1 **Epistatic effects between amino acid insertions and substitutions mediate toxin-resistance of**
2 **vertebrate Na[+],K[+]-ATPases**

3 **Authors:** Shabnam Mohammadi[1,2,], Halil İbrahim Özdemir[3], Pemra Ozbek[3], Fidan Sumbul[4], Josefin
4 Stiller[5], Yuan Deng[5,6] Andrew J. Crawford[7], Hannah M. Rowland[2], Jay F. Storz[8], Peter Andolfatto[9] &
5 Susanne Dobler[1]

6 **Affiliations:**

7 [1]Molecular Evolutionary Biology, Institute of Cell and Systems Biology of Animals, Universität
8 Hamburg, Hamburg, Germany

9 [2]Max Planck Institute for Chemical Ecology, Research Group Predators and Toxic Prey, Jena, Germany

10 [3]Department of Bioengineering, Marmara University, Göztepe, İstanbul, Turkey

11 [4]Aix-Marseille Université, Inserm, CNRS, Marseille, France

12 [5]Villum Centre for Biodiversity Genomics, University of Copenhagen, Denmark

13 [6]BGI-Shenzhen, Shenzhen 518083, China

14 [7]Department of Biological Sciences, Universidad de los Andes, Bogotá, Colombia

15 [8]School of Biological Sciences, University of Nebraska, Lincoln, NE, USA

16 [9]Department of Biological Sciences, Columbia University, New York, NY, USA

17

18 Correspondence to: shabnam.mohammadi@uni-hamburg.de
19
20

21 **Abstract:** The recurrent evolution of resistance to cardiotonic steroids (CTS) across diverse animals most
22 frequently involves convergent amino-acid substitutions in the H1-H2 extracellular loop of Na[+],K[+]-
23 ATPase (NKA). Previous work revealed that hystricognath rodents (e.g. chinchilla) and pterocliform
24 birds (sandgrouse) have convergently evolved amino-acid insertions in the H1-H2 loop, but their
25 functional significance was not known. Using protein engineering, we show that these insertions have
26 distinct effects on CTS resistance in homologs of each of the two species that strongly depend on
27 intramolecular interactions with other residues. Removing the insertion in the chinchilla NKA
28 unexpectedly increases CTS resistance and decreases NKA activity. In the sandgrouse NKA, the amino
29 acid insertion and substitution Q111R both contribute to an augmented CTS resistance without
30 compromising ATPase activity levels. Molecular docking simulations provide additional insight into the
31 biophysical mechanisms responsible for the context-specific mutational effects on CTS insensitivity of
32 the enzyme. Our results highlight the diversity of genetic substrates that underlie CTS insensitivity in
33 vertebrate NKA and reveal how amino-acid insertions can alter the phenotypic effects of point mutations
34 at key sites in the same protein domain.

1  **Introduction:**

2  The evolution of toxin resistance in animals provides one of the best studied examples of adaptive

3  molecular adaptation (Brodie 2009). In many cases, diverse animals have convergently evolved resistance

4  to the same toxin, thereby permitting insights into the functional effects of identical mutations on different

5  genetic backgrounds (McGlothlin et al. 2016; Mohammadi et al. 2022). Recent studies have characterized

6  the epistatic and pleiotropic effects of amino acid mutations and have provided insights into how such

7  effects influence the process of adaptive protein evolution (Weinreich et al. 2006; Gong and Bloom 2014;

8  Storz 2018). Substantially less is known about the functional effects of amino acid insertions and

9  deletions and their potential evolutionary significance (de la Chaux et al. 2007; Miton and Tokuriki 2022;

10 Savino et al. 2022). Here, we investigate the impact of convergently evolved insertions in the context of

11 the evolution of cardiotonic steroids (CTS) resistance in animals.

12

13 CTS comprise a diverse group of plant- and animal-derived secondary compounds that are often used as a

14 means of chemical defense against herbivores and predators (Krenn and Kopp 1998; Hutchinson et al.

15 2007; Agrawal et al. 2012). CTS are toxic to animals because they inhibit $Na^+,K^+$-ATPase (NKA), a

16 heterodimeric transmembrane protein that consists of a catalytic α-subunit, encoded by members of the

17 ATP1A gene family, and a glycoprotein β-subunit, encoded by members of the ATP1B gene family (Fig.

18 1A; (Köksoy 2002; Aperia 2007)). NKAs play a critical role in maintaining membrane potential via trans-

19 membrane exchange of $Na^+$ and $K^+$ ions and are consequently vital for the maintenance of diverse

20 physiological functions including neural signal transduction, muscle contraction, and cellular homeostasis

21 (Blanco and Mercer 1998; Mobasheri et al. 2000; Bagrov et al. 2009). CTSs bind to a specific domain of

22 the α-subunit of NKA (ATP1A), and binding affinity is strongly determined by the H1-H2 extracellular

23 loop, which forms part of that domain (Fig. 1A; (Laursen et al. 2015)).

24

25 Convergent evolution of resistance to CTS in diverse animals is often mediated, at least in part, by the

26 evolution of NKA 'target-site insensitivity'. Experimental studies have revealed a surprising consistency

27 in the underlying molecular mechanisms of CTS insensitivity of NKA (Ujvari et al. 2015; Storz 2016;

28 Karageorgi et al. 2019; Taverner et al. 2019; Mohammadi et al. 2021; Mohammadi et al. 2022). Different

29 combinations of amino-acid substitutions have been reported in animals that have evolved CTS-

30 resistance, and available evidence suggests that some substitutions impede CTS binding by preventing the

31 formation of hydrogen bonds between hydroxyl groups of CTS and polar residues of the H1-H2

32 extracellular loop (Laursen et al. 2015). In tetrapods, different isoforms of the NKA α-subunit are

33 encoded by three paralogous members of a multigene family (ATP1A1-ATP1A3), with a fourth paralog

34 specific to mammals (ATP1A4). Resistance-conferring amino acid substitutions have been identified in

1    paralogs A1-A3, but are most common in ATP1A1, which is also the most ubiquitously expressed (Ujvari

2    et al. 2013; Ujvari et al. 2015; Mohammadi et al. 2016; Marshall et al. 2018; Mohammadi et al. 2022).

3

4    In taxa as diverse as milkweed-feeding insects and toad-eating vertebrates, the evolution of CTS

5    resistance is often associated with amino acid substitutions at sites 111 and 122 in the H1-H2 extracellular

6    loop of ATP1A (Price and Lingrel 1988; Dobler et al. 2012; Zhen et al. 2012; Groen and Whiteman 2021;

7    Mohammadi et al. 2021; Mohammadi et al. 2022). Consequently, these two sites have been the focus of

8    experimental efforts to characterize molecular mechanisms of target-site insensitivity in NKA. However,

9    results from several recent studies have revealed nonadditive interactions (i.e., intramolecular epistasis)

10   involving  sites 111, 122, and other sites in ATP1A that impinge on multiple aspects of protein function

11   and affect resistance to CTS (Karageorgi et al. 2019; Taverner et al. 2019; Mohammadi et al. 2021;

12   Mohammadi et al. 2022). Two recent studies reported a two-amino acid insertion in addition to multiple

13   amino acid substitutions in the H1-H2 loop of pyrgomorphid grasshopper species that feed on CTS-

14   defended plants (Dobler et al. 2019; Yang et al. 2019). Together, the insertion and substitutions at sites

15   111 and 122 were shown to result in a strong increase in CTS resistance when added to the wild-type

16   *Drosophila melanogaster* ATP1A ortholog (Dobler et al. 2019).

17

18   We recently reported similar insertions in two vertebrate lineages – mammals (hystricognath rodents) and

19   birds (sandgrouse) (Mohammadi et al. 2022). While our initial site-directed mutagenesis experiments on

20   the ATP1A1 of these two taxa focused only on substitutions at sites 111 and 122 of ATP1A1, they

21   yielded puzzling results (Mohammadi et al. 2022), which motivated us to investigate the functional

22   effects of the insertions in the present study. The wildtype ATP1A1 of chinchilla (*Chinchilla lanigera*, a

23   representative of hystricognath rodents) does not exhibit CTS resistance and the rodents are not known to

24   consume CTS. However, when we engineered the replacement N122D, which introduces an amino-acid

25   state known to confer CTS resistance in ATP1A1 of other tetrapod species (Price et al. 1990;

26   Mohammadi et al. 2021; Mohammadi et al. 2022), it conferred orders of magnitude higher CTS resistance

27   than that observed in other rodents with the same substitution. Conversely, the wildtype ATP1A1 of the

28   sandgrouse, (*Pterocles gutturalis*, Aves: Pterocliformes) exhibits substantial resistance, despite there

29   being no records of this species consuming CTS-defended plants or insects. However, when we reverted

30   the derived "resistant" R111 to the ancestral "sensitive" amino-acid state (Q111), resistance remained

31   undiminished, suggesting that one or more resistance-conferring substitutions occurred at other sites

32   (Mohammadi et al. 2022). We thus hypothesized that the proximate insertions observed in hystricognath

33   rodents and sandgrouse potentially impact resistance to CTS, but in different ways. To address this

34   hypothesis, we combined our previously published data on chinchilla and sandgrouse ATP1A1 with

3

1   newly generated constructs to test whether these insertions alter the effects of point mutations at sites 111

2   and 122 known to confer NKA target-site insensitivity.

