nature biomedical engineering

ARTICLES

https://doi.org/10.1038/s41551-021-00801-1

Mining for encrypted peptide antibiotics in the human proteome

Marcelo D. T. Torres[1,2,3], Marcelo C. R. Melo[1,2,3], Orlando Crescenzi[4], Eugenio Notomista[5] and Cesar de la Fuente-Nunez[1,2,3]✉

The emergence of drug-resistant bacteria calls for the discovery of new antibiotics. Yet, for decades, traditional discovery strategies have not yielded new classes of antimicrobial. Here, by mining the human proteome via an algorithm that relies on the sequence length, net charge, average hydrophobicity and other physicochemical properties of antimicrobial peptides, we report the identification of 2,603 encrypted peptide antibiotics that are encoded in proteins with biological function unrelated to the immune system. We show that the encrypted peptides kill pathogenic bacteria by targeting their membrane, modulate gut and skin commensals, do not readily select for bacterial resistance, and possess anti-infective activity in skin abscess and thigh infection mouse models. We also show, in vitro and in the two mouse models of infection, that encrypted antibiotic peptides from the same biogeographical area display synergistic antimicrobial activity. Our algorithmic strategy allows for the rapid mining of proteomic data and opens up new routes for the discovery of candidate antibiotics.

According to the Centers for Disease Control and Prevention, in 2019, 2.8 million antibiotic-resistant infections occurred in the United States, leading to approximately 35,000 deaths[1]. Such untreatable infections are projected to reach 10 million people per year worldwide, becoming the leading cause of death in our society[2]. This daunting scenario coincides with the lack of innovation in antibiotic discovery. Most antibiotics available today have been used for over 30 years. These drugs often have unintended side effects, readily select for antibiotic resistance and, in the face of this resistance, are losing effectiveness[3]. Thus, there is an urgent need to discover new antimicrobial agents to target drug-resistant infections[4].

A broad array of computational methods have been developed to expedite drug development, usually focusing on small-molecule docking and optimization[5]. However, the application of such methods for antibiotic discovery is still in its infancy[6]. The computer-aided design of antimicrobial peptides (AMPs)[7,8] has surged as a promising source of new bioactive compounds, which could provide alternatives to conventional antibiotics. AMPs are small molecules (8–50 amino acid residues in length) produced by virtually all living organisms[9], usually presenting amphipathic and cationic sequences. Here, we used certain physicochemical features of amino acids to inform a scoring function that comprehensively searched the human proteome for novel antimicrobials (Fig. 1a). To find novel antibiotic scaffolds, we intentionally avoided the use of known conserved amino acid sequences (that is, patterns and motifs) found in previously described AMPs. Instead, we utilized physicochemical properties to guide our discovery approach. This allowed us to focus on the balance of key physicochemical features to discover a new region of peptide sequence space populated with antimicrobial sequences (Supplementary Fig. 1). Our search algorithm was based on key physicochemical determinants of AMPs, including net charge, average hydrophobicity, and sequence length, integrated into a fitness function[10] that selects for antimicrobial sequences. Antimicrobial potency is modelled as being linearly dependent on physicochemical properties raised to exponents previously fitted using known AMPs[10]. Using this scoring function, the entire human proteome, including protein isoforms, was scanned to find peptides ranging from 8 to 50 residues in length, with predicted antimicrobial activity. From the hundreds of millions of possible peptides within 42,361 protein sequences, we identified a total of 43,000 peptide candidates for further study (Supplementary Information). The Panther Classification System[11] was then applied to analyse genes coding for peptides with predicted antimicrobial activity, where we found an overrepresentation of genes for secreted or membrane-bound proteins, and an underrepresentation of DNA-bound and gene regulatory proteins (Fig. 1b). This observation supports our hypothesis that encrypted peptides may have been evolutionarily selected over millions of years of evolution to minimize genomic expansion and maintained in proteins that are exposed to the extracellular environment, where they are more likely to display their antimicrobial activity when encountering pathogens upon injury and infection.

Results and discussion

Focusing our search on the most potent and promising encrypted peptides in the human proteome, we thoroughly scanned the sequences of secreted proteins to discover 2,603 predicted peptide antibiotics (Supplementary Table 1). The identified molecules belonged to two major classes: (1) previously undescribed encrypted peptides derived from proteins with various biological functions, such as plasma proteins, coagulation factors, protein inhibitors, enzymes and signalling cascade factors; (2) peptide hormones with well-described putative functions but previously undisclosed antibiotic properties, such as neuropeptides, regulators of G-protein-coupled receptors (GPCRs), and diuretic hormones

[1]Machine Biology Group, Departments of Psychiatry and Microbiology, Institute for Biomedical Informatics, Institute for Translational Medicine and Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [2]Departments of Bioengineering and Chemical and Biomolecular Engineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA, USA. [3]Penn Institute for Computational Science, University of Pennsylvania, Philadelphia, PA, USA. [4]Department of Chemical Sciences, University of Naples Federico II, Naples, Italy. [5]Department of Biology, University of Naples Federico II, Naples, Italy. ✉e-mail: cfuente@upenn.edu

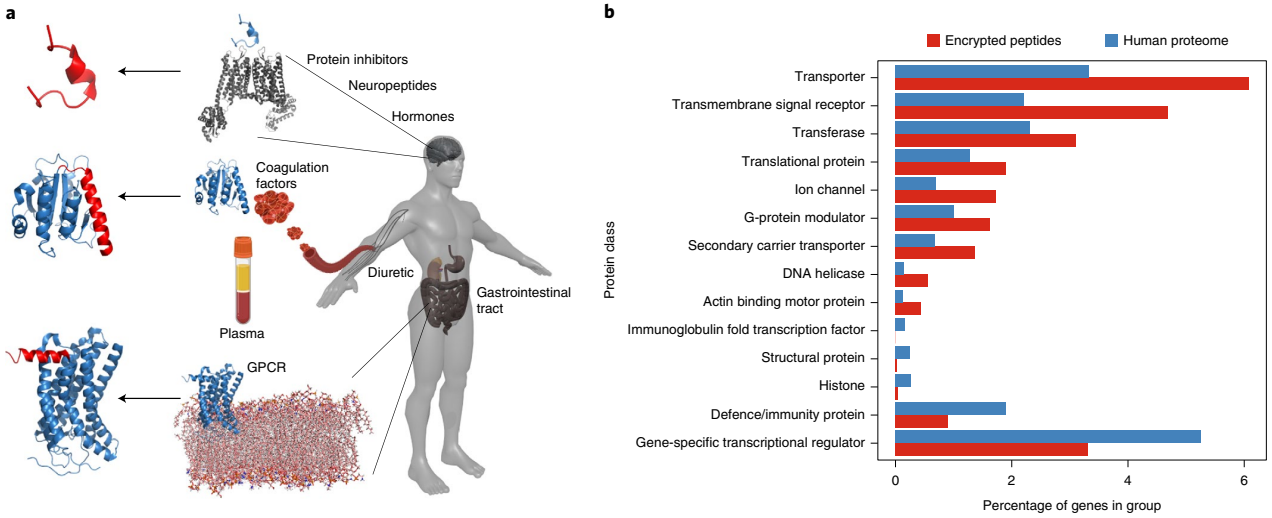


Fig. 1 | Discovery of hidden peptide antibiotics encoded in the human proteome. a, Identification of encrypted peptides within protein sequences from the human proteome using a physicochemically guided scoring function that considered the main physicochemical features of AMPs, that is, length, charge and hydrophobicity. **b**, Normalized abundance of genes encoding different protein classes across the two groups, proteins containing predicted encrypted peptides, and proteins in the entire human genome. The analysis was performed using Panther Proteins Classification system with a false discovery rate cutoff of 0.05.

(Fig. 1b). To validate the predictions made by our algorithm, 55 representative peptides were synthesized and characterized in detail; these peptides exhibited a range of predicted scores given by the fitness function and were derived from the cardiovascular, nervous, renal, haematopoietic and digestive systems.

Importantly, the amino acid patterns among the 55 encrypted peptides were substantially different from those of AMPs present in the broadly used Database of Antimicrobial Activity and Structure of Peptides (DBAASP)[12] (Fig. 2a), even when searches were based on the physicochemical properties displayed by these classical AMPs (Supplementary Fig. 2). Previously identified AMPs contained motifs and repeats of cationic and hydrophobic amino acid residues, creating an amphipathic structure that plays a key role in their mechanism of action. In contrast, the encrypted peptides we identified displayed heterogeneous sequences that diverged from those of the AMPs. While most AMPs described in the literature[13] present sequences with 40–60% hydrophobic residues (for example, leucine, isoleucine, valine, proline, glycine and alanine) and cationic residues (for example, histidine, lysine and arginine), the encrypted peptides described here displayed a substantially higher content of hydrophobic (16.3%) and basic (8.1%) amino acid residues and a lower content of polar (5.8%) and acidic (68.0%) residues than reported for AMPs, while the content of aliphatic residues was similar (Supplementary Table 1). For example, out of the nine most common amino acid residues found in AMPs (alanine, arginine, glycine, isoleucine, leucine, lysine, phenylalanine, proline and valine)[12], four hydrophobic residues (phenylalanine, isoleucine, leucine and valine) and the basic amino acid residue arginine were more frequently represented in encrypted peptides than in AMPs. AMPs, on the other hand, display small aliphatic and cationic residues (that is, alanine, glycine and lysine) at higher frequency than encrypted peptides (Fig. 2a). These data suggest that the encrypted peptides identified and described here may represent a novel class of natural peptide antibiotics, which does not necessarily rely on amphipathic structures but instead consists of arginine-rich, slightly more hydrophobic sequences.

We synthesized a library composed of 55 encrypted peptides and assessed their antimicrobial activity against eight clinically relevant pathogens (*Escherichia coli* ATCC11775, *Pseudomonas aeruginosa* PAO1, *P. aeruginosa* PA14, *Staphylococcus aureus* ATCC12600, *E. coli* AIG221, *E. coli* AIG222, *Klebsiella pneumoniae* ATCC133883 and *Acinetobacter baumannii* ATCC19606; Fig. 2b and Supplementary Fig. 3), all of which play a major role in infectious diseases and are ranked in the World Health Organization's watchlist[2]. We identified lead encrypted peptides that completely sterilized bacterial cultures (~10^6 bacterial cells per ml) of at least one of these eight pathogens (Fig. 2b and Supplementary Fig. 3). The majority (63.6%) of the encrypted peptides synthesized and tested displayed antimicrobial activity against pathogens (Fig. 2b and Supplementary Fig. 3), thus validating our algorithm and indicating their potential role in host defence, even though they are derived from proteins involved in processes unrelated to the immune system. A synthetic peptide whose sequence was randomly generated and not part of the encrypted peptide dataset (Supplementary Dataset 1) served as an additional negative control in the antimicrobial assays (Supplementary Fig. 4).

