Hedgehog Patterning Activity:
Role of a Lipophilic Modification Mediated
by the Carboxy-Terminal Autoprocessing Domain

Jeffery A. Porter,* Stephen C. Ekker,*§‖
Woo-Jin Park,*§ Doris P. von Kessler,*§
Keith E. Young,*§ Chien-Huan Chen,* Yong Ma,*
Amina S. Woods,† Robert J. Cotter,†
Eugene V. Koonin,‡ Philip A. Beachy*
*Howard Hughes Medical Institute
Department of Molecular Biology and Genetics
†Department of Pharmacology and Molecular Sciences
The Johns Hopkins School of Medicine
Baltimore, Maryland 21205
‡National Center for Biotechnology Information
National Library of Medicine
National Institutes of Health
Bethesda, Maryland 20894

Summary

Autocatalytic processing mediated by the carboxy-terminal domain of the *hedgehog (hh)* protein precursor (Hh) generates an amino-terminal product that accounts for all known signaling activity. The role of autoprocessing in biogenesis of the *hh* signal has been unclear, since a truncated unprocessed protein lacking all carboxy-terminal domain sequences retains signaling activity. Here, we present evidence that the autoprocessing reaction proceeds via an internal thioester intermediate and results in a covalent modification that increases the hydrophobic character of the signaling domain and influences its spatial and subcellular distribution. We demonstrate that truncated unprocessed amino-terminal protein causes embryonic mispatterning, even when expression is localized to cells that normally express Hh, thus suggesting a role for autoprocessing in spatial regulation of *hh* signaling. This type of processing also appears to operate in the biogenesis of other novel secreted proteins.

Introduction

The *hedgehog (hh)* gene, initially identified and isolated in Drosophila, functions in the patterning of embryonic segments and imaginal discs. Vertebrate *hh* gene products also have been implicated as the primary signals responsible for patterning of a number of embryonic structures, including the neural tube and its derivatives, the axial skeleton, and the appendages. The Drosophila *hh* protein (Hh) is synthesized as a precursor that undergoes autocatalytically mediated internal cleavage between Gly-257 and Cys-258 to generate an ~19 kDa amino-terminal and an ~25 kDa carboxy-terminal domain. The carboxy-terminal domain contains the determinants required for autoprocessing of the precursor

§These authors contributed equally to this work and are listed in alphabetical order.
‖Present address: University of Minnesota Medical School, Department of Biochemistry, 435 Delaware Street SE, Minneapolis, Minnesota 55455.

(Lee et al., 1994; Porter et al., 1995) and the amino-terminal product accounts for signaling activity, as demonstrated using ubiquitously expressed constructs lacking all (Porter et al., 1995) or a portion (Fietz et al., 1995) of the carboxy-terminal domain. These and other similar results in vertebrates raised the general question of whether the carboxy-terminal domain and its associated autoprocessing activity play any significant role in *hh* patterning functions.

The amino-terminal signaling domain remains tightly cell-associated when expressed in cultured cells (Lee et al., 1994; Porter et al., 1995), and in embryos is found predominantly associated with large punctate structures in the basolateral domain of expressing epithelial cells (Taylor et al., 1993; Lee et al., 1994; Tabata and Kornberg, 1994). The carboxy-terminal domain, in contrast, diffuses freely upon secretion in embryos and in cell cultures (Lee et al., 1994). Interestingly, although the amino-terminal signaling domain expressed from a construct lacking all carboxy-terminal domain sequences suffices to carry out *hh* signaling activities in embryos, this truncated protein is not tightly cell-associated when expressed in cultured cells but, instead, is found predominantly in the culture medium (Porter et al., 1995). This observation suggested that the autoprocessing activity of the carboxy-terminal domain influences the cellular localization of the amino-terminal domain and suggested the possibility of a role for autoprocessing in regulation of its distribution.

The relevance of cell surface localization in vertebrates is suggested by in vitro experiments in which treatment of neural plate explants with low concentrations of purified amino-terminal signaling domain of the *Sonic hedgehog* protein (Shh-N) induces motor neurons but not floor plate cells, whereas high concentrations induce floor plate cells at the expense of motor neurons (Roelink et al., 1995). In these explant experiments, induction of floor plate cells by the natural inducing tissue, the notochord, is contact-dependent. The ability to circumvent contact dependence with high concentrations of Shh-N protein suggests that one role for surface association of the signaling domain is to generate large concentration differences betweeen local and distant sites with consequent sharp distinctions between the cell types induced; this idea is consistent with the presence of signaling domain predominantly on the surface of notochord cells and with the formation of embryonic floor plate only in close proximity to the notochord. In Drosophila, *hh* function is required for maintenance of *wingless (wg)* gene expression, which normally occurs within a single cell-wide stripe adjacent to the *hh* stripe. Ubiquitous expression of *hh* causes expansion of *wg* beyond its normal domain of expression, suggesting that restricted diffusion of the *hh* signaling domain from its localized site of synthesis is important in preserving spatially appropriate expression of *wg.* Spatial relationships between targets and inducers thus appear to influence the cellular outcomes of *hh* signaling in Drosophila

and in vertebrate embryos, suggesting that the cell association conferred upon *hh* signaling domains by autoprocessing in cultured cells is likely to have relevance to embryonic patterning.

To gain a better understanding of the consequences of autoprocessing, we investigated its biochemical mechanisms and tested its role in Drosophila embryos. We found that the processing reaction extends beyond simple peptide bond cleavage, proceeding through a thioester intermediate to covalent modification of the amino-terminal signaling domain. This modification dramatically increases the hydrophobic character of the signaling domain, suggesting the operation of a novel lipid modification pathway. The unmodified amino-terminal domain displays abnormal embryonic patterning activity and altered spatial and subcellular distribution, indicating a critical regulatory role for the autoprocessing event. The presence of an *hh*-like autoprocessing domain in proteins encoded by at least six novel nematode genes suggests that the type of autoprocessing observed for Hh represents a more broadly utilized pathway for biogenesis of secreted proteins.

Results

Autoprocessing and Spatial Patterning Activity

To test the role of autoprocessing in Hh patterning activity, we utilized the Drosophila two-element system based on the GAL4 transcriptional activator of yeast (Brand and Perrimon, 1993) to control expression of constructs encoding the following: first, full-length Hh ; and second, Hh-N, a protein truncated immediately following the Gly at the cleavage site. Both constructs previously were shown to encode signaling activity, but these prior experiments involved ubiquitous expression and did not distinguish between the spatial patterning properties of the two proteins. In the current experiments, one element (*en*GAL4) contained GAL4 coding sequences under control of the *engrailed* (*en*) promoter, expressed coincidently with *hh* in segments of the trunk, and the second element contained a GAL4-responsive promoter upstream of sequences encoding either the intact Hh precursor (upstream activating sequence [UAS] Hh) or the truncated Hh-N protein (UAS Hh-N). Homozygotes were mated and the resulting embryos analyzed with respect to *wg* mRNA expression and patterning of ventral and dorsal cuticles, which normally are under Hh control.

Ubiquitous expression of Hh or Hh-N under heat shock promoter control previously was shown to cause expansion of *wg* expression and to generate characteristic changes in dorsal and ventral cuticle patterns (Ingham, 1993; Heemskerk and DiNardo, 1994; Porter et al., 1995). Localized expression of Hh driven by the *en*GAL4 element, in contrast, causes only a small increase in *wg* expression and generates dorsal and ventral cuticle patterns only slightly different from wild type (compare Figures 1B, 1E, and 1H with Figures 1A, 1D, and 1G). These embryos hatch and go on to form essentially normal adults except for wing venation defects (compare Figure 1J with 1K) and supernumerary scutellar bristles (data not shown). Localized expression of Hh-N, in contrast,

causes the following: first, maximal expansion of *wg* expression; second, ventral cuticle defects including a rectangular rather than trapezoidal shape for the denticle belts and a loss of denticle diversity; third, dorsal cuticle defects with loss of tertiary and an expansion in secondary cell types; and fourth, embryonic lethality (Figures 1C, 1F, and 1I). This phenotype is essentially identical to that caused by ubiquitous expression of Hh or Hh-N. Hh and Hh-N under *en*GAL4 control were expressed at comparable levels (see Figures 2 and 3 below), and similar mispatterning effects were observed with three independent Hh-N lines.

Fietz et al. (1995) reported no difference in patterning effects for *en*-driven expression of Hh or a truncated construct that includes all of Hh-N plus 51 additional residues from the carboxy-terminal domain (N+51). We examined the properties of the N+51 protein in cultured cells and found that, in addition to reduced protein levels as reported by Fietz et al. (1995), the N+51 protein also is less soluble and is poorly secreted (data not shown), all of which might account for its reduced activity relative to the precisely truncated Hh-N construct we analyzed.

Autoprocessing and the Distribution of the Amino-Terminal Domain

To investigate further the role of autoprocessing, we visualized by immunofluorescence the localization of amino-terminal epitopes derived from expression of Hh in the Schneider 2 (S2) Drosophila cultured cell line (Schneider, 1972). Confocal microscopy showed that amino-terminal epitopes are prominently localized at the cell surface (data not shown), thus accounting for the tight cell association observed previously by immunoblotting (Porter et al., 1995). The cell surface reactivity was observed either with or without detergent permeabilization. In contrast with cells expressing Hh, cells expressing Hh-N or the N+51 construct showed no surface localization for amino-terminal epitopes (data not shown). These results are similar to those previously reported for amino-terminal protein fragments derived from full-length and truncated constructs from the *Shh* gene (Bumcrot et al., 1995; Roelink et al., 1995). Taken together these experiments indicate that autoprocessing facilitates and is required for the cell surface association of the amino-terminal signaling domain in cultured cells.

