A predictive model for the identification of diketopiperazines using UPLC-TOF-MS

Eric Dossin, Adrian Haiduc, Philippe Guy, Federico Mora, Arne Glabasnia, Valérie Leloup, Sunil Kochhar, and IMRE BLANK
2,5-Diketopiperazines are cyclic dipeptides formed by dehydration of proteins or peptide fragments.

The common scaffold is believed to be responsible for the different biological properties of DKPs.

- Antibacterial
- Antiviral
- Phytotoxic
- Taste active

References:

- Peptides, 1995, 16, 151-164
Diketopiperazines (DKPs) and their diversity

**Cyclo(Gly-Leu)**
Muscle stimulation
*(Peptides, 1995, 16, 151-164)*

**Cyclo(Pro-Leu)**

**Cyclo(His-Pro)**
Pain perception
(endogenous in mammals)

**Cyclo(Arg-Tyr)**
Antinociceptive activity
*(Peptides, 1995, 16, 151-164)*

**Cyclo(Asp-Pro)**
Decreased caloric intake
*(Peptides, 1995, 16, 151-164)*

**Cyclo(Pro-Trp)**
Antimicrobial
DKPs are widely found in nature

<table>
<thead>
<tr>
<th>BIOLOGICAL SYSTEMS</th>
<th>FOODS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central nervous system</td>
<td>Beer</td>
</tr>
<tr>
<td>Gastrointestinal tract</td>
<td>Milk</td>
</tr>
<tr>
<td>Hypothalamus</td>
<td>Coffee</td>
</tr>
<tr>
<td>Cerebrospinal fluid</td>
<td>Cocoa</td>
</tr>
<tr>
<td>Blood</td>
<td>Cereals</td>
</tr>
<tr>
<td>Urine</td>
<td>Noodles</td>
</tr>
<tr>
<td>Amniotic fluid</td>
<td>Potted Meat</td>
</tr>
<tr>
<td>Breast Milk</td>
<td>Ham</td>
</tr>
<tr>
<td>Semen</td>
<td>Egg</td>
</tr>
</tbody>
</table>

- J. Food Chem. 2009, 74, 2, C100-C105
DKPs found in roasted coffee

Sensory notes: Bitter, astringent, metallic, mouth-coating (taste threshold: 10-50 ppm)

 cis / trans isomerisation

(Ginz & Engelhardt, 2000, 2001)
Energy dependant reactions are required for the formation of DKPs

Natural protein /peptide

At pH and/or temperatures DKPs epimerize

Unnatural protein /peptide
DKPs found in roasted cocoa

(Pickenhagen et al., 1975)
DKPs found in roasted cocoa

Concentration: <0.2 – 1700 ppm
Taste threshold: 30 – >8000 μmol/l

(Stark & Hofmann, 2005)
Objective of the work

- Optimize chromatographic conditions for the analysis of diketopiperazines (DKPs) by UPLC
- Develop predictive models supporting the identification of DKPs based on TOF-MS detection
- Apply the predictive model for the identification of unknown DKPs in foods (coffee and cocoa)
Two chromatographic conditions used to build the prediction model

- **HILIC**
  - Column: BEH HILIC (2.1 x 150 mm, 1.7 µm), from Waters Corp.
  - Mobile Phase:
    - A: 5 mM ammonium acetate in water/acetonitrile (5:95, v/v)
    - B: 5 mM ammonium acetate in water
  - Gradient used:

- **C18**
  - Column: Kinetex C18 (2.1 x 100 mm, 2.6 µm), from Phenomenex
  - Mobile Phase:
    - A: 0.1% formic acid in water
    - B: 0.1% formic acid in methanol
  - Gradient used:

- MS Detection: ESI+
Chromatographic behaviour of the DKP standards

|     | Ala | Arg | Asn | Asp | Cys | Glu | Gln | Gin | Gly | His | Ile | Leu | Lys | Met | Phe | Pro | Ser | Thr | Trp | Tyr | Val |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     |     |     |     | HILIC |     |     |     |     |     |     | C18 |     |     |     |     |     |     |     |     |     |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |

17 DKPs are separated on both columns

HILIC: Observed on the HILIC column (n=40)
C18: Observed on the C18 column (n=35)
Both: Observed on both columns (n=17)
92 DKP standards were used to build the predictive model

