Molecular Adaptation of a Leaf-Eating Bird: Stomach Lysozyme of the Hoatzin

Janet R. Kornegay, James W. Schilling,† and Allan C. Wilson,*,[1]*

*Department of Molecular and Cell Biology, University of California, Berkeley, and †Protein Structure Facility, University of California, Davis

This report describes a lysozyme expressed at high levels in the stomach of the hoatzin, the only known foregut-fermenting bird. Evolutionary comparison places it among the calcium-binding lysozymes rather than among the conventional types. Conventional lysozymes were recruited as digestive enzymes twice in the evolution of mammalian foregut fermenters, and these independently recruited lysozymes share convergent structural changes attributed to selective pressures in the stomach. Biochemical convergence and parallel amino acid replacements are observed in the hoatzin stomach lysozyme even though it has a different genetic origin from the mammalian examples and has undergone more than 300 million years of independent evolution.

Introduction

Lysozymes (EC 3.2.1.17) are ubiquitous bacteriolytic enzymes found in virtually all animals and normally expressed in macrophages, tears, saliva, avian egg white, and mammalian milk (Imoto et al. 1972; Osserman et al. 1974, pp. 1–641; Jollès and Jollès 1984). There are two major groups in the lysozyme *c* gene family: a conventional type and a calcium-binding type, which appear to have arisen from an ancient gene duplication preceding the divergence of birds and mammals some 300 million years ago (Prager and Wilson 1988; Nitta and Sugai 1989; Dautigny et al. 1991). A conventional lysozyme was recruited independently for stomach function in the ruminants (Dobson et al. 1984; Jollès et al. 1990) and in leaf-eating colobine monkeys (Stewart et al. 1987; Stewart and Wilson 1987). This enables lysis and digestion of the fermentative bacteria as they pass through the gut, which thus prevents the loss of valuable nutrients assimilated by the microorganisms. In this rare demonstration of adaptive evolution at the molecular level, convergent or parallel replacements at specific amino acid positions were identified and attributed to selective pressures in the stomach (Dobson et al. 1984; Jollès et al. 1984, 1990; Stewart et al. 1987; Stewart and Wilson 1987). Biochemically, the stomach lysozymes are characterized by low pH optima, protease resistance, and a bias against arginine residues. In addition, regulatory changes resulting in high levels of expression in the gut and gene duplications that may have facilitated recruitment for a digestive function have occurred (Dobson et al. 1984; Stewart et al. 1987; Stewart and Wilson 1987; Irwin et al. 1989, 1992; Irwin and Wilson 1990; Jollès et al. 1990). The hoatzin (*Opisthocomus hoazin*) has the distinction of being the only known avian foregut fermenter. This neotropical bird has an enlarged crop that contains resident microflora and acts as a fermentative chamber, allowing it to thrive on a diet of leaves (Grajal et al. 1989). Working from the hypothesis that bacteriolytic activity in the true stomach is an integral feature of this type of digestive strategy, we predicted the presence of lysozyme in the proventriculus of the hoatzin.

Material and Methods

Enzyme Extraction and Activity

The first proventriculus tissue from a single bird was provided by M. G. Dominguez-Bello and F. Michelangeli at the Instituto Venezolano de Investigaciones Científicas (IVIC). Crude hoatzin stomach extracts were prepared by mincing ~0.5 g of frozen proventriculus and vortexing in 2–3 volumes of 2% acetic acid + 2% 2-phenoxyethanol, then kept for 12–16 h at 4°C. The solids were thereafter removed by centrifugation and filtration. Extracts prepared in this way were stored at −20°C and thawed repeatedly without any apparent loss in bacteriolytic activity. Enzymatic activity using *Micrococcus luteus* as the substrate was measured with lysoplate and liquid turbidity assays (Dobson et al. 1984; Stewart 1986; Jollès et al. 1990), as well as following gel

1. Deceased July 21, 1991.

Key words: biochemical convergence, calcium-binding lysozyme, foregut fermentation, hoatzin, molecular adaptation, stomach lysozyme.

Address for correspondence and reprints: Janet R. Kornegay, Department of Molecular and Cell Biology, 401 Barker Hall, University of California, Berkeley, California 94720.

Mol. Biol. Evol. 11(6):921–928. 1994.

0737-4038/94/1106-0010$02.00

electrophoresis (see below). The estimated yield based on activity and visualization in polyacrylamide gels was ⩾ 1 mg of lysozyme per gram of starting tissue.

Native Gel Electrophoresis

Nondenaturing polyacrylamide gel electrophoresis at pH 4.3 and pH 8.9, protein detection with Coomassie G250, and detection of lysozyme activity using *Micrococcus luteus* slab gel overlays were performed essentially as described (Dobson et al. 1984; Jollès et al. 1989, 1990).

