Amino Acid Sequence and Sites of Phosphorylation in a Highly Acidic Region of Nucleolar Nonhistone Protein C23†

Mark D. Mamrack,‡ Mark O. J. Olson,* and Harris Busch

ABSTRACT: Novikoff hepatoma nucleolar nonhistone proteins, C23 and B23, contain highly acidic phosphorylated regions (Mamrack, M. D., et al. (1977) *Biochem. Biophys. Res. Commun. 76*, 150–157). Tryptic peptides from protein C23 containing these regions were purified by DEAE-Sephadex columns and paper electrophoresis at pH 1.8. One of these, peptide C23-Ca, was sequenced by combined automated and conventional methods. The proposed amino acid sequence is shown in eq 1. This peptide was found in three ^{32}P-labeled forms with phosphoryl groups at positions 8 and 25, and probably 28. The highly acidic sequences adjacent to the phosphorylation sites represent a unique class of phosphorylation sites different from those in histones or substrates for cytoplasmic cAMP-dependent kinases.

Ala-Ala-Pro-Ala-Ala-Pro-Ala-Ser-Glu-Asp-Glu-
Asp-Glu-Glu-Asp-Asp-Asp-Asp-Glu-Asp-Asp-Asp-Asp-
Asp-Ser-Gln-Glu-Ser-Glu-Glu-Glu-Asp-Glu-Glu-
Val-Met-Glu-Ile-Thr-Pro-Ala-Lys (1)

In a variety of systems, phosphorylation of nuclear acidic proteins has been correlated with increased chromatin template activity (Kleinsmith et al., 1976; Kleinsmith, 1978). Phosphorylation of nuclear proteins is catalyzed by cyclic nucleotide dependent and independent phosphoprotein kinases (Takeda et al., 1971; Rikans & Ruddon, 1976; Kish & Kleinsmith, 1974; Thomson et al., 1977; Cohen et al., 1975). The nucleolus, which contains both rDNA and systems for assembly of preribosomal particles (Busch & Smetana, 1970), is a useful model system for examining the role of protein phosphorylation. The activity of protein kinases associated with nucleoli (Grummt, 1974; Kang et al., 1974) correlates with nucleolar synthetic activity in liver regeneration (Ballal et al., 1975), neoplasia (Kang et al., 1975; Olson et al., 1978), or hormone treatment (Wilson & Ahmed, 1975), suggesting that protein phosphorylation is an important component of the process of ribosome biogenesis.

Recently it was found that protein C23 was the major protein phosphorylated by isolated nucleolar protein kinase (Olson et al., 1978). Protein C23 and an additional nonhistone protein, B23, were isolated from nucleoli and found to contain phosphorylated highly acidic tryptic peptides (Mamrack et al., 1977). The present study was initiated to determine the positions of the acidic and phosphorylated residues in such peptides. One purified peptide from protein C23, C23-Ca, contains a 27-residue sequence in which 24–27 of the amino acids bear negative charge as glutamic acid, aspartic acid or phosphoserine.

Materials and Methods

Isolation and Labeling of Nucleoli. Freshly prepared nucleoli (Matsui et al., 1977) from 6-day-old Novikoff hepatoma ascites cells grown in Holtzman rats were labeled in vitro with [γ-^{32}P]ATP as previously described (Mamrack et al., 1977). The proteins soluble in 0.4 N H_2SO_4 were separated on acid/urea–polyacrylamide gels (Mamrack et al., 1977; Knecht & Busch, 1971; Busch et al., 1974). Protein band C23 was eluted from the gels and applied to a column of hydroxylapatite (1 × 10 cm) and eluted stepwise according to the procedure described by Olson & Guetzow (1978). The fraction eluting at 250 mM phosphate was used for further studies. The protein at this stage was analyzed for homogeneity by two-dimensional gel electrophoresis (Busch et al., 1974).

Isolation of Phosphopeptides. Protein C23 was digested with Worthington trypsin (4%, w/w) for 4 h at 37 °C in 0.2 M *N*-ethylmorpholine acetate (pH 7.8) and the digest was chromatographed on DEAE-Sephadex A-25 columns (Mamrack et al., 1977). Of the three main peaks, peak C, as previously described by Mamrack et al. (1977), was used for the studies below.

The phosphopeptides of peak C were separated by high voltage paper electrophoresis at pH 1.8 (2% formic acid, 8% acetic acid) at 3000 V for 1.5 h and localized by autoradiography (Kodak XR5 film). After elution with 0.1 M pyridine–acetate buffer (pH 7.5), the peptides were further purified on Sephadex G-25 (0.5 × 95 cm) in 1 mM HCl.

