PATTERNS IN THE AMINO ACID COMPOSITIONS OF AVIAN EPIDERMAL PROTEINS

ALAN H. BRUSH
Biological Sciences Group, University of Connecticut, Storrs, Connecticut 06268 USA

ABSTRACT.—I used statistical analysis to compare the amino acid compositions of avian epidermal proteins. Comparisons were made of individual monomers from a single tissue of the morphological parts of various tissues, and of unfractionated tissues, at both the intra- and interspecific levels. The results imply a chemical basis for tissue protein structural heterogeneity, document tissue specificity, and indicate design relationships between protein distribution and tissue function. The amino acid composition of keratins provides information on the evolutionary relationships of natal down, feather, scale, claw, and beak. *Received 11 June 1979, accepted 3 June 1980.*

THE hard, keratinaceous structures of the avian epidermis differ morphologically and in their functional requirements. Nevertheless, the structural proteins of feathers, claw, down, scale, and beak commonly referred to as keratins are the products of closely related genes (Brush 1974, 1975) and share a common molecular configuration. Physical evidence indicates that the polypeptides that form these structures have a common β-pleated sheet structure and are organized into microfilaments (Fraser et al. 1972, Brush 1978). These in turn form filaments that consist of interacting chains, as modeled by Fraser and his colleagues (Fraser et al. 1971, Fraser and MacRae 1976). The supramolecular organization of the filaments has been studied with the electron microscope. The solubilized proteins are of two basic sizes, are tissue specific, and are extremely heterogeneous in electrophoresis (Brush 1975, Brush and Wyld MS). The feather proteins have an average molecular weight of 10,500 and those of the claw, scale, and beak of 14,500. The latter are enriched dramatically in glycine and show increased amounts of hydrophobic residues (e.g. Tyr, Leu, Phe) when compared to the lower molecular weight monomer in the feather and down (Walker and Bridgen 1976). The keratins of all the structures are relatively high in Cys, Pro, and Ser, and deficient in His, Met, and Lys.

The presence of a large number of similar protein monomers in a given tissue makes amino acid sequence studies difficult. This is especially true in keratins, where a family of closely related structural genes is involved. Simple separations based on charge (i.e. electrophoresis or chromatography) or solubility (i.e. Zn-acetate fractionation) are almost always incomplete. Absolute purification is difficult for those proteins whose sequence similarity is probably the result of only a few nucleotide changes. Yet full sequence data are available for pennacious feather keratin polypeptides of two species and for chicken down. Perhaps more progress has been made in the analysis of tryptic peptide maps and amino acid compositon of feather keratin monomers, techniques that require less preparation and are more rapid than sequence studies (Busch and Brush 1979). Relatively few species, however, have been examined, and only feathers of the contour plumage have been investigated. More extensive fractionations based on relative solubility of related groups of monomers have been carried out for a broad range of tissues in a half-dozen species (Brush in press). The fractions, which presumably share certain chemical properties, were not compared extensively by amino acid analysis. Thus the available amino

TABLE 1. Comparison of amino acid compositions of avian keratins. Values are given as average SΔQ. The range of values in a given comparison is included in parentheses. The large difference in the *Dacelo* beak value is discussed further in Brush (in press). The major tissues are feather (FKM), scale (SKM), and down (DKM). Details of intraspecific tissue differences and species differences in contour feather appear in Tables 3 and 4.

	SΔQ (range)	Reference
A. Intraspecific-Chemical		
1. Monomers		
Tern-FKM	48.8 (5.3–185)	Busch & Brush 1979
Turkey-FKM	35.7 (4–86)	Busch & Brush 1979
Gallus-FKM	44.5 (2–89)	Kemp & Rogers 1972
Gallus-FKM	9.2 (1–24)	Akahane et al. 1977
Gallus-SKM	11.5 (5.3–13)	Walker & Bridgen 1976
Gallus-DKM	25.3 (4.8–68)	Walker & Rogers 1976
2. Chemical fractionation		
Gallus-feather: ethanol ppt.	27.4 (13.3–55.2)	Harrap & Woods 1964
Gallus-scale: Zn ppt.	76	Brush in press
Dacelo-beak: Zn ppt.	421	Frankel & Gillespie 1976
B. Intraspecific-morphological		
1. Feather parts (rachis, calamus, barb, medulla)		
Gallus contour feather	15	Kemp & Rogers 1972
Gallus contour feather	15.4	Crewether et al. 1965
Anser contour feather	14	Crewether et al. 1965
Turkey contour feather	23	Crewether et al. 1965

acid composition data of avian epidermal structures are scattered and incomplete. Analyses from several laboratories are available for selected tissues of common species. Until now, little effort has been directed toward specific comparative morphological or systematic studies. The use of the amino acid compositions of the entire complex of solubilized proteins or partly purified preparations may provide significant information regarding the evolutionary relationships of the proteins, although this approach lacks the fine resolution of sequence data. Further, in concert with other data it may provide insights into structural problems, functional requirements, and evolution of these tissues. In this study I used a number of statistical methods to analyze and compare the amino acid compositions of avian epidermal proteins. This information was used to construct dendograms that, in association with other biochemical and structural information, provide a basis for speculation on the design and evolution of the epidermal structures of birds.

