Glycomimetics as a Therapeutic Strategy

<table>
<thead>
<tr>
<th>Compound</th>
<th>$K_i$ [nM]</th>
<th>Relative activity</th>
<th>$R^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>69000</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>43000</td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>2a</td>
<td>6700</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>2b</td>
<td>25000</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>2c</td>
<td>11000</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>2d</td>
<td>9500</td>
<td>73</td>
<td></td>
</tr>
<tr>
<td>4a</td>
<td>3000</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>4b</td>
<td>61</td>
<td>1130</td>
<td></td>
</tr>
<tr>
<td>4c</td>
<td>50</td>
<td>1380</td>
<td></td>
</tr>
<tr>
<td>4d</td>
<td>33</td>
<td>2090</td>
<td></td>
</tr>
</tbody>
</table>

- Exploits the specificity of the endogenous carbohydrate ligand
- Employ the native carbohydrate ligand as a basis for rational design
- Examples: Relenza® and Tamiflu®

How to choose the “$R$” groups?

Cumstey et al. Angew. Chem. Int. Ed. 44, 5110-5112
Many pathogens exploit binding to glycans as the first stage of infection:

- Influenza, parainfluenza, mumps, corona, noro, rota, and DNA tumor viruses,
- chikungunya (?), zika (?), many bacteria

Can We Block Pathogen Adhesion with Glycomimetics?

How do we make the glycomimetic bind more strongly to the protein receptor?

• Stronger binding – smaller dose, higher efficiency, less change of side effects

Make chemical modifications that create new contacts between the small molecule and the protein surface “rational design”
Molecular Dynamics Can Discriminate Good from Bad Binders

But – which inhibitors should we simulate?

Putative Inhibitor 1

Putative Inhibitor 2
Graft R groups onto carbohydrates and conformational sampling.

- All rotatable bonds in the chemical moiety are identified and rotated.
- A genetic algorithm is employed for conformational sampling.
- \[ \Delta G_{\text{rotamer}} = \Delta G_{\text{Vina}} + \Delta G_{\text{CH–π}} \]
Virtual Glycomimetic Screening

1. Upload protein-carbohydrate PDB file
2. Select position in ligand to modify
3. Select library of analogs to screen

4. Perform conformational search for each analog
   - rotamer generation
   - clash checking
   - energy minimization

5. Pass MD snapshots at 5 ns intervals

6. Perform MD of receptor-carbohydrate co-complex
   - explicit water
   - AMBER GAFF force field
   - 50 ns

7. Calculate GIST desolvation energy and docking scoring for all rotamers of each analog in each MD snapshot

8. Compute ligand conformational entropy

9. Return best pose and estimated binding free energy for each analog
Virtual Glycomimetic Screening

Moiety Library:
- Designed for addition to NH/OH atoms in carbohydrates
- Moieties scraped from chemical catalogs and PubChem
- Converted from SMILES to 3D
- Currently ~600 moieties
Virtual Glycomimetic Screening

Binding Energies:
- Multiple methods implemented
  - AutoDock VINA-Carb [1]
  - AMBER/GLYCAM MM-GBSA
  - AMBER/GLYCAM MM-GIST
- New functional forms added (CH-π)

Proposed Online Webtool

Step 1: Upload/Fetch PDB

Goals:
- Provide an online platform for non-specialists
- User uploads a protein-carbohydrate complex
- We perform a free Library screening, MD simulation, and binding energy prediction
- User downloads the predicted best ligand structures and energies
Proposed Online Webtool

Step 2: Choose options

We detected the following positions available for modification:

<table>
<thead>
<tr>
<th>Residue Index</th>
<th>Atom Name</th>
<th>Atom To Replace</th>
<th>Select Position</th>
<th>Select R Group Library</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td>O3</td>
<td>HO3</td>
<td></td>
<td>Aldehydes: 0</td>
</tr>
<tr>
<td>112</td>
<td>N2</td>
<td>C7</td>
<td>×</td>
<td>Sulfonyl Halides: 0</td>
</tr>
<tr>
<td>113</td>
<td>O6</td>
<td>HO6</td>
<td></td>
<td>Ketones: 0</td>
</tr>
<tr>
<td>114</td>
<td>O1</td>
<td>HO1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Submit
Step 3: Download Results

Top 5 Analogs

<table>
<thead>
<tr>
<th>R Group</th>
<th>Rank</th>
<th>Affinity Change (kcal/mol)</th>
<th>RIP*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>-5.67</td>
<td>12,708</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-4.56</td>
<td>1,998</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>-2.34</td>
<td>49.4</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>-1.23</td>
<td>7.77</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>-0.53</td>
<td>2.42</td>
</tr>
</tbody>
</table>

