}), 708.7 (y_{15}^{2+}), and 744.2 (y_{16}^{2+}) Da/e.

Analysis of P36: $EAS_{260}AS_{262}QS_{264}S_{265}DS_{267}EGHSGAHA-GIGQGQTSTTHR$. Analysis of fractions from Synchropak C18 HPLC containing P36 revealed phosphorylated forms with one to five phosphates, as well as a series of sodium adducts and a series of dehydrated peptides (Figure 6A); their identity was confirmed by MS/MS (not shown). The MS spectrum shown was obtained at an orifice voltage of 80 V; when the orifice voltage was reduced to 60 V, the dehydrated series nearly disappeared, suggesting that low energy collision prior to mass analysis induced the neutral loss of water. MS/MS of the P36 ions produced neutral loss of the predicted number of H_3PO_4 molecules and up to two additional water molecules (Figure 6B). Sequence-specific fragmentation of each of these phosphorylated peptides was also observed; however, the spectra were difficult to interpret and did not unequivocally localize the phosphates. After treatment with $Ba(OH)_2$, MS/MS of the resulting $(P36^{\Delta 5})^{4+}$ allowed unambiguous assignment of the phosphates to the first five serines in the peptide (Figure 6C). There was a large increase in fragmentation near the dehydroalanines in comparison with the fragmentation of the unphosphorylated F36 (not shown). In addition, extensive secondary cleavage occurred in the amino-terminal region of this peptide (where the dehydroalanines were located) leading to an overall reduction in b ions in P36 compared to F36 (Figure 6D) and an increase in internal cleavages products from this region. Apparently the b ions containing several dehydroalanines are unstable.

Analysis of P46: $GS_{368}S_{369}ES_{371}QAS_{374}DS_{376}EGHS_{380}-DYS_{383}EAHTQGAHGGIQTSQR$. Multiply phosphorylated forms of P46 were observed with four to seven phosphates (Figure 7A), as was the incomplete proteolytic product P45/46 (RPRGSSES...) with five to seven phosphates. P45/46 was not observed in similar digests of filaggrin, suggesting that the phosphorylation of P46 interferes with proteolysis. MS/MS of these ions yielded neutral loss of the expected number of H_3PO_4 molecules, along with some sequence specific fragment ions (not shown). When profilaggrin was rapidly purified from urea, predominantly P46*^{7} was recovered, although this peptide slowly lost phosphate over time (Figure 7A).

Peptides containing dehydroalanine or dehydrothreonine are chemically unstable and will either hydrate or produce cross-links over a period of hours, even when frozen. This was more obvious with P46 than with the other phosphorylated peptides in this study. For this reason, analysis of $P46^{\Delta 7}$ was carried out by LC/MS/MS, where the MS/MS spectrum was acquired as the peptide eluted from a reverse-phase capillary column. A fraction from Synchropak C18 chromatography containing the P46 peptides was treated with