METHODS

Several statistics may be used to compare the similarities of protein amino acid compositions. These are designated as the difference index (DI) (Metzger et al. 1968), the composition divergence (D) or deviation function (Harris and Teller 1973), and the composition coefficient (SΔQ) (Marchalonis and Weltman 1971, Dedman et al. 1974). The value of the coefficient lies in its ability to estimate similarity in sequence from similarity in composition (Black and Harkins 1977; Woodward 1978; Cornish-Bowden 1978, 1979). Not all the available methods have been evaluated with equal rigor. The DI was assessed independently by Woodward (1978). He compared the distribution of DI values among both related and unrelated protein pairs and concluded that a DI of less than 10 indicated relatedness (e.g. homology), that values over 27 indicated unrelatedness, and that DI's of pairs in the region 10–27 could not be judged reliably. Unfortunately, this presents an ambiguity in an area of intense interest to systematists. These studies conclude that each index has a reasonable predictive value and is a relatively reliable screening test for the detection of protein similarities. A major problem occurs in comparing molecules of different chain length, and various statistical compromises have been provided. In regard to most

TABLE 2. SΔQ values from feather keratin monomers of emu (*Dromaius novaehollandiae*) and gull (*Larus novaehollandiae*) that have been fully sequenced (O'Donnell and Inglis 1974). The average SΔQ value for all available monomers from tern (*Sterna hirundo*) and turkey (*Meleagris gallopavo*) was 86 (Busch and Brush 1979). The average SΔQ for all values in this table is 140.6 ± 41. These values represent minimal values to estimate interspecific differences in presumably homologous monomers.

	Larus	*Sterna*	*Dromaius*	*Meleagris*
Larus novaehollandiae	—	150	109	105
Sterna hirundo		—	181	101
Dromaius novaehollandiae			—	96
Meleagris gallopavo				—

parameters it is difficult to choose among them. The SΔQ value of Marchalonis and Weltman (1971) has been widely applied and is empirically useful.

I applied all these techniques to an extensive sampling of the available amino acid compositions of avian keratins from various structures. Amino acid compositions were obtained from the literature or from analysis in my laboratory. When necessary, data were converted to residues percent (e.g. residues/100 residues) as a common unit value. I attempted to identify problems such as the reproducibility of analysis by comparing data on similar material from different laboratories (Table 1). Technical differences among the various laboratories may be reflected in variation and reproducibility of reported values. All values for amino acid residues were rounded to a single decimal place for calculations.

The SΔQ is basically the sum of the difference between each amino acid squared:

$$S\Delta Q = \Sigma_j(X_{ij} - X_{kj})^2 \quad (1)$$

where X_j is the content of a particular amino acid of type j, and the subscripts i and k identify the particular proteins that were compared. SΔQ gives a direct estimate of the number of sequence differences from composition and is most accurate in comparisons of polypeptides of equal length. It is based on residues/100. The other method of choice is the SΔn statistic, which is more sensitive to differences in polypeptide length (Cornish-Bowden 1977). It is defined as one-half the individual differences among residues squared. In this case the comparison is based on total chain length rather than a normalized value:

$$S\Delta n = \tfrac{1}{2}\,\Sigma(N_{iA} - N_{iB})^2 \quad (2)$$

where N_{iA} and N_{iB} are the numbers of amino acid residues of the ith type in proteins A and B, respectively. The summation is carried out over each type of amino acid distinguished in most composition measurements. A correction factor is applied (Cornish-Bowden 1979; equation 5) if the proteins differ by approximately 20 or more residues. The SΔn is an estimator equal in reliability to SΔQ; indeed the two are interconvertible (see equation 3). The SΔn statistic has probably the most rigorous theoretical development, and is superior in comparisons of proteins of unequal length. There is good agreement between knowledge of how it ought to behave and its function in practice when applied to related proteins. There are not as many empirical tests of the SΔn as exist for the SΔQ. The values are related by the factor:

$$S\Delta Q = \frac{2 \times 10^4\, S\Delta n}{N^2} \quad (3)$$

where N is the total polypeptide length (Cornish-Bowden 1977). It is possible to apply either statistic to a given data set with internal consistency as the molecular weights of the major monomers are known and their distribution is tissue specific (Brush and Wyld in press). The values of the SΔn calculations were used in all studies of tissues with polypeptides of different lengths.

