Sonic Hedgehog is a member of the Hh/DD-peptidase family that spans the eukaryotic and bacterial domains of life.

Henk Roelink

Roelink@berkeley.edu

Abstract: Sonic Hedgehog (Shh) coordinates Zn^{2+} in a manner that resembles that of peptidases. The ability of Shh to undergo autoproteolytic processing is impaired in mutants that affect the Zn^{2+} coordination, while mutating residues essential for catalytic activity results in more stable forms of Shh. The residues involved in Zn^{2+} coordination in Shh are found mutated in some individuals with the congenital birth defect holoprosencephaly, demonstrating their importance in development. Highly conserved Shh domains are found a parts of some bacterial proteins that are members of the larger family of DD-peptidases, supporting the notion that Shh acts as a peptidase. Whereas this Hh/DD-peptidase motif is present in Hedgehog (Hh) proteins of nearly all animals, it is not present in *Drosophila* Hh, indicating that Hh signaling in fruit flies is derived, and perhaps not a good model for vertebrate Shh signaling. Sequence analysis of Hh proteins and their possible evolutionary precursors suggest that the evolution of modern Hh might have involved horizontal transfer of a bacterial gene coding of a Hh/DD-peptidase into a *Cnidarian* ancestor, recombining to give rise to modern Hh.

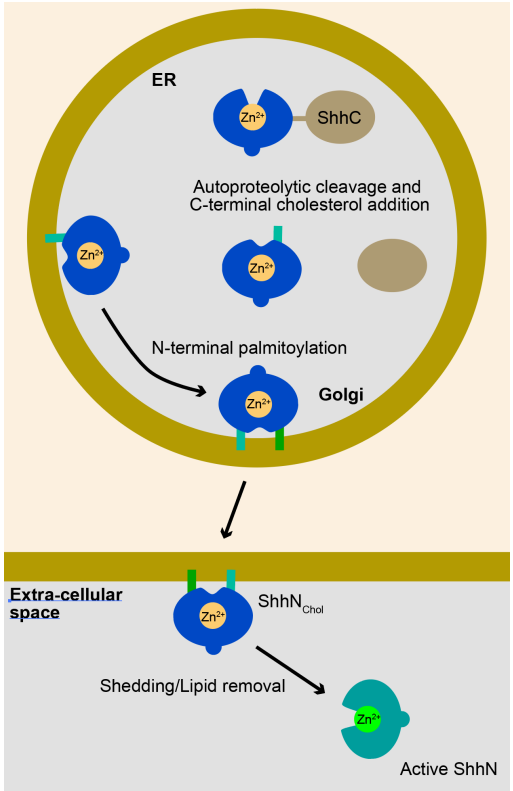
Keywords: Sonic Hedgehog; bacterial Hedgehog; *Drosophila* Hedgehog; Zn^{2+} peptidase; DD-peptidase; Hedgehog evolution

---

The Sonic Hedgehog pro-protein gives rise to the mature ligand after an autoproteolytic cleavage event

The *Hedgehog* (*Hh*) gene was first identified in the now famous developmental *Drosophila melanogaster* screen performed by Christiane Nüsslein-Volhard and Eric Wieshaus in the late 1970s. The screen used a technique known as “saturation mutagenesis" to isolate genes involved in the formation of the Drosophila body plan [1]. *Hh* mutant larvae have a solid lawn of denticles on the cuticle rather than stripes of denticles on the anterior half of each segment, hence the name “hedgehog”. Like other segment polarity genes found in this screen, also *Hh* genes are widely conserved among animals, and mammals have three Hh paralogs (Sonic, Indian, and Desert Hedgehog) that, like in *Drosophila*, play central roles in development [2].

Sonic Hedgehog (Shh) coordinates a Zn^{2+} metal ion with residues H141, D148, and H183 (mouse numbering, Figure 1) typical for Zn^{2+} peptidases, [3] such as the bacterial peptidase thermolysin [4]. Like many peptidases, the Shh undergoes an intramolecular auto-processing reaction resulting in cleavage between G198 and C199. As a consequence of this cleavage event the N-terminal product of this cleavage (ShhN), is modified with cholesterol [5,6]. Subsequently Shh is modified by an N-terminal acylation [7-9], rendering $ShhN_{chol}$ obligatory membrane bound. Secretion of this form of Shh requires Disp1and Scube2 [10] and ADAM-type metalloproteases, yielding a form that is stripped of its lipid modifications



and

Figure 1. Diagram of the Shh processing/maturation steps. Shh is translated into a pro-protein, consisting of a ShhN and ShhC domain, that undergoes an autoproteolytic processing event prior to entry into the Golgi. This results in the attachment of a cholesterol moiety the ShhN domain, likely shielding the Zn^{2+} coordination domain. In the Golgi, $ShhN_{chol}$ is further lipidated by a palmitoyl chain (green bars), further forcing its membrane association. Release and shedding is a regulated process involving Disp1, Scube2 and ADAM metalloproteases, resulting in the release of Shh devoid of its lipid moieties, and its Zn^{2+} coordination domain exposed.

active in signaling [11,12] (Figure 1). Whereas ShhN harbors the Zn^{2+} peptidase motif, the carboxyterminal domain has similarities to self-splicing bacterial inteins [13]. These inteins typically cleave before a Cysteine residue through the resolution of a thio-ester intermediate [6]. The G198/C199 site of cleavage is consistent with this idea. Single amino acid changes in C-terminal domains can prevent auto-processing, resulting in the perdurance of the Shh pro-protein [14]. Similarly, mutations of the residues that directly mediate Zn^{2+} coordination prevent the autoproteolytic processing of the Shh [15] and thermolysin pro-proteins [16]. There thus appear to be structural requirements in both the ShhC and ShhN domains for auto-processing to proceed, perhaps indicating that overlapping/complementary endopeptidase activities associated with the ShhN and -C domains are involved in the autocatalytic processing. The inherent endopeptidase activity of Shh is coupled to the addition of a cholesterol moiety to ShhN ($ShhN_{Chol}$), which affects distribution of the ligand [17]. Interestingly, many of the Zn^{2+} peptidase catalytic residues are not required for signaling by ShhN [18], and consequently the Zn^{2+} coordination domain of Shh has been referred to as the "pseudo active" site [19,20]. Nevertheless, the importance of the Zn^{2+} coordination domain has become apparent as mutations of this domain are associated with a congenital malformation, demonstrating a role in development.

