A Single Mutation Unlocks Cascading Exaptations in the Origin of a Potent Pitviper Neurotoxin

A. Carl Whittington,*[,1] Andrew J. Mason,[†,2] and Darin R. Rokyta[1]

[1]Department of Biological Science, Florida State University, Tallahassee, FL

[2]Department of Biology, University of Central Florida, Orlando, FL

[†]Present address: Department of Biological Sciences, Clemson University, Clemson, SC

***Corresponding author:** E-mail: awhittington@bio.fsu.edu.

Associate editor: Claus Wilke

Abstract

Evolutionary innovations and complex phenotypes seemingly require an improbable amount of genetic change to evolve. Rattlesnakes display two dramatically different venom phenotypes. Type I venoms are hemorrhagic with low systemic toxicity and high expression of tissue-destroying snake venom metalloproteinases. Type II venoms are highly neurotoxic and lack snake venom metalloproteinase expression and associated hemorrhagic activity. This dichotomy hinges on Mojave toxin (MTx), a phospholipase A_2 (PLA2) based β-neurotoxin expressed in Type II venoms. MTx is comprised of a nontoxic acidic subunit that undergoes extensive proteolytic processing and allosterically regulates activity of a neurotoxic basic subunit. Evolution of the acidic subunit presents an evolutionary challenge because the need for high expression of a nontoxic venom component and the proteolytic machinery required for processing suggests genetic changes of seemingly little immediate benefit to fitness. We showed that MTx evolved through a cascading series of exaptations unlocked by a single nucleotide change. The evolution of one new cleavage site in the acidic subunit unmasked buried cleavage sites already present in ancestral PLA2s, enabling proteolytic processing. Snake venom serine proteases, already present in the venom to disrupt prey hemostasis, possess the requisite specificities for MTx acidic subunit proteolysis. The dimerization interface between MTx subunits evolved by exploiting a latent, but masked, hydrophobic interaction between ancestral PLA2s. The evolution of MTx through exaptation of existing functional and structural features suggests complex phenotypes that depend on evolutionary innovations can arise from minimal genetic change enabled by prior evolution.

Key words: **gene duplication, evolutionary biophysics, evolutionary innovation, exaptation, venom.**

Introduction

Complex phenotypes and evolutionary innovations appear to necessitate passage through periods of low fitness to realize novel, fitness-improving functions (Fisher 1930; Orr 2005), but paths circumventing apparent low-fitness intermediates must, in some instances, be possible (Tawfik 2010; Wagner 2011; Pal 2017). The animal eye (Darwin 1859; Lamb et al. 2007) evolved by means of many incremental changes, each favored by selection (Nilsson and Pelger 1994), rendering the journey through a supposed fitness valley into a gradual climb up a hill. Large-effect mutations, in contrast, could bypass low-fitness intermediates altogether. In insect Batesian mimicry, for example, some phenotypes may evolve through single, large-effect mutations, followed by small-effect mutations that optimize the new phenotype (Norrström et al. 2006; Rubinoff et al. 2008; Gamberale-Stille et al. 2012; Leimar et al. 2012). Although such large-effect mutations are predicted to be rare and most likely deleterious (Fisher 1930; Orr 2005), in some cases, prior neutral evolution may render the crossing of a fitness valley via a large-effect mutation more likely.

Large-effect, function-shifting mutations that are deleterious in isolation can become beneficial in the presence of prior, permissive mutations (epistasis) that are neutral themselves, but shape the subsequent evolutionary histories of proteins and organisms (Dean and Thornton 2007; Breen et al. 2012; Harms and Thornton 2013; Boucher et al. 2014; Shah et al. 2015; Tufts et al. 2015; Starr and Thornton 2016; Steindel et al. 2016; Storz 2016; Hochberg and Thornton 2017; Kumar et al. 2017). Such a process occurred during the early evolution of the vertebrate glucocorticoid receptor (GR) from an ancestral mineralocorticoid receptor. After a duplication of the ancestral receptor gene, a mutation that was functionally neutral arose that permitted a subsequent, large-effect mutation that altered GR structure and shifted ligand specificity from aldosterone in the ancestor to cortisol in GR (Bridgham et al. 2006, 2009; Ortlund et al. 2007; Harms and Thornton 2014). Similarly, historical, potentiating mutations in a lineage of *Escherichia coli* in the long-term evolution experiment allowed a dramatic shift in resource acquisition ability in later generations (Blount et al. 2008, 2012). These experimental examples suggest certain protein-based evolutionary innovations and complex phenotypes may be contingent on, and even predisposed by, their evolutionary history.

Snake venoms are complex, protein-based phenotypes that can show extreme functional specialization (Mackessy 1988, 2008, 2009; Fry 2005; Castoe and Parkinson 2006;

Article

Sanz et al. 2006; Fry et al. 2007; Gibbs and Rossiter 2008; Gibbs and Mackessy 2009; Calvete et al. 2010; Rokyta et al. 2011, 2012, 2013, 2017; Castoe et al. 2012; Calvete 2013, 2017; Margres et al. 2013, 2014, 2015, 2017; Martínez-Romero et al. 2013; McGivern et al. 2014; Rokyta, Margres, et al. 2015; Rokyta, Wray, et al. 2015; Wray et al. 2015; Dowell et al. 2016; Margres, Walls, et al. 2016; Margres, Wray, et al. 2016). Rattlesnakes (genera *Crotalus* and *Sistrurus*) show a dramatic dichotomy in venom phenotype, which is usually polymorphic within species (Glenn et al. 1994; Wooldridge et al. 2001; Mackessy 2008; Calvete et al. 2010; Rokyta et al. 2013; Rokyta, Margres, et al. 2015), and appears to have arisen through rapid genetic change. Most rattlesnake venoms, Type I, are predominantly hemorrhagic, with low systemic toxicity and high expression of tissue-destroying snake venom metalloproteinases (SVMPs; Mackessy 2008). Type II venoms are highly neurotoxic with completely different patterns of gene expression, including loss of most SVMP activity (Mackessy 2008). Type I venoms have high metalloproteinase activity and high median lethal dose (LD_{50}) values (>1.0 μg/g mouse body weight), whereas type II venoms have low metalloproteinase activity and low LD_{50} values (<1.0 μg/g mouse body weight; Mackessy 2008). The active, featured component of Type II venoms is Mojave toxin (MTx) from *Crotalus scutulatus* and its homologs in other species (hereafter all referred to as MTx; fig. 1*A*). These β-neurotoxins (Doley et al. 2009) lead to rapid onset of paresthesia, weakness, fasciculation Clark et al. (1997), and even respiratory paralysis in envenomated humans (Jansen et al. 1992). MTx consists of two phospholipase A_2 (PLA2; Six and Dennis 2000; Berg et al. 2001) subunits, a negatively charged acidic subunit and positively charged basic subunit (Faure et al. 2011). Type II venoms are epitomized by the extremely simple tiger rattlesnake (*Crotalus tigris*) venom that contains high concentrations of MTx (66% of the total venom protein), forming the most potent rattlesnake venom known (Calvete et al. 2012). In contrast, individuals displaying a rare venom phenotype found among some populations of the timber rattlesnake (*Crotalus horridus*) lack expression of SVMPs and the two MTx subunits (Glenn et al. 1994). The venom from these individuals is nonhemorrhagic and non-neurotoxic with an extremely high LD_{50} value (>20 μg/g mouse body weight). These two extreme examples underscore the importance of the role of MTx in rattlesnake venom phenotypes.

The transcriptomic, proteomic, and subsequent phenotypic differences between Type I and Type II venoms appear to hinge on the expression and processing of the MTx acidic subunit (Glenn et al. 1994; Wooldridge et al. 2001; Mackessy 2008; Calvete et al. 2010; Rokyta et al. 2013; Rokyta, Margres, et al. 2015). The acidic subunit of MTx is a dual allosteric regulator of the weakly toxic basic subunit, an active lipase. The acidic subunit is both an agonist of neurotoxicity, chaperoning the basic subunit after prey envenomation to ensure interaction with target receptors and toxicity, and an antagonist of the lipase activity of the basic subunit, which interferes with the neurotoxic function (Radvanyi and Bon 1982). Without the acidic subunit, the basic subunit is only weakly toxic, and stability of the interaction between the two subunits correlates with neurotoxicity (Faure et al. 1993). The acidic subunit is non-neurotoxic in isolation and has lost its ancestral, hemorrhagic lipase activity due to proteolytic processing at five sites on the PLA2 precursor. The origin of MTx presents an evolutionary challenge because it would appear to require independent genetic changes to produce large amounts of acidic subunit that, in isolation, do not contribute to venom toxicity. These changes include multiple mutations to yield five different proteolytic cleavage sites in the acidic subunit and production of the proteases to cleave at the sites. This profusion of genetic change and production of large amounts of a nonfunctional protein conferring little benefit to venom function suggests passage through a deep fitness valley in the origin of MTx, the main driver of the rattlesnake venom dichotomy.

We used transcriptomics to assess the diversity of PLA2 venom toxins in North American pitvipers. This allowed us to place the MTx subunits into a phylogenetic context and enabled ancestral sequence reconstruction of extinct PLA2 venom toxins. We analyzed extinct and extant sequences to elucidate the series of genetic events that occurred during MTx evolution and to determine the role prior genetic variation may have played in MTx evolution. We modeled structures of extinct and extant PLA2 toxins to understand the biophysical mechanisms underlying MTx evolution. We used a series of selection tests to ascertain the role of positive and episodic selection in MTx evolution. Integration of these various bioinformatic approaches allowed us to reconstruct the evolutionary origin and history of MTx, the most potent of all rattlesnake venom toxins.