In addition to their ability to kill pathogenic organisms, several peptides targeted human commensals from the gut and skin microbiota (Fig. 2b,c), an interesting observation given that some previously described natural peptides are inactive against indigenous microbiota members[14]. The 13 most abundant members of the human gut microbiota[15] were exposed to increasing concentrations of the peptides. Species from four different phyla were used to determine susceptibility to the encrypted peptides: *Akkermansia muciniphila* (Verrucomicrobia); *Bacteroides fragilis*, *B. thetaiotaomicron*, *B. vulgatus*, *B. uniformis*, *B. eggerthi*, *Parabacteroidetes distasonis* and *Prevotella copri* (Bacteroidetes); *Collinsella aerofaciens* (Actinobacteria); and *Clostridium scindens* and *Clostridium spiroforme* (Firmicutes). The peptides displayed low micromolar antimicrobial activity against Gram-positive commensals, and five lead peptides were also able to target *A. muciniphila* and Bacteroidetes species.

One of the five lead antimicrobial candidates, CPX1-HVR25, derives from the probable carboxypeptidase X1, a protein expressed in many organs that may be involved in cell–cell interactions, collagen binding[16] and regulation of adipogenesis[17]; this peptide is classified as a metallocarboxypeptidase based on sequence similarity,

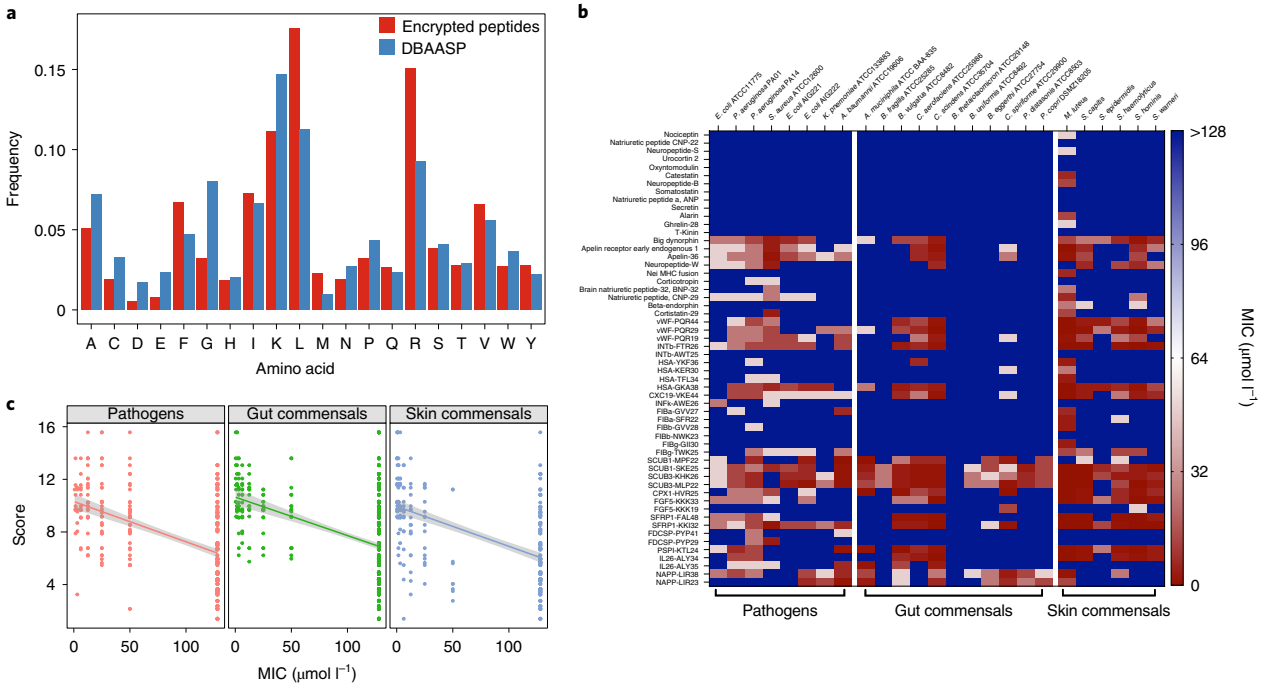


Fig. 2 | Composition and bacteria-targeting properties of encrypted peptides. a, Amino acid frequency in encrypted peptides compared with known AMPs. Amino acid usage was calculated from candidate encrypted peptides from the human proteome and from the DBAASP database. Encrypted peptides had overrepresentation of phenylalanine (F), isoleucine (I), leucine (L) and valine (V) compared with classical AMPs present in the DBAASP database. **b**, Antimicrobial activity of the encrypted peptides. Briefly, 10^6 bacterial cells and serially diluted encrypted peptides (0–128 μmol l^{-1}) were added to a 96-well plate and incubated at 37 °C. One day post-treatment, the solution in each well was measured in a microplate reader (600 nm) to check for inhibition of bacteria compared with the untreated controls and presented as a heat map of antimicrobial activities (μmol l^{-1}) against 8 pathogenic, 11 gut commensal and 6 skin commensal bacterial strains. Assays were performed in three independent replicates, and heat map OD_{600} values are the arithmetic mean of the replicates in each condition. **c**, Predictive power of the physicochemically based scoring function. The figure shows how the experimentally determined MIC of selected peptides correlates with their predicted scores. Multiple microbes were grouped in this test, ranging from pathogenic strains to gut and skin commensals. Higher predicted scores correlate with lower MICs (more potent antimicrobial activity). In **c**, data are the mean ± s.d.

although no carboxypeptidase activity has been reported[18]. PSPI-KTL24, another peptide with activity against *A. muciniphila*, a beneficial member of the gut microbiome that has been shown to contribute to obesity, glucose metabolism and intestinal immunity[19], derives from the human plasma serine protease inhibitor and is present throughout the body in the bloodstream. This peptide is responsible for the inactivation of serine proteases by irreversibly binding to their serine activation site and is involved in the regulation of intravascular and extravascular proteolytic activities. In addition, it plays haemostatic roles in blood plasma, acting as a procoagulant and pro-inflammatory factor by activating and inhibiting the anticoagulant protein C factor[20]. The peptides SCUB1-SKE25, SCUB3-KHK26 and SCUB3-MLP22 also showed potent activity against Gram-negative gut commensals. These are fragments from the terminal portions of two different complement *C1r*/*C1s*, *Uegf*, *Bmp1* (CUB) domains, which are structural motifs found in extracellular and plasma membrane-associated proteins. Proteins containing the CUB domains are involved in a diverse range of functions, including developmental patterning, tissue repair, axon guidance and angiogenesis, cell signalling, fertilization, haemostasis, inflammation, neurotransmission, receptor-mediated endocytosis, tumour suppression[21] and complement activation, the latter of which has been shown to contribute to innate immunity by generating antibacterial peptides[22]. Altogether, our findings suggest that encrypted peptides may be involved in shaping the gut microbiota, with implications for human health as disruptions in gut commensal communities have been associated with numerous diseases, including obesity, diabetes, inflammatory bowel disease, colitis, cancer and neurodegenerative disorders[23,24]. Encrypted peptides also targeted skin commensals, including Gram-positive strains isolated from healthy patients such as *Micrococcus luteus*, *S. capitis*, *S. epidermidis*, *S. hominis*, *S. haemolyticus* and *S. warneri* (Fig. 2b). Perturbations of the skin microbiota have been associated with conditions such as atopic dermatitis, rosacea, psoriasis and acne[25,26]; again, our results revealed novel microbiota-modulating functions of natural peptides. Similar to the gut microbiota experiments, the peptides were active against skin commensals at low micromolar doses (Fig. 2b). To assess the predictive power of our physicochemically based scoring function, we correlated experimentally determined minimal inhibitory concentrations (MICs) of encrypted peptides with their predicted scores. Generally, the antimicrobial activity of the peptides correlated with their predicted scores against pathogens, and gut and skin commensals (Fig. 2c), and at least 80% of the 55 tested encrypted peptides targeted either pathogens, gut commensals or skin commensals.

Next, we wondered whether peptides from the same biogeographical area could synergize to target bacteria at levels that are physiologically relevant (that is, in most cases from picomolar to millimolar concentrations) (Supplementary Table 2 and Extended Data Fig. 1). Remarkably, one pair of encrypted peptides synergized to kill pathogens at low micromolar to nanomolar concentrations both in vitro and in animal models (Figs. 3a and 4b and Extended

Data Fig. 2), displaying activity comparable to, and with even higher potency in some cases than, the most potent venom-derived peptides[27,28] and defensins from the human immune system[29]. To investigate synergy, we looked for the activity between combinations of 12 peptides: apelin receptor early endogenous 1, apelin-36, big dynorphin, natriuretic peptide and encrypted peptides from the coagulation factors, blood glycoproteins, serine proteases, cytokines and the CUB domains. These peptides were assessed for their ability to synergize with each other to inhibit the growth of the Gram-negative pathogenic bacterium *P. aeruginosa* PAO1. The 12 peptides tested in synergy assays were selected on the basis of their potency against *P. aeruginosa*, an intrinsically resistant bacterium[30] that infects the urinary tract, gastrointestinal tissue, and skin and soft tissues; *P. aeruginosa* also causes pneumonia and is one of the most common opportunistic pathogens in patients with cystic fibrosis[31]. To quantify synergistic interactions between peptides, we determined their fractional inhibitory concentration (FIC) index as previously described[32,33]. For most of the peptide combinations tested (82%), the peptides interacted additively, with FIC values ranging from 0.6 to 1 (Extended Data Fig. 2). Highly synergistic interactions (FIC ≤0.5) (Extended Data Fig. 2) were observed between the peptide hormones big dynorphin and apelin receptor early endogenous 1. Big dynorphin is the most potent endogenous opioid peptide, having affinity for κ-opioid receptors produced in the brain, and is involved in pain response, control of appetite, circadian rhythms and temperature regulation[34]. Apelin receptor early endogenous 1, on the other hand, is responsible for mesendodermal differentiation, blood vessel formation and heart morphogenesis[35]. To test synergistic interactions involving more than two encrypted peptides derived from the same protein, we designed and performed three-way synergy experiments. We selected three representatives from the CUB domains 1 and 3 (SCUB1-SKE25, SCUB1-MPF22 and SCUB3-MLP22), which were active against the Gram-negative pathogen *A. baumannii* at extremely low concentrations (1.56–3.12 μmol l^{-1}; 4.9–8.2 μg ml^{-1}) comparable to the inhibitory concentrations of standard-of-care antibiotics (for example, polymyxin B)[36]. The combination of the three peptides increased their individual antimicrobial activity by 100-fold, yielding MIC values of 40–90 nmol l^{-1} (125–280 ng ml^{-1}) and a combined FIC index of 0.2 (Fig. 3a). Overall, these results are notable because a number of parent proteins containing several of the encrypted peptides reported here are produced physiologically at levels at which the peptides present antimicrobial activity (Supplementary Table 2 and Extended Data Fig. 1)[37]. The ability of these agents to potentiate each other's antibiotic properties further underscores their potential involvement in innate defence.

