

Evolutionary Shift in the Site of Cleavage of Prelysozyme*

(Received for publication, August 30, 1985)

Lois S. Weisman‡, Barbara M. Krummel, and Allan C. Wilson

From the Department of Biochemistry, University of California, Berkeley, California 94720

Sequences are presented for the signal peptides of prelysozymes from 6 species of birds and compared to the known sequence for chicken prelysozyme *c*. The sequencing was done with synthetic oligonucleotides as primers and oviduct mRNA as the template, obviating the need to clone DNA from these species. Ring-necked pheasant prelysozyme *c* differs from all other prelysozymes *c* and pre- α -lactalbumins examined by being cleaved *in vivo* between amino acid residues 17 and 18 instead of between residues 18 and 19. The feature unique to the signal peptide of pheasant prelysozyme *c* is proline at position 17. Besides showing that proline is acceptable as the carboxyl-terminal amino acid of the signal peptide, our finding implies that it cannot occur as the penultimate amino acid in the signal peptide. This supports the view that unless a polypeptide has the proper secondary structure, signal peptidase will not cleave it, and that this secondary structure is a β -turn. Another outcome of this comparative study is an estimate that the mean rate of sequence evolution in the prelysozyme signal peptide is 1%/two million years of divergence, similar to that calculated for the insulin signal peptide. Because this rate is a third of the silent substitution rate, it is likely that one out of every three amino acid substitutions is compatible with signal peptide function.

The nature of signal peptidase specificity remains elusive. This enzyme, which removes the signal peptide from a preprotein as it is being translocated across a membrane, differs from other proteases in one important respect. Although many proteases exhibit a secondary specificity that extends to areas away from the cleavage site, the primary specificity resides in the amino acids surrounding that site (1). For signal peptidase, the primary specificity appears to include amino acids away from the cleavage site (2, 3). The residue preceding the cleavage site may be any one of six small amino acids, and the other side of the cleavage site can accommodate almost any residue. Despite this observation, the action of signal peptidase must be specific. As secreted proteins diverge over evolutionary time, the site of signal peptidase cleavage rarely changes; furthermore, there have been no reports of a signal peptidase acting upon a site within the mature portion of any secreted protein. Although the signal peptides of substrate proteins vary widely in length and amino acid sequence (4), it is notable that, in analogy with protein synthesis, components (including signal peptidase) from a heterologous system

will work together to process secretory proteins *in vitro*. In an *in vivo* example, rat preproinsulin, cloned into *Escherichia coli*, is correctly processed and secreted as proinsulin (5).

Mutations of preproteins which affect processing by signal peptidase fall into two categories. First, mutations in or near the signal peptide can alter the rate at which it is cleaved or prevent cleavage completely (6-8). These mutations give insight into the specificity of signal peptidase. The second type of mutation shifts the site of cleavage. Unfortunately, although this kind of mutation is expected to be equally informative, it is not obvious how to select for it, and few mutations of this type have been identified. A comparative approach is, therefore, warranted.

This paper describes a naturally occurring mutant of the latter type. Amino-terminal sequences are now known for 43 lysozymes *c*, and 42 of these begin with lysine. In ring-necked pheasant lysozyme, a glycine precedes the lysine (9). Comparison of the amino acid sequence of chicken prelysozyme with pheasant lysozyme suggested that the extra glycine may be due to a shift in the signal peptidase cleavage site, even though the original cleavage site, Gly-Lys, is still present. By comparison of the signal peptides of prelysozyme from several species of birds we were able to determine that a single amino acid change in the signal peptide can account for the altered processing of pheasant lysozyme. These sequences also contain information about the rate of signal peptide evolution, which is relevant to the functional constraints operating on them.

In order to elucidate the amino acid sequence of the signal peptide, our approach was to sequence lysozyme mRNA directly using oligonucleotide primers and reverse transcriptase. This strategy, which avoided both cDNA cloning and extensive purification of prelysozyme mRNA, greatly simplified the sequence determination of signal peptides from several bird species.

EXPERIMENTAL PROCEDURES

Birds—Laying hens were obtained from the following sources: domestic chicken (*Gallus gallus*) from W. Olsen, University of California, Berkeley; ring-necked pheasant (*Phasianus colchicus*) from Webb's Ranch, Livermore, CA; chukar partridge (*Alectoris graeca*), Japanese quail (*Coturnix coturnix*), and domestic turkey (*Meleagris gallopavo*) from Dr. H. Abplanalp, University of California, Davis. Pekin duck (*Anas platyrhynchos*) oviducts were from the Reichardt Duck Farm, Petaluma, CA.

