A GENERAL PROTOCOL FOR THE CONSTRUCTION OF STRUCTURE-ACTIVITY LANDSCAPES OF NON-CANONICAL PEPTIDES

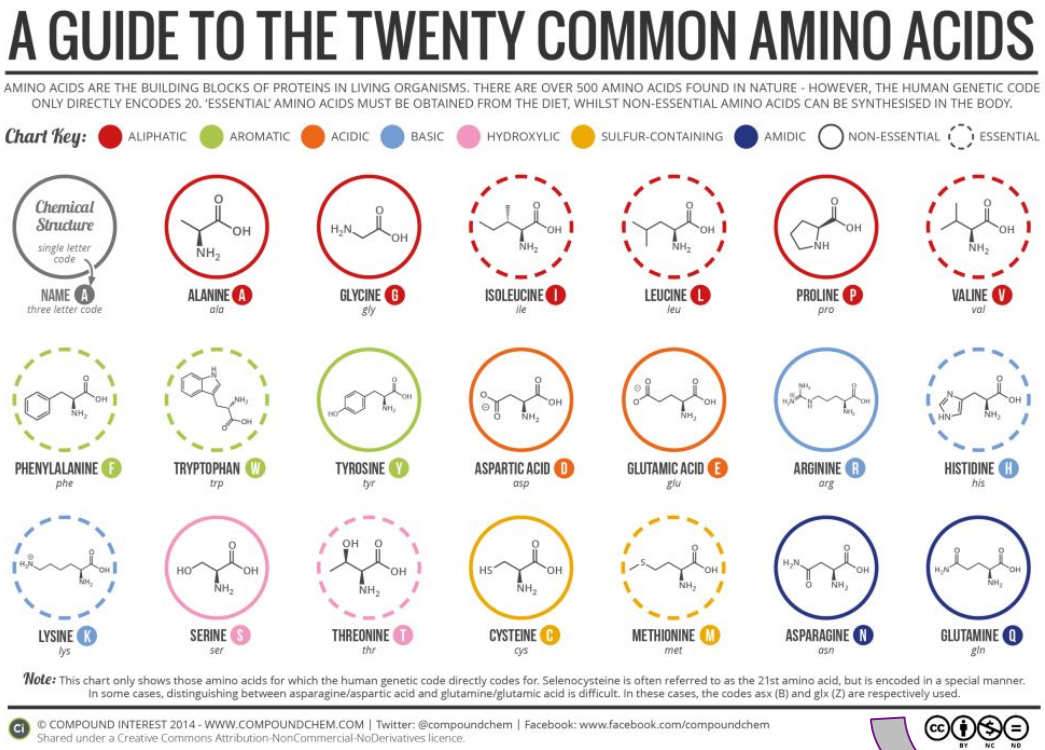
Edgar López-López [1,2,*]

[1] DIFACQUIM Research Group, Departamento de Farmacia, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad de México 04510, México.

[2] Departamento de Química y Programa de Posgrado en Farmacología, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Apartado 14-740, Ciudad de México 07000, México.

[*] **elopez.lopez@cinvestav.mx**

Introduction: What is a peptide and non-canonical peptide? What are their importance?



Biomolecules 2023, 13(6), 981.

Non-canonical (uncommon) peptides

- Improved 3D stability
- Enhanced bioactivity (e.g. affinity, ADMET).
- Improved specificity

Methodology

Database generation

We collected the peptide sequences and bioactivity data from APD3 database of natural or synthetic peptides, with anti-MRSA reported activity (mg mL−1), with <70 residues.

.

SMILES calculation

The SMILES strings were obtained manually using each peptide sequence and their non-canonical modifications.

Descriptors calculation

Using Python programming, the fingerprint MAP4 was calculated for each peptide.