Results and Discussion

PLA2 Toxins in North American Pitvipers

Recent work suggests a single origin for MTx homologs (fig. 1*A*; Wüster et al. 2008; Yang, Yang, et al. 2015; Dowell et al. 2016; Calvete 2017). The evolutionary events after this origination that led to proteolytic processing of the acidic subunit, interaction of MTx subunits, and potent neurotoxicity are unclear. To determine the molecular-evolutionary history of the MTx subunits and to facilitate estimation of ancestral PLA2 sequences, we assessed PLA2 diversity in venom gland transcriptomes of North American pitvipers, including three MTx positive species and ten MTx negative species. This included one individual each from previously described Type I and Type II populations of *C. horridus* (Rokyta et al. 2013; Rokyta, Margres, et al. 2015). Our estimated PLA2 gene phylogeny showed five well-defined clades corresponding to the set of ancestral rattlesnake genes (fig. 1*B*).

Our phylogeny reflects the structural and functional diversity in PLA2 toxins. We estimated and mapped isoelectric points of both extant and reconstructed extinct PLA2s onto the tree to show the dichotomy between acidic and basic PLA2s resulting from the gene duplication events (fig. 1*B*). In general, basic PLA2s have lower lipolytic activity, but are much more toxic than the acidic PLA2s, which show little systemic toxicity with high catalytic activity that may

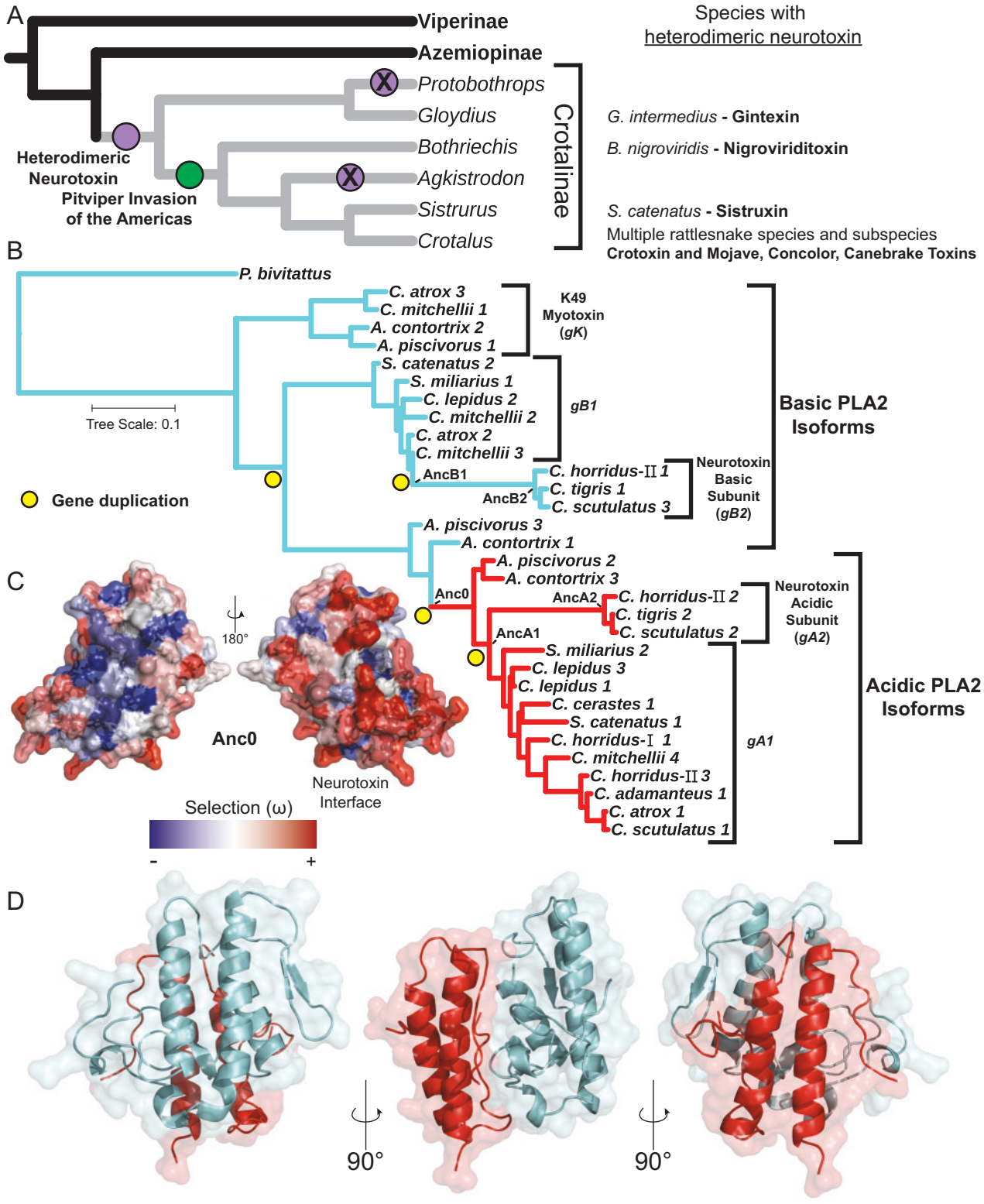

Fig. 1. The origin and evolution of MTx. (*A*) MTx evolved in an Asian pitviper ancestor (purple circle) before the invasion of the Americas (green circle) with subsequent, lineage-specific loss (purple circles with X). (*B*) Gene-duplication events (yellow spheres) and naming correspond to Dowell et al. (2016). The first event gave rise to the basic PLA2s, including *gK* with an Asp to Lys substitution at position 49 referred to as myotoxin (Lomonte et al. 2003; Lomonte and Rangel 2012). A second duplication yielded two basic lipases, *gB1* with potential edema producing activity (Tsai et al. 2003; Gibbs and Rossiter 2008), and *gB2*, the basic subunit of MTx. A third and fourth duplication gave rise to the acidic PLA2s, *gA*, which include *gA1* with anti-coagulant/hemolytic activity (Tsai et al. 2003), and *gA2*, the acidic subunit of MTx. Ancestral PLA2s are noted. (*C*) Site-specific d*N*/d*S* ratios (ω) mapped to the ancestral PLA2 (Anc0) structure. The surface that evolved into the MTx interface showed evidence of positive selection. The solvent-exposed side showed purifying selection. (*D*) The ancestral acidic PLA2 (AncA2; red) is proteolytically cleaved into three peptides linked by seven disulfide bonds. The acidic subunit binds noncovalently to, and acts as an allosteric regulator of, the ancestral basic PLA2 (AncB2; cyan).

play a digestive role (Rosenberg 1986; Križaj et al. 1993; Lomonte et al. 2009; Jiménez-Charris et al. 2016). Our reconstructed, extinct PLA2s had predicted isoelectric points that matched the descendant sequences, suggesting these ancestral sequences are realistic models of ancestral PLA2s. The first duplication event gave rise to the basic PLA2s, including a PLA2 toxin, *gK*, with a characteristic Asp to Lys substitution at position 49 referred to as myotoxin (Lomonte et al. 2003; Lomonte and Rangel 2012). A second duplication yielded two basic lipases, one of which, *gB2*, evolved to form the basic subunit of MTx. A third and fourth duplication gave rise to the acidic PLA2s, *gA*, including the acidic subunit of MTx, *gA2*. Our phylogenetic analyses demonstrated that PLA2 genetic diversity correlates with functional diversity (fig. 1*B*). In particular, the long branches leading to the MTx acidic subunit (31 amino acid substitutions) and basic subunit (32 amino acid substitutions) highlight the large amount of genetic change that was required for the evolution of extant MTx. This presumably included multiple mutations to yield the correct proteolytic processing of the acidic subunit while leaving the basic subunit intact, and the formation of a binding site between the processed, nonhemorrhagic, and nonneurotoxic acidic subunit and the basic subunit.

A MTx ortholog was present in the ancestral Asian pitviper (Yang, Yang, et al. 2015; Dowell et al. 2016; Calvete 2017) that invaded North America ca. 22 Ma (Wüster et al. 2008) and in the rattlesnake ancestor that arose ca. 12–14 Ma in the Sierra Occidental Mountains in the north-central Mexican Plateau (fig. 1*A*; Castoe and Parkinson 2006). The ancestral rattlesnake had five venom-expressed PLA2 loci resulting from a series of duplications of an ancestral PLA2 gene, and the pattern in extant species is likely due to lineage specific gene loss (Dowell et al. 2016). Our results for *Crotalus adamanteus*, *Crotalus atrox*, and *C. scutulatus* (fig. 1*B*) recapitulated the genomic-based results of Dowell et al. (2016), giving confidence our transcriptomic analysis fully captured PLA2 genetic diversity. Additionally, our phylogeny showed monophyletic clades for the acidic and basic MTx subunits, which is consistent with a single origin of MTx and its homologs.

Proteolytic Processing of MTx Acidic Subunit Originated from a Single Mutation

Correct proteolytic processing of the acidic subunit may have resulted from exaptation of snake venom serine proteases (SVSPs) already present in snake venoms that interfere with the hemostatic system in prey (fig. 2*B*; Serrano and Maroun 2005). None of the PLA2s identified in our transcriptomic analysis showed bioinformatic evidence of post-translational modifications (Materials and Methods) which could provide protection from proteolytic cleavage, nor did they show evidence of cleavage by proprotein convertase enzymes, indicating they are not specifically processed by endogenous Golgi complex enzymes. Additionally, SVMPs are not typically expressed at high levels in venoms that express MTx (Glenn et al. 1994; Mackessy 2008, 2009; Rokyta et al. 2013; Rokyta, Margres, et al. 2015), eliminating SVMPs as the proteolytic machinery for acidic subunit processing.