To determine whether autoprocessing also influences protein localization in embryos, we examined the distribution of amino-terminal domains derived from Hh and Hh-N constructs using the *en* promoter to drive expression, as described above (see Figure 1). Figures 2A–2C show that considerably higher levels of amino-terminal protein are detected in embryos expressing the UAS Hh and UAS Hh-N elements. Higher magnification views show that in normal embryos reactivity is concentrated within large, punctate structures localized within the basolateral domain of ectodermal cells (Figure 2H, arrowhead), consistent with previous descriptions of embryonic *hh* protein expression (Taylor et al., 1993; Lee et al., 1994; Tabata and Kornberg, 1994). Discrete punctate structures with a basolateral localization were also observed for UAS Hh when expressed under control of *en*GAL4, although the number and staining intensity of
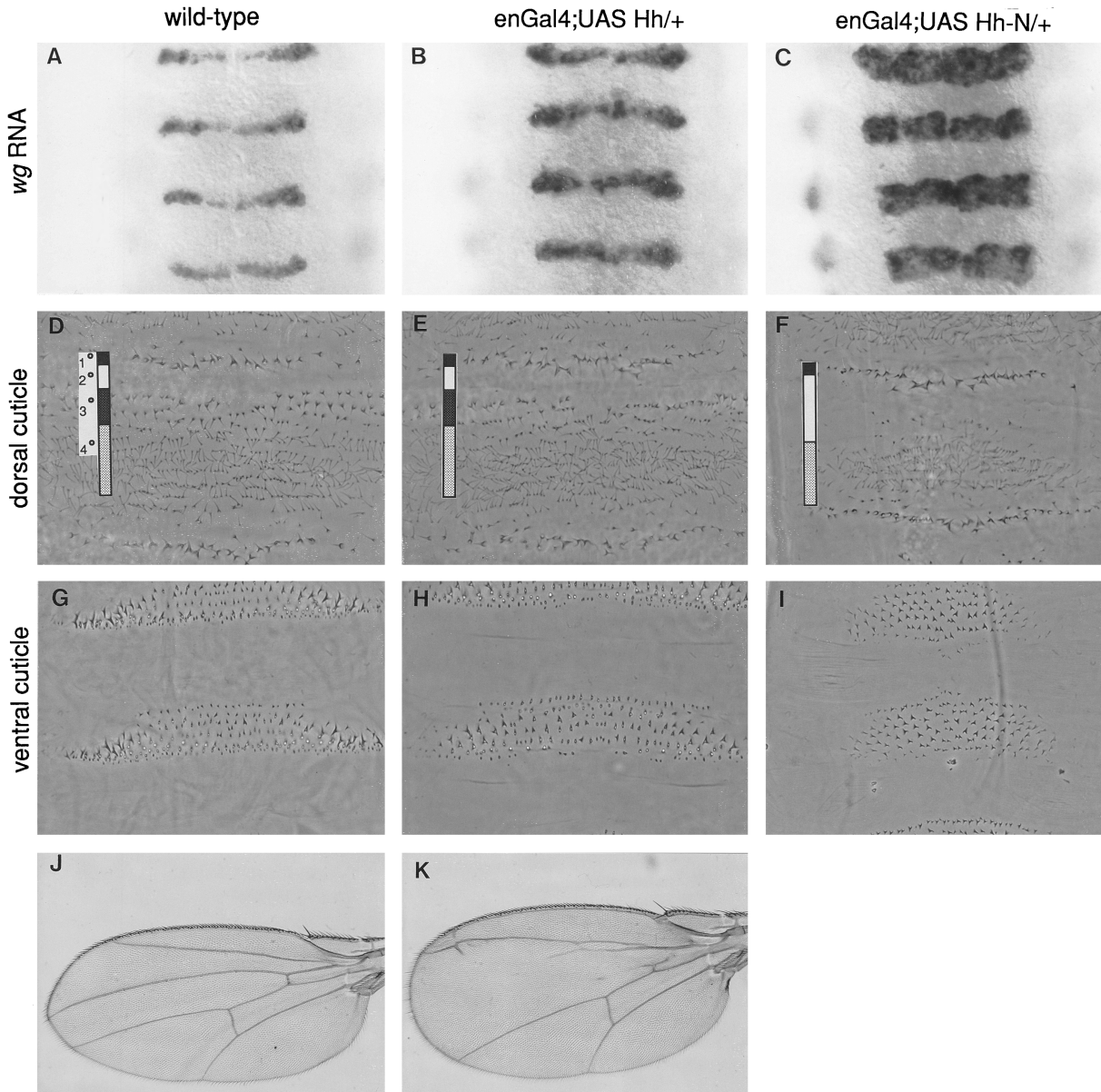
Figure 1. Patterning Effects of Localized Hh and Hh-N Expression

The pattern of *wg* mRNA expression (A–C) and ventral (D–F) and dorsal cuticle patterns (G–I) were analyzed in wild-type embryos (A, D, and G) and in embryos *trans*-heterozygous for the *en*GAL4 element (producing GAL4 under *engrailed* promoter control) and for a GAL4-responsive element containing either the intact *hh* open reading frame (UAS Hh) (B, E, H, and K) or a precisely truncated amino-terminal domain construct encoding residues 1–257 (UAS Hh-N) (C, F, and I). Adult wing structures are shown for wild type and UAS Hh (K and I) but not for UAS Hh-N since the few embryos that hatch (less than 1%) die as larvae.

these punctate structures are considerably higher than observed with the endogenous *hh* gene alone (Figures 2D, 2F, and 2I). In contrast, the UAS Hh-N encoded protein is not found in large punctate structures but is predominantly localized to a more apical domain in a more diffuse distribution (see Figures 2E, 2G, and 2J). The Hh-N protein also appears more broadly distributed with prominent concentrations of signal at locations up to several cell diameters away from the parasegment boundary (Figure 2J). We note that about the same number of punctate basolateral structures appear as typically are seen in wild-type embryos, suggesting that the

reactivity in these structures derives from expression of the endogenous wild-type gene.

Autoprocessing Confers Hydrophobic Character
The differences in spatial patterning properties and subcellular localizations of amino-terminal domain proteins encoded by Hh and Hh-N, shown here in embryos and previously in cultured cells, suggest the possibility of a physical difference between these proteins. To facilitate reference to these two proteins, the amino-terminal domain protein derived by autoprocessing henceforth will be referred to as Hh-N$_p$ and the amino-terminal domain
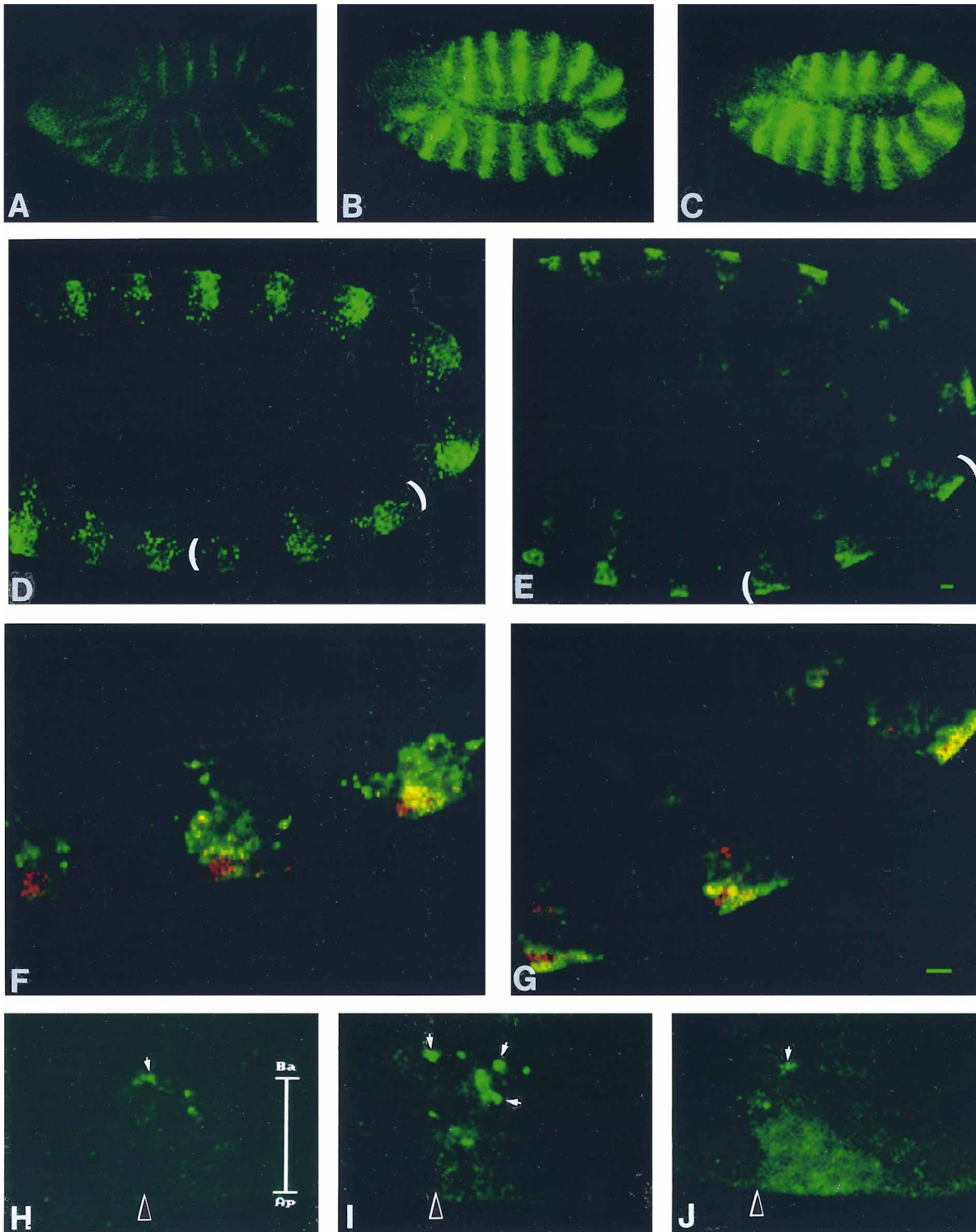
Figure 2. Signaling Domain Localization in Embryos Expressing Hh and Hh-N

Embryos expressing Hh or Hh-N under control of the *engrailed* promoter (see text; Figure 1) were stained for Hh amino-terminal reactivity (all panels, green) and for *engrailed* protein (panels F and G, red) and analyzed by confocal microscopy.