Additionally, paired calculations:

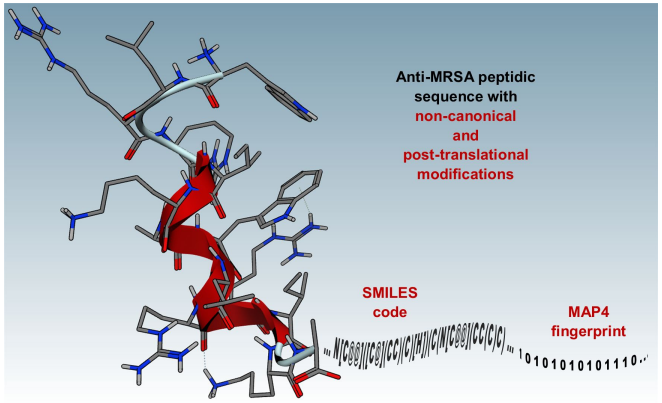
- Molecular similarity
- Paired difference activity
- Paired difference molecular weight.

SAS map visualization

The paired data was visualized and the Structure-Activity Landscape Index (SALI) values were calculated.

Chemical space analysis

Identification of peptide activity cliffs.



J Cheminform 2020, 12 (43), 1758-2946

Discussion: Structure-Activity(Propertie) Landscape

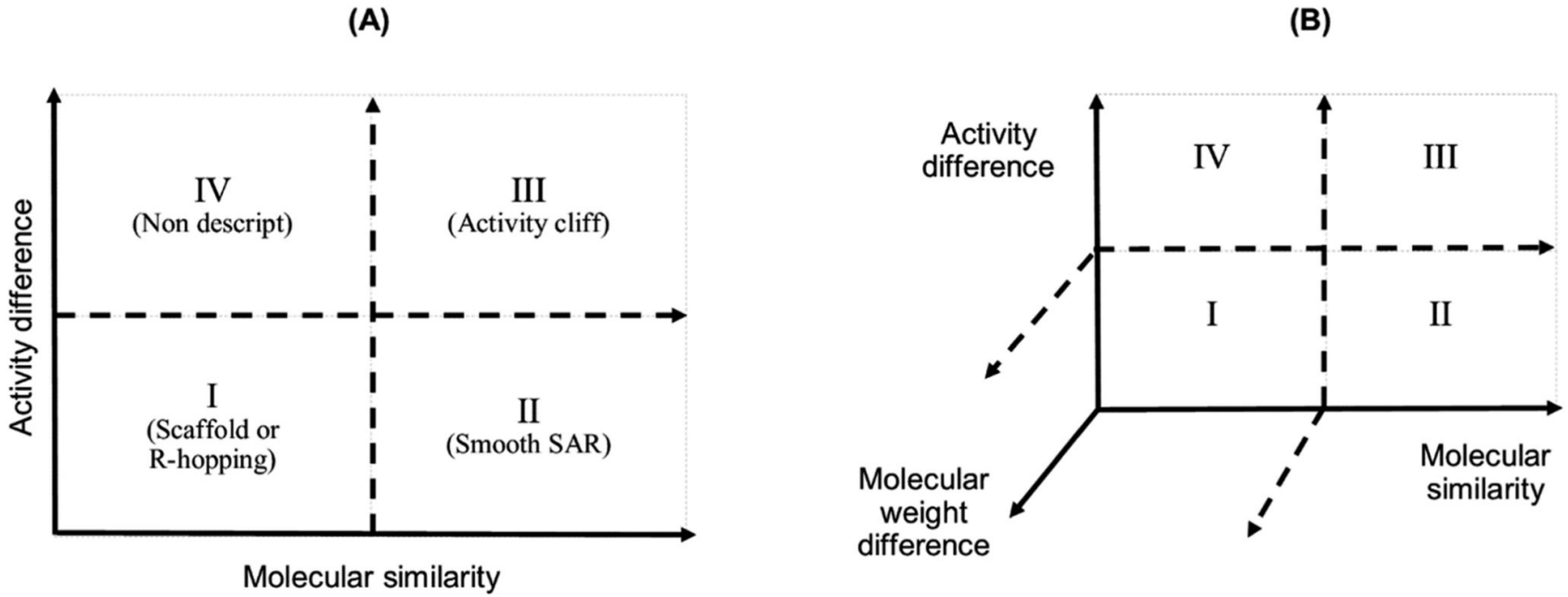
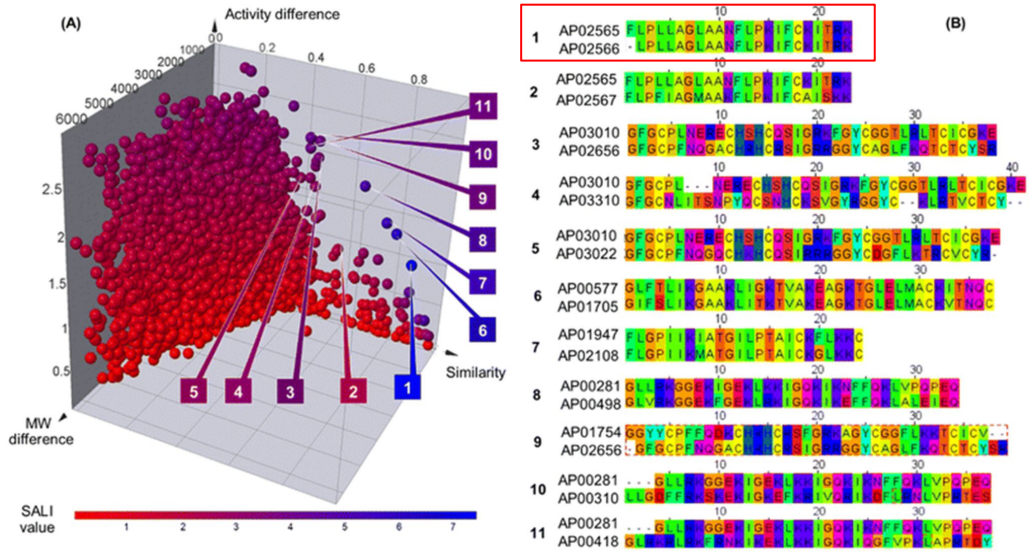