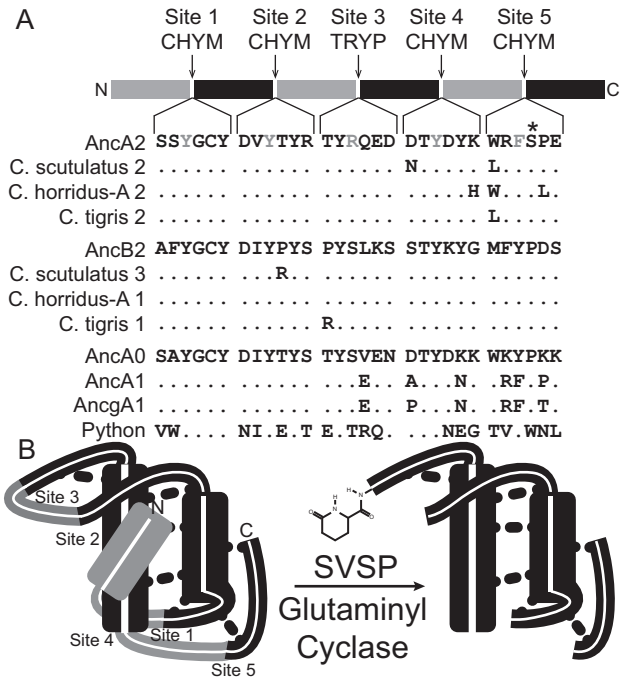

Fig. 2. A series of exaptations facilitated proteolytic processing of the MTx acidic subunit. (*A*) Cleavage site recognition residues (gray letters) are highly conserved across snake PLA2 evolution and cleavage sites were present prior to the evolution of MTx. Cleavage required a CHYM and TRYP specificity protease. Along the acidic subunit lineage—AncA0 → AncA1 → AncA2—there was a critical mutation after the gene duplication of AncA1 at site 5 (asterisk). In the ancestor of the MTx acidic subunits, AncA2, this mutation led to the replacement of a conserved proline with a serine adjacent to the recognition residue. The proline is conserved in the ancestral (AncB2) and extant MTx basic subunits. In each column, dots represent an identical residue to the ancestral reference sequence. (*B*) Processing was facilitated by two other exaptations. Prior to the evolution of the MTx acidic subunit, snake venoms already contained proteases, SVSPs, that had the necessary and sufficient specificity for correct processing of the acidic subunit. Processing is likely completed by cyclization of the N-terminal Gln at site 3 by glutiminyl cyclase already present in rattlesnake venoms. Dashed lines represent disulfide bonds. N and C indicate the amino- and carboxy-terminus of the protein, respectively.

To test whether the existing SVSPs in snake venoms provided the requisite specificity to cleave the acidic MTx subunit, we determined the full diversity of SVSPs in our venom gland transcriptomes and grouped them into classes based on canonical serine protease specificity determined by the protease S1 residue (for review, see Kay 1993; Perona and Craik 2008; Hedstrom 2002). Based on sequence specificity (fig. 2*A*), two different proteases with orthogonal specificity were necessary for acidic subunit proteolytic processing. Trypsin-like proteases (TRYP) with an aspartate at the S1 position cleave after positively charged recognition residues (arginine or lysine) in the target protein. Chymotrypsin-like proteases (CHYM-A) with a serine in the S1 position cleave after bulky hydrophobic recognition residues (phenylalanine, tryptophan, tyrosine), but not before proline (Kay 1993). A third type of SVSPs (CHYM-B) contains a glycine or

asparagine at the S1 position and likely have similar specificity for cleavage after bulky hydrophobic recognition residues. Sites 1, 2, 4, and 5 in the ancestral acidic subunit (AncA2) could all be cleaved by a CHYM-A or CHYM-B specificity protease, hereafter collectively referred to as CHYM, and site 3 could be cleaved by a TRYP SVSP (fig. 2A). A complement of TRYP and CHYM SVSPs, present in all of the snakes sampled in this study (supplementary table S1, Supplementary Material online), is sufficient for acidic subunit processing. The expression levels of SVSPs (supplementary fig. S1, Supplementary Material online) appear to be independent of MTx expression. Type I and Type II *C. horridus* individuals have similar levels of TRYP and CHYM-A transcripts whether MTx is expressed (I: ~96% and ~4%), or not (II: ~95% and ~5%). These results suggest that, in addition to being present before the evolution of MTx, any rattlesnake venom (of those we characterized) could correctly and efficiently process the extant acidic subunit of MTx and its homologs.

Among ancestral and extant PLA2s, including the PLA2 from the nonvenomous *Python bivittatus*, the SVSP recognition residues at the cleavage sites are highly conserved (gray residues; fig. 2A), suggesting that they have an adaptive role in PLA2 structure and function and that their role as cleavage sites during acidic subunit processing was an exaptation. However, the high sequence conservation at these cleavage sites also suggests there would be minimal differences in proteolytic processing of venom PLA2s, including the basic subunit of MTx and other PLA2s from Type I venoms. Our structural model of the ancestral MTx acidic subunit, AncA2, pro-protein (figs. 2A and 3A) shares the common venom PLA2 structure (Marchi-Salvador et al. 2008; Faure et al. 2011). Our AncA2 model showed that the physical arrangement of sites 1–5, and, specifically, the sequence of site 5 in PLA2s other than the MTx acidic subunit formed a physical barrier to cleavage (fig. 2A). The results of our phylogeny-based ancestral sequences and structural modeling suggested a single nonsynonymous mutation in the acidic subunit controlled its cleavage pattern while leaving other venom PLA2 toxins intact.

The critical mutation occurred at a site adjacent to the SVSP recognition residue at site 5 (fig. 2A). This mutation caused the substitution of a cleavage-blocking proline (Kay 1993), present in ancestral PLA2s and extant basic MTx subunits, with a serine in the acidic MTx subunit. Estimation of site-specific solvent accessible surface area (SASA; supplementary table S2, Supplementary Material online; fig. 3A) showed that site 1 is physically buried behind a region containing sites 4 and 5. Site 5 has high SASA and is likely cleaved first followed by site 4, which exposes site 1 for cleavage. Both sites 1 and 4 were pre-existing cleavage sites that did not require mutations (fig. 2A), and the cleavage cascade is likely the result of the single mutation at site 5.

Cleavage at sites 1, 4, and 5 was likely sufficient for ancestral MTx formation. Sites 2 and 3 are located on a β-hairpin anchored by the N-terminal α-helix (fig. 3A) lost upon site 1 cleavage. Structural modeling suggests the β-hairpin undergoes local unfolding (fig. 3B–E), which would facilitate cleavage

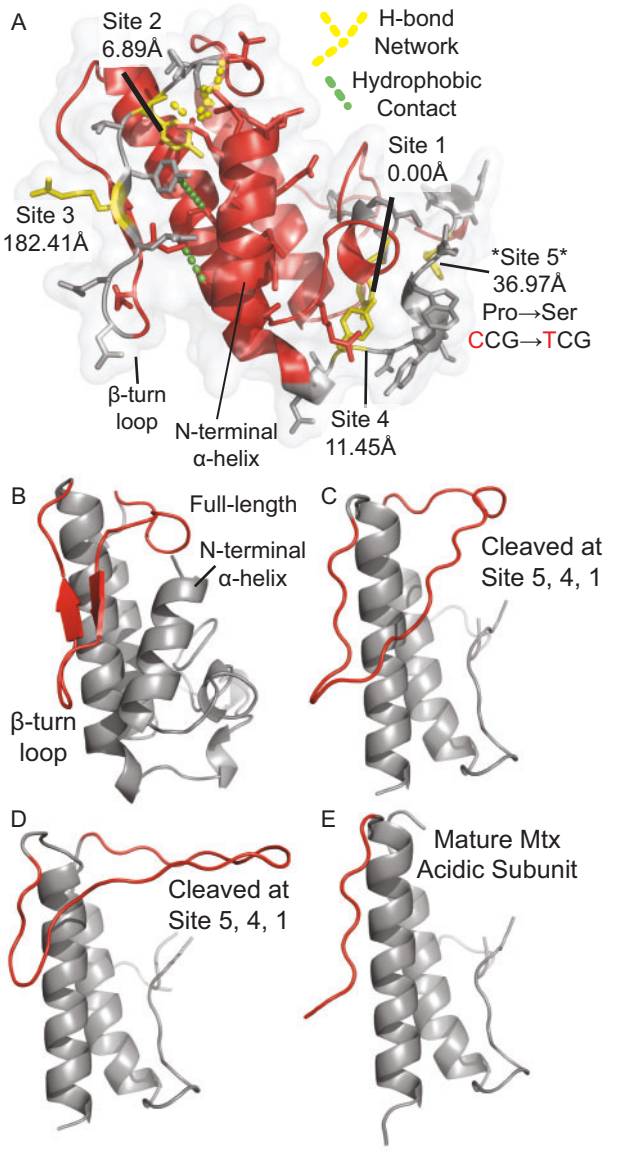

Fig. 3. A single nucelotide change in AncA2 relieved biophysical constraints to proteolytic processing. (*A*) Sites 1 and 4 are physically buried, as determined by solvent accessible surface area (see supplementary table S2, Supplementary Material online) with cleavage blocked by the proline at site 5. A substitution to serine opens this site to proteases. Sites 2 and 3 are located on a flexible, β-turn loop anchored in position by a network of noncovalent interactions with the N-terminal α-helix. Cleavage at sites 5, 4, then 1 removes the N-terminal α-helix leading to local unfolding of the loop. (*B*) Full-length, pro-AncA2 modeled as a monomer on the basic subunit of crotoxin with the loop (red) anchored by the N-terminal α-helix. (*C*) AncA2 with sites 2 and 3 intact modeled as a monomer. Here the structure in gray is modeled on the crotoxin A structure while the structure in red is simultaneously modeled on the corresponding structure in crotoxin B. (*D*) AncA2 with sites 2 and 3 intact modeled on the crotoxin A structure in complex with crotoxin B. Here, we did not provide a template for the red structure allowing molecular dynamics during the modeling procedure to find the structure of the loop. (*E*) Mature AncA2 modeled as a monomer on crotoxin A. Note: because of the highly locally unfolded nature of the loop in structures B and C, these structures will not form complexes with the basic subunit as the binding site is physically blocked by the loop structure.

