Structural composition analysis of approved peptide therapeutics and diagnostics as a guide for future peptide drug candidates

Vera D'Aloisio[1], Paolo Dognini[1], Danielle Liebnitz[2], Gillian A. Hutcheon[1], Christopher R. Coxon[2]

[1] School of Pharmacy & Biomolecular Sciences, Liverpool John Moores University, Liverpool, L3 3AF, UK

[2] Institute of Chemical Sciences, School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, EH14 4AS

Scan Me to access our open-source database

PepTherDia

Introduction

Small molecules (MW<500 Da)

Peptides

Chemical entities comprising:

- ✓ at least 2 amino acids linked by an amide bond
- ✓ maximum of 50 amino acids
- ✓ molar mass < 5000 g/mol
- ✓ They combine advantages of small molecules and proteins.

Large biologics (MW>5000 Da)

The peptide pharmaceuticals market is growing rapidly (**Fig 1**) but there is no general set of **guidelines** that can **increase** the **success rate** of peptide-based drug **approvals**.

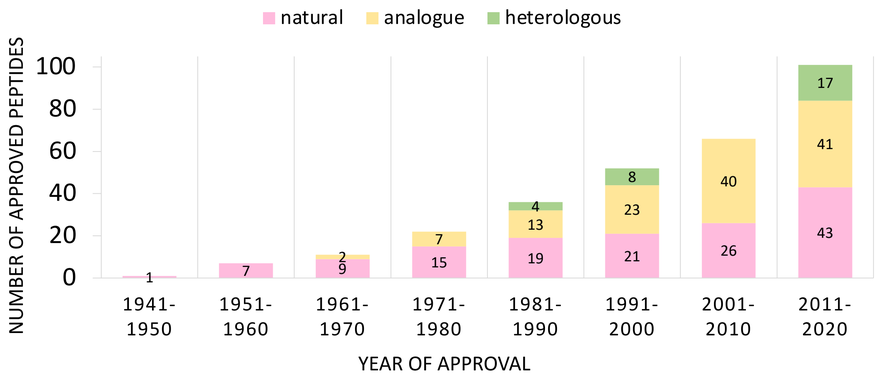

Fig. 1 Evolution of approved peptide medicines throughout the 20th and 21st centuries, colour-coded based on origin.

Aim

The aim of this study was to perform a detailed structural analysis on a database of approved peptide therapeutics and diagnostics and to evaluate their *in silico* physico-chemical properties. We aim to provide an overview of the key compositional trends of the peptides on the market to help guide the design of future peptide medicines.

Methods – data collection, structural analysis, and *in silico* predictions

1. Peptides search

- Peptide-based drugs and diagnostics were listed by searching in FDA, EMA, and PMDA web pages, DrugBank and Drug Central websites.
- Diversities in the so-called 'peptide' molecules made necessary to establish strict definition and selection criteria.

2. Criteria of selection

- Non-insulin peptide drugs
- Only human-use peptides
- Theragnostic and diagnostic peptides
- Peptides approved in at least in one country of the main pharmaceutical market areas (North America, Western Europe, and Japan)

3. Peptide properties

- *In silico* prediction software (i.e., Chemicalize, ACD-Labs) were used to calculate LogP, LogD, pI, tPSA, HBA, HBD of approved peptides.
- Data on terminal half-life, protein binding, therapeutic indication, routes of administration, design, and production methods were collected from DrugBank, NIH, Drugs.com, and pharmaceutical companies' websites.

4. Structural analysis

- Each peptide was divided into its basic building blocks, comprising natural and non-natural amino acids and modifications (e.g., lipophilic chains, sugars, or others).
- Each building block was further classified as aromatic, aliphatic, polar, acidic, basic.

5. PepTherDia

- A manually-curated repository of 105 approved peptide therapeutics and diagnostics was created.
- Information is stored in our database, **PepTherDia**.
- The database is freely available online and will be updated regularly.