Both the N- and C-terminal domains of Shh are the targets for point mutations found in holoprosencephaly

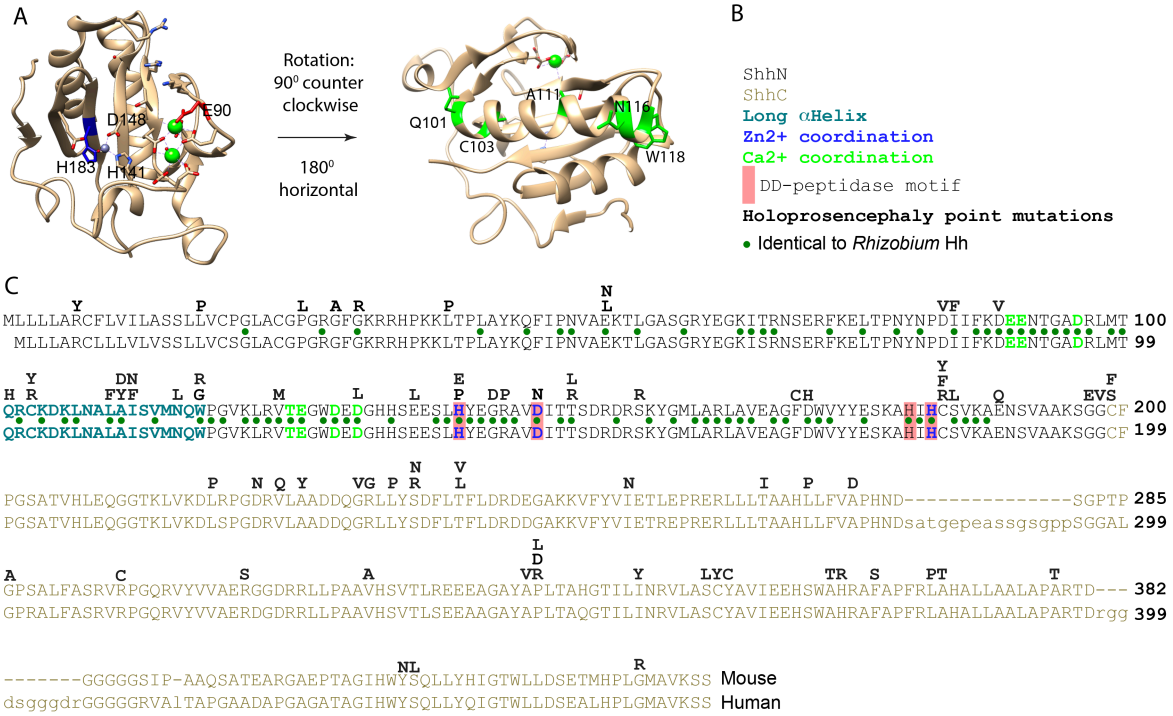


Figure 2. Salient features of the Shh protein.

(A) Crystal structure of ShhN. The Zn^{2+} (steel) and Ca^{2+} (green) coordination domains indicated. Residues in the large α-helix mutated in Holoprosencephaly are labeled in green. (B) Legend. (C) Lineup of mouse Shh (top) and human (bottom) SHH. Point mutations resulting in single amino acid changes in SHH found in holoprosencephalic individuals are above the lineup. Residues mediating Ca^{2+} and Zn^{2+} coordination are labeled in blue and green respectively. Green dots indicate conserved residues between *Rhizobium* Hh and mouse/human SHH.

SHH mutations are commonly found in holoprosencephaly a congenital syndrome that can be caused by aberrant SHH signaling [21-23] (Figure 2C). Single amino acid substitutions can be found in both the N- and C-domains of the SHH pro-protein, but are more prevalent in SHHN (40/181 in N, 38/266 in C, Z=-3.6, p=3.2x10^{-4}). Two of the Zn^{2+} coordination residues (H140 and D147 (Figure 2C, blue), have been found mutated in holoprosencephalic individuals, indicating that they are required for normal SHH function, consistent with the notion that the putative peptidase activity of Shh is important for signaling. The D148 equivalent is not conserved in *Drosophila* Hh, indicating it is not required for binding to Ptch. Traiffort *et al.* showed that SHH-H140P failed to undergo auto-processing, and was detected only as the SHH pro-protein [15]. This indicates that the perdurance of the SHH pro-protein might contribute to holoprosencephaly. It further shows that the correct Zn^{2+} coordination is necessary for processing the Shh pro-protein into $ShhN_{Chol}$. The face of ShhN opposite of the Zn^{2+} coordination domain is dominated by a large α–Helix (Figure 2A). This helix is enriched in point mutations found in holoprosencephalic individuals (Figure 2C, dark green). Two tested mutations, SHH-W117G and W117R, were unable to undergo auto-processing [15], further emphasizing structural requirements of the N-domain in auto-processing. Similarly, several mutations in the C-terminal domain prevent processing [14,15], emphasizing the central role that this domain play in processing the Shh pro-protein. The SHH mutations found in holoprosencephaly that thus likely affect SHH function indicate critical roles for both the N-terminal and C-terminal domains in auto-processing leaving the precise mechanisms and events by which the SHH pro-protein matures unresolved.

Shh has all hallmarks of a DD-peptidase

The notion that Hhs are pseudo-proteases is primarily based on studies claiming that the Zn^{2+} coordination site is required to maintain the correct and stable Shh structure, and for Shh binding to Ptch1 [24], but does not mediate protease activity [25]. This rejection of Shh protease function were based one experiments with negative outcomes, based on simple biochemical assays using artificial peptide substrates, common peptidase inhibitors, and *E.coli* expressed non-lipidated ShhN mutants that were not derived from the Shh pro-protein. Under such experimental conditions, even testing established proteases would possibly fail to reveal their proteolytic activitiesy. A possibly more useful approach would be to more precisely determine the effects of Robotnikinin, a small molecule that binds to the Shh Zn^{2+} coordinating domain [26] at high nanomolar concentrations [27], in cells that lack Ptch function.

In Zn^{2+} peptidases, the E177 equivalent abstracts a proton from the catalytic water at the Zn^{2+} coordination domain, which is followed by a nucleophilic attack of the OH^- on the peptide backbone. Shh-E177A is, therefore, predicted to be impaired for the intrinsic Zn^{2+} peptidase activity. Analysis of this mutant in has revealed two interesting properties. First, Shh-E177A is unable to mediate signaling from the notochord to the overlying neural tube (*in trans*), but is more capable than Shh to induce the Hh response when expressed in the developing neural tube (likely *in cis*) [28]. Second, purified ShhN-E177A is more stable in solution than ShhN, indicating a cannibalistic peptidase activity that is intrinsic to ShhN. This activity is inhibited by the binding of one or two Ca^{2+} ions [4] to ShhN (Figure 2A), using a binding motif that is conserved among Hh proteins, and present in thermolysin. It appears that interfering with the putative Zn^{2+} peptidase activity either via the Zn^{2+} coordination domain or E177 has negative consequences for Shh signaling during development, indicating a role for this peptidase activity-associated residue for normal Shh function.

