doi:10.2533/chimia.2024.648 Chimia 78 (2024) 648–653

Chemical Space for Peptide-based Antimicrobials

Markus Orsi, Hippolyte Personne, Etienne Bonvin, Thierry Paschoud, Basak Olcay, Xiaoling Hu, Sacha Javor, and Jean-Louis Reymond*

Abstract: Multidrug-resistant (MDR) bacteria represent a global public health threat, and antimicrobial peptides (AMPs), derived from naturally occurring linear or cyclic peptides, can provide the solution. However, most AMPs are sensitive to proteases and have poor pharmacokinetics. The EU-funded ERC Advanced Grant SPACE4AMPS aims to identify new AMPs by applying the concepts of chemical space and ligand-based virtual screening, which are well known for small molecule drug discovery, to the world of peptides. We create virtual libraries of peptides and related molecules and use these approaches to select a few tens of compounds for synthesis and detailed evaluation of antibacterial, toxicity and stability effects. Recent results and prospects of this ongoing project are presented in this review.

Keywords: Antimicrobial peptides · Antimicrobial resistance · Chemical space · Peptoids · Virtual screening

1. Introduction

Antimicrobial resistance (AMR) refers to the increasing occurrence of multidrug resistant (MDR) bacteria that cannot be treated with available antibiotics, which is most frequently encountered with ESKAPE pathogens (*Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa* and *Enterobacter spp*.).[1] One key problem of new antibiotics is that they are not economically profitable, in part because their use should be restricted to avoid the emergence of more resistance, such that their development has been abandoned by large pharmaceutical companies.[2]

Among new treatment options, antimicrobial peptides (AMPs) represent a promising opportunity because resistance does not easily arise against such compounds.[3] Most AMPs are polycationic and partly hydrophobic sequences of approximately 10–40 residues in length that kill bacteria by destabilizing or permeabilizing their membranes or by interacting with intracellular components such as the ribosome.[4] AMPs comprise linear peptides expressed as part of the innate immune system of most organisms, such as the human peptide LL-37,[5] as well as a diversity of cyclic peptides of microbial origin assembled by non-ribosomal peptide synthetases and containing non-proteinogenic amino acids, typically D-enantiomeric residues, for example the last resort antibiotic polymyxin B.[6] Furthermore, many AMP-like synthetic analogs ranging from short peptides and peptoids to large polymers are being investigated as new antimicrobials.[7]

Our first project in AMP discovery was designed as an application example for a combinatorial library decoding scheme applicable to cyclic peptides,[8] which we used to discover hybrids of the cyclic peptide antibiotics gramicidin S and tyrocidine A using an agar-plate antibacterial activity assay adapted for one-bead-one-compound libraries.[9] We later used this assay to test another combinatorial library consisting of peptide dendrimers,[10] where antimicrobial effects should be absent due to the lack of an organized amphiphilic structure. Contrary to expectations, these peptide dendrimers showed good antimicrobial effects and were later optimized by sequence adjustments.[11–13] Collaborations with clinical microbiologists, including Andrea Endimiani at the Institute of Infectious Diseases in Bern,[14] and Thilo Köhler and Christian van Delden at the Faculty of Medicine in Geneva,[15,16] were essential to establish that our peptide dendrimers demonstrated good activity against various MDR clinical isolates.

To identify truly interesting AMPs, we had to measure antimicrobial effects on multiple bacterial species and strains and keep cellular toxicity in check, also monitoring the influence of other key parameters such as the pH of the culture medium.[17] This meant that single-parameter high-throughput screening, as we initially performed with combinatorial libraries to discover active sequences against a background of inactive ones, was not suitable. This led to the idea of adapting small molecule cheminformatics and the notion of chemical space, an area of intense activity in our research group,[18–20] to the world of peptides. These computational tools would help us assemble virtual libraries and evaluate them *in silico* to select small sets of AMPs for detailed experimental evaluation. This avenue led to the program currently supported by the ERC Advanced Grant SPAC4AMPS and is reviewed in this article.

2. Chemical Space, Ligand-based Virtual Screening, and Visualization

A chemical space is a mathematical space occupied by molecules in which distances represent similarities between molecules. To identify interesting molecules for synthesis and evaluation, one constructs a virtual library of possible molecules of interest, which should be designed to be accessible by a selected synthetic approach. One then places this virtual library in a chemical space together with other known molecules of interest such as already known bioactive compounds possessing desirable or non-desirable profiles. One can then select compounds for synthesis and evaluation according to their similarity or diversity within the library and in relation to the known molecules. This simple procedure, known as ligand-based virtual screening (LBVS), is commonly used in small molecule drug discovery to improve the return on experimental testing, typically increasing the hit rate of approximately 0.1% for random screening up to hit rates sometimes as high as 50% for sets of compounds pre-selected by LBVS.[21]

Similarity between molecules can be quantified by comparing molecular properties, which are then used as dimensions for