by SVSPs (Hubbard 1998; Fontana et al. 2004; Park and Marqusee 2004; Chang and Park 2009) regardless of the sequence at site 3. MTx sequences from other, nonrattlesnake pitvipers support this evolutionary mechanism. MTx in the Asian pitviper, *Gloydius intermedius* (fig. 1A) has an acidic subunit with a site 5 sequence that is identical to rattlesnake MTx, but differs at site 3 (T and D instead of R and Q; Yang, Guo, et al. 2015). Despite this difference, its cleavage pattern is identical to rattlesnake MTx (Yang, Guo, et al. 2015). Site 3 and 5 are identical in all other MTx homologs (fig. 1A), suggesting the substitution at site 5 occurred before the substitution at site 3. The mutation at site 5 was likely sufficient to produce a primitive version of the MTx acidic subunit, and processing was later optimized by a site 3 mutation.

Acidic subunit processing was likely completed by venom glutaminyl cyclase (GC), which was already present in rattlesnake venoms where it is likely involved in structural alterations of toxin proteins (fig. 2B; Calvete et al. 2009; Wang et al. 2014). As mentioned above, MTx subunits are not processed by canonical Golgi mechanisms. The venom gland lumen has low pH which inhibits SVSP and GC function (Mackessy and Baxter 2006; Wang et al. 2014). Processing likely occurs in secretory vesicles during trafficking from the Golgi complex to the venom gland. This mechanism is used for SVMP processing (Moura-da Silva et al. 2016). Machinery adequate for acidic subunit processing (SVSPs and GC) was already present in snake venoms prior to the genetic change that allowed cleavage of the acidic proprotein (fig. 2B).

In summary, all of the components that would have been needed for proteolytic processing of the acidic subunit, which is the critical feature of the Type II, neurotoxic venom, were in place prior to the gene duplication events that produced the MTx subunits. The SVSPs had the necessary sequence specificity to cleave the acidic subunit. The five cleavage site sequences are conserved among all venom PLA2s. However, the biophysical arrangement of cleavage sites, and a cleavage blocking proline at site 5, prohibited proteolytic cleavage of ancestral PLA2s and extant PLA2s other than the MTx acidic subunit. A single nucleotide mutation at site 5 in the acidic MTx subunit lineage caused a proline to serine substitution and unlocked a proteolysis cascade to generate the mature acidic subunit. These results suggest that long-standing genetic features along an evolutionary trajectory may make a single, large-effect mutation more likely.

Latent Hydrophobic Interactions Predispose PLA2 to Dimerize

To determine the historical structural changes that occurred in the origin of the acidic and basic MTx subunit interaction, we utilized structural models of extinct PLA2s. We modeled the structures of ancestral PLA2s, AncA1 and AncB1 (fig. 4), corresponding to the acidic and basic ancestors prior to the gene duplication event that led to the subunits of MTx. We used the crystal structure of the MTx homolog crotoxin, from *Crotalus durissus* (fig. 1A) as the template structure, which consists of a proteolytically processed acidic subunit (template for AncA1) and a full-length basic subunit (template for AncB1). We computationally estimated interaction energies

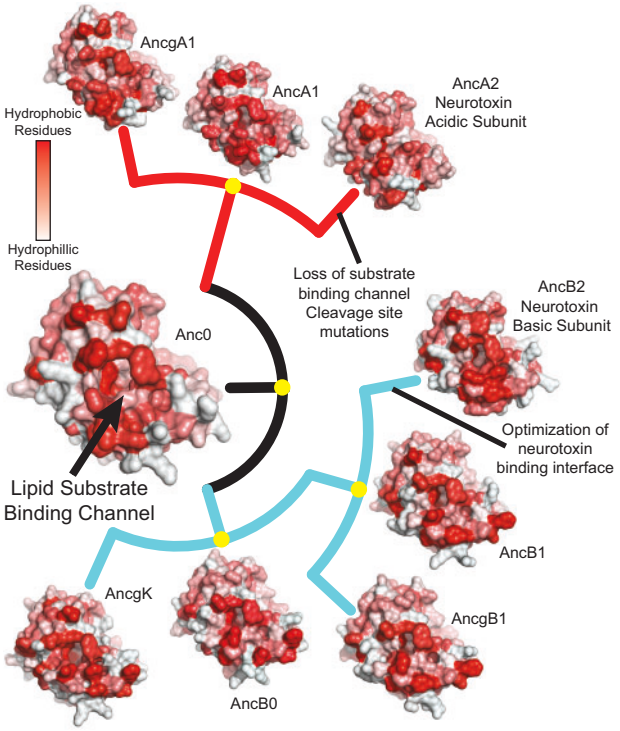

Fig. 4. The structural evolution underlying rattlesnake PLA2 toxin functional diversity. The ancestor of snake type II PLA2s (Anc0) was a catalytically active lipase with a hydrophobic surface that mediated lipase activity. Changes to this hydrophobic surface correlate with functional diversification in PLA2s, and it has been exploited to form the interaction surface of the basic subunit of MTx, and, in the acidic subunit, appears to have lost hydrophobicity correlating with loss of lipase activity and proteolytic processing. Hydrophobicity of the surface residues is shown by the color scale bar. Yellow circles denote gene duplication events. Black lines represent the ancestral PLA2, Anc0, prior to gene duplication events. Red denotes the acidic PLA2 lineage. Cyan denotes the basic lineage.

and found that these ancestral, non-MTx, PLA2s showed a favorable interaction energy (ΔG_{int}), which is similar to the AncA2–AncB2 complex, but the AncA1–AncB1 complex lacked thermodynamic stability (ΔG_{fold}; table 1). Here, the measure of thermodynamic stability represents the free energy of folding of the complex, and the interaction energy is a measure of binding strength, analogous to an experimentally derived association constant. A processed, ancestral acidic PLA2, AncA1, that did not natively undergo proteolytic processing, already had the requisite structural components to interact with an ancestral basic PLA2.

Extant MTx homologs and non-MTx PLA2 toxins appear to be quite promiscuous in their binding partners, adding further support to the observation that the interaction between MTx subunits may be exaptive in origin. The basic subunit of the MTx homolog crotoxin can form tetramers independent of the acidic subunit (Marchi-Salvador et al. 2008). This interaction is dominated by hydrophobic contacts at the interaction surface (Marchi-Salvador et al. 2008). Binding of basic subunits in the tetramer is disrupted by addition of the acidic subunit (Fernandes et al. 2017). Additionally, an in vitro interaction of the basic isoform

Table 1. Stabilities (ΔG_{fold}) and Interaction Energies (ΔG_{int}) between Complexes of Ancestral PLA2s.

Complex	ΔG_{fold} (kcal/mol)	ΔG_{int} (kcal/mol)
Uncleaved AncA2+AncB2	**1030.93**	**878.83**
Cleaved AncA2+AncB2	**−14.32**	**−20.20**
Cleaved AncA1+AncB1	**41.95**	**−3.51**

NOTE.—Stability and interaction energies were determined computationally with FoldX as described in Materials and Methods. A more negative energetic value indicates a more stable complex or more favorable interaction.

with calmodulin, which binds a wide range of protein targets via its hydrophobic domain, is disrupted by addition of the acidic subunit (Šribar et al. 2001). The acidic subunit of crotoxin can interact with monomeric, basic PLA2 neurotoxins from Asian viperids and pitvipers that do not natively form dimers (Choumet et al. 1993). Faure et al. (1994) demonstrated that the association of the subunits and stability of the interaction is directly correlated with crotoxin neurotoxicity. In support of this, we found that all of the extant MTx-positive rattlesnakes in this study show a similar, approximately 2:1 ratio of transcript abundance (supplementary table S3, Supplementary Material online) of acidic to basic MTx subunits. This ratio may promote association of the complex without the need to increase the binding affinity between MTx subunits past an optimum, which could result in slower dissociation of the subunits at the target receptor sites. Utilization of the majority of the basic subunit may help ensure maximum toxicity. The interaction of MTx subunits appears to originate from an exaptation of hydrophobic interaction interfaces involved in ancestral PLA2 lipase function that was subsequently optimized thought mutation of the hydrophobic interaction surface (fig. 4), and optimization of gene expression to promote the strong interaction in extant MTx.

