

In silico* identification of novel peptides with antibacterial activity against multidrug resistant *Staphylococcus aureus

Linda B Oyama^{*a}, Hamza Olleik^b, Ana Carolina Nery Teixeira^c, Matheus M Guidini^c, James A Pickup^a, Alan R Cookson^d, Hannah Vallin^d, Toby Wilkinson^e, Denise Bazzolli^c, Jennifer Richards^f, Mandy Wootton^f, Ralf Mikut^g, Kai Hilpert^h, Marc Maresca^b, Josette Perrier^b, Matthias Hessⁱ, Hilario C Mantovani^c, Narcis Fernandez-Fuentes^d, Christopher J Creevey^a, and Sharon A Huws^{*a}

Author Affiliation

^a Institute for Global Food Security, School of Biological Sciences, Medical Biology Centre, Queen's University Belfast, 97 Lisburn Road, Belfast, Northern Ireland, BT9 7BL, UK.

^b Aix Marseille Univ, CNRS, Centrale Marseille, iSm2, Marseille, France.

^c Departamento de Microbiologia, Universidade Federal de Viçosa, Viçosa, 36570-900, Brasil.

^d Institute of Biological Environmental and Rural Sciences, Aberystwyth University, Aberystwyth, Wales, SY23 3DA, UK.

^e The Roslin Institute and R(D)SVS, University of Edinburgh, Edinburgh, United Kingdom.

^f Specialist Antimicrobial Chemotherapy Unit, Public Health Wales, University Hospital of Wales, Heath Park, Cardiff, CF14 4XW, UK.

^g Karlsruhe Institute of Technology, Institute for Automation and Applied Informatics, Hermann-von-Helmholtz-Platz 1, 76344, Eggenstein, Leopoldshafen, Germany.

^h Institute of Infection and Immunity, St. George's University of London, Cranmer Terrace, London SW17 0RE, UK.

ⁱ UC Davis, College of Agricultural and Environmental Sciences, California 95616, USA.

*** Correspondence authors.** Dr Linda B Oyama (l.oyama@qub.ac.uk) and Dr Sharon A Huws (s.huws@qub.ac.uk). The Medical Biology Centre, 97 Lisburn Road, Institute for Global Food Security, School of Biological Sciences, Medical Biology Centre, Queen's University of Belfast, Belfast, United Kingdom, BT9 7BL. Telephone: 00442890972412.

Keywords

Antimicrobial peptides, MRSA, multidrug resistant infections, rumen microbiome

32 **Abstract**

33 Herein we report the identification and characterisation of two linear antimicrobial peptides
34 (AMPs), HG2 and HG4, with activity against a wide range of multidrug resistant (MDR)
35 bacteria, especially methicillin resistant *Staphylococcus aureus* (MRSA) strains, a highly
36 problematic group of Gram-positive bacteria in the hospital and community environment. To
37 identify the novel AMPs presented here, we employed the classifier model design, a feature
38 extraction method using molecular descriptors for amino acids for the analysis, visualization,
39 and interpretation of AMP activities from a rumen metagenomic dataset. This allowed for the *in*
40 *silico* discrimination of active and inactive peptides in order to define a small number of
41 promising novel lead AMP test candidates for chemical synthesis and experimental evaluation. *In*
42 *vitro* data suggest that the chosen AMPs are fast acting, show strong biofilm inhibition and
43 dispersal activity and are efficacious in an *in vivo* model of MRSA USA300 infection, whilst
44 showing little toxicity to human erythrocytes and human primary cell lines *ex vivo*.
45 Observations from biophysical AMP-lipid-interactions and electron microscopy suggest that
46 the newly identified peptides interact with the cell membrane and may be involved in the
47 inhibition of other cellular processes. Amphiphilic conformations associated with membrane
48 disruption are also observed in 3D molecular modelling of the peptides. HG2 and HG4 both
49 preferentially bind to MRSA total lipids rather than with human cell lipids indicating that HG4
50 may form superior templates for safer therapeutic candidates for MDR bacterial infections.

51

52 **Author Summary**

53 We are losing our ability to treat multidrug resistant (MDR) bacteria, otherwise known as
54 superbugs. This poses a serious global threat to human health as bacteria are increasingly
55 acquiring resistance to antibiotics. There is therefore urgent need to intensify our efforts to
56 develop new safer alternative drug candidates. We emphasise the usefulness of complementing
57 wet-lab and *in silico* techniques for the rapid identification of new drug candidates from
58 environmental samples, especially antimicrobial peptides (AMPs). HG2 and HG4, the AMPs
59 identified in our study show promise as effective therapies for the treatment of methicillin

- 60 resistant *Staphylococcus aureus* infections both *in vitro* and *in vivo* whilst having little
- 61 cytotoxicity against human primary cells, a step forward in the fight against MDR infections.

62 **Introduction**

63 The decline in effective treatment strategies for multidrug resistant (MDR) bacterial infections
64 due to the problem of antibacterial resistance threatens our ability to treat infections now and
65 in the future, and calls for an urgent need to explore new safe drug candidates and alternative
66 treatment strategies ¹. The MDR Gram positive bacteria, methicillin resistant *Staphylococcus*
67 *aureus* (MRSA), a human opportunistic pathogen, has become a leading causative agent of
68 hospital and community acquired infections over the past few decades, posing a number of
69 challenges for physicians ²⁻⁵. Due to its pathogenicity and potential impact on a large population
70 it is on the World Health Organization (WHO) list of priority pathogens ⁶. According to the
71 Centre for Disease Control and Prevention (CDC), MRSA represents a major burden on health
72 care as it can acquire resistance to almost any class of antibiotic ⁷, leading to more than 80,000
73 invasive infections and 11000 deaths each year in the USA alone ⁸. The prevalence of MRSA
74 infections in England and Northern Ireland increased for the first time since 2011 from 1.1 in
75 2016 to 1.3 reports per 100,000 population in 2017 ⁹. Moreover, treatment of MRSA
76 bacteraemia is a long-standing challenge for the healthcare profession, often complicated by
77 metastatic infections, treatment failure and mortality ⁴. Therefore, antimicrobial compounds
78 with new modes of action for the treatment of MRSA infections are urgently needed.

79

80 Research into identifying and optimising the use of antimicrobial peptides (AMPs) in infectious
81 disease treatment has been intensifying as they are favoured as a promising new class of
82 therapeutic agents ^{1, 10}. AMPs have broad spectrum of activity (bacteria, fungi, viruses,
83 parasites etc.), form amphipathic structures, which aid interaction with the cell membrane and
84 have multimodal mechanism of action, which contributes to delayed onset of resistance in host
85 cells against them ¹¹.

86

87 The complex microbial community of the rumen of cattle (*Bos taurus*) adapts to a wide array
88 of dietary feedstuffs and management strategies, and enzymes isolated from this ecosystem

89 have the potential to possess very unique biochemical properties with possible links to
90 economically or environmentally important traits¹². Indeed, the rumen microbiome has been
91 shown as an underexplored resource for antimicrobial peptide discovery¹³⁻¹⁵ and its potential
92 to contribute some of the much needed urgent alternative therapeutic candidates to tackle the
93 looming problem of difficult to treat multidrug resistant bacterial infections. Advances in
94 nucleic acid-based technology (second generation sequencing, meta ‘omic) and high-
95 throughput sequence analytic methods (advanced bioinformatic approaches) have created new
96 opportunities to investigate the complex relationships and niches within microbial
97 communities, redefining our understanding and improving our ability to describe various
98 microbiomes, including the rumen microbiome. Such an enhanced understanding enable the
99 identification and utilization of the beneficial traits in these microbiomes^{16, 17}. Several
100 metagenomic datasets from the rumen have been generated in the last few years, illustrating
101 some of the beneficial traits of the rumen microbiome including the presence of large numbers
102 of novel glycosyl hydrolases¹⁸⁻²⁷, esterases²⁸⁻³⁰, lysoytic enzymes^{28, 31, 32}, and more recently,
103 antimicrobial compounds^{15, 17}, with the latter group of compounds possessing therapeutic
104 potential for the treatment of multi-drug resistant bacteria.

105

106 Here we combined the application of metagenomics, using one of the largest rumen
107 metagenomic dataset¹⁹ available, with advanced computational analytic tools and chemical
108 models to identify and characterise AMP candidates for the treatment of MDR infections. This
109 metagenomic data set contains more than 268 Gb, or 1.5 billion read pairs, of metagenomic
110 DNA from one single sample with the DNA from microorganisms that colonized plant fiber
111 during incubation in the cow’s rumen. *De novo* assembly of reads resulted in more than 2.5
112 million predicted open reading frames at an average of 542 bp and 55% predicted full-length
113 genes. We employed the classifier model design, a feature extraction method using molecular
114 descriptors for amino acids for the analysis, visualization, and interpretation of AMP activities,
115 and for the *in silico* discrimination of active and inactive peptides in order to define a small

116 number of promising new lead AMP test candidates for chemical synthesis and experimental
117 evaluation from this dataset. We also show the innocuity *ex vivo*, and the anti-MRSA efficacy
118 of two of these AMPs both *in vitro* and *in vivo*.

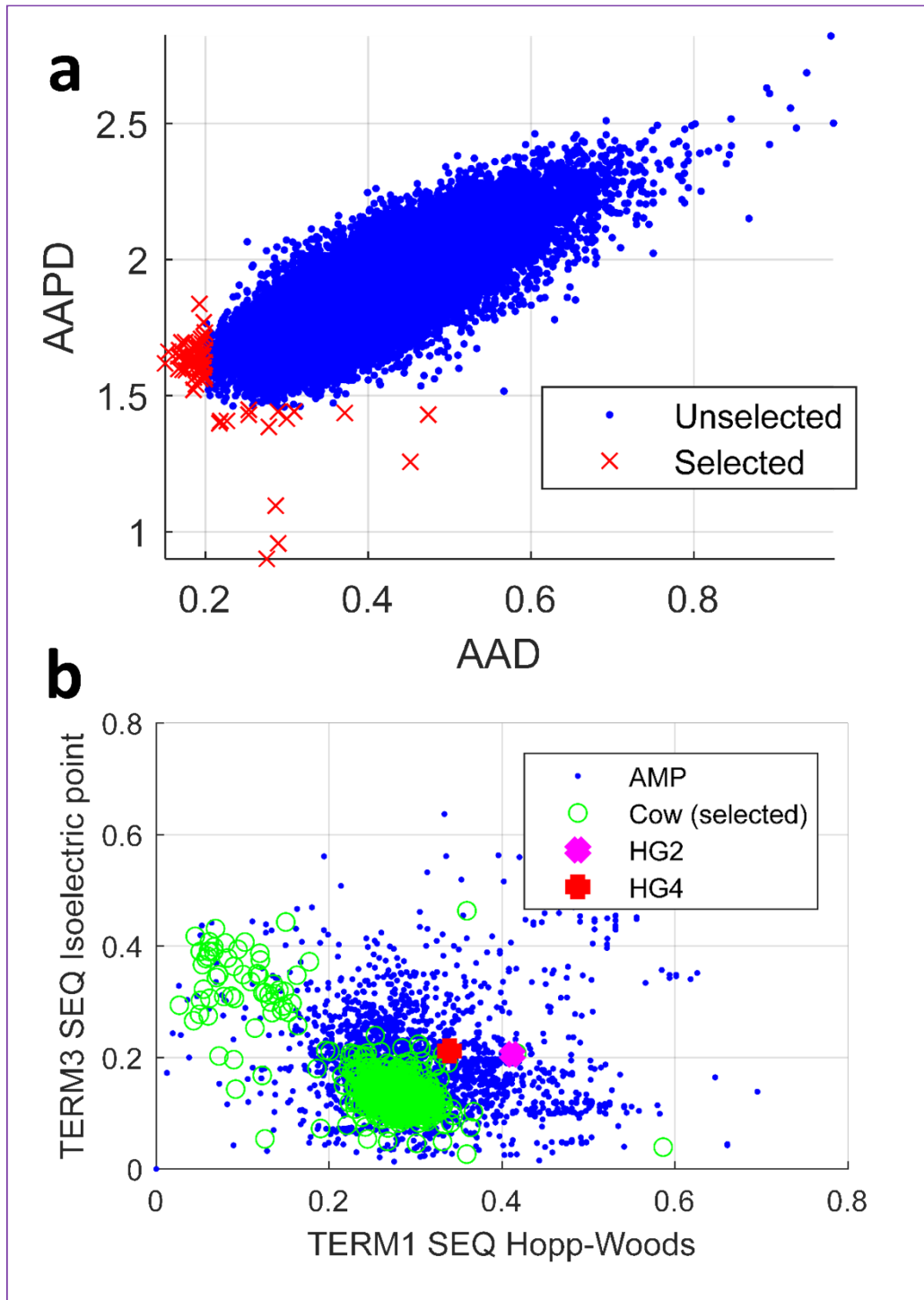
119

120 **Results and Discussion**

121

122 ***In silico* prediction and identification of AMPs using computational analysis**

123 Following, the first selection criteria in the computational analysis, 917,636 sequences (36%)
124 of the 2,547,270 predicted protein sequences in the Library “Cow” remained (i.e., protein
125 sequences with a maximum length of 200 amino acids (AAs)³³ and not more than 5% unknown
126 AAs (marked by X, *). Of these 917, 636 sequences, only 829 sequences fulfilled the criteria
127 of AA distances (AAD) <0.2 or a small AA pair distances (AAPD) <1.45, ensuring their
128 potential to be AMPs. For example, only 65 sequences met these criteria in the first 68, 274
129 sequences analysed, with isolated points outside a relatively dense distribution area as
130 illustrated in Fig. 1a. Descriptor computations generated positively charged loading -
131 hydrophobicity plots, indicating that the selected sequences from the Library “Cow” are
132 represented in only a small portion of all AMP regions from the Library “AMP” (see Fig. 1b).
133 Results from each computational step used in the identification of novel AMPs from the Hess
134 et al¹⁹ rumen metagenomic dataset is summarized in the supporting information (SI) Table S1.



135

136 **Fig. 1 a)** visualization of distances for AA acids (AAD) and AA pairs (AAPD) for the first
137 68,274 sequences from library "Cow"¹⁹ meeting the first selection criteria: candidates with
138 $AAD < 0.2$ or $AAPD < 1.45$ are selected as candidates (here: 65). **b)** standard hydrophobicity
139 (TERM1 SEQ Hopp-Woods) - loading (positively charged, TERM3 SEQ Isoelectric Point)
140 plot. Blue dots are known AMPs (library "AMP" consisting of AMPs from the APD2³⁴ and
141 Hilpert Library³⁵), green colored signs are AMP hits identified from library "Cow", and finally
142 selected peptides HG2 (magenta) and HG4 (red).