Complete sequence data are available for very few avian keratins. The complete sequences of individual feather keratin monomers (FKM) from a gull, *Larus novaehollandiae,* and Emu, *Dromaius novaehollandiae* (O'Donnell and Inglis 1974), chicken down (Walker and Rogers 1976), and a partial sequence of chicken scale monomer (SKM) (Walker and Bridgen 1976) were compared by diagonal index matching and sequence-nucleotide (REH) comparison (Holmquist et al. 1972, Moore et al. 1976). Comparisons were also made with the same samples using the SΔn statistic (Cornish-Bowden 1977).

TABLE 3. Amino acid composition differences among tissues of *Gallus gallus*, estimated with the SΔn values, which includes an appropriate correction for chain length differences. Values for similar tissues analyzed in different laboratories are included.

	Feather			Down				Scale (scute)			Reticulate scale	Beak	Claw
	a.	b.	c.	d.	e.	f.	g.	h.	i.	j.	k.	l.	m.
a. Feather (Akahane et al. 1977)	0	16.1	8.99	44.1	43.4	6.3	9.2	369.0	333.4	280.5	1,308.2	296.2	313.8
b. Feather (Harrap & Woods 1964)		0	1.5	34.8	34.3	14.8	17.9	307.9	265.5	224.8	1,213.2	230.6	240.8
c. Feather (Kemp & Rogers 1972)			0	30.5	29.0	8.1	11.0	318.6	275.3	223.2	1,220.1	241.2	253.3
d. Adult down				0	25.3	40.0	33.2	302.6	263.1	207.1	1,220.1	229.9	232.2
e. Natal down					0	24.1	37.2	310.7	259.4	208.0	1,072.1	215.8	230.0
f. Natal down (Walker & Rogers 1976)						0	10.5	348.3	299.7	249.7	1,187.9	260.3	286.0
g. Natal down (Kemp & Rogers 1972)							0	351.1	303.7	258.7	1,311.0	268.1	286.0
h. Scale (scute) (Walker & Bridgen 1976)								0	24.2	33.6	587.1	35.0	29.1
i. Scale (scute) (Kemp & Rogers 1972)									0	28.9	551.5	8.5	8.3
j. Scale (scute) (Brush & Wyld 1980)										0	566.3	26.6	23.0
k. Scale (reticulate) (Brush & Wyld 1980)											0	570.2	557.6
l. Beak (Frankel & Gillespie 1976)												0	4.0
m. Claw (Fraser et al. 1972)													0

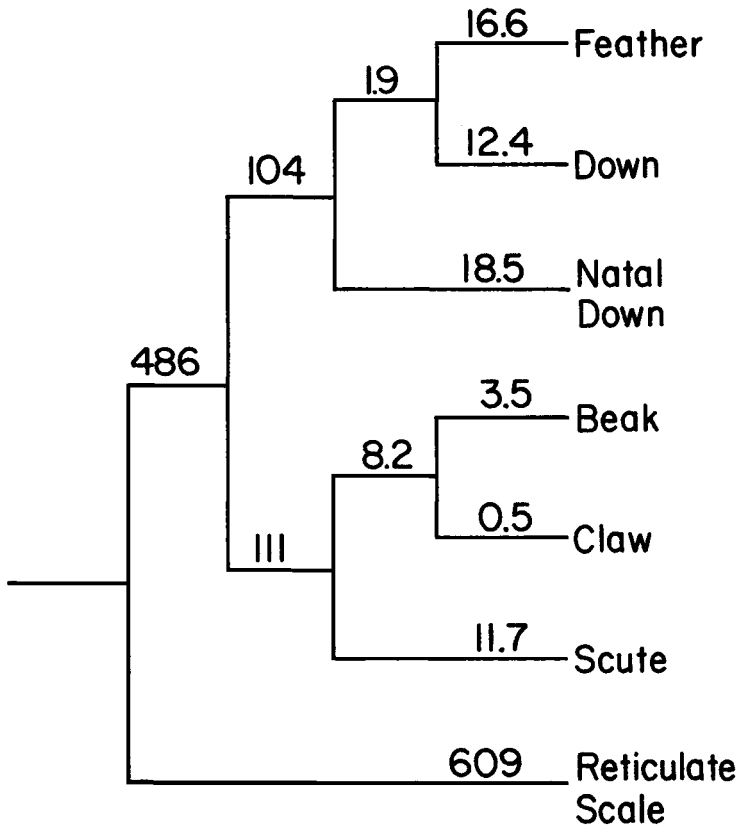

Fig. 1. Dendogram of avian keratin tissues based on amino acid compositional differences as estimated by SΔn. Branching sequence and limb lengths established by Fitch-Margoliash (1967) algorithm. This was the best-fit tree, with SD = 216, and F = 29. The high SD was a result of the large distances between the two tissue groups and the great distance to reticulate scale.

