Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 1043–1047, February 1983
Genetics

Silkmoth chorion multigene families constitute a superfamily: Comparison of *C* and *B* family sequences

(gene evolution)

JEROME C. REGIER*, FOTIS C. KAFATOS†‡, AND STAVROS J. HAMODRAKAS‡

*Department of Biochemistry, Molecular and Cellular Biology, Northwestern University, 2153 Sheridan Road, Evanston, Illinois 60201; †Department of Cellular and Developmental Biology, The Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138; and ‡Department of Biology, University of Athens, Panepistimiopolis, Athens 621, Greece

Communicated by E. Margoliash, May 17, 1982

ABSTRACT **We have characterized a new family of silkmoth chorion genes, called *C*, which is distinct from previously characterized *A* and *B* families. The amino acid compositions of 18 purified C proteins have been determined. Three subgroups are recognized on the basis of compositional similarities and may correspond to distinct gene families or subfamilies. The sequences of two overlapping cDNA clones have been determined in their entirety and shown to correspond to a *C*-specific sequence. Obvious homology is observed between the middle portions of the *C* sequence and previously characterized *B* sequences. By contrast, the arms of the *C* sequence share no significant similarities either with each other or with the corresponding arms of *B* sequences. Thus, the same tripartite structure originally observed in *A* and *B* family sequences is also present in the *C* family and may have functional significance. Secondary structure prediction of the *C* sequence is presented and supports this conclusion. The observed homology between *C* and *B* family sequences clearly establishes that silkmoth chorion multigene families constitute a superfamily.**

The silkmoth chorion (eggshell) is a paradigm of a eukaryotic structure built from the products of many genes. The mature chorion consists of more than 100 structural proteins, secreted by the follicular epithelium around the maturing oocyte (1). These proteins are produced in sequence, according to a precise developmental program (2), and assemble to form a complex structure that serves multiple physiological functions, including physical protection and embryonic respiration (3, 4). On the basis of size, most chorion proteins have been divided into four classes, A, B, C, and D in order of increasing molecular weight (5). The two predominant classes (88% of the total protein) are A (molecular weight range, 9,000–12,000) and B (molecular weight range, 12,000–14,000). These classes have been extensively characterized by protein (6–9) and DNA (10–12) sequence determinations; we have shown that each corresponds to a multigene family. Limited similarities between *A* and *B* sequences have suggested that the two families may possibly share ancient homology (10, 11).

Here we report on the third class, C, of chorion proteins. Three C subclasses are defined by amino acid compositional similarities and differences and may correspond to three additional gene families or subfamilies. Most of the sequence of one class C component has been determined by selection and sequence determination of cDNA clones. The results show localized but unambiguous homology to *B* family sequences and, thus, support the notion that the chorion gene families are related, constituting an evolutionarily ancient, developmentally regulated superfamily.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Purification and Analysis of Class C Proteins. Mature chorions from the silkmoth *Antheraea polyphemus* were dissolved and fractionated by differential precipitation (6). Fraction P (the precipitate, enriched in B and C proteins) was fractionated on a Bio-Gel P-150 column containing 7 M guanidinium chloride (9, 13). C protein-containing fractions were pooled, and the C proteins were resolved from each other by preparative isoelectric focusing (carrier ampholyte range, pH 3–5) in polyacrylamide slab gels (9). Individual bands were eluted and analyzed by $NaDodSO_4$ gel electrophoresis and amino acid analysis (14, 15).

Construction, Selection, and Sequence Determination of cDNA Clones. The construction of chorion-specific cDNA libraries has been described (16). Two libraries, one (89 clones) derived from early-stage choriogenic follicles (and therefore enriched for *C* sequences) and the other (532 clones) from all stages, used RSF 1030 and pML21, respectively, as plasmid vectors and *Escherichia coli* HB 101 as host. The methods for screening to select specific clones are described elsewhere (16–18). DNA sequence determination was by the method of Maxam and Gilbert (19).

RESULTS

Fractionation and Characterization of Class C Proteins. Beginning with total chorion proteins, a fraction that contains largely class C-size proteins (molecular weight, 16,000–20,000) could be obtained by differential precipitation and column chromatography (Fig. 1). This fraction was resolved by isoelectric focusing into 42 bands, 18 of which were preparatively isolated for subsequent analysis (Fig. 1, lane I). Many of the excised isoelectric focusing bands were heterogeneous with respect to size (e.g., band 5), although others (e.g., band 6) appeared to be homogeneous (Fig. 1, lanes 5 and 6, respectively). These results suggest the existence of 50–75 components in the C class; some of these may be allelic variants.