3

4   **Results:**

5   *Origin of the H1-H2 loop insertions*

6

7   A previous survey of ATP1A1 of 117 mammals and 70 birds established that chinchilla (a hystricognath

8   rodent) and sandgrouse (a pterocliform bird) have convergently evolved amino-acid insertions between

9   positions 114 and 115 of the H1-H2 extracellular loop (Mohammadi et al. 2022). In the chinchilla, the

10  inserted amino acid is methionine (insM), and in the sandgrouse it is aspartic acid (insD). To infer the

11  evolutionary origins of the insertions, we separately estimated maximum likelihood phylogenies of 26

12  rodent ATP1A1 sequences and 22 avian ATP1A1 sequences obtained from publicly available sources

13  (Table S1; Supplementary Datasets 1-2). The estimated phylogeny of rodent ATP1A1 sequences indicates

14  that the insertion in the H1-H2 loop evolved in a recent common ancestor of the hystricognath rodents,

15  which includes chinchillas, porcupines, guinea pigs, nutrias, mole rats, and allies (Marivaux et al. 2004)

16  (Fig. 2A). This implies that the insM is ancient, dating from 36 to 39 million years ago (Sallam et al.

17  2009). The estimated phylogeny of avian ATP1A1 sequences indicates that the insertion evolved in the

18  common ancestor of Pterocliformes, which includes all sandgrouse species (Fig. 2B). Despite its

19  relatively restricted phylogenetic distribution, insM of sandgrouse may also be quite ancient given that the

20  common ancestor of Pterocliformes diverged from its sister group, Mesitornithiformes, between 45 and

21  55 million years ago (Kuhl et al. 2021). In both cases the insertions are flanked by multiple additional

22  amino acid substitutions.

23

24  *The effect of insertions on CTS resistance*

25

26  To test the functional effects of the H1-H2 loop insertions, we used protein engineering to evaluate the

27  effects of various combinations of amino acid substitutions and insertions (Table 1). In the case of

28  chinchilla, we deleted insM on the wildtype ATP1A1 and on the ATP1A1 carrying a known resistance-

29  conferring mutation at site 122 (N122D). We had previously tested D122 because this amino-acid state is

30  shared by all members of the closely related clade of murid rodents (Price and Lingrel 1988; Mohammadi

31  et al. 2022); Fig. 2A). In the case of sandgrouse, we deleted insD on the wildtype ATP1A1 and on the

32  ATP1A1 carrying a known resistance-abolishing substitution at site 111 (R111Q) (Price and Lingrel

33  1988; Mohammadi et al. 2021).

34

1   For each recombinant protein, we quantified the level of CTS resistance as $IC_{50}$, which is the molar

2   concentration of CTS needed to reduce protein activity by 50%. We also quantified the rate of ATP

3   hydrolysis in the absence of CTS, which we used as a measure of native protein function (protein activity;

4   Table S2; Fig. 3). For ATP1A1 of both chinchilla and sandgrouse, we measured $IC_{50}$ of the wildtype

5   proteins and a combinatorially complete set of single- and double-mutant genotypes (Fig. 3; Table S3).

6   We found that insM in the chinchilla NKA causes a 19-fold decrease in $IC_{50}$ (linear regression, p = 0.002;

7   Table S3), while N122D causes a 90-fold decrease (linear regression, p = 1.37e-4; Table S3). Conversely,

8   N122D in the presence of insM increases $IC_{50}$ by over 700-fold (linear regression, p = 7.06e-6; Table S3).

9   These directionally contrasting nonadditive effects reveal that there is a significant epistatic interaction

10  involving the insertion and site 122 within the chinchilla ATP1A1 with regard to CTS resistance (2-way

11  ANOVA interaction term, $F_{1,8}$=145.4, p = 2e-6; Table S3).

12

13  In the case of sandgrouse, insD causes a six-fold increase in $IC_{50}$ (linear regression, p = 3.81e-4; Table

14  S3). The well-documented resistance producing substitution, Q111R, also significantly increases $IC_{50}$, by

15  11-fold (linear regression, p = 4.69e-5; Table S3). Together, however, insD and Q111R cause a six-fold

16  decrease in $IC_{50}$ (linear regreassion, p = 5.71e-4; Table S3). These results indicate that both insD and

17  Q111R have the potential to confer resistance, but that the two substitutions jointly exhibit "diminishing-

18  returns" epistasis (2-way ANOVA interaction term, $F_{1,8}$=30.287, p = 5.71e-4; Table S3).

19

20  *The effect of insertions on NKA activity*

21

22  Experiments on the various chinchilla and sandgrouse NKA mutants also reveal strong epistatic effects on

23  protein activity. The direction of these effects changed depending on the mutation combination. In

24  chinchilla, both insM and N122D cause a four-fold increase in activity (linear regression, p = 8.58e-4 and

25  p = 1.82e-3, respectivity; Table S3). However, together, they cause a three-fold decrease in activity

26  (linear regression, p = 5.68e-3; Table S3). In line with the nonadditive effects we observed, we found a

27  significant interaction between insM and N122D on protein activity (2-way ANOVA interaction term,

28  $F_{1,8}$=34.57, p <0.001; Table S3). In general, we found that protein activity in chinchilla ATP1A1

29  decreases as resistance increases (Fig. 3C-D).

30

31  Experiments on the sandgrouse ATP1A1 demonstrate similar multi-directional epistatic effects. We found

32  that insD causes a nine-fold decrease in activity (linear regression, p = 0.001; Table S3). Similarly,

33  Q111R causes a 19-fold decrease in protein activity (linear regression, p = 8.16e-6; Table S3). However,

34  when combined, they result in a 17-fold increase in activity (linear regression, p = 1.92e-4; Table S3).

5

1    These directionally constrasting effects indicate that there is a significant epistatic interaction between

2    insD and R111Q (2-way ANOVA interaction term, $F_{1,8}$=42.01, p <0.001; Table S3) with respect to

3    protein activity.

4

5    *The biophysical mechanism of insertion effects on resistance*

6

7    To investigate the biophysical mechanisms underlying the effects on resistance observed in recombinant

8    proteins, we used a homology model of a high-affinity structure of NKA to perform molecular docking

9    simulations (Zhen et al. 2012; Kanai et al. 2021) using Autodock Vina 1.1.2 (Trott and Olson 2010). We

10   modeled each recombinant protein and performed docking simulations using ouabain, the CTS used in

11   our functional experiments (see Figs. 3E-F). The trend in the docking scores for amino acid substitutions

12   and insertions is consistent with their observed effects on CTS resistance in functional experiments. The

13   amino acid substitutions and insertions altered the interaction network between ATP1A1 and ouabain,

14   thereby affecting ligand-binding affinity.

15

16   In the case of chinchilla ATP1A1, removing insM from the wildype chinchilla ATP1A1 results in loss of

17   H-bonds between the liganded ouabain and S119 and E908, despite the formation of new H-bonds with

18   E111 and D121 (Figs. 3E; 4B; 6A; S3A). The N122D mutation in the presence of insM is predicted to

19   alter the H-bonding network of the liganded ouabain and the receptor (Fig. 4C). This is largely

20   attributable to the loss of an H-bond between ouabain and N122. Although new H-bonds form between

21   C103, Y901 and ouabain, the loss of the H-bond with residue 122 is predicted to reduce binding affinity

22   (i.e., increasing resistance; Figs. 3E, 4C, 6A). N122D in the absence of insM results in loss of H-bonds

23   with S119 and E908 while recovering the H-bond with D122, as well as formation of additional H-bonds

24   with E116, E111, and D121 (Figs. 3E; 4D). The only difference between the two cases lies in the H-bond

25   between ouabain and the sidechain of E116, which seems crucial for stronger binding. Overall, across all

26   individual modelling replicates, N122D in the absence of insM results in the highest number of H-bonds.

27   In this case multiple bonds are formed by residues E111, E116 and D122 (Fig. S3A), which is predicted

28   to increase ouabain binding-affinity. The distances among H-bond donor-acceptor pairs are also shorter

29   for the N122D mutation in the absence of insM, leading to moderate interactions, which increases

30   binding-affinity (Fig. 6A).

31

32   In the case of sandgrouse ATP1A1, removal of insD from the wildtype ATP1A1 results in stronger H-

33   bonds with D884 but weakened H-bonds at R111 and D121 without changing overall affinity (Figs. 3F;

34   5B; 6B). R111Q is predicted to produce a slight alteration of the H-bond network, thereby altering the

6

docked conformation of ouabain. The R111Q mutation in the presence of insD results in the loss of an H-bond between ouabain and R111, while forming a new H-bond with K905 (Figs. 3F; 5C; S3B), but the net result is not expected to significantly alter ligand affinity. On the other hand, the R111Q substitution in the absence of insD recovers H-bonding at position 111. This leads to new H-bonds between ouabain and residues Q115 and G796 and stronger binding to ouabain (Figs. 3F; 5D; 6B; S3B). Thus, Q115 and G796 contribute significantly to the binding of oubain (Fig. 5D).

The predicted effects of amino acid substitutions and insertions follow a similar pattern for both species. In ATP1A1 of both chinchilla and sandgrouse, when insertions are removed from the H1-H2 loop containing mutagenized amino acid states at 111 and 122, respectively, the binding affinity is strengthened, resulting in less CTS-resistant enzymes. Conversely, when the insertions are removed from the wildtype ATP1A1 of both species, the binding affinity is weakened resulting in increased CTS resistance.

**Discussion:**

Previous work suggests that the evolution of CTS resistance is mediated by epistatic interactions between substitutions at a small number of key sites (Karageorgi et al. 2019; Taverner et al. 2019; Mohammadi et al. 2021; Mohammadi et al. 2022). Our protein engineering experiments and molecular docking simulations of chinchilla and sandgrouse NKA corroborate this pattern and reveal functional effects and epistatic interactions involving insertions in the H1-H2 extracellular loop of ATP1A1. Previous work also suggests that resistance-conferring substitutions of the H1-H2 loop are often associated with negative pleiotropy on enzyme activity (Zhen et al. 2012; Dalla et al. 2017; Taverner et al. 2019; Yang et al. 2019; Mohammadi et al. 2021; Mohammadi et al. 2022) and that these effects are often mitigated by other amino acid substitutions throughout the protein (Karageorgi et al. 2019; Taverner et al. 2019; Mohammadi et al. 2021). Our mutagenesis experiments reveal that, while amino acid substitutions at sites 111 and 122 can compromise NKA activity, amino acid insertions in the same subdomain can play a role in mitigating these effects. In the following paragraphs, we discuss how the results of our mutagenesis experiments and models contribute to our understanding of the evolution of target-site insensitivity of NKA to CTS and the potential importance of insertion-deletion substitutions to protein function and evolution.