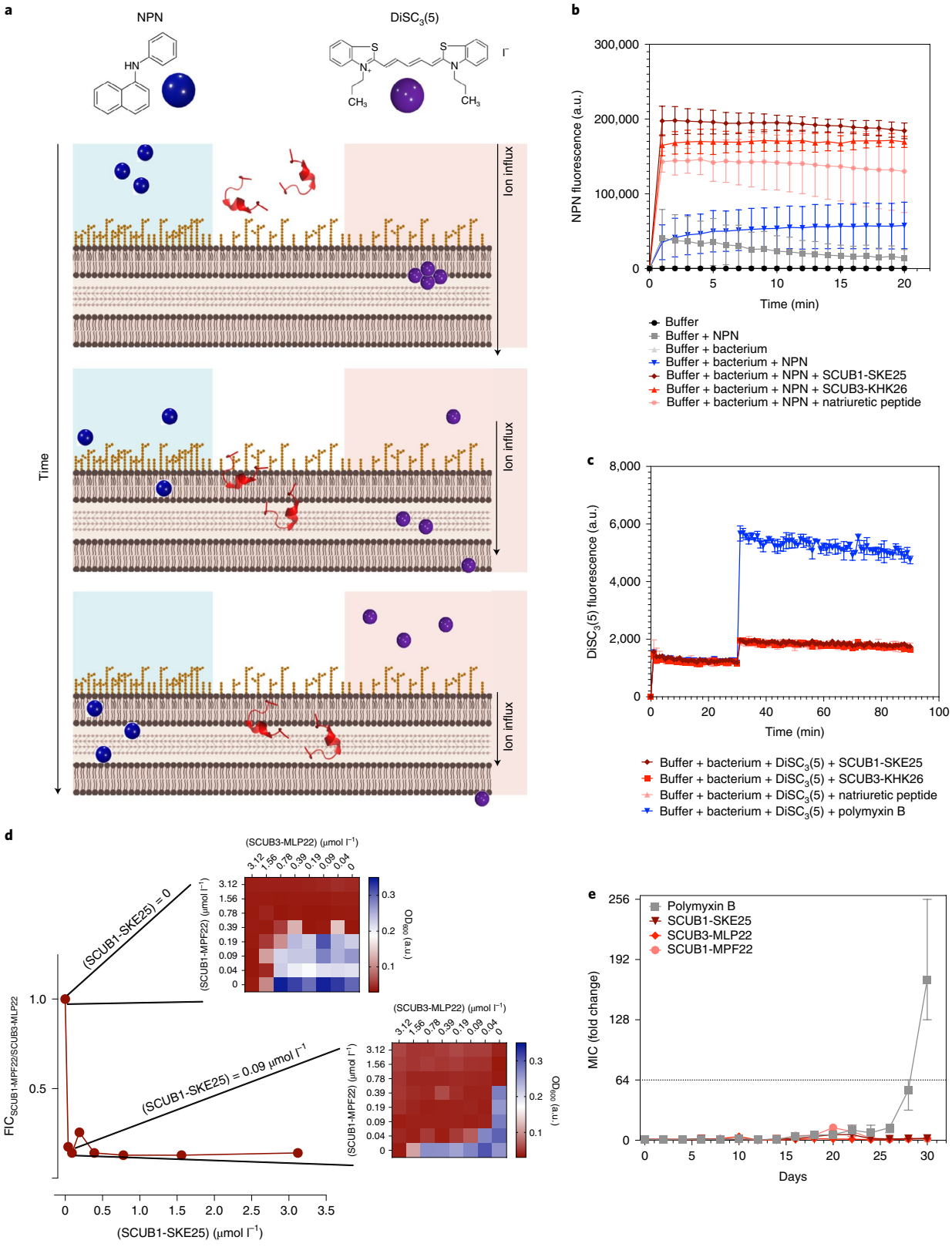
In any antibiotic discovery effort, it is crucial to assess how readily the molecules identified select for resistance mechanisms in bacteria, as the emergence of resistance has hampered the antibiotic field since its very inception with the discovery of penicillin in 1928 (ref. [38]). To determine whether bacteria become resistant to encrypted peptides, we performed longitudinal resistance assays with peptides derived from the CUB domains 1 and 3 against *A. baumannii*, using polymyxin B as a control (Fig. 3b). Bacterial cells became highly resistant to polymyxin B after 26 days, when concentrations needed to kill *A. baumannii* increased by as much as 256-fold (Fig. 3b). Conversely, treatment with the encrypted peptides did not lead to the evolution of spontaneous peptide-resistant *A. baumannii* mutant cells over the same period (Fig. 3b). Our results identify encrypted peptides as agents that do not readily select for bacterial resistance and point towards a different mechanism of action than that of the AMPs, such as polymyxin B, which have already been characterized.

To investigate the mode of action of encrypted peptides against bacterial cells, we performed assays to study their effects on the bacterial membrane. First, we assessed the ability of encrypted peptides to disrupt and depolarize the bacterial cytoplasmic membrane (Fig. 3c). We used 3,3′-dipropylthiadicarbocyanine iodide [$DiSC_3(5)$], a potentiometric fluorophore that accumulates and aggregates in the cytoplasmic membrane, quenching its fluorescence. Upon imbalances of the cytoplasmic membrane transmembrane potential, the fluorophore migrates to the outer environment, generating fluorescence. None of the peptides tested depolarized the cytoplasmic membrane of either pathogenic or commensal bacterial cells to the extent of the positive control group treated with polymyxin B, a well-known depolarizing peptide antibiotic that permeabilizes and disrupts bacterial membranes (Fig. 3d and Extended Data Fig. 3). These data reveal that encrypted peptides operate via a mechanism that is independent of cytoplasmic membrane depolarization and thus distinct from that of other AMPs characterized so far. To assess whether encrypted peptides permeabilized the outer membrane, we performed 1-(*N*-phenylamino)naphthalene (NPN) assays. NPN, a lipophilic dye that fluoresces weakly in aqueous environments but whose fluorescence is greatly increased upon contact with lipidic environments such as bacterial membranes, was added to bacterial solutions (Fig. 3e). NPN does not permeate the bacterial outer membrane unless the membrane is damaged, or its integrity has been compromised. Bacteria exposed to the most active encrypted peptides (that is, natriuretic peptide, which is responsible for inducing the excretion of sodium by the kidney, SCUB1-SKE25 and SCUB3-MLP22) at their MIC values (Fig. 3e) emitted increased fluorescence compared with either the untreated control group or cells treated with the positive control AMP polymyxin B. Thus, encrypted peptides exhibited increased ability to permeabilize the outer membrane compared with conventional AMPs. Overall, these data suggest that the encrypted peptides do not affect the cytoplasmic membrane of bacteria but instead exert their inhibitory effects by permeabilizing the outer membrane.

To test whether the peptides retain their antimicrobial activity in a complex living system, we probed their properties in two mouse models. The most active encrypted peptides, SCUB1-SKE25 and SCUB3-MLP22, were selected. Since these peptides are mostly found in plasma, we used a relevant skin abscess mouse model to test their anti-infective activity (Fig. 4a)[28]. Mice were infected with bacterial loads of 10^5 and 10^6 cells in 20 μl of the clinically relevant Gram-negative pathogens *P. aeruginosa* and *A. baumannii*, respectively (Fig. 4b). A single dose of each encrypted peptide (25 μmol l^{-1}, 77.9 μg ml^{-1} and 66.9 μg ml^{-1}) delivered to the infected area markedly reduced the bacterial load by three orders of magnitude, highlighting the anti-infective potential of these agents. To validate the synergistic effects obtained in vitro (Fig. 3d and Supplementary Fig. 6), we treated infected mice with combinations of encrypted

Fig. 3 | Synergy, resistance development and mechanism of action studies of encrypted peptides. a, Schematic showing increased fluorescence resulting from membrane destabilization (left panel, blue) and depolarization (right panel, red) caused by the peptides at their MIC over time. **b**, NPN assays showing the effect of encrypted peptides derived from the CUB domain and natriuretic peptide on permeabilization of the outer membrane of *A. baumannii*. **c**, Cytoplasmic membrane depolarization effects of the encrypted peptides against *A. baumannii*. **d**, Synergistic interactions of encrypted peptides from the CUB domain against *A. baumannii*, resulting in 100-fold lower concentrations of all the three peptides needed to completely inhibit bacterial growth. **e**, Evolution of resistance by *A. baumannii* to encrypted peptides derived from the CUB domain (red) or polymyxin B (grey) after 30 days of serial passaging in liquid NB. Peptides and antibiotics were used at subinhibitory concentrations. Cells were passaged every 48 h. The assays in **b**–**e** were performed in three independent replicates. In **b**–**e**, data are the mean ± s.d. Assays were performed in three independent replicates.

peptides (Fig. 4b). Treatment with a combination of SCUB1-SKE25 and SCUB3-MLP22 significantly decreased bacterial counts by five and six orders of magnitude for *A. baumannii* and *P. aeruginosa*, respectively (Fig. 4b). No damage or deleterious effects exerted by the peptides were detected in the mice in any of our experiments. Of note, the anti-infective activity displayed by the encrypted peptides