RESULTS

Avian epidermal structures can be categorized by a hierarchical sequence of chemical-morphological organization. Detailed comparisons of the amino acid compositions at each level can supplement other investigations. The degree of similarity in composition reflects both genetic relatedness and functional requirements.

Intraspecific differences in monomers.—The structural details of feather keratin monomers isolated from several species were given by Busch and Brush (1979). The compositional differences were used here as an indication of the range of values of variation within the primary molecular structural element among the available species (Table 1A). The isolated monomers represent the simplest structural element and each is the product of a separate gene. Individual monomers from other tissues showed similar difference values, which implies that sequence differences among monomers account for the electrophoretic and chemical heterogeneity. The chemical fractions, generally isolated by their solubility, are always electrophoretically heterogeneous (Brush in press). The chemical fractions are operational groups defined by specific physical-chemical properties (Table 1A and 1B).

TABLE 4. Comparison of amino acid compositions of adult contour feathers from a variety of species. Values are reported as SΔn. The mean interspecific value was 61.7 for 28 measurements. The average of all feather values to *Gallus* scute was 290 SΔn units. Data from references listed in Table 2; values for *Ardea* and *Tadorna* from this laboratory. Values for *Gallus a* from Akahane et al. (1977) and *b* from Harrap & Woods (1964).

	Sterna	*Melea-gris*	*Gallus a*	*Gallus b*	*Larus*	*Dromaius*	*Ardea*	*Tadorna*
Sterna	0	43.7	53.3	37.5	58.6	114.7	69.6	12.8
Meleagris		0	21.0	3.1	43.6	45.4	93.5	61.7
Gallus a			0	21.0	52.1	33.7	98.7	83.0
Gallus b				0	41.4	56.8	92.0	52.4
Larus					0	56.2	97.0	73.7
Dromaius						0	118.5	118.4
Ardea							0	73.9
Tadorna								0

At the morphological level the elements of the feather, such as rachis, calamus, and barbs, can be separated mechanically, recognized unambiguously, and associated with specific functions. Comparisons of amino acid compositions of feather parts are presented in Table 1B. The values were similar in magnitude to those obtained from chemically produced fractions. This indicates again a molecular heterogeneity associated with structure and, presumably, with functional requirements.

Monomer differences between species.—Complete sequence data are available for isolated monomers of only two species (O'Donnell and Inglis 1974). These can be compared with the presumably homologous monomer isolated from other species (Table 2). The values are greater than within-species differences. Presumably these differences represent sequence divergence in the monomers accumulated since these taxa shared a common ancestor. Thus, differences among monomers of each tissue within species are structurally or functionally derived, while those in homologous monomers in different species reflect genetic divergence.

Tissue differences.—The keratin tissues of an individual take several morphologically distinctive shapes. The most obvious are contour feather, bills, claws, and scales. Further, some represent different phases in development—natal and adult down, for example. The amino acid compositions of these structures can differ significantly (Table 3). In this table the composition values were calculated with the SΔn formula. For the previous data the SΔQ was used. Values for the comparisons of all tissues of *Gallus gallus* are given in Table 3. A dendogram based on these values is presented in Fig. 1. Note the close clustering of the feather-down complex and the small distances among the scale-claw-beak complex. The distance that separates these two clusters are vast by comparison. The reticulate scale is even more distant, which probably reflects the evolutionary origin of this structure (Brush and Wyld 1980).

Interspecific comparison of feather composition.—Average values among comparisons of the same tissue performed in different laboratories give an estimate of the reproducibility of the amino acid composition analysis. The average SΔn value for unfractionate *Gallus* contour feather keratin was 8.8, for down 23.9, and 28.9 for scute. The magnitude of these differences approached those measured among chemical fractions or gross morphological elements.