N122D is shared by all Muroid rodents (Muroidea) and by neotropical grass frogs (*Leptodactylus*), and always co-occurs with the substitution Q111R (Mohammadi et al. 2022). While N122D has been shown to confer resistance individually in ATP1A1 of mouse (Price and Lingrel 1988) and neotropical grass frog

7

(Mohammadi et al. 2021), the protein's resistance to CTS increases by orders of magnitude when N122D is combined with Q111R. While not native to the chinchilla lineage, we show that N122D also has the potential to contribute to target-site insensitivity in chinchilla ATP1A1 when combined with the insM insertion. Surprisingly, N122D decreases resistance in the absence of the insertion. Our *in silico* modeling analysis shows that insM affects the network of H-bonds between ouabain and several residues within the CTS binding pocket. These bonds are weakened by the removal of insM but can also be strengthened if its removal is combined with N122D.

The wildtype sandgrouse ATP1A1, which has R111, exhibits a level of CTS resistance on par with a CTS-sensitive ATP1A1 paralog (ATP1A1S) of neotropical grass frogs engineered to carry the Q111R mutation (Mohammadi et al. 2022). However, unlike the frog, in the sandgrouse ATP1A1, there is no difference in levels of resistance between amino acid states Q111 and R111. We found instead that sandgrouse ATP1A1 retains resistance with R111 alone or with insD alone; resistance is only lost when both states are reverted. Our protein structure models reveal that mutating R111Q in the absence of insD leads to the formation of multiple strong hydrogen bonds, enabling high affinity complex formation

Given that the wildtype chinchilla ATP1A1 lacks target-site insensitivity, our results also illustrate how epistasis can represent a source of contingency in the evolution of CTS resistance. For example, in the case of chinchilla ATP1A1, the insM mutation influences which amino acid substitutions could potentially contribute to the evolution of target-site insensitivity. If chinchillas or other hystrigonath rodents with insM were subject to selection for CTS resistance, the presence of insM would potentiate the effect of N122D, an amino acid substitution that confers target-site insensitivity in the NKA of other rodents. Thus, although the insM mutation might have been neutral in the ancestor of hystrigonaths, it could influence the selection coefficients associated with future mutations in the same protein.

1   **Materials and Methods:**

2

3   **Data sources.** Sequence data for ATP1A1 of the common chinchilla and 25 related rodents, and of the
4   yellow-throated sandgrouse and 22 related birds were collected from publicly available sources (Table
5   S1). Seventeen of the bird ATP1A1 sequences were obtained from genome assemblies under BioProject
6   PRJNA545868 (Feng et al. 2020). ATP1B1 sequences for *Chinchilla lanigera* (Genbank: XM005398203)
7   and *Pterocles gutturalis* (Genbank: XM010081314) were also collected from publicly available sources.
8   Previously characterized functional properties of two recombinant chinchilla and two recombinant
9   yellow-throated sandgrouse NKAs were obtained from a previous study (Mohammadi et al. 2022) (see
10  Table S2).

11

12  **Phylogenetic tree inference.** Nucleotide sequences representing ATP1A1 cDNA for rodents and birds
13  were aligned separately using the MAFFT function in Geneious Prime v 2021.0.3 (Biomatters Ltd). The
14  alignments were analyzed with maximum likelihood (ML) inference using IQ-TREE v 2.1.2 (Minh et al.
15  2020). The IQ-TREE analysis was run using codon-based sequences (-st) and the best-fit model for each
16  partition with 1000 ultrafast bootstrap (UFB) replicates. All IQ-TREE analyses were performed using the
17  CIPRES Science Gateway online server (Miller et al. 2012). To describe ancestral states and
18  substitutions, we used standardized numbering of residues based on the sheep (*Ovis ares*) ATP1A1
19  sequence (Genbank: NC019458.2) minus 5 residues from the 5' end.

20

21  **Construction of expression vectors**. To supplement previous data (Mohammadi et al. 2022), we created
22  four additional mutagenized versions of ATP1A1 sequences (Invitrogen$^{TM}$ GeneArt) that were codon
23  optimized for *Spodoptera frugiperda* (Table 1). Newly generated plasmid constructs were deposited at
24  Addgene repository under accession numbers 191226-191229. We used recombinant NKA protein
25  constructs for chinchilla and sandgrouse generated by Mohammadi et al. (Mohammadi et al. 2022). First,
26  ATP1B1 genes were inserted into pFastBac ™ Dual Expression Vector (Thermo Scientific$^{TM}$;
27  Cat#10712024) at the p10 promoter with XhoI and PaeI (FastDigest; Thermo Scientific$^{TM}$; Cat#FD0694
28  and Cat#FD0593, respectively) and confirmed by sequencing. The ATP1A1 genes were inserted at the
29  P$_{PH}$ promoter of vectors already containing the corresponding ATP1B1 genes using In-Fusion® HD
30  Cloning Kit (Takara Bio; Cat#638910) and confirmed by sequencing. All resulting vectors had the
31  ATP1A1 gene under the control of the P$_{PH}$ promoter and a ATP1B1 gene under the p10 promoter. The
32  resulting two vectors were then subjected to site-directed mutagenesis (QuickChange II XL Site-Directed
33  Mutagenesis Kit; Agilent Technologies, La Jolla, CA, USA; Cat#200521) to introduce the amino acid
34  codons of interest.

9

1

**Generation of recombinant viruses and transfection into Sf9 cells (*Spodoptera frugiperda*).**
*Escherichia coli* DH10bac cells harboring the baculovirus genome (bacmid) and a transposition helper
vector (Thermo Fisher Scientific™; Cat#10361012) were transformed according to the manufacturer's
protocol with expression vectors containing the different gene constructs. Recombinant bacmids were
selected through PCR screening, grown, and isolated. Subsequently, Sf9 cells ($4 \times 10^5$ cells*ml) in 2 ml
of Insect-Xpress medium (Lonza; Cat#BE12-730P10) were transfected with recombinant bacmids using
Cellfectin II reagent (Gibco-Thermo Fisher Scientific™; Cat#10362100). After a three-day incubation
period, recombinant baculoviruses were isolated (P1) and used to infect fresh Sf9 cells ($1.2 \times 10^6$
cells*ml) in 10 ml of Insect-Xpress medium with 15 mg/ml gentamycin (Roth; Cat#0233.1) at a
multiplicity of infection of 0.1. Five days after infection, the amplified viruses were harvested (P2 stock).

**Preparation of Sf9 membranes.** For production of recombinant NKA, Sf9 cells were infected with the
P2 viral stock at a multiplicity of infection of 1e3. The cells ($1.6 \times 10^6$ cells*ml) were grown in 50 ml of
Insect-Xpress medium with 15 mg/ml gentamycin at 27°C in 500 ml flasks (Dalla et al. 2017). After 3
days, Sf9 cells were harvested by centrifugation at 3,000 x g for 10 min. The cells were stored at -80 °C,
and then resuspended at 0 °C in 15 ml of homogenization buffer (0.25 M sucrose, 2 mM EDTA, and 25
mM HEPES/Tris; pH 7.0). The resuspended cells were sonicated at 60 W (Sonopuls 2070; Bandelin
Electronic Company, Berlin, Germany) for three 45 s intervals at 0 °C. The cell suspension was then
subjected to centrifugation for 30 min at 10,000 x g (J2-21 centrifuge, Beckmann-Coulter, Krefeld,
Germany). The supernatant was collected and further centrifuged for 60 min at 100,000 x g at 4 °C
(Ultra- Centrifuge L-80, Beckmann-Coulter) to pellet the cell membranes. The pelleted membranes were
washed twice and resuspended in ROTIPURAN® p.a., ACS water (Roth; Cat#HN68.2) and stored at -20
°C. Protein concentrations were determined by Bradford assays using bovine serum albumin as a
standard. Three biological replicates were produced for each NKA.

**Verification by SDS-PAGE/western blotting.** For each biological replicate, 10 ug of protein were
solubilized in 4x SDS-polyacrylamide gel electrophoresis sample buffer and separated on SDS gels
containing 10% acrylamide. Subsequently, they were blotted on nitrocellulose membrane (Roth;
Cat#HP42.1). To block non-specific binding sites after blotting, the membrane was incubated with 5%
dried milk in TBS-Tween 20 for 1 h. After blocking, the membranes were incubated overnight at 4 °C
with the primary monoclonal antibody α5 (Developmental Studies Hybridoma Bank, University of Iowa,
Iowa City, IA, USA; RRID:AB_2166869). Because only membrane proteins were isolated from
transfected cells, detection of the α subunit also indicates the presence of the β subunit. The primary

10

1  antibody was detected using a goat-anti-mouse secondary antibody conjugated with horseradish

2  peroxidase (Dianova, Hamburg, Germany; Cat#115-035-003; RRID:AB_2617176). The staining of the

3  precipitated polypeptide-antibody complexes was performed by addition of 60 mg 4-chloro-1 naphtol

4  (Merck/Sigma-Aldrich; Cat#C8890) in 20 ml ice-cold methanol to 100 ml phosphate buffered saline

5  (PBS) containing 60 ul 30% $H_2O_2$. See Figure S1.