Cleavage of Tryptic Phosphopeptides. For sequence analysis, the peptides were further digested with Pronase (Calbiochem, San Diego, CA, 10% w/w) for 2 h at 37 °C in 100 μL of 0.1 M pyridine–acetate (pH 7.5) or with thermolysin (Boehringer-Mannheim, Indianapolis, IN) or proteinase K (Beckman) under the same conditions for 20 h. Carboxypeptidase Y (Boehringer-Mannheim) digestion was done at 37 °C for various times in 0.1 M pyridine–acetate (pH 5.5); the reaction was stopped by heating at 100 °C for 2 min. Partial acid hydrolysis was done (Light, 1967) with: (a) 0.03 N HCl, 110 °C, 90 min; (b) 2 N HCl, 110 °C, 2 h; or (c) 5.7 N HCl, 37 °C, 24 h. The hydrolysates were analyzed by paper electrophoresis at pH 1.8.

Amino Acid and Sequence Analysis. After hydrolysis of the proteins and peptides in 5.7 N HCl at 110 °C for 22 h,

† From the Nuclear Protein Laboratory, Department of Pharmacology, Baylor College of Medicine, Houston, Texas 77030. *Received December 20, 1978.* These studies were supported by Cancer Program Project Grant CA-10893, P. 3, awarded by the National Cancer Institute, Department of Health, Education and Welfare, the Pauline Sterne Wolff Memorial Foundation, the Davidson Fund, and a generous gift from Mrs. Jack Hutchins.

‡ Predoctoral trainee supported by the Public Health Service Training Award GM 07405-1, Department of Health, Education and Welfare. Present address: Biology Division, Oak Ridge National Laboratory, Oak Ridge, TN 37830.

0006-2960/79/0418-3381$01.00/0

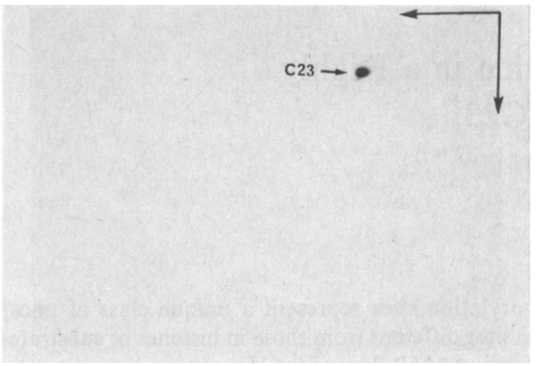

FIGURE 1: For analysis of the purity of isolated protein C23, two-dimensional electrophoresis of the product obtained from hydroxylapatite chromatography was carried out; the horizontal dimension was a 6% polyacrylamide acid–urea gel and the vertical dimension was an 8% polyacrylamide–$NaDodSO_4$ slab gel at pH 7.4 (Busch et al., 1974). The gel was stained in 0.25% Coomassie Blue/9% acetic acid/45% methanol for 6 h and destained in 10% acetic acid/10% methanol.

amino acid composition was determined with a Beckman 121-M amino acid analyzer. Milder conditions of hydrolysis (2 N HCl, 8 h, 110 °C, in vacuo) were used to release phosphorylated amino acids. Analysis of the hydrolysis products by pH 1.8 electrophoresis revealed the presence of phosphoserine but not phosphothreonine in peptide C23-Ca. Since yields of phosphoserine were variable, precise quantitation by the amino acid analyzer was not possible. For analysis of tryptophan, samples were hydrolyzed with 3 N mercaptoethanesulfonic acid for 22 h at 110 °C in vacuo (Penke et al., 1974).

The dansyl chloride procedure of Hartley (1970) was used for NH_2-terminal analysis of peptides. Sequence analysis was performed on a Beckman 890-B sequenator by utilizing a 0.1 M Quadrol program (Brauer et al., 1975). For small peptides, 3 mg of polybrene (Pierce, Rockford, IL) was added to the sequencer cup to minimize extractive losses (Klapper et al., 1978). The anilinothiazolinone derivatives were converted to the phenylthiohydantoins (Pth)[1] of amino acids by treatment with 1 M HCl at 80 °C for 10 min. The Pth-amino acids were identified by gas chromatography (Niall, 1973), one- and two-dimensional polyamide thin-layer chromatography (Summers et al., 1973), or amino acid analysis after hydrolysis in 57% HI for 18 h at 130 °C (Smithies et al., 1971) or in 5.7 HCl for 20 h at 150 °C (Van Orden & Carpenter, 1964). Since chromatography of Pth-phosphoserine was not practical with any of the above systems of analysis, phosphorylated amino acids in the sequence were placed by analyses of ^{32}P radioactivity in the water layer after conversion and ethyl acetate extraction. To reduce the background of ^{32}P radioactivity, prior to sequencing the samples were subjected to one or two preliminary cycles in the sequencer in which phenyl isothiocyanate was absent.

