CHEMBIOCHEM

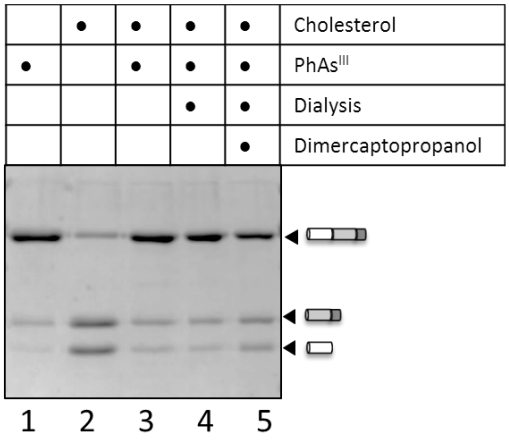
Supporting Information

Active Site Targeting of Hedgehog Precursor Protein with Phenylarsine Oxide

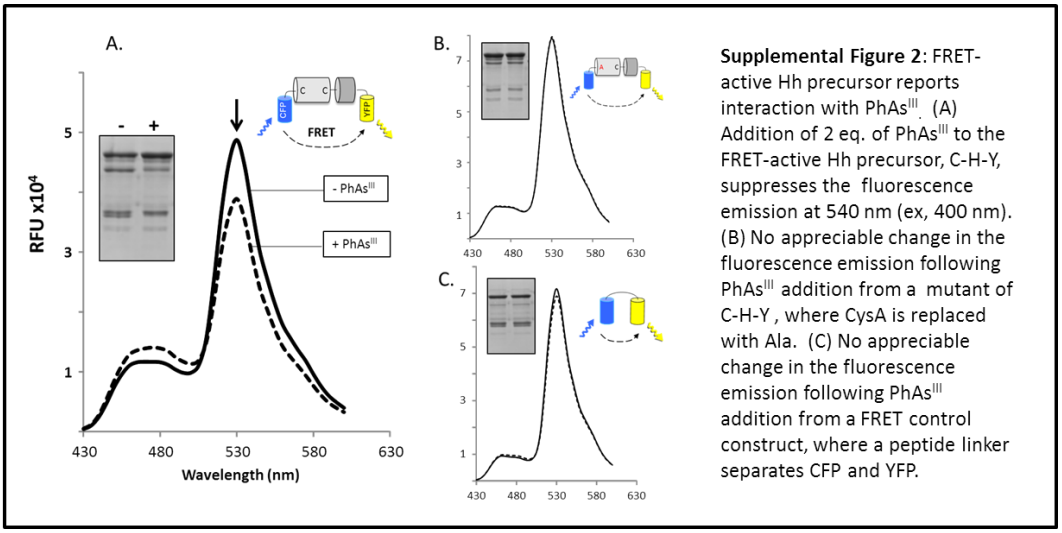
Timothy S. Owen,[a] Xie Jian Xie,[b] Benjamin Laraway,[a] George Ngoje,[a] Chunyu Wang,[b] and Brian P. Callahan*[a]

cbic_201402421_sm_miscellaneous_information.pdf

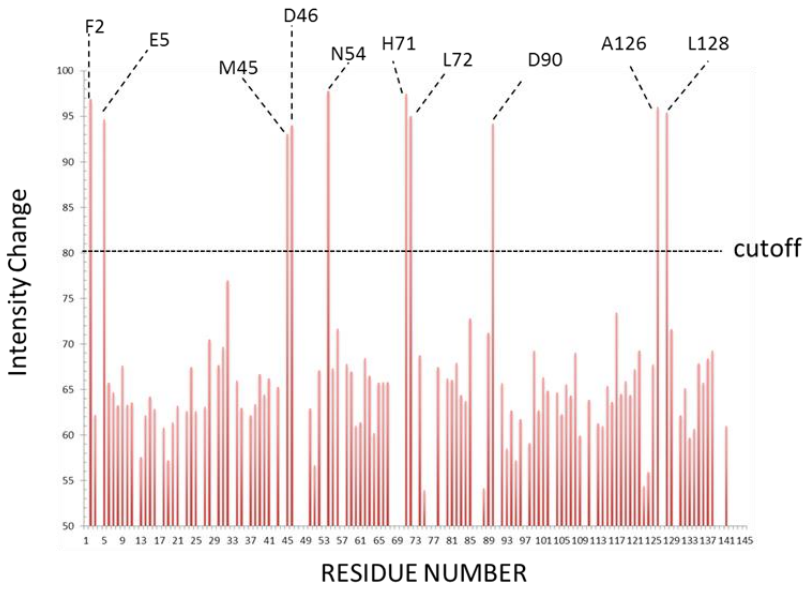
SUPPLEMENTAL FIGURES:



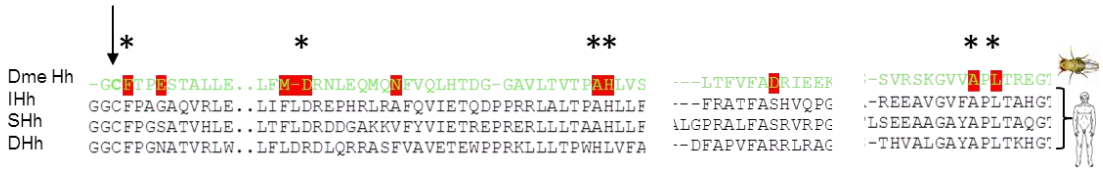
Supplemental Figure 1: Inhibition of Hh cholesterolysis by $PhAs^{III}$ can not be reversed by dialysis. Hedgehog precursor protein (4 µM), ShhN-DHhC, was incubated for 3 h with cholesterol (500 µM) in BENTT buffer in the absence (lane 2), and in the presence of added PhAsIII (80 µM) (lanes . The inhibitory effect of PhAsIII (lane 3), could not be reversed d by dialysis (lane 4). Partial activity was restored by adding the AsIII binding agent, dimercaptopropanol (250 µM), to the dialyzed protein (lane 5).



Supplemental Figure 2: FRET-active Hh precursor reports interaction with $PhAs^{III}$. (A) Addition of 2 eq. of $PhAs^{III}$ to the FRET-active Hh precursor, C-H-Y, suppresses the fluorescence emission at 540 nm (ex, 400 nm). (B) No appreciable change in the fluorescence emission following $PhAs^{III}$ addition from a mutant of C-H-Y , where CysA is replaced with Ala. (C) No appreciable change in the fluorescence emission following $PhAs^{III}$ addition from a FRET control construct, where a peptide linker separates CFP and YFP.



Supplemental Figure 3: Addition of 4 eq. of $PhAs^{III}$ to Dme-HINT domain (250 µM) markedly suppresses the ^{15}N chemical shift intensity for a subset of residues. Numbering arbitrarily begins with CysA at 1; equivalent to Cys258 in the full-length sequence Drosophila melanogaster Hh protein.



Supporting Figure 4. Residues at the $PhAs^{III}$ binding site are conserved in Human Hh proteins. Selected fragments of Drosophila melanogaster HINT domain align with HINT domains from Human Indian Hh (IHH), sonic Hh (Shh) and Desert Hh (DHh). Asterisks indicate conserved residues at the $PhAs^{III}$ binding site; arrow indicated CysA

TABLE 1

	CysA			CysB		
	^{13}C δ(ppm)	$^{1}H_A$ δ(ppm)	$^{1}H_B$ δ(ppm)	^{13}C δ(ppm)	^{1}H δ(ppm)	$^{1}H_B$ δ(ppm)
HINT	30.04	3.044	2.690	33.57	2.275	2.171
HINT – PAO	37.77	3.431	2.753	33.54	2.260	2.154
free cysteine	25.13	3.132	3.105	-	-	-

^{13}C NMR resonances of cysteine residues in the Dme HINT domain, along with ^{1}H NMR resonances of their diastereotopic protons. Experiments were performed in phosphate buffer (0.05 M, pH 7.1) with NaCl (0.1 M) and TCEP (0.002 M). Values for free cysteine obtained under identical conditions are provided for reference.