(A)–(C) Lateral views of stage 9/10 embryos showing relative expression levels of amino-terminal reactivity in wild-type control (A), UAS Hh (B), and UAS Hh-N (C) embryos.

(D)–(G) Superimposed confocal images (five lateral 2 μm sections) from embryos expressing Hh (D and F) and Hh-N (E and G). (F) and (G) are enlarged views of the regions indicated in (D) and (E). The nuclei of Hh-expressing cells (F and H, red) are located primarily in the apical region of the cytoplasm, and approximately delimit the apical and basolateral membrane domains. Scale bars in (E) and (G), 10 mm.
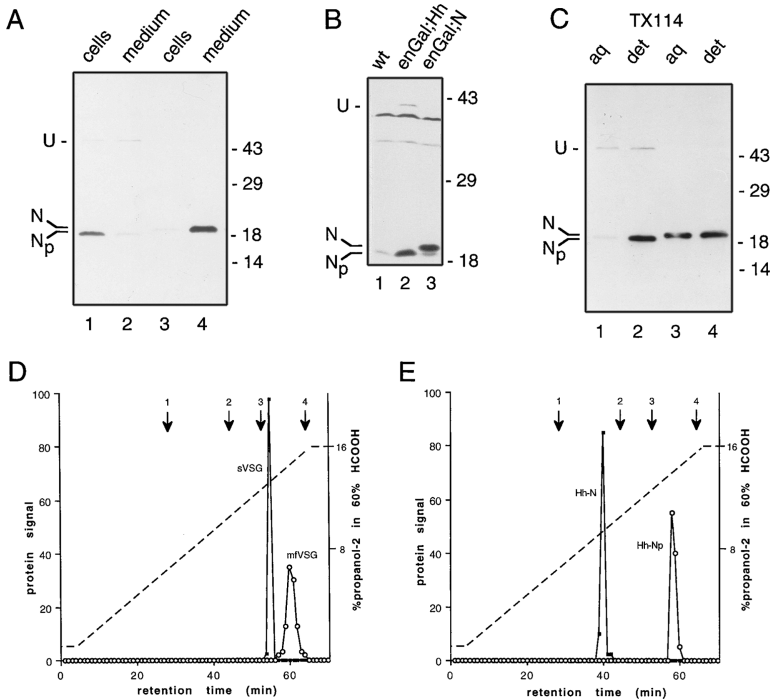
Figure 3. Physical Differences between Processed (Hh-N$_p$) and Unprocessed (Hh-N) Amino-Terminal Domains

(A) Immunoblot of cells and medium from S2 cell cultures expressing Hh-N$_p$ (cell-associated) and Hh-N (in the medium). Note that Hh-N$_p$ migrates faster than Hh-N (also in B and C).

(B) Immunoblot of extracts from wild-type control (lane 1), *en*GAL4/+;UAS Hh/+ (lane 2), and *en*GAL4/+;UAS Hh-N/+ (lane 3) embryos 6–8 hr after egg laying. The majority of protein from the UAS Hh construct is processed.

(C) Immunoblot of detergent (det) and aqueous (aq) phases from Triton X-114 extractions of S2 cell cultures expressing Hh (lanes 1 and 2) or Hh-N (lanes 3 and 4).

(D) Soluble (sVSG) and membrane form (mfVSG) of trypanosome variant surface glycoprotein are eluted with increasing isopropanol in 60% formic acid from a C4 HPLC column. Control proteins of increasing hydrophobic character elute as indicated: 1, cytochrome c; 2, α-lactalbumin; 3, carbonic anhydrase; 4, ovalbumin.

(E) Elution profile of Hh-N$_p$ and Hh-N under identical conditions as (D).

derived from the truncated construct will be designated Hh-N. The existence of a physical difference is supported by immunoblots of extracts from cultured cells (Figure 3A) and from embryos (Figure 3B), which show that under conditions of optimal electrophoretic separation the mobility of Hh-N$_p$ is greater than that of Hh-N. Such differences in mobility are commonly observed for modified proteins. In particular, proteins such as the trypanosome variant surface glycoprotein (VSG) (reviewed by Englund, 1993) display greater mobility when a glycolipid is covalently attached. The faster migration is attributed to an increase in overall SDS binding by the hydrophobic glycosyl phosphatidyl inositol (GPI) anchor in the modified protein as compared with VSG that has been treated with lipase to remove most of the anchor (Cardoso de Almeida and Turner, 1983). GPI-linked VSG is also tightly cell-associated in vivo whereas its lipase trimmed counterpart is freely released into the medium (Cardoso de Almeida and Turner, 1983), suggesting that Hh-N$_p$ may be modified in an analogous manner.

The possibility of a hydrophobic modification of Hh-N$_p$ prompted further studies with Triton X-114, a detergent that forms an aqueous solution at 4°C but separates into hydrophobic and aqueous phases when the temperature is raised to 30°C (Bordier, 1981). Triton X-114 extraction at 4°C followed by phase separation at 30°C–37°C has been used as a general indicator of the hydrophobic properties of proteins, with integral membrane proteins and highly hydrophilic proteins partitioning exclusively into the detergent and aqueous phases, respectively, and with amphiphilic proteins more equally distributed between the two phases. Such

experiments with VSG show that the GPI-linked protein partitions exclusively into the detergent phase whereas the lipase trimmed protein partitions into the aqueous phase (Ausubel et al., 1995, and references therein). Similar experiments with Triton X-114 extracts from S2 cells show that Hh-N$_p$ partitions exclusively into the detergent phase (Figure 3C, lanes 1 and 2) whereas Hh-N partitions evenly between the detergent and the aqueous phases (Figure 3C, lanes 3 and 4).

To characterize further the differences in hydrophobic character between Hh-N and Hh-N$_p$, we examined the behavior of these proteins by a high pressure liquid chromatography (HPLC) method that was established for the fractionation of hydrophobic proteins (Heukeshoven and Dernick, 1985). Figure 3D demonstrates the elution from a C4 HPLC column of four standard proteins in order of increasing hydrophobicity by a gradient of increasing isopropanol concentration. The peak of lipid-modified VSG protein appears 7 min later than that of the unmodified VSG protein (54 min versus 61 min for sVSG versus mfVSG, respectively). The elution difference between Hh-N and Hh-N$_p$ is even greater (Figure 3E), with protein peaks eluting at 40 min and 59 min, respectively. These data clearly demonstrate a markedly increased hydrophobic character for Hh-N$_p$ as compared with Hh-N.

Location and Mass of the Hh-N$_p$ Modification
To examine more directly the physical differences between Hh-N$_p$ and Hh-N, we subjected these proteins to mass spectral analysis (Chevrier and Cotter, 1991; Blackledge and Alexander, 1995) after cyanogen bro-

