

Evolution of multifunctionality through a pleiotropic substitution in the innate immune protein S100A9

*Joseph L. Harman^{1,2}, Andrea N. Loes^{1,2}, Gus D. Warren^{1,2}, Maureen C. Heaphy^{1,2}, Kirsten J. Lampi³,
Michael J. Harms^{1,2*}*

¹Department of Chemistry and Biochemistry, University of Oregon, Eugene, OR, United States

²Institute of Molecular Biology, University of Oregon, Eugene, OR, United States

³Oregon Health & Science University, Portland, Oregon 97239, USA

Funding sources:

This research was funded by grants from the American Heart Association (AHA16 15BGIA22830013, MJH) and the National Institutes of Health (NIH-3R01GM117140-03S1, MJH; NIH-T32GM007413, JLH, ANL). MJH is a Pew Scholar in the Biomedical Sciences, supported by The Pew Charitable Trusts. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ABSTRACT

Multifunctional proteins are evolutionary puzzles: how do proteins evolve to satisfy multiple functional constraints? S100A9 is one such multifunctional protein. It potently amplifies inflammation via Toll-like receptor 4 and is antimicrobial as part of a heterocomplex with S100A8. These two functions are seemingly regulated by proteolysis: S100A9 is readily degraded, while S100A8/S100A9 is resistant. We take an evolutionary biochemical approach to show that S100A9 evolved both functions and lost proteolytic resistance from a weakly proinflammatory, proteolytically resistant amniote ancestor. We identify a historical substitution that has pleiotropic effects on S100A9 proinflammatory activity and proteolytic resistance but has little effect on S100A8/S100A9 antimicrobial activity. We thus propose that mammals evolved S100A8/S100A9 antimicrobial and S100A9 proinflammatory activities concomitantly with a proteolytic “timer” to selectively regulate S100A9. This highlights how the same mutation can have pleiotropic effects on one functional state of a protein but not another, thus facilitating the evolution of multifunctionality.

INTRODUCTION

The innate immune system uses a small number of multifunctional proteins to respond to diverse immune challenges.¹⁻⁴ Multifunctional immune proteins are critical for pathogen defense,²⁻⁴ shaping host-associated microbial communities,⁵ and well-regulated tissue growth.⁶⁻⁸ They also drive pathological inflammation in disease, including autoimmune disorders, cancer, and cardiovascular disease.⁹⁻¹³ These multifunctional proteins raise both mechanistic and evolutionary questions. How can one protein sequence satisfy the multiple constraints imposed by having multiple functions? How can multiple functions evolve in one protein when, as a result of multifunctionality, each mutation likely has pleiotropic effects?^{1,14-16}

One such multifunctional protein is S100A9 (A9), a small, soluble protein found at high concentrations in the extracellular space during an inflammatory response.¹⁷ It has at least two key immune functions. As a homodimer, A9 potently activates inflammation via Toll-like receptor 4 (TLR4).¹⁸⁻³⁰ As a heterocomplex with S100A8 (A8/A9, also known as calprotectin – figure 1a) it is antimicrobial.³¹⁻⁴⁴ A9 exacerbates endotoxin-induced shock in mice.⁴⁵ Both A9 and A8/A9 are primary biomarkers for many human inflammatory diseases.⁴⁶⁻⁵⁰ Further, dysregulation of A9 is associated with various cancers, pulmonary disorders, and Alzheimer's disease.^{19,46-52} Understanding the mechanisms by which A9 performs its innate immune functions is critical for developing treatments for A9-mediated diseases.

The mechanism of A8/A9 antimicrobial activity is well established: it sequesters transition metals through a unique hexahistidine metal binding site at the A8/A9 heterodimer interface, thereby limiting the concentrations of essential microbial nutrients in the extracellular space.³¹⁻⁴⁴ In contrast, the proinflammatory mechanism of A9 is not well understood. A9 acts as a Damage-Associated Molecular Pattern (DAMP), activating NF- κ B and other cytokines through Toll-like receptor 4 (TLR4).¹⁸⁻³⁰ The interaction interface(s), affinity, and stoichiometry for the A9/TLR4 interaction are not known. A small

region of A9 has been suggested to form part of the A9/TLR4 binding surface,⁵³ but no mutant of A9 has been identified that substantially compromises its activation of TLR4.

An additional layer of A9 immune function is that A9 and A8/A9 are thought to be regulated in the extracellular milieu by proteases. A9 is very susceptible to proteolytic degradation, while A8/A9 is highly resistant (figure 1a).^{54,55} Proteolysis may serve to purge proinflammatory A9 from the extracellular space and thus selectively enrich for antimicrobial A8/A9. There may even be a direct, functional link between inflammation and proteolysis, as proteolytic fragments of A9 have been suggested to activate TLR4.⁵³ Testing these ideas, however, has been challenging. It has been proposed that the origin of proteolytic resistance of A8/A9 is through the formation of a tetrameric dimer of heterodimers—(A8/A9)₂, which forms in the presence of calcium.⁵⁶ There is no obvious way, however, to selectively alter the proteolytic resistance of A9, making it difficult to test for correlation between A9 proteolytic resistance and A9 proinflammatory activity.

We took an evolutionary biochemical approach to mechanistically dissect the evolution of A9 innate immune functions. Using phylogenetics, ancestral sequence reconstruction (ASR), and biochemical studies, we show that A9s evolved to form proteolytically resistant, antimicrobial A8/A9 complexes in early mammals. We find that A9s gained proinflammatory activity and lost proteolytic resistance in the ancestor of therian mammals from a weakly proinflammatory, proteolytically resistant amniote ancestor. We identify a pleiotropic substitution that is necessary for A9 activation of TLR4, sufficient to increase TLR4 activation by the A9 amniote ancestor and played a role in loss of A9 proteolytic resistance. Mutating this site has minimal effect on A8/A9 antimicrobial activity or proteolytic resistance. Lastly, we show that proteolysis is not required for A9 activation of TLR4. Taken together, this work reveals that mammals concomitantly evolved A8/A9 antimicrobial activity, A9 proinflammatory activity, and a way to selectively regulate A9 inflammation via loss of A9 proteolytic resistance. These findings provide unprecedented mechanistic and evolutionary insight into A9 function

and show how a single mutation can have pleiotropic effects in one functional state of a protein while not impacting another, thus facilitating the evolution of multifunctionality.

RESULTS

We first set out to establish when A9 evolved three innate immune properties: antimicrobial activity via formation of the A8/A9 complex, proinflammatory activation of TLR4 by A9 alone, and the differential proteolytic susceptibility of A9 and A8/A9.

A9s evolved to form antimicrobial A8/A9 complexes early in mammals

We sought to determine when A9 evolved to form the antimicrobial A8/A9 complex. We hypothesized that A8/A9 antimicrobial activity evolved in the ancestor of therian mammals (the shared ancestor of marsupials and placental mammals) for several reasons. First, the broad-spectrum antimicrobial activity of human and mouse A8/A9 is well established.^{32–44} Second, A9 and A8 genes are only found together in therian mammals (figure 1b);⁵⁷ therefore the A8/A9 complex could not have arisen earlier than in the ancestor of therian mammals. Lastly, the residues composing the antimicrobial hexahistidine metal binding site are fully conserved across therian mammals (figure S1).

To determine whether the antimicrobial A8/A9 complex arose in the ancestor of therian mammals, we compared human A8/A9 to two previously uncharacterized A8/A9 complexes. We first tested the antimicrobial activity of A8/A9 from opossum, which is one of the earliest-diverging mammals relative to humans that possesses both of the S100A8 and S100A9 genes. Following previous work,⁴¹ we produced a cysteine-free variant of the complex to avoid the use of reducing agents in the antimicrobial assay. We confirmed that cysteine-free opossum A8/A9 formed a heterotetramer (46.8 ± 0.7 kDa) in the presence of calcium – like the human and mouse proteins⁵⁸ – using size exclusion chromatography coupled with multi-angle laser light scattering (SEC MALS, figure S2).

We measured cysteine-free opossum A8/A9 antimicrobial activity against *Staphylococcus epidermidis* using a previously established *in vitro* antimicrobial assay that monitors bacterial growth in the absence or presence of S100 proteins (figure 1c, S3).⁴¹ To compare the activity of different proteins, we quantified inhibition at seven hours (figure 1d). We observed a dose-dependent decrease in *S. epidermidis* growth in the presence of low micromolar concentrations of cysteine-free opossum A8/A9 (figure 1c-d, S3). The antimicrobial activity of opossum A8/A9 was weaker than that of human A8/A9: opossum A8/A9 delayed bacterial growth, while human A8/A9 both delayed growth and decreased bacterial carrying capacity (figure 1c). It was previously found that cysteine-free human A8/A9 was potentially antimicrobial,⁴¹ while cysteine-free mouse A8/A9 exhibits weaker antimicrobial activity than wildtype mouse A8/A9.³² To determine whether the weaker activity of opossum A8/A9 was due to the removal of cysteines, we also measured the activity of wildtype opossum A8/A9 against *S. epidermidis*. We found that wildtype opossum A8/A9 had higher activity than cysteine-free opossum A8/A9 over the concentration range tested (figure S3). This suggests that the cysteines present in mouse and opossum A8/A9 play a role in their antimicrobial activity, unlike in human A8/A9. The antimicrobial activity of the opossum A8/A9 complex thus appears to be more similar to that of mouse A8/A9 than human A8/A9.

The shared antimicrobial activity of human, mouse, and opossum A8/A9 strongly suggests that the antimicrobial A8/A9 complex evolved in the ancestor of therian mammals. To test this further, we measured the antimicrobial activity of ancestrally reconstructed therian mammalian A8/A9 (ancA8/A9 – figure 1b). We used our previously published phylogenetic tree⁵⁷ consisting of 172 S100 sequences to reconstruct therian mammalian ancestral A8 and A9 (ancA9 and ancA8 – figure S4), which were used to form the ancA8/A9 complex. These proteins had average posterior probabilities of 0.83 and 0.88, respectively (table S4). We confirmed that each protein was folded and had secondary structure content similar to that of human A8/A9 using far-UV circular dichroism (CD) spectroscopy (figure S5).

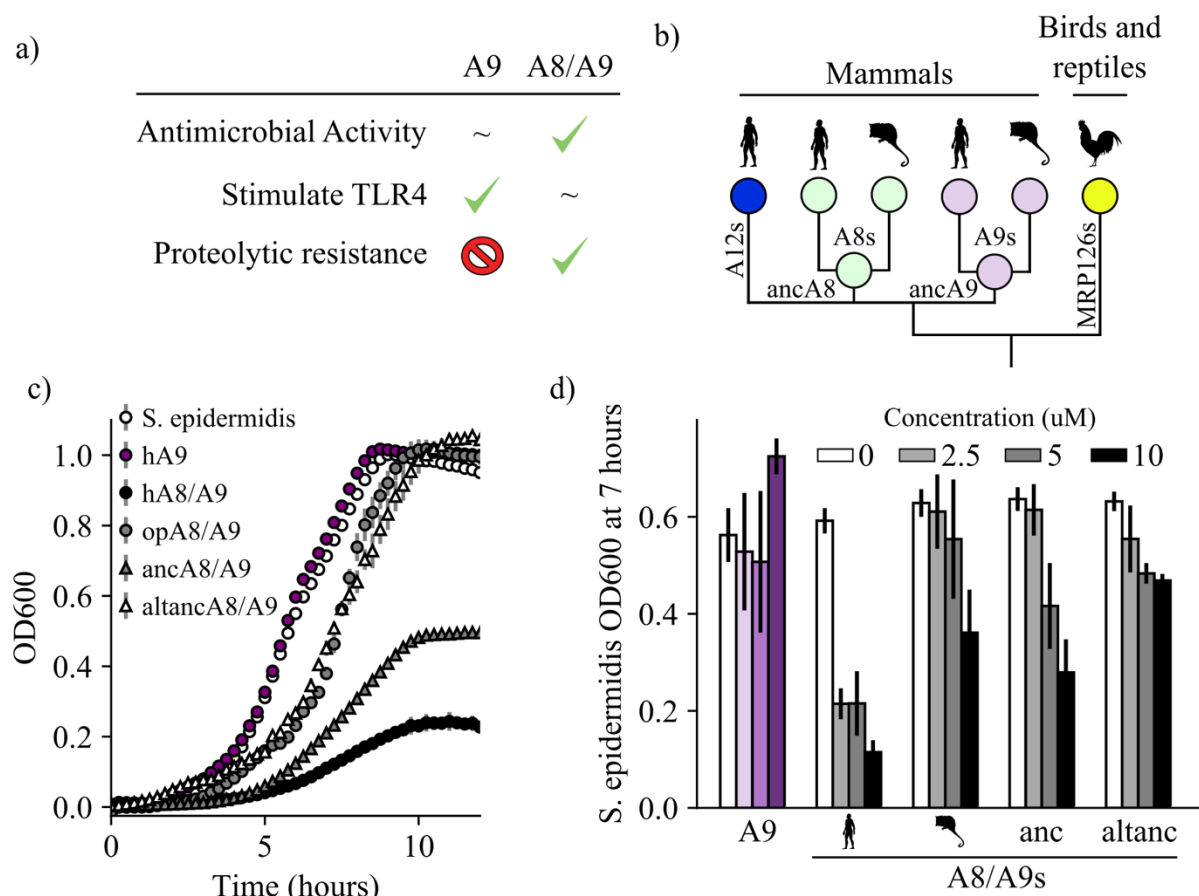


Figure 1. A9s evolved to form the antimicrobial A8/A9 complex early in mammals.

(a) Table of A9 and A8/A9 properties. “~” represents weak or ambiguously characterized function, check marks and red “X” represent confirmed property (check) or lack thereof (“x”). (b) Schematic of previously published S100 protein tree. Colored nodes represent single protein sequences. Species cartoons shown are human, opossum, and chicken. Ancestrally reconstructed protein nodes are labeled. Branch lengths not to scale. (c) Representative growth curves for *Staphylococcus epidermidis* in the presence or absence of 10 μ M S100 proteins. Each point represents optical density at 600 nm. *S. epidermidis* growth alone and in the presence of modern proteins are shown as circles, growth in the presence of ancestrally reconstructed proteins shown as triangles. Error bars are standard deviation of three technical replicates. (d) Percent of untreated *S. epidermidis* growth at 12 hours with S100 protein treatments. Data are average of three biological replicates. Error bars are standard error of the mean. Species cartoon labels are the same as in (b).

We then measured the antimicrobial activity of ancA8/A9 against *S. epidermidis*. We observed a potent reduction in *S. epidermidis* growth comparable to that of human A8/A9 (figure 1c-d, S3). To test for the robustness of this finding to phylogenetic uncertainty, we also tested the antimicrobial activity of an AltAll⁵⁹ reconstruction of ancA8/A9 against *S. epidermidis* (altancA8/A9, figure 1c-d, S3). In this reconstruction, we swapped all sites with ambiguous reconstructions for their second-most likely state

(see methods). AncA8/A9 and altancA8/A9 differ by 24 amino acids. AltancA8/A9 exhibited antimicrobial activity against *S. epidermidis* similar to opossum A8/A9: it delayed growth but did not ultimately limit bacterial carrying capacity. While the hexahistidine site residues are conserved in ancA8/A9 and altancA8/A9 (figure S1), it appears that a subset of the ambiguously reconstructed 24 residues are important for A8/A9 antimicrobial activity, perhaps affecting the orientation and/or affinity of the hexahistidine metal binding site.

Taken together, the antimicrobial activity of modern mammalian A8/A9 complexes (human, mouse, and opossum) and the antimicrobial activity of the reconstructed ancA8/A9 complex suggest that A9s evolved to form the antimicrobial A8/A9 complex in the ancestor of mammals.

A9s evolved potent proinflammatory activity from a weakly proinflammatory amniote ancestor

We next sought to determine when A9s evolved potent proinflammatory activity via activation of TLR4. Our previous work revealed that human A9 potently activates not only human TLR4 in functional assays, but also opossum and chicken TLR4 (figure 2a).⁵⁷ In contrast, chicken MRP126, the sauropsid ortholog of A9s, was found to be a weak activator of all TLR4s, including chicken TLR4. Both human and opossum A9 activate chicken TLR4 better than chicken MRP126 does. Two possibilities are consistent with these observations. Either mammalian A9s evolved enhanced proinflammatory activity from a less active amniote ancestral state, or A9s maintained a potent ancestral activity that was lost by chicken MRP126.

To differentiate between these two possibilities, we determined the ancestral proinflammatory activity of these proteins. We used ASR to reconstruct the shared amniote ancestor of A9s, A8s, A12s, and MRP126s. This group of proteins is known collectively as “calgranulins”, so we will refer to this ancestral protein as ancCG (ancestor of calgranulins). We also constructed an alternate, “alt All” version of this ancestor (altancCG, S4). The average posterior probability of ancCG was 0.86 (table S4). We also expressed and purified ancA9 and altancA9 – the A9 subunits from the ancestral A8/A9 complexes

described above. We confirmed that each protein was folded and had secondary structure content similar to that of modern S100s using far-UV CD spectroscopy (figure S5).

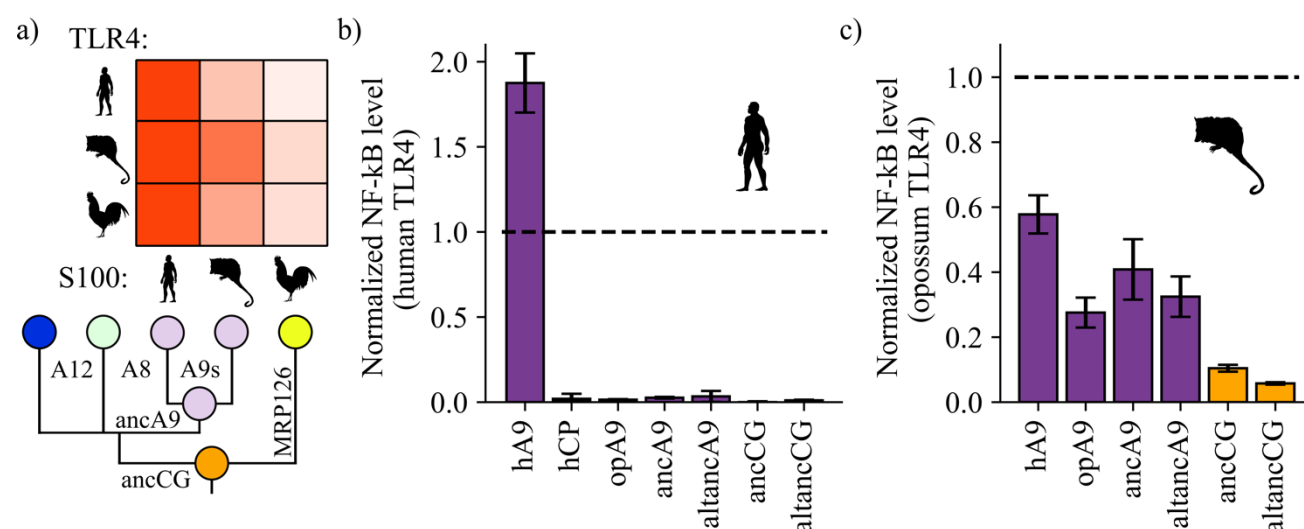


Figure 2. A9s gained proinflammatory activity from a weakly proinflammatory ancestor.

(a) Schematic of previously measured proinflammatory activity of S100s against various TLR4s. Species labels on x and y-axes of heatmap are the same as figure 1. Heatmap coloring is scaled to match 2 μM S100 activity levels measured in supplementary figure S2 of Loes et al. 2018.⁵⁷ (b) and (c) NF-κB production of human and opossum TLR4 in response to treatment with modern and ancestral S100 proteins. Bars represent average of >3 biological replicates, error bars are standard error of the mean. All values are background-subtracted and normalized to LPS positive control (see methods).

We then tested modern and ancestral S100s for activity against human TLR4. Following previous work,^{45,57,60,61} we transiently transfected HEK293T cells with plasmids encoding TLR4 and its species-matched cofactors MD-2 and CD14, added purified S100 proteins to the growth media, and then measured output of luciferase under control of an NF-κB promoter. Consistent with previous results,⁵⁷ we found that human A9 potently activated human TLR4, resulting in high levels of NF-κB production (figure 2b, S6-7). Human A8/A9 and opossum A9 exhibited much weaker activity against human TLR4. Lastly, we tested ancA9, altancA9, ancCG, and altancCG for activity against human TLR4 and observed weak or no activation for each ancestral protein. This result is unsurprising, as we previously found that human TLR4 is more specific than other amniote TLR4s: human TLR4 is activated much more potently by human A9 than by any other S100 protein (figure 2a-b).⁵⁷ In contrast, TLR4s from other species

(mouse, opossum, and chicken) appear to be more promiscuous and can be activated similarly by S100s from various species (figure 2a).⁵⁷ This is consistent with lineage-specific coevolution between human TLR4 and human A9 – a confounding variable that makes assessment of ancestral S100 protein proinflammatory activity difficult using human TLR4.

We predicted that opossum TLR4 would be a better protein to probe ancestral S100 proinflammatory function because opossum TLR4 is broadly activated by A9s across mammals and gives little indication of lineage-specific coevolution.⁵⁷ We therefore tested the proinflammatory activity of ancA9, ancCG, and their corresponding alternate reconstructions against opossum TLR4. Corroborating previous results, human A9 strongly activated opossum TLR4, while opossum A9 activity was approximately half that of human A9 (figure 2c, S6-8). AncA9 and altancA9 activated opossum TLR4 to the same extent as opossum A9. AncCG and altanCG were the weakest activators of opossum TLR4, with activity approximately 25% or less than that of human A9.

These findings suggest that A9s evolved enhanced proinflammatory activity early in mammals from a weakly proinflammatory amniote ancestor, while A8/A9s and chicken MRP126 maintained weak, ancestral proinflammatory activity.

A9s evolved proteolytic susceptibility from a proteolytically resistant amniote ancestor

We next sought to determine when the differential proteolytic susceptibility of A9 and A8/A9 evolved. We used a simple *in vitro* assay to monitor S100 protein degradation over time in the presence of proteinase K, a potent non-specific serine protease (figure 3a). Proteinase K was chosen both because of its low specificity and to mimic other serine proteases that A9 and A8/A9 encounter when released from neutrophils during an inflammatory response.^{62–66} Proteolytic decay rates were estimated by fitting a single exponential decay function to the data (figure 3b, S9-12).

