Unifying structural signature of eukaryotic α-helical host defense peptides

Nannette Y. Yount[a,b,c], David C. Weaver[d], Ernest Y. Lee[e], Michelle W. Lee[e], Huiyuan Wang[a,b,c], Liana C. Chan[a,b,c], Gerard C. L. Wong[e,f,g], and Michael R. Yeaman[a,b,c,h,1]

[a]Division of Molecular Medicine, Department of Medicine, Los Angeles County Harbor–University of California, Los Angeles, Medical Center, Torrance, CA 90509; [b]Division of Infectious Diseases, Department of Medicine, Los Angeles County Harbor–University of California, Los Angeles, Medical Center, Torrance, CA 90509; [c]Los Angeles Biomedical Research Institute, Department of Medicine, Harbor–University of California, Los Angeles, Medical Center, Torrance, CA 90502; [d]Department of Mathematics, University of California, Berkeley, CA 94720; [e]Department of Bioengineering, University of California, Los Angeles, CA 90095; [f]Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095; [g]California NanoSystems Institute, University of California, Los Angeles, CA 90095; and [h]Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA 90024

Edited by Lawrence Steinman, Stanford University School of Medicine, Stanford, CA, and approved February 14, 2019 (received for review November 9, 2018)

Diversity of α-helical host defense peptides (αHDPs) contributes to immunity against a broad spectrum of pathogens via multiple functions. Thus, resolving common structure–function relationships among αHDPs is inherently difficult, even for artificial-intelligence–based methods that seek multifactorial trends rather than foundational principles. Here, bioinformatic and pattern recognition methods were applied to identify a unifying signature of eukaryotic αHDPs derived from amino acid sequence, biochemical, and three-dimensional properties of known αHDPs. The signature formula contains a helical domain of 12 residues with a mean hydrophobic moment of 0.50 and favoring aliphatic over aromatic hydrophobes in 18-aa windows of peptides or proteins matching its semantic definition. The holistic α-core signature subsumes existing physicochemical properties of αHDPs, and converged strongly with predictions of an independent machine-learning–based classifier recognizing sequences inducing negative Gaussian curvature in target membranes. Queries using the α-core formula identified 93% of all annotated αHDPs in proteomic databases and retrieved all major αHDP families. Synthesis and antimicrobial assays confirmed efficacies of predicted sequences having no previously known antimicrobial activity. The unifying α-core signature establishes a foundational framework for discovering and understanding αHDPs encompassing diverse structural and mechanistic variations, and affords possibilities for deterministic design of antiinfectives.

antimicrobial | host defense | antiinfective | amphipathic | bioinformatics

Antimicrobial host defense peptides (HDPs) are an evolutionarily ancient arm of host immunity that first arose in prokaryotes as a means to counter microbial competitors. Subsequently, such peptides evolved through adaptive radiation to exist in all classes of eukaryotes, where they continue to act in first-line defense against infection (1). Extensive studies have established that such peptides are not indiscriminant detergents, but rather have complex and multimodal mechanisms of action (2–4).

As a group, α-helical HDPs (αHDPs) are among the most rapidly evolving molecules characterized to date. Moreover, genes encoding αHDPs are under strong positive selection, affording a high degree of mutational tolerance despite being limited by the biophysical constraints of an amphipathic helix (5–8). When compounded over an evolutionary timescale, this process has generated an exceptionally diverse repertoire of peptides capable of exerting multiple antimicrobial mechanisms. However, such diversity and context-dependent activity have also presented challenges to identifying common αHDP structure–activity relationships (SARs). While a number of groups have used computational or quantitative SAR (QSAR) methods, these efforts have largely focused on drug candidate optimization (9–12). As a result, other than charge or amphipathicity, resolving the unifying physicochemical requisites in three-dimensional space that confer antimicrobial activity to native αHDPs has remained elusive. Moreover, results from machine-learning–based methods are difficult to translate into first principles. Thus, whereas machine-learning–based classifiers can discriminate between HDPs and non-HDPs, none to date have yielded more generalizable definitions of αHDPs in a formulaic manner.

Here, knowledge-based annotation and iterative pattern recognition analyses of bioinformatic databases yielded a unifying definition of structural elements common to native eukaryotic αHDPs. Termed the α-core signature, this sequence formula positions specific amino acid residue patterns on polar or nonpolar peptide facets in the context of a three-dimensional amphipathic helix. This signature formula retrieved all known families of αHDPs when queried against well-established proteomic databases. The veracity of the α-core signature to recognize cognate semantic patterns was validated by machine-learning methods in silico. Furthermore, a recently developed support vector machine (SVM) classifier affirmed the α-core signature corresponds with peptide ability to induce negative Gaussian curvature (NGC) in microbial membranes. Topologically, this function enables a wide spectrum of membrane perturbing and ensuing mechanisms that exert direct antimicrobial functions (13–15). As proof of concept, sequences scoring highest relative to multiple parameters retrieved by the α-core

Significance

The emergence of multidrug-resistant pathogens is outpacing development of new antibiotics, which remains a slow and empirical process. Although machine-learning–based metrics exist, no principle-based formulae have been applied to define α-helical host defense peptides (αHDPs). Here, knowledge-based annotation of bioinformatic databases was validated with artificial intelligence methods to converge on a common structural framework for eukaryotic αHDPs. This multidimensional structural formula successfully predicted antimicrobial activity in α-helical peptides and domains embedded in larger proteins that had no previously known antimicrobial activity. The α-core signature provides a foundational definition for evolutionary diversity among αHDPs, sheds light on host defense immunology, and may accelerate the process of antiinfective discovery and development.

Author contributions: N.Y.Y., G.C.L.W., and M.R.Y. designed research; N.Y.Y., E.Y.L., and H.W. performed research; N.Y.Y., D.C.W., and M.R.Y. contributed new reagents/analytic tools; N.Y.Y., E.Y.L., M.W.L., H.W., L.C.C., G.C.L.W., and M.R.Y. analyzed data; and N.Y.Y., E.Y.L., G.C.L.W., and M.R.Y. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: mryeaman@ucla.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1819250116/-/DCSupplemental.

signature formula, but having no known or anticipated antimicrobial function, were synthesized and assayed against a panel of bacterial and fungal pathogens. All of the predicted peptides exerted significant antimicrobial activity in vitro.

Beyond resolving a unifying structural signature among diverse αHDPs, and its use to discover previously unknown antimicrobial peptides, the current findings address the more foundational question of how to define HDPs. Understanding functions mediated by the α-core signature can offer insights into the structural commonalities among HDPs that allow pleiotropic mechanistic profiles of diverse αHDPs. Furthermore, the α-core signature advances knowledge of host defense immunology and its evolution, and enables possibilities for deterministic design of antiinfectives to meet the challenge of drug-resistant infections.

Results

Derivation of the α-Core Sequence Formula. Proteomic databases searched using the logic-based strategy identified consensus motifs in proteins or peptides having α-helical sequence and three-dimensional structure (Fig. 1*A*). A mathematical formula was then derived and iteratively refined to fulfill consensus positional and physicochemical patterns of amino acid residues characteristic of known αHDPs (Fig. 1*B*). These analyses revealed that the inclusion of proline interrupted helical spans, while glycine and alanine are typically well tolerated in helices in membrane mimetic environments (refs. 16–18 and Fig. 1*C*). For this reason, proline was excluded from the explicit α-core formula, although degeneracy at positions 1 and 10 allows for this residue, whereas glycine and alanine were allowed at all positions (Fig. 1).

Next, a systematic approach enabled optimization of the linear α-core sequence formula (Fig. 2). This process iteratively refined 18 successive residue positions of a canonical right-handed α-helix consistent with consensus patterns of antimicrobial sequences. When translated into three dimensions, this formula describes an idealized amphipathic helix, with distinct hydrophobic and hydrophilic facets (Fig. 1*A*). Positions interposing the polar and nonpolar faces were assigned a value of "X" within the linear formula, yielding a polar angle (θ) minimum of 120° and maximum of 180°. Through this iterative pattern recognition and refinement process, an 18-residue generalized amphipathic formula was resolved that fulfilled the α-core consensus sequence and integrated a positional, hydrophobic, and hydrophilic residue pattern as distributed along a three-dimensional amphipathic helix. Sequences adherent to this signature ranged from 11 to 16 residues in length, corresponding to a span of 3–4.5 turns across an α-helix. This pattern was ultimately termed the α-core signature formula and represented all major classes and families of eukaryotic αHDPs.

Primary Database Searches Using the α-Core Sequence Formula. The refined α-core sequence formula was used as a query against public domain sequence databases (UniProtKB, Swiss-Prot). The formula was implemented using successive 12-residue scanning windows (Fig. 1*A*), and searches were further refined to sequences eukaryotic in origin, within proteins of less than 200 residues in length, containing a signal sequence, and within 50 residues of the C terminus. These characteristics enriched biological plausibility, as the vast majority of eukaryotic αHDPs are secreted into the extracellular environment.