6

7  **Ouabain inhibition assay**. To determine the sensitivity of each NKA construct against the water-soluble

8  cardiotonic steroid, ouabain (Ouabain octahydrate 96%; Acrōs Organic; Cat#AC161730010s), 100 ug of

9  each protein was pipetted into each well in a nine-well row on a 96-well flat-bottom microplate

10 containing stabilizing buffers (see buffer formulas in (Petschenka et al., 2013)). The first six wells in a

11 nine-well row were exposed to exponentially decreasing concentrations of ouabain ($10^{-3}$ M, $10^{-4}$ M, $10^{-5}$

12 M, $10^{-6}$ M, $10^{-7}$ M, $10^{-8}$ M, dissolved in distilled $H_2O$). The seventh well was exposed to HPLC-grade

13 water only (experimental control). The eighth and ninth wells were exposed to a combination of an

14 inhibition buffer lacking KCl and $10^{-2}$ M ouabain (ouabain octahydrate 96%; Acrōs Organic;

15 Cat#AC161730010s) that completely inhibit the NKA's ATPase activity and allow us to measure

16 background ATPase activity (see (Petschenka et al., 2013)). The proteins were incubated at 37°C and 200

17 rpms for 10 minutes on a microplate shaker (BioShake iQ; Quantifoil Instruments, Jena, Germany;

18 Cat#1808-0506). Next, ATP (Adenosin-5-triphosphat Bis-(Tris)-salt hydrate; Merck/Sigma-Aldrich;

19 CAS#102047-34-7) was added to each well and the proteins were incubated again at 37°C and 200 rpms

20 for 20 minutes. The activity of $Na^+/K^+$-ATPases following ouabain exposure was determined by

21 quantification of inorganic phosphate (Pi) released from enzymatically hydrolyzed ATP. Reaction Pi

22 levels were measured according to the procedure described by (Taussky and Shorr 1953) (see (Petschenka

23 et al. 2013)). All assays were run in duplicate and the average of the two technical replicates was used for

24 subsequent statistical analyses. Absorbance for each well was measured at 650 nm with a plate

25 absorbance reader (BioRad Model 680 spectrophotometer and software package).

26

27 **ATP hydrolysis assay.** To determine the functional efficiency of different $Na^+/K^+$-ATPase constructs, we

28 calculated the amount of Pi hydrolyzed from ATP per mg of protein per minute. The measurements were

29 obtained from the same assay as described above. In brief, absorbance from the experimental control

30 reactions, in which 100 ug of protein was incubated without any inhibiting factors (i.e., ouabain or buffer

31 excluding KCl), were measured and translated to mM Pi from a standard curve that was run in parallel

32 (1.2 mM Pi, 1 mM Pi, 0.8 mM Pi, 0.6 mM Pi, 0.4 mM Pi, 0.2 mM Pi, 0 mM Pi). Raw assay data available

33 on Dryad at https://doi.org/10.5061/dryad.ngf1vhhxc.

34

11

1   **Statistical analyses.** Background phosphate absorbance levels from reactions with inhibiting factors were
2   used to calibrate phosphate absorbance in wells measuring ouabain inhibition and in the control wells
3   (Petschenka et al. 2013). For ouabain sensitivity measurements, calibrated absorbance values were
4   converted to percentage non-inhibited $Na^+,K^+$-ATPases activity based on measurements from the control
5   wells (Petschenka et al. 2013). These data were plotted and $\log_{10}$ $IC_{50}$ values were obtained for each
6   biological replicate from nonlinear fitting using a four-parameter logistic curve, with the top asymptote
7   set to 100 and the bottom asymptote set to zero. Curve fitting was performed with the nlsLM function of
8   the minipack.lm library in R. For comparisons of recombinant protein activity, the calculated Pi
9   concentrations of 100 ug of protein assayed in the absence of ouabain were converted to nmol Pi/mg
10  protein/min (Table S2). $IC_{50}$ values were $\log_{10}$-transformed prior to analysis to better meet the
11  assumptions of normality and homogeneity of variance. We used a 2-way ANOVA to assess interaction
12  effects of point mutations and insertion and followed with a Tukey's test to identify significant
13  differences between the different recombinant proteins (Table S3; Levene's Test for Homogeneity of
14  Variance for chinchilla $IC_{50}$: $F_{3,8}= 0.3525$ p=0.7888 and protein activity: $F_{3,8}=0.1622$ p=0.9188 and for
15  sandgrouse $IC_{50}$: $F_{3,8}=0.0243$ p=0.9945 and protein activity: $F_{3,8}=0.4561$ p=0.7203). We then used a linear
16  regression to infer the effect size of the insertions and point-mutations, individually and together, and
17  assess the additive/nonadditive effects of these mutant sites (Table S3). All statistical analyses were
18  implemented in R. Data were plotted using the ggplot2 package in R.
19
20  **Homology modelling and *in silico* mutagenesis**. The structures of the chinchilla (*Chinchilla lanigera*)
21  and sandgrouse (*Pterocles gutturalis*) NKAs are not available in the protein data bank (PDB) and were
22  thus obtained via homology modelling. The template structures required to perform homology modeling
23  were searched using the BLAST search tool (Altschul et al. 1997) implemented in PyMod 3 (Janson and
24  Paiardini 2021). Crystal structure of high affinity NKA from *Sus scrofa* (PDB ID: 7DDJ with 94.46% and
25  85.44% sequence identities with chinchilla and sandgrouse respectively) was used as template due to its
26  high homology and higher resolution.
27
28  Alignment of template and target sequences for homology modeling was performed using MUSCLE
29  (Edgar 2004) software via PyMod 3 graphical interface. Homology models of the structures were
30  performed using Modeller (Sali and Blundell 1993) implemented in PyMod 3. In addition, regions with
31  low DOPE scores (Shen and Sali 2006), including the loops in the binding region, were further refined
32  after initial modeling using Modeller. Following standard procedures, ligand molecules and N-terminal
33  amino acids were deleted and disulfide bonds were patched for more accurate modeling (Gray et al.
34  2003). The modeled structures were energetically minimized with the 1000-step Steepest Descent

12

1  algorithm using the AMBER99SB-ILDN force field (Lindorff-Larsen et al. 2010) in the OpenMM toolkit

2  (Eastman et al. 2017). Finally, the required point mutations were performed using PyMol (Schrödinger

3  2015) mutation wizard, where the romaters with a probability greater than 20% are chosen. After

4  mutations, minimization is performed under the same conditions for another 1000 steps using the

5  OpenMM toolkit to optimize amino acid side chain orientations.

6

7  **Molecular docking**. Docking calculations for the modeled ATPase structures were performed using

8  Autodock Vina 1.1.2 (Trott and Olson 2010). For docking simulations, ouabain (OBN) ligand molecule

9  was extracted from the NKA-ouabain co-crystallized structure (PDB ID: 7DDJ) from Protein Data Bank

10  (PDB) (Berman et al. 2002). Hydrogen atoms were added and the point charges were corrected using

11  AutoDock Tools (ADT) graphical interface software included in MGLTools 1.5.7 (Sanner 1999). A grid

12  box of dimensions 35x35x40 Å was constructed to include the binding pocket of the ligand for all

13  docking experiments based on the co-crystal structure of ouabain and ATPase complex (PDB ID: 7DDJ)

14  and exhaustiveness value was taken as 10. In addition, interacting residues were selected from the

15  ATPase and OBN complex co-crystal structure (PDB ID: 7DDJ) using LIGPLOT (Wallace et al. 1995)

16  (Fig. 3E-F). In particular, residues involved in hydrogen bonding (Q111, Q119, E312 and T797) were

17  selected to be flexible in the docking process (Table S4) (Ravindranath et al. 2015). For each docking

18  calculation, 10 repetitions were performed and poses with lowest docking scores (low scores correspond

19  to best structures—the ones with highest affinity) were extracted (Fig. 3, Fig. S2). The pose with the

20  lowest docking score corresponds to the best binding ligand. PyMOL was used for visual inspection of

21  the docked structures and Discovery Studio was used for hydrogen binding determinations.

22

23  **Figures:**

24

25  **Figure 1.** (A) Crystal structure of an Na$^+$,K$^+$-ATPase with a bound cardiotonic steroid (bufalin) in red

26  (*Sus scrofa*; PDB 4RES). The α-subunit is colored in light grey tones and the β-subunit is colored in dark

27  grey. The zoomed-in panel shows the H1-H2 extracellular loop, highlighted in blue. Two sites, 111 and

28  122, at which substitutions have been repeatedly implicated in CTS resistance are labeled in blue. (B) The

29  ancestral tetrapod amino acid sequence of the H1-H2 extracellular loop (Mohammadi et al. 2022) is

30  indicated by blue text and the numbering follows the sheep ATP1A1 sequence convention. The

31  corresponding wildtype sequences for the yellow-throated sandgrouse (*Pterocles gutturalis*) and the

32  common chinchilla (*Chinchilla lanigera*) are listed below. The inferred position of the insertions in both

33  species are highlighted in grey.

34
35

13

1   **Figure 2.** Maximum likelihood phylogenies inferred from ATP1A1 nucleotide sequences using IQ-TREE

2   v 2.1.2 (Minh et al. 2020). Amino acid sequences of the H1-H2 loop (positions 111 to 122) are aligned to

3   the right of each phylogeny. (A) Rodent protein tree inferred from an alignment of 26 protein-coding

4   DNA sequences. Branch-tip labels in blue denote species with the insM insertion and includes all

5   members of the clade Hystricognathi. All five species also share several other amino acid substitutions,

6   indicating that they originated in the common ancestor of Hystricognathi. (B) Bird tree inferred from an

7   alignment of 22 protein-coding DNA sequences. Branch-tip labels in blue denote species with the insD

8   insertion, which include all sampled sandgrouse (Pterocliformes). Similar to the hystricognath rodents,

9   the insertion (insD in this case) is accompanied by several other substitutions shared by all sandgrouse.