Isolation and Amino-terminal Sequencing of Pheasant Lysozyme from a Single Egg—Lysozyme was purified from the total white (11.8 ml) of a single egg, as described by Prager and Wilson (10), including the Sephadex G-50 (Pharmacia) chromatography step. As a final step the lysozyme was applied to a column of Bio-Gel P-6 (1.5 × 24 cm), eluted with 20 mM formic acid, and lyophilized. The purified lysozyme was reduced with dithiothreitol and alkylated with iodoacetamide (11), after which it was rechromatographed on Bio-Gel P-6. The reduced and alkylated protein was sequenced by automated Edman degradation on a Beckman 890C sequencer, and the phenylthiohydantoin were identified by reverse-phase high pressure liquid chromatography.

* This research was supported by National Institutes of Health Grant GM-21509 to A. C. W. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Supported in part by American Cancer Society, California Division, Postdoctoral Fellowship J-55-82.

Immunoblot Analysis of Oviduct Extracts for Lysozyme and Prelysozyme—Oviducts were excised from both laying pheasants and laying chickens and immediately frozen in liquid nitrogen. One gram of tissue from each oviduct was thawed, rinsed thoroughly in 0.9% NaCl, and homogenized in 100 mM Tris-HCl, pH 7.5, buffer containing 10 mM EDTA, 10 mM EGTA¹, and the following protease inhibitors: 10 mM *o*-phenanthroline, 2 μ M leupeptin, 2 μ M pepstatin, 2 μ M aprotinin, and 2.5 mM phenylmethanesulfonyl fluoride. A range of 1–50 μ g of protein from chicken oviduct homogenate and 10–500 μ g of protein from pheasant were added to sample buffer (12) at 95 °C, heated for 5 min more, and then loaded on a 15% SDS-polyacrylamide gel (12). Chicken lysozyme standards ranging from 24 to 96 ng were also applied. After electrophoresis the samples were transferred to nitrocellulose (13). Lysozyme antigen was detected with a sandwich technique (13) using a rabbit antiserum to chicken lysozyme and goat antibodies to rabbit IgG; the goat antibodies were covalently linked to horseradish peroxidase (Bio-Rad).

Choice of Duck Oviduct Which Produces One Allele of Lysozyme—Multiple alleles of lysozyme *c* have been sequenced from the Pekin duck (14). To ensure that the mRNA sequenced coded for only one allele, 26 oviducts were screened by subjecting the cell extracts to electrophoresis in a 10% polyacrylamide gel at pH 4.3 and detecting the lysozymes with an activity overlay (15). For mRNA sequencing an oviduct which produced only lysozyme B (10, 14) was chosen.

Purification of mRNA from Oviduct—Total RNA was isolated from freshly excised oviduct using 5 M guanidinium thiocyanate (Fluka) (16) and purified by centrifugation through a 5.7 M CsCl pad (17). When the RNA was not isolated immediately, the freshly excised oviducts were frozen in liquid nitrogen and stored at –80 °C. The frozen oviducts were ground under liquid nitrogen with a porcelain mortar and pestle, and total RNA was extracted with 4 M guanidinium thiocyanate and phenol at 60 °C (18). Poly(A⁺) RNA was isolated from total RNA by chromatography on oligo (dT)-cellulose (Collaborative Research) (19).

Synthesis and Purification of Oligonucleotide Primers—Automated oligonucleotide synthesis (Applied Biosystems) was performed by the phosphite triester method (20). The oligonucleotides were purified either by electrophoresis through a 20% polyacrylamide gel or by thin layer chromatography on silica gel (20).

Sequencing the 5' End of Lysozyme mRNAs by Primer Extension—Amino acids 13–17 (Met-Lys-Arg-Met-Gly) of ring-necked pheasant lysozyme formed the basis for the synthesis of a mixture of primers 14 nucleotides long. In order to reduce the number of oligonucleotides in the mixture to eight, two of the six arginine codons, AGR, were not included since CGT is the codon utilized at that position in chicken (21). A new single primer, 5' GCTGCCAGCTCACATC 3', was synthesized based on the pheasant sequence determined. This primer was satisfactory for sequencing the 5' ends of all the bird lysozyme mRNAs.