Purification

Purification of the hoatzin stomach lysozyme from the extracted material described above was done by carboxymethyl-cellulose cation-exchange chromatography (CM-52, Whatman) with an ammonium acetate step gradient. Elution of the lytic activity occurred at pH 5.8 from 25 to 50 mM. Purity of the fractions was assessed by denaturing gel electrophoresis, which showed a slightly retarded mobility for this lysozyme relative to chicken egg-white lysozyme (M.W. 14,400), a phenomenon previously observed for mammalian stomach lysozymes (Jollès et al. 1990).

Amino Acid Sequence

A partial amino acid sequence was obtained from the purified protein using automated Edman degradation on the Applied Biosystems model 477A sequencer employing standard methods and reagents from the manufacturer. The first 49 residues were determined directly from the whole molecule. A lysyl-endopeptidase (Wako) digest was performed and fragments sequenced that confirmed residues 10–49, partially identified residues 67–92, and identified residues 96–124. Amino acid digestion and sequencing were performed by or under the direction of J.W.S. The remaining amino acids were predicted from nucleic acid sequence data generated by the polymerase chain reaction (PCR) (Saiki et al. 1988) with two approaches.

Genomic DNA Sequence

Genomic DNA was extracted from a frozen hoatzin tissue sample provided by the Louisiana State University Museum of Vertebrate Zoology collection (*Opisthocomus hoazin* B-10753), using a proteinase K digestion method for the extraction of high molecular weight DNA (Sambrook et al. 1989, pp. 9.16–9.19). The fragment encoding amino acid residues 39–67 was amplified using degenerate oligonucleotide primers based on the amino acid sequence data from the hoatzin and the most conserved regions of pigeon egg-white lysozyme when compared with other lysozyme sequences (Exon2C and Exon2E; see table 1). The PCR product was cloned into a bacteriophage M13 sequencing vector, and 12 clones were sequenced by the dideoxy chain-termination method.

Table 1
Primers Used for cDNA Synthesis, PCR Amplification, and Sequencing

Name	Sequence (5′–3′)
Exon2C	TTNCCRTCRTYRCACCA
Exon2E	AAYACNGARGCNTACAA
Poly-dT/Adapter	GACTCGAGTCGACATCGAT_{17}
RACE Adapter	GACTCGAGTCGACATCG
Exon2-EcoRI	GATCGAATTCGTCCAAGCAGGGACTAC
Anchor	CGATGAATTCTCGAGTC
Anti-Anchor	GACTCGAGAATTCATCG
AHSLX2-5′	CCCTGCTTGGACCATTGTTG
5′Leader[a]	ATGAGAAAACTCAATGCTCT
3′Flanker[a]	GAGATCAGTGCTGCTTGCGGA

[a] Used for sequencing only.

cDNA Sequence

To complete the amino acid sequence, a fragment encoding the entire carboxy-terminal end of the lysozyme starting at residue 47 was obtained by rapid amplification of cDNA ends (RACE) (Frohman et al. 1988). Briefly, total RNA was isolated from hoatzin proventriculus by guanidine hydrochloride extraction followed by a cesium chloride cushion (Sambrook et al. 1989, pp. 7.19–7.21). For this preparation, frozen tissue from a single bird was used; this animal was collected and dissected in the field and diverse tissues frozen immediately in liquid nitrogen. Total RNA (≈1.5 μg) was treated with DNase and then used to synthesize the first strand of cDNA using an oligo (dT)-adapter hybrid primer (see table 1 for all primer sequences) with BRL Superscript reverse transcriptase. An internal primer (Exon2-EcoRI) was used together with the adapter primer in the PCR carried out according to the following protocol: 1 cycle with 5 min at 94°C, 5 min at 55°C, and 10 min at 72°C; 39 cycles with 40 s at 94°C, 1 min at 55°C, and 3 min at 72°C; and a final incubation at 72°C for 10 min. The double-stranded product was purified by agarose gel electrophoresis and used to generate single-stranded DNA for direct sequencing (Gyllensten and Erlich 1988). In order to verify completely the amino acid sequence determined at the protein level with that predicted from cDNA sequence data, the 5′ end of the stomach lysozyme was amplified using ligation-anchored PCR (Troutt et al. 1992). In this method, the first strand of cDNA was generated with an oligo-dT primer as described above. An anchor oligonucleotide was blocked with dideoxy-ATP using terminal (d)-transferase and then ligated to the single-stranded cDNA at the 5′ end

using T4 RNA ligase (Boehringer Mannheim). PCR was performed using an anti-anchor primer and internal gene-specific primers under the following conditions: 1 cycle with 5 min at 93°C, 1 min at 50°C, and 2 min at 72°C; 35 cycles with 40 s at 93°C, 1 min at 50°C, and 2 min at 72°C; and a final extension of 5 min at 72°C. Double-stranded DNA purification and direct sequencing were performed as described above. The complete cDNA sequence has been submitted to the GenBank™ with accession number L36032.