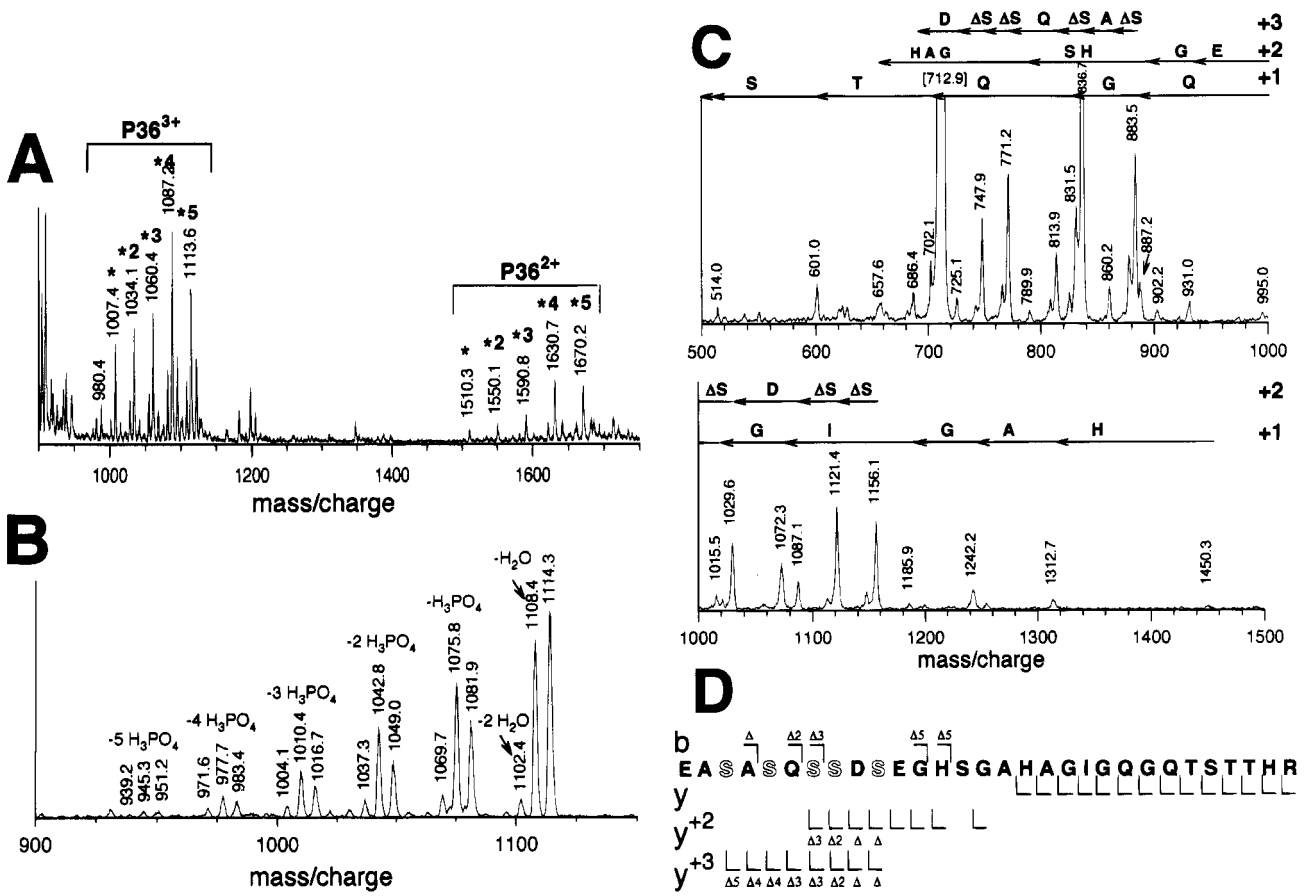

FIGURE 6: Analysis of P36 phosphopeptides. (A) MS of P36 phosphopeptides showing the peptide MH_3^{3+} ions and MH_2^{2+} ions (summed scans 168−219) representing a complete series from one phosphate (*) to five phosphates (*5), where the unphosphorylated peptide mass would be 2937.3 Da (calculated from the observed peptide ions). Two other series of phosphorylated ions (unlabeled) were observed, e.g., peptide ions of 1121.6/1094.6/1068.2/1041.8/1014.8 Da/e represent the MH_3^{3+} ions of sodium adducts of P36 and its phosphoforms (unphosphorylated mass of 2959.8 Da), and 1107.8/1081.3/1054.9/1027.9/1001.2 Da/e represent the MH_3^{3+} ions where dehydration of P36 and its phosphoforms has occurred (unphosphorylated mass of 2919.3 Da). The identity of all of these ions was confirmed by MS/MS. (B) MS/MS of the $(P36^{*5})^{3+}$ ion of *m/z* 1113.6 from panel A demonstrating the neutral loss of five H_3PO_4 and two H_2O molecules. For clarity, only part of the spectrum is shown, and the dehydration products of the dephosphorylated ions are not labeled. Sequence-specific fragmentation ions were also observed, which included all of the ions seen in the completely β-eliminated peptide (cf. C, below) as well as those which contain 1−5 phosphates; the presence of so many possible variations made unambiguous assignment of many ions impossible. (C) MS/MS spectrum of the β-eliminated $(P36^{\Delta 5})^{3+}$ ion. For clarity only part of the sequence is shown and only the y ions are labeled. The three lines above the spectra correspond to the singly, doubly, and triply charged ions from bottom to top. ΔS indicates dehydroalanine. The unlabeled peaks generally represent dehydration products of the nearby peak; dehydration ions of similar intensity were observed in the MS/MS spectrum of F36, the unphosphorylated form of this peptide derived from filaggrin (not shown). The unlabeled ion at 1199 Da/e and those between 500 and 700 Da/e represent internal fragment ions. (D) Summary of the observed y ions (⌈) in the P36 sequence, including those not shown in panel C. See the legend to Figure 2 for a description of the symbols in this panel.

