Aust. J. Biol. Sci., 1982, **35**, 263–9

Two Low-molecular-weight Apoproteins (Apovitellenins I and II) from a Lipoprotein of Goose's Egg Yolk: A Comparison with Related Species

A. S. Inglis,[A] *P. M. Strike*[A] *and R. W. Burley*[B]

[A] Division of Protein Chemistry, CSIRO, Parkville, Vic. 3052.
[B] Division of Food Research, CSIRO, North Ryde, N.S.W. 2113.

Abstract

As part of a comparative study of egg yolk from different avian species, the major lipoprotein and its mixed apoproteins from the egg yolk of the chinese goose (*Anser cygnoides*) have been prepared. From the apoprotein mixture, two new proteins, of molecular weight approximately 10000 and 22000 according to gel electrophoresis in detergent, have been isolated by gel-filtration chromatography in urea. The protein of lower molecular weight corresponds in amino acid sequence to apovitellenin I, a protein previously isolated from other avian species. As a comparison with other members of the same avian family (Anatidae), the amino acid sequence of apovitellenin I from the pekin duck (*Anas platyrhynchos*) was re-investigated and that of the muscovy duck (*Cairina moschata*) investigated. These were found to be identical to the sequence of goose's apovitellenin I. The second new protein is similar in composition, molecular weight, and solubility to apovitellenin II, a protein present in small amount in hen's egg yolk. A protein corresponding to apovitellenin II could not, however, be detected in the egg yolk of either species of duck.

Extra keywords: protein evolution.

Introduction

Apoproteins of low molecular weight have previously been isolated from the major (i.e. low density, high lipid) lipoprotein (density 0·95 g/ml) of yolk from eggs of the emu, hen, duck, and turkey; and the amino acid sequence of one apoprotein (apovitellenin I) has been determined for each of these species (Inglis and Burley 1977; Inglis *et al.* 1979*a*; Dugaiczyk *et al.* 1981). The study of apovitellenin I is of current interest for several reasons: (1) its synthesis is readily inducible by hormones, even in male birds, so it is of use to those concerned with gene expression in birds (Dugaiczyk *et al.* 1981; Wieringa *et al.* 1981); (2) it may provide an example of the way in which proteins are transferred from the blood to the yolk—a process that is still largely unknown although it is now clear that apovitellenin I is transported to the yolk with no change in amino acid sequence (Dugaiczyk *et al.* 1981); (3) the amino acid sequence of this protein shows considerable interspecies variation, involving about 20% of the residues throughout the sequence. This large variation implies that apovitellenin I may be useful as an indicator of evolutionary relationships. We were primarily interested in (3); in particular, in the variations between closely related species because of our previous observation that hen's and turkey's apovitellenin I differ in 12 positions out of 82 (Inglis *et al.* 1979*a*). Accordingly, we have prepared and sequenced apovitellenin I from the goose because of its close relationship to the duck, a species studied previously. We found no difference between these species

after making a slight correction to the previously reported sequence for the duck. During this work, a second protein of low molecular weight was isolated in the apoprotein mixture from the goose. This protein was not detected in the duck's lipoprotein but it is similar to apovitellenin II previously isolated from the lipoprotein of hen's egg yolk and also from the livetin fraction, i.e. the proteins of yolk that are soluble in aqueous solutions (Burley 1975; Bengtsson *et al.* 1977; Burley and Vadehra 1979).

A preliminary report of the sequence of goose's apovitellenin I has already been given with details of one aspect of its determination, namely cleavage with dilute acid (Inglis *et al.* 1980).

Materials and Methods

Proteins

These were isolated from the yolk low-density lipoprotein of eggs of the brown chinese goose (*Anser cygnoides*), the pekin duck (*Anas platyrhynchos*), and the muscovy duck (*Cairina moschata*) following procedures described for hen's egg yolk (Burley 1978). In brief, the lipoprotein was isolated by centrifuging and the lipid-free protein dissolved in acid urea for gel-filtration chromatography on columns of Sephadex G100 and G75.

