Biochimica et Biophysica Acta, 700 (1982) 42–51
Elsevier Biomedical Press

BBA 31011

SECONDARY STRUCTURE PREDICTIONS FOR SILKMOTH CHORION PROTEINS

S.J. HAMODRAKAS [a], C.W. JONES [b,*] and F.C. KAFATOS [a,b]

[a] *Department of Biology, University of Athens, Panepistimiopolis, Athens 621 (Greece) and* [b] *Department of Cellular and Developmental Biology, The Biological Laboratories, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138 (U.S.A.)*

(Received May 12th, 1981)

Key words: Secondary structure prediction; Chorion protein; (Silkmoth)

The complete primary structure of many silkmoth chorion proteins is now known. These proteins are products of two major multigene families, A and B. Here we predict the secondary structure of representative A and B chorion proteins, using six different predictive methods. In both families of proteins β-sheet structure predominates. The proteins can be divided into a number of distinct regions or domains, according to the degree of evolutionary constancy in sequence, the amino acid composition and secondary structure features. Both families are characterized by a central 'core' region which is highly structured and enriched in valine and alanine. 'Arms' are more variable, presumably reflecting protein-specific functions. Cysteines are found predominantly near the extremes of the molecules. β-turns are predicted frequently, and may often connect short anti-parallel β-sheet strands.

Introduction

The silkmoth eggshell or chorion has been studied extensively in recent years (reviewed in Ref. 1). It is a lamellar structure consisting of helicoidally arranged fibrils [1,2]. Two-dimensional gel electrophoresis has revealed a surprising biochemical complexity: as many as 186 components have been resolved from the chorion of an individual *Antheraea polyphemus* moth [3]. Most of the dry mass (approx. 88%) is accounted for by two molecular weight classes of proteins. A (approx. M_r 9000 to 12000) and B (M_r 12000 to 14000). Sequencing studies have shown that members of a class are highly similar to each other, but are encoded by distinct DNA sequences, since they differ by amino acid replacements, insertions or deletions: thus they correspond to a family of evolutionarily related genes (multigene family; [4,5,6,7]). The A and B families have many compositional and limited sequence features in common; it is not certain as yet whether these similarities represent ancient homology or convergent evolution.

Considerable information on chorion protein sequences is currently available, derived both from direct protein sequencing and from sequencing of recombinant DNA clones corresponding to individual proteins. We now wish to consider the implications of these sequences, both for the structure and morphogenesis of the chorion and for the evolution of the multigene families. As an initial step, we have investigated the secondary structure of predominant chorion components. In this paper we report theoretical predictions of secondary structure, based on complete sequences of both A and B components; in the future we will report the results of experimental analysis of protein secondary structure in the chorion.

* Present address: Department of Biochemistry, Stanford University School of Medicine, Stanford, CA 94305, U.S.A.

Methods

Computer programs making use of the predictive method of Chou and Fasman [8,9], Burgess et al. [10] and Lim [11,12] were kindly supplied by Dr. J.A. Lenstra and were modified for use in the CDC 6000 computer of the Greek Ministry of Agriculture. The algorithm of Chou and Fasman [8,9] was modified as described by J.A. Lenstra [13,14]. The conformational parameters used were those reported by Chou and Fasman [8,9], based on a set of 15 proteins with known conformation. The program written for the method of Burgess et al. [10] utilizes their constants and simultaneously predicts helix, extended structure and bends according to their nonamer model. The algorithm of Lim [11,12] is based on the relative positions of the hydrophobic and large hydrophilic residues along a sequence.

The prediction schemes of Garnier et al. [15] and Nagano [16,17] were kindly supplied by Drs. Garnier and Nagano as computer programs. These were also modified for use in the same computer mentioned above. The method of Garnier et al. [15] unambiguously assigns one of four conformational states to each residue in an amino acid sequence, and is based on information statistics [18]. Their directional method was used by choosing run constants of one and decision constants of zero. The directional information measures were taken from Garnier et al. [15]. For the statistical method of Nagano, the latest set of threshold parameters and coefficients were employed [16,17].

The method of Dufton and Hider [19], which is a modification of the Chou and Fasman method, was computerized and also used in the predictions. Conformational parameters for α-helix and β-sheet were taken from Chou and Fasman [8,9] and for β-turn from Chou and Fasman [20].

In general, rather consistent predictions were obtained with most methods (see Figs. 1 and 2). As might be expected, the most deviant results were obtained by the method of Lim [11,12], which is a stereo-chemical method developed for water-soluble globular proteins.

Chorion cDNA and genomic DNA sequences were determined and converted to protein sequences as described elsewhere [21,22]. They are presented according to the one-letter code (see legend to Fig. 3).