$Ba(OH)_2$ and immediately applied to a 500 μm capillary column with all of the eluate directed to the mass spectrometer. MS/MS was conducted on-line as the peptides eluted (LC/MS/MS). After treatment of $P46^{*7}$ with $Ba(OH)_2$, the mass of the resulting ion, $(P46^{\Delta 7})$, was 1.5 Da larger than predicted. The fragment ions corresponding to y_4 and y_7 through y_{11} were all shifted by +1 Da, indicating that Gln_{395} was converted to Glu; other glutamines in the peptide were not affected. Comparison of this spectrum with that of the unphosphorylated F46 suggested that there was suppression of b type ions in the MS/MS spectrum of $(P46^{\Delta 7})^{4+}$, as well as an increase in internal fragment ions derived from the amino terminus of the peptide. Despite these changes in ion chemistry, the large yield of y ions produced by cleavage of the peptide bonds involving the dehydroalanines allowed assignment of the seven phosphates to the first seven serines in the peptide. The MS/MS spectra of $(P46^{\Delta 4})^{4+}$ and $(P46^{\Delta 6})^{4+}$ (not shown) indicated that $P46^{*4}$ and $P46^{*6}$ were also phosphorylated within these seven serines; in these two cases, certain residues were completely phosphorylated, while others were partially phosphorylated.

DISCUSSION

The ability of electrospray MS to analyze mixtures of large peptides provides a powerful tool for analyzing the large, multiply phosphorylated profilaggrin peptides, without the necessity of radiolabeling the protein. In this study, six high stoichiometry phosphopeptides were detected in LC/MS of total tryptic digests of profilaggrin by their altered masses of 80 Da (or multiples thereof), and the phosphorylated residues were located by MS/MS experiments, without purifying the peptides. Three phosphorylated residues (S_{22}, S_{199}, and S_{242}) were located by MS/MS of the phosphopeptides, and five more were identified because all serines in the peptides were phosphorylated (S_{171}, S_{173}, S_{175}, S_{178}, and S_{236}). The remaining phosphorylated residues were in clusters which could not be unambiguously sequenced. By treatment of these phosphorylated peptides with $Ba(OH)_2$,