Amino Acid Analyses

A Beckman 120C analyser was used after the protein had been hydrolysed with hydrochloric acid as described for the hen's apoproteins (Burley 1975), a correction being applied for loss of some residues during hydrolysis. In addition, methanesulfonic acid (Simpson *et al.* 1976) and hydriodic acid (Inglis *et al.* 1979*a*) were used for goose's apovitellenin I to confirm the values for tryptophan and valine-isoleucine respectively. Values for amide were confirmed by the method of Inglis *et al.* (1974).

Amino Acid Sequence of Goose's Apovitellenin I

The sequence was determined for the first 40 residues of the protein according to the procedure of Inglis *et al.* (1979*a*). The remaining sequence was derived from fragments isolated chromatographically after (i) dilute acid cleavage (Inglis *et al.* 1979*b*, 1980) and (ii) cyanogen bromide cleavage, as described for hen's apovitellenin I (Dugaiczyk *et al.* 1981).

Electrophoretic Determination of the Molecular Weights of the Goose's Apoproteins

The procedure of Weber and Osborn (1969) was used with the minor difference that the electrode buffer was Tris hydrochloride (0·03 M), sodium dodecyl sulfate (0·1% w/v), pH 7·6. Approximate molecular weights were determined from the graph of the electrophoretic mobility against log molecular weight for standard proteins of low molecular weight including: emu's apovitellenin I (9740), cytochrome *c* (12400), trypsinogen (24000), ovalbumin (43000).

Results and Discussion

Fig. 1(*a*) shows the gel-filtration chromatography in acidic urea of the total apoprotein mixture from the major lipoprotein of goose's egg yolk. The large low-molecular-weight peak (A) contained 68% of the weight of protein eluted from the column. The relative size of this peak decreased with the age of the eggs. After they had been stored for 3 months at 2°C peak A contained only 25% of the total. The reasons for this behaviour have not been determined. Similar behaviour was observed with turkey's eggs (Inglis *et al.* 1979*a*).

The protein in peak A (Fig. 1*a*) was isolated by precipitation with trichloroacetic acid and rechromatographed on another column (Fig. 1*b*), according to which three main constituents were present (I, II, X). The constituent of lowest molecular weight (X) has not been investigated. Possibly this consisted of lipase cofactors of low molecular weight analogous to those isolated by Bengtsson *et al.* (1977) from hen's egg yolk.