MTx Evolved through Episodic, Diversifying Selection

We sought to determine the role of positive, negative, and episodic selection in the evolution of the structural and functional features of MTx and other PLA2 venom toxins. To understand a structural link between snake venom PLA2 genetic and functional diversity, we tested for and found evidence of positive selection across the snake venom PLA2 tree using codeml (supplementary table S4, Supplementary Material online; Yang 2007). We mapped site-specific d*N*/d*S* ratios onto the ancestral PLA2 (Anc0) structure which revealed a region of positively selected residues on the surface of the molecule (fig. 1*C*; supplementary table S5, Supplementary Material online). In extant lipolytic PLA2s, this surface mediates lipase activity. Also, this region evolved into the interface of both MTx subunits (fig. 1*D*). We performed several other tests for selection to understand the adaptive correlates of MTx structural features. We used the fixed effects likelihood (FEL) and the single likelihood ancestor counting (SLAC) methods (Kosakovsky Pond and Frost 2005b) to further test for sites under positive and negative selection. SLAC and FEL methods (Kosakovsky Pond and Frost 2005b) showed good agreement with codeml, finding eight and 14 sites in common, respectively (supplementary table S5, Supplementary Material online). None of the sites under negative selection according to SLAC or FEL were found to be under positive selection by codeml (supplementary table S6, Supplementary Material online). Twelve interfacial residues showed evidence of positive selection across the PLA2 tree (fig. 1C; supplementary table S5, Supplementary Material online). This positive selection signal may be a result of the interface being a modular region altered to yield various PLA2 functions (fig. 1*B*), and our selection analysis may reflect functional diversification of PLA2s. In support of this, the opposite face of the molecule showed strong purifying selection, suggesting this region may act as a molecular scaffold upon which the modular interface is supported (fig. 1C; supplementary table S6, Supplementary Material online). The modular interface is formed by a cluster of hydrophobic residues that, when mapped onto the unrooted PLA2 phylogeny (fig. 4), showed changes correlating with PLA2 evolution.

We also tested whether some sites experienced more complex selection patterns using the mixed-effects maximum likelihood model (MEME; Murrell et al. 2012) that tests for episodic, diversifying evolution. We found that 31 sites showed evidence of diversifying selection during PLA2 evolution, including a number of interface residues (supplementary table S8, Supplementary Material online). Under episodic evolution, specific sites in the protein could have undergone variation of selection over time. For example, after a gene duplication event, sites in a protein or certain protein lineages may experience a burst of positive selection during neofunctionalization followed by purifying selection once the new function is established (Studer and Robinson-Rechavi 2009; Sironi et al. 2015). These results support our structural analysis (fig. 4) that showed that changes to the hydrophobic interface during PLA2 evolution correlate with PLA2 functional diversity.

To understand the dynamics of evolution along the long branches leading to the acidic and basic MTx subunits, we performed a series of branch-site model tests in codeml to test for different scenarios of positive selection along select branches of a tree (Yang 2007). Although selection was slightly elevated, it was not significantly different than the rest of the tree under several test variations (supplementary table S7, Supplementary Material online). This may be a result of the evolutionary mechanism that yielded MTx. We demonstrated above how a single mutation can reveal a series of exaptations to generate an ancestral MTx. This primitive MTx could have been optimized by positive selection, then, once strong MTx neurotoxicity was established, it could have been maintained by strong purifying selection. This scenario would be consistent with variation in selection that would not be evident in models that only test for positive selection.

To test for episodic, diversifying selection along the long branches leading to the MTx subunits we utilized branch-specific diversifying selection tests (Branch-site REL; Kosakovsky Pond et al. 2011). We found seven branches of our PLA2 phylogeny under episodic evolution (supplementary fig. S2, Supplementary Material online). The acidic subunit branch was not one of them, which may be due to the

unique structure of the mature acidic MTx subunit. For example, altered selective pressures across the molecule as a result of the unique proteolytic processing that yields a mature protein that is only 64% of the full length precursor suggests that 36% of the protein may be experiencing relaxed selective pressures. Also, a lack of episodic evolution may indicate a large-effect mutation in the acidic subunit that would not require accelerated evolution (Rubinoff et al. 2008), which fits with our model of evolution through exaptation described above. In contrast, the basic subunit branch did show evidence of episodic selection (supplementary fig. S2, Supplementary Material online) including three interface residues (positions 39, 77, and 121) along this branch. After the single mutation in the acidic subunit unlocked the cleavage cascade and exposed the interaction interface, the basic subunit likely underwent a burst of positive selection that optimized the interaction interface. This was likely followed by purifying selection to maintain the interaction and, thus, neurotoxicity. Our selection analyses suggest that the subunits of MTx, and, more generally, the structural and functional diversity among venom PLA2 toxins, reflect complex selection pressures.

Conclusions

Snake venoms evolve with remarkable efficiency and display variation at all levels of biological organization (Mackessy 1988, 2009, 2008; Fry 2005; Castoe and Parkinson 2006; Sanz et al. 2006; Fry et al. 2007; Gibbs and Rossiter 2008; Gibbs and Mackessy 2009; Calvete et al. 2010; Rokyta et al. 2011, 2012, 2013, 2017; Castoe et al. 2012; Calvete 2013; Margres et al. 2013, 2014, 2015, 2017; Martínez-Romero et al. 2013; McGivern et al. 2014; Rokyta, Margres, et al. 2015; Wray et al. 2015; Dowell et al. 2016; Margres, Walls, et al. 2016; Margres, Wray, et al. 2016; Rokyta, Wray et al. 2015; Calvete 2017). One source of toxin diversity is exaptation, as illustrated by MTx. Existing snake venom machinery, the SVSPs, provided the proteolytic processing of the acidic subunit of MTx. Proteolytic cleavage sites were present before the evolution of the MTx acidic subunit, but were blocked by biophysical constraints. These constraints were relieved by a single mutation. A latent interaction that was present between ancestral PLA2s prior to the evolution of the MTx subunits was exploited and optimized during MTx evolution.

Evolutionary innovations and complex phenotypes appear more likely when there is prior, standing genetic variation (Bridgham et al. 2006, 2009; Dean and Thornton 2007; Ortlund et al. 2007; Blount et al. 2008, 2012; Breen et al. 2012; Harms and Thornton 2013, 2014; Boucher et al. 2014; Shah et al. 2015; Tufts et al. 2015; Starr and Thornton 2016; Steindel et al. 2016; Hochberg and Thornton 2017; Kumar et al. 2017), underscoring the role of historical contingency in evolution (Blount et al. 2008; Harms and Thornton 2014; Hochberg and Thornton 2017). This phenomenon is exemplified by the evolutionary emergence of MTx and Type II rattlesnake venom. Our results suggest that an evolutionary innovation, and subsequent complex phenotype, can evolve through a large-effect mutation that depends on historical evolutionary events. In elucidating the origin of MTx, we have demonstrated the evolutionary and biophysical emergence of an innovation and complex phenotype via the smallest increment of genetic change—a single nucleotide substitution.

Materials and Methods

Transcriptomics

Initial transcriptome assembly and identification of full-length PLA2 and SVSP transcripts was performed for each sample as described by Rokyta, Margres, et al. (2015). To ensure the most complete representation of these toxins possible, we performed an additional set of assemblies using new assemblers or new versions of previously used assemblers. Overlapping reads were merged with PEAR version 0.9.10 (Zhang et al. 2014). The resulting merged and unmerged reads were assembled with SeqMan NGen version 14.0 using the default de novo transcriptome assembly settings, BinPacker version 1.0 (Liu et al. 2016) with $k = 25$, SOAPdenovo-Trans version 1.03 (Xie et al. 2014) with $k = 75$, and Trinity version 2.4.0 (Grabherr et al. 2011). To identify PLA2s and SVSPs, we downloaded the viperid representatives of these two families from the UniProt animal toxins database and searched our assembled contigs against these sequences using BlastX version 2.6.0, retaining only contigs with full-length coding sequences. We screened for chimeric transcripts by aligning the merged reads against the identified full-length coding sequences using bowtie2 version 2.3.1 (Langmead et al. 2009) and checking for multimodal coverage distributions. The coding sequences of the resulting transcripts were aligned with ClustalW, and sequences within 1% nucleotide divergence were clustered together and treated as a single gene.

Phylogenetics

Pitviper PLA2 coding sequences were isolated from transcripts, translated to amino acid sequence, and aligned with ClustalW (Thompson et al. 1994). The corresponding nucleotide alignment was partitioned by codon position, and nucleotide substitution models were fitted with Partitionfinder v2.1.1 (Lanfear et al. 2017) with AICc as the selection criteria. The appropriate substitution models were implemented for PLA2 gene tree reconstruction by both maximum likelihood and Bayesian inference with RAxML v.8.2.8 (Stamatakis 2014), and Mr. Bayes v.3.2.6 (Ronquist et al. 2012), respectively. Nodal support for RAxML tree was assessed via 1000 bootstrap replicates implemented with the rapid bootstrap option (Stamatakis et al. 2008). The Mr. Bayes reconstruction consisted of two independent runs with one cold and three heated chains run for 1 million generations sampling every 200 generations with the first 100,000 generations discarded as burn-in. Chain stationarity was verified in Tracer 1.6.0 (Rambaut and Drummond 2013) and ESS values for all parameters were greater than 500. Tests of positive selection, determination of site-specific dN/dS ratios, and reconstruction of ancestral PLA2 sequences on the unrooted PLA2 tree were performed with codeml in

PAML v4.8 (Yang 2007). This same unrooted PLA2 tree was used for tests for positive, negative, and episodic selection with tools on the Datamonkey server (Kosakovsky Pond and Frost 2005a) using the general reversible model (REV) of nucleotide substitution, and the default cutoff of 0.1 for significance.