The conservation of the Zn^{2+} coordinating, Ca^{2+} coordinating and other catalytic residues throughout evolution supports the possibility that Hhs are not pseudo-proteases but, using the properly matured form with the endogenous substrate, may indeed act as peptidases. Furthermore, structural analysis of interactions between Shh and the Hh interacting protein (Hhip) show that they resemble molecular interactions between matrix metalloproteases (MMPs) and their tissue inhibitor of metalloproteinase (TIMP). In both cases a loop present in Hhip and TIMP interacts closely with the ShhN-coordinated Zn^{2+} thus blocking catalytic activity. This striking similarity between unrelated protease/inhibitor pairs further supports the notion that ShhN is a Zn^{2+} peptidase: not only is an active site present, but interactions of this site with Hhip and possibly Ptch1 also resemble established metalloprotease/inhibitor interactions [19].

Typical for a protease active site, the Zn^{2+} in ShhN sits at the bottom of a cleft and is exposed to solvent, and not hidden inside of the molecule. This configuration that is conserved in the mature form of well-characterized protease thermolysin [29]. ShhN homologs in bacteria are characterized by a conserved Zn^{2+} coordination motif, that defines a family of prokaryotic proteins that are characterized by the DD-peptidase fold [30]. The signature peptidase fold is a central, five-stranded antiparallel β-sheet, separating the Zn^{2+} coordination domain from several α-helices, as found in ShhN (Figure 2A). Members of this family include murein endopeptidase (penicillin resistant enzymes) and peptidase M15 (bacterial D-alanyl-D-alanine carboxypeptidases (DD-peptidase), the target for penicillin) [30]. Both these peptidases play critical roles in bacterial cell wall modification. Shh and peptidase M15 share the two Histidine residues and an Aspartic acid residue that mediate Zn^{2+} coordination (H141 D148, H183 for Shh). The overall Zn^{2+} coordination motif is also found in lysostaphin, another peptidase that cleaves peptide bonds in bacterial peptidoglycan, and is referred to as the "LAS" (**L**ysostaphin, D-**a**lanyl-D-alanine carboxypeptidase, **S**hh) arrangement [31]. DD-peptidases characteristically contain a H-X(6)-D and H-X-H motif that coordinates Zn^{2+} in an accessible catalytic cleft. By all these criteria Shh is a member of this class of peptidases [30], further supporting the notion that Shh can function as a protease, perhaps even targeting a glycoprotein. Characterized bacterial DD-peptidases use a catalytic mechanism in which the acyl-linked peptide-enzyme intermediate is resolved by a nucleophilic attack of the adjacent peptidoglycan chain, thus cross-linking the peptidoglycan chains in the bacterial cell wall, and it remains to be determined if Shh has a similar activity.

Several bacterial species have highly conserved HhN domains

Remarkably, several species of bacteria carry a highly conserved Hedgehog proteins of unknown function (Figure 3, 4, Appendix A). An about 145 amino acid domain comprising the bulk of ShhN (Figure 3, green box) and containing the typical DD-peptidase motif has over 50% identity to some bacterial proteins (Figure 4B), and I'll refer to this class of protein domains as bacHhs. All the residues involved in Zn^{2+} and Ca^{2+} coordination are identical, as is E177, a residue required for non-autonomous Shh signaling. In bacteria, the Hh/DD-peptidases domain is the C-terminal part of larger proteins. Homologs of the N-terminal domains of the bacterial Hh proteins can be found associated with different subtypes of Zn^{2+} peptidases throughout bacteria (Figure 4A). In a few bacterial species, including some *Rhizobium* and

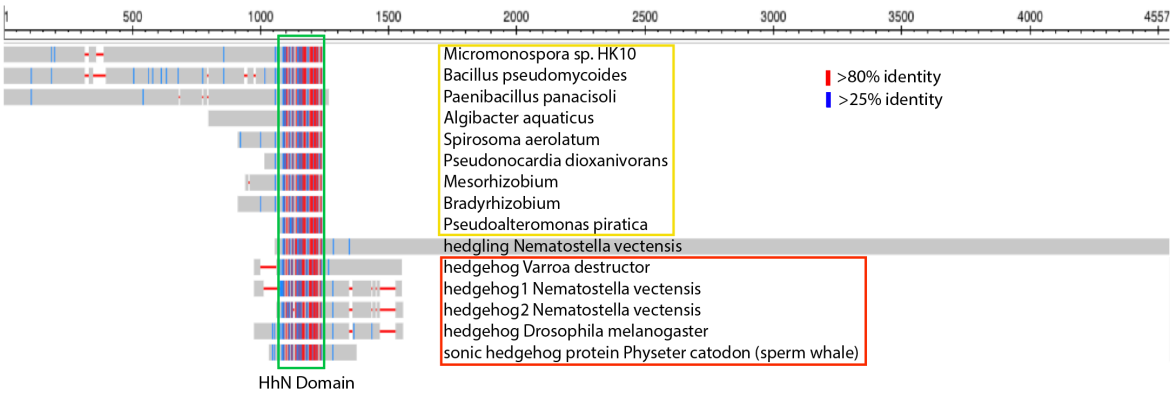


Figure 3. HhN domains are present in Bacteria.

Lineup of several hypothetical proteins in bacteria, Hedgling and metazoan Hhs. The conserved domain (HhN, green box, about 175 residues) is flanked by other sequences in the various species. In bacteria (yellow box) the HhN domain is the C-terminal end of the hypothetical proteins. The HhN domain is the N-terminal part of hedgling and Hhs (red box). Blue and red lines indicate medium and high levels of conservation.

Bradyrhizobium species, the N-terminal part preceding the Hh/DD-peptidase domain is predicted to contain three transmembrane regions, placing the Hh domain in the periplasmic space (Figure 4A). In this configuration, the third transmembrane domain od *Rhizobium* bacHh lines up perfectly with the signal sequence of Shh. No clear homologs of the Hh receptors Patched and Smoothened are present in any of these bacteria, or their *Rhizobium Legume* hosts, further supporting the idea that bacHhs serve as peptidases, rather than as ligands. The similarity of ShhN to its bacterial counterparts is as high as to other distant metazoan Hhs, and higher than to the *Cnidarian* hedgling N-domain, and *Drosophila* HhN (Figure 4B). Given the high degree of identity between Shh and its bacterial counterparts, it is likely that they share a specific function. The relatively small number of sequenced bacteria with a highly conserved HhN domain (I found about a dozen) do not share any obvious characteristics or ecological niches, and the precise function of this bacHh peptidase activity remains unknown.