Resulting raw data consisted of more than 70,000 sequences, often representing multiple target hits shifted along amphipathic helical spans of a single protein. Hits were compiled and duplications extracted to yield a nonredundant dataset of ∼13,000 unique sequences for subsequent studies. Dataset proteins lacking a signal sequence motif were excluded to generate a final dataset of ∼5,200 sequences (Table 1). Inclusion of the signal peptide allowed for enrichment of sequences that are secreted for potential activity toward extracellular targets. To retain a sharp focus, protein sequences not containing a signal sequence are the topic of our separate studies.

Efficiency of the α-Core Sequence Formula. The α-core signature exhibited excellent sensitivity and specificity, retrieving all major classes of αHDPs and 106 distinct helical peptide families, including 93% of all known individual αHDPs (more than 800 peptides). Further specificity of the α-core signature was affirmed by the observation that peptides not retrieved by the formula were typically very short mature peptides of <12–14 residues, or containing incomplete or interrupted helical domains, such as amphibian "ranabox" peptides categorized as helix–turn–helix in structure.

Beyond its considerable sensitivity, the formula was also tunable in terms of specificity or biochemical classification. For example, in selected searches, sequences were gated as having an isoelectric point (pI) of ≥8.5, and carried out in stages of either 0–50 or 51–200 residues. For the 0- to 50-residue set, ∼71% of retrieved sequences were known αHDPs; whereas ∼27% of the identified sequences were α-helical antimicrobial proteins in the 51- to 200-residue set. Illustrating the veracity of this approach, the α-core formula retrieved 827 of the ∼885 known αHDPs represented in the 555,594-sequence UniProtKB/Swiss-Prot database (93% of known sequences). This outcome is equivalent to an in silico enrichment of ∼90-fold. Moreover, the formula retrieved an additional 4,412 sequences representing peptides having hypothetical antimicrobial functions. The α-core algorithm retrieved only 0.17% of all sequences in the known Swiss-Prot/UniProt database, attesting to its high specificity.

Residue Composition Frequency Within Search Hits. Sequences retrieved using the α-core signature formula were diverse, but logically categorized and analyzed as belonging to one of three distinct groups: αHDPS, Toxins, or Other (Table 1).

αHDPs. Within the αHDP group, the most abundant residues were the low–molecular-mass amino acids, glycine and alanine, present at frequencies of 14% and 13%, respectively. Of the polar amino acids, lysine was the most abundant residue (14%), preferred over the cationic residue arginine at a 5:1 ratio (Fig. 3*A*). In a more detailed analysis (Fig. 3 *B* and *C*), this preference was greatest in amphibians (10:1), and somewhat less in arthropods and mammals (3:1). Of the hydrophobic residues, leucine was most abundant and preferred over isoleucine or valine at a ratio of ∼2:1.

Toxins and Other peptides. In addition to αHDPs, Toxins, and Other peptide sequences were also retrieved using the α-core formula (Tables 1 and 2). One highly represented family was that of toxin/venom peptides, making up ∼17% of the dataset. In the Toxin group, glycine (14%), cysteine (12%), and alanine (10%) were the predominant residues, whereas cationic lysine (10%) and arginine (6%) residue frequencies were lower compared with αHDPs (Fig. 3*A*). The abundance of cysteine within the Toxin group reflects the fact that the vast majority of these proteins are stabilized by disulfide arrays and have multiple structural domains including β-sheet, γ-core, as well as α-helical domains in their modular topology (19). Within the Other group, residue frequency trends were generally similar to αHDPs and Toxins; however, this group had comparatively lower frequencies of glycine.

Spatial Distribution of Residues in αHDPs. The distribution of residues in three-dimensional space of the α-core signature placed polar amino acids at positions 2, 5, 6, 9, 12, 13, 16, and 17, and nonpolar residues at positions 3, 4, 7, 8, 11, 14, 15, and 18 (Fig. 3*C*). Molecules returned using this signature revealed distinct patterns recurring within phylogenetic classes. For example, at positions 5, 6, and 9 in arthropods, the most common polar residue is lysine, with positions 5 and 6 typically bracketed by

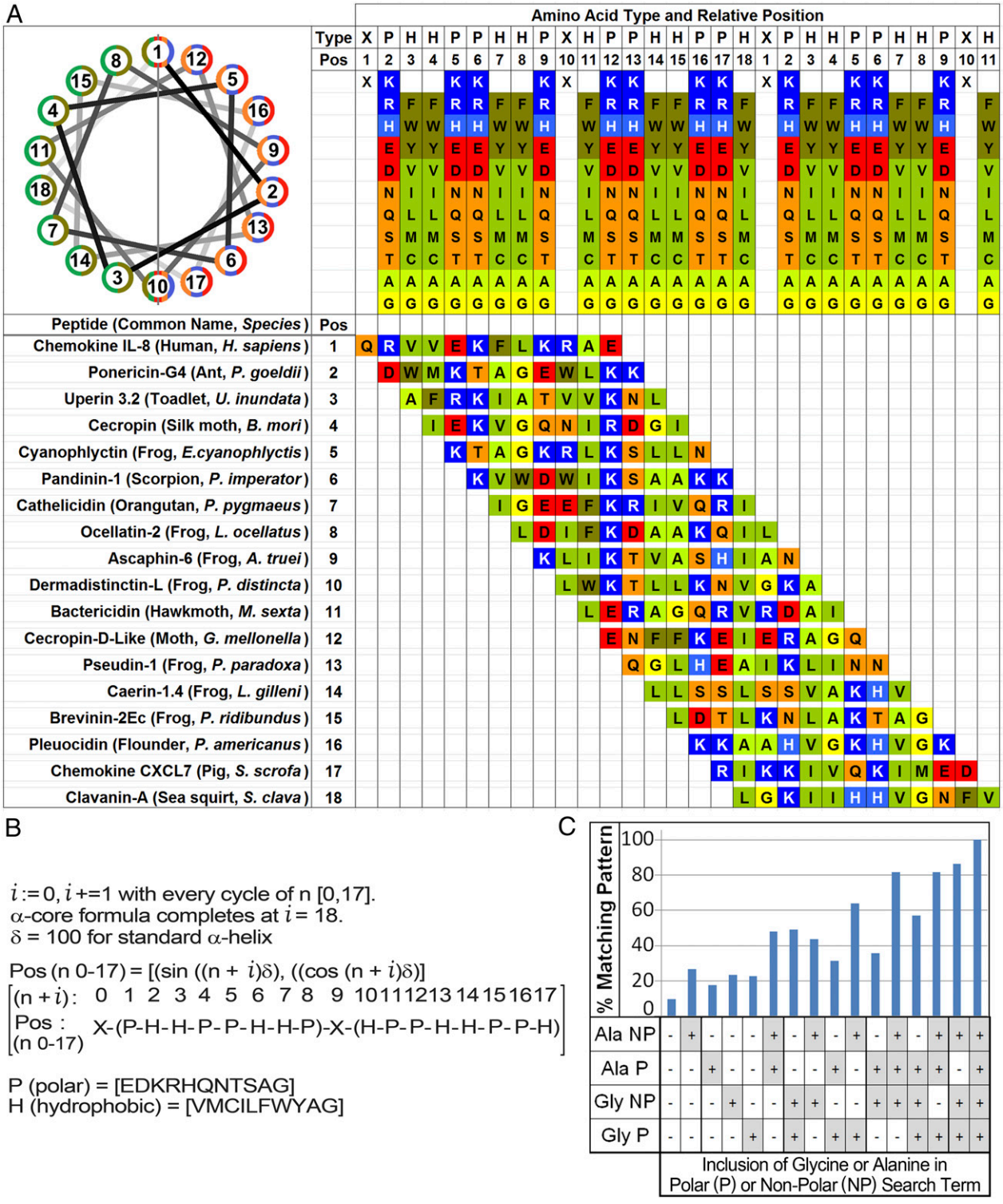

Fig. 1. The α-core consensus signature and scanning window of prototypic αHDPs. (*A*) Helical wheel and linear depiction of the 18-residue α-core sequence formula. For alignment with αHDP prototypes, 18 initiation points for the 12-residue scanning window of the formula are shown, representing the process by which the ProSite pattern search tool was utilized to query UniProtKB databases. Hydrophobic residues (F, W, Y, V, I, L, M, and C) are in various shades of green, with greater hydrophobicity indicated by increasingly darker hues. Hydrophilic residues (K, R, H, E, D, N, Q, S, and T) are represented by the following: blue, cationic; red, anionic; and orange, uncharged polar. Alanine (light green) and glycine (yellow) are included with both hydrophilic and hydrophobic groups. H, hydrophobic; P, polar. (*B*) Mathematical representation of the α-core formula. (*C*) Optimization of α-core sequence formula by localization of glycine and alanine residues. Iterative refinement of the α-core formula was carried out to assess the requirement for glycine and/or alanine as a component of either the polar (hydrophilic) or nonpolar (hydrophobic) residue set. Percentage of returned sequences from a control αHDP dataset of more than 600 peptides are shown.