*Correspondence: Prof. Dr. J.-L. Reymond, E-mail: jean-louis.reymond@unibe.ch
Department of Chemistry, Biochemistry and Pharmaceutical Sciences, University of Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

the chemical space. For example, molecular weight and polarity define a 2-dimensional chemical space in which a specific area defines oral bioavailability for small molecule drugs.[22,23] In a related approach, we have used simple descriptors of molecular structure called MQN (Molecular Quantum Numbers), including counts for different atom types, bond types, polar groups and ring structures, to construct a 42-dimensional chemical space suitable for global classification of chemical structures.[24–26] Although such simple chemical spaces may be used for LBVS,[27] more detailed similarity comparisons are in most cases necessary, typically using extended connectivity fingerprints,[28,29] or by comparing molecular shape and pharmacophores either *via* 3D-models,[30,31] or *via* molecular fingerprints computed from the structural formula.[32–34]

In our first implementation of cheminformatics to discover AMPs, we created small virtual libraries of bicyclic peptides consisting of lysines and leucines as functional residues and a pair of cysteines to close the rings by thioether ligation to *N*-terminal electrophiles. We used a chemical space computed from a pharmacophore fingerprint to sample these virtual libraries for sequences featuring diverse distributions of lysines and leucines, a parameter expected to be key for activity, and later to refine initial hits by identifying nearest neighbors.[35,36] This project led to the discovery of several interesting bicyclic peptides, for example **bp65**, a short bicyclic AMP with good activity against various Gram-negative bacteria. This bicyclic peptide featured a membrane disruptive amphiphilic α-helical structure as evidenced by circular dichroism and X-ray crystallography (Fig. 1a).[37] In a parallel project, we used the same approach to optimize the amino acid sequence of an AMP dendrimer, similarly sampling various distributions of lysines and leucines as functional residues across dendrimer branches.[38,39]

Visualization by dimensionality reduction usually helps to understand a chemical space. For simple descriptor spaces such as MQN, principal component analysis (PCA) can be used since the first two principal components usually cover over 70% of the data variability. This method was also suitable in the above mentioned AMP projects to visualize the distribution of our bicyclic peptides and peptide dendrimers in their chemical space, as exemplified for the similarity map representing a virtual library of 4.6 million bicyclic peptides from which **bp65** was identified (Fig. 1b).[35,36,38] For higher dimensional spaces such as those corresponding to substructure fingerprints, non-linear dimensionality reduction methods such as similarity mapping,[40,41] t-SNE (t-distributed stochastic neighbor embedding),[42] UMAP (uniform manifold approximation and projection),[43] or TMAP (treemap),[44] are required to obtain interpretable representations. For example, we used similarity mapping to produce an interactive 3D-map of the chemical space of large molecules such as peptides and macrocycles in the ChEMBL database reflecting simi-