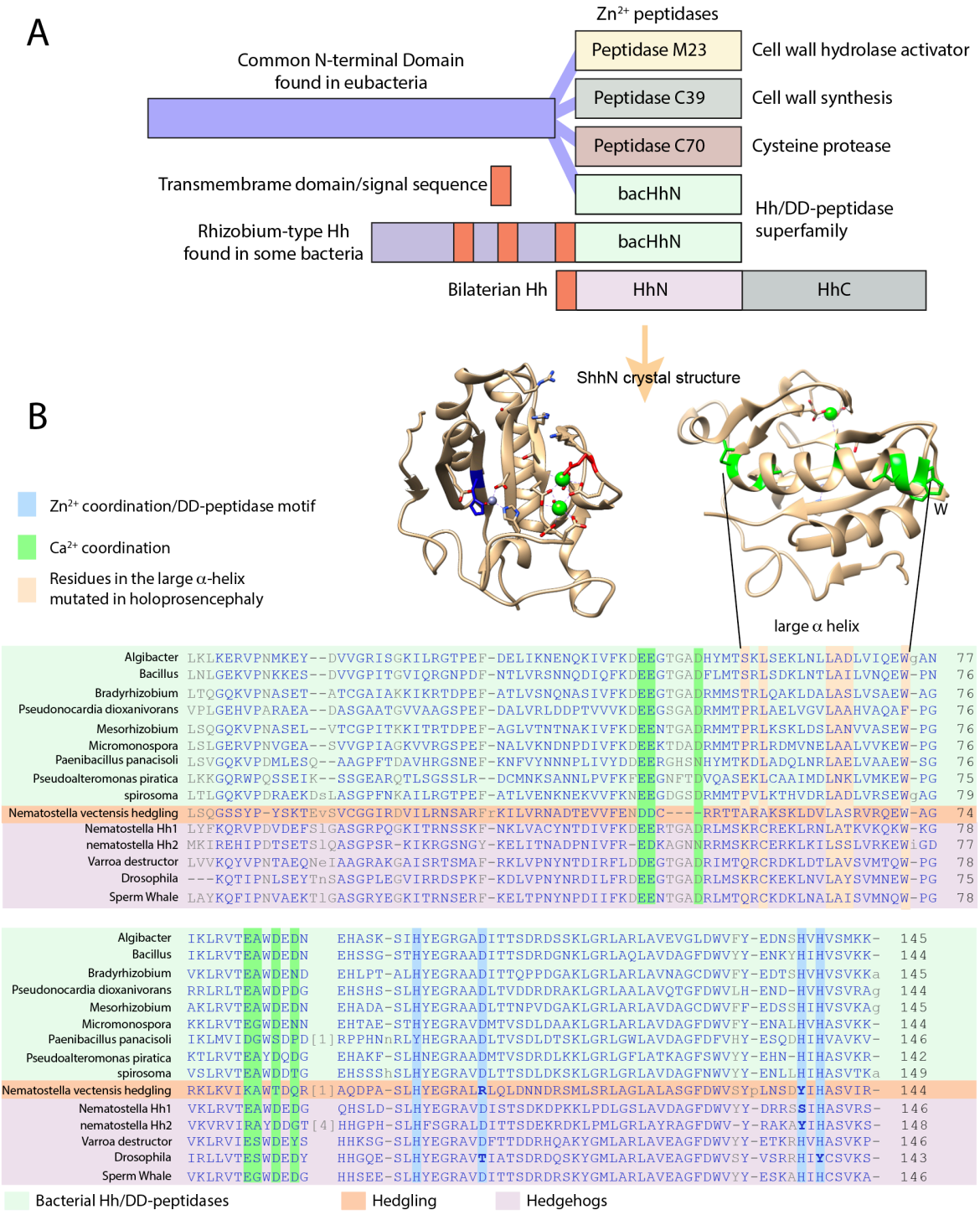


Figure 4. The Hh/DD-peptidase family

(A) Diagram of the structure of the DD-peptidases in bacteria and animals. In bacteria the peptidases are the C-terminal domain of larger proteins. In all cases the peptidase domain is predicted to be located in the periplasmic space. In *Rhizobium* Hh the third transmembrane domain (red) is at the same position as the signal sequence in metazoan Hh (red). (B) Lineup of the Hh/DD-peptidase domains of bacteria (green background), Hedgling (salmon background) and Hhs (purple background). Mutations in the Zn^{2+} coordinating/DD-peptidase motif defining residues in Hedgling and *Drosophila* Hh are indicated in bold. Blue columns indicate residues involved in Zn^{2+} coordination and define the DD-peptidase motif, green columns indicate residues involved in Ca2+ binding. *Varroa* is typical for all non-*Drosophilid* insects.

What are the possible substrates for the Shh peptidase activity?

The general properties of bacterial DD-peptidases and lysostaphins as modifiers of the bacterial cell wall is likely shared with the bacHhs. The peptidoglycans that are a major component of the bacterial periplasmic space have some similarities to the proteoglycans that are common in the insect and vertebrate extracellular matrix. Both bacterial cell wall peptidoglycans and animal matrix proteoglycans are large molecules in which polypeptides are covalently attached to chains of glycans. In particular, the matrix heparin sulfate proteoglycans (HSPGs) bind Shh and can both negatively and positively affect the Shh response [32-34]. Furthermore, mutations in *ext* genes that code for glycosyltransferases that catalyze glycosaminoglycan addition to the core protein, disrupt Hh signaling in vertebrates [35] and insects [36]. It is thus possible that functional conservation between bacHhs and Shh is reflected in the ability of Shh to cleave or modify proteoglycans, thus affecting the Shh response or distribution, independent of binding to the canonical receptors.

Although any Shh antagonist could be a possible target for Shh peptidase activity, both Hhip and Ptch1 are unlikely substrates, as they have characteristics of metalloprotease inhibitors [19].

***Drosophila* Hh is not a member of the Hh/DD-peptidase family**

Drosophila Hh (dmHh) has been the guide molecule for all Hh signaling since its discovery as necessary for embryogenesis [1]. However, dmHh is unusual in that is does not have the core sequences H-X(6)-D, H-X-H that define the Hh/DD-peptidase motif, but instead dmHh has H-X(6)-**T**, H-X-**Y**. Furthermore, the catalytically important E177 (mouse numbering) is mutated in dmHh into a Valine residue. The absence of the Hh/DD-peptidase motif is unique to all sequenced *Drosophilids*; searching with the *Drosophila* sequence that encompasses the DD-peptidase motif (ESL**H**YEGRAV**T**IATSDRDQSKYGMLARLAVEAGFDWVSY**V**SRR**H**I**Y**CSVKS) places *Drosophilids* as an outgroup to all other arthropods and deuterostomes. *Drosophilids* are the outlier because Hh proteins in all other protostomes and deuterostomes retain the conserved Hh/DD-peptidase Zn^{2+} coordination motif. Searching with a hypothetical *dmHh* peptide that conforms with the Hh/DD-peptidase motif and E177 does no longer uniquely group with the *Drosophilids*, and now has similar homology to several vertebrate and insect Hhs. Given the high degree of HhN conservation between bacteria and eukaryotes (Figure 3), the most plausible explanation is that *Drosophilids* have lost some of the activities associated with the ancestral Zn^{2+} coordination domain. The loss of Zn^{2+} coordinating residues in conjunction with the E177 homolog is no surprise in the light of the likely loss of peptidase activity in *Drosophila* Hh. The observation that vertebrate Shh is active in *Drosophila* embryos [37], but *Drosophila Hh* is not active in vertebrates, further supports the notion that some aspects of Hh signaling present in most animals are lost in fruit flies, but that both proteins can bind to Ptch. These observations question to what extent the lessons learned in *Drosophila* embryos regarding Hh processing and signaling can be directly applied to vertebrates, or even other arthropods, as all these animals maintain an intact Hh/DD-peptidase domain. Study of Hh signaling in another insect, such a *Tribolium* [38] might help resolve how *Drosophila* Hh signaling is impacted by the loss of the DD-peptidase motif in the ligand.

