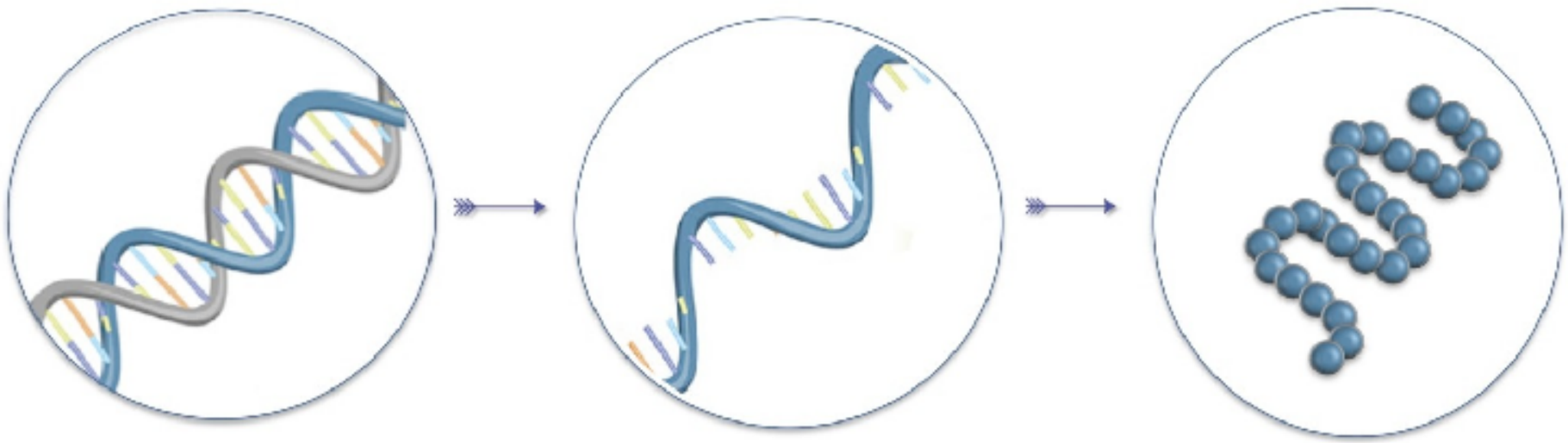
RNA dynamics using molecular modeling and experimental data

Giovanni Bussi
Molecular and Statistical Biophysics
SISSA, Trieste, Italy

bussi@sissa.it
http://people.sissa.it/~bussi
http://srnas.sissa.it

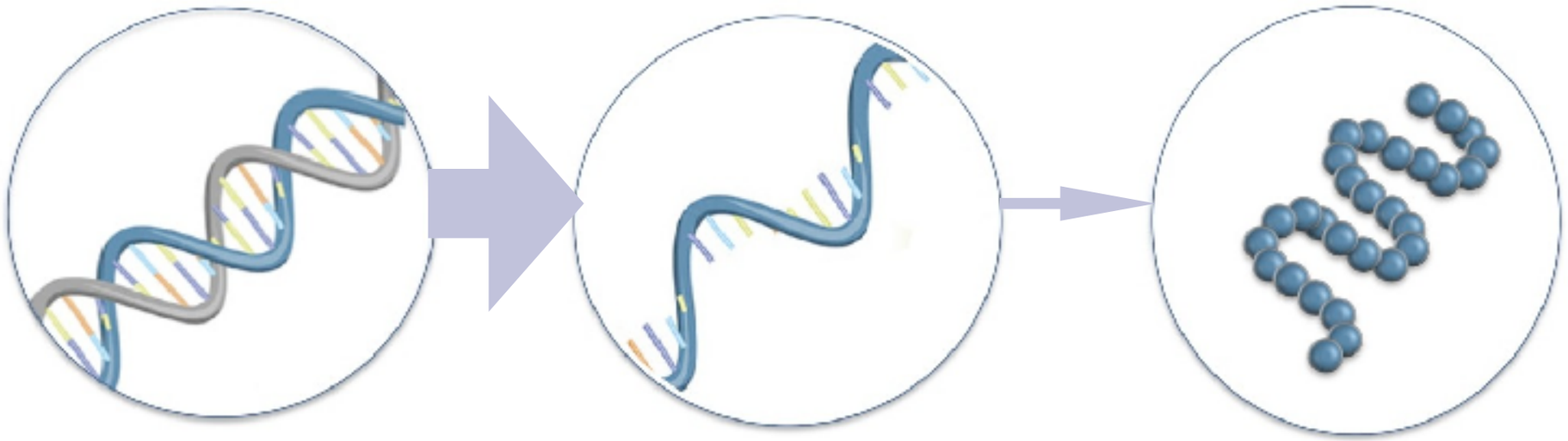
RNA



DNA → RNA → proteins

(central dogma of molecular biology)

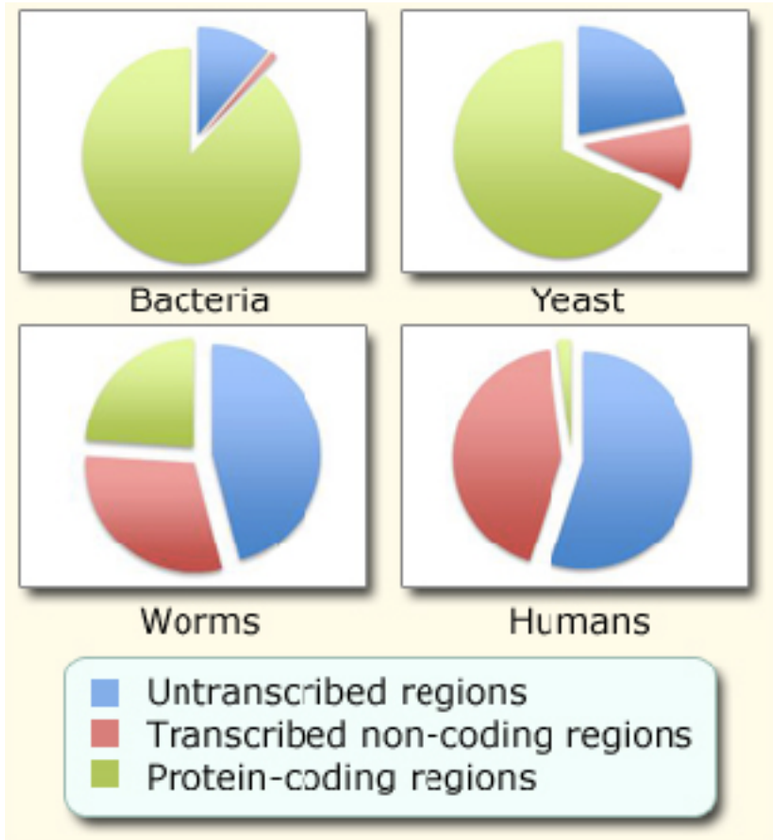
RNA



DNA

DNA → RNA

DNA → RNA → proteins



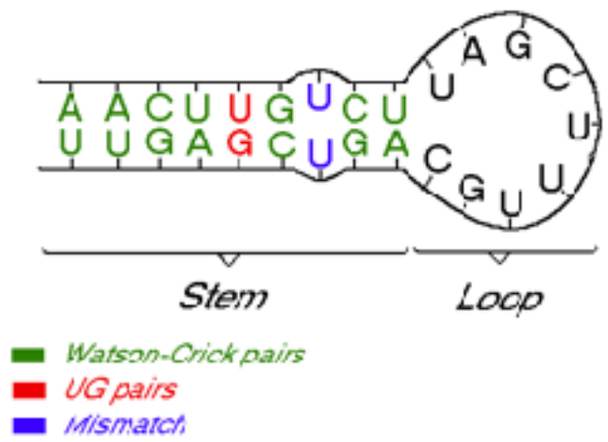
http://ocw.mit.edu/courses/biology/7-345-non-coding-rnas-junk-or-critical-regulators-in-health-and-disease-spring-2012/

RNA Structure ...

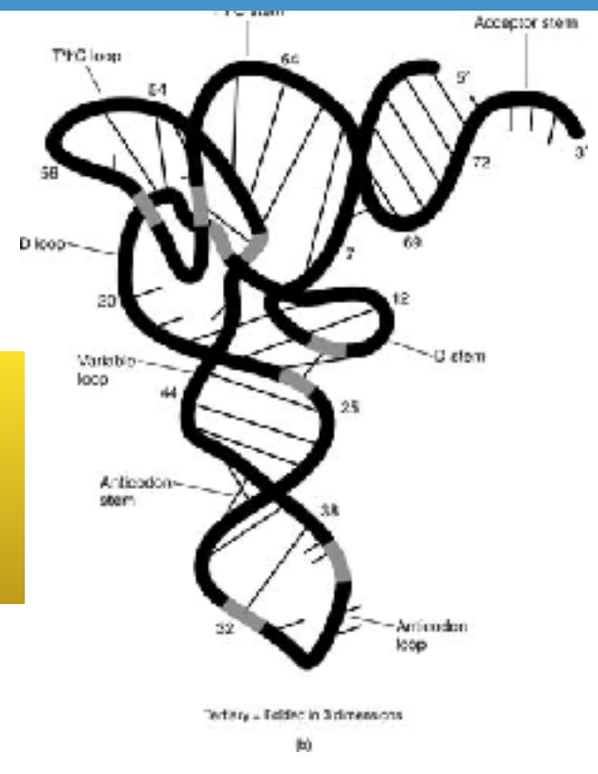
primary (1D)

...AUUCGGGCUUCUAUUC...

secondary (2D)



tertiary (3D)

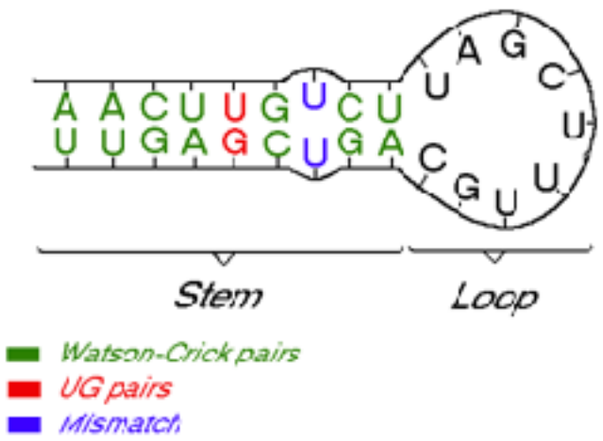


RNA Structure ...

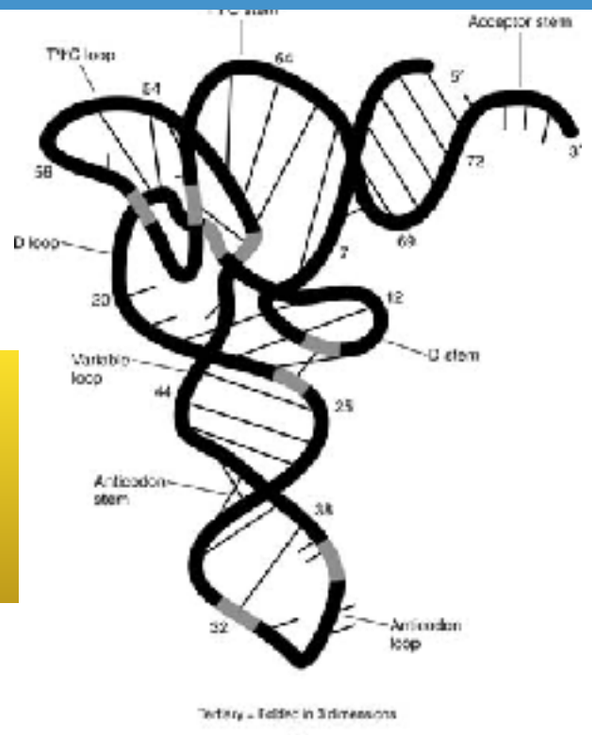
primary (1D)

...AUUCGGGCUUCUAUUC...

secondary (2D)



tertiary (3D)



... and function

Coding RNAs: function depends mostly on 1D*
Non-coding RNAs: function depends on 1D, 2D, and 3D%
Dynamics is often fundamental$

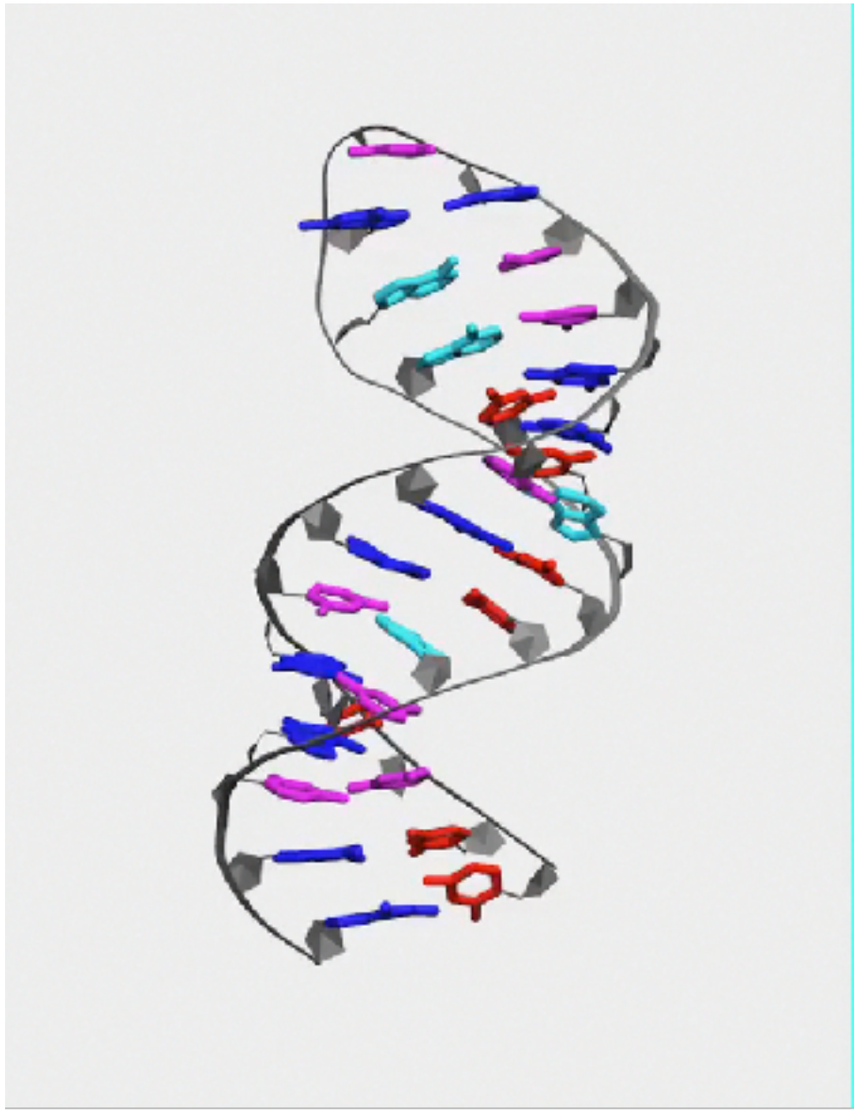
*and at least partly on 2D, Faure et al, NAR (2016); Langdon et al, Science (2018)
%ribosomes, ribozymes, riboswitches,...
$binding with proteins/ligands/ions, catalysis, etc

Molecular dynamics

$$E_{\text{total}} = \sum_{\text{bonds}} k_{\text{b}}\left(\ell - \ell_0\right)^2 + \sum_{\text{angles}} k_{\text{a}}\left(\theta - \theta_0\right)^2$$

$$+ \sum_{\text{torsions}} \frac{1}{2} V_{\text{n}}[1 + \cos(n\omega - \gamma)]^2$$

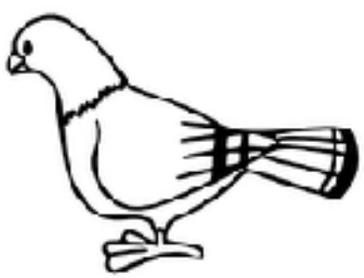
$$+ \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} \left\{ \varepsilon_{i,j} \left[\left(\frac{r_{0ij}}{r_{ij}} \right)^{12} - 2\left(\frac{r_{0ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right\}$$



Empirical force field*:

- Chemically motivated interactions
- Atomistic details
- Explicit water and ions
- No polarization
- No chemical reactivity

Approx ~20-200 ns/day



*AMBER (ff99+parmbsc0+ChiOL3+TIP3P)

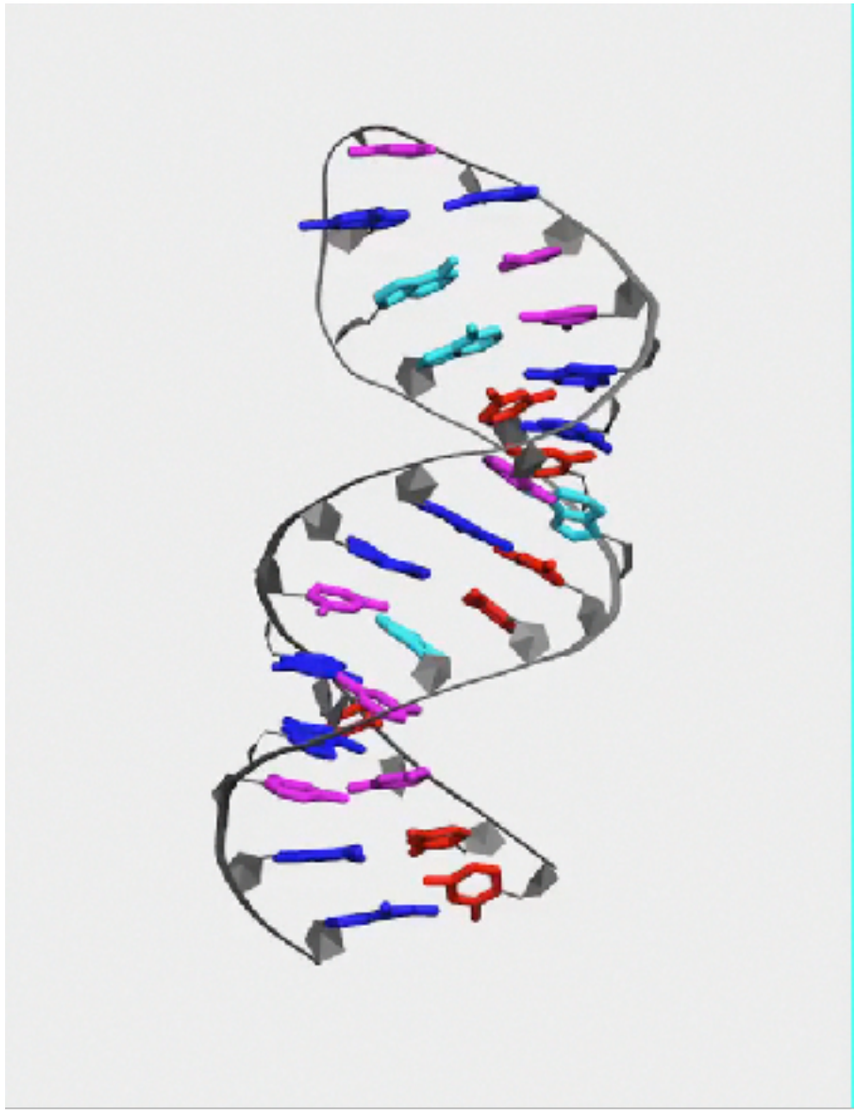
gromacs.org + plumed.org

Molecular dynamics

$$E_{\text{total}} = \sum_{\text{bonds}} k_{\text{b}}\left(\ell - \ell_0\right)^2 + \sum_{\text{angles}} k_{\text{a}}\left(\theta - \theta_0\right)^2$$

$$+ \sum_{\text{torsions}} \frac{1}{2} V_{\text{n}}[1 + \cos(n\omega - \gamma)]^2$$

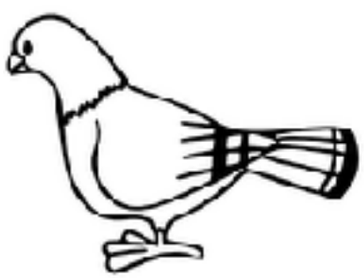
$$+ \sum_{j=1}^{N-1} \sum_{i=j+1}^{N} \left\{ \varepsilon_{i,j} \left[\left(\frac{r_{0ij}}{r_{ij}} \right)^{12} - 2\left(\frac{r_{0ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right\}$$



Empirical force field*:

- Chemically motivated interactions
- Atomistic details
- Explicit water and ions
- No polarization
- No chemical reactivity

Approx ~20-200 ns/day



*AMBER (ff99+parmbsc0+ChiOL3+TIP3P)

gromacs.org + plumed.org

Pros vs Cons

High (unlimited) time and space resolution

Access to dynamics

(Relatively) cheap 🙂

Pros vs Cons

High (unlimited) time and space resolution

Access to dynamics

(Relatively) cheap 🙂

Pros vs Cons

High (unlimited) time and space resolution

Access to dynamics

(Relatively) cheap 🙂

Short timescales

Force-field inaccuracies

It’s just a model 😞

Pros vs Cons

High (unlimited) time and space resolution

Access to dynamics

(Relatively) cheap 🙂

Short timescales

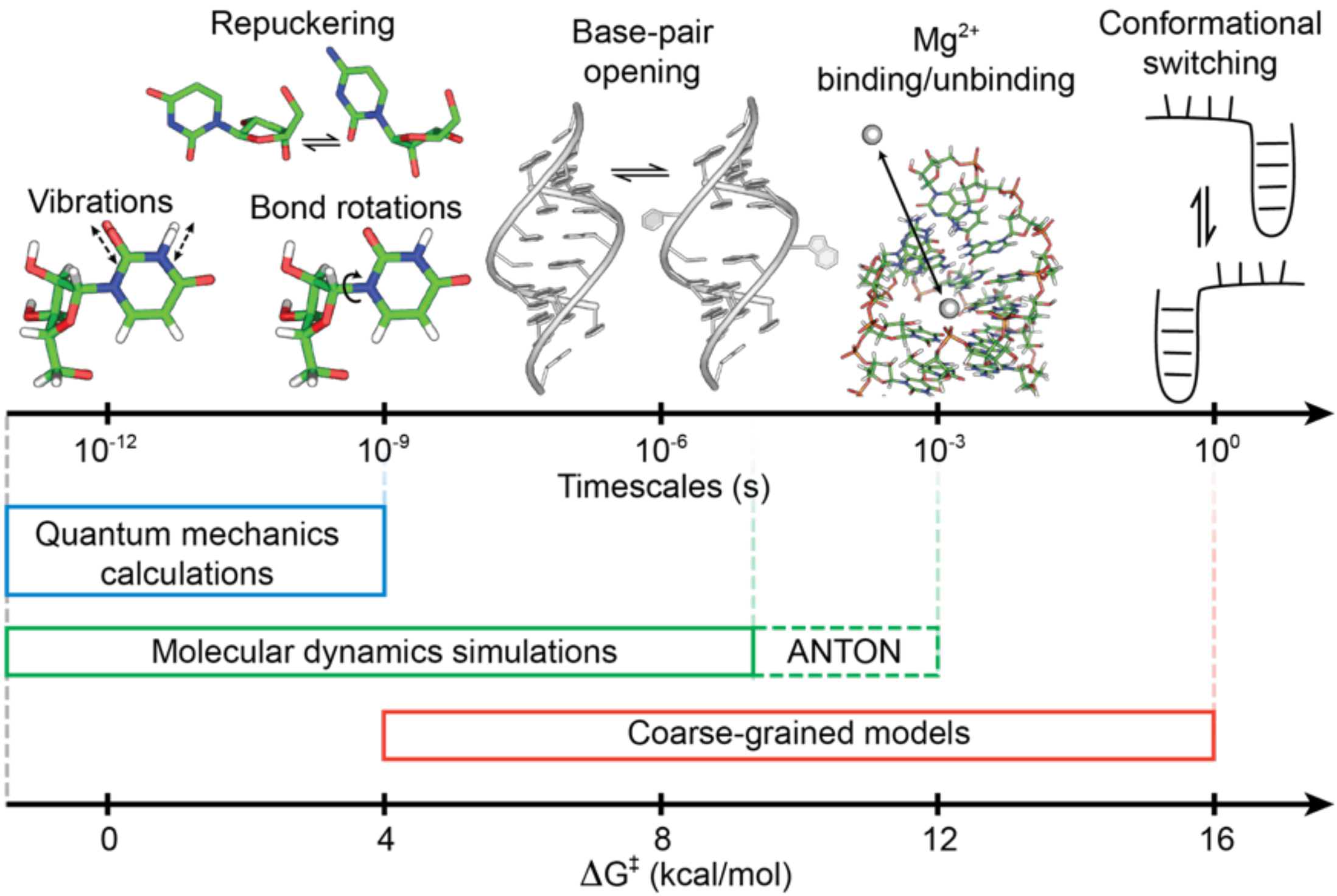
Force-field inaccuracies

It's just a model 😞

In science, as in life, it is very dangerous to fall in love with beautiful models