143

144 Six most promising AMP sequences (termed Hess-Gene 1-6 (HG1- HG6) (SI Table S2) were
 145 identified from the 829 sequences we utilized *in silico* approaches (Materials and Methods)
 146 sequences. The corresponding nucleotide sequence and additional information for the
 147 identified genes can be found in supporting information S3.

148 Two of these six promising candidates, HG2 (MKKLLLILFCLALALAGCKKAP) and HG4
 149 (VLGLALIVGGALLIKKKQAKS) containing 22 and 21 AA residues respectively, were
 150 selected for subsequent characterisation. Sequence homology analysis using NCBI's BLASTP
 151 ³⁶ against the non-redundant (nr) protein sequences suggests that HG2 is most similar to
 152 hypothetical or uncharacterised proteins with unknown functions from *Treponema*
 153 *maltophilum* ATCC 51939, *Pedobacter soli* and *Bacteroides* sp., while HG4 is most similar to
 154 Na⁺/H⁺ antiporter NhaC family protein from *Fibrobacter* sp. UWR2, and a hypothetical protein
 155 from *Bifidobacterium adolescentis* (Table 1).

156
 157

Table 1. Homology of antimicrobial peptides HG2 and HG4 to known sequences

Assigned AMP name		HG2	HG4
Sequence		MKKLLLILFCLALALAGCKKAP*	VLGLALIVGGALLIKKKQAKS*
Amino acid (AA) (length)		22	21
Location on cow dataset		NODE_664976_length_19740_cov_2.0 33485_orf_00810_19724..19789	NODE_3958153_length_85376_cov_8.525 382_orf_203250_82784..82849
Most similar homolog on APD2 (stop codon '**' removed)	APD ID	AP00494 ³⁷	AP01737 ³⁸
	Similarity %	40%	48%
Most similar homolog on NCBI blastp (stop codon '**' removed)	Accession number	EPF31931.1	WP_088636977.1
	Description	Hypothetical protein HMPREF9194_02286 [<i>Treponema maltophilum</i> ATCC 51939]	Na ⁺ /H ⁺ antiporter NhaC family protein [<i>Fibrobacter</i> sp. UWR2]
	Score bits/Identities %/ E-value	43.5 bits (95)/ 17/24(71%) / 0.001	44.8 bits (98)/ 16/19(84%) / 0.001

158

159 Similarly, homology analysis of the AMP sequences against the proteins from WGS
 160 metagenomic projects (env_nr) suggests that HG2 is most similar to hypothetical proteins from
 161 marine metagenomes, while HG4 is most similar to permease of the drug/metabolite transporter

162 (dmt) superfamily identified from a hydrocarbon rich ditch metagenome and hypothetical
163 proteins from marine metagenomes. It is important to note that the AMPs match to only a small
164 part of their homologous sequences. The high e-values and low bit scores and coverage of the
165 homologous sequences to HG2 and HG4 (as the best hits still had relatively high e-values of
166 0.001), as well as the fact that they are automatically curated unreviewed sequences should
167 also be noted, as this indicates the potential novelty of the peptides. Therefore, this study has
168 not only identified HG2 and HG4 sequences from the rumen metagenomic dataset as
169 antimicrobial proteins but may also provide insight into the function of the full proteins
170 carrying their homologous sequences currently annotated as hypothetical or uncharacterised
171 proteins for the first time.

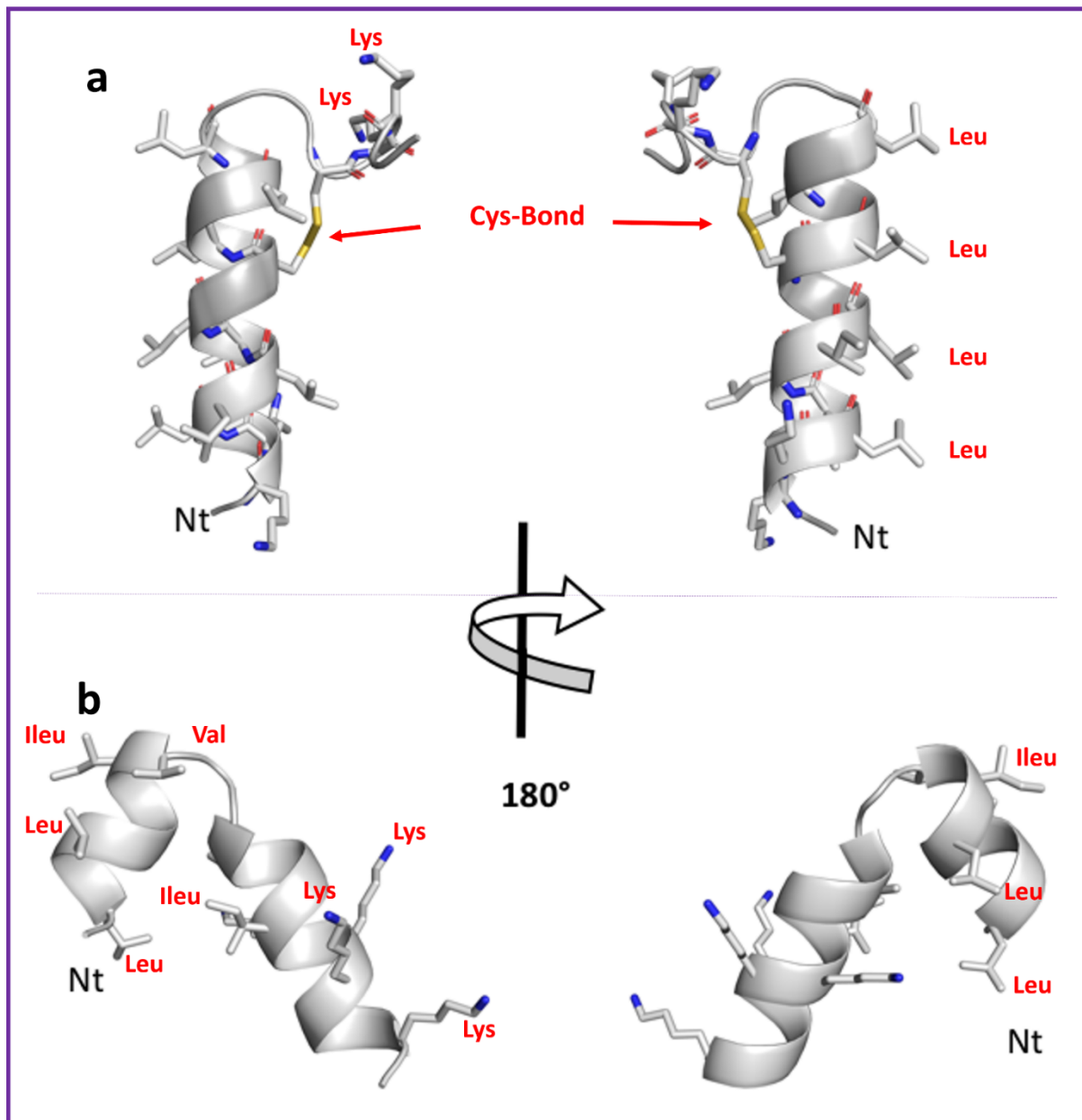
172

173 HG2 and HG4 were chemically synthesized as linear, C-terminal amidated peptides on resin
174 ($\geq 95\%$ purity, see SI Fig. S1 for mass spectrometry analysis and peptide synthesis reports)
175 using solid phase Fmoc peptide chemistry³⁹ before their antimicrobial activity was
176 investigated. It should be noted that peptide HG2 was synthesised with a disulphide bond
177 linking cysteine residues at positions 10 and 18, since HG2 showed little antimicrobial activity
178 when lacking this disulphide bond (results not shown). Similar to previously reported
179 antimicrobial peptides⁴⁰⁻⁴², HG2 ($C_{111}H_{196}N_{26}O_{23}S_3$; MW=2359.12 Da) and HG4
180 ($C_{99}H_{182}N_{26}O_{24}$; MW=2120.69 Da) are cationic both having a net positive charge of +4.
181 Whereas a hydrophobicity ratio of 57% was calculated for HG4 using ExPASy's ProtParam
182 tool⁴³, a hydrophobicity ratio of 72% was predicted for HG2, which is unusually high
183 compared to the ratio that has been reported for most AMPs⁴⁴⁻⁴⁷. This puts HG2 into the small
184 group of AMPs, representing <1% of AMPs deposited in the APD3 database³³, for which a
185 hydrophobicity ratio of $\geq 72\%$ ³³ has been reported. The positive charge and hydrophobicity of
186 AMPs are known to contribute of their antimicrobial activity as they play a role in their ability
187 to interact with the bacterial cell membrane⁴⁸.

188 **3-Dimensional molecular modelling of peptide structures**

189 Three dimensional structural modelling of HG2 and HG4 using PEPFOLD ⁴⁹ suggests that
190 these peptides have a high proportion of helical content (Fig. 2). In the case of HG2, the
191 cysteine bond stabilises the capping of the helix and the C-terminus region. Noteworthy is the
192 clear amphipathic nature of the helix with the hydrophobic residues, particularly, Leu, aligned
193 in a typical Leu-zipper motif. The Nt- C-termini includes a high proportion of the charge
194 residues (Lys), which also contribute to the segregation of charges along the peptide. As shown
195 for other examples ⁵⁰, the amphipathicity of peptides, and in particular, peptides with helical
196 conformation is an important feature of antimicrobial peptides that explains their membrane
197 disruptive mechanism of action.

198 Structural modelling also shows a high content of helical conformation in HG4 (Fig. 2). The
199 peptide forms a helix-turn-helix motif with the C-terminal helix capping stabilised by
200 hydrophobic interactions between the helices. The distribution of charges is asymmetrical as
201 expected, given the sequence of the peptide with the C-terminal part including all charged
202 residues (Lys mainly). The N-terminal helix contain mainly hydrophobic residues and the C-
203 terminal helix all charged residues with the exception of the first turn of the helix, containing
204 a high proportion of hydrophobic residues that form a mini core with the previous helix
205 possibly stabilising the conformation of the motif. The resulting conformation of the peptide is
206 therefore an amphipathic molecule, albeit different from HG2, could also point to a mechanism
207 of action on membranes.



208
209
210
211
212
213

Fig. 2 Predicted 3D structures for peptides: a) HG2, b) HG4. Main-chain and side chains depicted in ribbon and stick representation respectively and coloured according to atom type: Carbon, Oxygen and Nitrogen in green, red and blue respective. Two orientations are shown rotated about the shown axis. Ct and Nt as well as selected residues are depicted in the figure. Figures were rendered using PyMol.

214
215
216

Antimicrobial susceptibility studies

Determination of Minimum inhibitory concentrations (MIC)

219 We determined the antibacterial activity of HG2 and HG4 against various clinically important
220 multidrug-resistant pathogens including strains of *Acinetobacter baumannii*, *Klebsiella*
221 *pneumoniae*, *Pseudomonas aeruginosa*, *Escherichia coli*, *Salmonella enterica* serovar
222 Typhimurium, *Staphylococcus aureus*, *Bacillus cereus*, *Enterococcus faecalis* and *Listeria*

223 *monocytogenes* (Table 2). HG2 and HG4 had favourable antibacterial activity mostly against
224 Gram-positive pathogens, including multidrug resistant (MDR) strains (Table 2), and were
225 most potent against methicillin resistant *Staphylococcus aureus* (MRSA) strains. HG2 had a
226 minimum inhibitory concentration (MIC) range of 16-32 $\mu\text{g/ml}$ while HG4's MIC was 32-64
227 $\mu\text{g/ml}$ depending on MRSA strain, falling within the range of MICs for other rumen-derived
228 AMPs we identified previously ¹⁵, and for commercially available AMPs from isolates ^{51, 52}.
229 The peptides also showed activity against some Gram-negative bacteria strains, specifically
230 some non-resistant *A. baumannii* strains and *P. aeruginosa* strains C3719 and LES400 isolated
231 from cystic fibrosis patients (Table 2).

232

Table 2: MDR bacteria susceptibility to HG2, HG4 and comparator antibiotics measured by MIC

Organisms		MIC for peptides and comparator antibiotics (µg/ml)						
Lab no./Strain ID	Organism	Resistances	Cip/Levof (L)	Polymyxin B	HG2	H-G4	Vancomycin	Mupirocin
EMRSA-15	<i>S. aureus</i>	MRSA, Cip	>256, 1(L)	256	32	32	2	-
ATCC 33591		MRSA	0.015 (L)	-	64	32	2	-
USA300 BAA-1717		MRSA	0.0075 (L)	-	16	32	2	0.12
RN4220		Sensitive	>256	256	256	32	1	-
JH2-2	<i>Ent. faecalis</i>		64	32	256	128	64	-
NCTC 11994	<i>L. monocytogenes</i>		64	64	-	512	64	-
518842	<i>K. pneumoniae</i>	CTX-M	>128	2	512	512	-	-
ATCC 700603		SHV-18	0.25	2	>512	>512	-	-
NCTC 13442		OXA-48	64	8	512	512	-	-
526903		Sensitive	0.03	4	>512	512	-	-
515785	<i>A. baumannii</i>	IMI, MER	16	0.25	128	64	-	-
515908		OXA-23, OXA-50	>128	0.5	256	128	-	-
515722		Sensitive	16	0.5	32	64	-	-
515722		Sensitive	16	0.25	16	32	-	-
K12	<i>E. coli</i>		0.06	2	256	512	128	-
SL1344	<i>Sal. typhimurium</i>		0.12	2	256	512	256	-
	<i>B. cereus</i>		0.015	-	256	512	-	-
PA01	<i>P. aeruginosa</i>		0.5	2	>512	>512	64	-
AMT0060			0.12	0.5	256	256	-	-
C3719	<i>P. aeruginosa (CF)</i>		4	1	64	128	-	-
LES400			4	1	64	128	-	-

Cip Ciprofloxacin, Lev (L) Levofloxacin, CF isolates from cystic fibrosis infections, OXA Oxacillin, CTX-M extended spectrum β-Lactamase, SHV-18 β-Lactamase, IMI (imipenem), MER (meropenem).

