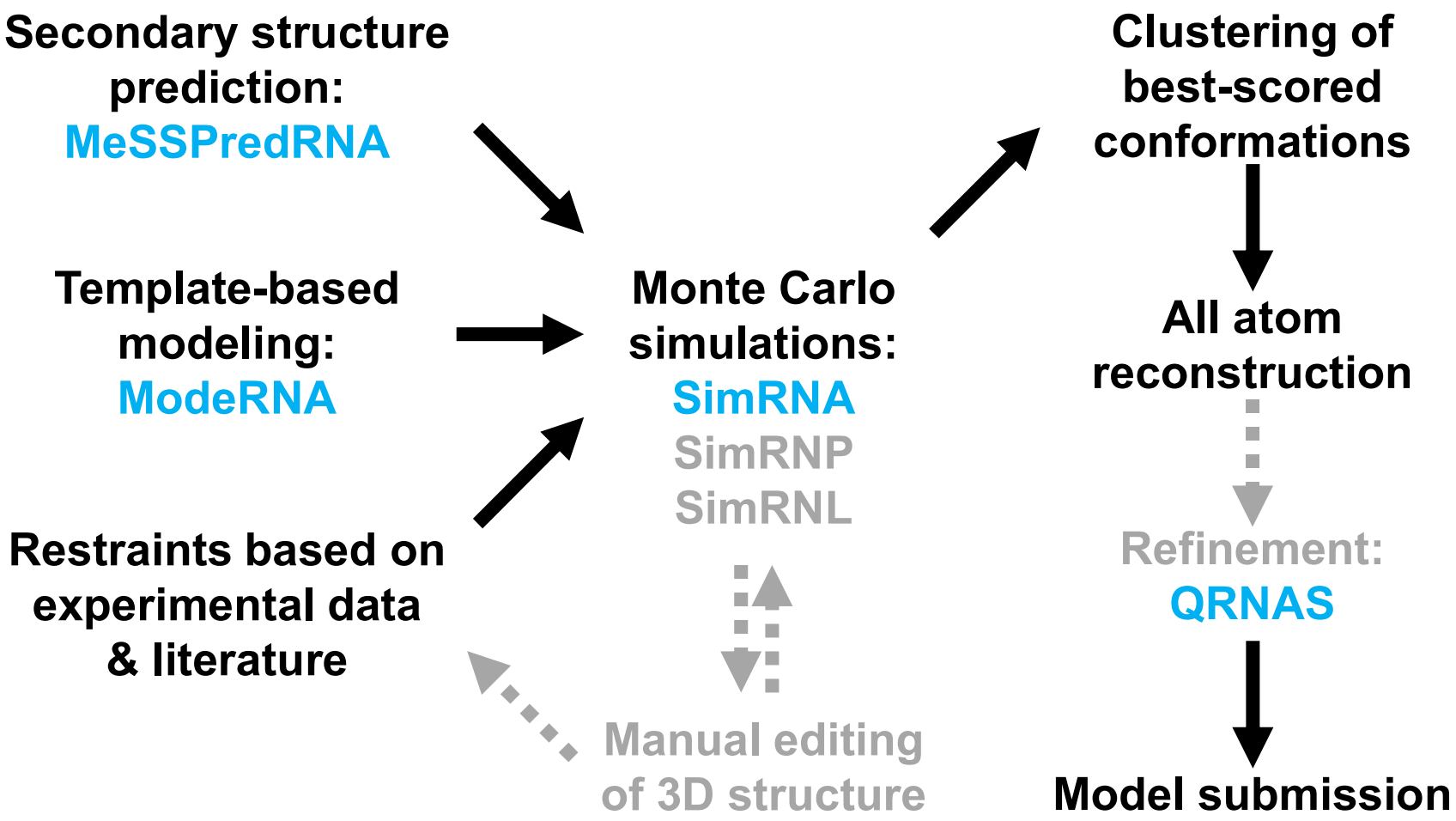
RNA 3D structure prediction
GeneSilico group (128) in CASP15