Results – trends in peptide structure and *in silico* properties

From PepTherDia database (Fig 2):

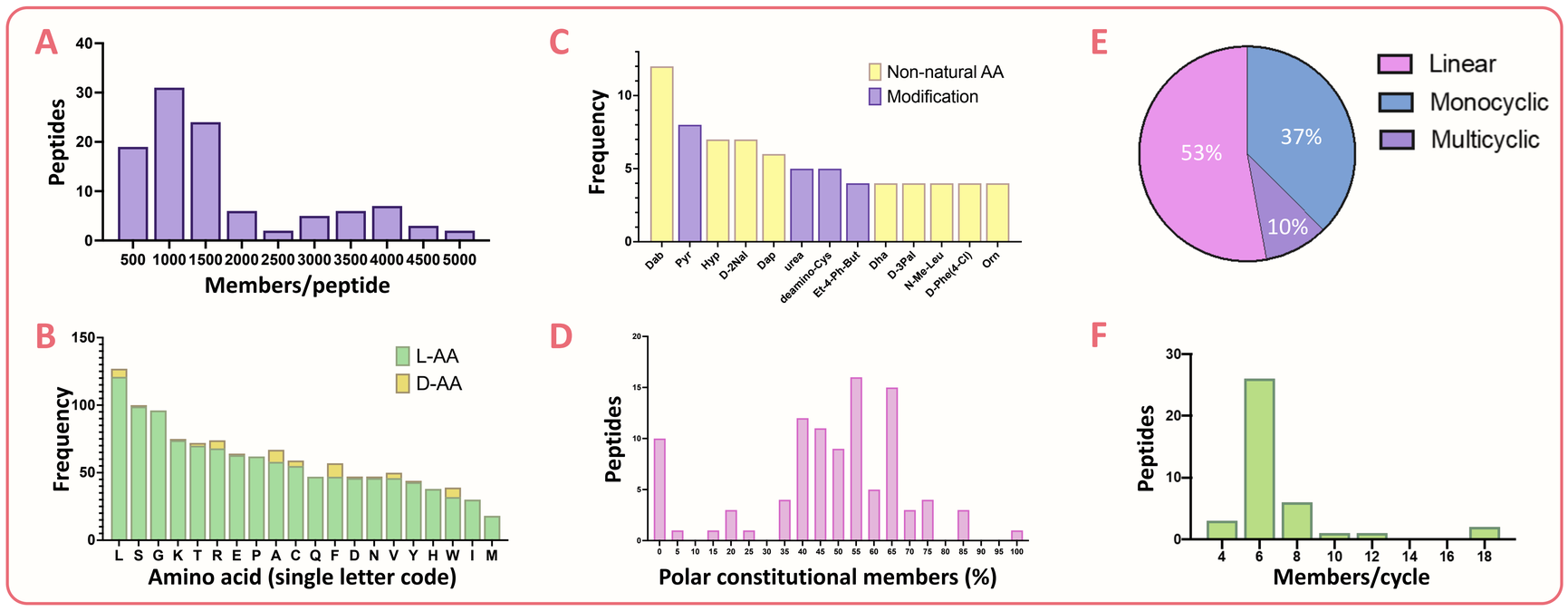

Fig. 2 Structural trends in the pool of approved peptide therapeutics and diagnostics. (A) Number of constitutional members distribution; **(B)** occurrence of L-amino acids (AAs) (light green) and D-AAs (yellow); **(C)** most frequently encountered non-natural AAs, in yellow, and modifications, in purple; Dab for **(D)** polarity distribution within the pool of approved peptides **(E)** peptide structure, divided into linear, monocyclic, and multicyclic; **(F)** macrocycle size, shown as the number of constitutional members per cycle [1].