Structural Modeling

Isoelectric points of rattlesnake PLA2 sequences were determined using ProtParam (web.expasy.org/protparam/; last accessed June 27, 2017; Gasteiger et al. 2005). Site-specific solvent accessible surface areas were determined using GetArea (http://curie.utmb.edu/getarea.html; last accessed August 12, 2017; Fraczkiewicz and Braun 1998). We used the X-ray crystal structure of the heterodimeric PLA2 neurotoxin from *C. durrisus terrificus*, known as crotoxin (pdb: 3R0L; Faure et al. 1991) as the template structure for modeling of PLA2s from rattlesnake venoms. Homology models of extant and extinct PLA2 sequences were built using Modeller v 9.18 (Sali and Blundell 1993). For each model, 10 independent structures were built using the automodel method with slow VFTM optimization for 500 iterations, and slow MD refinement. Refinement was repeated four times. Models were evaluated using the DOPE-HR assessment. The model with the lowest DOPE-HR score of the ten models was chosen as the structure used for subsequent analysis. Basic PLA2 and uncleaved acidic sequences were modeled using the B subunit of crotoxin. Mature acidic structures were modeled using the A subunit of crotoxin. In FoldX (Schymkowitz et al. 2005), PDB structures of the models were repaired using the Repair function followed by measurement of folding stability and interaction energies with the Stability and AnalyseComplex functions, respectively. The following post-translational processing mechanisms of PLA2s were assessed bioinformatically: C- (NetCGlyc 1.0 Server—http://www.cbs.dtu.dk/services/NetCGlyc/; last accessed May 17, 2017; Julenius 2007), N- (NetNGlyc 1.0 Server—http://www.cbs.dtu.dk/services/NetNGlyc/; last accessed May 17, 2017; Gupta et al. 2004), and O-linked (NetOGlyc 4.0 Server—http://www.cbs.dtu.dk/services/NetOGlyc/; last accessed May 17, 2017; Steentoft et al. 2013) glycosylation; lysine glycation (NetGlycate 1.0 Server—http://www.cbs.dtu.dk/services/NetGlycate/; last accessed May 17, 2017; Johansen et al. 2006); phosphorylation (NetPhos 3.1 Server—http://www.cbs.dtu.dk/services/NetPhos/; last accessed May 17, 2017; Blom et al. 1999); proprotein convertase cleavage (ProP 1.0 Server—http://www.cbs.dtu.dk/services/ProP/; last accessed May 17, 2017; Duckert et al. 2004).

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

This work was supported by the National Science Foundation (DEB 1145987 to D.R.R). Transcriptomes are available as raw sequence data (SRA) under BioProject PRJNA88989: *Agkistrodon contortrix*—SRR2032114, *Agkistrodon piscivoros*—SRR2032118, *Crotalus adamanteus*—SRR441163, *Crotalus atrox*—SRR5270430, *Crotalus cerastes*—SRR5270834, *Crotalus horridus* A—SRR575168, *Crotalus horridus* B—SRR1554232, *Crotalus lepidus*—SRR5270851, *Crotalus mitchellii*—SRR5270850, *Crotalus molossus*—SRR5270852, *Crotalus scutulatus*—SRR5270449, *Crotalus tigris*—SRR5270853, *Sistrurus catenatus*—SRR2029826, *Sistrurus miliarius*—SRR2031930. The accession numbers for the sequences used in this work are provided in Supplementary Material online.

References

Berg OG, Gelb MH, Tsai MD, Jain MK. 2001. Interfacial enzymology: The secreted phospholipase A2-paradigm. *Chem Rev*. 101(9): 2613–2653.

Blom N, Gammeltoft S, Brunak S. 1999. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *J Mol Biol*. 294(5): 1351–1362.

Blount ZD, Borland CZ, Lenski RE. 2008. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*. *Proc Natl Acad Sci*. 105(23): 7899–7906.

Blount ZD, Barrick JE, Davidson CJ, Lenski RE. 2012. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature* 488: 513–518.

Boucher JI, Jacobowitz JR, Beckett BC, Classen S, Theobald DL. 2014. An atomic-resolution view of neofunctionalization in the evolution of apicomplexan lactate dehydrogenases. *Elife* 3: e02304.

Breen MS, Kemena C, Vlasov PK, Notredame C, Kondrashov FA. 2012. Epistasis as the primary factor in molecular evolution. *Nature* 490(7421): 535.

Bridgham JT, Carroll SM, Thornton JW. 2006. Evolution of hormone-receptor complexity by molecular exploitation. *Science* 312(5770): 97–101.

Bridgham JT, Ortlund EA, Thornton JW. 2009. An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. *Nature* 461(7263): 515–519.

Calvete JJ. 2013. Snake venomics: From the inventory of toxins to biology. *Toxicon* 75:44–62.

Calvete JJ. 2017. Venomics: integrative venom proteomics and beyond. *Biochem J*. 474(5): 611–634.

Calvete JJ, Fasoli E, Sanz L, Boschetti E, Righetti PG. 2009. Exploring the venom proteome of the western diamondback rattlesnake, *Crotalus atrox*, via snake venomics and combinatorial peptide ligand library approaches. *J Proteome Res*. 8(6): 3055–3067.

Calvete JJ, Sanz L, Cid P, de la Torre P, Flores-Díaz M, Dos Santos MC, Borges A, Bremo A, Angulo Y, Lomonte B, et al. 2010. Snake venomics of the Central American Rattlesnake *Crotalus simus* and the South American *Crotalus durissus* complex points to neurotoxicity as an adaptive paedomorphic trend along *Crotalus* dispersal in South America. *J Proteome Res*. 9(1): 528–544.

Calvete JJ, Pérez A, Lomonte B, Sánchez EE, Sanz L. 2012. Snake venomics of *Crotalus tigris*: The minimalist toxin arsenal of the deadliest neartic rattlesnake venom. Evolutionary clues for generating a pan-specific antivenom against crotalid type II venoms. *J Proteome Res*. 11(2): 1382–1390.

Castoe TA, Parkinson CL. 2006. Bayesian mixed models and the phylogeny of pitvipers (Viperidae: Serpentes). *Mol Phylogenet Evol*. 39(1): 91–110.

Castoe TA, Braun EL, Bronikowski AM, Cox CL, Rabosky ARD, Jason de Koning AP, Dobry J, Fujita MK, Giorgianni MW, Hargreaves A, et al. 2012. Report from the first snake genomics and integrative biology meeting. *Stand Genomic Sci*. 7(1): 150–152.

Chang Y, Park C. 2009. Mapping transient partial unfolding by protein engineering and native-state proteolysis. *J Mol Biol*. 393(2): 543–556.

Choumet V, Saliou B, Fideler L, Chen Y, Gubensek F, Bon C, Delot E. 1993. Snakevenom phospholipase A2 neurotoxins: Potentiation of a