Janusz M. Bujnicki

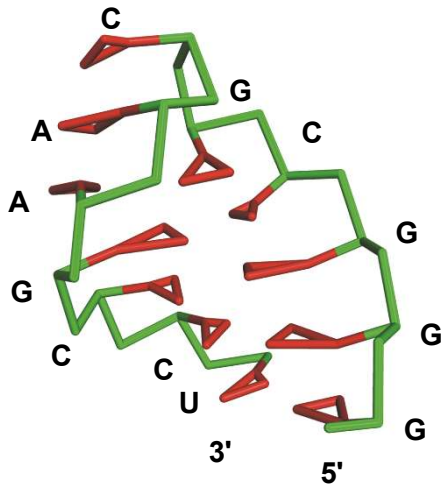
International Institute of Molecular and Cell Biology in Warsaw

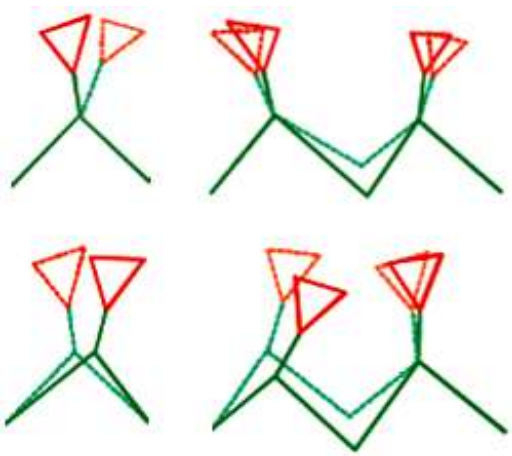
CASP15 conference, Antalya 2022.12.12

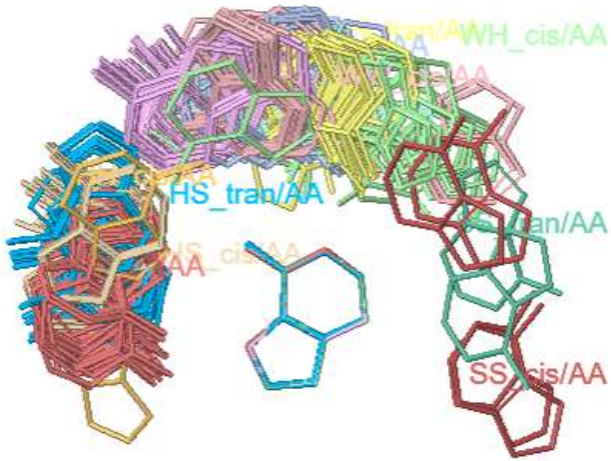
Overall methodological workflow



Central methodology:
RNA 3D structure modeling with SimRNA







Representation:
coarse-grained

3 atoms per base
2 atoms for backbone
3D grids

Move set:
random conformational changes

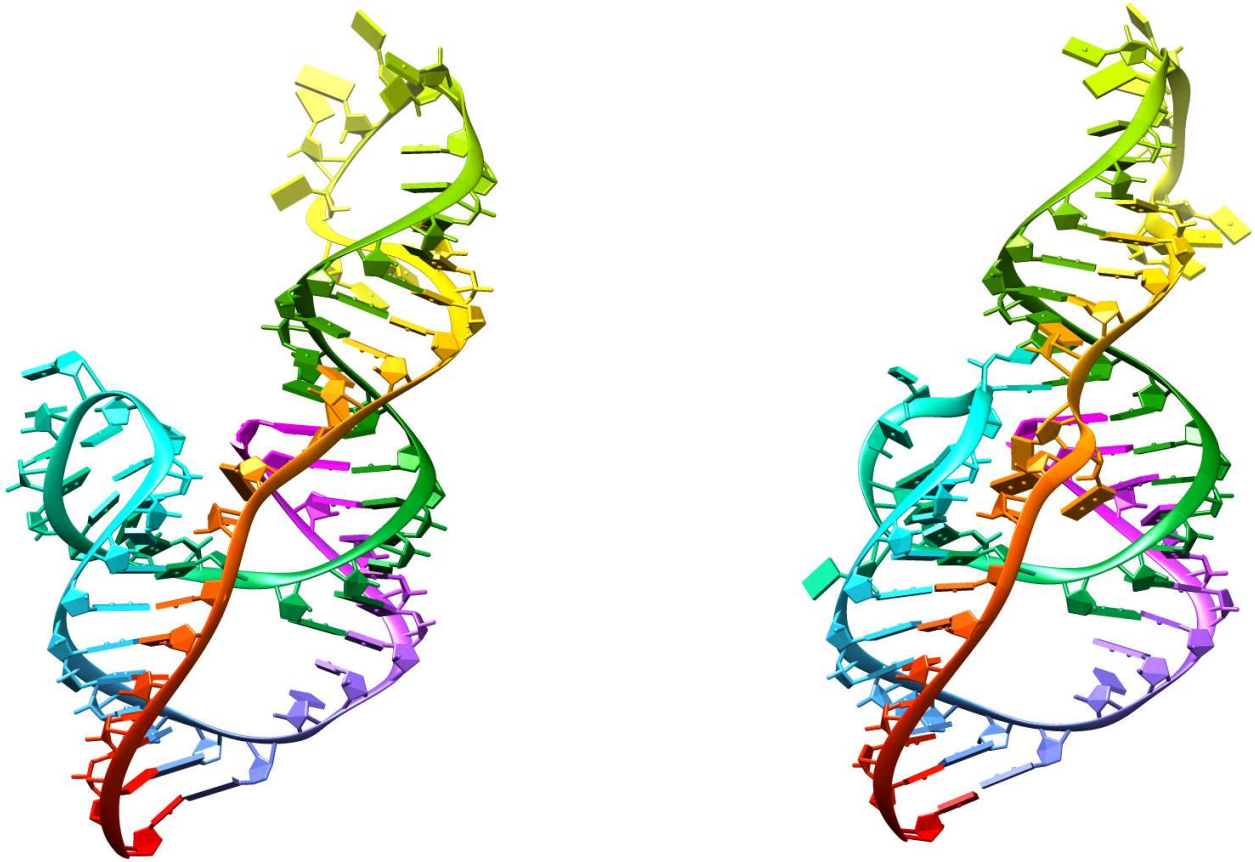
Monte Carlo approach
moves accepted or rejected
depending on energy and temperature

Scoring:
statistical potential

calculated from
experimentally determined
RNA 3D structures

Boniecki MJ, Lach G, Dawson WK, Tomala K, Lukasz P, Soltysinski T, Rother KM, Bujnicki JM
SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction
Nucleic Acids Res. 2016 Apr 20;44(7):e63.