The presence of bacHhs suggest an alternative hypothesis regarding the evolution of modern Hh

Hh genes are present in *Cnidarians* (corals and jellyfish), but not in sponges and protozoa [39]. However, another class of proteins that contain an N-terminal HhN-like domain are the Hedgling proteins that can be found in *Cnidarians,* sponges and *Choanoflagellates* [40], which are protozoans. In these Hedgling proteins the HhN domain is followed by a large C-terminal domain that is related to cell adhesion molecules. A plausible evolutionary path is that the HhN domain of Hedgling was recombined to a Hog

domain that is present in most genomes (Figure 5 green arrows). However, the HhN domain of Hedgling does not contain a complete Hh/DD-peptidase motif, whereas it is likely that its putative ancestral form in the last universal common ancestor (LUCA) had an intact Hh/DD-peptidase domain. Although this evolutionary path from a LUCA DD-peptidase via Hedgling to Hh could conceivably occur without

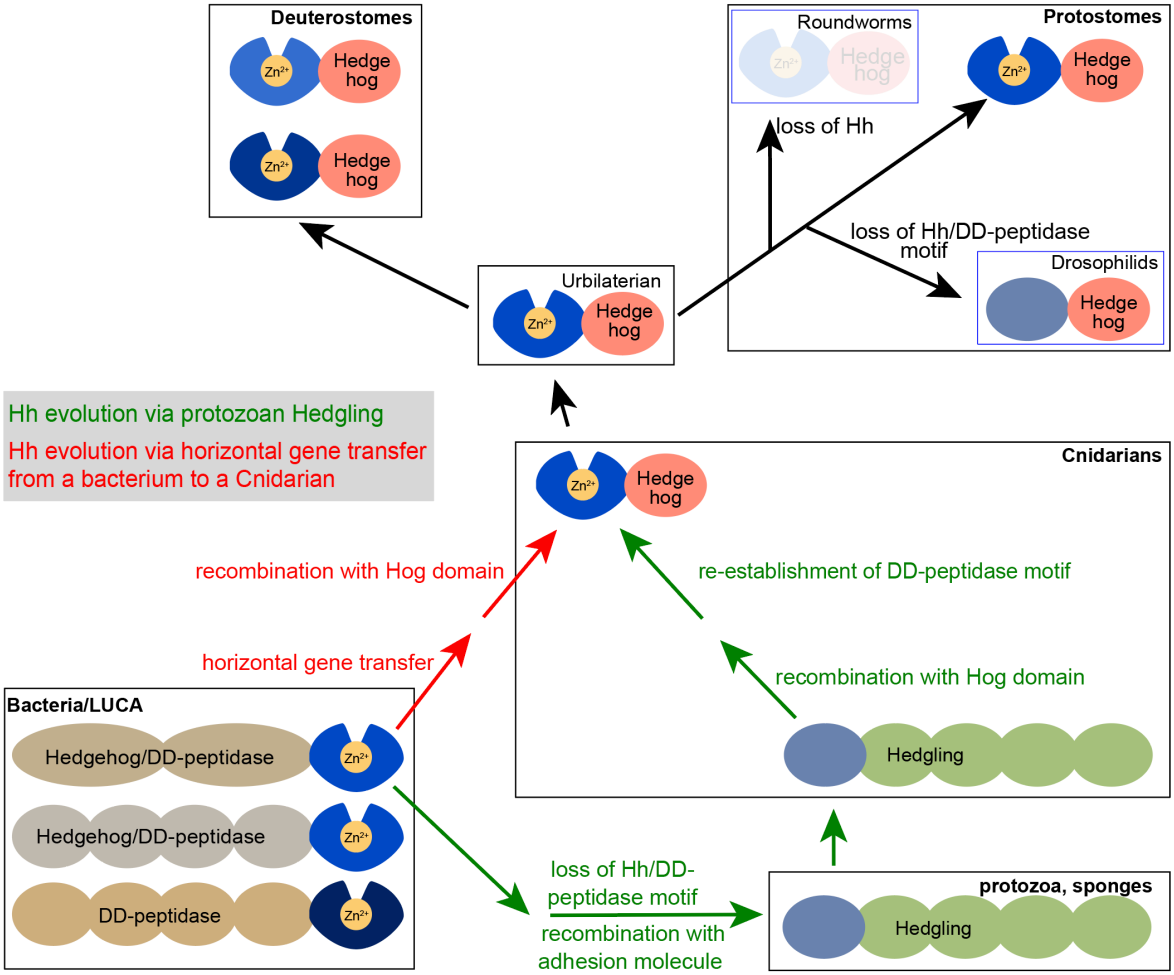


Figure 5. Models for the evolution of modern Hh proteins.

The presence of both Hhs and Hedglings in Cnidarians support the model that the HhN of Hh domain arose via a recombination of the Hedgling N-domain. As Hedgling is present in at least some protozoa, it is plausible that is was derived from a DD-peptidase in the last universal common ancestor (LUCA). Alternatively, a Hh/DD-peptidase could have entered a Cnidarian ancestor from a bacterium via horizontal gene transfer, never losing the Zn^{2+} coordination in the process. Modern Hh was present in the Urbilaterian (the last common ancestor of protostomes and deuterostomes) and was retained in most of its offspring. The loss of the Hh/DD-peptidase motif in *Drosophilids* is derived, as is the loss of all Hh in some bilaterian protostomes (such as roundworms).

horizontal gene transfer (Figure 5 green arrows), it does require several recombination events. The first event would entail the recombination of the bacterial-like Hh domain from a DD-peptidase to form the Hedgling protein, the second event would recombine the HhN domain of Hedgling into modern Hh. Following this potential evolutionary path the Hh/DD-peptidase motif, is lost in Hedgling, and re-established in Hh.

The presence of highly conserved HhN protein domains in bacteria suggests an alternate hypothesis, in which horizontal *bacHh//DD-peptidase* gene transfer from a bacterium to a *Cnidarian* ancestor was followed by a recombination of the HhN domain with a *Cnidarian Hog* domain, resulting in a modern *Hh* gene (Figure 5, red arrows). This possible evolutionary path requires only one recombination event,

and does not involve the subsequent loss and re-establishment of the Hh/DD-peptidase domain as is indicated for the evolutionary path via Hedgling, and is thus parsimonious path. The higher conservation between ShhN and bacterial Hh/DD-peptidases than between ShhN and the N-domain of Hedgling lends further support to the horizontal gene transfer model.