Results

Seven complete sequences from each family were analyzed by computer programs for secondary structure predictions, as explained in Methods. The A sequences were pc609, pc18, 18b, 18c, pc292, 292a, 292b, and the B sequences were pc408, pc10, 10a, 10b, pc401, 401a, 401b [6,7,21,22]. By convention, 'pc' numbers refer to sequences derived from cDNA clones, and plain numbers to sequences derived from genomic DNA clones. Identical numbers refer to very similar (but non-identical) sequences, corresponding to different copies of a gene; chorion genes are repetitive as well as multigenic [22,23]. Since no significant differences were observed in predicted secondary structure, we present below the data corresponding to only one 'copy' of each type of gene.

Figs. 1 and 2 present detailed results for the A and B families, respectively. For each protein, individual predictions of α-helix (α), β-plated sheet (β) and β-turns (T) were made, according to the methods [8–12,15–17,19] of Nagano (N), Garnier et al. (G), Burgess et al. (B), Chou and Fasman (F), Lim (L; α and β only) and Dufton and Hider (D); predicted structures are indicated by corresponding horizontal lines in Figs. 1 and 2. Joint prediction histograms (JP) were then constructed, since they are more dependable than individual prediction schemes [24,25]. A structure predicted by three or more methods (out of six, or out of five in the case of turns) was considered probable and is shaded in Figs. 1 and 2.

To make the plots most comparable, the sequences within each family were aligned to maximize sequence homologies. Necessary gaps introduced in the aligned sequences are indicated by dots (one per missing residue). Thus corresponding parts of proteins in the same family can be identified by nominal (aligned) residue position. The analyzed proteins of the A family have 120 nominal positions (99 to 113 actual residues), and B proteins have 160 nominal positions (142 to 153 actual residues).

The joint histograms of the A family are summarized in Fig. 3. In addition to the probable structures which are indicated by shaded bars (3 or more predictions; solid for β, hatched for α, stippled for turns), we indicate by open bars corre-