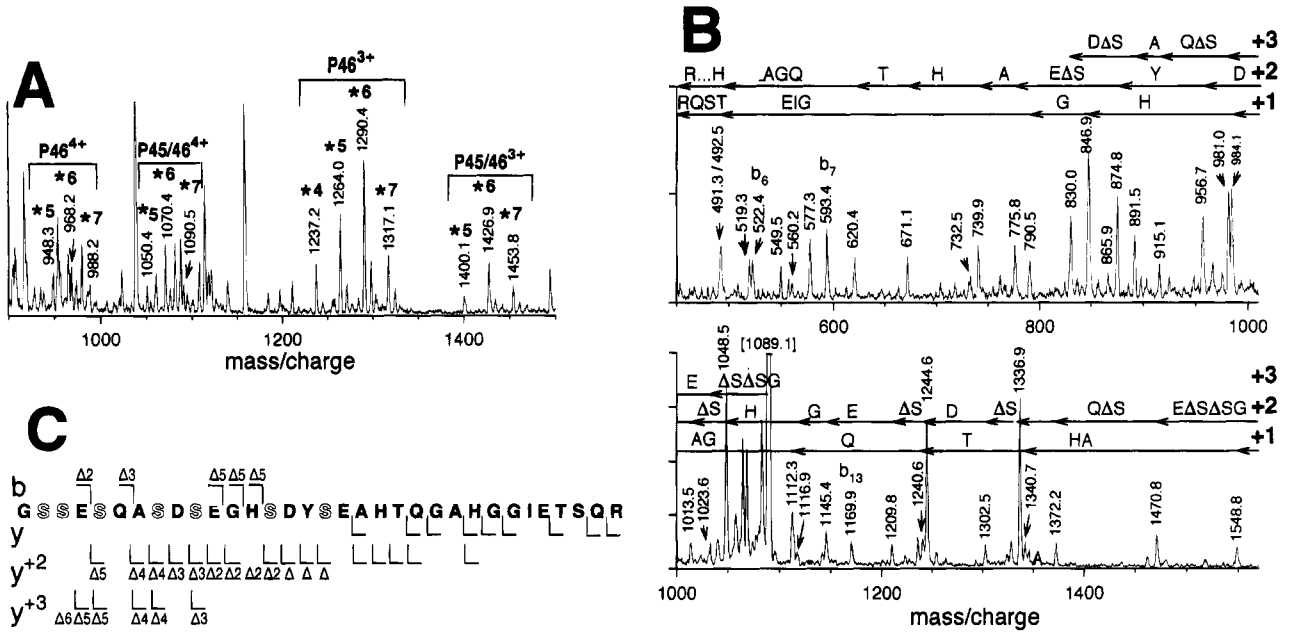

FIGURE 7: Analysis of P46 phosphopeptides. (A) Mass spectrum of a Synchropak HPLC fraction containing P46 phosphopeptides. A series of peptide ions differing by 26.7 and 20 Da/e (the MH_3^{3+} and MH_4^{4+} peptide ions, respectively) are observed which correspond to those predicted for P46*4 (predicted *m/z* 1237.5/927.9), P46*5 (predicted *m/z* 1264.3/947.9), P46*6 (predicted *m/z* 1291.1/967.9), P46*7 (predicted *m/z* 1317.3/987.9) P45/46*5 (predicted *m/z* 1400.0/1050.2), P45/46*6 (predicted *m/z* 1426.8/1070.2), P45/45*7 (predicted *m/z* 1453.5/1090.3). These peptide ion masses correspond to peptide masses of 3708.6, 3788.1, 3867.9, and 3948.4 Da, where the unphosphorylated mass of P46 is predicted to be 3387.4 Da. There are also phosphorylated proteolytic products of P1 in this fraction. The identity of all ions were confirmed by MS/MS. Note the altered ratio of the various forms in this sample after storage compared to the LC/MS results of Table 1. (B) MS/MS spectrum of $P46^{\Delta 7}$ (MH_3^{+3} precursor ion of 1089.1 Da/e). Predicted masses were calculated assuming conversion of Q_{395} to glutamic (see Results); above 1200 Da, the predicted masses were calculated using average residue masses. Nearly all observed ions were y ions and the sequences that they represent are indicated by the arrows above the data panels; some of the prominant b ions are labeled individually. The three lines of arrows are as described in Figure 6. The symbol ΔS indicates dehydroalanine. Ions representing loss of water or ammonia from the more intense y ions were no greater than in the MS/MS spectrum of F46. For clarity, internal fragment ions at 549.5, 557.0, and 577.3 Da/e, the parent ion − H_2O at 1083.1 Da/e, and unidentified ions at 1057.6, 1064.2, and 1068.2 Da/e are not labeled (these are apparently triply charged ions in that each has an associated ion that is 6 Da smaller, presumably due to the loss of water); furthermore, the data from 100 to 450 Da/e are not shown, including the major ions 174.9 (y_1), 325.0 (b_4), 109.9 ([H]), 198.2 (ΔSQ), 169.9 9 ([ΔSQ]), 268.9 (ΔSQA), 251.2 (not identified), 365.4 (HGGI), 337/3 (ΔSΔSEΔS), 348.3 (ΔSΔSY), 320.2 ([ΔSΔSY]), 393.5 (ΔSEGH), 448.6 (not identified) Da/e, where brackets indicate immonium or a-type ions. (C) Summary of the observed sequence specific b and y ions of $P46^{7\Delta}$, including those not shown in panel B. The b^+ ions are displayed above the sequence (┐) and the y^+, y^{2+}, and y^{3+} ions are below (└). Only ions whose intensity was greater than 5 times the background are shown. Q_{395} has been replaced with an E in this sequence. See the legend to Figure 2 for a description of the other symbols in this panel.