Phylogenetic Analysis

The tree in figure 4 was constructed from the amino acid sequence data using the PAUP computer package (Swofford 1990). The PROTPARS substitution matrix was used in order to consider the number of nucleic acid substitutions required for each amino acid replacement, and the character states were optimized using the delayed transformation (DELTRANS) option. Gaps in the sequence alignments were each counted as a single event. The root was assigned using the midpoint rooting option in PAUP, and its placement supports separation of the calcium-binding and conventional lysozyme groups as shown. Bootstrap analysis was performed using PAUP with random resampling of the data and parsimony reconstruction reiterated 1,000 times. Winning-sites tests, in which the number of phylogenetically informative sites favoring each of three four-taxon networks are compared, were done with MacClade (Maddison and Maddison 1992) as previously (Prager and Wilson 1988; Dautigny et al. 1991); the scores were tested for significance with a chi-square test. Pairwise minimal mutation distances were used in neighbor-joining phylogenetic analyses with the NJ TREE computer program (Saitou and Nei 1987).

Results

Biochemical Properties of Hoatzin Stomach Lysozyme

Bacteriolytic activity in crude hoatzin stomach extracts was initially demonstrated using standard *Micrococcus luteus* lysoplate and liquid turbidity assays. The results of nondenaturing gel electrophoretic analyses under contrasting pH conditions with bacteriolytic activity overlays (fig. 1) are consistent with those observed for mammalian stomach lysozymes (Jollès et al. 1990). Subsequent to these initial findings, an independent group has also reported bacteriolytic activity in hoatzin proventriculus (Ruiz et al. 1994). Purification by cation-exchange chromatography and denaturing gel electrophoresis (data not shown) confirm a putative hoatzin stomach lysozyme with a low pH optimum, a less basic surface charge, and a predicted molecular weight consistent with other vertebrate lysozymes *c*. In the cathodal

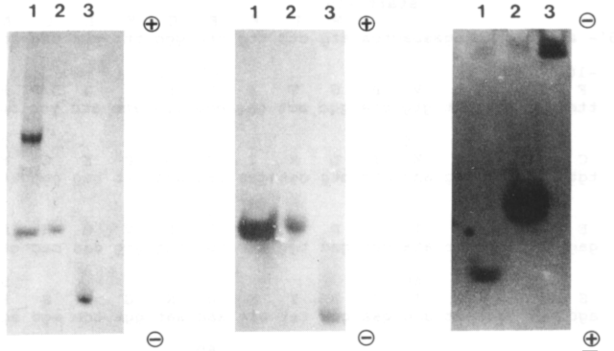

FIG. 1.—Native gel electrophoresis of hoatzin stomach lysozyme. *Left panel,* Nondenaturing gel electrophoresed in the cathodal direction at pH 4.3 and stained for protein. Lanes 1–3 contain, respectively, 5 µl of crude hoatzin stomach extract, 1.5 µg of purified axis deer stomach lysozyme, and 2 µg of chicken lysozyme. *Middle panel,* Activity overlay of the pH 4.3 gel shown to the left in which the lysozymes can be identified by the dark zones of lysis. *Right panel,* Activity overlay of a pH 8.9 nondenaturing gel electrophoresed in the anodal direction. Samples in each lane are identical to those shown in the left and middle panels. The chicken lysozyme does not migrate in the anodal direction at pH 8.9, as shown by the lytic zone in the well of lane 3.

direction at pH 4.3, the hoatzin stomach extract yields a clearly separated bacteriolytic band with a mobility similar to axis deer stomach lysozyme, while chicken lysozyme migrates about 35% further. At pH 8.9, where chicken lysozyme still migrates toward the cathode, the bacteriolytic activity in the hoatzin extract migrates further in the anodal direction than does axis deer lysozyme, which indicates a surface charge even less basic than most of the ruminant stomach lysozymes analyzed under these conditions (Jollès et al. 1990). Consistent with this observation is an isoelectric point of around 6 estimated for the hoatzin stomach lysozyme using previous assumptions (Jollès et al. 1984, 1989, 1990)—that is, a value below the 6.7 computed for axis deer lysozyme. The low pH optimum of a stomach lysozyme is characterized by a narrow activity profile with a peak near pH 5 and was inferred for hoatzin lysozyme from its predicted isoelectric point and an increase in lytic zones for the gel overlays of pH 4.3 native gels compared to those for pH 8.9 gels (fig. 1).