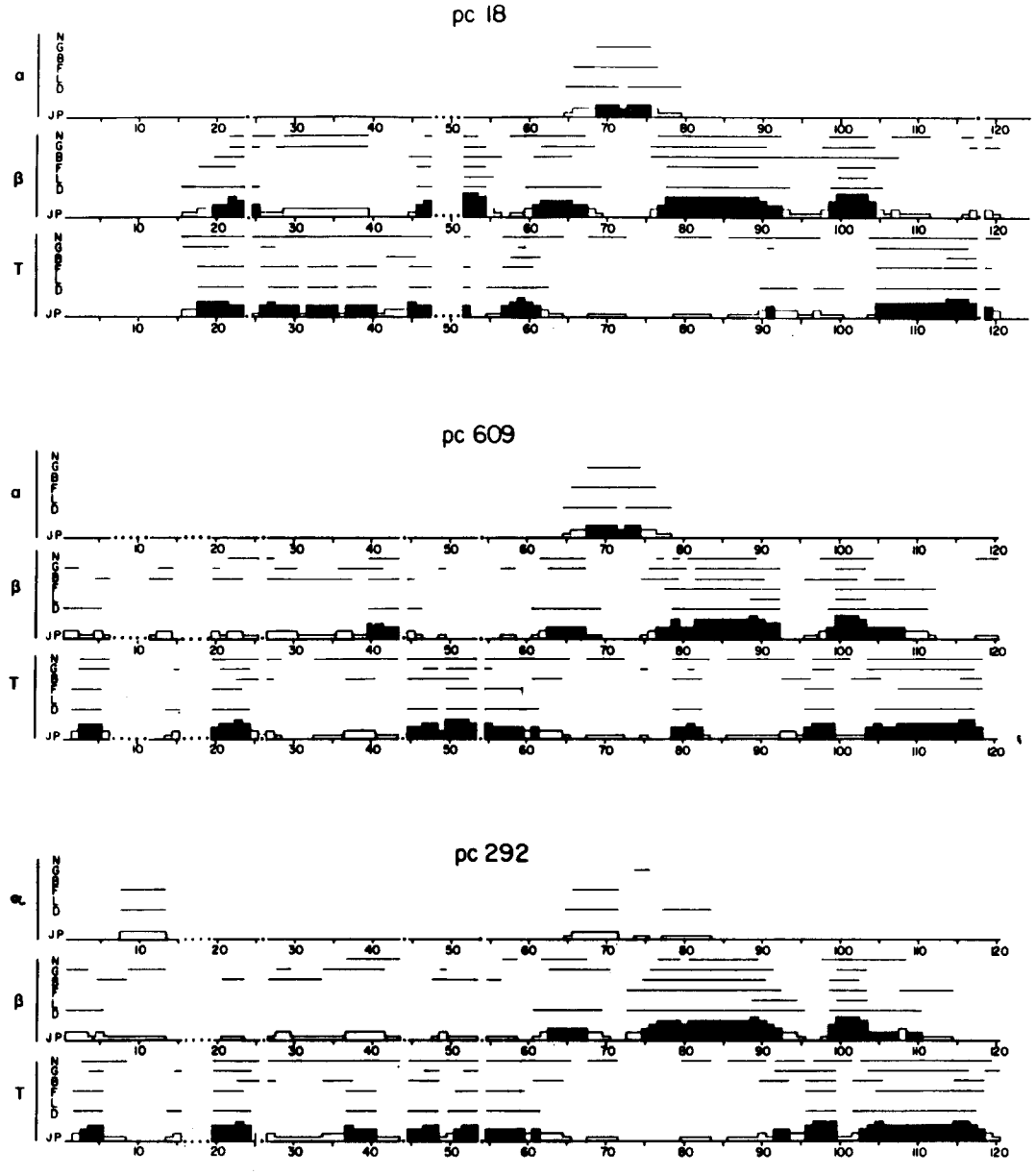

Fig. 1. Secondary structure predictions for three chorion proteins of the A family, encoded by clones pc18, pc609 and pc292. Individual predictions for α-helix (α), β-sheet (β) or β-turn (T) are shown by horizontal lines, as derived according to Nagano, (N, [16,17]); Garnier et al., (G, [15]); Burgess et al. (B, [10]); Chou and Fasman, (F, [8,9]); Lim (L, [11,12]); and Dufton and Hider (D, [19]). Joint prediction histograms (JP), constructed by tallying the individual predictions, are also shown. Sequences are numbered from amino to carboxyl-terminus after alignment (see Fig. 3); dots indicate gaps necessary for alignment. The most probable structures, predicted by three or more methods, are shaded.

sponding structures which are somewhat less probable (prediction by two methods each). The diagrams are accompanied by the actual protein sequences, presented in the one-letter code. To identify evolutionarily conservative regions of the sequences, blocks of two or more residues which are identical in all seven analyzed sequences of the family are boxed, while individual invariant re-

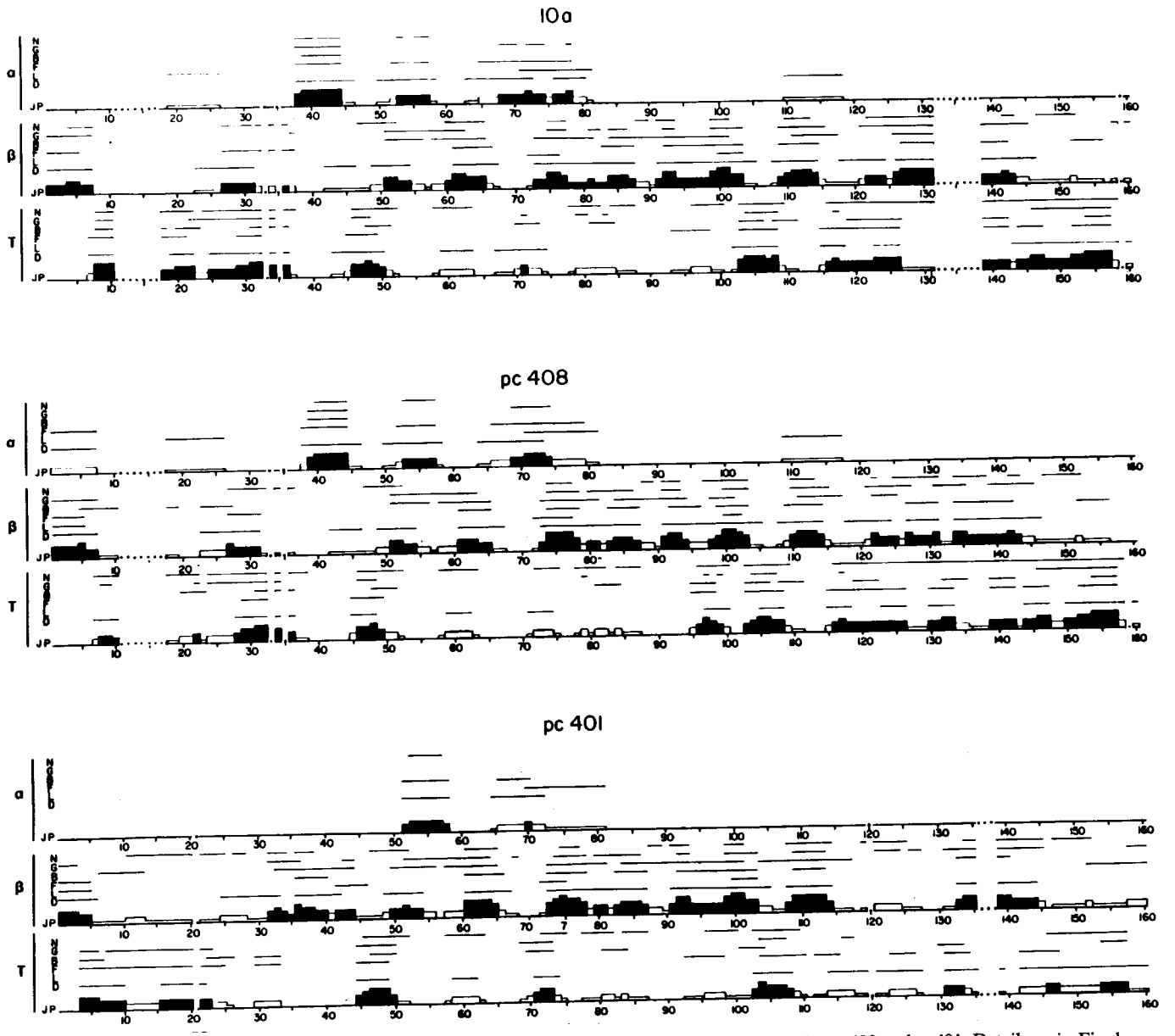

Fig. 2. Secondary structure predictions for three chorion proteins of the B family, encoded by clones 10a, pc408 and pc401. Details as in Fig. 1.