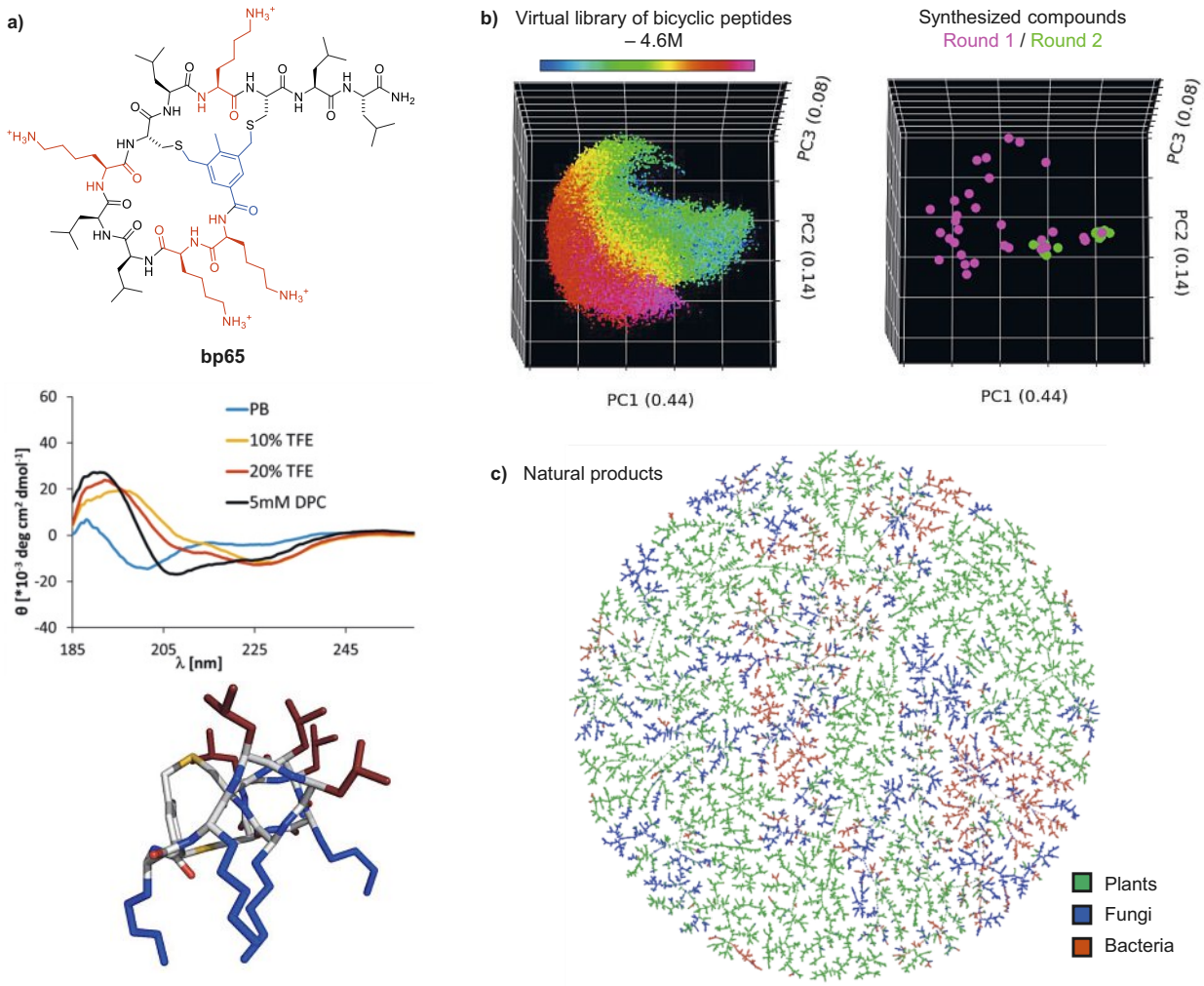

Fig. 1. Chemical space guided exploration of peptides and natural products. (**a**) Structure, CD spectrum and X-ray crystal structure of the bicyclic AMP **bp65** showing its α-helical amphiphilic conformation. (**b**) PCA (Principal Component Analysis) map of the pharmacophore chemical space representing the virtual library of 4.6 million bicyclic peptides from which **bp65** was identified, color-coded by the number of cyclic residues from 2 (blue) to 18 (magenta), and focused visualization of the 1st and 2nd round of compounds that were synthesized and tested in this library. (**c**) A TMAP of natural products from the COCONUT database computed by MAP4 similarity and color-coded by the origin of the natural products.

larities computed with a shape and pharmacophore fingerprint.[45] In a related project, we produced a map of natural products using the TMAP algorithm computed from similarities calculated with MAP4, a molecular fingerprint performing particularly well for virtual screening (Fig.1c).[46–48] One can also visualize similarities between pairs of drug molecules according to different similarity measures simultaneously in form of a TMAP by using methods of reaction informatics.[49]

3. Peptide Diastereomers and Peptide-peptoid Hybrids

The studies mentioned above provided convincing evidence that a chemical space analysis based on simple molecular fingerprints could guide the discovery and understanding of AMPs and natural products and comforted us to plan a broader discovery campaign along these lines. In one of the SPACE4AMPS projects, we are exploring the chemical space of peptide diastereomers. In an initial study, we had found that solid-phase peptide synthesis using racemic amino acids provided stereorandomized (*sr*-) peptides in excellent yields and purities, which were indistinguishable for the homochiral sequence except in terms of optical and biological properties, providing insights into their mechanism of action.[50] For example, stereorandomization allowed us to show that AMP dendrimers have an intrinsically disordered bioactive conformation.[39] On the other hand, we found that typical amphiphilic α-helical AMPs became inactive when stereorandomized.

Surprisingly, we found that **ln65**, the linear AMP analog of **bp65**, which is α-helical and amphiphilic, retained part of its antibacterial activity but showed strongly reduced hemolysis as the stereorandomized sequence *sr*-**ln65**. Therefore we set out to investigate single diastereomers of this sequence, an approach which has been investigated in part for other AMPs to tune their properties, however not in a systematic manner.[51] We found that α-helicity and antibacterial activity are maintained and hemolysis reduced in selected diastereomers of **ln65** with up to 40% of the residues with mixed chirality, such as **ln69**, **dln69**, **HP5** and **HP7**, while serum stability is very often improved, opening an obvious avenue for optimizing the activity/toxicity ratio and stability (Table 1).[52] We are currently collecting additional data to understand the stereochemistry-property profile of AMP **ln65**, as well as of other related natural AMPs. To support these studies, we have recently reported MAP4C, a chiral version of our MAP4 molecular fingerprint which encodes stereochemistry and allows the analysis of the chemical space of peptide diastereomers.[53] Interestingly, none of the existing models of AMP folding and activity considers residue chirality as a variable and reliable data are almost entirely missing, rendering the problem both challenging and groundbreaking.

In a second project we are investigating the replacement of amino acids by their peptoid equivalents to form peptide-peptoid hybrids. In a peptoid residue, the side chain is attached to the α-amino group and the α-carbon is free, which suppresses both chirality and the hydrogen bond donor of the peptide bond. Peptoids follow different folding rules compared to peptides and may also exhibit antibacterial properties.[54,55] Introducing one or several peptoid units in an AMP is expected to alter its folding and potentially its mechanism of action. In a preliminary study dedicated to linear and dendritic peptoids, we have discovered hybrids such as **EB1**, **EB5** or **EB6** which retain the membrane

Table 1. Helicity, antimicrobial activity and hemolysis of **bp65** and its linear analogs.