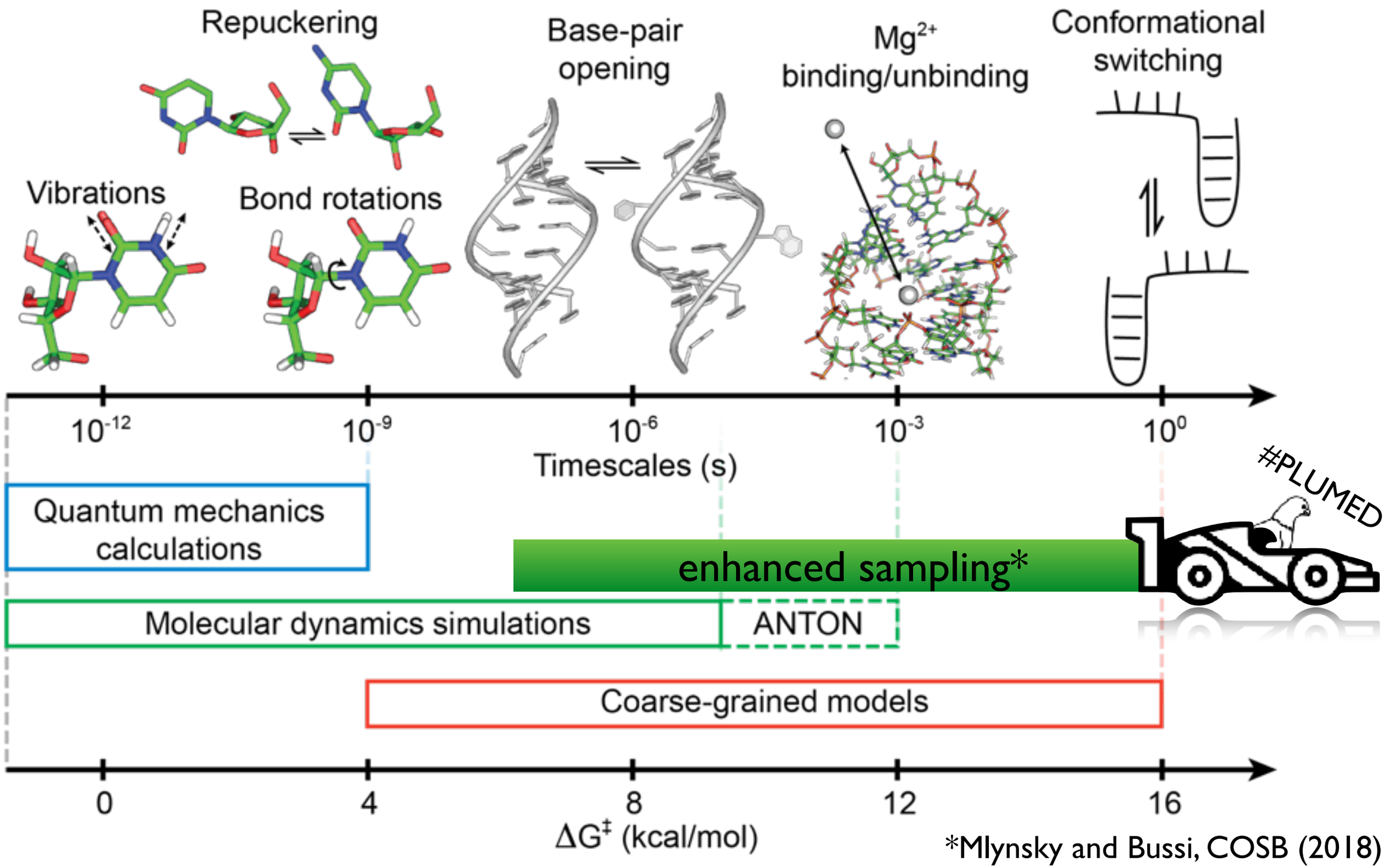
V. J. Pande

RNA timescales



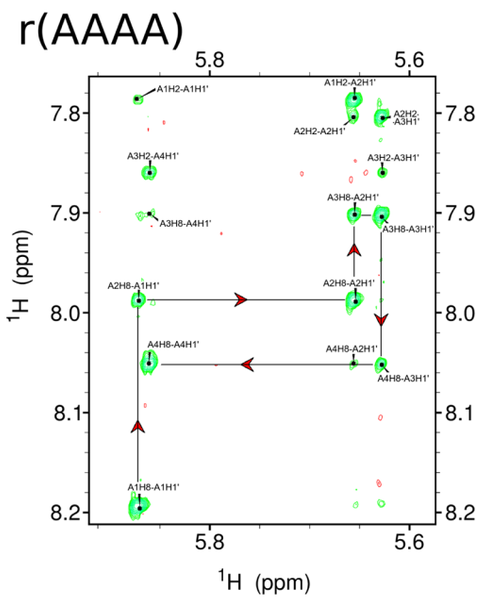
Sponer, Bussi, et al, Chem Rev (2018)

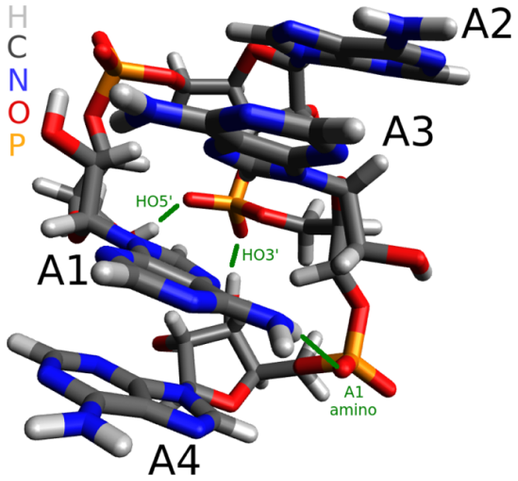
RNA timescales

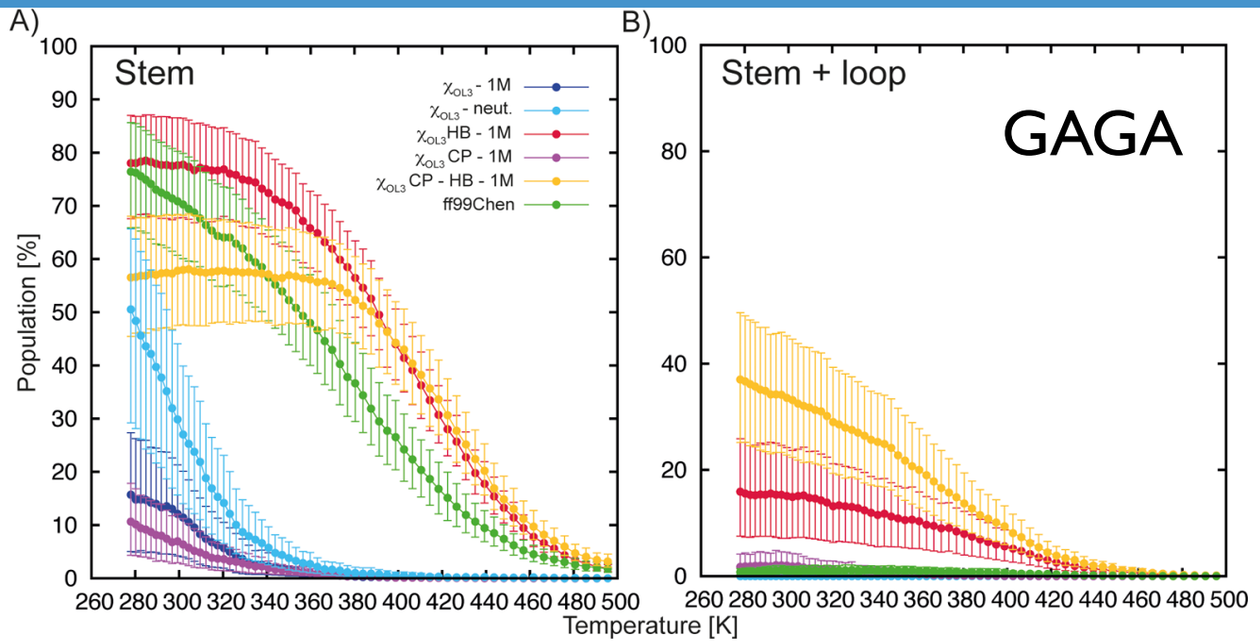


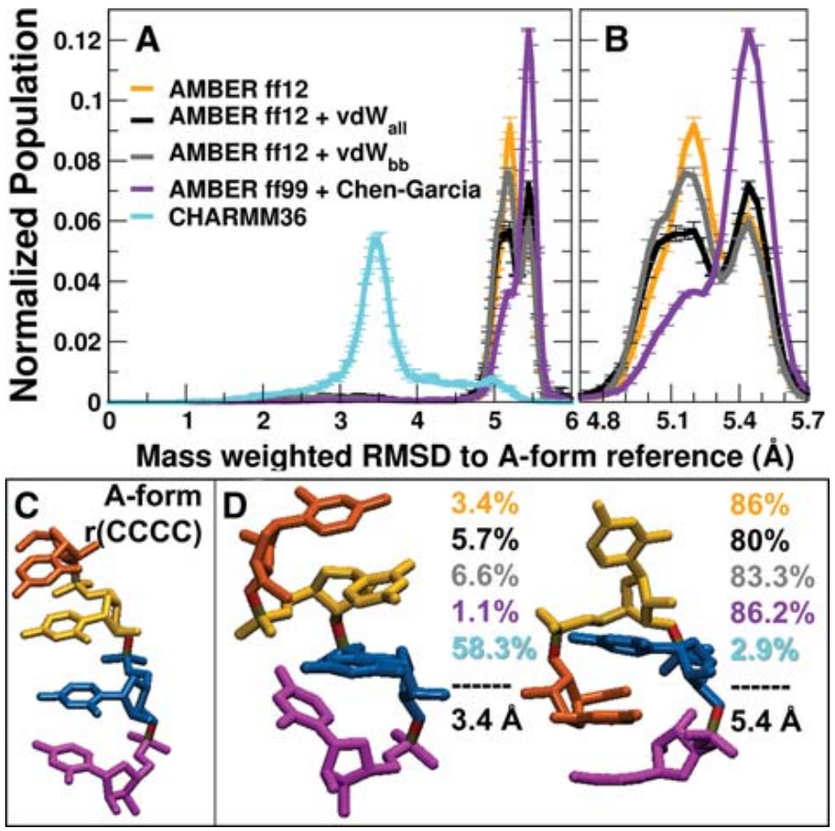
*Mlynsky and Bussi, COSB (2018)
Sponer, Bussi, et al, Chem Rev (2018)

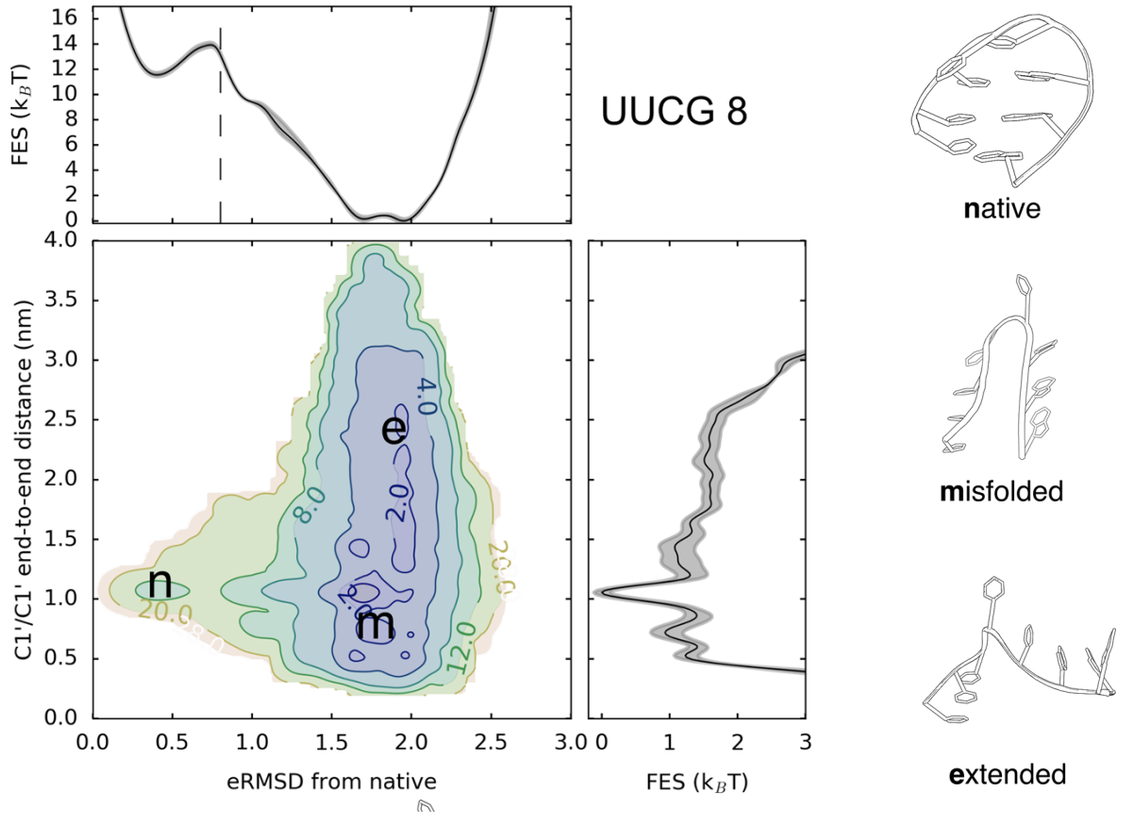
Issues with force fields







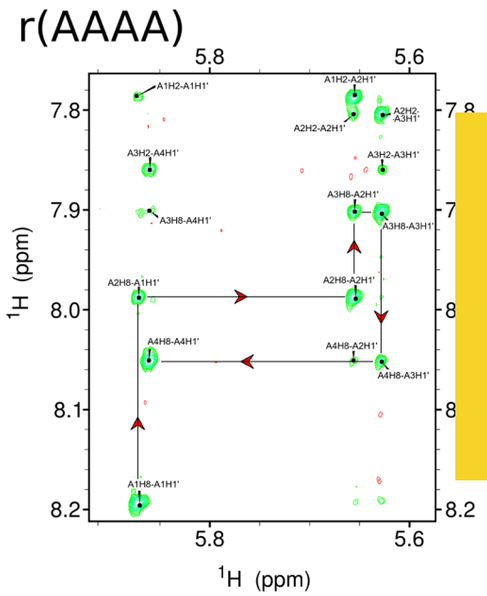


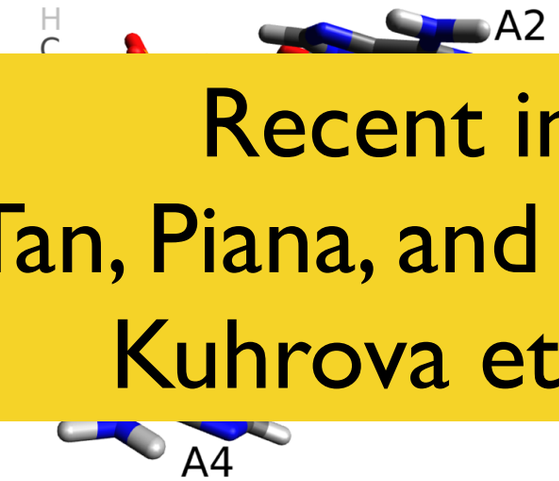


Condon et al, JCTC (2015)
Bergonzo et al, RNA (2015)

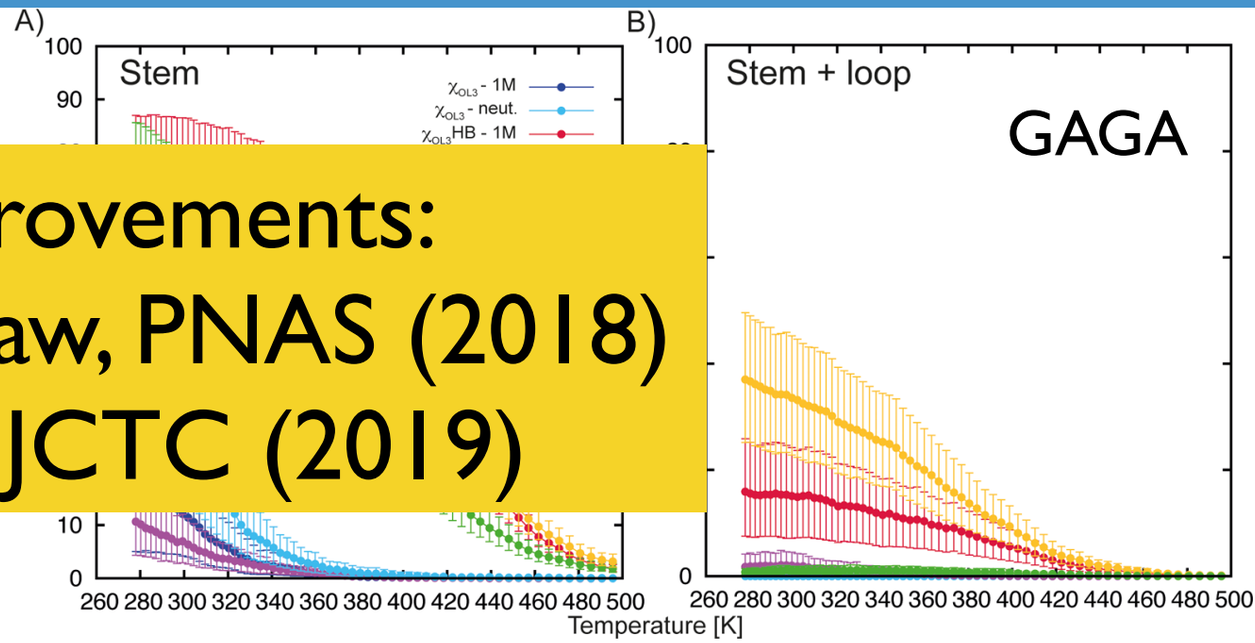
Kuhrova et al, JCTC (2016)
Bottaro et al, JPCL (2016)

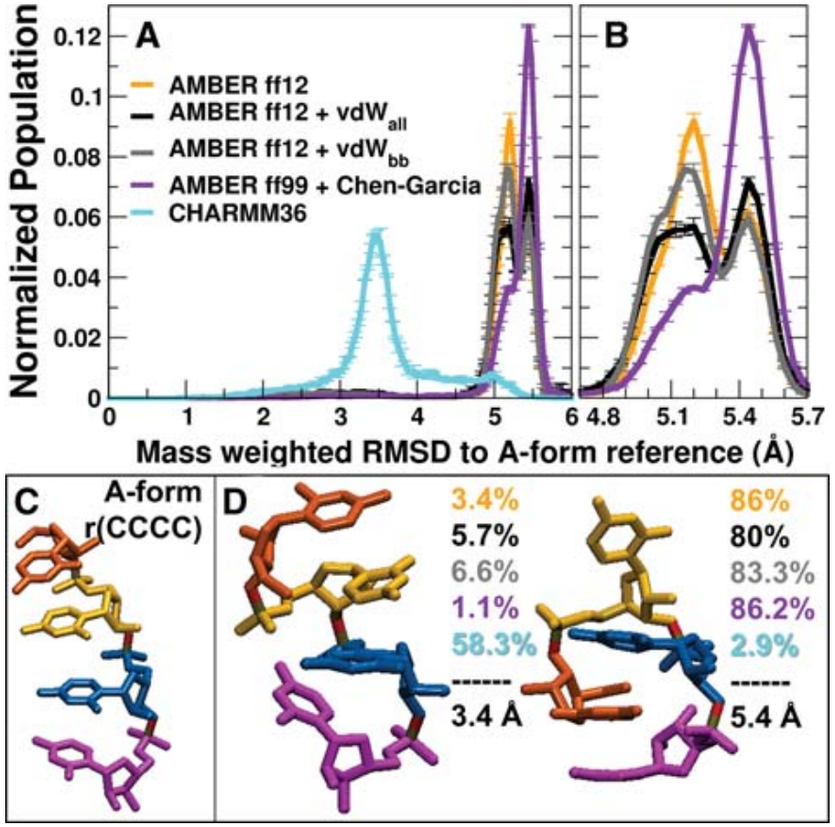
Issues with force fields

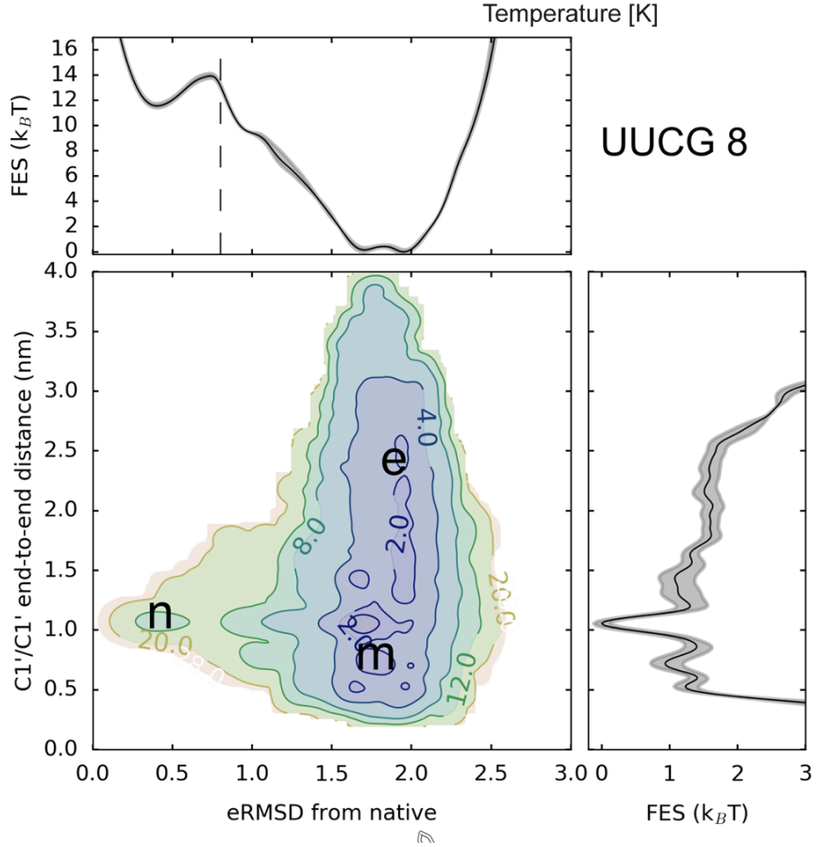




Recent improvements:
Tan, Piana, and Shaw, PNAS (2018)
Kuhrova et al JCTC (2019)



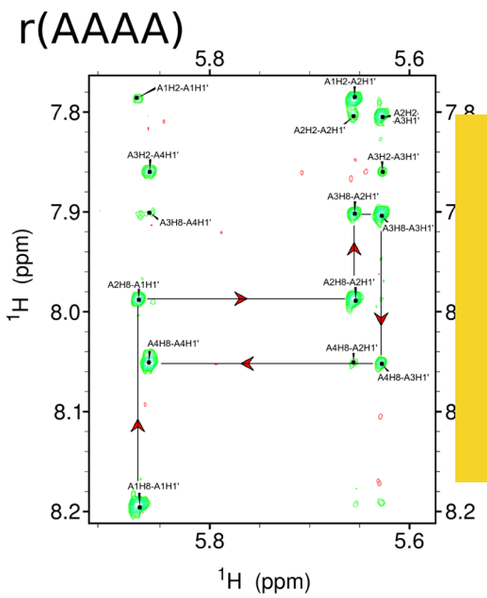


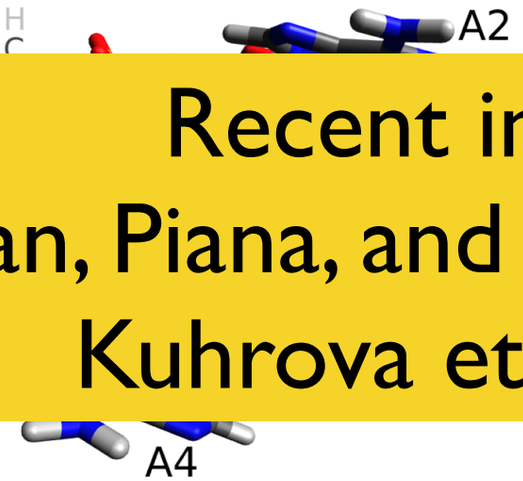


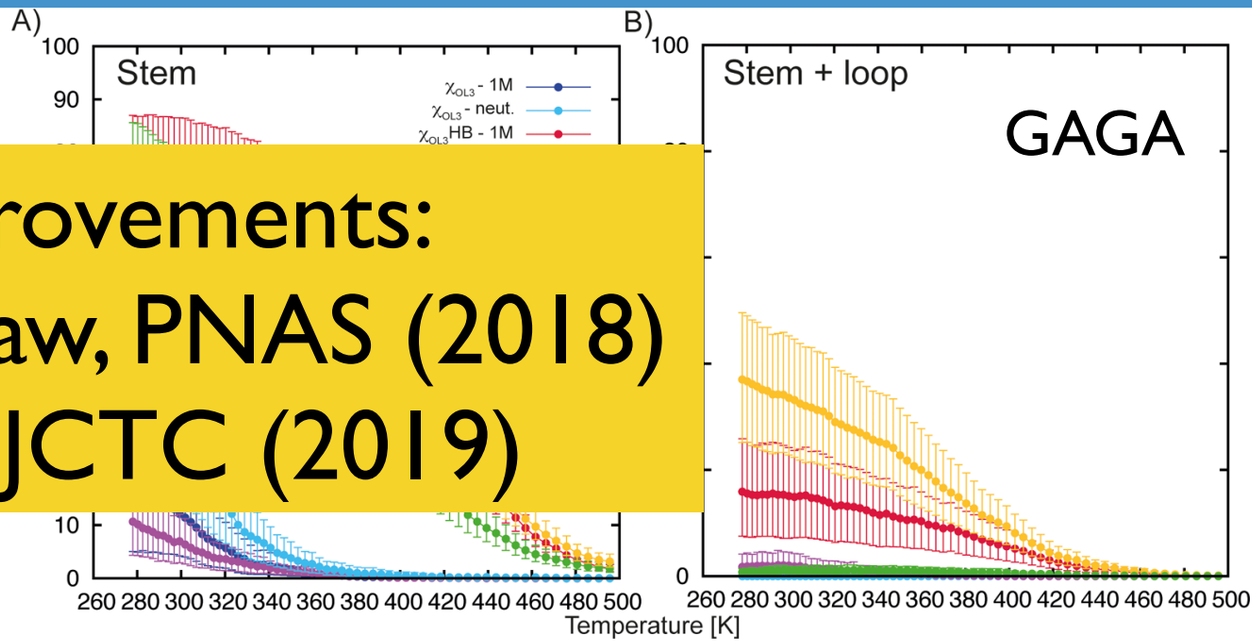
Condon et al, JCTC (2015)
Bergonzo et al, RNA (2015)

Kuhrova et al, JCTC (2016)
Bottaro et al, JPCL (2016)

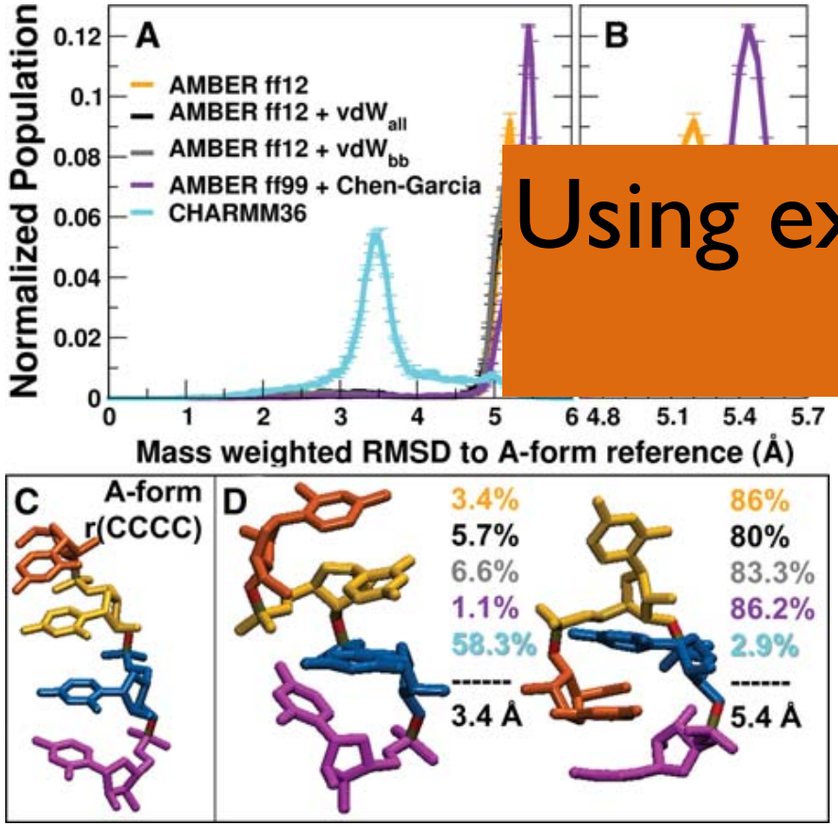
Issues with force fields

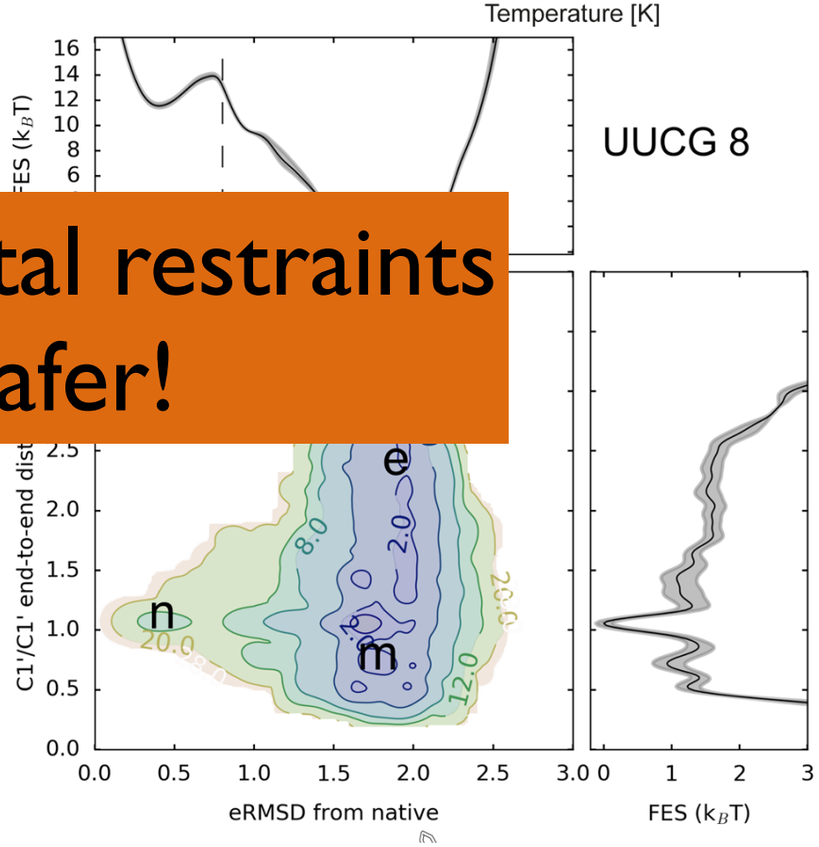


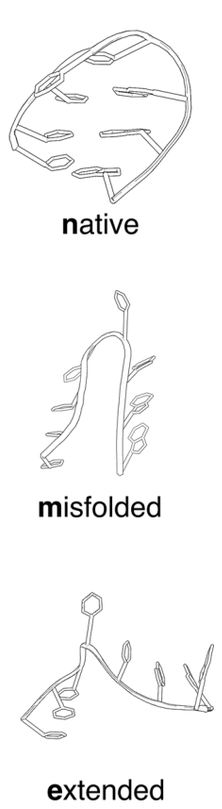




Recent improvements:
Tan, Piana, and Shaw, PNAS (2018)
Kuhrova et al JCTC (2019)







Using experimental restraints is much safer!