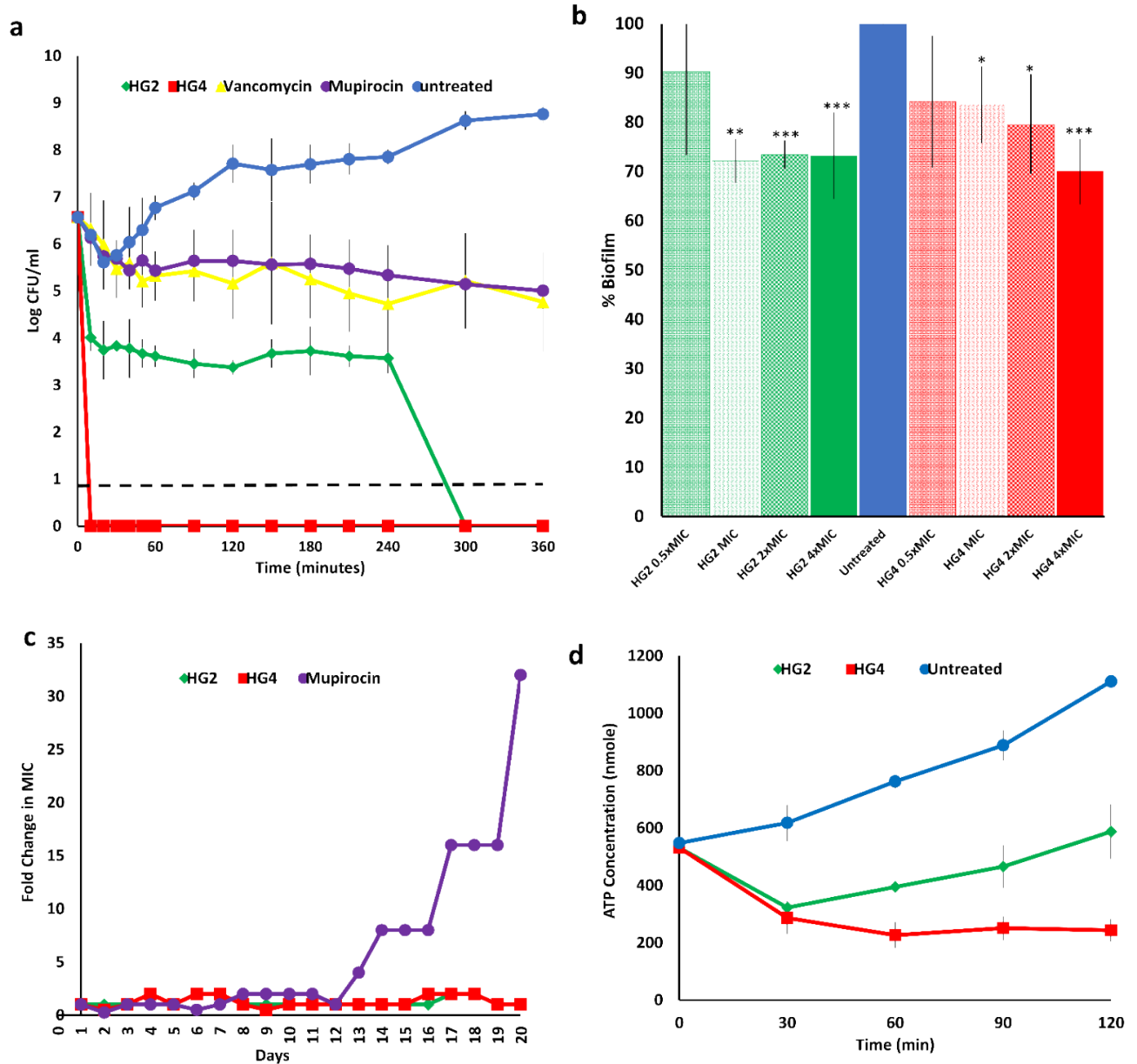
236 **Time kill kinetics**

237 The bactericidal activity of HG2 and HG4 against logarithmic-phase MRSA USA300 cells was
238 investigated by time kill kinetic studies. Compared with vancomycin and mupirocin, HG2 and
239 HG4 (at 3x MIC concentration) had a rapid bactericidal activity against MRSA USA300 strain
240 (Fig. 3a), causing reductions of $>3 \log_{10}$ CFU/ml and $>6 \log_{10}$ CFU/ml, respectively, within
241 the first 10 min. HG2 and HG4 induced complete cell death within 5 hours and 10 min of
242 treatment respectively, with no recovery observed after 24 hours of incubation. This rapid and
243 total loss in bacteria cell viability is similar to the killing kinetics that have been reported for
244 many fast acting antimicrobial peptides^{15, 53}. As expected, vancomycin and mupirocin at 3x
245 MIC produced $\geq 2 \log_{10}$ CFU/ml reductions attributable to differences in kill kinetics and mode
246 of action⁵⁴.

247 **Anti-biofilm activity**

248 There have been numerous AMPs that have been reported to be capable of inhibiting biofilm
249 formation for difficult to treat pathogens, favouring their application as antimicrobial agents in
250 medical implants and other biomaterials⁵⁵⁻⁶¹. This prompted us to utilize a 96-well biofilm
251 model¹⁵ to investigate the ability of HG2 and HG4 to dislodge/disrupt and disperse already
252 formed and established MRSA USA300 biofilms. In general, all AMP treatments showed
253 activity against established biofilms at MIC, 2x MIC and 4x MIC concentrations. No
254 statistically significant anti-biofilm activity was observed in biofilms treated with 0.5x MIC
255 AMP concentrations (Fig. 3b). The anti-biofilm activities of HG2 and HG4 indicate their
256 suitability potential agents for the disinfection of medical devices as well as in the treatment of
257 biofilm infections such as wounds.

258



259

260 **Fig. 3 Antimicrobial susceptibility and activity of HG2 and HG4.** **a)** Time dependent kill of MRSA
 261 USA300 cells by AMPs at 3x MIC concentration. Dashed lines indicate limit of detection. **b)** Anti-
 262 biofilm activity against MRSA USA300 biofilms: *, ** and *** ($P \leq 0.05$, 0.01 and 0.001 respectively-
 263 significantly different from untreated cells (positive). **c)** Resistance acquisition during serial passaging
 264 of MRSA USA300 cells in the presence of sub-MIC levels of antimicrobials. The y axis is the fold
 265 change in MIC during passaging. For mupirocin, 32x MIC was the highest concentration tested. The
 266 figure is representative of 3 independent experiments. **d)** ATP depletion activity in MRSA USA300
 267 cells.

268

269

270 **Selection for resistance (serial passage)**

271 Although relatively uncommon, bacterial resistance to cationic antimicrobial peptides is an
272 evolving phenomenon^{62, 63}. Resistance to many AMPs including polymyxin B has recently
273 been reported⁶⁴⁻⁶⁶, and it is therefore important to understand bacterial resistance to AMPs and
274 to identify and design more robust AMPs. Mechanisms of resistance to AMPs, which are
275 mostly non-specific and confer moderate levels of resistance^{67, 68}, are mainly based on changes
276 in the physicochemical properties of surface molecules and the cytoplasmic membrane^{62, 63}.
277 For therapeutic AMP candidates, it is important that bacterial AMP resistance, which may
278 develop due to selective pressure is not based on mutations or acquisition of specific resistance
279 genes, which can then be horizontally transferred between bacteria species as with conventional
280 antibiotics^{69, 70}. Here, we assessed the likelihood of resistant mutants and/or resistance arising
281 when MRSA cells are exposed to sub-MIC levels of HG2 and HG4. Continuous exposure of
282 bacteria cells to sub-lethal doses of the AMPs over a period of 20 days did not produce resistant
283 mutants (Fig. 3c), and MICs remained within 1-2 fold increases compared to mupirocin treated
284 cells, which had a 32-fold MIC increase within the same period. The observed increase in MIC
285 is rather common for many AMP-based molecules as a small change in the MIC after exposure
286 to the AMP is to be expected^{71, 72}. Our inability to recover resistant mutants in this experiment
287 suggests that the HG2 and HG4 may have non-specific or multiple cellular targets as has been
288 described previously for peptides⁷³.

289
290 **Biochemical mode of action studies**

291 **ATP depletion assay**

292 Adenosine triphosphate (ATP) is a high-energy nucleoside triphosphate molecule formed in
293 the cytosol of bacteria and mitochondria of eukaryotes and drives most cellular and metabolic
294 processes in microbial cells⁷⁴⁻⁷⁶. Changes in the concentration of ATP can be used as an
295 indicator of cell viability and competence. We tested the effect of HG2 and HG4 on ATP
296 concentration levels in *S. aureus* MRSA USA300. As expected, untreated bacteria cells

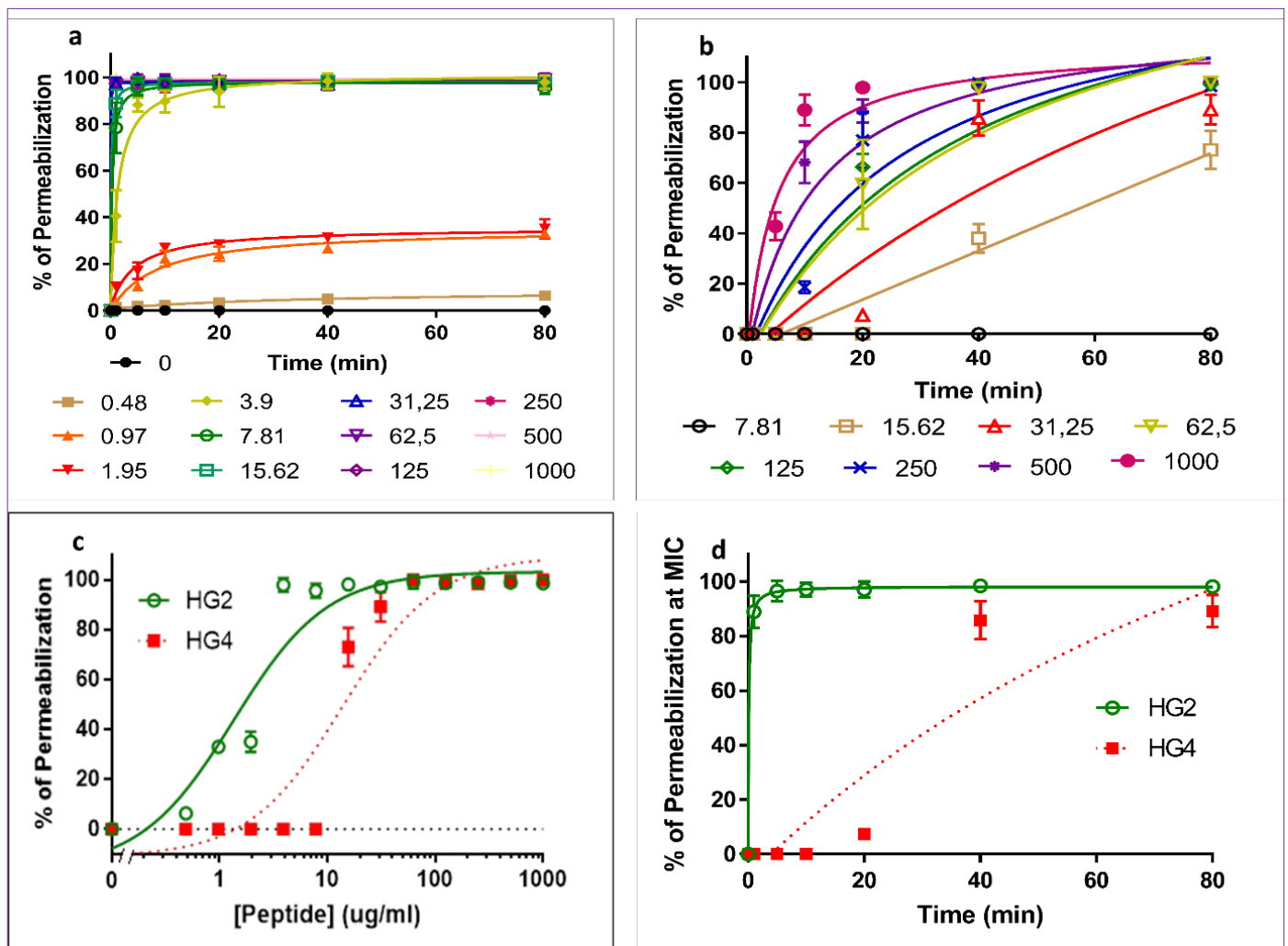
297 generated increasing amounts of ATP over time, whereas significantly lower concentrations of
298 ATP were observed in HG2 and HG4 treated cells (Fig. 3d). This decrease in ATP
299 concentrations may be an indication of ATP depletion, limiting cellular energy and thus other
300 related cellular processes (such as substrate transport, homeostasis and anabolism) likely linked
301 to cell membrane disturbance. HG2 and HG4 induced a significant ($P = 0.018$ and 0.003 ,
302 respectively) decrease in ATP concentration in *S. aureus* MRSA USA300 cells, which is
303 similar to what was reported for other cationic AMPs^{47, 77}. Hilpert et al⁴⁷ demonstrated that
304 many short AMPs had a strong effect on ATP concentration whereas several 26mer α -helical
305 peptides did not.

306

307 **Bacterial membrane permeabilisation assay**

308 Since HG2 and HG4 peptides possessed rapid bactericidal effect, we used the propidium iodide
309 (PI) method^{15, 78} to determine if these AMPs were able to permeabilise MRSA USA300
310 cytoplasmic membrane similar to what was observed for other rumen derived AMPs¹⁵. MRSA
311 USA300 cells exposed to increasing concentrations of HG2 or HG4 both showed increase in
312 PI entry/fluorescence over time, demonstrating that they were able to permeate the cytoplasmic
313 membrane and therefore indicating that they may possess pore-forming activity (Fig. 4a, b).
314 Indeed, significant permeabilisation ($p < 0.01$) of MRSA USA300 cytoplasmic membrane was
315 observed even at sub-MIC concentrations, e.g. as low as 1 and 16 $\mu\text{g/ml}$ for HG2 and HG4,
316 respectively (MIC values being 16 and 32 $\mu\text{g/ml}$ for HG2 and HG4 respectively). The Effective
317 Concentration 50 (EC50) (defined as the concentration of a drug at which the drug is half-
318 maximally effective) of HG2 and HG4 measured after 80 min of incubation were 1.351 (± 0.27)
319 and 13.85 (± 3.22) $\mu\text{g/ml}$, while total membrane permeabilisation was observed at 3.9 and 62.4
320 $\mu\text{g/ml}$, respectively (Fig. 4c). Membrane permeabilisation kinetics of HG2 and HG4 at their
321 MIC concentration, showed that HG2 was able to permeabilise the membrane faster (80%

322 permeabilisation at 1 min and maximal effect after 5 min) than HG4 (minor permeabilisation
 323 at 20 min, maximal effect after 40 min) (Fig. 4d).