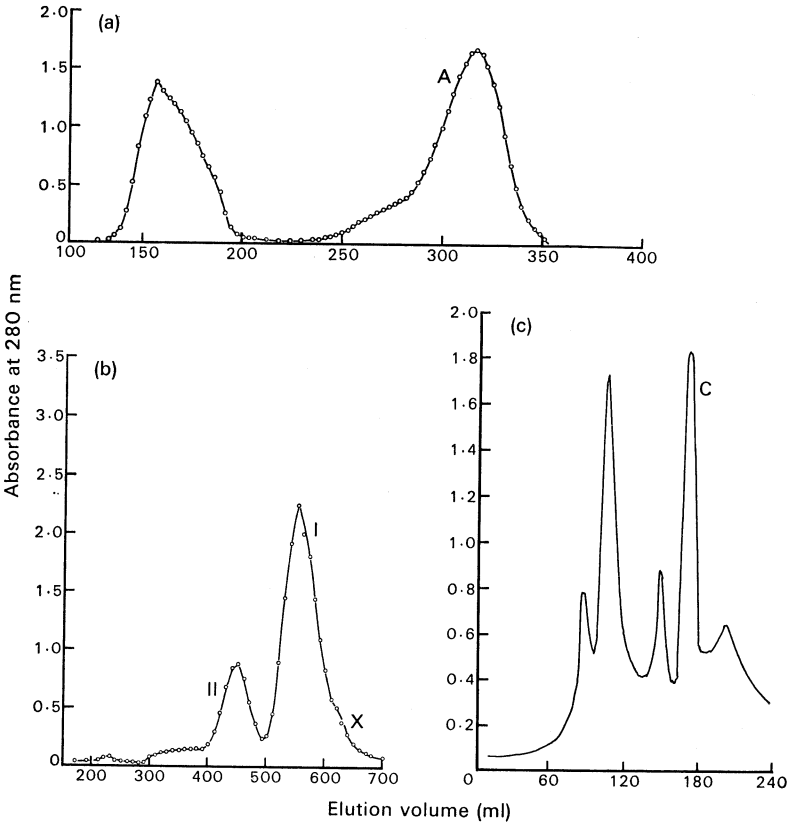

Fig. 1. (*a*) Gel-filtration chromatography of the total apoprotein mixture from goose's egg yolk lipoprotein on a column (3·3 by 44 cm) of Sephadex G100 (fine) in 6 M urea, pH 3·3, 20°C, eluting at 20 ml/h. Peak A contained two proteins of low molecular weight. The other peak contained a complicated mixture of proteins of high molecular weight. (*b*) Gel-filtration chromatography of proteins isolated from peak A on a column (120 by 2·7 cm) of Sephadex G75 (fine) in the solvent used in (*a*), 20°C. Peaks I and II were later identified as apovitellenins I and II. 'X' was not identified. (*c*) Gel-filtration chromatography on a column (110 by 1·2 cm) of Sephadex G50 (fine) of fragments of goose's apovitellenin formed by cyanogen bromide treatment. The column was eluted at 6 ml/h with 75% (v/v) formic acid and the sample was applied in 98% (v/v) formic acid. Peak 'C' represents the *C*-terminal peptide.

Fractions I and II (Fig. 1*b*) were rechromatographed on the same column to give proteins that were homogeneous according to gel electrophoresis in detergent (see Materials and Methods). Their amino acid compositions and approximate molecular weights are given in Table 1.

According to Table 1, fraction II resembles hen's apovitellenin II, the amino acid composition of which is included for comparison. It is therefore probable that it is the second of the apovitellenin II group to be isolated. The points of similarity include: molecular weight, absence of methionine, high levels of some residues (Gly, Leu, Asp) and low levels of others (Ile, Lys). Goose's apovitellenin II was soluble in salt solutions, a characteristic of hen's apovitellenin II but not of other avian apoproteins examined so far. A notable difference is that goose's apovitellenin II lacks histidine whereas the hen's protein contains histidine but has no tyrosine. It is also notable that this protein could not be detected under the same conditions in the lipoproteins of either of the duck species studied.

Table 1. Amino acid compositions of goose's apovitellenins I and II and of fragments used for sequence analyses

Values are expressed as moles per mole of protein with values found by sequence analysis given in brackets. Approximate molecular weights from gel electrophoresis in detergent are also shown

Amino acid residue	Apovitellenin I		Apovitell-enin I	Apovitell-enin II	Apovitell-enin II from hen's egg yolk[A]
	CNBr fragment, residues 69–82	Dilute HCl fragment residues 40–82			
Lys	2·5 (3)	3·5 (4)	5·7 (6)	5·4	2
His	0	0	0	0	1
Arg	0	2·0 (2)	6·1 (6)	21·3	14
Trp	0·8 (1)	0·8 (1)	1·8 (2)	n.d.	7
Asp	1·1 (1)	1·2 (1)	6·3 (6)	24·5	20
Thr	1·2 (1)	5·5 (6)	6·6 (7)	13·9	12
Ser	0	1·0 (1)	3·0 (3)	16·1	13
Glu	1·2 (1)	5·0 (5)	7·9 (8)	18·8	20
Pro	0	1·1 (1)	3·4 (3)	21·7	10
Gly	1·2 (1)	2·1 (2)	3·9 (3)	31·1	26
Ala	0·2	3·2 (3)	7·2 (7)	10·3	10
½Cys	0	0	0·3	8·0	9
Val	1·0 (1)	3·4 (4)	6·9 (7)	11·2	12
Met	0	1·0 (1)	1·7 (2)	0	0
Ile	0·9 (1)	2·1 (2)	5·7 (6)	2·3	4
Leu	2·0 (2)	7·0 (7)	9·0 (9)	20·9	17
Tyr	1·9 (2)	2·1 (2)	3·0 (3)	1·6	0
Phe	0·1	1·1 (1)	3·9 (4)	8·4	8
Amide	4·9 (5)				18
Mol. wt			10000	22000[B]	20000

[A] From Burley (1975).
[B] Minimum molecular weight from amino acid composition 22900.

