FEBS Letters 463 (1999) 12–14 FEBS 23017

Isolation of a glycosylated form of the chicken eggshell protein ovocleidin and determination of the glycosylation site. Alternative glycosylation/phosphorylation at an *N*-glycosylation sequon

Karlheinz Mann*

Max-Planck-Institut für Biochemie, Am Klopferspitz 18, D-82152 Martinsried, Germany

Received 27 October 1999

Edited by Pierre Jolles

Abstract Ovocleidin, a major protein of the avian eggshell calcified layer, occurs in the eggshell soluble organic matrix in at least two forms. The major form is a phosphoprotein with two phosphorylated serines (OC-17) which was sequenced recently. A minor form is a glycosylated protein with identical sequence and only one phosphorylated serine (OC-23). The site of glycosylation is Asn^{59}, the only asparagine in the amino acid sequence contained in the *N*-glycosylation site consensus sequence, N-A-S. Ser^{61}, which is part of this site, is phosphorylated in OC-17 but not in OC-23 indicating that the two modifications are mutually exclusive. This is the first example of alternative glycosylation/phosphorylation occurring at an *N*-glycosylation site.

Key words: Ovocleidin; Eggshell; Glycosylation; Phosphorylation

1. Introduction

The avian eggshell consists of three layers: the most internal shell membranes which contain collagens as major components, the intermediate calcified zone and the outermost cuticle [1,2]. Electron microscopic analysis of the shell has shown the calcified layer to consist of at least two components, the ∼100 μm thick mamillary layer which is in close contact to the shell membranes and is thought to provide the starting sites for calcification, and the ∼300 μm thick palisade or crystalline layer, the thickest layer of the shell. The organic matrix of this layer can be isolated after decalcification. A major component of the soluble fraction of the matrix is a protein with M_r ∼17 000 called ovocleidin 17 (OC-17) [3]. This protein is produced by cells of the shell gland mucosa, tends to form larger aggregates and has been shown by immunohistochemical methods to be localised in the calcified layer. It consists of 142 amino acids including two phosphorylated serines [4]. The amino acid sequence forms a single C-type animal lectin domain and is significantly similar to a group of pancreas proteins, the pancreatic stone protein (lithostathin) and pancreatitis-associated proteins. Here we report the isolation of a glycosylated form of ovocleidin (OC-23) and the identification of the glycosylation site.

2. Materials and methods

2.1. Isolation of ovocleidins

Eggshells were treated with 5% EDTA for 30 min at 4–6°C to facilitate membrane removal. The powdered calcified layer was extracted with 10% acetic acid as described [4]. The extract was centrifuged at 14 000×g_{av} for 1 h at 4°C. The supernatant was dialysed against 5×10 vol. of 5% acetic acid. The slightly turbid solution was lyophilised. For HPLC on a Vydac C4 reversed phase column (50×4.6 mm) the lyophilised protein mixture was suspended in 10% acetic acid, centrifuged, filtered [4] and the soluble proteins were separated using a gradient of 14–46% acetonitrile in 0.1% trifluoroacetic acid in 40 min at a flow rate of 0.5 ml/min. The ovocleidin-containing fractions from C4 reversed phase HPLC runs were dried and dissolved in 0.1 M ammonium acetate buffer, pH 5, containing 4 M guanidinium and chromatographed on two TSK SW_{XL} 2000 columns (300×7.8 mm) connected in tandem using the same buffer at a flow rate of 0.3 ml/min. The separated proteins were dialysed against 5% acetic acid and lyophilised.

2.2. Analytical methods

SDS-polyacrylamide gel electrophoresis (SDS-PAGE) [5] was performed using 10–15% polyacrylamide gradient gels. Amino acid and hexosamine analyses were done on a Biotronik LC3000 after hydrolysis at 110°C in 6 M HCl for 20 h and 3 M HCl for 8 h, respectively. Carboxymethylation, cleavage with chymotrypsin, separation of the peptides, blotting onto polyvinylidene difluoride (PVDF) membrane and sequence analysis were done as previously described [4,6]. Protein-bound phosphate was determined using a malachite green assay [7] with 1 h hydrolysis at 60°C with ribonuclease A as a negative control. OC-17, which has been shown by amino acid sequence analysis and mass spectrometry to contain two phosphoserines [4], was used as a positive control. The amount of protein was determined by amino acid analysis as above. Automated Edman degradation was performed on an Applied Biosystems sequencer model 473A. Deglycosylation was done with peptide *N*-glycosidase F (Boehringer Mannheim) in 0.05 M sodium phosphate buffer, pH 8, containing 10 mM EDTA and 10 mM phenylmethylsulphonyl fluoride for 24 h at 23°C using 1 U/10 μg of carboxymethylated protein or 5 nmol of peptide.