324
 325 **Fig. 4 Membrane permeabilisation action of HG2 and HG4 against MRSA:** a) Membrane
 326 permeabilization activity of HG2 at different concentrations ($\mu\text{g/ml}$) against MRSA USA300 cells
 327 measured by propidium iodide assay over time. b) Membrane permeabilization activity of HG4 at
 328 different concentrations ($\mu\text{g/ml}$) against MRSA USA300 cells measured by propidium iodide assay over
 329 time. c) Determination of EC₅₀ (Effective Concentration 50) of HG2 and HG4 membrane
 330 permeabilisation measured after 80 min. d) Membrane permeabilisation kinetics of HG2 and HG4 at their
 331 MIC concentration. In all cases, values are from three independent replicates; results are expressed as
 332 means \pm standard deviation).

333

334 Transmission electron microscopy

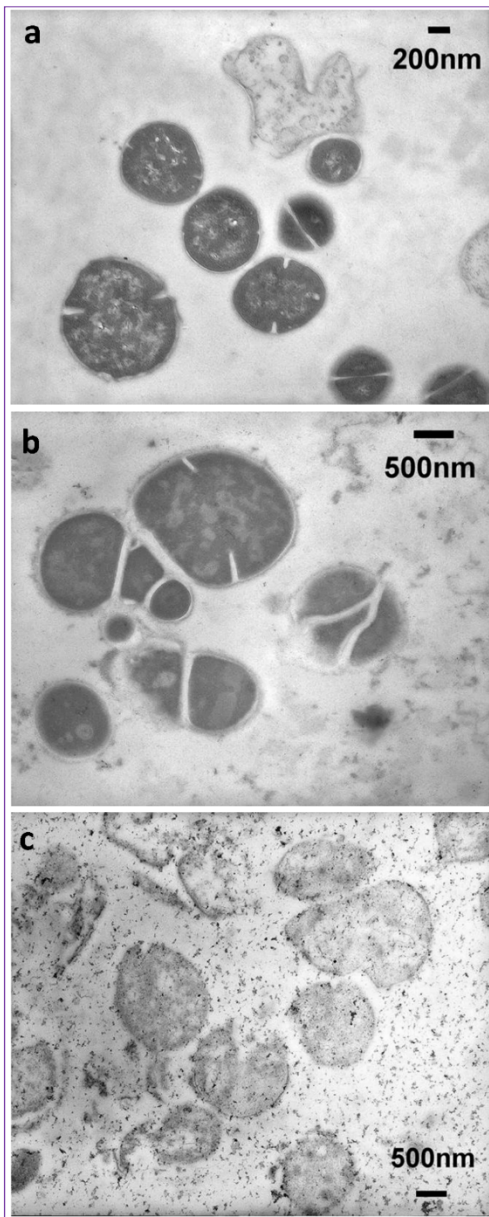
335 Transmission electron micrographs (TEM) of cells treated with HG2 and HG4 at 3x MICs for

336 1 h revealed changes in cell morphology and some cytoplasmic damage (Fig. 5). The

337 morphological changes observed in the HG2 and HG4 treated MRSA USA300 cells correspond

338 with the membrane permeabilisation activity of the peptides. The semi-quantitative nature of

339 TEM analysis means that investigation into events leading up to changes in cell morphology
340 may be necessary.



341 **Fig. 5 Representative transmission electron micrographs of MRSA cells. a)** micrographs
342 **untreated MRSA USA300 cells. b)** HG2 treated (3x MIC for 1 h) MRSA USA300 cells. **c)**
343 **HG4 treated (3x MIC for 1 h) MRSA USA300 cells. Scale bars are 200 or 500 nm as shown**
344 **on micrographs.**
345

346

347 ***In vitro* and *ex vivo* innocuity and cytotoxicity studies**

348 **Haemolytic activity**

349 To establish the potential of HG2 and HG4 as therapeutic agents, the haemolytic effect of HG2
350 and HG4 were tested on human red blood cells. HG2 and HG4 induced low haemolysis, with

351 HC₅₀ (i.e. the concentration of peptide causing 50 % haemolysis) of 409 ±67 and 458 ±101
 352 µg/ml, respectively, with safety factors of 26.2 and 14.6X MIC respectively (Table 3).

353 **Cytotoxicity studies on human primary cells and cells lines**

354 Cytotoxicity of HG2 and HG4 against different human cell types was evaluated by measuring
 355 the IC₅₀, which is the concentration of peptide inhibiting 50% of the cell viability. Lung
 356 fibroblast (IMR-90) cells were found to be the most sensitive to HG2 and HG4 with IC₅₀ of 96
 357 ± 21 and 294 ± 42 µg/ml for HG2 and HG4. Lung epithelial (BEAS-2B) and liver (HepG2)
 358 cells) cells were the least sensitive to HG2 and HG4 with IC₅₀ of 120 ± 25 and 359 ± 76 µg/ml
 359 for HG2 and IC₅₀ >1000 µg/ml for HG4 (Table 3). As a whole, cytotoxicity data showed that
 360 HG4 was less toxic than HG2 for human cells and that epithelial cell types (BEAS-2B and
 361 HEPG2) were the least sensitive, while fibroblast and endothelial cells were more susceptible
 362 to the peptides (Table 3). The high hydrophobicity of HG2 may contribute to its higher toxicity
 363 to human erythrocytes and cell lines. This is similar to Gramicidins that possess high
 364 hydrophobicity ratios and are exclusively used topically due to their haemolytic side-effects⁷⁹⁻
 365 ⁸¹, it is possible that the future application of HG2 might be restricted to topical application to
 366 treat superficial infections, unless modified derivatives/analogues of HG2 with improved
 367 cytotoxicity become available.

368 **Table 3. Cytotoxicity and haemolytic activities of HG2 and HG4 on human cells.**
 369 Cytotoxicity is expressed as IC₅₀ (i.e. the concentration of peptide in µg/ml causing a reduction
 370 of 50% of the cell viability). Haemolytic activity is expressed as HC₅₀ (i.e. the concentration
 371 of peptide causing 50% haemolysis). IC₅₀ and HC₅₀ are expressed in µg/ml concentrations.
 372 Therapeutic Indexes (T.I) corresponding to the fold difference between IC₅₀ or HC₅₀ and MIC
 373 values (for MRSA USA300) are given in brackets.

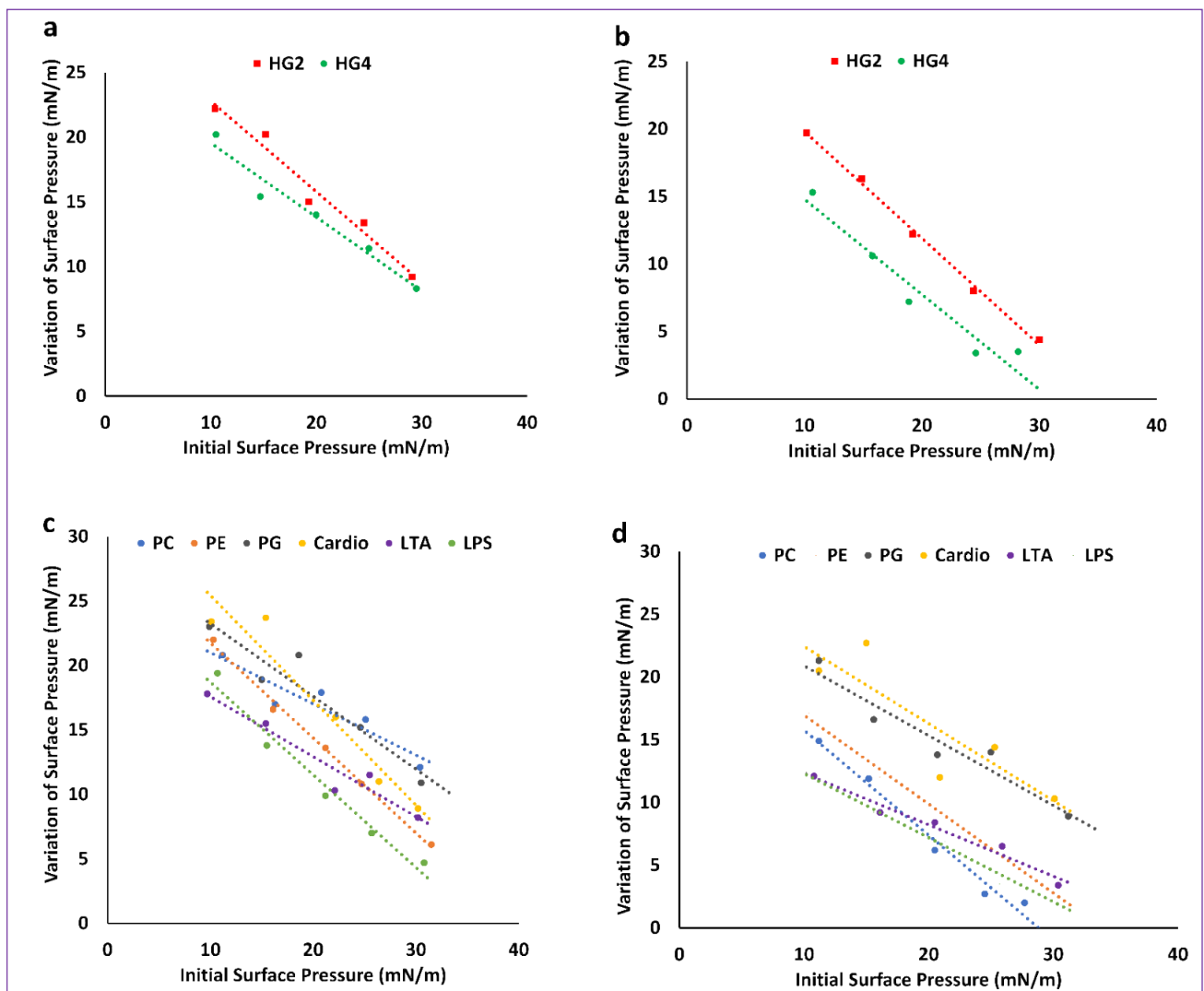
Cell type (human)	HG2 (X MIC)	HG4 (X MIC)
BEAS-2B	120 +/- 25 (X 7.7)	>1000 (> 32)
HEPG2	359 +/- 76 (X 23)	>1000 (> 32)
IMR-90	96 +/- 21 (X 6.1)	294 +/- 42 (X 9.4)
Erythrocytes	409 +/- 67 (X 26.2)	458 +/- 101 (X 14.6)

374 **Peptide-lipid interaction and insertion assay**

375 The interaction of HG2 and HG4 with lipid monolayers was measured by the critical pressure
376 of insertion, reflecting the affinity of the peptides for specific lipids. Insertion capacity was
377 first measured using total lipids extracts obtained from MRSA USA300 cells or human
378 erythrocytes (Fig. 6a, b, and SI Table S4) and obtained data suggested that HG2 and HG4 had
379 higher affinity and insertion ability in MRSA lipids, with critical pressure of insertion of 35.07
380 and 30.99 mN/m and 42.59 and 44.18 mN/m for HG2 and HG4 in MRSA and erythrocyte
381 lipids, respectively. These results showing that HG4 is less able than HG2 to insert into
382 erythrocyte lipids is in accordance with the lower haemolytic activity observed in HG4
383 compared to HG2 (HC50 of 409 and 458 $\mu\text{g/ml}$ for HG2 and HG4, respectively).

384 To identify the lipid partner(s) recognized by HG2 and HG4 in bacterial and eukaryotic
385 membranes measurement of the critical pressure of insertion in pure lipids were performed
386 (Fig. 6c, d and SI Table S4). Results indicated that HG2 and HG4 had different affinities and
387 insertion capacities in pure lipids from bacteria or eukaryotes. Whereas HG4 interacts
388 preferentially with bacterial lipids expressed in the outer leaflet of the membrane (1-palmitoyl-
389 2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol) (POPG or PG), 1-palmitoyl-2-oleoyl-sn-
390 glycero-3-phosphoethanolamine (POPE or PE), cardiolipin, lipoteichoic acid from *S. aureus*
391 (LTA), lipopolysaccharide from *E. coli* (LPS) over eukaryotic lipids (1-palmitoyl-2-oleoyl-
392 glycero-3-phosphocholine POPC or PC) (i.e. PG > cardiolipin > LTA > LPS > PE > PC), HG2
393 displayed the opposite order of selectivity, with first the major lipid present in the outer leaflet
394 of the eukaryotic membrane (PC) followed by bacterial membrane lipids (i.e. PC > PG > LTA
395 > Cardio > PE > LPS). These observations are in accordance with the higher toxicity of HG2
396 (high susceptibility of human cells to HG2) against human cell lines compared to HG4 (Fig.
397 6c, d and SI Table S4). Interestingly, neither HG2 nor HG4 insert efficiently into LPS. This
398 corresponds to the low antimicrobial activity of HG2 and HG4 against Gram negative bacteria
399 compared to their potent activity against Gram positive bacteria). Measurement of the speed of

400 insertion of HG2 and HG4 in total lipid extracts or pure lipids at an initial surface pressure of
401 30 ± 0.5 mN/m (corresponding to the theoretical surface pressure of eukaryotic and prokaryotic
402 membranes) showed that HG2 inserts faster into all lipid monolayers compared to HG4 (Table
403 2 in supporting information), confirming results obtained in the membrane permeabilisation
404 assay, which indicated a faster bacterial membrane permeabilisation of HG2 when compared
405 to HG4.



406
407 **Fig. 6 Peptide lipid interaction and insertion measurements:** a) Interaction of HG2 and HG4 (at 1
408 $\mu\text{g}/\text{mL}$ final concentration) with lipids (either total lipid extracts or pure lipids) was measured using
409 lipid monolayers. a) interaction HG2 and HG4 with total MRSA lipid extract. b) interaction HG2 and
410 HG4 with total lipid extract from human erythrocytes. c) interaction of HG2 with pure lipids and c)
411 interaction of HG4 with pure lipids. 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
412 (PG), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (PE), Cardiolipin (Cardio),
413 Lipoteichoic acid (LTA) from *S. aureus*, Lipopolysaccharide (LPS) from *E. coli* and (1-palmitoyl-2-
414 oleoyl-glycero-3-phosphocholine (PC).