10

11

12   **Figure 3.** Joint in vitro functional properties of eight engineered $Na^+,K^+$-ATPases (NKAs) from the

13   common chinchilla (*Chinchilla lanigera*; panels A and C) and yellow-throated sandgrouse (*Pterocles*

14   *gutturalis*; panels B and D). For each recombinant protein, amino-acid states at positions 111, 122, and

15   the insertion are denoted, and mutagenesis-derived changes at these states are underlined. For each

16   species, recombinant proteins consist of (from left to right), the NKA without the insertion or derived

17   mutation, the NKA with the insertion, the NKA with the derived mutation (N122D for chinchilla and

18   Q111R for sandgrouse), and the NKA with the derived mutation and the insertion. Mean $\pm$ SEM $\log_{10}$

19   $IC_{50}$ (i.e., CTS resistance) of three biological replicates for each recombinant NKA is plotted on the y-axis

20   of panels A and B. Mean $\pm$ SEM ATP hydrolysis rate (i.e., protein activity) for the same proteins is

21   plotted on the y-axis of panels C and D. Raw data from three biological replicates of each protein are

22   shown as open circles and jittered with respect to the x-axis. Significant differences (Tukey's HSD; Table

23   S3) between proteins for each panel are indicated by different letters. Corresponding in silico docking

24   scores from the best docking position of ouabain docked in the common chinchilla (panel E) and yellow-

25   throated sandgrouse (panel F) recombinant NKAs. The best docking member for each case is defined as

26   the docking position closest to the position of ouabain in the published co-crystal structure (PDB id:

27   7DDJ). The predicted binding energy of 10 individual docking simulations is represented by open circles,

28   jittered with respect to the x-axis, and the mean values ($\pm$ SEM) by colored spheres.

29

30

31   **Figure 4.** The docked structure of ouabain in the binding pocket of chinchilla ATP1A1 modeled together

32   with hydrogen bonds obtained from molecular docking simulations. The dark gray ligand in panel A

33   belongs to the co-crystal structure of ouabain with ATP1A1 (PDB ID: 7DDJ), and the green ligand in all

34   panels represents docked ouabain position to the chinchilla wildtype (WT) model. The yellow ligand in

35   panels B-D shows docked ouabain conformations in different mutant ATP1A1 structures. The interacting

14

1  residues are labeled and shown in stick form and the H-bonds between the ATP1A1 and ouabain are

2  represented as dashed lines. The transmembrane helices of the homology models are superimposed to the

3  same region of the co-crystal structure. The extracellular region of the α-subunit is removed for

4  simplicity.

5

6

7  **Figure 5**. The docked structure of ouabain in the binding pocket of sandgrouse ATP1A1 modeled

8  together with hydrogen bonds obtained from molecular docking simulations. The dark gray ligand in

9  panel A belongs to the co-crystal structure of ouabain with ATP1A1 (PDB ID: 7DDJ), and the green

10 ligand in all panels represents docked ouabain pose to the Sandgrouse wildtype (WT) model. The yellow

11 ligand in panels B-D shows docked ouabain conformations in different mutant ATP1A1 structures. The

12 interacting residues are labeled and shown in sticks and the H-bonds between the ATP1A1 and ouabain

13 are represented as dashed lines. The transmembrane helices of the homology models are superimposed to

14 the same region of the co-crystal structure. The extracellular region of the α-subunit is removed for

15 simplicity.

16

17 **Figure 6**. The distribution of distances between the donor-acceptor pairs making H-bonds among

18 corresponding ATP1A1 residues and ouabain for A) chinchilla and B) sandgrouse. The distances among

19 bond forming atoms depict the strength of corresponding H-bonds. The lower the bond distance, the

20 stronger the bond is. The H-bond lengths, as the distances between the H-atoms and the corresponding

21 acceptor atoms, are calculated via Discovery Studio Visualizer v21.1.0.20298 (Visualizer 2021) with its

22 default settings.

23

24 **[PLEASE PLACE THIS TABLE NEAR FIGURE 3]**

25 **Table 1.** List of gene constructs used to test functional effects of amino acid mutations and species-

26 specific insertions in ATP1A1 of the common chinchilla (*Chinchilla lanigera*) and yellow-throated

27 sandgrouse (*Pterocles gutturalis*). The corresponding wildtype ATP1B1 gene of each recombinant protein

28 construct was co-expressed with ATP1A1. Asterisks indicate constructs generated by (Mohammadi et al.

29 2022). The abbreviations "wt" and "mut" indicate wildtype or derived states, respectively.

30

| Construct Name | Engineered substitution | AA state at 111 | AA state at 122 | Insertion state |
|---|---|---|---|---|
| *Chinchilla | none | E111 (wt) | N122 (wt) | insM (wt) |
| Chinchilla–insM | –insM | E111 (wt) | N122 (wt) | none (mut) |

| | | | | |
|---|---|---|---|---|
| *Chinchilla+N122D | N122D | E111 (wt) | D122 (mut) | insM (wt) |
| Chinchilla+N122D–insM | N122D –insM | E111 (wt) | D122 (mut) | none (mut) |
| *Sandgrouse | none | R111 (wt) | N122 (wt) | insD (wt) |
| Sandgrouse–insD | –insD | R111 (wt) | N122 (wt) | none (mut) |
| *Sandgrouse+R111Q | R111Q | Q111 (mut) | N122 (wt) | insD (wt) |
| Sandgrouse+R111Q–insD | R111Q –insD | Q111 (mut) | N122 (wt) | none (mut) |

**References**

Agrawal AA, Petschenka G, Bingham RA, Weber MG, Rasmann S. 2012. Toxic cardenolides: chemical ecology and coevolution of specialized plant-herbivore interactions. *New Phytol.* 194:28–45.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25:3389–3402.

Aperia A. 2007. New roles for an old enzyme: Na, K-ATPase emerges as an interesting drug target. *J. Intern. Med.* 261:44–52.

Bagrov AY, Shapiro JI, Fedorova OV. 2009. Endogenous cardiotonic steroids: physiology, pharmacology, and novel therapeutic targets. *Pharmacol. Rev.* 61:9–38.

Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S. 2002. The protein data bank. *Acta Crystallogr. D Biol. Crystallogr.* 58:899–907.

Blanco G, Mercer RW. 1998. Isozymes of the Na-K-ATPase: heterogeneity in structure, diversity in function. *Am. J. Physiol.-Ren. Physiol.* 275:F633–F650.

Brodie ED. 2009. Toxins and venoms. *Curr. Biol.* 19:R931–R935.

de la Chaux N, Messer PW, Arndt PF. 2007. DNA indels in coding regions reveal selective constraints on protein evolution in the human lineage. *BMC Evol. Biol.* 7:1–12.

Dalla S, Baum M, Dobler S. 2017. Substitutions in the cardenolide binding site and interaction of subunits affect kinetics besides cardenolide sensitivity of insect Na, K-ATPase. *Insect Biochem. Mol. Biol.* 89:43–50.

Dobler S, Dalla S, Wagschal V, Agrawal AA. 2012. Community-wide convergent evolution in insect adaptation to toxic cardenolides by substitutions in the Na, K-ATPase. *Proc. Natl. Acad. Sci.* 109:13040–13045.

Dobler S, Wagschal V, Pietsch N, Dahdouli N, Meinzer F, Romey-Glüsing R, Schütte K. 2019. New ways to acquire resistance: imperfect convergence in insect adaptations to a potent plant toxin. *Proc. Biol. Sci.* 286:20190883.

16

1  Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, Wang L-P, Simmonett
2        AC, Harrigan MP, Stern CD. 2017. OpenMM 7: Rapid development of high performance
3        algorithms for molecular dynamics. *PLoS Comput. Biol.* 13:e1005659.

4  Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
5        *Nucleic Acids Res.* 32:1792–1797.

6  Feng S, Stiller J, Deng Y, Armstrong J, Fang Q, Reeve AH, Xie D, Chen G, Guo C, Faircloth BC.
7        2020. Dense sampling of bird diversity increases power of comparative genomics.
8        *Nature* 587:252–257.

9  Gong LI, Bloom JD. 2014. Epistatically interacting substitutions are enriched during adaptive
10        protein evolution. *PLoS Genet.* 10:e1004328.

11 Gray JJ, Moughon S, Wang C, Schueler-Furman O, Kuhlman B, Rohl CA, Baker D. 2003. Protein–
12        protein docking with simultaneous optimization of rigid-body displacement and side-
13        chain conformations. *J. Mol. Biol.* 331:281–299.

14 Groen SC, Whiteman NK. 2021. Convergent evolution of cardiac-glycoside resistance in
15        predators and parasites of milkweed herbivores. *Curr. Biol. CB* 31:R1465–R1466.

16 Hutchinson DA, Mori A, Savitzky AH, Burghardt GM, Wu X, Meinwald J, Schroeder FC. 2007.
17        Dietary sequestration of defensive steroids in nuchal glands of the Asian snake
18        Rhabdophis tigrinus. *Proc. Natl. Acad. Sci. U. S. A.* 104:2265–2270.

19 Janson G, Paiardini A. 2021. PyMod 3: a complete suite for structural bioinformatics in PyMOL.
20        *Bioinformatics* 37:1471–1472.

21 Kanai R, Cornelius F, Ogawa H, Motoyama K, Vilsen B, Toyoshima C. 2021. Binding of
22        cardiotonic steroids to Na+, K+-ATPase in the E2P state. *Proc. Natl. Acad. Sci.* 118.

23 Karageorgi M, Groen SC, Sumbul F, Pelaez JN, Verster KI, Aguilar JM, Hastings AP, Bernstein SL,
24        Matsunaga T, Astourian M, et al. 2019. Genome editing retraces the evolution of toxin
25        resistance in the monarch butterfly. *Nature* 574:409–412.