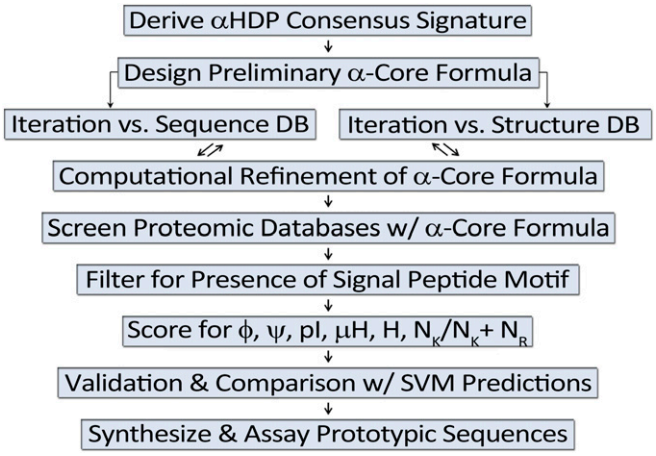

Fig. 2. The α-core signature iterative optimization algorithm. Development and implementation of the α-core signature formula and its application proceeded through quantitative and logic-driven steps to optimize sensitivity and specificity in retrieval of known and hypothetical αHDPs.

leucine residues. By comparison, amphibian αHDPs share a high frequency of lysine at position 5, but much less so at positions 6 or 9. This trend is extended in mammals, with position 4 dominated by glutamine. On the contrary, lysine occurs much more frequently at positions 2 and 12 in mammals than in amphibians or arthropods. Furthermore, the observed enrichment of tryptophan at the first nonpolar residue at helix position 3 is in agreement with prior findings that documented an increased level of tryptophan near the beginning of α-helical spans (20). Thus, the α-core signature identified unforeseen patterns of spatial distribution and periodicity among seemingly diverse αHDPs.

Physicochemical Parameters Distinguishing αHDPs. Study proteins were analyzed for individual or composite parameters distinguishing αHDPs: hydrophobic moment (μH), net charge (Q), hydrophobicity (H), pI, and lysine-to-arginine ratio ($N_K/N_K + N_R$; Table 1). Univariate analyses revealed that μH and either Q or pI were important parameters discriminating αHDPs from other sequences (Fig. 4). Multivariate analyses revealed that one of the most distinguishing characteristics of αHDPs results from an integration of hydrophobic moment and charge (μHQ). When μHQ data were binned, rankings representing the top 25th (μHQ_{25}) and 50th (μHQ_{50}) percentiles were most predictive of sequences having antimicrobial activity (Table 2). Similarly, direct positive correlations between $N_K/N_K + N_R$ and mean hydrophobicity were observed in sequences retrieved by the α-core signature (Fig. 5), consistent with sequence trends deduced from geometric and topological requirements for membrane permeation (21, 22).

Correlation of αHDP Signature and NGC Propensity. Datasets returned by the α-core formula were categorized by μHQ (μHQ_{25}; μHQ_{50}) and screened with a recently developed SVM-based machine-learning classifier to assess propensity to induce NGC in target membranes (13–15). A strong, statistically robust correlation was seen between sequences identified by the α-core formula and those predicted by the SVM protocol to be membrane active (σ score; Fig. 6 and Table 2). Comparison of the two methods in relative monotonic ranking of the datasets achieved highly significant correlations ($P = 2.2 \times 10^{-6}$ to 1.2×10^{-46}) for nearly all groups (Fig. 6). The αHDPs were predicted to be membrane active with an average probability of 0.93 (scale 0–1) for the μHQ_{25} dataset, and 0.91 for the μHQ_{50} subset by the SVM. Similarly, the Toxin and Other datasets were also predicted to be membrane active by the SVM, with average probabilities of 0.94 (μHQ_{25}) or 0.88 (μHQ_{50}) for the Toxins, and 0.95 (μHQ_{25}) or 0.89 (μHQ_{50}) for the Other sequences. As a negative control, a nonsense version of the α-core formula that prevented amphipathic helicity was used to probe a randomized sample of the UniProtKB database. On SVM analysis, the probability of this sequence ensemble to have antimicrobial activity was only 0.06 (Table 2 and *SI Appendix*).

αHDP Prediction and Proof of Concept. The α-core signature formula retrieved many sequences having no prior known or anticipated antimicrobial function. To test predictive accuracy, 10 candidate sequences representing diverse functional families retrieved by the formula were synthesized by standard F-moc chain assembly, purified, authenticated, and assessed for antimicrobial efficacy in vitro (ref. 23 and Fig. 7*A*). All test peptides exerted potent activity against one or more of the prototypic human pathogens studied (Fig. 7*B*). Many of these peptides exerted activities greater than a prototypic control αHDP (LL-37) at either pH 7.5 (simulating bloodstream) or pH 5.5 (simulating abscess). Of great interest, several peptides exhibited preferential pH or target organism activity. For example, anti-staphylococcal activity of the sequence from human IFN-γ was essentially specific to pH 7.5, while minimal at pH 5.5. Alternatively, the anti-candidal activities of sequences from human IL-13 or IL-21 were minimal at pH 7.5 but significant at pH 5.5. By comparison, sequences such as Pp2 from *Phytophthora* demonstrated striking activity against all organisms tested at both pH 7.5 and 5.5. Other interesting patterns included *Acinetobacter* being susceptible to all test peptides under either pH condition. A comparative statistical analysis of the activities of study peptides is included in *SI Appendix*, Fig. S1.

Discussion

The present findings define the α-core signature as a multidimensional pattern integrating physicochemical and three-dimensional relationships common to evolutionarily diverse αHDPs. Furthermore, this signature represents a foundational structure–activity framework for delineating necessary or sufficient determinants of antimicrobial activity. For example, the present data identify key physicochemical and structural patterns that appear requisite for antimicrobial activity of αHDPs, including the following: (*i*) a periodic enrichment of low–molecular-weight amino acids glycine, alanine, and to a lesser extent serine; (*ii*) a pI typically >8 with a preference for lysine over arginine; by comparison, β-sheet HDPs tend to use arginine vs. lysine (21); (*iii*) a minimum helical core domain spanning ∼12 residues (∼18 Å; ref. 20), with a mean hydrophobic moment of ∼0.50; and (*iv*) propensity for aliphatic over aromatic hydrophobic residues. Collectively, these signature elements enable the quantitative definition of αHDPs, which have previously been described generally as cationic, amphiphilic peptides. Therefore, the α-core signature affords unique insights into evolutionary, phylogenetic, and molecular effectors of antiinfective immunology, as well as fundamental strategies for rapid discovery, prototyping, and optimization of antiinfective peptide or biologic designs.

The current findings also extend beyond prior understanding of properties that correlate with microbicidal activities of αHDPs, such as small mass, cationic charge, and amphipathicity (2, 4, 24). While they can exhibit varied effects on distinct pathogens, all αHDPs initiate their antimicrobial mechanisms by engaging, perturbing, and/or transcending the target cell membrane. Importantly, the α-core signature correlates with the

Table 1. Biophysical properties of retrieved dataset proteins

Group	*n*	*Q*	$N_K/N_K + N_R$	μH	*H*	pI
αHDPs	907	2.0	0.82	0.50	0.42	9.23
Toxins	787	1.1	0.62	0.48	0.32	7.44
Other	3,539	0.5	0.57	0.49	0.41	6.98