415 ***In vivo* efficacy studies**

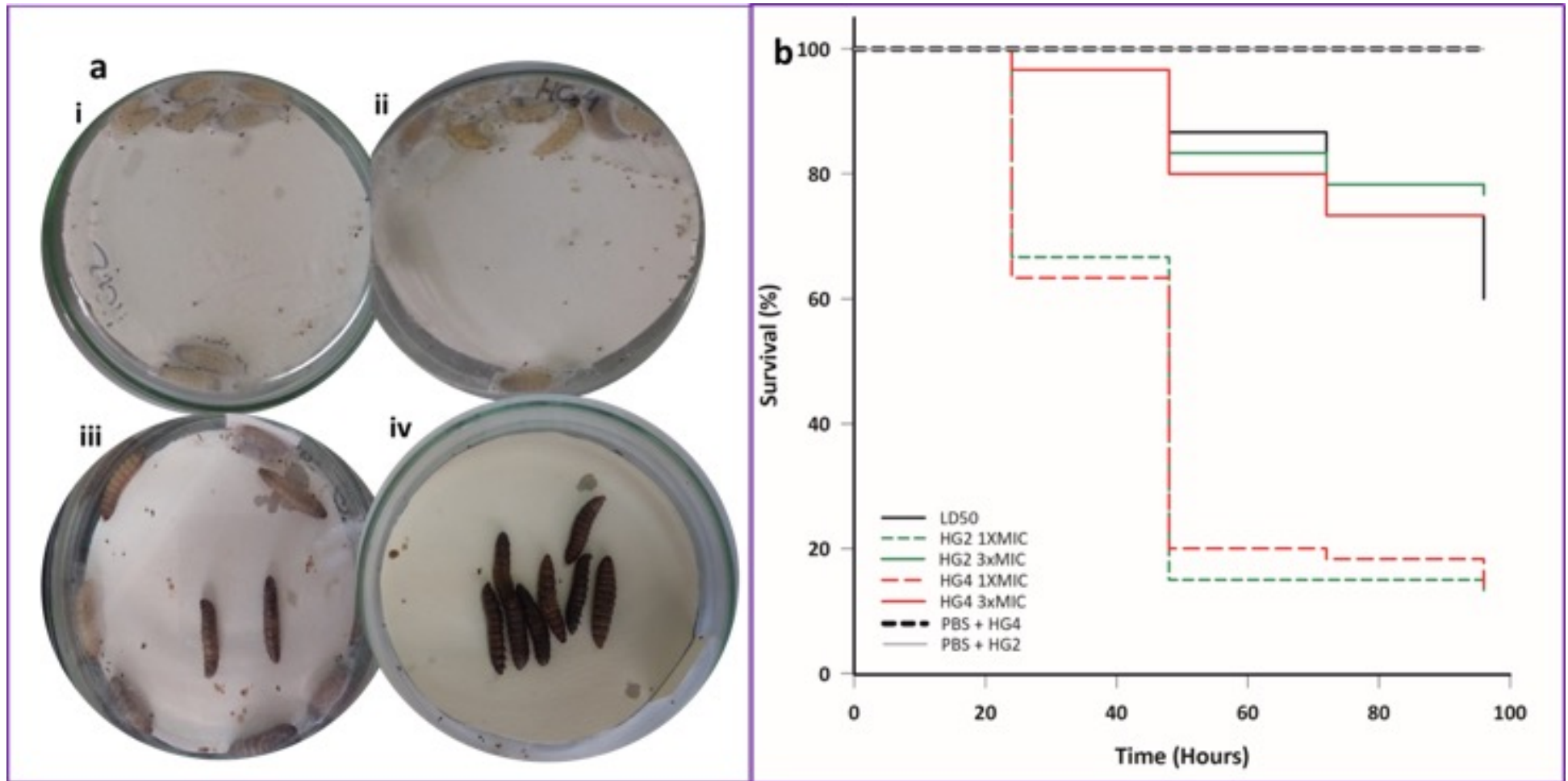
416 *Galleria mellonella* has been used widely as an effective model for testing antimicrobials drugs
417 and their toxicity⁸²⁻⁸⁵. We investigated the toxicity of HG2 and HG4 against *G. mellonella*
418 larvae as well as their ability to protect the larvae from a lethal dose of an MRSA USA300
419 infection. The peptides HG2 and HG4 showed no toxic effect in *G. mellonella* at either 1x MIC
420 or 3x MIC concentrations (Fig. 7a). As with the control group of larvae, no phenotypic
421 modification such as melanisation and/or motility reduction was observed in larvae injected
422 with HG2 and HG4.

423

424 The MRSA USA300 infective dose (LD₅₀) was determined as being 10⁵ CFU/larvae, while the
425 lethal dose (LD) was determined as 2.25x10⁶ CFU/larvae (caused melanisation and death of all
426 larvae within 24 h (Fig. 7a). Larvae infected with MRSA USA300 LD and treated with peptides
427 HG2 and HG4, at 1x MIC concentration had increased survival rate by ~20% (Fig.7b). In
428 comparison to larvae in the untreated control group, larvae infected with MRSA USA300 LD,
429 followed by treatment with peptides 3x MIC increased survival by 4.6-fold and 4.4-fold for
430 HG2 and HG4 with a survival rate of 78% and 75%, respectively (Fig. 7b).

431

432 We were able to show that both peptides, especially at 3x MIC, concentration can effectively
433 control MRSA USA300 infection *in vivo* in *G. mellonella*. Other pharmacological aspects of
434 the peptides need to be investigated in order to improve the efficacy of HG2 and HG4, as
435 survival rate of AMP-treated larvae was comparable to the LD₅₀ survival rate, probability due
436 to distribution and/or bioavailability of the peptides *in vivo* (in *G. mellonella* model).
437 Nonetheless, our results show that the peptides HG2 and HG4 at 3x MIC concentration are
438 capable of significantly improving the survival of larvae infected with a lethal dose MRSA
439 USA300 and are active against this clinically important drug resistant pathogen in an *in vivo*
440 model.



441

442 **Fig. 7 In vivo efficacy assessment in *G. mellonella* MRSA infection model:** a) representative images of toxicity assay of peptides- i) HG2, and ii)
 443 HG4 in *G. mellonella* 120 h post treatment with 3x MIC concentrations. The larvae remained alive and without melanisation. iii) virulence assay of
 444 MRSA USA300 in *G. mellonella* using a lethal dose inoculum of 10^6 CFU/per larvae- iii) 24 hours post infection: some larvae were dead and partial
 445 melanisation was observed. iv). 48 hours post infection: most larvae were dead and complete melanisation was evident. The experiment was done with
 446 three experimental replicates, each containing groups of 10 larvae. b) Kaplan-Meier survival curves of *G. mellonella* infected with a lethal dose of *S.*
 447 *aureus* (2.25×10^6 CFU/larvae) and treated with placebo (showing a 100% larvae survival rate) or peptides HG2 and HG4 at a 1x and 3x MIC
 448 concentrations.

449 **Conclusion**

450 The two AMPs, HG2 and HG4, identified using *in silico* approaches from a rumen
451 metagenomic dataset have further confirms the rumen as an invaluable resource for urgently
452 needed alternatives to currently available antibiotics. Furthermore, study presented here
453 emphasises the usefulness of complementing wet-lab and *in silico* techniques for the rapid
454 identification of new AMP candidates from environmental samples. The low similarity of the
455 newly identified AMPs to previously known sequences suggest their novelty from an
456 evolutionary point of view. Experimental evaluation and characterisation of the antimicrobial
457 properties of HG2 and HG4, two of the identified AMP candidates revealed their antimicrobial
458 activity against Gram positive bacteria. Their activity against MRSA suggest that membrane
459 permeabilisation and decrease in intracellular ATP concentration might play a role in their
460 antimicrobial activity. HG2 and HG4 both preferentially bind to MRSA total lipids rather than
461 with human erythrocyte lipids. HG4 was less cytotoxic against all cell lines tested and was
462 observed to bind more specifically to pure bacterial membrane lipids, indicating that HG4 may
463 form a more superior template for a safer therapeutic candidate than HG2. The non-toxic effect
464 of the peptides against *G. mellonella* larvae, and their *in vivo* efficacy against MRSA USA300
465 infection in the *G. mellonella* infection model suggests that these peptides might possess
466 potential as safe alternative therapeutic agents with anti-biofilm activity for the treatment of
467 bacterial infections. Given the technological advances, improvements in genomic methods and
468 computational analytic approaches as well as the growing abundance of omics data, it is likely
469 that the approach developed and presented here for the identification of novel AMPs, might
470 facilitate the discovery of a growing number of other AMPs and other bioactives from
471 environments where conventional isolation and cultivation of microorganisms is challenging.

472

473 **Materials and Methods**

474 ***In silico* prediction and identification of AMPs using computational analysis**

475 Antimicrobial peptide prediction and similarity analysis was performed on the rumen
476 metagenomic dataset from the study by Hess et al, ¹⁹. The dataset was termed the Library

477 “Cow” dataset and contains 2,547,270 predicted protein sequences
478 ('metagenemark_predictions.faa.gz') which were downloaded from the weblink
479 (<http://portal.nersc.gov/project/jgimg/CowRumenRawData/submission/>). All other datasets
480 (libraries) used for similarity analysis prediction/identification of novel AMP candidates from
481 the “Cow” dataset and their respective sources are as follows: The Library “AMP1”, which
482 contained a list of 2308 known antimicrobial peptides (AMPs) downloaded from APD2 ³⁴
483 (downloaded on November 10, 2013 and available at <http://aps.unmc.edu/AP/main.php>) and
484 the Library “AMP2”, which contains of a list of 48 synthetic AMPs (Hilpert Library) identified
485 by Ramon-Garcia et al ³⁵. The MATLAB toolbox Gait-CAD and its successor SciXMiner
486 (<http://sourceforge.net/projects/scixminer/>) ^{86, 87} including the Peptides Extension Package ⁸⁸
487 were mostly used for the computational analysis.

488 The "fastread" function of MATLAB Bioinformatics toolbox was used to import the “Cow”
489 dataset. For easier computational analysis, the imported dataset was then split into 26 parts
490 with ~100,000 sequences each. Following recommendations that small antimicrobial proteins
491 should have a length <200 amino acids (AAs), with most AMPs (>90%) on the APD2 database
492 having a length of <60 AAs ³⁴, only protein sequences with a maximum length of 200 AAs³³
493 and not more than 5% unknown AAs (marked by X, *) were selected from the “Cow” dataset
494 predicted protein sequences (metagenemark_predictions.faa.gz). Libraries "AMP1" and
495 "AMP2" were combined to produce Library "AMP" composing of a total of 2356 peptides.
496 Thereafter, AA distribution and AA dimer (pair) distribution were computed, resulting in
497 proportion for 20 AAs (and 20x20 = 400 AA dimers) for Libraries "Cow" and "AMP". Pairwise
498 distances of AA acid distributions between two peptides (termed “AAD”) were computed with
499 a minimal value of 0 for identical and increasing values for different AAs. Pairwise distances
500 of AA acid and AA acid pair distributions between two peptides were computed respectively
501 (distance for 400 + 20 features), termed "AAPD". For each candidate of Library "Cow", the
502 values of “AAD” and “AAPD” to each peptide in Library "AMP" were computed. Again, for
503 each sequence in Library "Cow", minimal distance values from the previous computational

504 step and the number of most similar peptide from Library "AMP" were saved as separate
505 features. To select only promising candidates of AMP predictions, only sequences from the
506 Library "Cow" with small AA distances $AAD < 0.2$ or a small AA and AA pair distances $AAPD$
507 < 1.45 were saved. The distances were computed using the 1-norm (Manhattan norm). The
508 thresholds for AAD and AAPD were heuristically chosen to balance the trade-off between too
509 many weak candidates (too high values of AAD and AAPD) vs. the loss of promising
510 candidates (too low values of AAD and AAPD). All sequences that fulfilled the conditions in
511 the preceding step were collected and some interesting hits were used to check similarity in
512 APD2 database (criteria: small values of AAD or AAPD, different neighbours to explore the
513 variety of the candidates found, short peptide length). Finally, descriptors were computed
514 following procedures described by Mikut et al ^{88, 89} to check the expected balance between
515 hydrophobicity and positively charged AAs as a typical design criterion for AMPs.

516

517 **Peptide synthesis and 3-dimensional molecular modelling of peptide structures**

518 Pure peptides ($\geq 95\%$ purity) were synthesized on resin using solid phase Fmoc peptide
519 chemistry ³⁹ by GenScript Inc. USA. A *de novo* structural prediction method, PEP-FOLD ⁴⁹
520 was used to model the 3D conformation of peptides HG2 and HG4. 200 simulations were ran
521 for each peptide and resulting conformations were clustered and ranked using the sOPEP
522 coarse grained force field ⁹⁰. In the case of HG2 the formation of the cysteine bond between
523 Cys10 and Cys18 was imposed as restraint to the simulation. The best 3D models for each
524 peptide was chosen and manually analysed using PyMOL v1.7.6 ⁹¹.

525

526 **Antimicrobial susceptibility of bacterial cells**

527 To determine the antimicrobial activity of new antimicrobial peptides, HG2 and HG4, their
528 minimum inhibitory concentrations (MICs) were determined by broth microdilution method ⁹²
529 in cation adjusted Mueller Hinton broth (MHB) following the International Organization for
530 Standardization (ISO) 20776-1 standard for MIC testing using a final bacterial inoculum of 5
531 $\times 10^5$ CFU/ml ⁹³. The lowest concentration of the AMPs that inhibited the visible growth of
532 the bacteria tested after an overnight incubation at the appropriate temperatures (37°C for all

533 organisms except for *Listeria monocytogenes* 30°C) and growth conditions was taken as the
534 MIC. The peptides dissolved in sterile distilled water were added to bacteria culture and
535 incubated overnight at appropriate conditions. The MICs of the peptides and comparator
536 antibiotics were recorded after 18-24 h. The minimum bactericidal concentrations (MBCs)
537 were also determined and were taken as the lowest concentration of the antimicrobial that
538 prevented the growth of bacterial cells after subculture of cells (from MIC treatment) onto
539 antibiotic-free media.