Results

Protein Purity and Composition. Protein C23 isolated by preparative gel electrophoresis and hydroxylapatite chromatography was found to be of homogeneity greater than 90% by two-dimensional polyacrylamide gel electrophoresis (Figure 1). Amino acid analysis indicated that protein C23 was rich in acidic amino acids (Table I). However, the protein also contained a relatively high percentage of lysine, glycine, and alanine. The acidic to basic amino acid ratio was calculated to be two, suggesting a negative net charge on the protein.

[1] Abbreviations used: Pth, phenylthiohydantoin; $NaDodSO_4$, sodium dodecyl sulfate.

Table I: Amino Acid Composition of Protein C23

amino acid	mol %[a]
Asx	12.0
Thr[b]	5.6
Ser[b]	5.4
Glx	18.5
Pro	5.3
Gly	10.0
Ala	10.3
$^1/_2$-cystine[c]	0.0
Val	6.0
Met	1.2
Ile	2.5
Leu	5.6
Tyr	0.8
Phe	3.5
Lys	11.8
His	0.5
Trp[d]	0.0
Arg	2.9
Asx, Glx / Lys, His, Arg	2.0

[a] Average of two analyses of protein C23 purified by preparative polyacrylamide gel electrophoresis followed by hydroxylapatite chromatography. The protein eluting at 250 mM phosphate was used for analysis. [b] Threonine and serine were corrected for 7% and 10% destruction, respectively. [c] Not detected on chromatogram. [d] Tryptophan was not found after hydrolysis in 3 N mercaptoethanesulfonic acid.

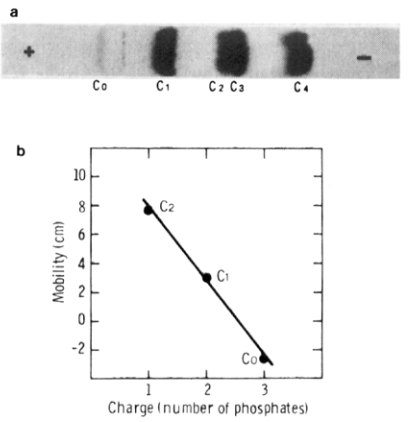

FIGURE 2: (a) Autoradiogram of fraction C separated by paper electrophoresis. The sample was desalted, placed on Whatman 3MM paper, and electrophoresed at pH 1.8 for 1.5 h at 3000 V; the paper was placed on Kodak XR5 film for 24 h. The major phosphopeptides are indicated. Peptide C_0 was the only peptide that migrated toward the anode. (b) Plot of relative mobility vs. presumed number of phosphate residues per peptide molecule. The data from Figure 2a were plotted (Offord, 1977) to test the linearity of the relationship between electrophoretic mobility and integral multiples of the negative charge carried by each species of the Ca tryptic peptide. The peptides were assumed to have identical molecular weights as indicated by amino acid composition (Table II).

Preliminary Characterization of Phosphopeptide Fraction C23-C. Paper electrophoresis at pH 1.8 (Figure 2a) showed that five radioactive components were present in phosphopeptide fraction C from DEAE-Sephadex (Mamrack et al., 1977). One peptide (C_0) moved slightly off the origin toward the positive pole. The other four peptides (C_1–C_4) were positively charged and were numbered in order of increasing mobility. On the basis of amino acid analysis (Table II), two groups of phosphopeptides were present. Group Ca (C_0, C_1, and C_2) contained 1 methionine, 1 isoleucine, and 6 alanine residues which were absent from group Cb; 25 of the 42 amino acids were aspartic or glutamic acid residues. In addition to

Table II: Amino Acid Ratios of Fraction C23-C Peptides

	group Ca			group Cb	
	C0	C1	C2	C3	C4
Lys	1.0	1.0	1.0	2.0	2.0
His					
Arg					
Asx	11.7	11.8	11.9	11.2	11.8
Thr	0.9	0.9	0.9		
Ser[a]	1.2	1.3	1.3	1.3	1.4
Glx	12.6	13.0	12.9	13.8	13.7
Pro	3.0	3.1	2.9	2.2	2.1
Gly					
Ala	5.3	5.6	5.4	0.8	0.7
Val	1.0	1.0	1.0	1.7	1.9
Met	0.8	0.8	1.0		
Ile	0.9	0.9	1.1		
Leu	0.9	0.9	1.1		
Phe				1.0	1.0
Tyr					

[a] The value for serine has not been corrected for destruction during hydrolysis inasmuch as the phosphoserine recovery as serine was variable (30–60%).