(H)–(J) High resolution micrographs (obtained with a 100× objective) of cells at the parasegment boundary in wild type (H) and UAS Hh (I) and UAS Hh-N (J) embryos. The apical (Ap) and basolateral (Ba) domains of the ectodermal epithelium are as indicated in (H). Punctate structures are denoted by arrows and the parasegment boundaries by an arrowhead outline.
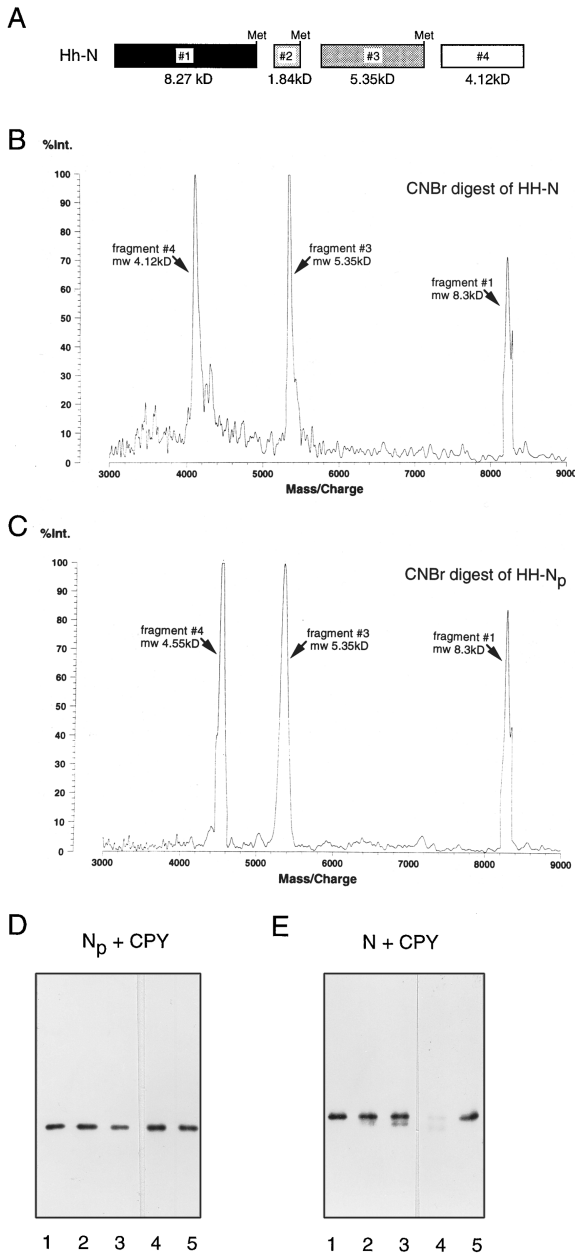
are shown in Figures 4B and 4C. Fragments 1, 2, and 3 were identical for each digest (fragment 2 is not shown). Fragment 4, however, of predicted mass 4.12 kDa, appeared to contain ∼430 daltons of additional mass in digests of Hh-N$_p$. These results are consistent with modification of Hh-N$_p$ as a consequence of the autoprocessing reaction. At present, we cannot rule out the possibility that CNBr digestion alters the modifying group, and we are unable to obtain mass spectrometry data on Hh-N$_p$ in the absence of CNBr fragmentation.

The modifying group associated with Hh-N$_p$ appears to be covalently linked, since the harsh conditions utilized for electrophoresis (high concentrations of SDS and β-mercaptoethanol [β-ME]), for the HPLC method (60% formic acid), and for mass spectral analysis (overnight incubation in 70% formic acid and washing of membrane-bound protein samples in 50% methanol) would be expected to disrupt noncovalent interactions such as hydrogen bonding, hydrophobic or van der Waals interactions, and even certain types of covalent interactions, such as disulfide and thioester linkages. The presence of additional mass in combination with faster electrophoretic mobility in SDS–polyacrylamide gel electrophoresis (SDS–PAGE) (see Figures 3A and 3B) suggest that this modifying group does not consist exclusively of carbohydrates or amino acids, since these modifications on their own would be expected to decrease mobility.

The location of the modifying group from mass spectral analysis appears to be somewhere within the amino acid residues comprised by the carboxy-terminal product of CNBr fragmentation. We tested the sensitivity of these two proteins to digestion by carboxypeptidase Y and observed that Hh-N$_p$ is resistant to digestion (Figure 4D) whereas Hh-N is degraded (Figure 4E). This result confirms the physical difference between Hh-N$_p$ and Hh-N and suggests that the modification occurs on the carboxyl terminus of the amino-terminal domain, thus rendering Hh-N$_p$ resistant to the activity of carboxypeptidase Y. A carboxy-terminal location for the modification is also consistent with the proposed mechanism of modification (see below).

Proposed Chemical Mechanism of Autoprocessing

To gain a better understanding of the regulation and biological consequences of the autoprocessing event, we initiated a biochemical study of its mechanism. Much of what we know about the biochemistry of Hh autoprocessing derives from in vitro studies of a homogeneous bacterially expressed protein containing a hexahistidine purification tag and nine residues from the amino-terminal domain followed by the carboxy-terminal domain in its entirety (His$_6$Hh-C) (see Experimental Procedures; Porter et al., 1995). Previous work had demonstrated that cleavage of this protein in vitro occurs between residues corresponding to Gly-257 and Cys-258 of intact Hh. An intramolecular mechanism for this in vitro reaction was indicated by its concentration-independent kinetics (Porter et al., 1995) and by the inability of exogenously added cleavage site peptides to affect it (data not shown). Similarities of this cleavage reaction

Figure 4. Mass and Location of the Hh-N$_p$ Modification

(A) CNBr fragmentation pattern of an amino-terminal domain protein truncated at the cleavage site. Methionines are marked above and predicted sizes below.

(B) MALDI mass spectral analysis of Hh-N after CNBr digestion. Relative to the predicted values, experimental mass determinations (fragment 2 not shown) are within the potential 0.2% error inherent in this technique.

(C) MALDI mass spectral analysis of Hh-N$_p$ after CNBr digestion. The experimentally determined mass of Hh-N$_p$ fragment 4 exceeds by 430 daltons that obtained for Hh-N.

(D) and (E) Immunoblots of purified Hh-N$_p$ (D) and Hh-N (E) incubated with carboxypeptidase Y. Incubation times were 0, 1, 3, 15, and 15 hr for lanes 1–5, with no carboxypeptidase in lane 5.

mide (CNBr) digestion. The fragments predicted from the amino acid sequence of Hh-N are shown in Figure 4A and mass spectra for the Hh-N$_p$ and Hh-N digests
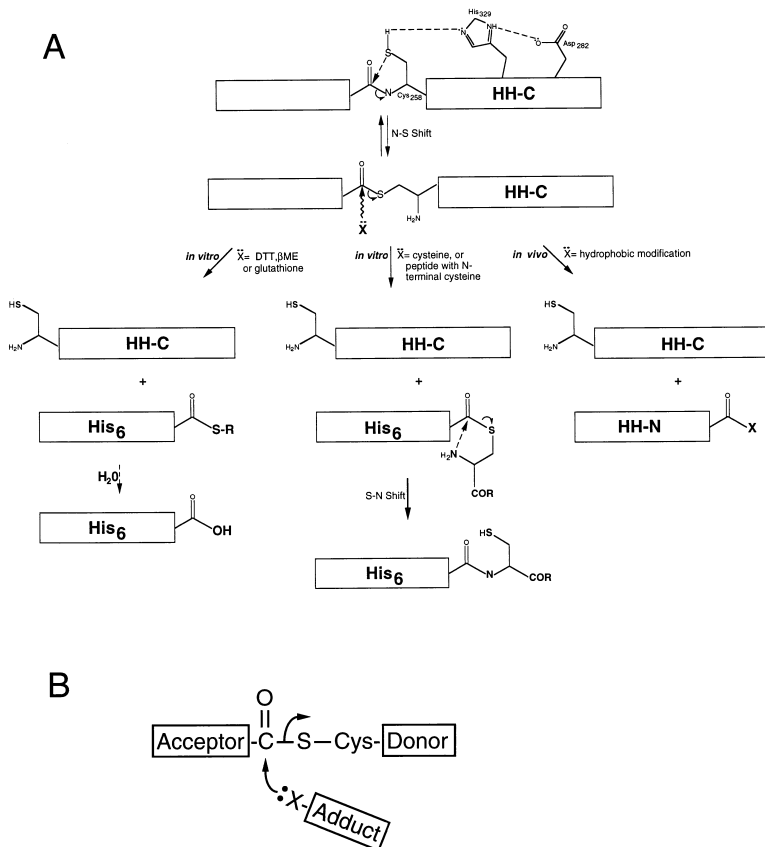
A

Figure 5. Proposed Chemical Mechanism of Hh Autoprocessing

(A) An internal rearrangement converts the peptide bond linking Hh-N and Hh-C (between the Gly-257 and the Cys-258 residues) to a thioester. The three branches away from this thioester intermediate represent the presumed in vitro and in vivo reaction pathways. All involve the displacement of Hh-C by a nucleophilic attack on the carbonyl of Gly-257 where X represents the nucleophile.

(B) A general mechanism for protein modification reactions involving a cysteine-derived thioester intermediate (see Discussion).

B

(see below) to the autoprocessing of prohistidine decarboxylase (proHisDcase) (van Poelje and Snell, 1990) led us to the following mechanistic proposal: the Cys-258 thiol initiates a nucleophilic attack on the carbonyl of the Gly-257 residue, leading to displacement of the amine of Cys-258 and formation of a thioester bond between the Gly-257 carbonyl and the Cys-258 side chain (Figure 5). In the standard in vitro reaction, dithiothreitol is added and is thought to act as a nucleophile for attack of the thioester (Rose and Warns, 1983), resulting in a free carboxy-terminal fragment and an amino-terminal fragment with a dithiothreitol adduct in thioester linkage.

Experimental Studies of the Autoprocessing Mechanism: Thiol Attack and a Thioester Intermediate

If His_6Hh-C cleavage proceeds via Cys-258 attack on the carbonyl of Gly-257 to generate a thioester intermediate, then the reaction would be expected to require the thiol side chain of Cys-258 and should be inhibited by thiol-attacking reagents such as N-ethylmaleimide (NEM). Mutagenesis of Cys-258 to Ala (Figure 6A, lane 3) completely abolishes in vitro cleavage, and pretreatment of His_6Hh-C with NEM (Figure 6A, lane 5) inhibits subsequent autocleavage by ~60%. Furthermore, by analogy to proHisDcase, it might be expected that either a thiol or a hydroxyl could function as the initial attacking group. Indeed, although reaction efficiency is greatly reduced, replacement of Cys-258 by a Ser still results in some cleavage (Figure 6A, lane 4). In addition, we find that

the requirement for dithiothreitol is not absolute, and that the small nucleophile hydroxylamine can substitute (data not shown).