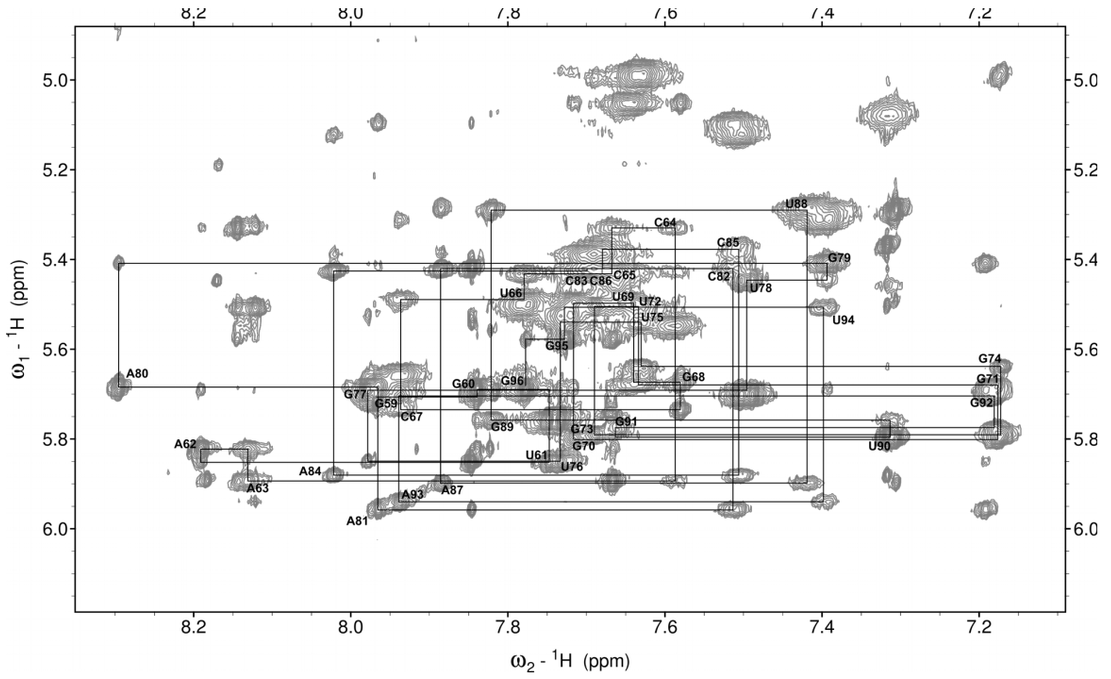
Condon et al, JCTC (2015)
Bergonzo et al, RNA (2015)

Kuhrova et al, JCTC (2016)
Bottaro et al, JPCL (2016)

Solution-phase experiments

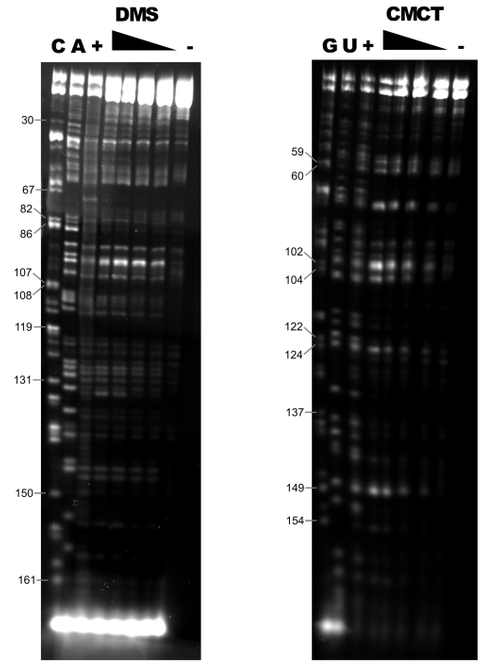
<u>NMR</u>

- Proton distances, angles, etc
- Many different techniques
- Standard *forward models**



<u>Chemical probing</u>

- Identify reactive nucleotides
- Different probes (e.g. DMS/CMCT/SHAPE)
- Non-standard *forward models***



*i.e. "formula to compute experiment from structure" (e.g. Karplus formulas for J_c)

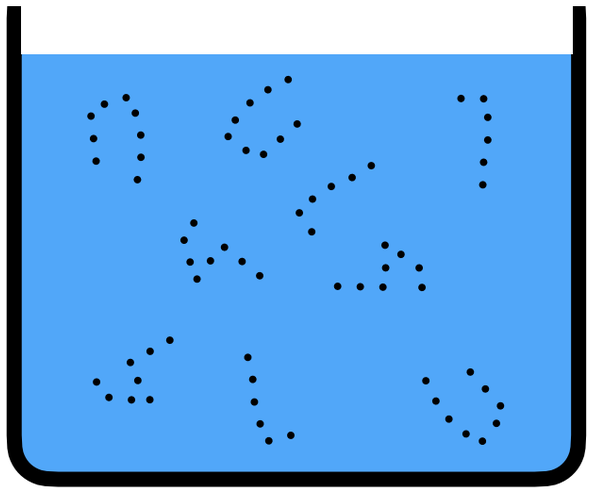
**usually reactive nucleotides are interpreted as "non WC paired"

(pictures from Podbevsek et al, Sci Rep 2018)

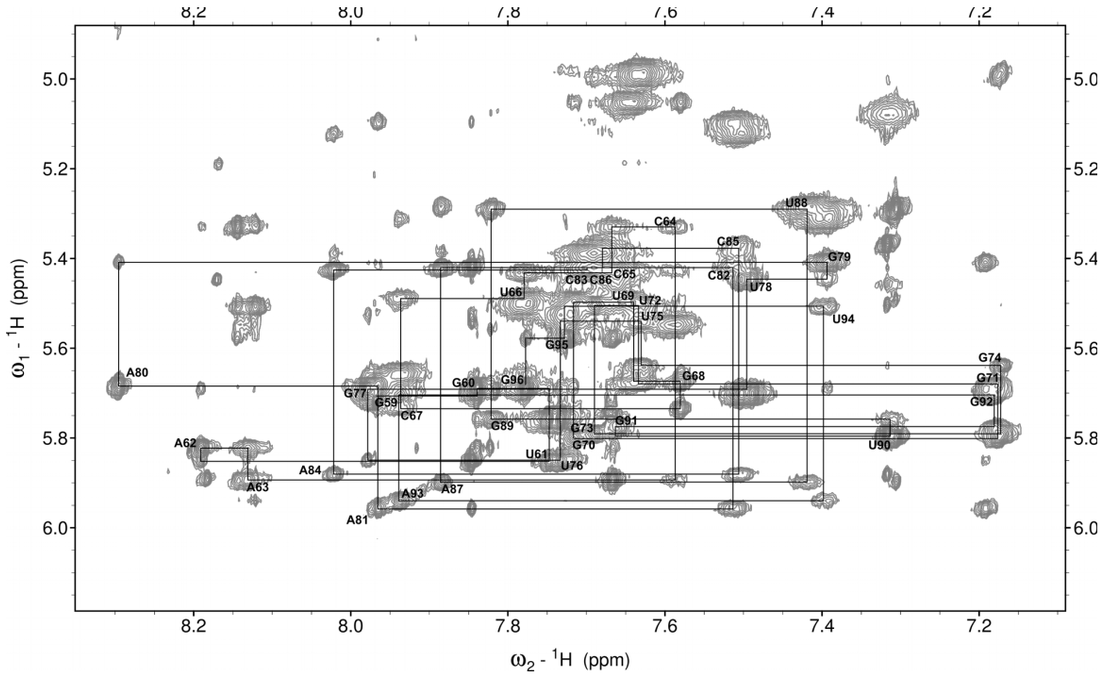
Solution-phase experiments

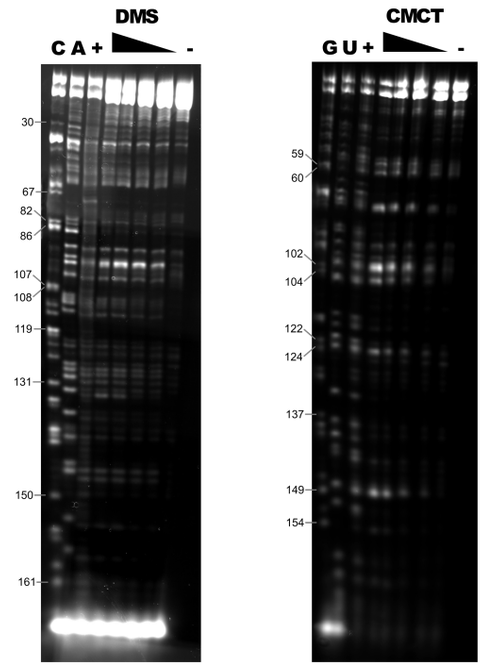
<u>NMR</u>

- Proton distances, angles, etc
- Many different techniques
- Standard *forward models**



average





<u>Chemical probing</u>

- Identify reactive nucleotides
- Different probes (e.g. DMS/CMCT/SHAPE)
- Non-standard *forward models***

*i.e. "formula to compute experiment from structure" (e.g. Karplus formulas for J_c)

**usually reactive nucleotides are interpreted as "non WC paired"

(pictures from Podbevsek et al, Sci Rep 2018)

Agenda

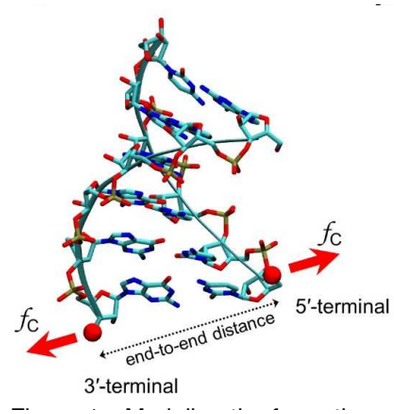
<u>Combine experiment (NMR) and MD</u>

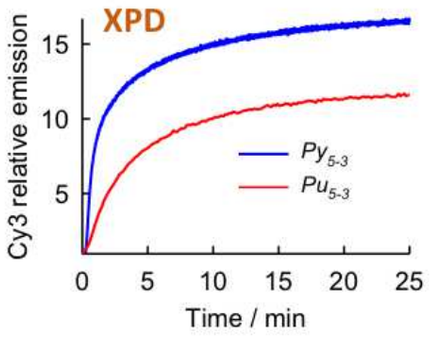
- enforce averages*
- dynamics of a RNA hairpin%

$$S[P] = -\sum_x P(x) \log \frac{P(x)}{Q(x)}$$

<u>RNA (and DNA) unzipping dynamics&</u>

- Go model
- Helicase mediated unwinding
- Experimental validation





*Cesari, Reisser, and Bussi, Computation (2018)
%Podbevsek et al, Sci. Rep. (2018); Reisser et al, NAR (2020)
&Colizzi et al, PNAS (2019)

Enforcing averages using MaxEnt

Enforce expected value
(underdetermined)

$$\langle f \rangle = \sum_x P(x) f(x) = f_{data}$$

Enforcing averages using MaxEnt

Enforce expected value (underdetermined)

$$\langle f \rangle = \sum_x P(x) f(x) = f_{data}$$

Entropy*

$$S[P] = -\sum_x P(x) \log \frac{P(x)}{Q(x)}$$

P=posterior
Q=prior

Jaynes Proc IEEE (1982)
Chodera and Pitera, JCTC (2014)
Cesari, Reisser, Bussi, Computation (2018)

*Entropy measures "how much extra information is needed" besides *Q(x)*

Enforcing averages using MaxEnt

Enforce expected value (underdetermined)

$$\langle f \rangle = \sum_x P(x) f(x) = f_{data}$$

Entropy*

$$S[P] = -\sum_x P(x) \log \frac{P(x)}{Q(x)}$$

P=posterior
Q=prior

MaxEnt

$$\frac{\delta S}{\delta P(x)} = 0 \longrightarrow \begin{aligned} P(x) &\propto Q(x) e^{-\lambda f(x)} \\ U' &= U + \lambda k_B T f(x) \end{aligned}$$

Jaynes Proc IEEE (1982)
Chodera and Pitera, JCTC (2014)
Cesari, Reisser, Bussi, Computation (2018)

*Entropy measures "how much extra information is needed" besides *Q(x)*

Enforcing averages using MaxEnt

Enforce expected value (underdetermined)

$$\langle f \rangle = \sum_x P(x) f(x) = f_{data}$$

Entropy*

$$S[P] = -\sum_x P(x) \log \frac{P(x)}{Q(x)}$$

P=posterior
Q=prior

MaxEnt

$$\frac{\delta S}{\delta P(x)} = 0 \longrightarrow \begin{aligned} P(x) &\propto Q(x) e^{-\lambda f(x)} \\ U' &= U + \lambda k_B T f(x) \end{aligned}$$

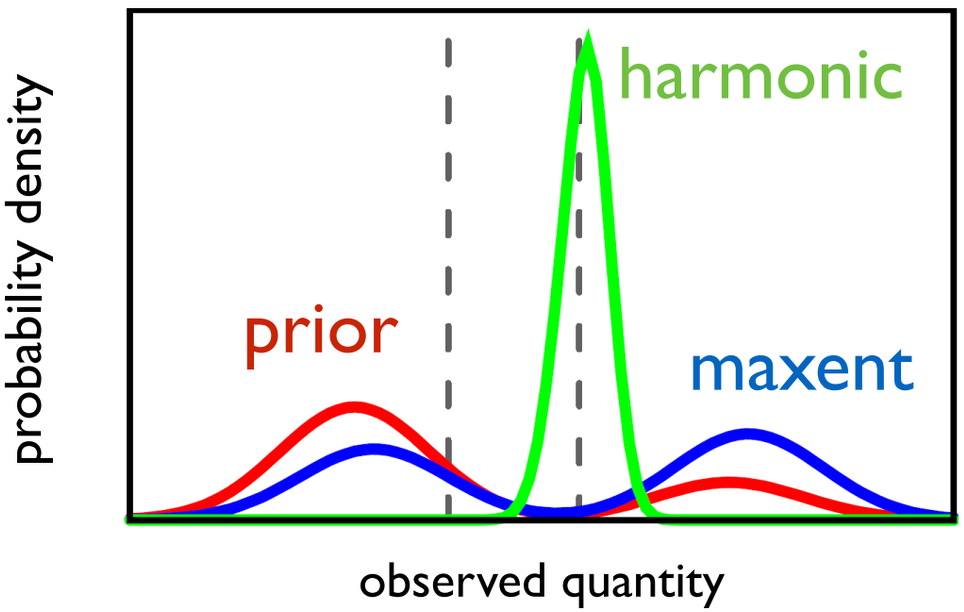
λ to be chosen such that average on *P* agrees with exp.

$$\frac{\sum_x Q(x) e^{-\lambda f(x)} f(x)}{\sum_x Q(x) e^{-\lambda f(x)}} = f_{data}$$

Jaynes Proc IEEE (1982)
Chodera and Pitera, JCTC (2014)
Cesari, Reisser, Bussi, Computation (2018)

*Entropy measures "how much extra information is needed" besides *Q(x)*

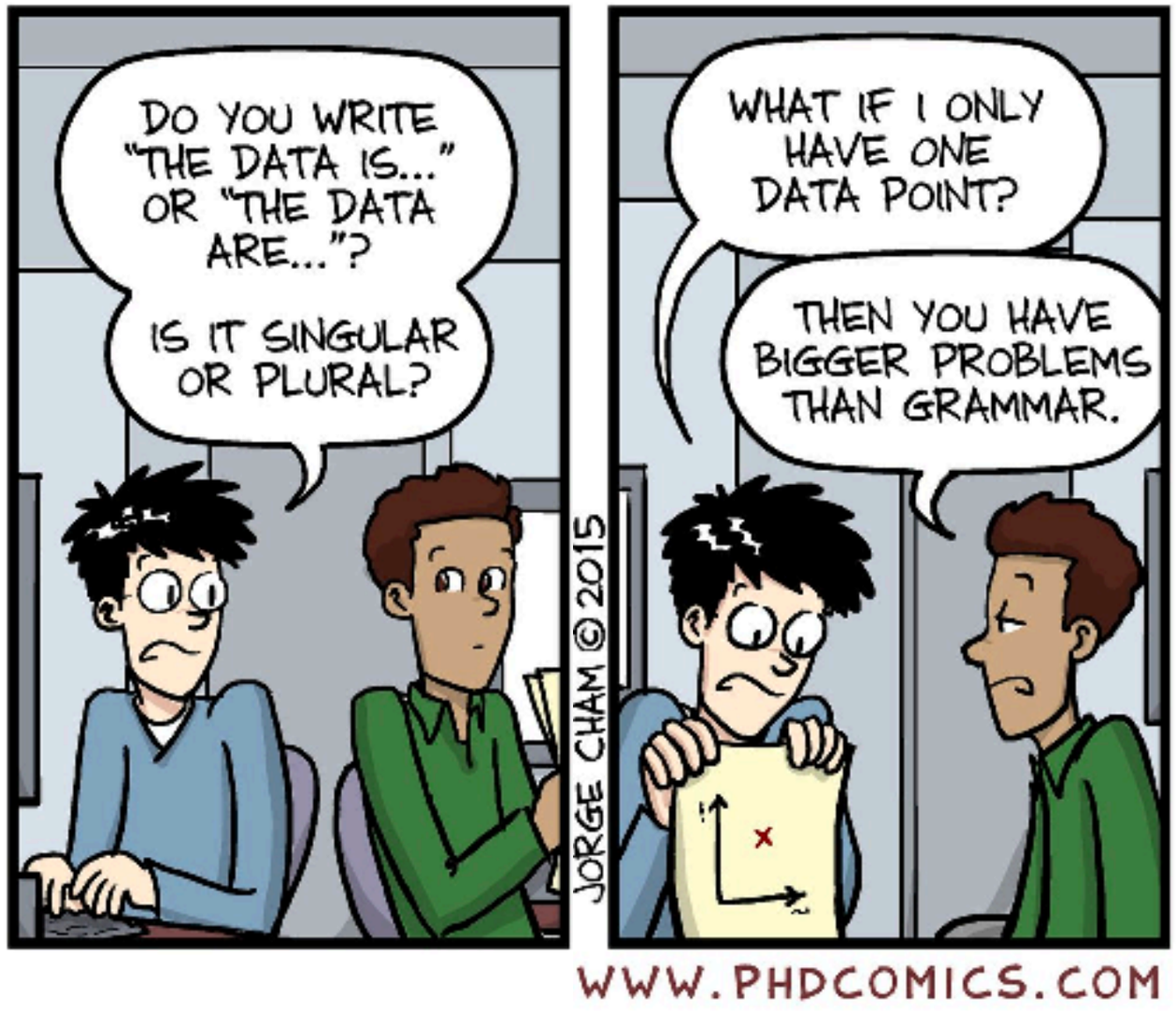
Comments on MaxEnt



Much information from prior is retained

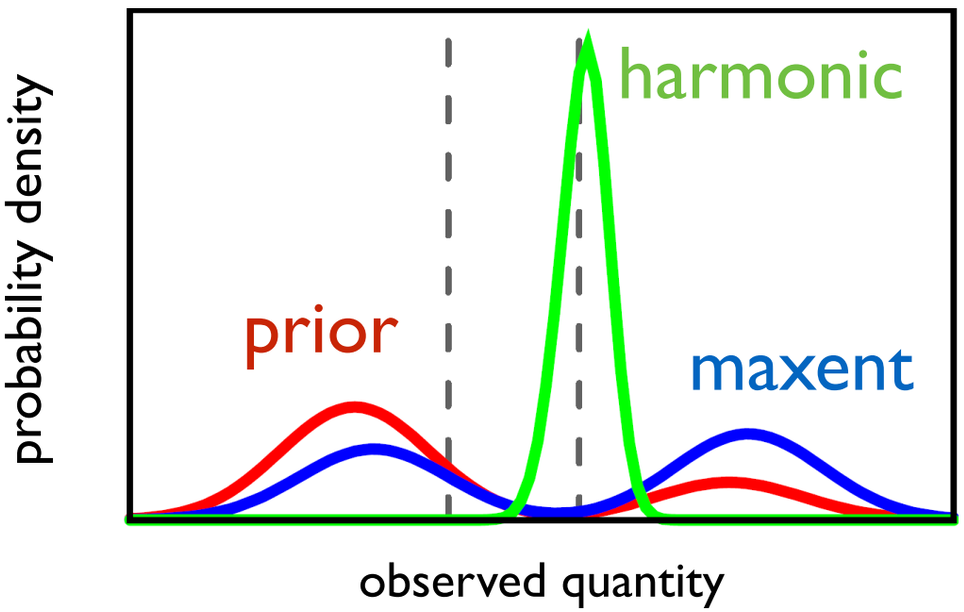
↓

Very good prior required
(when compared with harmonic restraint)



see Bonomi et al COSB (2017) for a review on "MaxEnt" and "Maximum Parsimony" methods

Comments on MaxEnt

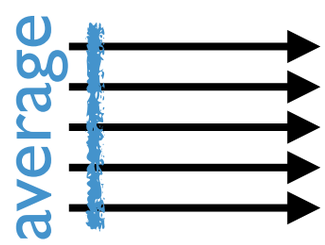


Much information from prior is retained

↓

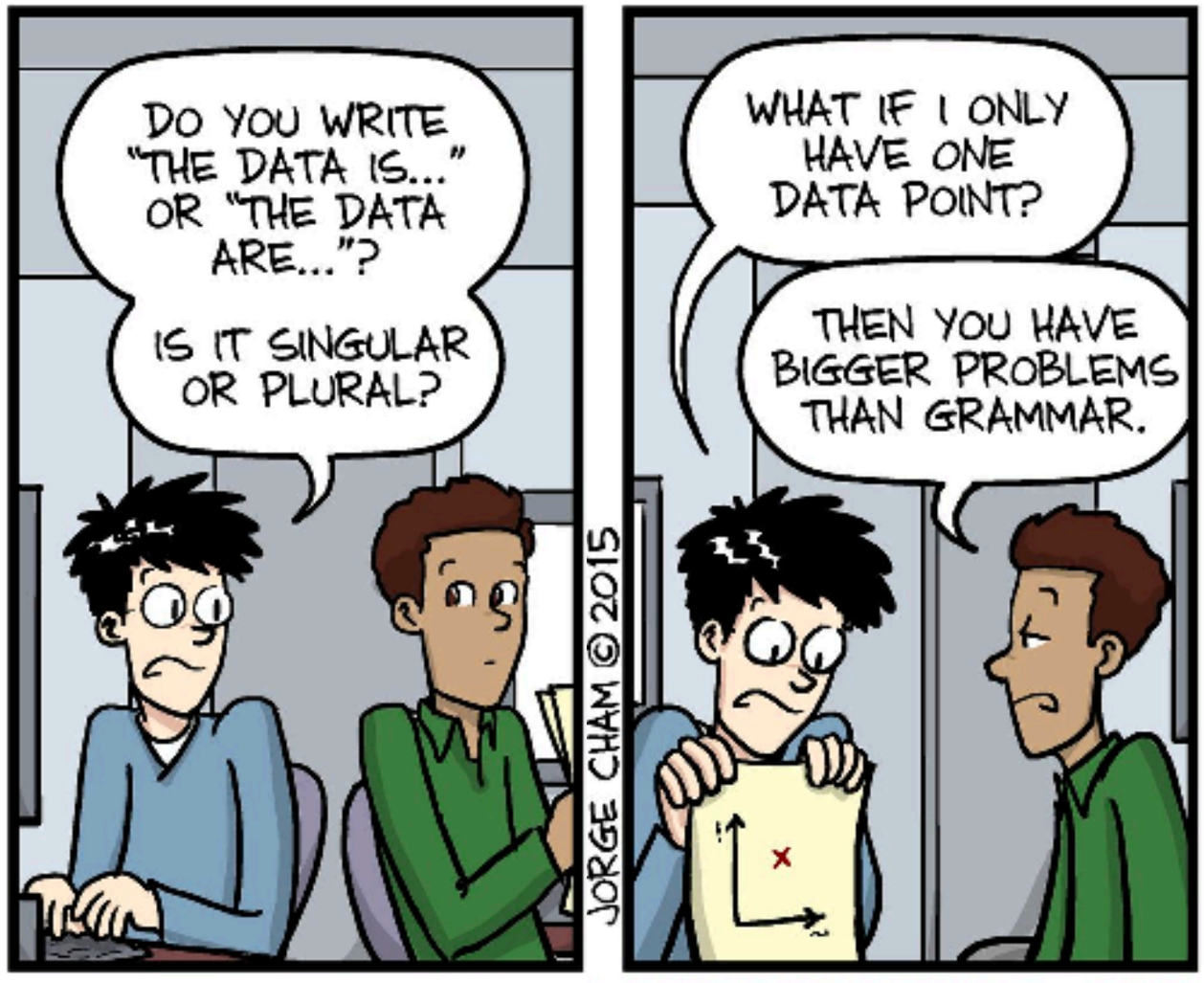
Very good prior required
(when compared with harmonic restraint)