*RIP = relative inhibitory potential
Virtual Library Creation: Automated Ligand Parameterization

**GAFF Forcefield**
- R group
  - ESP Calculation (QM)
  - RESP Charge Fitting
- Glycan
  - Subgraph Isomorphism Matching
  - Force Field Parameter Assignment

**Forcefield Mixing**
- New Bond, Angle, and Torsional Terms
- Charge Adjustment

**Glycomimetic Analog**
Automated MD Simulation

Preprocessing

- Remove Co-crystallization Reagents
- Detect Disulfide Bonds
- Detect Missing Segments

Glycomimetic Ligand

- Repair (Modeler, AlphaFold, etc)
- Place High Restraints

Simulation

- Heating, Equilibration, MD
- Input Files Created Automatically Without Human Intervention
MD Post-processing and Analysis

- Extract Representative Structures
- Binding Energy Calculation
- Other Advanced Analyses
- MM-PB/GBSA
- Autodock Vina Scoring Function
- Develop New Force Field Terms?
- CH-π
CH-π Interactions in Carbohydrate Binding

PDB 2UVO
Wheat Germ Agglutinin + GlcNAc
A Molecular Mechanical Model for CH-π Interaction
Experimental CH-π Interactions

- All the CH bonds on top aromatic amino acids.
- Diffuse distribution with the average matching the canonical geometry.

<table>
<thead>
<tr>
<th>n = 173</th>
<th>n = 169</th>
<th>n = 192</th>
</tr>
</thead>
<tbody>
<tr>
<td>![PHE side]</td>
<td>![PHE top]</td>
<td>![PHE avg]</td>
</tr>
<tr>
<td>![TRP side]</td>
<td>![TRP top]</td>
<td>![TRP avg]</td>
</tr>
<tr>
<td>![n = 84]</td>
<td>![HIS side]</td>
<td>![HIS top]</td>
</tr>
<tr>
<td>![HIS avg]</td>
<td>![TYR side]</td>
<td>![TYR avg]</td>
</tr>
</tbody>
</table>

Empirical CH-π Functional Form

\[ E_{CH-\pi} = \sum \left( f(\theta)E_{HC} \cdot e^{-\frac{(R_{HC}-R_{HC}^0)^2}{2C_{HC}^2}} \right) \]

- \( f(\theta) = \cos\theta \) \( (0 \leq \theta < 90^\circ) \)

- \( E_{CH-Aromatic} = E_{CH-\pi} + E_{Vina} \)

\[ \Delta E_{total} = w_{CH-\pi} \Delta E_{CH-\pi} + \Delta E_{Vina} \]

36 systems containing carbohydrate ligands (Vina-Carb[1]).

Significantly improved percentages of acceptable poses and average RMSDs.

Blind Docking Example (PDB 5V6F)

$w_{CH-\pi} = 0.0$

$w_{CH-\pi} = 0.8$

Crystal Structure of the Second beta-Prism Domain of RbmC from *V. cholerae* bound to Mannotriose
Systems Where Our Method Worked Well (DC-SIGN)

• A lectin involved in immunity. Exploited for infection by HIV³ and COVID⁴.

Mannose (control)

Representative Analog


Vina with CH-π significantly outperformed MM-GBSA in this system.

Statistical Correlation to Experimental Affinity

![Graphs showing correlation between Vina and MM-GBSA](image)

\[ y = 0.63x - 0.28 \quad \text{R}^2 = 0.52 \]

\[ y = 0.04x - 1.54 \quad \text{R}^2 = 0.07 \]

Conclusions

Glycomimetic design is amenable to automation
  ➢ Expect to see it at glycam.org in 2024

Online Webtools ensure consistency and ease of use

Predicted binding energies need to be improved
  ➢ Parameterization of existing scoring functions
  ➢ Introduction of new physics (CH-π, water, entropy)
  ➢ Need to introduce CH-π into AMBER force field for MD

Need beta test users
## Acknowledgements

<table>
<thead>
<tr>
<th>Infrastructure</th>
<th>Online Modeling Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lachele Foley</td>
<td>Dave Montgomery</td>
</tr>
<tr>
<td>Dan Wentworth</td>
<td>Spandana Makeneni</td>
</tr>
<tr>
<td>Yao Xiao</td>
<td>Amika Sood</td>
</tr>
<tr>
<td>Oliver C. Grant</td>
<td></td>
</tr>
</tbody>
</table>