Analysis of the reaction products also supports our model. First, as previously shown, the carboxy-terminal fragment contains an amino-terminal Cys (Porter et al., 1995). Second, mass spectral analysis (Chevrier and Cotter, 1991; Blackledge and Alexander, 1995) of the dithiothreitol-driven reaction reveals a species corresponding to the mass expected from a dithiothreitol adduct in thioester linkage with the carboxy-terminal Gly of the amino-terminal His_6 fragment (corresponding to Gly-257 of Hh) (see Figure 6B). This species is observed in admixture with a species of mass expected from the His_6 fragment with no dithiothreitol adduct, and this lower-mass species is presumed to result from hydrolysis of the dithiothreitol adduct. Third, other thiol-containing molecules such as β-ME, glutathione, and free cysteine can also drive the His_6Hh-C autocleavage reaction (Figure 6A, lanes 8, 9, and 10), and mass spectral analysis of the products resulting from these reactions revealed species of mass expected from thioester-linked adducts to the His_6 fragment (data not shown).

Interestingly, whereas reactions with β-ME and glutathione yielded species corresponding to the hydrolysis product in addition to species containing an adduct, reactions driven by free cysteine yielded primarily a species of mass indicative of a cysteine adduct with no hydrolysis product evident. To account for the absence of hydrolysis product in the cysteine-driven reaction,
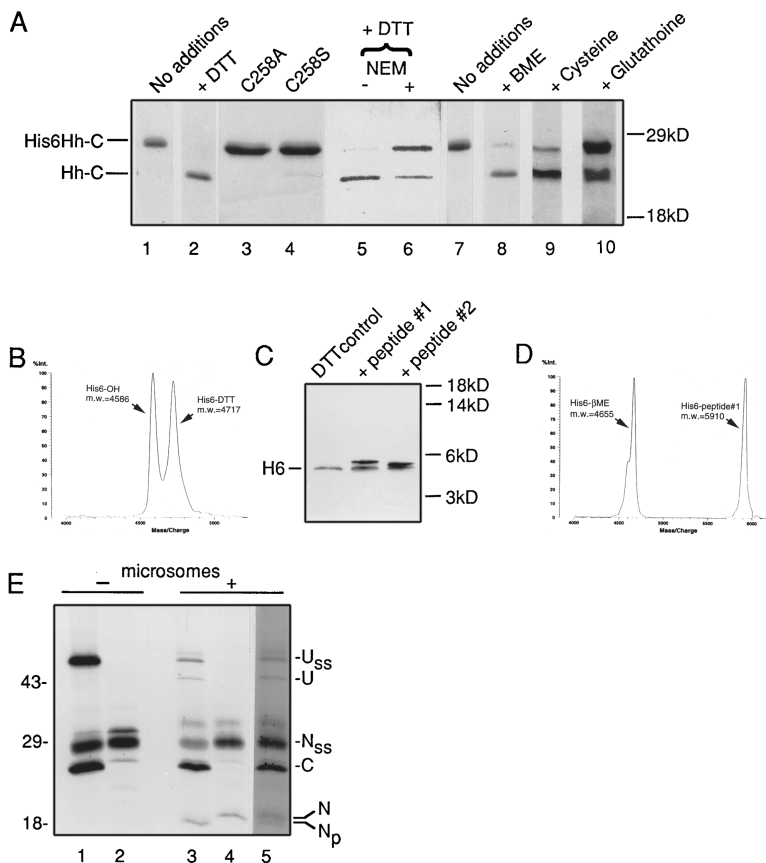
A

His6Hh-C —
Hh-C —

No additions | + DTT | C258A | C258S | + DTT { NEM - + } | No additions | + BME | + Cysteine | + Glutathoine

— 29kD
— 18kD

1   2   3   4   5   6   7   8   9   10

B

His6-OH m.w.=4586
His6-DTT m.w.=4717

C

DTTcontrol | + peptide #1 | + peptide #2

— 18kD
— 14kD
— 6kD
— 3kD

H6 —

D

His6-βME m.w.=4655
His6-peptide#1 m.w.=5910

E

microsomes
−   +

43-
29-
18-

-Uss
-U
-Nss
-C
-N
-Np

1   2   3   4   5

Figure 6. In Vitro Characterization of Hh Autoprocessing

(A)–(D) His6Hh-C, a bacterially expressed protein in which a His6 purification tag replaces most of the amino-terminal domain, was used to characterize the processing reaction in vitro. Except where specified, reactions were incubated for 16 hr at 23°C in the presence of 50 mM DTT.

(A) Coomassie-stained SDS–PAGE gels of His6Hh-C reactions without (lane 1) and with (lane 2) DTT. The amino-terminal product, labeled His6 in panels (B)–(D), is not shown. The effects of mutating Cys-258 to Ala (C258A) or to Ser (C258S) are shown in lanes 3 and 4. The small amount of C258S product accumulates independently of DTT addition (data not shown). Reactions in lanes 5 and 6 were pretreated with or without NEM. Reactions in lanes 7,8,9, and 10 were incubated for 15 hr with DTT eliminated or replaced by β-ME (100 mM), free cysteine (50 mM), or glutathione (50 mM).

(B) Mass spectral analysis of the amino-terminal fragment of a His6Hh-C processing reaction. The peak labeled His6-OH represents the presumed hydrolysis product of the peak labeled His6-DTT, which represents the presumed thioester-linked product of DTT attack (see Figure 5). The predicted mass of His6-OH is 4587 mass units (mu) and for His6-DTT is 4724 mu.

(C) Immunoblot of His6Hh-C reaction products electrophoresed in an SDS Tris–Tricine polyacrylamide gel and detected with anti-His6 monoclonal antibody. Lane 1 shows a control His6Hh-C reaction, and lanes 2 and 3 show reactions in the presence of reduced levels of DTT (5 mM) and Cys-initiated peptides 1 or 2 (2 mM), containing 13 and 10 residues, respectively.

(D) Mass spectral analysis of the amino-terminal fragment of a His6Hh-C cleavage reaction performed with β-ME and peptide 1. The predicted mass values of His6–β-ME and His6–peptide 1 fusion are 4647 and 5910 mu, respectively.

(E) In vitro translation reactions of Hh (lanes 1, 3, and 5) and the Hh-N construct (lanes 2, 4, and 5) without (lanes 1 and 2) and with microsomes (lanes 3, 4, and 5). N, C, and U denote the amino-terminal fragment from a truncated construct, the carboxy-terminal fragment, and the uncleaved precursor, respectively. In the absence of microsomes, the uncleaved and amino-terminal domain proteins retain signal sequences and are labeled Uss or Nss. In the presence of microsomes, the processed amino-terminal domain (Np) migrates faster than the truncated protein (N).

we propose that the thioester adduct with cysteine undergoes a rapid S to N shift to form an amide bond that is stable to hydrolysis. This phenomenon, observed previously when hydroxyl or thiol containing amino acids were used as nucleophiles to attack thioester intermediates, has recently been exploited to facilitate protein ligation in vitro (Levine and Dodds, 1990; Dawson et al., 1994). In these ligation studies, peptides containing amino-terminal cysteines were utilized as thioester-attacking nucleophiles capable of rearranging to form stable peptide bonds at neutral pH. We asked whether our in vitro cleavage reaction could catalyze protein ligation to peptides containing amino-terminal cysteines by adding two different cysteine-initiated peptides, 13 or 10 residues in length, to in vitro cleavage reactions. As assayed by immunoblotting (Figure 6C) and by mass spectrometry (Figure 6D; data not shown), each peptide tested was able to participate in a ligation reaction.

The Modification Reaction In Vitro
Our in vitro studies with purified protein do not recapitulate the full modification reaction, since the adducts

observed derive from exogenous nucleophiles added at high concentration. To gain a better understanding of the naturally occurring modification reaction, we asked whether the altered physical properties of Hh-Np could be conferred by in vitro translation of the Hh precursor in the presence of microsomes. Indeed, in transcription/translation extracts to which dog pancreatic microsomes were added, faster migration was observed for Hh-Np relative to Hh-N (Figure 6E, lanes 3, 4, and 5). This alteration in mobility of the amino-terminal fragment translated and processed in the presence of microsomes suggests that the endoplasmic reticulum may be the cellular site of the Hh-C-mediated modification reaction.

Modularity of Hh-C Function
Although our in vitro studies with His6Hh-C indicated that amino-terminal domain sequences are not required for thioester formation, a possible role of amino-terminal domain sequences for the full modification reaction was not ruled out. Accordingly, we tested the ability to undergo modification in vivo of a construct in which amino-

Figure 7. Modular Function of Hh-C and of Related Protein Domains Encoded by Novel Genes from C. elegans

(A) Immunoblot of extracts from S2 cultured cells expressing a flutagHh-C chimera and a corresponding construct lacking Hh-C sequences. Anti-flutag antibodies detect both processed and truncated proteins (lanes 1 and 2, FT$_p$ and FT, respectively). Two nonspecific bands are seen in each lane.

(B) Alignment of conserved blocks of predicted amino acid sequences was essentially as described (Koonin et al., 1996). Short variable spacers between conserved blocks are shown in lower case; longer spacers and sequences extending to the protein termini are replaced by the number of amino acid residues. Alignment of the carboxy-terminal autoprocessing domain is shown for two Hh proteins and six homologues from C. elegans; for the C. elegans proteins ZK, M75, and M89 the alignment of the amino-terminal domain that has no counterpart in Hh is also presented. The consensus line shows residues that are conserved in all three of the aligned sequences of the amino-terminal domain, or at least seven of eight of the aligned sequences in the carboxy-terminal autoprocessing domain. U indicates bulky hydrophobic residues (I, L, V, M, F, Y, and W, yellow); O indicates a small residue (G, A, and S, green); conserved negatively charged residues are highlighted in red, conserved positively charged residues in blue, and conserved cysteines in magenta. We marked eight conserved cysteines that may form S–S bonds by colons. The conserved Cys, Asp, and His residues in the autoprocessing domain that form a putative catalytic triad are indicated by asterisks.