Amino acid analysis of the 18 class C protein fractions revealed some general similarities to previously characterized chorion proteins of the A and B classes: high percentage of nonpolar amino acids, especially high content of glycine and alanine, low content of charged amino acids (Table 1). However, relative to proteins of classes A and B, class C proteins were depleted in cysteine and glycine (except bands 7, 10, 12, 17, and 18) and enriched in proline (except band 10). Most C proteins showed additional differences for amino acids such as lysine, threonine, serine, and phenylalanine. Their compositional similarities but nonidentities, together with their restricted size range, were consistent with the interpretation that chorion class C proteins are encoded by related genes.

Class C protein compositions were more heterogeneous than

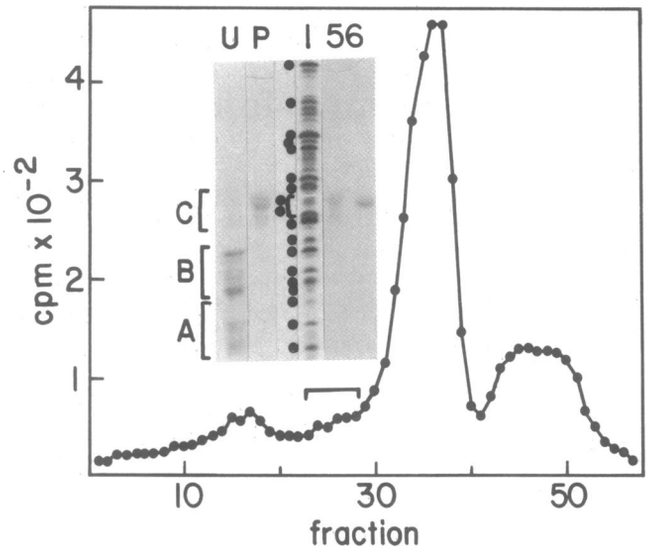

FIG. 1. Fractionation of class C proteins. Total chorion proteins were solubilized, [^{14}C]carboxamidomethylated, and fractionated by differential precipitation (1). The precipitate was further fractionated by column chromatography. C protein-containing fractions were pooled (see horizontal bracket) and analyzed by NaDodSO$_4$ gel electrophoresis (lane U, unfractionated; lane P, pooled column fractions). The pooled C protein-containing fractions were fractionated on an isoelectric focusing gel (lane I). Eighteen bands, indicated by dots to the left, were isolated and further analyzed. They are referred to by number, beginning with 1 at the top (basic end). There is uncertainty within the range indicated for bands 8 and 9. Each band was then analyzed on NaDodSO$_4$ gels to assay size and heterogeneity (e.g., lane 5, band 5; lane 6, band 6).

those of classes A and B. In particular, the glycine/alanine ratio varied from 0.82 (band 3) to 4.59 (band 10). Within this range, three subgroups, each with similar ratios, could be recognized (CI, $\bar{x} = 1.10$, $s = 0.16$; CII, $\bar{x} = 1.98$, $s = 0.27$; CIII, $\bar{x} = 4.53$, $s = 0.08$).

Isolation and Characterization of Class C Protein-Encoding cDNA Clones. To circumvent the difficulties imposed by high complexity and low abundance, we turned to cDNA clones to obtain the sequence of a class C component. To identify appropriate clones, we took advantage of three features that generally distinguish C proteins and their encoding nucleic acids from the major A and B sequences—larger size, earlier developmental specificity, and lower abundance (5, 20). One clone, H12, clearly hybridized to a large, early, and nonabundant RNA and, thus, was selected for further characterization as a likely class C protein-specific clone. Since H12 proved to have only a short insert (201 bp; see below), it in turn was used as a probe to obtain a similar clone, pc404, with a larger insert size (412 bp).

The nature of H12 and pc404 sequences was confirmed by RNA blot analysis (21). Both clones hybridized preferentially to RNA that is developmentally early and larger than RNAs encoding A and B proteins (unpublished data). Hybrid-selected translation also established that both clones hybridized to RNA that yielded a class C-size polypeptide upon cell-free translation (ref. 22; unpublished data).