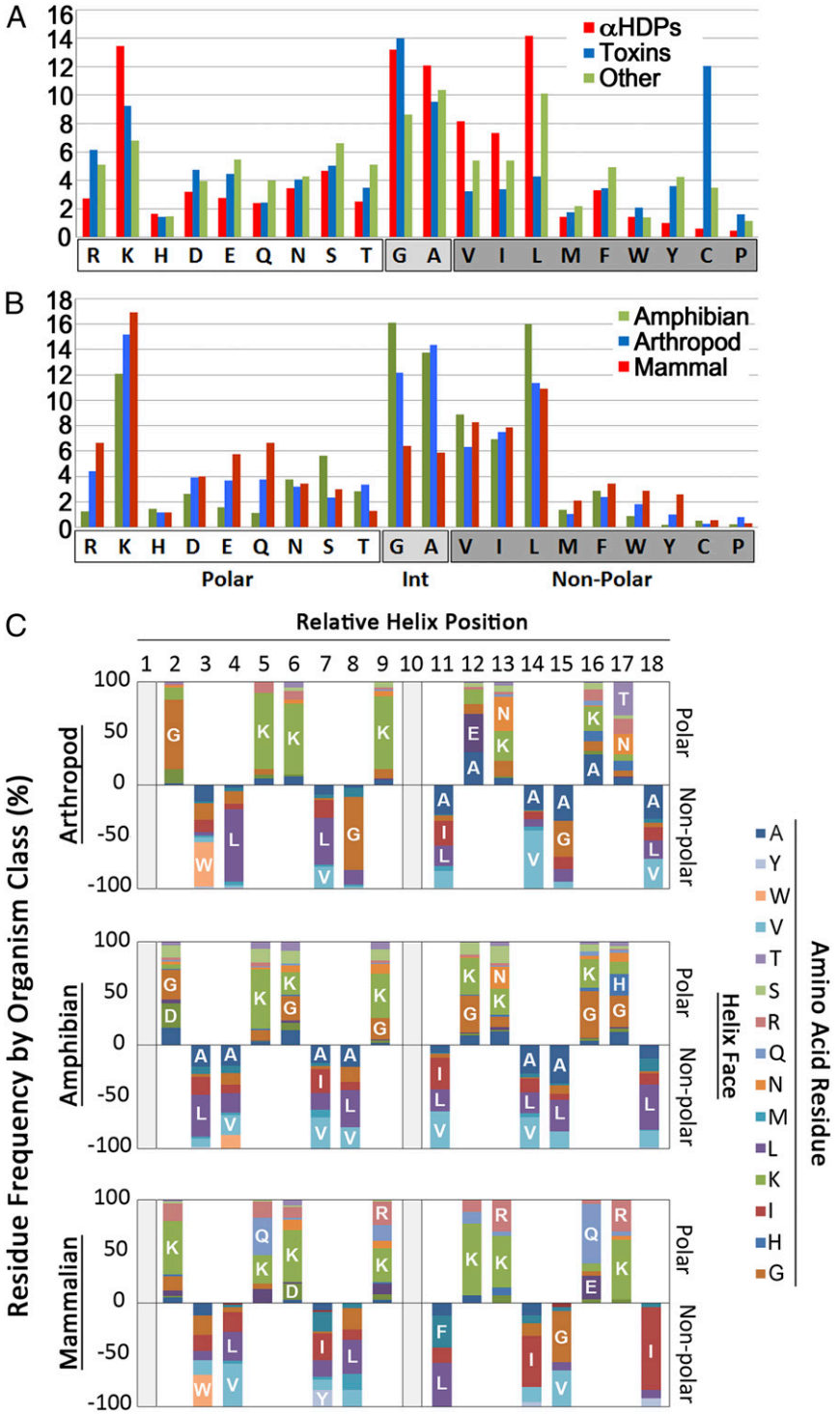

Fig. 3. Positional and spatial amphipathic residue frequency by organism class and functional group. (*A*) αHDP residue frequency by functional group. Relative amino acid percentages are displayed for the αHDP (red), Toxin (blue), and Other (green) groups. (*B*) αHDP residue frequency by organism class. Relative amino acid percentages for amphibians (green), arthropods (blue), and mammals (red). Int, intermediate. (*C*) Percentages of individual residues on either the polar or nonpolar peptide face of study peptides are represented as various color blocks. Residues above the *x*-axis are found on the polar face of retrieved peptides, and residues below the axis are found on the nonpolar face.

ability of a peptide to induce NGC in microbial target membranes, such that deviations away from the signature diminishes their ability to do so. Positive charge is thought to contribute to antimicrobial selectivity, given relative electronegative charge and inward-rectified potential ($\delta\psi$) of prokaryotic cell membranes or mitochondria of eukaryotic pathogens (2, 25–28). Moreover, the high concentration of lipids with negative intrinsic curvature in a target membrane, as is the case for most bacteria, also contributes to specific activity of αHDPs (29–31) by modifying the Gaussian modulus. The present finding that lysine is

Table 2. Biophysical properties of retrieved dataset proteins

	Total	μH*Q > 1.5						SVM
Group	*n*	*n*	%	μH	*Q*	*H*	pI	%
αHDPs	907	237	26	0.57	4.5	0.29	10.3	93
Toxins	787	123	16	0.54	4.0	0.21	8.9	94
Other	3,539	376	11	0.65	3.8	0.37	8.7	95
		μH*Q > 1.0						
αHDPs	907	426	47	0.54	3.6	0.36	10.1	91
Toxins	787	200	25	0.54	3.4	0.24	8.7	88
Other	3,539	659	19	0.61	3.2	0.36	8.4	89
Nonsense	1,820							6

preferred in αHDPs corroborates the observation that there exist regions in amino acid sequence space with specific ratios of lysine, arginine, and hydrophobic residues that are conducive to topological remodeling of the membrane (21, 32, 33). Serine residues also contribute, but appear to play a more supporting role as spacers and mediators of hydrogen bonding (19, 34).

Cross-validation using ML suggests the α-core signature reveals basic principles not previously accessible regarding essential structural elements of αHDPs. For example, these peptides have periodic distribution of low–molecular-weight and low-steric bulk residues glycine, alanine, and to a lesser extent serine. The systematic interposition of these residues among larger charged and hydrophobic residues likely accounts for the three-dimensional conformational optimization of αHDPs in membrane target environments (17, 18) to achieve local charge density (e.g., +1/70 $Å^2$ to +1/350 $Å^2$) necessary to maximize entropy gain for initial membrane binding (29). Consistent with this view, proline residues disrupt requisite helical periodicity in the α-core signature, and native αHDPs lack proline in the helical domains (positions 2–9 and 11–18 within the α-core signature). This efficiency in identifying helical spans compares favorably with algorithms reported to correctly identify helices, with accuracies ranging from 81.5 to 94% (35–37). Previous QSAR and other in silico methods have been useful to assess SARs in specific peptide-based libraries (9–12). However, these approaches leave the unique biophysical and three-dimensional properties conferring antimicrobial mechanisms in the native peptide realm relatively unexplored. Similarly, artificial-intelligence–based methods are typically optimized to identify multifactorial trends rather than common principles. In contrast, the present investigation integrated knowledge-driven as well as stochastic bioinformatic and pattern recognition strategies to resolve the α-core signature as it is conserved throughout the eukaryotic proteome. Thus, the α-core signature overcomes historic limitations that no general model had defined what αHDPs are, especially given their complex and diverse mechanistic profile.

Adaptation of the α-core signature within specific phylogenetic groups revealed interesting insights into evolution of αHDPs and other host defense molecules. While an abundance of lysine as a preferred cationic residue in αHDPs has been reported (21), the present data revealed frequencies of lysine differ among αHDPs in distinct phylogenetic groups. For example, lysine–arginine ratios in arthropods or mammals (3:1) differ significantly from those of amphibians (10:1). In general, arginine-rich HDPs tend to have β-sheet secondary structure (e.g., β-sheet defensins of higher mammals), whereas lysine-rich peptides tend to be α-helical. There are interesting functional consequences of lysine vs. arginine substitution as well. For example, as the arginine-to-lysine ratio is increased in a α-helical peptides, less hydrophobicity is needed to generate NGC required for membrane perturbation. Consequently, the membrane residence time of the peptide is decreased, and it begins to function progressively more like a cell-penetrating peptide that transcends the membrane without killing the cell, rather than a cidal peptide that kills the target cell via long-lived pores (32).

Moreover, glycine and alanine were highly represented in αHDP helices from arthropods or amphibians (G + A content, 30% or 26%, respectively), whereas their frequency is much lower in mammals (12%). The evolutionary basis for such residue preferences is unknown; however, there are intriguing hypothetical possibilities. Given their propensity to allow conformational freedom that deforms α-helical structure in aqueous environments, it is conceivable that glycine and alanine may allow peptides to remain in relatively unstructured, inactive (e.g., nontoxic) conformations before interacting with the microbial target (38). A number of studies have demonstrated that αHDPs can be unstructured in aqueous environments, adopting their strong α-helical conformation only upon encountering hydrophobic lipid membrane environments characteristic of microorganisms (2, 17, 18, 39, 40). Likewise, the fact that amphibian, arthropodian, and mammalian αHDP sequences all fulfill the saddle-splay selection rule for lysine, arginine, and hydrophobes (ref. 21 and Fig. 5) suggests different sequence selection pressures imparted structural adaptations that optimized antimicrobial function in context of the specific physiology of the host.