Equivalent to
multi-replica approaches
(e.g. Vendruscolo group)



Closely related to VES*
(Valsson & Parrinello)

Equivalent to EDS*
(White, Hocky, and Voth)



*see our review on Cesari, Reisser, and Bussi, Computation (2018) for a detailed comparison
see Bonomi et al COSB (2017) for a review on "MaxEnt" and "Maximum Parsimony" methods

Comments on MaxEnt

Two possible routes:

<u>Enforce</u>

run MD subject to (average) restraints

<u>Reweighting</u>

a posteriori assign weights to simulated snapshots

(see Rangan et al, JCTC 2018 for a comparison)

rom prior is retained

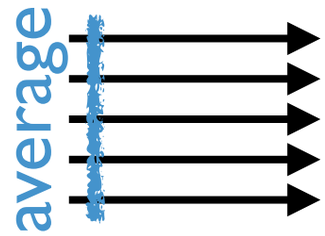
↓

prior required

th harmonic restraint)

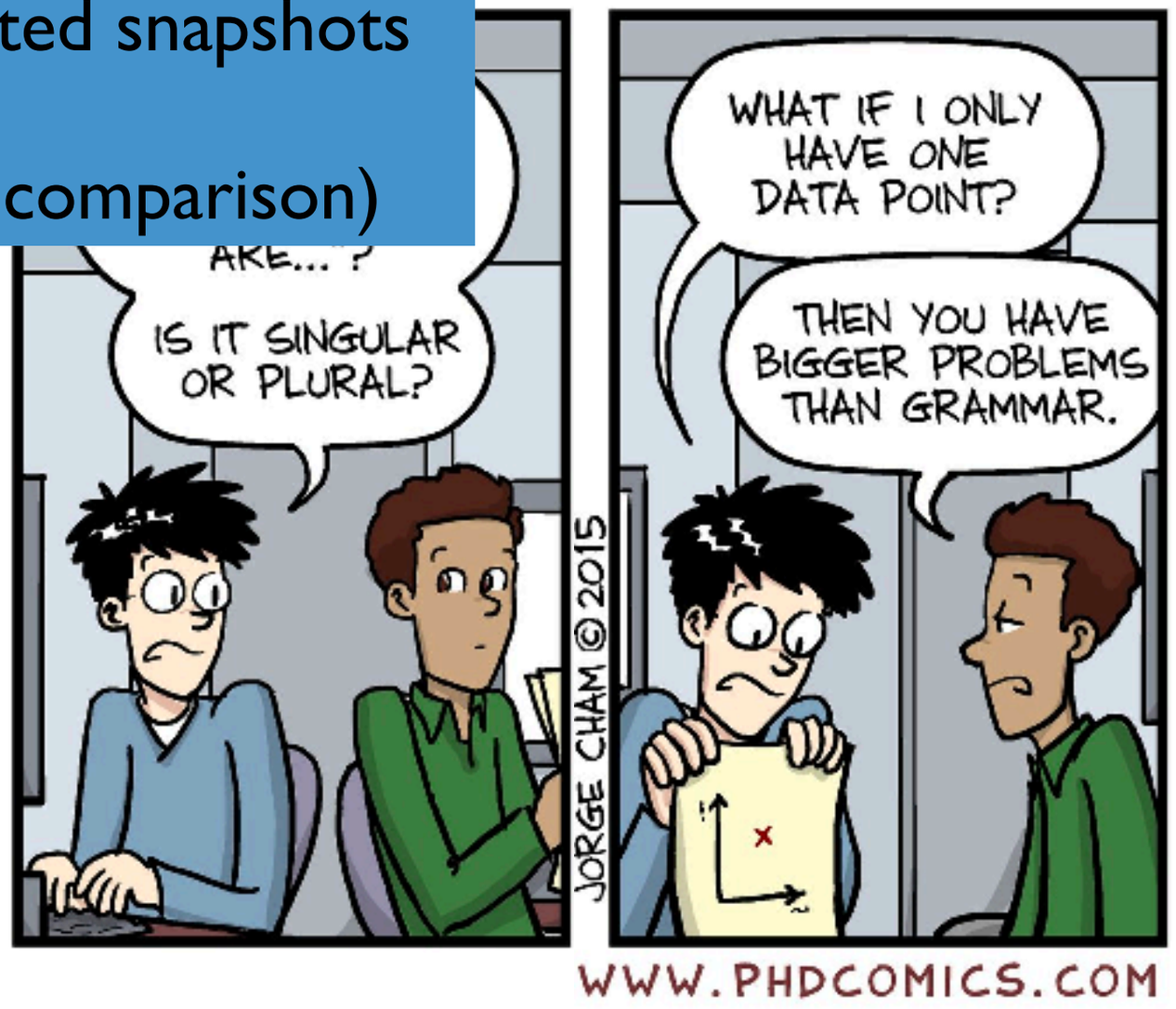
probability density

Equivalent to
multi-replica approaches
(e.g. Vendruscolo group)



Closely related to VES*
(Valsson & Parrinello)

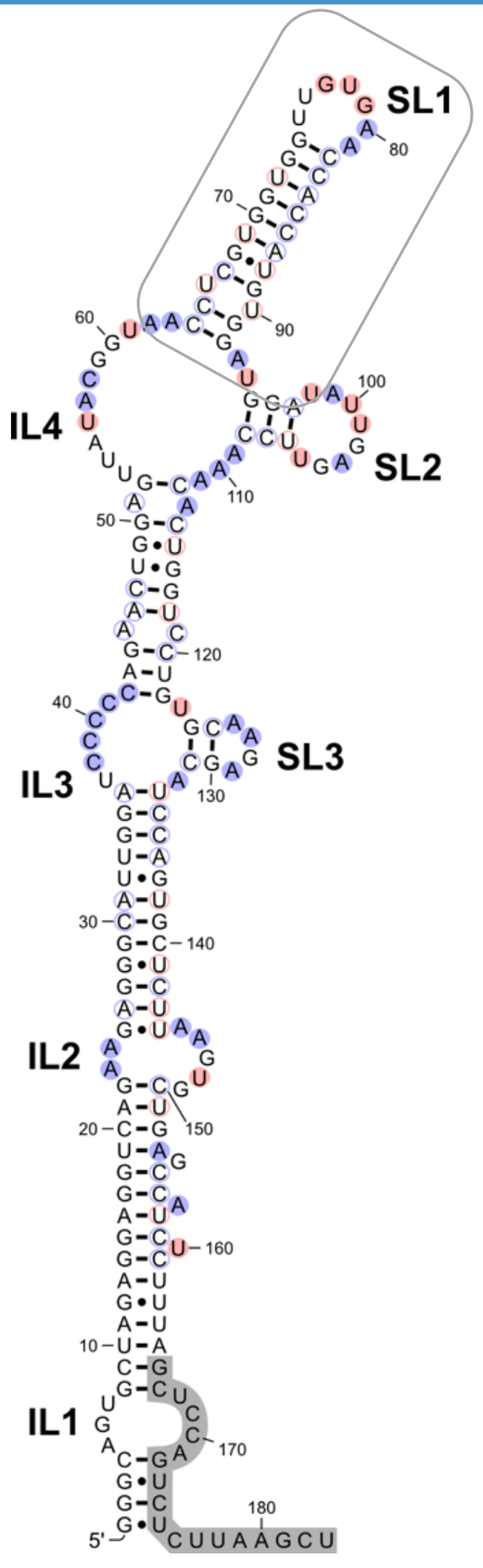
Equivalent to EDS*
(White, Hocky, and Voth)

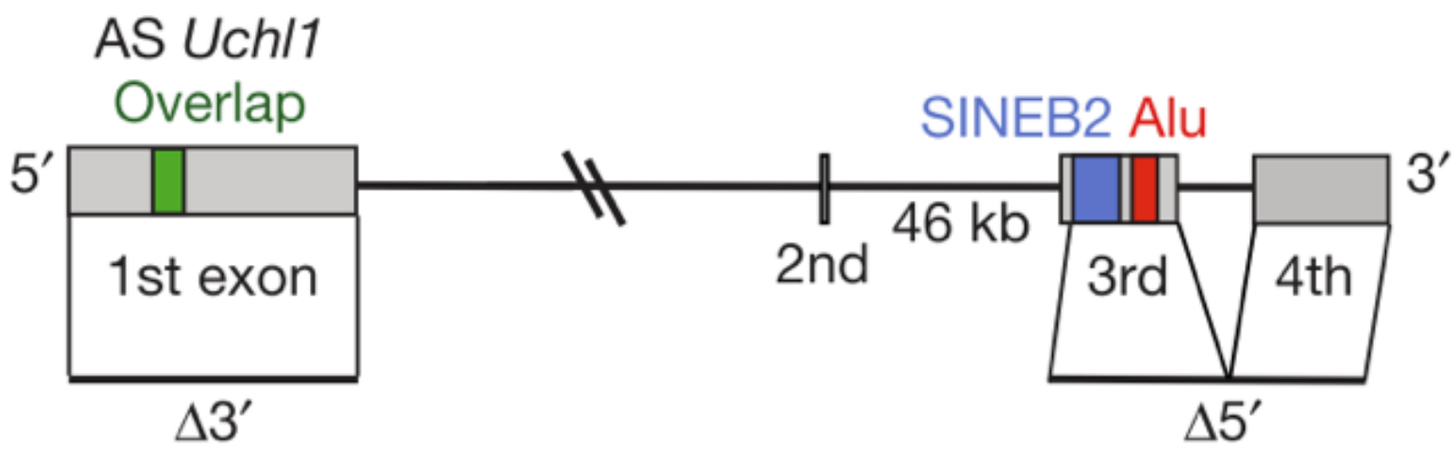


*see our review on Cesari, Reisser, and Bussi, Computation (2018) for a detailed comparison

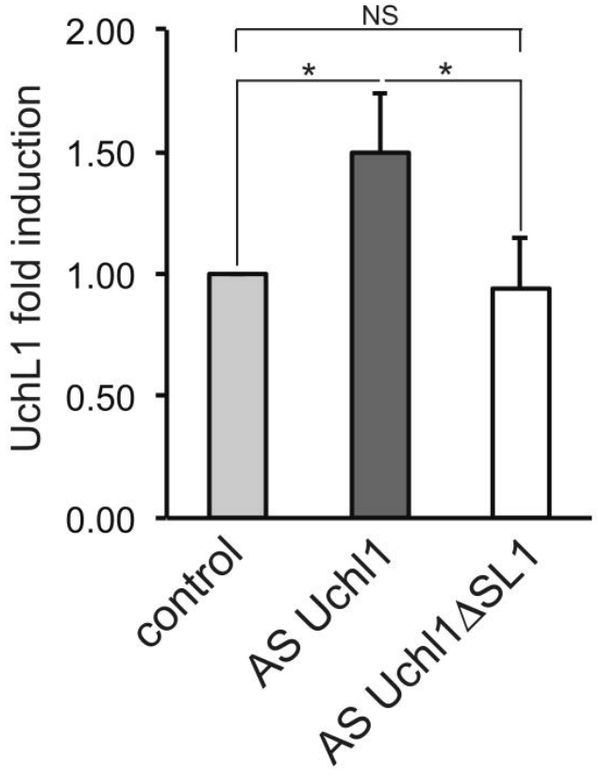
see Bonomi et al COSB (2017) for a review on "MaxEnt" and "Maximum Parsimony" methods

SINE B2 elements





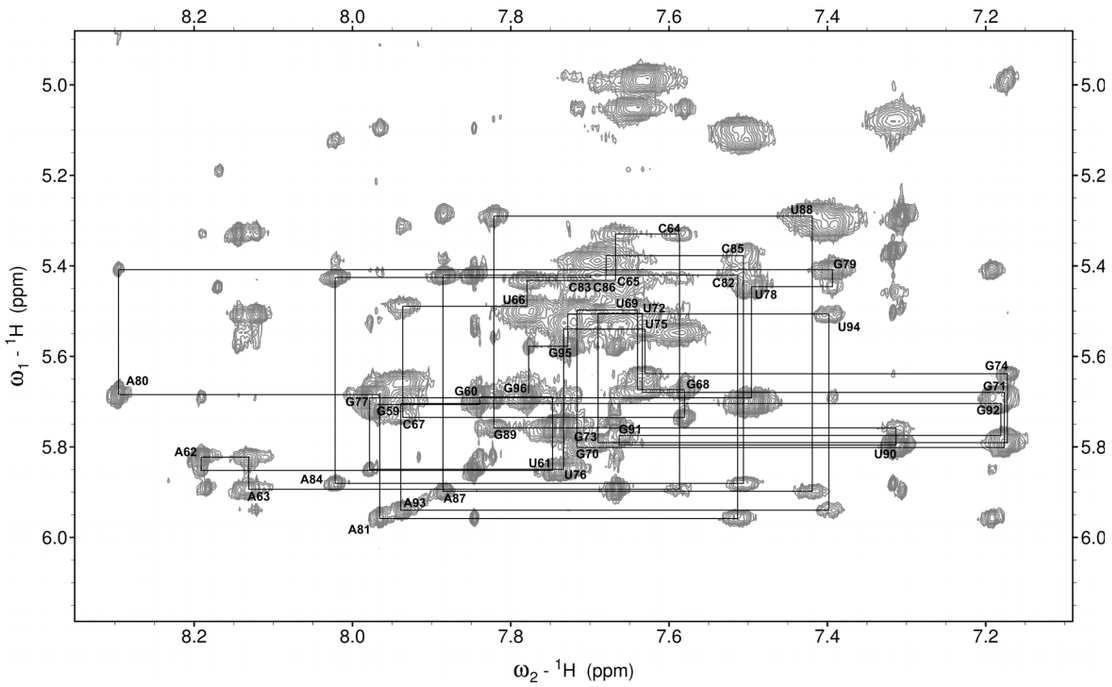
Inverted SINEB2 element enhances protein translation*



Deletion studies identified the 29-nt terminal hairpin SL1 as essential for protein synthesis enhancing%

*Carrieri et al, Nature (2014)
%Podbevsek, et al Sci. Rep. (2018)

Solution data (hairpin)



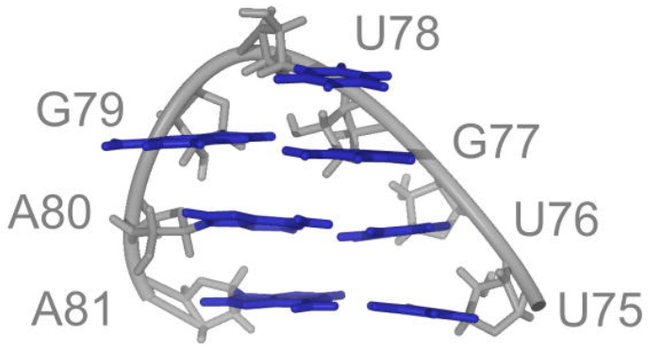
$$d_{NOE_{i,exp}}(\text{strong}) = 3.6\text{Å} \rightarrow \langle 1/d_{H-H}^6 \rangle^{-1/6} < 3.6\text{Å}$$

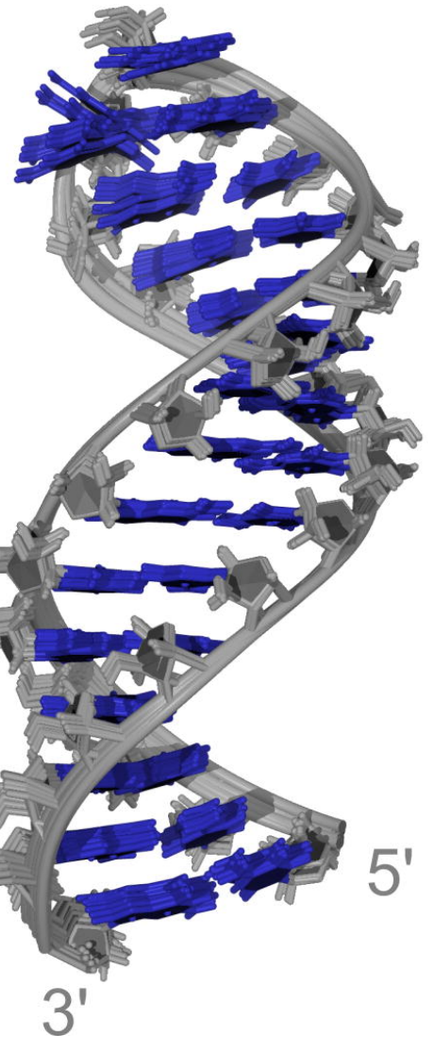
$$d_{NOE_{i,exp}}(\text{medium}) = 5.0\text{Å} \rightarrow \langle 1/d_{H-H}^6 \rangle^{-1/6} < 5.0\text{Å}$$

$$d_{NOE_{i,exp}}(\text{weak}) = 6.5\text{Å} \rightarrow \langle 1/d_{H-H}^6 \rangle^{-1/6} < 6.5\text{Å}$$

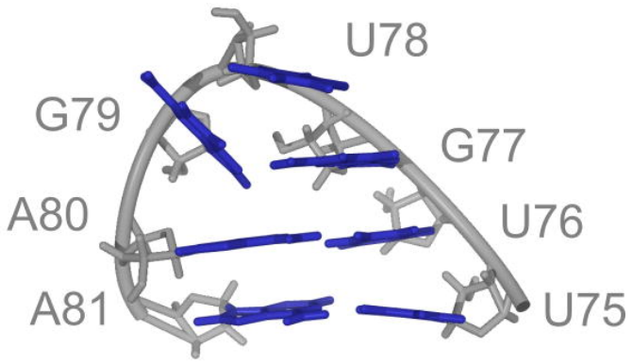
2D NOESY

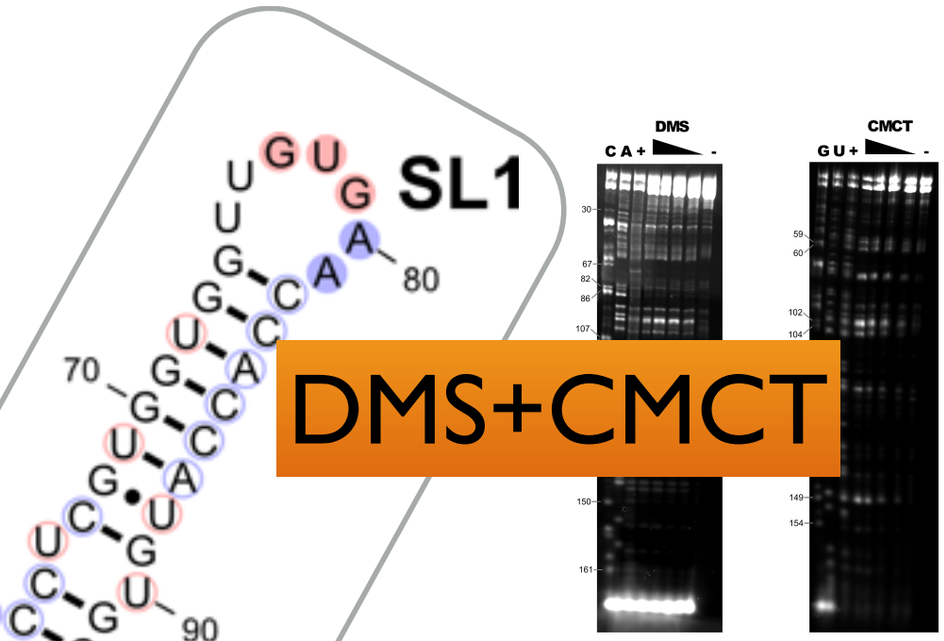
125 signals
(strong/medium/weak)





2 similar models
resolved PDB: 5LSN%





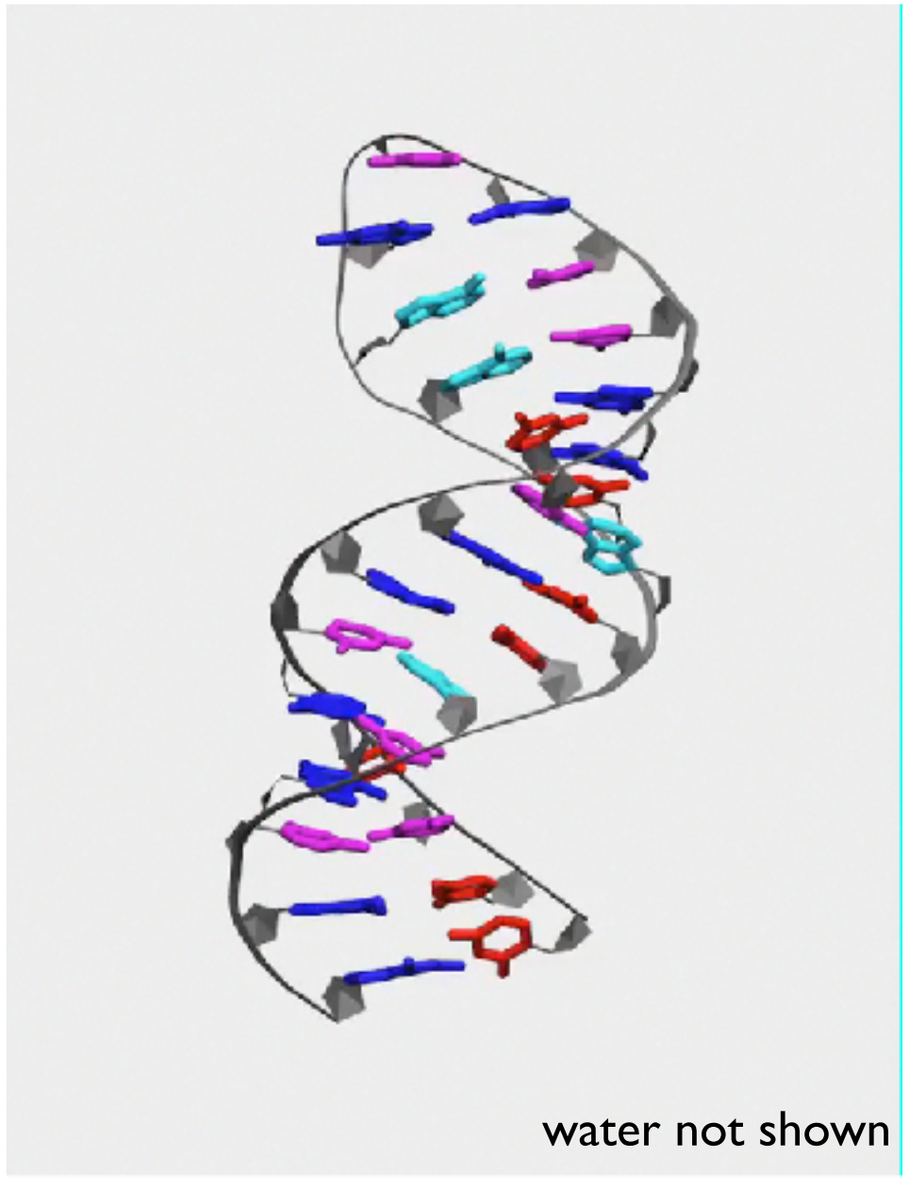
%Podbevsek, et al Sci. Rep. (2018)

A posteriori reweighting

Run a long MD from NMR structure
(just dynamics close to native)

Several NOEs are violated

Reweight to enforce NOE signals



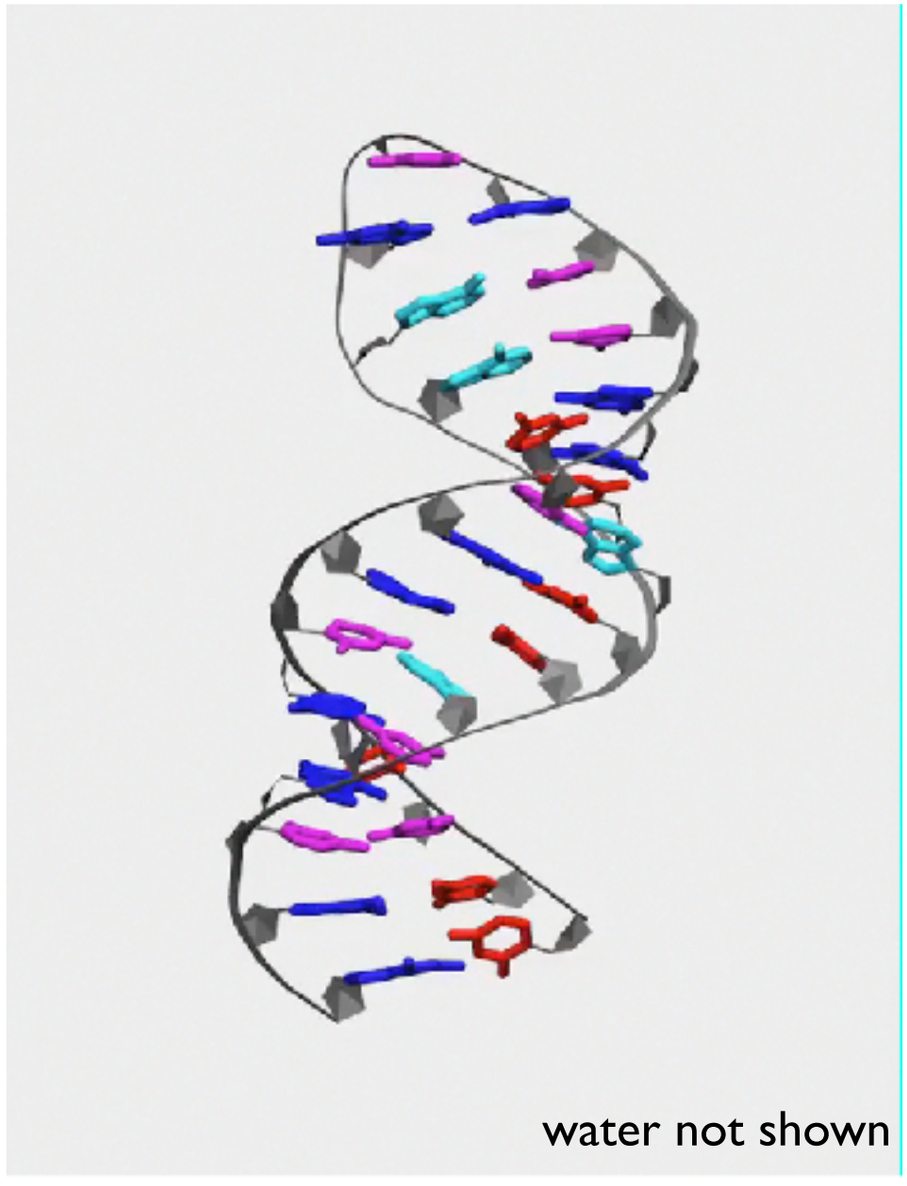
Podbevsek, et al Sci. Rep. (2018)