The sequences of H12 and pc404 were determined and shown to share a stretch of 112 base pairs with 94% homology (Fig. 2). Thus, for purposes of defining the main features of a class C component, H12 and pc404 were combined to generate a hybrid sequence of 501 base pairs, which we call pc404–H12. The reading frame for this sequence was determined from the observed homology with other chorion protein sequences and from the fact that only one frame contained no termination codons. (The other five frames contained 3–8 terminators, generally distributed over much of the sequence.) The sequence was incomplete at both ends. It appeared to contain neither an initiator methionine, a signal-like sequence of the type known to be present in all other chorion proteins (12, 22), nor a termination codon.

pc404–H12 has a translatable sequence of 167 amino acid residues with a molecular weight of 16,189 (Fig. 3). This can be

Table 1. Amino acid compositions of chorion sequences (residues per 100 residues)*

	CI								CII									CIII		As‡	Bs‡
	1	2	3	4	5	6	8	9	7	11	13	14	15	16	pc404 –H12†	17	18	10	12	(s/s)	(p/p)
Cys§	2.7	2.3	2.4	2.7	2.2	2.7	2.0	2.1	2.6	1.7	1.6	1.6	1.9	2.0	1.8	2.0	2.1	2.8	2.2	8.4	5.6
Asp¶	5.1	5.0	5.2	5.0	5.5	5.6	4.3	4.4	5.9	4.6	3.6	4.8	5.0	4.6	5.4	5.4	5.3	6.0	5.7	2.5	4.0
Thr	4.5	4.2	4.5	4.6	5.6	5.6	4.3	4.5	3.9	4.9	5.0	5.1	4.4	4.6	5.4 (6.6)	4.7	4.7	2.7	3.2	3.4	3.0
Ser	6.2	5.6	6.2	6.1	7.0	7.0	7.0	7.2	5.8	5.2	5.8	4.8	5.6	5.7	8.4 (7.8)	5.4	4.7	4.4	4.2	3.0	3.8
Glu¶	4.0	4.0	3.6	3.7	3.4	4.4	4.2	4.3	3.4	3.9	4.3	3.9	4.0	4.2	3.6	4.0	3.6	3.4	3.4	4.2	4.7
Pro	6.7	7.9	9.5	9.0	9.3	8.0	8.3	8.4	6.1	10.4	11.4	12.1	11.6	11.5	10.2	8.3	8.2	4.3	6.1	4.1	4.5
Gly	21.8	21.0	16.9	18.2	19.8	19.6	20.5	19.6	28.0	25.0	24.0	23.0	23.1	22.3	21.6	27.2	27.4	31.2	34.4	32.4	32.2
Ala	16.8	18.5	20.6	19.9	16.8	16.4	17.6	18.0	11.6	12.0	13.5	13.4	13.4	13.0	12.0	12.3	12.6	6.8	7.7	13.8	11.4
Val	8.3	6.0	5.6	5.9	6.0	6.5	7.5	7.8	5.9	7.1	6.5	7.5	7.3	6.6	8.4 (7.8)	5.8	5.7	5.5	6.4	7.2	6.1
Met	1.0	1.3	1.6	1.4	1.5	0.9	1.3	1.3	1.4	1.2	0.8	0.9	0.8	0.7	1.2	0.3	0.2	1.4	1.4	0.1	0.5
Ile	3.3	5.6	5.8	5.5	5.7	5.7	4.8	4.7	4.6	5.4	4.8	5.9	5.1	4.9	4.8	4.5	4.6	3.7	4.3	3.8	3.6
Leu	7.2	6.1	6.2	6.4	6.1	6.5	6.4	6.2	8.0	6.6	6.9	6.2	7.3	7.2	4.8	3.9	3.7	9.7	8.6	6.4	7.9
Tyr	5.6	5.7	5.3	5.2	4.8	5.2	4.9	4.9	7.3	5.8	6.1	5.2	6.2	6.5	6.6	7.2	7.7	9.4	8.2	6.2	6.9
Phe	1.9	2.2	2.4	2.3	2.7	2.3	2.9	2.9	2.3	3.1	3.6	3.5	3.3	3.9	4.2	6.4	6.6	2.3	2.8	1.0	1.6
His	0.0	0.0	0.1	0.0	0.1	0.1	0.2	0.2	0.3	0.2	0.1	0.1	0.2	0.1	0.0	0.1	0.0	0.5	0.4	0.0	0.0
Lys	0.5	0.1	0.2	0.3	0.6	0.1	0.2	0.2	0.3	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.0	0.0	1.0	0.4
Trp	1.1	1.2	0.8	1.2	1.0	0.8	1.6	1.3	0.7	1.8	1.2	1.6	0.1	1.5	1.2	2.0	2.3	0.1	0.1	0.5	1.2
Arg	3.1	3.1	3.1	2.7	1.8	2.5	1.8	2.0	1.9	1.1	0.6	0.6	0.6	0.6	0.6	0.7	0.6	1.9	1.2	1.8	2.4

* Threonine, serine, and tryptophan values are corrected for losses during hydrolysis (see ref. 6). Glucosamine and galactosamine are absent from all 18 class C protein fractions.