26 Köksoy AA. 2002. Na+, K+-ATPase: a review. *J Ank. Med Sch* 24:73–82.

27 Krenn L, Kopp B. 1998. Bufadienolides from animal and plant sources. *Phytochemistry* 48:1–29.

28 Kuhl H, Frankl-Vilches C, Bakker A, Mayr G, Nikolaus G, Boerno ST, Klages S, Timmermann B,
29        Gahr M. 2021. An unbiased molecular approach using 3'-UTRs resolves the avian family-
30        level tree of life. *Mol. Biol. Evol.* 38:108–127.

31 Laursen M, Gregersen JL, Yatime L, Nissen P, Fedosova NU. 2015. Structures and
32        characterization of digoxin- and bufalin-bound Na+,K+-ATPase compared with the
33        ouabain-bound complex. *Proc. Natl. Acad. Sci. U. S. A.* 112:1755–1760.

1   Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. 2010. Improved
2       side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 78:1950–
3       1958.

4   Marivaux L, Vianey-Liaud M, Jaeger J-J. 2004. High-level phylogeny of early Tertiary rodents:
5       dental evidence. *Zool. J. Linn. Soc.* 142:105–134.

6   Marshall BM, Casewell NR, Vences M, Glaw F, Andreone F, Rakotoarison A, Zancolli G, Woog F,
7       Wüster W. 2018. Widespread vulnerability of Malagasy predators to the toxins of an
8       introduced toad. *Curr. Biol. CB* 28:2194.

9   McGlothlin JW, Kobiela ME, Feldman CR, Castoe TA, Geffeney SL, Hanifin CT, Toledo G, Vonk FJ,
10      Richardson MK, Brodie Jr ED. 2016. Historical contingency in a multigene family
11      facilitates adaptive evolution of toxin resistance. *Curr. Biol.* 26:1616–1621.

12  Miller MA, Pfeiffer W, Schwartz T. 2012. The CIPRES science gateway: enabling high-impact
13      science for phylogenetics researchers with limited resources. In: p. 1–8.

14  Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R.
15      2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the
16      genomic era. *Mol. Biol. Evol.* 37:1530–1534.

17  Miton CM, Tokuriki N. 2022. Insertions and Deletions (Indels): A Missing Piece of the Protein
18      Engineering Jigsaw. *Biochemistry*.

19  Mobasheri A, Avila J, Cózar-Castellano I, Brownleader MD, Trevan M, Francis MJ, Lamb JF,
20      Martín-Vasallo P. 2000. Na+, K+-ATPase isozyme diversity; comparative biochemistry
21      and physiological implications of novel functional interactions. *Biosci. Rep.* 20:51–91.

22  Mohammadi S, Gompert Z, Gonzalez J, Takeuchi H, Mori A, Savitzky AH. 2016. Toxin-resistant
23      isoforms of Na+/K+-ATPase in snakes do not closely track dietary specialization on toads.
24      *Proc. Biol. Sci.* 283:20162111.

25  Mohammadi S, Herrera-Álvarez S, Yang L, del Pilar Rodríguez-Ordoñez M, Zhang K, Storz JF,
26      Dobler S, Crawford AJ, Andolfatto P. 2022. Constraints on the evolution of toxin-
27      resistant Na,K-ATPases have limited dependence on sequence divergence. *PLoS Genet.*
28      18:e1010323.

29  Mohammadi S, Yang L, Harpak A, Herrera-Álvarez S, Del Pilar Rodríguez-Ordoñez M, Peng J,
30      Zhang K, Storz JF, Dobler S, Crawford AJ, et al. 2021. Concerted evolution reveals co-
31      adapted amino acid substitutions in Na+K+-ATPase of frogs that prey on toxic toads.
32      *Curr. Biol. CB* 31:2530-2538.e10.

33  Petschenka G, Fandrich S, Sander N, Wagschal V, Boppré M, Dobler S. 2013. Stepwise evolution
34      of resistance to toxic cardenolides via genetic substitutions in the Na+/K+ -ATPase of
35      milkweed butterflies (lepidoptera: Danaini). *Evol. Int. J. Org. Evol.* 67:2753–2761.

Price EM, Lingrel JB. 1988. Structure-function relationships in the Na,K-ATPase alpha subunit: site-directed mutagenesis of glutamine-111 to arginine and asparagine-122 to aspartic acid generates a ouabain-resistant enzyme. *Biochemistry* 27:8400–8408.

Price EM, Rice DA, Lingrel JB. 1990. Structure-function studies of Na, K-ATPase. Site-directed mutagenesis of the border residues from the H1-H2 extracellular domain of the alpha subunit. *J. Biol. Chem.* 265:6638–6641.

Ravindranath PA, Forli S, Goodsell DS, Olson AJ, Sanner MF. 2015. AutoDockFR: advances in protein-ligand docking with explicitly specified binding site flexibility. *PLoS Comput. Biol.* 11:e1004586.

Sali A, Blundell TL. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234:779–815.

Sallam HM, Seiffert ER, Steiper ME, Simons EL. 2009. Fossil and molecular evidence constrain scenarios for the early evolutionary and biogeographic history of hystricognathous rodents. *Proc. Natl. Acad. Sci. U. S. A.* 106:16722–16727.

Sanner MF. 1999. Python: a programming language for software integration and development. *J. Mol. Graph. Model.* 17:57–61.

Savino S, Desmet T, Franceus J. 2022. Insertions and deletions in protein evolution and engineering. *Biotechnol. Adv.* 60:108010.

Schrödinger L. 2015. The PyMOL molecular graphics system, version 1.8.

Shen M-Y, Sali A. 2006. Statistical potential for assessment and prediction of protein structures. *Protein Sci. Publ. Protein Soc.* 15:2507–2524.

Storz JF. 2016. Causes of molecular convergence and parallelism in protein evolution. *Nat. Rev. Genet.* 17:239–250.

Storz JF. 2018. Compensatory mutations and epistasis for protein function. *Curr. Opin. Struct. Biol.* 50:18–25.

Taussky HH, Shorr E. 1953. A microcolorimetric method for the determination of inorganic phosphorus. *J. Biol. Chem.* 202:675–685.

Taverner AM, Yang L, Barile ZJ, Lin B, Peng J, Pinharanda AP, Rao AS, Roland BP, Talsma AD, Wei D, et al. 2019. Adaptive substitutions underlying cardiac glycoside insensitivity in insects exhibit epistasis in vivo. *eLife* 8:e48224.

Trott O, Olson AJ. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31:455–461.

19

1    Ujvari B, Casewell NR, Sunagar K, Arbuckle K, Wüster W, Lo N, O'Meally D, Beckmann C, King
2         GF, Deplazes E, et al. 2015. Widespread convergence in toxin resistance by predictable
3         molecular evolution. *Proc. Natl. Acad. Sci. U. S. A.* 112:11911–11916.

4    Ujvari B, Mun H, Conigrave AD, Bray A, Osterkamp J, Halling P, Madsen T. 2013. Isolation breeds
5         naivety: island living robs Australian varanid lizards of toad-toxin immunity via four-
6         base-pair mutation. *Evol. Int. J. Org. Evol.* 67:289–294.

7    Visualizer BDS. 2021. Dassault Systemes; San Diego, CA, USA: 2021. *Version V21* 1:20298.

8    Wallace AC, Laskowski RA, Thornton JM. 1995. LIGPLOT: a program to generate schematic
9         diagrams of protein-ligand interactions. *Protein Eng. Des. Sel.* 8:127–134.

10   Weinreich DM, Delaney NF, DePristo MA, Hartl DL. 2006. Darwinian evolution can follow only
11        very few mutational paths to fitter proteins. *science* 312:111–114.

12   Yang L, Ravikanthachari N, Mariño-Pérez R, Deshmukh R, Wu M, Rosenstein A, Kunte K, Song H,
13        Andolfatto P. 2019. Predictability in the evolution of Orthopteran cardenolide
14        insensitivity. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 374:20180246.

15   Zhen Y, Aardema ML, Medina EM, Schumer M, Andolfatto P. 2012. Parallel molecular evolution
16        in an herbivore community. *Science* 337:1634–1637.

17

**A**

extracellular

intracellular

ins

111

122

**B**

| | 1 1 1 | 1 1 2 | 1 1 3 | 1 1 4 | — | 1 1 5 | 1 1 6 | 1 1 7 | 1 1 8 | 1 1 9 | 1 2 0 | 1 2 1 | 1 2 2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ancestral | Q | A | A | T | — | E | E | E | P | Q | N | D | N |
| | R | S | . | V | D | Q | . | A | A | D | . | . | . |
| | E | S | L | S | M | . | . | G | . | S | . | . | . |
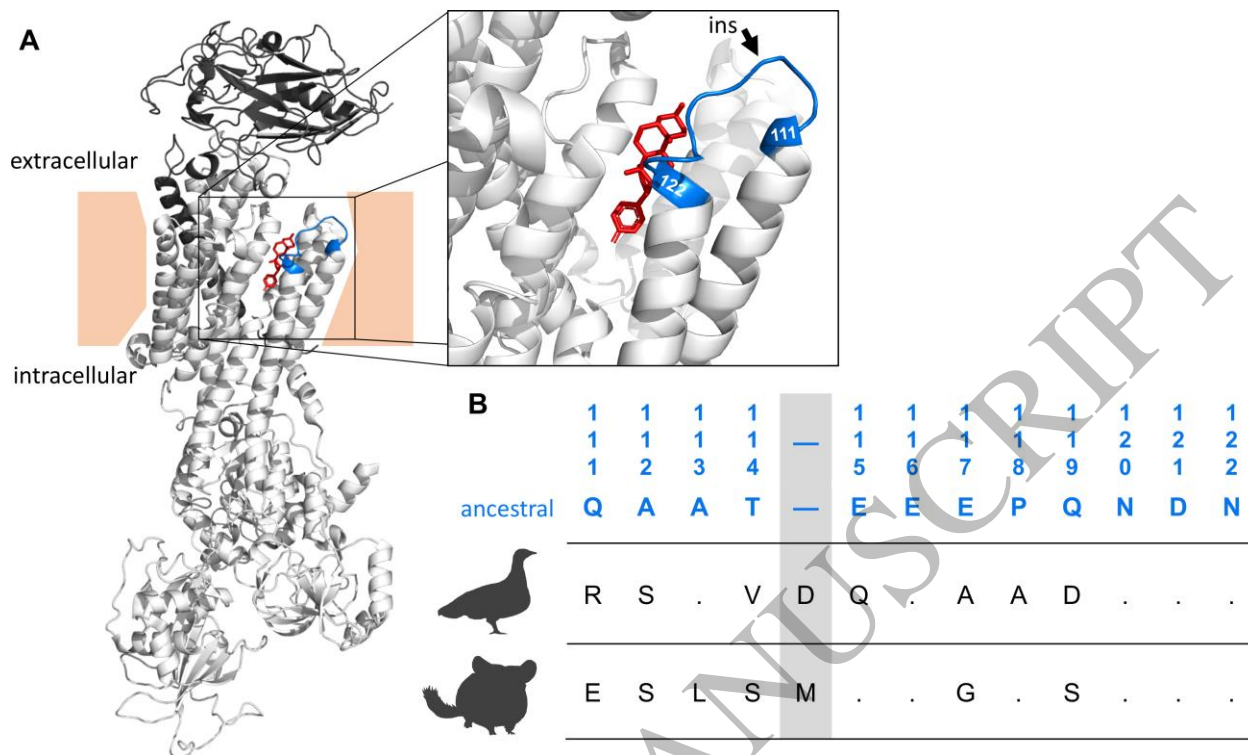
**Figure 1**
**165x99 mm (4.5 x  DPI)**