(C) Domain relationships between Hh-C and the novel nematode genes.

(D) Three constructs to test ZK processing.

(E) Processing of the ZK protein. Immunoblots of extracts from S2 Drosophila cell cultures expressing zkWT-Flu (lane 1), zkU-Flu (lane 2), and zkN-Flu (lane 3). Anti-flutag antibodies detect the amino-terminal fragment derived by processing or from the zkN-Flu (lanes 1 and 3, ZK-N). Unprocessed precursor is also detected (lanes 1 and 2, ZK-U).

terminal domain sequences are replaced by an influenza virus epitope tag such that the epitope tag is located upstream of the cleavage site (see Experimental Procedures). As seen in Figure 7A (lanes 1 and 2), the processed fragment (FT$_p$) migrates faster in SDS–PAGE than the corresponding fragment derived from a truncated construct (FT), suggesting that the full auto-

processing reaction, including hydrophobic modification, can be initiated by the carboxy-terminal domain with no specific requirement for amino-terminal domain sequences.

The sufficiency of Hh-C sequences to initiate autoprocessing and modification raises the possibility that the Hh-C autoprocessing domain might be found in as-

sociation with proteins that do not contain a counterpart to the Hh-N domain. Comparison of the Hh sequences with protein sequence and expressed sequence tag (EST) databases indeed revealed highly significant similarity (probability of matching by chance between 10^{-5} and 10^{-13}) between Hh-C and six novel Hh-C-related Caenorhabditis elegans proteins (see Figures 7B and 7C, sequence alignment and schematic diagram). Conservation throughout the autoprocessing domain includes residues corresponding to Hh residues Cys-258, His-329, and Asp-282 that may constitute a conserved catalytic triad. Sequences from cDNAs representing four of these six (one incomplete) include extensive amino-terminal regions, none of which showed significant similarity to Hh-N or to any other proteins in current databases. Of the four, however, three are related to each other, and each of these three is predicted to contain an amino-terminal cleavable signal peptide (von Heijne, 1986).

We tested whether one of these proteins, ZK (overlapping but distinct from the theoretical protein ZK1290.5, which appears to have been incorrectly assembled from genomic sequence), was cleaved in a manner analogous to Hh. An influenza virus epitope tag was inserted into the amino-terminal region of the coding sequence and introduced into S2 cells in the following three forms: first, flu-tagged ZK (zkWT-Flu); second, flu-tagged ZK with Cys-291 replaced by an alanine codon (zkU-Flu); and third, flu-tagged ZK with Cys-291 replaced by a chain termination codon (zkN-Flu) (see Figure 7D). Cys-291 in ZK corresponds to Cys-258 in Hh-C (Figure 7B), and thus represents the first residue following the predicted cleavage site. As seen in Figure 7E, a Western blot of extracts from transfected cells shows that zkWT-Flu produces a product that comigrates with zkU-Flu and corresponds to unprocessed precursor, but in addition yields a cleaved product that comigrates with the product of zkN-Flu. The fact that the processed flu-tagged amino-terminal cleavage product comigrates with zkN demonstrates cleavage at the site predicted by homology, and the requirement for Cys-291 at this cleavage site indicated by the zkU-Flu construct suggests that cleavage proceeds by a mechanism similar to that of Hh-C. The identification of a modifying group will require further characterization, since these nematode proteins may not be accurately processed in Drosophila cells. The sequence data and the cleavage of ZK in cultured cells suggest that these genes encode a novel class of secreted proteins related to Hh in their ability to autoprocess.

Discussion

Previous work established that Hh processing is autocatalytic, identified the site of cleavage, and demonstrated that the amino-terminal product of autoprocessing suffices to account for *hh*-associated signaling activities in Drosophila and in vertebrates. Here, we report that autoprocessing influences the cell and tissue distribution of the amino-terminal signaling domain and its spatial patterning properties in embryos. Biochemical characterization of the processed protein revealed striking physical differences from the unprocessed protein,

indicative of a covalent modification and of an associated increase in hydrophobic character; this modification thus appears to be responsible for the normal cellular localization of the signaling domain and for the normal spatial pattern of embryonic signaling activity.

The Mechanism of the Modification Reaction
The chemical mechanism of Hh autoprocessing and modification that emerges from our biochemical studies involves an internal rearrangement to form a thioester between the carbonyl of Gly-257 and the thiol of Cys-258. This thioester then undergoes nucleophilic attack, resulting in the release of Hh-C and modification of the amino-terminal fragment on Gly-257 by an adduct derived from the attacking nucleophile (Figure 5B). This general scheme is common to a number of other protein modification reactions that also carry out acyl transfers via cysteine-derived thioester intermediates. Such reactions include the ubiquitin transfer cascade (Hershko and Ciechanover, 1992), complement fixation and the related action of the α-macroglobulin proteinase inhibitor (Chu and Pizzo, 1994), and the subset of self-splicing proteins that bear a sequence resemblance to Hh-C at the cleavage site (Koonin, 1995; Cooper and Stevens, 1995; Porter et al., 1995). This mechanism is also likely to operate in the biogenesis of other secreted proteins that contain sequences like those in Hh-C.

The Nature of the Modification
The physical properties of Hh-N_p that distinguish it from Hh-N include greater electrophoretic mobility and increased hydrophobic character, i.e., quantitative detergent extraction and dramatically prolonged retention in hydrophobic interaction chromatography. Hh-N_p also possesses greater mass than Hh-N, and is insensitive to carboxypeptidase digestion. As argued above (see Results), the unique chemical behavior of Hh-N_p conferred by the autoprocessing reaction is indicative of modification by covalent linkage to the carboxyl terminus of the amino-terminal signaling domain. In addition, the greater electrophoretic mobility rules out modifying groups consisting exclusively of carbohydrates or of amino acid residues, since these modifications would be expected to decrease mobility.

The significant increase in hydrophobic character of Hh-N_p relative to Hh-N suggests that the modifying group may be a lipid, given that lipids are broadly defined as hydrophobic cellular components requiring organic solvents for extraction (Alberts et al., 1994). We attempted to identify or rule out the known lipid modifications associated with other proteins (myristoylation, palmitoylation, prenylation, and GPI addition) (reviewed by Casey, 1995). No incorporation of label into immunoprecipitated Hh-N_p was observed when we pulse-labeled S2 cells with radioactive phosphate or with the radioactive fatty acids mevalonate, palmitate, myristate, and stearate (data not shown). To rule out more specifically the GPI anchor, we treated Hh-N_p with the enzymes phosphatidylinositol-specific phospholipase C (PI-PLC) and PI-PLD, or with nitrous acid, and observed no change in its electrophoretic mobility or hydrophobic character (data not shown). A GPI anchor is also ruled out by

failure to incorporate phosphate or labeled fatty acids using standard protocols, and by the small mass of the Hh-N$_p$ modification, since GPI anchors are generally $\sim$2 kDa. Given that Hh does not contain the hydrophobic carboxy-terminal sequences characteristic of GPI-modified proteins (Udenfriend and Kodukula, 1995), these results are not particularly surprising. Since the GPI anchor is the only known lipophilic modification associated with other cell surface proteins, these data suggest that the Hh-C-mediated autoprocessing reaction represents a novel pathway for lipophilic modification of secreted proteins. Our fatty acid labeling attempts appear to eliminate the known lipid modification pathways, but we cannot conclude that these fatty acids are not components of the Hh-N$_p$ modification, since we do not know the metabolic dynamics of the lipid pools we are attempting to label.

Cell Membrane Association as a Consequence of Signaling Domain Modification

Lipophilic modifications of other proteins are known to cause cell membrane association, presumably through hydrophobic interactions of the modifying group with the lipid bilayer (Casey, 1995). This mechanism provides the simplest explanation for the cell surface localization of Hh-N$_p$ in S2 cultured cells. Consistent with this idea, the unprocessed Hh-N protein does not receive the lipophilic modification and is not tethered to the cell surface, whereas Hh-N$_p$ remains cell-associated at extremely high salt concentrations and can only be extracted from cells with the use of detergent (data not shown). We cannot rule out contributions to membrane association by other interactions as noted for other lipid-modified proteins (McLaughlin and Aderem, 1995), one intriguing suggestion being that interaction with the extracellular matrix influences cell surface association or tissue distribution (Lee et al., 1994; Bumcrot et al., 1995; Roelink et al., 1995). These other interactions, however, are not sufficient for cell surface association, since truncated amino-terminal domains of vertebrate and Drosophila Hh proteins expressed in cultured cells are released into the medium (Porter et al., 1995; Bumcrot et al., 1995; Roelink et al., 1995).