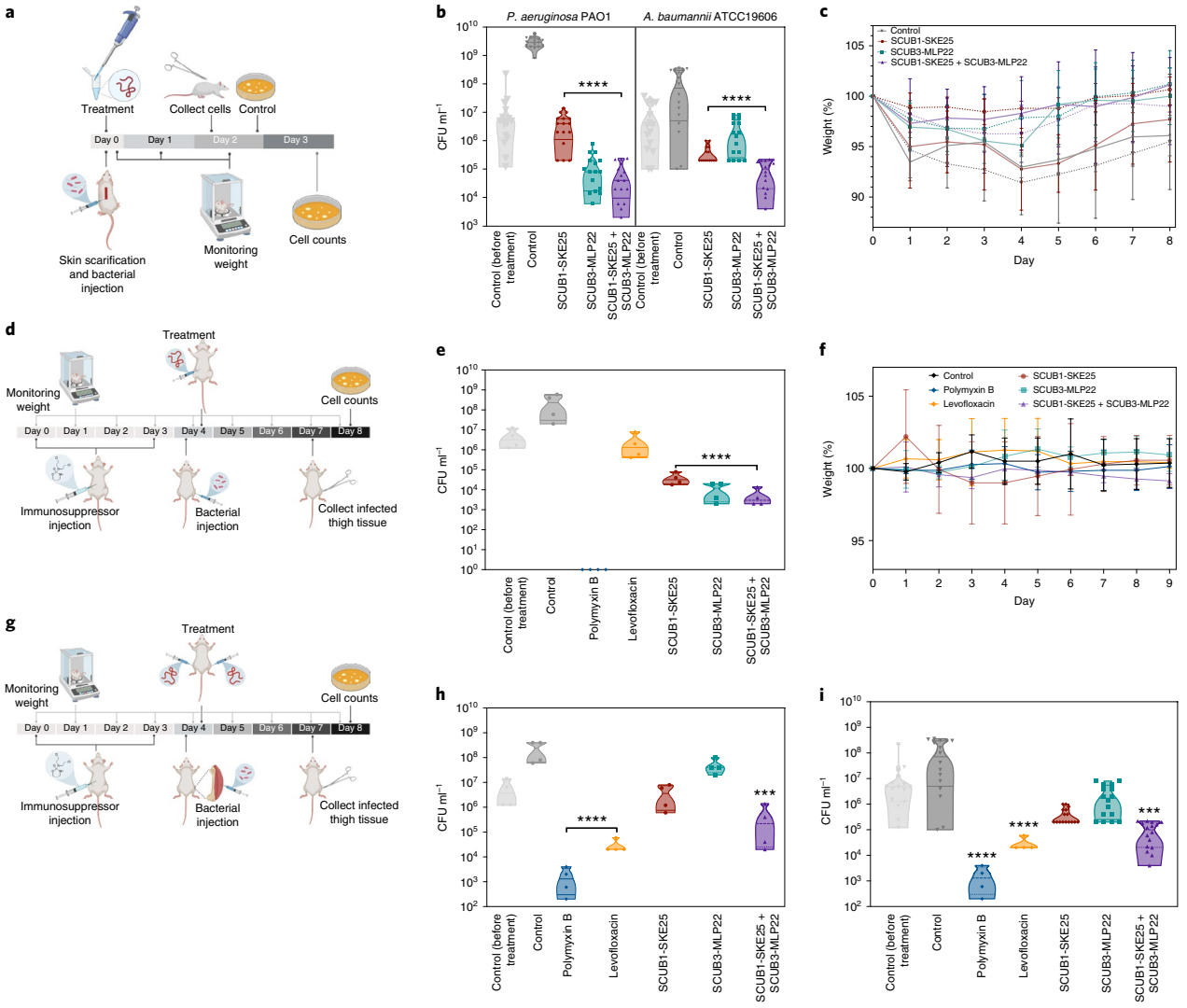


Fig. 4 | Anti-infective activity and synergistic interactions of encrypted peptides in vivo. a, Schematic of the skin abscess mouse model used to assess the anti-infective activity of selected encrypted peptides from plasma proteins ($n = 8$). **b**, SCUB1-SKE25 (25 μmol l^{-1}; 77.9 μg ml^{-1}) and SCUB3-MLP22 (25 μmol l^{-1}; 66.9 μg ml^{-1}) showed inhibitory activity, especially when used in combination (at 3.12 and 6.25 μmol l^{-1}; 9.74 and 16.73 μg ml^{-1}; for SCUB1-SKE25 and SCUB3-MLP22, respectively), against both *A. baumannii* ATCC19606 and *P. aeruginosa* PAO1. **c**, Mouse weight was monitored throughout the experiment (2 days) to rule out potential toxic effects of the encrypted peptides (solid lines, *P. aeruginosa* PAO1; dashed lines, *A. baumannii* ATCC19606). **d**, Schematic of the neutropenic thigh infection mouse model in which encrypted peptides from plasma proteins were injected intraperitoneally to assess their anti-infective activity when administered systemically ($n = 8$). **e**, Treatment with SCUB1-SKE25 and SCUB3-MLP22 either alone or in combination, at the same concentrations as those used in **b**, reduced infections caused by *A. baumannii* ATCC19606. **f**, Mouse weight was monitored throughout the duration of the neutropenic thigh infection model (8 days total) to rule out potential toxic effects of the encrypted peptides. **g**, Schematic of the neutropenic thigh infection mouse model injected with the encrypted peptides both in the opposite (left leg) and in the same infected area where the bacterial load was injected (right leg). **h**, Combination treatment with SCUB1-SKE25 and SCUB3-MLP22, at their MIC obtained from in vitro synergy experiments (that is, 3.12 and 6.25 μmol l^{-1}, respectively), led to decreased *A. baumannii* ATCC19606 bacterial counts. Both peptides were administered intramuscularly in the left thigh at a distance from the infection site (that is, the right thigh). **i**, Intramuscular peptide administration at the infection site (both in mono- and combination therapies) also led to reduced bacterial counts compared with the untreated control group. Polymyxin B and levofloxacin were used as positive controls in the neutropenic thigh infection mouse model. Statistical significance in **b**,**e**,**h**,**i** was determined using one-way ANOVA, $^{***}P < 0.001$, $^{****}P < 0.0001$; features on the violin plots represent median and upper and lower quartiles. Data in **c**,**f** are the mean ± s.d.

was comparable to that of some of the most potent AMPs described to date, such as polybia-CP and PaDBS1R6[28,39–42]. Next, we assessed the efficacy of the same encrypted peptides (SCUB1-SKE25 and SCUB3-MLP22) in monotherapy and in combination therapy in a murine thigh infection model (Fig. 4d,g)[43]. The conventional antibiotics polymyxin B and levofloxacin were used as controls (Supplementary Fig. 5). This preclinical model is sensitive, standardized and widely used for assessing antibiotic translatability[43,44]. Briefly, mice were rendered neutropenic by cyclophosphamide treatment before intramuscular infection with 10^6 cells per ml of

the bacterial pathogens *P. aeruginosa* (Extended Data Fig. 4) and *A. baumannii* (Fig. 4e,f,h,i). Either a single dose of each encrypted peptide (at 25 μmol l^{-1}) or their combination (also at 25 μmol l^{-1}) was delivered to the infected site via intramuscular administration at the infection site (Fig. 4l), in the opposite thigh (Fig. 4h) and intraperitoneally (Fig. 4e). Three days post-treatment, the peptides reduced the bacterial load by two to three orders of magnitude when administered intraperitoneally, presenting higher antimicrobial activity than levofloxacin (also administered at its MIC) but less so than polymyxin B, which cleared the infection (Fig. 4e). The experiments were performed also using a higher concentration of the peptides (fourfold MIC, 100 μmol l^{-1} alone and 12.5 and 25 μmol l^{-1} for SCUB1-SKE25 and SCUB3-MLP22, respectively, in combination), revealing similar results (Extended Data Fig. 5) as those obtained when using the peptides at their MIC (Fig. 4i). When administered intramuscularly at the opposite thigh (left thigh) away from the infection (right thigh), the peptides alone and in combination displayed a bacteriostatic effect (Fig. 4h). Administration of the peptides at the infection site (right thigh) decreased the bacterial load of *P. aeruginosa* and *A. baumannii* by four and two orders of magnitude, respectively, compared with untreated mice. The combination treatment reduced by five and three orders of magnitude infections caused by *P. aeruginosa* (Extended Data Fig. 4) and *A. baumannii*, respectively (Fig. 4i). Importantly, no significant changes in weight, a proxy for toxicity, were observed in our experiments (Fig. 4c,f), thus further demonstrating the lack of toxicity of the encrypted peptides in relevant mouse models. Our in vivo results demonstrate the antibiotic and synergistic properties of encrypted peptides under physiological conditions and strongly suggest that the peptides constitute excellent candidates for the future development of antibacterial agents.

Although several approaches have been previously taken[45–47] to identify novel molecules in the human microbiota and soil bacteria[48], to date, a systematic study exploring the human body as a potential source of drugs has not been performed. Here, we report the discovery of 2,603 previously undescribed antibiotics encrypted within the human proteome. At least 55 of these agents, tested here, were found to target both pathogens and commensals, and the lead peptides operated by permeabilizing the bacterial outer membrane displayed synergistic interactions that potentiated their activity and effectively reduced infections in relevant rodent models. We show that encrypted peptides differ from conventional AMPs because they (1) possess a different amino acid frequency profile (Fig. 2a) and (2) target commensals (Fig. 2b,c). Our results point to the proteome as a previously untapped source of novel antibiotics and reveal the multifunctional nature of numerous proteins that were traditionally thought to have only a single biological function. Because our approach computationally identifies antibiotics already optimized by nature and produced in our own bodies, we expect they will serve as excellent candidates for antibiotic development. The data presented also suggest connections between host defence and other systems in the body that were previously unrecognized as arms of host immunity, such as the nervous, endocrine, digestive and cardiovascular systems. More broadly, we speculate that the existence of proteins with multiple functions, enabled by encrypted fragments, reflects an evolutionary mechanism that expanded protein capabilities while minimizing genomic expansion. This reduced the number of protein-coding genes that perform all the functionalities necessary to operate and defend the human body.

Methods

Human proteome screening for encrypted peptides. To scan the human proteome, all canonical and isoform sequences of *Homo sapiens* proteins were downloaded from UniProt[49]. A Python script was used to scan all proteins using multiple moving windows with lengths ranging from 8 to 50 residues, leading to millions of peptide sequences. Each sequence was scored using the method described by Pane et al.[10], which uses peptide length, charged residues and hydrophobic residues to create a score for the propensity of the peptide to present antimicrobial activity. After the scan, identical peptide sequences were filtered out (even if present in multiple proteins or isoforms) so that only unique sequences were present in the final dataset. Finally, the top 1,000 scoring peptides per sequence length were combined to form a 43,000-long dataset of candidates from across the human genome. An analogous search was carried out to generate a final list of top candidates, but this time focusing solely on secreted proteins, as those were predicted to provide more biologically relevant encrypted peptides.

Peptide synthesis. The peptides were synthesized using solid-phase peptide synthesis and *N*-9-fluoromethyloxycarbonyl strategy. They were purified by high-performance liquid chromatography. The peptide purity used in all assays was higher than 95%.