Sequence Data

The complete cDNA and amino acid sequences of the hoatzin stomach lysozyme are shown in figure 2, and the protein sequence has been aligned with several other vertebrate lysozymes for comparative analyses (fig. 3). Conserved features that are critical for the fundamental structure and function of lysozymes are present in the hoatzin sequence (e.g., the eight half-cystine residues and catalytic residues glutamic acid 35 and aspartic acid 53). Unusual features of this lysozyme *c* sequence

```
                    start -19
                                 M   L   F   F   G   F   L   L   A
5'- actgcactatgagaaaactca atg ctc ttc ttc ggc ttt ctt ctg gcc

  -10                                   -1  +1
   F   L   S   A   V   P   G   T   E   G   E   I   I  S/P  R
  ttc ctc tcg gct gtg cca ggc act gag gga gaa atc atc ycc cga

                  10                                       20
   C   E   L   V   K   I   L   R   E   H   G   F   E   G   F
  tgt gag ttg gtg aag atc ctg cgt gag cat ggc ttt gag ggc ttc

                                      30
   E   G   T   T   I   A   D   W   I   C   L   V   Q   H   E
  gag ggc aca acc ata gct gac tgg atc tgc ctg gtg caa cac gag

                  40                                       50
   S   D   Y   N   T   E   A   Y   N   N   N   G   P   S   R
  agc gat tac aac acc gaa gcg tat aac aac aat ggt cca agc agg

                                      60
   D   Y   G   I   F   Q   I   N   S   K   Y   W   C   N   D
  gac tac ggg atc ttt cag atc aac agc aag tac tgg tgt aat gat

                  70                                       80
   G   K   T   S   G   A   V   D   G   C   H   I   S   C   S
  ggc aag acc agt gga gcc gtc gac ggc tgc cac atc agt tgc tca

                                      90
   E   L   M   T   N   D   L   E   D   D   I   K   C   A   K
  gag ctt atg aca aat gat ctt gag gay gat att aag tgt gcc aag

                 100                                      110
   K   I   A   R   D   A   H   G   L   T   P   W   Y   G   W
  aag att gct aga gat gct cac ggc ctc act ccc tgg tac ggc tgg

                                     120
   K   N   H   C   E   G   R   D   L   S   S   Y   V   K   G
  aaa aac cat tgc gag ggc aga gac ctg agt tcc tac gtc aag ggc

  126
   C   *
  tgc taa atgactccgcaagcagcactgatct -3'
```

FIG. 2.—Complete cDNA sequence of hoatzin stomach lysozyme obtained via the RACE and ligation-anchored PCR methods. The predicted amino acid sequence, including the putative signal peptide, is shown above the nucleic acid data. Amino acids are numbered above, with position +1 being the beginning of the mature protein sequence. The signal peptide is numbered backward from the cleavage site to the methionine that marks the translation start site at position −19. A partial 3′-untranslated region following the termination codon is also shown. Boldface type highlights the two apparent heterozygosities in the cDNA sequence. One of these results in an amino acid difference (proline to serine) at position 4, which is discussed in the text.

include an amino acid deletion at position 50 and the unique occurrence of a glutamic acid residue in the first position, rather than lysine. The amino-terminal lysine was until now considered a highly conserved residue in lysozymes, participating in structural bonding interactions (Imoto et al. 1972); however, glutamic acid is occasionally found at the N terminus of the related α-lactalbumins (Prager and Wilson 1988; Nitta and Sugai 1989; Dautigny et al. 1991). The global properties of hoatzin stomach lysozyme in comparison with other stomach and non-stomach lysozymes are summarized in table 2. Hoatzin lysozyme exhibits the low arginine content postulated to confer increased resistance to inactivation by fermentation products such as diacetyl and to proteolytic digestion by trypsin (Stewart and Wilson 1987; Jollès et al. 1989). This bias against arginines is also illustrated by the low ratio of arginine to lysine residues. Of interest, pigeon lysozyme exhibits a relatively low ratio of arginine to lysine as well, but the total number of these basic residues is almost twice that found in hoatzin stomach lysozyme. Hoatzin lysozyme contains three additional histidine residues in comparison to the pigeon enzyme, perhaps to compensate in part for the reduced number of other basic residues. However, this avian stomach lysozyme exhibits the general loss of basicity in stomach lysozymes that is reflected in their calculated isoelectric points.

The complete cDNA data revealed two apparent heterozygosities (fig. 2). One of these, at amino acid position 4, results in two possible amino acids (proline and serine), which are predicted to be present in approximately equal amounts on the basis of band intensities in the cDNA sequence. There are two possibilities to account for this result: allelic differences at a single locus or expression of at least two different genes in the stomach of the hoatzin. Molecular genetic data for hoatzin lysozyme genes indicate the existence of at least three genes that in sum encode the two stomach lysozyme species (Kornegay 1994). Serine was noted as a variant at position 4 in the first direct amino-terminal sequencing experiment, consistent with the cDNA sequence data shown in figure 2. Subsequent protein sequencing, as well as the multiple gene data for these lysozymes, indicated a predominance of proline at that position. For simplicity, we therefore chose proline rather than serine (or a heterozygous position) for the sequence alignment in figure 3 and the phylogenetic analysis.