From its molecular weight and amino acid composition (Table 1), especially the absence of histidine, fraction I is clearly a member of the apovitellenin I group (see Inglis *et al.* 1979*a*). Its amino acid sequence is given later in Table 2. This was established from a sequenator analysis of three molecules; the intact protein, a dilute HCl cleavage fragment, and a cyanogen bromide fragment—the *C*-terminal sequence (Fig. 1*c*). All amino acid derivatives from the sequenator were initially identified by thin-layer chromatography and, where necessary, confirmed by high-pressure liquid chromatography (Dugaiczyk *et al.* 1981). In addition, residues

40–82 were confirmed by amino acid analysis of the hydrolysed derivatives. Except for a high glycine value and a slightly low threonine in the intact protein, the analyses of the three molecules correlate well with the sequence data (see Table 1).

Table 2. Amino acid sequence of goose's apovitellenin I

Residues 1–40 of the sequence were determined by the sequenator on the whole protein, residues 41–70 on peptides isolated after dilute acid hydrolysis and residues 69–82 on peptides isolated after cyanogen bromide cleavage

Lys	Ser	Ile	Phe	Glu	Arg	Asp	Arg	Arg	Asp^{10}
Trp	Leu	Val	Ile	Pro	Asp	Ala	Ile	Ala	Ala^{20}
Tyr	Ile	Tyr	Glu	Thr	Val	Asn	Lys	Met	Ser^{30}
Pro	Arg	Val	Gly	Gln	Phe	Leu	Ala	Asp	Ala^{40}
Ala	Gln	Thr	Pro	Val	Val	Val	Gly	Thr	Arg^{50}
Thr	Phe	Leu	Ile	Arg	Glu	Thr	Thr	Lys	Leu^{60}
Ser	Leu	Leu	Ala	Glu	Gln	Leu	Met	Glu	Lys^{70}
Ile	Lys	Asn	Leu	Trp	Tyr	Thr	Lys	Val	Leu^{80}
Gly	Tyr^{82}								

The sequence for goose's apovitellenin I (Table 2) differs from that reported for the pekin duck (*Anas platyrhynchos*, tribe Anatini) (Inglis and Burley 1977) in only two residues: the duck has a serine instead of threonine at residue 58 and a threonine instead of serine at residue 61. In view of this similarity, and the labile nature of the derivatives of these residues, this region of the molecule was redetermined on a fresh preparation of apovitellenin I from the pekin duck. Furthermore the sequence of this region was also determined for muscovy duck (*Cairina moschata*, tribe Cainini). Both sequences were identical to that of goose's apovitellenin I (Table 2). Apparently, in the earlier work, after 58 degradation cycles on the intact protein, the method was not sufficiently sensitive to distinguish serine and threonine. This problem has now been overcome by using a fragment, obtained by dilute acid hydrolysis, starting

Table 3. Amino acid sequences of apovitellenin I from egg yolk of various avian species