RNA sequencing was performed by modification of the dideoxynucleotide chain termination method (22) for DNA sequencing, with reverse transcriptase in place of the Klenow fragment of DNA polymerase I. The RNA was heated for 3 min at 90 °C and then cooled to 0 °C immediately before annealing. Approximately 2.5 μ g of poly(A⁺) with 1–2 pmol of primer was used in each chain termination reaction. Annealing was performed at 42 °C for 30 min. Primer extension was done in a buffer containing 20 mM Tris-HCl, pH 8.6, 7.5 mM KCl, 7.5 mM MgCl₂, and 2 mM dithiothreitol. Radioactively labeled dCTP, 20 μ Ci of [α -³²P]dCTP (Amersham Corp.), was added at a final concentration of 0.083 mM. The final concentration of the other deoxynucleotides was 0.83 mM. Dideoxynucleotides were present at the following concentrations²: 0.5 mM ddGTP, 0.7 mM ddATP, 1.0 mM ddTTP, and 0.013 mM ddCTP. Reactions (6 μ l) were initiated with 10 units of reverse transcriptase (Seikagaku, America) and incubated at 42 °C for 10 min. All deoxynucleotide concentrations then were brought to 2 mM and the reactions allowed to continue at 50 °C for another 10 min. The reactions were stopped by the addition of 4 μ l of formamide, heated at 90 °C for 3 min, and 6- μ l samples were subjected to electrophoresis through a 9% polyacrylamide gel containing 8 M urea. Any empty lanes were loaded with the reaction

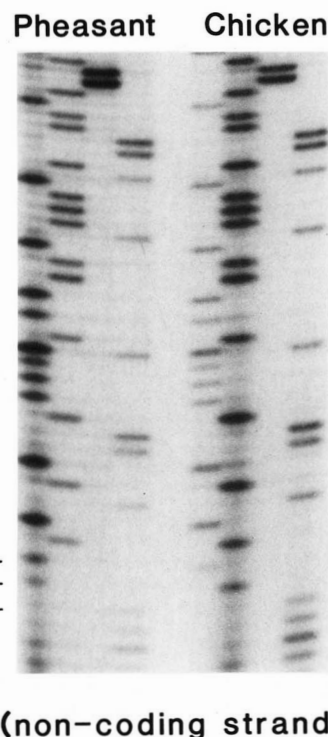


FIG. 1. Autoradiogram illustrating sequencing of the cDNA transcribed from the 5' end of chicken and pheasant lysozyme mRNAs. The portion of the sequencing gel corresponding to base 83 at the bottom (for numbering see Fig. 2) through base 43 at the top is shown. The sequencing reactions were loaded from left to right in the following order: G, A, T, C. The sequence read from the cDNA is complementary to the mRNA. The only base change in the signal peptide-coding region occurs at position 79, where A in chicken corresponds to a G in pheasant. (There are two anomalous features in the gel. First, bases 83–80 show some shadowing in the darkest lanes, G for pheasant, A for chicken, caused by a premature stop at base 84. Second, proceeding from base 70 to 64, one would read AGGG and then G or C followed by A. From previous sequencing of chicken lysozyme genomic DNA (24) and from amino acid sequencing of the signal peptide (25), we know the sequence must read AGGGGCA and the ambiguity in the gel is a result of G compression.)

mixture buffer; this prevented the lanes from converging during electrophoresis.

RESULTS

Amino-terminal Sequence of Lysozyme from an Individual Pheasant—Because ring-necked pheasants exhibit a high level of heterozygosity in the region of the lysozyme gene (23),³ and since sequencing the signal peptide from the nucleic acids will not indicate the site of signal peptidase cleavage, lysozyme from the egg of a single individual was partially sequenced. The first 12 amino acids of the published sequence, with the confirmed residues underlined, are: NH₂-Gly-Lys-Val-Tyr-Gly-Arg-Cys-Glu-Leu-Ala-Ala-Ala. In the Edman degradation, valine was found as the third residue, and no tyrosine was detected at this position. Likewise, the ninth residue was found to be leucine with no contaminating alanine. From these results it was concluded that at least 90% of the lysozyme begins with glycine.

¹ The abbreviations used are: EGTA, ethylene glycol bis(β -aminoethyl ether)-N,N',N',N'-tetraacetic acid; SDS, sodium dodecyl sulfate; UEP, unit evolutionary period.

² In our experience using this sequencing method, appropriate dideoxynucleotide concentrations must be derived empirically for the RNA of interest.