Human A8/A9 has been described as extremely resistant to proteases.⁵⁴ however, it has not been compared to S100 proteins besides human S100A8 (A8) and A9. To establish a baseline expectation for S100 protein proteolytic resistance, we characterized the proteolytic resistance of a broad set of human S100s against proteinase K. As previously shown,⁵⁴ human A9 and A8 alone were rapidly proteolytically degraded while the human A8/A9 complex exhibited strong resistance (figure 3c, S9-10). Under our conditions, the degradation rates for human A8 and A9 were approximately three orders of magnitude faster than that of the human A8/A9 heterocomplex. We then characterized closely related protein human S100A12 (A12), the chicken ortholog MRP126, and six distantly related human S100s.⁶⁷ Human A12, chicken MRP126, and five out of six more distantly related human S100s exhibited intermediate to strong proteolytic resistance, each degrading 1-2 orders of magnitude slower than human A8 or A9 but, on average, one order of magnitude faster than human A8/A9 (figure 3c, S9). Notably, human A12 and chicken MRP126 formed predominantly homodimers by SEC MALS under these conditions (figure S2), indicating that higher-order oligomerization (> 2 subunits) isn't required for S100 proteolytic resistance. Lastly, human A14 degraded faster than A9 or A8. This protein is evolutionarily distant⁶⁷ and therefore likely reflects independent evolution of this property. Taken together, these data show that the A8/A9 complex, A9, and A8 indeed fall at the extremes of human S100 proteolytic resistance; human A9 and A8 are among the fastest-degrading S100s tested, while human A8/A9 is one of the slowest.

To test whether A9 and A8 proteolytic susceptibility and A8/A9 resistance are conserved across mammals, we characterized mouse and opossum A9, A8, and A8/A9 for proteolytic resistance. Mouse A9 and A8 were found to be highly proteolytically susceptible and mouse A8/A9 strongly proteolytically resistant, matching the pattern observed for their human counterparts (figure 3c and S10). Opossum A9 and A8 were also highly proteolytically susceptible, while opossum A8/A9 was resistant (figure 3c, S10). This indicates that the susceptibility of A9s and A8s and the resistance of A8/A9 complexes is conserved across mammals.

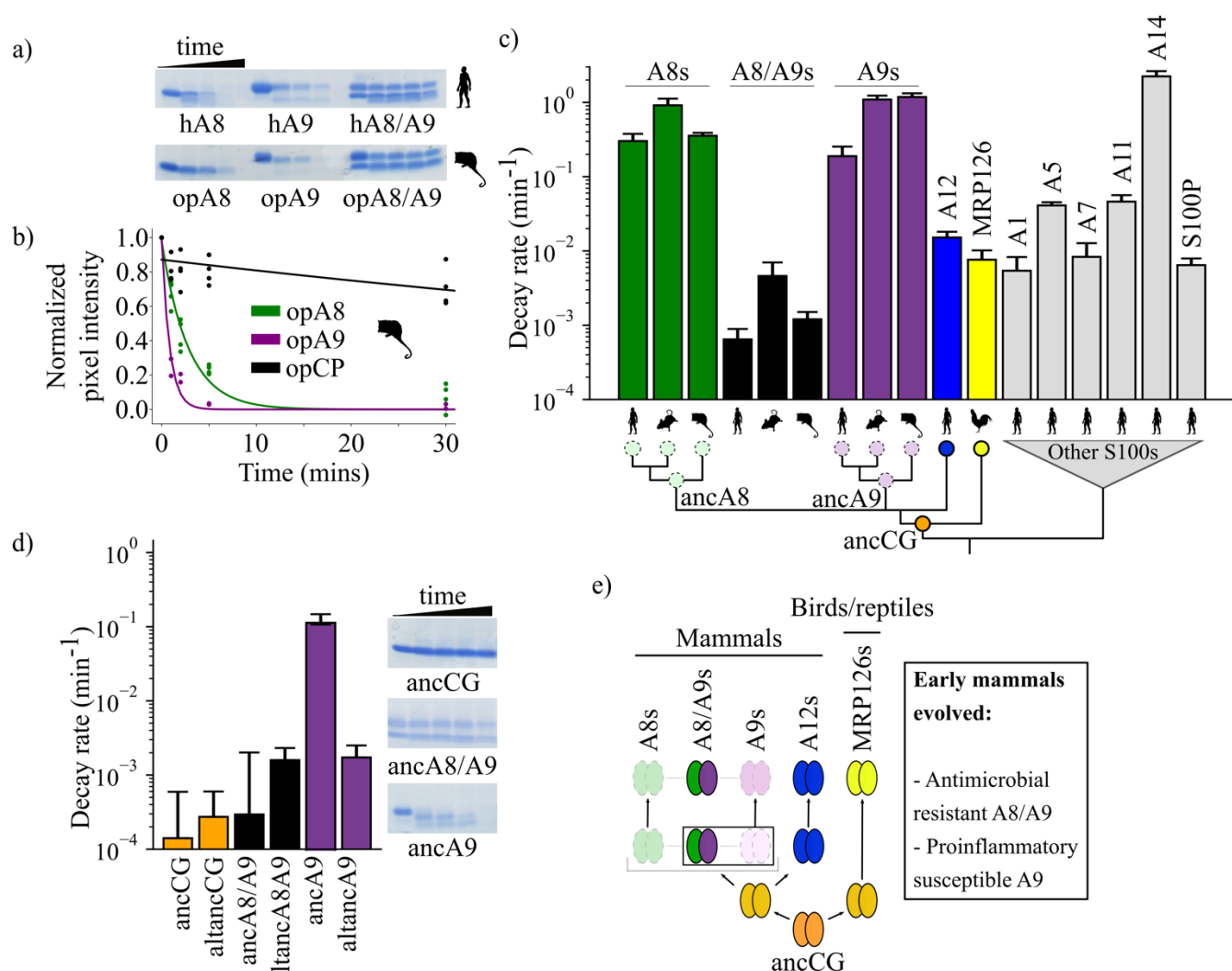


Figure 3. A9s lost proteolytic resistance from a proteolytically resistant amniote ancestor.

(a) *In vitro* proteolytic resistance assay showing SDS-PAGE gel of S100 protein degradation via proteinase K over time. Gels were quantified using densitometry and normalized to the undigested protein band intensity. (b) A single exponential decay model was globally fit to the data to quantify decay rates. Points are biological replicates, lines are model fit to data. (c) S100 protein proteolysis rates mapped onto schematized S100 phylogeny. X-axis cartoon labels same as in figure 1. Circles indicate proteolytic susceptibility (faded/dashed) and resistance (solid), with predicted resistance shown for ancA8, ancA9, and ancCG nodes. (d) Decay rates for ancestrally reconstructed proteins, with gels shown on the right. For panels (c) and (d), error bars are the square root of the diagonalized covariance matrix from the fit and the y-axis is in log scale. (e) Summary model for proposed evolution of A9 and A8/A9 innate immune properties. Box around A8/A9 and A9 indicate location in tree (ancestor of therian mammals) where immune functions evolved.

When mapped onto the S100 phylogeny, the most parsimonious explanation for these data is that the shared amniote ancestor—ancCG—was proteolytically resistant (figure 3c). In this scenario, A12s, MRP126s, and A8/A9s conserved ancestral resistance, while A9s and A8s independently lost resistance

early in mammals. Alternatively, ancCG could have been proteolytically susceptible. This would mean that A9s and A8s maintained an ancestral susceptibility, while MRP126s, A12s, and A8/A9s each evolved novel proteolytic resistance.

To distinguish between these possibilities, we characterized ancestrally reconstructed S100s for proteolytic resistance. AncCG and altancCG exhibited extremely high proteolytic resistance (figure 3d, S11), with degradation rates 3-4 orders of magnitude slower than modern A8s or A9s and approximately one order of magnitude slower than modern A8/A9s. AncA8/A9 and altancA8/A9 also demonstrated high proteolytic resistance, with degradation rates approximately 2-3 orders of magnitude slower than A8s and A9s and comparable to modern A8/A9 complexes (figure 3d, 2f). Together, these data paint a consistent picture: the amniote ancestor of A9s, ancCG, was strongly resistant to proteolytic degradation. Modern A9s and A8s lost proteolytic resistance from an ancestrally resistant state, while modern A12s, A8/A9 complexes, and MRP126s maintained the ancestral proteolytic resistance (figure 3e).

Finally, we sought to better resolve when A9s acquired proteolytic susceptibility. We hypothesized that this occurred in the ancestor of mammalian A9s before the divergence of therian mammals and marsupials. To test this hypothesis, we measured the proteolytic susceptibility of therian mammalian ancA9 and found that it degraded rapidly. However, its alternative reconstruction (altancA9), was slow to degrade, with a rate two orders of magnitude slower than ancA9 and comparable to other highly resistant S100s. Because the descendants of ancA9 all exhibit proteolytic susceptibility, the simplest explanation is that altancA9 is a low-quality reconstruction that does not capture the properties of the historical protein. Alternatively, proteolytic susceptibility could have been independently acquired along marsupial and placental mammal lineages.

A single substitution had pleiotropic effects on TLR4 activity and proteolytic resistance

280 We found above that A9 evolved to form the antimicrobial A8/A9 complex, gained potent
281 proinflammatory activity, and lost proteolytic resistance over the narrow evolutionary interval after the
282 divergence of mammals and sauropsids but before the divergence of placental mammals and marsupials.
283 We next sought to determine how A9 evolved its antimicrobial and proinflammatory activities and lost
284 proteolytic resistance.

285 The mechanism by which A9 evolved to form the antimicrobial A8/A9 complex appears
286 straightforward. After the duplication that gave rise to A8 and A9, mutations accumulated in both
287 molecules that created the antimicrobial hexahistidine metal binding site in the A8/A9 complex (figure
288 3d). Mammalian A9s also have a conserved C-terminal extension that contributes two histidines to the
289 hexahistidine metal binding site. The quantitative difference between ancA8/A9 and altancA8/A9
290 antimicrobial activity suggests that other amino acid changes tuned the antimicrobial activity of the
291 molecule, but the core functionality is determined by whether the six histidine residues were present.

292 The mechanisms by which A9s gained proinflammatory activity and lost proteolytic resistance
293 are less obvious, particularly because the mechanism by which A9 activates TLR4 is not well understood.
294 We reasoned that we could identify functionally important amino acid substitutions by focusing on the
295 evolutionary interval over which these properties evolved. We therefore compared the sequences of
296 ancCG (weakly proinflammatory and resistant to proteolytic degradation) and ancA9 (potently
297 proinflammatory and susceptible to proteolytic degradation). We further narrowed down sequence
298 changes of interest by looking for residues conserved in modern A9s (figure S13). Finally, we focused
299 on amino acid changes in helix III of A9, as this region is thought to be important for A9 activation of
300 TLR4 based on *in vitro* binding studies and *in silico* docking studies.⁵³ Only one historical amino acid
301 substitution met all three criteria: position 63 (human A9 numbering). This residue is a phenylalanine in
302 both ancCG and altancCG, is conserved as a phenylalanine in 95% of modern A8s and A12s and has
303 been substituted for a methionine or leucine (M/L) in 97% of A9s (figure 4a).

304 We hypothesized that reverting this site to its amniote ancestral state—M63F—might affect A9
305 proinflammatory activity. We mutated this position to a phenylalanine in human A9 and opossum A9 and
306 tested each protein for TLR4 activation. Strikingly, we found that introducing M63F into human A9
307 severely compromised its ability to activate human TLR4 (figure 4b, S7). This was also true for opossum
308 A9: introduction of M63F strongly decreased opossum A9 activation of opossum TLR4 (figure 4c, S8).
309 We next introduced the forward substitution, F63M, into ancCG and tested its proinflammatory activity
310 against opossum TLR4. We observed a modest increase in ancCG activity with the F63M substitution,
311 with activity comparable to that of opossum A9 (figure 4c, S8).

312 For most proteins we studied, the amino acid at position 63 did indeed play an important role in
313 determining the pro-inflammatory activity of A9. The effects of toggling position 63 between Met and
314 Phe were not, however, universal. We introduced M63F into ancA9 and observed no change in
315 proinflammatory activity (figure 4c, S8). Further, altancA9 has a Phe at position 63 but activates TLR4
316 in the assay (figure 2c, S8). Thus, while position 63 is an important contributor to activity in modern A9s,
317 other substitutions were also important for the transition from a weakly pro-inflammatory ancestor to the
318 modern set of potentially pro-inflammatory A9s.

319 Because A9s lost proteolytic resistance and gained proinflammatory activity over the same
320 evolutionary time interval, we reasoned that the F63M substitution might have also played a role in A9
321 loss of proteolytic resistance. To test this, we characterized the proteolytic resistance of human A9 M63F
322 and ancA9 M63F. Strikingly, reversion of this single mutation rendered both ancA9 and human A9
323 strongly resistant to proteolytic degradation, decreasing their respective degradation rates by 1-2 orders
324 of magnitude and approaching the degradation rates of ancCG and various A8/A9 complexes (figure 4d,
325 S11). We also tested the effect of the forward mutation – F63M – on ancCG proteolytic resistance. We
326 observed no change in resistance for ancCG F63M, indicating that additional substitutions were required
327 to render ancA9 proteolytically susceptible. Together these data show that a single historical reversion is

sufficient to render A9s proteolytically resistant, indicating that this position played a role in the loss of A9 proteolytic resistance early in therian mammals.

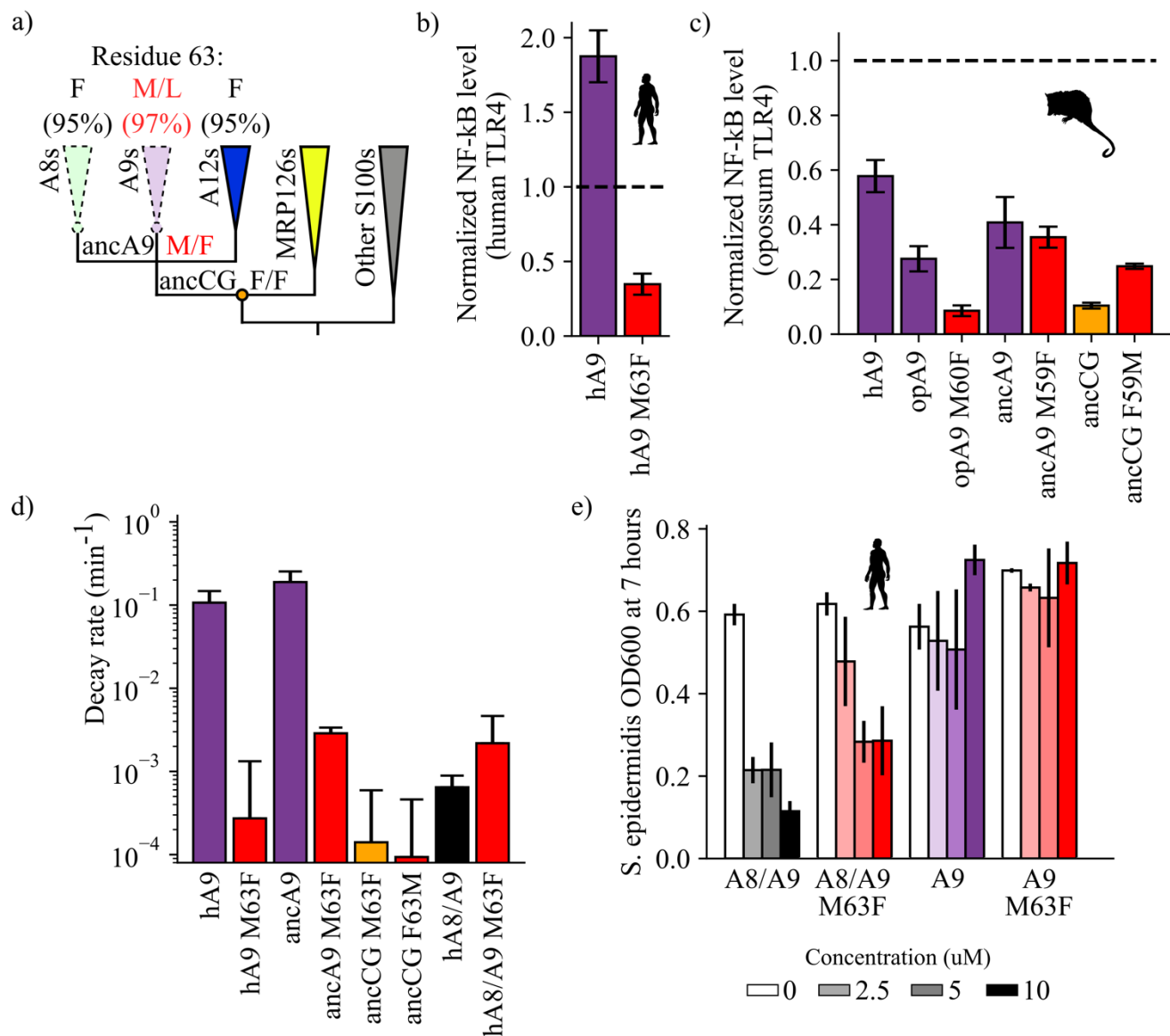


Figure 4. A single historical substitution affects A9 proinflammatory activity and proteolytic resistance without affecting A8/A9 proteolytic resistance or antimicrobial activity.

(a) Schematic S100 phylogenetic tree with the amino acid state of position 63 shown at key nodes. Wedges represent clades, colored as in figure 1. Lines indicate proteolytic susceptibility (faded/dashed) and resistance (solid). Circles indicate characterized ancestors. Amino acid labels represent maximum likelihood state/alternate amino acid state for position 63 at ancestral nodes, while labels at clade tips represent percent conservation across modern S100 protein sequences. (b-c) NF-κB production of S100 point mutants at position 63 against human (b) and opossum (c) TLR4. (d) Proteolysis rates for S100 point mutants at position 63. Error bars and y-axis are the same as in figure 1. (e) Antimicrobial activity of hA9 and hA8/A9 with and without M63F mutation against *S. epidermidis*. Axes and error bars same as in figure 1d.

The pleiotropic substitution does not affect antimicrobial activity

We then investigated whether introducing M63F affects the properties of the antimicrobial A8/A9 complex. As A9 is a critical subunit of A8/A9, we reasoned that perhaps position 63 was altered in A9s due to some incompatibility between Phe63 and A8/A9 formation or function. To this end, we introduced M63F into human A8/A9 and tested it for oligomeric state, proteolytic resistance, and antimicrobial activity against *S. epidermidis*. Human A8/A9 M63F predominantly formed a heterotetramer in the presence of calcium by SEC MALS with a molecular weight similar to that of wildtype human A8/A9 (48.7 ± 4.2 kDa – figure S2). We found that human A8/A9 M63F was also strongly resistant to proteolytic degradation, similar to human A8/A9 (figure 4d). Lastly, M63F had minimal impact on human A8/A9 antimicrobial activity against *S. epidermidis*, retaining potent antimicrobial activity (figure 4e). In contrast, neither human A9 nor human A9 M63F were antimicrobial against *S. epidermidis* (figure 4e). These findings suggest that this single amino acid position had important effects on the evolution of A9 activation of TLR4 and loss of proteolytic resistance without significantly impacting A8/A9 oligomeric state, proteolytic resistance, or antimicrobial activity.

M63F increases protein thermodynamic stability and decreases unfolding rate of human A9

We next asked what effect M63F has on the biophysical properties of human A9. Residue 63 sits in the middle of helix III of A9, pointing inward toward helix II, and is neither a core residue nor fully surface-exposed (figure 5a).⁶⁸ Based on the published structure of human A9,⁶⁸ a Phe at position 63 could be plausibly tolerated without a steric clash. Using circular dichroism (CD) spectroscopy, we found that the bulk secondary structure content of human A9 M63F was similar to that of hA9 (figure 5b). We measured the oligomeric state of human A9 M63F by SEC MALS and found that it predominantly forms a homodimer in solution similarly to human A9, with no detectable monomers or larger oligomers (figure 5c). These data together indicate that M63F doesn't significantly alter human A9's secondary structure or oligomeric state.

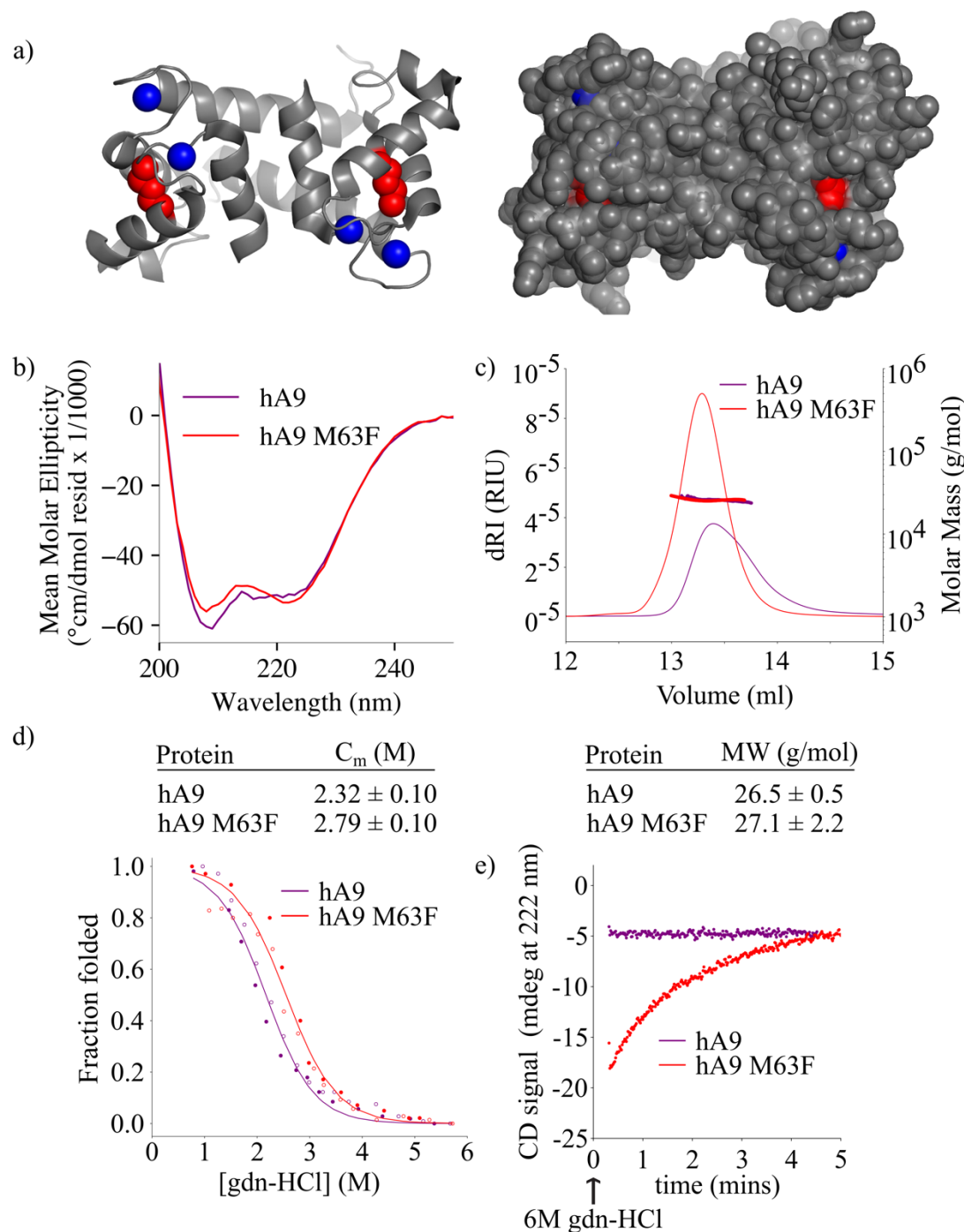


Figure 5. M63F increases human A9 apparent stability by decreasing its unfolding rate.