21
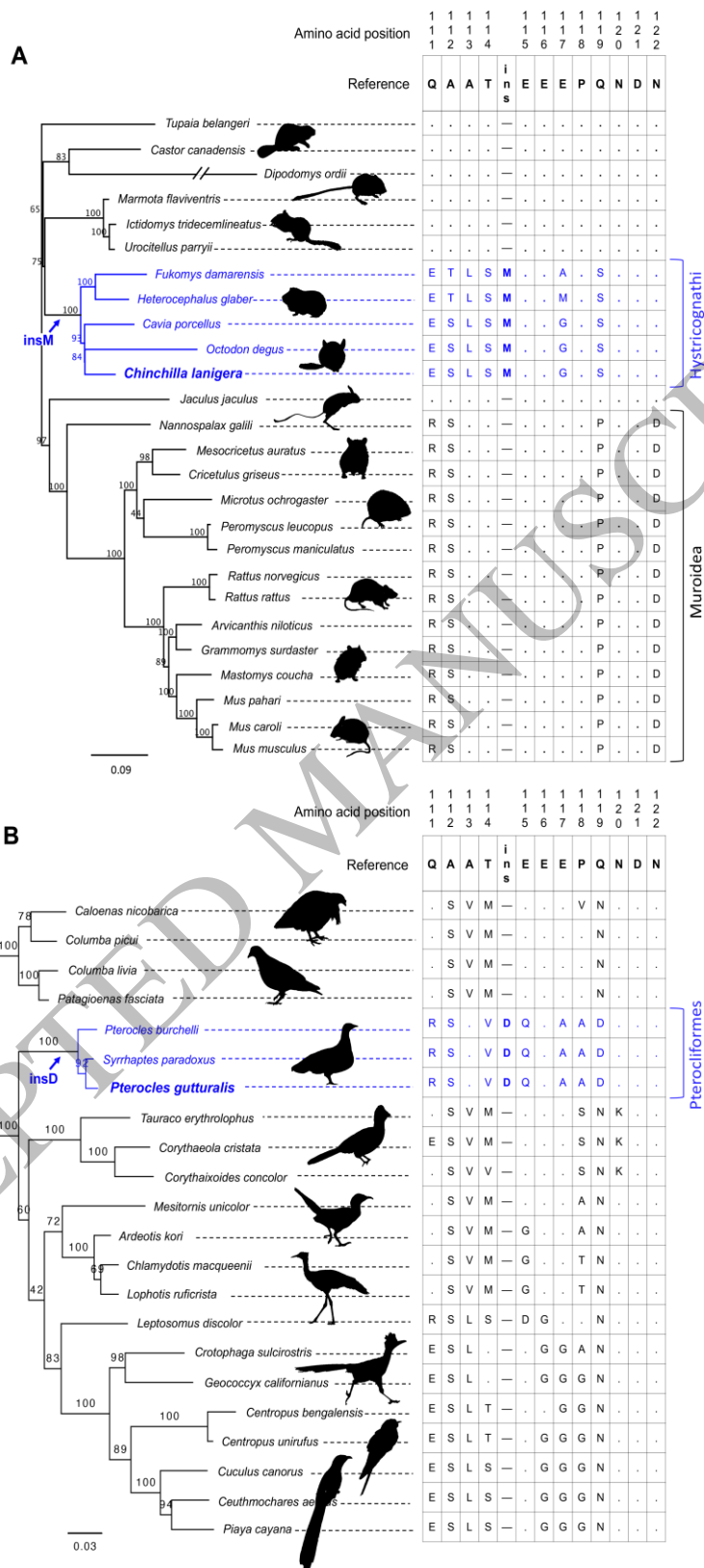
**Figure 2**
111x229 mm (4.5 x DPI)

1
2
3
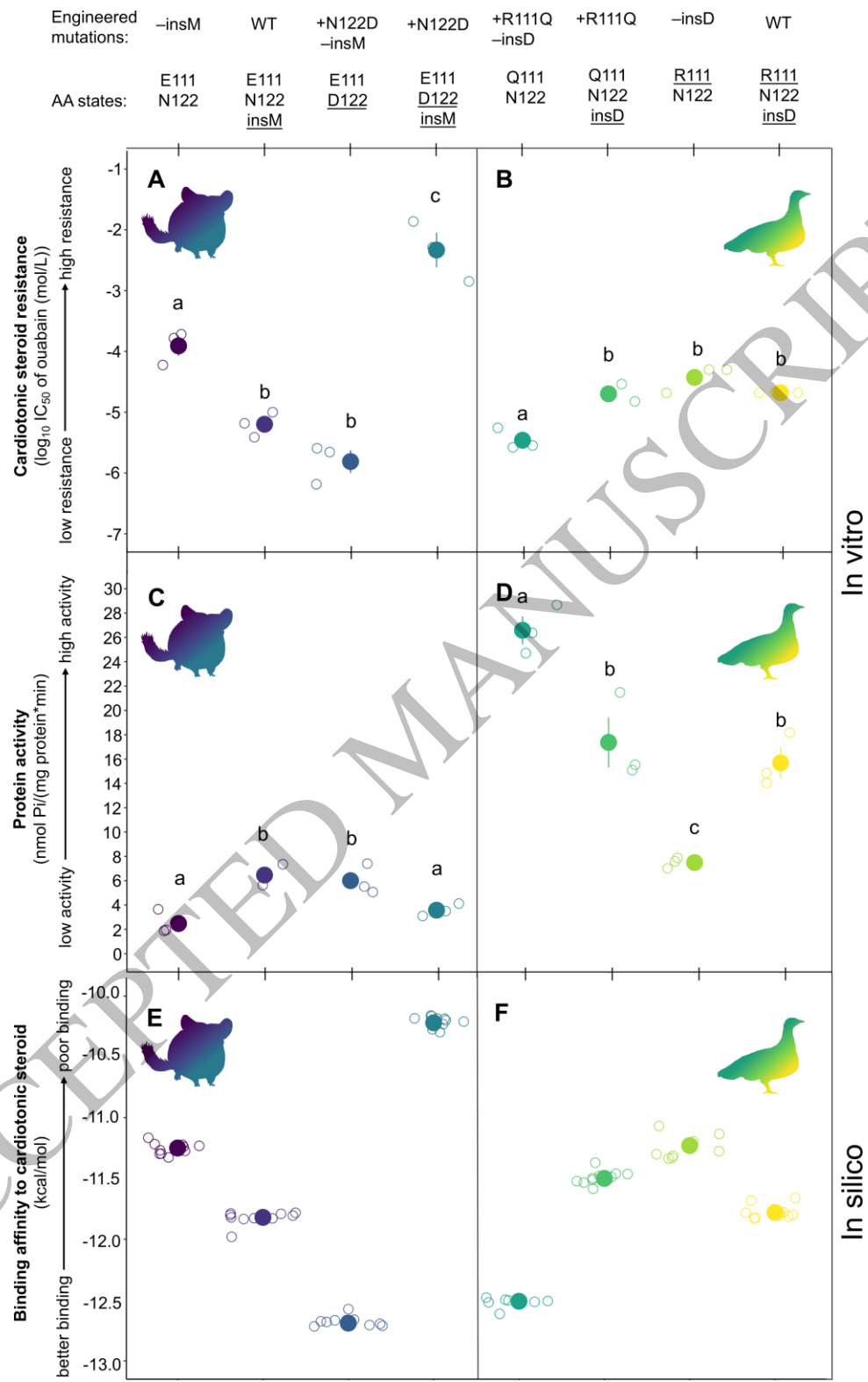
**Figure 3**
**142x229 mm (4.5 x  DPI)**
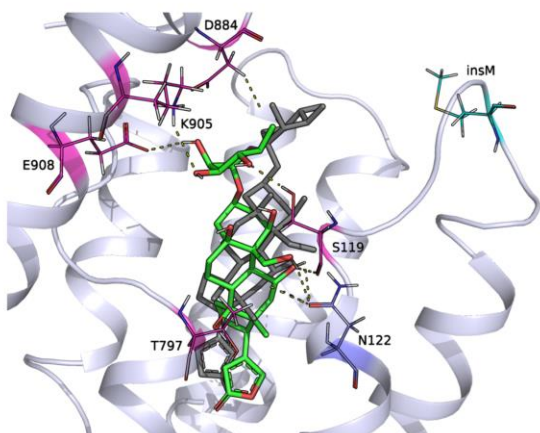
**A**  Chinchilla WT