A posteriori reweighting

Run a long MD from NMR structure
(just dynamics close to native)

Several NOEs are violated

Reweight to enforce NOE signals



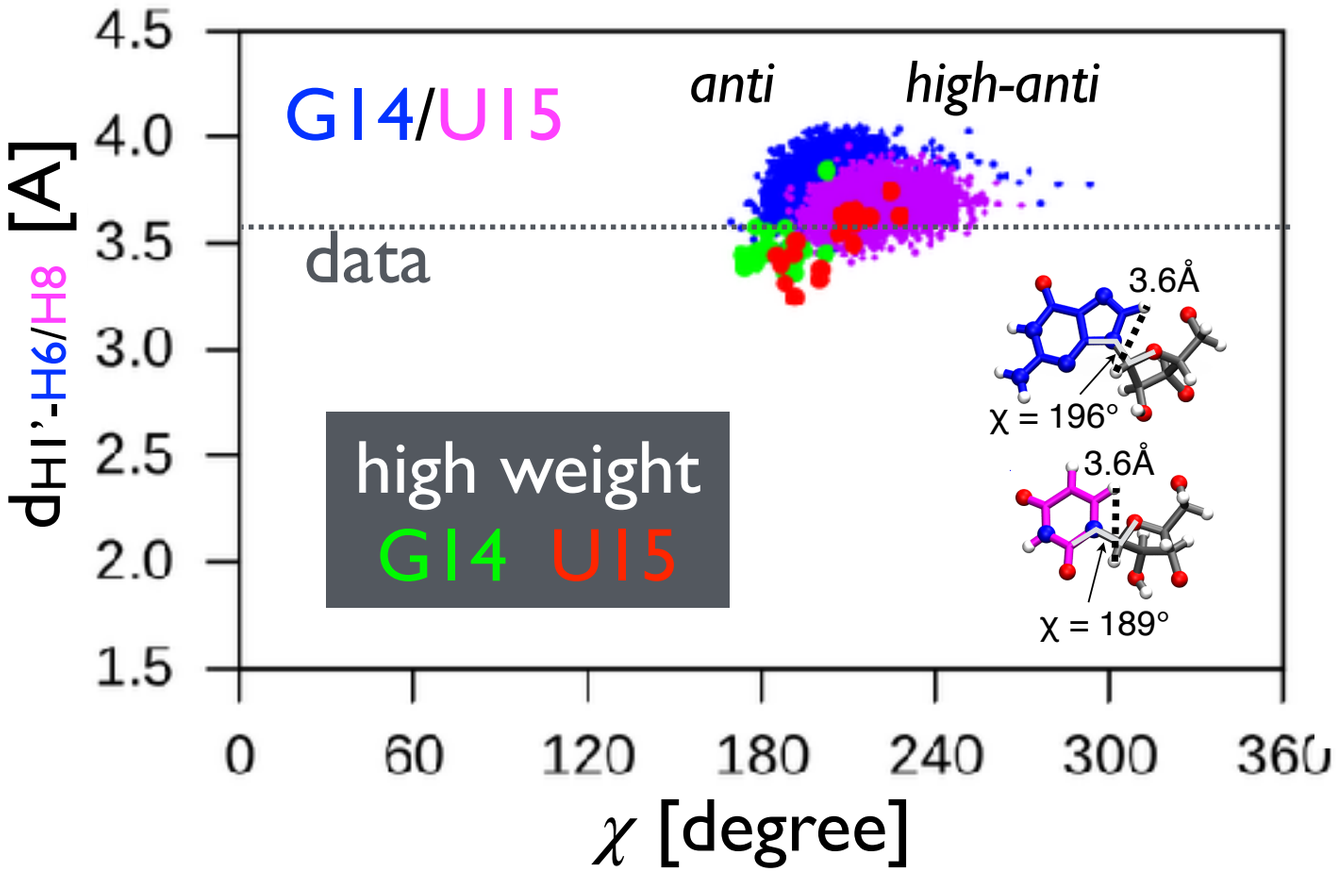
Podbevsek, et al Sci. Rep. (2018)

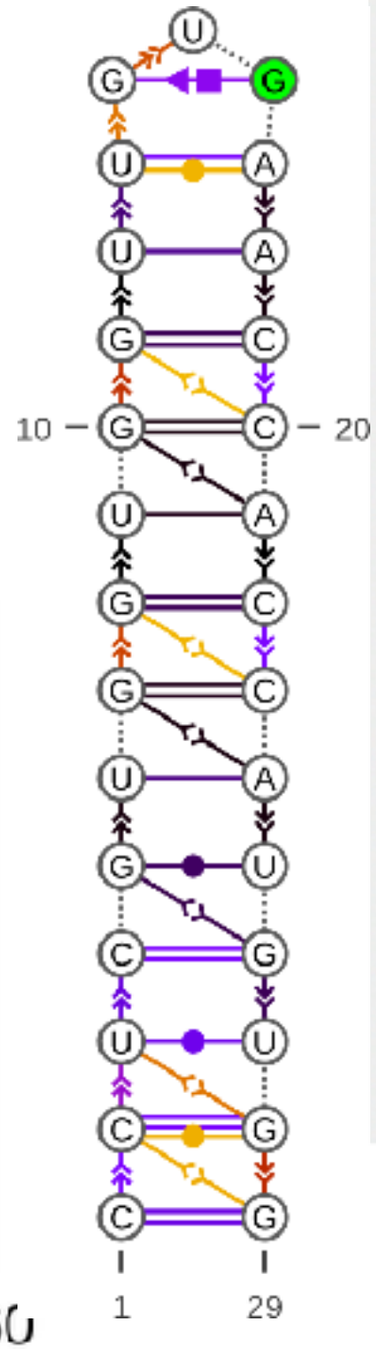
A posteriori reweighting

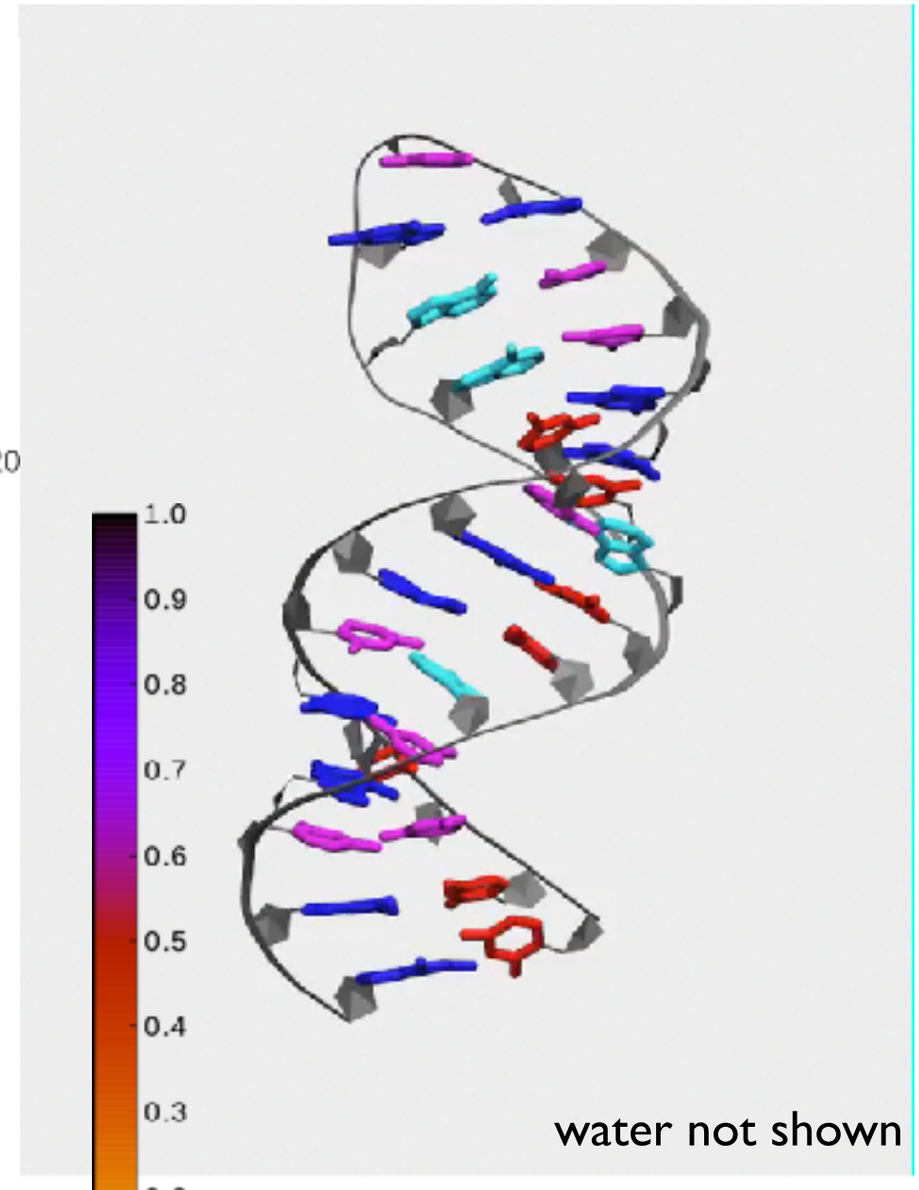
Run a long MD from NMR structure
(just dynamics close to native)

Several NOEs are violated

Reweight to enforce NOE signals







Kish size ~ 0.002

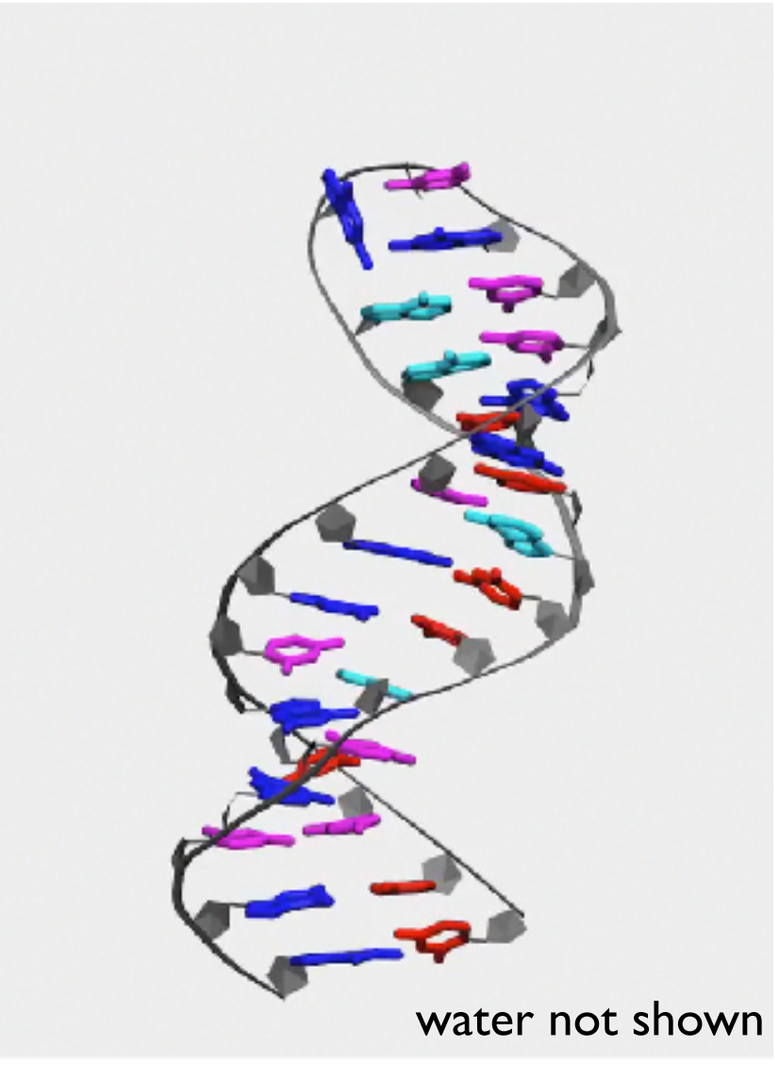
Podbevsek, et al Sci. Rep. (2018)

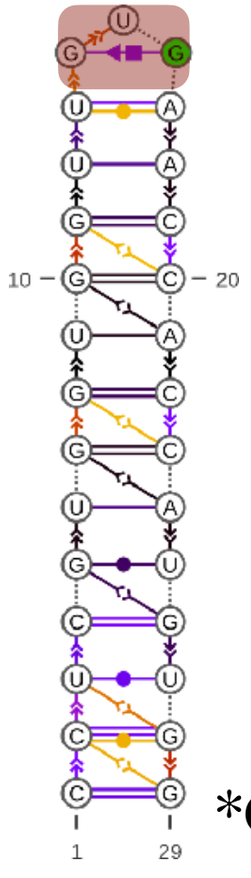
Enhanced sampling on SINE hairpin

8 Replicas (RECT*)

Accelerated degrees of freedom:
χ14-16 and coordination numbers 14-16

MaxEnt for 125 NOEs throughout the whole hairpin





Kish's sample size ~ 0.7

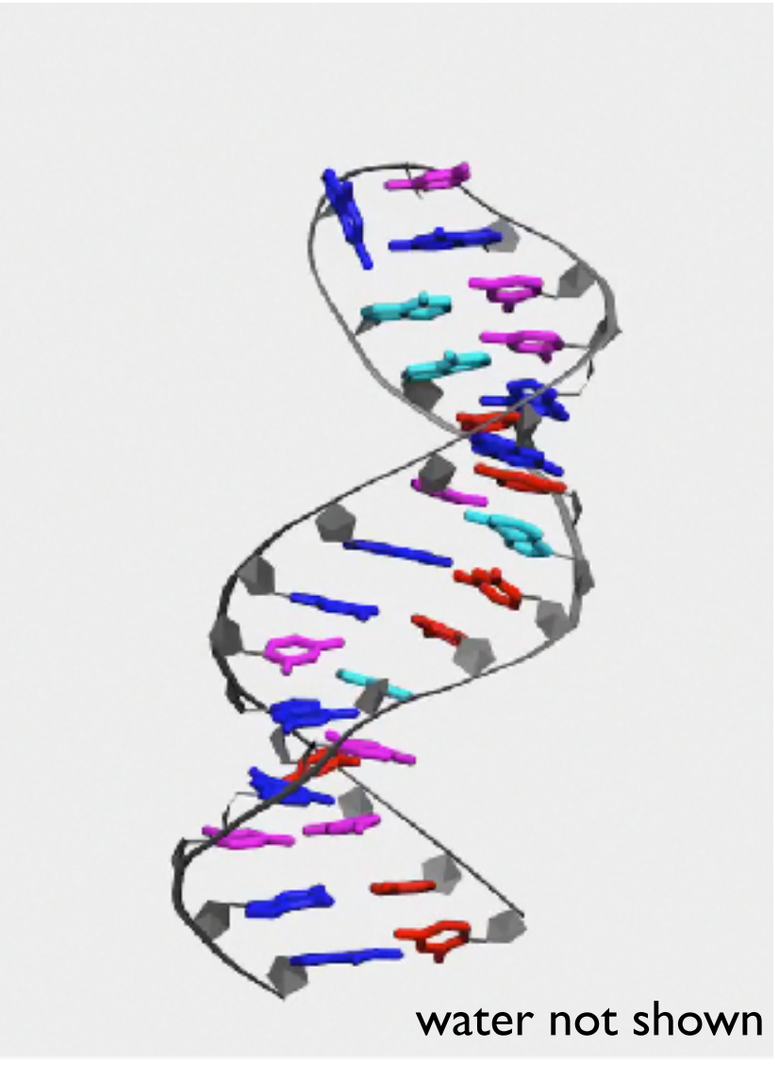
*Gil-Ley and Bussi, JCTC (2015)
Reisser et al, NAR (2020)

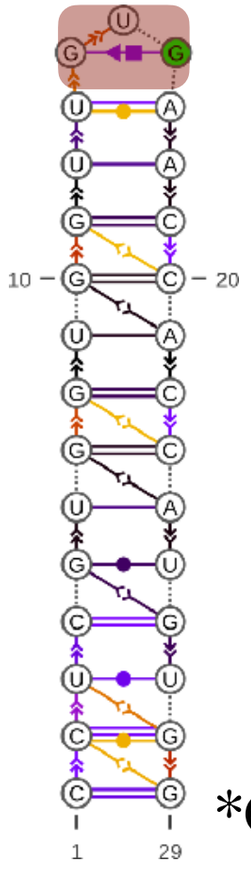
Enhanced sampling on SINE hairpin

8 Replicas (RECT*)

Accelerated degrees of freedom:
χ14-16 and coordination numbers 14-16

MaxEnt for 125 NOEs throughout the whole hairpin





Kish's sample size ~ 0.7

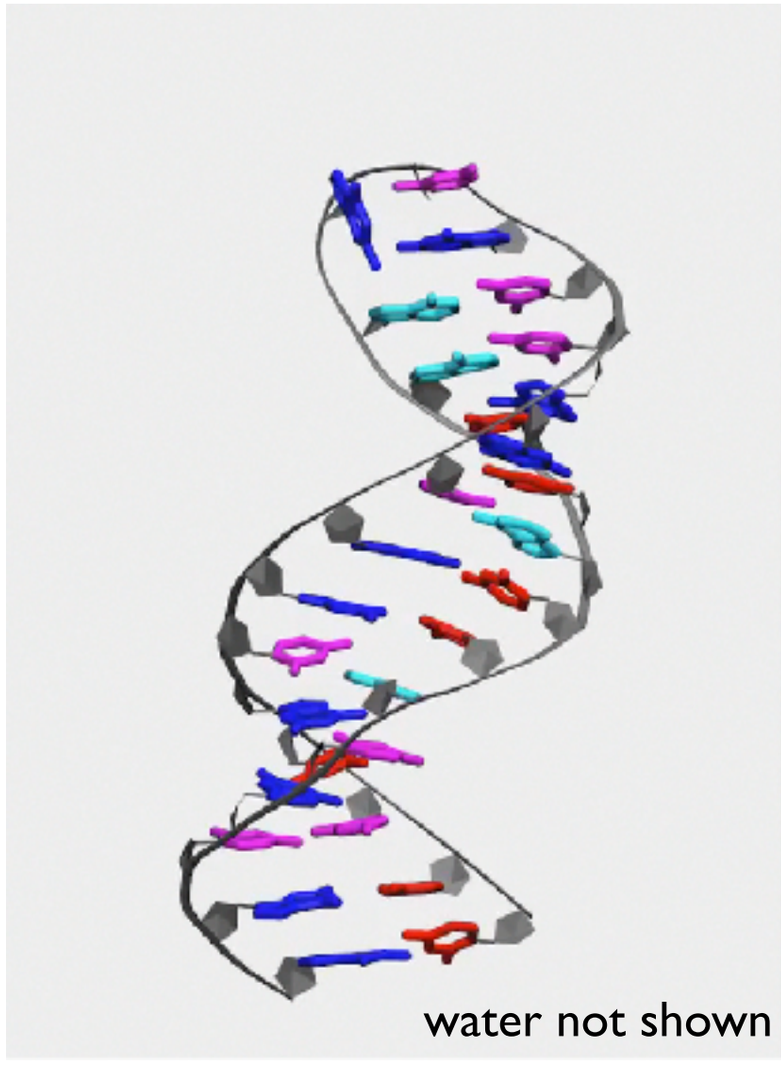
*Gil-Ley and Bussi, JCTC (2015)
Reisser et al, NAR (2020)

Enhanced sampling on SINE hairpin

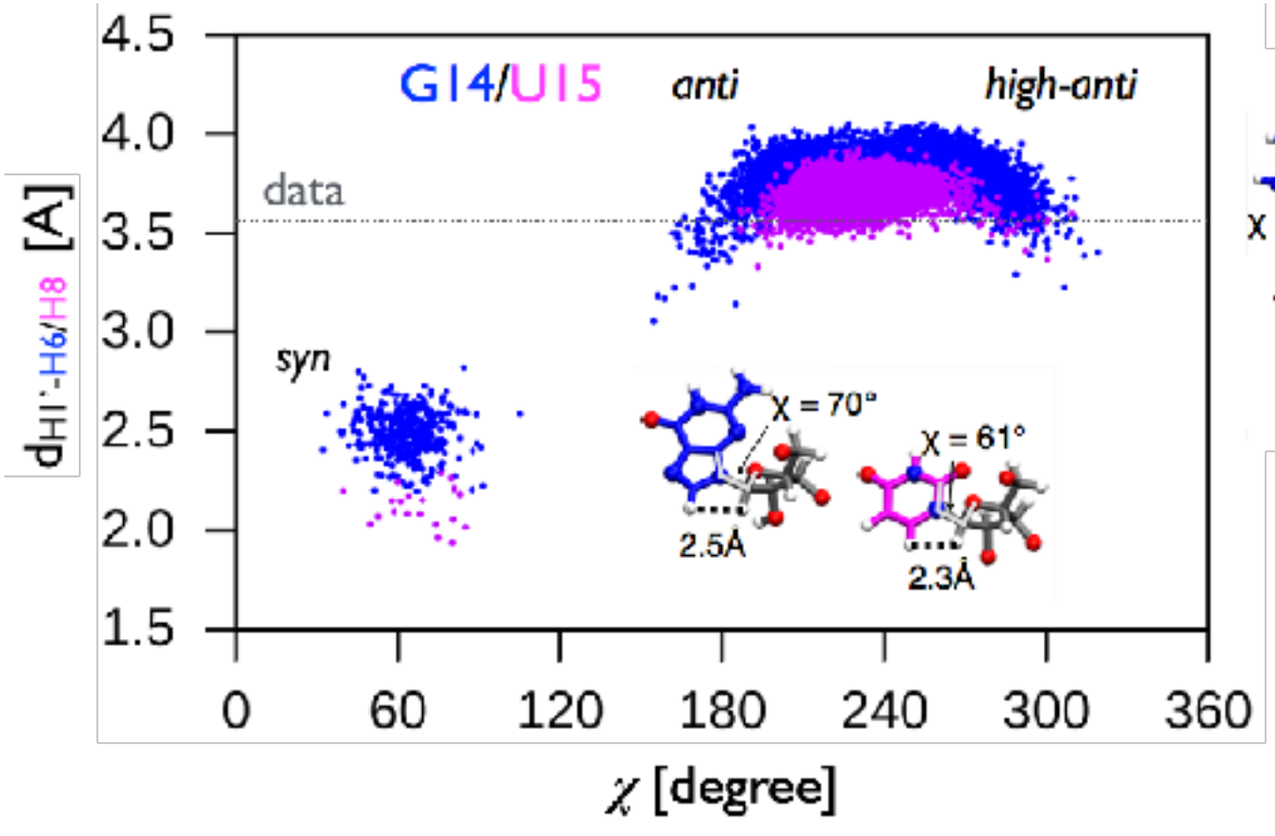
8 Replicas (RECT*)

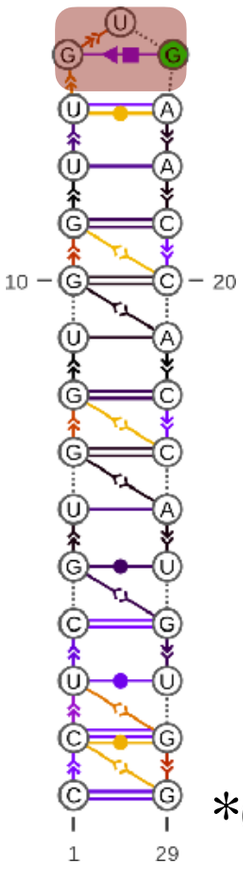
Accelerated degrees of freedom:
χ14-16 and coordination numbers 14-16

MaxEnt for 125 NOEs throughout the whole hairpin



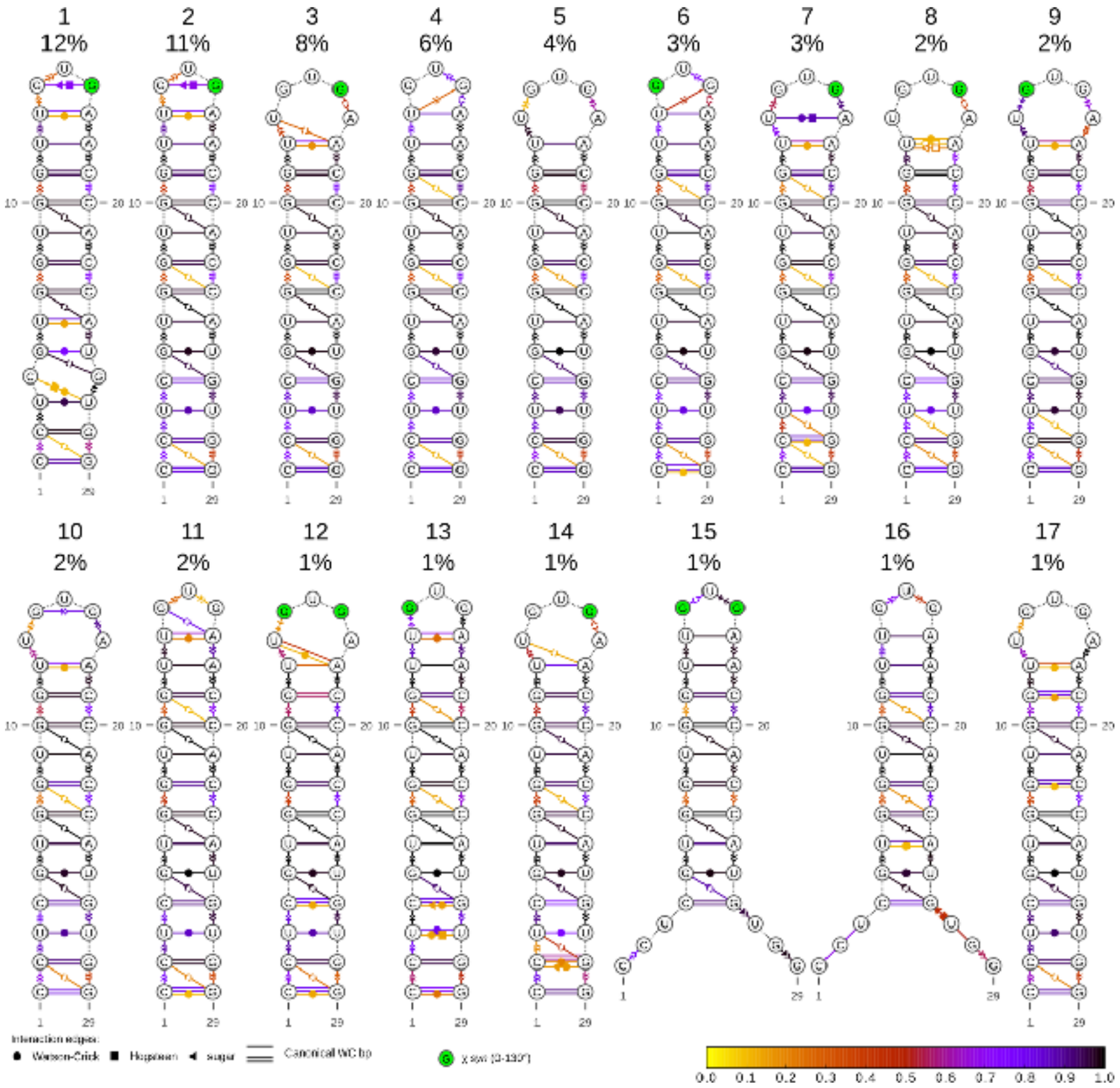
Kish's sample size ~ 0.7





*Gil-Ley and Bussi, JCTC (2015)
Reisser et al, NAR (2020)