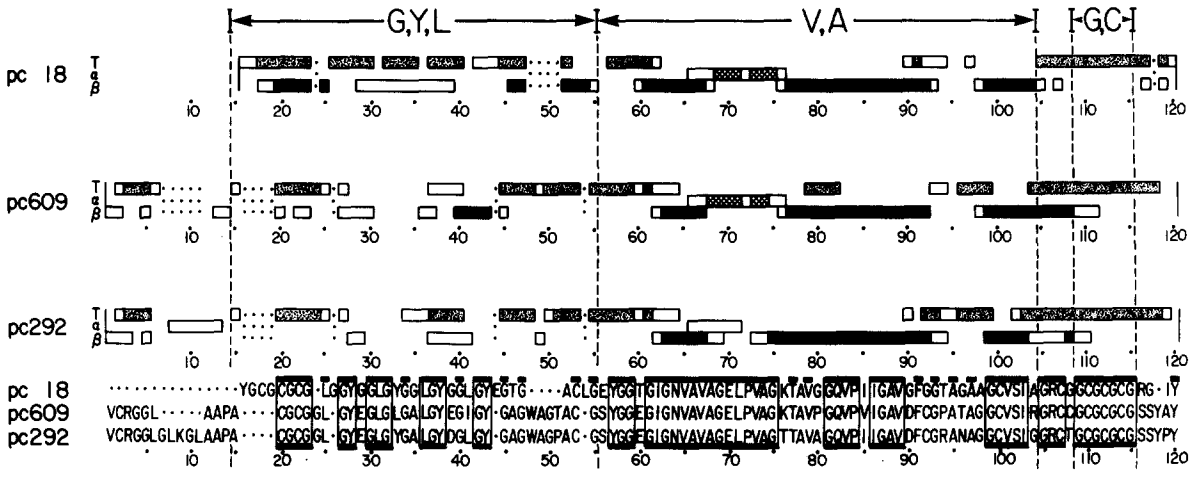

Fig. 3. Summary of joint prediction histograms for secondary structure in proteins of the A family (see Fig. 1). Structures predicted by two methods are shown as open rectangles, while the most probable structures predicted by three or more methods are shown solid (β-sheet), hatched (α-helix) or stippled (β-turn, T). The amino- and carboxyl-termini are indicated by short vertical lines, and regions enriched in certain amino acids are outlined by dashed lines and named accordingly (G, Y, L; V, A; G, C). Sequences are aligned and numbered, and necessary gaps are indicated by dots. The actual sequences are presented at the bottom, according to the IUPAC-IUB one-letter code, as follows: A=Ala, C=Cys, D=Asp, E=Glu, F=Phe, G=Gly, I=Ile, K=Lys, L=Leu, M=Met, N=Asn, P=Pro, Q=Gln, R=Arg, S=Ser, T=Thr, Y=Tyr, V=Val, W=Trp. Blocks outline two or more residues which are invariant in these and four additional related sequences (18b, 18c, 292a, 292b; [21,22]); individual invariant residues are overlined.

sidues are overlined. The sequences and diagrams are divided into regions enriched in certain amino acids (G, Y, L; V, A; G, C) as explained in the Discussion.

Fig. 4 includes the summarized joint prediction histograms and actual sequences for the B family; it is completely analogous to Fig. 3.

Discussion

General features of A and B chorion proteins

The computer predictions clearly indicate that both A and B chorion proteins (two families which jointly account for approximately 88% of the chorion dry weight) have a considerable proportion of

TABLE I

AMINO ACID COMPOSITIONS OF CHORION PROTEINS AND SELECTED REGIONS

Compositions are shown as molar percent, relative to all amino acids in the protein or in the region. Regions are defined as in Figs. 3 and 4, according to the amino acids in which they are relatively enriched (bold type). The regions near the amino terminus (G, Y, L in the A family; G in the B family) are also relatively enriched in cysteine, but for simplicity this feature is not used in the designations. Data are derived from the sequences shown in Figs. 3 and 4.