Evolutionary Analysis

Despite the biochemical similarities to mammalian stomach lysozymes, a phylogenetic analysis clearly shows that the hoatzin enzyme falls among the calcium-binding

```
                  10        20        30        40        50        60
Hoatzin  EIIPRCELVKILREHGFEGFEGTTIADWICLVQHESDYNTEAYNNNG-P-SRDYGIFQINSKYWC
Pigeon   KD..........R......V.K.V.N.V...K...G.R.T.F....-.N..........S...
Horse    KVFSK...AHK.KAQEMD..G.YSL.N.V.MAEY..NF..R.F.GKNANG.S...L..L.N.W..
Cow      KVFE....ART.KKL.LD.YK.VSL.N.L..TKW..S...K.T.Y.PSSE.T.........GW..
Langur   K.FE....ART.KKL.LD.YK.VSL.N.V..AKW..G.....T.Y.PGDE.T.........E...
Human    KVFE....ART.KRL.MD.YR.ISL.N.M..AKW..G...R.T.Y.AGDR.T.........R...
Chicken  KVFG....AAAMKR..LDNYR.YSLGN.V.AAKF..NF..Q.T.R.T-DG.T....L.....RW..

                  70        80        90        100       110       120      130
Hoatzin  NDGKTSGAVDGCHISCSELMTWDLEDDIKCAKKIARDAHGLTPWYGWKNHCEGRDLSSYVKGC--
Pigeon   .....R.SKNA.N.N..K.RDDNIA...Q.......E.R.....VA..KY.Q.K.....R.....--
Horse    K.N.-RSSSNA.N.M..K.LDENID...S...RVV..PK.MSA.KA.VK..KDK...E.LAS.NL
Cow      .....PN..D...V......EM.IAKAVA...H.VSE-Q.I.A.VA..S..RDH.V....E.....TL
Langur   .N...P...DA......A.LQWNIA.AVA...RVVS.PQ.IRA.VA.R...QNK.V.Q..K..GV
Human    .....P...NA..L...A.LQDNIA.AVA...RVV..PQ.IRA.VA.R.R.QN..VRQ..Q..GV
Chicken  ...R.P.SRNL.N.P..A.LSS.ITASVN.....VS.GN.MNA.VA.R.R.K.T.VQAWIR..RL
```

FIG. 3.—Amino acid sequence of the hoatzin stomach lysozyme aligned with six other vertebrate lysozymes. Sequences are shown in the single-letter code, with identities to hoatzin indicated by dots and length variations due to deletion or insertion of amino acids denoted by dashes (—). Boldface type highlights those residues identified as adaptive for stomach function (table 2). Sequences shown are from pigeon (*Columba livia*) egg white, horse (*Equus caballus*) milk, cow (*Bos taurus*) stomach 2, langur (*Presbytis entellus*) stomach, human (*Homo sapiens*), and chicken (*Gallus gallus*) egg white. The occurrence of indels accounts for the slightly different residue numbering systems here and in fig. 2.

lysozymes (fig. 4). This result implies that the hoatzin stomach lysozyme arose from a different branch of the lysozyme gene family from the mammalian stomach lysozymes, separated by over 300 million years of independent evolution. Of the 130 amino acid positions in the alignment of figure 3, 98 are variant and 70 of those variant sites are phylogenetically informative for a parsimony analysis. The grouping of hoatzin with pigeon is statistically significant in this parsimony bootstrap analysis ($P < 0.05$) and also by winning-sites tests of four-taxon networks using hoatzin, pigeon, chicken, and either human or horse ($P < 0.001$). Pairwise sequence comparisons are shown in table 3. The extent of amino acid difference observed between the hoatzin stomach and pigeon egg-white lysozymes is around 30%, a value comparable to that observed between the ruminant and primate lysozymes. In contrast, the average amino acid difference between the calcium-binding and conventional types is around 50%, and the amino acid difference between chicken and hoatzin lysozymes is nearly 60%.

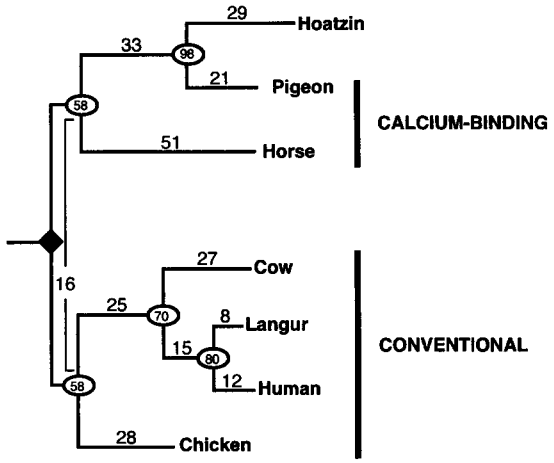

FIG. 4.—Phylogenetic relationship of the hoatzin stomach lysozyme to the other vertebrate lysozymes listed in fig. 3. The large diamond indicates an ancient gene duplication event that occurred at least 300 million years ago, preceding the divergence of birds and mammals (Dautigny et al. 1991). This event gave rise to the conventional lysozymes, which include the mammalian stomach lysozymes, and to the calcium-binding lysozymes. Numbers on lineages are the assigned branch lengths in terms of inferred replacement changes in a parsimony analysis for a total tree length of 265 events. Numbers within circles are the percentage of times the branches were found in 1,000 bootstrap replicates. The neighbor-joining analysis using minimal mutation distances resulted in the same topology as the one shown here.