(a) Crystal structure of hA9 (PDB entry 1irj).⁶⁸ Cartoon depiction left, surface view right. Calcium ions are blue spheres. M63 is highlighted in red – two total for homodimeric A9. (b) Far-UV circular dichroism (CD) spectroscopy scans of hA9 and hA9 M63F. Data represent average of 3 scans. (c) SEC MALS analysis of hA9 and hA9 M63F oligomeric state. Solid lines are refractive index (left y-axis), points and molecular weights in table below represent molar mass calculated from light scattering detectors using ASTRA software (right y-axis - see methods). (d) Equilibrium chemical denaturation of 0.5 uM hA9 and hA9 M63F monitored by CD at 222 nm. Filled circles are protein unfolding, empty circles are refolding. A two-state denaturation model was fit to the data to extract the C_m – errors in table below represent standard error of the mean of three biological replicates ($p = 0.02$; paired t-test) (figure S14). (e) Kinetics of hA9 and hA9 M63F unfolding via chemical denaturation. Graph depicts one representative unfolding experiment (figure S15).

We then examined whether M63F alters the stability of human A9. We measured unfolding and refolding curves for human A9 and human A9 M63F using CD spectroscopy and chemical denaturation via guanidine hydrochloride (gdn-HCl). We found that M63F appears to stabilize human A9, shifting the C_m by approximately 0.5 M (figure 5d, S14). We also measured the unfolding kinetics of human A9 and human A9 M63F in the presence of calcium by spiking protein directly into 6M gdn-HCl and monitoring its unfolding rate by CD spectroscopy. Strikingly, human A9 M63F takes several minutes to unfold under these conditions, while human A9 unfolds immediately within the dead time of the experiment (figure 5e, S15). We note that the folding pathway for A9 is complex and almost certainly not two-state—calcium binding, monomer folding, and dimerization all contribute—and thus we cannot reliably determine how M63F affects the stability of each of these potential folding intermediates. The large increase in the apparent unfolding rate reveals, however, that the mutation stabilizes some aspect of the folded structure rather than simply increasing the rate of folding, oligomerization and/or calcium binding.

Proteolysis is not required for A9 activation of TLR4

The work above identified a mutation that, when introduced into human A9, increases the stability of the protein structure while also potentially compromising the ability to activate TLR4. The mutation is not at a surface position and is therefore not likely a direct participant in the A9/TLR4 protein/protein interface. Further, the same mutation dramatically decreases the proteolytic susceptibility of the protein. One simple way to explain these observations would be if the proteolytic susceptibility itself was the feature that evolved to allow activation of TLR4. This would be consistent with a previous observation that proteolytic products of A9 activate TLR4.⁵³

To test whether proteolysis itself was sufficient for activity, we engineered an alternate variant of A9 that was proteolytically susceptible. We introduced the M63A mutation into human A9, anticipating that the short alanine sidechain would not have the stabilizing effect of M63F. As expected, human A9 M63A was highly susceptible to proteolytic degradation, similar to wildtype human A9 (figure 6a, S12).

400 We reasoned that if proteolysis is the primary determinant of A9 activation of TLR4, then proteolytically
 401 susceptible human A9 M63A should potently activate TLR4. Human A9 M63A, however, exhibited
 402 diminished proinflammatory activity, similar to human A9 M63F (figure 6b, S7). This indicates that the
 403 methionine at position 63 is important for A9 activation of TLR4. Further, we quantified the amount of
 404 human A9, human A9 M63F, and human A9 M63A before and after measuring TLR4 activity and
 405 observed no decrease in the amount of full-length protein remaining for wildtype human A9 or either
 406 mutant by western blot (figure 6c). This indicates that A9 is not digested by extracellular proteases over
 407 the course of the *ex vivo* assay and that proteolysis is likely not necessary for A9 activation of TLR4.

408 Although proteolysis does not appear to be a requirement for TLR4 activation, this does not rule
 409 out that proteolysis could increase A9 proinflammatory activity by releasing proinflammatory fragments
 410 of A9. To test for this possibility, we treated human A9 with agarose-immobilized proteinase K for
 411 increasing amounts of time, removed the protease, and then measured the proinflammatory activity of
 412 A9 degradation products (figure 6d). If proteolytic products of A9 are the most proinflammatory form of
 413 the protein, we might expect to observe a spike in TLR4 activation upon A9 digestion. Instead, we
 414 observed a steady decrease in human A9 activity with increasing digestion time. This suggests that full-
 415 length human A9 is the most potent activator of TLR4.

416 We did observe moderate activity for proteolytic products of human A9, as previously shown.⁵³
 417 After 30 minutes of digestion, no detectable full-length A9 remains by western blot (< 30 ng, figure 6d),
 418 but NF-κB production is still quite high, revealing that smaller fragments of A9 are sufficient to provide
 419 some degree of activation of TLR4. This raised the possibility that part of M63F's deleterious effect on
 420 proinflammatory activity could be to limit the release of active proteolytic fragments of A9. To test this,
 421 we also measured human A9 M63F activation of TLR4 after digestion for multiple hours (figure 6d).
 422 Unlike wildtype, however, fragments of human A9 M63F did not activate TLR4—even after being
 423 liberated by the protease. This strongly suggests that the historical mutation induced some change in the

native structure or dynamics of the molecule to bring about increased activity, independent of its effect on proteolytic susceptibility.

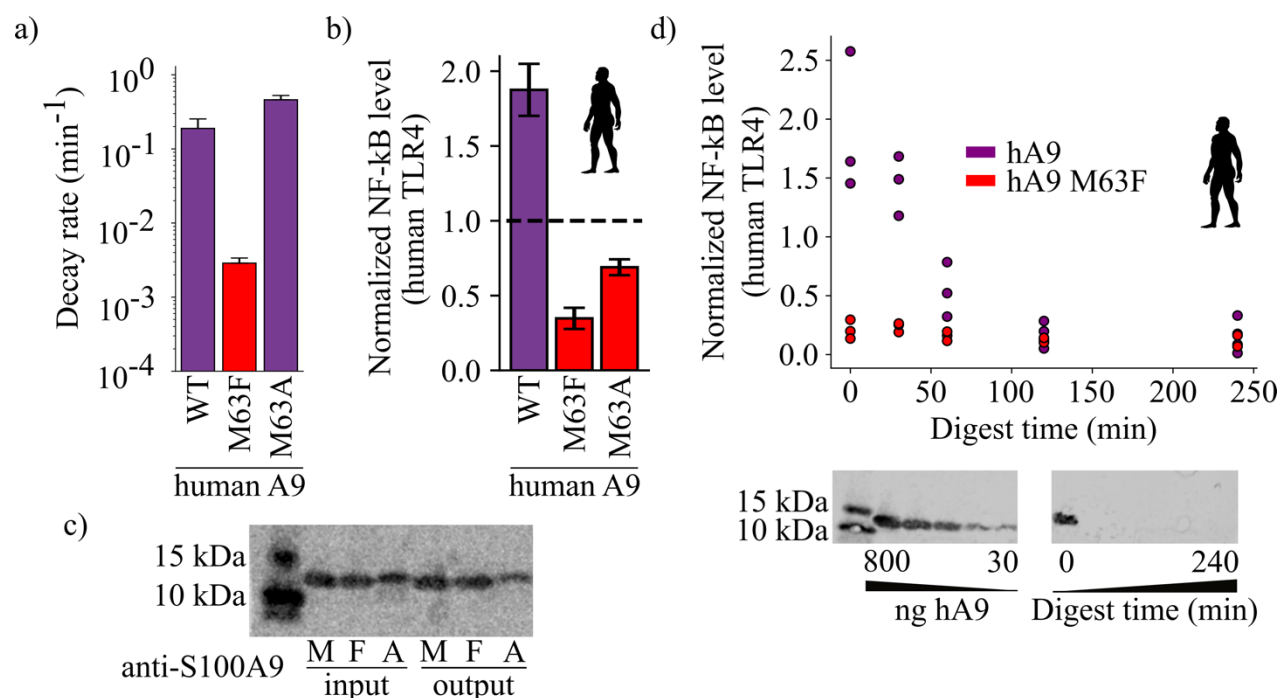


Figure 6. Proteolysis is not required for A9 activation of TLR4.

(a) Proteolytic decay rates for human point mutants at position 63. Error bars and axes are the same as in figure 3. (b) NF-κB production of human TLR4 in response to treatment with hA9, hA9 M63F, and hA9 M63A. Error bars the same as in figure 2. (c) Western blot of hA9 and position 63 point mutants before and after proinflammatory activity assay. Left bands represent 10 and 15 kDa ladder. (d) NF-κB production of human TLR4 in response to hA9 and hA9 M63F pre-proteolyzed with proteinase K for increasing amounts of time. Points are biological replicates and are the average of three technical replicates. Western blots below depict the amount of full-length A9 remaining over time. Left blot shows antibody sensitivity to A9, right shows digestion time course samples. Ladder and antibody same as in (c).

DISCUSSION

The work presented here provides insight into how the multifunctional protein A9 evolved critical innate immune functions. We find that mammalian A9s gained enhanced proinflammatory activity and lost proteolytic resistance from a resistant, weakly proinflammatory amniote ancestor. A single substitution played a key role in the evolution of these properties without significantly affecting the antimicrobial activity of the A8/A9 heterocomplex. This work contributes to our mechanistic

understanding of how A9 activates TLR4 to drive inflammation and clarifies the role of proteolysis in A9 innate immune function.

An evolutionary lens provides unique insight into the evolution of A9 innate immune functions

How did A9s evolve multifunctionality over a short evolutionary window? One might expect that the protein was under pleiotropic constraint to optimize each function.^{1,15,16,69–71} However, we show that one amino acid position is pleiotropic for one functional state of A9 (the A9 homodimer) but not the other (the A8/A9 heterocomplex). Reverting M63F in human A9 has a pleiotropic effect on its proinflammatory activity and proteolytic resistance but has little effect on human A8/A9 antimicrobial activity or proteolytic resistance. This suggests that mammals benefitted from this pleiotropic substitution early in A9s by acquiring a proinflammatory molecule that can be regulated via proteolytic degradation without altering the critical antimicrobial function of the A8/A9 complex. This reveals that the same mutation can have pleiotropic effects on one functional state of a protein while leaving another untouched, thus providing an advantageous route for evolving multifunctionality.

This study highlights the utility of taking an evolutionary approach to study protein function.⁷² Previous work examining the proteolytic resistance of S100s focused on the extreme proteolytic resistance of A8/A9, which was first observed over ten years ago.⁵⁴ Here we show that many S100s are, in fact, proteolytically resistant (figure 3c). By examining the proteolytic resistance of S100s through an evolutionary lens, we find that it is the loss of A9 resistance—not a gain of resistance in A8/A9—that was the important evolutionary change that occurred in these proteins. This approach facilitated identification of a single site (position 63)—out of more than 100—that strongly correlates with both A9 proteolytic susceptibility and A9 activation of TLR4. Isolating this position would have been very challenging without an evolutionary approach, as it does not stand out in a simple multiple sequence alignment or as part of the protein structure. We note that while reverting this site to its ancestral state is sufficient to render A9s resistant, further mutations were required between the amniote (ancCG) and

466 mammalian A9 (ancA9) ancestors to render A9s susceptible, as introduction of F63M into ancCG did
467 not decrease this protein's proteolytic resistance (figure 4d).

468 Why did A9s lose proteolytic resistance? We suggest three possibilities. The first is that loss of
469 proteolytic resistance in A9s was simply a byproduct of evolving proinflammatory activity. No A9
470 characterized in this study, with the exception of the alternate reconstruction of ancA9, is both
471 proteolytically resistant and potentially proinflammatory. This indicates that the molecular requirements for
472 A9 proteolytic resistance may be incompatible with those required for A9 activation of TLR4, and A9s
473 may have gained proinflammatory activity at the expense of proteolytic resistance. A second possibility
474 is that A9 proteolytic susceptibility is being maintained to actively remove proinflammatory A9 from the
475 cell and retain the antimicrobial A8/A9 complex. The last possibility for A9 loss of proteolytic resistance
476 is adaptive constraint. There could be selection for some property of A9 or A8/A9 that we did not measure
477 that is incompatible with A9 proteolytic resistance.

478 While we cannot explicitly distinguish between each of these possibilities, the end result is that
479 A9s evolved potent proinflammatory activity while simultaneously evolving a proteolytic “timer” to
480 regulate it, all without affecting A8/A9 function.

481 **Innate immune functions of A9 continued to evolve within the mammals**

482 Our data suggest that the proinflammatory and antimicrobial activities of A9 and the A8/A9
483 complex have undergone further optimization in placental mammals since these functions evolved. While
484 the histidines composing the high-affinity metal binding site of A8/A9 complexes are conserved, we
485 observed differences in antimicrobial potency for different A8/A9 complexes. In particular, human
486 A8/A9 is one of the most potentially antimicrobial A8/A9 complexes characterized. This suggests that
487 further optimization of the metal binding site has occurred in along the human lineage within mammals.
488 We also observed differences in activation of TLR4 by different A9s—human A9 is a potent,

promiscuous activator of TLR4s from multiple species, while earlier-diverging A9s and other S100s exhibit weaker proinflammatory activity.⁵⁷ Future studies are necessary to understand how, mechanistically, later-diverging A9s and A8/A9 complexes have optimized these critical innate immune functions.

Novel mechanistic insight into A9 activation of TLR4

Our findings suggest new directions for understanding how A9 potently activates TLR4. TLR4-driven inflammation has been the focus of intense study for over 20 years,^{19,23–25,27,60,73–76} and the structural basis of TLR4 activation by exogenous agonists, such as bacterial cell wall component lipopolysaccharide (LPS), is well understood.⁷⁷ In contrast, very little is known about how A9 activates TLR4. We have shown here that proteolytic degradation appears dispensable for activation; however, smaller fragments of the protein are sufficient activate TLR4 (figure 6). Given the effect of mutating position 63 on A9 proinflammatory activity, we propose that the region surrounding it—helix III—is important for activity. This is independently supported by Vogl et al., who identified four pairs of double mutants within helix III (amino acids 64, 65, 73, and 77) that, when mutated to alanines in pairs, decrease A9 binding to TLR4 *in vitro*.⁵³ Biophysical characterization of hA9 M63F (figure 5) indicates that it is more stable and unfolds more slowly, yet it maintains its bulk secondary structure and oligomeric state. The simplest explanation for these data is that M63F is affecting some functionally important dynamic process of the protein, possibly mediated by helix III, that is critical for A9 activation of TLR4. The proteolytic susceptibility of A9s also supports this hypothesis, as proteolysis is a dynamic process that often relies on substrate flexibility and local unfolding events to proceed.^{78–82} Damage-Associated Molecular Patterns (DAMPs) often interact with their targets via hydrophobic surfaces;^{83–85} one possibility is that A9 undergoes a local unfolding event that exposes a hydrophobic surface to interact with TLR4. This would mean that studies of the native structure of A9 might not be sufficient to gain

mechanistic understanding of how it activates TLR4. Further work is required to understand the nature of the active functional state of A9.

Pleiotropic mutations can facilitate the evolution of multifunctionality

Finally, our results suggest a positive role for pleiotropy in the evolution of protein function. Pleiotropy is often viewed as a constraint on evolution: as functional complexity is added to a polypeptide sequence, it becomes increasingly challenging to introduce substitutions—and new functions—without perturbing existing ones.^{1,15,16,69,70,86} Here, however, we find a single mutation that had beneficial pleiotropic effects on two important properties of A9: proinflammatory activity and proteolytic susceptibility. If A9 evolved potent proinflammatory activity without gaining susceptibility, it could potentially overstimulate inflammation simply by lingering in the extracellular milieu. Since both properties evolved at once, however, mammals evolved a proinflammatory molecule with a built-in “timer”: they gained a new inflammatory signal while avoiding potentially deleterious effects. This shows how pleiotropy can positively contribute to the evolution of new functions.

This same mutation, in contrast, had little pleiotropic effect on another functional state of A9: the A8/A9 complex. The antimicrobial activity of the A8/A9 complex was insulated from any pleiotropic effects from the mutation because proinflammatory and antimicrobial activities were partitioned between A9 and the A8/A9 complex, respectively. A mutation arose in the A9 amino acid sequence and is thus present in both A9 and A8/A9 states, but we only observe effects on the A9 state. This shows that pleiotropic constraint can be reduced when protein functions are partitioned amongst different protein states.

These findings reveal the diversity of pleiotropic roles that a single mutation can play. It further shows how the deleterious pleiotropic effects of mutations can be reduced by partitioning protein functions and properties into different functional states, thus enabling the acquisition, optimization, and

expansion of new protein functions. Given the vast diversity of protein functional domains and protein-protein interactions in biology, we suspect that this is a common occurrence in the evolution of protein multifunctionality.

Conclusions

This study demonstrates the utility of placing biological findings into an evolutionary context to study the origin and mechanism of protein functions. By isolating and characterizing key intervals of evolutionary change in A9s, we have contributed to our understanding of how mammals evolved to activate and regulate inflammation and provided fresh insight into how these functions work. Our findings reveal a positive role for pleiotropy in evolution: pleiotropic substitutions can render proteins multifunctional by altering one functional state and leaving another intact, in this case leading to critical functional changes in mammalian innate immunity.

Acknowledgements

We thank current and former members of the Harms lab for helpful discussion and input. We thank the Barber lab for use of their cell culture facility. We thank Karen Guillemin for helpful comments on the manuscript.

Competing interests

None.

MATERIALS AND METHODS

Phylogenetics and ancestral sequence reconstruction

We reconstructed ancestral sequences using a previously published a phylogenetic tree of S100 proteins containing 172 sequences from 30 amniote taxa (File S2).⁵⁷ We used PAML4 to generate maximum likelihood ancestors (marginal probability method)^{87,88} using the previously-identified maximum likelihood (ML) substitution model (LG+ Γ_8)⁸⁹ on the ML tree. To account for reconstruction uncertainty, we also generated “altAll” versions of each ancestor.⁵⁹ We took every site in which the alternate reconstruction had a posterior probability > 0.20 and substituted that amino acid into the maximum-likelihood ancestor. These alternate reconstructions had an average of 12 sequence differences relative to the maximum-likelihood ancestors (Table S4). They represent a “worst case” reconstruction relative to our best, maximum likelihood reconstruction.

We also investigated the effect of topological uncertainty on our reconstructed ancestors. In the published phylogenetic analysis, A8s, A9s, A12s, and MRP126s all formed distinct and well-supported clades; however, the branching pattern between these four clades could not be resolved with high confidence.⁵⁷ To explore how this uncertainty altered our reconstructed ancestral proteins, we constructed all 15 possible topologies for the A8, A9, A12, and MRP126 clades—i.e ((A8,A9),(A12,MRP126)), ((A8,A12),(MRP126,A9)), etc.—while maintaining species-corrected, within-clade topologies. We then optimized the tree branch lengths and substitution rates for each tree using PhyML.⁹⁰ Finally, we used PAML to reconstruct ancA9, ancCG, and ancA8 for all 15 possible arrangements of the MRP126, A12, A8, and A9 clades. The average number of sequence differences for ancestors reconstructed using different topologies was less than or equal to the number of sequence differences between the ML and altAll reconstructions (Table S1). Further, the sites that differed were a subset of those that differed between the ML and altAll reconstructions. Thus, the altAll reconstructions account for sequence changes due to both uncertainty given the ML tree and uncertainty due to topological uncertainty.

Cloning and mutagenesis

All S100 genes in this study were purchased as synthetic constructs in pUC57 vectors from Genscript. S100 genes (A8s, A9s, A12s, MRP126s, and ancestrally reconstructed genes) were sub-cloned into a pETDuet-1 (pD) vector (Millipore). A8s, A12s, MRP126s, and ancCGs were cloned into multiple cloning site #1 (MCS1) of the pD vector, while A9s were cloned into MCS2. For expression and purification of A8/A9 heterocomplexes (A8/A9s), pD plasmids containing an A8 gene in MCS1 and an A9 gene in MCS2 were used as previously described.⁹¹ Opossum A8 was sub-cloned into an MBP-LIC vector to yield a His-MBP-TEV-opA8 construct. For opossum A8/A9, the entire His-MBP-TEV-A8 construct was then sub-cloned into MCS1 of a pD vector containing a marsupial A9 in MCS2. Other S100s (A1, A5, A7, A11, A14, and P) were previously cloned into a pET28/30 vector to yield a TEV-cleavable N-terminal His tag.⁶⁷ Cysteine-free versions of all S100 genes, as well as point mutants, were prepared using site-directed mutagenesis (Agilent).

Protein expression and purification

Recombinant protein overexpression was conducted in *E. coli* BL21 (DE3) pLysS Rosetta cells. Cultures were inoculated in luria broth overnight at 37°C, shaking at 250 rpm, in the presence of ampicillin and chloramphenicol. The following day, 10 ml of saturated culture was diluted into 1.5 L of media with antibiotics, grown to OD₆₀₀ = 0.6 – 1, and then induced overnight at 16°C using 1 mM IPTG. Cells were pelleted at 3,000 rpm for 20 min and stored at -20°C for no more than three months.

Lysates were prepared by vortexing pellets (3-5 g) in tris buffer (25 mM tris, 100 mM NaCl, pH 7.4) and incubating for 20 min at RT with DNase I and lysozyme (ThermoFisher Scientific). Lysates were sonicated and cell debris was pelleted by centrifugation at 15,000 rpm at 4°C for > 20 min. All proteins were purified on an Äkta PrimePlus FPLC using various 5 ml HiTrap columns (HisTrap FF (Ni-affinity), Q HP (anion exchange), SP FF (cation exchange), and MBPTrap HP (MBP) - GE Health Science). A1, A5, A7, A11, A14, and S100P were purified using a TEV-cleavable His tag strategy used by our lab previously^{1,3,4}. All other S100s, except for opossum A8 and opossum A8/A9, were purified in

three steps using Ni-affinity chromatography in the presence of calcium followed by two rounds of anion exchange chromatography at different pHs. For Ni-affinity chromatography, proteins were eluted over a 50 ml gradient from 25-1000 mM imidazole in tris buffer. Peak elution fractions were pooled and placed in dialysis overnight at 4°C in 4 L of tris buffer (calcium-free) adjusted to pH 8. Anion exchange chromatography was then performed the following day over a 50 ml gradient from 100-1000 mM NaCl in pH 8 tris buffer. Fractions containing majority S100 were pooled and analyzed for purity on an SDS-PAGE gel. If trace contaminants remained, an additional anion exchange step was performed at pH 6 using the same elution strategy as for the previous anion exchange step.