Based on the sensitivity and specificity with which the α-core signature formula retrieved αHDPs, it was applied as a discovery tool to seek: (*i*) unknown or uncharacterized antimicrobial proteins; (*ii*) sequences with an assigned function that also have a previously unidentified antimicrobial function; or (*iii*) helical antimicrobial peptide domains within larger proteins. This analysis uncovered a large number of sequences having no previously known direct antimicrobial functions and exhibited high scores on the ML classifier. Confirming the predictive accuracy of the α-core signature formula, all study sequences synthesized and assayed were found to have robust antimicrobial activity in vitro. Examples included members of the γ-chain–dependent IL family, along with type II IFNs. These results are consistent with unforeseen antimicrobial activities discovered in chemokines (2, 4, 26) and selected ILs (4, 17, 18, 41, 42). Of the ILs, sequences derived from turkey IL-5 (tIL-5) exhibited striking efficacy against all pathogens at pH simulating bloodstream (pH 7.5) and abscess (pH 5.5) environments. Sequences derived from human ILs (hIL-7, hIL-13, hIL-21) also exerted significant activity, enhanced at pH 5.5.

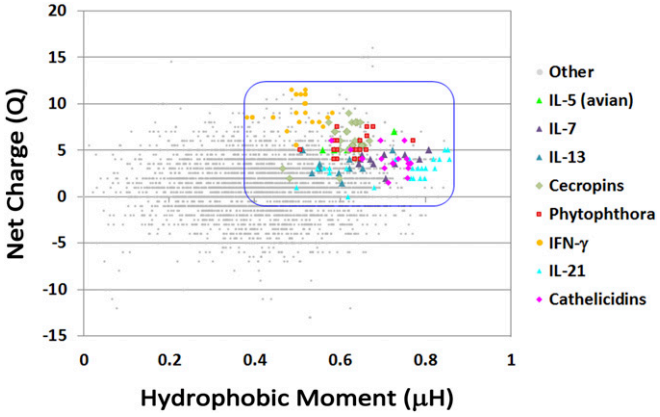

Fig. 4. Prioritization of predicted α-helical peptide sequences selected for experimental assessment. Values for net charge (*Q*) versus hydrophobic moment (μH) are shown for the retrieved peptide dataset. All retrieved sequences are shown in gray. Peptide groups among which prototypic sequences were selected for further characterization are shown in color. For comparison, prototypic αHDPs are shown in pink and olive.

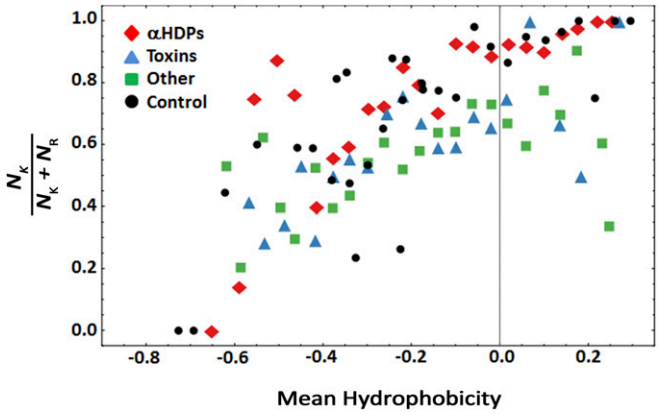

Fig. 5. Comparative physicochemical properties of α-core helices. Percentage of lysine (N_K) relative to arginine (N_R) expressed as ($N_K/N_K + N_R$) versus hydrophobicity (H) in study αHDPs and toxins. Preference of lysine compared with arginine is reflected in an increased value of H for peptides capable of generating NGC in membranes as predicted by the saddle-splay rule (21). Binned along x-axis into 31 bins based on mean hydrophobicity; helical AMPs in black are from the APD2 database.

Likewise, peptides derived from human or bat IFN-γ had significant activity against most test organisms, with an efficacy profile somewhat greater at pH 7.5 than 5.5. Unanticipated αHDP sequences were also discovered in nonvertebrate species. For example, among the highest scoring peptides were discovered in the ubiquitous plant pathogen, *Phytophthora parasitica*. Of these, Pp-2 was particularly active against bacteria and fungi at pH 7.5 and pH 5.5.

It should be noted that methods in which peptides are assessed for antimicrobial activity in artificial media significantly underestimate the potent efficacy of many HDPs. Therefore, the antiinfective roles of these peptides may be even greater in vivo. For example, we have previously shown the efficacy of representative HDPs to be logarithmically greater in whole human blood, plasma, or serum compared with artificial media (43). Similarly, ionic strength can significantly influence HDP efficacy (44). Divalent and monovalent ions behave very differently in their interactions with peptide sequences that form electrostatically distinct domains as is evident in the α-core. These effects may extend beyond contributions imposed by the Poisson–Boltzmann formalism (45).

Current findings also suggest insights into evolution and multifunctionality of HDPs across the phylogenetic spectrum. For example, we have shown that amphiphilic α-helical HDPs having specific sequence patterns can assemble into superhelical coiled-coil protofibrils that organize dsDNA into ligands that optimally engage TLR9 receptors (46, 47). Therefore, beyond their direct antimicrobial activities, such αHDPs appear to activate synergistic immune responses that promote holistic host defense. In these ways, the α-core signature discovered in the present study may enable precise definition of these SARs, which, along with target membrane perturbation, are likely integrated into HDP sequences and their multiplex superstructures.

The α-core signature, and its demonstrated ability to identify antimicrobial sequences (Fig. 8), underscores its practical implications to mine proteomic databases for antimicrobial sequences. In this respect, it reveals structural determinants that are essential for membrane-specific antimicrobial activity, compared with those which have been adapted to a given host. In turn, this knowledge should enable peptide or biologic designs that optimize antimicrobial activity and pharmacology relative to toxicity. As there have been historical challenges in translating HDPs into antiinfectives, a clearer understanding of their structure–mechanism relationships is of high priority. For example, as discovered in the current work, enhanced insights into pharmacophores of HDP sequences—beyond simply positive charge and amphiphilicity—will be invaluable to drive next steps toward development of antiinfectives that address the threat of pathogen-resistant conventional antibiotics. In this respect, of special interest is the ability of the α-core signature formula to recognize cryptic or antimicrobial sequences embedded in larger proteins. In these ways, the α-core signature may accelerate the identification or engineering of innovative antiinfectives that are urgently needed to meet the rising challenge of antibiotic resistance.

Materials and Methods

Bioinformatic Identification of the α-Core Formula. To identify a consensus formula that encompassed all major classes of αHDPs, successive sequence alignments were carried out using prototypical α-helical antimicrobial peptides in CLUSTAL W (https://www.ebi.ac.uk/Tools/msa/clustalw2/). Alignments were refined for consensus optimization using MEGA 6 (48). Through iterative analyses, an 18-residue generalized amphipathic formula was resolved that fulfilled the overall consensus sequence and integrated the positional hydrophobic

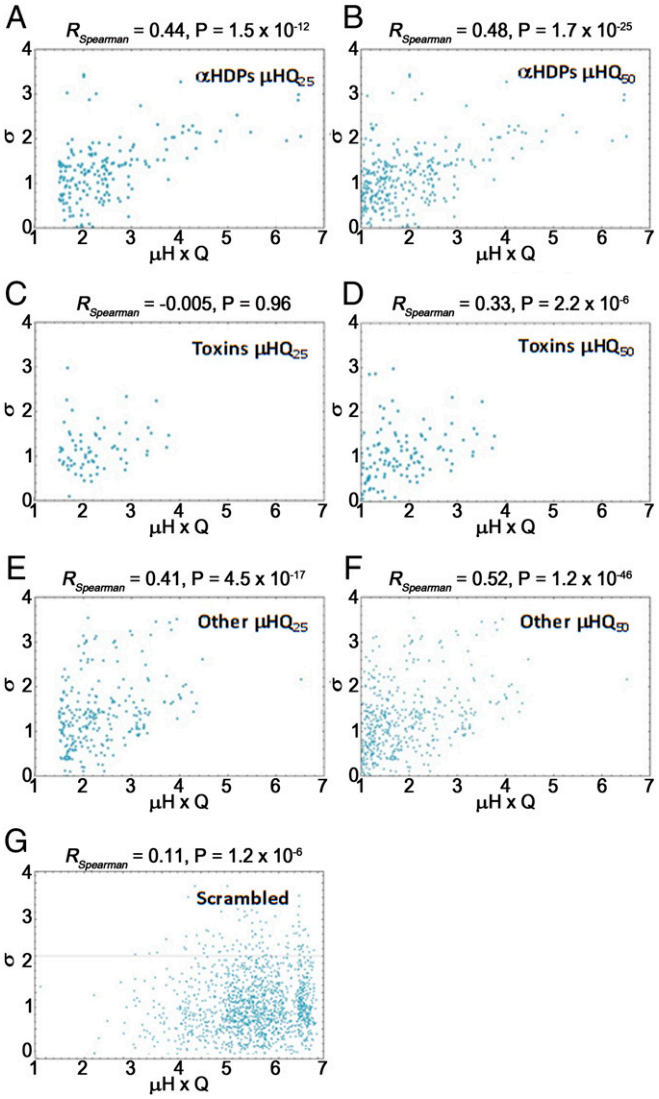

Fig. 6. Spearman correlation analysis of study peptide μHQ and SVM σ. Spearman correlations for the α-core μHQ_{50}, μHQ_{25}, and scrambled peptide groups and SVM σ are shown. Panels depict μHQ_{25} and μHQ_{50} for αHDP (*A* and *B*), Toxin (*C* and *D*), Other (*E* and *F*), and scramble groups, respectively. *P* values indicate level of significance.

