DBAASP: A COMPREHENSIVE REPOSITORY OF NATURAL MULTIFUNCTIONAL CYCLIC ANTIMICROBIAL PEPTIDES

XXIX Symposium on Bioinformatics and Computer-Aided Drug Discovery _ BCADD-2023

N. Samsonidze, E. Alimbarashvili, M. Grigolava, M. Pirtskhalava
Ivane Beritashvili Center of Experimental Biomedicine, Tbilisi, Georgia

18/09/2023
Cyclic peptides

Why Cyclic peptides?

Macrocyclization improves the pharmacological properties and bioactivity of peptides.

- High metabolic stability
- Oral availability
- Selective affinity for receptors
- Low cytotoxicity
DBAASP is a repository of data on antimicrobial/cytotoxic activity and structure of more than 21,000 peptides.

https://dbaasp.org/home
DBAASP stores data on 4357 Cyclic peptides

1759 peptides are natural

They have various targets

Target groups of Cyclic peptides

- Virus
- Protista
- Parasite
- Nematode
- Mollicute
- Mammalian cell
- Insect
- Gram+
- Gram-
- Fungus
- Cancer
- Biofilm
Which bonds are used by nature to cyclize peptides?

- **TIE** = Thioether
- **IMN** = Imine
- **ETH** = Ether
- **EST** = Ester
- **DSB** = Disulfide
- **CAR** = Carbon
- **AMD** = Amide
- **AMN** = Amine
Small macrocyclic peptides represent more valuable drug candidates

Small Cyclic peptides defined as a peptides with the length 1-25 aa

Ribosomal cyclic peptides length $<25$ AA

Non-ribosomal cyclic peptides length $<25$ amino acids (AA)

Ultra-short length 2-5aa and short length 6-25aa

Ultra-short peptides (USP) satisfy Rule of 5 (Ro5)

Short peptides (SP) does not. They have beyond Ro5 (bRo5) targets

---

Small cyclic peptides

Ultra-short peptides (USCPs)

Short cyclic peptides (SCPs)
Natural USCP

Total number of natural USPs in DBAASP is 197, comprising 7 ribosomal and 190 non-ribosomal peptides. Among non-ribosomal 145 are cyclic. Among ribosomal only 2 are cyclized.

Majority of USCPs are cyclized by amide bonds

- **AMD** = Amide
- **AMN** = Amine
- **DSB** = Disulfide
- **EST** = Ester
- **TIE** = Thioether

### Bonds of short cyclic peptides

<table>
<thead>
<tr>
<th>Bonds</th>
<th>Nonribosomal</th>
<th>Ribosomal</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMD</td>
<td>149</td>
<td>2</td>
</tr>
<tr>
<td>AMN</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>DSB</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>EST</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>TIE</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

### Nonribosomal USCPs target objects

- Cytoplasmic protein: 42%
- Membrane protein: 27%
- DNA/RNA: 9%
- Mitochondrial membrane: 20%
- Virus replication: 1%
- Quorum sensing: 1%
- Lipid bilayer: 1%
- Membrane protein: 1%

Majority of USCPs are cyclized by amide bonds.
# Cyclic types of Natural USCP bonds

<table>
<thead>
<tr>
<th>Type of bond</th>
<th>Type of Cycle formed by:</th>
<th>Nonribosomal AMPs</th>
<th>Ribosomal AMPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amide</td>
<td>LAC-Lactam&lt;br&gt;NCP (head to tail cyclized)&lt;br&gt;DKP (2,5 diketopiperazine)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Amine</td>
<td>LAC</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Disulfide</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Ester</td>
<td>LCN-Lactone</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Thioether</td>
<td>THZD (Thiazolidine)</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>
Pro is abundant in the natural USCPs

Amino acid composition of nonribosomal linear peptides length of 1-5 (USPs). Total 29 entries.

Amino acid composition of nonribosomal cyclic peptides length of 1-5 (USCPs). Total 145 entries.
Natural SCPs of DBAASP

DBAASP stores data on 1482 ribosomal and 355 nonribosomal SPs of length in the interval of 6-25 aa. Among Ribosomal SP, 419 are cyclic. Among nonribosomal SP cyclized 295 peptides.
Bonds used for cyclization of Natural SCPs

Majority of ribosomal SCPs are cyclized by disulfide bonds, while in nonribosomales such bonds are not appeared.

**Bonds of short cyclic peptides**

<table>
<thead>
<tr>
<th>Bonds</th>
<th>Nonribosomal</th>
<th>Ribosomal</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMD</td>
<td>202</td>
<td>75</td>
</tr>
<tr>
<td>AMN</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>DSB</td>
<td>0</td>
<td>356</td>
</tr>
<tr>
<td>EST</td>
<td>71</td>
<td>2</td>
</tr>
<tr>
<td>TIE</td>
<td>9</td>
<td>19</td>
</tr>
<tr>
<td>CAR</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>ETH</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>IMN</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

**Codes**

- AMD = Amide
- AMN = Amine
- DSB = Disulfide
- EST = Ester
- TIE = Thioether
- CAR = Carbon
- ETH = Ether
- IMN = Imine

Majority of ribosomal SCPs are cyclized by disulfide bonds, while in nonribosomales such bonds are not appeared.
### Cyclic types of Natural SCP bonds

<table>
<thead>
<tr>
<th>Type of bond</th>
<th>Type of Cycle formed by:</th>
<th>Nonribosomal AMPs</th>
<th>Ribosomal AMPs</th>
<th>Type of bond</th>
<th>Type of Cycle formed by:</th>
<th>Nonribosomal AMPs</th>
<th>Ribosomal AMPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amide</td>
<td>Macrolactams</td>
<td></td>
<td></td>
<td>Thioether</td>
<td>THZD (Thiazolidine)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NCB (head to tail cyclized)</td>
<td></td>
<td></td>
<td>THZ (Thiazoline)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amine</td>
<td>Macrolactams</td>
<td>+</td>
<td>+</td>
<td>LNT (Lanthionine)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disulfide</td>
<td>CST (Cystine)</td>
<td>+</td>
<td></td>
<td>Carbon</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ester</td>
<td>LCN - Lactone</td>
<td>+</td>
<td>+</td>
<td>Ether</td>
<td>OXZ (Oxazoline)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>THZ (Thiazoline)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NCB, Macrolactames, Lactones, Thiazolines Thiazolidinines are presented in both ribosomal and nonribosomal SCPs.

Oxazolines, Lanthionines and Cystines _in ribosomal SCPs._
Natural short cyclic peptides amino acid composition relative to UniProt

Ribosomal short cyclic peptides Amino acid composition relative to ‘average protein’

RSCPs contain more hydrophobic (with a higher abundance of phenylalanine, isoleucine, leucine, and tryptophan) and more basic (due to higher abundance of lysine) amino acids compared to the ‘average protein’.
Conclusion

- Bonds used to cyclize structure in both ribosomal and nonribosomal peptides are: amide, ether, ester, thioether, amine.
- Disulfide bonds are formed mostly in ribosomal peptides.
- Thiazoline, Thiazolidine and Lactams and Lactones have been appeared in both, ribosomal and nonribosomal peptides.
- Different system of synthesis creates similar rings, the proteins engaged in these systems are not homologous, and consequently can be supposed, that have arisen through convergent evolution.
Thanks for attention!