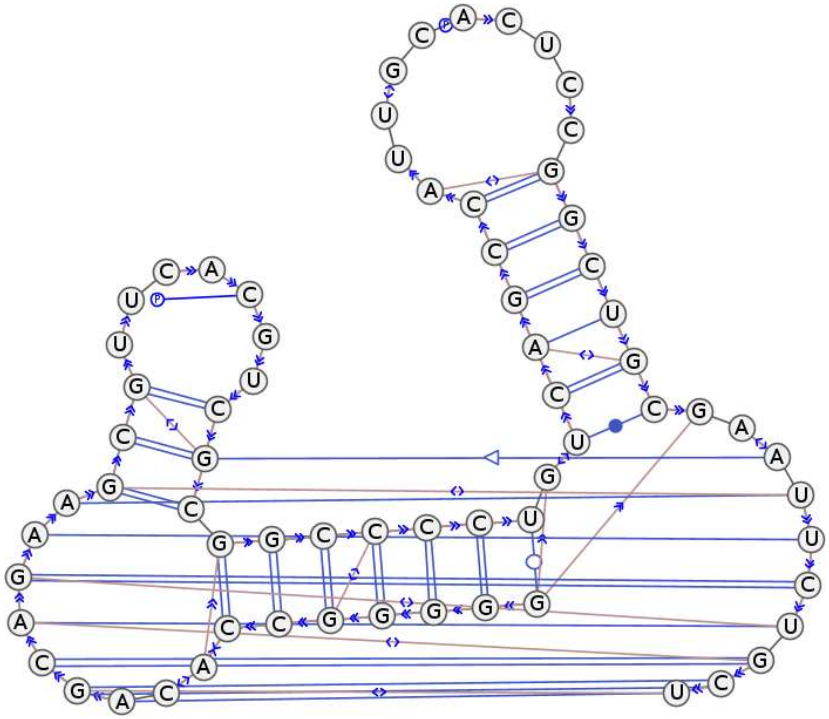
R1108 – model 3, ranked 1st

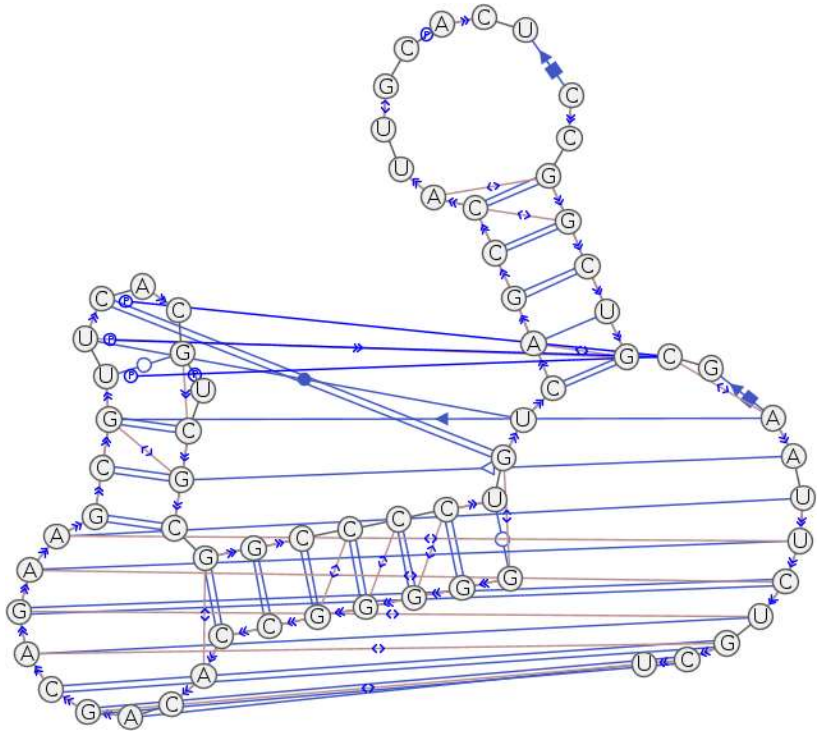


reference **model**

RMSD 5.48, LDDT 0.742, TMscore 0.544, GDT_TS 64.86