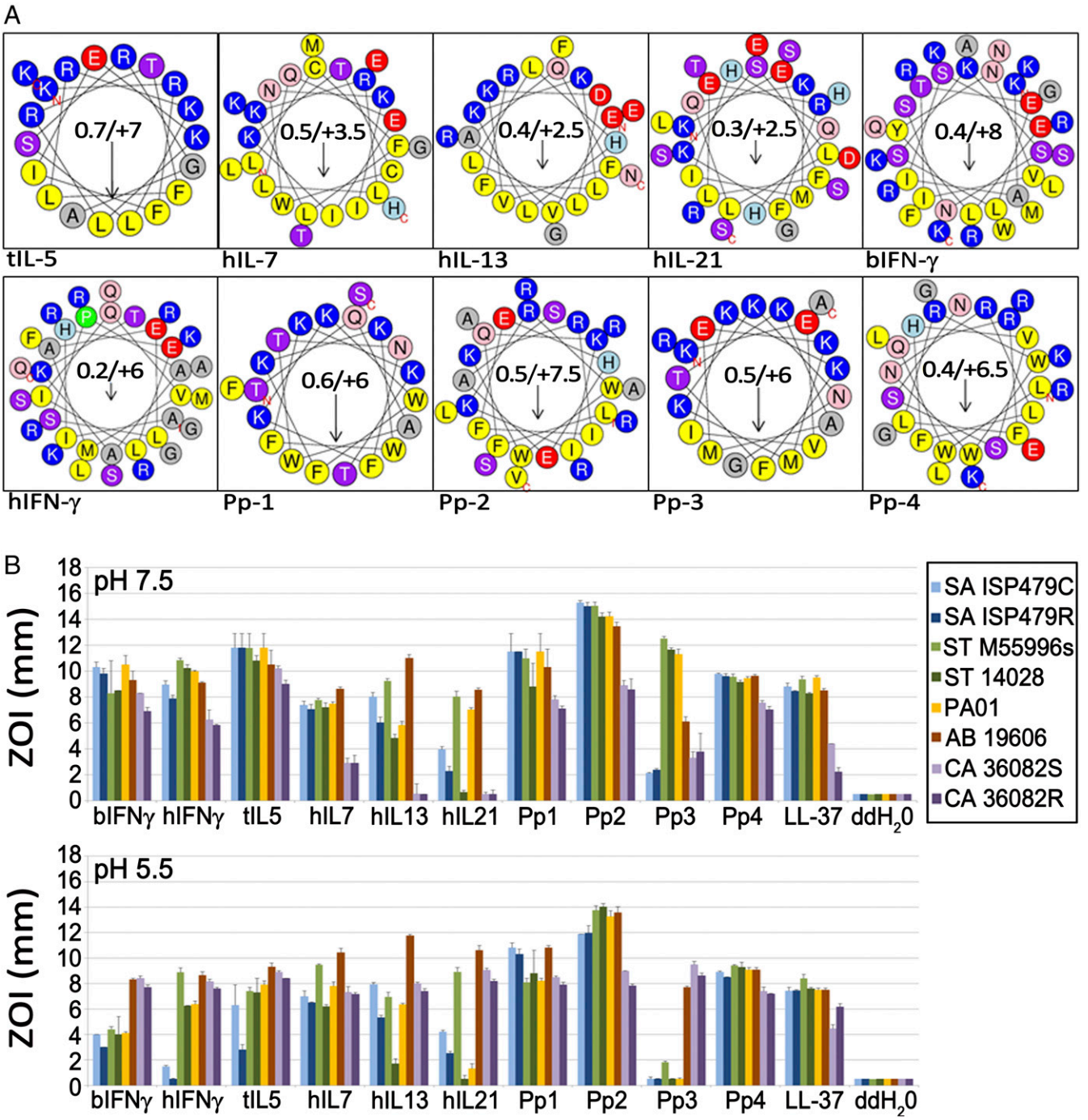

Fig. 7. Efficacy of predicted antimicrobial peptides retrieved by the α-core signature formula. (*A*) Helical wheel depiction of synthesized study peptides. Peptides derived from the following proteins were assayed for antimicrobial activity: ILs [IL-5, *Meleagris gallopavo* (common turkey); IL-7, human; IL-13, human; IL-21, human]; IFNs [IFN-γ, *Myotis davidii* (vesper bat); IFN-γ, human]; *Phytophthora parasitica* uncharacterized sequences (Pp-1, Pp-2, Pp-3, and Pp-4). Hydrophobic moment (μH), charge (Q), and moment magnitude and direction are indicated. Coloration: cationic full charge (KR), blue; partial charge (H), light blue; anionic, red; polar, yellow; tiny, gray; polar (NQ), pink, and (TS), purple. (*B*) Microbicidal activity of study test peptides versus a panel of prototypic Gram-positive (*S. aureus*), Gram-negative (*S. typhimurium*, *P. aeruginosa*, and *A. baumannii*) and fungal (*C. albicans*) pathogens at two pH conditions simulating bloodstream (pH 7.5) or phagolysosomal (pH 5.5) environments.

and hydrophilic residue pattern. This pattern was ultimately termed the α-core signature formula, representing all major classes and families of eukaryotic αHDPs.

Assignment of Residue Polarity in the α-Core Formula. Within the α-core formula, individual residues were categorized as either hydrophobic or hydrophilic as per the Wimley–White hydrophobicity scale (49) that has been empirically derived and includes contributions from the peptide bond. One exception was for alanine (A), which was also included with the hydrophobic residues as per the Eisenberg (50) and Kyte and Doolittle (51) hydrophobicity scales. This assignment was made in part due to preliminary data that localized alanine to hydrophobic facets of many αHDPs.

Accuracy of the α-Core Formula in αHDP Recognition. The α-core formula was queried against the Protein Data Bank 3D database (www.wwpdb.org/) to

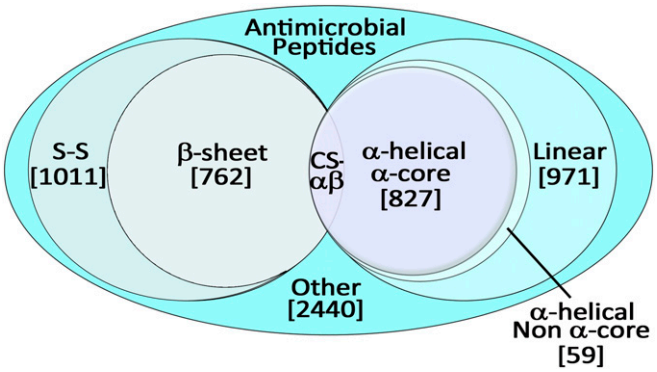

Fig. 8. Classification of αHDPs and other HDP classes retrieved by the α-core signature from the UniProtKB database. Data represent results from a query of all UniProtKB sequences retrieved with the search terms: keyword [antimicrobial] + length [0–200] + organism class [Eukaryota]. Peptides manually scored for structure and type and sequences within each group are enumerated. Sequences were categorized as follows: (*i*) α-helical returned by the α-core formula (α-helical α-core); (*ii*) α-helical not returned by the α-core formula (non–α-core); (*iii*) linear cysteine-free sequences with many being enriched for individual amino acids (linear peptides); (*iv*) β-sheet defensin or defensin-like peptides (β-sheet peptides); (*v*) other disulfide-containing peptides (S–S peptides); (*vi*) Cys-stabilized peptides containing α-helix and β-sheet domains (CS–αβ); and (*vii*) other group including proteins with no known structure or homology to other known sequences, as well as groups that are not traditionally categorized as HDPs such as enzymes, DNA binding proteins, protease inhibitors, lectins, and toxins.

assess the fidelity of the formula in recognizing the helical domain signature within proteins or peptides. One hundred nonredundant retrieved structures were evaluated for helicity of the target sequence. Hits were considered to be positive if the target sequence was ≥75% helical based on known or predicted structure (www.wwpdb.org/).

In addition to optimizations for the retrieval of three-dimensional structural elements, the α-core formula was also iteratively refined to assess the requirement for glycine and/or alanine as a component of either the polar (hydrophilic) or nonpolar (hydrophobic) residue set. Variant versions of the formula containing all possible G/A combinations localized to either the polar, nonpolar, or both residue groups were queried against a control αHDP dataset of more than 600 peptides that was manually collated from documented αHDPs from the UniProtKB database.