540

541 **Time kill kinetics**

542 The bactericidal activity of HG2, HG4 and comparator antimicrobial compounds was assessed
543 as previously described⁹⁴ using exponential-phase cultures of MRSA USA300 grown in MHB
544 ($1 \times 10^{6-8}$ CFU/ml). Cells were treated with antimicrobial compounds at 3 times their MIC
545 concentrations (final concentrations), incubated 37°C with gentle shaking at 110 rpm. Samples
546 were taken at different time points and inoculated onto MH agar plates using the spread plate
547 technique. After overnight incubation, the colony forming units per millilitre of cell culture
548 (CFU/ml) was calculated. Experiments were performed in quadruplicates.

549

550 **Anti-biofilm activity**

551 The ability of HG2 and HG4 to disrupt established *S. aureus* biofilms was measured using a
552 96 well format as described by¹⁵. MRSA USA300 cultures grown overnight in Brain Heart
553 Infusion (BHI) broth was re-suspended to an $OD_{600nm} = 0.02$ and grown without shaking at
554 37°C in 96 well tissue culture plates for another 24 h. The planktonic cells were removed by
555 three PBS (phosphate buffered saline) washes. Thereafter, fresh BHI broth containing peptides
556 HG2 or HG4 at sub- and supra MIC concentrations (0.5X, 1X, 2X and 4X MIC) was added to
557 wells containing adherent cells and incubated without shaking for another 24 h. Planktonic
558 cells were again removed by three PBS washes. Biofilms were fixed with methanol for 20 min,
559 stained with 0.4% (w/v) crystal violet solution for 20 min and re-solubilised with 33% (v/v)
560 acetic acid. The optical density of re-solubilised biofilms was measured at 570_{nm} in a microplate

561 spectrophotometer. The growth of biofilm per treatment was calculated as a percentage of the
562 untreated cells and anti-biofilm activity was determined for statistically significant treatments.

563

564 **Serial passage/resistance assay**

565 *In vitro* evaluation of the potential for MRSA USA3000 cells to develop resistance to HG2 and
566 HG4 was performed as previously described⁹⁵. Briefly, on Day 1, overnight cultures grown
567 from a single colony of MRSA USA300 strain in MHB at 37°C with shaking at 225 rpm was
568 subjected to microbroth dilution susceptibility testing performed using a standard doubling-
569 dilution series of AMP concentrations as described for MIC determination. Cultures from the
570 highest concentration that supported growth were diluted 1:1000 in MHB and used to provide
571 the inoculum for the next passage day. This process was continued for 20 days. Any putative
572 mutants recovered were colony purified for three generations on MHA, prior to further
573 characterization.

574

575 **ATP determination assay**

576 Adenosine triphosphate (ATP) drives many cellular and metabolic processes and can be used
577 to ascertain the integrity of cells^{75,76}. To determine whether HG2 and HG4 affected ATP levels
578 in *S. aureus* treated cells, we used the ATP colorimetric/fluorometric assay kit (Sigma Aldrich)
579 which determines ATP concentration by phosphorylating glycerol, resulting in a colorimetric
580 product that shows absorbance at 570nm. Briefly, in a clear 96 well plate, samples and ATP
581 standard provided with the kit were added in a reaction mix containing ATP assay buffer, ATP
582 Probe and Converter and Developer mix. The reaction mix was incubated in the dark at room
583 temperature for 30 min. The absorbance at 570nm was then measured in a microplate
584 spectrophotometer. The background ATP levels obtained from samples and ATP standards
585 were autocorrected by subtracting ATP levels from blank treatments. The amount of ATP in
586 unknown samples were determined from the ATP standard curve and calculated using the
587 formula in the instruction manual. All samples were tested in triplicates.

588

589 **Membrane permeabilisation assay**

590 Permeabilisation of the bacterial cytoplasmic membrane by HG2 and HG4 peptide was
591 evaluated using the cell-impermeable DNA probe propidium iodide as previously described¹⁵,
592 ⁷⁸. Logarithmic phase bacterial suspension of MRSA USA300 was prepared by diluting over-
593 night cultures in fresh MHA broth (1 in 10 dilution), incubated 3 h at 37°C, 200 rpm, and
594 pelleted by centrifugation for 5 min at 3000 g. Bacterial cell pellet was then resuspended in
595 sterile PBS at about 10⁹ bacteria/ml. Propidium iodide (at 1 mg/ml, Sigma Aldrich) was added
596 to the bacterial suspension at a final concentration of 60 µM. This suspension (100 µl) was then
597 transferred into black 96-well plates already containing 100 µl of serially diluted HG2 or HG4
598 peptide in PBS. Kinetics of fluorescence variations (excitation at 530nm / emission at 590 nm)
599 were then recorded using a microplate reader over 80 min period with incubation at 37°C. Cetyl
600 trimethylammonium bromide (CTAB) (at 300 µM) served as positive control giving 100%
601 permeabilisation. The permeabilisation effect of HG2 and HG4 were expressed as percentage
602 of total permeabilisation.

603

604 **Transmission electron microscopy**

605 Transmission electron microscopy was used to investigate the effects of HG2 and HG4 on *S.*
606 *aureus* cell morphology as described by¹⁵. Mid-log phase *S. aureus*
607 cultures treated with HG2 and HG4 (at 3× MIC for 1 h) were fixed with 2.5% (v/v)
608 glutaraldehyde and post-fixed with 1% osmium tetroxide (w/v). They were then
609 stained with 2% (w/v) uranyl acetate and Reynold's lead citrate after which they were
610 observed using a JEOL JEM1010 transmission electron microscope (JEOL
611 Ltd, Tokyo, Japan) at 80 kV.

612

613 **Haemolytic activity**

614 The ability of HG2 and HG4 to cause leakage of erythrocytes from human whole red blood
615 cells was determined to ascertain probable cytotoxicity to mammalian cells and the suitability
616 of peptides for use as therapeutic agents. The haemolytic activity of HG2 and HG4 was
617 determined as previously described¹⁵. Briefly, fresh human erythrocytes (obtained from Divbio

618 Science Europe, NL) were washed 3 times by centrifugation at 800 g for 5 min with sterile
619 phosphate buffer saline (PBS, pH 7.4). The washed erythrocytes were resuspended in PBS to
620 a final concentration of 8%. 100 µl of human erythrocytes were then added per well into sterile
621 96 well microplate already containing serial dilutions of the peptides in 100 µl of PBS. The
622 treated red blood cells were incubated at 37°C for 1 h and centrifuged at 800 g for 5 min. The
623 supernatant (100 µl) were carefully transferred to a new 96 well microplate and absorbance
624 OD_{450nm} measured using microplate reader. Triton-X100 at 0.1% (v/v) was used as positive
625 control giving 100% haemolysis and haemolysis caused by HG2 and HG4 was expressed as
626 percentage of total haemolysis. The HC₅₀ values for HG2 and HG4 (i.e. the concentration of
627 peptide causing either 50% of haemolysis) were calculated using GraphPad® Prism 7 software.

628 **Peptide-lipid interaction and insertion assay**

629 Peptide-lipid interaction was measured using lipid monolayer formed at the air:water interface
630 with total lipid extracts and pure lipids. Total lipids were extracted from overnight cultures of
631 MRSA or human erythrocytes using Folch extraction procedure as previously described^{15, 78,}
632⁹⁶. Extracted total lipids were dried, resolubilised in chloroform:methanol (2:1, v/v) and stored
633 at -20 °C under nitrogen. Pure prokaryotic and eukaryotic lipids used were: cardiolipin, POPC
634 (1-palmitoyl-2-oleoyl-glycero-3-phosphocholine), POPE (1-palmitoyl-2-oleoyl-sn-glycero-3-
635 phosphoethanolamine) and POPG (1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-(1'-rac-
636 glycerol) (Avanti Polar Lipid). LTA (Lipoteichoic acid from *S. aureus*) and LPS
637 (lipopolysaccharide from *E. coli*) (both obtained from Invitrogen) were also tested. Pure lipids
638 were reconstituted in chloroform:methanol (2:1, v/v) at 1 mg/ml and stored at -20 °C under
639 nitrogen. For peptide-lipid interaction assay, lipid monolayers at the air:water interface were
640 formed by spreading total lipid extract or pure lipids at the surface of 800 µl of sterile PBS
641 using a 50 µl Hamilton's syringe. Lipids were added until the surface pressure reached the
642 desired value. After 5-10 min of incubation allowing the evaporation of the solvent and
643 stabilization of the initial surface pressure, 8 µl of HG2 or HG4 diluted in sterile PBS at 100
644 µg/ml were injected into the 800 µl sub-phase of PBS under the lipid monolayer (pH 7.4,
645 volume 800 µl) using a 10 µl Hamilton's syringe giving a final concentration of peptide of 1

646 $\mu\text{g/ml}$, as preliminary experiments had shown that this concentration was optimal. The
647 variation of the surface pressure caused by peptide insertion was then continuously monitored
648 using a fully automated microtensiometer ($\mu\text{TROUGH SX}$, Kibron Inc) until it reached
649 equilibrium (maximal surface pressure increase being usually obtained after 15-25 min). To
650 reflect physiological situations, the initial surface pressure was fixed at $30 \pm 0.5 \text{ mN/m}$ in some
651 experiments as this value corresponds to a lipid packing density theoretically equivalent to that
652 of the outer leaflet of the eukaryotic and prokaryotic cell membrane ⁹⁷. In other experiments,
653 the critical pressure of insertion of HG2 or HG4 in the total lipid extracts and pure lipids was
654 measured as previously described ^{15, 96}. Briefly, in these experiments, the initial pressure of
655 lipid monolayer was set-up at different values (between 10 and 30 mN/m) and the variation of
656 pressure caused by the injection of peptide was measured. Critical pressure of insertion was
657 calculated by plotting the variation of surface pressure caused by peptide insertion as a function
658 of the initial surface pressure, and corresponds to the theoretical value of initial pressure of
659 lipid monolayer that does not allow the insertion of the peptide, i.e. a variation of pressure
660 equal to 0 mN/m. All experiments were carried out in a controlled atmosphere at $20 \text{ }^\circ\text{C} \pm 1 \text{ }^\circ\text{C}$
661 and data were analyzed using the Filmware 2.5 program (Kibron Inc.). The accuracy of the
662 system under our experimental conditions was determined to be $\pm 0.25 \text{ mN/m}$ for surface
663 pressure measurements.

664

665 **Human cell culture and cytotoxicity studies**

666 The toxicity of HG2 and HG4 was tested as previously described ⁹⁸⁻¹⁰⁰. The following human
667 cells were used: BEAS-2B (normal airway epithelial cells, ATCC CRL-9609), IMR-90
668 (normal fibroblasts, ATCC CCL-186) and HepG2 (liver cell line, ATCC HB-8065). BEAS-
669 2B, IMR-90 and HepG2 cells were cultured in Dulbecco's modified essential medium (DMEM)
670 supplemented with 10% fetal calf serum (FCS), 1% L-glutamine and 1% antibiotics (all from
671 Invitrogen). Cells were routinely grown onto 25 cm² flasks maintained in a 5% CO₂ incubator
672 at 37°C. Briefly, cells grown on 25 cm² flasks were detached using trypsin-EDTA solution
673 (Thermofisher) and seeded into 96-well cell culture plates (Greiner Bio-one) at approximately

674 10^4 cells per well (counted using Mallasez's chamber). The cells were grown at 37°C in a 5%
675 CO₂ incubator until they reached confluence (approximately 48-72 h of seeding). Wells were
676 then aspirated and increasing concentrations of HG2 or HG4 were added to the cells and
677 incubated for a further 48 h at 37°C in a 5% CO₂ incubator. The wells were then emptied, and
678 cell viability was evaluated using resazurin based *in vitro* toxicity assay kit (Sigma-Aldrich)
679 following manufacturer's instructions. Briefly, resazurin stock solution was diluted 1:100 in
680 sterile PBS containing calcium and magnesium (PBS⁺⁺, pH 7.4) and emptied wells were filled
681 with 100 µl of the resazurin diluted solution. After 4 h incubation at 37°C with the peptide
682 treated cells, fluorescence intensity was measured using microplate reader (excitation
683 wavelength of 530 nm/emission wavelength of 590 nm). The fluorescence values were
684 normalized by the controls and expressed as percentage of cell viability. The IC₅₀ values of
685 HG2 or HG4 on cell viability (i.e. the concentration of peptides causing a reduction of 50% of
686 the cell viability) were calculated using GraphPad® Prism 7 software.

687

688 ***In vivo* efficacy in *Galleria* model**

689 All procedures of larvae rearing, injection and *G. mellonella* killing assays were conducted as
690 previously described¹⁰¹. In each assay, ten (10) larvae weighing between 280 - 300 mg each
691 were randomly selected. Larvae with previous melanisation of the cuticle were not used in the
692 experiments. All the experiments were designed in at least four experimental and biologic
693 replicates.

694 Firstly, we evaluated the putative toxic effect of the peptides HG2 and HG4 in *G. mellonella*.
695 Each peptide solution prepared in sterile H₂O was injected in larvae at the concentration of 1x
696 MIC, as 16 mg/kg of larvae body weight (LBW) for HG2 and 32 mg/kg LBW for HG4; and
697 3x MIC 64 mg/kg LBW HG2 and 98 mg/kg LBW HG4 as previously determined. After
698 injection with peptides, the larvae were maintained at 37°C in the dark. Phenotypic aspects
699 such as melanisation and mobility of the larvae as well as survival were monitored every 24
700 hours for 96 hours. Larvae not inoculated with APMs were used as controls.

701 To determine the LD₅₀ and LD of *S. aureus* MRSA USA300 in *G. mellonella*, an inoculum of
702 10 µl of MRSA USA300 suspension in PBS 1X (10³ to 10⁶ CFU/larva) was injected into the
703 larvae hemocoel using insulin syringes (Becton Dickinson, USA). Larvae inoculated with PBS
704 and larvae not inoculated were used as negative controls. After the injections, the larvae were
705 maintained at 37°C in the dark, the survival was recorded every 24 hours for 96 hours and the
706 LD₅₀ and LD were determined.

707 The efficacy of the peptides HG2 and HG4 to control MRSA USA300 infection *in vivo*, was
708 evaluated following the protocol by Peleg et al.,⁸⁴. The groups of *G. mellonella* larvae were
709 infected with the predetermined lethal dose of MRSA USA300. After 30 minutes of the larvae
710 injection with the bacteria, the peptides were injected at 1x MIC and 3x MIC concentrations
711 respectively. Larvae injected with PBS solution, and HG2 or HG4 peptides solution (1x MIC
712 or 3x MIC) alone were used as negative controls. Injected larvae were maintained at 37°C in
713 the dark and their survival was monitored and analyzed as above.

