Aust. J. Biol. Sci., 1978, **31**, 443–46

A Correction and Extension of the Acetylated Amino Terminal Sequence of Ovalbumin

E. O. P. Thompson and W. K. Fisher

School of Biochemistry, University of New South Wales,
P.O. Box 1, Kensington, N.S.W. 2033.

Abstract

The acetylpeptides derived from *S*-carboxymethylovalbumin by cyanogen bromide and chymotrypsin have been isolated and shown by enzyme digestion and the dansyl–Edman method to fit the sequence acetyl-Gly-Ser-Ile-Gly-Ala-Ala-Ser-Met-Glu-Phe. This corrects the order of the third and fourth residues in the five-residue sequence given by Narita and Ishii [*J. Biochem.* (*Tokyo*), 1962, **52**, 367–73].

The overlap of the *C*-terminal sequence of this extended sequence with the six-residue *N*-terminal sequence surrounding a half-cystine residue in ovalbumin gives the *N*-terminal sequence for ovalbumin as acetyl-Gly-Ser-Ile-Gly-Ala-Ala-Ser-Met-Glu-Phe-Cys-Phe-Asp-Val-Phe-Lys with residue 11 a cysteine residue.

Introduction

An acetylated amino terminal sequence for ovalbumin was determined by Narita and Ishii (1962) as *N*-acetyl-Gly-Ser-Gly-Ile-Ala-. The acetyl-Gly-Ser peptide was isolated from a pronase digest and its structure was confirmed by hydrazinolysis and by comparison with a synthetic sample. A peptide obtained by peptic digestion which contained an additional residue of glycine as well as isoleucine was given the sequence acetyl-Gly-Ser-Gly-Ile on the basis of amino acids released by carboxypeptidase digestion. Data from partial acid hydrolysis of a longer peptic peptide which was given the sequence acetyl-Gly-Ser-Gly-Ile-Ala was noted to be contrary to expectation in that little glycine was liberated, in contrast to large amounts of serine and alanine. Normally isoleucyl and valyl bonds are more stable to acid hydrolysis than are bonds involving other amino terminal residues. The *C*-terminal alanine was confirmed by hydrazinolysis.

In examination of acidic peptides obtained from performic acid-oxidized ovalbumin (Thompson and Fisher 1978) evidence was obtained for a longer acetylated peptide sequence containing methionine in chymotryptic digests. In experiments designed to isolate larger acetylated peptides and to place the additional amino acids in sequence it became clear that the acetylated amino terminal sequence has the sequence acetyl-Gly-Ser-Ile-Gly-Ala- with a reversal of the order of isoleucine and glycine given by Narita and Ishii (1962).

In the present paper an extension to the corrected terminal sequence was deduced from acetylated peptides formed by cyanogen bromide cleavage of *S*-carboxymethyl (SCM)-ovalbumin or chymotryptic digestion of SCM-ovalbumin. The extended sequences were determined from peptides liberated by other enzymes and sequenced by the dansyl–Edman procedure.

Materials and Methods

Ovalbumin was prepared and the thiol groups carboxymethylated as described by Smith (1968).

Cyanogen bromide cleavage of SCM-ovalbumin (35 mg) was effected at room temperature in 70% (v/v) formic acid (1 ml) with 35 mg CNBr. After freeze-drying the residue was extracted with water and applied to a column (5 by 2 cm) of sulfonated polystyrene (Bio-rad AG50 X2, 200–400 mesh) in the hydrogen form and eluted with water (20 ml). A sample was hydrolysed with 6 M HCl at 110°C and the hydrolysate subjected to amino acid analysis.

Enzyme Digestions

After digestion of the remaining peptide fraction with thermolysin as previously described (Thompson *et al.* 1971), the digest was fractionated as above on sulfonated polystyrene. The non-adsorbed acidic fraction was eluted with water. The bound peptides were eluted with 1 M NH_3. After freeze-drying, samples of each fraction were hydrolysed for amino acid analysis.

Chymotryptic digestion of SCM-ovalbumin (100 mg) was carried out in 1% (w/v) ammonium bicarbonate for 4 h using 1 mg enzyme. After acidification with acetic acid to approximately pH 3 the mixture was applied to a column of sulfonated polystyrene (12 by 2 cm) in the hydrogen form and eluted with water. The eluate was freeze-dried and oxidized with preformed performic acid reagent [30% (v/v) H_2O_2–formic acid, 1:9, 1 h at 25°C, 0·5 ml] at 0°C for 1 h. After freeze-drying a small sample was hydrolysed for amino acid analysis. Samples (1 μmol) of the remaining chymotryptic acidic peptide fraction were separately hydrolysed with pepsin (0·05 mg) in 5% (v/v) formic acid and subtilisin BPN′ (0·05 mg) in 1% (w/v) ammonium bicarbonate for 4 h. After freeze-drying, the digests were dissolved in pH 6·5 pyridine–acetic acid buffer.

Peptide Fractionation and Identification

The mixtures of peptides were fractionated by ionophoresis at pH 6·5 followed by chromatography in butanol–pyridine–acetic acid–water (14:10:3:12, v/v) as previously described (Thompson *et al.* 1971). After drying, peptides were revealed by 0·02% (w/v) ninhydrin. Each peptide zone was washed with acetone and a portion eluted with 6 M HCl for amino acid analysis. A portion was eluted wth 60% (v/v) pyridine and samples used for *N*-terminal sequence identification by the dansyl–Edman method (Hartley 1970).

