CONTROL OF RNA FUNCTION BY CONFORMATIONAL DESIGN

CYCLING IN ENERGY AND DESIGN LANDSCAPES

Stefan Badelt

Department of Theoretical Chemistry
Theoretical Biochemistry Group (tbi)
University of Vienna

March 1st, 2016

OUTLINE

RNA modeling and RNA design

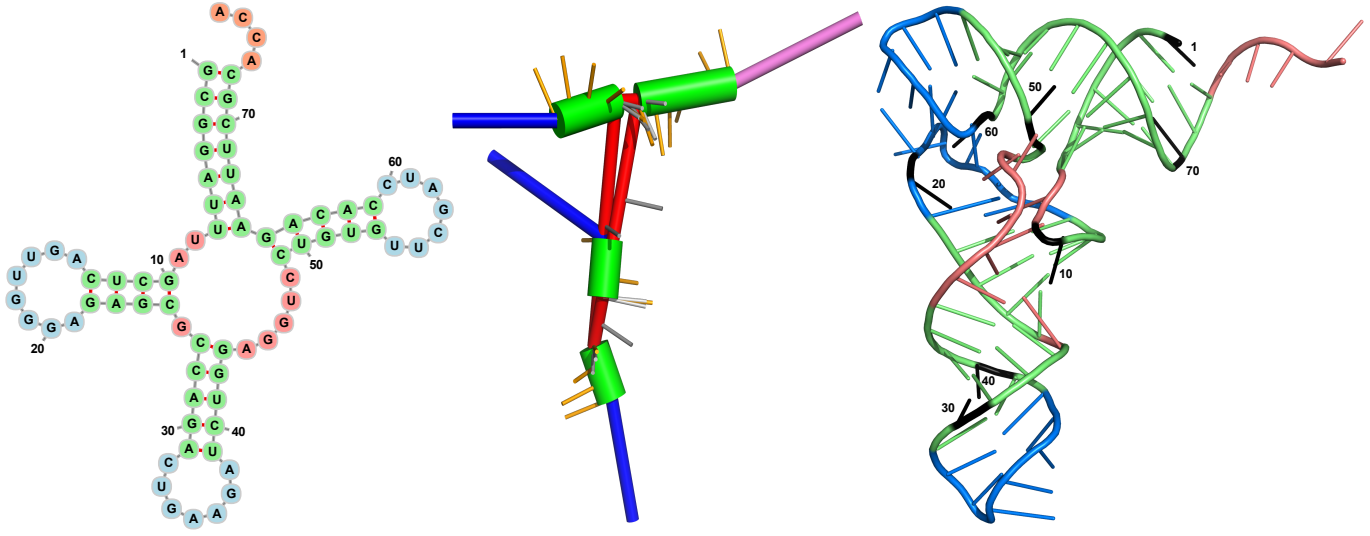
Design of RNAs with prion-like behavior

Design of self-processing ribozymes

Kinetics of RNA-RNA interactions

RNA STRUCTURE

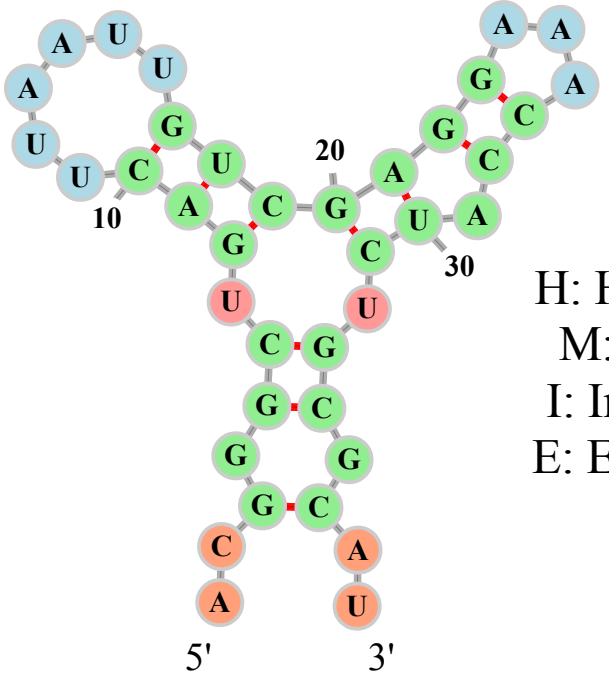
GCGGAUUUAGCUCAGUUGGGAGAGCGCCAGACUGAAGAUCUGGAGGUCCUGUGUUCGAUCCACAGAAUUCGCACCA



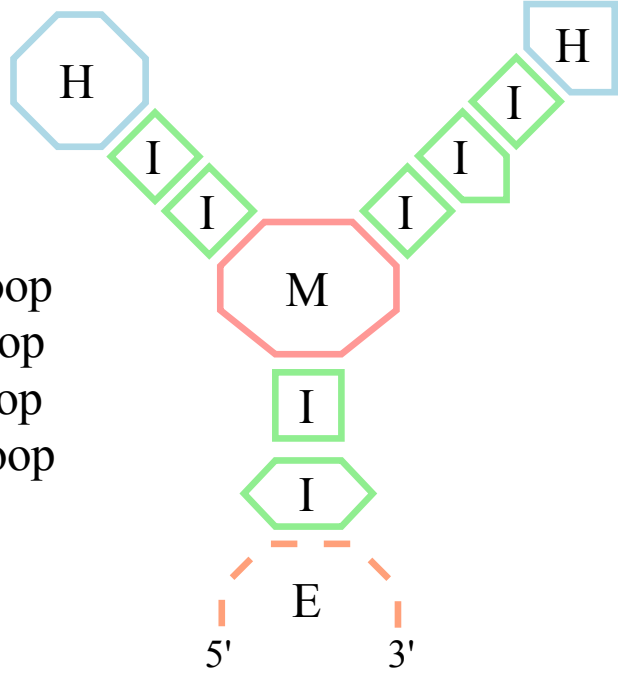
A secondary structure is a list of base pairs, where:

- A base may participate in at most one base pair
- Base pairs must not cross (no pseudoknots)
- Only isosteric base-pairs (GC, AU, GU) are allowed.

THE NEAREST NEIGHBOR ENERGY MODEL



H: Hairpin loop
M: Multi loop
I: Interior loop
E: Exterior loop



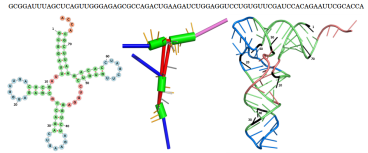
$$E(s) = \sum_{l \in s} e(l)$$

ENERGY LANDSCAPES

An energy landscape is defined by

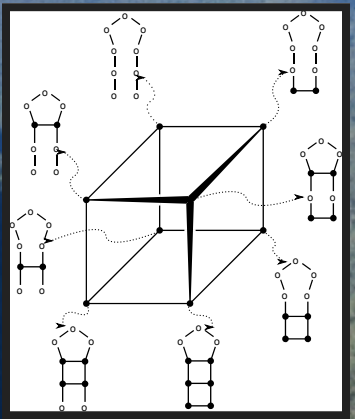
- Conformation space $s \in \Omega$
- Neighborhood relation [Move set] $M(s)$
- Energy function $E(s)$

RNA STRUCTURE REPRESENTATION

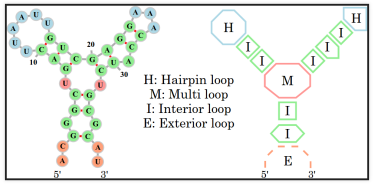


A secondary structure is a list of base pairs, where:

- A base may participate in at most one base pair
- Base pairs must not cross (no pseudoknots)
- Only isosteric base-pairs (GC, AU, GU) are allowed.