3. Results and discussion

3.1. Isolation of ovocleidins

Ovocleidin was isolated from decalcified chicken eggshells as described [4] with the modification that dialysis was against 5% acetic acid, which prevented the precipitation of ovocleidin previously described with dialysis against water. Concomitantly, a protein band with M_r ∼23 000 and strong tailing, which had occurred only in small amounts in previous preparations, became more prominent in the ovocleidin pool obtained after reversed phase HPLC (Fig. 1). N-terminal amino acid sequence analysis of this protein (henceforth termed OC-23) blotted onto PVDF membrane after SDS-PAGE showed that it had the same N-terminal sequence as the major form of

*Fax: (49)-89-85782401.
E-mail: mann@biochem.mpg.de

PII: S0014-5793(99)01586-0

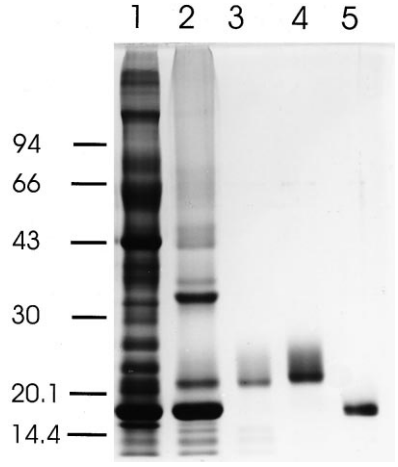

Fig. 1. Purification of ovocleidin monitored by PAGE. Lane 1, whole extract; lane 2, ovocleidin mixture after C4 reversed phase chromatography; lane 3, glycosylated ovocleidin (OC-23) after size exclusion chromatography (Fig. 2, peak A); lane 4, carboxymethylated OC-23 after rechromatography; lane 5, OC-23 after deglycosylation with peptide *N*-glycosidase F. The position of marker proteins is indicated in kDa at the left.

ovocleidin described previously [4]. Reversed phase HPLC on a short C4 column was used to enrich both forms of ovocleidin which appeared in the same broad peak with a slight enrichment of the larger form in the front fractions. The two forms of the protein were separated from each other and the remaining other proteins by size exclusion HPLC in a denaturing solvent (Fig. 2). The ratio of the larger to the smaller form was estimated to be 1:10 by comparison of the peak areas. OC-23, but not OC-17, was degraded partially by contaminating protease(s) when kept in denaturing solvent around pH 7. Buffering at pH 5 decreased this degradation and was used for size exclusion chromatography. To further minimise degradation, the samples were dissolved immediately before this purification step and dialysed against 5% acetic acid immediately after separation. Smaller fragments were removed by rechromatography after carboxymethylation of the protein (Fig. 1). The yield of OC-23 was 2–4 μg/g of eggshell powder. Amino acid analysis and hexosamine analysis showed that OC-23 had the same amino acid composition (later confirmed by sequence analysis) as OC-17 but contained in addition 3–4 glucosamines (3.4 ± 0.6) and 1 galactosamine (1.1 ± 0.3) per polypeptide chain (mean of two analyses ± maximal deviation) demonstrating that the different mobility in SDS-PAGE was due to glycosylation of site(s) which had not been modified in the major form of the protein. A molybdate/malachite green assay for phosphate after alkaline hydrolysis yielded 2.9 ± 0.5 phosphates/polypeptide chain for OC-17 and 1.5 ± 0.1 for OC-23 (mean of five determinations ± S.D.). The previous sequence analysis of OC-17 [4] had shown the presence of two phosphates/protein molecule indicating that the values found by the colour assay were too high for unknown reasons. Nevertheless, the data clearly indicated that OC-23 was a phosphoprotein and that it contained only half the amount of protein-bound phosphate as compared to OC-17.

3.2. Determination of the glycosylation site

While galactosamine is in most cases indicative of *O*-glycosylation of a protein, the presence of an *N*-glycan could not be ruled out from the experimentally determined ratio of galactosamine to glucosamine. Deglycosylation of OC-23 with peptide *N*-glycosidase F decreased the M_r of this form of ovocleidin to the M_r of OC-17 (Fig. 1) indicating that the carbohydrate was bound to asparagine. The amino acid sequence of ovocleidin contains only two asparagines, Asn^{59} and Asn^{134} [4]. Only one of them, Asn^{59}, was found in a typical *N*-glycosylation site consensus sequence, N-X-T/S. However, in the non-glycosylated form of ovocleidin Ser^{61}, which was part of this site, has been identified as one of the two phosphoserines of the protein [4]. To identify the possible site of *N*-glycosylation experimentally, the reduced and carboxymethylated protein was cleaved with chymotrypsin, the peptides separated by size exclusion chromatography on a Superdex Peptide column and by reversed phase HPLC using a C18 column, and then subjected to Edman sequence analysis [4]. The only peptide which showed differences compared to the corresponding sequences of OC-17 started at Arg^{52} and ended at Trp^{76}: Asn^{59} was not observed during sequence analysis and Ser^{61}, which was phosphorylated in the non-glycosylated protein, was detected in a concentration comparable to that of surrounding residues indicating that it was not phosphorylated. When this peptide was treated with peptide *N*-glycosidase F, rechromatographed and sequenced again, position 59 was determined to be aspartic acid. Together these