Cpd.	Sequence[a]	α-helix content	*E. coli* W3110	*P. aeruginosa* PAO1	*A. baumannii* ATCC19606	*K. pneumoniae* NCTC418	*S. aureus* COL MRSA	MHC[d]
		(%)[b]	MIC (μg/mL)[c]					(μg/mL)
bp65	*****KKLLK**C**LK**C**LL	46	n.d.	8	4	16	8	16.5
ln65	KKLLKLLKLLL	73	4	2-4	2-4	4	4	125
***sr*-ln65**	KKLLKLLKLLL	10	4	4	4	16-32	8	1000
ln69	**kk**LL**k**LL**k**LLL	61	4	8	2-4	8	16	1000
dln69	KK**ll**K**ll**K**lll**	59	0.5-1	2-4	2	4	2	250
HP1	K**k**LLKLLKLLL	73	2	4	4	4-8	2	< 15.6
HP2	**kk**LLKLLKLLL	69	4	4	4	2-4	4	< 15.6
HP3	K**k**LL**k**LLKLLL	69	2	4	2	8	2-4	< 15.6
HP4	K**k**L**l**KLLKLLL	46	2-4	2-4	2-4	2	2-4	< 15.6
HP5	**k**KLLKLLKLL**l**	90	0.5	2	0.5	2	2	62.5
HP6	KKLLK**ll**KLLL	29	2	8	4	64	16	500
HP7	**kk**LLKLLKLL**l**	60	0.5	2	0.5	4	2	125
EB1	***k***KLLKLLKLL***l***	30	2	4	2	8	8	1000
EB5	***kkll k***LLKLLL	24	2	4	2	>32	8	1000
EB6	***kk***LL***k***LLKLLL	41	2	4	2	8	8	250
EB9[e]	K***k***L***l***K***l***L***k***L***l***L	11	2	4	16	>32	32	>2000

[a]One letter code for amino acids, L-amino acids in capital letters, D-amino acid in lowercase boldface, stereo randomized positions are underlined, and peptoid residues in lowercase italics boldface. For **bp65**, * is the 3,5-dimethyltolyl group thioether ligated to the cysteine residues, see Fig. 1. [b]Values are corresponding to data recorded by circular dichroism for the condition 5mM DPC in 7 mM PB buffer pH 7.4. [c]Minimum Inhibitory Concentrations (MIC) were determined after incubation in Mueller-Hinton (MH) broth pH 7.4 for 16-20 h at 37 °C. Values represent two independent duplicates MIC determinations by 2-fold serial dilution. n.d. = not determined. [d]Minimum Hemolytic Concentration (MHC) measured on human red blood cells in PBS (pH 7.4) after 4 h incubation at room temperature, by serial 2-fold serial dilution. **ln65** and related peptides have a molecular weight of MW = 1321 Da, such that for this series μg/mL, which is the standard unit for MIC or MHC values in microbiology, corresponds approximately to μM. [e]MIC values for **EB9** were measured in diluted (12.5 %) MH medium.

disruptive and partially the folding character of the parent peptide. This study also uncovered hybrids such as **EB9** which do not act by membrane disruption but rather kill bacteria by entering the cells and aggregating their intracellular content (Table 1).[56–58] We are currently investigating macrocyclic peptide-peptoid sequences designed as analogs of antibacterial natural products using a genetic algorithm[59] which exhibit promising antibacterial effects and no detectable toxicity on human cells.

4. Machine Learning

In the AMP discovery projects discussed above, we used a pharmacophore chemical space to analyze focused virtual libraries of up to several million possible bicyclic or dendritic peptides and selects a few tens of sequences for synthesis and evaluation. The only prior information taken from known AMPs was the choice of lysine and leucine as functional residues and the approximate overall size expected for an active AMP. However, a much larger and more detailed source of information is available in the form of the sequence and activity of thousands of documented AMPs, which have been collected in various databases. Such information can be used to train neural networks, which can then generate and/or classify peptides as potential AMPs.[60,61]

We have tested the feasibility of this machine learning (ML) approach to discover non-hemolytic antimicrobial peptides,[62] as well as non-hemolytic anticancer peptides,[63] exploiting activity and hemolysis data from the database of antimicrobial activity and structure of peptides (DBAASP).[64] While both projects delivered the expected active and non-hemolytic sequences, they also revealed the limit of this approach, namely the difficulty to identify sequences that are significantly different from the database sequences used for training, and more surprisingly the inability of current ML models, including alphafold2, to correctly predict the degree of α-helicity of short peptides. This is disappointing because the degree of α-helicity of short amphiphilic peptides, as measured experimentally by circular dichroism (CD), correlates nicely with their membrane disruptive activity in our various projects.[52,62,63] Most recently we also evaluated the large language model GPT-3.5 for its ability to predict AMP activity and toxicity, as measured by hemolysis. Our study showed that this model does not perform as well as a simple chemical space classification using our MAP4C fingerprint (Fig. 2).[65] Currently, we are exploiting ML models to support the optimization of several AMP series, training our models with our own experimental data. This approach follows a general trend in chemistry to use ML on