† Calculated from hybrid clone sequence. Alternate values are due to polymorphism between pc404 and H12. See Fig. 3.

‡ Proteins of classes A and B; ref. 6.

§ Determined as carboxymethylcysteine.

¶ Includes the amidic forms.

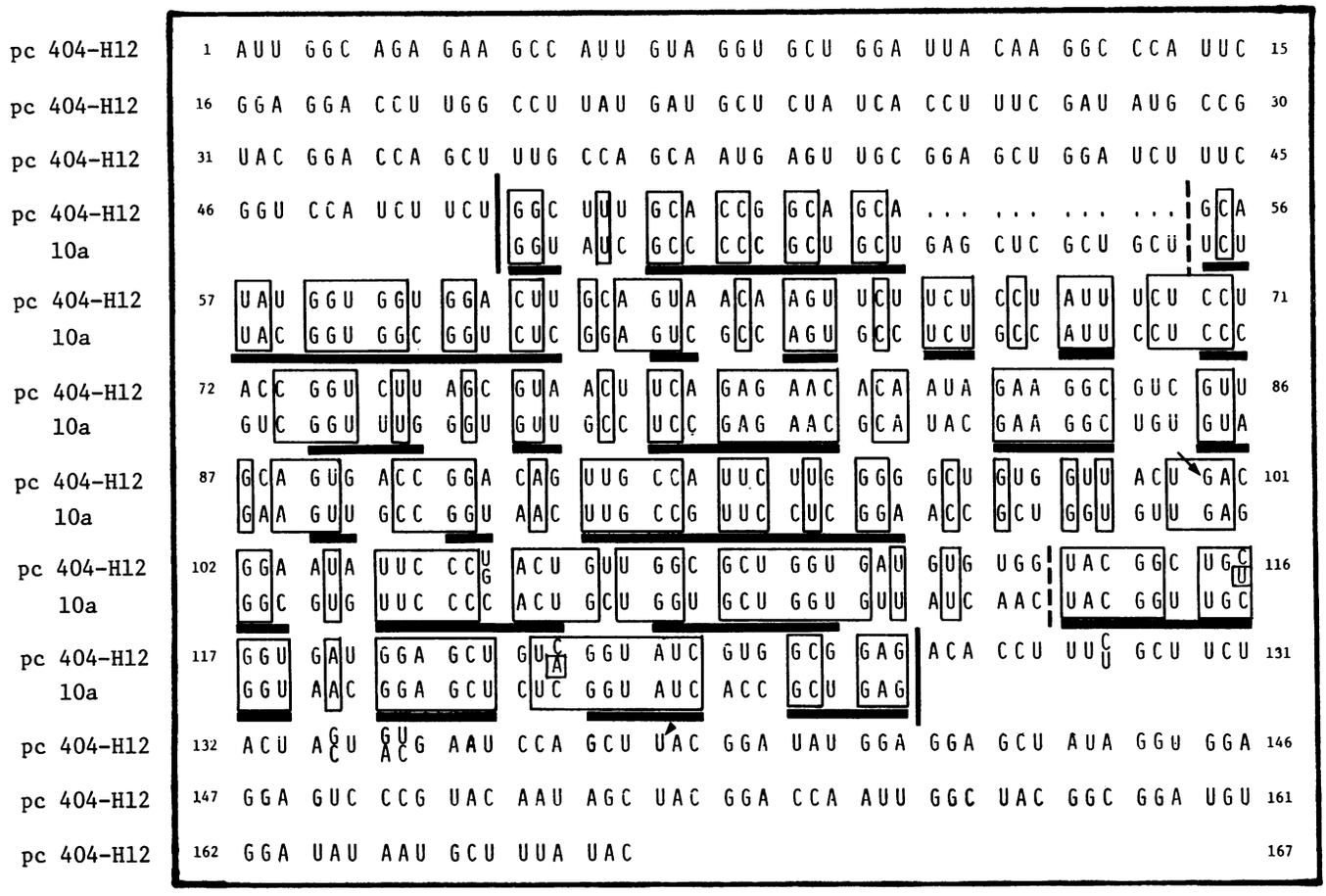

FIG. 2. mRNA sequence of the pc404–H12 chorion component and comparison with the homologous region of the 10a nucleotide sequence. Identical nucleotides are enclosed in boxes. Codons are numbered consecutively for pc404–H12 from the 5′ end of the insert. Double codons at positions 105, 116, and 121 result from polymorphism between pc404 (upper alternative) and H12. Those encoding identical amino acids are underlined. The dashed vertical lines indicate the conservative central region defined previously from B protein sequence comparisons (23), and the solid vertical lines indicate the extent of homology between B and C protein sequences; see also Fig. 4. The arrow indicates the first nucleotide (in codon 101) represented in clone H12. The arrowhead indicates the last nucleotide in clone pc404 (in codon 138). All mRNA and protein sequences (see also Figs. 3 and 4) were originally determined as cloned DNA sequences and then converted to their mRNA and protein counterparts.

compared with pc401, which encodes the largest known sequence of the class B protein family, and includes 153 residues (mature translated sequence only) with a molecular weight of 14,245 (10). The large size of pc-404–H12 relative to pc401 establishes that it cannot encode a class A or B protein sequence.

Given the distinctive nature of class C protein compositions, we calculated from the nucleotide sequence of pc404–H12 the amino acid composition of its translation product for comparison (Table 1). As is typical of most C protein sequences, pc404–H12 protein is low in glycine (relative to class *A* and *B* sequences) and rich in proline. Its glycine/alanine ratio (1.80) is typical of proteins in the CII subgroup (range, 1.72–2.41). Most striking is its close similarity with a particular C protein fraction (no. 16). Protein 16 was estimated to have 167 residues, if one assumes that it contains one arginine. This coincidence