Fig. 1 Graphical representation of a Structure-Activity Similarity (SAS) map (A) and an extension of a SAS map (B). A SAS map is based on a pairwise comparison of each compound on a data set. Each data point in the graph in the map represents a pair of compounds. SAS map is based on the activity differences of the pair of compounds against a specific biological endpoint and their molecular distance.

Results



Pair	Peptide pair	Peptide activity 1	Peptide activity 2	Activity Difference ($pMIC_{50}$)	Fingerprint-based similarity	SALI value	MW difference	Percentage of identity
1	AP02565 — AP02566	318.4 μM	2.1 μM	0.96	0.866	1.11	147.2	100.00
2	AP02565 — AP02567	318.4 μM	2.4 μM	0.90	0.561	1.61	33.0	69.57
3	AP03010 — AP02656	32 μM	0.4 μM	1.90	0.453	4.19	38.0	55.26
4	AP03010 — AP03022	32 μM	0.63 μM	1.70	0.447	3.96	121.1	56.78
5	AP03010 — AP03311	32 μM	0.68 μM	1.66	0.387	4.31	90.2	52.78
6	AP01705 — AP00577	25 μM	1.33 μM	1.27	0.790	6.06	15.9	89.19
7	AP01947 — AP02108	24.14 μM	0.99 μM	1.38	0.750	5.56	72.0	91.67
8	AP00281 — AP00498	64 μM	1.12 μM	1.75	0.651	5.03	67.9	76.47
9	AP01754 — AP02656	50 μM	0.4 μM	2.14	0.454	3.93	38.0	61.11
10	AP00281 — AP00310	64 μM	0.33 μM	2.28	0.452	4.16	614.2	47.06
11	AP00281 — AP00418	64 μM	0.45 μM	2.15	0.452	3.93	574.4	55.8

Fig. 2 Structural and sequence similarity between the 223 anti-MRSA peptides: (A) modified (extended) structure-activity similarity map; each sphere represents a pairwise comparison of the chemical structure (quantified utilizing MinHassed distance/MAP4 fingerprints), activity difference, and molecular weight difference. The spheres are colored according to the SALI values using a continuous scale from low (blue) to high (red) values. (B) Sequence alignment and (C) summary characterization of 11 representative peptide activity cliffs (pairs). SALI: Structural–Activity Landscape Index.