Bacterial strains and media. The strains used in this study were the pathogens *E. coli* ATCC11775, *A. baumannii* ATCC19606, *P. aeruginosa* PAO1, *P. aeruginosa* PA14, *S. aureus* ATCC12600, *E. coli* AIG221, *E. coli* AIG222 (colistin-resistant strain) and *Klebsiella pneumoniae* ATCC13883, and the gut commensals *A. muciniphila* ATCCBAA-635, *B. fragilis* ATCC25285, *B. thetaiotaomicron* ATCC29148, *B. eggerthi* ATCC27754, *B. uniformis* ATCC8492, *B vulgatus* ATCC8482, *Parabacteroidetes distasonis* ATCC8503, *Prevotella copri* DSMZ18205, *Colinisella aerofaciens* ATCC25986 and *Clostriudium scindens* ATCC35704. In addition, we used the following skin commensals in this study: *Micrococcus luteus*, *S. capitis*, *S. epidermidis*, *S. haemolyticus*, *S. hominis* and *S. warneri*. Pathogenic bacteria were grown and plated on Luria–Bertani (LB) and Pseudomonas Isolation (*P. aeruginosa* strains) agar plates incubated overnight at 37 °C, gut microbiome commensals were grown and plated on brain–heart infusion (BHI) agar plates supplemented with vitamin K3, haemin and L-cysteine from frozen stocks incubated overnight at 37 °C, and skin microbiome commensals were grown in nutrient broth (NB) and plated in tryptic soy broth (TSB) from 4 °C stocks which were incubated overnight at 37 °C. Following the incubation period, one isolated colony was transferred to 5 ml of medium (LB, NB or supplemented BHI broth), which was incubated overnight (12–16 h) at 37 °C. On the following day, inoculums were prepared by diluting the bacterial overnight solutions 1:100 in 5 ml of the respective media and incubated at 37 °C until logarithmic phase (OD_{600} 0.3–0.5).

Antibacterial assays. MICs of peptides were determined using the broth microdilution technique in LB, NB and supplemented BHI with an initial inoculum of 5×10^6 cells in nontreated polystyrene microtiter plates (Corning). Peptides were added to the plate as solutions in LB, NB and supplemented BHI broth in concentrations ranging from 0 to 128 μmol l^{-1}. The MIC was considered as the lowest concentration of peptide that completely inhibited the visible growth (readings were made in a spectrophotometer at 600 nm) of bacteria after 24 h of incubation of the plates at 37 °C. All assays were done in three independent replicates.

Synergy assays. *P. aeruginosa* PAO1 and *A. baumannii* ATCC19606 were chosen for the synergy assays because of their relevance as pathogens that are intrinsically resistant to antimicrobials[50] and their ability to infect the urinary tract, gastrointestinal tissue, and skin and soft tissues. After determination of the MIC for each peptide, the most active encrypted peptides against *P. aeruginosa* and *A. baumannii* were orthogonally diluted using the microdilution technique with concentrations ranging from 2-fold MIC to 0.03-fold MIC. Plates were incubated for 24 h at 37 °C. All assays were done in three independent replicates.

Bacterial resistance development assays. The evolution of resistance by *A. baumannii* cells to encrypted peptides was monitored for 30 days of serial passaging in liquid NB. The experiment was designed to avoid population extinction. Thus, peptides and polymyxin B, which was used as control, were plated in chequerboard pattern to avoid cross-contamination at concentrations ranging from 8- to 0.25-fold MIC. Bacterial loads of 10^6 cells were added in each passage, and the subinhibitory concentrations (0.5-fold MIC) were considered as the first concentration of peptide that inhibited cell growth by ~50%. Bacteria remaining upon treatment with subinhibitory concentrations were re-grown overnight, and after each incubation period, a new inoculum of 10^6 cells was prepared for inoculation of the following passage containing fresh medium and increased doses of the antimicrobial agent. To ensure reproducibility, three parallel experiments with two technical replicates per compound were performed. Cells were passaged every 48 h.

Membrane depolarization assays. The cytoplasmic membrane depolarization activity of the peptides was determined by measurements of fluorescence of the membrane-potential-sensitive dye $DiSC_3(5)$. Briefly, *A. baumannii* ATCC19606, *P. aeruginosa* PAO1 and *B. fragilis* ATCC25285 were grown at 37 °C with agitation until they reached mid-log phase (OD_{600} 0.5). The cells were then centrifuged and washed twice with washing buffer (20 mmol l^{-1} glucose, 5 mmol l^{-1} HEPES, pH 7.2) and re-suspended to an OD_{600} of 0.05 in the same buffer (20 mmol l^{-1} glucose, 5 mmol l^{-1} HEPES, pH 7.2) but containing 0.1 mol l^{-1} KCl. Thereafter, the cells (100 μl) were incubated for 15 min with 20 nmol l^{-1} of $DiSC_3(5)$ until

a stable reduction of fluorescence was achieved, indicating the incorporation of the dye into the bacterial membrane. Membrane depolarization was then monitored by observing the change in the fluorescence emission intensity of the membrane-potential-sensitive dye $DiSC_3(5)$ (λ_{ex} = 622 nm, λ_{em} = 670 nm), after the addition of peptides (100 µl solution at MIC values).

Membrane permeabilization assay. The membrane permeability of the peptides was determined by using the NPN uptake assay. *A. baumannii* ATCC19606, *P. aeruginosa* PAO1, *S. epidermidis* and *B. fragilis* ATCC25285 were grown to an OD_{600} of 0.4, centrifuged (10,000 r.p.m. at 4 °C for 10 min), washed and re-suspended in buffer (5 mmol l^{-1} HEPES, 5 mmol l^{-1} glucose, pH 7.4). Then, 4 µl of NPN solution (0.5 mmol l^{-1}; working concentration of 10 µmol l^{-1} after dilutions) was added to 100 µl of the bacterial solution in a white 96-well plate. The background fluorescence was recorded at λ_{ex} = 350 nm and λ_{em} = 420 nm. Peptides' solutions in water (100 µl solution at their MIC values) were added to the 96-well plate, and fluorescence was recorded as a function of time until no further increase in fluorescence was observed (20 min).

Skin abscess infection mouse model. *A. baumannii* ATCC19606 and *P. aeruginosa* strain PAO1 were used to infect the murine skin. Briefly, bacteria were grown in TSB medium. Subsequently, cells were washed twice with sterile PBS (pH 7.4, 13,000 r.p.m. for 1 min), and re-suspended to a final concentration of 1×10^5 and 1×10^6 CFU per 20 µl for *A. baumannii* and *P. aeruginosa*, respectively. Female CD-1 mice (6 weeks old) were anaesthetized with isoflurane and had their backs shaved, and a superficial linear skin abrasion was made with a needle to damage the stratum corneum and upper layer of the epidermis. An aliquot of 20 µl containing the bacterial load re-suspended in PBS was inoculated using a pipette tip over each defined scratched area. One hour after the infection, peptides also in PBS (pH 7.4) at their MIC value were administered to the infected area. Animals were euthanized, and the area of scarified skin was excised 2 and 4 days post-infection, homogenized using a bead beater for 20 min (25 Hz), and tenfold serially diluted for CFU quantification. Two independent experiments were performed with eight mice per group in each condition.

Neutropenic thigh infection mouse model. Two doses of cyclophosphamide (150 mg kg^{-1}) applied intraperitoneally with an interval of 3 days were used to render the mice neutropenic. One day later, the mice were infected intramuscularly in their right thigh with a bacterial load of 10^6 CFU ml^{-1} of the pathogens *A. baumannii* ATCC19606 and *P. aeruginosa* PAO1, which were previously grown in TSB, washed twice with PBS (pH 7.4) and re-suspended to the desired concentration. Two hours later, peptides also re-suspended in PBS (pH 7.4) were administered at the infection site (right thigh), in the opposite thigh (left thigh) and intraperitoneally. Before each injection, mice were anaesthetized with isoflurane and monitored (respiratory rate and pedal reflexes). Next, we monitored the established infection and euthanized the mice. The infected area was excised 3 days post-infection, homogenized using a bead beater for 20 min (25 Hz), and tenfold serially diluted for CFU quantification. The experiments were performed using 8 (antibiotics and untreated groups before treatment) and 16 mice per group (untreated controls, mice treated with a single peptide and mice treated with combination therapy of two peptides).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The main data supporting the results in this study are available within the paper and its Supplementary Information. The list of the identified encrypted peptides, and source data for the normalized abundance of genes encoding different protein classes, for amino acid frequency in encrypted peptides compared with known AMPs, for expression levels of proteins containing encrypted peptides, and for antimicrobial activity (in vitro and in vivo), as well as data on synergistic interactions, evolution of resistance and mechanism of action, are provided with this paper. Source data are provided with this paper.

Code availability

The custom Python code for scanning the human proteome to detect candidate encrypted peptides is available as Supplementary Information.

Received: 7 October 2020; Accepted: 25 August 2021;
Published online: 4 November 2021