in lengths between proteins 16 and pc404–H12, combined with their very similar compositions, suggests that pc404–H12 DNA contains almost the entire mature coding sequence. The minor differences in composition are easily explained by the small missing ends of the pc404–H12 sequence and by slight differences between related C proteins.

Sequence Comparison of pc404–H12 and a Class B Protein. A comparison of pc404–H12 DNA with a typical member of the *B* gene family, clone 10a, revealed that their translation products have extensive similarity from amino acid position 50 to position 126 (Fig. 3). In this region, 48/77 or 62% of the residues are identical (excluding a single gap of four residues). The area of homology encompasses (but is not limited to) the "central region" of class B protein sequences (23); that region, corresponding to positions 56–113 in pc404–H12 protein, is conserved among all B sequences (12) and is thought to form a highly structured "core," largely in the β-sheet conformation (24). Between positions 50 and 126, sequence homology is also obvious at the nucleotide level (Fig. 2): 65% of all nucleotides are identical, including 43% of the nucleotides in the third codon position. Clearly, these two sequences are homologous.

No unambiguous homology was detectable in the flanking sequences (matching of glycines near the COOH-terminus is probably not significant because glycines are so abundant in this region). However, when several gaps were introduced, the COOH-terminal sequence (residues 139–167 in pc404–H12) could be matched with the corresponding sequence of 10a (13/21 residues are identical, or 62%, excluding four gaps totaling 10 residues).

Secondary Structure Prediction for pc404–H12. We have predicted the secondary structure (α helix, β sheet, β turn) of pc404–H12, using six different predictive methods in a manner as described (see Fig. 4 and ref. 23). In the interior region of homology with class B sequences (positions 50–126), extensive β sheet is strongly predicted, possibly in the form of short strands that alternate with β turn. Alternatively, one and, less likely, two short strands may be in α-helical conformation.