dehydrated forms were produced which could be sequenced by MS/MS, without the derivatization required for Edman degradation. In this way, 13 more phosphorylation sites were located—for a total of one threonine and 20 serine phosphorylation sites in these six peptides. Because profilaggrin contains multiple repeats of filaggrin [approximately 20 in Sprague-Dawley rat (Haydock & Dale, 1990)], the fractional occupancy at any given site can be calculated (Table 2), predicting an average phosphorylation level of 355 phosphates in each profilaggrin molecule, neglecting the contribution of the calcium binding amino-terminal domain (Presland et al., 1992; Markova et al., 1993) and assuming 20 filaggrin repeats per molecule. This is an average of 17.8 phosphorylated residues in each repeat, which is consistent with the estimate of 20 in each repeat measured by Lonsdale-Eccles et al. (1982) using a colorimetric method.

Sufficient data were obtained for P25 and P31 (Figures 3 and 4) in both phosphorylated and $Ba(OH)_2$ treated forms to test for side reactions of $Ba(OH)_2$ treatment. Recovery of the various forms of the peptides before and after $Ba(OH)_2$ treatment was similar (for example, see Figure 2), indicating that the β-eliminated peptides were reasonably stable if analyzed within 2 h, although reduced yields were observed upon storage of lyophilized HPLC purified peptides at −20 °C for several days. The only change detected in the profilaggrin peptides was the occasional conversion of Gln to Glu, although it should be noted that not all amino acids were present in these peptides. Because derivatization was not necessary, a source of variability due to alternative addition products was eliminated (Byford, 1991). Furthermore, the anomalous behavior of the highly phosphorylated peptides during HPLC was avoided, making recovery of these peptides more reliable. Peptides containing dehydroalanine exhibited somewhat different gas phase chemistry, compared to the *un*phosphorylated forms, although the overall cleavage pattern of each was similar. There was often increased cleavage at both peptide bonds of the dehydroalanine residues, a phenomenon which was useful in analyzing low yield peptides. In some cases when low yield parent ions were analyzed, the only fragment ions observed were generated from cleavage at the peptide bonds of the dehydroalanine residues (not shown).

When the sequences of the rat and mouse profilaggrin are compared (Haydock et al., 1992), the rat phosphopeptides align with the previously identified mouse phosphopeptides (Figure 8C) with the exception of P23 and P25, which are part of a large insert that is not present in mouse profilaggrin

Table 2: Calculation of Phosphate Occupancy in Each Peptide[a]

peptide number	number of repeats in that form[b]	number of phosphates[c]
P1	3.2	0
P1*	8.5	9
P1*2	8.3	16
P23	4.8	0
P23*	5.5	6
P23*2	5.1	10
P23*3	1.3	4
P23*4	3.3	13
P25	5.4	0
P25*	14.6	15
P31*	2.5	3
P31*2	17.5	35
P36*	0.7	1
P36*2	0.9	2
P36*3	0.8	2
P36*4	5.0	20
P36*5	12.6	63
P46*4	0.7	3
P46*5	1.5	8
P46*6	8.7	52
P46*7	9.1	64

[a] Calculated from the fifth column of Table 1. [b] Assuming 20 filaggrin segments per profilaggrin. [c] Number of phosphates in each molecule of profilaggrin that are derived from each phosphoeptide, calculated by multiplying column 2 by the phosphate content of each peptide, rounded to the nearest whole number.