714

715

716 **Statistical analysis**

717 All biological experiments were repeated at least three times and three biological replicates
718 were used wherever applicable. Results are expressed as mean ± standard error. The MRSA
719 USA300 LD₅₀ value was calculated by linear regression using software R v.2.13.0¹⁰². The
720 Kaplan–Meier method was used to plot the survival curves. Differences in survival were
721 calculated using the log-rank test with the software SigmaPlot from Systat Software Inc., San
722 Jose, California¹⁰³. A *P*-value of 0.05 was considered to be statistically significant.

723

724 **Author Contributions**

725 LO and SH, and CC conceived the project. LO, with help from HV, TW and MW, completed
726 the laboratory work under supervision of SH and CC. AC and NF assisted LO with
727 transmission electron microscopy and 3D structural modelling respectively. LO, HO, MM and
728 JP completed the membrane permeabilisation, cytotoxicity and lipid interaction assays. LO,

729 and JAP completed the ATP assays. KH and RM assisted LO with screening of metagenomic
730 library and identification of AMP candidates. AT, MG completed *in vivo* work in *G. mellonella*
731 with input from LO and supervision by DB, HM and SH. MH and HM have provided valuable
732 ideas into the project from the time of conception. LO wrote the paper with input from all co-
733 authors.

734 **Acknowledgement**

735 This project was funded partly by the Cross River State Government of Nigeria, the Life
736 Sciences Research Network Wales, RCUK Newton Institutional Link Fund (172629373), and
737 the BBSRC UK (BB/L026716/1). HM thanks the Coordination for the Improvement of Higher
738 Education Personnel (CAPES) for providing the Joint Institutional Links grant. KH thanks the
739 Institute of Infection and Immunity of St. George's University of London for a start-up grant.
740 We are also grateful to Dr Colin Greengrass, for his advice. The authors declare no competing
741 financial or other interests.

742 **Competing Interests**

743 The authors declare no competing interests.

744

745 References

- 746
- 747 1. J. O'Neill, *Tackling Drug-Resistant Infections Globally: final report and*
748 *recommendations*, United Kingdom, 2016.
 - 749 2. K. Becker, F. Schaumburg, C. Fegeler, A. W. Friedrich and R. Köck, *International*
750 *Journal of Medical Microbiology*, 2017, **307**, 21-27.
 - 751 3. A. S. Lee, H. de Lencastre, J. Garau, J. Kluytmans, S. Malhotra-Kumar, A. Peschel
752 and S. Harbarth, *Nature Reviews Disease Primers*, 2018, **4**, 18033.
 - 753 4. P. O. Lewis, E. L. Heil, K. L. Covert and D. B. Cluck, 2018, **43**, 614-625.
 - 754 5. H. F. L. Wertheim, D. C. Melles, M. C. Vos, W. van Leeuwen, A. van Belkum, H. A.
755 Verbrugh and J. L. Nouwen, *The Lancet Infectious Diseases*, 2005, **5**, 751-762.
 - 756 6. WHO, *Global priority list of antibiotic-resistant bacteria to guide research,*
757 *discovery, and development of new antibiotics*, Geneva, 2017.
 - 758 7. H. F. Chambers and F. R. DeLeo, *Nature Reviews Microbiology*, 2009, **7**, 629.
 - 759 8. C. f. D. C. a. P. CDC, *Journal*, 2013.
 - 760 9. P. H. E. PHE, *Laboratory surveillance of Staphylococcus aureus bacteraemia in*
761 *England, Wales and Northern Ireland: 2017*, United Kingdom, 2018.
 - 762 10. J. D. Steckbeck, B. Deslouches and R. C. Montelaro, *Expert opinion on biological*
763 *therapy*, 2014, **14**, 11-14.
 - 764 11. S. A. Baltzer and M. H. Brown, *J Mol Microbiol Biotechnol*, 2011, **20**, 228-235.
 - 765 12. E. M. Ross, P. J. Moate, C. R. Bath, S. E. Davidson, T. I. Sawbridge, K. M.
766 Guthridge, B. G. Cocks and B. J. Hayes, *BMC genetics*, 2012, **13**, 53.
 - 767 13. A. C. Azevedo, C. B. Bento, J. C. Ruiz, M. V. Queiroz and H. C. Mantovani, *Appl*
768 *Environ Microbiol*, 2015, **81**, 7290-7304.
 - 769 14. L. B. Oyama, J. A. Crochet, J. E. Edwards, S. E. Girdwood, A. R. Cookson, N.
770 Fernandez-Fuentes, K. Hilpert, P. N. Golyshin, O. V. Golyshina, F. Prive, M. Hess,
771 H. C. Mantovani, C. J. Creevey and S. A. Huws, *Front Chem*, 2017, **5**, 51.
 - 772 15. L. B. Oyama, S. E. Girdwood, A. R. Cookson, N. Fernandez-Fuentes, F. Prive, H. E.
773 Vallin, T. J. Wilkinson, P. N. Golyshin, O. V. Golyshina, R. Mikut, K. Hilpert, J.
774 Richards, M. Wootton, J. E. Edwards, M. Maresca, J. Perrier, F. T. Lundy, Y. Luo, M.
775 Zhou, M. Hess, H. C. Mantovani, C. J. Creevey and S. A. Huws, *NPJ Biofilms*
776 *Microbiomes*, 2017, **3**, 33.
 - 777 16. J. C. McCann, T. A. Wickersham and J. J. Loor, *Bioinformatics and biology insights*,
778 2014, **8**, 109-125.
 - 779 17. R. Rezaei Javan, A. J. van Tonder, J. P. King, C. L. Harrold and A. B. Brueggemann,
780 2018, **9**.
 - 781 18. J. M. Brulc, D. A. Antonopoulos, M. E. Miller, M. K. Wilson, A. C. Yannarell, E. A.
782 Dinsdale, R. E. Edwards, E. D. Frank, J. B. Emerson, P. Wacklin, P. M. Coutinho, B.
783 Henrissat, K. E. Nelson and B. A. White, *Proc Natl Acad Sci U S A*, 2009, **106**, 1948-
784 1953.
 - 785 19. M. Hess, A. Sczyrba, R. Egan, T. W. Kim, H. Chokhawala, G. Schroth, S. Luo, D. S.
786 Clark, F. Chen, T. Zhang, R. I. Mackie, L. A. Pennacchio, S. G. Tringe, A. Visel, T.
787 Woyke, Z. Wang and E. M. Rubin, *Science*, 2011, **331**, 463-467.
 - 788 20. Y. Cheng, Y. Wang, Y. Li, Y. Zhang, T. Liu, Y. Wang, T. J. Sharpton and W. Zhu,
789 *Frontiers in Microbiology*, 2017, **8**, 2165.
 - 790 21. M. Ferrer, A. Ghazi, A. Beloqui, J. M. Vieites, N. López-Cortés, J. Marín-Navarro, T.
791 Y. Nechitaylo, M.-E. Guazzaroni, J. Polaina, A. Waliczek, T. N. Chernikova, O. N.
792 Reva, O. V. Golyshina and P. N. Golyshin, *PLoS ONE*, 2012, **7**, e38134.
 - 793 22. K.-C. Ko, J. H. Lee, Y. Han, J. H. Choi and J. J. Song, *Biochemical and Biophysical*
794 *Research Communications*, 2013, **441**, 567-572.
 - 795 23. H. J. Lee, J. Y. Jung, Y. K. Oh, S. S. Lee, E. L. Madsen and C. O. Jeon, *Appl Environ*
796 *Microbiol*, 2012, **78**, 5983-5993.

- 797 24. L. D. Lopes, A. O. de Souza Lima, R. G. Taketani, P. Darias, L. R. F. da Silva, E. M.
798 Romagnoli, H. Louvandini, A. L. Abdalla and R. Mendes, *Antonie van Leeuwenhoek*,
799 2015, **108**, 15-30.
- 800 25. K. M. Singh, B. Reddy, D. Patel, A. K. Patel, N. Parmar, A. Patel, J. B. Patel and C.
801 G. Joshi, *BioMed Research International*, 2014, **2014**, 267189.
- 802 26. Y.-H. Song, K.-T. Lee, J.-Y. Baek, M.-J. Kim, M.-R. Kwon, Y.-J. Kim, M.-R. Park,
803 H. Ko, J.-S. Lee and K.-S. J. F. M. Kim, 2017, **62**, 175-181.
- 804 27. K.-T. Lee, S. H. Tousehik, J.-Y. Baek, J.-E. Kim, J.-S. Lee and K.-S. Kim, *Journal of*
805 *Agricultural and Food Chemistry*, 2018, **66**, 9034-9041.
- 806 28. F. Privé, C. J. Newbold, N. N. Kaderbhai, S. G. Girdwood, O. V. Golyshina, P. N.
807 Golyshin, N. D. Scollan and S. A. Huws, *Applied Microbiology and Biotechnology*,
808 2015, **99**, 5475-5485.
- 809 29. M. C. Rodríguez, I. Loaces, V. Amarelle, D. Senatore, A. Iriarte, E. Fabiano and F.
810 Noya, *PLoS ONE*, 2015, **10**, e0126651.
- 811 30. L. Ufarté, E. Laville, S. Duquesne, D. Morgavi, P. Robe, C. Klopp, A. Rizzo, S.
812 Pizzut-Serin and G. Potocki-Veronese, *PLoS ONE*, 2017, **12**, e0189201.
- 813 31. S. Bayer, A. Kunert, M. Ballschmiter and T. Greiner-Stoeffele, *Journal of Molecular*
814 *Microbiology and Biotechnology*, 2010, **18**, 181-187.
- 815 32. S. Zhao, J. Wang, K. Liu, Y. Zhu, D. Bu, D. Li and P. Yu, *Sheng Wu Gong Cheng*
816 *Xue Bao*, 2009, **25**, 869-874.
- 817 33. G. Wang, X. Li and Z. Wang, *Nucleic Acids Res*, 2016, **44**, D1087-1093.
- 818 34. G. Wang, X. Li and Z. Wang, *Nucleic Acids Research*, 2009, **37**, D933-D937.
- 819 35. S. Ramon-Garcia, R. Mikut, C. Ng, S. Ruden, R. Volkmer, M. Reischl, K. Hilpert and
820 C. J. Thompson, *Antimicrob Agents Chemother*, 2013, **57**, 2295-2303.
- 821 36. S. F. Altschul, W. Gish, W. Miller, E. W. Myers and D. J. Lipman, *Journal of*
822 *Molecular Biology*, 1990, **215**, 403-410.
- 823 37. S. H. Lee, D. G. Lee, S. T. Yang, Y. Kim, J. I. Kim, K. S. Hahm and S. Y. Shin,
824 *Protein Pept Lett*, 2002, **9**, 395-402.
- 825 38. R. Liu, H. Liu, Y. Ma, J. Wu, H. Yang, H. Ye and R. Lai, *J Proteome Res*, 2011, **10**,
826 1806-1815.
- 827 39. K. Hilpert, D. F. H. Winkler and R. E. W. Hancock, *Nat. Protocols*, 2007, **2**, 1333-
828 1349.
- 829 40. D. A. Devine and R. E. Hancock, *Curr Pharm Des*, 2002, **8**, 703-714.
- 830 41. R. E. Hancock, *Expert Opin Investig Drugs*, 2000, **9**, 1723-1729.
- 831 42. T. Anunthawan, C. de la Fuente-Nunez, R. E. Hancock and S. Klaynongsruang,
832 *Biochim Biophys Acta*, 2015, **1848**, 1352-1358.
- 833 43. E. Gasteiger, C. Hoogland, A. Gattiker, S. e. Duvaud, M. R. Wilkins, R. D. Appel and
834 A. Bairoch, in *The Proteomics Protocols Handbook*, ed. J. M. Walker, Humana Press,
835 Totowa, NJ, 2005, DOI: 10.1385/1-59259-890-0:571, pp. 571-607.
- 836 44. R. E. Hancock, *Lancet*, 1997, **349**, 418-422.
- 837 45. R. E. Hancock, K. L. Brown and N. Mookherjee, *Immunobiology*, 2006, **211**, 315-
838 322.
- 839 46. R. E. Hancock, T. Falla and M. Brown, *Adv Microb Physiol*, 1995, **37**, 135-175.
- 840 47. K. Hilpert, B. McLeod, J. Yu, M. R. Elliott, M. Rautenbach, S. Ruden, J. Burck, C.
841 Muhle-Goll, A. S. Ulrich, S. Keller and R. E. Hancock, *Antimicrob Agents*
842 *Chemother*, 2010, **54**, 4480-4483.
- 843 48. L. Tzong-Hsien, N. H. Kristopher and A. Marie-Isabel, *Current Topics in Medicinal*
844 *Chemistry*, 2016, **16**, 25-39.
- 845 49. J. Maupetit, P. Derreumaux and P. Tuffery, *Nucleic Acids Res*, 2009, **37**, W498-503.
- 846 50. K. A. Brogden, *Nat Rev Microbiol*, 2005, **3**, 238-250.
- 847 51. M. Gough, R. E. Hancock and N. M. Kelly, *Infect Immun*, 1996, **64**, 4922-4927.
- 848 52. K. L. Piers, M. H. Brown and R. E. Hancock, *Antimicrob Agents Chemother*, 1994,
849 **38**, 2311-2316.