R1108 – model 3, ranked 1st





reference **model**

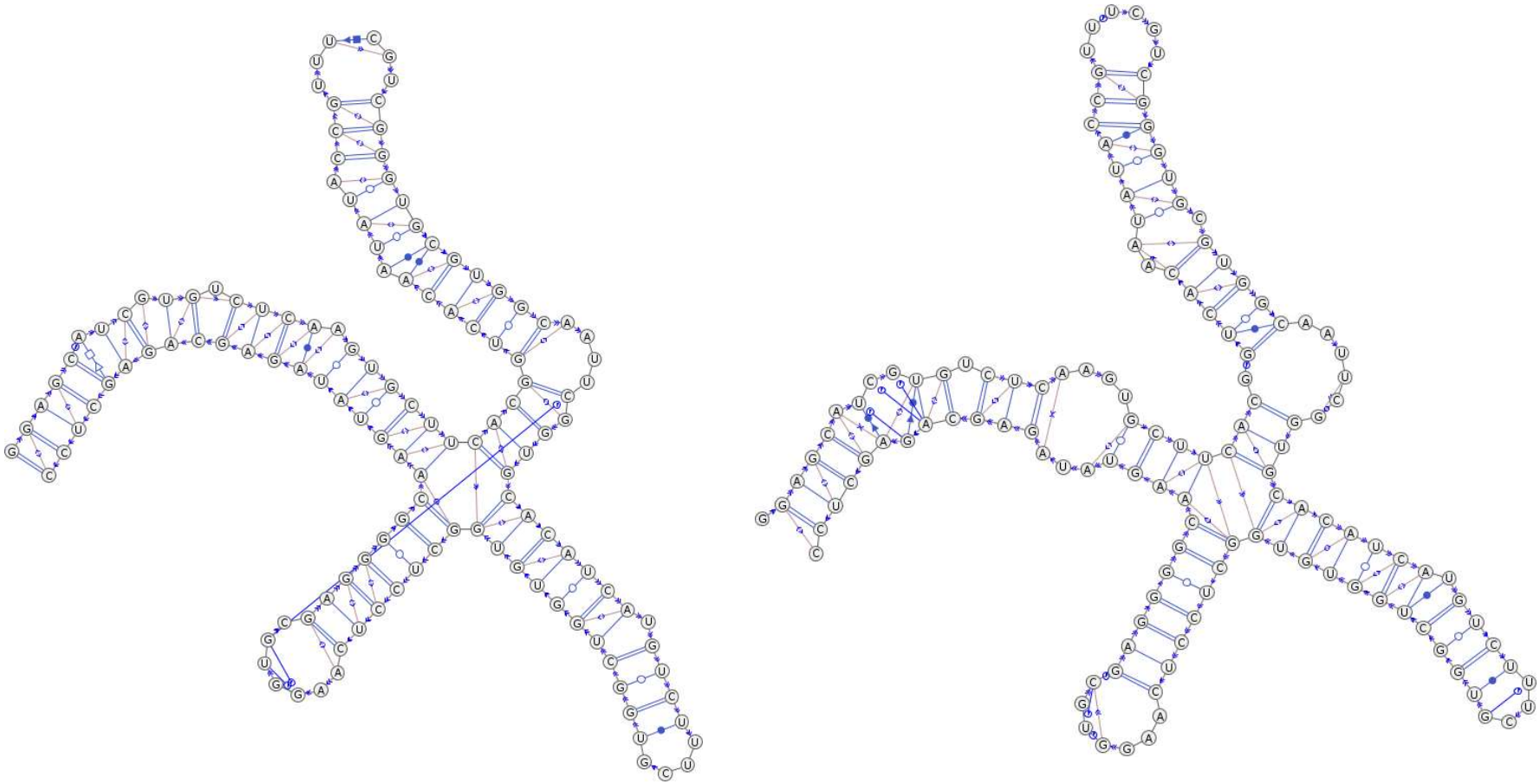
MCC: can 0.92, ncan 0.42, stack 0.74, all 0.70