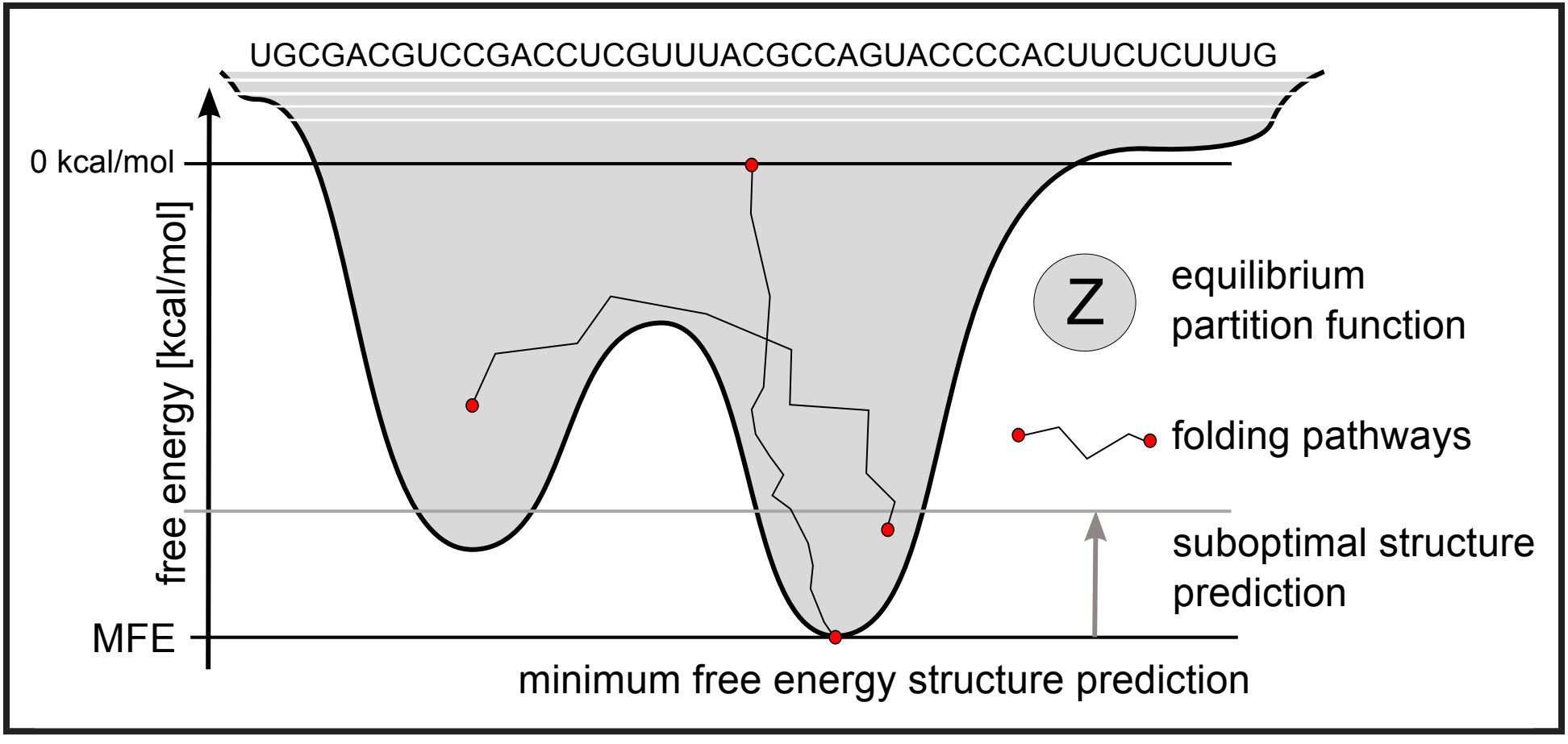


THE NEAREST NEIGHBOR ENERGY MODEL



$$E(s) = \sum_{l \in s} e(l)$$

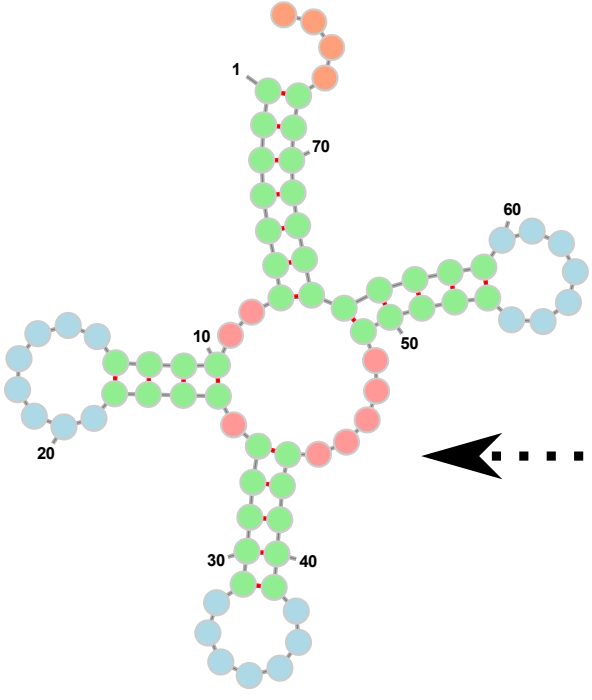
ENERGY LANDSCAPES



$$Z = \sum_{s \in \Omega} e^{\frac{-E(s)}{kT}} \qquad G = -kT \ln Z \qquad P(s) = \frac{e^{-E(s)/kT}}{Z}$$

COMPUTATIONAL RNA DESIGN

FROM STRUCTURE(S) TO SEQUENCE



GCGGAUUUAGC
UCAGUUGGGAG
AGCGCCAGACU
GAAGAUCUGGA
GGUCCUGUGUU
CGAUCCACAGA
AUUCGCACCA

$\Phi(\sigma)$

The objective function $\Phi(\sigma)$ can vary

structure is MFE of sequence, maximize probability of structure, ...

RNA DESIGN IS FORMALLY HARD, BUT EASY IN PRACTICE

sequences

(bigger)

mapping

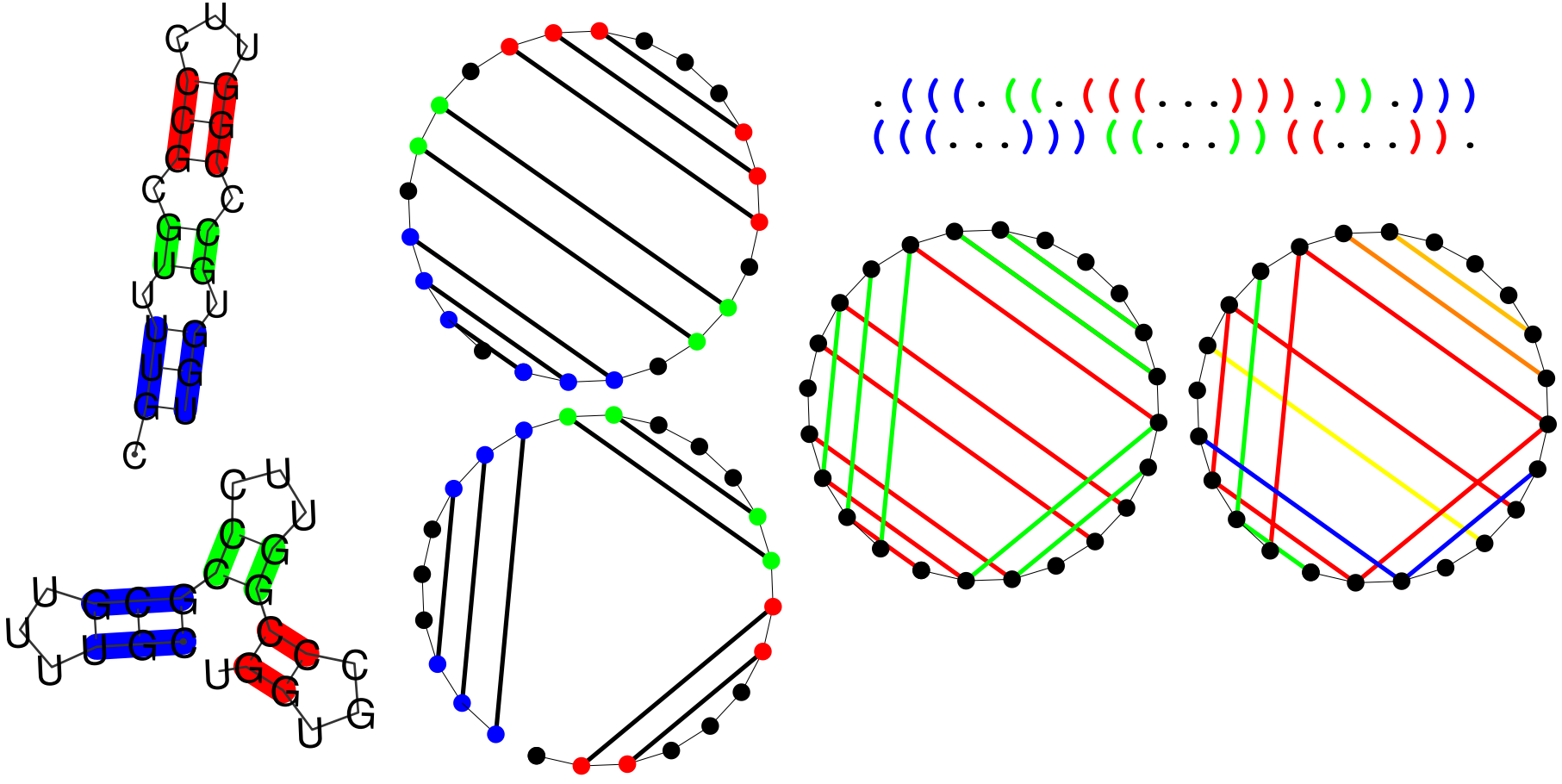
redundant,
sensitive

minimum free energy structures

(common motifs realized more often)

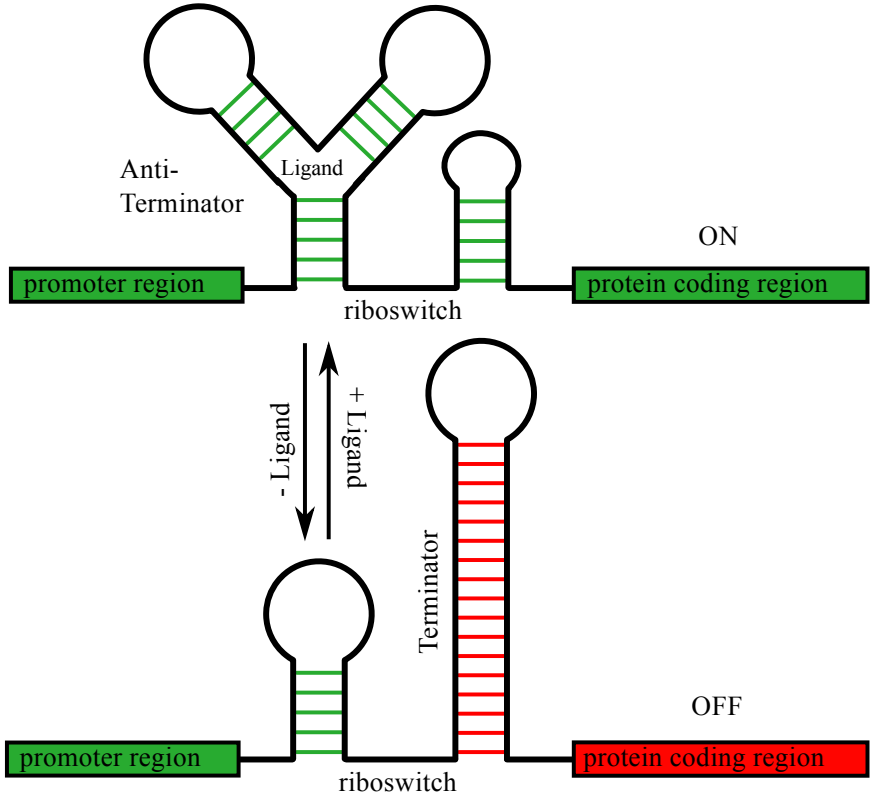
... simple adaptive walks usally lead to satisfactory solutions