The only SΔn values available in quantity were for adult contour feathers (Table 4). These values are comparable with previously published SΔQ values, but are considerably more extensive. These values correspond roughly with generally ac-

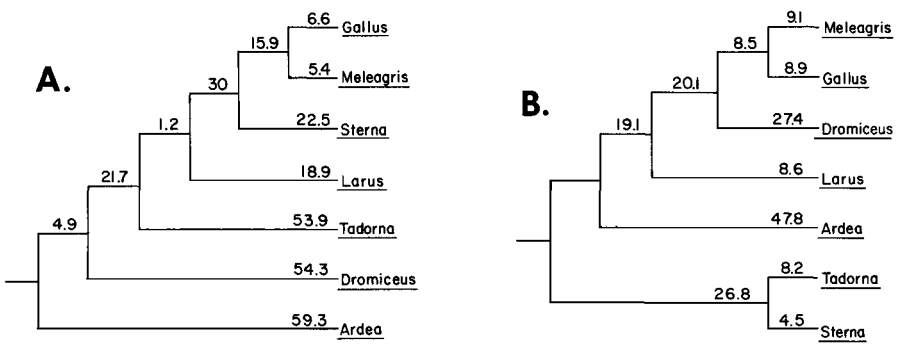

Fig. 2. Phylogeny based on amino acid composition analysis of unfractionated feather keratins. **A.** The Fitch-Margoliash techniques gave a branching order that was more like the generally accepted classification. Tests of "goodness-of-fit" gave an SD = 21, F = 51. **B.** The Farris-Wagner (Farris 1970) algorithm had an SD = 21, F = 17. The origins, significance, and use of these values is discussed by Prager and Wilson (1978).

cepted taxonomic opinion (Fig. 2). The possibility of basing phylogenies on amino acid composition analysis was recently explored (Cornish-Bowden 1979). Clearly, amino acid composition is only one dimension useful in interspecific comparisons. Electrophoretic comparisons of solubilized feather proteins are also of considerable promise (Brush 1976, Knox 1979).

DISCUSSION

The solubilized SCM-keratin monomers of avian epidermal structures can be characterized by their electrophoretic pattern in polyacrylamide gels. When compared under various conditions, the morphologically diverse tissues within an individual have many bands in common, but each contains unique polypeptide units. That is, recognizable tissue-specific patterns exist that involve both common and unique bands and differences in concentration. The distribution and relative amounts of each monomer unit reflect structural and functional differences. There are also species-specific influences that modify the tissue pattern. There is a significant difference between the size of the basic keratin polypeptide in feathers and that in scale, beak, and claw, but, unlike the keratins of the integument of other amniotic vertebrates (Fraser et al. 1972, Wyld and Brush 1979), no significant size heterogeneity of polypeptides occurs within single avian structures. Overall, the keratin structures share many biochemical and organizational properties (Brush and Wyld MS).

In order to establish the existence of a hierarchical sequence of amino acid difference values, I began with the system in which the component molecules were presumed to be most closely related. The isolated SCM-monomers of a single tissue such as feathers or scales within species provided such information (Table 1). There are 8–10 major structural genes, and probably an equal number of minor genes, that produce the feather and scale keratins (Rogers 1978). Comparison of the polypeptides within species indicates the degree of sequence variation among the products of a single genome (Table 1). These data were then expanded to include other fractionation procedures, individual monomers isolated from different species, the unfractionated solubilized proteins from morphologically distinct portions of the

same structure in various species, and the solubilized proteins from different tissues within and among species. Several patterns emerged from the amino acid composition analysis based on these comparisons.

Average differences among the isolated monomers from tissues within species were small and presumably represent a limit of minimal sequence differences. By all criteria they represent groups of closely related proteins but differ among tissues. Presumably this reflects differences in the functional requirements of each structure.

Although the average difference values for monomers varied among the tissues, each monomer in an electrophoretic or chromatographic series was most like its immediate neighbor. Distances between neighbors increased along the series. Simultaneously, the cumulative differences increased as one progressed through the pattern. This supports the hypothesis that the electrophoretic heterogeneity of the elements is based on structural differences and agrees with differences produced by other types of fractionation.