Opossum A8 and A8/A9 lysates were prepared as above and then flowed over a nickel column, eluting over a 50 ml gradient from 25-1000 mM imidazole in tris buffer. Peak elution was pooled and the MBP tag was cleaved by incubation with ~1:5 TEV protease at 4°C overnight in 4 L of tris buffer. The MBP tag was then removed by flowing the sample over an MBPTrap column, step-eluting with 10 mM maltose. Additional MBP columns were run until all MBP was removed from the purified protein, assessed by SDS-PAGE. If necessary, an additional anion exchange step at pH 8 was performed to complete purification. All purified proteins were dialyzed overnight at 4°C in tris buffer + 2 g/L Chelex-100 resin (Biorad), flash-frozen the following day in liquid nitrogen, and stored at -80°C.

Biophysical and biochemical characterization

For all experiments, protein aliquots were thawed fresh from freezer stocks and were either dialyzed in the appropriate experimental buffer overnight at 4°C or exchanged 3X into experimental buffer using 3K microsep spin concentrator columns (Pall Corporation). All samples were filter-sterilized using 0.1 um spin filters (EMD Millipore) prior to measuring concentration and using in experiments. Thawed aliquots were used for no more than one week before discarding. All concentrations were measured by Bradford assay and correspond to micromolar dimeric protein.

For *in vitro* proteolytic susceptibility experiments, proteins were dialyzed or exchanged into tris buffer + 1 mM CaCl₂. 12.5 uM S100 protein was treated with 5 uM monomeric Proteinase K from *Tritirachium album* (Sigma Aldrich) in thin-walled PCR tubes, which were held at a constant temperature of 25°C over the course of the experiment using a thermal cycler. Proteinase K activity was quenched at different time points by directly pipetting an aliquot of the reaction into an equal volume of 95% Laemmli SDS-PAGE loading buffer + 5% BME at 95°C in a separate thermal cycler. Time points were analyzed via SDS-PAGE, and gels were quantified by densitometry using in-house gel analysis software (<https://github.com/jharman25/gelquant>). An exponential decay function ($A_0 e^{-kt}$) was fit to the data to extract the decay rate, floating A_0 and k .

Oligomeric states were measured using a superose 12 10/300 GL size exclusion column (Amersham Biosciences) with in-line concentration detection using refractive index (RI) and particle mass measured using a multiangle laser light scattering (MALS) instrument (Dawn Heleos, Wyatt Technology). Samples were concentrated to 0.5-2 mg/ml in tris buffer + 0.5 mM CaCl₂, 0.1 um sterile-filtered, and analyzed at a flow rate of 0.2 ml/min. Data were processed using manufacturer's software (Astra).

Circular dichroism (CD) and chemical denaturation experiments were performed using a Jasco J-815 CD spectrometer and spectroscopy-grade guanidine hydrochloride (gdn-HCl). Chemical denaturation was performed using 25 uM dimeric protein in tris buffer with CaCl₂, with tris substituted for spectroscopy-grade trizma. Reversible unfolding and refolding curves were constructed by making concentrated 100 uM protein stocks in either buffer or 6M gdm and then preparing protein dilutions in various concentrations of gdn-HCl in buffer. Samples were left to equilibrate in denaturant for a minimum of three hours to allow for equilibration and were then analyzed by CD. Unfolding/refolding equilibration was confirmed by comparing unfolded vs. refolded protein at the same concentration. CD signal was quantified at 222 nm in a 1 mm cuvette using a 1 nm bandwidth, standard sensitivity, and 2

second D.I.T. HT voltage was < 600 V for all measurements. Apparent unfolding kinetics studies were performed using the above conditions by spiking concentrated protein stock directly into 6M gdm and immediately monitoring CD signal at 222 nm.

Functional Assays

The antimicrobial activity of S100s was measured against *S. epidermidis* using a well-established assay.^{32,36,40,41,92} The day before, a 5 ml starter culture of *S. epidermidis* in tryptic soy broth (TSB) was grown overnight. The next day, the culture was diluted ~1:100 in TSB and grown for approximately 2 hours to an OD600 of ~0.8. Immediately prior to experiment, the *S. epidermidis* culture was again diluted 1:100 at a ratio of 62:38 experimental buffer (25 mM tris, 100 mM NaCl, 3 mM CaCl₂, pH 7.4):TSB. S100 proteins were exchanged into experimental buffer. Each well of a sterile 96-well plate was prepared with 40 ul of *S. epidermidis* diluted in experimental buffer + TSB, S100 protein at the desired concentration in experimental buffer, and then filled to 200 ul, maintaining a ratio of 62:38 experimental buffer:TB. *S. epidermidis* growth was monitored on a plate reader, measuring OD600 every 15 minutes for 13 hours. Each measurement was collected in technical triplicate and background-subtracted using a blank containing experimental buffer and TSB alone. Protein samples were confirmed to lack bacterial contamination by measuring S100 protein growth in experimental buffer and TSB lacking *S. epidermidis*.

All plasmids, cell culture conditions, and transfections for measuring the activity of S100s against TLR4s were identical to those previously described.^{38,39,57,60} Lipopolysaccharide *E. coli* K-12 LPS (LPS - tlrl-eklps, Invivogen) aliquots were prepared at 5 mg/ml in endotoxin-free water and stored at -20°C. Working solutions were prepared at 10 ug/ml and stored at 4°C to avoid freeze-thaw cycles. S100 proteins were prepared by exchanging into endotoxin-free PBS and incubating with an endotoxin removal column (Thermo Fisher Scientific) for 2 hours. S100 LPS contamination was assessed by measuring activity with and without Polymixin B, an LPS chelating agent (Figure S6). LPS (200 ng per 100 ul well) or S100 (0.8, 0.4, 2, 4, or 5 uM dimer) treatments were prepared by diluting in 25:75

674 endotoxin-free PBS:serum-free Dulbecco's Modified Eagle Media (DMEM – Thermo Fisher
675 Scientific). Polymixin B (PB, 200 ug per 100 ul well) was added to all S100 experimental samples to
676 limit background endotoxin contamination activity from recombinant protein preps. Cells were
677 incubated with treatments for 3 hours prior to assaying activity. The Dual-Glo Luciferase Assay
678 System (Promega) was used to assay Firefly and Renilla luciferase activity of individual wells. Each
679 NF- κ B induction value shown represents the Firefly luciferase activity divided by the Renilla luciferase
680 activity, background-subtracted using the LPS + PB activity for each TLR4 species and normalized to
681 the activity of LPS alone for each TLR4 species to normalize between plates. All measurements were
682 performed using three technical replicates per plate, a minimum of three biological replicates total, and
683 a minimum of two separate protein preps.

684 For TLR4 activation measurements by A9 proteolytic products, 12.5 uM hA9 or hA9 M63F
685 were incubated with 2.5 mg/ml Proteinase K immobilized to agarose at 37°C for increasing amounts of
686 time. The reaction was quenched by spin-filtering the sample to remove Proteinase K. 2 uM A9
687 proteolysis treatments were then added to cells as outlined above. Western blots were performed by
688 running an SDS-PAGE gel and transferring to a nitrocellulose membrane. Membranes were blocked
689 using Odyssey Blocking Buffer for 1 hour, incubated with 1:1000 mouse anti-S100A9 primary
690 antibody (M13 clone 1CD22, Abnova) for 1 hour, and incubated with 1:10,000 IRDye Goat anti-mouse
691 800CW IgG (H+L, Licor) for 1 hour, with 3x5 min TBST washes in between each step. Blots were
692 imaged using the Licor Odyssey Fc imaging system.

693 References

- 694 (1) Lee, M. W.; Lee, E. Y.; Wong, G. C. L. What Can Pleiotropic Proteins in Innate Immunity
695 Teach Us about Bioconjugation and Molecular Design? *Bioconjug. Chem.* **2018**, 29 (7), 2127–
696 2139. <https://doi.org/10.1021/acs.bioconjchem.8b00176>.
697 (2) Auvynet, C.; Rosenstein, Y. Multifunctional Host Defense Peptides: Antimicrobial Peptides, the
698 Small yet Big Players in Innate and Adaptive Immunity: AMPs, the Small yet Big Players of
699 Immunity. *FEBS J.* **2009**, 276 (22), 6497–6508. [https://doi.org/10.1111/j.1742-](https://doi.org/10.1111/j.1742-4658.2009.07360.x)
700 4658.2009.07360.x.

- 701 (3) Postel, S.; Küfner, I.; Beuter, C.; Mazzotta, S.; Schwedt, A.; Borlotti, A.; Halter, T.;
702 Kemmerling, B.; Nürnberger, T. The Multifunctional Leucine-Rich Repeat Receptor Kinase
703 BAK1 Is Implicated in Arabidopsis Development and Immunity. *Eur. J. Cell Biol.* **2010**, *89* (2–
704 3), 169–174. <https://doi.org/10.1016/j.ejcb.2009.11.001>.
- 705 (4) Gudmundsson, G. H.; Agerberth, B. Neutrophil Antibacterial Peptides, Multifunctional Effector
706 Molecules in the Mammalian Immune System. *J. Immunol. Methods* **1999**, *232* (1–2), 45–54.
707 [https://doi.org/10.1016/S0022-1759\(99\)00152-0](https://doi.org/10.1016/S0022-1759(99)00152-0).
- 708 (5) Singh, V.; Yeoh, B. S.; Chassaing, B.; Zhang, B.; Saha, P.; Xiao, X.; Awasthi, D.;
709 Shashidharamurthy, R.; Dikshit, M.; Gewirtz, A.; et al. Microbiota-Inducible Innate Immune
710 Siderophore Binding Protein Lipocalin 2 Is Critical for Intestinal Homeostasis. *Cell. Mol.*
711 *Gastroenterol. Hepatol.* **2016**, *2* (4), 482–498.e6. <https://doi.org/10.1016/j.jcmgh.2016.03.007>.
- 712 (6) Säemann, M. D.; Haidinger, M.; Hecking, M.; Hörl, W. H.; Weichhart, T. The Multifunctional
713 Role of MTOR in Innate Immunity: Implications for Transplant Immunity. *Am. J. Transplant.*
714 **2009**, *9* (12), 2655–2661. <https://doi.org/10.1111/j.1600-6143.2009.02832.x>.
- 715 (7) Bals, R.; Wilson, J. M. Cathelicidins - a Family of Multifunctional Antimicrobial Peptides. *Cell.*
716 *Mol. Life Sci. CMLS* **2003**, *60* (4), 711–720. <https://doi.org/10.1007/s00018-003-2186-9>.
- 717 (8) Leask, A.; Abraham, D. J. The Role of Connective Tissue Growth Factor, a Multifunctional
718 Matricellular Protein, in Fibroblast Biology. *Biochem. Cell Biol.* **2003**, *81* (6), 355–363.
719 <https://doi.org/10.1139/o03-069>.
- 720 (9) Ehrchen, J. M.; Sunderkötter, C.; Foell, D.; Vogl, T.; Roth, J. The Endogenous Toll-like
721 Receptor 4 Agonist S100A8/S100A9 (Calprotectin) as Innate Amplifier of Infection,
722 Autoimmunity, and Cancer. *J. Leukoc. Biol.* **2009**, *86* (3), 557–566.
723 <https://doi.org/10.1189/jlb.1008647>.
- 724 (10) Salama, I.; Malone, P. S.; Mihaimeed, F.; Jones, J. L. A Review of the S100 Proteins in Cancer.
725 *Eur. J. Surg. Oncol. EJSO* **2008**, *34* (4), 357–364. <https://doi.org/10.1016/j.ejso.2007.04.009>.
- 726 (11) Shabani, F.; Farasat, A.; Mahdavi, M.; Gheibi, N. Calprotectin (S100A8/S100A9): A Key
727 Protein between Inflammation and Cancer. *Inflamm. Res.* **2018**. [https://doi.org/10.1007/s00011-](https://doi.org/10.1007/s00011-018-1173-4)
728 *018-1173-4*.
- 729 (12) Gruden, M. A.; Davydova, T. V.; Wang, C.; Narkevich, V. B.; Fomina, V. G.; Kudrin, V. S.;
730 Morozova-Roche, L. A.; Sewell, R. D. E. The Misfolded Pro-Inflammatory Protein S100A9
731 Disrupts Memory via Neurochemical Remodelling Instigating an Alzheimer's Disease-like
732 Cognitive Deficit. *Behav. Brain Res.* **2016**, *306*, 106–116.
733 <https://doi.org/10.1016/j.bbr.2016.03.016>.
- 734 (13) Peng, S.; Sun, X.; Wang, X.; Wang, H.; Shan, Z.; Teng, W.; Li, C. Myeloid Related Proteins Are
735 Up-Regulated in Autoimmune Thyroid Diseases and Activate Toll-like Receptor 4 and pro-
736 Inflammatory Cytokines in Vitro. *Int. Immunopharmacol.* **2018**, *59*, 217–226.
737 <https://doi.org/10.1016/j.intimp.2018.04.009>.
- 738 (14) Bombliès, K. Pleiotropic Effects of the Duplicate Maize FLORICAULA/LEAFY Genes Zfl1 and
739 Zfl2 on Traits Under Selection During Maize Domestication. *Genetics* **2005**, *172* (1), 519–531.
740 <https://doi.org/10.1534/genetics.105.048595>.
- 741 (15) Stearns, F. W. One Hundred Years of Pleiotropy: A Retrospective. *Genetics* **2010**, *186* (3), 767–
742 773. <https://doi.org/10.1534/genetics.110.122549>.
- 743 (16) Paaby, A. B.; Rockman, M. V. The Many Faces of Pleiotropy. *Trends Genet.* **2013**, *29* (2), 66–
744 73. <https://doi.org/10.1016/j.tig.2012.10.010>.
- 745 (17) Berntzen, H. B.; Fagerhol, M. K. L1, a Major Granulocyte Protein; Isolation of High Quantities
746 of Its Subunits. *Scand. J. Clin. Lab. Invest.* **1990**, *50* (7), 769–774.
747 <https://doi.org/10.1080/00365519009091071>.

- (18) Vogl, T.; Gharibyan, A. L.; Morozova-Roche, L. A. Pro-Inflammatory S100A8 and S100A9 Proteins: Self-Assembly into Multifunctional Native and Amyloid Complexes. *Int. J. Mol. Sci.* **2012**, *13* (12), 2893–2917. <https://doi.org/10.3390/ijms13032893>.
- (19) Källberg, E.; Vogl, T.; Liberg, D.; Olsson, A.; Björk, P.; Wikström, P.; Bergh, A.; Roth, J.; Ivars, F.; Leanderson, T. S100A9 Interaction with TLR4 Promotes Tumor Growth. *PLoS ONE* **2012**, *7* (3), e34207. <https://doi.org/10.1371/journal.pone.0034207>.
- (20) Duan, L.; Wu, R.; Zhang, X.; Wang, D.; You, Y.; Zhang, Y.; Zhou, L.; Chen, W. HBx-Induced S100A9 in NF-KB Dependent Manner Promotes Growth and Metastasis of Hepatocellular Carcinoma Cells. *Cell Death Dis.* **2018**, *9* (6). <https://doi.org/10.1038/s41419-018-0512-2>.
- (21) Shepherd, C. E.; Goyette, J.; Utter, V.; Rahimi, F.; Yang, Z.; Geczy, C. L.; Halliday, G. M. Inflammatory S100A9 and S100A12 Proteins in Alzheimer's Disease. *Neurobiol. Aging* **2006**, *27* (11), 1554–1563. <https://doi.org/10.1016/j.neurobiolaging.2005.09.033>.
- (22) Schiopu, A.; Cotoi, O. S. S100A8 and S100A9: DAMPs at the Crossroads between Innate Immunity, Traditional Risk Factors, and Cardiovascular Disease. *Mediators Inflamm.* **2013**, *2013*, 1–10. <https://doi.org/10.1155/2013/828354>.
- (23) Laouedj, M.; Tardif, M. R.; Gil, L.; Raquil, M.-A.; Lachhab, A.; Pelletier, M.; Tessier, P. A.; Barabé, F. S100A9 Induces Differentiation of Acute Myeloid Leukemia Cells through TLR4. *Blood* **2017**, *129* (14), 1980–1990. <https://doi.org/10.1182/blood-2016-09-738005>.
- (24) He, Z.; Riva, M.; Björk, P.; Swärd, K.; Mörgelin, M.; Leanderson, T.; Ivars, F. CD14 Is a Co-Receptor for TLR4 in the S100A9-Induced Pro-Inflammatory Response in Monocytes. *PLOS ONE* **2016**, *11* (5), e0156377. <https://doi.org/10.1371/journal.pone.0156377>.
- (25) Gao, H.; Zhang, X.; Zheng, Y.; Peng, L.; Hou, J.; Meng, H. S100A9-Induced Release of Interleukin (IL)-6 and IL-8 through Toll-like Receptor 4 (TLR4) in Human Periodontal Ligament Cells. *Mol. Immunol.* **2015**, *67* (2), 223–232. <https://doi.org/10.1016/j.molimm.2015.05.014>.
- (26) Tsai, S.-Y.; Segovia, J. A.; Chang, T.-H.; Morris, I. R.; Berton, M. T.; Tessier, P. A.; Tardif, M. R.; Cesaro, A.; Bose, S. DAMP Molecule S100A9 Acts as a Molecular Pattern to Enhance Inflammation during Influenza A Virus Infection: Role of DDX21-TRIF-TLR4-MyD88 Pathway. *PLoS Pathog.* **2014**, *10* (1), e1003848. <https://doi.org/10.1371/journal.ppat.1003848>.
- (27) Lee, N. R.; Park, B. S.; Kim, S. Y.; Gu, A.; Kim, D. H.; Lee, J.-S.; Kim, I. S. Cytokine Secreted by S100A9 via TLR4 in Monocytes Delays Neutrophil Apoptosis by Inhibition of Caspase 9/3 Pathway. *Cytokine* **2016**, *86*, 53–63. <https://doi.org/10.1016/j.cyto.2016.07.005>.
- (28) Björk, P.; Björk, A.; Vogl, T.; Stenström, M.; Liberg, D.; Olsson, A.; Roth, J.; Ivars, F.; Leanderson, T. Identification of Human S100A9 as a Novel Target for Treatment of Autoimmune Disease via Binding to Quinoline-3-Carboxamides. *PLoS Biol.* **2009**, *7* (4), e1000097. <https://doi.org/10.1371/journal.pbio.1000097>.
- (29) Kang, J. H.; Hwang, S. M.; Chung, I. Y. S100A8, S100A9 and S100A12 Activate Airway Epithelial Cells to Produce MUC5AC via Extracellular Signal-Regulated Kinase and Nuclear Factor- κ B Pathways. *Immunology* **2015**, *144* (1), 79–90. <https://doi.org/10.1111/imm.12352>.
- (30) Chen, B.; Miller, A. L.; Rebelatto, M.; Brewah, Y.; Rowe, D. C.; Clarke, L.; Czapiga, M.; Rosenthal, K.; Imamichi, T.; Chen, Y.; et al. S100A9 Induced Inflammatory Responses Are Mediated by Distinct Damage Associated Molecular Patterns (DAMP) Receptors In Vitro and In Vivo. *PLOS ONE* **2015**, *10* (2), e0115828. <https://doi.org/10.1371/journal.pone.0115828>.
- (31) Besold, A. N.; Culbertson, E. M.; Nam, L.; Hobbs, R. P.; Boyko, A.; Maxwell, C. N.; Chazin, W. J.; Marques, A. R.; Culotta, V. C. Antimicrobial Action of Calprotectin That Does Not Involve Metal Withholding. *Metallomics* **2018**. <https://doi.org/10.1039/C8MT00133B>.
- (32) Hadley, R. C.; Gu, Y.; Nolan, E. M. Initial Biochemical and Functional Evaluation of Murine Calprotectin Reveals Ca(II)-Dependence and Its Ability to Chelate Multiple Nutrient Transition Metal Ions. *Biochemistry* **2018**. <https://doi.org/10.1021/acs.biochem.8b00309>.