Discussion

A distinct genetic origin for hoatzin lysozyme lowered the probability that one would observe the specific adaptive amino acid replacements predicted from the mammalian stomach lysozymes. However, this stringent evolutionary comparison has actually allowed us to suggest additional important adaptive structural changes and to weed out irrelevant ones. Seven "convergent" residues were implicated in the comparison of langur and cow lysozymes (Stewart et al. 1987; Stewart and Wilson 1987). Two of these, leucine 17 and serine 101, can be removed from the list because not only are both absent from hoatzin lysozyme, but both are present in chicken egg-white lysozyme and in other bird and mammal lysozymes for which a digestive role has not been proposed. It has already been suggested that serine 101 is less important for stomach function because this replacement was not observed in additional ruminant ex-

Table 2
Structural Adaptations Evident in Hoatzin Stomach Lysozyme

Characteristic	Lysozyme Type			
	Hoatzin Stomach	Mammalian Stomach	Chicken Egg-White	Pigeon Egg-White
Low pH optimum	+	+	−	−
Isoelectric point	~6	6.2–7.7[a]	11.2	~10.6
Total arginines	5	3–6	11	10
Arginine-to-lysine ratio	0.63	0.27–0.67	1.83	0.77
Adaptive residues:				
E/K14	+	+	−	−
E/K21	+	+	−	−
D75	+	+	−	−
N87	+	+	−	−
E/K126	+	+	−	−

NOTE.—Mammalian stomach lysozymes used in this comparison were from three true ruminants (cow, goat, and axis deer) and the langur monkey. Adaptive replacements are summarized in the single-letter amino acid code and numbered according to fig. 3. Properties of chicken and pigeon lysozymes are shown to represent nonstomach lysozymes.

[a] The isoelectric point for langur lysozyme is somewhat higher (Stewart et al. 1987; Stewart and Wilson 1987).

Table 3
Pairwise Amino Acid Sequence Differences among Vertebrate Lysozymes

Lysozymes Compared	1	2	3	4	5	6	7
1. Hoatzin	...						
2. Pigeon	38	...					
3. Horse	79	66	...				
4. Cow	60	64	71	...			
5. Langur	61	61	65	32	...		
6. Human	66	60	64	41	18	...	
7. Chicken	76	68	67	59	54	52	...

NOTE.—A deletion or addition at any position was counted as an amino acid difference.

amples (Jollès et al. 1989, 1990). At position 50 in the alignment (fig. 3) there is a gap in the hoatzin lysozyme rather than the glutamic acid found in other stomach lysozymes. The four remaining convergent replacements are included among the five adaptive residues summarized in table 2. Two of these replacements, aspartic acid 75 and asparagine 87, occur in the hoatzin as well as in the mammals. We can include the predicted replacements at positions 14 and 21 and add one at position 126 by making the assertion that glutamic acid at each of those three positions is as adaptive for stomach function as is lysine. Position 14 is, in fact, occupied by glutamic acid in most ruminant stomach lysozymes other than those of the cow (Irwin and Wilson 1990; Jollès et al. 1990). The residue at position 126 is a glutamic acid in nearly all ruminant stomach lysozymes and the hoatzin and is a lysine in the langur.

The previous comparisons of the langur and cow enzymes identified lysine residues as the primary adaptive alternative to arginine at specific sites (Stewart et al. 1987; Stewart and Wilson 1987); hence, an adaptive role for glutamic acid was not recognized. According to structural models based on chicken lysozyme, residues at these three positions (14, 21, and 126) are located on the surface of the enzyme and would therefore be exposed to the acidic, aqueous environment of the stomach. These positions are often occupied by arginine in nonstomach lysozymes, which would render the enzyme susceptible to inactivation due to covalent modification by digestive products such as diacetyl. The less basic nature and lower pH profile of stomach lysozymes suggest that fewer basic residues are required for optimal function. Because the glutamic acid residues are hydrophilic and will be in a neutral, protonated state, it is likely that they function at least as well as lysines at these three positions. In other words, the two adaptive requirements exhibited at those three sites, hydrophilicity and nonarginine, can be met by either glutamic acid or by lysine. An equivalent acceptability of glutamic acid and lysine is exhibited at additional positions of lysozyme as well. Examples can be observed at positions 1, 41, 83, and 117 in the alignment presented in figure 3.

In contrast to the hoatzin, pigeon lysozyme, which is phylogenetically closest to hoatzin lysozyme (fig. 4), lacks all of the adaptive residues and has 10 arginines and a computed isoelectric point nearly as high as that of the chicken enzyme. The hoatzin lineage has accumulated eight more replacement changes than the pigeon lineage, suggestive of accelerated evolution that can facilitate a biochemical recruitment event. Such an acceleration has been observed previously in the mammalian stomach lysozymes. Although not reflected in our representative phylogenetic analysis in figure 4, previous analyses that included clusters of more closely related lysozyme sequences showed that the langur stomach lysozyme lineage has accumulated twice as many replacement changes as lineages leading to non-leaf-eating monkeys (Stewart et al. 1987; Stewart and Wilson 1987; Jollès et al. 1990; Swanson et al. 1991; E. M. Prager, personal communication). Similar analyses demonstrated that the ruminant stomach enzymes underwent a period of accelerated evolution followed by a pronounced slowing down (Jollès et al. 1989, 1990; Irwin et al. 1992). Testing the possibility of an acceleration along the lineage leading to hoatzin stomach lysozyme thus must await the availability of lysozyme sequences from birds more closely related to the hoatzin. It is possible that other replacements on the hoatzin lineage may be identified as adaptive once additional avian calcium-binding lysozymes with nondigestive functions have been examined.