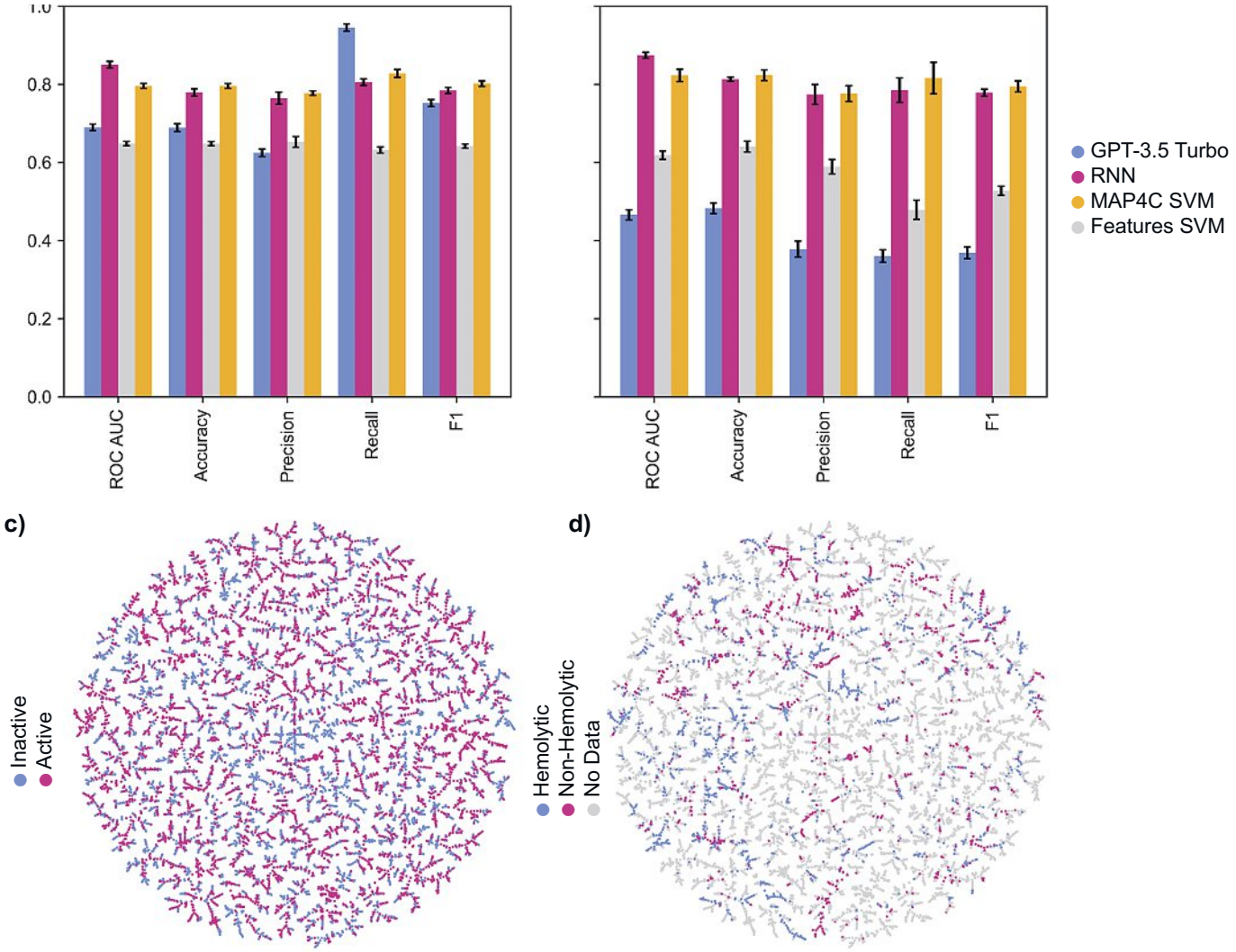

Fig. 2. Classification and chemical space coverage of AMPs from the DBAASP database. Results of the 5-fold cross-validation study aimed at validating the following models: MAP4C SVM: support vector machine trained with the molecular fingerprint MAP4C, Features SVM: support vector machine trained with the fraction of helical residues and hydrophobic moment of the peptides, RNN and GPT-3.5 turbo: recurrent neural network and large language model trained with the linear sequence of the peptides written as single letters. The performance measures, given in fraction (maximum = 1), are measured for the predictions of (**a**) antimicrobial activity and (**b**) hemolysis. An AUC of 0.5 indicates a model performing no better than random chance, whereas the higher the value of each performance metric, the better the model performance. TMAP computed using MAP4C color-coded by antimicrobial activity (**c**) or hemolysis (**d**) data.

relatively small but well-defined sets of compounds and their associated data.[66]

5. Conclusion and Outlook

Our current SPACE4AMPS project exploits the concept of chemical space as well as machine learning to guide the discovery of new antimicrobial peptides and peptoids inspired from natural products. These computational tools are essential to properly exploit the available structural diversity of peptide-like compounds, which is often neglected, such as the chemical space of diastereomers, sequence isomers, and peptide-peptoid hybrids under investigation here. The project's success depends on our ability to synthesize and test the molecules identified in our virtual chemical spaces, which for this project is straightforward using solid-phase peptide synthesis, a robust technology also available for GMP manufacturing due to the current popularity of peptide drugs.[67] Any active AMP can be readily resynthesized, produced in gram scale, and made available to collaborating groups for implementation in projects, several of which are currently ongoing.[68]

Acknowledgements

This work was supported financially by the European Research Council, Grant no. 885076.

Received: May 31, 2024

[1] D. M. P. De Oliveira, B. M. Forde, T. J. Kidd, P. N. A. Harris, M. A. Schembri, S. A. Beatson, D. L. Paterson, M. J. Walker, *Clin. Microbiol. Rev*. **2020**, *33*, 10.1128/cmr.00181, https://doi.org/10.1128/cmr.00181-19.

