Predicting Peptide Structural Feasibility with CNNs

Owen Krueger1, Mohit Malu3, Rebecca Sutcliffe4, Nicholas Dunne2, You Zhou1, Giulia Pedrielli1, Andreas Spanias3

1. SCAI at Arizona State University 2. Biodesign Europe at Dublin City University 3. School of ECEE at Arizona State University 4. Queens University Belfast

Project Goal
Predict the structural feasibility of peptides utilizing CNNs

Task 1: Training Set Creation
Create a labeled training set by comparing the structural similarity of generated peptides to their real counterparts

Motivation
Provide researchers with a predictive tool to accelerate the analysis of peptide model stability.

Sequence:
SPLGEEMRDRARAHVDALRTHLAPYS DELRQRGLARLEALKENGGA

Preliminary Results

<table>
<thead>
<tr>
<th>Peptide ID</th>
<th>Cosine Similarity</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shu3</td>
<td>0.9974</td>
<td>Positive</td>
</tr>
<tr>
<td>N7I</td>
<td>0.9913</td>
<td>Positive</td>
</tr>
<tr>
<td>e0a</td>
<td>0.2074</td>
<td>Negative</td>
</tr>
<tr>
<td>M8l</td>
<td>0.2635</td>
<td>Negative</td>
</tr>
<tr>
<td>1CEU</td>
<td>0.3680</td>
<td>Negative</td>
</tr>
<tr>
<td>1Hzn</td>
<td>0.4200</td>
<td>Negative</td>
</tr>
<tr>
<td>Gw4</td>
<td>0.1675</td>
<td>Negative</td>
</tr>
<tr>
<td>1Loi</td>
<td>0.1848</td>
<td>Negative</td>
</tr>
</tbody>
</table>

Sources
4. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. David Weininger Journal of Chemical Information and Computer Sciences 1988 28 (1), 31-36 DOI: 10.1021/ci00057a005

Task 2: Adapt NU-Resnet for Peptide Sequences
Adapt current RNA based CNN to take in and predict peptide sequences

Task 3: Adapt NUMO-Resnet to utilize Peptide Motifs
Implement motif input data for NUMO-Resnet CNN to utilize

Future Key Tasks
- Task 2: Adapt NU-Resnet for Peptide Sequences
- Task 3: Adapt NUMO-Resnet to utilize Peptide Motifs