Cyclo(Ala-Ala), cyclo(Ala-Gly), cyclo(Ala-His), cyclo(Ala-Ile), cyclo(Ala-Leu), cyclo(Ala-Lys), cyclo(Ala-Met), cyclo(Ala-Phe), cyclo(Ala-Ser), cyclo(Ala-Thr), cyclo(Ala-Trp), cyclo(Ala-Val), cyclo(Asn-Gly), cyclo(Asn-Lys), cyclo(Asn-Ser), cyclo(Asn-Thr), cyclo(Asp-Asp), cyclo(Asp-Gly), cyclo(Asp-Met), cyclo(Asp-Phe), cyclo(Asp-Ser), cyclo(Asp-Thr), cyclo(Asp-Trp), cyclo(Asp-Tyr), cyclo(Glu-Glu), cyclo(Glu-Gly), cyclo(Glu-Leu), cyclo(Glu-Lys), cyclo(Glu-Met), cyclo(Glu-Phe), cyclo(Glu-Trp), cyclo(Glu-Val), cyclo(Gln-Leu), cyclo(Gln-Phe), cyclo(Gln-Pro), cyclo(Gln-Ser), cyclo(Gln-Thr), cyclo(Gln-Trp), cyclo(Gly-Gly), cyclo(Gly-His), cyclo(Gly-Leu), cyclo(Gly-Lys), cyclo(Gly-Phe), cyclo(Gly-Pro), cyclo(Gly-Ser), cyclo(Gly-Trp), cyclo(Gly-Val), cyclo(His-Ile), cyclo(His-Phe), cyclo(His-Thr), cyclo(His-Trp), cyclo(Ile-Ile), cyclo(Ile-Lys), cyclo(Ile-Met), cyclo(Ile-Phe), cyclo(Ile-Pro), cyclo(Ile-Thr), cyclo(Ile-Trp), cyclo(Ile-Tyr), cyclo(Leu-Met), cyclo(Leu-Phe), cyclo(Leu-Pro), cyclo(Leu-Ser), cyclo(Leu-Tyr), cyclo(Leu-Val), cyclo(Lys-Thr), cyclo(Lys-Trp), cyclo(Met-Met), cyclo(Met-Phe), cyclo(Met-Pro), cyclo(Met-Trp), cyclo(Met-Tyr), cyclo(Met-Val), cyclo(Phe-Phe), cyclo(Phe-Pro), cyclo(Phe-Ser), cyclo(Phe-Trp), cyclo(Phe-Tyr), cyclo(Phe-Val), cyclo(Pro-Trp), cyclo(Pro-Tyr), cyclo(Pro-Val), cyclo(Ser-Ser), cyclo(Ser-Tyr), cyclo(Thr-Trp), cyclo(Trp-Val), cyclo(Tyr-Val), cyclo(Val-Val), Cyclo(Pro-Thr), cyclo(Leu-Trp), cyclo(Ala-Pro), cyclo(Thr-Thr).
CID MS/MS allow the unambiguous characterization of the DKPs

Cyclo (Pro-Val)

RT = 3.99 min

CID of $m/z$ 197.1284 [M+H]$^+$

**MS data acquisition:**
- Full scan: $m/z$ 25-1200
- CID: 20 eV collision energy (peak width 4 amu: isotopic pattern)
- MS/MS spectra of 3 parent compounds (most intense MS response)

**Mass accuracy**
0.69 ppm
CID MS/MS also allows the unambiguous detection of the constitutional isomers

$m/z$ 171.1124 [M+H]$^+$

RP18 column

(immonium ions)
Physico-chemical properties were used as discrimination factors to build the predictive model

• DKPs are characterized in terms of:
  – Elementary molecular formula: # of C, H, N, O
  – Functional groups
    • Example: -NH2 = 0 not present
    • -NH2 = 1 one group
    • -NH2 = 2 two groups…..etc
  – Ratio of C/H, N/C, O/C, S/C
  – Hydrophobicity index at pH 7 and log P

• A predictive model is built by regression analysis of known retention times against characterized DKPs

Retention time = f (C,H,N,O, functional groups, C/H,N/C, O/C, …)
Supervised statistical methods were generated using SIMCA+ software (version 12.0.1, Umetrics, Umea, Sweden). A Partial Least Square Discriminant Analysis (PLS-DA) plot was generated to optimize the correlation between observed retention times and physicochemical parameters. From the score and loading plots, the models were then validated using the classical “leave-one-out” statistical approach. This iteration was extensively repeated to ensure the robustness of the model. Once the validation was completed, the expected retention times of the not available diketopiperazines were calculated from the correlation equation (using defined physico-chemical data).
The HILIC correlation model

The model is very strong resulting in a slope of 1 and $R^2 = 0.9755$ as coefficient of correlation.

The presence of amines and carboxylic acid moieties in the DKP play a major role in the retention.
The C18 model requires a split due to the gradient applied.