Protein and Region	Amino acid					
	G	Y	L	V	A	C
A family, total	33.4	6.3	6.3	7.5	12.2	9.4
Region G, Y, L	**45.6**	**11.7**	**12.6**	–	9.7	9.7
Region V, A	25.2	2.0	2.0	**15.0**	16.3	3.4
Region G, C	**57.1**	–	–	–	–	**42.9**
B family, total	32.5	6.7	9.0	7.0	11.2	5.8
Region G	**43.9**	5.3	5.3	–	5.3	12.3
Region V, A	23.0	3.5	6.9	**14.4**	14.2	1.7
Region C, Y	**44.3**	**15.3**	9.9	2.3	6.9	2.3
Region G, C	**57.1**	–	–	–	–	**42.9**

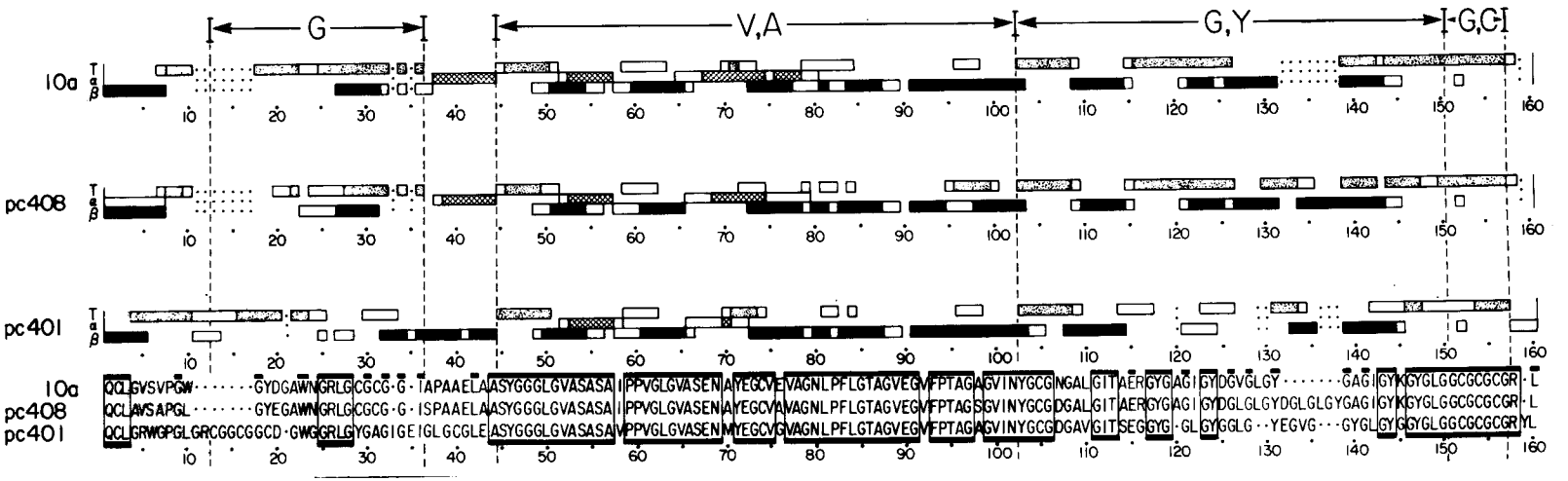

Fig. 4. Summary of joint prediction histograms for secondary structure in proteins of the B family (see Fig. 2). Details as in Fig. 3. Blocks and overlinings refer to four additional sequences (pc10, 10b, 401a, 401b; [21,22]).

β-pleated sheets, but only a minor amount of α-helix. These predictions were verified by laser Raman spectroscopic studies (unpublished results), and by X-ray diffraction studies (S.J. Hamodrakas, unpublished data).

Numerous β-turns are also predicted, especially in localized regions (see below), in agreement with certain features of the laser Raman spectra (unpublished results). The exact prevalence of turns is unclear, however, since some of them correspond to sequences which are plausibly interpreted as β-sheet or α-helix, and others may simply be random coils.

Sequence comparisons between chorion proteins, in both the A and B families, suggested the existence of evolutionary 'domains': in intrafamily comparisons, parts of the polypeptide appeared to be especially resistant to sequence divergence, and certain parts showed limited similarities even between families [6,7]. It now appears (Figs. 3 and 4) that the evolutionarily conservative sequence 'domains' also have structural significance, since they correspond to regions of extensive and characteristic secondary structure and are enriched in specific amino acids. Nevertheless, until these domains are further characterized, we shall refer to them as 'regions'. For convenience, we shall name them according to the amino acids in which they are enriched (Table I). The boundaries of the regions shown in Figs. 3 and 4 are slightly revised from those used previously [6,7], to take into account structural features as well as evolutionary and compositional properties.