The single available calibration that relates $S\Delta Q$ to actual sequence difference is in the comparison of Band III for Silver Gull and Emu contour feathers (Table 2). The data for monomers of known sequence were compared with bands of similar electrophoretic mobility from other species. The estimated sequence difference (average = 141) was large relative to intraspecific comparisons of isolated feather keratin monomers. Thus, it appears that intraspecific keratin monomers were more alike than the monomer from another species with an identical electrophoretic mobility. The $S\Delta Q$ values indicated that the proteins were homologous. The major differences in sequence are presumably the result of divergence. The second point illustrated by this comparison is that proteins with similar electrophoretic mobilities may indeed have different sequences. This can be a source of confusion in electrophoretic comparison of complex protein mixtures (Brush 1979). Extrapolation of the sequence data indicates an average difference of about nine residues among the monomers within a species. The differences are cumulative, as indicated by matrix analysis along a chromatographic or electrophoretic series. One can conclude that, although differences among the monomers that form a single tissue are minimal, those monomers whose chemical behavior is most similar have smaller differences than polypeptides with greater chemical differences (e.g. position in elution sequence). This inference is supported further by observations based on sequence (Walker and Rogers 1976) and peptide map studies (Busch and Brush 1979). The latter have been especially informative because fine differences are resolved with a minimum of material. The compositional differences imply significant sequence differences among closely related polypeptides.

The conservative nature of change in closely related monomers is supported by the intraspecific comparisons of the various parts of feathers (Table 1B). The barb, rachis, medulla, and calamus within several species have average $S\Delta Q$ values lower than among the isolated feather monomers. These structures were also distinguished by their electrophoretic pattern. Thus, the unfractionated protein mixtures from different morphological parts appear more similar than the fractionated monomers prepared from the whole extract of the same structure. Values for structures from different stages varied intraspecifically. Chick and adult scale, for example, were very similar, while natal down and adult contour feathers differed by a magnitude similar to the interspecific isolated feather keratin monomer values. Pennaceous feather and natal down proteins differ in other parameters such as electrophoretic pattern and solubility as well (Brush in press, Brush and Wyld MS). The implication

is that morphologically different parts within tissues retain a high portion of their structural elements in common. Further, the combined sequence difference among these units is small. The proportional distribution of monomers or related groups of monomers must affect gross compositional differences. Such distributional differences would tend to minimize apparent differences in unfractionated samples. These arguments, and the reproducibility of values for similar structures analyzed in different laboratories, imply that the difference in SΔQ values for tissue monomers may approach the limits of resolution of the analytical system.

The difference values for tissues within species were greater than values for the same tissue between species. This reflects the conservative nature of keratins in terms of taxonomic divergence over time and the specific requirements of function and design of the tissues within a species. The compositional differences for unfractionated solubilized keratin monomers of homologous structures from different species averaged about twice that for the parts within species. The differences among tissues of an individual and the feather generations from the same species were larger. This implies that the tissue-specific differences are related to both functional and morphogenetic differences. The molecular heterogeneity provides a basis for the morphological diversity and functional requirements of the epidermal structures.

This system of analysis is not free of problems. For example, comparisons among tern (*Sterna hirundo*) feather keratins showed that monomers number VI and VII differed by almost 185 SΔQ units, enough to be considered unrelated proteins. Closer examination of the data indicated that the difference was due mainly to the enrichment of a single residue (Gly) in FKM-VII. This monomer is also likely to be the one found in the high-Gly portion of the Zn-Acetate fractionation procedure (Brush in press). If this is the case, then the system is sensitive to differences in individual residues. Enrichment by repeats in the sequence may be taxonomically uninformative but a structural necessity. Similarly, there are technical problems associated with the detection and quantification of specific amino acid residues. Fortunately, none of these occurs in significant numbers in the keratins. It is also apparent that the statistical analyses were sensitive to differences in only a small number of residues. Together, these factors could produce inaccurate comparisons and may account for lowered reproducibility among laboratories. Nevertheless, adequate internal consistencies exist in the data from various laboratories and in patterns of the results to make the data useful. Further, the behavior of fractions in separation by solubility was consistent with the amino acid compositional data. Chain length differences can be accommodated by the SΔn formulation.

The phylogeny of the keratin tissues constructed from amino acid compositional data provides a basis for studies of their relationship (Fig. 1). The values for contour feather and two types of down cluster together. There is little indication that natal down is signficantly different enough to be considered a primitive structure or a precursor to adult feathers. One can only conclude that all feathers appear to be derived structures unique to birds. The "harder" structures, e.g. scale, beak, and claws, cluster together and are relatively distant from the feathers. Again this reflects structural differences related to function. Both these groups are vastly separated from reticulate scale. On the basis of these and additional data, we suggest that the reticulate scale is homologous to the reptilian scale and the other hard tissues and feathers are derived in birds (Brush and Wyld 1980).

Computer analysis of the scanty sequence data show the proteins of different tissues to be distantly related in comparison to polypeptides within structures. Se-

quence changes are mostly the product of conservative nucleotide changes and are not unexpected, given the present models of the evolution of these structures (e.g. Maderson 1972a,b). Neither computer method was used to establish limb lengths in the relationship dendograms. Despite limited data, the hypothesis that homologous morphological structures are more alike at the biochemical level than morphologically different structures was supported.