Table III: Carboxypeptidase Y Digestion of Peptide C23-C1 and -C2

	C23-C1				C23-C2
	15 min[a]	30 min[a]	120 min[a]	24 h[b]	30 min[a]
Lys	1.0 (52%)	1.0 (86%)	1.0 (60%)	1.0 (65%)	1.0 (75%)
Ala		1.0	0.96	0.7	0.93
Pro		0.89	0.96	0.8	0.93
Thr		0.75	0.96	0.9	0.86
Ile		0.45	0.91	0.8	0.52
Glu		1.13	1.13	3.7	
Val			0.74	1.0	
Met			0.70[c]	0.7	
Asp				1.3	

[a] Peptides C23-C1 and C23-C2 (0.5 nmol) were digested for various times at 37 °C as described under Materials and Methods. Results are expressed as ratios of released amino acids to lysine with the percent yield given in parentheses. [b] Peptide C23-C1 (4 nmol) was digested with carboxypeptidase Y for 24 h. Half the sample was applied directly to the amino analyzer after separation of free amino acids on Sephadex G-25 and the other half was hydrolyzed in 5.7 N HCl for 22 h at 110 °C. In the hydrolyzed sample, methionine was detected. In the unhydrolyzed sample, a peak appeared which comigrated with methionine sulfoxide. The results are for the hydrolyzed sample. [c] Determined as methionine sulfoxide.

amino acid analysis, further evidence for the similarity of the polypeptides in group Ca was the similarity of digestion products of carboxypeptidase Y treatment (Table III, 30 min). In group Cb (C_3, C_4), 1 phenylalanine residue was present. These peptides contained 2 lysines as compared with 1 in group Ca; in this group, 26 of the 37 amino acids were glutamic or aspartic acid residues.

Amino Acid Sequence of the Amino-Terminal Portion of Phosphopeptide C23-Ca. Automated sequence analysis identified the amino acids in positions 1–31 (Table IV and Figure 3) with the exception of amino acids 25, 27, and 28. The first seven amino acids consisted of the repeating sequence Ala-Ala-Pro-Ala-Ala-Pro-Ala. Position 8 was identified as phosphoserine; the serine in this position was found in the form of alanine by the amino acid analyzer after HI hydrolysis (Smithies et al., 1971), and ^{32}P was detected in the water layer of step 8. The presence of a serine residue in position 8 was confirmed by peptides C1H2 and C2P4 (Figure 3 and Table V) produced by partial digestion with acid and Pronase, respectively.

Positions 9–24 were sequences of glutamic and aspartic acid

Table IV: Summary of Automated Sequence Analysis of Peptide C23-C1

	amino acid (nmol)[a]							
step no.	Ala/ Ser	Pro	Asx	Glx	Asx/ Glx[b]	GC[c]	polya-mide[d]	step assign.
1	1.04					Ala		Ala
2	1.08					Ala		Ala
3	0.24	0.50				Pro		Pro
4	0.92	0.17				Ala		Ala
5	1.00	0.12				Ala		Ala
6	0.24	0.32				Pro		Pro
7	0.78	0.12				Ala		Ala
8	0.48					(–)		Ser(P)[e]
9	0.48			1.22	0.1	(–)	Glu	Glu
10	0.40		1.18	0.70	1.7		Asp	Asp
11			0.68	1.28	0.5		Glu	Glu
12			1.16	0.78	1.5		Asp	Asp
13			0.58	1.22	0.5		Glu	Glu
14			0.40	1.42	0.3	(–)		Glu
15			0.84	0.80	1.1		Asp	Asp
16			0.98	0.40	2.5		Asp	Asp
17			1.16	0.30	3.9		Asp	Asp
18			1.00	0.24	4.2		Asp	Asp
19			0.60	0.74	0.8	(–)	Glu	Glu
20			0.60	0.42	1.4	(–)	Asp	Asp
21			0.62	0.30	2.1		Asp	Asp
22			0.62	0.26	2.4		Asp	Asp
23			0.5	0.16	3.1		Asp	Asp
24			0.48	0.14	3.4		Asp	Asp
25			0.30	0.14	2.1	(–)	?	
26			0.26	0.17	1.5	Gln	Gln	Gln
27			0.13	0.17	0.8	(–)		Glu
28			0.13	0.16	0.8	(–)		
29			0.12	0.18	0.7	(–)	Glu	Glu
30			0.12	0.20	0.6	(–)	Glu	Glu
31			0.13	0.20	0.7	(–)	Glu	Glu

[a] Pth-amino acids were hydrolyzed in HI for 18 h at 130 °C and analyzed on the amino acid analyzer. Serine is converted to alanine under these conditions and methionine is destroyed (Smithies et al., 1971). Three separate sequencer runs were made. [b] The ratio of asparatic to glutamic acid was used to identify steps in which overlap was a problem. The ratio for each step was consistent in three separate sequencer runs. [c] Gas chromatography was used to distinguish alanine from serine in the initial steps as well as glutamine from glutamic acid in other steps as indicated. [d] Thin-layer chromatography was used to identify amidated amino acids on polyamide sheets (Summers et al., 1973). [e] The ^{32}P radioactivity above background was determined from the water layers after the Pth-amino acids had been extracted with ethyl acetate; cpm of ^{32}P in steps 8, 9, 10, and 11 were 150, 180, 120, and 130, respectively.