**B**  Chinchilla - insM
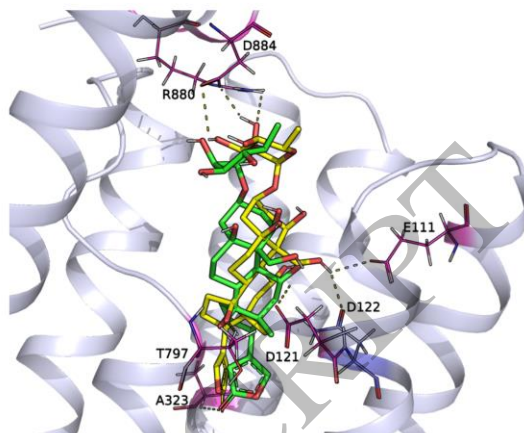
**C**  Chinchilla + N122D
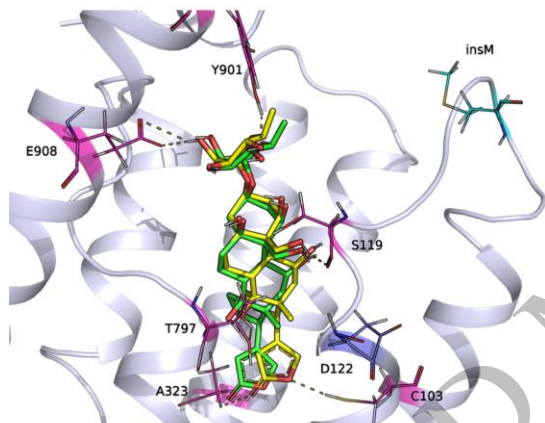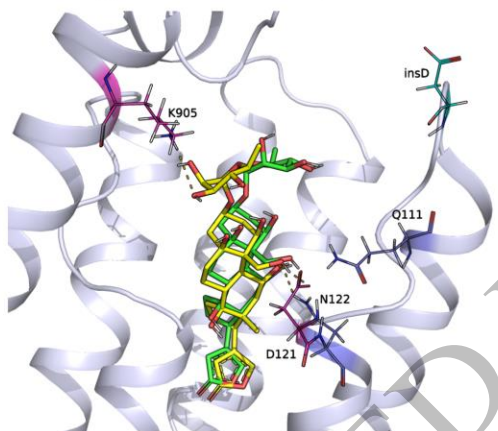
**D**  Chinchilla + N122D - insM
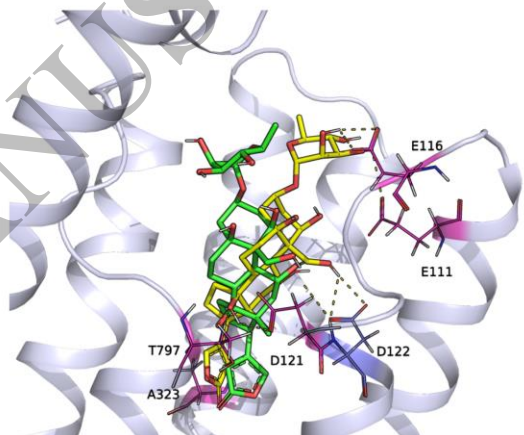
**Figure 4**
**165x128 mm (4.5 x  DPI)**
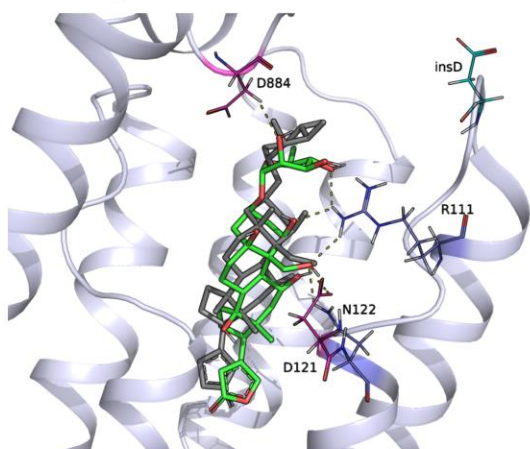
**A** Sandgrouse WT

**B** Sandgrouse - insD
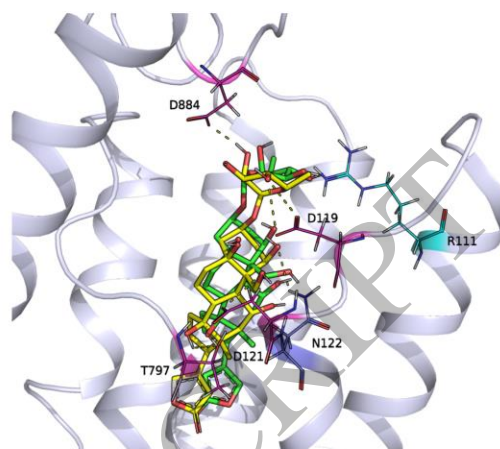
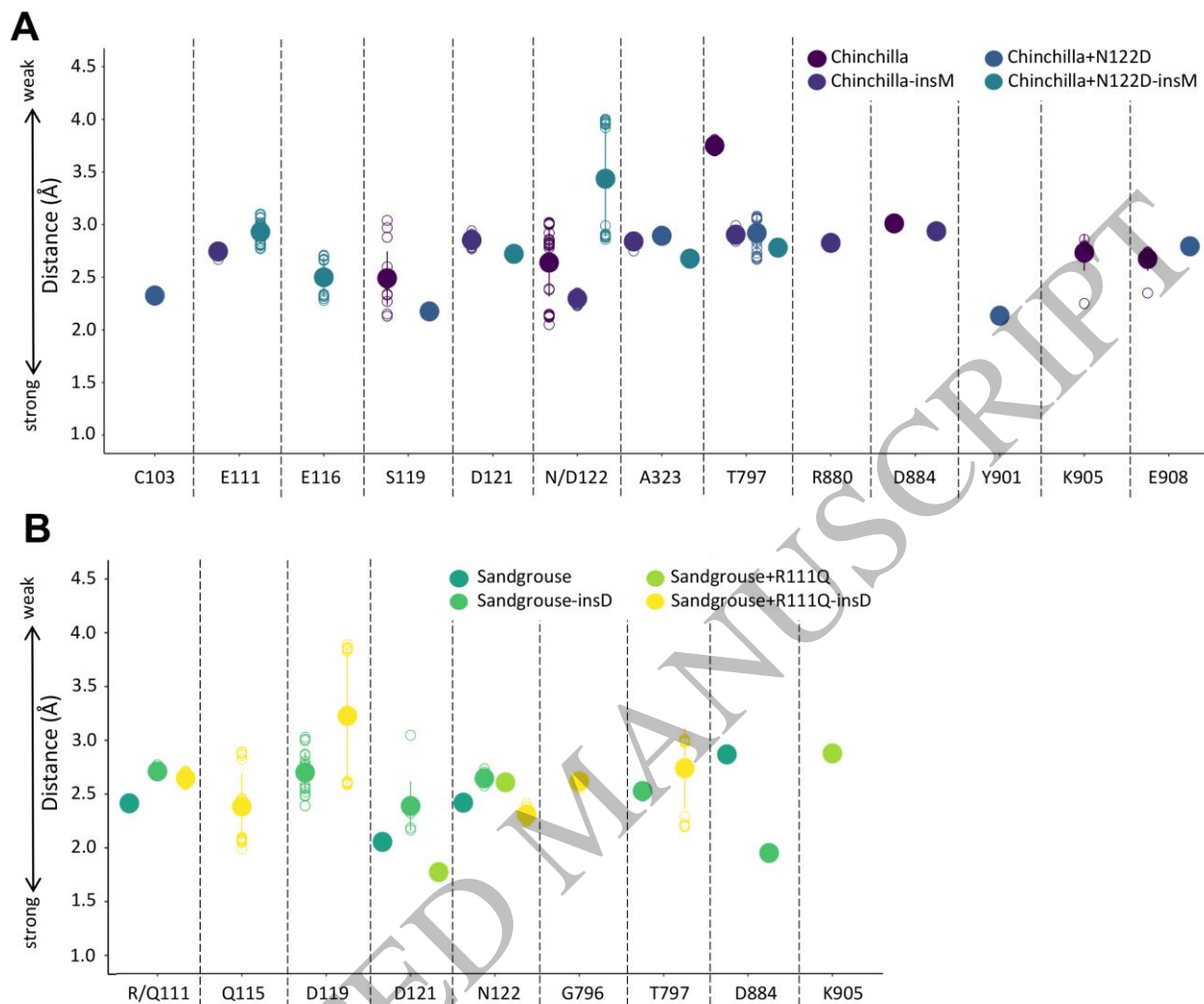**C** Sandgrouse + R111Q

**D** Sandgrouse + R111Q - insD
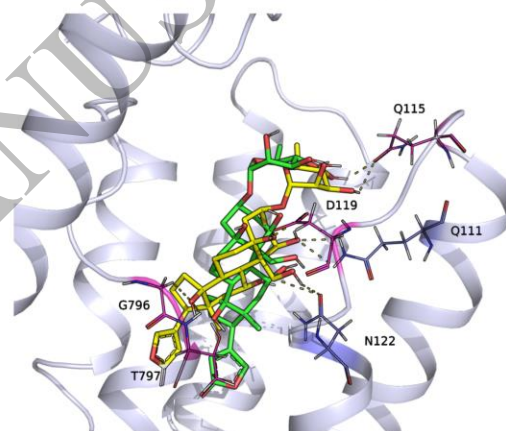
**Figure 5**
**165x131 mm (4.5 x  DPI)**

**Figure 6**
**165x138 mm (4.5 x  DPI)**