- (33) Damo, S. M.; Kehl-Fie, T. E.; Sugitani, N.; Holt, M. E.; Rathi, S.; Murphy, W. J.; Zhang, Y.; Betz, C.; Hench, L.; Fritz, G. Molecular Basis for Manganese Sequestration by Calprotectin and Roles in the Innate Immune Response to Invading Bacterial Pathogens. *Proc. Natl. Acad. Sci.* **2013**, *110* (10), 3841–3846.
- (34) Clark, H. L.; Jhingran, A.; Sun, Y.; Vareechon, C.; de Jesus Carrion, S.; Skaar, E. P.; Chazin, W. J.; Calera, J. A.; Hohl, T. M.; Pearlman, E. Zinc and Manganese Chelation by Neutrophil S100A8/A9 (Calprotectin) Limits Extracellular *Aspergillus Fumigatus* Hyphal Growth and Corneal Infection. *J. Immunol.* **2016**, *196* (1), 336–344.
<https://doi.org/10.4049/jimmunol.1502037>.
- (35) Besold, A. N.; Gilston, B. A.; Radin, J. N.; Ramsoomair, C.; Culbertson, E. M.; Li, C. X.; Cormack, B. P.; Chazin, W. J.; Kehl-Fie, T. E.; Culotta, V. C. The Role of Calprotectin in Withholding Zinc and Copper from *Candida Albicans*. *Infect. Immun.* **2017**, IAI.00779-17.
<https://doi.org/10.1128/IAI.00779-17>.
- (36) Nakashige, T. G.; Stephan, J. R.; Cunden, L. S.; Brophy, M. B.; Wommack, A. J.; Keegan, B. C.; Shearer, J. M.; Nolan, E. M. The Hexahistidine Motif of Host-Defense Protein Human Calprotectin Contributes to Zinc Withholding and Its Functional Versatility. *J. Am. Chem. Soc.* **2016**, *138* (37), 12243–12251. <https://doi.org/10.1021/jacs.6b06845>.
- (37) Hayden, J. A.; Brophy, M. B.; Cunden, L. S.; Nolan, E. M. High-Affinity Manganese Coordination by Human Calprotectin Is Calcium-Dependent and Requires the Histidine-Rich Site Formed at the Dimer Interface. *J. Am. Chem. Soc.* **2013**, *135* (2), 775–787.
<https://doi.org/10.1021/ja3096416>.
- (38) Nakashige, T. G.; Zhang, B.; Krebs, C.; Nolan, E. M. Human Calprotectin Is an Iron-Sequestering Host-Defense Protein. *Nat. Chem. Biol.* **2015**, *11* (10), 765–771.
<https://doi.org/10.1038/nchembio.1891>.
- (39) Liu, J. Z.; Jellbauer, S.; Poe, A. J.; Ton, V.; Pesciaroli, M.; Kehl-Fie, T. E.; Restrepo, N. A.; Hosking, M. P.; Edwards, R. A.; Battistoni, A.; et al. Zinc Sequestration by the Neutrophil Protein Calprotectin Enhances *Salmonella* Growth in the Inflamed Gut. *Cell Host Microbe* **2012**, *11* (3), 227–239. <https://doi.org/10.1016/j.chom.2012.01.017>.
- (40) Brunjes Brophy, M.; Nakashige, T. G.; Gaillard, A.; Nolan, E. M. Contributions of the S100A9 C-Terminal Tail to High-Affinity Mn(II) Chelation by the Host-Defense Protein Human Calprotectin. *J. Am. Chem. Soc.* **2013**, *135* (47), 17804–17817.
<https://doi.org/10.1021/ja407147d>.
- (41) Brophy, M. B.; Hayden, J. A.; Nolan, E. M. Calcium Ion Gradients Modulate the Zinc Affinity and Antibacterial Activity of Human Calprotectin. *J. Am. Chem. Soc.* **2012**, *134* (43), 18089–18100. <https://doi.org/10.1021/ja307974e>.
- (42) Baker, T. M.; Nakashige, T. G.; Nolan, E. M.; Neidig, M. L. Magnetic Circular Dichroism Studies of Iron(II) Binding to Human Calprotectin. *Chem Sci* **2017**, *8* (2), 1369–1377.
<https://doi.org/10.1039/C6SC03487J>.
- (43) Gagnon, D. M.; Brophy, M. B.; Bowman, S. E. J.; Stich, T. A.; Drennan, C. L.; Britt, R. D.; Nolan, E. M. Manganese Binding Properties of Human Calprotectin under Conditions of High and Low Calcium: X-Ray Crystallographic and Advanced Electron Paramagnetic Resonance Spectroscopic Analysis. *J. Am. Chem. Soc.* **2015**, *137* (8), 3004–3016.
<https://doi.org/10.1021/ja512204s>.
- (44) Nisapakultorn, K.; Ross, K. F.; Herzberg, M. C. Calprotectin Expression Inhibits Bacterial Binding to Mucosal Epithelial Cells. *Infect. Immun.* **2001**, *69* (6), 3692–3696.
<https://doi.org/10.1128/IAI.69.6.3692-3696.2001>.
- (45) Vogl, T.; Tenbrock, K.; Ludwig, S.; Leukert, N.; Ehrhardt, C.; van Zoelen, M. A. D.; Nacken, W.; Foell, D.; van der Poll, T.; Sorg, C.; et al. Mrp8 and Mrp14 Are Endogenous Activators of

- 845 Toll-like Receptor 4, Promoting Lethal, Endotoxin-Induced Shock. *Nat. Med.* **2007**, *13* (9),
846 1042–1049. <https://doi.org/10.1038/nm1638>.
- 847 (46) Vogl, T.; Eisenblätter, M.; Völler, T.; Zenker, S.; Hermann, S.; van Lent, P.; Faust, A.; Geyer,
848 C.; Petersen, B.; Roebrock, K.; et al. Alarmin S100A8/S100A9 as a Biomarker for Molecular
849 Imaging of Local Inflammatory Activity. *Nat. Commun.* **2014**, *5* (1), 4593.
850 <https://doi.org/10.1038/ncomms5593>.
- 851 (47) Hara, A.; Sakamoto, N.; Ishimatsu, Y.; Kakugawa, T.; Nakashima, S.; Hara, S.; Adachi, M.;
852 Fujita, H.; Mukae, H.; Kohno, S. S100A9 in BALF Is a Candidate Biomarker of Idiopathic
853 Pulmonary Fibrosis. *Respir. Med.* **2012**, *106* (4), 571–580.
854 <https://doi.org/10.1016/j.rmed.2011.12.010>.
- 855 (48) Obry, A.; Lequerré, T.; Hardouin, J.; Boyer, O.; Fardellone, P.; Philippe, P.; Le Loët, X.;
856 Cosette, P.; Vittecoq, O. Identification of S100A9 as Biomarker of Responsiveness to the
857 Methotrexate/Etanercept Combination in Rheumatoid Arthritis Using a Proteomic Approach.
858 *PLoS ONE* **2014**, *9* (12), e115800. <https://doi.org/10.1371/journal.pone.0115800>.
- 859 (49) Horvath, I.; Jia, X.; Johansson, P.; Wang, C.; Moskalenko, R.; Steinau, A.; Forsgren, L.;
860 Wågberg, T.; Svensson, J.; Zetterberg, H.; et al. Pro-Inflammatory S100A9 Protein as a Robust
861 Biomarker Differentiating Early Stages of Cognitive Impairment in Alzheimer's Disease. *ACS*
862 *Chem. Neurosci.* **2016**, *7* (1), 34–39. <https://doi.org/10.1021/acscchemneuro.5b00265>.
- 863 (50) Huang, C.-H.; Kuo, C.-J.; Liang, S.-S.; Chi, S.-W.; Hsi, E.; Chen, C.-C.; Lee, K.-T.; Chiou, S.-
864 H. Onco-Proteogenomics Identifies Urinary S100A9 and GRN as Potential Combinatorial
865 Biomarkers for Early Diagnosis of Hepatocellular Carcinoma. *BBA Clin.* **2015**, *3*, 205–213.
866 <https://doi.org/10.1016/j.bbacli.2015.02.004>.
- 867 (51) Kim, H.-J.; Kang, H. J.; Lee, H.; Lee, S.-T.; Yu, M.-H.; Kim, H.; Lee, C. Identification of
868 S100A8 and S100A9 as Serological Markers for Colorectal Cancer. *J. Proteome Res.* **2009**, *8*
869 (3), 1368–1379. <https://doi.org/10.1021/pr8007573>.
- 870 (52) Averill, M. M.; Kerkhoff, C.; Bornfeldt, K. E. S100A8 and S100A9 in Cardiovascular Biology
871 and Disease. *Arterioscler. Thromb. Vasc. Biol.* **2012**, *32* (2), 223–229.
872 <https://doi.org/10.1161/ATVBAHA.111.236927>.
- 873 (53) Vogl, T.; Stratis, A.; Wixler, V.; Völler, T.; Thurainayagam, S.; Jorch, S. K.; Zenker, S.;
874 Dreiling, A.; Chakraborty, D.; Fröhling, M.; et al. Autoinhibitory Regulation of S100A8/S100A9
875 Alarmin Activity Locally Restricts Sterile Inflammation. *J. Clin. Invest.* **2018**.
876 <https://doi.org/10.1172/JCI89867>.
- 877 (54) Nacken, W.; Kerkhoff, C. The Hetero-Oligomeric Complex of the S100A8/S100A9 Protein Is
878 Extremely Protease Resistant. *FEBS Lett.* **2007**, *581* (26), 5127–5130.
879 <https://doi.org/10.1016/j.febslet.2007.09.060>.
- 880 (55) Riva, M.; He, Z.; Källberg, E.; Ivars, F.; Leanderson, T. Human S100A9 Protein Is Stabilized by
881 Inflammatory Stimuli via the Formation of Proteolytically-Resistant Homodimers. *PLoS ONE*
882 **2013**, *8* (4), e61832. <https://doi.org/10.1371/journal.pone.0061832>.
- 883 (56) Stephan, J. R.; Nolan, E. M. Calcium-Induced Tetramerization and Zinc Chelation Shield
884 Human Calprotectin from Degradation by Host and Bacterial Extracellular Proteases. *Chem Sci*
885 **2016**, *7* (3), 1962–1975. <https://doi.org/10.1039/C5SC03287C>.
- 886 (57) Loes, A. N.; Bridgham, J. T.; Harms, M. J. Coevolution of the Toll-Like Receptor 4 Complex
887 with Calgranulins and Lipopolysaccharide. *Front. Immunol.* **2018**, *9*.
888 <https://doi.org/10.3389/fimmu.2018.00304>.
- 889 (58) Streicher, W. W.; Lopez, M. M.; Makhatadze, G. I. Modulation of Quaternary Structure of S100
890 Proteins by Calcium Ions. *Biophys. Chem.* **2010**, *151* (3), 181–186.
891 <https://doi.org/10.1016/j.bpc.2010.06.003>.

- (59) Eick, G. N.; Bridgham, J. T.; Anderson, D. P.; Harms, M. J.; Thornton, J. W. Robustness of Reconstructed Ancestral Protein Functions to Statistical Uncertainty. *Mol. Biol. Evol.* **2016**, msw223. <https://doi.org/10.1093/molbev/msw223>.
- (60) Anderson, J. A.; Loes, A. N.; Waddell, G. L.; Harms, M. J. Tracing the Evolution of Novel Features of Human Toll-like Receptor 4. *Protein Sci.* **2019**, pro.3644. <https://doi.org/10.1002/pro.3644>.
- (61) Ehrhardt, C.; Schmolke, M.; Matzke, A.; Knoblauch, A.; Will, C.; Wixler, V.; Ludwig, S. Polyethylenimine, a Cost-Effective Transfection Reagent. *Signal Transduct.* **2006**, 6 (3), 179–184. <https://doi.org/10.1002/sita.200500073>.
- (62) Henry, C. M.; Sullivan, G. P.; Clancy, D. M.; Afonina, I. S.; Kulms, D.; Martin, S. J. Neutrophil-Derived Proteases Escalate Inflammation through Activation of IL-36 Family Cytokines. *Cell Rep.* **2016**, 14 (4), 708–722. <https://doi.org/10.1016/j.celrep.2015.12.072>.
- (63) Heutinck, K. M.; ten Berge, I. J. M.; Hack, C. E.; Hamann, J.; Rowshani, A. T. Serine Proteases of the Human Immune System in Health and Disease. *Mol. Immunol.* **2010**, 47 (11–12), 1943–1955. <https://doi.org/10.1016/j.molimm.2010.04.020>.
- (64) Kessenbrock, K.; Dau, T.; Jenne, D. E. Tailor-Made Inflammation: How Neutrophil Serine Proteases Modulate the Inflammatory Response. *J. Mol. Med.* **2011**, 89 (1), 23–28. <https://doi.org/10.1007/s00109-010-0677-3>.
- (65) Stapels, D. A.; Geisbrecht, B. V.; Rooijackers, S. H. Neutrophil Serine Proteases in Antibacterial Defense. *Curr. Opin. Microbiol.* **2015**, 23, 42–48. <https://doi.org/10.1016/j.mib.2014.11.002>.
- (66) Fu, Z.; Thorpe, M.; Akula, S.; Chahal, G.; Hellman, L. T. Extended Cleavage Specificity of Human Neutrophil Elastase, Human Proteinase 3, and Their Distant Ortholog Clawed Frog PR3—Three Elastases With Similar Primary but Different Extended Specificities and Stability. *Front. Immunol.* **2018**, 9, 2387. <https://doi.org/10.3389/fimmu.2018.02387>.
- (67) Wheeler, L. C.; Donor, M. T.; Prell, J. S.; Harms, M. J. Multiple Evolutionary Origins of Ubiquitous Cu²⁺ and Zn²⁺ Binding in the S100 Protein Family. *PLOS ONE* **2016**, 11 (10), e0164740. <https://doi.org/10.1371/journal.pone.0164740>.
- (68) Itou, H.; Yao, M.; Fujita, I.; Watanabe, N.; Suzuki, M.; Nishihira, J.; Tanaka, I. The Crystal Structure of Human MRP14 (S100A9), a Ca²⁺-Dependent Regulator Protein in Inflammatory Process. *J. Mol. Biol.* **2002**, 316 (2), 265–276. <https://doi.org/10.1006/jmbi.2001.5340>.
- (69) He, X. Toward a Molecular Understanding of Pleiotropy. *Genetics* **2006**, 173 (4), 1885–1891. <https://doi.org/10.1534/genetics.106.060269>.
- (70) Pavličev, M.; Cheverud, J. M. Constraints Evolve: Context Dependency of Gene Effects Allows Evolution of Pleiotropy. *Annu. Rev. Ecol. Evol. Syst.* **2015**, 46 (1), 413–434. <https://doi.org/10.1146/annurev-ecolsys-120213-091721>.
- (71) ST, I. Calprotectin – a Pleiotropic Molecule in Acute and Chronic Inflammation. **2004**, 53, 9.
- (72) Harms, M. J.; Thornton, J. W. Analyzing Protein Structure and Function Using Ancestral Gene Reconstruction. *Curr. Opin. Struct. Biol.* **2010**, 20 (3), 360–366. <https://doi.org/10.1016/j.sbi.2010.03.005>.
- (73) Ibrahim, Z. A.; Armour, C. L.; Phipps, S.; Sukkar, M. B. RAGE and TLRs: Relatives, Friends or Neighbours? *Mol. Immunol.* **2013**, 56 (4), 739–744. <https://doi.org/10.1016/j.molimm.2013.07.008>.
- (74) Nagai, Y.; Akashi, S.; Nagafuku, M.; Ogata, M.; Iwakura, Y.; Akira, S.; Kitamura, T.; Kosugi, A.; Kimoto, M.; Miyake, K. Essential Role of MD-2 in LPS Responsiveness and TLR4 Distribution. *Nat. Immunol.* **2002**, 3 (7), 667–672. <https://doi.org/10.1038/ni809>.
- (75) Poltorak, A. Defective LPS Signaling in C3H/HeJ and C57BL/10ScCr Mice: Mutations in Tlr4 Gene. *Science* **1998**, 282 (5396), 2085–2088. <https://doi.org/10.1126/science.282.5396.2085>.

- (76) Prince, L. R.; Whyte, M. K.; Sabroe, I.; Parker, L. C. The Role of TLRs in Neutrophil Activation. *Curr. Opin. Pharmacol.* **2011**, *11* (4), 397–403. <https://doi.org/10.1016/j.coph.2011.06.007>.
- (77) Park, B. S.; Song, D. H.; Kim, H. M.; Choi, B.-S.; Lee, H.; Lee, J.-O. The Structural Basis of Lipopolysaccharide Recognition by the TLR4–MD-2 Complex. *Nature* **2009**, *458* (7242), 1191–1195. <https://doi.org/10.1038/nature07830>.
- (78) Guharoy, M.; Bhowmick, P.; Tompa, P. Design Principles Involving Protein Disorder Facilitate Specific Substrate Selection and Degradation by the Ubiquitin-Proteasome System. *J. Biol. Chem.* **2016**, *291* (13), 6723–6731. <https://doi.org/10.1074/jbc.R115.692665>.
- (79) Fontana, A.; Polverino de Laureto, P.; De Filippis, V.; Scaramella, E.; Zambonin, M. Probing the Partly Folded States of Proteins by Limited Proteolysis. *Fold. Des.* **1997**, *2* (2), R17–R26. [https://doi.org/10.1016/S1359-0278\(97\)00010-2](https://doi.org/10.1016/S1359-0278(97)00010-2).
- (80) Hubbard, S. J. The Structural Aspects of Limited Proteolysis of Native Proteins. *Biochim. Biophys. Acta BBA - Protein Struct. Mol. Enzymol.* **1998**, *1382* (2), 191–206. [https://doi.org/10.1016/S0167-4838\(97\)00175-1](https://doi.org/10.1016/S0167-4838(97)00175-1).
- (81) Ottesen, M. Induction of Biological Activity by Limited Proteolysis. *Annu. Rev. Biochem.* **1967**, *36* (1), 55–76. <https://doi.org/10.1146/annurev.bi.36.070167.000415>.
- (82) Imoto, T.; Yamada, H.; Ueda, T. Unfolding Rates of Globular Proteins Determined by Kinetics of Proteolysis. *J. Mol. Biol.* **1986**, *190* (4), 647–649. [https://doi.org/10.1016/0022-2836\(86\)90250-0](https://doi.org/10.1016/0022-2836(86)90250-0).
- (83) Garg, A. D.; Nowis, D.; Golab, J.; Vandenabeele, P.; Krysko, D. V.; Agostinis, P. Immunogenic Cell Death, DAMPs and Anticancer Therapeutics: An Emerging Amalgamation. *Biochim. Biophys. Acta BBA - Rev. Cancer* **2010**, *1805* (1), 53–71. <https://doi.org/10.1016/j.bbcan.2009.08.003>.
- (84) Rubartelli, A.; Lotze, M. T. Inside, Outside, Upside down: Damage-Associated Molecular-Pattern Molecules (DAMPs) and Redox. *Trends Immunol.* **2007**, *28* (10), 429–436. <https://doi.org/10.1016/j.it.2007.08.004>.
- (85) Bianchi, M. E. DAMPs, PAMPs and Alarmins: All We Need to Know about Danger. *J. Leukoc. Biol.* **2007**, *81* (1), 1–5. <https://doi.org/10.1189/jlb.0306164>.
- (86) Hirano, T. Molecular Basis Underlying Functional Pleiotropy of Cytokines and Growth Factors. *Biochem. Biophys. Res. Commun.* **1999**, *260* (2), 303–308. <https://doi.org/10.1006/bbrc.1999.0609>.
- (87) Ziheng Yang; Sudhir Kumar; Masatoshi Nei. A New Method of Inference of Ancestral Nucleotide and Amino Acid Sequences. *Genetics* **1995**, No. 141, 1641–1650.
- (88) Yang, Z. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Mol. Biol. Evol.* **2007**, *24* (8), 1586–1591. <https://doi.org/10.1093/molbev/msm088>.
- (89) Jones, D. T.; Taylor, W. R.; Thornton, J. M. The Rapid Generation of Mutation Data Matrices from Protein Sequences. *Bioinformatics* **1992**, *8* (3), 275–282. <https://doi.org/10.1093/bioinformatics/8.3.275>.
- (90) Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst. Biol.* **2010**, *59* (3), 307–321. <https://doi.org/10.1093/sysbio/syq010>.
- (91) Futami, J.; Atago, Y.; Azuma, A.; Putranto, E. W.; Kinoshita, R.; Murata, H.; Sakaguchi, M. An Efficient Method for the Preparation of Preferentially Heterodimerized Recombinant S100A8/A9 Coexpressed in Escherichia Coli. *Biochem. Biophys. Rep.* **2016**, *6*, 94–100. <https://doi.org/10.1016/j.bbrep.2016.03.009>.

(92) Cunden, L. S.; Gaillard, A.; Nolan, E. M. Calcium Ions Tune the Zinc-Sequestering Properties and Antimicrobial Activity of Human S100A12. *Chem Sci* **2016**, 7 (2), 1338–1348. <https://doi.org/10.1039/C5SC03655K>.

All species cartoons were taken from the following websites: <http://phylopic.org/image/c089caae-43ef-4e4e-bf26-973dd4cb65c5/>, <http://phylopic.org/image/aff847b0-ecbd-4d41-98ce-665921a6d96e/>, <http://phylopic.org/image/0f6af3d8-49d2-4d75-8edf-08598387afde/http://phylopic.org/image/dde4f926-c04c-47ef-a337-927ceb36e7ef/>. We acknowledge Sarah Werning and David Liao as authors of the opossum and mouse cartoons respectively, which were made publicly available through the creative commons attributions 3.0 unported license (<https://creativecommons.org/licenses/by/3.0/>).

SUPPLEMENTAL MATERIAL

Figure S1: Hexahistidine site conservation of modern and ancestrally reconstructed A8s and A9s

Figure S2: Oligomeric state analysis of S100 proteins by SEC-MALS

Figure S3: *S. epidermidis* growth curves in the presence of S100 proteins

File S1: Alignment of S100 proteins used in this study

Table S1: Average posterior probabilities for ancestrally reconstructed proteins

File S2: S100 protein tree used for ASR

Figure S5: CD of ancestors

Figure S6: Validation of polymyxin B treatment for endotoxin contamination

Figure S7: Human TLR4 dose curves

Figure S8: Opossum TLR4 dose curves

Figure S9: Extant S100 protein proteolysis fits

Figure S10: Mammalian A8, A9, and A8/A9 complex proteolysis fits

Figure S11: Ancestral S100 proteolysis fits

Figure S12: Mutant S100 proteolysis fits

Figure S13: Sequence alignment highlighting identification of position 63

Figure S14: A9 equilibrium chemical denaturation experiments

Figure S15: A9 kinetic chemical denaturation experiments

Figure S1. The residues composing the A8/A9 hexastidine site are conserved across mammalian and ancestrally reconstructed A8s and A9s. Human (h), mouse (m), opossum (op), maximum likelihood therian mammalian ancestors (anc), and AltAll ancestors (altanc) shown. Alignment truncated to show conservation of key hexahistidine site metal binding residues (boxed + arrows). A8s conserve two (positions 17 and 27 in human A8), while A9s conserve four (positions 91, 95, 103, and 105 of human A9). Consensus residues for alignment are highlighted.