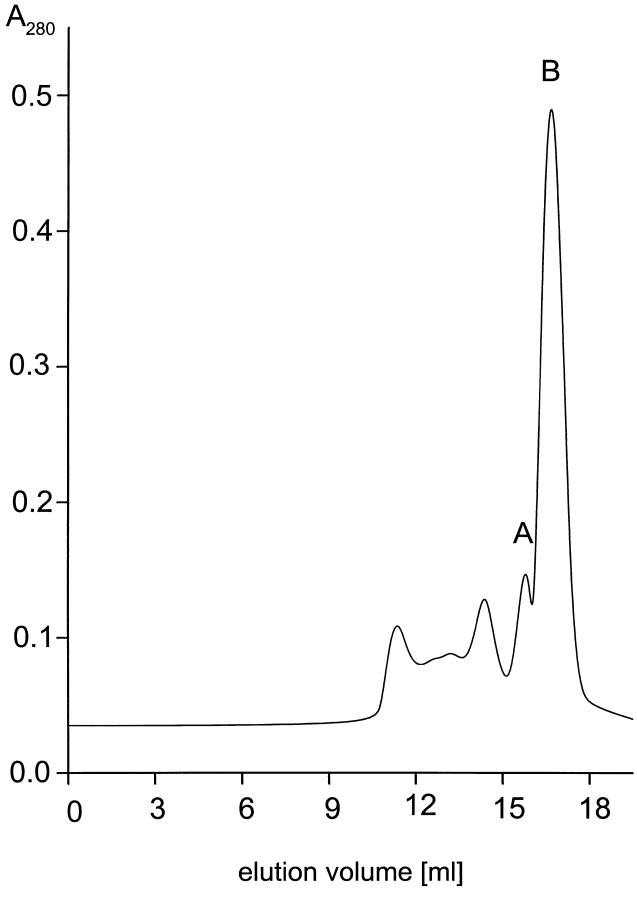

Fig. 2. Separation of the ovocleidin forms. OC-23 (peak A) was separated from OC-17 (peak B) by size exclusion chromatography at pH 5 in 4 M guanidinium. The peak areas were approximately 1:10.

results identified Asn[59] as an *N*-glycosylation site. Because Ser[61], which also formed part of the glycosylation site, was not phosphorylated, glycosylation of Asn[59] and phosphorylation of Ser[61] were apparently mutually exclusive.

Alternative glycosylation/phosphorylation is known to occur in many nuclear and cytoskeletal proteins [8] but at present has not been observed in extracellular proteins. In intracellular proteins this modification can involve *O*-glycosylation or alternative phosphorylation of a single serine [9,10] or residues in the close vicinity [11]. Ovocleidin seems to be the first example for the alternative glycosylation/phosphorylation of a protein involving *N*-glycosylation of an asparagine and alternative phosphorylation of the corresponding hydroxyamino acid forming an *N*-glycosylation sequon, N-X-T/S. Because *N*-glycosylation at such sites occurs only in extracellular proteins, this type of competitive posttranslational modification must be confined to extracellular proteins. *N*-Glycosylation is a very early event taking place cotranslationally in the endoplasmic reticulum. A direct competition between glycosylation and phosphorylation of ovocleidin would require a kinase in the same compartment. For many extracellular phosphoproteins the cellular location of the phosphorylation is not known. An exception is casein which has been reported to be phosphorylated in the Golgi apparatus of mammary gland cells of lactating mammals [12,13]. Furthermore, a casein kinase-like activity has been identified in the endoplasmic reticulum of cells producing dentin phosphoproteins and it has been proposed that phosphorylation of these extracellular proteins was a cotranslational event [14,15]. Therefore it may not be too far-fetched to suppose the presence of a kinase in the endoplasmic reticulum of chicken shell gland cells which produce ovocleidin.