R1156 – model 5, ranked 1st

reference **model**

RMSD 5.37, LDDT 0.654, TMscore 0.664, GDT_TS 46.67

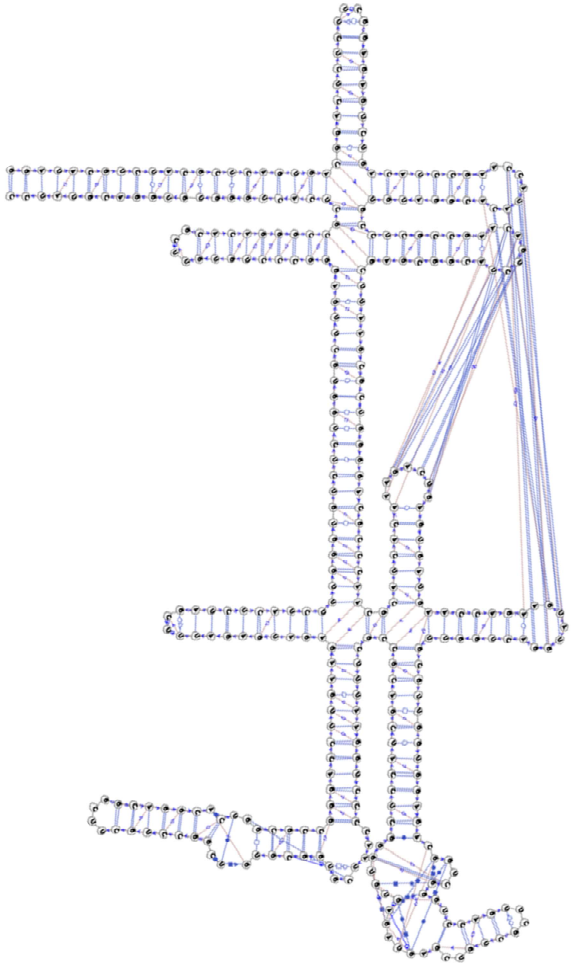
R1156 – model 5, ranked 1st

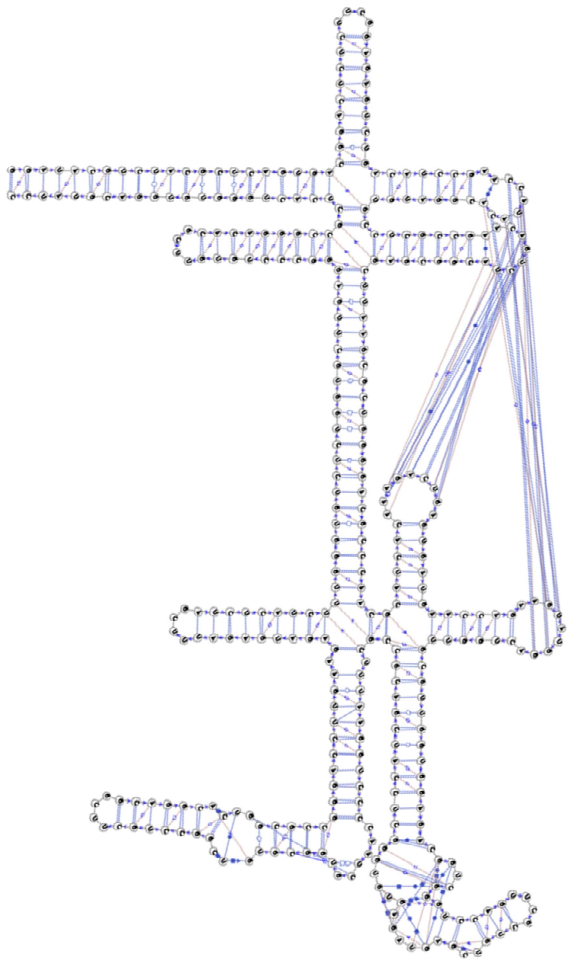


reference **model**

MCC: can 0.93, ncan 0.55, stack 0.71, all 0.64

R1136 – model 1, ranked ….