Clustering

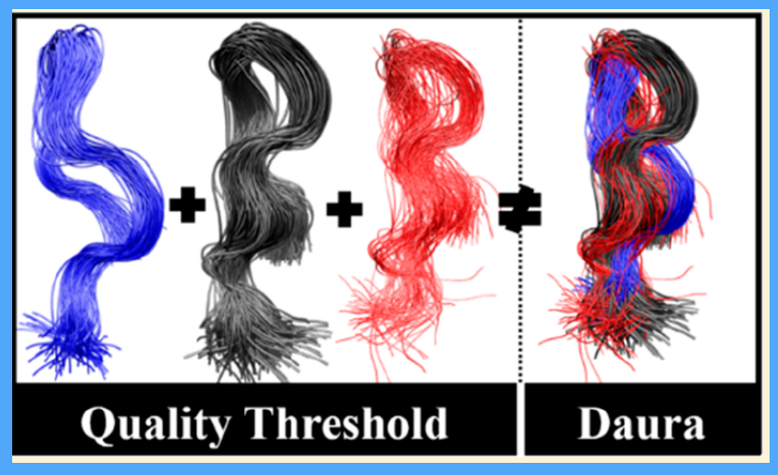


Annotated and plotted with Barnaba, Bottaro et al, RNA (2019) `(pip install barnaba)`

Clustering

Clusters are homogeneous (low eRMSD* and same χ-pattern) thanks to an ad-hoc (expensive) clustering method based on maximum cliques@

See González-Alemán et al, JCIM (2020) for a quality-threshold clustering& with similar properties and for a comparison with the popular "GROMOS" method$



*Bottaro et al, NAR (2014)
@Reisser et al, NAR (2020)
&Heyer et al, Genome Res (1999)
$Daura et al, ACIE (1999)

Annotated and plotted with Barnaba, Bottaro et al, RNA (2019) `(pip install barnaba)`

Maximum parsimony

Select the smallest number of conformational clusters that can explain exp data

$$\langle f_{NOE}(d_i)\rangle_{\text{set}} = \sum_{y \in \text{set}}^{Y} w'_y \langle f_{NOE}(d_i)\rangle_y$$

$$D_{KL}(w'_y||P_y) = \sum_{y \in \text{set}} w'_y \ln \frac{w'_y}{P_y}$$

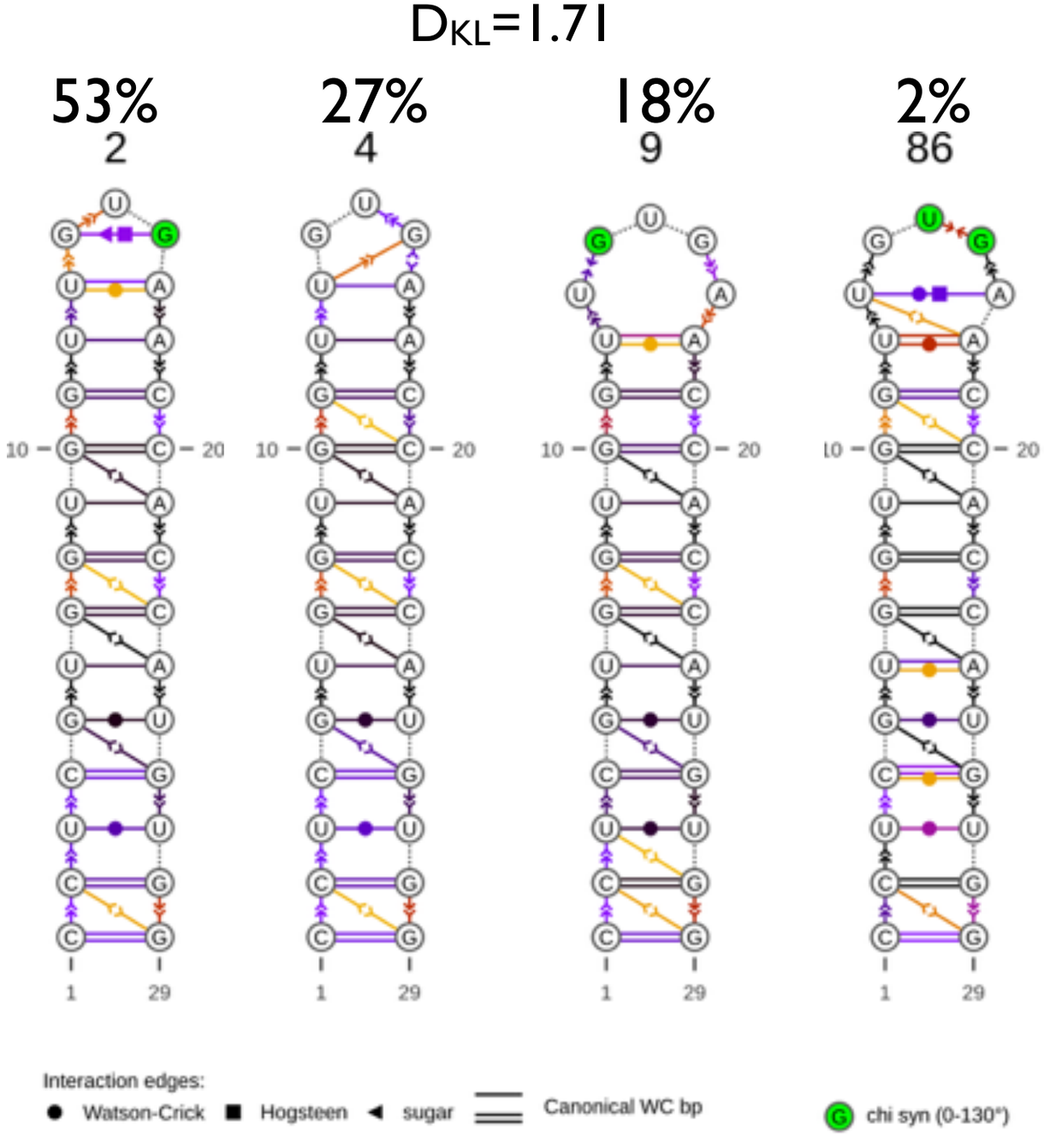
%Podbevsek, et al Sci. Rep. (2018)
Reisser et al, NAR (2020)

Maximum parsimony

Select the smallest number of conformational clusters that can explain exp data

$$\langle f_{NOE}(d_i)\rangle_{\text{set}} = \sum_{y\in\text{set}}^{Y} w'_y \langle f_{NOE}(d_i)\rangle_y$$

$$D_{KL}(w'_y||P_y) = \sum_{y\in\text{set}} w'_y \ln\frac{w'_y}{P_y}$$



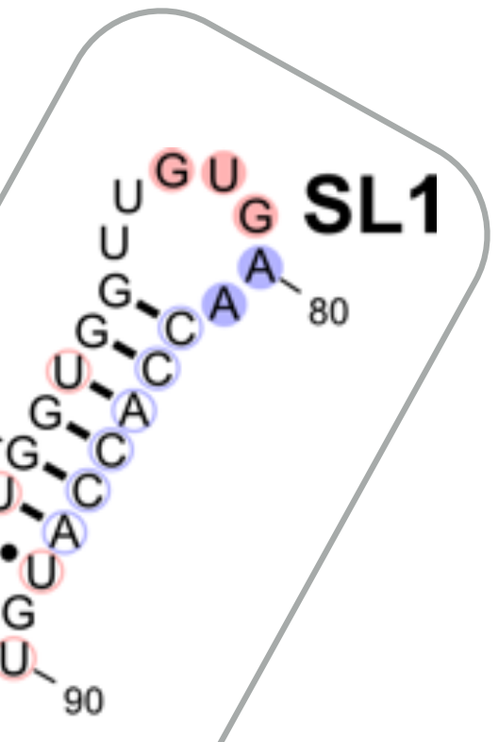
%Podbevsek, et al Sci. Rep. (2018)
Reisser et al, NAR (2020)

Maximum parsimony

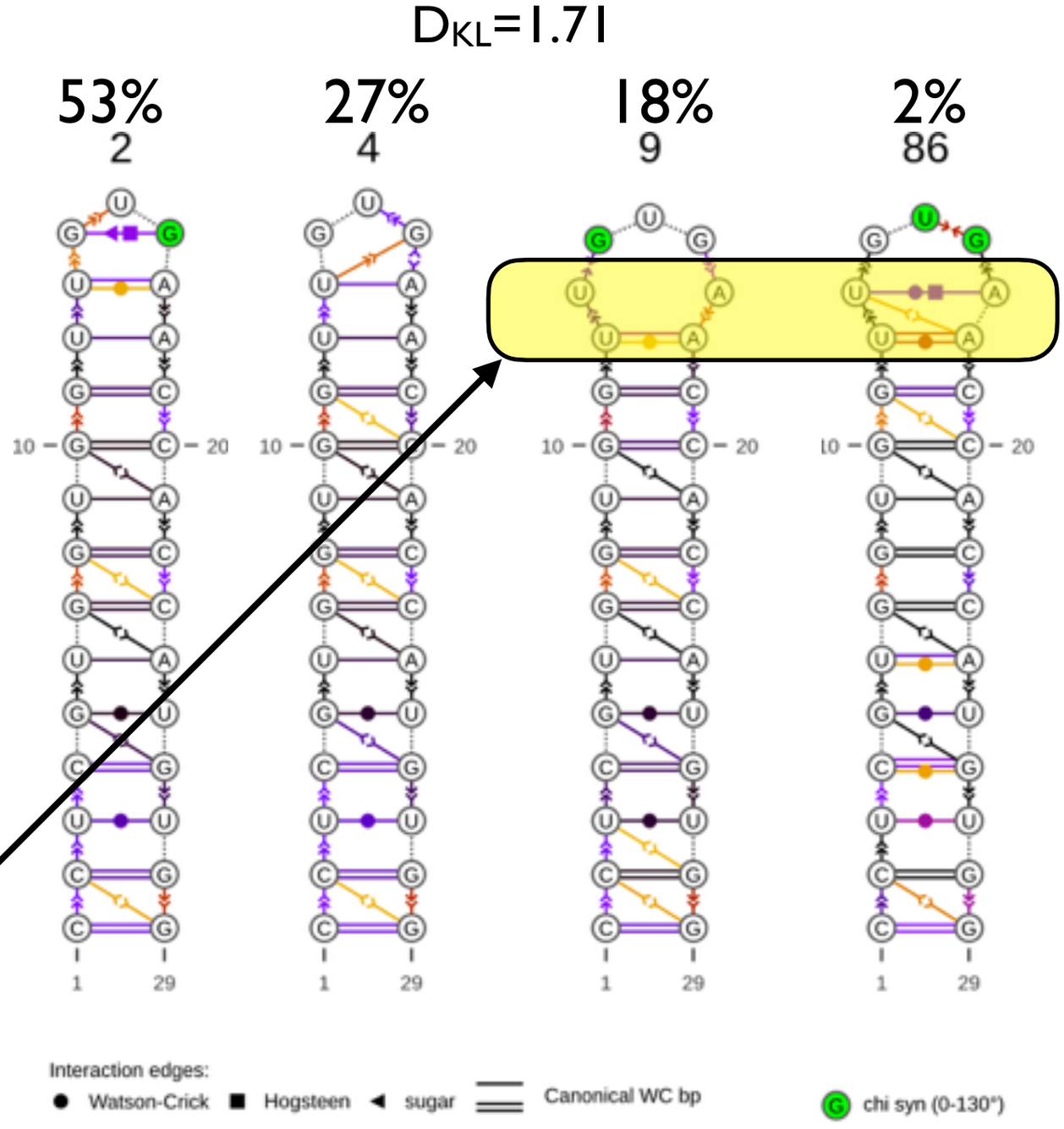
Select the smallest number of conformational clusters that can explain exp data

$$\langle f_{NOE}(d_i)\rangle_{\text{set}} = \sum_{y\in\text{set}}^{Y} w'_y \langle f_{NOE}(d_i)\rangle_y$$

$$D_{KL}(w'_y||P_y) = \sum_{y\in\text{set}} w'_y \ln\frac{w'_y}{P_y}$$



Consistent with DMS reactivity% (validation)



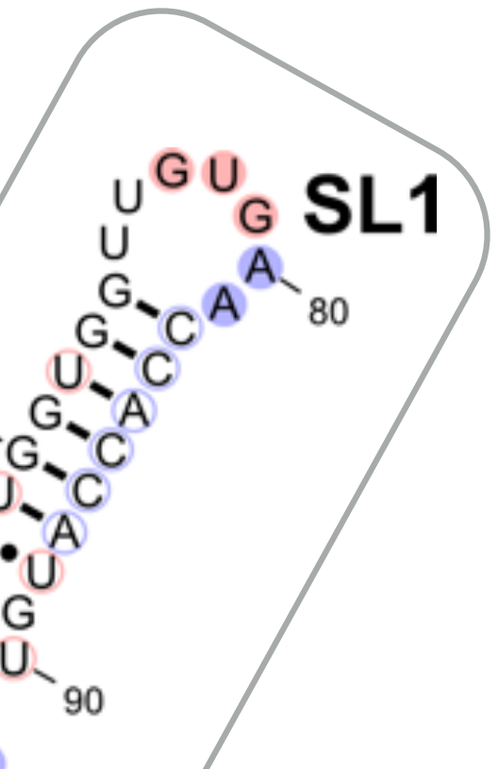
%Podbevsek, et al Sci. Rep. (2018)
Reisser et al, NAR (2020)

Maximum parsimony

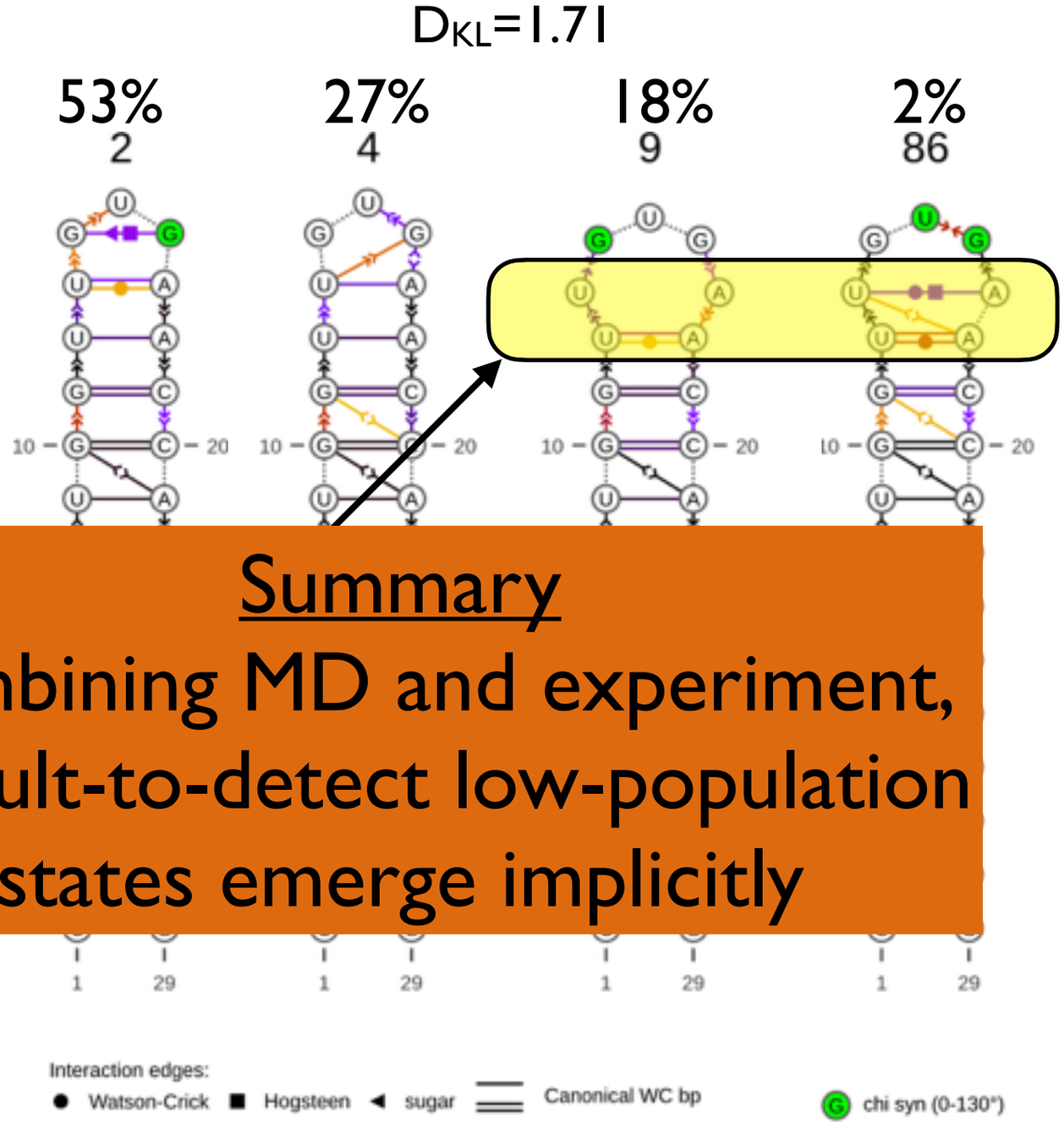
Select the smallest number of conformational clusters that can explain exp data

$$\langle f_{NOE}(d_i)\rangle_{\text{set}} = \sum_{y \in \text{set}}^{Y} w'_y \langle f_{NOE}(d_i)\rangle_y$$

$$D_{KL}(w'_y||P_y) = \sum_{y \in \text{set}} w'_y \ln \frac{w'_y}{P_y}$$



Consistent with DMS reactivity% (validation)



Summary

Combining MD and experiment, difficult-to-detect low-population states emerge implicitly

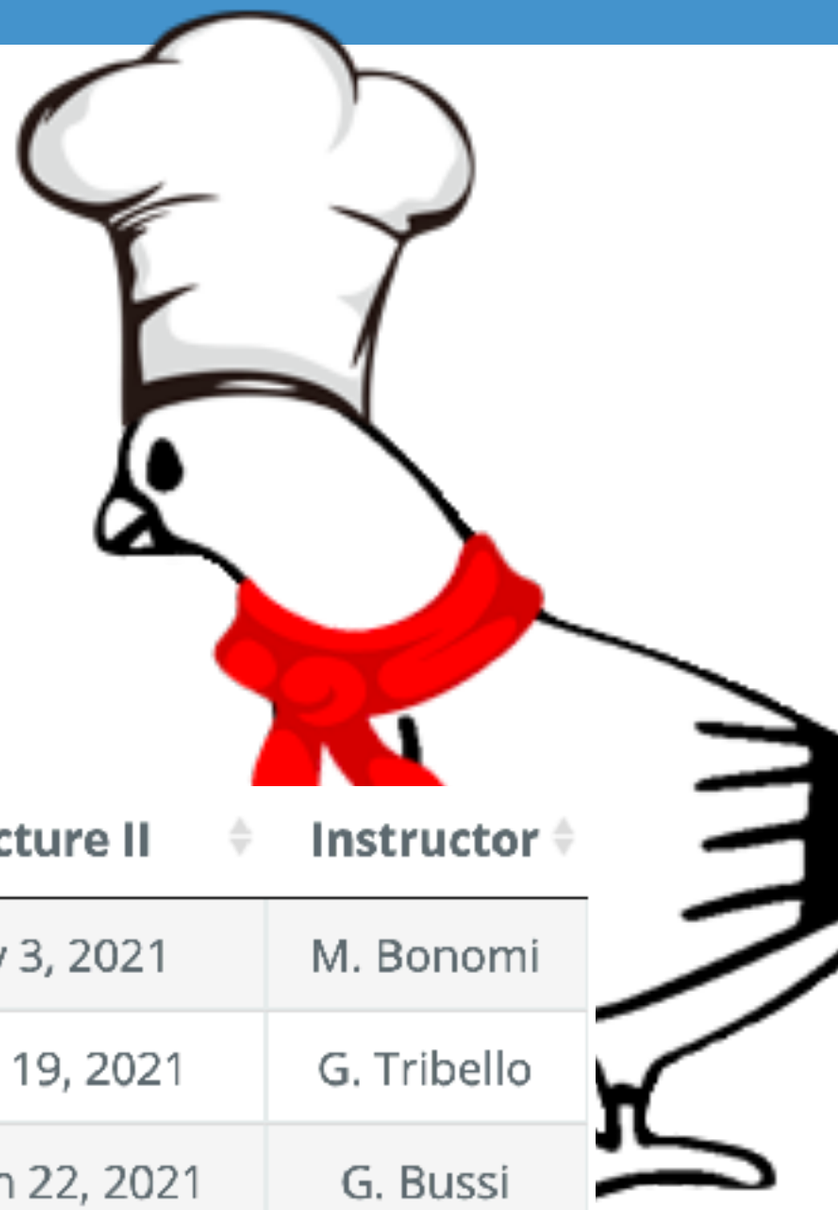
%Podbevsek, et al Sci. Rep. (2018)
Reisser et al, NAR (2020)

plumed.org/masterclass

Zoom lectures with a limited number of participants

Dedicated Slack workspace

Deadline Nov 18, 2020



Class	Topic	Lecture I	Lecture II	Instructor
21.VII	Performance optimization	April 26, 2021	May 3, 2021	M. Bonomi
21.VI	Dimensionality reduction	April 12, 2021	April 19, 2021	G. Tribello
21.V	Replica exchange methods	March 15, 2021	March 22, 2021	G. Bussi
21.IV	Metadynamics	March 1, 2021	March 8, 2021	M. Bonomi
21.III	Umbrella sampling	February 15, 2021	February 22, 2021	G. Bussi
21.II	Statistical errors in MD	February 1, 2021	February 8, 2021	G. Tribello
21.I	PLUMED syntax and analysis	January 18, 2021	January 25, 2021	M. Bonomi

www.plumed-nest.org

PLUMED-NEST

The public repository of the PLUMED consortium

Home | News | PLUMED | Consortium | Contribute | Cite | Browse

- Repository of the data needed to reproduce PLUMED-enhanced simulations
- PLUMED input files tested for compatibility with current version of the code
- Hyperlinks to PLUMED documentation to learn from real-life examples
- Promote scientific reproducibility - create educational material
- 110 eggs in the nest as of today

Bonomi, Bussi, Camilloni, Tribello, et al, Nature Methods (2019)

www.plumed-nest.org

PLUMED-NEST

The public repository of the PLUMED consortium

- Repository of the data needed to reproduce PLUMED-enhanced simulations
- PLUMED input files tested for compatibility with current version of the code
- Hyperlinks to PLUMED documentation to learn from real-life examples
- Promote scientific reproducibility - create educational material
- 110 eggs in the nest as of today

Bonomi, Bussi, Camilloni, Tribello, et al, Nature Methods (2019)

www.plumed-nest.org

Project ID: plumID:19.072
Name: SINE hairpin MD+NMR
Archive: https://github.com/bussilab/plumed-nest-sine/archive/4c288794ee470cd694231ff2976607991df11649.zip (browse)
Checksum (md5): 67f64ac04c91977cfacfd1579c4539bf
Category: bio
Keywords: metadynamics, RNA, NMR
PLUMED version: 2.4
Contributor: Giovanni Bussi
Submitted on: 26 Sep 2019
Publication: unpublished

PLUMED input files

File	Compatible with
plumed.0.dat	v2.6 passing master passing
plumed.1.dat	v2.6 passing master passing

- Repository of th ns
- PLUMED input fi de
- Hyperlinks to PL documentation to learn from real-life examples
- Promote scientific reproducibility - create educational material
- 110 eggs in the nest as of today

Bonomi, Bussi, Camilloni, Tribello, et al, Nature Methods (2019)

www.plumed-nest.org

Project ID: plumID:19.072
Source: plumed.0.dat
Originally used with PLUMED version: 2.4
Stable: raw zipped stdout - stderr
Master: raw zipped stdout - stderr

```
RESTART
INCLUDE FILE=cv.dat
INCLUDE FILE=plu_maxent_NOE_RECT.dat

############################
#### Metadynamics
############################

METAD ARG=c77 SIGMA=0.25 HEIGHT=0 PACE=500 BIASFACTOR=1.001 TEMP=300 GRID_MIN=-pi GRID_MAX=pi
METAD ARG=c78 SIGMA=0.25 HEIGHT=0 PACE=500 BIASFACTOR=1.001 TEMP=300 GRID_MIN=-pi GRID_MAX=pi
METAD ARG=c79 SIGMA=0.25 HEIGHT=0 PACE=500 BIASFACTOR=1.001 TEMP=300 GRID_MIN=-pi GRID_MAX=pi
METAD ARG=co77 SIGMA=0.05 HEIGHT=0 PACE=500 BIASFACTOR=1.001 TEMP=300 GRID_MIN=0 GRID_MAX=30 F
METAD ARG=co78 SIGMA=0.05 HEIGHT=0 PACE=500 BIASFACTOR=1.001 TEMP=300 GRID_MIN=0 GRID_MAX=30 F
METAD ARG=co79 SIGMA=0.05 HEIGHT=0 PACE=500 BIASFACTOR=1.001 TEMP=300 GRID_MIN=0 GRID_MAX=30 F
```

- Reposito
- PLUMED
- Hyperlink
- Promote scientific reproducibility - create educational material
- 110 eggs in the nest as of today

Bonomi, Bussi, Camilloni, Tribello, et al, Nature Methods (2019)

www.plumed-nest.org

Project ID: plumID:19.072

METAD

This is part of the bias module

Used to performed metadynamics on one or more collective variables.