DEPENDENCY GRAPHS FOR BISTABLE RNA DESIGN



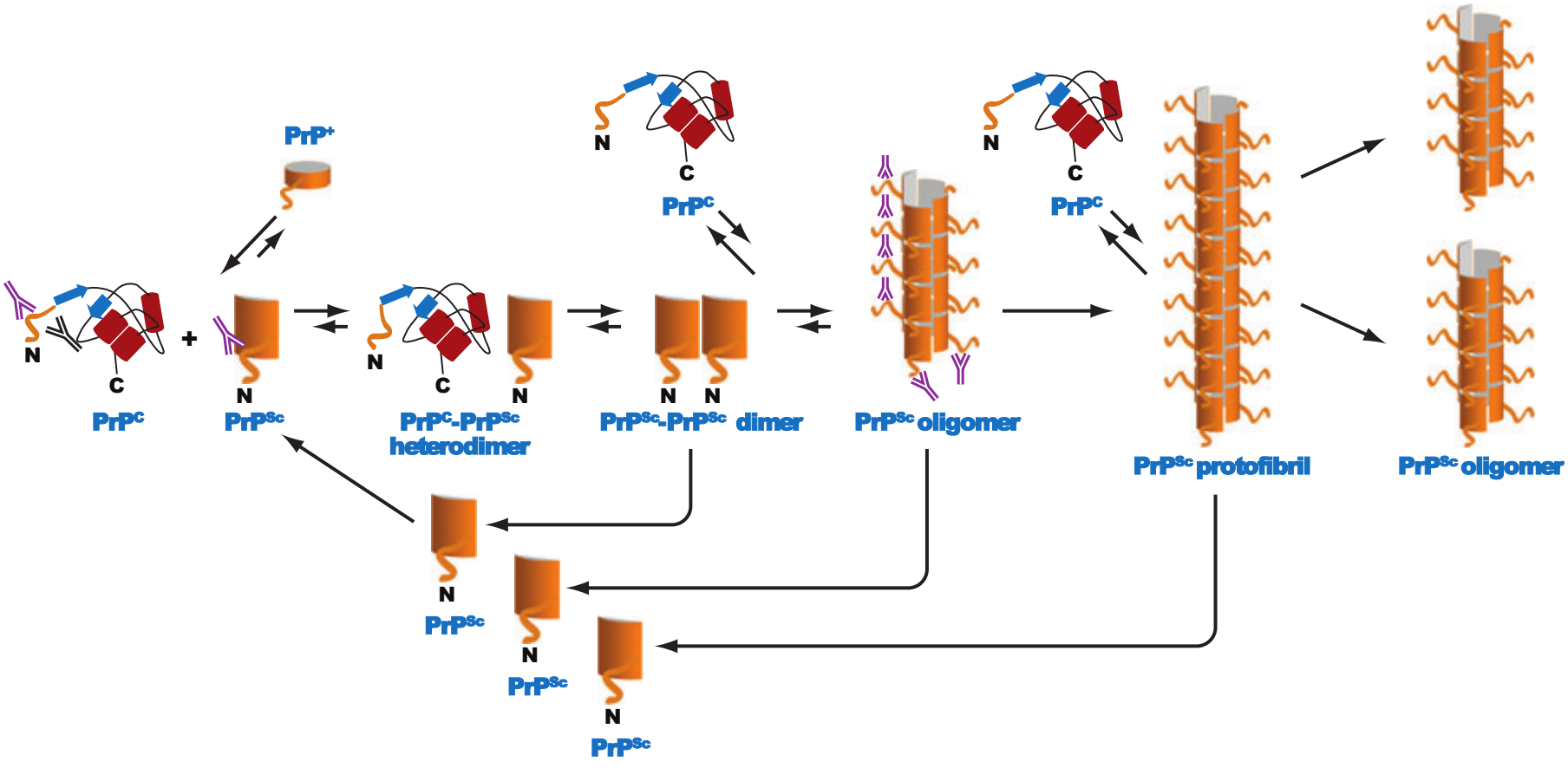
C. Flamm, I.L. Hofacker, S. Maurer-Stroh, P.F. Stadler, and M. Zehl.
Design of multi-stable RNA molecules. RNA, 7:254–265, 2001

ON THE FUNCTION OF RIBOSWITCHES



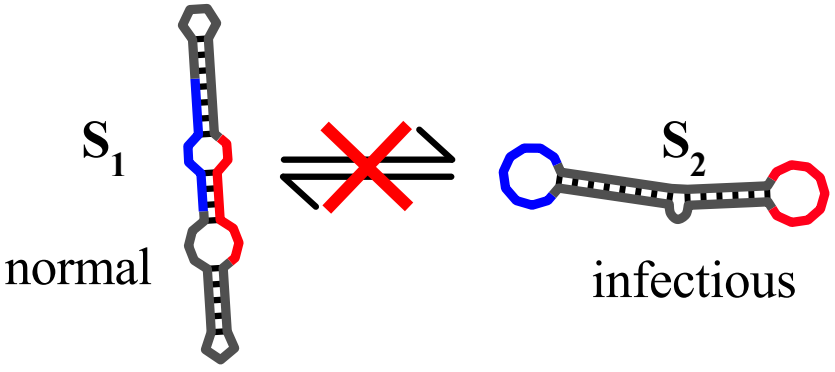
CAN WE DEMONSTRATE AUTOCATALTYIC COFORMATIONAL SELF-REPLICATION?

THE MECHANISM OF A PRION

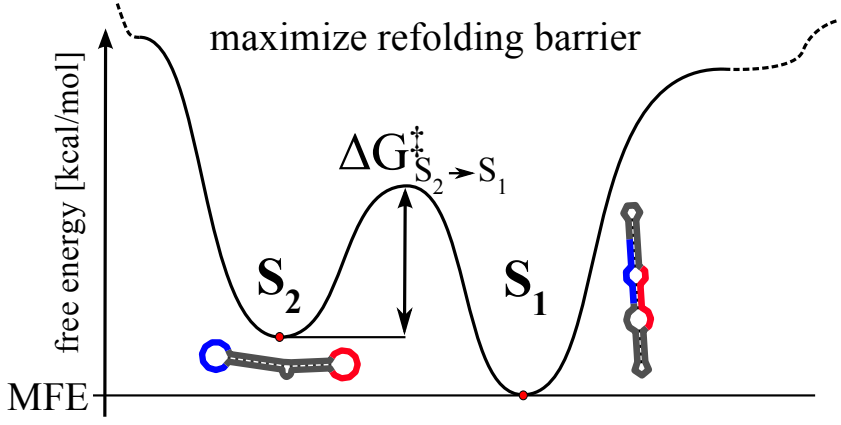


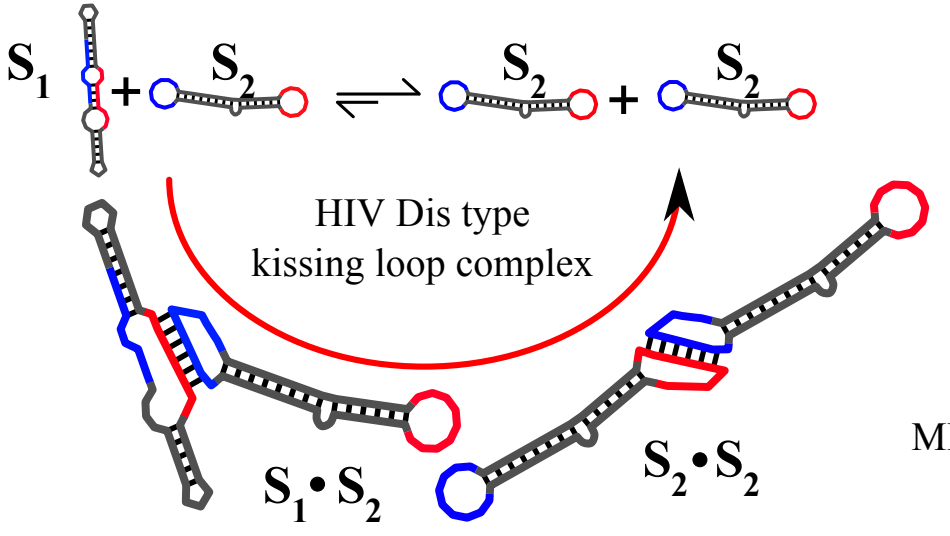
Aguzzi, A., Sigurdson, C., and Heikenwaelder, M. (2008). **Molecular mechanisms of prion pathogenesis.** Annual Review of Pathology: Mechanisms of Disease, 3, 11–40.

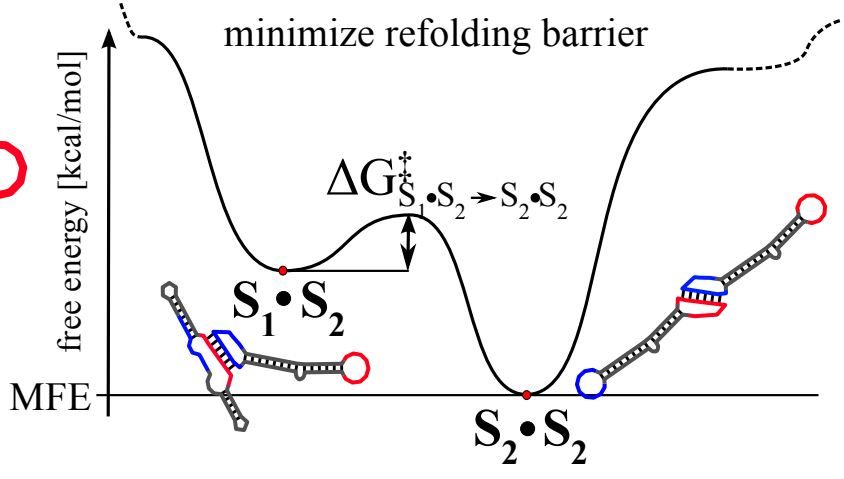
Requirements for an RNA prion



Energy Landscape







PRION DESIGN PIPELINE

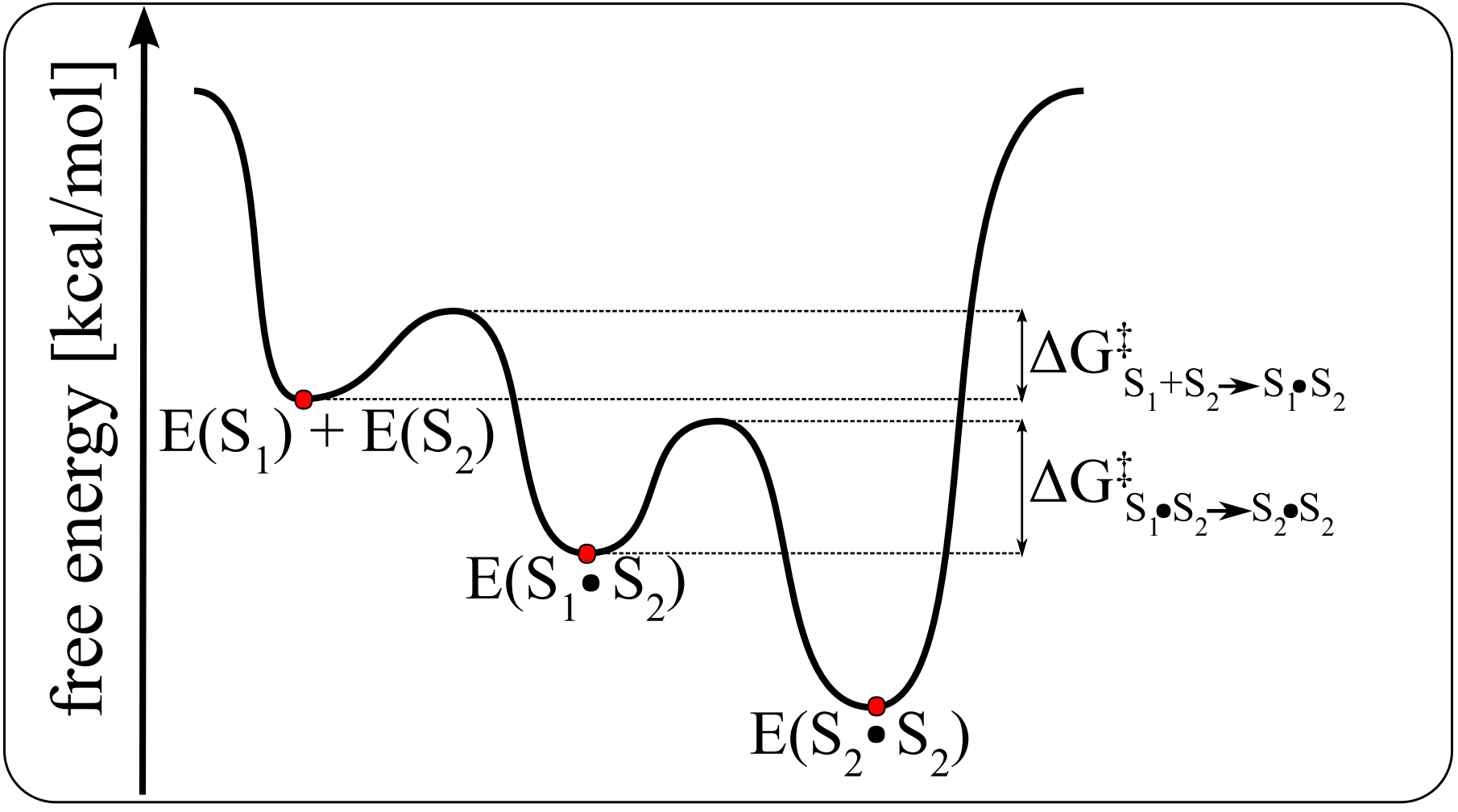
```
....((((((..((((...((((...)))))...))))..)))))))
((((((.........))))))....((((.........))))))).
NNNNNNNAACCGACGANNNNNNNNNNNNNNNNNAACGUCGGANNNNNNN
```