SUPPLEMENTAL METHODS:

Chemicals and plasmids: Arsenicals and cholesterol were purchased from Sigma Aldrich. CAUTION: *Arsenical compounds are toxic and should be handled with care.* TCEP (tris(2-carboxyethyl)phosphine) hydrochloride was from Hampton Research . All other chemicals (buffers, antibiotics, bacterial growth media, salts, and Triton-X 100), DNA polymerase, and restriction enzymes were obtained from Thermo Fisher Scientific unless otherwise indicated. Plasmids used in preparing the FRET reporter constructs, pBAD33 C-I-Y & pBAD33 C-Y, were kind gifts from Dr. Marlene Belfort (University at Albany).

DNA Constructs: The expression plasmid, pet-22b-ShhN-DhhC, encoding HhN of human Sonic HhN fused to HhC of *D. melanogaster* was prepared by ligating the synthetic gene fusion (Genscript) into pET-22b (Novagen) using unique NdeI and HindIII sites. For pet45b His_6-HhC, the coding sequence for *D. melanogaster* HhC fragment was PCR-amplified from genomic DNA (kindly provided by Dr. Robert Glaser, Wadsworth Center), digested with KpnI and HindIII, and ligated into pET-45b to generate an N-terminal His-tagged construct. Primers sequences, with restriction sites underlined were as follows: FWD: TTTCACGTGGGTACCGGTGAAAACCTGTATTTTCAGGGCTCGACGGTGCATGGCTGCTTC; REV: TTTAAGCTTAATCGTGGCGCCAGCTCTGCGGCAGAACG.
The pBAD33 C-H-Y construct was prepared from pBAD33 C-I-Y [1], by replacing the "I" fragment with *D. melanogaster* HhC, (H), using engineered SalI and PstI sites flanking HhC and the vector's unique XhoI and PstI sites. Primer sequences are as follows: FWD-SalI- 5'-AAAGTCGACGGTGCATGGCTGCTTCACGCCGGAGAGCACAGCGC; REV-PstI-5' AAACTGCAGGCCGCCATCGTGGCGCCAGCTCTGCGGCAGAACG. A control construct, where Cys1 is replaced with Ala was amplified using the mutagenic forward primer: 2896-SalI- 5'-AAAGTCGACGGTGCATGGCGCCTTCACGCCGGAGAGCACAGCGC and the same reverse primer.

Protein production: All proteins were expressed in soluble form in *E. coli* and purified by Ni^{2+}-NTA chromatography. pET vectors encoding SHhN-DHhC, Dme His_6-HhC, and Dme His_6-HINT were expressed at 16 $^{\circ}C$ in *E. coli* strain BL21-DE3, cultured in TB-MGB [2], terrific broth, and minimal media, respectively. FRET constructs, pBAD33 C-H-Y, C-(C1A)H-Y, and C-Y, were expressed in *E. coli* strain LMG194 (Invitrogen) grown in LB media as described [1]. In general, expression cultures were harvested 22 h after induction at 18 $^{\circ}C$. Pellets were resuspended in lysis buffer (0.05 M KH_2PO_4, 0.4 M NaCl, 0.1 M KCl, 10% Glycerol, 0.5% Triton X-100, 0.01 M imidazole) at pH 7.8 supplemented with lysozyme 10 µg/mL and DNAase 1 µg/mL, and subjected to 2-3 freeze-thaw cycles. Purification from clarified lysate was carried out with a Ni^{2+}-NTA column (3 mL) according to manufacturer's protocol. Yields from 1 L culture were ~ 14 mgs, with 80% purity. Proteins were quantified by Bradford Assay (Pierce Biotech).