In a metadynamics simulations a history dependent bias composed of intermittently added Gaussian functions is added to the potential [64].

$$V(\vec{s},t) = \sum_{k\tau<t} W(k\tau) \exp\left(-\sum_{i=1}^{d} \frac{(s_i - s_i^{(0)}(k\tau))^2}{2\sigma_i^2} \right).$$

This potential forces the system away from the kinetic traps in the potential energy surface and out into the unexplored parts of the energy landscape. Information on the Gaussian functions from which this potential is composed is output to a file called HILLS, which is used both

```
METAD ARG=co78 SIGMA=0.05 HEIGHT=0 PACE=500 BIASFACTOR=1.001 TEMP=300 GRID_MIN=0 GRID_MAX=30 F
METAD ARG=co79 SIGMA=0.05 HEIGHT=0 PACE=500 BIASFACTOR=1.001 TEMP=300 GRID_MIN=0 GRID_MAX=30 F
```

- PLUMED
- Hyperlinks to PLUMED documentation to learn from real life examples
- Promote scientific reproducibility - create educational material
- 110 eggs in the nest as of today

Bonomi, Bussi, Camilloni, Tribello, et al, Nature Methods (2019)

www.plumed-nest.org

PLUMED-NEST

The public repository of the PLUMED consortium

Home | News | PLUMED | Consortium | Contribute | Cite | Browse

- Repository of the data needed to reproduce PLUMED-enhanced simulations
- PLUMED input files tested for compatibility with current version of the code
- Hyperlinks to PLUMED documentation to learn from real-life examples
- Promote scientific reproducibility - create educational material
- 110 eggs in the nest as of today

Bonomi, Bussi, Camilloni, Tribello, et al, Nature Methods (2019)

www.plumed-nest.org

PLUMED-NEST

The public repository of the PLUMED consortium

- Repository of the data needed to reproduce PLUMED-enhanced simulations
- PLUMED input files tested for compatibility with current version of the code
- Hyperlinks to PLUMED documentation to learn from real-life examples
- Promote scientific reproducibility - create educational material
- 110 eggs in the nest as of today

Bonomi, Bussi, Camilloni, Tribello, et al, Nature Methods (2019)

www.plumed-nest.org

Home

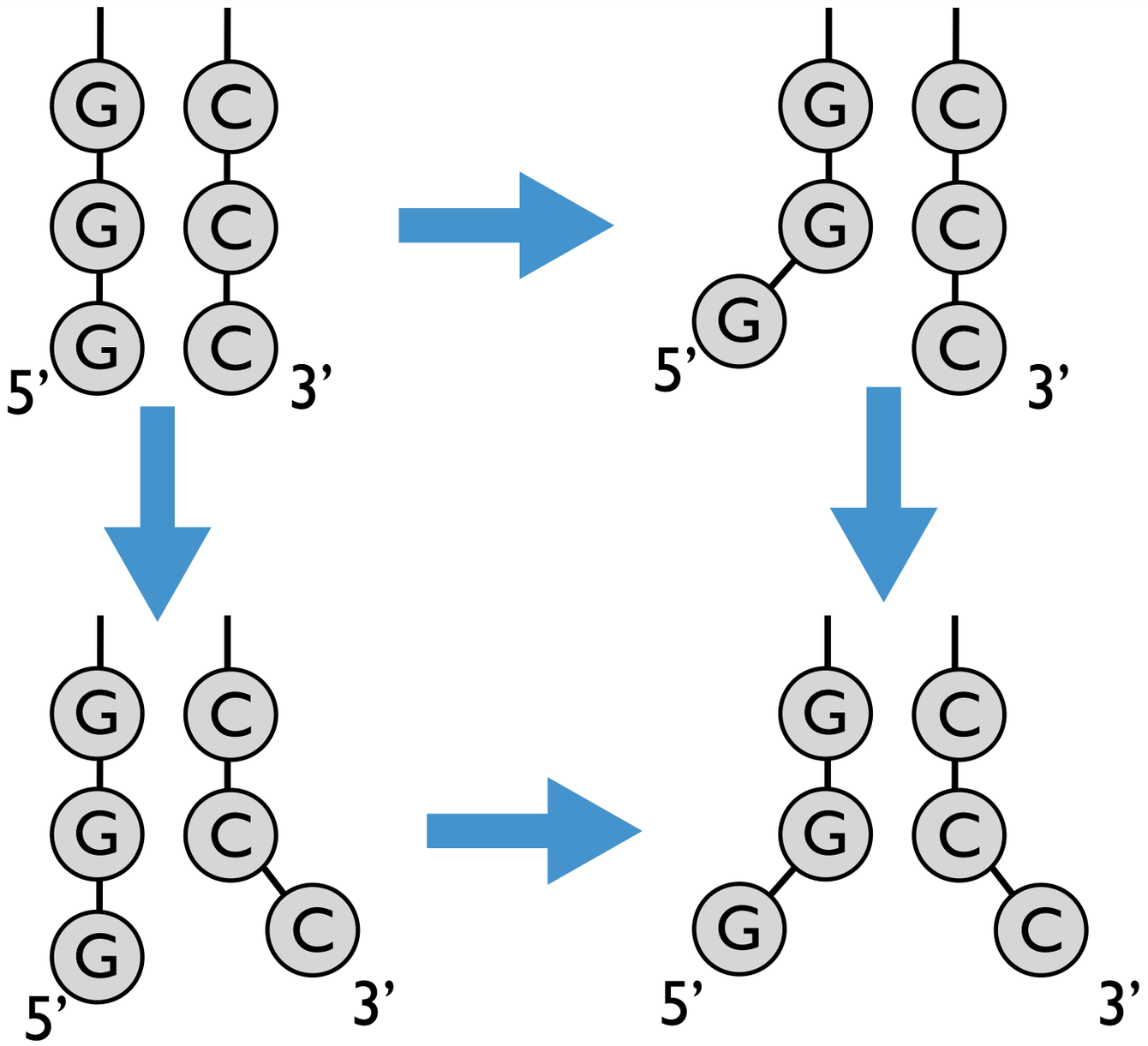
Fields marked with "*" are optional

plumID	new (plumID to be assigned)
Project name	
URL	
PLUMED input files*	examples: colvar.dat, bias.dat, ...
Category	bio
Keywords	examples: metadynamics, RNA, protein folding, small m
Instructions	Please explain how to use the deposited input files and provide a list of other software used (i.e. GROMACS) along with the specific version (i.e. 2018.6)
PLUMED version	examples: 2.4, 2.5-dev
Contributor	
Publication	examples: 10.1016/j.cpc.2013.09.018, unpublished
Contact	
Contact email	
Comments*	

Submit

Bonomi, Bussi, Camilloni, Tribello, et al, Nature Methods (2019)

RNA (and DNA) unzipping

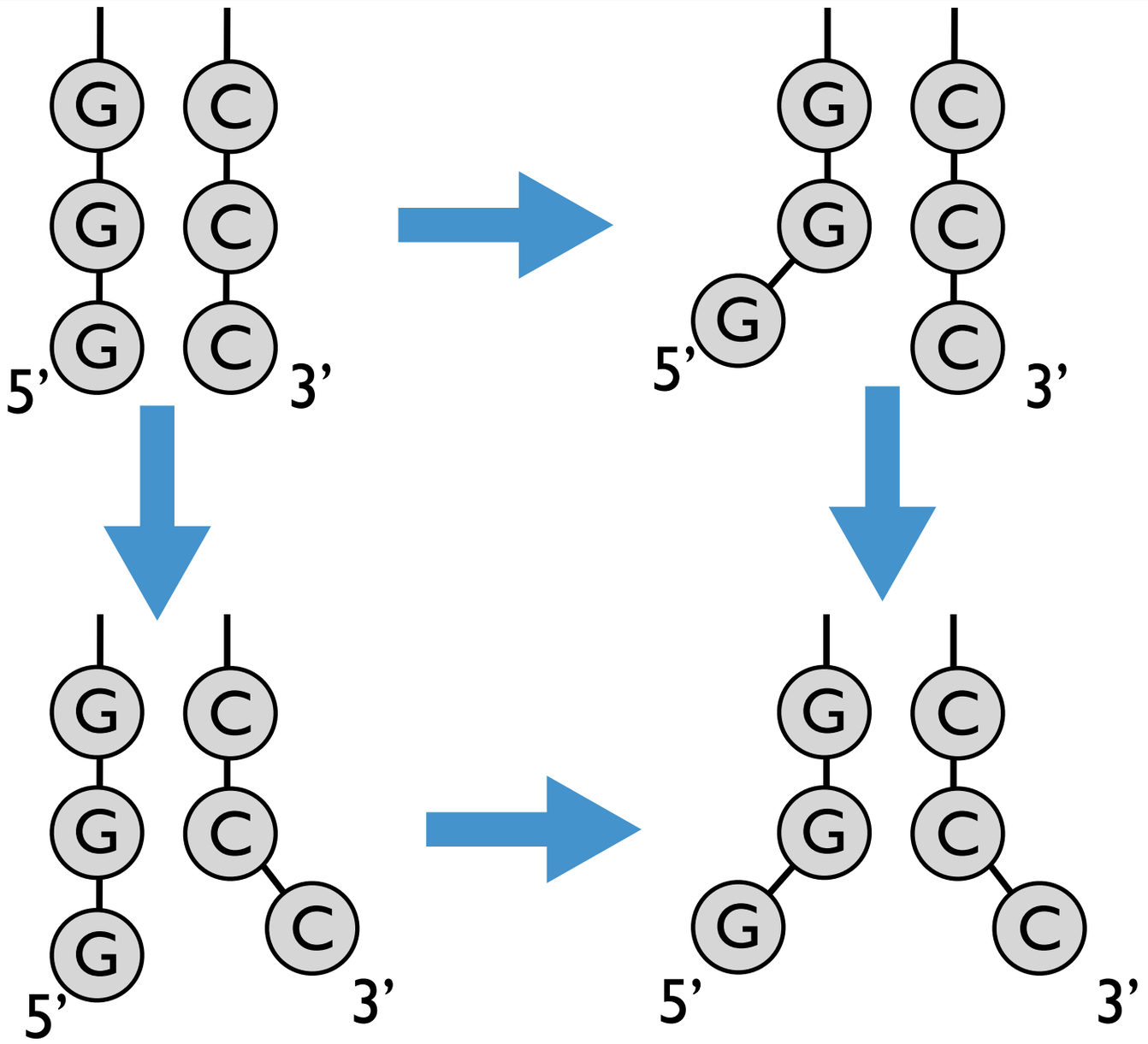


Preferential path?

Terminal asymmetry? (3’-5’)

Sequence dependence? (GC-CG)

RNA (and DNA) unzipping



Preferential path?

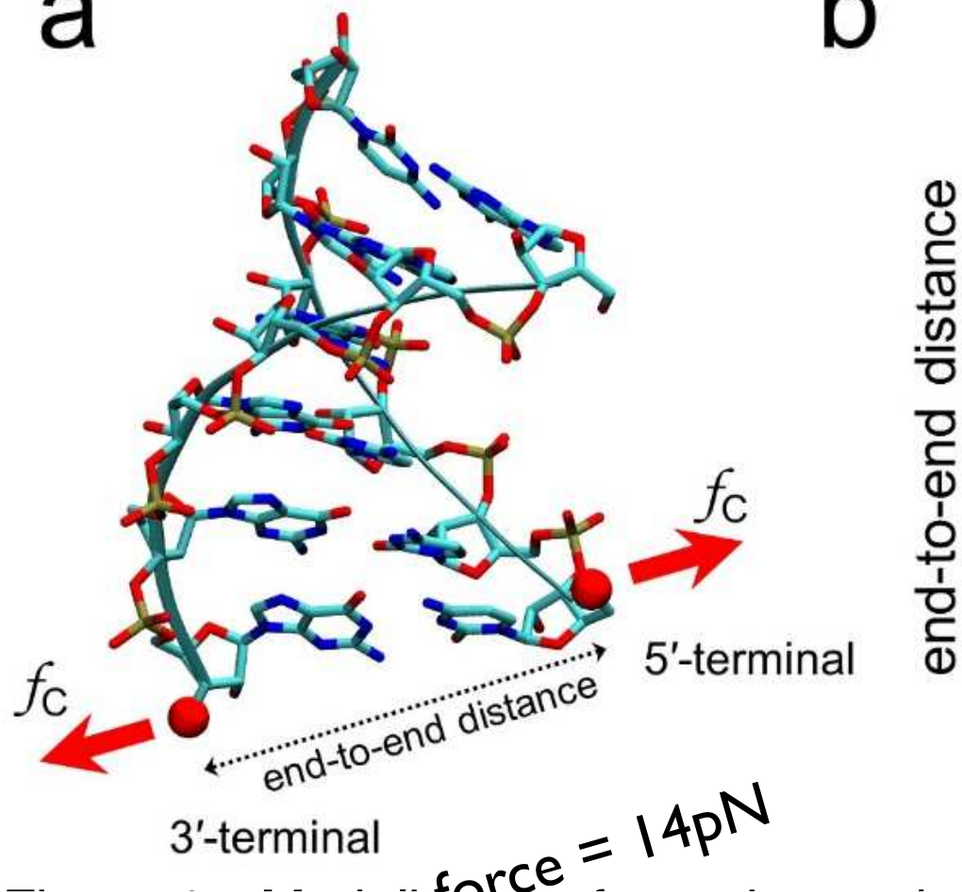
Terminal asymmetry? (3’-5’)

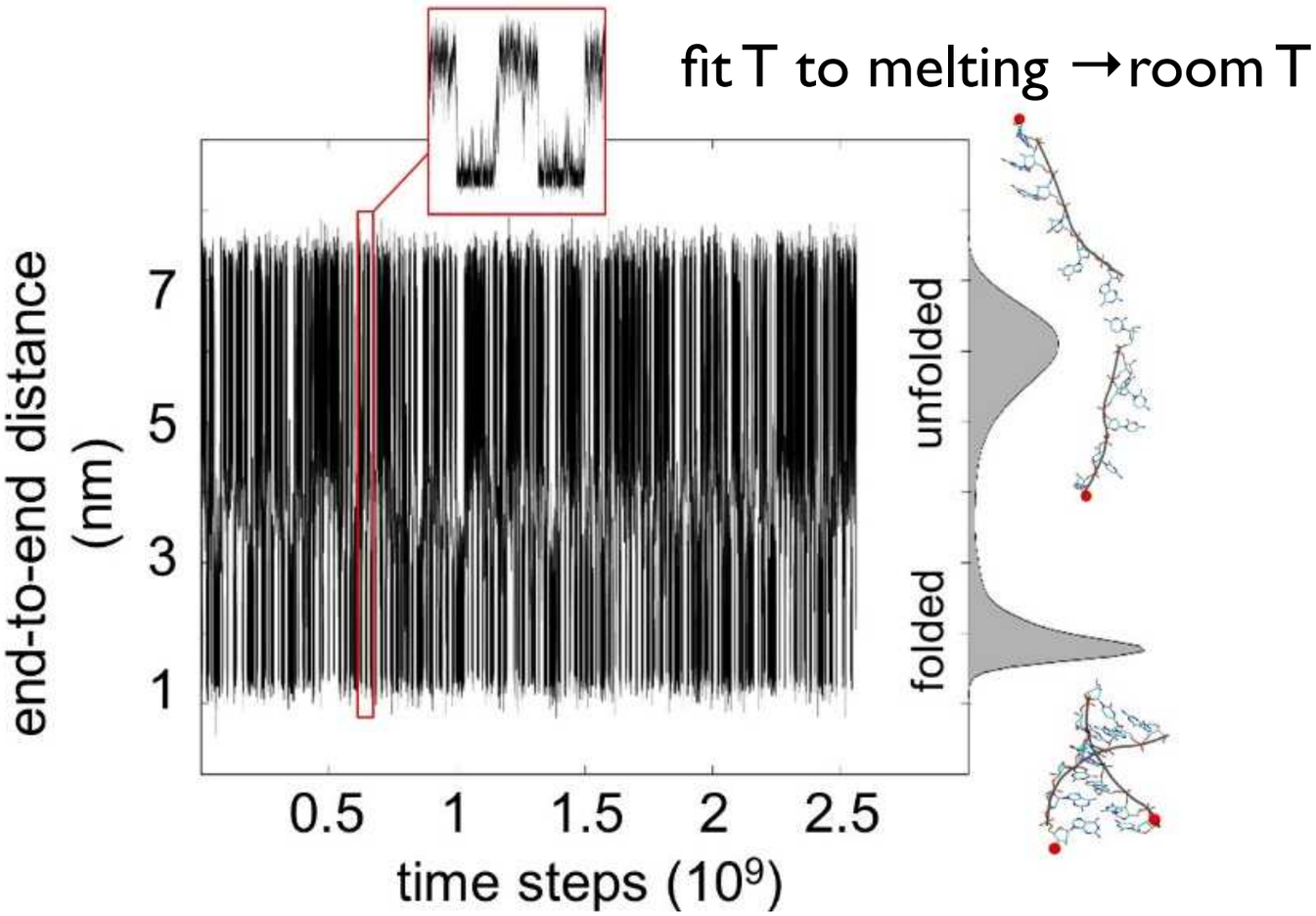
Sequence dependence? (GC-CG)

Unzipping is ubiquitously required in RNA metabolism (transcription, translation, etc)

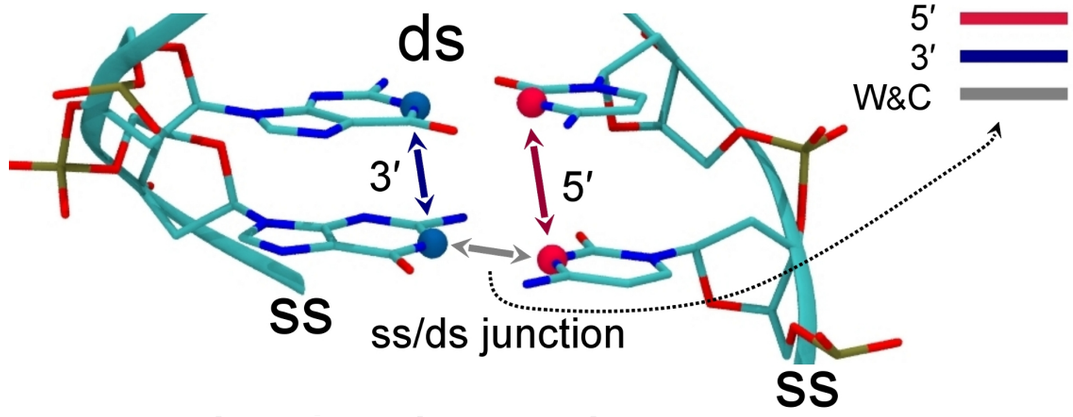
In vivo: helicases*

Structure-based model





water not there!



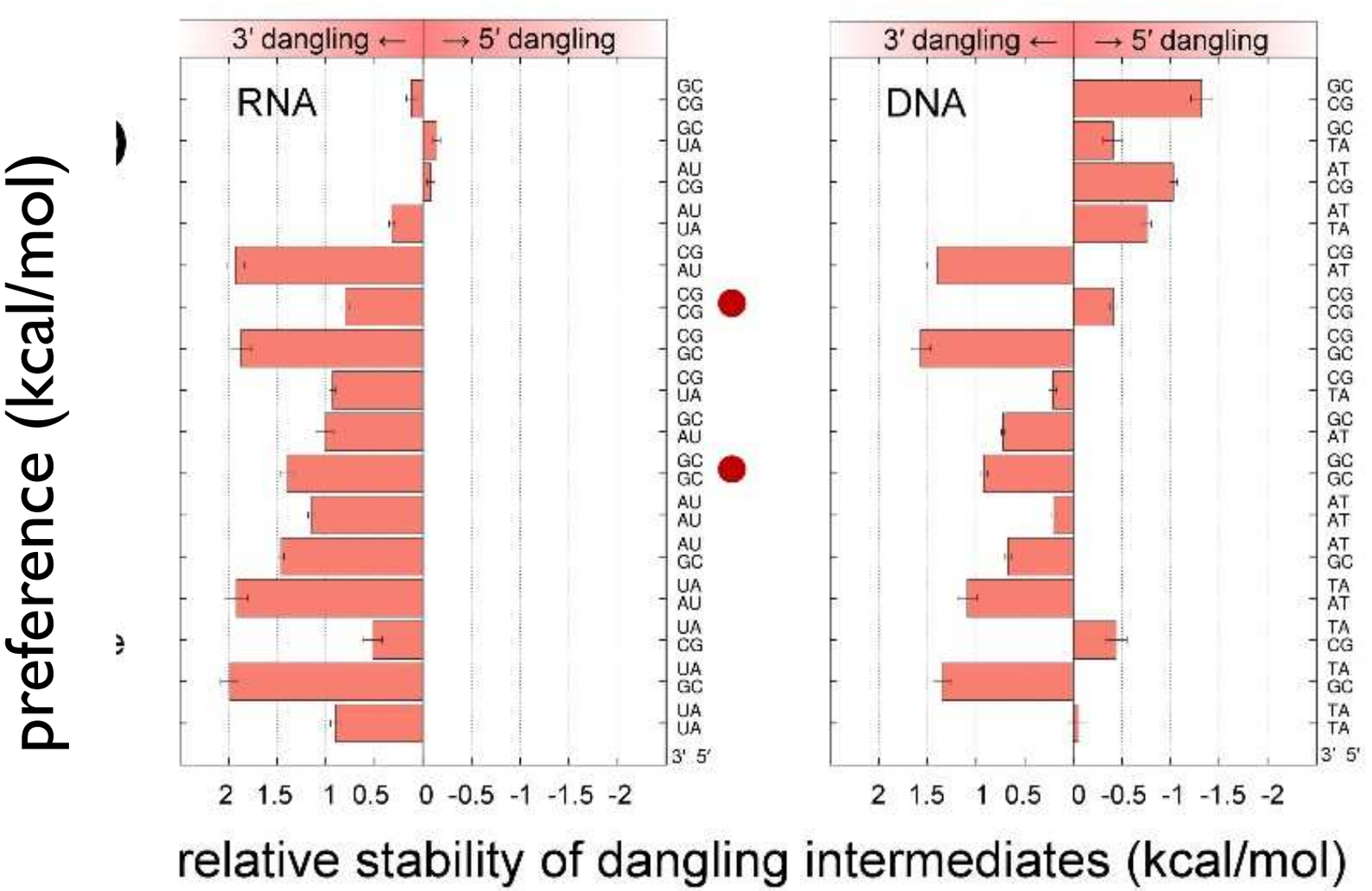
dangling bases detection

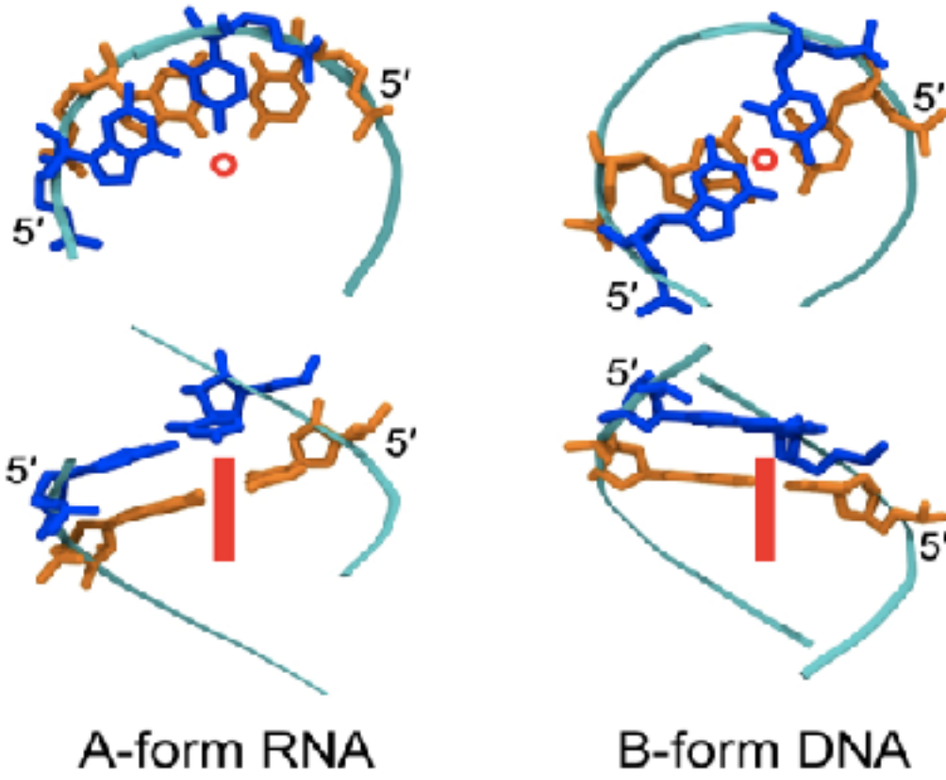
SMOG (atomistic Go model)#
6 sequences x 6 bp (all possible steps)
DNA (B-helix) vs RNA (A-helix)

#http://smog-server.org/
Colizzi et al, PNAS (2019)

MD results on a limited testset: Colizzi and Bussi, JACS (2012)

RNA (A-helix) vs DNA (B-helix)