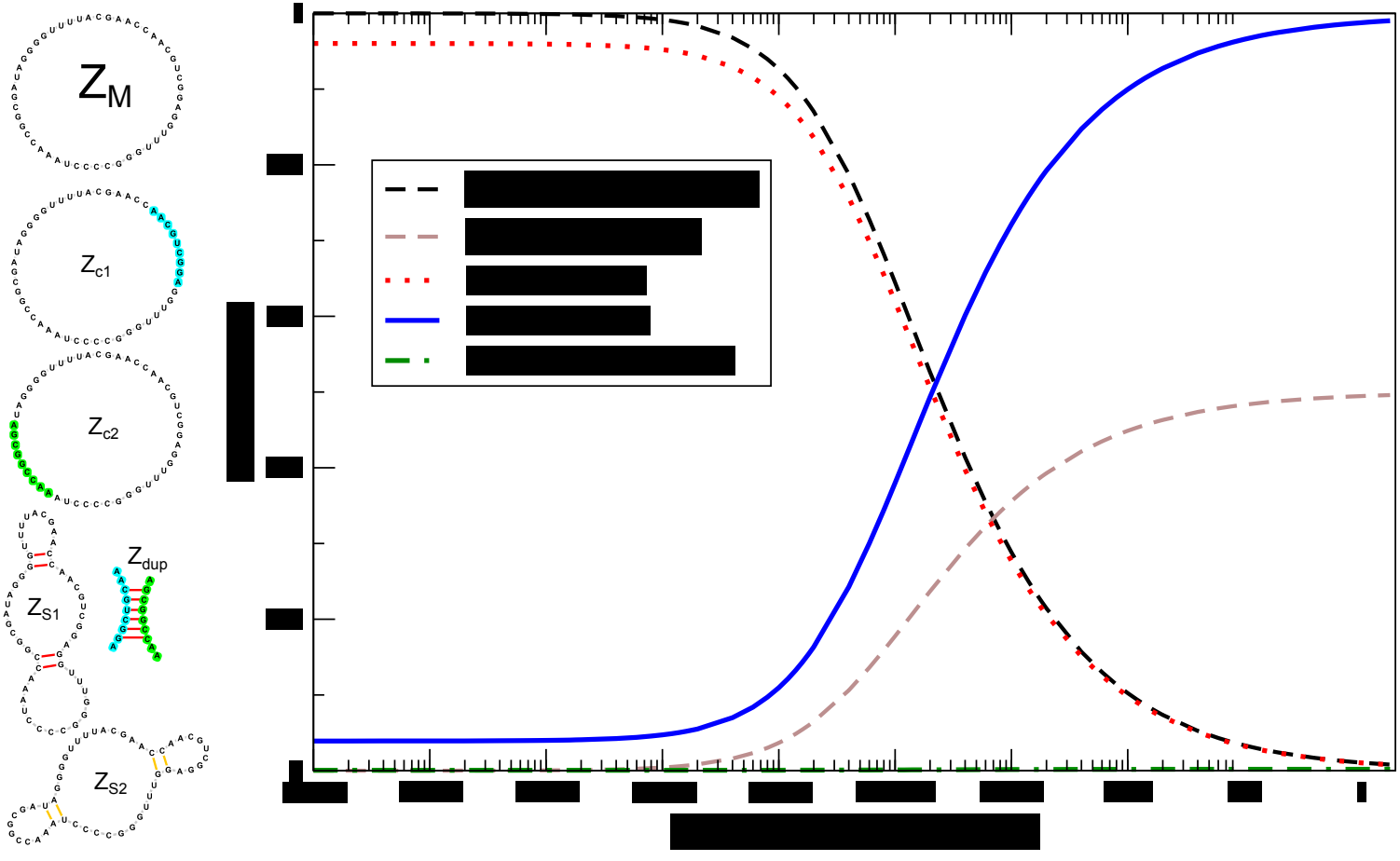
[switch.pl]

thermodynamic candidate molecules

[RNAsubopt + barriers] (M)
[findpath] (H)
[RNAsubopt + barriers] (D)

final ranking of molecules

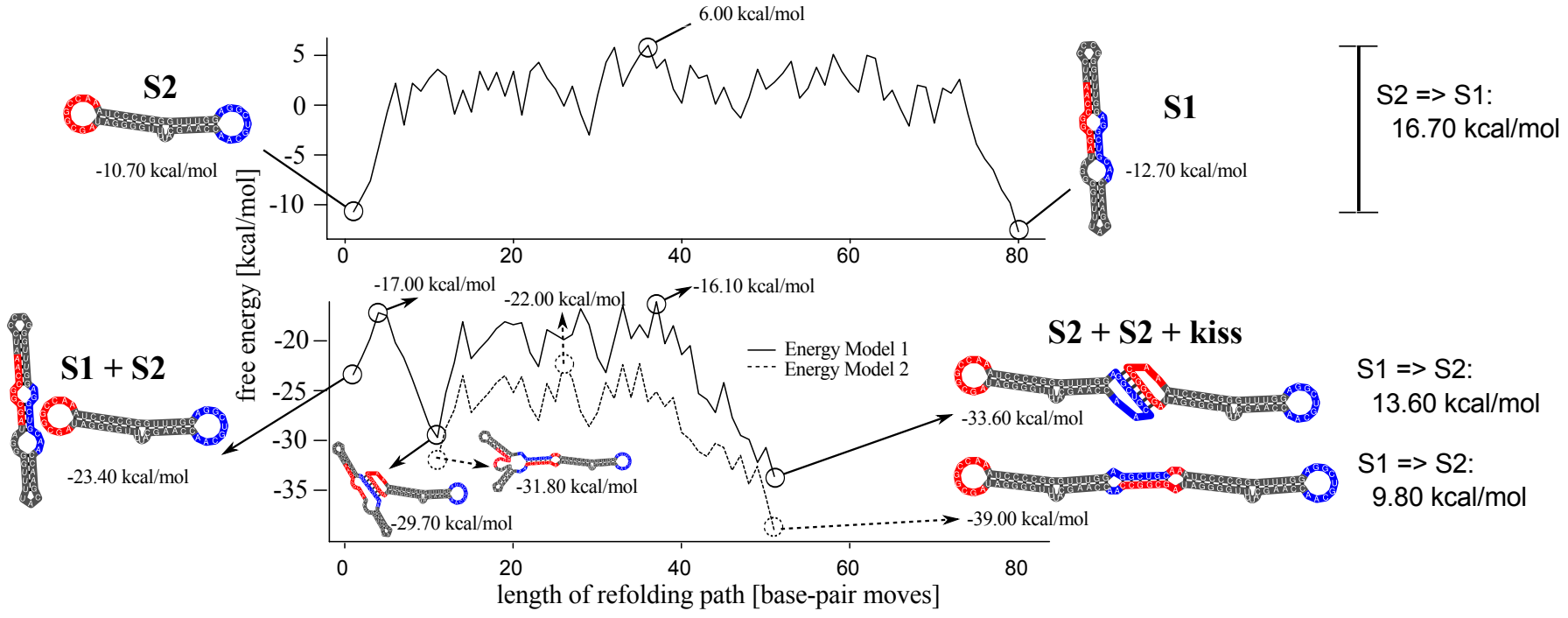

free energy [kcal/mol]
E(S1) + E(S2)
E(S1•S2)
E(S2•S2)
ΔG‡S1+S2→S1•S2
ΔG‡S1•S2→S2•S2



$$[S1] = \frac{Z_{S1}}{Z_M}[M] + \left(\frac{Z_{S1+c1}}{Z_{c1}} + \frac{Z_{S1+c2}}{Z_{c2}} \right)[D]$$

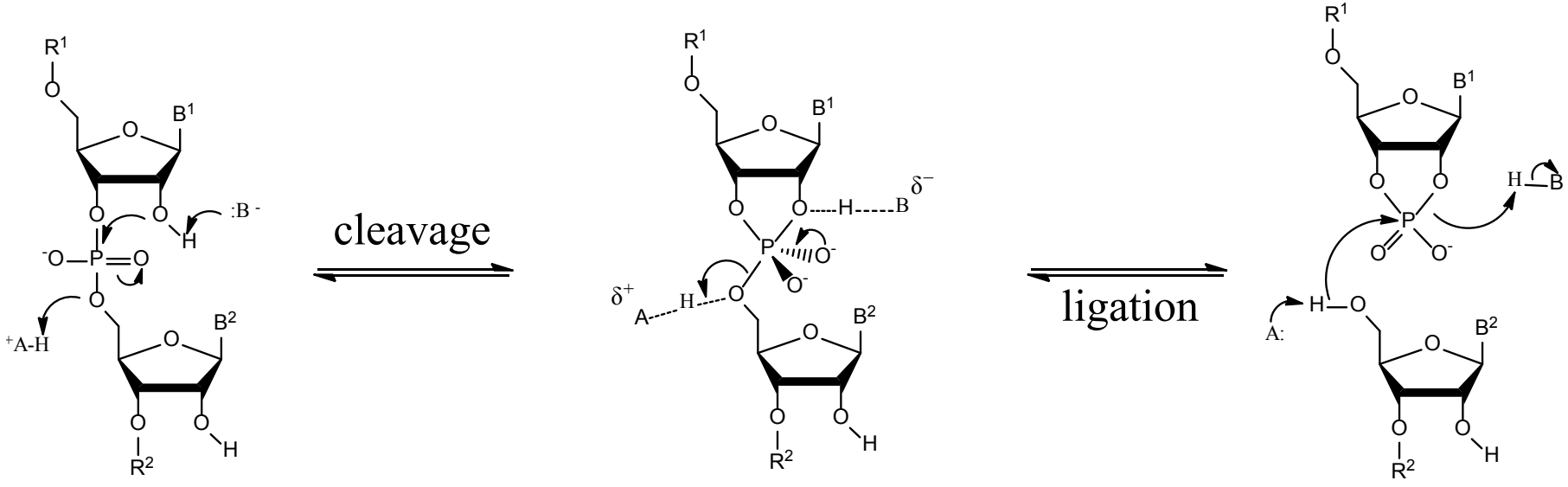
FOLDING BARRIERS WITH AND WITHOUT AUTO-CATALYSIS

S.