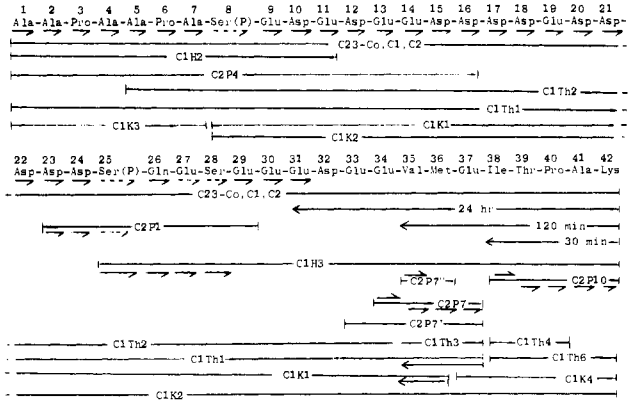

FIGURE 3: Composite diagram of sequence analysis of peptide C23-Ca. The symbols used are: (→) sequencer steps, broken arrows indicate inconclusive steps; (⇀) dansyl N terminals; (←) carboxypeptidase Y, unbroken lines indicate placement of groups of amino acids; (—) fragments placed by composition.

(Figure 3). Because of the partial overlap between steps, the basis for the assignment was the ratio as aspartic to glutamic

Table V: Fragments of Peptide C23-Ca Used in Sequence Analysis[a]

	C1H2 1-11	C1H3 25-42	C2P1 23-29	C2P4 1-15	C2P7 34-37	C2P7′ 35-36	C2P7′ 33-37	C2P10 38-42
Asp	1.1	2.3	2.0	4.3				
Thr		0.8						1.0
Ser[b]	0.6 (1)	0.8 (2)	0.8 (2)	1.1				
Glx	2.0	7.6	2.9	3.8	2.3	3.4		
Pro	1.9	0.8		2.3				1.1
Ala	5.0	0.8		5.0				1.0
Val		0.9			1.0	1.0	1.0	
Met		0.9			1.0	1.0	1.0	
Ile		0.8						0.9
Lys		1.0						0.9
^{32}P[c]	+	+	+					

[a] The values are expressed as residues per peptide. The nomenclature of these peptides is as follows: The first two symbols represent the peptide that was digested (i.e., C1 = C23-C1 and C2 = C23-C2). The next letter designates the procedure used: H for partial acid hydrolysis, P for Pronase treatment. The numbers directly below the peptide numbers indicate the residue numbers encompassed by the peptide. [b] The uncorrected value is given as well as the theoretical value for serine in parentheses; the recovery of phosphoserine as serine was variable (30–60%). [c] + sign denotes the peptide fragments that were radioactively labeled with ^{32}P following the incubation.

Table VI: Summary of the Sequence Analysis of Peptide C1H3

	amino acid released (nmol)[a]				
step no.	Asx	Glx	Ala/Ser	CPM	assignment
1	0.37	0.28	0.75	238	Ser P 25
2	0.23	0.66	0.48	188	Glx 26
3	0.08	0.48	0.15	138	Glx 27
4	0.14	0.35	0.27	113	Ser 28

[a] Peptide C1H3 (1 nmol; Table V and Figure 3) was applied to the sequencer cup and subjected to Edman degradation in the presence of polybrene. The high value of Asx in the first step represents a contaminating peptide containing several aspartic acid residues. One sequencer cycle was run without PITC prior to beginning the degradation to minimize background amino acids and cpm.

acid (Table IV). These ratios were consistent for three separate sequencer runs.

The sequence assignments in positions 25–28 were confirmed by sequence analysis of peptide C1H3 (Table V). Fragment C1H3 (Table VI) contained 2 serine residues which were separated by 2 Glx residues, Ser-Glx-Glx-Ser-; ^{32}P was found in the aqueous phase of the first serine position (Table VI).[2] The glutamine at position 26 was identified in the initial sequencer runs by isothermal gas chromatography at 235 °C and by polyamide layer chromatography (Table IV).

Amino Acid Sequence of the C-Terminal Portion of the Peptide C23-Ca. Lysine was confirmed at the C terminus of peptide C23-C1 by carboxypeptidase digestion (Table III). Positions 38–42 (Ile-Thr-Pro-Ala-Lys) were defined by analysis of peptide C2P10 (Table V) in the sequencer with polybrene as carrier. The sequence Glu-Glu-Val-Met-Glu in positions 33–37 was determined by the composition of peptide C2P7′ (Table V), sequence analysis of peptide C2P7, and N-terminal analysis of the dipeptide C2P7″. In addition, carboxypeptidase Y digestion (30 min) of thermolysin peptide C1Th1 (Figure 3 and Table VII) released equimolar quantities of glutamic acid, valine, and methionine, which indicated that the Val-Met sequence was near or at the C-terminal end of C1Th1. Cleavage by thermolysin at the amino group of isoleucine-38 was in agreement with the known specificity of the enzyme and further supported the placement of the Val-Met-Glu sequence adjacent to isoleucine-38. All 42 residues of peptide C1 were accounted for by thermolysin peptides and by proteinase K peptides (Table VII).