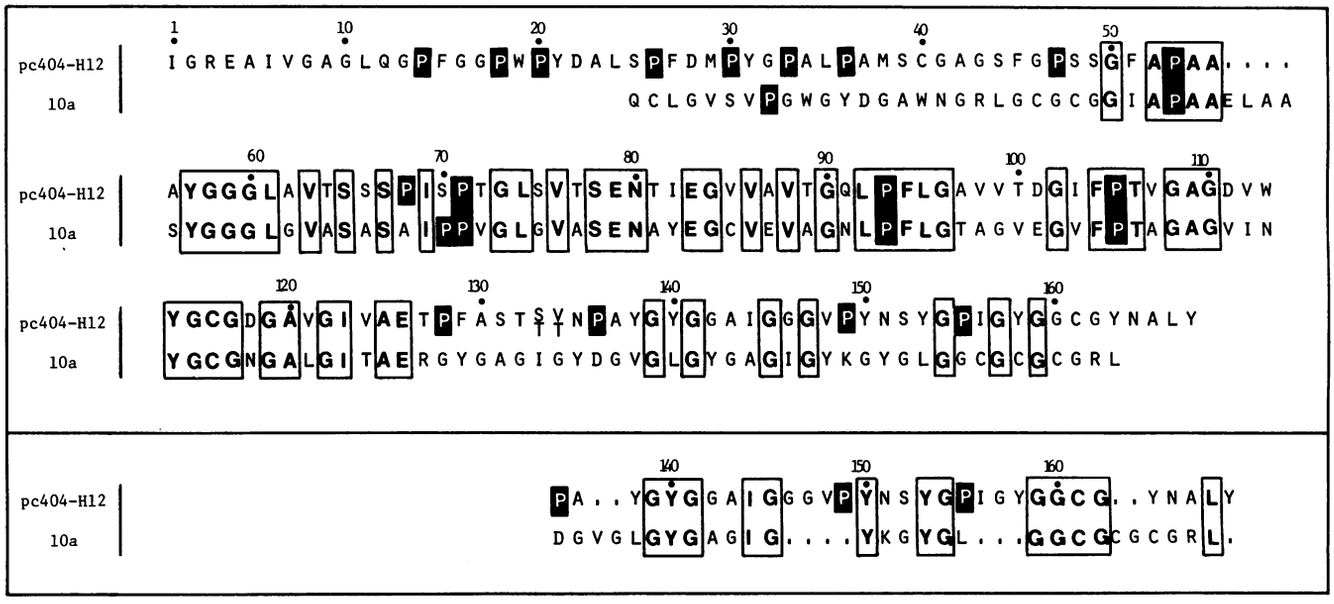

FIG. 3. (*Upper*) Comparison of protein sequences of the pc404–H12 and 10a chorion components. Residues that are identical are in bold-faced letters and are enclosed in boxes. Proline residues are indicated by white letters on black backgrounds. Gaps that improve sequence alignment are shown by dots. Double residues at positions 133 and 134 in the pc404–H12 sequence result from polymorphism between pc404 (upper alternative) and H12. (*Lower*) An alternative alignment for positions 136–167.

In the flanking sequences, most of the predicted secondary structure is β turn. This is particularly true in the NH_2-terminal region, where only one β-sheet strand is prominent in the predictions (positions 4–10). In the COOH-terminal region, predicted β turns are dominant, but β-sheet strands may also exist.

DISCUSSION

The *C* Family of Chorion Genes Is a Branch of a Superfamily. In this report we have partially characterized the C family of chorion proteins and shown that it belongs to a larger superfamily. Among the multiple but distinct C proteins, which share a restricted molecular weight range and important compositional features, three subgroups are recognized, chiefly by differences in glycine/alanine ratios; these could correspond to distinct families or subfamilies (Table 1). CI and CII, representing the majority of class C proteins, are synthesized exclusively during the early period of choriogenesis (2), whereas CIII proteins are synthesized only during the very late period (2, 25). It seems reasonable that these subgroups will turn out to be distinct both developmentally and structurally, as already has been shown for classes A and B subfamilies (16).

pc404–H12 is a component typical of many others in the C family (CII subgroup), according to amino acid composition,