In summary, it is apparent that tissues considered homologous by morphological criteria have strong compositional similarities. Compositional differences are greater between tissue types within species than between the homologous tissue in other species. I conclude that protein composition is determined largely by the structural and functional requirements of a tissue and less by species differences due to anagenic changes. Future emphasis should be placed on studies of tissue differences rather than phylogenetic relations based on a single factor in a tissue. One remarkable aspect of the keratins is the tremendous organizational flexibility, as shown in the morphological manifestation of structure produced by a group of related gene products whose chemistry and supramolecular interactions are only now being perceived.

Although amino acid compositions have been useful in indicating trends in the structure and evolution of avian keratins, some unanswered questions remain. Composition data are valuable in probing the nature of the structural heterogeneity of polypeptides within tissues. They provide insight into possible relationships and homologies among structures, and the relative contribution of functional requirements and species difference. Finally, they are an important factor in addressing problems in design and evolution of epidermal structures. For example, amino acid compositions were instrumental in demonstrating differences in the protein of reticulate scales and scutes and their relationships to other structures (Brush and Wyld 1980). Further, we have established that the scale and claw are very similar in amino acid composition and may be primitive. Beak proteins are more like claw and scale than feathers, but differ significantly and are probably derived. In each case, some further definitive test is possible. The molecular approach may provide a framework for explaining the function-structure relationships of morphological entities at the protein level. Consistencies exist among the amino acid composition, molecular morphology, chemical fractionation behavior, and other molecular data (e.g. electrophoretic heterogeneity and molecular sizing) that, when considered together, are mutually supportive. These parameters reflect trends in the design and evolution of the structures that relate to traditional questions of structural design, homology, and taxonomic relationships. The heterogeneity and tissue-specific nature of avian epidermal keratins is now well documented. The available data indicate that the proteins have significant consequences for the functional design of tissues. Questions of the relationships of synthesis, distribution, and organization of the proteins in relation to tissue morphology and functional requirements can be approached further on both the molecular and ultrastructural levels.

ACKNOWLEDGMENTS

Ms. April Ladden performed preliminary statistical analysis of the composition data. Ms. Erika Kares has provided continued laboratory support and had considerable input into the phylogenetic reconstructions. Dr. Jean Wyld read and commented on several drafts of the manuscript. Dr. Pam Todd of the Biochemistry Department, University of Melbourne, assisted with the computer comparisons of sequence data. Several anonymous reviewers provided careful, positive criticism; I am particularly grateful for

their time and efforts. Support was provided from a grant from the National Science Foundation (DEB-76-20967). Funding for the amino acid analyses was made available from the University of Connecticut Research Foundation.

LITERATURE CITED

AKAHANE, K., S. MUROZONO, & K. MURAYAMA. 1977. Soluble proteins from fowl feather keratin. J. Biochem. 81: 11–18.

BLACK, J. A., & R. N. HARKINS. 1977. Amino acid compositions and evolutionary relationships within protein families. J. Theoret. Biol. 66: 281–295.

BRUSH, A. H. 1974. Feather keratin—analysis of subunit heterogeneity. Comp. Biochem. Physiol. 48B: 661–670.

———. 1975. Molecular heterogeneity and the structure of feathers. Pp. 901–914 *in* Isozymes IV. (C. L. Markert, Ed.) New York, Academic Press.

———. 1976. Waterfowl feather proteins: analysis of use in taxonomic studies. J. Zool., London 179: 467–498.

——— (Ed.). 1978. Feather keratins. Chemical Zoology, X (Aves). New York, Academic Press.

———. 1979. Comparison of egg-white proteins: effect of electrophoretic conditions. Biochem. Syst. Ecol. 7: 155–165.

———. In press. Chemical heterogeneity in keratin proteins of avian epidermal structures: possible relations to structure and function. *In* The skin of vertebrates (R. I. C. Spearman, Ed.) Symp. Linnean Soc. (London). London, Academic Press.

———, & J. A. WYLD. 1980. Molecular correlates of morphological differentiation: avian scutes and scales. J. Exp. Zool. 212: 153–157.

BUSCH, N. E., & A. H. BRUSH. 1979. Avian feather keratins: molecular aspects of structural heterogeneity. J. Exp. Zool. 210: 39–48.

CORNISH-BOWDEN, A. 1977. Assessment of protein sequence identity from amino acid composition data. J. Theoret. Biol. 65: 735–742.

———. 1978. Interpretation of the difference index as a guide to protein sequence identity. J. Theoret. Biol. 74: 155–162.