Expression of Hh in embryos occurs in cells of the polarized epithelial layer of surface ectoderm. As previously noted in this context (Taylor et al., 1993; Lee et al., 1994; Tabata and Kornberg, 1994), Hh-N$_p$ is predominantly localized to large punctate structures in the basolateral domain of this epithelium. When the *en*GAL4 element is used to increase significantly the level of Hh-N$_p$ produced, this punctate basolateral distribution is maintained. In contrast, Hh-N protein expressed under control of the *en*GAL4 element is predominantly detected apically. Thus, beyond simple association with the cell membrane, processing, and its associated modification in embryos, appears to influence apical/basal localization of the signaling domain. Interestingly, the involvement of a lipophilic modification in subcellular sorting has precedent in the predominantly apical localization of GPI-linked proteins (Rodriguez-Boulan and Powell, 1992).

Effects of Modification upon the Tissue Distribution of the Signaling Domain

The induction of floor plate in vertebrate embryos occurs only in close proximity to *Shh*-expressing cells of the notochord; in Drosophila embryos, *wg*-expressing cells are found only adjacent to *hh*-expressing cells. The local nature of these embryonic effects corresponds to the sharp differences in concentration of signaling domain observed by immunostaining of expressing versus non-expressing cells in Drosophila and vertebrate embryos. This study suggests that these differences result from autoprocessing, lipophilic modification, and consequent cell membrane association of the signaling domain.

Other *hh*-dependent fates, in contrast, are formed at some distance from the site of *hh* synthesis. One mechanism that could account for this long-range action is the local induction by *hh* of other secondary signals that are then more directly responsible for induction of distant fates. Alternatively, induction of distant cell fates may occur by a more direct mechanism, requiring diffusion of the amino-terminal signaling domain from Hh-expressing cells. The possibility that low levels of processed modified amino-terminal domain may diffuse in vivo is consistent with the observation that several cytoplasmic proteins carrying lipophilic modifications are able to dissociate from the cell membrane (McLaughlin and Aderem, 1995). Modified amino-terminal domain proteins, furthermore, can be detected in the medium of Drosophila and mammalian cell cultures (data not shown).

The extent of intercellular protein diffusion is almost certainly determined by the hydrophobic character of the modification. Thus far, we have characterized only the adduct for Drosophila Hh-N$_p$ produced in S2 cultured cells of embryonic origin. During biosynthesis, the identity or presence of an adduct may be regulated, thus governing the tissue profile of the signaling domain. As an alternative, cell surface ligands, receptors, and enzymes have been shown to be enzymatically released from the membrane (Ehlers and Riordan, 1991). One candidate for such a releasing activity for the *Shh* signaling domain is the putative zinc hydrolase revealed by the crystal structure of the amino-terminal domain of *Shh* itself (Hall et al., 1995). There is as yet no evidence regarding a role or a substrate for this putative catalytic site.

A Role beyond Autoprocessing for the Cleaved Hh-C Domain?

We have considered the carboxy-terminal domain exclusively in the context of its role in the autoprocessing and modification of Hh-N$_p$. The possibility of an additional independent role is suggested by a recently defined family of four molecules, the amino-terminal nucleophile (Ntn) hydrolases (Brannigan et al., 1995). None of these four proteins are similar in sequence, but they share an unusual common fold that defines a structural superfamily. These proteins all undergo an autocatalytic cleavage reaction to generate mature enzymes with active site nucleophiles from the side chains of their amino-terminal residues (Ser, Thr, or Cys); the same side chain

is proposed to be active in the autoprocessing event. Since all of the Ntn hydrolases are enzymes independent of their autoprocessing roles, the suggestion is that Hh-C could have an independent enzymatic function that involves the amino-terminal residue (Cys-258) as nucleophile.

Hh-C-Like Domains in the Biogenesis of Other Secreted Proteins

We have identified a group of six proteins from C. elegans that contain domains related to Hh-C. Sequence comparisons indicate that this is a conserved domain that meets the definition of an ancient conserved region antedating at least the metazoan radiation (Green et al., 1993). By contrast, a comparison with the complete Saccharomyces cerevisiae genome showed that yeast does not encode Hh-C homologs, suggesting that this domain is restricted to multicellular organisms. We have demonstrated that at least one of the C. elegans proteins containing the domain related to Hh-C undergoes in vivo cleavage at a site like that in Hh. The cleavage, furthermore, is dependent upon a Cys residue homologous to the Cys that forms the thioester intermediate in Hh autoprocessing. Apparently, this conserved autoprocessing domain is modular and can be combined with different amino-terminal domains. We have shown in Hh that the carboxy-terminal autoprocessing domain governs the spatial pattern of Hh signaling activity by effecting a lipophilic modification of the amino-terminal signaling domain. The occurrence of this autoprocessing domain in other secreted proteins suggests that patterned protein distributions are required in other novel cell communication pathways that remain to be discovered.

Experimental Procedures

Analysis of *en*GAL4; UAS Hh or *en*GAL4; UAS Hh-N Flies

Homozygous *en*GAL4 flies were crossed to homozygous UAS*hh* or UAS*hh*N flies. The resulting embryos were subjected either to in situ hybridization to analyze *wg* RNA expression or to a cuticle preparation to analyze ventral and dorsal cuticle pattern (Porter et al., 1995). Relative levels of protein expression were examined by immunoblotting extracts from equivalent numbers of embryos collected 5–8 hr after egg laying. Embryos were immunostained as described (Lee et al., 1994) and viewed by confocal microscopy.

Protein Expression in Drosophila S2 Cultured Cells

Hh expression constructs used for making Hh-N$_p$ and Hh-N were generated by subcloning into pMK33, and establishing stably transformed S2 cell lines as described previously (Porter et al., 1995). Expression was induced by adding CuSO$_4$ to the culture medium to a final concentration of 0.3–1.0 mM. Two constructs (flutag-C and flutag) were generated in pMK33 by replacing Hh residues 89–254 with a trimer of the influenza hemagglutinin antigen (HA) in both the wild-type *hh* cDNA and in the truncated cDNA containing only N. The ZK cDNA expression constructs were generated by subcloning zk coding sequences, supplemented with a hemagglutinin epitope tag, into pMK33. The C to A mutation of the zkU-Flu construct and the amino-terminal truncation construct (zkN-Flu) were made with the use of the polymerase chain reaction. These constructs were transiently transfected into S2 cells and induced, and cell extracts were immunoblotted.

Protein Purification

Hh-N and Hh-N$_p$ proteins were purified from the stable Schneider S2 cell lines described above and previously (Porter et al., 1995).

Sixteen 245 mm × 245 mm tissue culture plates (Nunc) containing N$_p$-expressing cells at 80% confluence were induced with CuSO$_4$ at 1 mM final concentration. Cells were collected after 18 hr and subjected to a hypotonic lysis and dounce homogenization. Nuclei were removed with a low speed centrifugation and membranes were isolated by high speed centrifugation of the postnuclear supernatant. The N$_p$ membrane pellet was solubilized in buffer A (phosphate buffered saline (PBS) + 1 mM DTT) plus 1% Triton X-100. The solution was then subjected to centrifugation at 100,000 × g for 1 hr. The supernatant was passed over an anti-Hh-N immuno-affinity column, washed with buffer A plus 0.1% Triton X-100, and eluted with 100 mM glycine–HCl (pH 2.5) plus 0.1% Triton X-100. Fractions containing Hh-N were neutralized with 0.15 vol of 1 M Tris–HCl (pH 8.0). For purification of Hh-N, culture medium was collected from expressing cells, incubated for 2 hr at 4°C with 10 ml of heparin agarose beads, beads were washed with buffer A plus 0.1% Triton X-100, and elution was with buffer A plus 500 mM NaCl plus 0.1% Triton X-100. The Hh-N-containing fractions were pooled, diluted with 10 mM phosphate (pH 7.4) to decrease the NaCl concentration, and subjected to anti-Hh-N immuno-affinity chromatography as described above.

Biochemical Characterization of Hh-N and Hh-N$_p$

Triton X-114 extractions were performed on cultures expressing Hh-N and Hh-N$_p$ as described (Bordier, 1981), and samples were analyzed by immunoblotting.

For carboxypeptidase digestion purified Hh-N and Hh-N$_p$ proteins (~10 ng/μl) were diluted 1:1 with 0.2% SDS in 100 mM sodium citrate (pH 6.0), heated to 37°C for 10 min, and then diluted again with 9 vol of 50 mM citrate buffer. To 100 μl reactions 1 μl of sequencing grade carboxypeptidase Y (Boehringer Mannheim) was added, and the reactions were incubated at 30°C. Aliquots were removed at various times and analyzed by immunoblotting.

For CNBr digestions, purified Hh-N and Hh-N$_p$ proteins (250–500 μl at ~10 ng/μl) were precipitated with trichloroacetic acid (TCA) and resuspended in 70% formic acid. One mg of CNBr was added; the tubes were flushed with nitrogen, capped, and incubated in the dark at room temperature for 18–24 hr. For mass spectral analysis, the samples were either spotted directly onto polyethylene membrane or first diluted with 10 vol of H$_2$O, frozen, lyopholyzed, resuspended in 10 μl of 0.5% SDS, 100 mM Tris–HCl (pH 8.0), and then spotted.

HPLC analysis of Hh-N$_p$ and Hh-N was performed as described previously (Heukeshoven and Dernick, 1985). In brief, purified protein samples were combined with 88% formic acid at a ratio of 1:4 and loaded onto a C4 narrowbore HPLC column (Rainin) equilibrated with 60% formic acid and 0.8% isopropanol. At a flow rate of 0.5 ml/min, the proteins were eluted from the column by increasing the isopropanol concentration to 16% using a 50 min linear gradient.