Conventional lysozymes have been thoroughly studied, while relatively little is known about the calcium-binding ones, of which pigeon egg-white lysozyme is the only completely characterized avian example (Rodríguez et al. 1985, 1987; Nitta et al. 1988). It is imaginable that calcium binding provided additional structural stability to lysozyme along the lineage leading to the hoatzin, which facilitated its recruitment as a digestive enzyme. The small amount of purified hoatzin

lysozyme available after amino acid sequencing procedures precluded direct assessment of calcium binding; however, a contemporary functional role for calcium is questionable because the hoatzin enzyme lacks an aspartic acid at position 86 (fig. 3), a key ligand in the putative calcium-binding loop (Stuart et al. 1986), and the pH measured in the hoatzin proventriculus is 2.1 (Grajal et al. 1989), considerably lower than the presumed pKs of the three aspartic acid residues that would bind calcium only when in a deprotonated state.

Given lysozyme's small size, the identification of a combination of five amino acid replacements unique to the evolution of stomach lysozymes from completely different genetic origins is significant. Comparative biochemistry has demonstrated that sequence evolution is predominantly divergent, even in cases of functional convergence. Stewart et al. (1987) discussed the a priori likelihood that sequence convergence in the mammalian stomach lysozymes would accompany the structural and regulatory convergent evolution, and they suggested that there may be a limited number of ways to convert a conventional mammalian lysozyme to a stomach lysozyme. It appears that there is indeed a set of specific structural changes that are necessary and sufficient to convert a lysozyme for stomach function even when different genes have been separated by more than 300 million years of independent evolution. Just as this independently derived adaptation in the hoatzin has provided a fruitful comparison with the mammalian data, examples from a broad phylogenetic base will continue to refine our understanding of the adaptive evolution of biochemical systems.

Acknowledgments

We thank I. Matsumura and E. M. Prager for technical advice; D. M. Irwin, T. D. Kocher, A. Sidow, W. K. Thomas, and especially E. M. Prager for valuable discussion and comments; all individuals who generously provided hoatzin tissues; and M. G. Dominguez and F. Michelangeli for assistance in collection of additional samples. This work received support from a National Institutes of Health grant to A.C.W. and a National Science Foundation predoctoral fellowship to J.R.K. Travel support for collection of hoatzin tissues was provided to J.R.K. by the American Ornithologists' Union and the Frank Chapman Memorial fund.

LITERATURE CITED

DAUTIGNY, A., E. M. PRAGER, D. PHAM-DINH, J. JOLLÈS, F. PAKDEL, B. GRINDE, and P. JOLLÈS. 1991. cDNA and amino acid sequences of rainbow trout (*Oncorhynchus mykiss*) lysozymes and their implications for the evolution of lysozyme and lactalbumin. J. Mol. Evol. **32**:187–198.

DOBSON, D. E., E. M. PRAGER, and A. C. WILSON. 1984. Stomach lysozymes of ruminants: I. Distribution and catalytic properties. J. Biol. Chem. **259**:11607–11616.

FROHMAN, M. A., M. K. DUSH, and G. R. MARTIN. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA **85**:8998–9002.

GRAJAL, A., S. D. STRAHL, R. PARRA, M. G. DOMINGUEZ, and A. NEHER. 1989. Foregut fermentation in the hoatzin, a neotropical leaf-eating bird. Science **245**:1236–1238.

GYLLENSTEN, U. B., and H. A. ERLICH. 1988. Generation of single-stranded DNA by the polymerase chain reaction and its application to direct sequencing of the *HLA-DQA* locus. Proc. Natl. Acad. Sci. USA **85**:7652–7656.

IMOTO, T., L. N. JOHNSON, A. C. T. NORTH, D. C. PHILLIPS, and J. A. RUPLEY. 1972. Vertebrate lysozymes. Vol. 7 *in* P. D. BOYER, ed. The enzymes, 3d ed. Academic Press, New York.

IRWIN, D. M., E. M. PRAGER, and A. C. WILSON. 1992. Evolutionary genetics of ruminant lysozymes. Anim. Genet. **23**:193–202.

IRWIN, D. M., A. SIDOW, R. T. WHITE, and A. C. WILSON. 1989. Multiple genes for ruminant lysozymes. Pp 73–85. *in* S. J. SMITH-GILL and E. E. SERCARZ, eds. The immune response to structurally defined proteins: the lysozyme model. Adenine, Schenectady, N.Y.