singlechain neurotoxin by the chaperon subunit of a twocomponent neurotoxin. *Eur J Biochem*. 211(1–2): 57–62.
Clark RF, Williams SR, Nordt SP, Boyer-Hassen LV. 1997. Successful treatment of crotalid-induced neurotoxicity with a new polyspecific crotalid Fab antivenom. *Ann Emerg Med*. 30(1): 54–57.
Darwin C. 1859. On the origin of species, 1859. London: John Murray.
Dean AM, Thornton JW. 2007. Mechanistic approaches to the study of evolution: the functional synthesis. *Nat Rev Genet*. 8(9): 675–688.
Doley R, Zhou X, Kini R. 2009. Snake venom phospholipase A2 enzymes. In: Mackessy SP, editor. Handbook of venoms and toxins of reptiles. Boca Raton, FL: CRC Press Fr. Group. p. 173–205.
Dowell NL, Giorgianni MW, Kassner VA, Selegue JE, Sanchez EE, Carroll SB. 2016. The deep origin and recent loss of venom toxin genes in rattlesnakes. *Curr Biol*. 26(18): 2434–2445.
Duckert P, Brunak S, Blom N. 2004. Prediction of proprotein convertase cleavage sites. *Protein Eng Des Sel*. 17(1): 107–112.
Faure G, Guillaume JL, Camoin L, Saliou B, Bon C. 1991. Multiplicity of acidic subunit isoforms of crotoxin, the phospholipase A2 neurotoxin from *Crotalus durissus terrificus* venom, results from posttranslational modifications. *Biochemistry* 30(32): 8074–8083.
Faure G, Harvey AL, Thomson E, Saliou B, Radvanyi F, Bon C. 1993. Comparison of crotoxin isoforms reveals that stability of the complex plays a major role in its pharmacological action. *Eur J Biochem*. 214(2): 491–496.
Faure G, Choumet V, Bouchier C, Camoin L, Guillaume JL, Monegier B, Vuilhorgne M, Bon C. 1994. The origin of the diversity of crotoxin isoforms in the venom of *Crotalus durissus terrificus*. *Eur J Biochem*. 223(1): 161–164.
Faure G, Xu H, Saul FA. 2011. Crystal structure of crotoxin reveals key residues involved in the stability and toxicity of this potent heterodimeric β-neurotoxin. *J Mol Biol*. 412(2): 176–191.
Fernandes CAH, Pazin WM, Dreyer TR, Bicev RN, Cavalcante WLG, Fortes-Dias CL, Ito AS, Oliveira CLP, Fernandez RM, Fontes MRM. 2017. Biophysical studies suggest a new structural arrangement of crotoxin and provide insights into its toxic mechanism. *Sci Rep*. 7: 43885.
Fisher RA. 1930. The genetical theory of natural selection. Oxford, UK: Oxford University Press.
Fontana A, De Laureto PP, Spolaore B, Frare E, Picotti P, Zambonin M. 2004. Probing protein structure by limited proteolysis. *Acta Biochim Pol*. 51(2): 299–321.
Fraczkiewicz R, Braun W. 1998. Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J Comput Chem*. 19(3): 319–333.
Fry BG. 2005. From genome to "venome": molecular origin and evolution of the snake venom proteome inferred from phylogenetic analysis of toxin sequences and related body proteins. *Genome Res*. 15(3): 403–420.
Fry BG, Scheib H, van der Weerd L, Young B, McNaughtan J, Ramjan SFR, Vidal N, Poelmann RE, Norman JA. 2007. Evolution of an arsenal: structural and functional diversification of the venom system in the advanced snakes (Caenophidia). *Mol Cell Proteomics*. 7(2): 215–246.
Gamberale-Stille G, Balogh ACV, Tullberg BS, Leimar O. 2012. Feature saltation and the evolution of mimicry. *Evolution (N Y)* 66(3): 807–817.
Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. Proteomics Protocols Handbook. p. 571–607. Totowa (NJ): Humana Press.
Gibbs HL, Mackessy SP. 2009. Functional basis of a molecular adaptation: prey-specific toxic effects of venom from Sistrurus rattlesnakes. *Toxicon* 53(6): 672–679.
Gibbs HL, Rossiter W. 2008. Rapid evolution by positive selection and gene gain and loss: PLA 2 venom genes in closely related *Sistrurus* rattlesnakes with divergent diets. *J Mol Evol*. 66(2): 151–166.
Glenn JL, Straight RC, Wolt TB. 1994. Regional variation in the presence of canebrake toxin in *Crotalus horridus* venom. *Comp Biochem Physiol C Pharmacol*. 107(3): 337–346.
Grabherr MG, Haas BJ, Yassour M, Levin J, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 29(7): 644–652.
Gupta R, Jung E, Brunak S. 2004. NetNGlyc: prediction of N-glycosylation sites in human proteins. Available from: http://www.cbs.dtu.dk/services/NetNGlyc/
Harms MJ, Thornton JW. 2013. Evolutionary biochemistry: revealing the historical and physical causes of protein properties. *Nat Rev Genet*. 14(8): 559–571.
Harms MJ, Thornton JW. 2014. Historical contingency and its biophysical basis in glucocorticoid receptor evolution. *Nature* 512(7513): 203–207.
Hedstrom L. 2002. Serine protease mechanism and specificity. *Chem Rev*. 102(12): 4501–4523.
Hochberg GKA, Thornton JW. 2017. Reconstructing ancient proteins to understand the causes of structure and function. *Annu Rev Biophys*. 46(1): 247–269.
Hubbard S. 1998. The structural aspects of limited proteolysis of native proteins. *Biochim Biophys Acta Protein Struct Mol Enzymol*. 1382(2): 191–206.
Jansen PW, Perkin RM, Stralen DV. 1992. Mojave rattlesnake envenomation: prolonged neurotoxicity and rhabdomyolysis. *Ann Emerg Med*. 21(3): 322–325.
Jiménez-Charris E, Montealegre-Sánchez L, Solano-Redondo L, Castro-Herrera F, Fierro-Pérez L, Lomonte B. 2016. Divergent functional profiles of acidic and basic phospholipases A2 in the venom of the snake *Porthidium lansbergii* lansbergii. *Toxicon* 119:289–298.
Johansen MB, Kiemer L, Brunak S. 2006. Analysis and prediction of mammalian protein glycation. *Glycobiology* 16(9): 844–853.
Julenius K. 2007. NetCGlyc 1. 0: prediction of mammalian C-mannosylation sites. *Glycobiology* 17(8): 868–876.
Kay J. 1993. Specificity of proteolysis. *FEBS Lett*. 331(1–2): 201.
Kosakovsky Pond SL, Frost SDW. 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* 21(10): 2531–2533.
Kosakovsky Pond SL, Frost SDW. 2005. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol*. 22(5): 1208–1222.
Kosakovsky Pond SL, Murrell B, Fourment M, Frost SDW, Delport W, Scheffler K. 2011. A random effects branch-site model for detecting episodic diversifying selection. *Mol Biol Evol*. 28(11): 3033–3043.
Križaj I, Siigur J, Samel M, Cotič V, Gubenšek F. 1993. Isolation, partial characterization and complete amino acid sequence of the toxic phospholipase A2 from the venom of the common viper, *Vipera berus berus*. *Biochim Biophys Acta*. 1157(1): 81–85.
Kumar A, Natarajan C, Moriyama H, Witt CC, Weber RE, Fago A, Storz JF. 2017. Stability-mediated epistasis restricts accessible mutational pathways in the functional evolution of avian hemoglobin. *Mol Biol Evol*. 34(5): 1240–1251.
Lamb TD, Collin SP, Pugh EN. 2007. Evolution of the vertebrate eye: opsins, photoreceptors, retina and eye cup. *Nat Rev Neurosci*. 8(12): 960–976.
Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2017. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol Biol Evol*. 34(3): 772–773.
Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 10(3): R25.
Leimar O, Tullberg BS, Mallet J. 2012. Mimicry, saltational evolution, and the crossing of fitness valleys. In: Svensson EI and Calsbeek R, editors. The Adaptive Landscape in Evolutionary Biology. Oxford (UK): Oxford University Press. p. 259–270.
Liu J, Li G, Chang Z, Yu T, Liu B, McMullen R, Chen P, Huang X, Lengauer T. 2016. BinPacker: packing-based de novo transcriptome assembly from RNA-seq data. *PLoS Comp Biol*. 12(2): 1–15.
Lomonte B, Rangel J. 2012. Snake venom Lys49 myotoxins: from phospholipases A 2 to non-enzymatic membrane disruptors. *Toxicon* 60(4): 520–530.

