

## USING SEQUENCE PATTERNS TO IDENTIFY AND DESIGN NEW ANTIMICROBIAL PEPTIDES

*I.E. Eliseev<sup>1</sup>, I.N. Terterov<sup>1</sup>, O.V. Shamova<sup>2</sup>*

<sup>1</sup> Nanobiotechnology lab, Saint Petersburg National Research Academic University RAS, Saint Petersburg, Russia;

<sup>2</sup> Department of General Pathology and Pathophysiology, Institute of Experimental Medicine, Saint Petersburg, Russia

## ИСПОЛЬЗОВАНИЕ ПАТТЕРНОВ АМИНОКИСЛОТНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ДЛЯ ПОИСКА И ДИЗАЙНА АНТИМИКРОБНЫХ ПЕПТИДОВ

*И.Е. Елисеев<sup>1</sup>, И.Н. Тертеров<sup>1</sup>, О.В. Шамова<sup>2</sup>*

<sup>1</sup> Лаборатория нанобиотехнологий, ФГБУ ВОН «Санкт-Петербургский национальный исследовательский академический университет РАН», Санкт-Петербург;

<sup>2</sup> Отдел общей патологии и патологической физиологии,

ФГБНУ «Институт экспериментальной медицины», Санкт-Петербург

Natural antimicrobial peptides (AMPs) are remarkably diverse, yet they all share some common structural and functional features. In an attempt to find what determines similar activities of non-homologous molecules, we performed a comprehensive analysis of sequence patterns in AMPs. We found that natural AMPs possess characteristic sequence patterns, and these patterns differ for peptides with  $\alpha$ -helical and  $\beta$ -sheet structure. We showed that the patterns facilitate computational identification of AMPs in databases. We then used patterns to design new peptides, synthesized them and assayed for antibacterial activity. The most active among synthetic peptides exhibited activity against Gram(+) and Gram(-) pathogens comparable to best natural AMPs.

**Keywords:** antimicrobial peptides; structure-activity relationship; peptide design; pattern discovery.

Природные антимикробные пептиды (АМП) чрезвычайно разнообразны, однако все они имеют общие структурные и функциональные характеристики. В попытке определить, что именно лежит в основе схожего механизма действия негомологичных молекул, мы провели детальный анализ паттернов в аминокислотных последовательностях АМП. Мы обнаружили что аминокислотные последовательности природных АМП обладают характерными паттернами, и эти паттерны различны для пептидов с  $\alpha$ -спиральной и  $\beta$ -листовой структурой. Мы показали, как паттерны могут быть эффективно использованы для поиска последовательностей новых АМП в базах данных. Затем мы использовали паттерны для дизайна новых пептидов, которые были синтезированы и экспериментально исследованы. Наиболее активный из синтетических пептидов проявил антимикробную активность по отношению к Грам(+) и Грам(-) патогенам на уровне лучших природных АМП.

**Ключевые слова:** антимикробные пептиды; связь структура – активность; дизайн пептидов; поиск паттернов.

**Introduction.** Natural antimicrobial peptides (AMPs) are remarkably diverse, yet they all share some common structural and functional features. Most AMPs target bacterial membrane, where they associate into dimers and oligomers and form channels or pores, which ultimately leads to bacterial killing. This general mechanism relies on the ability of AMPs to adopt well-defined amphipathic structures upon binding to membranes.

Despite the similarity in their mode of action, AMPs from different organisms have virtually no sequence homology. The elements in their sequences that underlie similar activities of non-homologous molecules remain obscure. Many efforts have been made to identify these elements, sometimes named “templates” [1] or “signatures” [2].

Modern computational methods greatly facilitate this sequence analysis. Perhaps the most powerful algorithm for the identification of patterns in non-homologous sequences is *Teiresias* [3]. A great example of its successful application to the design

of AMPs is the “linguistic” approach [4]. In this method, patterns derived from natural AMPs governed the construction of new peptide sequences from individual residues, analogously to grammar rules in a natural language. Inspired by this elegant approach, we aimed to understand whether the patterns in natural AMPs encode certain structural features, and how to use them to identify or design new active peptides.

**Materials and methods.** For the computational analysis of patterns, we selected sequences of natural AMPs from APD2 database [5] with confirmed activity against Gram(+) and Gram(-) bacteria, structural annotation, and non-negative charge. We extracted patterns by running *Teiresias* on the entire AMP set, and different subsets:  $\alpha$ -helical peptides,  $\beta$ -sheet peptides, and  $\alpha$ + $\beta$  peptides. To analyze pattern specificity, we performed the same analysis on sets of *shuffled* sequences generated by random permutations of residues in natural AMPs.