Mass Spectrometry: Ni^{2+}-NTA purified Dme HhC protein (10 µM) plus/minus $PhAs^{III}$ (1 eq.) were analyzed at the Cornell Proteomics Facility using a Waters Synapt HDMS mass spectrometer. Samples were desalted with a Zeba spin column (7 kDa, cutoff), acidified with formic acid to pH ~3.0, then injected onto a Waters MassPREP Micro Desalting Column (P/N 186004032) installed on a Waters Nano Acquity UPLC system. Following elution with a binary water/acetonitrile/formic acid gradient, proteins were introduced into the Waters Synapt HDMS mass spectrometer. Deconvoluted mass is expected to be within ~1Da of the actual molecular weight based on calibrations with bovine trypsinogen.

Arsenic inhibition of Hh Cholesterolysis: Activity of the hedgehog precursor, ShhN-DhhC, was assayed in BENTT buffer, pH 7.1 (BisTris 0.05 M; EDTA, 0.005 M; NaCl, 0.1 M; Triton X-100, 0.4 %; TCEP, 0.002 M). For the initial

screen, representative pentavalent and trivalent arsenicals were added (80 µM, final) to precursor (4 µM, final) in BENTT buffer at 25 °C. Following 1 h incubation, cholesterol was added (500 µM, final) to initiate processing. At selected intervals aliquots were mixed with SDS-PAGE loading buffer, and boiled to stop the reaction. Following separation by 15% SDS-PAGE and Coomassie blue staining, the extent of reaction was determined by the intensity of product bands (SHhN and DHhC) relative to precursor. Gels were imaged and analyzed using Gel Doc™ EZ imager (BioRad) equipped with Image Lab software.

$$Ratio\ Processed = \frac{SHhN + DHhC}{Precursor}$$

Controls reactions included samples of unreacted precursor, and precursor reacted with hydroxylamine (0.3 M).

Titration of ShhN-DhhC with the inhibitory organoarsenical, $PhAs^{III}$, was carried out and analyzed in a similar manner. Cholesterolysis activity was calculated by dividing the extent of processing at a given [$PhAs^{III}$] divided by the extent of processing in the absence of $PhAs^{III}$. To calculate an apparent EC_{50} value, fractional activity as a function of [$PhAs^{III}$] was fit to the dose-response equation below:

$$Activity_{calc} = max + \frac{(min - max) * [\mathrm{PhAsIII}]}{EC_{50} + [\mathrm{PhAsIII}]}$$

The same data was also fit using the quadratic equation for tight binding inhibition, where [E] and K_i represent the precursor concentration and apparent inhibition constant respectively.

$$Fractional\ Activity = Min + \left\{Max * \left[\frac{[E] - [\mathrm{PhAsIII}] - K_i + \sqrt{([E] - [PAO] - K_i)^2 + 4[E]K_i}}{2[E]}\right]\right\}$$

Values for EC50 and K_i were obtained by nonlinear regression using X-cell solver.

$PhAs^{III}$ Binding by FRET quenching: Purified FRET active precursor C-H-Y (2 µM), and controls, C-(C1A)H-Y and C-Y in BENT2 buffer (98 µl) were loaded into wells of a Corning NBS 96-well Microplate, then mixed with 2 ul of ethanol +/- phenylarsenic oxide (4 µM, final). Following 1 h incubation, fluorescence from the samples was measured using a BioTek Synergy H1 plate reader with an excitation wavelength of 400 nm, and emission from 420 to 600 nm.

Kinetics of $PhAs^{III}$ binding by FRET quenching were measured using a similar experimental setup except that fluorescent measurements were recorded immediately after addition of $PhAs^{III}$ to C-H-Y (0.2 µM), with an excitation wavelength of 400 nm and emission wavelengths were fixed at 460 and 540. FRET values were calculated as follows:

$$FRET = \frac{Fluorescence\ at\ 540nm}{Fluorescence\ at\ 460nm}$$

Experiments were carried out in triplicate. Pseudo first order rate constants (k_{obs}) at each [$PhAs^{III}$] were calculated as follows:

$$FRET_{time} = A * e^{-k_{obs} * t} * C$$

with best fit values for A, k_{obs} and C obtained by nonlinear regression using Excel solver. The curvature apparent in plots of the rate constants (k_{obs}) as a function of increasing [$PhAs^{III}$] suggested the following kinetic scheme:

$$\text{C-H-Y} + \text{PhAs}^{\text{III}} \underset{}{\overset{K_d}{\rightleftharpoons}} \left\{\text{C-H-Y : PhAs}^{\text{III}}\right\} \xrightarrow{k_{inact}} \text{C-H-Y} - \text{PhAs}^{\text{III}}$$