***In silico* studies (Fig 3):**

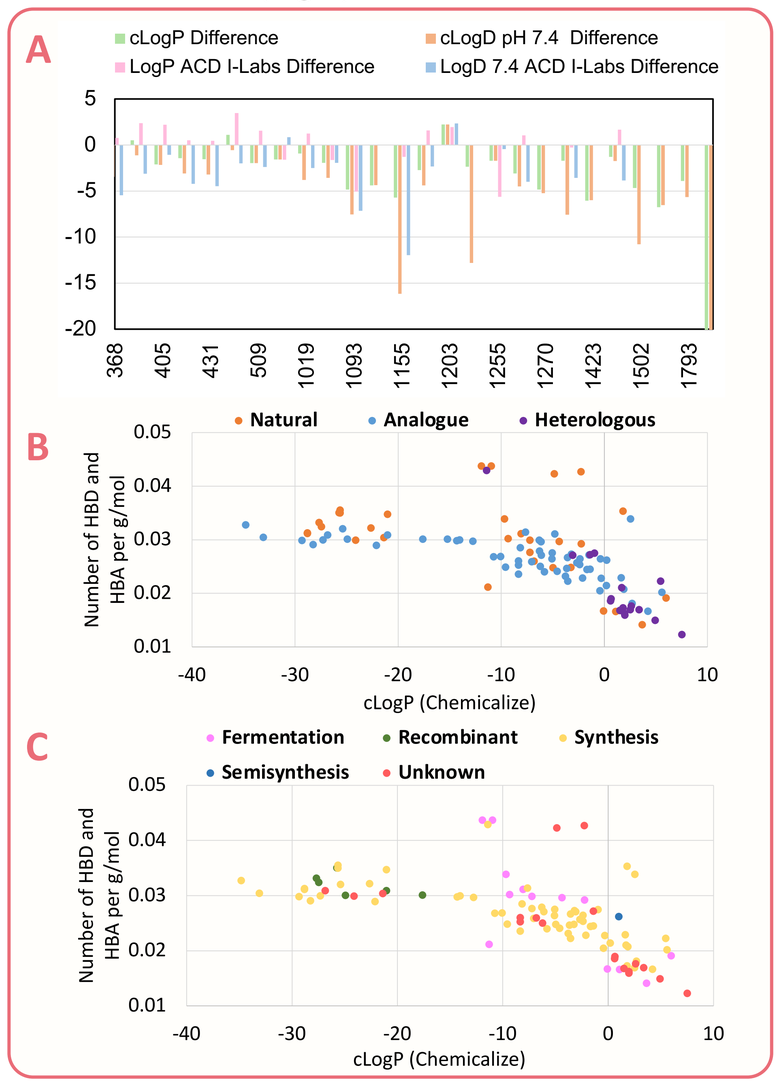

Fig. 3 *In Silico* trends in the pool of approved peptides. (A) Comparison of lipophilicity predicted values with literature values ranked by molar mass (g/mol); **(B)** density of hydrogen bond forming versus cLogP predictions colour-coded by peptide; **(C)** density of hydrogen bond forming versus cLogP predictions colour-coded by production method.

Conclusions

In this study, we provide an overview of the strategies most commonly used in peptide drug design, which have successfully brought these peptides to the market and we propose the use of *in silico* tools for small-size peptide ranking and evaluation.

A peptide most likely to become a drug or a diagnostic agent will have a molar mass < 2000 g/mol; comprise mainly natural amino acids (around 81%), with a balance between hydrophobic and polar building blocks; present mainly free or amidated C-terminal, free or acetylated N-terminal; and, finally, have small-size cycle (5-7 members), if present (47% of the cases).

***In silico* experiments** showed good accuracy and proved to be useful tools for molecules lipophilicity ranking within a molar mass range from 300 to 1500 g/mol. In general, heterologous peptides appear to tend towards more lipophilic values of cLogP and lower density of hydrogen bonding groups, purposefully designed to increase chances of oral availability.

With expected **upwards trend** in peptide **approvals** there will be no lack of data to update and improve these guidelines to make them more and more accurate.

References

[1] Vera D'Aloisio, Paolo Dognini, Gillian A. Hutcheon, Christopher R. Coxon, PepTherDia: database and structural composition analysis of approved peptide therapeutics and diagnostics. Drug Discovery Today, 2021; https://doi.org/10.1016/j.drudis.2021.02.019.