From MRSA persisters to the full ESKAPE panel

Interpretable, precision-first machine learning for *de novo* antimicrobial-peptide design

ESKAPE-AMP: Nasopoulou E[1,†], Felix L[2,†], Shehadeh F[1,2], Mishra B[2], Sotiriadis PP[1], Mylonakis E[2]
CAMPER: Shehadeh F[†], Mishra B[†], Ferrer-Espada R, Basu A, Felix L, Dellis C, Ganesan N, Zhang L, Martens AT, Goulev Y, Sherman MB, Paulsson J, Naik MT, Sotiriadis PP, Mylonakis E
[1]National Technical University of Athens, Greece · [2]Houston Methodist Research Institute, Houston, TX, USA · [†]Equal contribution

CAMPER (Shehadeh *et al.*, *Nat. Commun.* 2026) proved that an interpretable ML + biophysical-scoring model designs potent membrane-targeting peptides against *S. aureus* persisters/biofilms.
ESKAPE-AMP scales the same principle to *all six* ESKAPE pathogens with Gram-specific models, SHAP interpretability, and wet-lab validation.

The problem

ESKAPE pathogens drive multidrug-resistant hospital infection, and the antibiotic pipeline against them is shrinking.

1.27M	**~10M**
deaths, 2019	per yr, 2050

Antimicrobial peptides (AMPs) disrupt bacterial *membranes* directly ⇒ refractory to single-residue resistance, and cheap to synthesise (SPPS).

But ML predictors have three gaps:

1. black-box models lack *interpretability*;
2. they optimise balanced accuracy, not *top-k precision*, the metric that matters when synthesis is the bottleneck;
3. *Gram-specific* selectivity is rarely modelled directly.

Foundation: CAMPER , *Nat. Commun. 2026*

A **mechanistic-AI** framework that pairs a Random-Forest classifier with an explicit **biophysical scoring function** over four membrane properties (**charge, helicity, hydrophobicity, amphipathicity**), combined by *cooperative* rules:

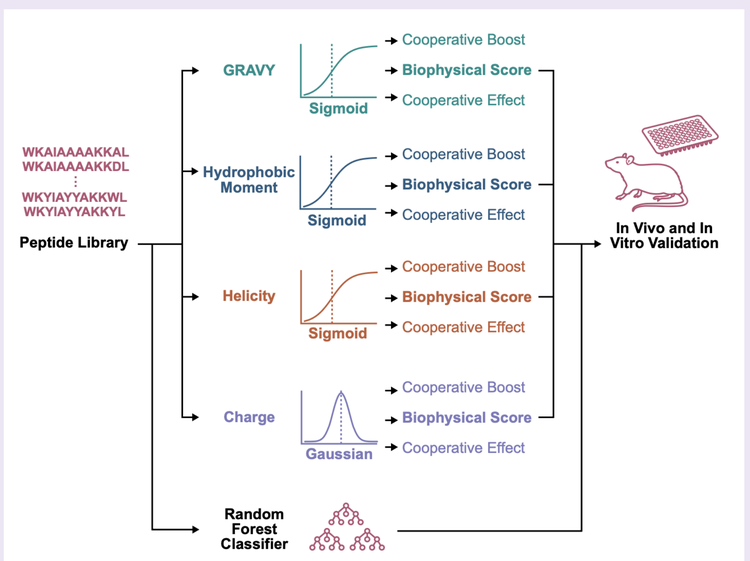
$$\text{coop} = \min(\text{charge, helicity, GRAVY, } \mu_H) \quad \textit{weakest link}$$
$$\text{boost} = \beta \cdot \mathbb{1}(\text{all scores} > \tau) \quad \textit{synergy}$$

Adding the biophysical layer:

- precision **81 %** → **86 %**
- false positives **−77 %** (92 → 21)

Result: WP-CAMPER1 (`WKLIALLAKKLL`), from a **160,000-peptide** mastoparan library:

- **MIC 4 µg/mL** against *S. aureus*
- protease-stable D-enantiomer clears **persisters & biofilms** *in vivo*
- **no resistance** across 21 serial passages



CAMPER pipeline: four biophysical scorers + a Random-Forest classifier.

ESKAPE-AMP pipeline

This **interpretable, calibrated, synthesis-bound** pipeline spans the full ESKAPE panel, with MIC data labelled per Gram class and a direct *biophysical reading* favoured over black-box embeddings.

1. **DBAASP v3 MIC data:** Gram-split, active/inactive at MIC $\leq 32\,\mu g/mL$ (2,165 GP, 1,686 GN).
2. **3 models × 10 feature sets** (30 combinations), homology-aware CV.
3. **Random Forest + physicochemical fingerprint** as the chosen model.
4. **Platt-calibrated probabilities** feed the *de novo* virtual screen.

Precision-first performance

	Gram-positive	Gram-negative
Test AUROC	0.767	0.827
Test AUPRC	0.798	0.836
P@10%	**90.2%**	**90.6%**
Base rate	57.5%	52.6%

- Top-10% precision reaches **~90 %** (~1.6–1.7× over base rate), the regime that matters when synthesis capacity is scarce.
- **Broad-spectrum** set (263 peptides): precision **0.781**, recall **0.258**.