Lomonte B, Angulo Y, Calderón L. 2003. An overview of lysine-49 phospholipase A2 myotoxins from crotalid snake venoms and their structural determinants of myotoxic action. *Toxicon* 42(8): 885–901.
Lomonte B, Angulo Y, Sasa M, Gutierrez JM. 2009. The phospholipase A2 homologues of snake venoms: biological activities and their possible adaptive roles. *Protein Pept Lett*. 16(8): 860–876.
Mackessy S. 2009. The field of reptile toxinology: snakes, lizards and their venoms. In: Mackessy SP, editor. Handbook of venoms and toxins of reptiles. Boca Raton, FL: CRC Press. Fr. Group. p. 3–23.
Mackessy SP. 1988. Venom ontogeny in the Pacific Rattlesnakes *Crotalus viridis helleri* and *C. v. oreganus*. *Copeia* 1988(1): 92–101.
Mackessy SP. 2008. Venom composition in rattlesnakes: trends and biological significance. In: Hayes WK, editor. Biology of rattlesnakes. Loma Linda (CA): Loma Linda University Press. p. 495–510.
Mackessy SP, Baxter LM. 2006. Bioweapons synthesis and storage: the venom gland of front-fanged snakes. *Zool Anz*. 245(3–4): 147–159.
Marchi-Salvador DP, Corrêa LC, Magro AJ, Oliveira CZ, Soares AM, Fontes MRM. 2008. Insights into the role of oligomeric state on the biological activities of crotoxin: crystal structure of a tetrameric phospholipase A2 formed by two isoforms of crotoxin B from *Crotalus durissus* terrificus venom. *Proteins Struct Funct Genet*. 72(3): 883–891.
Margres MJ, Aronow K, Loyacano J, Rokyta DR. 2013. The venom-gland transcriptome of the eastern coral snake (*Micrurus fulvius*) reveals high venom complexity in the intragenomic evolution of venoms. *BMC Genomics* 14(1): 531.
Margres MJ, McGivern JJ, Wray KP, Seavy M, Calvin K, Rokyta DR. 2014. Linking the transcriptome and proteome to characterize the venom of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *J Proteomics*. 96:145–158.
Margres MJ, Wray KP, Seavy M, McGivern JJ, Sanader D, Rokyta DR. 2015. Phenotypic integration in the feeding system of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *Mol Ecol*. 24(13): 3405–3420.
Margres MJ, Wray KP, Seavy M, McGivern JJ, Herrera ND, Rokyta DR. 2016. Expression differentiation is constrained to low-expression proteins over ecological timescales. *Genetics* 202(1): 273–283.
Margres MJ, Walls R, Suntravat M, Lucena S, Sánchez EE, Rokyta DR. 2016. Functional characterizations of venom phenotypes in the eastern diamondback rattlesnake (*Crotalus adamanteus*) and evidence for expression-driven divergence in toxic activities among populations. *Toxicon* 119:28–38.
Margres MJ, Bigelow AT, Lemmon EM, Lemmon AR, Rokyta DR. 2017. Selection to increase expression, not sequence diversity, precedes gene family origin and expansion in rattlesnake venom. *Genetics* 206(3): 1569–1580.
Martínez-Romero G, Rucavado A, Lazcano D, Gutiérrez JM, Borja M, Lomonte B, Garza-García Y, Zugasti-Cruz A. 2013. Comparison of venom composition and biological activities of the subspecies *Crotalus lepidus lepidus*, *Crotalus lepidus klauberi* and *Crotalus lepidus morulus* from Mexico. *Toxicon* 71:84–95.
McGivern JJ, Wray KP, Margres MJ, Couch ME, Mackessy SP, Rokyta DR. 2014. RNA-seq and high-definition mass spectrometry reveal the complex and divergent venoms of two rear-fanged colubrid snakes. *BMC Genomics* 15(1): 1061.
Moura-da Silva AM, Almeida MT, Portes-Junior JA, Nicolau CA, Gomes-Neto F, Valente RH. 2016. Processing of snake venom metalloproteinases: generation of toxin diversity and enzyme inactivation. *Toxins (Basel)* 8(6):183.
Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL, Malik HS. 2012. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet* 8(7): e1002764.
Nilsson D-E, Pelger S. 1994. A pessimistic estimate of the time required for an eye to evolve. *Proc R Soc Lond Ser B Biol Sci*. 256(1345): 53–58.
Norrström N, Getz WM, Holmgren NH. 2006. Coevolution of exploiter specialization and victim mimicry can be cyclic and saltational. *Evol Bioinform Online*. 2(0): 35–43.
Orr HA. 2005. The genetic theory of adaptation: a brief history. *Nat Rev Genet*. 6(2): 119–127.
Ortlund EA, Bridgham JT, Redinbo MR, Thornton JW. 2007. Crystal structure of an ancient protein: evolution by conformational epistasis. *Science* 317(5844): 1544–1548.
Pal CPB. 2017. Evolution of complex adaptations in molecular systems. *Nat Ecol Evol*. 1(August): 1084–1092.
Park C, Marqusee S. 2004. Probing the high energy states in proteins by proteolysis. *J Mol Biol*. 343(5): 1467–1476.
Perona JJ, Craik CS. 2008. Structural basis of substrate specificity in the serine proteases. *Protein Sci*. 4(3): 337–360.
Radvanyi FR, Bon C. 1982. Catalytic activity and reactivity with p-bromophenacyl bromide of the phospholipase subunit of crotoxin. Influence of dimerization and association with the noncatalytic subunit. *J Biol Chem*. 257(21): 12616–12623.
Rambaut A, Drummond AJ. 2013. Tracer v1.6. Available from http://tree.bio.ed.ac.uk/software/tracer/.
Rokyta DR, Joyce P, Caudle SB, Miller C, Beisel CJ, Wichman HA, Malik HS. 2011. Epistasis between beneficial mutations and the phenotype-to-fitness map for a ssDNA virus. *PLoS Genet*. 7(6): e1002075.
Rokyta DR, Lemmon AR, Margres MJ, Aronow K. 2012. The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *BMC Genomics* 13(1): 312.
Rokyta DR, Wray KP, Margres MJ. 2013. The genesis of an exceptionally lethal venom in the timber rattlesnake (*Crotalus horridus*) revealed through comparative venom-gland transcriptomics. *BMC Genomics* 14(1): 394.
Rokyta DR, Margres MJ, Calvin K. 2015. Post-transcriptional mechanisms contribute little to phenotypic variation in snake venoms. *G3* 5(11): 2375–2382.
Rokyta DR, Wray KP, McGivern JJ, Margres MJ. 2015. The transcriptomic and proteomic basis for the evolution of a novel venom phenotype within the Timber Rattlesnake (*Crotalus horridus*). *Toxicon* 98:34–48.
Rokyta DR, Margres MJ, Ward MJ, Sanchez EE. 2017. The genetics of venom ontogeny in the eastern diamondback rattlesnake (*Crotalus adamanteus*). *PeerJ* 5:e3249.
Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. Mrbayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 61(3): 539–542.
Rosenberg P. 1986. The relationship between enzymatic activity and pharmacological properties of phospholipases in natural poisons. In: Harris JB, editor. Natural Toxins—Animal, Plant and Microbial. Oxford (UK): Clarendon Press. p. 129–174.
Rubinoff D, Le Roux JJ, Freckleton RP. 2008. Evidence of repeated and independent saltational evolution in a peculiar genus of sphinx moths (*Proserpinus*: Sphingidae). *PLoS One* 3(12): e4035.
Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*. 234(3): 779–815.
Sanz L, Lisle Gibbs H, Mackessy SP, Calvete JJ. 2006. Venom proteomes of closely related Sistrurus rattlesnakes with divergent diets. *J Proteome Res*. 5(9): 2098–2112.
Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. 2005. The FoldX web server: an online force field. *Nucleic Acids Res*. 33(Web Server): 382–388.
Serrano SMT, Maroun RC. 2005. Snake venom serine proteinases: sequence homology vs. substrate specificity, a paradox to be solved. *Toxicon* 45(8): 1115–1132.
Shah P, McCandlish DM, Plotkin JB. 2015. Contingency and entrenchment in protein evolution under purifying selection. *Proc Natl Acad Sci U S A*. 112(25): E3226–E3235.
Sironi M, Cagliani R, Forni D, Clerici M. 2015. Evolutionary insights into hostpathogen interactions from mammalian sequence data. *Nat Rev Genet*. 16(4): 224–236.
Six DA, Dennis EA. 2000. The expanding superfamily of phospholipase A 2 enzymes: classification and characterization. *Biochim Biophys Acta*. 1488(1–2): 1–19.
Šribar J, Čopič A, Pariš A, Sherman NE, Gubenšek F, Fox JW, Križaj I. 2001. A high affinity acceptor for phospholipase A2 with neurotoxic activity is a calmodulin. *J Biol Chem*. 276(16): 12493–12496.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9): 1312–1313.
Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol*. 57(5): 758–771.
Starr TN, Thornton JW. 2016. Epistasis in protein evolution. *Protein Sci*. 25(7): 1204–1218.
Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT-BG, Lavrsen K, Dabelsteen S, Pedersen NB, Marcos-Silva L, et al. 2013. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J*. 32(10): 1478–1488.
Steindel PA, Chen EH, Wirth JD, Theobald DL. 2016. Gradual neofunctionalization in the convergent evolution of trichomonad lactate and malate dehydrogenases. *Protein Sci*. 25(7): 1319–1331.
Storz JF. 2016. Causes of molecular convergence and parallelism in protein evolution. *Nat Rev Genet*. 17(4): 239–250.
Studer R, Robinson-Rechavi M. 2009. Evidence for an episodic model of protein sequence evolution. *Biochem Soc Trans*. 37(4): 783–786.
Tawfik DS. 2010. Messy biology and the origins of evolutionary innovations. *Nat Chem Biol*. 6(10): 692–696.
Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 22(22): 4673–4680.
Tina KG, Bhadra R, Srinivasan N. 2007. PIC: protein interactions calculator. *Nucleic Acids Res*. 35(Web Server): W473.
Tsai IH, Wang YM, Chen YH, Tu AT. 2003. Geographic variations, cloning, and functional analyses of the venom acidic phospholipases A2 of *Crotalus viridis viridis*. *Arch Biochem Biophys*. 411(2): 289–296.
Tufts DM, Natarajan C, Revsbech IG, Projecto-Garcia J, Hoffmann FG, Weber RE, Fago A, Moriyama H, Storz JF. 2015. Epistasis constrains mutational pathways of hemoglobin adaptation in high-altitude Pikas. *Mol Biol Evol*. 32(2): 287–298.
Wagner A. 2011. The molecular origins of evolutionary innovations. *Trends Genet*. 27(10): 397–410.
Wang YM, Huang KF, Tsai IH. 2014. Snake venom glutaminyl cyclases: purification, cloning, kinetic study, recombinant expression, and comparison with the human enzyme. *Toxicon* 86:40–50.
Wooldridge BJ, Pineda G, Banuelas-Ornelas JJ, Dagda RK, Gasanov SE, Rael ED, Lieb CS. 2001. Mojave rattlesnakes (*Crotalus scutulatus scutulatus*) lacking the acidic subunit DNA sequence lack Mojave toxin in their venom. *Comp Biochem Physiol B Biochem Mol Biol*. 130(2): 169–179.
Wray KP, Margres MJ, Seavy M, Rokyta DR. 2015. Early significant ontogenetic changes in snake venoms. *Toxicon* 96:74–81.
Wüster W, Peppin L, Pook CE, Walker DE. 2008. A nesting of vipers: phylogeny and historical biogeography of the Viperidae (Squamata: Serpentes). *Mol Phylogenet Evol*. 49(2): 445–459.
Xie Y, Wu G, Tang J, Luo R, Patterson J, Liu S, Huang W, He G, Gu S, Li S, et al. 2014. SOAPdenovo-Trans: de novo transcriptome assembly with short RNA-Seq reads. *Bioinformatics* 30(12): 1660–1666.
Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 24(8): 1586–1591.
Yang ZM, Guo Q, Ma ZR, Chen Y, Wang ZZ, Wang XM, Wang YM, Tsai IH. 2015. Structures and functions of crotoxin-like heterodimers and acidic phospholipases A2 from *Gloydius intermedius* venom: insights into the origin of neurotoxic-type rattlesnakes. *J Proteomics*. 112:210–223.
Yang ZM, Yang YE, Chen Y, Cao J, Zhang C, Liu LL, Wang ZZ, Wang XM, Wang YM, Tsai IH. 2015. Transcriptome and proteome of the highly neurotoxic venom of *Gloydius intermedius*. *Toxicon* 107(Pt B): 175–186.
Zhang J, Kobert K, Flouri T, Stamatakis A. 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30(5): 614–620.