Inasmuch as the above data provided evidence for positions 1–31 and 33–42, the aspartic acid, which was released with extended (24 h) carboxypeptidase digestion (Table III and Figure 3), was placed at position 32 by difference. This assignment was consistent with the difference between the 120-min and 24-h carboxypeptidase data (Table III) and analyses of thermolysin fragments C1Th1 and C1Th2 (Table VII and Figure 3).

Phosphorylation States of Phosphopeptides C23-Ca. The difference in mobility observed for peptides C_0, C_1, and C_2 (Figure 2a) suggested that they differed only in phosphate content: (1) all three peptides had the same amino acid composition; (2) the products of carboxypeptidase digestion of peptides C_2 and C_1 were very similar (Table III); (3) partial acid hydrolysis and Pronase fragments were essentially identical for peptides C_2 and C_1 with the exception of the nonphosphorylated N-terminal fragment (C2P4, Table V) present only in digests of C_2. By assuming identical sequences for each form (C_0, C_1, C_2), the mobility upon paper electrophoresis at pH 1.8 was directly proportional to phosphate content (Figure 2b) according to the procedure of Offord (1977). These results are consistent with peptides C_2, C_1, and C_0 having 1, 2, and 3 phosphoryl groups, respectively. Peptide C_0 accounted for only 5% and peptide C_1 for 60–70% of the total mass of the Ca fraction as isolated.

Direct evidence for phosphate at positions 8 and 25 was obtained from sequencer runs (Tables IV and VI). Since serines were found only at positions 8, 25, and 28, and only phosphoserine but not phosphothreonine was found in peptides C_0, C_1, and C_2, the most probable remaining site for the third phosphoryl group was at position 28. In addition, a ^{32}P-labeled peptide obtained in small amounts from partial acid hydrolysates of peptide C_1 contained glutamic acid and migrated faster than inorganic phosphate on pH 1.8 electrophoresis. Similar peptides from casein and phosvitin have been shown to contain 2 or more phosphoserines (Williams & Sanger, 1959). Rehydrolysis of the fast moving spot in 2 N HCl at 110 °C for 3 h liberated radioactive phosphoserine. This experiment suggested that a subpopulation of the peptide C_1 contained phosphoserines at both positions 25 and 28.

Methionine at Position 36. The methionine residue at position 36 was placed in the sequence by the peptide C2P7 and its fragments (Figure 3). The products of carboxypeptidase Y digestion (37 °C, 24 h) of peptide C23-C1 were

[2] Radioactive peptide C2P1 (Table V and Figure 3) confirmed the presence of phosphoserine at position 25. During automatic sequencing, cycles 1 and 2 indicated aspartic acid. The radioactivity in the water layer increased from 64 cpm at cycle 2 to 126 cpm at cycle 3 and decreased to 85 cpm at cycle 4. Background was subtracted from the cpm.

Table VII: Thermolysin and Proteinase K Fragments of Peptide C23-C1[a]

	C1Th1 1-37	C1Th2 1-34	C1Th3 35-37	C1Th4 38-40	C1Th6 38-42	C1K1[b] 8-36	C1K2 8-42	C1K3 1-7	C1K4 37-42
Asx	12.0	12.0				11.6	12.5		
Thr				1.0	0.9		0.6		1.0
Ser	2.8	1.3				1.1	1.2		
Glx	13.4	12.0	1.1			12.0	13.2		1.1
Pro	2.1	1.3		1.0	1.0		1.2	2.0	1.0
Ala	5.1	2.5			1.0		1.1	5.0	1.1
Val	1.0		1.0			1.1	1.0		
Met	0.9		1.0			0.8	0.6		
Ileu				0.9	0.9		0.6		1.0
Lys					1.0		1.0		1.1
^{32}P	+	+	−	−	−	+	+	−	−

[a] Values expressed as residues per peptide. See Table V footnotes for explanation of abbreviations. [b] Peptide C1K1 was digested for 30 min with carboxypeptidase Y. Equimolar quantities of Val (0.20 nmol) and Met (0.23 nmol, as the sulfoxide) were liberated.

run on a Sephadex G-25 column. The included volume was divided into two equal parts. Half of the sample was hydrolyzed in 5.7 N HCl for 22 h at 110 °C. In the hydrolyzed sample, methionine was recovered in a 70% molar yield along with the other amino acids (Table III). However, in the unhydrolyzed sample, a peak was found in the position of a known sample of methionine sulfoxide. Inasmuch as methionine sulfoxide is converted to methionine in a 90% yield under conditions used for hydrolysis of proteins (Means & Feeney, 1971), the predominant form of methionine in this peptide appears to be the sulfoxide in these tryptic digests.