Experiments were performed with peptides synthesized by Fmoc chemistry, purified by chromatography on a C18 column and characterized by mass spectrometry. Measurements of minimal inhibitory concentrations (MICs) were performed by standard broth microdilution assay with two Gram(+) pathogens, *S. aureus* and *L. monocytogenes*, and two Gram(-) pathogens, *E. coli* and *P. aeruginosa*. Structural analysis of peptides was made by circular dichroism spectroscopy in buffer solutions and membrane-mimicking environments, modeled by anionic (SDS) or zwitterionic (DPC) micelles.

**Results and discussion.** Here we briefly present our main findings, a more detailed description of which appears in our very recent paper [6]. We first questioned, whether the presence of characteristic sequence patterns is an inherent property of natural AMPs, not just a coincidence. We compared patterns derived from natural AMPs with patterns from corresponding shuffled sets. The analysis revealed that natural sequences always had more patterns with high *support*, which is defined as the number of times a pattern appears in the sequence set, than their shuffled analogues. This result demonstrates the existence of common ordered motifs in the sequences of natural AMPs, which is not a consequence of amino acid composition.

We then compared patterns derived from peptides with different structure. It appeared that  $\alpha$ -helical and  $\beta$ -sheet peptides shared very few patterns, and all common patterns had low support, i.e. they were not statistically significant. Interestingly,  $\alpha$ + $\beta$  peptides shared more patterns with  $\alpha$ -helical and  $\beta$ -sheet peptides than these two groups with each other. These observations suggest that patterns may carry certain structural information.

Since natural AMPs possess characteristic sequence patterns, we hypothesized that these pat-

terns can be used to identify new antimicrobials in databases. We took short peptides from Swiss-Prot and selected 115 candidates having  $> 5\alpha$ -helical patterns. From the selected sequences, 65% had experimentally confirmed antimicrobial activity, and additional 15% are presumably antimicrobial based on homology. Moreover, most of the peptides with known structure identified in this computational experiment were  $\alpha$ -helical.

Having established the connection between certain sequence patterns and antimicrobial activity, we used patterns from natural  $\alpha$ -helical AMPs to design new synthetic peptides. Patterns were merged successively to yield sequence templates of length 20. The sequence templates were filled with random residues and filtered based on their physicochemical properties. Seven sequences were selected, synthesized, and then assayed for antimicrobial activity. Surprisingly, 5 of 7 designed peptides were active ( $\text{MIC} \leq 128 \mu\text{M}$ ), and one peptide had a low micromolar activity (1–8  $\mu\text{M}$ ) against all pathogens tested, which is comparable to best natural AMPs. Its scrambled versions were also synthesized, and demonstrated 4 to 128 fold reduced activity. Finally, we analyzed the structure of the designed peptides. Based on circular dichroism spectra, all peptides were unstructured in solution but adopted an  $\alpha$ -helical conformation upon binding to SDS micelles, and 6 of 7 were helical in complex with DPC micelles. The structural data support our hypothesis that sequence patterns encode certain structural information.

In conclusion, our results suggest that natural AMPs possess characteristic sequence patterns, which encode their structural and functional features. We demonstrate that this approach may be useful in the identification and computational design of new active peptides.

## References

1. Zelezetsky I, Tossi A. Alpha-helical antimicrobial peptides — using a sequence template to guide structure-activity relationship studies. *Biochim. Biophys. Acta.* 2006;1758:1436-1449.
2. Yount NY, Yeaman MR. Multidimensional signatures in antimicrobial peptides. *Proc. Natl. Acad. Sci. U.S.A.* 2004;101:7363-7368.
3. Rigoutsos I, Floratos A. Combinatorial pattern discovery in biological sequences: the TEIRESIAS algorithm. *Bioinformatics.* 1998;14:55-67.
4. Loose C, Jensen K, Rigoutsos I, Stephanopoulos G. A linguistic model for the rational design of antimicrobial peptides. *Nature.* 2006;443:867-869.
5. Wang G, Li X, Wang Z. APD2: the updated antimicrobial peptide database and its application in peptide design. *Nucleic Acids Res.* 2009;37:D933-D937.
6. Eliseev IE, Terterov IN, Yudenko AN, Shamova OV. Linking sequence patterns and functionality of alpha-helical antimicrobial peptides. *Bioinformatics.* 2018. bty1048 (epub ahead of print).