(the variable domain in Figure 8B). The placement of these phosphopeptides in the PIR unit is diagramed in Figure 8A. The phosphorylation sites are found clustered within previously identified domains (Haydock & Dale, 1990), each cluster including nearby acidic residues. The phosphorylated sequences include consensus phosphorylation sites of known kinases, although so far only casein kinase II has been shown to phosphorylate these peptides.[2]

The similarity of the profilaggrin phosphorylation clusters to each other may be more consistent with convergent evolutionary pressure from the processing enzymes than with gene duplication events. P25 and P31 resemble each other and the mouse sequence in that each contains a QSPD sequence, with high stoichiometry of phosphorylation at this serine (Figure 8C), while the phosphorylated regions of P36 and P46 are similar to each other in their first 15 residues, which encompass most of the phosphate. In neither case does the similarity extend much beyond the phosphorylated region. The multiply phosphorylated P23 does not show obvious similarity to any mouse or rat sequence, except that it also is an acidic, multiply phosphorylated peptide. The similarity between the diphosphorylated P31 sequence and the mouse sequence became more obvious when the mouse peptide SGSGGR was considered. This latter peptide was originally identified by amino acid analysis of a ^{32}P-labeled tryptic peptide of profilaggrin which had a specific activity indicating that there was one phosphate on the peptide (Resing et al., 1985). In an LC/MS analysis of mouse profilaggrin, the incomplete proteolytic peptide SGSGGRGQSPDGSGR was observed as a high stoichiometry diphosphorylated peptide (data not shown); the mouse peptide GQSPDGSGR was known to be singly phosphorylated on the first serine as a result of the previous study, suggesting that the other phosphate was in the SGSGGR peptide. Other than the spacing between the serines, there was little sequence similarity between the SGSGGR and the GESPAG of P31. Further studies with profilaggrin from

[2] K. A. Resing, E. Kam, and B. A. Dale, unpublished data.