³ The lysozyme gene was mapped in five pheasant individuals with several restriction enzymes. The extent of variation observed between individuals was much greater than that usually seen among individuals of the same species. Helm-Bychowski and Wilson (in preparation) propose that this heterozygosity results from interbreeding of pheasant subspecies.

	10	20	30	40	50	60	70	80
Chicken	agucccgugugugacgacacuggaac	AUG AGG UCU UUG CUA AUC UUG GUG CUU UGC UUC CUG CCC CUG GCU GCU CUG GGG						
Pheasant	C.....a.....							C.....
Turkey	-a.....a.....							
Partridge	a.....				C.....	U.....		
Quail	g.....				G.....	U.....		
Duck	g.....cu.....g.....a.....		A G.....	G C.....C U.....	C.....	A.....		A.....

FIG. 2. Comparison of sequences at the 5' end of bird prelysozyme mRNAs. The sequences are numbered according to the published chicken sequence (24). All the mRNAs except that of the duck were the same length as that of the chicken and demonstrated the same size heterogeneity reported for chicken (24). The duck sequence was longer than the other mRNAs but was not read past the point shown. Dots denote identity with chicken. Dashes indicate portions of the sequence that were not read. The codons are capitalized.

Sequence of the 5' End of Pheasant Lysozyme mRNA—Primer extension with a mixture of 8 oligonucleotides and in the presence of dideoxynucleotides yielded readable sequence starting four bases from the 3' end of the primer (data not shown). A new unique primer, 5' GCTGCCAGCTCACATC 3', was synthesized based on the pheasant sequence read. RNA from the same pheasant whose protein had been sequenced was used. Both chicken and pheasant lysozyme mRNA were primed with the same oligonucleotide in the presence of dideoxynucleotides (Fig. 1). Note that reverse transcriptase reads up the message, i.e. in the 3'-to-5' direction, and that the sequence produced is that of the noncoding strand. The chicken and pheasant mRNA sequences are of identical lengths, have the same bases at 79 of the 83 positions examined, and are shown in Fig. 2. The only base change in the signal peptide-coding region, U to C, changes a leucine to a proline at position 17 in prelysozyme (Fig. 3).⁴

Sequences of the 5' End of Other Bird Lysozyme mRNAs—Due to the conservation of sequence in the region of the single primer described above, the 5' end of lysozyme mRNA for turkey, partridge, quail, and duck could be sequenced with the same primer. As is evident from the base sequences presented in Fig. 2 and the corresponding amino acid sequences presented in Fig. 3, all have signal peptide-coding regions of the same length as the chicken. The leucine to proline change observed between chicken and pheasant is not seen in any of the other birds. Turkey, partridge, and quail all have a leucine at position 17 in prelysozyme. Duck has a glutamine instead of a leucine and, as would be predicted by von Heijne's rule (3), the cleavage site is not altered; according to this rule, glutamine is an acceptable amino acid at position -2 in a signal peptide.

Is Processing of Pheasant Lysozyme Slower Than Processing of Chicken Lysozyme?—Being aware that proline has not often been found to occupy position -1 in a signal peptide (3, 26-28), we asked whether processing is so impaired that prelysozyme could be detected in the pheasant oviduct. Homogenates of both pheasant and chicken oviduct were electrophoresed in a polyacrylamide gel and transferred to nitrocellulose, and lysozyme was detected as described under "Experimental Procedures." No prelysozyme was detectable in either oviduct extract, even when 1000 times the amount of extract needed to see lysozyme was applied (data not shown).

Calculation of the Rate of Signal Peptide Evolution—By making the assumption that all signal peptides evolve at a fixed rate which is independent of the structural protein's

	1	5	10	15
Chicken	met arg ser leu leu ile leu val leu cys phe leu pro leu ala ala leu gly			
Pheasant	- - - - -	- - - - -	- - - - -	- - - - -
Turkey	- - - - -	- - - - -	- - - - -	- - - - -
Partridge	- - - - -	- - - - -	- - - - -	- - - - -
Quail	- - - - -	- - - - -	- - - - -	- - - - -
Duck	- lys ala - - - - -	- thr - - - - -	- phe - leu - - - - -	- - - - - gln -

FIG. 3. Comparison of amino acid sequences for the signal peptides for six bird lysozymes. These sequences are inferred from the mRNA sequences. The amino-terminal residue of the signal peptide is numbered 1. By this numbering system, the first amino acid of mature chicken lysozyme would be 19, whereas the first amino acid of mature pheasant lysozyme would be 18.