Brief discussion

Pair	Activity Difference ($pMIC_{50}$)	Fingerprint -based similarity	Percentage of identity
12	0.12	0.962	92.31
13	0.90	0.562	80.00
14	1.31	0.371	47.37
15	0.20	0.237	22.22

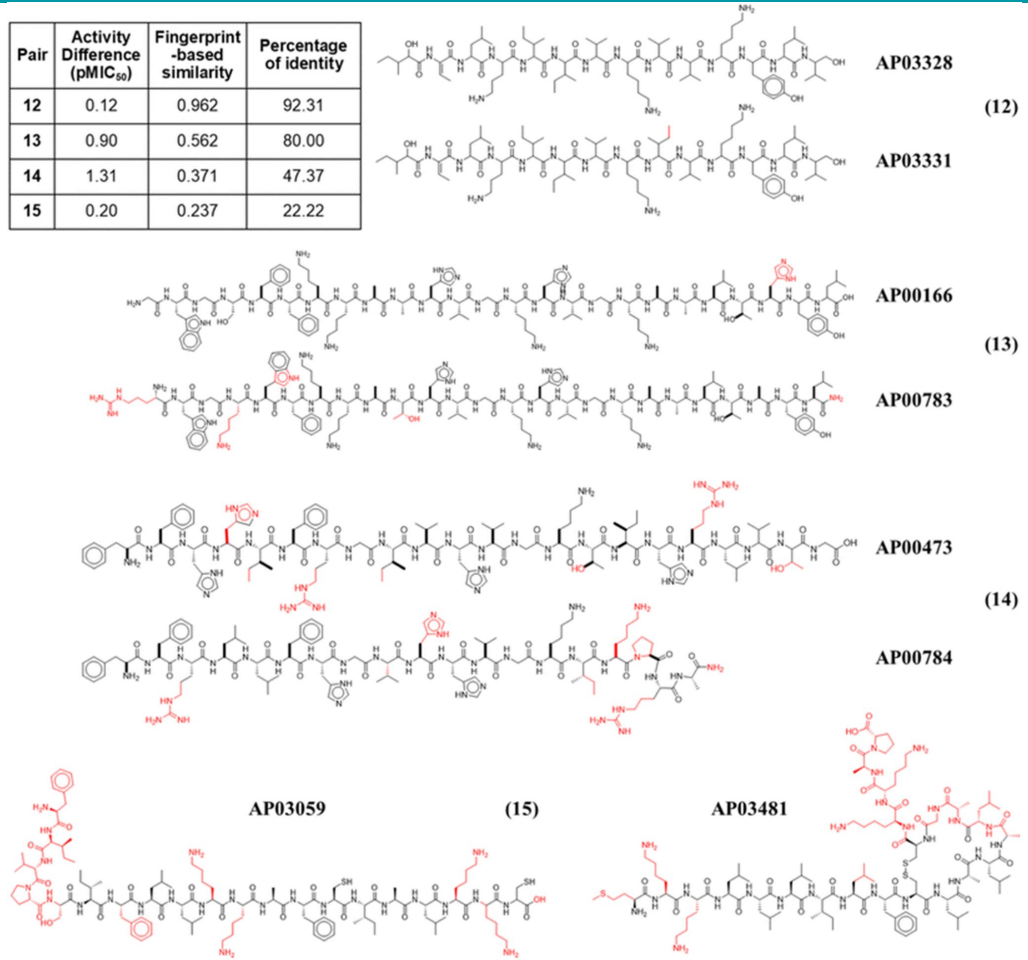


Fig. 3 Representative anti-MRSA peptide pairs 12–15. Chemical changes observed between each peptide pair are colored in red, whereas shared chemical structures are depicted in black.

What can we identify?

- Small structural changes
- Changes in the -COOH and -NH ends
- Non-canonical modifications
- Changes in linear and cyclic peptides

Conclusions and future perspectives

- The fingerprint-based similarity (FP-Sim) values positively correlate with the sequence-based identity values (R^2 = 0.31), **suggesting that FP-Sim measures complement the information derived from sequence alignments, but do not replace them.**
- We introduced the **extended SAS map** (using MW differences values of each peptide pair) that facilitated the **rapid identification of peptide activity cliffs.**
- This protocol has a **very broad domain of applicability**.

