RNA as drug target: docking studies

Florent Barbault
ITODYS CNRS (UMR7086) Univ. Paris7
Overview

- Why targeting RNA
- Drug design strategy
- Parametrisation of the scoring function
- First Virtual screening
- Flexibility of target
- Conclusions & Prospects
Why targeting RNA?

RNA and DNA are negatively charged molecules.

RNA structures
- single strand
  => complex 3D structures

DNA structures
- double strand
  => mainly Helix
Why targeting RNA?

Biological role of RNA:

- Protein synthesis (mRNA, Ribosome, tRNA)
- Enzyme (Ribozyme)
- RNA is the genome of all retrovirus (HIV, HCV,...)
- RNA can control gene regulation (siRNA)

We need new and original targets

*RNA can be one of them*

RNA is involved in a lot of biological functions

*Different opportunities and effects*

RNA have structural domains that are more highly conserved

*Slower development of drug resistance*

RNA is upstream in translation pathway (protein synthesis)

*Inhibiting 1 RNA (ribosome) could prevent ~1000 proteins*
Drug-design strategy

Purpose: design antimicrobial compounds

Protein synthesis scheme

DNA → mRNA → ribosome → PROTEIN

Same global scheme for procaryote & eucaryote but 2 different RNA ribosomal fragments:
  RNA 16S (bacteria)
  RNA 18S (human)

We want to design selective ligands against 16S and not 18S
Drug-design strategy

RNA ligands?

Aminoglycosides!

- Natural products
- Good affinity
- Low selectivity
- Ammonium groups
- Difficult to synthetize
- Flexible molecules

Goal:
- Take these molecules as scaffold for chemical modifications
- Add a nucleoside to make a triple-base-pair or intercalation
Drug-design strategy

Two questions:

<table>
<thead>
<tr>
<th>3D structure of the target?</th>
<th>List of compounds with bioactivity?</th>
</tr>
</thead>
<tbody>
<tr>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td></td>
<td>All</td>
</tr>
<tr>
<td></td>
<td>Docking</td>
</tr>
<tr>
<td>NO</td>
<td>QSAR-2D</td>
</tr>
<tr>
<td></td>
<td>QSAR-3D</td>
</tr>
<tr>
<td></td>
<td>Pharmacophoric screening</td>
</tr>
<tr>
<td></td>
<td>NOTHING</td>
</tr>
</tbody>
</table>
Parametrisation of the scoring function

Autodock (3.0) pre-calculate energy grid maps:
1 for each atoms + electrostatic
Ligands are flexible whereas target is rigid
Genetic Algorithm + Local Search to find ligand position
Parametrisation of the scoring function

\[ \Delta G = f_{vdw} \sum \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) + f_{hbond} \left[ \sum E(t) \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{10}} \right) + E_{Hbond} \right] + f_{elec} \sum \left( \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} \right) + \Delta G_{tors} N_{tors} + f_{solv} \sum \left( S_i V_j + S_j V_i \right) e^{-r_{ij}^2} \]

5 empirical parameters derived from protein/ligand complexes

NOT SUITABLE TO RNA! We need these 5 parameters for RNA
Parametrisation of the scoring function

We need RNA/Ligands structures with experimental $\Delta G$

*ONLY 8!*

Tobramycin with RNA aptamer I  
Tobramycin with RNA aptamer II  
Neomycin-B with RNA Tau exon  
Neomycin-B with RNA aptamer  
Neomycin-B with RNA HIV-1 Tar  
Gentamicin C1A with A-site rRNA  
Paramomycin with RNA-16S  
Acetylpromazine with RNA HIV-1 Tar

aminoglycosides
Parametrisation of the scoring function

For each structures:
   31 docking calculations with different parameters.
   Values are ranged from 0 to 2x of their default value.
   At the end: 248 results

Correct docking:
   Free energy of binding similar to the experimental
   RMSD between experimental and calculated is low

ANN is used to correlate $\Delta G$ and RMSD with parameters.