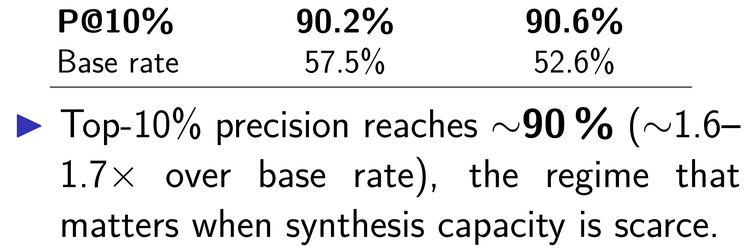
SHAP: the model rediscovers biology

With *no* mechanism engineering, SHAP rediscovers the **canonical Gram-specific AMP mechanism**:

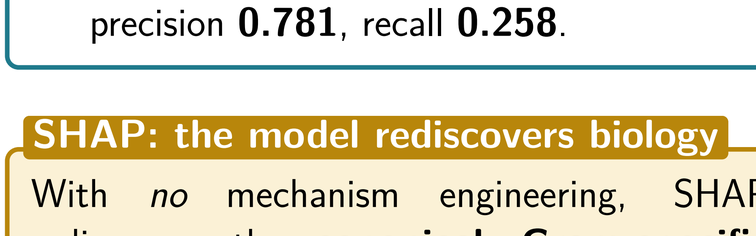
- **Hydrophobicity** drives Gram-positive predictions.
- **Charge & helical amphipathicity** drive Gram-negative ones.

The *same* properties CAMPER scores by hand.

Gram-positive *hydrophobicity-driven*



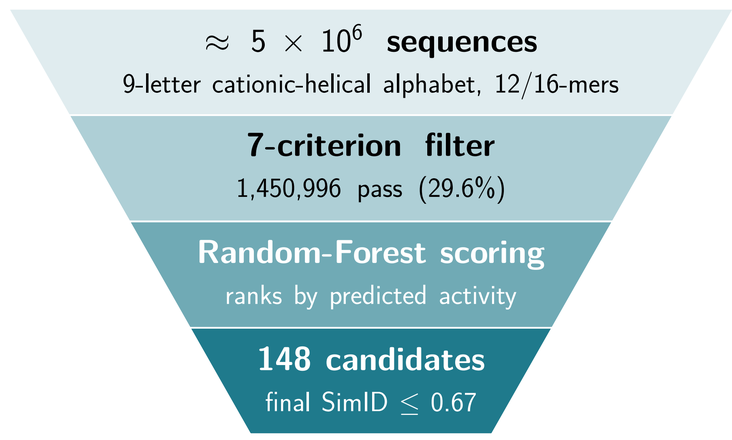
Gram-negative *charge & amphipathicity*



Top-5 features by mean |SHAP| on the held-out test set.

De novo virtual screen

The calibrated models screen a large *de novo* library down to synthesis-ready leads:



The 148 candidates split into three design groups, **Gram-positive-biased**, **Gram-negative-biased**, and **broad-spectrum**. Top-5 per group (**15 total**) proceeded to wet-lab validation.

Wet-lab validation

The **15 synthesised peptides** were assayed by broth microdilution against **all six ESKAPE species**.

ID	*Ec*	*Sa*	*Kp*	*Ab*	*Pa*	*Ef*	HL50
GP-4	>32	4	>32	16	>32	8	16[†]
GP-5	>32	8	>32	32	>32	8	32
GN-1	16	32	>32	8	>32	32	>32
GN-3	16	8	8	8	>32	8	>32
GN-4	32	>32	>32	8	>32	>32	>32
GN-5	32	16	16	4	>32	16	>32
BR-2	32	16	16	32	>32	8	>32
BR-3	16	8	16	16	>32	16	>32
BR-5	16	32	16	16	>32	16	>32

MIC (µg/mL): ≤8 16 32 >32 . **Bold** = "golden four".
[†]No therapeutic window. (6 moderate/inactive candidates omitted.)

- **Hit rate:** 13 / 15 candidates active vs ≥1 species; 11 / 15 reached MIC ≤ 16.
- **Golden four** (GN-3, GN-5, BR-3, BR-5): ≥4 pathogens at MIC < 16, no haemolysis.
- **Caveat:** all 15 candidates are inactive against *P. aeruginosa*, likely due to its intrinsic outer-membrane barriers.

Conclusions

- Interpretable, **precision-first** ML delivers *lab-confirmed* ESKAPE leads, validating end-to-end the principle CAMPER established for *S. aureus*.
- SHAP gives every decision a **mechanistic, regulator-friendly reading**.

CAMPER: Shehadeh *et al.*, *Nat. Commun.* 2026;17:3689.

Presented at **Greeks in AI 2026**

Funding: NIH grants P01AI083214 and P20GM121344 (NIGMS COBRE Center for AMR & Therapeutic Discovery) to E. Mylonakis.
F. Shehadeh (PhD student) and E. Nasopoulou (Diploma thesis student) of P.P. Sotiriadis.