On the basis of partial sequencing data, Regier et al. [26] pointed out that, as in avian keratin, cysteines are preferentially localized near the two ends of the chorion polypeptide chain. The complete chorion protein sequences now available confirm this feature, and further emphasize the analogy with avian keratins. In both feather and scale keratin, amino-terminal and carboxyl-terminal segments which are rich in cysteine are thought to provide a disulfide-linked matrix, in which fibrillar elements are situated, corresponding to internal polypeptide segments rich in neutral and hydrophobic residues [27–35]. Infrared spectroscopic studies [31] and comparisons of complete sequences for feather keratins from Emu and Silver Gull [29,30] show that the internal, hydrophobic segment tends to be evolutionarily conservative and contains the crystalline, β-sheet portion of keratin. By contrast, sequence variations are predominantly located at or near the cysteine-rich tails (matrix). In the chorion proteins, also, the sequences show the existance of an evolutionarily conservative, highly β-structured 'core', which we call the V, A region (Figs. 3 and 4). Greater variability in sequence and secondary structure is seen in the remainder of the molecule ('arms'), although the cysteines themselves, especially those of the carboxyl end, also tend to be found in conservative sequences.

The central conservative region (V, A)

In the B proteins, the most pronounced conservative region is approximately 40% of the total molecule in length, and is centrally located. Of 58 contiguous positions, between positions 45 and 102, only five show replacements in any of the seven analyzed proteins of this family, and none show deletions (Fig. 4) [21]. Since it is enriched in valine and alanine relative to the rest of the molecule (Table I), this region will be referred to as the V, A region.

A somewhat similar V, A region is also found in the A proteins. In addition to being enriched in valine and alanine, this region is also long (49 residues, or nearly half the total molecule); it is conservative (with no deletions, and replacements in only 12 residues, from seven fully sequenced proteins of this family); and it is internal, although closer to the carboxyl terminus. Limited sequence similarities exist between these two family-specific conservative regions [7].

For both A and B proteins, β-sheet strands are dominant in the central conservative region (V, A). In addition, limited but probably significant turns and α-helical segments are predicted.

For both families, the region putatively begins with a β-turn (SYGG in all cases except one, where it is EYGG). A short sequence follows (12 residues in the A family, 10 in the B), in which sequential β-sheet and α-helical segments are predicted; the predictions overlap, however, and the β- and α-segments are so short as to be perhaps mutually exclusive. In the B family, after two proline residues, a second sequence of β- and α-segments is predicted; these segments are longer

and only slightly overlapping, and thus are more likely to exist as such. In the A family, only a single proline exists, and is not followed by a similar $\beta\alpha$ unit. In summary, it is possible that, after the initial β-turn, the V, A region shows in the A family a single $\beta\alpha$ unit followed by a proline, and in the B family two $\beta\alpha$ units separated by two prolines. $\beta\alpha\beta$ super-secondary structures are known to be generally very stable [34,35]. It should be noted that here and elsewhere (see below) the putative β-strands often contain glycine alternating with relatively bulky residues (GLGV; VGLGV; GIGN). The remainder of the V, A region (31 residues in the A family; 30 in the B) is predominantly in the β-sheet conformation.

An intriguing feature of the detailed predictions (Figs. 1 and 2), which is obscured in the summaries (Figs. 3 and 4), is an apparent periodicity of β-sheet strands, evidenced by nearly periodic occurrence of maxima for β-sheet predictions. For example, in the pc401 sequence, clear maxima occur approximately 10 residues apart, at or around positions 52, 63, 75, 85, 93, 101 and 111. In between these maxima, β-turns are often predicted. These features are reminiscent of Silver Gull feather keratin [36,37], where the β-sheet in the unit of structure is comprised of four segments with eight residues each, in antiparallel arrangement; that structure is revealed by an 8-residue periodicity in β-sheet propensities, and a similar but out-of-phase periodicity of random coil propensities. Thus, it is not unlikely that a significant proportion of the major chorion proteins, and in particular much of the conservative V, A region, is in the 'cross-β' conformation [38], i.e. is folded into short, antiparallel β-sheet strands, connected by β-turns. Near the beginning of the V, A region, antiparallel β-sheet strands may possibly be connected by one or two α-helical segments.

The carboxyl-terminal conservative region (G, C)

An invariant sequence, $G(CG)_3$, is found near the carboxyl terminus in all A and B proteins sequenced to date. It is strongly predicted to participate in β-turns, possibly two contiguous ones, and probably serves to cross-link chorion fibrils via disulfide bridges in the mature eggshell [1,26]. A more variable segment of similar composition is also found near the amino terminus (around approximately residue 20 in the A family, residue 15 in pc401, and residue 30 in 10a and pc408); that segment might also serve a similar cross-linking function; it is predicted to exist in a β-turn, although in some cases an alternative β-sheet prediction is possible.

The tyrosine-enriched regions (G, Y and G, Y, L)

Most tyrosine residues tend to be clustered, in a region (G, Y, L or G, Y in Figs. 3 and 4) which either precedes (A family) or follows (B family) the central conservative region, V, A. The tyrosine-enriched regions (Table I) are also very rich in glycine; for the A family it is enriched in leucine as well. Despite the differences in location and composition, the G, Y, L region of the A family and the G, Y region of the B family show sufficient sequence similarities to suggest the possibility of an ancient homology [6,7].