To find the maximum rate of covalent modification (k_{max}) and the apparent dissociation constant for $PhAs^{III}$ / CHY encounter complex (K_d) using the equation:

$$k_{obs} = \frac{k_{max} * [\text{PhAsIII}]}{K_d + [\text{PhAsIII}]}$$

The change in FRET from C-H-Y with added $PhAs^{III}$ was also used to calculate dissociation constant using a modified quadratic equation for tight binding, where [E] and K_i represent the precursor concentration and apparent inhibition constant respectively, and $\Delta FRET_{max}$ represents the maximum change in FRET at saturating $PhAs^{III}$.

$$\Delta FRET_{obs} = (\Delta FRET_{max}) * \frac{[E] - [PAO] - EC_{50} + \sqrt{([E] - [PAO] - K_i)^2 + 4[E]EC_{50}}}{2[E]}$$

Enrichment with ^{13}C at Cysteine: The HINT domain from *Drosophila melanogaster* Hh was expressed [3] in a cysteine-auxotroph E. coli strain, BL21-(DE3)cysE [4], kindly provided by Dr. Marie-Paule Strub (NIH), to achieve enrichment with ^{13}C at CysA and CysB. Cultures were seeded from a single colony into LB media with ampicillin (100 µg/ml) and grown at 37 °C to an optical density (OD) of 0.1. Cells were collected by centrifugation (5000 xg for 15 min), washed with M9 minimal media three times, then resuspended in minimal media with glucose (0.4%) and 19 unlabeled amino acids (100 µg/ml), plus C-3-^{13}C cysteine (50 µg/ml) (Cambridge Isotopes). The culture was incubated at 37 °C until reaching OD_{600} of 0.6, and then cooled to 16 °C before inducing with IPTG (2 mM). Cells were harvested after 18 hrs, and the protein was purified by Ni^{2+}-NTA chromatography, as described above.

$PhAs^{III}$ binding by NMR: Samples were prepared from recombinant ^{15}N-labeled, or ^{13}C-Cys labelled HINT domain from *Drosophila melanogaster* as described above and [3]. Prior to analysis, an N-terminal peptide containing the His_6 purification tag was cleaved off from the HINT domain via thiolysis of the internal thioester at CysA. Cleaved HINT domain was then dialyzed into 50 mM Tris buffer (pH 7.1), 100 mM NaCl, 2 mM TCEP, concentrated to ~250 µM, then D_2O was added to 10%. $PhAs^{III}$ was solubilized in Ethanol-d5 (Cambridge isotopes) to create a stock solution of 25 mM. Chemical shift changes were monitored using ^{15}N–^{1}H-HSQC or ^{13}C–^{1}H-HSQC spectra at increasing molar ratios of $PhAs^{III}$ to HINT protein. A control sample of Ethanol-d5 without $PhAs^{III}$ was also titrated into HINT samples as a solvent control. Chemical shift perturbation was calculated as a percent using the change in normalized intensity with added $PhAs^{III}$, as follows, (1-Intensity$^{+PhAsIII}$/ Intensity$^{-PhAsIII}$)*100. All NMR experiments were performed on an 800 or 600 MHz spectrometer equipped with a cryogenic probe at 298 K.

Sequences of proteins used in this study:

(1) ShhN-DhhC

MIIGPGRGFGKRRHPKKLTPLAYKQFIPNVAEKTLGASGRYEGKISRNSERFKELTPNYNPDIIFKDEENTGADRLMTQRCKDKLNALAISVMNQWPGVKLRVTEGWDEDGHHSEESLHYEGRAVDITTSDRDRSKYGMLARLAVEAGFDWVYYESKAHIHCSVKAENSVAAKSGGHG^**CFTPESTALLESGVRKPLGELSIGDRVLSMTANGQAVYSEVILFMDRNLEQMQNFVQLHTDGGAVLTVTPAHLVSVWQPESQKLTFVFADRIEEKNQVLVRDVETGELRPQRVVKVGSVRSKGVVAPLTREGTIVVNSVAASCYAVINSQSLAHWGLAPMRLLSTLEAWLPAKEQLHSSPKVVSSAQQQNGIHWYANALYKVKDYVLPQSWRHD***

*Notes: ShhN is underlined; DHhC is in bold; ^ indicates autocleavage site.