Discussion

The approach to understanding possible functional roles of nucleolar phosphoproteins C23 and B23 was initially undertaken by analysis of changes in their concentration and phosphorylation during various physiological changes (Kang et al., 1974, 1975; Ballal et al., 1975; Olson et al., 1978). The apparent high levels of phosphorylation prompted the study of the localization of the phosphoserine residues in these proteins. Tryptic peptides were recently isolated and found to contain many acidic amino acids (Mamrack et al., 1977). In peptide C23-Ca, remarkable features of the sequence include both the presence of phosphoserine residues and the clusters of numerous aspartic and glutamic acid residues. As noted above, this sequence is unique and represents the first nonhistone nuclear phosphopeptide sequence reported. Walker et al. (1976, 1978) and Goodwin et al. (1978) found aspartic and glutamic acids in a partially defined sequence of 41 amino acids in the nuclear protein HMG 1, but their peptide did not contain phosphoserine.

Although peptide C23-C1 was the first of the acidic phosphopeptides purified in sufficient yield for sequence analysis, others are now being analyzed. During the sequencing, the lability of the phosphoserine residues caused difficulty in determining peptide ratios and identification of sequencer steps. Identification was also hindered by the repetitive nature of the sequence in which overlap between steps was not easily distinguished from the newly liberated residues. However, the Asp/Glu ratio at each step was consistent in three different sequencer runs. The internal region of the peptide was relatively resistant to protease digestion. Partial acid hydrolysis and Pronase were used to cleave the clusters of acidic amino acids. Although initial analyses indicated a methionine residue in at least 80% molar yields, CNBr treatment was ineffective. The methionine was subsequently demonstrated to be in the oxidized form, methionine sulfoxide.

The phosphorylation of C23-Ca appears to partially follow the pattern of phosphorylation of some sites in casein. Casein contains several serine residues which are phosphorylated when followed two positions in the sequence by a glutamic acid or another phosphoserine (Mercier et al., 1971; Brignon et al., 1977). The serine at position 8 in peptide C23-Ca is followed by Glu-Asp-Glu. A relatively small fraction of the peptide is phosphorylated at both positions 25 and 28. Both serines are potential acceptor sites of the casein type with glutamic acid residues at positions 27 and 30. The phosphorylation of the serine at position 28 may affect the phosphorylation of the serine at position 25, as in phosphorylation of casein (Brignon et al., 1977). In contrast, serine residues in histones and cytoplasmic proteins that are phosphorylated by cAMP-dependent kinase are generally preceded by a basic amino acid two to three positions (Kemp et al., 1974; Williams, 1976).

At this time, none of the functions of such structures can be defined, although it is clear that they are potential binding sites for cations such as metal ions or for clusters of basic amino acids in histones (Busch, 1977). In connection with metal ion binding, it was recently demonstrated that silver selectively stains the nucleolus organizer regions (NORs) on chromosomes (Howell et al., 1975; Bloom & Goodpasture, 1976). In this laboratory, specific granules within nucleoli of Novikoff hepatoma and other tumors were shown to be silver-stained (Busch et al., 1979). The number of argyrophilic granules per nucleolus appears to depend on the rate of preribosomal RNA synthesis. In preliminary studies (M. Lischwe and H. Busch, 1978, in preparation) protein C23 stained with silver on polyacrylamide gels. It remains to be determined whether protein C23 is a nucleolus organizer protein or if the acidic regions of the protein are argyrophilic. However, if protein C23 is involved in organizing nucleolar structure, the variation in phosphorylation may affect binding of this peptide to cationic regions or be related to changes in three-dimensional structure.

On the basis of the rules of Chou & Fasman (1974), no regions in this peptide can maintain α helical or β conformations. However, peptide C23-Ca may fit the β-turn conformation where the polypeptide chain folds back on itself. Many of the phosphorylation sites reported in substrates of the cAMP-dependent kinase are in regions predicted to be β turns (Small et al., 1977). In peptide C23-Ca, the proline residues at either end in positions 6 and 40 are within regions with high β-turn potential, suggesting that the cluster of acidic residues may extend or loop out into space; the phosphorylation at position 8 may affect the extension of the loop

Acknowledgments

The authors thank Rose Busch for transplating the tumor cells and Karl Guetzow for technical assistance in amino acid and sequence analysis.

References

Ballal, N. R., Kang, Y. J., Olson, M. O. J., & Busch, H. (1975) *J. Biol. Chem. 250*, 5921–5925.

Bloom, S. E., & Goodpasture, C. (1976) *Hum. Genet. 34*, 199–206.

Brauer, A. W., Margolies, M. N., & Haber, E. (1975) *Biochemistry 14*, 3029–3035.

Brignon, G., Ribadeau-Dumas, B., Mercier, J. C., Pelissier, J. P., & Das, B. C. (1977) *FEBS Lett. 76*, 274–279.