[2] M. A. Cook, G. D. Wright, *Sci. Transl. Med*. **2022**, *14*, eabo7793, https://doi.org/10.1126/scitranslmed.abo7793.

[3] C. A. Roque-Borda, L. M. D. G. Primo, H. Franzyk, P. R. Hansen, F. R. Pavan, *Heliyon* **2024**, *10*, https://doi.org/10.1016/j.heliyon.2024.e31958.

[4] L. T. Nguyen, E. F. Haney, H. J. Vogel, *Trends Biotechnol*. **2011**, *29*, 464, https://doi.org/10.1016/j.tibtech.2011.05.001.

[5] M. R. Scheenstra, M. van den Belt, J. L. M. T. Bokhoven, V. A. F. Schneider, S. R. Ordonez, A. van Dijk, E. J. A. Veldhuizen, H. P. Haagsman, *Sci. Rep*. **2019**, *9*, 1, https://doi.org/10.1038/s41598-019-41246-6.

[6] L. Poirel, A. Jayol, P. Nordmann, *Clin. Microbiol. Rev*. **2017**, *30*, 557, https://doi.org/10.1128/CMR.00064-16.

[7] B. H. Gan, J. Gaynord, S. M. Rowe, T. Deingruber, D. R. Spring, *Chem. Soc. Rev*. **2021**, 50, 7820, https://doi.org/10.1039/D0CS00729C.

[8] J. Kofoed, J. L. Reymond, *J. Comb. Chem*. **2007**, *9*, 1046, https://doi.org/10.1021/cc7001155.

[9] V. S. Fluxa, N. Maillard, M. G. Page, J. L. Reymond, *Chem. Commun*. **2011**, *47*, 1434, https://doi.org/10.1039/C0CC04670A.

[10] N. Maillard, R. Biswas, T. Darbre, J.-L. Reymond, *ACS Comb. Sci*. **2011**, *13*, 310, https://doi.org/10.1021/co200006z.

[11] M. Stach, N. Maillard, R. U. Kadam, D. Kalbermatter, M. Meury, M. G. P. Page, D. Fotiadis, T. Darbre, J.-L. Reymond, *MedChemComm* **2012**, *3*, 86, https://doi.org/10.1039/c1md00272d.

[12] M. Stach, T. N. Siriwardena, T. Kohler, C. van Delden, T. Darbre, J. L. Reymond, *Angew. Chem., Int. Ed*. **2014**, *53*, 12827, https://doi.org/10.1002/anie.201409270.

[13] T. N. Siriwardena, M. Stach, R. He, B.-H. Gan, S. Javor, M. Heitz, L. Ma, X. Cai, P. Chen, D. Wei, H. Li, J. Ma, T. Köhler, C. van Delden, T. Darbre, J.-L. Reymond, *J. Am. Chem. Soc*. **2018**, *140*, 423, https://doi.org/10.1021/jacs.7b11037.

[14] J. Pires, T. N. Siriwardena, M. Stach, R. Tinguely, S. Kasraian, F. Luzzaro, S. L. Leib, T. Darbre, J. L. Reymond, A. Endimiani, *Antimicrob. Agents Chemother*. **2015**, *59*, 7915, https://doi.org/10.1128/aac.01853-15.

[15] B.-H. Gan, T. N. Siriwardena, S. Javor, T. Darbre, J.-L. Reymond, *ACS Infect. Dis.* **2019**, *5*, 2164, https://doi.org/10.1021/acsinfecdis.9b00299.

[16] F. B. Jeddou, L. Falconnet, A. Luscher, T. Siriwardena, J.-L. Reymond, C. van Delden, T. Köhler, *Antimicrob. Agents Chemother*. **2020**, *64*, e02040, https://doi.org/10.1128/AAC.02040-19.

[17] X. Cai, S. Javor, B. H. Gan, T. Köhler, J.-L. Reymond, *Chem. Commun*. **2021**, *57*, 5654, https://doi.org/10.1039/D1CC01838H.

[18] J.-L. Reymond, *Acc. Chem. Res*. **2015**, *48*, 722, https://doi.org/10.1021/ar500432k.

[19] K. Meier, S. Bühlmann, J. Arús-Pous, J.-L. Reymond, *CHIMIA* **2020**, *74*, 241, https://doi.org/10.2533/chimia.2020.241.

[20] J.-L. Reymond, *CHIMIA* **2022**, *76*, 1045, https://doi.org/10.2533/chimia.2022.1045.

[21] J. Vázquez, M. López, E. Gibert, E. Herrero, F. J. Luque, *Molecules* **2020**, *25*, 4723, https://doi.org/10.3390/molecules25204723.

[22] C. Lipinski, A. Hopkins, *Nature* **2004**, *432*, 855, https://doi.org/10.1038/nature03193.

[23] G. B. Santos, A. Ganesan, F. S. Emery, *ChemMedChem* **2016**, *11*, 2245, https://doi.org/10.1002/cmdc.201600288.

[24] K. T. Nguyen, L. C. Blum, R. van Deursen, J.-L. Reymond, *ChemMedChem* **2009**, *4*, 1803, https://doi.org/10.1002/cmdc.200900317.