rate of evolution, one can calculate the evolutionary rate for any translocated protein whose signal peptide sequence is known for more than one organism. In the case of lysozyme we can compare the chicken, pheasant, turkey, partridge, quail, and duck signal peptides. All 15 possible pairs were compared, and the per cent sequence differences were calculated. These values were corrected to account for multiple hits, and the time required for a 1% sequence difference was estimated, as described in Table I. The unit evolutionary period (UEP), the average time required in millions of years for a 1% sequence difference, is 2.3 ± 0.4 (S.E.) for the lysozyme signal peptide. This rate is similar to the rate of change calculated from the same species for the structural part of lysozyme, $2.6 \text{ UEP} \pm 0.2$ (S.E.).

DISCUSSION

Proline at the Cleavage Site—Empirical rules formulated by Perlman and Halvorson (2) and by von Heijne (3, 26) often correctly predict signal peptidase cleavage sites in secreted proteins. Both groups identify a hydrophobic core, which defines a window of about six amino acids in which one expects the residue at position -1 to be small and uncharged, that at the -2 position to be almost anything, and that at the -3 position not to be aromatic, charged, or large and polar.

Two mutations which result in a new cleavage site both support the above model. In one case, when a threonine analogue, β -norhydroxyvaline, is substituted for threonine at position -1 in rat prolactin, the cleavage site is shifted, presumably because the extra carbon in the analogue's side chain makes it too bulky to fit into the signal peptidase active site (30). In the second case, when alanine, amino acid -1 of yeast invertase, is changed to valine, the cleavage site shifts by one amino acid, again presumably because valine is too large to fit into the active site (31).

By sequencing mRNA corresponding to the signal peptides of several bird lysozymes, our work has demonstrated that changing the amino acid at position -2 from leucine in chicken, turkey, partridge, and quail to a proline in pheasant results in a shift in the cleavage site, even though the original cleavage site, Gly-Lys, is still present in the pheasant protein. Our findings suggest that proline cannot be tolerated at position -2 but can be accommodated at position -1 in a

⁴ According to conventional numbering of signal peptides, the leucine in chicken at position 17 occurs at position -2 from the cleavage site, and the proline in pheasant is at position -1. Since this numbering system is too awkward for describing a cleavage site shift, we number the prelysozyme molecules beginning with the first amino acid of the signal peptide. However, when comparing the prelysozyme sequences with other signal peptides we use the conventional numbering.

TABLE I
Rate of evolution of the signal peptide of prelysozyme

Sequences compared ^a	Divergence time ^b	Amino acid differences ^c	Difference	Corrected value ^d	UEP
	million years		%		million years
Chicken <i>versus</i> quail	17	1	5.6	5.7	3.0
Chicken <i>versus</i> turkey	19	0	0	0	
Chicken <i>versus</i> pheasant	19	1	5.6	5.7	3.3
Chicken <i>versus</i> partridge	17	0	0	0	
Chicken <i>versus</i> duck	53	6	33.3	40.5	1.3
Quail <i>versus</i> turkey	19	1	5.6	5.7	3.3
Quail <i>versus</i> pheasant	19	2	11.1	11.8	1.6
Quail <i>versus</i> partridge	17	1	5.6	5.7	3.0
Quail <i>versus</i> duck	53	6	33.3	40.5	1.3
Turkey <i>versus</i> pheasant	11	1	5.6	5.7	1.9
Turkey <i>versus</i> partridge	19	0	0	0	
Turkey <i>versus</i> duck	53	6	33.3	40.5	1.3
Pheasant <i>versus</i> partridge	19	1	5.6	5.7	3.3
Pheasant <i>versus</i> duck	53	6	33.3	40.5	1.3
Partridge <i>versus</i> duck	53	6	33.3	40.5	1.3

^a Sequences shown in Fig. 3.

^b Divergence times based on molecular and fossil data ((23) and K. M. Helm-Bychowski (personal communication)).

^c Amino acid differences out of the 18 residues in the signal peptide.

^d Corrected for multiple hits at an amino acid site (29).

signal peptide.⁵ This suggestion fits with the view that secondary structure is one of the criteria involved in signal peptidase recognition (33–37).

The observation that mature pheasant lysozyme begins with glycine demonstrates that the signal peptidase cleavage site has been shifted but does not rule out the possibility that the new cleavage site is closer to the amino terminus of prelysozyme with a second peptidase acting on lysozyme further along in the secretory pathway. Since examination of the signal peptide sequence in light of the processing rules (2, 3, 26) does not reveal a reasonable alternative site, we consider this possibility unlikely.