Busch, G. I., Yeoman, L. C., Taylor, C. W., & Busch, H. (1974) *Physiol. Chem. Phys. 6*, 1–10.

Busch, H. (1977) in *Receptors and Hormone Action* (O'-Malley, B. W., & Birnbaumer, L., Eds.) Vol. I, pp 31–102, Academic Press, New York.

Busch, H., & Smetana, K. (1970) *The Nucleolus*, Academic Press, New York.

Busch, H., Daskal, Y., Gyorkey, F., & Smetana, K. (1979) *Cancer Res. 39*, 857–863.

Cohen, P., Watson, D. C., & Dixon, G. H. (1975) *Eur. J. Biochem. 51*, 79–92.

Goodwin, G. H., Walker, J. M., & Johns, E. W. (1978) in *The Cell Nucleus* (Busch, H., Ed.) Vol. VI, pp 181–219, Academic Press, New York.

Grummt, I. (1974) *FEBS Lett. 39*, 125–128.

Hartley, B. S. (1970) *Biochem. J. 119*, 805–822.

Howell, W. N., Denton, T. E., & Diamond, J. R. (1975) *Experientia 31*, 260–262.

Kang, Y. J., Olson, M. O. J., & Busch, H. (1974) *J. Biol. Chem. 249*, 5580–5585.

Kang, Y. J., Olson, M. O. J., Jones, C. E., & Busch, H. (1975) *Cancer Res. 35*, 1470–1475.

Kemp, B. E., Bylund, D. B., Huang, T. S., & Krebs, E. G. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3448–3452.

Kish, V. M., & Kleinsmith, L. J. (1974) *J. Biol. Chem. 249*, 750–761.

Klapper, D. G., Wilde, C. E., & Capra, J. D. (1978) *Anal. Biochem. 85*, 126–131.

Kleinsmith, L. J. (1978) in *The Cell Nucleus* (Busch, H., Ed.) Vol. VI, pp 221–261, Academic Press, New York.

Kleinsmith, L. J., Stein, J., & Stein, G. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 1174–1178.

Knecht, M. E., & Busch, H. (1971) *Life Sci. 10*, 1297–1309.

Light, A. (1967) *Methods Enzymol. 11*, 417–240.

Mamrack, M. D., Olson, M. O. J., & Busch, H. (1977) *Biochem. Biophys. Res. Commun. 76*, 150–157.

Matsui, S., Fuke, M., & Busch, H. (1977) *Biochemistry 16*, 39–45.

Means, G. E., & Feeney, R. E. (1971) in *Chemical Modification of Proteins*, p 164, Holden-Day, San Francisco, CA.

Mercier, J. C., Grosclaude, F., & Ribadeau-Dumas, B. (1971) *Eur. J. Biochem. 23*, 41–51.

Niall, H. D. (1973) *Methods Enzymol. 27*, 942–1010.

Offord, R. E. (1977) *Methods Enzymol. 47*, 51–69.

Olson, M. O. J., & Guetzow, K. A. (1978) *Biochim. Biophys. Acta 526*, 174–185.

Olson, M. O. J., Hatchett, S., Guetzow, K., Ballal, N. R., & Busch, H. (1978) *Cancer Res. 19*, 755.

Penke, B., Ferencz, R., & Kovacs, K. (1974) *Anal. Biochem. 60*, 45.

Rikans, L. F., & Ruddon, R. W. (1976) *Biochim. Biophys. Acta 433*, 73–86.

Small, D., Chou, P. Y., & Fasman, G. D. (1977) *Biochem. Biophys. Res. Commun. 79*, 341–345.

Smithies, O., Gibson, D., Fanning, E. M., Goodfliesh, R. M., Gilman, J. G., & Ballantyne, D. L. (1971) *Biochemistry 10*, 4912–4921.

Summers, B. R., Smythers, G. W., & Oroszlan, S. (1973) *Anal. Biochem. 53*, 624–628.

Takeda, M., Yamamura, H., & Ohga, Y. (1971) *Biochem. Biophys. Res. Commun. 42*, 103–110.

Thomson, J. A., Chiu, J. F., Sakuma, K., & Hnilica, L. S. (1977) *Cancer Res. 37*, 3266–3273.

Van Orden, H. O., & Carpenter, F. H. (1964) *Biochem. Biophys. Res. Commun. 14*, 399–403.

Walker, J. M., Shooter, K. V., Goodwin, G. H., & Johns, E. W. (1976) *Biochem. Biophys. Res. Commun. 70*, 88–93.

Walker, J. M., Hastings, J. R. B., & Johns, E. W. (1978) *Nature (London) 271*, 281–282.

Williams, J., & Sanger, F. (1959) *Biochim. Biophys. Acta 33*, 294–296.

Williams, R. E. (1976) *Science 192*, 473–474.

Wilson, M. J., & Ahmed, K. (1975) *Exp. Cell Res. 93*, 261–266.