———. 1979. How well do amino acid composition comparisons predict sequence similarities between proteins? J. Theoret. Biol. 76: 369–386.

CREWETHER, W. G., R. D. B. FRASER, F. G. LENNOX, & H. LINDLEY. 1965. The chemistry of keratins. Adv. Prot. Chem. 20: 191–346.

DEDMAN, J. R., R. W. GRACY, & B. G. HARRIS. 1974. A method for estimating sequence homology from amino acid composition. The evolution of *Ascaris* employing aldolase and glyceraldehyde-3-phosphate dehydrogenase. Comp. Biochem. Physiol. 49B: 715–731.

FARRIS, J. S. 1970. Methods for computing Wagner trees. Syst. Zool. 19: 83–92.

FITCH, W. M., & E. MARGOLIASH. 1967. Construction of phylogenetic trees. Science 155: 279–284.

FRANKEL, M. J., & J. M. GILLESPIE. 1976. The proteins of the keratin component of birds' beaks. Aust. J. Biol. Sci. 29: 467–479.

FRASER, R. D. B., T. P. MACRAE, & G. E. ROGERS. 1972. Keratins. Their composition, structure and biosynthesis. Springfield, Illinois, C. C Thomas.

———, ———, D. A. D. PERRY, & E. SUZUKI. 1971. The structure of feather keratin. Polymer. 12: 35–56.

———, & ———. 1976. The molecular structure of feather keratin. Proc. 16th Intern. Ornithol. Cong.

HARRAP, B. S., & E. F. WOODS. 1964. Soluble derivatives of feather keratin 1. Isolation, fractionation and amino acid composition. Biochem. J. 92: 8–18.

HARRIS, C. E., & D. C. TELLER. 1973. Estimation of primary sequence homology from amino acid composition of evolutionary related proteins. J. Theoret. Biol. 38: 347–362.

HOLMQUIST, R., C. CANTOR, & T. JUKES. 1972. Improved procedures for comparing homologous sequences in molecules of proteins and nucleic acids. J. Mol. Biol. 64: 145–161.

KEMP, D. J., & G. E. ROGERS. 1972. Differentiation of avian keratinocytes. Characterization and relationships of keratin proteins of adult and embryonic feathers and scales. Biochemistry 11: 969–975.

KNOX, A. G. 1979. Feather protein as a source of avian taxonomic information. Comp. Biochem. Physiol. 65B: 45–54.

MADERSON, P. F. A. 1972a. When? and How?: some speculations on the evolution of the vertebrate integument. Amer. Zool. 12: 159–171.

———. 1972b. On how an archosaurian scale might have given rise to an avian feather. Amer. Natur. 106: 424–428.

MARCHALONIS, J. J., & J. K. WELTMAN. 1971. Relatedness among proteins: a new method of estimation and its application to immunoglobins. Comp. Biochem. Physiol. 38B: 609–625.

METZGER, H., M. B. SHAPIRO, J. E. MOSIMANN, & J. E. VINTON. 1968. Assessment of compositional relatedness between proteins. Nature 219: 1166–1168.

MOORE, G. W., M. GOODMAN, C. CALLAHAN, R. HOLMQUIST, & H. MORSE. 1976. Stochastic *versus* augmented maximum parsimony method for estimating superimposed mutations in the divergence of protein sequence. J. Mol. Biol. 105: 15–37.

O'DONNELL, I. J., & A. S. INGLIS. 1974. Amino acid sequence of a feather keratin from Silver Gull (*Larus novae-hollandiae*) and comparison with one from Emu (*Dromaius novae-hollandiae*). Aust. J. Biol. Sci. 27: 369–382.

PRAGER, E. M., & A. C. WILSON. 1978. Construction of phylogenetic trees for proteins and nucleic acids: empirical evaluation of alternative matrix methods. J. Mol. Evol. 11: 129–142.

ROGERS, G. E. 1978. Keratins viewed at the nucleic acid level. Trends in Biochem. Sci. 3: 131–133.

WALKER, I. E., & J. BRIDGEN. 1976. The keratin chains of avian scale tissue. Eur. J. Biochem. 67: 283–293.

———, & G. E. ROGERS. 1976. The structural basis for heterogeneity of chick down feather keratin. Eur. J. Biochem. 69: 341–350.

WOODWARD, D. R. 1978. Is the "Difference Index" useful for assessment of protein relatedness? J. Theoret. Biol. 72: 732–749.

WYLD, J. A., & A. H. BRUSH. 1979. The molecular heterogeneity and diversity of reptilian keratins. J. Mol. Evol. 12: 331–347.