References

1. *2019 Antimicrobial Resistant Threats Report* (Centers for Disease Control and Prevention, 2019).
2. World Health Organization. Antimicrobial resistance: Global Health Report on Surveillance. *Bull. World Health Organ*. https://doi.org/10.1007/s13312-014-0374-3 (2014).
3. Lepore, C., Silver, L., Theuretzbacher, U., Thomas, J. & Visi, D. The small-molecule antibiotics pipeline: 2014–2018. *Nat. Rev. Drug Discov*. https://doi.org/10.1038/d41573-019-00130-8 (2019).
4. de la Fuente-Nunez, C., Torres, M. D., Mojica, F. J. & Lu, T. K. Next-generation precision antimicrobials: towards personalized treatment of infectious diseases. *Curr. Opin. Microbiol.* **37**, 95–102 (2017).
5. Torres, P. H. M., Sodero, A. C. R., Jofily, P. & Silva-Jr, F. P. Key topics in molecular docking for drug design. *Int. J. Mol. Sci.* **20**, 4574 (2019).
6. Torres, M. D. T., Cao, J., Franco, O. L., Lu, T. K. & de la Fuente-Nunez, C. Synthetic biology and computer-based frameworks for antimicrobial peptide discovery. *ACS Nano* **15**, 2143–2164 (2021).
7. Torres, M. D. T. & de la Fuente-Nunez, C. Toward computer-made artificial antibiotics. *Curr. Opin. Microbiol.* **51**, 30–38 (2019).
8. Porto, W. F. et al. In silico optimization of a guava antimicrobial peptide enables combinatorial exploration for peptide design. *Nat. Commun.* **9**, 1490 (2018).
9. Zasloff, M. Antimicrobial peptides of multicellular organisms. *Nature* **415**, 389–395 (2002).
10. Pane, K. et al. Antimicrobial potency of cationic antimicrobial peptides can be predicted from their amino acid composition: application to the detection of "cryptic" antimicrobial peptides. *J. Theor. Biol.* **419**, 254–265 (2017).
11. Mi, H., Muruganujan, A., Ebert, D., Huang, X. & Thomas, P. D. PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* **47**, D419–D426 (2019).
12. Pirtskhalava, M. et al. DBAASP v.2: an enhanced database of structure and antimicrobial/cytotoxic activity of natural and synthetic peptides. *Nucleic Acids Res.* **44**, D1104–D1112 (2016).
13. Kang, X. et al. DRAMP 2.0, an updated data repository of antimicrobial peptides. *Sci. Data* **6**, 148 (2019).
14. Cullen, T. W. et al. Antimicrobial peptide resistance mediates resilience of prominent gut commensals during inflammation. *Science* **347**, 170–175 (2015).
15. Almeida, A. et al. A new genomic blueprint of the human gut microbiota. *Nature* **568**, 499–504 (2019).
16. Kim, Y.-H., O'Neill, H. M. & Whitehead, J. P. Carboxypeptidase X-1 (CPX-1) is a secreted collagen-binding glycoprotein. *Biochem. Biophys. Res. Commun.* **468**, 894–899 (2015).
17. Kim, Y. et al. Identification of carboxypeptidase X (CPX)-1 as a positive regulator of adipogenesis. *FASEB J.* **30**, 2528–2540 (2016).
18. Lei, Y., Xin, X., Morgan, D., Pintar, J. E. & Fricker, L. D. Identification of mouse CPX-1, a novel member of the metallocarboxypeptidase gene family with highest similarity to CPX-2. *DNA Cell Biol.* **18**, 175–185 (1999).
19. Cani, P. D. & de Vos, W. M. Next-generation beneficial microbes: the case of *Akkermansia muciniphila*. *Front. Microbiol.* **8**, 1765 (2017).
20. Brenner, A. V. et al. Common single nucleotide polymorphisms in genes related to immune function and risk of papillary thyroid cancer. *PLoS ONE* **8**, e57243 (2013).
21. Bork, P. & Beckmann, G. The CUB domain. *J. Mol. Biol.* **231**, 539–545 (1993).
22. Nordahl, E. A. et al. Activation of the complement system generates antibacterial peptides. *Proc. Natl Acad. Sci. USA* **101**, 16879–16884 (2004).
23. Durack, J. & Lynch, S. V. The gut microbiome: relationships with disease and opportunities for therapy. *J. Exp. Med.* **216**, 20–40 (2019).
24. Shreiner, A. B., Kao, J. Y. & Young, V. B. The gut microbiome in health and in disease. *Curr. Opin. Gastroenterol.* **31**, 69–75 (2015).
25. Grice, E. The skin microbiome: potential for novel diagnostic and therapeutic approaches to cutaneous disease. *Semin. Cutan. Med. Surg.* **33**, 98–103 (2014).
26. Sanford, J. A. & Gallo, R. L. Functions of the skin microbiota in health and disease. *Semin. Immunol.* **25**, 370–377 (2013).
27. Robinson, S. D. et al. A comprehensive portrait of the venom of the giant red bull ant, *Myrmecia gulosa*, reveals a hyperdiverse hymenopteran toxin gene family. *Sci. Adv.* **4**, eaau4640 (2018).
28. Torres, M. D. T. et al. Structure-function-guided exploration of the antimicrobial peptide polybia-CP identifies activity determinants and generates synthetic therapeutic candidates. *Commun. Biol.* **1**, 221 (2018).
29. Magana, M. et al. The value of antimicrobial peptides in the age of resistance. *Lancet Infect. Dis*. https://doi.org/10.1016/S1473-3099(20)30327-3 (2020).
30. Breidenstein, E. B. M., de la Fuente-Núñez, C. & Hancock, R. E. W. *Pseudomonas aeruginosa*: all roads lead to resistance. *Trends Microbiol.* **19**, 419–426 (2011).
31. Pachori, P., Gothalwal, R. & Gandhi, P. Emergence of antibiotic resistance *Pseudomonas aeruginosa* in intensive care unit; a critical review. *Genes Dis.* **6**, 109–119 (2019).
32. Tyers, M. & Wright, G. D. Drug combinations: a strategy to extend the life of antibiotics in the 21st century. *Nat. Rev. Microbiol.* **17**, 141–155 (2019).
33. Reffuveille, F., de la Fuente-Núñez, C., Mansour, S. & Hancock, R. E. W. A broad-spectrum antibiofilm peptide enhances antibiotic action against bacterial biofilms. *Antimicrob. Agents Chemother.* **58**, 5363–5371 (2014).

34. Merg, F. et al. Big dynorphin as a putative endogenous ligand for the kappa-opioid receptor. *J. Neurochem.* **97**, 292–301 (2006).
35. Wang, Z. et al. Elabela-apelin receptor signaling pathway is functional in mammalian systems. *Sci. Rep.* **5**, 8170 (2015).
36. Cheah, S.-E. et al. Polymyxin resistance in *Acinetobacter baumannii*: genetic mutations and transcriptomic changes in response to clinically relevant dosage regimens. *Sci. Rep.* **6**, 26233 (2016).
37. Samaras, P. et al. ProteomicsDB: a multi-omics and multi-organism resource for life science research. *Nucleic Acids Res.* **48**, D1153–D1163 (2019).
38. Fleming, A. Penicillin. Nobel lecture. *Nobel Foundation* https://www.nobelprize.org/uploads/2018/06/fleming-lecture.pdf (1945).
39. Fensterseifer, I. C. M. et al. Selective antibacterial activity of the cationic peptide PaDBS1R6 against Gram-negative bacteria. *Biochim. Biophys. Acta Biomembr.* **1861**, 1375–1387 (2019).
40. Cardoso, M. H. et al. A computationally designed peptide derived from *Escherichia coli* as a potential drug template for antibacterial and antibiofilm therapies. *ACS Infect. Dis.* **4**, 1727–1736 (2018).
41. Cândido, E. S. et al. Short cationic peptide derived from archaea with dual antibacterial properties and anti-infective potential. *ACS Infect. Dis.* **5**, 1081–1086 (2019).
42. Pane, K. et al. Identification of novel cryptic multifunctional antimicrobial peptides from the human stomach enabled by a computational–experimental platform. *ACS Synth. Biol.* **7**, 2105–2115 (2018).
43. Ling, L. L. et al. A new antibiotic kills pathogens without detectable resistance. *Nature* **517**, 455–459 (2015).
44. Gudmundsson, S. & Erlendsdóttir, H. in *Handbook of Animal Models of Infection* (eds Zak, O. & Sande M. A.) 137–144 (Elsevier, 1999).
45. Sberro, H. et al. Large-scale analyses of human microbiomes reveal thousands of small, novel genes. *Cell* **178**, 1245–1259.e14 (2019).
46. Donia, M. S. & Fischbach, M. A. Small molecules from the human microbiota. *Science* **349**, 1254766 (2015).
47. Sugimoto, Y. et al. A metagenomic strategy for harnessing the chemical repertoire of the human microbiome. *Science* **366**, eaax9176 (2019).
48. Culp, E. J. et al. Hidden antibiotics in actinomycetes can be identified by inactivation of gene clusters for common antibiotics. *Nat. Biotechnol.* **37**, 1149–1154 (2019).
49. The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).

Acknowledgements

C.d.l.F.-N. holds a Presidential Professorship at the University of Pennsylvania, is a recipient of the Langer Prize by the AIChE Foundation and acknowledges funding from the Institute for Diabetes, Obesity, and Metabolism, the Penn Mental Health AIDS Research Center of the University of Pennsylvania, the Nemirovsky Prize, and the Dean's Innovation Fund from the Perelman School of Medicine at the University of Pennsylvania. Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number R35GM138201 and the Defense Threat Reduction Agency (DTRA; HDTRA11810041 and HDTRA1-21-1-0014). All figures were prepared using the Biorender drawing toolkit. The authors thank E. Broset and all members of the de la Fuente Lab for insightful discussions.

Author contributions

M.D.T.T. and C.d.l.F.N. designed the experiments. M.D.T.T. performed all biological activity experiments, mechanism of action assays and animal models. M.C.R.M., O.C. and E.N. performed algorithm development and database searches. M.D.T.T. and C.d.l.F.N. wrote the first draft of the manuscript. M.C.R.M., O.C. and E.N. revised the manuscript. C.d.l.F.N. supervised and administered the work.

Competing interests

The authors declare no competing interests.

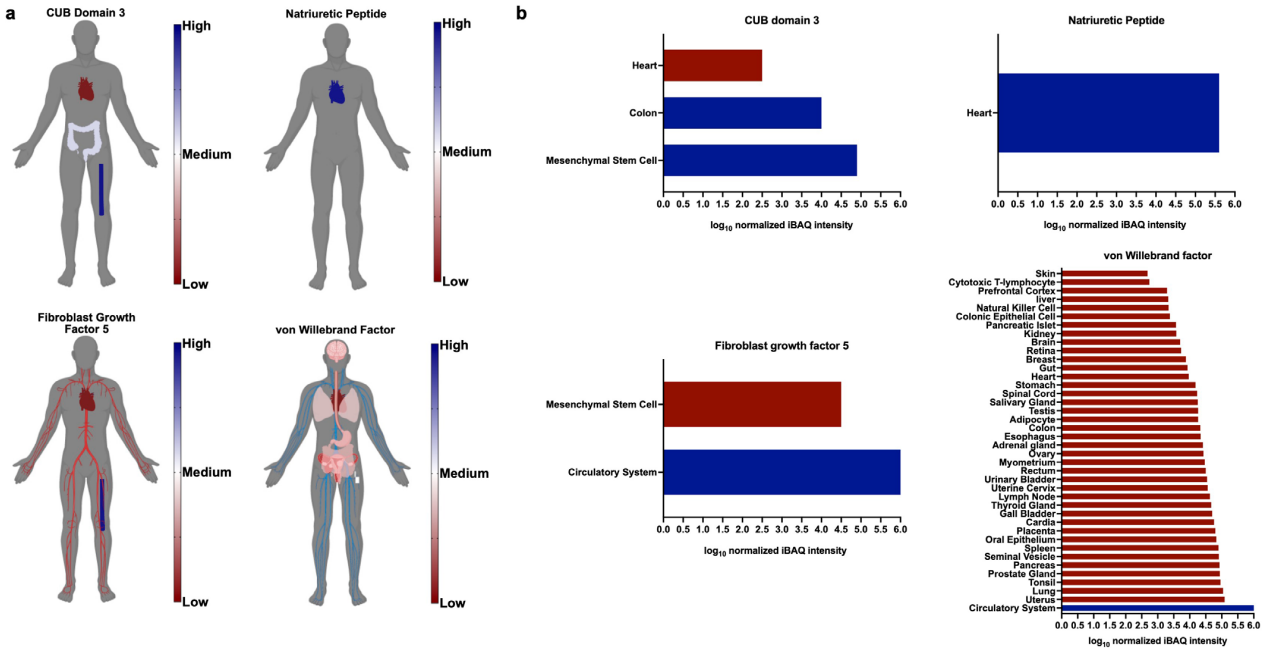
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41551-021-00801-1.