MALDI Mass Spectrometry

Matrix-assisted laser desorption ionization (MALDI) mass spectral analysis (Chevrier and Cotter, 1991) was carried out on a Kratos Kompact MALDI/III mass spectrometer. A polyethylene (PE) membrane manufactured by 3M was used as a sample support as previously described (Blackledge and Alexander, 1995) to facilitate the removal of detergents required for handling proteins prior to mass analysis. Matrices used were 3,5-dimethoxy-4-hydroxycinnaminic or α-cyano-4-hydroxycinnaminic acid (Aldrich Chemicals) prepared as saturated solutions in 1:1 H$_2$0/acetonitrile and 1:1 acetonitrile/ 2% trifluoroacetic acid (TFA), respectively.

In Vitro Cleavage Assays

Bacterial Hh-fusion protein (His$_6$Hh-C) was produced and purified as described previously (Porter et al., 1995). This fusion protein (M$_r$-29K) contains Hh residues 83–471 with residues 89–254 deleted, all fused downstream of the hexahistidine tag from the pRSET B vector (Invitrogen). Standard cleavage reactions were incubated at 23°C in 150 mM NaCl, 100 mM Tris–HCl (pH 7.4), 0.05% Triton X-100, 1.25 mM β-ME, 2.5% glycerol in the presence or absence of small thiols: DTT (2–50 mM), β-ME (100 mM), cysteine (50 mM), and gluta-thione (50 mM). Reactions involving peptides were incubated at 30°C with 5 mM DTT or 25 mM β-ME and peptides at 2 mM. The

peptide sequences were: CAFSMPHGGSLHV (#1) and CPFGMHP
VLQ (#2).

Reactions with NEM were done with 2 mM NEM (0.1 M NaPO$_4$ (pH 6.9), for 15 min, at 25°C) while the protein was bound to Ni^{++} resin (Qiagen). After unreacted NEM was washed away, the protein was released from the beads with imidazole hydrochloride and the cleavage reaction initiated with DTT.

Identification and Analysis of Protein Sequences Related to Those of Hh-C

Sequence database searches were performed essentially as described (Koonin et al., 1996). cDNAs corresponding to homologous sequences were obtained by screening C. elegans embryonic and mixed-stage larval libraries; cDNAs corresponding to ESTs were obtained from L. Fulton at Washington University, St. Louis, MO, and from A. R. Kerlavage at The Institute for Genomic Research, Rockville, MD. In Figure 7, the labels ZK, M75, M89, and CE refer to proteins encoded by ZK, M75796, M89293, and CE cDNAs. WO6 is an alternatively assembled theoretical protein based on WO6B11.4, and M110 is the TO5C12.10 theoretical protein.

References

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J.D. (1994). Molecular Biology of the Cell (New York: Garland Publishing, Incorporated).

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K., eds. (1995). Current Protocols in Molecular Biology (New York: Greene Publishing Associates and Wiley-Interscience).

Blackledge, J.A., and Alexander, J.A. (1995). Polyethylene membrane as a sample support for direct matrix-assisted laser desorption/ionization mass spectrometric analysis of high mass proteins. Anal. Chem. *67*, 843–848.

Bordier, C. (1981). Phase separation of integral membrane proteins in Triton X-114 solution. J. Biol. Chem. *256*, 1604–1607.

Brand, A.H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development *118*, 401–415.

Brannigan, J.A., Dobson, G., Duggleby, H.J., Moody, P.C. E., Smith, J.L., Tomchick, D.R., and Murzin, A.G. (1995). A protein catalytic framework with an N-terminal nucleophile is capable of self-activation. Nature *378*, 416–419.

Bumcrot, D.A., Takada, R., and McMahon, A.P. (1995). Proteolytic processing yields two secreted forms of *sonic hedgehog*. Mol. Cell. Biol. *15*, 2294–2303.

Cardoso de Almeida, M.L., and Turner, M.J. (1983). The membrane form of variant surface glycoproteins of Trypanosoma brucei. Nature *302*, 349–352.

Casey, P.J. (1995). Protein lipidation in cell signaling. Science *268*, 221–225.

Chevrier, M.R., and Cotter, R.J. (1991). A matrix-assisted laser desorption time-of-flight mass spectrometer based on a 600 ps, 1.2 mJ nitrogen laser. Rapid Commun. Mass Spectrom. *5*, 611–617.

Chu, C.T., and Pizzo, S.V. (1994). α-2 macroglobulin, complement,

and biological defense: antigens, growth factors, microbial proteases, and receptor ligation. Lab. Invest. *71*, 792–812.

Cooper, A.A., and Stevens, T.H. (1995). Protein splicing: self-splicing of genetically mobile elements at the protein level. Trends Biochem. Sci. *20*, 351–356.

Dawson, P.E., Muir, T.W., Clark-Lewis, I., and Kent, S.B.H. (1994). Synthesis of proteins by native chemical ligation. Science *266*, 776–779.

Ehlers, M.R.W., and Riordan, J.F. (1991). Membrane proteins with soluble counterparts: role of proteolysis in the release of transmembrane proteins. Biochemistry *30*, 10065–10074.

Englund, P.T. (1993). The structure and biosynthesis of glycosyl phosphatidylinositol protein anchors. Annu. Rev. Biochem. *62*, 121–138.

Fietz, M.J., Jacinto, A., Taylor, A.M., Alexandre, C., and Ingham, P.W. (1995). Secretion of the amino-terminal fragment of the Hedgehog protein is necessary and sufficient for *hedgehog* signalling in Drosophila. Curr. Biol. *5*, 643–650.

Green, P., Lipman, D.J., Hillier, L., Waterston, R., States, D.J., and Claverie, J.M. (1993). Ancient conserved regions in new gene sequences and the protein databases. Science *259*, 1711–1716.

Hall, T.M.T., Porter, J.A., Beachy, P.A., and Leahy, D.J. (1995). A potential catalytic site revealed by the 1.7-Å crystal structure of the amino-terminal signalling domain of Sonic hedgehog. Nature *378*, 212–216.

Heemskerk, J., and DiNardo, S. (1994). Drosophila *hedgehog* acts as a morphogen in cellular patterning. Cell *76*, 449–460.

Hershko, A., and Ciechanover, A. (1992). The ubiquitin system for protein degradation. Annu. Rev. Biochem. *61*, 761–807.

Heukeshoven, J., and Dernick, R. (1985). Characterization of a solvent system for separation of water-insoluble poliovirus proteins by reversed-phase high-performance liquid chromatography. J. Chromatogr. *326*, 91–101.

Ingham, P.W. (1993). Localized *hedgehog* activity controls spatial limits of *wingless* transcription in the Drosophila embryo. Nature *366*, 560–562.

Koonin, E.V. (1995). A protein splice–junction motif in hedgehog family proteins. Trends Biochem. Sci. *20*, 141–142.

Koonin, E.V., Tatusov, R.L., and Rudd, K.E. (1996). Protein sequence comparison at a genome scale. Meth. Enzymol. *266*, 295–322.

Lee, J.J., Ekker, S., von Kessler, D.P., Porter, J.A., Sun, B.I., and Beachy, P.A. (1994). Autoproteolysis in *hedgehog* protein biogenesis. Science *266*, 1528–1537.

Levine, R.P., and Dodds, A.W. (1990). The thioester bond of C3. Curr. Topics Microbiol. Immunol. *153*, 73–82.

McLaughlin, S., and Aderem, A. (1995). The myristoyl–electrostatic switch: a modulator of reversible protein–membrane interactions. Trends Biochem. Sci. *20*, 272–276.

Porter, J.A., von Kessler, D.P., Ekker, S.C., Young, K.E., Lee, J.J., Moses, K., and Beachy, P.A. (1995). The product of *hedgehog* autoproteolytic cleavage active in local and long-range signalling. Nature *374*, 363–366.

Rodriguez-Boulan, E., and Powell, S.K. (1992). Polarity of epithelial and neuronal cells. Annu. Rev. Cell Biol. *8*, 395–427.

Roelink, H., Porter, J.A., Chiang, C., Tanabe, Y., Chang, D.T., Beachy, P. A., and Jessell, T.M. (1995). Floor plate and motor neuron induction by different concentrations of the amino-terminal cleavage product of *sonic hedgehog* autoproteolysis. Cell *81*, 445–455.

Rose, I.A., and Warns, J.V.B. (1983). An enzyme with ubiquitin carboxy-terminal esterase activity from reticulocytes. Biochemistry *22*, 4234–4237.

Schneider, I. (1972). Cell lines derived from late embryonic stages of *Drosophila melanogaster*. J. Embryol. Exp. Morphol. *27*, 353–365.

Tabata, T., and Kornberg, T.B. (1994). Hedgehog is a signalling protein with a key role in patterning Drosophila imaginal discs. Cell *76*, 89–102.

Taylor, A.M., Nakano, Y., Mohler, J., and Ingham, P.W. (1993). Contrasting distributions of patched and hedgehog proteins in the Drosophila embryo. Mech. Dev. *42*, 89–96.

Udenfriend, S., and Kodukula, K. (1995). How glycosyl-phosphatidyl-inositol-anchored membrane proteins are made. Annu. Rev. Biochem. *64*, 563–591.

van Poelje, P.D., and Snell, E.E. (1990). Pyruvoyl-dependent enzymes. Annu. Rev. Biochem. *59*, 29–59.

von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. Nucleic Acids Res. *14*, 4683–4690.

GenBank Accession Numbers

Accession numbers for the cDNAs reported in this paper are U61235 for ZK, U61236 for M75, U61237 for M89, and U61288 for CE.