Before examining how proline at position –1 can modify the cleavage site selection rules, it is important to examine how well proline is accommodated in the active site of signal peptidase. The prolactin and invertase cleavage site shifts cited above were accompanied by a slower processing rate and the secretion of some unprocessed protein, presumably because the new cleavage site was only partially acceptable. However, amino-terminal sequencing of pheasant lysozyme did not suggest that any of the lysozyme contained a signal peptide.⁶ Also, the level of pheasant lysozyme found in the egg white is very similar to the level of lysozyme found in chicken eggs.⁷

As an alternative approach, chicken and pheasant oviduct extracts were examined by immunoblots of SDS gels. If the peptidase works as poorly as it does on mutant preinvertase or on β -hydroxynorvaline-containing rat preprolactin, then

⁵ Macaque proinsulin may have a shifted cleavage site like pheasant lysozyme. It is reported to have a proline at position –2 (32), but its amino acid sequence was deduced from the sequence of a cDNA clone, and its processing site was decided in analogy with human and rat preproinsulins. Macaque proinsulin should be sequenced to determine whether, as predicted by the pheasant lysozyme case, proline is now found at position –1.

⁶ We are assuming that the isolation procedures (chromatography on CM-cellulose, Sephadex G-50, and Bio-Gel P-6) would not have removed any contaminating prelysozyme. Pheasant prelysozyme contains an extra 17 amino acids in addition to the 130 amino acids it shares with mature lysozyme.

⁷ We recognize that because protein secretion is a multistep phenomenon, it would be possible to have slower than normal translocation and processing rates and still see normal levels of lysozyme in egg white.

one would expect to see 50% of the total lysozyme in the cell as the unprocessed preprotein. Instead, only processed lysozyme was observed in both chicken and pheasant oviduct. By this method, 0.1% of unprocessed lysozyme would have been detectable. Because the tissue was homogenized in buffer containing several protease inhibitors, added to hot 3% SDS, and rapidly loaded onto a gel, it is unlikely that some endogenous proteases cleaved the signal peptide before the sample could be electrophoresed.

That proline is acceptable at position –1 suggests that size of the amino acid side chains is not the only parameter involved in the choice of a cleavage site but that secondary structure is also important. According to Chou and Fasman (38), proline generally disrupts both α -helix and β -pleated sheet structures, and their rules predict an α -helix from amino acid 1 to amino acid 17 in the bird lysozyme signal peptides. The extra proline in the pheasant signal peptide disrupts the predicted α -helix four amino acids earlier. This may be the type of effect responsible for the shift in cleavage site.

Alternatively, the proline may cause the shift by its involvement in a specific structure. Whereas proline usually disrupts α -helices and β -sheets, the secondary structure that proline promotes is a β -turn. Furthermore, 67% of Pro-Gly dipeptides, the sequence at the pheasant lysozyme cleavage site, occur as the middle two residues of β -turns (39). Indeed, the predicted structure of pheasant prelysozyme, according to Chou and Fasman rules (38), does have a β -turn at the cleavage site.

Predictions of signal peptide secondary structure have been reported by many investigators. Rosenblatt *et al.* (34) studied the signal peptide plus seven amino acids into the structural portion of preproparathyroid hormone. They predicted a β -turn at the cleavage site. Other investigators have predicted the secondary structures of many signal peptides by a variety of methods (2, 35, 37), and many of the structures (although not the majority) were predicted to have a β -turn at the cleavage site. In our own analysis of 24 signal peptides (chosen such that only unrelated proteins were examined), a β -turn at the cleavage site was predicted by Chou and Fasman rules in 14 of them. Secondary structure analysis, however, is still not as predictive as one would like (40). Thus, our finding that Pro-Gly works as a cleavage site for signal peptidase adds to these earlier observations and suggests that a β -turn may

always be the polypeptide conformation at the cleavage site. Ultimately, physical analysis of signal peptide conformation in this region will be essential for understanding signal peptidase recognition.