Conclusions

Although the Zn^{2+} coordination domain of Shh is often referred to as its "pseudo catalytic" domain [19,41,42], the remarkable similarity of most Hhs to bacterial Hh/DD-peptidases further supports the notion that Shh functions as a peptidase during development. Some mutations found in holoprosencephaly patients break the Hh/DD-peptidase motif, and negatively affect Hh signaling possibly by preventing autoproteolytic cleavage, indicating that an intrinsic Zn^{2+} peptidase activity of Shh is critical to its function. The lack of the Hh/DD-peptidase domain in *Drosophila* Hh demonstrates that it is evolutionary derived, and perhaps not the best model for Hh signaling in animals with the ancestral Hh/DD-peptidase motif, like humans and mice. Finally, the presence of HhN protein domains in some bacteria might support an alternate pathway for the evolution of Hh, via horizontal gene transfer from bacteria into am Urbilaterian ancestor, resulting in modern Hh that retains an ancestral peptidase activity.

Funding: This research was funded by NIGMS grant number R01GM117090.

Conflicts of Interest: None

Appendix A

Accession numbers for the proteins used for the lineup in Figure 3.

WP_082159544.1	hypothetical protein [Micromonospora sp. HK10]	
WP_098188151.1	hypothetical protein [Bacillus pseudomycoides]	Related Information
WP_028591995.1	hypothetical protein [Paenibacillus panacisoli]	
WP_069829738.1	hypothetical protein [Algibacter aquaticus]	Related Information
WP_080055297.1	hypothetical protein [Spirosoma aerolatum]	
WP_103381118.1	hypothetical protein [Pseudonocardia dioxanivorans]	
ESZ39647.1	hypothetical protein X732_15330 [Mesorhizobium sp. L2C066B000]	
WP_029077494.1	hypothetical protein [Bradyrhizobium sp. th.b2]	
WP_040135141.1	hypothetical protein [Pseudoalteromonas piratica]	Related Information
ABX84114.1	hedgling [Nematostella vectensis]	
XP_022667503.1	indian hedgehog protein-like [Varroa destructor]	Related Information
ABX89897.1	hedgehog1 [Nematostella vectensis]	Related Information
ABX89898.1	hedgehog2 [Nematostella vectensis]	Related Information
NP_001034065.1	hedgehog [Drosophila melanogaster]	Related Information
XP_007124731.1	sonic hedgehog protein [Physeter catodon]	

Appendix B

BLAST Multiple alignment input (ncbi.nlm.nih.gov) for Hh/DD-peptidase domains used for the lineup in Figure 4

>Algibacter
LKLKERVPNMKEYDVVGRISGKILRGTPEFDELIKNENQKIVFKDEEGTGADHYMTSKLSEKLNLLADLVIQEWGA
NIKLRVTEAWDEDNEHASKSIHYEGRGADITTSDRDSSKLGRLARLAVEVGLDWVFYEDNSHVHVSMKK
>Bacillus
LNLGEKVPNKKESDVVGPITGVIQRGNPDFNTLVRSNNQDIQFKDEEGTGADFLMTSRLSDKLNTLAILVNQEWP
NIKLRVTEAWDEDNEHSSGSTHYEGRAADITTSDRDGNKLGRLAQLAVDAGFDWVYYENKYHIHVSVKK
>Bradyrhizobium
LTQGQKVPNASETATCGAIAKKIKRTDPEFATLVSNQNASIVFKDEEGTGADRMMSTRLQAKLDALASLVSAEWA
GVKLRVTEAWDENDEHLPTALHYEGRAADITTQPPDGAKLGRLARLAVNAGCDWVFYEDTSHVHVSVKKA
>Pseudonocardia dioxanivorans
VPLGEH
VPARAEADASGAATGVVAAGSPEFDALVRLDDPTVVVKDEEGSGADRMMTPRLAELVGVLAAHVAQAFPGRRL
RLTEAWDPDGEHSHSSLHYEGRAADLTVDDRDRAKLGRLAALAVQTGFDWVLHENDHVHVSVRAG
>Mesorhizobium
LSQGQKVPNASELVTCGPITKKITRTDPEFAGLVTNTNAKIVFKDEENTGADRMMTPRLKSKLDSLANVVASEWPG
AKLRVTEAWDEDNEHADASLHYEGRAADLTTNPVDGAKLGRLARLAVDAGCDWVFFEDSSHIHVSVKAG
>Micromonospora
LSLGERVPNVGEASVVGPIAGKVVRGSPEFNALVKNDNPDIVFKDEEKTDADRMMTPRLRDMVNELAALVVKEW
PGKKLRVTEGWDENNEHTAESTHYEGRAVDMTVSDLDAAKLGRLARLAVDAGFDWVFYENALHVHASVKK
>Nematostella Hh1
LYFKQRVPDVDEFSLGASGRPQGKITRNSSKFNKLVACYNTDIVFKDEERTGADRLMSKRCREKLRNLATKVKQK
WKGVKLRVTEAWDEDGQHSLDSLHYEGRAVDISTSDKDPKKLPDLGSLAVDAGFDWVYYDRRSSIHASVRS
>Nematostella Hh2
MKIREHIPDTSETSLQASGPSRKIKRGSNGYKELITNADPNIVFREDKAGNNRRMSKRCERKLKILSSLVRKEWIGDV
KVRVIRAYDDGTSKKRHHGPHSLHFSGRALDITTSDEKRDKLPMLGRLAYRAGFDWVYRAKAYIHASVKS
>Nematostella vectensis hedgling
LSQGSSYPYSKTEVSVCGGIRDVILRNSARFRKILVRNADTEVVFENDDCRRTTARAKSKLDVLASRVRQEWAGRKL
KVIKAWTDQRTAQDPASLHYEGRALRLQLDNNDRSMLSRLAGLALASGFDWVSYPLNSDYIHASVIR
>Paenibacillus panacisoli
LSVGQKVPDMLESQAAGPFTDAVHRGSNEFKNFVYNNNPLIVYDDEERGHSNHYMTKDLADQLNRLAELVAAE
WSGIKLMVIDGWSDPDVRPPHNNRLYHEGRAADLTVSDLDTSKLGRLGWLAVDAGFDFVHYESQDHIHVAVKV
>Pseudoalteromonas piratica
LKKGQRWPQSSEIKSSGEARQTLSGSSLRDCMNKSANNLPVFKFEEGNFTDVQASEKLCAAIMDLNKLVMKEWPG
KTLRVTEAYDQDGEHAKFSLHNEGRAADMTVSDRDLKKLGRLGFLATKAGFSWVYYEHNHIHASVKR
>spirosoma
LTLGQKVPDRAEKDSLASGPFNKAILRGTPEFATLVENKNEKVVFKNEEGDGSDRMMTPVLKTHVDRLADLVRSE
WGAGVSLRVTEAWDDTGEHSSSHSLHYEGRAVDLTTSDLDKSKLGRLGRLAVDAGFNWVYYENLLHIHASVTKA
>Varroa destructor Hh
LVVKQYVPNTAEQNEIAAGRAKGAISRTSMAFRKLVPNYNTDIRFLDDEGTGADRIMTQRCRDKLDTLAVSVMTQ
WPGVKLRVIESWDEYSHHKSGSLHYEGRAVDFTTDDRHQAKYGMLARLAVEAGFDWVYYETKRHVHASVKP
>Drosophila Hh