Monitoring: UPLC-ESI-TOF-MS(ESI)
C18 split model shows that hydrophilic moieties in the DKP play a major role in the retention time.

\[ \text{RT} = 0-5 \text{ min} \]

(RMSEE = root mean square error of estimation)

\[ \text{RT} = 5-16 \text{ min} \]
DKPs found in Coffee on C18: In red the DKPs identified using the model

Identification criteria:
• Mass accuracy: \([M+H]^+\) (< 10 ppm)
• Mass accuracy: Immonium ion
• Expected RT within min/max shift

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Cyclo(Glu-Glu)</td>
<td>16. Cyclo(Pro-Val)</td>
<td>31. Cyclo(Leu-Val)</td>
</tr>
<tr>
<td>2. Cyclo(Ala-Ala)</td>
<td>17. Cyclo(t) (Pro-Val)</td>
<td>32. Cyclo(Phe-Val)</td>
</tr>
<tr>
<td>3. Cyclo(Gly-Pro)</td>
<td>18. Cyclo(Ala-Ile)</td>
<td>33. Cyclo(Leu-Leu)</td>
</tr>
<tr>
<td>4. Cyclo(Pro-Thr)</td>
<td>19. Cyclo(Glu-Leu)</td>
<td>34. Cyclo(Ile-Leu)</td>
</tr>
<tr>
<td>5. Cyclo(Glu-Pro)</td>
<td>20. Cyclo(Glu-Phe)</td>
<td>35. Cyclo(Phe-Trp)</td>
</tr>
<tr>
<td>7. Cyclo(Glu-Met)</td>
<td>22. Cyclo(Glu-Ile)</td>
<td>37. Cyclo(Ile-Ile)</td>
</tr>
<tr>
<td>8. Cyclo(Ala-Val)</td>
<td>23. Cyclo(Ala-Phe)</td>
<td>38. Cyclo(Ile-Phe)</td>
</tr>
<tr>
<td>9. Cyclo(Gly-Leu)</td>
<td>24. Cyclo(Tyr-Val)</td>
<td></td>
</tr>
<tr>
<td>10. Cyclo(Gin-Ile)</td>
<td>25. Cyclo(Ile-Pro)</td>
<td></td>
</tr>
<tr>
<td>11. Cyclo(Phe-Ser)</td>
<td>26. Cyclo(t) (Ile-Pro)</td>
<td></td>
</tr>
<tr>
<td>12. Cyclo(Glu-Val)</td>
<td>27. Cyclo(t) (Leu-Pro)</td>
<td></td>
</tr>
<tr>
<td>13. Cyclo(t) (Pro-Tyr)</td>
<td>28. Cyclo(Leu-Pro)</td>
<td></td>
</tr>
<tr>
<td>14. Cyclo(Phe-Thr)</td>
<td>29. Cyclo(t) (Phe-Pro)</td>
<td></td>
</tr>
<tr>
<td>15. Cyclo(t) (Pro-Tyr)</td>
<td>30. Cyclo(Phe-Pro)</td>
<td></td>
</tr>
</tbody>
</table>
DKPs found in Coffee on HILIC:
In red the DKPs identified using the model

<table>
<thead>
<tr>
<th>#</th>
<th>DKP</th>
<th>RT (min)</th>
<th>RT (Tgt) Standard</th>
<th>RT Diff (Std) (min)</th>
<th>RT (Tgt) Model</th>
<th>RT Diff (Tgt) (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.93</td>
<td>1.62</td>
<td>0.31</td>
<td>1.65</td>
<td>+0.28</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2.01</td>
<td>-</td>
<td>-</td>
<td>0.09</td>
<td>+1.92</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2.07</td>
<td>2.03</td>
<td>0.04</td>
<td>2.31</td>
<td>-0.24</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3.40</td>
<td>-</td>
<td>-</td>
<td>3.77</td>
<td>-0.37</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>3.54</td>
<td>3.52</td>
<td>0.02</td>
<td>3.66</td>
<td>-0.12</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3.99</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>4.25</td>
<td>4.30</td>
<td>-0.05</td>
<td>3.81</td>
<td>+0.44</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>5.79</td>
<td>-</td>
<td>-</td>
<td>5.30</td>
<td>+0.49</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>5.95</td>
<td>4.97</td>
<td>0.98</td>
<td>5.33</td>
<td>+0.62</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>9.01</td>
<td>-</td>
<td>-</td>
<td>8.58</td>
<td>+0.43</td>
<td></td>
</tr>
</tbody>
</table>
DKPs found in Cocoa on HILIC: In red the DKPs identified using the model