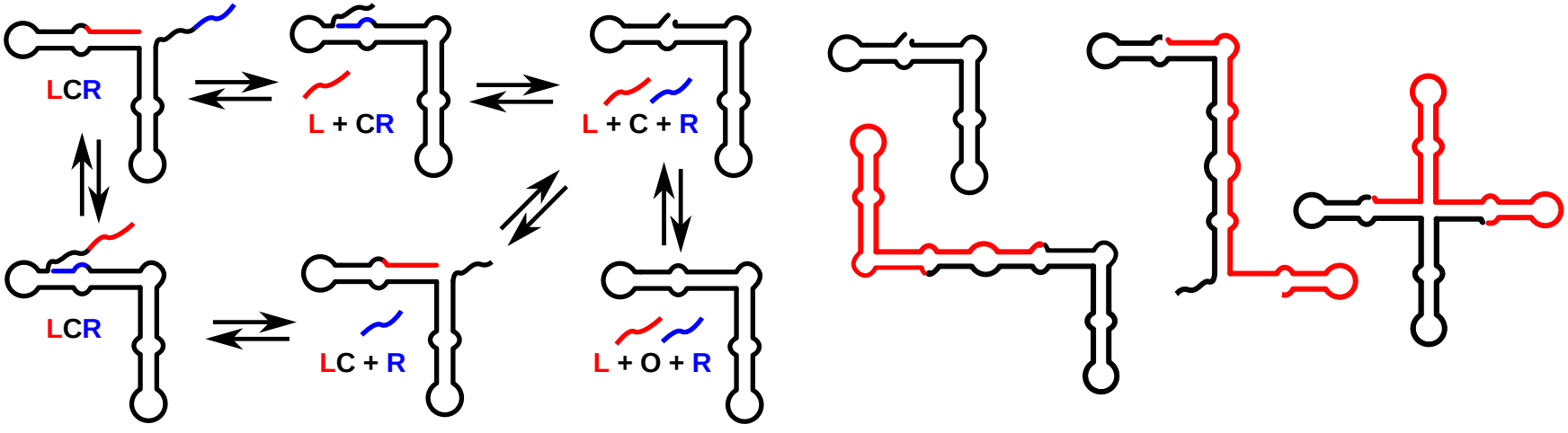


Badelt, C. Flamm, and I.L. Hofacker.
Computational design of a circular RNA with prionlike behavior.
Artificial Life 22, pages 1–14, 2016

CAN AN RNA CIRCULARIZE ITSELF?

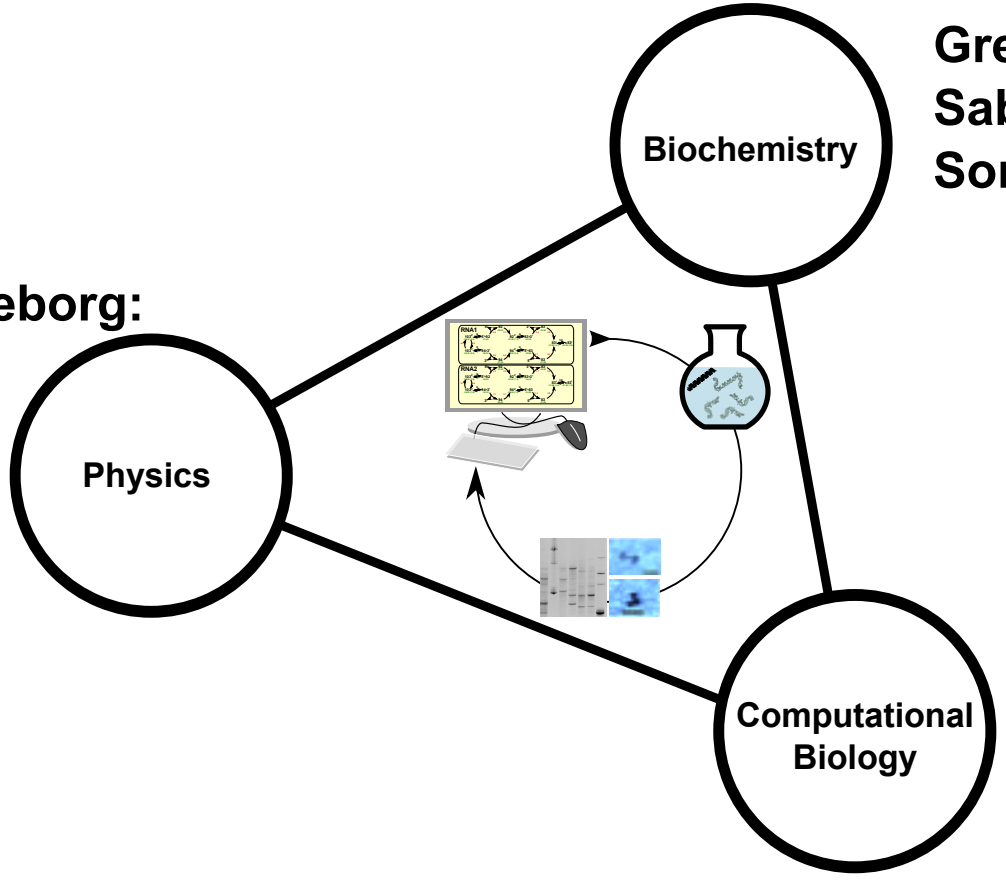


DESIGN OF SELF-PROCESSING RNA

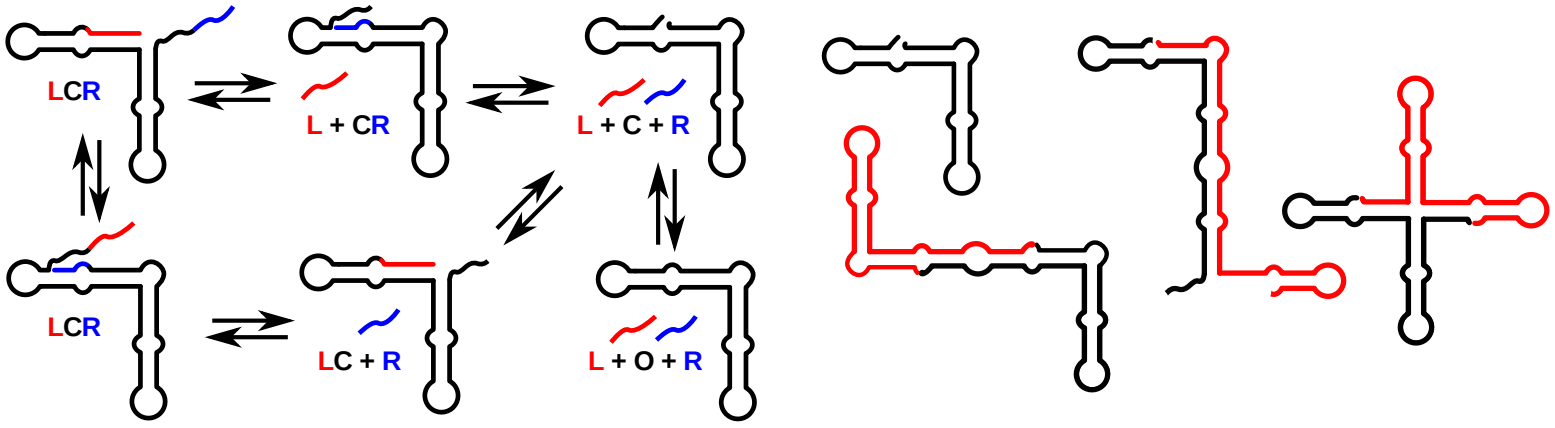


Greifswald:
Sabine Müller
Sonja Petkovic

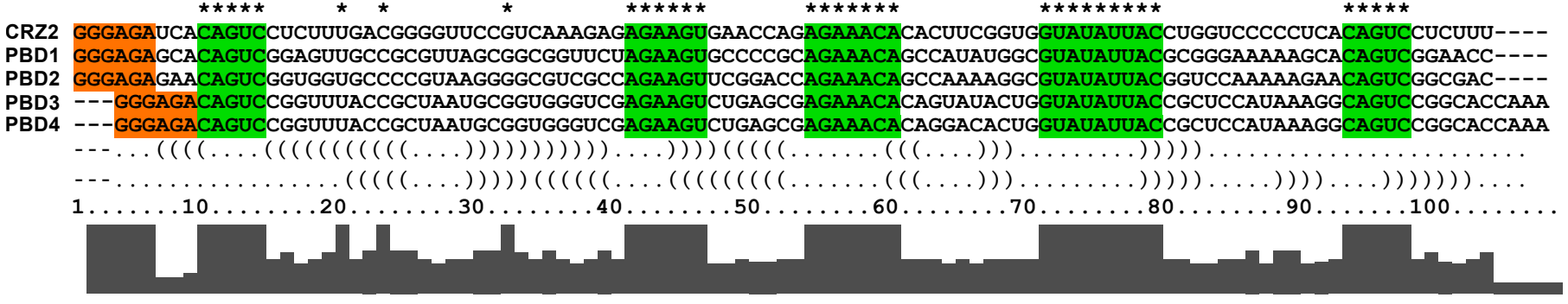
Greifswald/Göteborg:
Mihaela Delcea
Stephan Block



Wien:
Ivo Hofacker
Christoph Flamm
Stefan Badelt



- compute a set of candidate molecules (switch.pl)
- maximize probabilities to form reactive conformations
- optimize towards reactive monomers and/or dimers