One-letter code follows Dayoff (1972). Variable regions are set in *italic* type

Emu[A]	K S I *F* *E* R *D* *N* R R D W L V I P D A *V* A A Y *V* Y E *T* V N
Goose[B]	K S I *F* *E* R *D* — R R D W L V I P D A *I* A A Y *I* Y E *T* V N
Duck[B]	K S I *F* *E* R *D* — R R D W L V I P D A *I* A A Y *I* Y E *T* V N
Hen[C]	K S I *I* *D* R *E* — R R D W L V I P D A *A* A A Y *I* Y E *A* V N
Turkey[D]	K S I *F* *E* R *D* — R R D W L V I P D A *V* A A Y *I* Y E *A* V N
Emu	K *M* *F* P *K* *V* G Q F L *A* D *A* *A* Q *I* *P* V *I* *V* G *T* R *N* F L I *R*
Goose	K *M* *S* P *R* *V* G Q F L *A* D *A* *A* Q *T* *P* V *V* *V* G *T* R *T* F L I *R*
Duck	K *M* *S* P *R* *V* G Q F L *A* D *A* *A* Q *T* *P* V *V* *V* G *T* R *T* F L I *R*
Hen	K *V* *S* P *R* *A* G Q F L *L* D *V* *S* Q *T* *T* V *V* *S* G *I* R *N* F L I *N*
Turkey	K *M* *S* P *R* *A* G Q F L *V* D *I* *S* Q *T* *T* *V* *V* *S* G *T* R *N* F L I *R*
Emu	E T *S* *K* L *S* I L A E Q *M* M E K *V* K *T* L *W* *N* T K V *L* G Y *Y*
Goose	E T *T* *K* L *S* L L A E Q *L* M E K *I* K *N* L *W* *Y* T K V *L* G Y —
Duck	E T *T* *K* L *S* L L A E Q *L* M E K *I* K *N* L *W* *Y* T K V *L* G Y —
Hen	E T *A* *R* L *T* K L A E Q *L* M E K *I* K *N* L *C* *Y* T K V *L* G Y —
Turkey	E T *A* *R* L *T* L L A E Q *L* M E K *I* K *N* L *N* *Y* T K V *Q* G Y —

[A] Dopheide and Inglis (1974). [B] This work. [C] Dugaiczyk *et al.* (1981). [D] Inglis *et al.* (1979*a*).

at residue 40, and by confirming the identification of residues 58–61 by high-pressure liquid chromatography.

The identity of the apovitellenin I sequences from different duck and goose species is in agreement with their classification in the same family (Anatidae). The fossil record for birds is incomplete but it is clear that the ancestors of the goose and duck diverged relatively recently, possibly in the Oligocene (Howard 1964). All available apovitellenin I sequences are given in Table 3, from which it is apparent that more sequences would be desirable; nevertheless the duck and goose sequences emphasize the anomalous difference between the hen and turkey. The hen and turkey are also classed in one family (Phasianidae) and are also closely related by several criteria including the pattern of their egg-white proteins (Silbey and Ahlquist 1972). We do not at present have an explanation for the difference (Inglis *et al.* 1979*a*).

Table 4. Comparison of amino acid compositions of apovitellenin I from five avian species

Values are expressed as moles per mole of protein

Amino acid side chain	Goose	Duck	Turkey	Hen	Emu
Basic	12	12	12	12	12
Acidic	9	9	9	9	9
Amide	5	5	7	7	7
Hydrocarbon[A]	37	37	35	37	38
Hydroxyl[B]	10	10	10	10	8
Aromatic	9	9	9	7	10
Total	82	82	82	82	84

[A] Includes Pro and Met, not Phe.
[B] Includes Cys (–SH instead of –OH), not Tyr.

In considering the possible role of apovitellenin I in the lipoprotein, sequence data on other apoproteins, such as apovitellenin II, would be of interest. The relatively large number of apparent mutations in apovitellenin I suggests that it does not need a unique sequence to function properly. Nevertheless, the changes are largely confined to like amino acid residues. As shown in Table 4, there is a marked similarity in composition when amino acids with similar side chains are grouped together. Furthermore, when the amino acid sequences are aligned as in Table 3, many invariable positions are apparent, especially at the *N*- and *C*-terminal positions.

Acknowledgments

We wish to thank Mrs Janet Mediati, Warriewood, N.S.W., for giving us a supply of chinese goose's eggs, and Dr C. M. Roxburgh for amino acid analyses.