The tyrosine-enriched regions tend to be somewhat variable, in sequence and in secondary structure, even within the same family. Their sequence divergence may be related to structures and functions which are specific for specific proteins. In the B family, the G, Y region appears to consist largely of β-sheet strands alternating with β-turns; perhaps it serves as an extension of the V, A core, with variations which may be significant for the specific functions of particular B family proteins. In the A family, the G, Y, L region appears less structured; β-turns may predominate, although β-sheet strands may also exist.

Both tyrosine-enriched regions may be partly derived from tandem repeats of a consensus pentapeptide, GYGGL [4–7]. A prototype is the pc18 sequence, in which this repeat array might correspond to a series of β-turns (GYGG), separated by a single residue (L). Alternative predictions of β-sheet structure are possible, however. In scale keratin, a repetitive sequence of the type GGX (where X = F, L or Y) exists [32], and is thought to form β-sheet structures, possibly facilitating the characteristic hexagonal packing of scale keratin [28]. Furthermore, the synthetic polypeptide, $(GGA)_n$, assumes a β-conformation under certain conditions and a polyglycine II structure under others [39].

In several cases, in these and other parts of the molecule, relatively small residues (G, but also A

or T) tend to alternate with bulky residues (e.g., V, L, I or Y), especially in sequences which are predicted as β-strands (e.g., residues 61 to 65, 109 to 113, 121 to 131 and 139 to 144 in the B family). This is reminiscent of the alternation of 'small' (G) and 'bulky' (A, S) residues to opposite sides of the β-sheet structure in silk fibroin [40,41], and may be similarly important in the chorion for packing of successive β-sheets, with the side groups of one sheet fitting between similar (in terms of volume), groups of an adjacent sheet [42].

Other family-specific regions

A segment of the B proteins close to the amino terminus is labeled 'G' in the figures. The features which it shares with the similarly located G, Y, L region of the A family include variable combinations of β-strands and β-turns, a variable G, C-enriched portion (see above), and short, identical but variably placed sequences (GYEG, GYDG or GYGA; not repeated). The G and G, Y, L regions, however, are substantially different in total length, internal arrangement, and amino acid composition (other than high glycine content).

Between the G and the V, A regions of the B proteins, a short region is found, which is highly structured but variable: in 10a and pc408 it is clearly an α-helical segment, whereas in pc401 it is predicted as a β-strand.

The extreme amino-terminal region of the B proteins is constant in sequence only for the first three residues, but shows a consistent structural pattern: a short β-strand followed by a β-turn.

In the A family, as already mentioned, the central V, A region and the carboxyl-terminal G, C segment are not separated by a G, Y region. A short connecting segment (GRCX) exists instead, which is predicted as part of either β-sheet or β-turn structures.

In the AII subfamily, exemplified by pc609 and pc292, the extreme amino-terminal region has a peptide which is totally missing in the AI subfamily, exemplified by pc18 [4–6]. Except for the probable presence of a β-turn, the structure of this AII-specific peptide is unclear.

Summary and Conclusions

The characteristic features of both A and B chorion proteins include: (1) a long, more or less central, V, A-enriched region which forms a highly structured, predominantly β-sheet polypeptide 'core', and (2) a short, invariant segment near the carboxyl-terminus, which is enriched in cysteine and presumably functions for cross-linking.

The V, A core regions differ somewhat between the A and B families, presumably corresponding to distinct (although related), family-specific structures and functions. Similarly, family-specific functions might be ascribed to limited constant features of the amino terminal region of the B's: a short β-sheet segment followed by a β-turn. The amino terminal residue of all characterized B proteins is a cyclized glutamine (pyrrolidone carboxylic acid) [43].

Conversely, features which vary within a family (e.g., between subfamilies) presumably serve more specific functions. According to this interpretation, protein-specific or subfamily-specific functions would be served by the amino terminal peptide and the G, Y, L region of the A proteins, the G region and the adjacent α-helical or β-strand connecting segment of the B's, and to a lesser extent the G, Y region of the B's.

The validity of our structural predictions, and of their interpretations, remains to be tested by more refined theoretical work as well as by experimental analysis. As a first step, we have carried out a laser Raman spectroscopic study, which confirms the preponderance of β-pleated sheets in the silkmoth chorion (unpublished results).

Acknowledgements

We are grateful to Drs. J. Garnier, J.A. Lenstra and Prof. K. Nagano for computer programs, Dr. J.C. Regier for reading the manuscript, M.D. Koehler and G. Meyer for drawing the figures and S. Foy for secretarial assistance. We thank the Greek Ministry of Agriculture Computer Laboratory for computing facilities, and the University of Athens and the NIH for financial support.