Results

The acidic peptide fraction from the cyanogen bromide digest of SCM-ovalbumin gave the following amino acid composition: $Ser_{2.0}\,Hse_{1.1}\,Gly_{2.3}\,Ala_{1.7}\,Ile_{0.6}$. The absence of glutamic acid suggested that pyroglutamyl peptide contamination was not a problem. The composition corresponded with that expected from the sequence of Narita and Ishii (1962) with additional residues of alanine, serine and homoserine. Since the acetylated glycyl residue could not be detected by the dansyl method the fraction was digested with thermolysin which would be expected to split the peptide bond preceding the isoleucyl residue. The dansyl–Edman procedure then gave the sequence *Ile-Gly-Ala-Ala-Ser*. When fractionated into an acidic non-adsorbed fraction and an adsorbed fraction on sulfonated polystyrene followed by hydrolysis, amino acid composition studies of the two gave for the non-adsorbed acetylated fraction $Ser_{0.8}\,Gly_{1.0}$, and for the adsorbed fraction $Ser_{1.1}\,Gly_{1.2}\,Ala_{2.1}\,Ile_{1.0}\,Hse_{1.1}$, suggesting the sequence acetyl-Gly-Ser-Ile-Gly-Ala-Ala-Ser-Met in ovalbumin.

A chymotryptic digest of SCM-ovalbumin gave an acidic fraction which, after performic acid oxidation to give the stable methionine sulfone, had the amino acid composition $Mes_{1.0}\,Ser_{1.5}\,Glu_{0.9}\,Gly_{2.0}\,Ala_{1.7}\,Ile_{1.0}\,Phe_{0.7}$ and only traces of other amino acids. The specificity of chymotrypsin suggested a longer peptide with *C*-terminal phenylalanine.

When subjected to digestion with subtilisin BPN′ and pepsin the peptides shown below were obtained. They were identified from their amino acid compositions and *N*-terminal sequences, shown in *italic* fount.

Acetyl-Gly-Ser-Ile-Gly-Ala-Ala-Ser-Met-Glu-Phe	
Subtilisin (neutral)	Ala
Subtilisin (neutral)	*Ala-Ser-Met*
Subtilisin (acidic)	*Ser-Met-Glu*-Phe
Subtilisin (acidic)	*Met-Glu-Phe*
Pepsin (neutral)	*Ala-Ala-Ser*-Met + Phe
Pepsin (acidic)	*Gly-Ala*-Ala-Ser-Met-Glu-Phe
Pepsin (acidic)	*Ala-Ala-Ser*-Met-Glu

Discussion

The evidence presented reverses the order of the glycylisoleucyl sequence deduced by Narita and Ishii (1962). This change is supported by their data from partial acid hydrolysis experiments. The carboxypeptidase digestion results in their work can only be explained if their analysis by the dinitrophenyl method gave misleading quantitation.

The *N*-terminal sequence of ovalbumin is extended by the data presented to acetyl-Gly-Ser-Ile-Gly-Ala-Ala-Ser-Met-Glu-Phe-.

The sequences of amino acids surrounding half-cystine residues in ovalbumin (Thompson and Fisher 1978) include a sequence -Ala-Ala-Ser-Met-Glu-Phe-Cys-Phe-Asp-Val-Phe-Lys- which was detected as a [2-^{14}C]carboxymethylcysteine peptide in tryptic digests of labelled SCM-ovalbumin. The *N*-terminal sequence of six residues is identical to the *C*-terminal sequence in the acetyl sequence given above, so that the *N*-terminal sequence of ovalbumin can be deduced as

acetyl-Gly-Ser-Ile-Gly-Ala-Ala-Ser-Met-Glu-Phe-Cys-Phe-Asp-Val-Phe-Lys-.

The evidence previously given (Thompson and Fisher 1978) suggests that residue 11 is a cysteine residue in ovalbumin.

Acknowledgments

The authors wish to thank Mr R. G. Mann, Mrs D. D. Koureas and Miss P. A. Davidson for skilled technical assistance. The work was supported in part by the Australian Research Grants Committee.

References

Hartley, B. S. (1970). Strategy and tactics in protein chemistry. *Biochem. J.* **119**, 805–22.

Narita, K., and Ishii, J. (1962). *N*-terminal sequence in ovalbumin. *J. Biochem.* (*Tokyo*) **52**, 367–73.

Smith, M. B. (1968). The isolation of a large peptide from denatured plakalbumin. *Biochim. Biophys. Acta* **154**, 263–6.

Thompson, E. O. P., and Fisher, W. K. (1978). Amino acid sequences containing half-cystine residues in ovalbumin. *Aust. J. Biol. Sci.* **31**, 433–42.

Thompson, E. O. P., Sleigh, R. W., and Smith, M. B. (1971). The amino acid sequence of the large plakalbumin peptide and the *C*-terminal sequence of ovalbumin. *Aust. J. Biol. Sci.* **24**, 525–34.

Manuscript received 23 March 1978

Note Added in Proof

Since this work was submitted the paper 'Ovalbumin: a secreted protein without a transient hydrophobic leader sequence' by R. D. Palmiter, J. Gagnon and K. A. Walsh (*Proc. Natl Acad. Sci. U.S.A.*, 1978, **75**, 94–8) has come to our notice. Our results, which confirm their data, have been obtained by a different sequencing strategy.