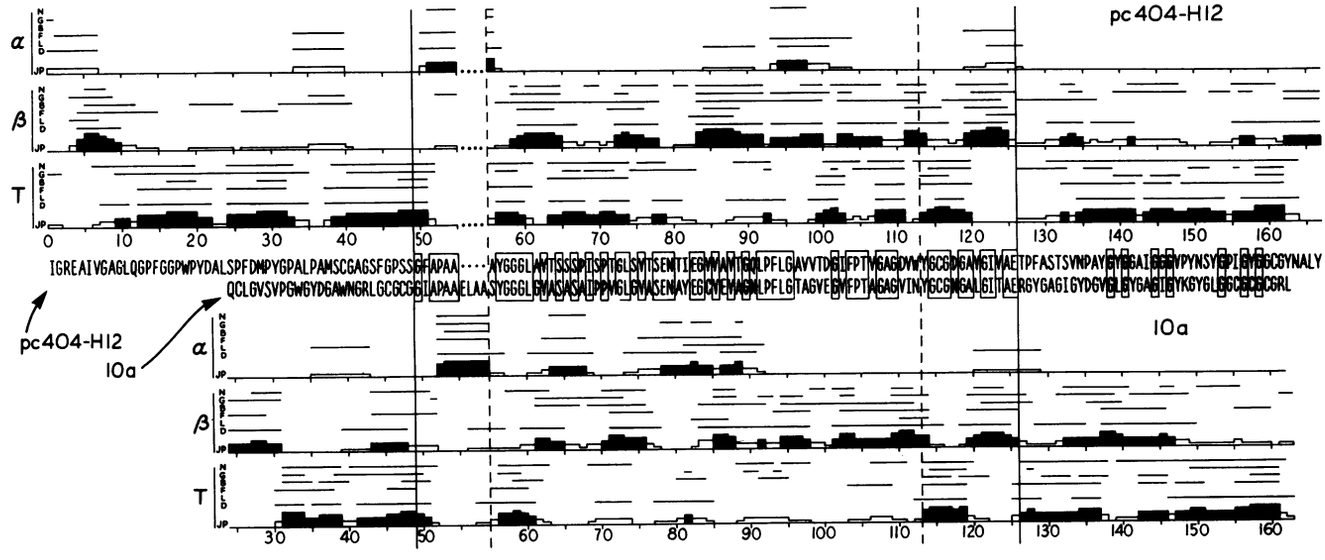

FIG. 4. Secondary structure predictions for protein sequences of the pc404–H12 and 10a chorion components. Individual predictions for α helix (α), β sheet (β), or β turn (T) are shown by horizontal lines, as derived according to Nagano (N), Garnier *et al.* (G), Burgess *et al.* (B), Chou and Fasman (F), Lim (L), and Dufton and Hider (D). See ref. 23 for a complete listing of references. Joint prediction histograms (JP), constructed by tallying the individual predictions, are also shown. The most probable structures, predicted by three or more methods, are shaded. Sequences are numbered from the NH_2- to the COOH-terminus of pc404–H12. Identical residues are enclosed in boxes. A gap between positions 55 and 56 in pc404–H12 is indicated by dots. As in Fig. 2, the solid vertical lines indicate the borders of the clear interfamily homologies; the dashed vertical lines, the conservative central region of B protein sequences (12, 23).

size, and developmental specificity. Thus, its clear homology with part of a typical B-family sequence unequivocally establishes that the C and B families are evolutionarily related. Although both the developmentally "middle" (lower molecular weight) and "late" (higher molecular weight) B subfamilies are homologous, the pc404–H12 component appears more related to the former subfamily (10, 12).

We have pointed out (11) limited similarities between A and B family sequences, which suggested (but were not in themselves adequate to establish) that A and B families are homologous. More recently, unequivocal homology was demonstrated between the A family and a family of high-cysteine (Hc) chorion proteins (26). Now that the C family has been shown to be homologous to the B family, the weight of the evidence leads us to the conclusion that these major chorion proteins constitute a superfamily.

Domain Structure. All chorion sequences characterized to date can be divided into three regions: a central region that is evolutionarily conservative within each family and highly structured, chiefly into β-pleated sheet, and two flanking arms, which evolve more rapidly and appear less structured (9, 12, 23, 26). It is significant that the homologous portion of the B and C sequences completely encompasses without gaps what was previously defined as the "central region" by detailed comparisons of multiple components of the B family (12, 23). From this and other comparisons (11, 26), it would appear that maintenance of the central region sequence and length is of special importance during the evolution of chorion proteins.

The arms are more variable, both within and between the families, and are subject to numerous segmental mutations (deletions, duplications, insertions) as well as base substitutions (12). Despite definite similarities in the arms of A, B, and Hc sequences (26), such similarities are not as obvious in the B versus C comparison. Pending comparisions with additional C family sequences, we suggest that the arms in the B and C families may be homologous but have sustained extensive sequence diversification. One aspect of this diversification would be the appearance of multiple proline residues in the C family arms, analogous to the enrichment for cysteine residues in the evolution of Hc proteins (26).

Although central regions and arms appear reasonably well-defined within each family by different modes of evolution, their borders do not seem unequivocal when different families are considered. For example, the G-F-A-P-A-A amino acid sequence of pc404–H12 protein (residues 50–55) is homologous with similar sequences in 10a protein and other members of the "middle period" B subfamily but is not conserved in the "late period" B subfamily (10). Similarly, positions 114–126 are shared between C and B families but not with A or Hc families (10, 35). Conservation of different marginal regions in different families and subgroups may be significant for protein structure (see Fig. 4 and ref. 23). It may be significant that in chorion genes, unlike immunoglobulin genes, borders of homology units do not correspond to intron locations (27).