- 850 53. W. F. Porto, L. Irazazabal, E. S. F. Alves, S. M. Ribeiro, C. O. Matos, Á. S. Pires, I.
851 C. M. Fensterseifer, V. J. Miranda, E. F. Haney, V. Humblot, M. D. T. Torres, R. E.
852 W. Hancock, L. M. Liao, A. Ladram, T. K. Lu, C. de la Fuente-Nunez and O. L.
853 Franco, *Nature Communications*, 2018, **9**, 1490.
- 854 54. J. Alder and B. Eisenstein, *Curr Infect Dis Rep*, 2004, **6**, 251-253.
- 855 55. P. Y. Chung and R. Khanum, *Journal of Microbiology, Immunology and Infection*,
856 2017, **50**, 405-410.
- 857 56. N. Delattin, K. Brucker, K. Cremer, B. P. Cammue and K. Thevissen, *Curr Top Med*
858 *Chem*, 2017, **17**, 604-612.
- 859 57. I. Di Bonaventura, X. Jin, R. Visini, D. Probst, S. Javor, B. H. Gan, G. Michaud, A.
860 Natalello, S. M. Doglia, T. Kohler, C. van Delden, A. Stocker, T. Darbre and J. L.
861 Reymond, *Chem Sci*, 2017, **8**, 6784-6798.
- 862 58. M. Di Luca, G. Maccari and R. Nifosi, *Pathogens and Disease*, 2014, **70**, 257-270.
- 863 59. M. K. Kim, H. K. Kang, S. J. Ko, M. J. Hong, J. K. Bang, C. H. Seo and Y. Park,
864 *Scientific Reports*, 2018, **8**, 1763.
- 865 60. G. Michaud, R. Visini, M. Bergmann, G. Salerno, R. Bosco, E. Gillon, B. Richichi, C.
866 Nativi, A. Imberty, A. Stocker, T. Darbre and J. L. Reymond, *Chem Sci*, 2016, **7**, 166-
867 182.
- 868 61. S.-G. Susana and M.-d.-T. Guillermo, *Current Topics in Medicinal Chemistry*, 2017,
869 **17**, 590-603.
- 870 62. O. Fleitas, C. M. Agbale and O. L. Franco, *Front Biosci (Landmark Ed)*, 2016, **21**,
871 1013-1038.
- 872 63. H.-S. Joo, C.-I. Fu and M. Otto, *Philosophical Transactions of the Royal Society B:*
873 *Biological Sciences*, 2016, **371**, 20150292.
- 874 64. J. Dobias, L. Poirel and P. Nordmann, *Clinical Microbiology and Infection*, 2017, **23**,
875 676.e671-676.e675.
- 876 65. S. A. Loutet and M. A. Valvano, *Frontiers in Cellular and Infection Microbiology*,
877 2011, **1**, 6.
- 878 66. S. Matamouros and S. I. Miller, *Biochimica et biophysica acta*, 2015, **1848**, 3021-
879 3025.
- 880 67. H.-S. Joo and M. Otto, *Biochimica et biophysica acta*, 2015, **1848**, 3055-3061.
- 881 68. S. Maria-Neto, K. C. de Almeida, M. L. R. Macedo and O. L. Franco, *Biochimica et*
882 *Biophysica Acta (BBA) - Biomembranes*, 2015, **1848**, 3078-3088.
- 883 69. M. Juhas, *Critical Reviews in Microbiology*, 2015, **41**, 101-108.
- 884 70. G. G. Perron, M. Zasloff and G. Bell, *Proceedings of the Royal Society B: Biological*
885 *Sciences*, 2006, **273**, 251-256.
- 886 71. R. H. Baltz, *Current Opinion in Chemical Biology*, 2009, **13**, 144-151.
- 887 72. C. M. Ernst and A. Peschel, 2011, **80**, 290-299.
- 888 73. L. L. Ling, T. Schneider, A. J. Peoples, A. L. Spoering, I. Engels, B. P. Conlon, A.
889 Mueller, T. F. Schaberle, D. E. Hughes, S. Epstein, M. Jones, L. Lazarides, V. A.
890 Steadman, D. R. Cohen, C. R. Felix, K. A. Fetterman, W. P. Millett, A. G. Nitti, A.
891 M. Zullo, C. Chen and K. Lewis, *Nature*, 2015, **517**, 455-459.
- 892 74. E. Alirol and J. C. Martinou, *Oncogene*, 2006, **25**, 4706-4716.
- 893 75. P. Mitchell, *Nature*, 1961, **191**, 144.
- 894 76. P. Mitchell, *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 2011, **1807**, 1507-
895 1538.
- 896 77. S. Gottschalk, C. T. Gottlieb, M. Vestergaard, P. R. Hansen, L. Gram, H. Ingmer and
897 L. E. Thomsen, *J Med Microbiol*, 2015, **64**, 1504-1513.
- 898 78. E. Di Pasquale, C. Salmi-Smail, J.-M. Brunel, P. Sanchez, J. Fantini and M. Maresca,
899 *Chemistry and Physics of Lipids*, 2010, **163**, 131-140.
- 900 79. W. E. Herrell and D. Heilman, *Journal of Clinical Investigation*, 1941, **20**, 583-591.
- 901 80. R. Mösges, C. M. Baues, T. Schröder and K. Sahin, *Current Medical Research and*
902 *Opinion*, 2011, **27**, 871-878.

- 903 81. J. Swierstra, V. Kapoerchan, A. Knijnenburg, A. van Belkum and M. Overhand,
904 *European Journal of Clinical Microbiology & Infectious Diseases*, 2016, **35**, 763-769.
- 905 82. A. P. Desbois and P. J. Coote, *Adv Appl Microbiol*, 2012, **78**, 25-53.
- 906 83. K. Ignasiak and A. Maxwell, *BMC Res Notes*, 2017, **10**, 428.
- 907 84. A. Y. Peleg, S. Jara, D. Monga, G. M. Eliopoulos, R. C. Moellering, Jr. and E.
908 Mylonakis, *Antimicrob Agents Chemother*, 2009, **53**, 2605-2609.
- 909 85. C. J. Tsai, J. M. Loh and T. Proft, *Virulence*, 2016, **7**, 214-229.
- 910 86. R. Mikut, O. Burmeister, S. Braun and M. P. Reischl, Potsdam, 2008.
- 911 87. R. Mikut, A. Bartschat, W. Doneit, J. Á. González Ordiano, B. Schott, J. Stegmaier,
912 S. Waczowicz and M. Reischl, *arXiv:1704.03298*, 2017.
- 913 88. R. Mikut, *Methods in molecular biology*, 2010, **618**, 287-299.
- 914 89. R. Mikut and K. Hilpert, *Int J Pept Res Ther*, 2009, **15**, 129-137.
- 915 90. J. Maupetit, P. Tuffery and P. Derreumaux, *Proteins: Structure, Function, and*
916 *Bioinformatics*, 2007, **69**, 394-408.
- 917 91. L. L. C. Schrödinger, *Journal*, 2010.
- 918 92. I. Wiegand, K. Hilpert and R. E. Hancock, *Nature protocols*, 2008, **3**, 163-175.
- 919 93. J. H. Jorgensen, *Clinical Microbiology Newsletter*, 2006, **28**, 153-157.
- 920 94. B. Oliva, K. Miller, N. Caggiano, A. J. O'Neill, G. D. Cuny, M. Z. Hoemann, J. R.
921 Hauske and I. Chopra, *Antimicrob Agents Chemother*, 2003, **47**, 458-466.
- 922 95. L. Friedman, J. D. Alder and J. A. Silverman, *Antimicrobial Agents and*
923 *Chemotherapy*, 2006, **50**, 2137-2145.
- 924 96. R. Mahfoud, M. Maresca, M. Santelli, A. Pfohl-Leszkowicz, A. Puigserver and J.
925 Fantini, *Journal of Agricultural and Food Chemistry*, 2002, **50**, 327-331.
- 926 97. S. R. Dennison, Y. S. Kim, H. J. Cha and D. A. Phoenix, *Eur Biophys J*, 2008, **38**, 37-
927 43.
- 928 98. E. H. Ajandouz, S. Berdah, V. Moutardier, T. Bege, D. J. Birnbaum, J. Perrier, E. Di
929 Pasquale and M. Maresca, *Toxins.*, 2016, **8**, 232.
- 930 99. R. Borie, C. Kannengiesser, F. Sicre de Fontbrune, D. Boutboul, L. Tabeze, F.
931 Brunet-Possenti, E. Lainey, M. P. Debray, A. Cazes and B. Crestani, *Eur Respir J*,
932 2017, **50**.
- 933 100. H. Razafimanjato, N. Garmy, X.-J. Guo, K. Varini, C. Di Scala, E. Di Pasquale, N.
934 Taïeb and M. Maresca, *NeuroToxicology*, 2010, **31**, 475-484.
- 935 101. M. F. Pereira, C. C. Rossi, M. V. de Queiroz, G. F. Martins, C. Isaac, J. T. Bosse, Y.
936 Li, B. W. Wren, V. S. Terra, J. Cuccui, P. R. Langford and D. M. Bazzolli,
937 *Microbiology*, 2015, **161**, 387-400.
- 938 102. R Core Team, *R: A language and environment for statistical computing. R*
939 *Foundation for Statistical Computing*, Vienna, Austria., 2013.
- 940 103. Systat Software Inc, *SYSTAT* San Jose, California USA.

941
942
943
944

945 **Figure Legends**

946 **Fig. 1 a)** visualization of distances for AA acids (AAD) and AA pairs (AAPD) for the first
947 68,274 sequences from library "Cow"¹⁹ meeting the first selection criteria: candidates with
948 AAD<0.2 or AAPD<1.45 are selected as candidates (here: 65). **b)** standard hydrophobicity
949 (TERM1 SEQ Hopp-Woods) - loading (positively charged, TERM3 SEQ Isoelectric Point)
950 plot. Blue dots are known AMPs (library "AMP" consisting of AMPs from the APD2³⁴ and
951 Hilpert Library³⁵), green colored signs are AMP hits identified from library "Cow", and finally
952 selected peptides HG2 (magenta) and HG4 (red).

953 **Fig. 2 Predicted 3D structures for peptides: a) HG2, b) HG4.** Main-chain and side chains
954 depicted in ribbon and stick representation respectively and coloured according to atom type:
955 Carbon, Oxygen and Nitrogen in green, red and blue respective. Two orientations are shown
956 rotated about the shown axis. Ct and Nt as well as selected residues are depicted in the figure.
957 Figures were rendered using PyMol.

958 **Fig. 3 Antimicrobial susceptibility and activity of HG2 and HG4.** **a)** Time dependent kill
959 of MRSA USA300 cells by AMPs at 3x MIC concentration. Dashed lines indicate limit of
960 detection. **b)** Anti-biofilm activity against MRSA USA300 biofilms: *, ** and *** ($P \leq 0.05$,
961 0.01 and 0.001 respectively- significantly different from untreated cells (positive). **c)**
962 Resistance acquisition during serial passaging of MRSA USA300 cells in the presence of sub-
963 MIC levels of antimicrobials. The y axis is the fold change in MIC during passaging. For
964 mupirocin, 32x MIC was the highest concentration tested. The figure is representative of 3
965 independent experiments. **d)** ATP depletion activity in MRSA USA300 cells.

966 **Fig. 4 Membrane permeabilisation action of HG2 and HG4 against MRSA: a)** Membrane
967 permeabilization activity of HG2 at different concentrations ($\mu\text{g/ml}$) against MRSA USA300
968 cells measured by propidium iodide assay over time. **b)** Membrane permeabilization activity
969 of HG4 at different concentrations ($\mu\text{g/ml}$) against MRSA USA300 cells measured by
970 propidium iodide assay over time. **c)** Determination of EC50 (Effective Concentration 50) of

971 HG2 and HG4 membrane permeabilisation measured after 80 min. **d)** Membrane
972 permeabilisation kinetics of HG2 and HG4 at their MIC concentration. In all cases, values are
973 from three independent replicates; results are expressed as means \pm standard deviation).

974 **Fig. 5 Representative transmission electron micrographs of MRSA cells.** **a)** micrographs
975 untreated MRSA USA300 cells. **b)** HG2 treated (3x MIC for 1 h) MRSA USA300 cells. **c)**
976 HG4 treated (3x MIC for 1 h) MRSA USA300 cells. Scale bars are 200 or 500 nm as shown
977 on micrographs.

978 **Fig. 6 Peptide lipid interaction and insertion measurements:** **a)** Interaction of HG2 and
979 HG4 (at 1 $\mu\text{g}/\text{mL}$ final concentration) with lipids (either total lipid extracts or pure lipids) was
980 measured using lipid monolayers. **a)** interaction HG2 and HG4 with total MRSA lipid extract.
981 **b)** interaction HG2 and HG4 with total lipid extract from human erythrocytes. **c)** interaction of
982 HG2 with pure lipids and **c)** interaction of HG4 with pure lipids. 1-palmitoyl-2-oleoyl-sn-
983 glycerol-3-phospho-(1'-rac-glycerol) (PG), 1-palmitoyl-2-oleoyl-sn-glycerol-3-
984 phosphoethanolamine (PE), Cardiolipin (Cardio), Lipoteichoic acid (LTA) from *S. aureus*,
985 Lipopolysaccharide (LPS) from *E. coli* and (1-palmitoyl-2-oleoyl-glycerol-3-phosphocholine
986 (PC).

987 **Fig. 7 In vivo efficacy assessment in *G. mellonella* MRSA infection model:** **a)** representative
988 images of toxicity assay of peptides- **i)** HG2, and **ii)** HG4 in *G. mellonella* 120 h post treatment
989 with 3x MIC concentrations. The larvae remained alive and without melanisation. **iii)** virulence
990 assay of MRSA USA300 in *G. mellonella* using a lethal dose inoculum of 10^6 CFU/per larvae-
991 **iii)** 24 hours post infection: some larvae were dead and partial melanisation was observed. **iv).**
992 48 hours post infection: most larvae were dead and complete melanisation was evident. The
993 experiment was done with three experimental replicates, each containing groups of 10 larvae.
994 **b)** Kaplan-Meier survival curves of *G. mellonella* infected with a lethal dose of *S. aureus* (2.25
995 $\times 10^6$ CFU/larvae) and treated with placebo (showing a 100% larvae survival rate) or peptides
996 HG2 and HG4 at a 1x and 3x MIC concentrations.

997 **Supporting Information Legends**

998
999 **Fig. S1. Mass spectrometry and certificate of analysis (COA) report for AMP chemical synthesis:**

1000 a) HG2 and b) HG4. Peptides were synthesised by GenScript Inc. USA. Mass spectrum analysis and
1001 COA report show correct molecular weight and purity grade for both peptides.

1002

1003 **Table. S1 Summary of results from computation steps**

1004

1005 **Table. S2 Six most promising antimicrobial peptide candidates (HG1-HG6) identified in**
1006 **cow rumen metagenome dataset**

1007

1008 **Table. S3 Scaffold nucleotide sequences from which putative antimicrobials H-G1- H-G6 were**
1009 **identified**

1010

1011 **Table. S4 Peptide lipid interaction and insertion measurements:** interaction of peptides, HG2 and
1012 HG4 with total MRSA and erythrocyte lipid extracts, and interaction of HG2 and HG4 (at 1 $\mu\text{g}/\text{mL}$ final
1013 concentration) with pure lipids the initial surface pressure of lipid monolayer. Maximal variation of
1014 surface pressure induced by the injection of peptide in lipid monolayer with initial surface pressure of
1015 30 \pm 0.5 mN/m.