Rate of Signal Peptide Evolution—The lysozyme signal peptide evolves at a high rate (UEP = 2.3 million years), similar to that calculated for the insulin signal peptide (41). This rate conforms with current observations of low constraints on signal peptide structure, *i.e.* that signal peptides vary widely in amino acid sequence. We can calculate an expected maximum rate of evolution for any peptide that is unconstrained functionally. This rate is expected to be equal to the rate of base substitution at the silent sites of codons, which is about 0.7%/million years (42). Converting that rate into change per codon and assuming that on average only 70% of the base substitutions will affect the amino acid sequence, we calculate a UEP of 0.7 million year, which differs from the UEP for signal peptides by a factor of three. From this, we infer tentatively that two out of every three amino acid substitutions in the signal peptide interfere with function.

Aside from examining signal peptides from bird lysozymes, we have also examined signal peptides from four α -lactalbumins (43–45), a related protein which diverged from lysozyme several hundred million years ago (46). Although only one invariant amino acid residue, excluding the initiating methionine, exists in both sets of signal peptides, the cleavage site is at the same position for both lactalbumin and lysozyme. This observation emphasizes how rarely evolutionary shifts occur in the site of cleavage.

Characteristics of Signal Peptidase—The rapid evolution of all signal peptides lessens the possibility that not only has pheasant lysozyme changed but that pheasant signal peptidase is different than most other signal peptidases. To date only two leader peptidases have been identified in *E. coli*. One is specific for the outer membrane lipoprotein (47), and the other has been shown to be active on the periplasmic, inner membrane, and outer membrane proteins tested (48). These results imply that only one or a few leader peptidases in *E. coli* process the several hundred secreted and membrane-bound proteins. Because dog pancreas microsomes can process secreted and membrane-bound proteins from any cell type tested (36), we assume that an analogous situation exists in mammals. It is not clear how one signal peptidase can recognize the many substrates encoded by one genome and further can recognize substrates from a heterologous system. Either signal peptidase itself is constrained not to evolve or the enzyme is very nonspecific. That signal peptidase is a rather nonspecific protease seems unlikely because it cleaves all its substrates once and at a specific site. Furthermore, when purified bacterial leader peptidase is incubated with a preprotein (in the absence of membranes) it cleaves once and properly, even at 1000 times excess (8). It is likely that signal peptidase evolution is highly constrained, in analogy with other systems which interact with many substrates, *e.g.* calmodulin, ribosomes, and histones. In this light it is improbable that pheasant signal peptidase is unique among avian signal peptidases.

Acknowledgments—The generosity of those who helped us obtain birds is gratefully acknowledged. We thank H. Sample and R. D. Cole

for sequencing the amino terminus of pheasant lysozyme, B. Malcolm and J. F. Kirsch for the oligonucleotide primers, and E. M. Prager for the antiserum to chicken lysozyme. Thanks are due also to D. R. Tolan, R. E. Cohen, R. Higuchi, and A. Larson for suggestions on RNA isolation and sequencing; and to E. M. Prager, P. Jollès, K. M. Helm-Bychowski, K. A. Vincent, C.-B. Stewart, and R. Jones for discussion.