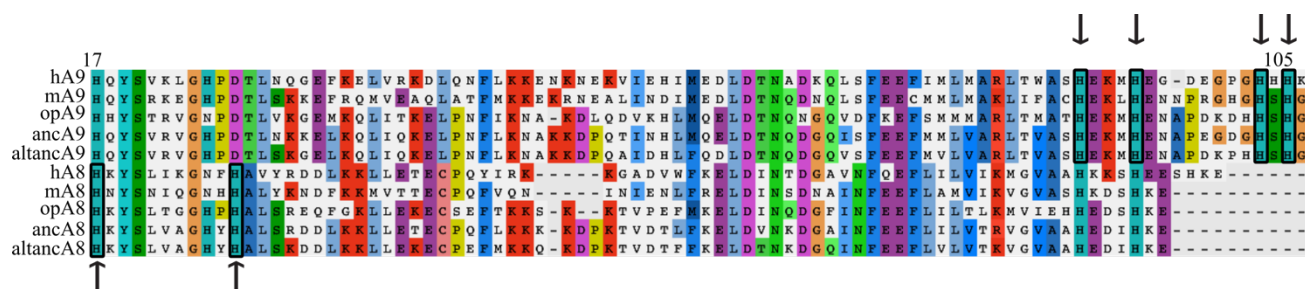
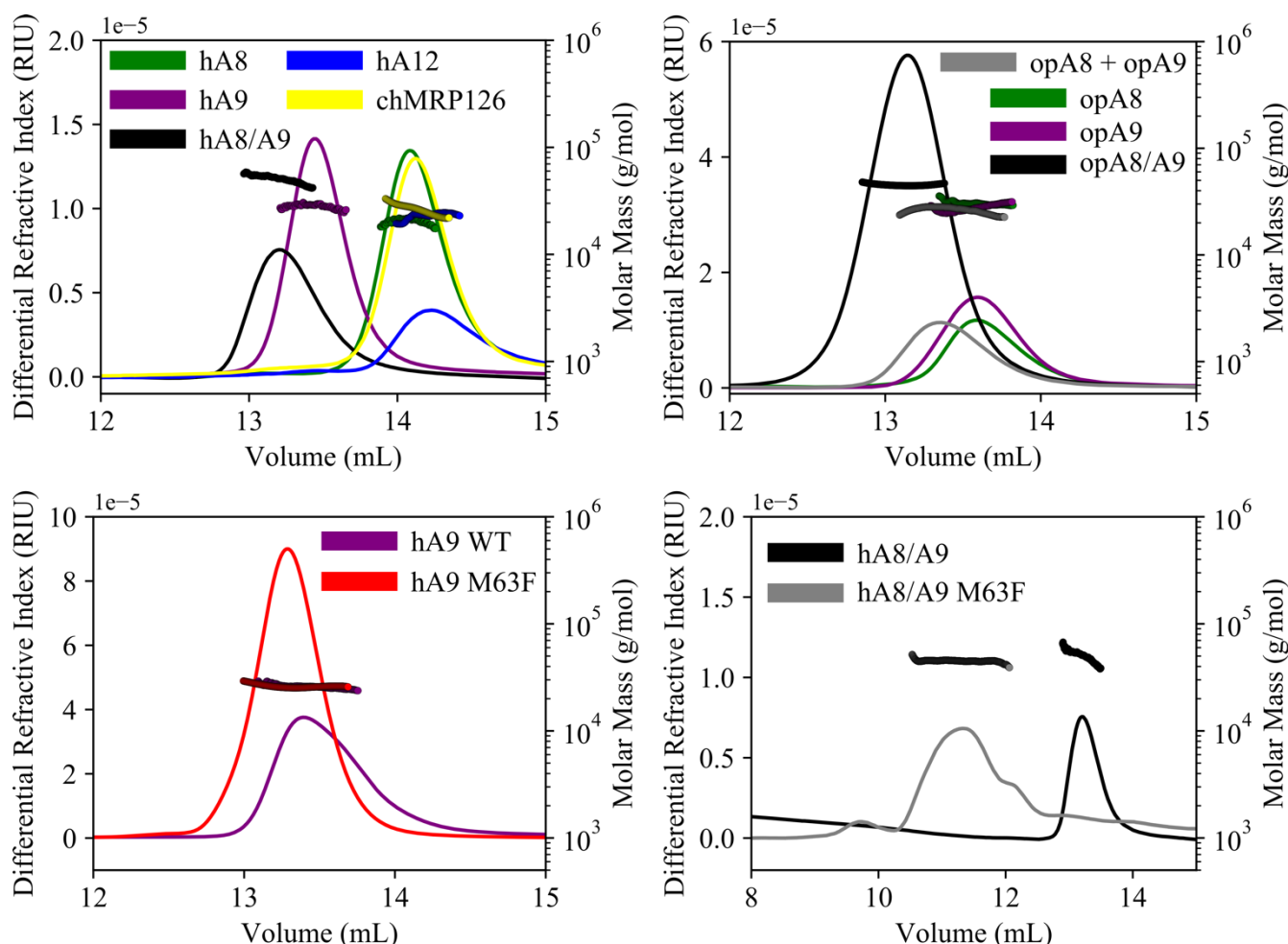


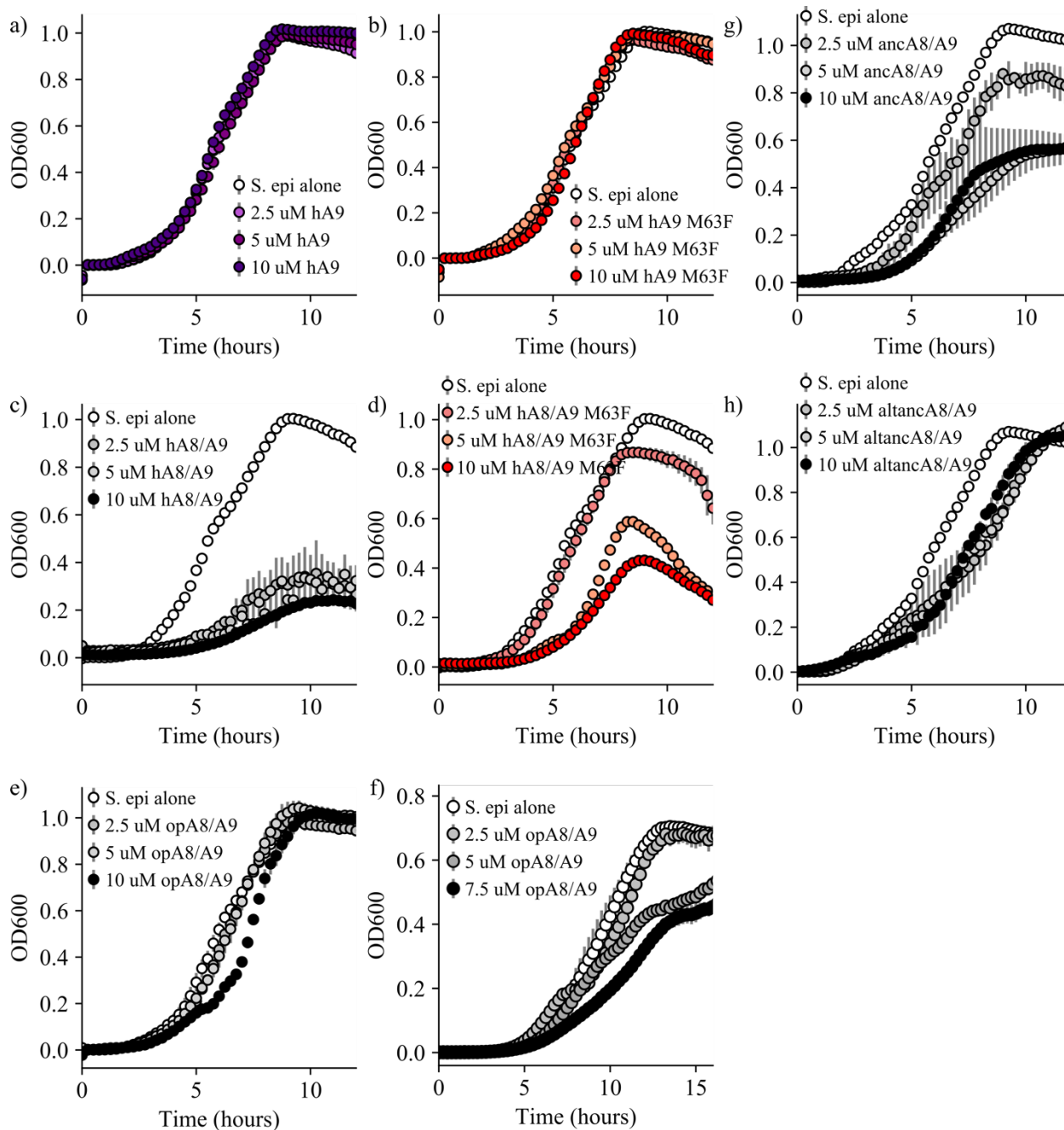
Figure S2. Analysis of protein oligomeric state using SEC-MALS. Differential refractive index (left y-axis, lines) and calculated molecular weights from light scattering detectors (right y-axis, points) for modern S100 proteins used in this study. h = human, op = opossum, ch = chicken species. Opossum A8 + A9 sample is an equimolar mixture of opossum A8 and A9 homodimers. Table below shows summary data calculated using Wyatt Astra software.



Protein	Mw (kDa)	Error	Polydispersity (Mw/Mn)	Error	Mass fraction (%)
hA9	26.5	2.00%	1.004	2.74%	96.5
hA9 M63F	27.1	8.20%	1.058	9.24%	90
hA8/A9	47.8	4.80%	1.046	7.28%	94.6
hA8/A9 M63F	48.7	8.70%	1.049	9.77%	88.7
hA8	18.5	7.40%	1.073	10.60%	89.6
hA9	27.4	6.10%	1.061	7.42%	91.6
chMRP126	28.1	15.60%	1.056	18.97%	84.9
hA12	23.5	14.20%	1.031	19.19%	80.2
opA8	31.7	9.3%	1.017	12.25%	87.5
opA9	29.7	7.6%	1.019	10.41%	91.9
opA8/A9	46.8	1.4%	1.003	1.93%	92.6
opA8 + opA9	25.7	9.6%	1.012	14.01%	93.2

Figure S3. Bacterial growth in the presence of S100 proteins.

Representative *S. epidermidis* growth curves in the presence of a) human A9, b) human A9 M63F, c) human A8/A9, d) human A8/A9 M63F, e) opossum A8/A9 (cysteine-free), f) opossum A8/A9 (containing cysteines), g) ancA8/A9, and h) altancA8/A9. Error bars are the standard deviation for 3 technical replicates, points show one representative biological replicate for each protein at four different concentrations.



1047 **File S1. Alignment of modern and ancestrally reconstructed S100 proteins used in this study.**

1048 >hA9_protein

1049 -----

1050 MTCKMSQLERNIETIINTFHQYSVKLGHPDTLNQGEFKELVRKDLQNFLKKENKNEKVIEHIM
1051 EDLDTNADKQLSFEEFIMLMARLTWASHEKMHEG--DEGPGHHHKPGLGEGTP-----

1052 >mA9_protein

1053 -----

1054 MANKAPSQMERSITTIIDTFHQYSRKEGHPDTLSKKEFRQMVEAQLATFMKKEKRNEALINDI
1055 MEDLDTNQDNQLSFEECMMLMAKLIFACHEKLHENN-PRGHGHSHGKGCGK-----

1056 >opA9_protein

1057 -----MENCTMEKALDIIVNTFHHYSTRVGNPDTLVKGEMKQLITKELPNFIKNA-
1058 KDLQDVKHLMQELDTNQNQGVDFKEFSMMMARLTMATHEKMHENA-
1059 PDKDHSHSGPGLEGKGGSSCGSGHGHSH

1060 >ancA9_protein

1061 -----

1062 MSQLEKAIETIINVFHQYSVRVGHDPDTLNKKELKQLIQKELPNFLKNAKKDPQTINHLMQELDT
1063 NQDGQISFEEFMMLVARLTVASHEKMHENA-PEGDGHSHGPGLGGG-----

1064 >altancA9_protein

1065 -----

1066 MSELEKSIETIINVFHQYSVRVGHDPDTLSKGELKQLIQKELPNFLKNAKKDPQAIDHLFQDLDT
1067 NQDGQVSFEEFMVLVARLTVASHEKMHENA-PDKPHSHSGPGLEEK-----

1068 >hA8_protein

1069 -----MLTELEKALNSIIDVYHKYSLIKGNFHAVYRDDLKLLTECPQYIRK----
1070 KGADVWFKELDINTDGAVNFQEFLILVIKMGVAAHKKSHESHKE-----

1071 >mA8_protein

1072 -----MPSELEKALSNIIDVYHNYSNIQGNHHALYKNDFKKMVTTECPQFVQN----
1073 INIENLFRELDINSNAINFEEFLAMVIKVGVAASHKDSHKE-----

1074 >opA8_protein

1075 -----MATKLECAINCLVEVFHKYSLTGGHPHALSREQFGKLEKECSEFTK---
1076 KSKKTVPFEMKELDINQDGFNFEEFLILTLKMVIEHHEDSHKE-----

1077 >ancA8_protein

1078 -----MTELEKAINSLIDVFHKYSLVAGHYHALSRDDLKLLTECPQFLKKK-
1079 KDPKTVDTLKFELDVNKGAINFEEFLILVTRVGVAAHEDIHKE-----

1080 >altancA8_protein

1081 -----MTELEKAINSLIDVFHKYSLVAGHYHALSKDDLKLLLEKECPEFMKKQ-
1082 KDPKTVDTFFKELDTNKGQINFEEFLVLVTKVGVA AHEDIHKE-----
1083 >hA12_protein
1084 -----MTKLEEHLEGIVNIFHQYSVRKGFDTLSKGELKQLLTKELANTIKNI-
1085 KDKAVIDEIFQGLDANQDEQVDFQEFISLVAIALKAAHYHTHKE-----
1086 >chMRP126_protein
1087 MSKGCQTQGPLSELEKAIDVIIDVFHQYSRREGDKDTLTRKELKLLIEKQLANYLKHV-
1088 KNQVSIDQIFKDLNNDQQLSFGEVMLLIIRVTVATHEHLHFCEDHQQQHQQHQQHNH--
1089 -----
1090 >ancCG_protein
1091 -----MSELEKAIETIINVFHQYSVRVGHPDTLSKKELKQLIQKELPNFLKNA-
1092 KDPATIDNLFQELDKNKDGQISFEEFMVLVARVTVACHEMHKE-----
1093 >altancCG_protein
1094 -----MSQLEKSIETIINVFHQYSVRVGHPDTLSKGELKQLIQKELPNFLKNA-
1095 KDPATIDNLFQDLDTNKGQVSFEEFMVLVARVTVASHEHIHKE-----
1096
1097
1098
1099

Table S1. Statistics for ancestrally reconstructed proteins.

Ancestrally reconstructed protein	Average posterior probability	# of sites that differ from the ML sequence	Average # sites that differ from ML sequence for reconstructions on alternate topologies
ancA9	0.83	--	5
altancA9	0.81	10	--
ancA8	0.88	--	1
altancA8	0.86	17	--
ancCG	0.86	--	8
altancCG	0.84	8	--

1108 **File S2: Species-corrected S100 tree used for ancestral sequence reconstruction.**

1109 TREE #1 with branch lengths:

1110 (((S100B-Elephant_00000774: 0.011851, (((S100B-Rat_00000790: 0.011435, S100B-
1111 Mouse_00000792: 0.000001): 0.000053, S100B-GuineaPig_00000756: 0.023169): 0.011662, S100B-
1112 Rabbit_00000772: 0.021407): 0.000104, ((S100B-Cow_00000764: 0.011306, S100B-Pig_00000782:
1113 0.011385): 0.012895, (S100B-Dog_00000788: 0.257680, S100B-Cat_00000770: 0.010143):
1114 0.000134): 0.000064): 0.012116): 0.094551, (S100B-Tasmaniandevil_00000794: 0.047966, S100B-
1115 Platypus_00000796: 0.156880): 0.056624): 0.023509, ((S100B-Turkey_00000800: 0.021217, S100B-
1116 Duck_00000808: 0.035275): 0.000090, S100B-Flycatcher_00000802: 0.012419): 0.018486,
1117 (((S100G-Opossum_00000898: 0.161245, S100G-Tasmaniandevil_00000900: 0.073097): 0.072758,
1118 (((S100G-Cow_00000862: 0.075218, S100G-Pig_00000864: 0.075304): 0.013547, (S100G-
1119 Cat_00000872: 0.028672, S100G-Dog_00000868: 0.058022): 0.000130): 0.013671, (((S100G-
1120 GuineaPig_00000876: 0.160619, (S100G-Rat_00000880: 0.012341, S100G-Mouse_00000886:
1121 0.057520): 0.116953): 0.074398, S100G-Rabbit_00000878: 0.108386): 0.013349, S100G-
1122 Human_00000827: 0.044740): 0.000068): 0.007312, S100G-Elephant_00000892: 0.168194):
1123 0.027686): 0.539077, (((((MRP-126-Greenseaturtle_00002085: 0.253840, (((S100A8-
1124 Turkey_00000478: 0.072090, S100A8-Chicken_00000476: 0.000020): 0.000264, (S100A8-
1125 Duck_00000480: 0.109166, MRP-126-Goose_00002001: 0.010919): 0.041843): 0.074079, (MRP-126-
1126 EmperorPenguin_00002077: 0.186574, (MRP-126-BarnOwl_00002030: 0.037136, MRP-126-
1127 Eagle_00002015: 0.035591): 0.010190, ((S100A8-Flycatcher_00000474: 0.020194, MRP-126-
1128 Zebrafinch_00002029: 0.031209): 0.093969, MRP-126-Budgerigar_00002048: 0.110229): 0.000062):
1129 0.000087): 0.105688): 0.300170, (MRP-126-Alligator_00002081: 0.063287, MRP-126-
1130 Saltwatercrocodyle_00002082: 0.079422): 0.125231): 0.081595): 0.076483, ((S100A12-
1131 Elephant_00000364: 0.375852, ((S100A12-Rabbit_00000366: 0.208679, (S100A12-
1132 Human_00000307: 0.000008, (S100A12-Gorilla_00000308: 0.011294, S100A12-
1133 Chimpanzee_00000310: 0.011316): 0.011289): 0.267716): 0.095317, ((S100A12-Cow_00000356:
1134 0.132593, (S100A12-Sheep_00000358: 0.108573, S100A12-Pig_00000360: 0.126117): 0.028511):
1135 0.033478, (S100A12-Cat_00000370: 0.142638, S100A12-Dog_00000338: 0.257857): 0.050457):
1136 0.000517): 0.139973): 0.270007, (S100A9-Tasmaniandevil_00000558: 0.376463, S100A9-
1137 Opossum_00000560: 0.457951): 0.518003): 0.002864): 0.051883, (S100A8-Platypus_00000472:
1138 0.118845, (((((S100A8-Human_00000391: 0.000001, S100A8-Gorilla_00000392: 0.023248):
1139 0.176912, (S100A8-Rabbit_00000416: 0.187531, ((S100A8-Rat_00000424: 0.038358, S100A8-
1140 Mouse_00000428: 0.199436): 0.355484, S100A8-GuineaPig_00000436: 0.297343): 0.025972):
1141 0.016180): 0.090034, ((S100A8-Cat_00000442: 0.071389, S100A8-Dog_00000452: 0.053183):
1142 0.000277, (S100A8-Pig_00000456: 0.103224, (S100A8-Cow_00000438: 0.057455, S100A8-
1143 Sheep_00000448: 0.041907): 0.077606): 0.090118): 0.012641): 0.025245, S100A8-
1144 Elephant_00000462: 0.299468): 0.253959, (S100A8-Tasmaniandevil_00000468: 0.460050, S100A8-
1145 Opossum_00000470: 0.441062): 0.259349): 0.257176): 0.377056): 0.100813, ((S100A12-
1146 Opossum_00000380: 0.171391, S100A12-Tasmaniandevil_00000382: 0.287175): 0.326228,
1147 (((((S100A9-Sheep_00000534: 0.128256, S100A9-Cow_00000538: 0.045919): 0.046648, S100A9-
1148 Pig_00000542: 0.131808): 0.046797, (S100A9-Cat_00000540: 0.121027, S100A9-Dog_00000544:
1149 0.023205): 0.114595): 0.090769, (((S100A9-GuineaPig_00000508: 0.315010, (S100A9-
1150 Mouse_00000514: 0.138646, S100A9-Rat_00000516: 0.120216): 0.182176): 0.113753, S100A9-
1151 Rabbit_00000506: 0.303617): 0.026172, (S100A9-Human_00000483: 0.009750, S100A9-

1152 Chimpanzee_00000486: 0.000001): 0.194039): 0.024606): 0.000327, S100A9-Elephant_00000526:
 1153 0.263248): 0.273405): 0.171144): 0.119685, ((((((S100A7-Eagle_00003007: 0.111865, S100A7-
 1154 BarnOwl_00003010: 0.091265): 0.036718, (S100A7-Zebrafinch_00003027: 0.345523, S100A7-
 1155 Budgerigar_00003002: 0.094486): 0.031823): 0.000200, S100A7-EmperorPenguin_00003023:
 1156 0.158408): 0.000450, ((S100A7-Goose_00003015: 0.025054, S100A7-Duck_00003021: 0.070377):
 1157 0.176006, S100A12-Turkey_00000386: 0.306748): 0.086695): 0.200304, (S100A7-
 1158 Saltwatercrocodile_00003038: 0.208248, S100A7-Alligator_00003036: 0.149850): 0.259960):
 1159 0.646359, ((((((S100A7-Cow_00000594: 0.063900, S100A7-Sheep_00000590: 0.126150): 0.166292,
 1160 S100A7-Pig_00000608: 0.358581): 0.006962, (((S100A7-Chimpanzee_00000570: 0.022260,
 1161 (S100A7-Human_00000569: 0.022139, (S100A7A-Chimpanzee_00000632: 0.010951, (S100A7A-
 1162 Human_00000631: 0.045504, S100A7A-Gorilla_00000634: 0.010950): 0.000001): 0.034351):
 1163 0.000001): 0.161014, (S100A7L2-Human_00000701: 0.000076, S100A7L2-Chimpanzee_00000702:
 1164 0.011065): 0.571687): 0.109920, S100A7-Rabbit_00000622: 1.129756): 0.000404): 0.104943,
 1165 S100A7L2-Elephant_00000722: 0.365579): 0.838659, (S100A7-Tasmaniandevil_00000628: 0.070559,
 1166 S100A7-Opossum_00000630: 0.031288): 0.685999): 0.127986): 0.444082): 0.054121): 0.183817,
 1167 (((((((CRNN-Pig_00000046: 0.021872, (CRNN-Sheep_00000058: 0.000012, CRNN-Cow_00000062:
 1168 0.012337): 0.049611): 0.053537, (CRNN-Cat_00000052: 0.062868, CRNN-Dog_00000044:
 1169 0.098552): 0.000113): 0.000129, (((CRNN-Mouse_00000030: 0.011248, CRNN-Rat_00000036:
 1170 0.037055): 0.104929, CRNN-GuineaPig_00000024: 0.081970): 0.017533, CRNN-Rabbit_00000038:
 1171 0.101269): 0.000147, (CRNN-Chimpanzee_00000002: 0.023821, (CRNN-Human_00000001:
 1172 0.000001, CRNN-Gorilla_00000008: 0.000001): 0.000001): 0.062164): 0.000053): 0.013720, CRNN-
 1173 Elephant_00000066: 0.046485): 0.164013, (CRNN-Opossum_00000072: 0.021379, CRNN-
 1174 Tasmaniandevil_00000074: 0.039367): 0.053794): 0.344728, (CRNN-Duck_00000080: 0.077283,
 1175 (CRNN-Turkey_00000076: 0.074076, CRNN-Chicken_00000078: 0.096950): 0.172464): 0.198220):
 1176 0.324597, (((((HRNR-Cat_00000210: 0.026197, HRNR-Pig_00000206: 0.099361): 0.028399,
 1177 ((HRNR-Gorilla_00000182: 0.011289, (HRNR-Human_00000179: 0.000001, HRNR-
 1178 Chimpanzee_00000180: 0.000001): 0.000001): 0.095248, (HRNR-GuineaPig_00000196: 0.117870,
 1179 (HRNR-Rat_00000192: 0.071187, HRNR-Mouse_00000194: 0.048424): 0.129632): 0.017313):
 1180 0.020017): 0.267241, ((FLG-Platypus_00000108: 0.166845, (((((FLG2-Rabbit_00000132: 0.061346,
 1181 ((FLG2-Rat_00000138: 0.012063, FLG2-Mouse_00000140: 0.036027): 0.067101, FLG2-
 1182 GuineaPig_00000150: 0.145605): 0.008646): 0.010452, (FLG2-Human_00000117: 0.023588, FLG2-
 1183 Chimpanzee_00000122: 0.000001): 0.048487): 0.000136, (((FLG2-Sheep_00000134: 0.011315,
 1184 FLG2-Cow_00000136: 0.012165): 0.023989, FLG2-Pig_00000146: 0.011764): 0.000110, FLG2-
 1185 Cat_00000148: 0.011761): 0.000095): 0.000444, FLG2-Elephant_00000160: 0.059260): 0.164766,
 1186 (FLG-Opossum_00000106: 0.461181, FLG2-Tasmaniandevil_00000168: 0.244192): 0.001459):
 1187 0.014822): 0.012030, (((((FLG-GuineaPig_00000092: 0.104274, FLG-Mouse_00000090: 0.549526):
 1188 0.031819, (FLG-Chimpanzee_00000084: 0.000001, (FLG-Human_00000081: 0.000001, FLG-
 1189 Gorilla_00000082: 0.000001): 0.000001): 0.146515): 0.000938, FLG-Cat_00000096: 0.213312):
 1190 0.000290, FLG-Elephant_00000102: 0.221474): 0.297499, (FLG2-Tasmaniandevil_00000162:
 1191 0.012310, FLG2-Opossum_00000164: 0.119507): 0.235223): 0.074421): 0.030743): 0.328552,
 1192 (RPTN-Platypus_00000298: 0.436291, (((((RPTN-Gorilla_00000234: 0.014272, (RPTN-
 1193 Human_00000233: 0.000001, RPTN-Chimpanzee_00000236: 0.000001): 0.009982): 0.132650,
 1194 (RPTN-Mouse_00000252: 0.061911, RPTN-Rat_00000254: 0.000001): 0.228536): 0.034794,
 1195 (((RPTN-Cow_00000258: 0.023455, RPTN-Sheep_00000288: 0.012142): 0.109764, RPTN-
 1196 Pig_00000286: 0.045747): 0.011461, (RPTN-Dog_00000260: 0.057191, RPTN-Cat_00000262:

1197 0.067589): 0.000130): 0.009278): 0.000254, RPTN-Elephant_00000266: 0.124335): 0.219785,
1198 (RPTN-Tasmaniandevil_00000294: 0.186345, RPTN-Opossum_00000296: 0.071637): 0.075981):
1199 0.106617): 0.084801): 0.231034, (FLG-AnoleLizard_00000110: 0.603813, (FLG-Turkey_00000112:
1200 0.087944, FLG-Flycatcher_00000116: 0.342666): 0.341688): 0.036398): 0.000240): 0.498390):
1201 0.757090);

1202 TREE #1 as a cladogram:

1203 (((S100B-Elephant_00000774, (((S100B-Rat_00000790, S100B-Mouse_00000792) 177 , S100B-
1204 GuineaPig_00000756) 176 , S100B-Rabbit_00000772) 175 , ((S100B-Cow_00000764, S100B-
1205 Pig_00000782) 179 , (S100B-Dog_00000788, S100B-Cat_00000770) 180) 178) 174) 173 , (S100B-
1206 Tasmaniandevil_00000794, S100B-Platypus_00000796) 181) 172 , ((S100B-Turkey_00000800,
1207 S100B-Duck_00000808) 183 , S100B-Flycatcher_00000802) 182 , (((S100G-Opossum_00000898,
1208 S100G-Tasmaniandevil_00000900) 187 , (((S100G-Cow_00000862, S100G-Pig_00000864) 191 ,
1209 (S100G-Cat_00000872, S100G-Dog_00000868) 192) 190 , (((S100G-GuineaPig_00000876, (S100G-
1210 Rat_00000880, S100G-Mouse_00000886) 196) 195 , S100G-Rabbit_00000878) 194 , S100G-
1211 Human_00000827) 193) 189 , S100G-Elephant_00000892) 188) 186 , (((((MRP-126-
1212 Greenseaturtle_00002085, (((S100A8-Turkey_00000478, S100A8-Chicken_00000476) 205 ,
1213 (S100A8-Duck_00000480, MRP-126-Goose_00002001) 206) 204 , (MRP-126-
1214 EmperorPenguin_00002077, ((MRP-126-BarnOwl_00002030, MRP-126-Eagle_00002015) 209 ,
1215 ((S100A8-Flycatcher_00000474, MRP-126-Zebrafinch_00002029) 211 , MRP-126-
1216 Budgerigar_00002048) 210) 208) 207) 203 , (MRP-126-Alligator_00002081, MRP-126-
1217 Saltwatercrocodile_00002082) 212) 202) 201 , ((S100A12-Elephant_00000364, ((S100A12-
1218 Rabbit_00000366, (S100A12-Human_00000307, (S100A12-Gorilla_00000308, S100A12-
1219 Chimpanzee_00000310) 218) 217) 216 , ((S100A12-Cow_00000356, (S100A12-Sheep_00000358,
1220 S100A12-Pig_00000360) 221) 220 , (S100A12-Cat_00000370, S100A12-Dog_00000338) 222) 219)
1221 215) 214 , (S100A9-Tasmaniandevil_00000558, S100A9-Opossum_00000560) 223) 213) 200 ,
1222 (S100A8-Platypus_00000472, (((((S100A8-Human_00000391, S100A8-Gorilla_00000392) 229 ,
1223 (S100A8-Rabbit_00000416, ((S100A8-Rat_00000424, S100A8-Mouse_00000428) 232 , S100A8-
1224 GuineaPig_00000436) 231) 230) 228 , ((S100A8-Cat_00000442, S100A8-Dog_00000452) 234 ,
1225 (S100A8-Pig_00000456, (S100A8-Cow_00000438, S100A8-Sheep_00000448) 236) 235) 233) 227 ,
1226 S100A8-Elephant_00000462) 226 , (S100A8-Tasmaniandevil_00000468, S100A8-
1227 Opossum_00000470) 237) 225) 224) 199 , ((S100A12-Opossum_00000380, S100A12-
1228 Tasmaniandevil_00000382) 239 , (((((S100A9-Sheep_00000534, S100A9-Cow_00000538) 244 ,
1229 S100A9-Pig_00000542) 243 , (S100A9-Cat_00000540, S100A9-Dog_00000544) 245) 242 ,
1230 (((S100A9-GuineaPig_00000508, (S100A9-Mouse_00000514, S100A9-Rat_00000516) 249) 248 ,
1231 S100A9-Rabbit_00000506) 247 , (S100A9-Human_00000483, S100A9-Chimpanzee_00000486) 250)
1232 246) 241 , S100A9-Elephant_00000526) 240) 238) 198 , ((((((S100A7-Eagle_00003007, S100A7-
1233 BarnOwl_00003010) 256 , (S100A7-Zebrafinch_00003027, S100A7-Budgerigar_00003002) 257) 255
1234 , S100A7-EmperorPenguin_00003023) 254 , ((S100A7-Goose_00003015, S100A7-Duck_00003021)
1235 259 , S100A12-Turkey_00000386) 258) 253 , (S100A7-Saltwatercrocodile_00003038, S100A7-
1236 Alligator_00003036) 260) 252 , ((((((S100A7-Cow_00000594, S100A7-Sheep_00000590) 265 ,
1237 S100A7-Pig_00000608) 264 , (((S100A7-Chimpanzee_00000570, (S100A7-Human_00000569,
1238 (S100A7A-Chimpanzee_00000632, (S100A7A-Human_00000631, S100A7A-Gorilla_00000634) 271
1239) 270) 269) 268 , (S100A7L2-Human_00000701, S100A7L2-Chimpanzee_00000702) 272) 267 ,
1240 S100A7-Rabbit_00000622) 266) 263 , S100A7L2-Elephant_00000722) 262 , (S100A7-

1241 Tasmaniandevil_00000628, S100A7-Opossum_00000630) 273) 261) 251) 197) 185 , (((((((CRNN-
1242 Pig_00000046, (CRNN-Sheep_00000058, CRNN-Cow_00000062) 281) 280 , (CRNN-Cat_00000052,
1243 CRNN-Dog_00000044) 282) 279 , (((CRNN-Mouse_00000030, CRNN-Rat_00000036) 286 ,
1244 CRNN-GuineaPig_00000024) 285 , CRNN-Rabbit_00000038) 284 , (CRNN-Chimpanzee_00000002,
1245 (CRNN-Human_00000001, CRNN-Gorilla_00000008) 288) 287) 283) 278 , CRNN-
1246 Elephant_00000066) 277 , (CRNN-Opossum_00000072, CRNN-Tasmaniandevil_00000074) 289) 276
1247 , (CRNN-Duck_00000080, (CRNN-Turkey_00000076, CRNN-Chicken_00000078) 291) 290) 275 ,
1248 (((((HRNR-Cat_00000210, HRNR-Pig_00000206) 296 , ((HRNR-Gorilla_00000182, (HRNR-
1249 Human_00000179, HRNR-Chimpanzee_00000180) 299) 298 , (HRNR-GuineaPig_00000196,
1250 (HRNR-Rat_00000192, HRNR-Mouse_00000194) 301) 300) 297) 295 , ((FLG-Platypus_00000108,
1251 (((((FLG2-Rabbit_00000132, ((FLG2-Rat_00000138, FLG2-Mouse_00000140) 310 , FLG2-
1252 GuineaPig_00000150) 309) 308 , (FLG2-Human_00000117, FLG2-Chimpanzee_00000122) 311) 307
1253 , (((FLG2-Sheep_00000134, FLG2-Cow_00000136) 314 , FLG2-Pig_00000146) 313 , FLG2-
1254 Cat_00000148) 312) 306 , FLG2-Elephant_00000160) 305 , (FLG-Opossum_00000106, FLG2-
1255 Tasmaniandevil_00000168) 315) 304) 303 , (((((FLG-GuineaPig_00000092, FLG-Mouse_00000090)
1256 320 , (FLG-Chimpanzee_00000084, (FLG-Human_00000081, FLG-Gorilla_00000082) 322) 321)
1257 319 , FLG-Cat_00000096) 318 , FLG-Elephant_00000102) 317 , (FLG2-Tasmaniandevil_00000162,
1258 FLG2-Opossum_00000164) 323) 316) 302) 294 , (RPTN-Platypus_00000298, (((((RPTN-
1259 Gorilla_00000234, (RPTN-Human_00000233, RPTN-Chimpanzee_00000236) 330) 329 , (RPTN-
1260 Mouse_00000252, RPTN-Rat_00000254) 331) 328 , (((RPTN-Cow_00000258, RPTN-
1261 Sheep_00000288) 334 , RPTN-Pig_00000286) 333 , (RPTN-Dog_00000260, RPTN-Cat_00000262)
1262 335) 332) 327 , RPTN-Elephant_00000266) 326 , (RPTN-Tasmaniandevil_00000294, RPTN-
1263 Opossum_00000296) 336) 325) 324) 293 , (FLG-AnoleLizard_00000110, (FLG-Turkey_00000112,
1264 FLG-Flycatcher_00000116) 338) 337) 292) 274) 184) 171 ;

1265

1266

1267

1268

1269

1270

1271

1272

1273

1274

1275

1276

1277

1278

Figure S5. Secondary structure characterization of ancestrally reconstructed S100s by CD spectroscopy. Data shown are the average of 3 scans. Solid lines are maximum likelihood ancestral proteins, dotted lines are alt-all ancestors (colored the same as matching maximum likelihood ancestor for comparison).

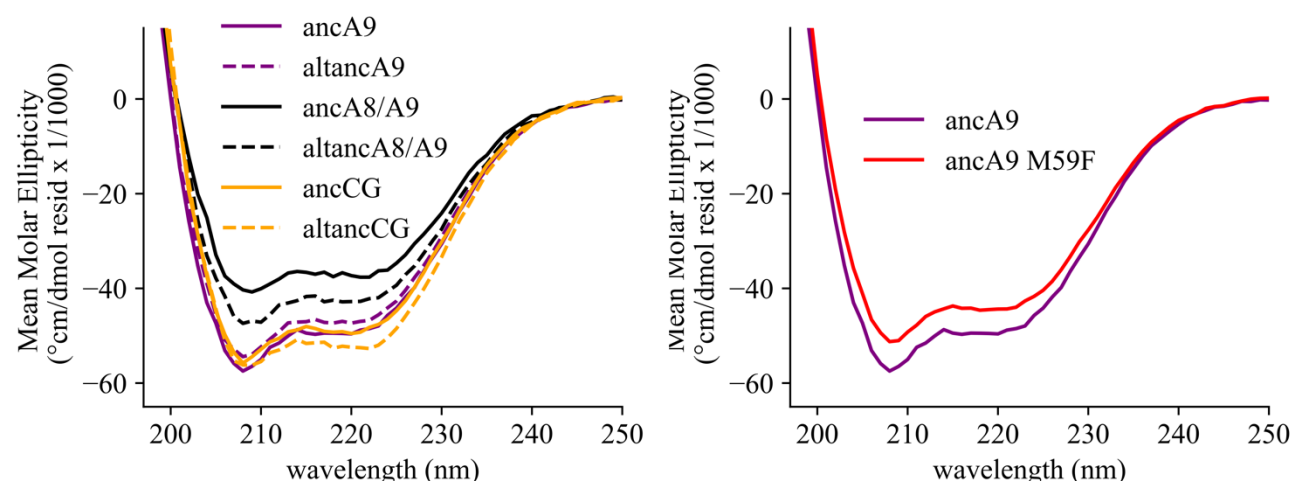


Figure S6. Analysis of S100 protein LPS contamination. Top panel: Activity of LPS (0.2 ng/ul) against human (left) and opossum TLR4 (right) is inhibited by the addition of Polymixin B (PB, 0.2 ug/ul). Middle and bottom panels: S100 activation of human (middle) and opossum (bottom) TLR4 with no PB, 0.2 ug/ul PB (+) and 0.25 ug/ul PB (++). No-PB data were not collected for hA9, opA9, and opA9 M60F against opossum TLR4 (bottom panel). Data were background-subtracted using the LPS + PB control and normalized to LPS activity against either human or opossum TLR4. Error bars are standard deviation.

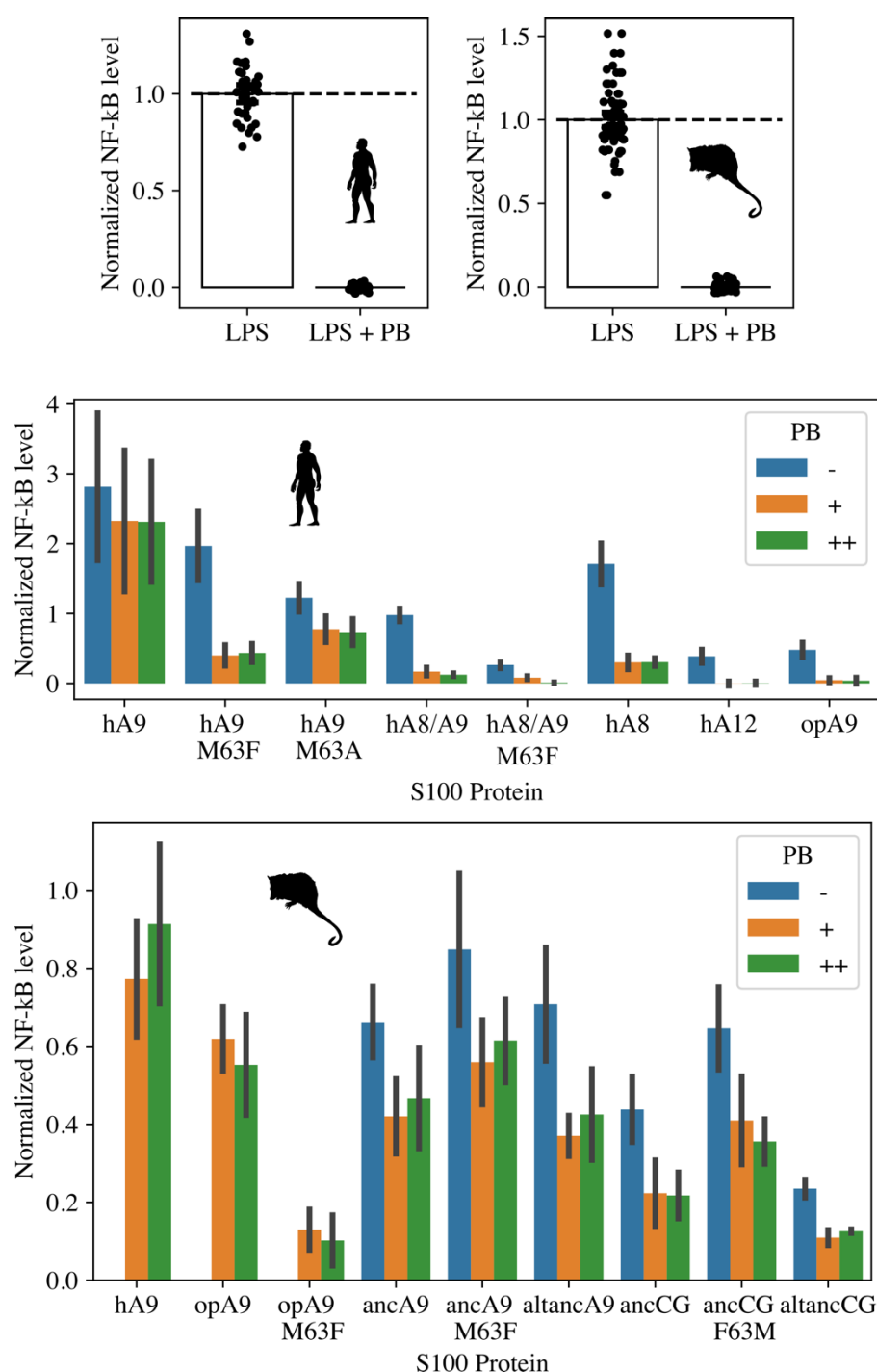


Figure S7. Dose curves for S100 activation of human TLR4. Points are the average of >3 biological replicates each consisting of 3 technical triplicates, error bars are standard error of the mean. An asterisk (*) indicates a concentration at which a single biological replicate was measured. Data were background-subtracted using the LPS + PB control and normalized to LPS activity against human TLR4.

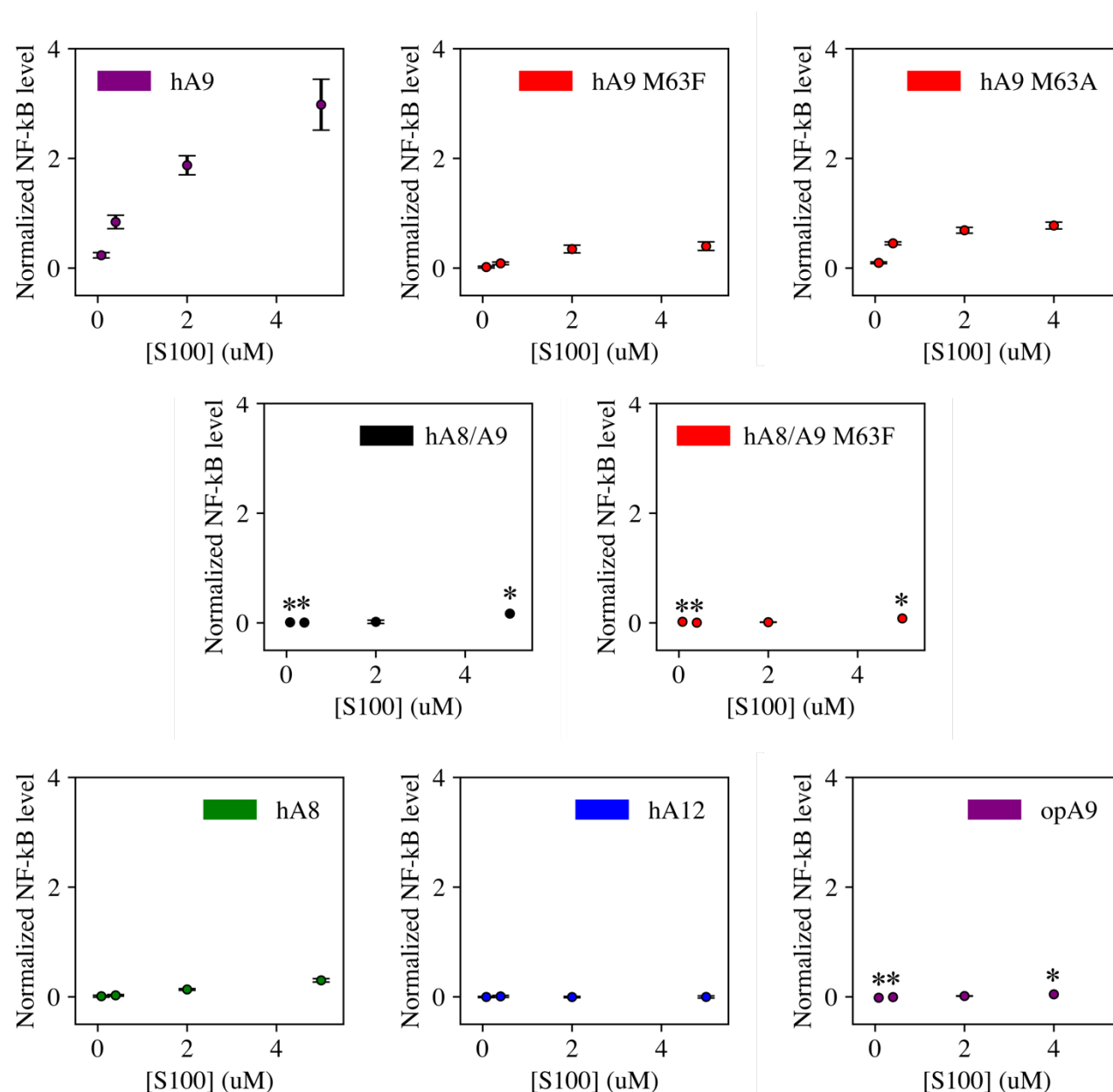


Figure S8. Dose curves for S100 activation of opossum TLR4. Points are the average of >3 biological replicates each consisting of 3 technical triplicates, error bars are standard error of the mean. An asterisk (*) indicates a concentration at which a single biological replicate was measured. Data were background-subtracted using the LPS + PB control and normalized to LPS activity against opossum TLR4.

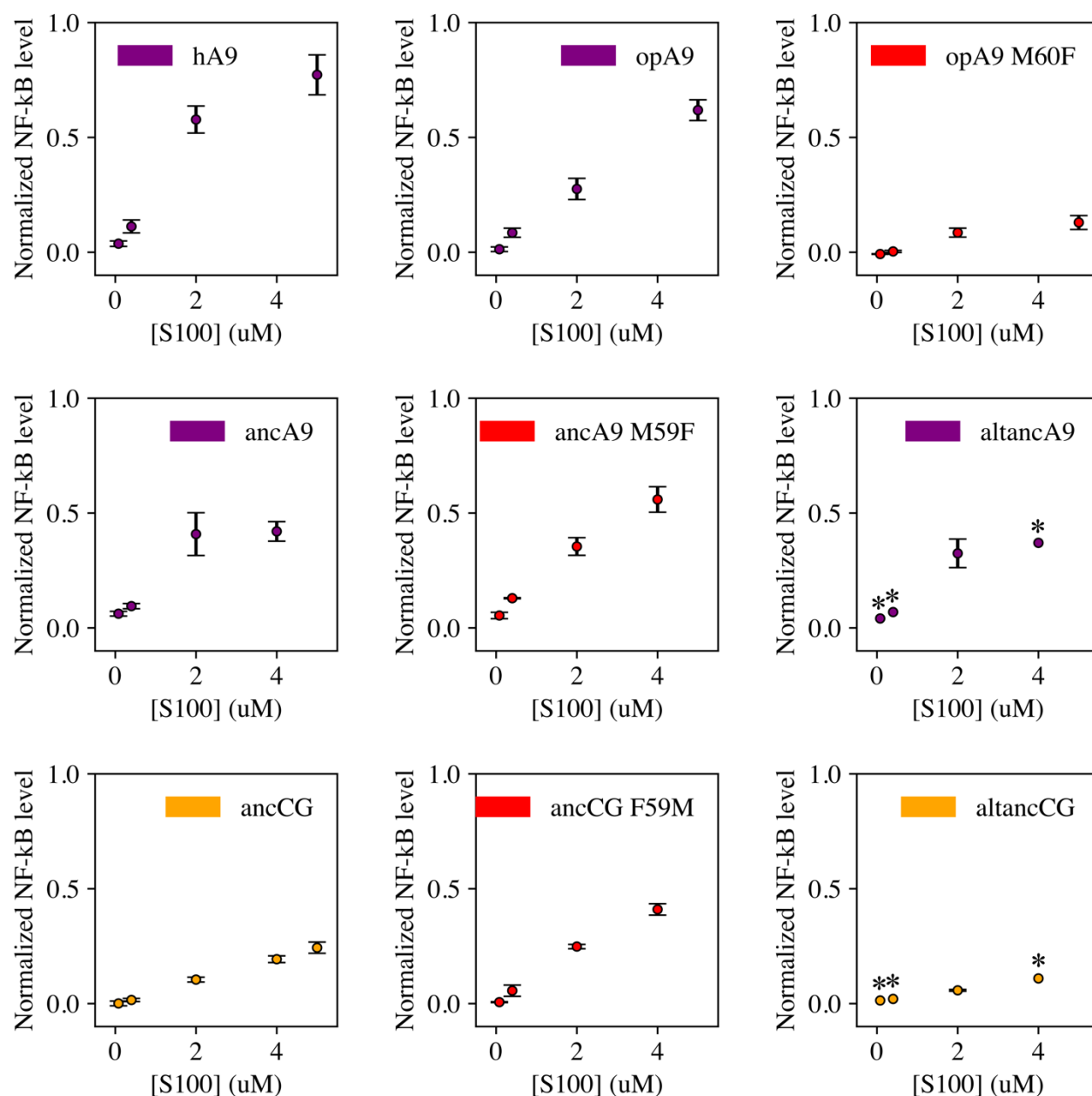


Figure S9. Survey of proteolytic susceptibility across modern S100 proteins. Blue dots are biological replicates, orange line is a single exponential decay fit (see methods). Protein is listed at the top. Pixel intensity was quantified by densitometry from SDS-PAGE gels.