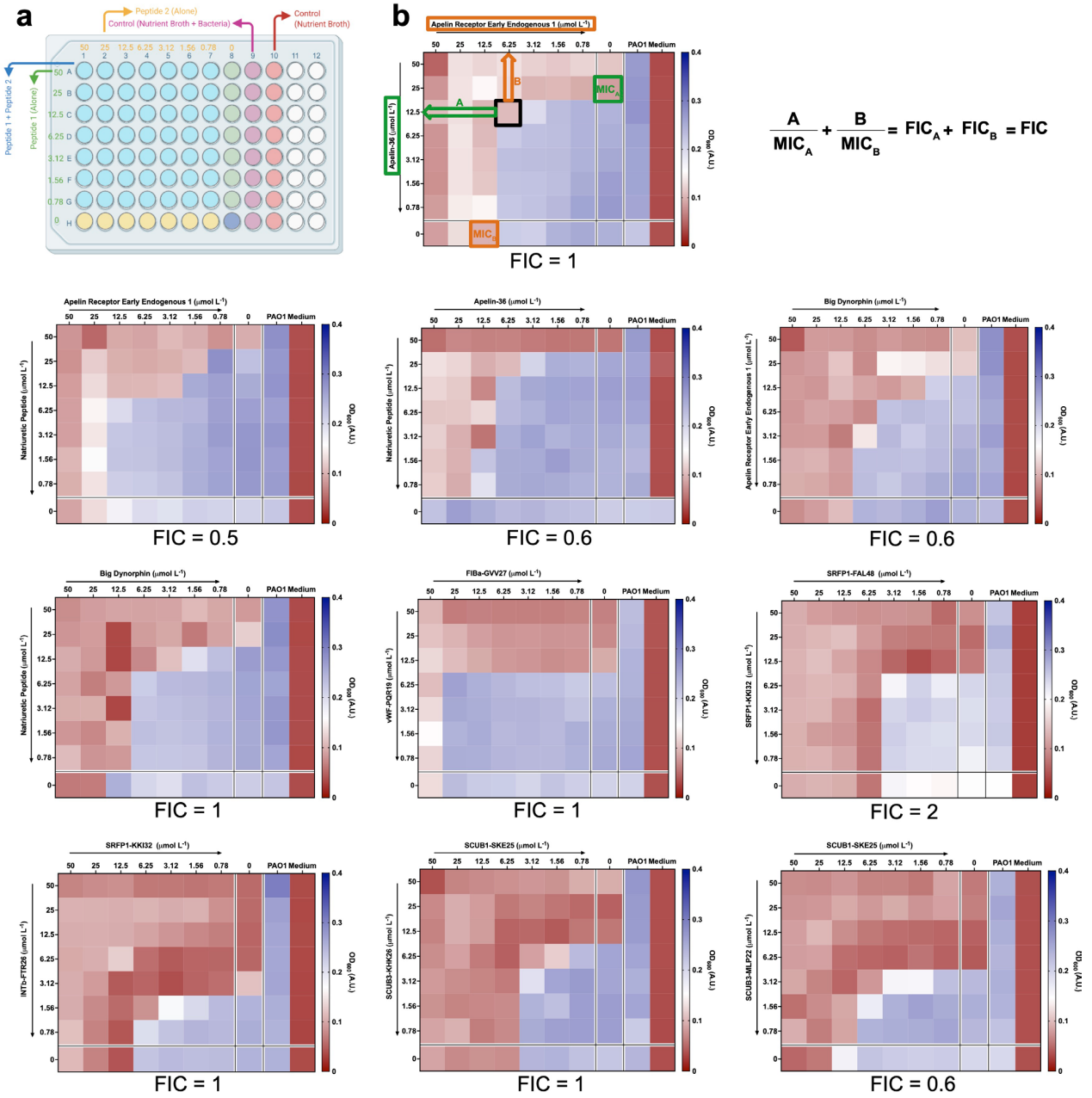
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41551-021-00801-1.

Correspondence and requests for materials should be addressed to Cesar de la Fuente-Nunez.

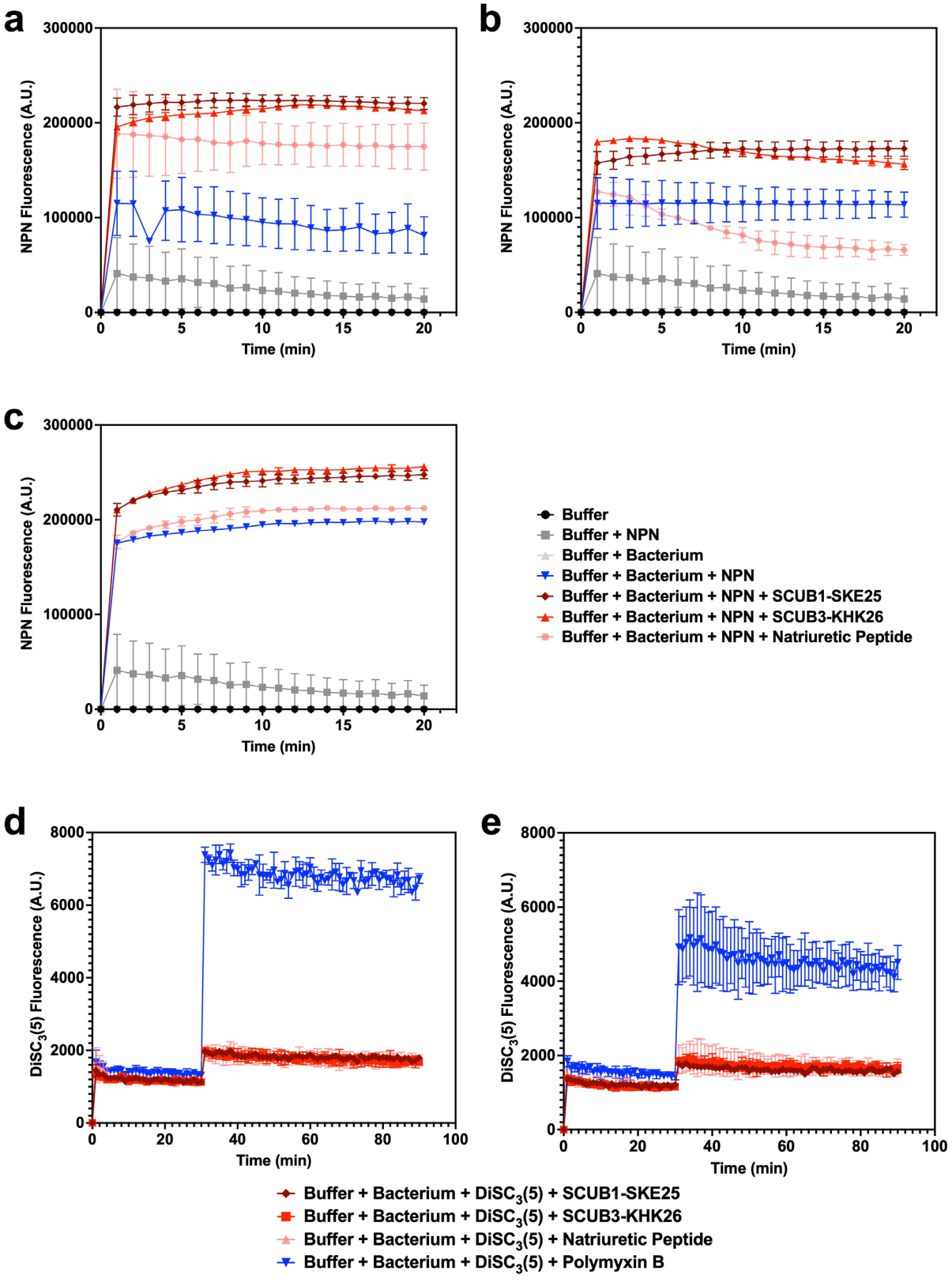
Peer review information *Nature Biomedical Engineering* thanks Mohan Babu, Kim Lewis and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



Extended Data Fig. 1 | Expression levels of proteins containing encrypted peptides. a, Schematic of the biogeographic region within the human body where proteins containing encrypted peptides are located. Expression levels are displayed in a gradient; organs in blue indicate high expression levels and organs in red, low expression levels. **b**, Normalized expression level values expressed in $\log_{10}$ intensity based absolute quantification (iBAQ), a commonly used metric for protein abundance[46].