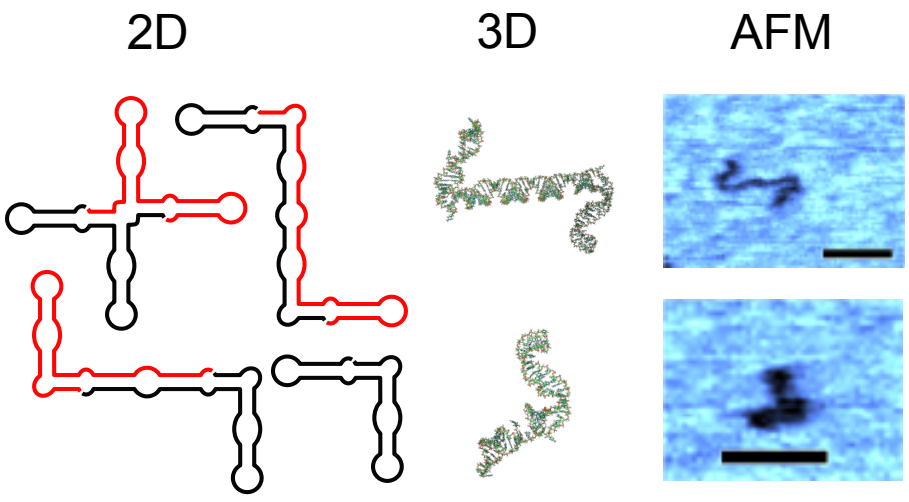


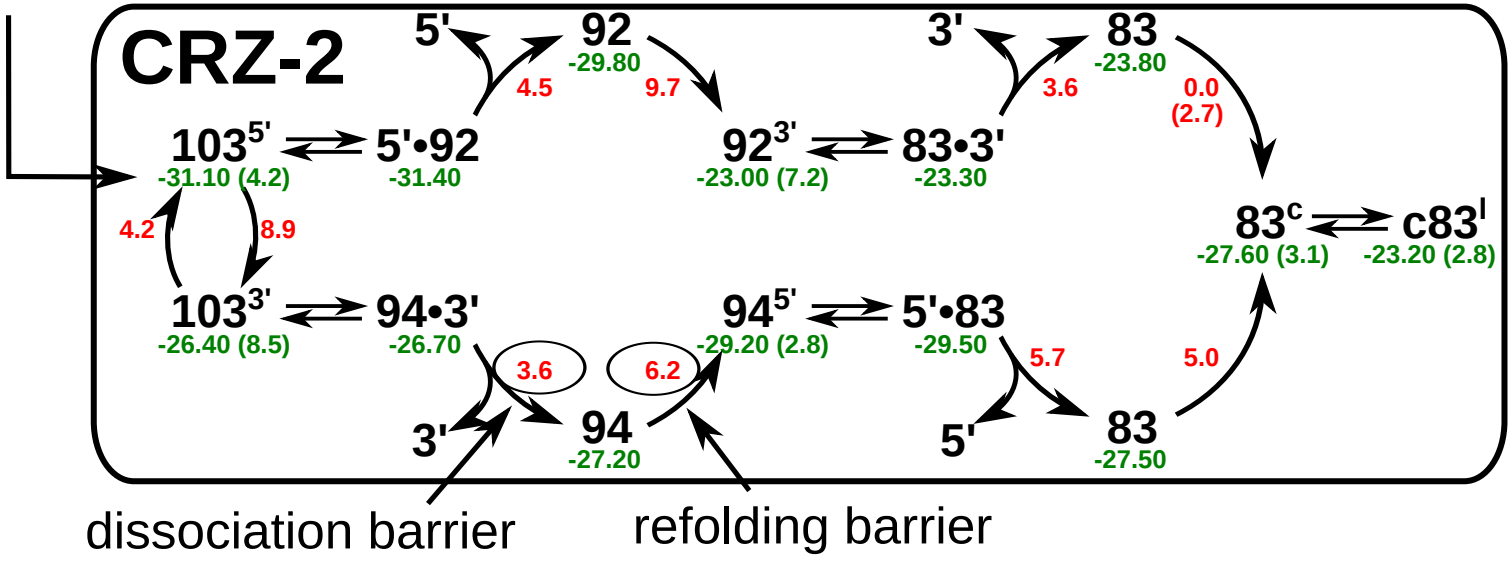
TABLE 3. AFM contour-length measurements and their implications on the ratio between monomers (M), dimers (D) and trimers (T)

Species length	AFM contour-length measurements (nm)			
Number of bases	CRZ-2 l-83mer	CRZ-2	PBD1	PBD4
M (83–103/105)	–	6.8 ± 0.7	6.2 ± 0.7	6.3 ± 0.7
	–	9.5 ± 0.4	9.3 ± 0.7	8.3 ± 0.9
	13.4 ± 1.9	**12.3 ± 1.1**	**13.8 ± 1.3**	**14.2 ± 1.0**
D (166–186/188)	–	17.2 ± 0.4	19.0 ± 1.7	20.0 ± 1.0
	24.5 ± 1.7	**22.7±0.6**	**24.5 ± 1.0**	**24.0 ± 0.9**
T (249–269/271)	–	–	31.7 ± 0.9	29.6 ± 0.7
	36.5 ± 3.2	–	**36.5 ± 0.7**	**34.8 ± 0.5**
Ratio (M:D:T)	2:1.5:1	7.4:1:0	10:2.8:1	4.5:1:1
Bold ratio only (M:D:T)	2:1.5:1	6:1:0	9.5:4:1	2.5:1:1

DISSOCIATION BARRIERS DETERMINE EFFICIENCY

RNA	103nt	91-97nt	83nt	circ83nt	> 83nt
CRZ2	-	+++	++	-	6×
PBD1	-	-	-	+++	2×
PBD2	+	+	-	++	4×
PBD3	+	+	-	+	2×
PBD4	++	+	-	+	2×

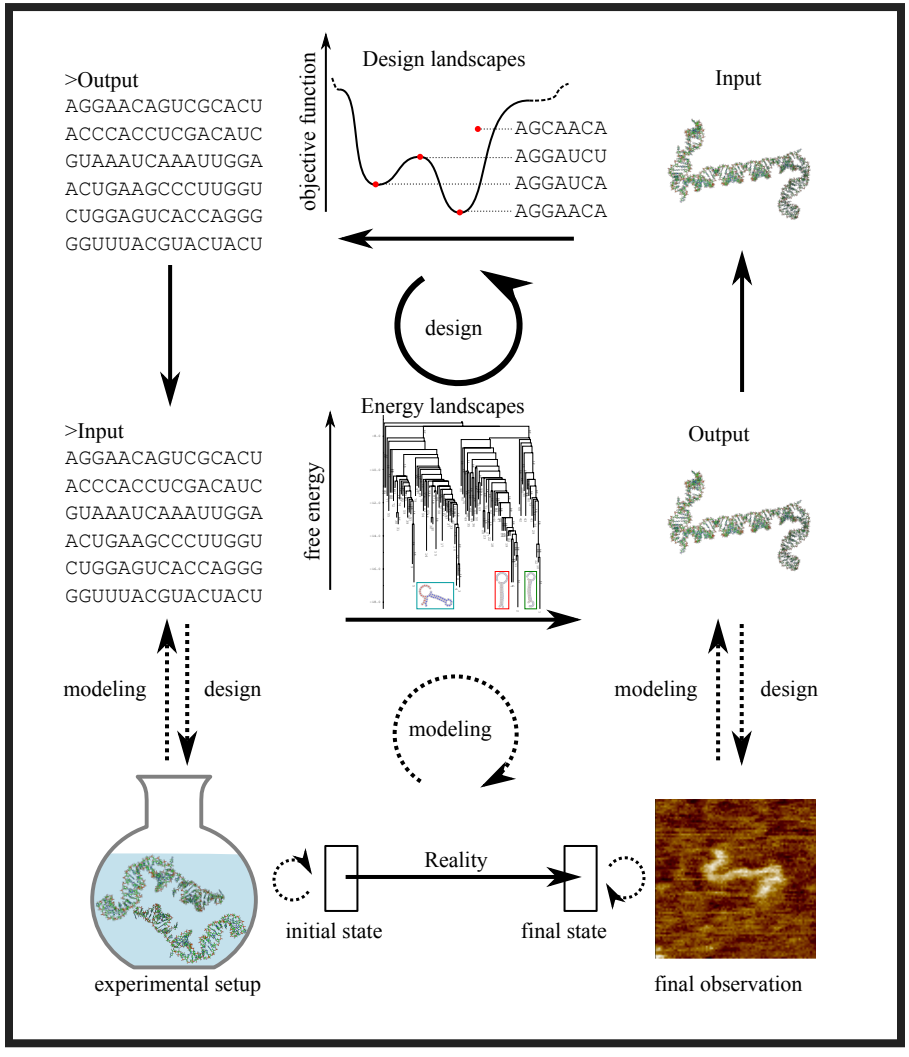
free energy (activation energy)



RESULTS

- RNA (in theory) capable of conformational self-replication
- ribozymes can be designed to ligate themselves
- energy barriers for dissociation are important for design
- experimental data to model interactions of ribozymes

SUMMARY


>Output
AGGAACAGUCGCACU
ACCCACCUCGACAUC
GUAAAUCAAAUUGGA
ACUGAAGCCCUUGGU
CUGGAGUCACCAGGG
GGUUUACGUACUACU
objective function
Design landscapes
AGCAACA
AGGAUCU
AGGAUCA
AGGAACA
Input
design
>Input
AGGAACAGUCGCACU
ACCCACCUCGACAUC
GUAAAUCAAAUUGGA
ACUGAAGCCCUUGGU
CUGGAGUCACCAGGG
GGUUUACGUACUACU
free energy
Energy landscapes
Output
modeling
design
modeling
modeling
design
Reality
initial state
final state
experimental setup
final observation