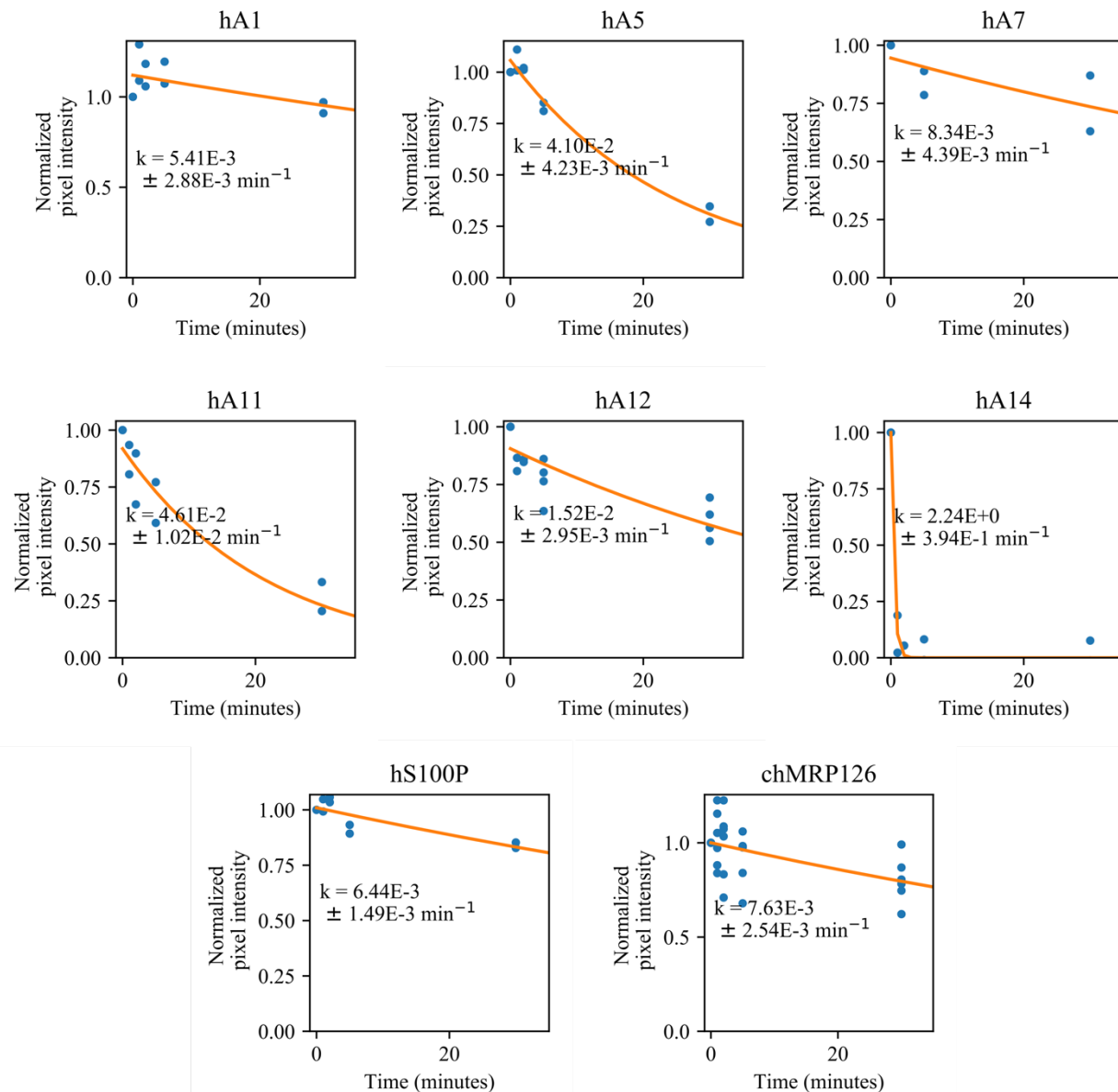


Figure S10. Comparison of proteolytic susceptibility for A8s, A9s, and A8/A9 complexes across mammals. Blue dots are biological replicates, orange line is a single exponential decay fit (see methods). Protein is listed at the top. Pixel intensity was quantified by densitometry from SDS-PAGE gels. Longer time points were collected for proteins with slower degradation rates (see x-axis).

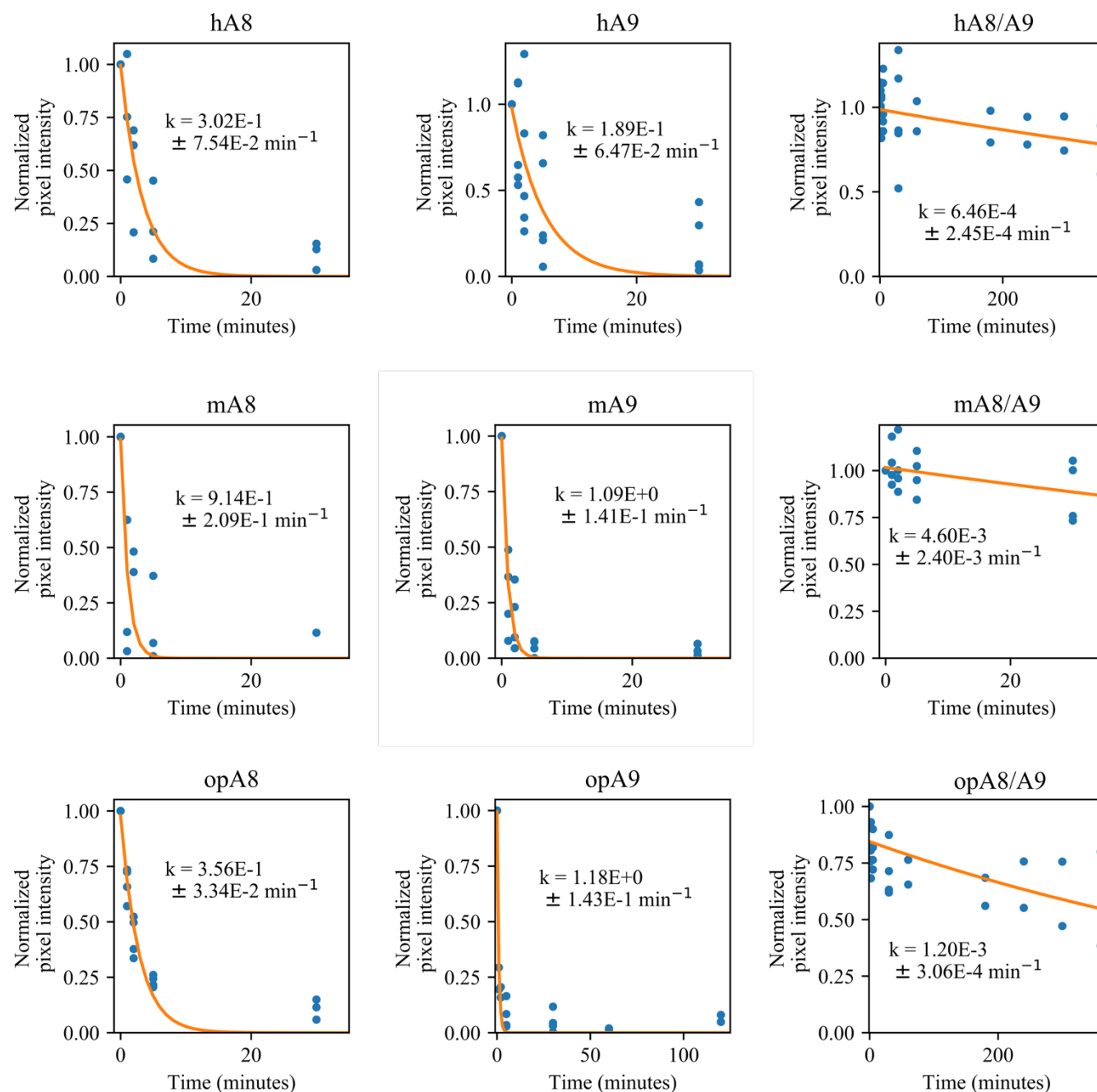


Figure S11. Comparison of proteolytic susceptibility for ancestrally reconstructed S100s. Blue dots are biological replicates, orange line is a single exponential decay fit (see methods). Protein is listed at the top. Pixel intensity was quantified by densitometry from SDS-PAGE gels. Longer time points were collected for proteins with slower degradation rates (see x-axis).

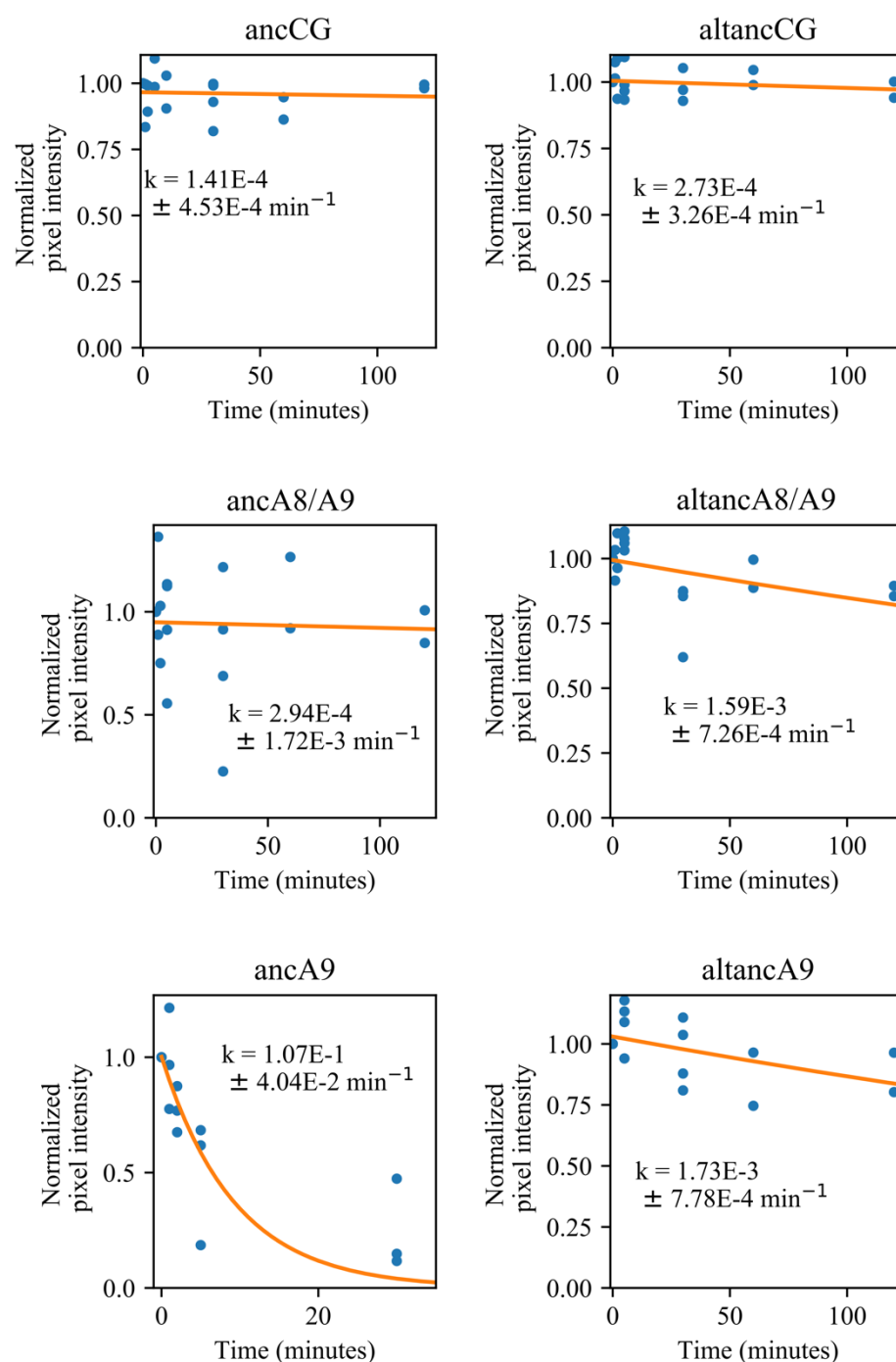


Figure S12. Comparison of proteolytic S100 protein mutants at position 63. Blue dots are biological replicates, orange line is a single exponential decay fit (see methods). Protein is listed at the top. Pixel intensity was quantified by densitometry from SDS-PAGE gels. Longer time points were collected for proteins with slower degradation rates (see x-axis).

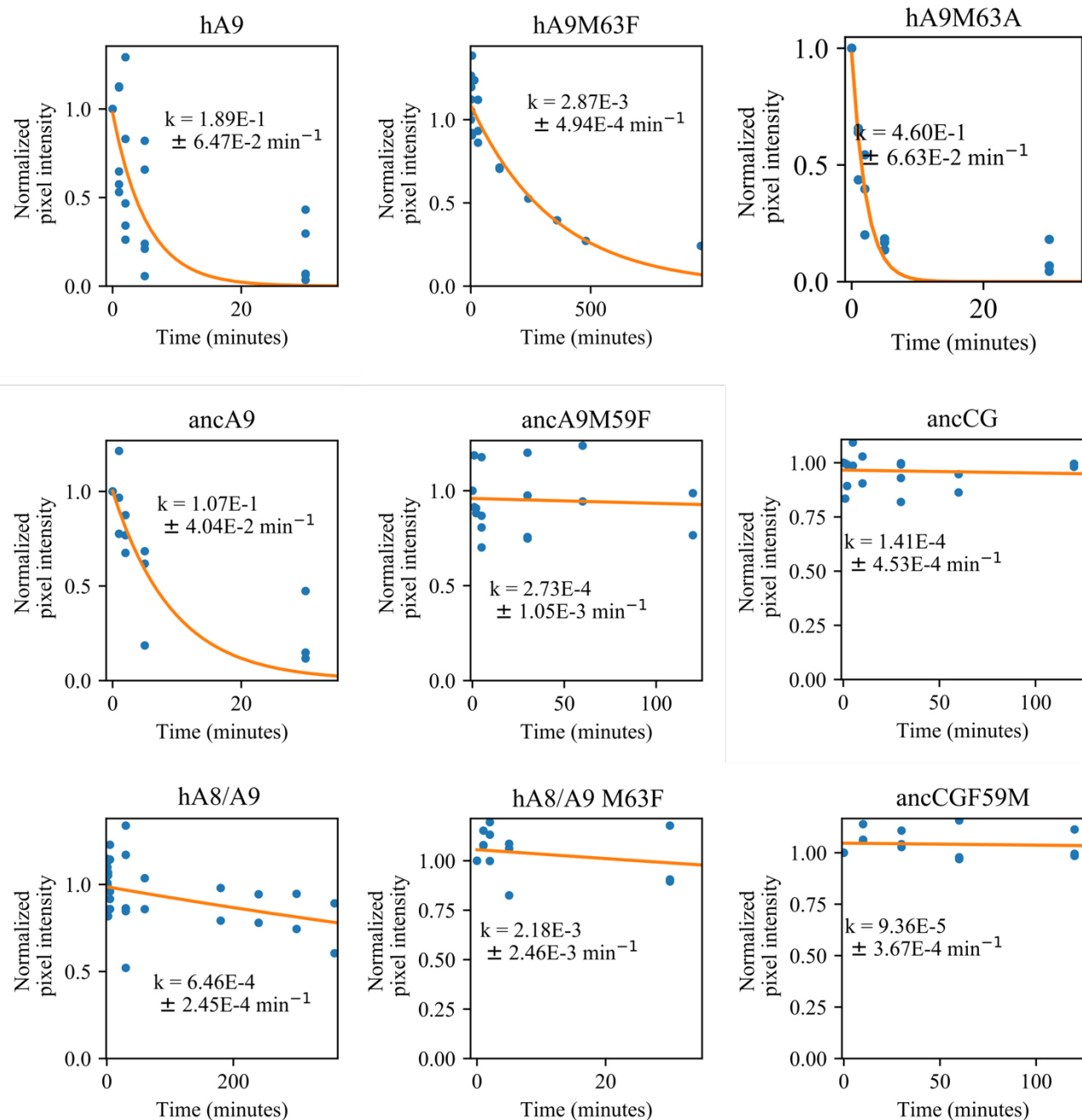


Figure S13. Changes at position 63 correlate with A9 proinflammatory activity and proteolytic resistance. S100 protein sequences are grouped into proteolytically susceptible (top) or resistant and potentially proinflammatory (red text) or not (black text). Only the first 90 residues out of 114 total were examined as the disordered A9 tail (residues ~93-114) are highly variable and the tail is dispensable for A9 proinflammatory activity. Residues are colored when found to be the consensus residue for a column.



Figure S14. Chemical denaturation of hA9 and hA9 M63F. Each graph is a single biological replicate. Filled circles are unfolding, empty circles are refolding. A two-state unfolding model was fit to the data, shown as a solid line. Because A9 is a dimer and could therefore exhibit concentration-dependent folding properties, we verified that the C_m for M63F was higher than A9 at 0.5, 2.0, and 5 μ M protein. Protein concentrations are noted on graphs.

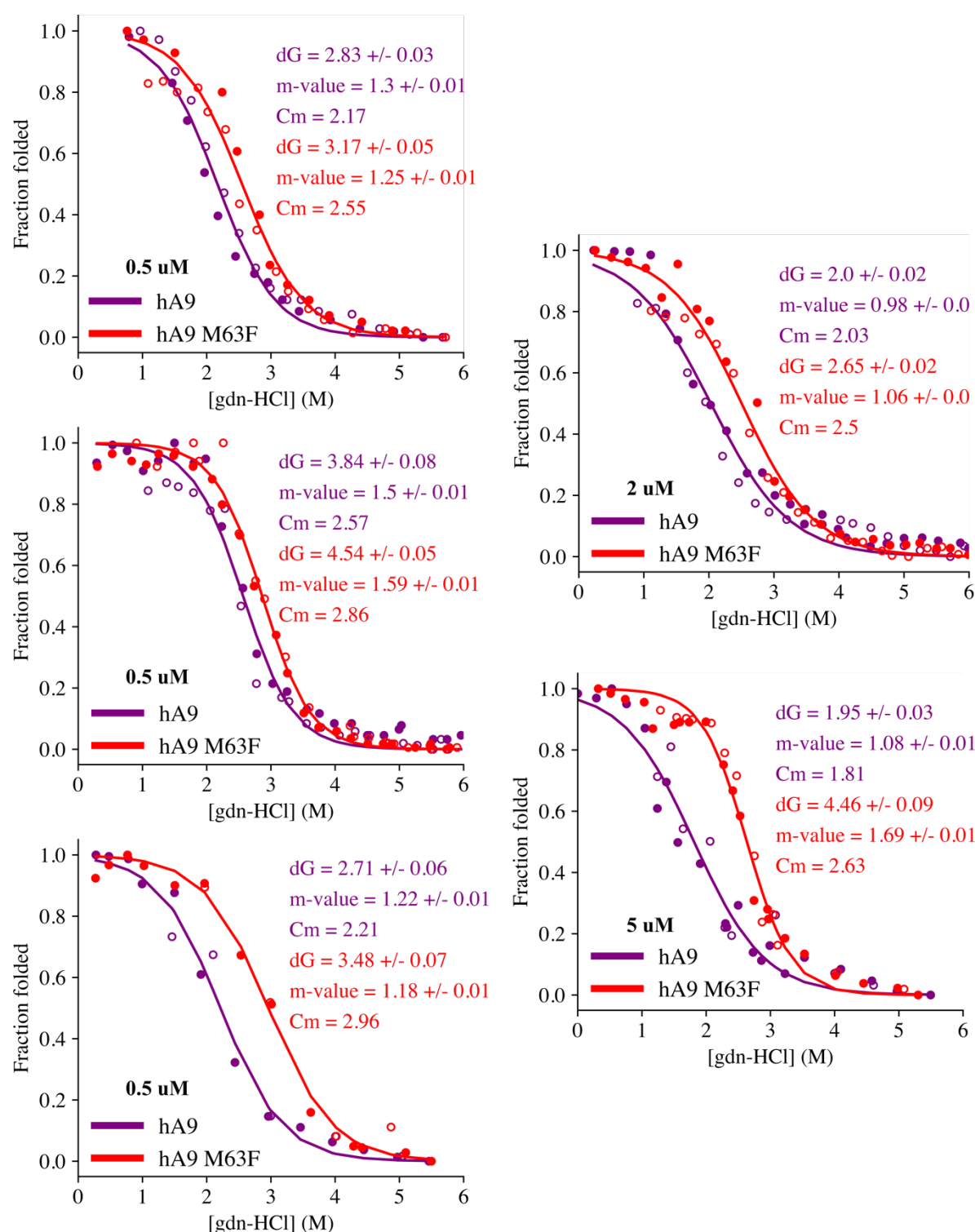


Figure S15. Unfolding kinetics of hA9 and hA9 M63F. Time course measurement of hA9 (top) and hA9 M63F (bottom) unfolding upon addition of 6M gdn-HCl. Each curve is a single replicate at one concentration, monitoring CD signal at 222 nm (y-axis).