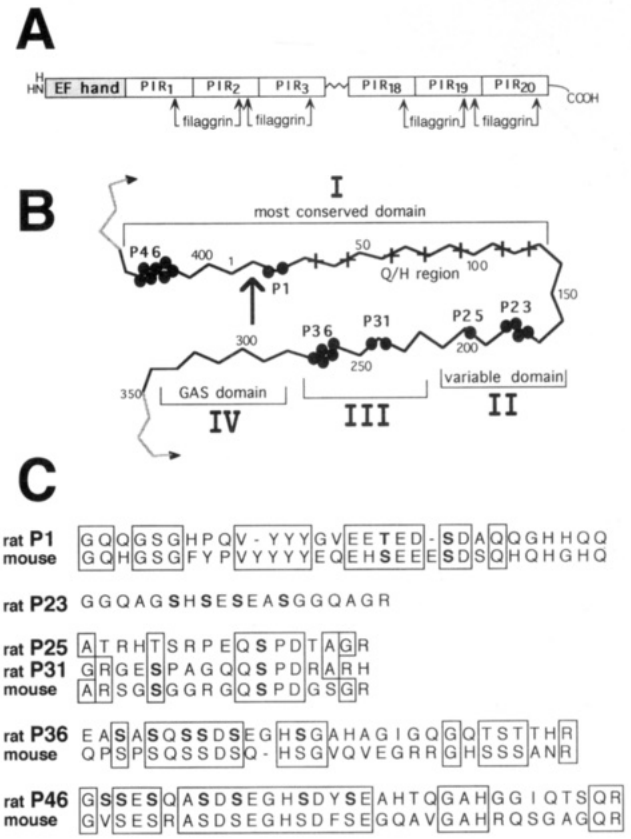

FIGURE 8: (A) Diagram of the profilaggrin structure. At the amino terminus is a 50 kDa domain which contains a Ca^{2+}-binding EF hand. Following that are a series of internal repeats (PIRs), the number and size of which varies with the animal. At the carboxyl terminus of rat profilaggrin is a unique 15-residue segment. There are approximately 20 PIRs of 405 residues in each molecule of rat profilaggrin utilized in this study. Processing to release filaggrin occurs at several amino acids approximately 65 residues from the carboxyl terminus of each PIR, so that the boundaries of filaggrin and of the PIR are offset from one another. (B) Diagram of the placement of phosphorylation sites within the PIR of rat profilaggrin. The model is drawn to reflect the boundaries of the PIR, rather than of the filaggrin sequence (Haydock & Dale, 1990), placing the proteolytic processing sites within the PIR (indicated by an arrow). The solid line represents one PIR, and the gray lines indicate the adjacent PIRs; each short segment represents 10 residues. Residue 1 is the amino-terminal residue of profilaggrin peptide P1. The cleavage site is between residues 12 and 14 in P1. The four domains are based on those identified in Haydock and Dale, except that several are grouped to form a highly conserved domain I containing the linker and the His/Gln-rich segments [as originally identified in McKinley-Grant et al. (1989)]. The phosphorylation sites of profilaggrin are clustered in three of these domains, with P46 and P1 in the conserved domain I, P23 and P25 in domain II, and P31 and P36 in domain III. Domain IV is Gly/Ala/Ser rich, possibly representing a flexible hinge; the variable domain II is missing in mouse profilaggrin. The "+" symbols indicate the His/Gln-rich basic segment of the domain I; the placement of residues 160 through 270 (domains II and III) next to this basic segment is speculative. The proposed interactions may occur between adjacent PIRs as well. (C) Alignment of the rat phosphopeptides with the corresponding mouse phosphopeptides. In some cases, the amino acid sequence beyond that of the peptide is shown to clarify the alignment. The phosphorylated residues are in bold type; only a few of the phosphorylated residues of mouse profilaggrin have been identified (Resing et al., 1985, and unpublished data). Residues identical in the two aligned sequences are boxed; conservative replacement of D/E and T/S are also boxed.

other species (such as guinea pig and rabbit, which have much larger PIRs than those in rat and mouse) will be necessary to clarify the evolution of these phosphorylated sequences.

The presence of a large number of basic residues in profilaggrin lends itself to the speculation that an important determinant of profilaggrin secondary and tertiary structure involves salt bridges between these residues and the phosphate. In support of this model, several peptides, both phosphorylated and unphosphorylated, were found with noncovalently bound H_3PO_4, even after HPLC. Furthermore, tryptic cleavage at several sites near the phosphophorylated residues was inhibited in comparison to their cleavage in filaggrin (for example, producing P30/31 and P45/46), suggesting that the phosphate may be masking the arginines. The possible role of salt bridges between the basic residues and the phosphorylated residues is indicated in Figure 8A by juxaposition of the conserved Q/H basic region next to the second and third domains, which feature the QSPD-containing phosphorylated sequences.

The stoichiometry of phosphorylation at the individual sites ranges from less than 15% to nearly 100%. The loss of 5% of the phosphate during purification cannot account for this heterogeneity, although selective dephosphorylation at a few sites cannot be excluded. In the monophosphorylated forms of P1 and P30 (which are diphosphorylated), only one residue was phosphorylated. This result suggests that ordered phosphorylation (or dephosphorylation) occurs at these sites, because random phosphorylation would produce two monophosphorylated forms. Analysis of P46*[4] and P46*[6] also indicated that a similar ordered phosphorylation (or dephosphorylation) occurs in this peptide (data not shown). Sequential addition of phosphate has been observed in other multiply phosphorylated proteins. For example, casein kinase II must phosphorylate certain residues in glycogen synthase before glycogen synthase kinase-3 can phosphorylate other residues (Roach, 1990a).