IRWIN, D. M., and A. C. WILSON. 1990. Concerted evolution of ruminant stomach lysozymes. Characterization of lysozyme cDNA clones from sheep and deer. J. Biol. Chem. **265**:4944–4952.

JOLLÈS, J., P. JOLLÈS, B. H. BOWMAN, E. M. PRAGER, C.-B. STEWART, and A. C. WILSON. 1989. Episodic evolution in the stomach lysozymes of ruminants. J. Mol. Evol. **28**:528–535.

JOLLÈS, J., E. M. PRAGER, E. S. ALNEMRI, P. JOLLÈS, I. M. IBRAHIMI, and A. C. WILSON. 1990. Amino acid sequences of stomach and nonstomach lysozymes of ruminants. J. Mol. Evol. **30**:370–382.

JOLLÈS, P., and J. JOLLÈS. 1984. What's new in lysozyme research? Mol. Cell. Biochem. **63**:165–189.

JOLLÈS, P., F. SCHOENTGEN, J. JOLLÈS, D. E. DOBSON, E. M. PRAGER, and A. C. WILSON. 1984. Stomach lysozymes of ruminants: II. Amino acid sequence of cow lysozyme 2 and immunological comparisons with other lysozymes. J. Biol. Chem. **259**:11617–11625.

KORNEGAY, J. R. 1994. Evolution of avian lysozymes. Ph.D. diss. University of California, Berkeley.

MADDISON, W. P., and D. R. MADDISON. 1992. MacClade: analysis of phylogeny and character evolution. Version 3.0. Sinauer, Sunderland, Mass.

NITTA, K., and S. SUGAI. 1989. The evolution of lysozyme and α-lactalbumin. Eur. J. Biochem. **182**:111–118.

NITTA K., H. TSUGE, K. SHIMAZAKI, and S. SUGAI. 1988. Calcium-binding lysozymes. Biol. Chem. Hoppe-Seyler **369**: 671–675.

OSSERMAN, E. F., R. E. CANFIELD, and S. BEYCHOK, eds. 1974. Lysozyme. Academic Press, New York.

PRAGER, E. M., and A. C. WILSON. 1988. Ancient origin of lactalbumin from lysozyme: analysis of DNA and amino acid sequences. J. Mol. Evol. **27**:326–335.

RODRÍGUEZ, R., L. MENÉNDEZ-ARIAS, G. GONZÁLEZ DE BUITRAGO, and J. G. GAVILANES. 1985. Amino acid sequence of pigeon egg-white lysozyme. Biochem. Int. **11**:841–843.

RODRÍGUEZ, R., L. MENÉNDEZ-ARIAS, G. GONZÁLEZ DE BUITRAGO, and J. G. GAVILANES. 1987. Structure of the pigeon lysozyme and its relationship with other type *c* lysozymes. Comp. Biochem. Physiol. **88B**:791–796.

RUIZ, M. C., M. G. DOMÍNGUEZ-BELLO, and F. MICHELANGELI. 1994.Gastric lysozyme as a digestive enzyme in the hoatzin (*Opisthocomus hoazin*), a ruminant-like folivorous bird. Experientia **50**:499–501.

SAIKI, R. K., D. H. GELFAND, S. STOFFEL, S. J. SCHARF, R. HIGUCHI, G. T. HORN, K. B. MULLIS, and H. A. ERLICH. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487–491.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

SAMBROOK, J., E. F. FRITSCH, and T. MANIATIS. 1989. Molecular cloning. 2d ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

STEWART, C.-B. 1986. Lysozyme evolution in Old World monkeys. Ph.D. diss. University of California, Berkeley.

STEWART, C.-B., J. W. SCHILLING, and A. C. WILSON. 1987. Adaptive evolution in the stomach lysozymes of foregut fermenters. Nature **330**:401–404.

STEWART, C.-B., and A. C. WILSON. 1987. Sequence convergence and functional adaptation of stomach lysozymes from foregut fermenters. Cold Spring Harb. Symp. Quant. Biol. **52**:891–899.

STUART, D. I., K. R. ACHARYA, N. P. C. WALKER, S. G. SMITH, M. LEWIS, and D. C. PHILLIPS. 1986. α-Lactalbumin possesses a novel calcium binding loop. Nature **324**:84–87.

SWANSON, K. W., D. M. IRWIN, and A. C. WILSON. 1991. Stomach lysozyme gene of the langur monkey: tests for convergence and positive selection. J. Mol. Evol. **33**:418–425.

SWOFFORD, D. L. 1990. PAUP: phylogenetic analysis using parsimony, version 3.0g. Illinois Natural History Survey, Champaign.

TROUTT, A. B., M. G. MCHEYZER-WILLIAMS, B. PULENDRAN, and G. J. V. NOSSAL. 1992. Ligation-anchored PCR: a simple amplification technique with single-sided specificity. Proc. Natl. Acad. Sci. USA **89**:9823–9825.

MITIKO GO, reviewing editor

Received December 10, 1993

Accepted July 13, 1994