References

1 Kafatos, F.C., Regier, J.C., Mazur, G.D., Nadel, M.R., Blau, H.M., Petri, W.H., Wyman, A.R., Gelinas, R.E., Moore, P.B., Paul, M., Efstratiadis, A., Vournakis, J.N., Goldsmith, M.R., Hunsley, J.R., Baker, B., Nardi, J. and

Koehler, M. (1977) in Results and Problems in Cell Differentiation, Vol. 8, pp. 54–145, Springer-Verlag, Berlin
2 Smith, D.S., Telfer, W.H. and Neville, A.C. (1971) Tissue Cell 3, 477–498
3 Regier, J.C., Mazur, G.D. and Kafatos, F.C. (1980) Dev. Biol. 76, 286–304
4 Regier, J.C., Kafatos, F.C., Goodfliesh, R. and Hood, L. (1978) Proc. Natl. Acad. Sci. USA 75, 390–394
5 Rodakis, G. (1978) Ph.D. Thesis, University of Athens, Athens, Greece
6 Jones, C.W., Rosenthal, N., Rodakis, G.C. and Kafatos, F.C. (1979) Cell 18, 1317–1322
7 Tsitilou, S.G., Regier, J.C. and Kafatos, F.C. (1980) Nucl. Acids Res. 8, 1987–1997
8 Chou, P.Y. And Fasman, G.D. (1974) Biochemistry 13, 221–222
9 Chou, P.Y. and Fasman, G.D. (1974) Biochemistry 13, 222–245
10 Burgess, A.W., Ponnuswamy, P.K. and Scheraga, H.A.(1974) Israel J. Chem. 12, 239–286
11 Lim, V.I. (1974) J. Mol. Biol. 88, 857–872
12 Lim, V.I. (1974) J. Mol. Biol. 88, 873–894
13 Lenstra, J.A., Hofsteenge, J. and Beintema, J.J. (1977) J. Mol. Biol. 109, 185–194
14 Lenstra, J.A. (1977) Biochim. Biophys. Acta 439, 261–273
15 Garnier, J., Osguthorpe, D.J. and Robson, R. (1978) J. Mol. Biol. 120, 97–120
16 Nagano, I. (1977) J. Mol. Biol. 109, 235–250
17 Nagano, K. (1977) J. Mol. Biol. 109, 251–274
18 Robson, B. and Suzuki, E. (1976) J. Mol. Biol. 107, 327–356
19 Dufton, M.J. and Hider, R.C. (1977) J. Mol. Biol. 115, 177–193
20 Chou, P.Y. and Fasman, G.D. (1977) J. Mol. Biol. 115, 135–175
21 Jones, C.W. (1980) Ph.D. Thesis, Harvard University, Cambridge, Massachusetts
22 Jones, C.W. and Kafatos, F.C. (1980) Cell 22, 855–867
23 Jones, C.W. and Kafatos, F.C. (1980) Nature 284, 635–638
24 Schulz, G.E., Barry, C.D., Friedman, J., Chou, P.Y., Fasman, G.D., Finkelstein, A.W., Lim, V.I., Ptitsyn, O.B., Kabat, E.A., Wu, T.T., Levitt, M., Robson, B. and Nagano, K. (1974) Nature 250, 140–142
25 Argos, P., Schwarz, J. and Schwarz, J. (1976) Biochim. Biophys. Acta 439, 261–273
26 Regier, J.C., Kafatos, F.C., Kramer, K.J., Heinrikson, R.L. and Keim, P.S. (1978) J. Biol. Chem. 253, 1305–1314
27 Filshie, B.K. and Rogers, G.E. (1962) J. Cell Biol. 13, 1–12
28 Stewart, M. (1977) J. Ultrastruct. Res. 60, 27–33
29 O'Donnell, I.J. (1973) Aust. J. Biol. Sci. 26, 415–437
30 O'Donnell, U.J. and Inglis, A.S. (1974) Aust. J. Biol. Sc. 27, 369–382
31 Suzuki, E. (1973) Aust. J. Biol. Sci. 26, 435–437
32 Walker, I.D. and Bridgen, J. (1976) Eur. J. Biochem. 67, 283–293
33 Walker, I.D. and Rogers, G.E. (1976) Eur. J. Biochem. 69, 329–339
34 Sternberg, M.J.E. and Thornton, J.M. (1976) J. Mol. Biol. 105, 367–382
35 Nagano, K. (1977) J. Mol. Biol. 109, 235–250
36 Fraser, R.D.B., MacRae, T.P., Parry, D.A.D. and Suzuki, E. (1971) Polymer 12, 35–56
37 Fraser, R.D.B. and MacRae, T.P. (1976) Proc. 16th International Ornithological Congress, Canberra pp. 443–451, Australian Academy of Science
38 Geddes, A.J., Parker, K.D., Atkins, E.D.T. and Beighton, E. (1968) J. Mol. Biol. 32, 343–358
39 andries, J.C., Anderson, J.M. and Walton, A.G. (1971) Biopolymers 10, 1049–1057
40 Lucas, F., Shaw, J.T.B. and Smith, S.G. (1957) Biochem. J., 66, 468–479
41 Lucas, F., Shaw, J.T.B. and Smith, S.G. (1962) Biochem. J., 83, 164–171
42 Marsh, R.E., Corey, R.B. and Pauling, L. (1955) Biochim. Biophys. Acta. 16, 1–34
43 Regier, J.C. and Kafatos, F.C. (1981) J. Biol. Chem. 265, 6444–6451