THANKS TO

Advisers: Ivo L. Hofacker, Christoph Flamm

PhD Committee: Sabine Müller, Peter F. Stadler

ViennaRNA package: Ronny Lorenz

self-processing RNA: Sonja Petkovic, Stephan Block

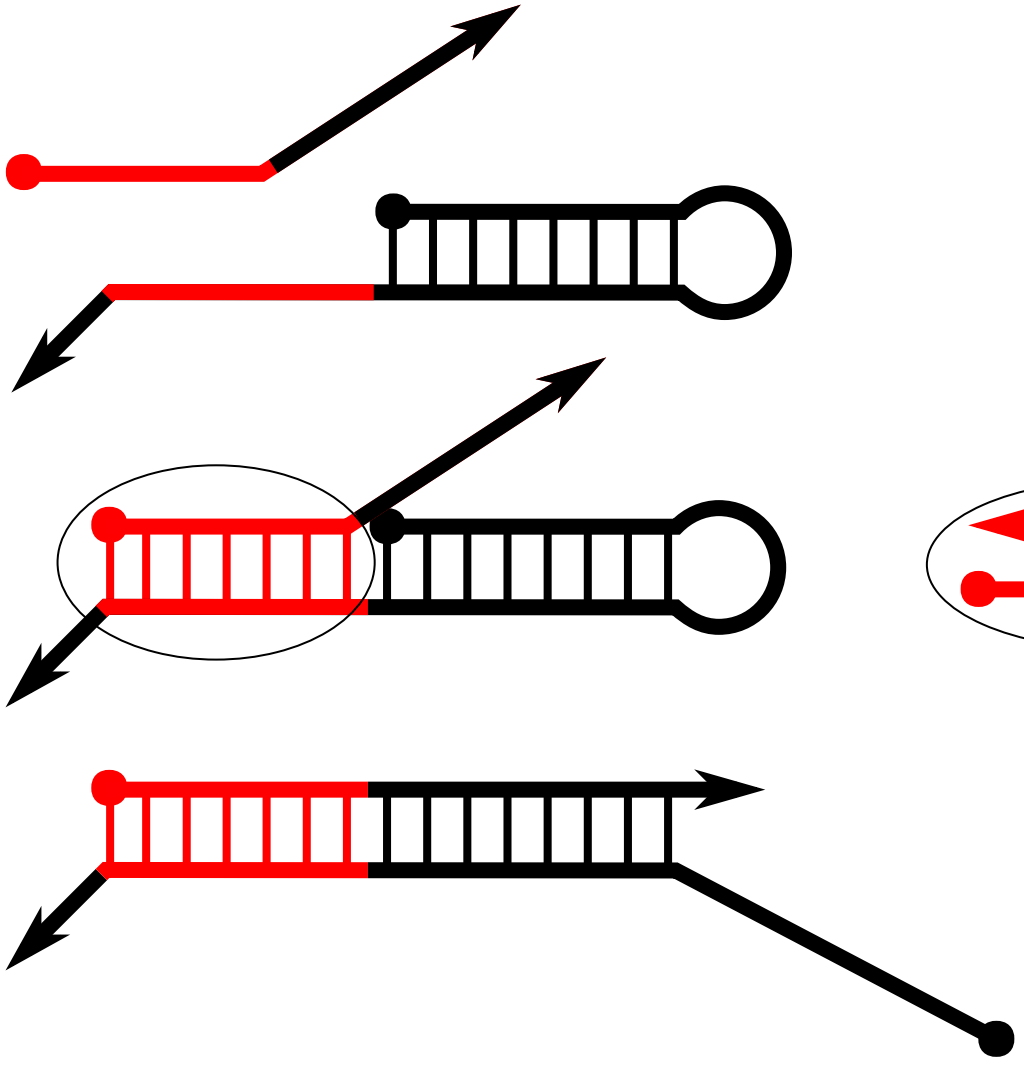
cotranscriptional folding/design: Peter Kerpedjiev, Stefan Hammer, Andrea Tanzer, Michael T. Wolfinger, Michael Jantsch, Konstantin Licht, Mansoureh Tajaddod

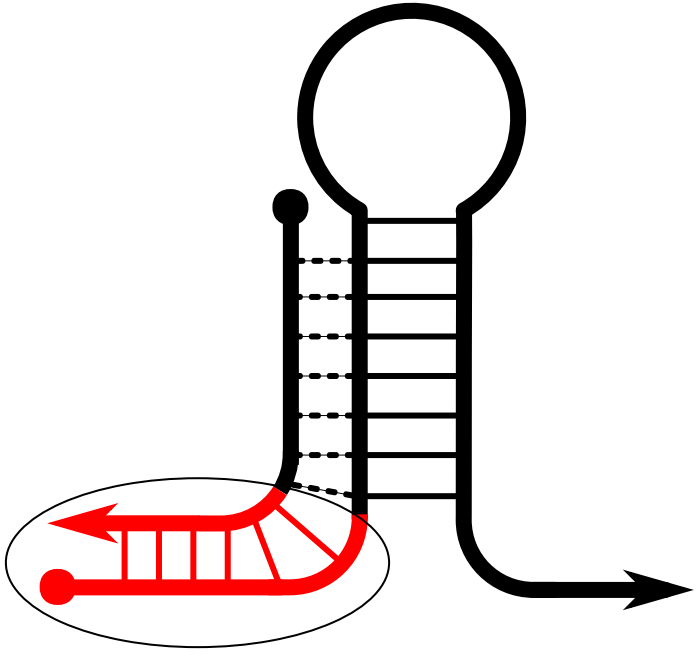
RNA design: Mario Mörl, Regula Arreger

This research was funded in parts by the FWF International Programme I670, the DK RNA program FG748004 and the FWF project "SFB F43 RNA regulation of the transcriptome".