Check it out!!

Issue 5, 2023

Previous Article | Next Article

From the journal:
Digital Discovery

Mapping the structure–activity landscape of non-canonical peptides with MAP4 fingerprinting†

Check for updates

Edgar López-López, *[ab] Oscar Robles, [c] Fabien Plisson [d] and José L. Medina-Franco *[a]

Author affiliations

Abstract

Peptide structure–activity/property relationship (P-SA/PR) studies focus on understanding how the structural variations of peptides influence their biological activities and other functional properties. This knowledge accelerates the rational design and optimisation of peptide-based drugs, biomaterials, or diagnostic agents. These studies examine peptide structures from their primary sequences, essentially encoded from the 20 amino acids. Current approaches often exclude peptide libraries with post-translational and synthetic modifications. The molecular fingerprint MAP4 was recently developed to map complex molecules' sequence/structure diversity, including peptides. This study used structure–activity landscape modelling to conduct the P-SA/PR studies of an exemplary dataset of 223 antimicrobial peptides against methicillin-resistant *Staphylococcus aureus* (MRSA). To this end, we employed the MAP4 fingerprint to represent the chemical structures of the peptides, study their relationship(s) with the antibacterial activity, and seek the potential activity cliff(s). We identified critical residues and structural motifs that play a crucial role in the anti-MRSA activity of the peptides. This is the first computational study to systematically explore the activity landscape of peptides with non-canonical residues, emphasising the quantification of structural similarity.

Click the following link to read more:
https://doi.org/10.1039/D3DD00098B

Acknowledgements

Collaborators

Dr. Carolina Barrientos-Salcedo (Universidad Veracruzana)
Dr. Karina Martínez-Mayorga (UNAM)
Dr. Michel M. Vichi-Ramírez (Universidad Veracruzana)
Dr. Carlos M. Cerda-García-Rojas (CINVESTAV)
Dr. Daniel Mucs (Karolinska Institute)
Dr. Fabien Plisson (CINVESTAV)
Dr. Gilberto Castañeda Hernández (CINVESTAV)
Dr. José L. Medina-Franco (UNAM)
Dr. Jürgen Bajorath (Rheinische Friedrich-Wilhelms-Universität)
Dr. Oscar Robles (Universidad Veracruzana)
Dr. Ramón A. Miranda- Quintana (University of Florida)
Dr. Ulf Norinder (Karolinska Institute)

Special thanks:

To the entire team, my colleagues and friends.

To the organizers of XXX Symposium on BCADD-2024.

Thanks all for your kind attention and time.

Working groups:

Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional
Química de Productos Naturales Orgánicos

DIFACQUIM
UNAM

DIFACQUIM
COMPUTER-AIDED DRUG DESIGN AT UNAM

https://www.difacquim.com/

Financing:

CONAHCYT
CONSEJO NACIONAL DE HUMANIDADES CIENCIAS Y TECNOLOGÍAS

- **PAPIME** - PE200118
- **PAPIIT**- IA203718 and IN201321
- **PAIP** -5000-9163
- **NUATEI-IIB-UNAM**
- **DGTIC-HUAWEI**

Collaboration is part of our scientific philosophy! Contact us!