While the presence of glucosamine in ovocleidin has not been reported, galactosamine has been identified using biotinylated lectins [3]. However, this was done with OC-17, which according to the sequencing results [4] was not glycosylated. Possibly these preparations contained glycosylated fragments of OC-23 with the same relative mobility as OC-17. The presence of galactosamine usually indicates *O*-glycosylation of a protein. However, sequence analysis of a set of peptides covering the entire sequence of ovocleidin showed only two modified amino acids identified by the lack of a detectable PTH-amino acid signal in the respective sequencing cycles. These were the glycosylated Asn[59] and the amino acid at position 61 of the ovocleidin sequence. The latter amino acid has been identified as phosphoserine in OC-17 by mass spectrometry [4]. Interestingly the sequence adjacent to this serine was identical to a recognition signal sequence for the attachment of glycosaminoglycans in some proteoglycan core proteins, D/E-X-S-G-X-G [16]. Although this signal was not used in OC-17 and OC-23, one cannot exclude the possibility that it may be used in other forms of ovocleidin. The presence of phosphate in OC-23 and the mobility shift in PAGE to the position of OC-17 after treatment with peptide *N*-glycosidase F may indicate that this amino acid is also phosphoserine in OC-23. Therefore, OC-23 may belong to a small group of proteins which have been shown to contain galactosamine (or rather *N*-acetylgalactosamine which is converted to galactosamine by hydrolysis) in their *N*-glycan. Other examples are bovine lactotransferrin [17], the snake venom enzyme batroxobin [18] and human urinary kallidinogenase [19]. OC-17 is apparently not derived from OC-23 because hydrolytic glycosidases convert glycosylated asparagine into aspartic acid. But OC-17 has asparagine in its *N*-glycosylation sequon. These data suggest that ovocleidin is alternatively modified and that these two forms of the protein may play different roles in the construction of the eggshell.

References

[1] Arias, J.L., Fink, D.J., Xiao, S., Heuer, A.H. and Caplan, A.I. (1993) Int. Rev. Cytol. 145, 217–250.
[2] Leach, R.M. (1982) Poult. Sci. 61, 2040–2047.
[3] Hincke, M.T., Tsang, C.P.W., Courtney, M., Hill, V. and Narbaitz, R. (1995) Calcif. Tissue Int. 56, 578–583.
[4] Mann, K. and Siedler, F. (1999) Biochem. Mol. Biol. Int. 47, 997–1007.
[5] Laemmli, U.K. (1970) Nature 227, 680–685.
[6] Matsudaira, P. (1987) J. Biol. Chem. 262, 10035–10038.
[7] Ekman, P. and Jäger, O. (1993) Anal. Biochem. 214, 138–141.
[8] Haltiwanger, R.S., Busby, S., Grove, K., Li, S., Mason, D., Medina, L., Moloney, D., Philipsberg, G. and Scartozzi, R. (1998) Biochem. Biophys. Res. Commun. 231, 237–242.
[9] Chou, T.-Y., Hart, G.W. and Dang, C.V. (1995) J. Biol. Chem. 270, 18961–18965.
[10] Medina, L., Grove, K. and Haltiwanger, R.S. (1998) Glycobiology 8, 383–391.
[11] Kelly, W.G., Dahmus, M.E. and Hart, G.W. (1993) J. Biol. Chem. 268, 10416–10424.
[12] Bingham, E.W., Farrell, H.M. and Basch, J.J. (1972) J. Biol. Chem. 247, 8193–8194.
[13] Pascall, J.C., Boulton, A.P. and Craig, R.K. (1981) Eur. J. Biochem. 119, 91–99.
[14] Sfeir, C. and Veis, A. (1995) J. Bone Miner. Res. 10, 607–615.
[15] Veis, A., Sfeir, C. and Wu, C.B. (1997) Crit. Rev. Oral Biol. Med. 8, 360–379.
[16] Bourdon, M.A., Krusius, T., Campbell, S., Schwartz, N.B. and Ruoslahti, E. (1987) Proc. Natl. Acad. Sci. USA 84, 3194–3198.
[17] Coddeville, B., Strecker, G., Wieruszeski, J.-M., Vliegenthart, J.F.G., van Halbeck, H., Peter-Katalinic, J., Egge, H. and Spik, G. (1992) Carbohydr. Res. 236, 145–164.
[18] Tanaka, N., Nakada, H., Itoh, N., Mizuno, Y., Takanishi, M., Kawasaki, T., Tate, S., Inagaki, F. and Yamashina, I. (1992) J. Biochem. 112, 68–74.
[19] Tomiya, N., Awaya, J., Kurono, M., Hanzawa, H., Shimada, I., Arata, Y., Yoshida, T. and Takahashi, N. (1993) J. Biol. Chem. 268, 113–126.