THE AWESOME TBI

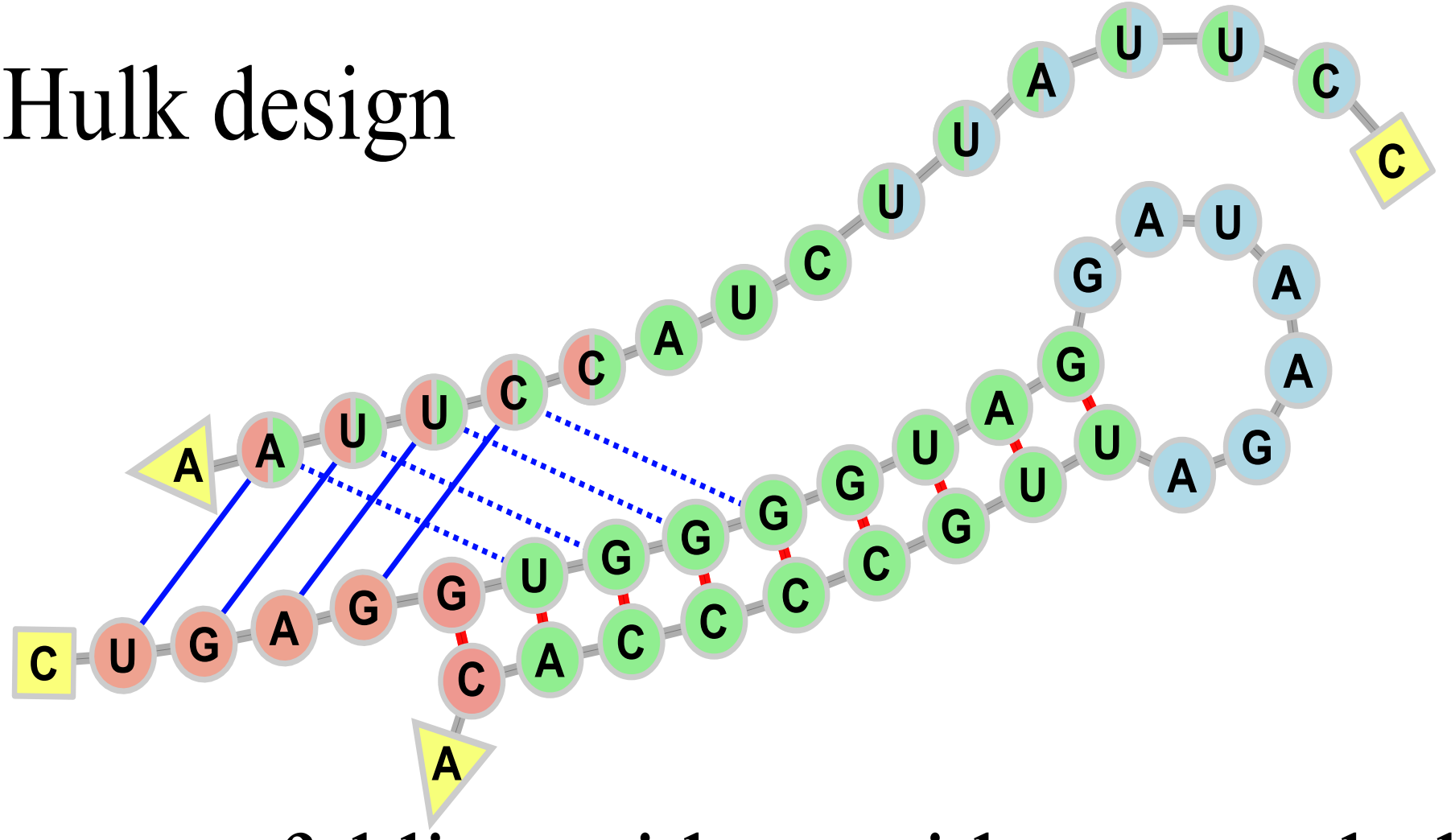
RNA-RNA INTERACTIONS



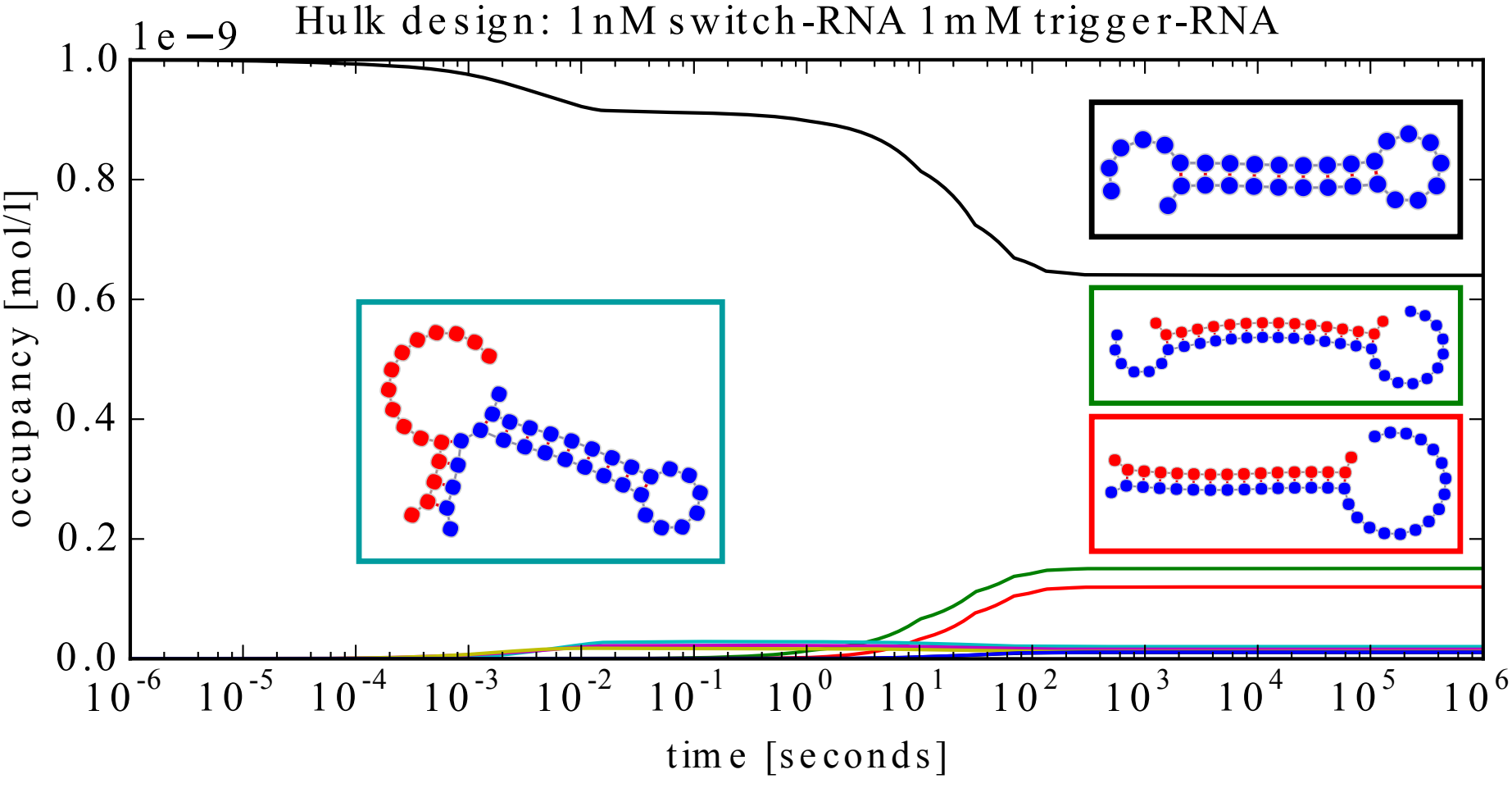


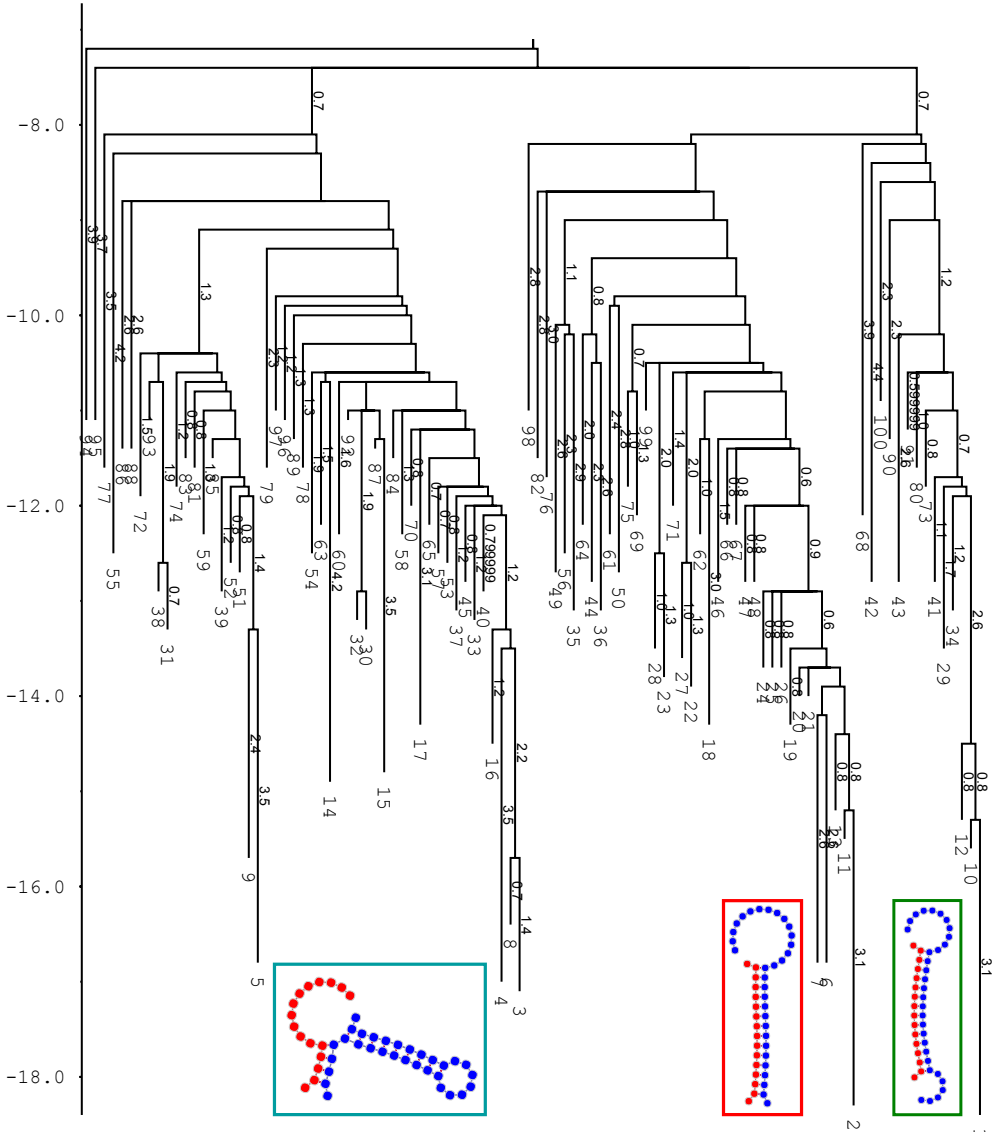
toehold
interactions

Hulk design

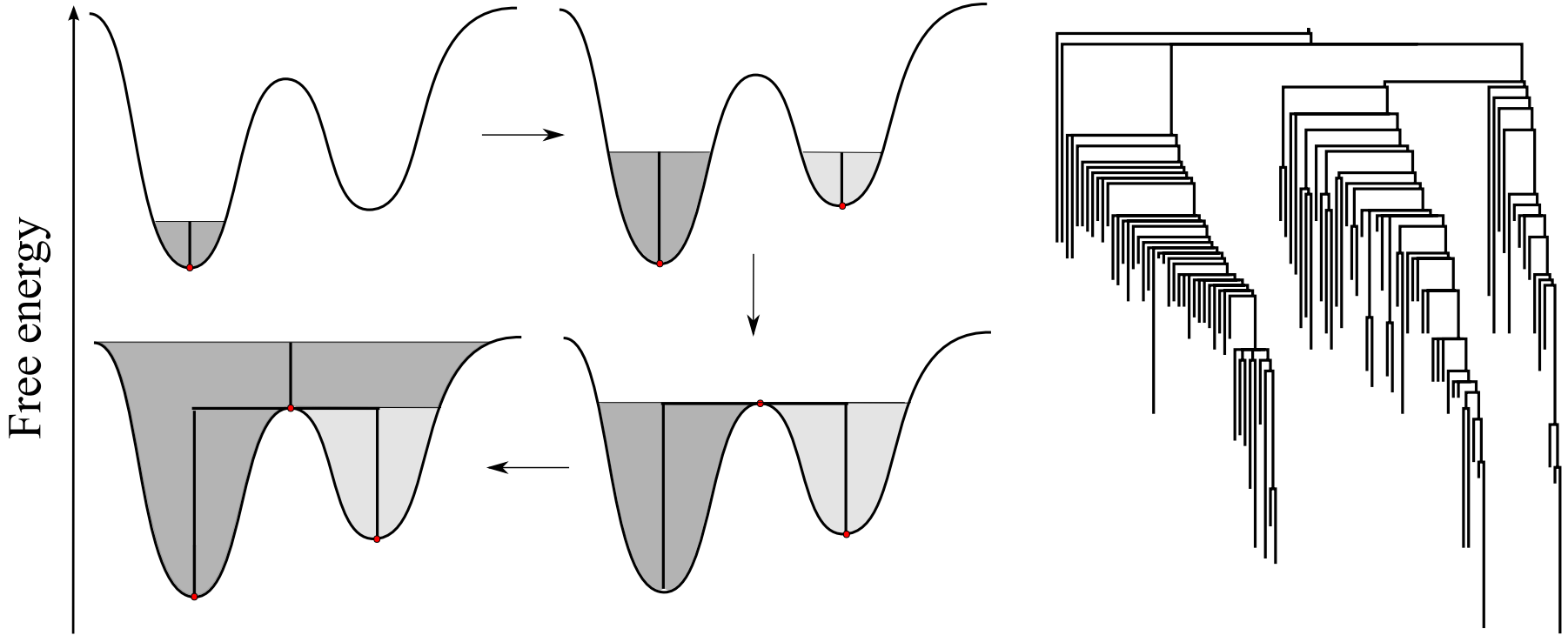


refolding with or without a toehold


Hulk design: 1nM switch-RNA 1mM trigger-RNA
1e−9
1.0
0.8
0.6
0.4
0.2
0.0
occupancy [mol/l]
10-6 10-5 10-4 10-3 10-2 10-1 100 101 102 103 104 105 106
time [seconds]


-8.0
-10.0
-12.0
-14.0
-16.0
-18.0

ENERGY LANDSCAPES



Christoph Flamm, Walter Fontana, Ivo L Hofacker, and Peter Schuster
RNA folding at elementary step resolution. RNA, 6:325–338, 2000.

Michael T. Wolfinger, Andreas Svrcek-Seiler, Christoph Flamm, Ivo L. Hofacker, and Peter F. Stadler.
Efficient computation of RNA folding dynamics. Journal of Physics A: Mathematical and General, 37:4731–4741, 2004.

KINETICS

Calculate transition rates from energy barriers

$$\Delta G^{\ddagger} = E(s_j) - E(s_i)$$

$$k_{ij} = \begin{cases} k_0 & \text{if } \Delta G^{\ddagger} \leq 0 \\ k_0 e^{-\frac{\Delta G^{\ddagger}}{RT}} & \text{otherwise} \end{cases}$$

... where k_0 is a constant to relate folding to wall-clock time

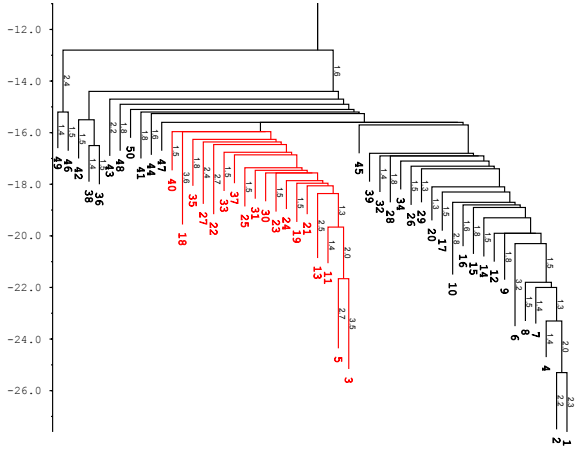
N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. **Equation of state calculations by fast computing machines.** The Journal of Chemical Physics, 21(6):1087–1092, 1953.

THE MASTER EQUATION

$$\frac{dP_i(t)}{dt} = \sum_{i \neq j} (P_j(t)k_{ji} - P_i(t)k_{ij})$$

... together with the rate of gradient basin transitions ...

$$k_{\alpha\beta} = \sum_{i \in \alpha} \sum_{j \in \beta} P(i|\alpha)k_{ij}$$



... can be solved for 60-80 nucleotides sequence length