(1) C-H-Y

MSKGEELFGGIVPILVELEGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTLTWGVQCFSRYPDHMKQHDFFKSVMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYISHNVYITADKQKNGIKANFKARHNITDGSVQLADHYQQNTPIGDGPVILPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITHGMDELYKGSGGSTVHG^**CFTPESTALLESGVRKPLGELSIGDRVLSMTANGQAVYSEVILFMDRNLEQMQNFVQLHTDGGAVLTVTPAHLVSVWQPESQKLTFVFADRIEEKSQVLVRDVETGELRPQRVVKVGSVRSKGVVAPLTREGTIVVNSVAASCYAVINSQSLAHWGLAPMRLLSTLEAWLPAKEQLHSSPKVVSSAQQQNGIHWYANALYKVKDYVLPQSWRHD**GGLQGSGGSMSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKLLCTTGKLPVPWPTLVTTLGYGVQCFARYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYITADKQKNGIKANFKIRHNIEDGGVQLADHYQQNTPIGDGPVLLPDNHYLSYQSALFKDPNEKRDHMVLLEFLTAAGITEGMNELYKHHHHHH

*Notes: CFP is in blue; DHhC is in bold; YFP is in orange; ^ indicates autocleavage site.

(2) Dme-HhC

MAHHHHHHVGTGENLYFQGSTVHG^**CFTPESTALLESGVRKPLGELSIGDRVLSMTANGQAVYSEVILFMDRNLEQMQNFVQLHTDGGAVLTVTPAHLVSVWQPESQKLTFVFADRIEEKNQVLVRDVETGELRPQRVVKVGSVRSKGVVAPLTREGTIVVNSVAASCYAVINSQSLAHWGLAPMRLLSTLEAWLPAKEQLHSSPKVVSSAQQQNGIHWYANALYKVKDYVLPQSWRHD**

* DHhC is in bold; ^ indicates autocleavage site.

(2) Dme-HINT

MAHHHHHHVHG^**CFTPESTALLESGVRKPLGELSIGDRVLSMTANGQAVYSEVILFMDRNLEQMQNFVQLHTDGGAVLTVTPAHLVSVWQPESQKLTFVFADRIEEKNQVLVRDVETGELRPQRVVKVGSVRSKGVVAPLTREGTIVVNSVAASCYAV**

* HINT is in bold; ^ indicates autocleavage site; amino-terminal sequence, MAHHHHHHVHG, was removed via thiolysis with dithiothreitol prior to NMR analysis.

REFERENCES

[1] G. Amitai, B. P. Callahan, M. J. Stanger, G. Belfort, M. Belfort, *Proceedings of the National Academy of Sciences, USA* **2009**, *106*, 11005-11010.

[2] D. P. Baker, F. R. Taylor, R. B. Pepinsky, *Methods Mol Biol* **2007**, *397*, 1-22.

[3] J. Xie, Z. Du, B. Callahan, M. Belfort, C. Wang, *Biomol NMR Assign* **2013**.

[4] aS. Muller, H. Senn, B. Gsell, W. Vetter, C. Baron, A. Bock, *Biochemistry* **1994**, *33*, 3404-3412; bM. P. Strub, F. Hoh, J. F. Sanchez, J. M. Strub, A. Bock, A. Aumelas, C. Dumas, *Structure* **2003**, *11*, 1359-1367; cF. Arnesano, L. Banci, I. Bertini, I. C. Felli, M. Losacco, G. Natile, *J Am Chem Soc* **2011**, *133*, 18361-18369.