elopez.lopez@cinvestav.mx

Edgar López-López

Edgar-Lopez-Lopez-4

@EdgarLpezLpez10

Edgar López-López

@edgar_lopez_2

A GENERAL PROTOCOL FOR THE CONSTRUCTION OF STRUCTURE-ACTIVITY LANDSCAPES OF NON-CANONICAL PEPTIDES

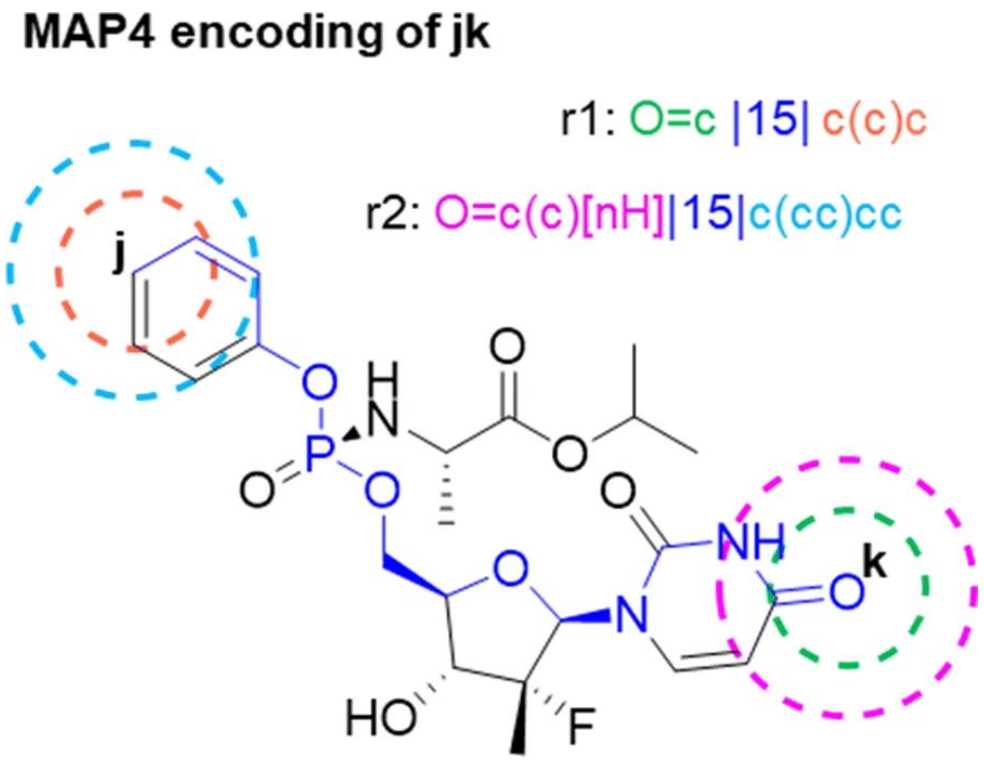
Edgar López-López [1,2,*]

[1] DIFACQUIM Research Group, Departamento de Farmacia, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad de México 04510, México.

[2] Departamento de Química y Programa de Posgrado en Farmacología, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Apartado 14-740, Ciudad de México 07000, México.

[*] **elopez.lopez@cinvestav.mx**

Supporting information



J Cheminform 2020, 12 (43), 1758-2946

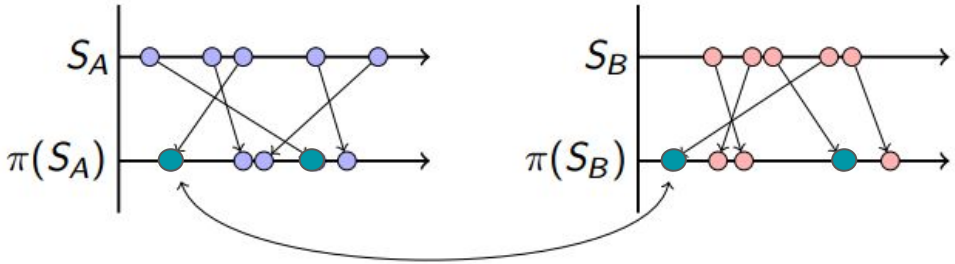
Example:

- For atoms *J* and *K*
- MAP4 = Radius 2
  - (Blue and Pink)

Notes:

- Topological FP (e.g. Morgan and ECFP)
- Each connectivity of *X* atom (with a radius *Y*) is represented by a unique bit (0, 1)
- 1024 bits … or more.
- Similarity quantification is based on MiniHash algoritm (inspired on Jaccard Similarity for large sets).

Supporting information



Notes:

- MAP encoding of A (SA) and B (SB)
- Similarity (A,B)
- π = Random perturbation of SA and SB
- πSA and πSB used by the fingerprint

Fingerprinting

SA = 10011110
SB = 11100101

= In A and B
= Only in A
= Only in B

1 = Presence
0 = Absence

J Cheminform 2020, 12 (43), 1758-2946

Adapted to: Approximating Jaccard similarity with MinHash. September, 2024.
https://aksakalli.github.io/2016/03/01/jaccard-similarity-with-minhash.html