KQTIPNLSEYTNSASGPLEGVIRRDSPKFKDLVPNYNRDILFRDEEGTGADRLMSKRCKEKLNVLAYSVMNEWPGIR
LLVTESWDEDYHHGQESLHYEGRAVTIATSDRDQSKYGMLARLAVEAGFDWVSYVSRRHIYCSVKS
>Sperm Whale Shh
LAYKQFIPNVAEKTLGASGRYEGKITRNSERFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQW
PGVKLRVTEGWDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYYESKAHIHCSVKA

References

1. Nüsslein-Volhard, C.; Wieshaus, E. Mutations affecting segment number and polarity in Drosophila. *Nature* **1980**, *287*, 795–801.

2. Echelard, Y.; Epstein, D. J.; St-Jacques, B.; Shen, L.; Mohler, J.; McMahon, J. A.; McMahon, A. P. Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* **1993**, *75*, 1417–1430.

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

4. Rebollido-Rios, R.; Bandari, S.; Wilms, C.; Jakuschev, S.; Vortkamp, A.; Grobe, K.; Hoffmann, D. Signaling domain of Sonic Hedgehog as cannibalistic calcium-regulated zinc-peptidase. *PLoS Comput. Biol.* **2014**, *10*, e1003707.

5. Porter, J. A.; Ekker, S. C.; Park, W. J.; Kessler, von, D. P.; Young, K. E.; Chen, C. H.; Ma, Y.; Woods, A. S.; Cotter, R. J.; Koonin, E. V.; Beachy, P. A. Hedgehog patterning activity: role of a lipophilic modification mediated by the carboxy-terminal autoprocessing domain. *Cell* **1996**, *86*, 21–34.

6. Porter, J. A.; Young, K. E.; Beachy, P. A. Cholesterol modification of hedgehog signaling proteins in animal development. *Science (New York, N.Y* **1996**, *274*, 255–259.

7. Taylor, F. R.; Wen, D.; Garber, E. A.; Carmillo, A. N.; Baker, D. P.; Arduini, R. M.; Williams, K. P.; Weinreb, P. H.; Rayhorn, P.; Hronowski, X.; Whitty, A.; Day, E. S.; Boriack-Sjodin, A.; Shapiro, R. I.; Galdes, A.; Pepinsky, R. B. Enhanced potency of human Sonic hedgehog by hydrophobic modification. *Biochemistry* **2001**, *40*, 4359–71.

8. Pepinsky, R. B.; Zeng, C.; Wen, D.; Rayhorn, P.; Baker, D. P.; Williams, K. P.; Bixler, S. A.; Ambrose, C. M.; Garber, E. A.; Miatkowski, K.; Taylor, F. R.; Wang, E. A.; Galdes, A. Identification of a palmitic acid-modified form of human Sonic hedgehog. *The Journal of biological chemistry* **1998**, *273*, 14037–14045.

9. Chamoun, Z.; Mann, R. K.; Nellen, D.; Kessler, von, D. P.; Bellotto, M.; Beachy, P. A.; Basler, K. Skinny hedgehog, an acyltransferase required for palmitoylation and activity of the hedgehog signal. *Science (New York, N.Y* **2001**, *293*, 2080–2084.

10. Tukachinsky, H.; Kuzmickas, R. P.; Jao, C. Y.; Liu, J.; Salic, A. Dispatched and scube mediate the efficient secretion of the cholesterol-modified hedgehog ligand. *Cell Rep* **2012**, *2*, 308–320.

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

12. Ohlig, S.; Farshi, P.; Pickhinke, U.; van den Boom, J.; Höing, S.; Jakuschev, S.; Hoffmann, D.; Dreier, R.; Schöler, H. R.; Dierker, T.; Bordych, C.; Grobe, K. Sonic hedgehog shedding results in functional activation of the solubilized protein. *Developmental cell* **2011**, *20*, 764–774.

13. Hall, T. M.; Porter, J. A.; Young, K. E.; Koonin, E. V.; Beachy, P. A.; Leahy, D. J. Crystal structure of a Hedgehog autoprocessing domain: homology between Hedgehog and self-splicing proteins. *Cell* **1997**, *91*, 85–97.

14. Tokhunts, R.; Singh, S.; Chu, T.; D'Angelo, G.; Baubet, V.; Goetz, J. A.; Huang, Z.; Yuan, Z.; Ascano, M.; Zavros, Y.; Thérond, P. P.; Kunes, S.; Dahmane, N.; Robbins, D. J. The full-length unprocessed hedgehog protein is an active signaling molecule. *The Journal of biological chemistry* **2010**, *285*, 2562–2568.

15. Traiffort, E.; Dubourg, C.; Faure, H.; Rognan, D.; Odent, S.; Durou, M.-R.; David, V.; Ruat, M. Functional characterization of sonic hedgehog mutations associated with holoprosencephaly. *J. Biol. Chem.* **2004**, *279*, 42889–42897.

16. Inouye, K.; Kusano, M.; Hashida, Y.; Minoda, M.; Yasukawa, K. Engineering, expression, purification, and production of recombinant thermolysin. *Biotechnol Annu Rev* **2007**, *13*, 43–64.

17. Zeng, X.; Goetz, J. A.; Suber, L. M.; Scott, W. J. J.; Schreiner, C. M.; Robbins, D. J. A freely diffusible form of Sonic hedgehog mediates long-range signalling. *Nature* **2001**, *411*, 716–720.

18. Fuse, N.; Maiti, T.; Wang, B.; Porter, J. A.; Hall, T. M.; Leahy, D. J.; Beachy, P. A. Sonic hedgehog protein signals not as a hydrolytic enzyme but as an apparent ligand for patched. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 10992–10999.

19. Bosanac, I.; Maun, H. R.; Scales, S. J.; Wen, X.; Lingel, A.; Bazan, J. F.; de Sauvage, F. J.; Hymowitz, S. G.; Lazarus, R. A. The structure of SHH in complex with HHIP reveals a recognition role for the Shh pseudo active site in signaling. *Nature structural & molecular biology* **2009**, *16*, 691–697.

20. Maun, H. R.; Wen, X.; Lingel, A.; de Sauvage, F. J.; Lazarus, R. A.; Scales, S. J.; Hymowitz, S. G. Hedgehog pathway antagonist 5E1 binds hedgehog at the pseudo-active site. *The Journal of biological chemistry* **2010**, *285*, 26570–26580.