REFERENCES

- Fruton, J. S. (1975) in *Proteases and Biological Control* (Reich, E., Rifkin, D. B., and Shaw, E., eds) pp. 33–50, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Perlman, D., and Halvorson, H. O. (1983) *J. Mol. Biol.* **167**, 391–409.
- von Heijne, G. (1983) *Eur. J. Biochem.* **133**, 17–21.
- Watson, M. E. E. (1984) *Nucleic Acids Res.* **12**, 5145–5164.
- Talmadge, K., Kaufman, J., and Gilbert, W. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 3988–3992.
- Boeke, J. D., Russel, M., and Model, P. (1980) *J. Mol. Biol.* **144**, 103–116.
- Lin, J. J. C., Kanazawa, H., Ozols, J., and Wu, H. C. (1978) *Proc. Natl. Acad. Sci. U. S. A.* **75**, 4891–4895.
- Kuhn, A., and Wickner, W. (1985) *J. Biol. Chem.* **260**, 15914–15918.
- Jollès, J., Ibrahim, I. M., Prager, E. M., Schoentgen, F., Jollès, P., and Wilson, A. C. (1979) *Biochemistry* **18**, 2744–2752.
- Prager, E. M., and Wilson, A. C. (1971) *J. Biol. Chem.* **246**, 523–530.
- Konigsberg, W. (1972) *Methods Enzymol.* **25**, 185–188.
- Laemmli, U. K. (1970) *Nature* **227**, 680–685.
- Towbin, H., Staehelin, T., and Gordon, J. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76**, 4350–4354.
- Kondo, K., Fujio, H., and Amano, T. (1982) *J. Biochem. (Tokyo)* **91**, 571–587.
- Hammer, M. F. (1984) Ph.D. thesis, University of California, Berkeley.
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
- Glisin, V., Crkvenjakov, R., and Byus, C. (1974) *Biochemistry* **13**, 2633–2637.
- Feramisco, J. R., Smart, J. E., Burrige, K., Helfman, D. M., and Thomas, G. P. (1982) *J. Biol. Chem.* **257**, 11024–11031.
- Aviv, H., and Leder, P. (1972) *Proc. Natl. Acad. Sci. U. S. A.* **69**, 1408–1412.
- Alvarado-Urbina, G., Sathe, G. M., Liu, W.-C., Gillen, M. F., Duck, P. D., Bender, R., and Ogilvie, K. K. (1981) *Science* **214**, 270–274.
- Jung, A., Sippel, A. E., Grez, M., and Schütz, G. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 5759–5763.
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467.
- Helm-Bychowski, K. M., and Wilson, A. C. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, in press.
- Grez, M., Land, H., Giesecke, K., and Schütz, G. (1981) *Cell* **25**, 743–752.
- Palmiter, R. D., Gagnon, J., Ericsson, L. H., and Walsh, K. A. (1977) *J. Biol. Chem.* **252**, 6386–6393.
- von Heijne, G. (1985) *J. Mol. Biol.* **184**, 99–105.
- Davidson, J. M., Leslie, B., Wolt, T., Crystal, R. G., and Sandberg, L. B. (1982) *Arch. Biochem. Biophys.* **218**, 31–37.
- Lin, Y., and Gross, J. K. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 2825–2829.
- King, M.-C., and Wilson, A. C. (1975) *Science* **188**, 107–116.
- Hortin, G., and Boime, I. (1981) *Cell* **24**, 453–461.
- Schauer, I., Emr, S., Gross, C., and Schekman, R. (1985) *J. Cell Biol.* **100**, 1664–1675.
- Wettkam, W., Groneberg, J., Leineweber, M., Wengenmayer, F., and Winnacker, E.-L. (1982) *Gene* **19**, 179–183.
- Burstein, Y., and Schechter, I. (1978) *Biochemistry* **17**, 2392–2400.
- Rosenblatt, M., Beaudette, N. V., and Fasman, G. D. (1980) *Proc. Natl. Acad. Sci. U. S. A.* **77**, 3983–3987.
- Austen, B. M. (1979) *FEBS Lett.* **103**, 308–313.
- Steiner, D. F., Quinn, P. S., Chan, S. J., Marsh, J., and Tager, H. S. (1980) *Ann. N. Y. Acad. Sci.* **343**, 1–16.
- Garnier, J., Gaye, P., Mercier, J.-C., and Robson, B. (1980) *Biochimie (Paris)* **62**, 231–239.
- Chou, P. Y., and Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol.* **47**, 45–148.
- Zimmerman, S. S., and Scheraga, H. A. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 4126–4129.
- Richardson, J. S. (1981) *Adv. Protein Chem.* **34**, 167–339.
- Li, W.-H., Wu, C.-I., and Luo, C.-C. (1985) *Mol. Biol. Evol.* **2**, 150–174.
- Perler, F., Efstratiadis, A., Lomedico, P., Gilbert, W., Kolodner, R., and Dodgson, J. (1980) *Cell* **20**, 555–566.
- Hall, L., Craig, R. K., Edbrooke, M. R., and Campbell, P. N. (1982) *Nucleic Acids Res.* **10**, 3503–3515.
- Mercier, J.-C., Haze, G., Gaye, P., Petrissant, G., Hue, D., and Boissard, M. (1978) *Biochem. Biophys. Res. Commun.* **85**, 662–670.
- Gaye, P., Hue, D., Raymond, M.-N., Haze, G., and Mercier, J.-C. (1982) *Biochimie (Paris)* **64**, 173–184.
- Wilson, A. C., Carlson, S. S., and White, T. J. (1977) *Annu. Rev. Biochem.* **46**, 573–639.
- Innis, M. A., Tokunaga, M., Williams, M. E., Loranger, J. M., Chang, S.-Y., Chang, S., and Wu, H. C. (1984) *Proc. Natl. Acad. Sci. U. S. A.* **81**, 3708–3712.
- Wolfe, P. B., Silver, P., and Wickner, W. (1982) *J. Biol. Chem.* **257**, 7898–7902.