The clustering of phosphorylated residues as seen in profilaggrin are often observed in other proteins and are common in cytoskeletal proteins (Roach, 1991b). In this case, phosphorylation may play a role in packing profilaggrin into keratohyalin, in preventing premature aggregation of keratin, in directing the association of other proteins with keratohyalin, and/or in directing proteolytic processing. It is probable that different sites provide different functions. The present study demonstrates that electrospray mass spectrometry can facilitate the exploration of the function of profilaggrin phosphorylation.

ACKNOWLEDGMENT

We acknowledge Lowell Ericcson for assistance with mass spectrometry, Santosh Kumar for advice on peptide purification, Robert Resing for preparation of Figure 1, and Dr. Natalie Ahn for many helpful discussions on phosphorylation and for critical reading of the manuscript.

REFERENCES

Biemann, K. (1990a) *Methods Enzymol. 193,* 455–79.
Biemann, K. (1990b) *Methods Enzymol. 193,* 886–897.
Burlet, O., Yars, C.-Y., & Gaskell, S. J. (1992) *J. Am. Soc. Mass Spectrom. 3,* 337–344.
Byford, M. F. (1991) *Biochem. J. 280,* 261–265.
Dale, B. A., Presland, R. B., P., F., Kam, E., & Resing, K. A. (1993) *Molecular Biology of the Skin,* pp 79–106, Academic Press, New York.
Gan, S.-Q., McBride, O. W., Idler, W. W., Markova, N., & Steinert, P. M. (1990) *Bichemistry 29,* 9432–9440.
Haugen-Scofield, J., Resing, K. A., & Dale, B. A. (1988) *J. Invest. Dermatol. 91,* 553–559.
Haydock, P. V., & Dale, B. A. (1990) *DNA Cell Biol. 9,* 251–261.
Hunt, D. F., Yates, J. R. I., Shabanowitz, J., Winston, S., & Hauer, C. R. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83,* 6233–6237.
Kam, E., Resing, K. A., Lim, S., & Dale, B. A. (1993) *J. Cell Sci. 106,* 219–226.
Lonsdale-Eccles, J. D., Teller, D. C., & Dale, B. A. (1982) *Biochemistry 21,* 5940–5948.
Markova, N. G., Marekov, L. N., Chipev., C. C., S.-Q. Gan, Idler, W. W., & Steinert, P. M. (1993) *Mol. Cell. Biol. 13,* 613–625.
McKinley-Grant, L. J., Idler, W. W., Bernstein, I. A., Parry, D. A., Cannizzaro, L., Croce, C. M., Huebner, K., Lessin, S. R., & Steinert, P. M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86,* 4848–4852.
Presland, R. B., Haydock, P. V., Fleckman, P., Nirunsuksiri, W., & Dale, B. A. (1992) *J. Biol. Chem. 267,* 23772–23781.
Resing, K. A., Dale, B. A., & Walsh, K. A. (1985) *Biochemistry 24,* 4167–4175.
Resing, K. A., Walsh, K. A., Haugen-Scofield, J., & Dale, B. A. (1989) *J. Biol. Chem. 264,* 1837–1846.
Resing, K. A., Al-Alawi, N., Blomquist, C., Fleckman, P., & Dale, B. A. (1993a) *J. Biol. Chem. 268,* 25139–25145.
Resing, K. A., Johnson, R. S., & Walsh, K. A. (1993b) *Biochemistry 32,* 10036–10045.
Resing, K. A., Mansour, S. J., Hermann, A. S., Johnson, R. S., M., C. J., Fukasawa, K., VandeWoude, G. F., & Ahn, N. A. (1995) *Biochemistry 34,* 2610–2620.
Roach, P. J. (1990) *FASEB J. 4,* 2961–2968.
Roach, P. J. (1991) *J. Biol. Chem. 266,* 14139–14142.
Rothnagel, J. A., & Steinert, P. M. (1990) *J. Biol. Chem. 265,* 1862–1865.
Siuzdak, G. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91,* 1290–11297.

BI9500382