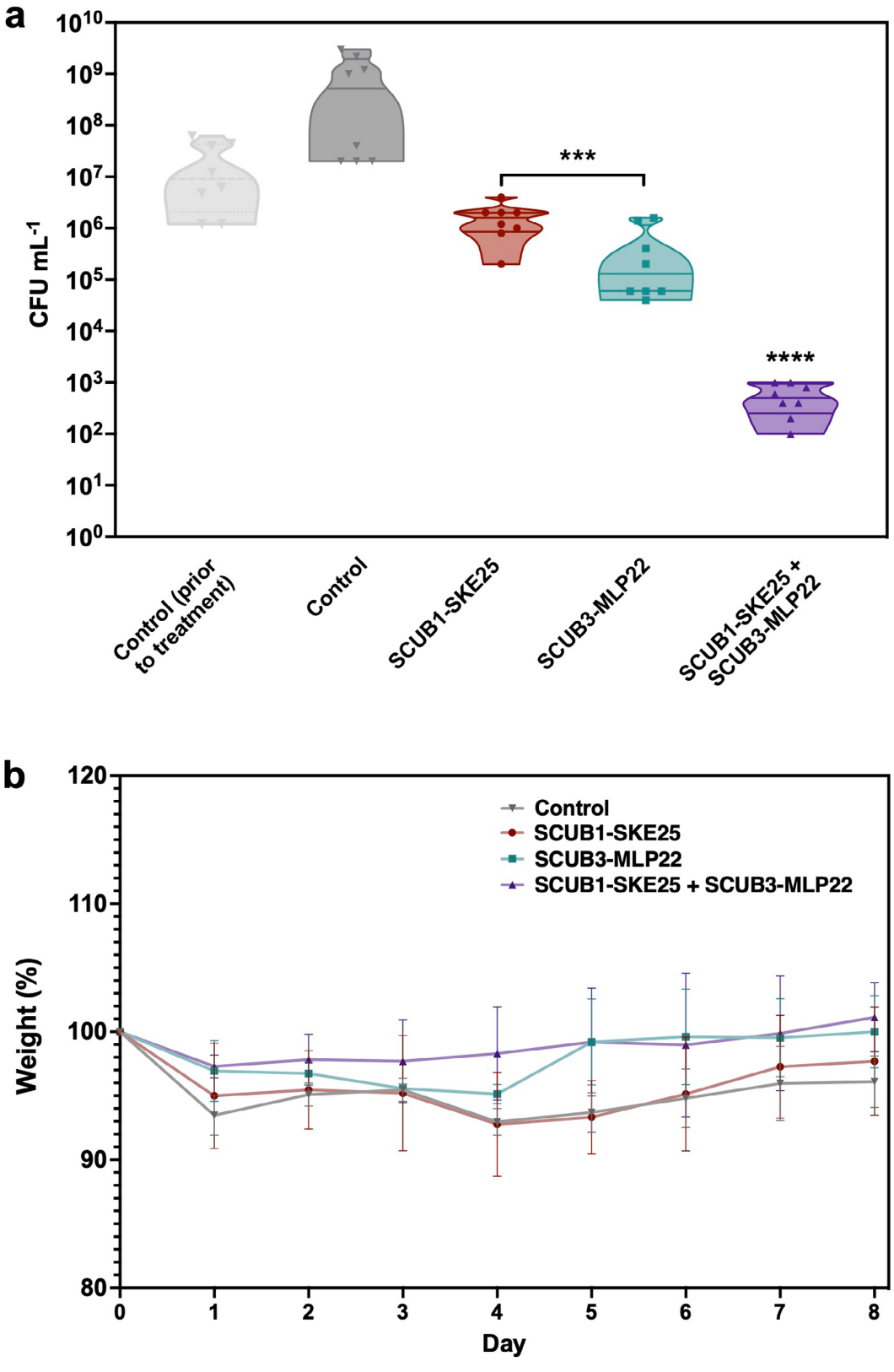


Extended Data Fig. 2 | Synergy between encrypted peptides found within the same area of the human body. a, Experimental layout of the 96-well plates used for two-way synergy experiments using pairs of encrypted peptides. The following encrypted peptides were used: apelin-36, apelin receptor early endogenous 1, natriuretic peptide, big dynorphin, FIBa-GVV27, vWF-PQR19, SRFP1-KKI32, SRFP1-FAL48, INTb-FTR26, SCUB1-SKE25, SCUB3-KHK26, and SCUB3-MLP22. Briefly, two-fold dilutions ranging from 0 to 50 μmol L^{-1} of the peptide solutions were plated in 96-well plates and 10^6 bacterial cells in NB were added to each well to reach a final volume of 200 μL. **b**, The FIC value, which indicates the degree of synergy between two antimicrobial agents against a target microorganism (in this case, *P. aeruginosa* PAO1) was calculated based on the MICs of the peptides used alone and in combination. FIC index values ≤0.5 indicate synergy; additive effects are captured by 0.5≤FIC≤1; 1≤FIC≤4 indicates indifference; and FIC index ≥4 represents antagonism. Assays were performed in three independent replicates.



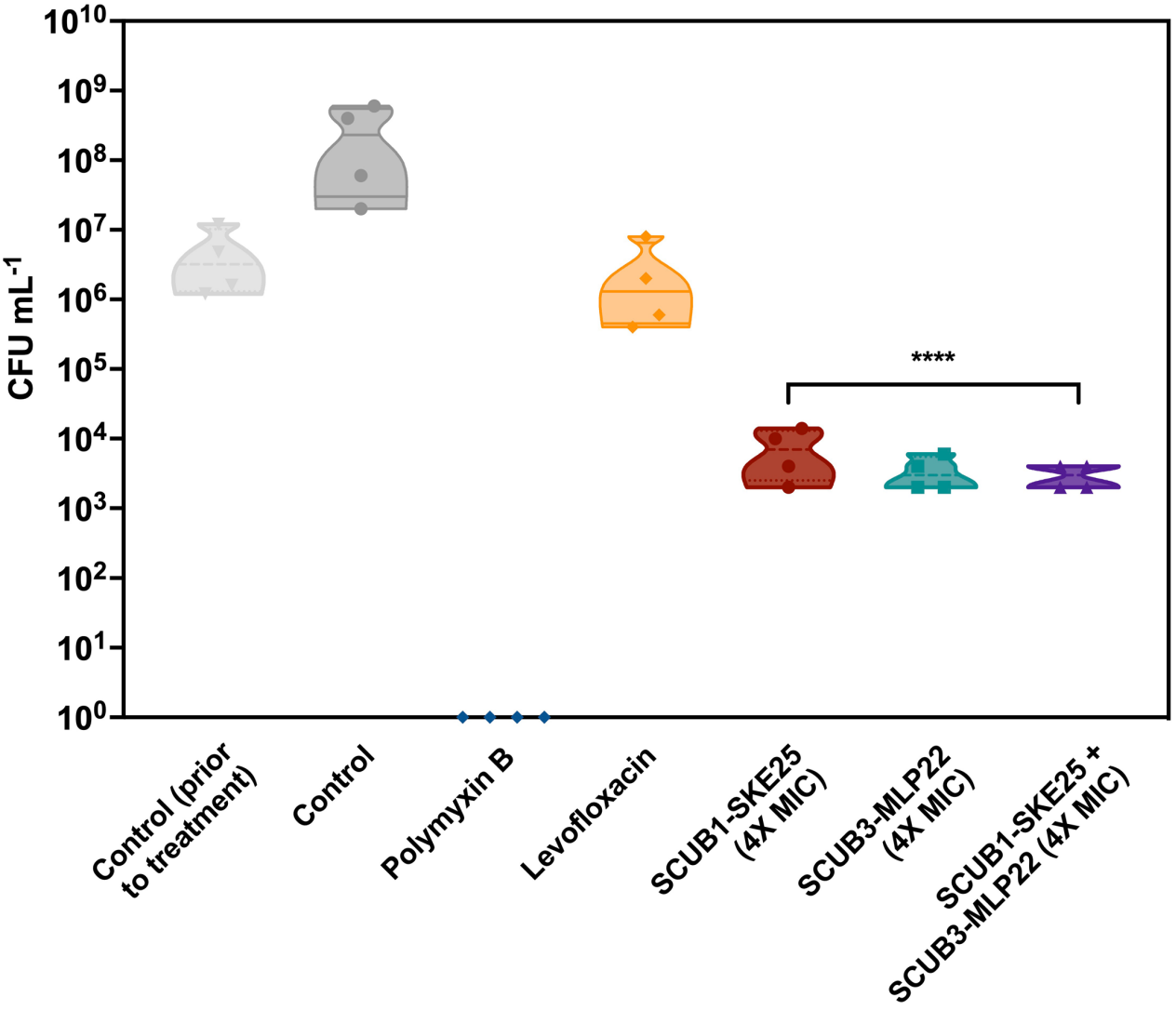
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Membrane permeabilization and depolarization assays for several bacterial strains. a, Outer membrane permeabilization experiments showed that encrypted peptides (SCUB1-SKE25, SCUB3-KHK26 and natriuretic peptide) permeabilized the outer membranes of *P. aeruginosa* PAO1, *B. fragilis* ATCC25285, and *S. epidermidis* as much as they permeabilized the *A. baumannii* ATCC19606 outer membrane (Fig. 3e). **b**, Cytoplasmic membrane depolarization assays performed against *P. aeruginosa* PAO1 and the gut commensal *B. fragilis* ATCC25285. As shown for *A. baumannii* ATCC19606 (Fig. 3d), the encrypted peptides did not depolarize the cytoplasmic membrane. Data in **a** and **b** are the mean ± s.d. Assays were performed in three independent replicates.



Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Anti-infective activity and synergistic interactions of encrypted peptides against *P. aeruginosa* PAO1 in a neutropenic thigh infection mouse model. a, SCUB1-SKE25 (25 μmol L^{-1}; 77.9 μg mL^{-1}) and SCUB3-MLP22 (25 μmol L^{-1}; 66.9 μg mL^{-1}) showed inhibitory activity against *P. aeruginosa* PAO1, especially when used in combination at their MIC (obtained from *in vitro* synergy experiments; 3.12 and 6.25 μmol L^{-1}, respectively). **b**, Mouse weight was monitored throughout the duration of the neutropenic thigh infection model (8 days) and under all conditions tested to rule out potential toxic effects mediated by the encrypted peptides. The statistical significance in **a** was determined using one-way ANOVA, ***$p < 0.001$, ****$p < 0.0001$, features on the violin plots represent median and upper and lower quartiles. The data in **b** are the mean ± s.d. Eight mice were used per group.



Extended Data Fig. 5 | Anti-infective activity of encrypted peptides in a neutropenic thigh infection mouse model. Treatment with 4-fold MIC of SCUB1-SKE25 (100 μmol L^{-1}; 311.6 μg mL^{-1}) and SCUB3-MLP22 (100 μmol L^{-1}; 267.6 μg mL^{-1}) alone and in combination (at 12.5 and 25 μmol L^{-1}, respectively). Both monotherapy and combination therapy displayed similar antimicrobial activity against *A. baumannii* ATCC19606 to treatment groups using the MIC of each peptide. Polymyxin B and levofloxacin were used as controls, the former of which completely cleared the infection. The statistical significance was determined using one-way ANOVA, ****$p < 0.0001$, features on the violin plots represent median and upper and lower quartiles. Four mice were used per group.

nature portfolio

Corresponding author(s): Cesar de la Fuente-Nunez

Last updated by author(s): Aug 25, 2021

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	EnVision multimode microplate reader.
Data analysis	Prism v9

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The main data supporting the results in this study are available within the paper and its Supplementary Information. The list of the identified encrypted peptides, and source data for the normalized abundance of genes encoding different protein classes, for amino acid frequency in encrypted peptides compared to known antimicrobial peptides, for expression levels of proteins containing encrypted peptides, for antimicrobial activity (in vitro and in vivo), as well as data on synergistic interactions, evolution of resistance and mechanism of action, are provided with this paper.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes were calculated on the basis of previous literature on similar experimental settings.
Data exclusions	No data points were excluded.
Replication	Experiments were in general performed three times (two replications) to confirm the results. The results are only reported if the replications were successful and all experimental results could be confirmed.
Randomization	Mice and cages were randomly allocated to the experimental groups.
Blinding	Peptides were solubilized and labelled with internal codes for completely unbiased experimentation by the same researcher.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	CD-1 mice (Charles River Laboratory).
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	The two animal model protocols, skin-abscess infection mouse model (806763) and neutropenic thigh-infection mouse model (807055), were revised and approved by the University Laboratory Animal Resources (ULAR) from the University of Pennsylvania.

Note that full information on the approval of the study protocol must also be provided in the manuscript.