Use of α-Core Formula as a Database Query Tool. The α-core sequence formula was used with the ScanProsite search tool (https://prosite.expasy.org/scanprosite/) to seek iterative patterns in the UniProtKB Swiss-Prot database. Iterative searches with varying lengths of the α-core formula found that a scanning sequence motif of 12 positions in length was most efficient at retrieving the majority of known αHDP sequences (Fig. 1*B*).

The minimized sequence motif was then used to probe the UniProtKB Swiss-Prot and TrEMBL protein databases. The formula was advanced one position at a time through 18 iterations to represent an entire 18-residue helical wheel span. To explore structural and biological relationships, the ScanProsite search results were also filtered by the following: (*i*) sequence length <200 residues; (*ii*) eukaryotic origin; and (*iii*) localization to the C-terminal region of the target protein using a "X(0,50)>" logical operator.

α-Core formula (first iteration of 18). [KREDNQSTHAG]-[VILMCFWYAG]-[VILMCFWYAG]-[KREDNQSTHAG]-[KREDNQSTHAG]-[VILMCFWYAG]-[VILMCFWYAG]-[KREDNQSTHAG]-X-[VILMCFWYAG]-[KREDNQSTHAG]-[KREDNQSTHAG]-X(0,50)>

Nonsense formula (first iteration of 18). [KREDNQSTHAG]-[VILMCFWYAG]-[KREDNQSTHAG]-[VILMCFWYAG]-[KREDNQSTHAG]-[VILMCFWYAG]-[KREDNQSTHAG]-[VILMCFWYAG]-X-[VILMCFWYAG]-[KREDNQSTHAG]-[VILMCFWYAG]-X(0,50)>

Signal Peptide and Physicochemical Parameter Assessment. Retrieved datasets were screened for the presence or absence of a signal peptide structural domain using SignalP 4.1 (www.cbs.dtu.dk/services/SignalP/). Additionally, hydrophobic moment (μH) (52), mean hydrophobicity (H) (53), net charge [Q, K, and R (+1); H (+0.5); D and E (−1)], and K and R residue frequency were determined in batch using Python algorithms specifically created for this study. The Wimley–White hydrophobicity scale values were used for hydrophobicity and hydrophobic moment calculations (49). Sequence pI was determined using ExPasy Compute PI (https://web.expasy.org/compute_pi/).

The equation for mean hydrophobicity (H) is as follows:

$$\langle H\rangle = \frac{1}{N}\sum_{n=1}^{N} H_n.$$

The equation for hydrophobic moment (μH), $\delta = 100°$, is as follows:

$$\langle \mu H\rangle = \frac{1}{N}\left[\left(\sum_{n=1}^{N} H_n \sin(n\delta)\right)^2 + \left(\sum_{n=1}^{N} H_n \cos(n\delta)\right)^2\right]^{\frac{1}{2}}.$$

Machine-Learning Validation of Datasets. To further characterize the datasets retrieved by the α-core formula, a previously developed SVM-based classifier (13–15) was used to screen the obtained sequences for antimicrobial activity. Briefly, the SVM classifier was trained to optimally partition 243 known α-helical sequences present in the Antimicrobial Peptide Database (ref. 54; APD, aps.unmc.edu/AP/main.php) from 243 decoy peptides with no reported antimicrobial activity. A minimal model consisting of 12 physicochemical descriptors was obtained using feature selection from 1,588 starting descriptors (55, 56). The SVM generated 12 descriptors from the peptide sequence and output a score σ specifying the distance of the peptide from the 11-dimensional hyperplane separating antimicrobial and nonantimicrobial sequences. Experimental validation of computational predictions of membrane activity was carried out using small-angle X-ray scattering (SAXS) experiments. Here, model membranes were incubated with synthesized α-helical test peptides, and induced NGC was quantified. Comparison of σ scores with calibrating SAXS data revealed a strong correlation between the ability to generate NGC in membranes and σ. Thus, a large, positive σ score correlates with the ability to induce NGC in membranes, whereas a negative σ score indicates a lack of membrane-permeating activity. These features also correspond with the presence or absence of antimicrobial activity, respectively (13–15). Sequences retrieved from the α-core search tool were screened using this algorithm, and σ scores calculated. Spearman correlations were quantified between σ and α-core metrics using Mathematica.

Synthesis of Candidate Peptides. Selected candidate αHDPs were synthesized using standard automated F-moc chain assembly, purified by reverse-phase HPLC, and authenticated by mass spectroscopy as previously detailed (43, 57). Lyophilized peptides were resuspended in ddIH_20 and stored in aliquots at −20 °C.

Assay for Antimicrobial Activity. Putative αHDP or like peptides were assayed for microbicidal activity using a well-established radial diffusion method adjusted to pH 5.5 or 7.5 (58). Sequences representing αHDP μHQ_{25} and μHQ_{50} prototypes and not known to have antimicrobial activity were prioritized for study. Sequences representing three diverse high-scoring groups were selected: ILs (tIL-5, hIL-7, hIL-13, and hIL-21); IFNs (bIFN-γ, hIFN-γ); and uncharacterized sequences from *P. parasitica.* These sequences were synthesized and tested for microbicidal activity against prototypic bacterial and fungal pathogens in vitro. Cathelicidin LL-37 (Peptides International), a prototypic human αHDP, was used as a comparator. The efficacy of these peptides was evaluated against a panel of human pathogens, including αHDP susceptible and resistant isogenic pairs: *Staphylococcus aureus* [ISP479C/ISP479R (59)]; *Salmonella typhimurium* [MS5996s/MS14028 (43)]; *Pseudomonas aeruginosa* (PA01); *Acinetobacter baumanni* (19606); and *Candida albicans* [36082S/36082R (60)]. In brief, logarithmic-phase organisms were inoculated (10^6 CFU/mL) into molten buffered molecular-grade agarose plates. Peptides (10 μg) were introduced into wells in the seeded matrix and incubated for 3 h at 37 °C. Nutrient overlay medium was applied, and assays incubated at 37 or 30 °C for bacteria or fungi, respectively. After 24 h, zones of inhibition were measured. Independent experiments were repeated a minimum of two times.

ACKNOWLEDGMENTS. These studies were supported in part by the NIH–National Institute of Allergy and Infectious Diseases (NIAID) Systems Immunobiology (Grant AI-124319) and Innovation (Grant AI-111661) (to M.R.Y.); the Systems and Integrative Biology Training Program (Grant T32GM008185), Medical Scientist Training Program (Grant T32GM008042), Dermatology Scientist Training Program (Grant T32AR071307) at University of California, Los Angeles, and an Early Career Research Grant from the National Psoriasis Foundation (to E.Y.L.); and National Science Foundation Grant DMR1808459 and NIH Grant 1R56AI125429-01A1 (to M.W.L. and G.C.L.W.).