21. Odent, S.; Atti Bitach, T.; Blayau, M.; Mathieu, M.; Aug, J.; Delezo de, A. L.; Gall, J. Y.; Le Marec, B.; Munnich, A.; David, V.; Vekemans, M. Expression of the Sonic hedgehog (SHH) gene during early human development and phenotypic expression of new mutations causing holoprosencephaly. *Hum. Mol. Genet.* **1999**, *8*, 1683–1689.

22. Hehr, U.; Pineda-Alvarez, D. E.; Uyanik, G.; Hu, P.; Zhou, N.; Hehr, A.; Schell-Apacik, C.; Altus, C.; Daumer-Haas, C.; Meiner, A.; Steuernagel, P.; Roessler, E.; Winkler, J.; Muenke, M. Heterozygous mutations in SIX3 and SHH are associated with schizencephaly and further expand the clinical spectrum of holoprosencephaly. *Hum. Genet.* **2010**, *127*, 555–561.

23. Roessler, E.; El-Jaick, K. B.; Dubourg, C.; Vélez, J. I.; Solomon, B. D.; Pineda-Alvarez, D. E.; Lacbawan, F.; Zhou, N.; Ouspenskaia, M.; Paulussen, A.; Smeets, H. J.; Hehr, U.; Bendavid, C.; Bale, S.; Odent, S.; David, V.; Muenke, M. The mutational spectrum of holoprosencephaly-associated changes within the SHH gene in humans predicts loss-of-function through either key structural alterations of the ligand or its altered synthesis. *Hum. Mutat.* **2009**, *30*, E921–35.

24. Day, E. S.; Wen, D.; Garber, E. A.; Hong, J.; Avedissian, L. S.; Rayhorn, P.; Shen, W.; Zeng, C.; Bailey, V. R.; Reilly, J. O.; Roden, J. A.; Moore, C. B.; Williams, K. P.; Galdes, A.; Whitty, A.; Baker, D. P. Zinc-dependent structural stability of human Sonic hedgehog. *Biochemistry* **1999**, *38*, 14868–14880.

25. Fuse, N.; Maiti, T.; Wang, B.; Porter, J. A.; Hall, T. M.; Leahy, D. J.; Beachy, P. A. Sonic hedgehog protein signals not as a hydrolytic enzyme but as an apparent ligand for patched. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 10992–10999.

26. Hitzenberger, M.; Schuster, D.; Hofer, T. S. The Binding Mode of the Sonic Hedgehog Inhibitor Robotnikinin, a Combined Docking and QM/MM MD Study. *Front Chem* **2017**, *5*, 76.

27. Stanton, B. Z.; Peng, L. F.; Maloof, N.; Nakai, K.; Wang, X.; Duffner, J. L.; Taveras, K. M.; Hyman, J. M.; Lee, S. W.; Koehler, A. N.; Chen, J. K.; Fox, J. L.; Mandinova, A.; Schreiber, S. L. A small molecule that binds Hedgehog and blocks its signaling in human cells. *Nature chemical biology* **2009**, *5*, 154–156.

28. Himmelstein, D. S.; Cajigas, I.; Bi, C.; Clark, B. S.; Van Der Voort, G.; Kohtz, J. D. SHH E176/E177-Zn(2+) conformation is required for signaling at endogenous sites. *Developmental biology* **2017**, *424*, 221–235.

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

30. van Heijenoort, J. Peptidoglycan hydrolases of Escherichia coli. *Microbiol. Mol. Biol. Rev.* **2011**, *75*, 636–663.

31. Bochtler, M.; Odintsov, S. G.; Marcyjaniak, M.; Sabala, I. Similar active sites in lysostaphins and D-Ala-D-Ala metallopeptidases. *Protein Sci.* **2004**, *13*, 854–861.

32. Carrasco, H.; Olivares, G. H.; Faunes, F.; Oliva, C.; Larraín, J. Heparan sulfate proteoglycans exert positive and negative effects in Shh activity. *J. Cell. Biochem.* **2005**, *96*, 831–838.

33. Capurro, M. I.; Xu, P.; Shi, W.; Li, F.; Jia, A.; Filmus, J. Glypican-3 inhibits Hedgehog signaling during development by competing with patched for Hedgehog binding. *Developmental cell* **2008**, *14*, 700–711.

34. Witt, R. M.; Hecht, M.-L.; Pazyra-Murphy, M. F.; Cohen, S. M.; Noti, C.; van Kuppevelt, T. H.; Fuller, M.; Chan, J. A.; Hopwood, J. J.; Seeberger, P. H.; Segal, R. A. Heparan sulfate proteoglycans containing a glypican 5 core and 2-O-sulfo-iduronic acid function as Sonic Hedgehog co-receptors to promote proliferation. *The Journal of biological chemistry* **2013**, *288*, 26275–26288.

35. Siekmann, A. F.; Brand, M. Distinct tissue-specificity of three zebrafish ext1 genes encoding proteoglycan modifying enzymes and their relationship to somitic Sonic hedgehog signaling. *Dev. Dyn.* **2005**, *232*, 498–505.

36. Bellaiche, Y.; The, I.; Perrimon, N. Tout-velu is a Drosophila homologue of the putative tumour suppressor EXT-1 and is needed for Hh diffusion. *Nature* **1998**, *394*, 85–88.

37. Krauss, S.; Concordet, J. P.; Ingham, P. W. A functionally conserved homolog of the Drosophila segment polarity gene hh is expressed in tissues with polarizing activity in zebrafish embryos. *Cell* **1993**, *75*, 1431–1444.

38. Villarreal, C. M.; Darakananda, K.; Wang, V. R.; Jayaprakash, P. M.; Suzuki, Y. Hedgehog signaling regulates imaginal cell differentiation in a basally branching holometabolous insect. *Developmental biology* **2015**, *404*, 125–135.

39. Adamska, M.; Matus, D. Q.; Adamski, M.; Green, K.; Rokhsar, D. S.; Martindale, M. Q.; Degnan, B. M. The evolutionary origin of hedgehog proteins. *Curr. Biol.* **2007**, *17*, R836–7.

40. Fairclough, S. R.; Chen, Z.; Kramer, E.; Zeng, Q.; Young, S.; Robertson, H. M.; Begovic, E.; Richter, D. J.; Russ, C.; Westbrook, M. J.; Manning, G.; Lang, B. F.; Haas, B.; Nusbaum, C.; King, N. Premetazoan genome evolution and the regulation of cell differentiation in the choanoflagellate Salpingoeca rosetta. *Genome biology* **2013**, *14*, R15.

41. Pettigrew, C. A.; Asp, E.; Emerson, C. P. A new role for Hedgehogs in juxtacrine signaling. *Mechanisms of development* **2014**, *131*, 137–149.

42. Ochi, H.; Pearson, B. J.; Chuang, P.-T.; Hammerschmidt, M.; Westerfield, M. Hhip regulates zebrafish muscle development by both sequestering Hedgehog and modulating localization of Smoothened. *Developmental biology* **2006**, *297*, 127–140.