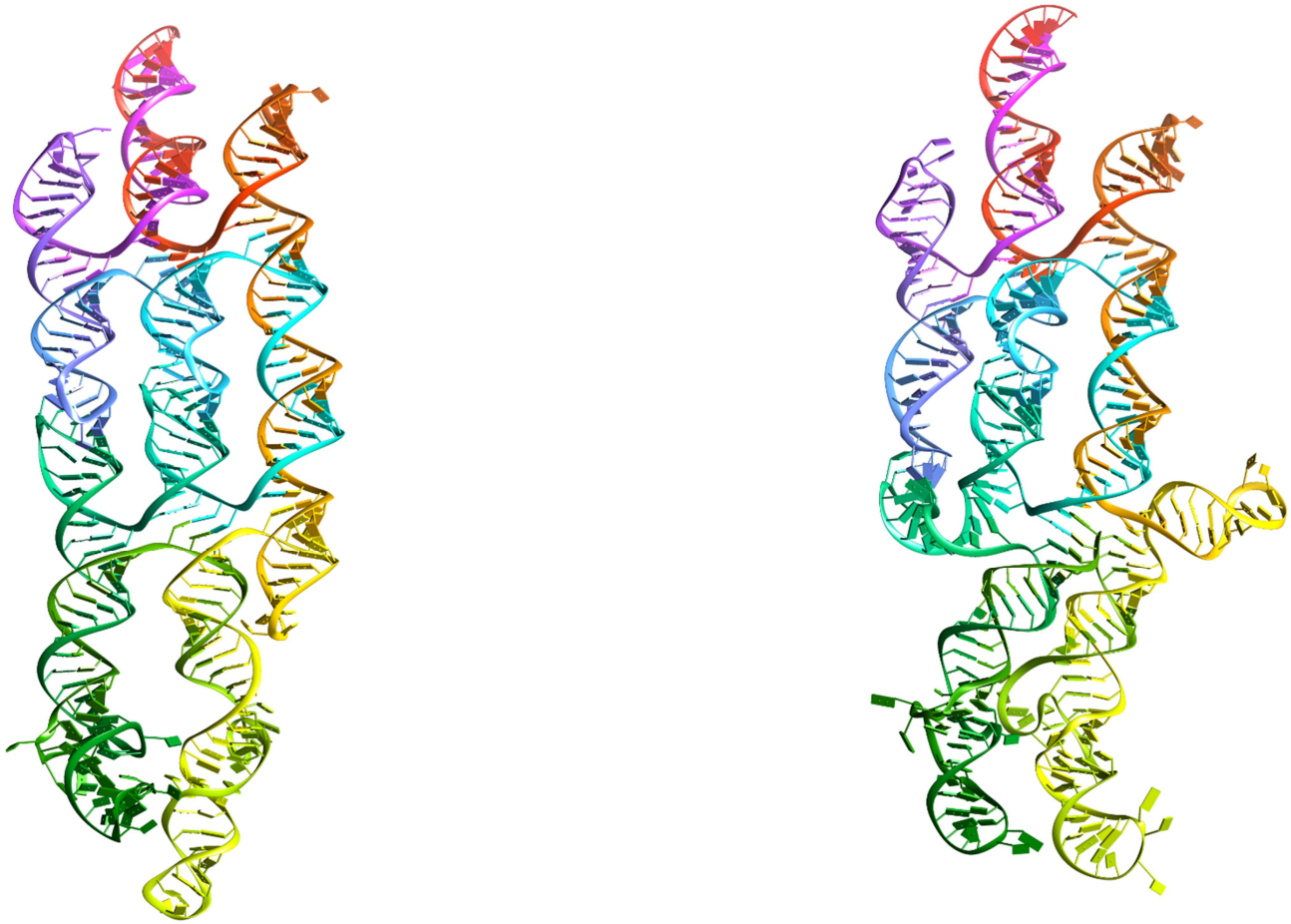




reference **model**

MCC: can 0.99, ncan 0.83, stack 0.88, all 0.86

R1136 – model 1, ranked 30th



reference model

RMSD 13.23, LDDT 0.701, TMscore 0.533, GDT_TS 20.99

What went right

Prediction of (nearly all) canonical and wobble pairs

- For all targets at least one model with $MCC_{canonical}$>0.9

- Always: Prediction of simple secondary structure elements

- Often: Prediction of pseudoknotted helices
  in particular kissing-loop interactions
  (important in artificially designed RNAs)

Prediction of stacking within helical segments

- For all targets at least one model with $MCC_{stacking}$>0.7

Sometimes:
- topology of 4-way junctions
- co-axial stacking of helices

What went wrong

Prediction of non-canonical pairs (who with whom and how)

- Long-range nc pairs are extremely hard to predict
- Average $MCC_{non\text{-}canonical}$ for our best models only 0.58
- Best prediction: $MCC_{non\text{-}canonical}$ 0.83 (R1136)
  but mostly based on short-range interactions within loops
- Successes are rare (e.g., R1117 model 1)

Prediction of stacking interactions outside of helices

- In helices and short loops it's easy, otherwise it's hard

Bending helices. In reality they break rather than bend.

Sometimes:
- topology of 4-way junctions
- co-axial stacking of helices

Acknowledgements

CASP15 modelers (GeneSilico group)

Eugene F. Baulin
Sunandan Mukherjee
Farhang Jaryani
Masoud A. Farsani
Tomasz Wirecki
S. Naeim Moafinejad
Filip Stefaniak
Nagendar G. Badepally

NARODOWE CENTRUM NAUKI

our software: http://genesilico.pl