   1 hidden layers
   Back propagation algorithm
   Leave-One-Out cross-validation

Final RMS error: 0.269 and Final max error: 0.883
Parametrisation of the scoring function

\[ R^2 = 0.8679 \]

\[ R^2 = 0.9419 \]
Parametrisation of the scoring function

With the Neural Network Model

- 80,000 parameters were randomly generated and tested

- Best values were selected
  as good as possible $\Delta G$ agreement
  as weak as possible possible RMSD value

<table>
<thead>
<tr>
<th></th>
<th>Vdw</th>
<th>Elec</th>
<th>Hbond</th>
<th>Tors</th>
<th>Sol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autodock</td>
<td>0.1485</td>
<td>0.1146</td>
<td>0.0656</td>
<td>0.3113</td>
<td>0.1711</td>
</tr>
<tr>
<td>RNA-Autodock</td>
<td>0.155</td>
<td>0.101</td>
<td>0.056</td>
<td>0.361</td>
<td>0.153</td>
</tr>
</tbody>
</table>
Parametrisation of the scoring function

NeomycinB with RNA Tau 10

White: experimental
Blue: autodock
Red: RNA-autodock
Parametrisation of the scoring function

Neomycin with RNA Tar

White: experimental
Blue: autodock
Red: RNA-autodock
Parametrisation of the scoring function

Gentamycin with RNA-16S

White: experimental
Blue: autodock
Red: RNA-autodock
Parametrisation of the scoring function

Paromomycin with RNA-30S

White: experimental
Blue: autodock
Red: RNA-autodock
First Virtual Screening

With the new scoring function:
First Virtual Screening with ~300 compounds against 16S Library created in agreement with the organic chemist group

A nucleoside is linked to the aminoglycoside to form a 3\textsuperscript{rd} base pair.
First Virtual Screening

40 compounds were proposed for the synthesis
25 were synthesized (not only belonging to the 40)
17 present micromolar affinity (16&18S) & 8 no interaction.
First Virtual Screening

Back docking calculation with Autodock for 16S & 18S:
  - Understand selectivity
  - Find Structure/Activity Relationships

Results:
  - All ligands without interaction are placed at the end
  - Except Neamine (bad pose?)
  - Correlation between $\Delta G$ calculated and experimental
  - $R^2= 0.72$ for RNA-16S and $R^2=0.7$ for RNA-18S

But:
  - Unable to explain selectivity of 16S vs 18S
  - Unable to find a 3rd base-pair or intercalation

TARGET FLEXIBILITY
First Virtual Screening

After a first docking with RNA as rigid target:

Small Molecular Dynamic simulations with GB implicit solvent (AMBER 8.0)

- MD of complexes (~2 hours for each complexes)
- MD of target alone (~2 hours)
- MD of Ligands alone (~5 minutes for each ligands)

\[ E_{\text{interaction}} = E_{\text{complex}} - E_{\text{target}} - E_{\text{ligand}} \]

Force-fields: parm03 for RNA and gaff for Ligands
Flexibility of the target

- Clustering and extraction of Autodock output file
  - In-house software

- Add hydrogens and atom types
  - Sybyl 7.2

- Calculation of AM1 charges
  - Divcon

- Creation of coordinates and parameters files for AMBER MD simulation
  - In-house software

- 1000 steps of minimisation
- 20ps of MD equilibration
- 50ps of MD production
  - AMBER 8.0

- Analysis
  - VMD & ptraj
Flexibility of the target

Correlation

Experimental free energy of binding

50ps average interaction energy

R²=0.8412
Flexibility of the target

Graphical analysis

Neamine compound who was difficult to correlate with classical docking is expelled with MD.
Flexibility of the target

Graphical analysis

Formation of an AU reverse hoogsten base pair

MD permit the adaptation of the target to form a triple base-pair
Conclusions & prospects

Conclusions:

• Design of a specific set of parameters for the scoring function
• Docking is able to discriminate "good" or "bad" ligands
• Docking is limited for RNA because of induced fit
• MD can predict more efficiently free energy of binding
• MD is able to remove "false" docking position
• MD need big computation times not yet applicable for VS

Prospects:

• Finish calculation for RNA 18S
• Extract structure activity relationships
• Limiting MD computation times with SA with restraint?
Aknowledgements

Peking University CHINA, Health science center
Pr. Lihe Zhang, Pr. Liangren Zhang & PhD Ron Bo

ITODYS, CNRS UMR 7086 University Paris7 FRANCE
Group of Molecular Modeling & Chemical Information
Thank you for your attention