References

Bengtsson, G., Marklund, S., and Olivecrona, R. (1977). Protein components of very low density lipoprotein from hen's egg yolk. *Eur. J. Biochem.* **79**, 211–33.

Burley, R. W. (1975). Studies on the apoproteins of the major lipoprotein of the yolk of hen's eggs. I. Isolation and properties of the low-molecular-weight apoproteins. *Aust. J. Biol. Sci.* **28**, 12–32.

Burley, R. W. (1978). Studies on the apoproteins of the major lipoprotein of the yolk of hen's eggs. III. Influence of salt concentration during isolation on the amount and composition of the apoproteins. *Aust. J. Biol. Sci.* **31**, 587–92.

Burley, R. W., and Vadehra, D. V. (1979). Chromatographic separation of the soluble proteins of hen's egg yolk: an analytical and preparative study. *Anal. Biochem.* **94**, 53–9.

Dayhoff, M. O. (1972). Protein data introduction. In 'Atlas of Protein Sequence and Structure'. Vol. 5. (Ed. M. O. Dayoff.) p. D-2. (National Biomedical Research Foundation: Silver Springs, Md.)

Dopheide, T. A. A., and Inglis, A. S. (1974). Amino-acid sequence of a protein (apovitellenin I) from the low-density lipoprotein of emu's egg yolk. *Aust. J. Biol. Sci.* **27**, 15–21.

Dugaiczyk, A., Inglis, A. S., Strike, P. M., Burley, R. W., Beattie, W. G., and Chan, L. (1981). Comparison of the nucleotide sequence of cloned DNA coding for an apolipoprotein (apo VLDL-II) from avian blood and the amino-acid sequence of an egg-yolk protein (apovitellenin I): equivalence of the two sequences. *Gene* **14**, 175–82.

Howard, H. (1964). Fossil anseriformes. In 'The Waterfowl of the World'. (Ed. J. Delacour.) Vol. 4. pp. 233–6. (Country Life Ltd: London.)

Inglis, A. S., Roxburgh, C. M., and Takayanagi, H. (1974). Analysis for amide nitrogen in proteins. *Anal. Biochem.* **61**, 25–31.

Inglis, A. S., and Burley, R. W. (1977). Determination of the amino-acid sequence of apovitellenin I from duck's egg yolk using an improved sequenator procedure: a comparison with other avian species. *FEBS Letters* **73**, 33–7.

Inglis, A. S., Strike, P. M., Osborne, W. C., and Burley, R. W. (1979*a*). Sequenator determination of the amino-acid sequence of apovitellenin I from turkey's egg yolk. *FEBS Letters* **97**, 179–82.

Inglis, A. S., McKern, N. M., and Strike, P. M. (1979*b*). Specific cleavage of proteins at aspartic acid residues. *Proc. Aust. Biochem. Soc.* **12**, 12.

Inglis, A., McKern, N., Roxburgh, C., and Strike, P. (1980). Cleavage of proteins at aspartic acid residues using dilute acid. In 'Methods in Peptide and Protein Sequence Analysis'. (Ed. C. Birr.) pp. 329–43. (Elsevier/North Holland Biomedical Press.)

Sibley, C. G., and Ahlquist, J. E. (1972). A comparative study of the egg white proteins of non-passerine birds. Bull. Peabody Museum of Natural History, Yale, Connecticutt, U.S.A. No. 39, pp. 93–113.

Simpson, R. J., Neuberger, M. R., and Liu, T. Y. (1976). Complete amino-acid analysis of proteins from a single hydrolysate. *J. Biol. Chem.* **251**, 1936–40.

Weber, K., and Osborn, M. (1969). The reliability of molecular weight determinations by dodecyl-sulphate polyacrylamide-gel electrophoresis. *J. Biol. Chem.* **244**, 4406–12.

Wieringa, B., Ab, G., and Gruber, M. (1981). Nucleotide sequence of the very low density lipoprotein II mRNA from chicken. *Nucleic Acids Res.* **9**, 489–501.

Manuscript received 16 February 1981, accepted 21 April 1982