[25] R. van Deursen, L. C. Blum, J. L. Reymond, *J. Chem. Inf. Model*. **2010**, *50*, 1924, https://doi.org/10.1021/ci100237q.

[26] D. Probst, J.-L. Reymond, *Bioinformatics* **2018**, *34*, 1433, https://doi.org/10.1093/bioinformatics/btx760.

[27] J. J. Burgi, M. Awale, S. D. Boss, T. Schaer, F. Marger, J. M. Viveros-Paredes, S. Bertrand, J. Gertsch, D. Bertrand, J. L. Reymond, *ACS Chem. Neurosci*. **2014**, *5*, 346, https://doi.org/10.1021/cn4002297.

[28] H. L. Morgan, *J. Chem. Doc*. **1965**, *5*, 107, https://doi.org/10.1021/c160017a018.

[29] D. Rogers, M. Hahn, *J. Chem. Inf. Model*. **2010**, *50*, 742, https://doi.org/10.1021/ci100050t.

[30] A. Nicholls, G. B. McGaughey, R. P. Sheridan, A. C. Good, G. Warren, M. Mathieu, S. W. Muchmore, S. P. Brown, J. A. Grant, J. A. Haigh, N. Nevins, A. N. Jain, B. Kelley, *J. Med. Chem*. **2010**, *53*, 3862, https://doi.org/10.1021/jm900818s.

[31] C. Simonin, M. Awale, M. Brand, R. van Deursen, J. Schwartz, M. Fine, G. Kovacs, P. Hafliger, G. Gyimesi, A. Sithampari, R. P. Charles, M. A. Hediger, J. L. Reymond, *Angew. Chem., Int. Ed*. **2015**, *54*, 14748, https://doi.org/10.1002/anie.201507320.

[32] R. E. Carhart, D. H. Smith, R. Venkataraghavan, *J. Chem. Inf. Comput. Sci*. **1985**, *25*, 64, https://doi.org/10.1021/ci00046a002.

[33] G. Schneider, W. Neidhart, T. Giller, G. Schmid, *Angew. Chem. Int. Ed. Engl.* **1999**, *38*, 2894, https://doi.org/10.1002/(SICI)1521-3773(19991004)38:19<2894::AID-ANIE2894>3.0.CO;2-F.

[34] M. Awale, J. L. Reymond, *J. Chem. Inf. Model*. **2014**, *54*, 1892, https://doi.org/10.1021/ci500232g.

[35] I. Di Bonaventura, X. Jin, R. Visini, D. Probst, S. Javor, B. H. Gan, G. Michaud, A. Natalello, S. M. Doglia, T. Kohler, C. van Delden, A. Stocker, T. Darbre, J. L. Reymond, *Chem. Sci*. **2017**, *8*, 6784, https://doi.org/10.1039/c7sc01314k.

[36] I. Di Bonaventura, S. Baeriswyl, A. Capecchi, B.-H. Gan, X. Jin, T. N. Siriwardena, R. He, T. Kohler, A. Pompilio, G. Di Bonaventura, C. van Delden, S. Javor, J.-L. Reymond, *Chem. Commun*. **2018**, *54*, 5130, https://doi.org/10.1039/c8cc02412j.

[37] S. Baeriswyl, H. Personne, I. D. Bonaventura, T. Köhler, C. van Delden, A. Stocker, S. Javor, J.-L. Reymond, *RSC Chem. Biol.* **2021**, *2*, 1608, https://doi.org/10.1039/D1CB00124H.

[38] T. N. Siriwardena, A. Capecchi, B. H. Gan, X. Jin, R. He, D. Wei, L. Ma, T. Kohler, C. van Delden, S. Javor, J. L. Reymond, *Angew. Chem., Int. Ed*. **2018**, *57*, 8483, https://doi.org/10.1002/anie.201802837.

[39] X. Cai, M. Orsi, A. Capecchi, T. Köhler, C. van Delden, S. Javor, J.-L. Reymond, *Cell Rep. Phys. Sci*. **2022**, *3*, https://doi.org/10.1016/j.xcrp.2022.101161.

[40] J. L. Medina-Franco, G. M. Maggiora, M. A. Giulianotti, C. Pinilla, R. A. Houghten, *Chem. Biol. Drug Des.* **2007**, *70*, 393, https://doi.org/10.1111/j.1747-0285.2007.00579.x.

[41] M. Awale, J. L. Reymond, *J. Chem. Inf. Model*. **2015**, *55*, 1509, https://doi.org/10.1021/acs.jcim.5b00182.

[42] L. Maaten, G. Hinton, *J. Mach. Learn. Res*. **2008**, *9*, 2579.

[43] L. McInnes, J. Healy, J. Melville, arXiv:1802.03426 [cs, stat] **2018**.

[44] D. Probst, J.-L. Reymond, *J. Cheminf.* **2020**, *12*, 12, https://doi.org/10.1186/s13321-020-0416-x.

[45] A. Capecchi, M. Awale, D. Probst, J. L. Reymond, *Mol. Inf.* **2019**, *38*, 1900016, https://doi.org/10.1002/minf.201900016.