Class C Protein Function. When the predicted secondary structures of B and C proteins are compared, both important similarities and differences are apparent (Fig. 4, ref. 23). Major similarities include the prevalence of putative β turns in the arms, especially in the NH_2-terminal arm, and the high degree of structure in the central region. This latter feature may be important for formation of fibrils, the predominant ultrastructural feature of silkmoth chorion (3, 4). Differences include the much greater length of the NH_2-terminal arm and the absence of α helix from the first half of the central region in the C sequence and may be related to the specific function of class C proteins. Elsewhere we have suggested that early class C proteins are major and necessary components of the initial chorion framework, which assembles extracellularly and is subsequently modified through expansion and densification by A and B proteins (4).

This work was supported by grants from the National Science Foundation and the National Institutes of Health to F.C.K., from the American Cancer Society (Illinois Division) and the National Institutes of Health to J.C.R., and from the University of Athens to S.J.H.

1. Regier, J. C., Mazur, G. D. & Kafatos, F. C. (1980) *Dev. Biol.* **76**, 286–304.
2. Paul, M. & Kafatos, F. C. (1975) *Dev. Biol.* **42**, 141–159.
3. Mazur, G. D., Regier, J. C. & Kafatos, F. C. (1982) in *Insect Ultrastructure*, eds. Akai, H. & King, R. C. (Plenum, New York), pp. 150–185.
4. Regier, J. C., Mazur, G. D., Kafatos, F. C. & Paul, M. (1982) *Dev. Biol.* **92**, 159–174.
5. Paul, M., Goldsmith, M. R., Hunsley, J. R. & Kafatos, F. C. (1972) *J. Cell Biol.* **55**, 653–680.
6. Regier, J. C., Kafatos, F. C., Kramer, K. J., Heinrikson, R. L. & Keim, P. S. (1978) *J. Biol. Chem.* **253**, 1305–1314.
7. Regier, J. C., Kafatos, F. C., Goodfliesh, R. & Hood, L. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 390–394.
8. Rodakis, G. C. (1978) Dissertation (Univ. of Athens, Athens, Greece).
9. Rodakis, G. C., Moschonas, N. K. & Kafatos, F. C. (1982) *Mol. Cell. Biol.* **2**, 554–563.
10. Jones, C. W., Rosenthal, N., Rodakis, G. C. & Kafatos, F. C. (1979) *Cell* **18**, 1317–1332.
11. Tsitilou, S. G., Regier, J. C. & Kafatos, F. C. (1980) *Nucleic Acids Res.* **8**, 1987–1997.
12. Jones, C. W. & Kafatos, F. C. (1982) *J. Mol. Evol.*, in press.
13. Regier, J. C. (1975) Dissertation (Harvard Univ., Cambridge, MA).
14. Goldsmith, M. R., Rattner, E. C., Koehler, M. D., Balikov, S. R. & Bock, S. C. (1979) *Anal. Biochem.* **99**, 33–40.
15. O'Farrell, P. H. (1975) *J. Biol. Chem.* **250**, 4007–4021.
16. Sim, G. K., Kafatos, F. C., Jones, C. W., Koehler, M. D., Efstratiadis, A. & Maniatis, T. (1979) *Cell* **18**, 1303–1316.
17. Grunstein, M. & Wallis, J. (1979) *Methods Enzymol.* **68**, 379–389.
18. Kafatos, F. C., Jones, C. W. & Efstratiadis, A. (1979) *Nucleic Acids Res.* **7**, 1541–1552.
19. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
20. Gelinas, R. E. & Kafatos, F. C. (1977) *Dev. Biol.* **55**, 179–190.
21. Thireos, G., Griffin-Shea, R. & Kafatos, F. C. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5789–5793.
22. Thireos, G. & Kafatos, F. C. (1980) *Dev. Biol.* **78**, 36–46.
23. Hamodrakas, S. J., Jones, C. W. & Kafatos, F. C. (1981) *Biochim. Biophys. Acta* **700**, 42–51.
24. Hamodrakas, S. J., Asher, S. A., Mazur, G. D., Regier, J. C. & Kafatos, F. C. (1982) *Biochim. Biophys. Acta* **703**, 216–222.
25. Mazur, G. D., Regier, J. C. & Kafatos, F. C. (1980) *Dev. Biol.* **76**, 305–321.
26. Rodakis, G. C. & Kafatos, F. C. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 3551–3555.
27. Jones, C. W. & Kafatos, F. C. (1980) *Cell* **22**, 855–867.