1. Cyclo(Asn-Pro)
2. Cyclo(Gly-Pro)
3. Cyclo(t)(Pro-Thr)
4. Cyclo(Ala-Pro)
5. Cyclo(Pro-Thr)
6. Cyclo(Asn-Leu)
7. Cyclo(Glu-Pro)
8. Cyclo(t)(Ala-Pro)
9. Cyclo(Asn-Ile)
10. Cyclo(Ala-Val)
11. Cyclo(Phe-Ser)
12. Cyclo(Pro-Tyr)
13. Cyclo(Pro-Val)
14. Cyclo(Ala-Trp)
15. Cyclo(Glu-Leu)
16. Cyclo(Ala-Ile)
17. Cyclo(Glu-Ile)
18. Cyclo(Ala-Phe)
19. Cyclo(Ala-Leu)
20. Cyclo(Ala-Phe)
21. Cyclo(Tyr-Val)
22. Cyclo(Ile-Pro)
23. Cyclo(Leu-Pro)
24. Cyclo(Val-Val)
25. Cyclo(Phe-Pro)
26. Cyclo(Trp-Val)
27. Cyclo(Leu-Val)
28. Cyclo(Phe-Ile)
29. Cyclo(Phe-Trp)
30. Cyclo(Ile-Ile)
31. Cyclo(Leu-Phe)
32. Cyclo(Ile-Phe)
33. Cyclo(Phe-Phe)
DKPs found in Cocoa on C18:
In red the DKPs identified using the model

1. Cyclo (Ile-Met) 12. Cyclo (Gln-Phe)
2. Cyclo (Phe-Tyr) 13. Cyclo (Asn-Ser)
3. Cyclo (Pro-Tyr) 14. Cyclo (Met-Ser)
4. Cyclo (Phe-Thr) 15. Cyclo (Gln-Ser)
5. Cyclo (Ala-Trp) 16. Cyclo (His-Pro)
6. Cyclo (Ala-Pro) 17. Cyclo (Ala-His)
7. Cyclo (Asn-Ile) 18. Cyclo (Glu-Ile)
8. Cyclo (Pro-Ser) 19. Cyclo (Glu-Leu)
9. Cyclo (Asn-Leu) 20. Cyclo (Lys-Phe)
10. Cyclo (Asn-Pro) 21. Cyclo (Lys-Pro)
11. Cyclo (Ala-Asn) 22. Cyclo (Ala-Lys)
Conclusions

• 92 DKPs analyzed by UPLC-TOF-MS
  – 57 identified on the HILIC column
  – 52 identified on the C18 column
  – 17 DKP’s were identified in both chromatographic systems

• A predictive model was built by regression analysis of known retention times against characterized DKPs, i.e.:
  – Retention time = f (C,H,N,O, functional groups, C/H,N/C, O/C,…)
  – Model showed good correlation on both chromatographic systems

• The predictive tool can be used to monitor DKPs in food products, incl. unknown DKPs, under similar analytical conditions

• The DKP fingerprint may correspond to process conditions and product attributes
Thanks to you for your kind attention
RT shifts of observed coffee and cocoa samples ... 

... as compared to the ones obtained during the analysis of pure standard DKPs and of the data calculated from the prediction models

<table>
<thead>
<tr>
<th></th>
<th>Observed RT shift from standard DKPs (n=92)</th>
<th>Coffee</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># DKPs</td>
<td></td>
<td>Compared to standard DKP</td>
<td>Compared to predicted models</td>
<td>Compared to standard DKP</td>
<td>Compared to predicted models</td>
</tr>
<tr>
<td>C18</td>
<td>Minimum</td>
<td>-2.11 min (n=52)</td>
<td>-0.07 min (n=27)</td>
<td>-1.19 min (n=33)</td>
<td>-0.17 min (n=26)</td>
<td>-1.74 min (n=31)</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>+1.22 min (n=52)</td>
<td>+0.75 min (n=27)</td>
<td>+1.05 min (n=33)</td>
<td>+0.41 min (n=26)</td>
<td>+0.87 min (n=31)</td>
</tr>
<tr>
<td></td>
<td>Mean RT&lt;sub&gt;shift&lt;/sub&gt;&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.46 min (n=52)</td>
<td>0.06 min (n=27)</td>
<td>0.37 min (n=33)</td>
<td>0.23 min (n=26)</td>
<td>0.55 min (n=31)</td>
</tr>
<tr>
<td>HILIC</td>
<td>Minimum</td>
<td>-1.39 min (n=57)</td>
<td>-0.05 min (n=5)</td>
<td>-0.37 min (n=9)</td>
<td>-0.53 min (n=11)</td>
<td>-0.91 min (n=22)</td>
</tr>
<tr>
<td></td>
<td>Maximum</td>
<td>+2.59 min (n=57)</td>
<td>+0.98 min (n=5)</td>
<td>+1.92 min (n=9)</td>
<td>+0.50 min (n=11)</td>
<td>+1.81 min (n=22)</td>
</tr>
<tr>
<td></td>
<td>Mean RT&lt;sub&gt;shift&lt;/sub&gt;&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.41 min (n=57)</td>
<td>0.28 min (n=5)</td>
<td>0.55 min (n=9)</td>
<td>0.14 min (n=11)</td>
<td>0.19 min (n=22)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Mean value RT shift has been calculated from absolute values.
cis / trans epimerisation in coffee