[46] A. Capecchi, J.-L. Reymond, *Biomolecules* **2020**, *10*, 1385, https://doi.org/10.3390/biom10101385.

[47] A. Capecchi, J.-L. Reymond, *J. Cheminf.* **2021**, *13*, 82, https://doi.org/10.1186/s13321-021-00559-3.

[48] A. Capecchi, D. Probst, J.-L. Reymond, *J. Cheminf.* **2020**, *12*, 43, https://doi.org/10.1186/s13321-020-00445-4.

[49] M. Orsi, D. Probst, P. Schwaller, J.-L. Reymond, *Digital Discovery* **2023**, *2*, 1289, https://doi.org/10.1039/D3DD00039G.

[50] T. N. Siriwardena, B.-H. Gan, T. Köhler, C. van Delden, S. Javor, J.-L. Reymond, *ACS Cent. Sci*. **2021**, *7*, 126, https://doi.org/10.1021/acscentsci.0c01135.

[51] D. Ben Hur, G. Kapach, N. A. Wani, E. Kiper, M. Ashkenazi, G. Smollan, N. Keller, O. Efrati, Y. Shai, *J. Med. Chem*. **2022**, *65*, 9050, https://doi.org/10.1021/acs.jmedchem.2c00270.

[52] H. Personne, T. Paschoud, S. Fulgencio, S. Baeriswyl, T. Köhler, C. van Delden, A. Stocker, S. Javor, J.-L. Reymond, *J. Med. Chem*. **2023**, *66*, 7570, https://doi.org/10.1021/acs.jmedchem.3c00460.

[53] M. Orsi, J.-L. Reymond, *J. Cheminform*. **2024**, *16*, 53, https://doi.org/10.1186/s13321-024-00849-6.

[54] R. N. Zuckermann, *Peptide Sci*. **2011**, *96*, 545, https://doi.org/10.1002/bip.21573.
[55] J. E. Nielsen, M. A. Alford, D. B. Y. Yung, N. Molchanova, J. A. Fortkort, J. S. Lin, G. Diamond, R. E. W. Hancock, H. Jenssen, D. Pletzer, R. Lund, A. E. Barron, *ACS Infect. Dis*. **2022**, *8*, 533, https://doi.org/10.1021/acsinfecdis.1c00536.
[56] E. Bonvin, H. Personne, T. Paschoud, J. Reusser, B.-H. Gan, A. Luscher, T. Köhler, C. van Delden, J.-L. Reymond, *ACS Infect. Dis*. **2023**, *9*, 2593, https://doi.org/10.1021/acsinfecdis.3c00421.
[57] E. Bonvin, J.-L. Reymond, *Helv. Chim. Acta* **2023**, *106*, e202300035, https://doi.org/10.1002/hlca.202300035.
[58] H. Personne, X. Hu, E. Bonvin, J. Reusser, J.-L. Reymond, *Helv. Chim. Acta* **2024**, *107*, e202400041, https://doi.org/10.1002/hlca.202400041.
[59] A. Capecchi, A. Zhang, J.-L. Reymond, *J. Chem. Inf. Model*. **2020**, *60*, 121, https://doi.org/10.1021/acs.jcim.9b01014.
[60] A. Capecchi, J.-L. Reymond, *Med. Drug Discovery* **2021**, *9*, 100081, https://doi.org/10.1016/j.medidd.2021.100081.
[61] M. D. C. Aguilera-Puga, F. Plisson, *Sci. Rep*. **2024**, *14*, 11995, https://doi.org/10.1038/s41598-024-62419-y.
[62] A. Capecchi, X. Cai, H. Personne, T. Köhler, C. van Delden, J.-L. Reymond, *Chem. Sci*. **2021**, *12*, 9221, https://doi.org/10.1039/D1SC01713F.
[63] E. Zakharova, M. Orsi, A. Capecchi, J.-L. Reymond, *ChemMedChem* **2022**, *17*, e202200291, https://doi.org/10.1002/cmdc.202200291.
[64] G. Gogoladze, M. Grigolava, B. Vishnepolsky, M. Chubinidze, P. Duroux, M.-P. Lefranc, M. Pirtskhalava, *FEMS Microbiol. Lett*. **2014**, *357*, 63, https://doi.org/10.1111/1574-6968.12489.
[65] M. Orsi, J.-L. Reymond, *RSC Med. Chem*. **2024**, 15, 2030, https://doi.org/10.1039/D4MD00159A.
[66] M. Orsi, B. Shing Loh, C. Weng, W. H. Ang, A. Frei, *Angew. Chem., Int. Ed*. **2024**, *63*, e202317901, https://doi.org/10.1002/anie.202317901.
[67] M. W. Pennington, B. Zell, C. J. Bai, *Medicine in Drug Discovery* **2021**, *9*, 100071, https://doi.org/10.1016/j.medidd.2020.100071.
[68] O. Jordan, B. H. Gan, S. Alwan, K. Perron, E. Sublet, V. Ducret, H. Ye, G. Borchard, J.-L. Reymond, V. Patrulea, *Adv. Healthcare Mat*. **2024**, *13*, 2304118, https://doi.org/10.1002/adhm.202304118.

The definitive version of this article is the electronic one that can be found at https://doi.org/10.2533/chimia.2024.648