1. Zasloff M (2002) Antimicrobial peptides of multicellular organisms. *Nature* 415: 389–395.
2. Yeaman MR, Yount NY (2003) Mechanisms of antimicrobial peptide action and resistance. *Pharmacol Rev* 55:27–55.
3. Yount NY, Yeaman MR (2004) Multidimensional signatures in antimicrobial peptides. *Proc Natl Acad Sci USA* 101:7363–7368.
4. Yeaman MR, Yount NY (2007) Unifying themes in host defence effector polypeptides. *Nat Rev Microbiol* 5:727–740.
5. Patil A, Hughes AL, Zhang G (2004) Rapid evolution and diversification of mammalian alpha-defensins as revealed by comparative analysis of rodent and primate genes. *Physiol Genomics* 20:1–11.
6. Tennessen JA, Blouin MS (2007) Selection for antimicrobial peptide diversity in frogs leads to gene duplication and low allelic variation. *J Mol Evol* 65:605–615.
7. Zelezetsky I, et al. (2006) Evolution of the primate cathelicidin. Correlation between structural variations and antimicrobial activity. *J Biol Chem* 281:19861–19871.
8. Tennessen JA (2005) Molecular evolution of animal antimicrobial peptides: Widespread moderate positive selection. *J Evol Biol* 18:1387–1394.
9. Jenssen H, et al. (2007) Evaluating different descriptors for model design of antimicrobial peptides with enhanced activity toward *P. aeruginosa*. *Chem Biol Drug Des* 70: 134–142.
10. Fjell CD, et al. (2009) Identification of novel antibacterial peptides by chemoinformatics and machine learning. *J Med Chem* 52:2006–2015.
11. Toropova MA, Veselinović AM, Veselinović JB, Stojanović DB, Toropov AA (2015) QSAR modeling of the antimicrobial activity of peptides as a mathematical function of a sequence of amino acids. *Comput Biol Chem* 59:126–130.
12. Cherkasov A, et al. (2009) Use of artificial intelligence in the design of small peptide antibiotics effective against a broad spectrum of highly antibiotic-resistant superbugs. *ACS Chem Biol* 4:65–74.
13. Lee EY, Fulan BM, Wong GCL, Ferguson AL (2016) Mapping membrane activity in undiscovered peptide sequence space using machine learning. *Proc Natl Acad Sci USA* 113:13588–13593.
14. Lee EY, Lee MW, Fulan BM, Ferguson AL, Wong GCL (2017) What can machine learning do for antimicrobial peptides, and what can antimicrobial peptides do for machine learning? *Interface Focus* 7:20160153.
15. Lee EY, Wong GCL, Ferguson AL (2018) Machine learning-enabled discovery and design of membrane-active peptides. *Bioorg Med Chem* 26:2708–2718.
16. Yount NY, et al. (2011) Context mediates antimicrobial efficacy of kinocidin congener peptide RP-1. *PLoS One* 6:e26727.
17. Bourbigot S, et al. (2009) Antimicrobial peptide RP-1 structure and interactions with anionic versus zwitterionic micelles. *Biopolymers* 91:1–13.
18. Bourbigot S, Fardy L, Waring AJ, Yeaman MR, Booth V (2009) Structure of chemokine-derived antimicrobial peptide interleukin-8alpha and interaction with detergent micelles and oriented lipid bilayers. *Biochemistry* 48:10509–10521.
19. Yount NY, et al. (2009) Selective reciprocity in antimicrobial activity versus cytotoxicity of hBD-2 and crotamine. *Proc Natl Acad Sci USA* 106:14972–14977.
20. Sharadadevi A, Sivakamasundari C, Nagaraj R (2005) Amphipathic alpha-helices in proteins: Results from analysis of protein structures. *Proteins* 59:791–801.
21. Schmidt NW, et al. (2011) Criterion for amino acid composition of defensins and antimicrobial peptides based on geometry of membrane destabilization. *J Am Chem Soc* 133:6720–6727.
22. Schmidt NW, et al. (2012) Molecular basis for nanoscopic membrane curvature generation from quantum mechanical models and synthetic transporter sequences. *J Am Chem Soc* 134:19207–19216.
23. Yount NY, et al. (2007) Structural correlates of antimicrobial efficacy in IL-8 and related human kinocidins. *Biochim Biophys Acta* 1768:598–608.
24. Matsuzaki K (2009) Control of cell selectivity of antimicrobial peptides. *Biochim Biophys Acta* 1788:1687–1692.
25. Matsuzaki K, Sugishita K, Harada M, Fujii N, Miyajima K (1997) Interactions of an antimicrobial peptide, magainin 2, with outer and inner membranes of Gram-negative bacteria. *Biochim Biophys Acta* 1327:119–130.
26. Yount NY, Yeaman MR (2012) Emerging themes and therapeutic prospects for anti-infective peptides. *Annu Rev Pharmacol Toxicol* 52:337–360.
27. Singh J, et al. (2016) Enhanced cationic charge is a key factor in promoting staphylocidal activity of α-melanocyte stimulating hormone via selective lipid affinity. *Sci Rep* 6:31492.
28. Yeaman MR, Büttner S, Thevissen K (2018) Regulated cell death as a therapeutic target for novel antifungal peptides and biologics. *Oxid Med Cell Longev* 2018: 5473817.
29. Schmidt NW, Wong GC (2013) Antimicrobial peptides and induced membrane curvature: Geometry, coordination chemistry, and molecular engineering. *Curr Opin Solid State Mater Sci* 17:151–163.
30. Dathe M, Wieprecht T (1999) Structural features of helical antimicrobial peptides: Their potential to modulate activity on model membranes and biological cells. *Biochim Biophys Acta* 1462:71–87.
31. Dathe M, et al. (1997) Hydrophobicity, hydrophobic moment and angle subtended by charged residues modulate antibacterial and haemolytic activity of amphipathic helical peptides. *FEBS Lett* 403:208–212.
32. Mishra A, et al. (2011) Translocation of HIV TAT peptide and analogues induced by multiplexed membrane and cytoskeletal interactions. *Proc Natl Acad Sci USA* 108: 16883–16888.
33. Wu Z, Cui Q, Yethiraj A (2013) Why do arginine and lysine organize lipids differently? Insights from coarse-grained and atomistic simulations. *J Phys Chem B* 117: 12145–12156.
34. Chakraborty S, et al. (2014) Ternary nylon-3 copolymers as host-defense peptide mimics: Beyond hydrophobic and cationic subunits. *J Am Chem Soc* 136:14530–14535.
35. Mirabello C, Pollastri G (2013) Porter, PaleAle 4.0: High-accuracy prediction of protein secondary structure and relative solvent accessibility. *Bioinformatics* 29:2056–2058.
36. Drozdetskiy A, Cole C, Procter J, Barton GJ (2015) JPred4: A protein secondary structure prediction server. *Nucleic Acids Res* 43:W389–W394.
37. Xie S, Li Z, Hu H (2018) Protein secondary structure prediction based on the fuzzy support vector machine with the hyperplane optimization. *Gene* 642:74–83.
38. Rončević T, et al. (2018) Antibacterial activity affected by the conformational flexibility in glycine-lysine based alpha-helical antimicrobial peptides. *J Med Chem* 61: 2924–2936.
39. Hwang PM, Vogel HJ (1998) Structure-function relationships of antimicrobial peptides. *Biochem Cell Biol* 76:235–246.
40. Nguyen LT, Haney EF, Vogel HJ (2011) The expanding scope of antimicrobial peptide structures and their modes of action. *Trends Biotechnol* 29:464–472.
41. Meller S, et al. (2015) T_H17 cells promote microbial killing and innate immune sensing of DNA via interleukin 26. *Nat Immunol* 16:970–979.
42. Kaplan A, et al. (2017) Direct antimicrobial activity of IFN-β. *J Immunol* 198: 4036–4045.
43. Yeaman MR, Gank KD, Bayer AS, Brass EP (2002) Synthetic peptides that exert antimicrobial activities in whole blood and blood-derived matrices. *Antimicrob Agents Chemother* 46:3883–3891.
44. Yeaman MR, Bayer AS, Koo SP, Foss W, Sullam PM (1998) Platelet microbicidal proteins and neutrophil defensin disrupt the *Staphylococcus aureus* cytoplasmic membrane by distinct mechanisms of action. *J Clin Invest* 101:178–187.
45. Angelini TE, Liang H, Wriggers W, Wong GC (2003) Like-charge attraction between polyelectrolytes induced by counterion charge density waves. *Proc Natl Acad Sci USA* 100:8634–8637.
46. Schmidt NW, et al. (2015) Liquid-crystalline ordering of antimicrobial peptide-DNA complexes controls TLR9 activation. *Nat Mater* 14:696–700.
47. Lee EY, et al. (2019) Helical antimicrobial peptides assemble into protofibril scaffolds that present ordered dsDNA to TLR9. *Nat Commun* 10:1012.
48. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis, Version 6.0. *Mol Biol Evol* 30:2725–2729.
49. Wimley WC, White SH (1996) Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat Struct Biol* 3:842–848.
50. Eisenberg D (1984) Three-dimensional structure of membrane and surface proteins. *Annu Rev Biochem* 53:595–623.
51. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157:105–132.
52. Eisenberg D, Weiss RM, Terwilliger TC (1982) The helical hydrophobic moment: A measure of the amphiphilicity of a helix. *Nature* 299:371–374.
53. Fauchere JL, Pliska V (1983) Hydrophobic parameters of pi amino-acid side chains from the partitioning of *N*-acetyl-amino-acid amides. *Eur J Med Chem Chim Ther* 18: 369–375.
54. Wang G, Li X, Wang Z (2016) APD3: The antimicrobial peptide database as a tool for research and education. *Nucleic Acids Res* 44:D1087–D1093.
55. Bi J, Bennett K, Embrechts M, Breneman C, Song M (2003) Dimensionality reduction via sparse support vector machines. *J Mach Learn Res* 3:1229–1243.
56. Cao DS, Xu QS, Liang YZ (2013) propy: A tool to generate various modes of Chou's PseAAC. *Bioinformatics* 29:960–962.
57. Yeaman MR, et al. (2007) Modular determinants of antimicrobial activity in platelet factor-4 family kinocidins. *Biochim Biophys Acta* 1768:609–619.
58. Chaili S, et al. (2015) The GraS sensor in *Staphylococcus aureus* mediates resistance to host defense peptides differing in mechanisms of action. *Infect Immun* 84:459–466.
59. Yount NY, Yeaman MR (2006) Structural congruence among membrane-active host defense polypeptides of diverse phylogeny. *Biochim Biophys Acta* 1758:1373–1386.
60. Gank KD, et al. (2008) *SSD1* is integral to host defense peptide resistance in *Candida albicans*. *Eukaryot Cell* 7:1318–1327.