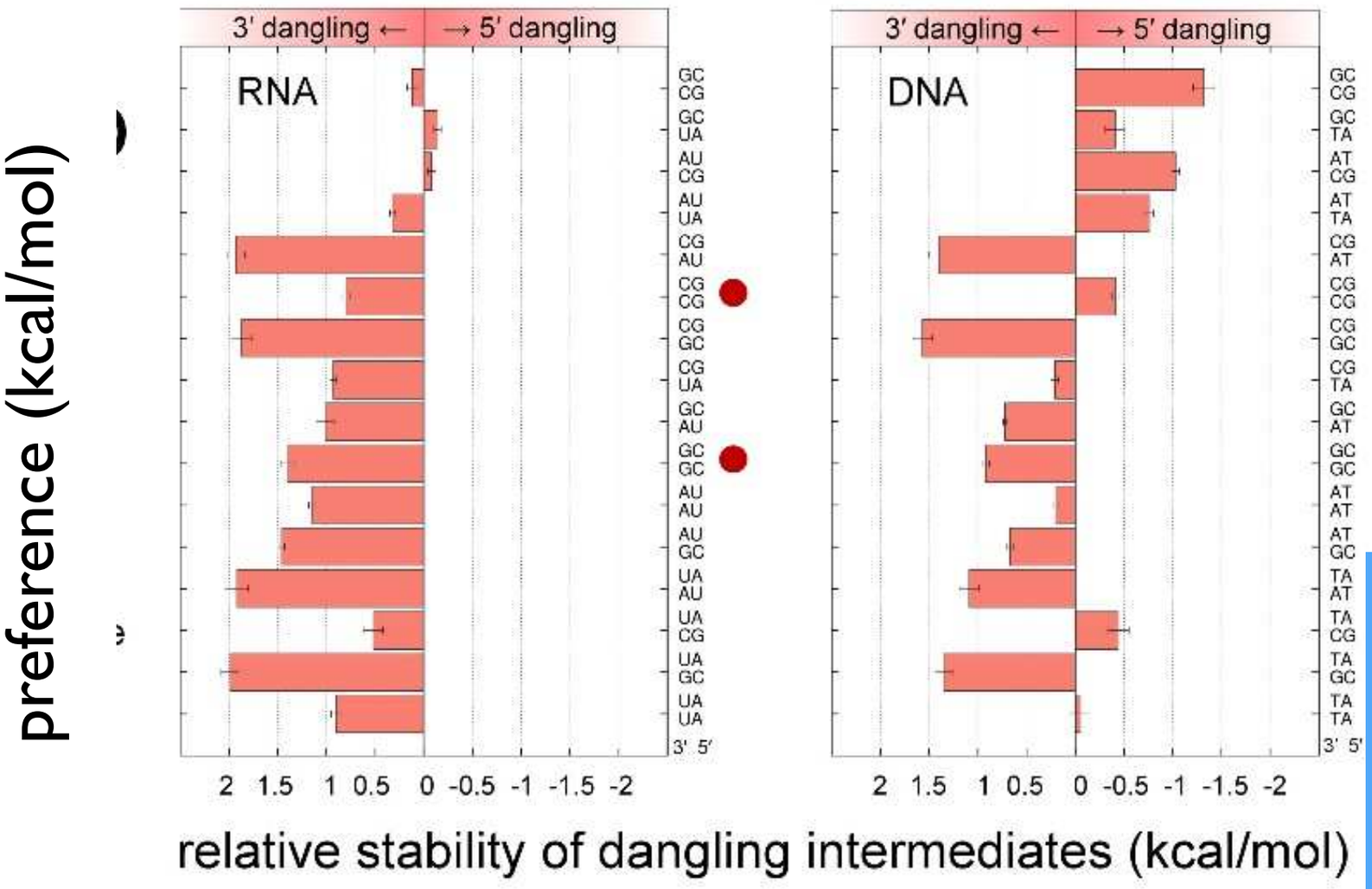


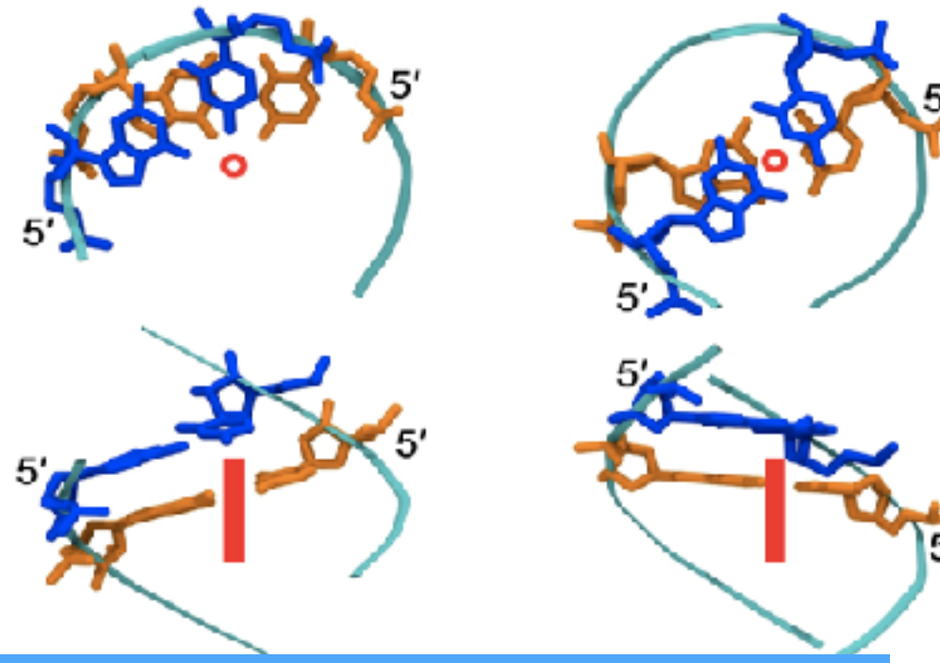


RNA: preference for 3’ stacking (consistent with explicit solvent MD)
DNA: sequence dependent

Colizzi et al, PNAS (2019)
Results are consistent with atomistic MD, Colizzi and Bussi, JACS (2012)

RNA (A-helix) vs DNA (B-helix)



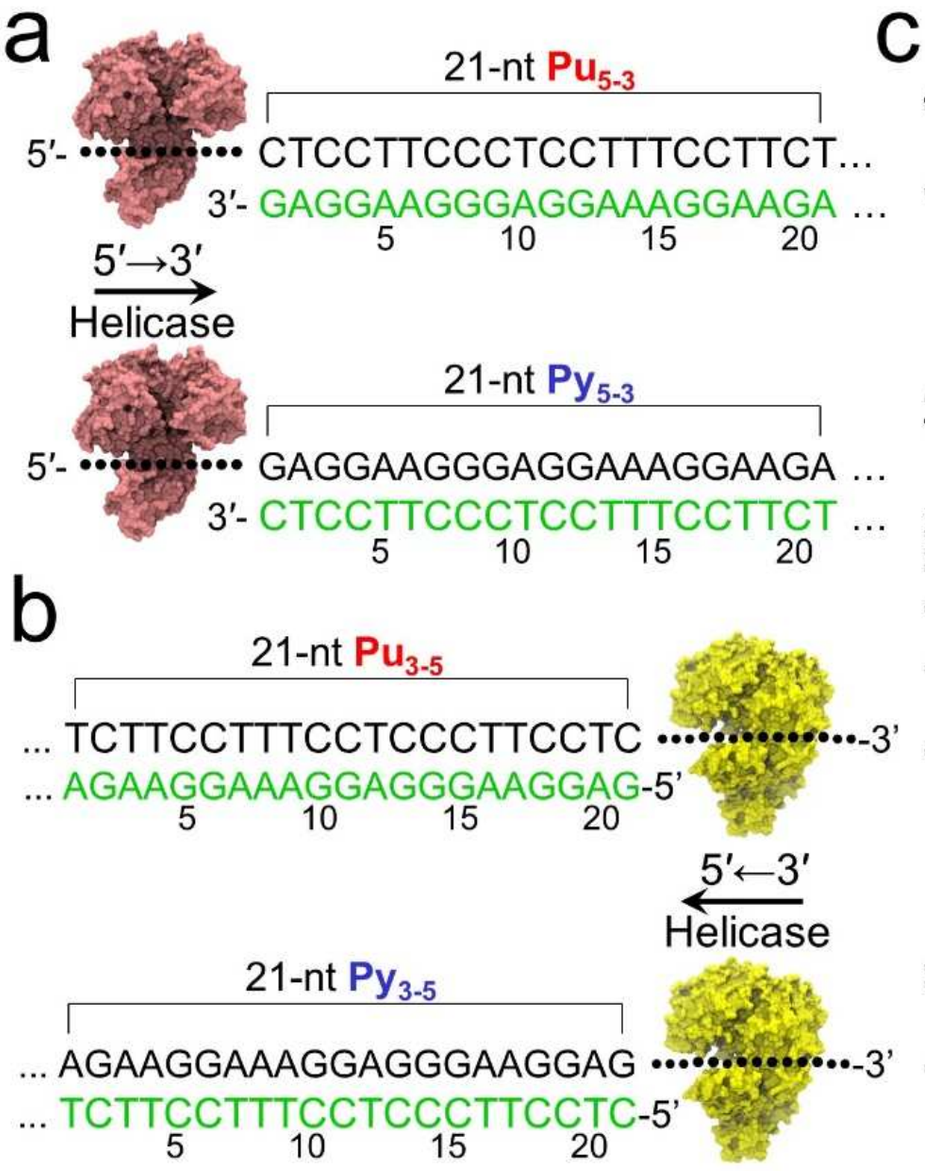


How does this difference affect helicase-mediated unzipping?

RNA: preference for 3’ stacking (consistent with explicit solvent MD)
DNA: sequence dependent

Colizzi et al, PNAS (2019)
Results are consistent with atomistic MD, Colizzi and Bussi, JACS (2012)

Helix "unwindability"

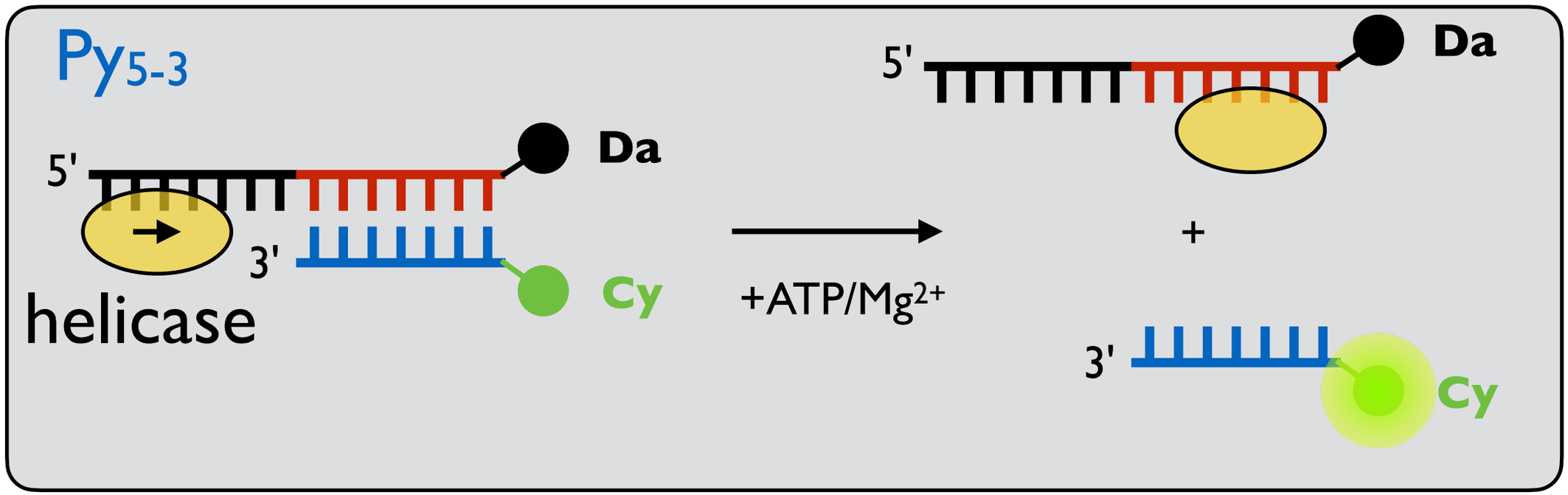


$$\frac{1}{\left\langle e^{-\frac{\Delta F}{k_B T}} \right\rangle}$$

- Proportional to "inverse of average waiting time"
- Depends on the unwinding direction

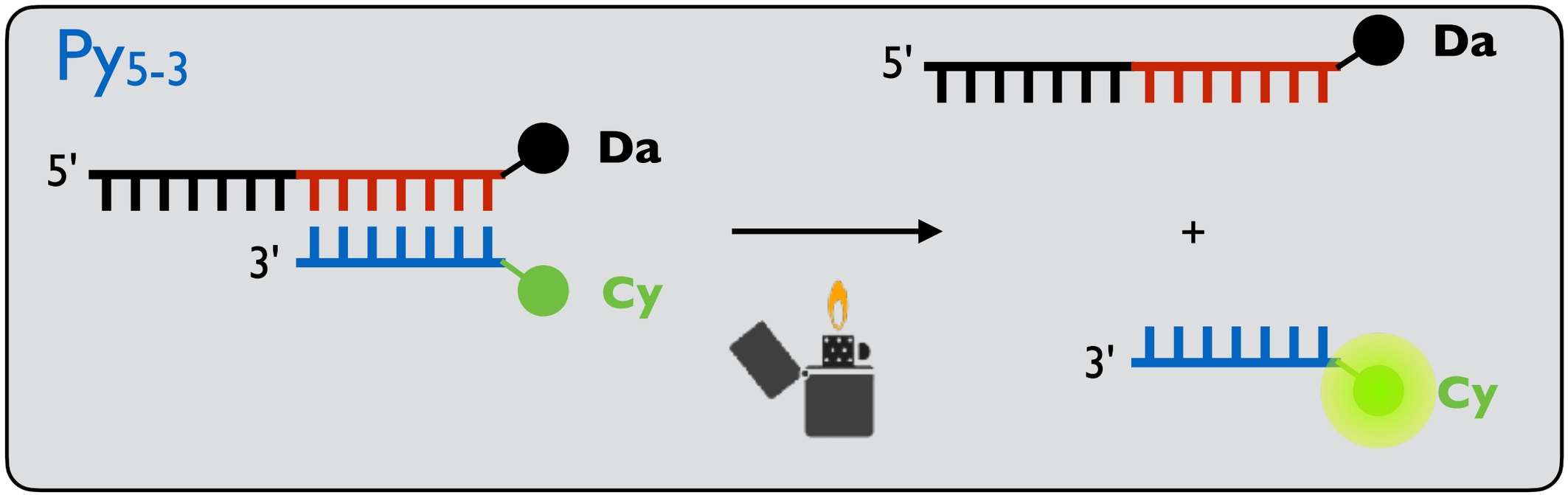
Colizzi et al, PNAS (2019)

Experimental validation on DNA

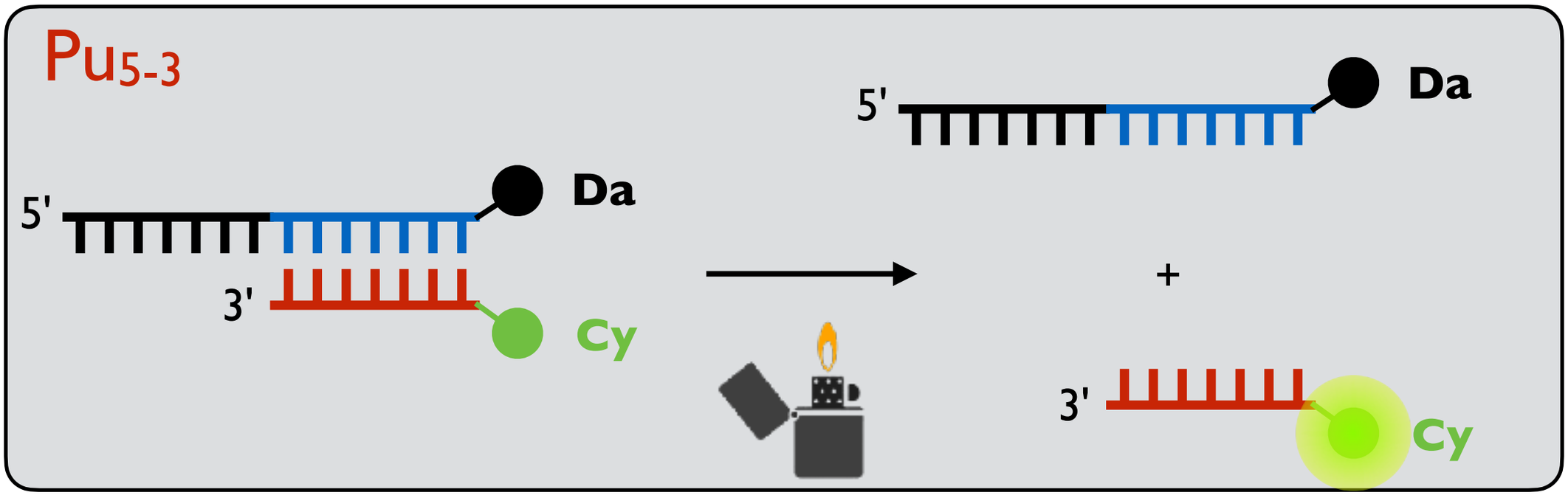


VS

First check thermal stability

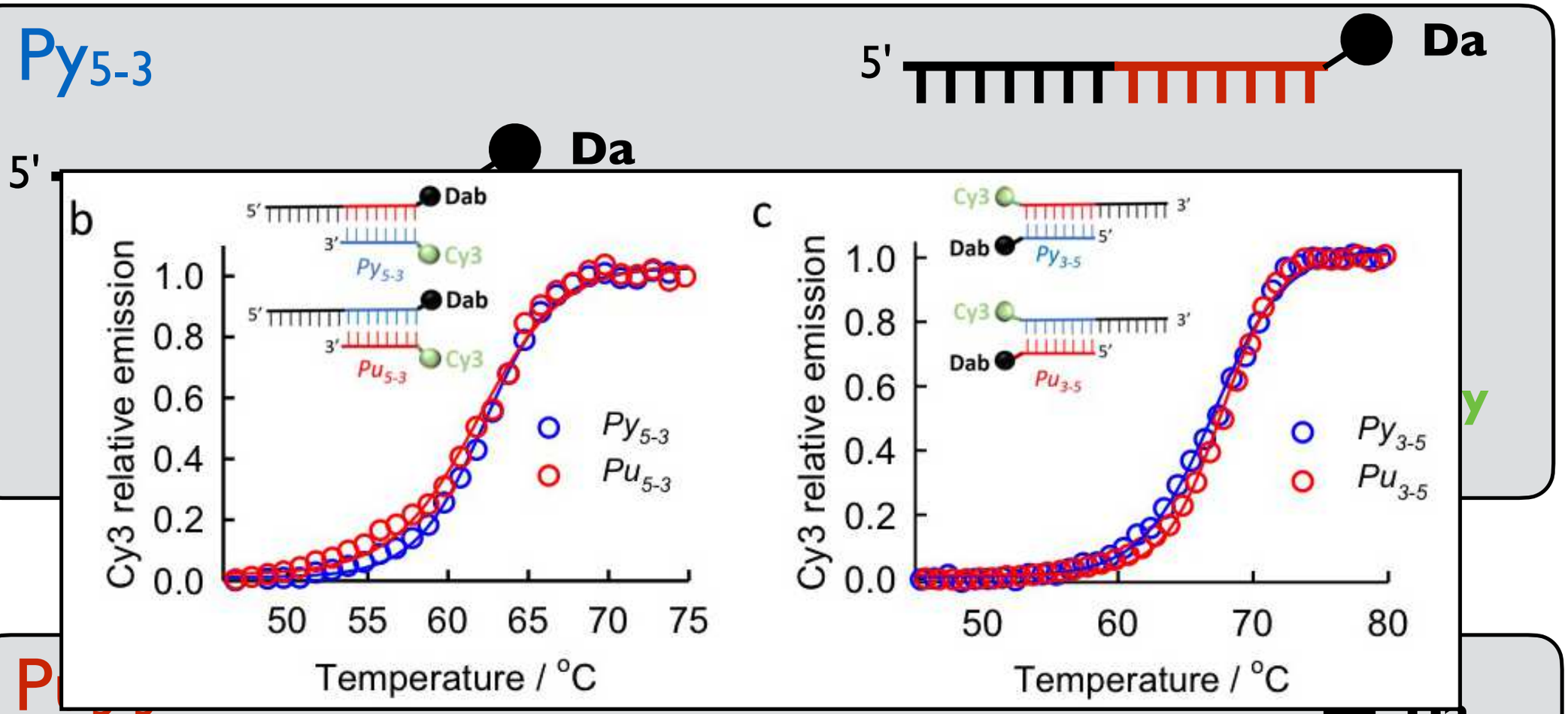


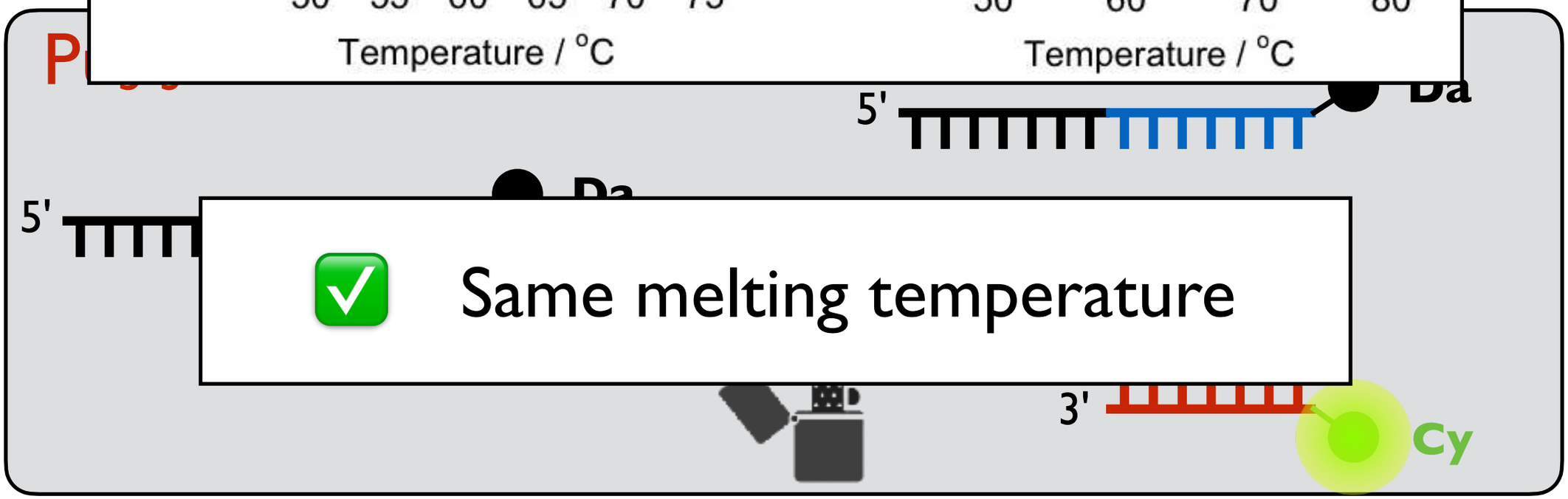
VS



First check thermal stability

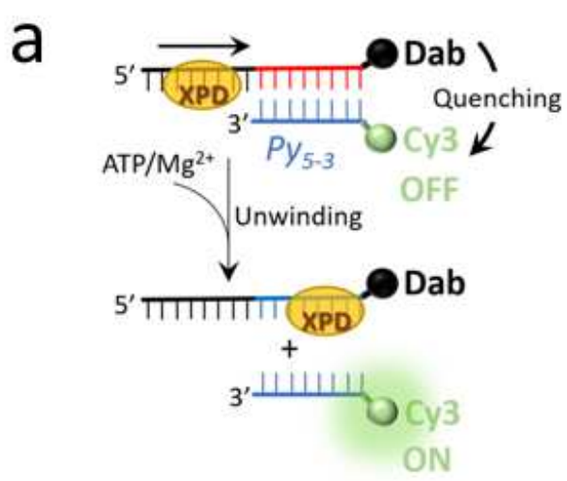
Py_{5-3}

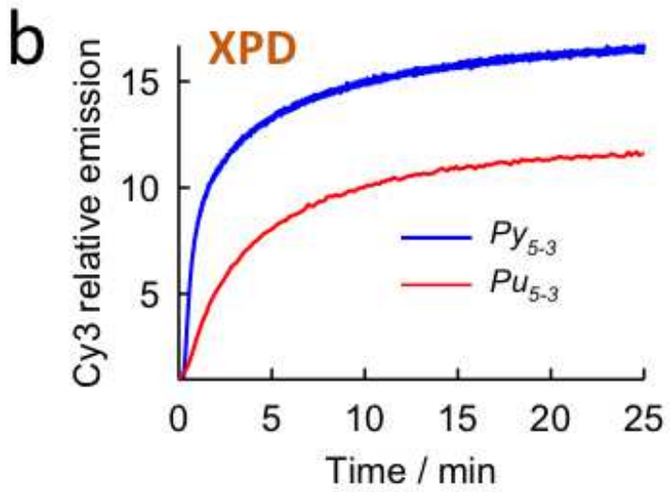


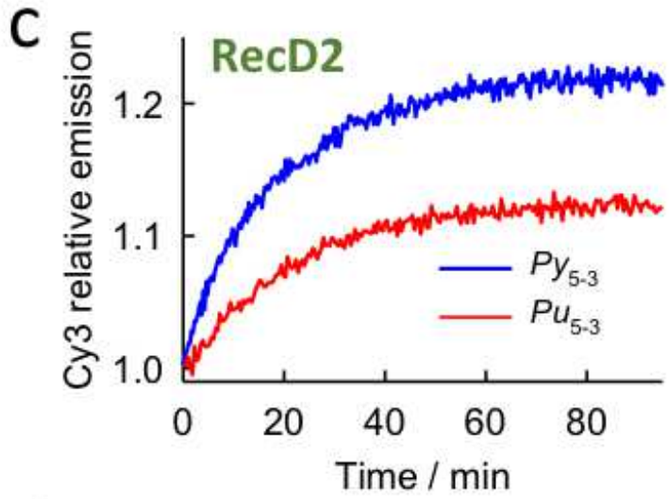


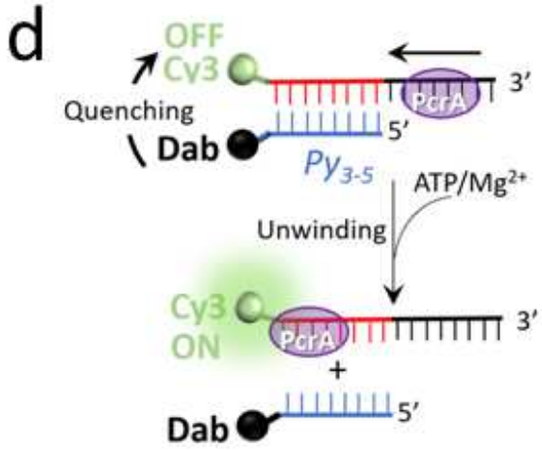
✅ Same melting temperature

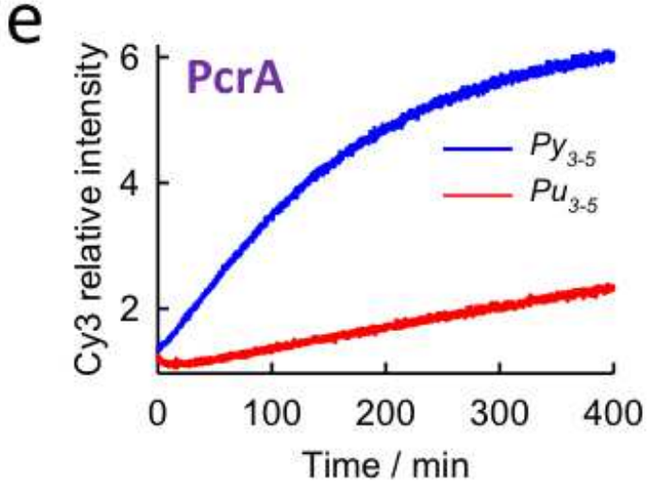
Experiments: efficiency

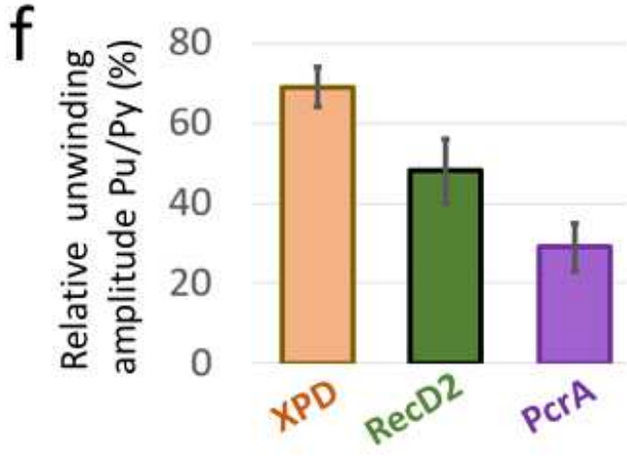












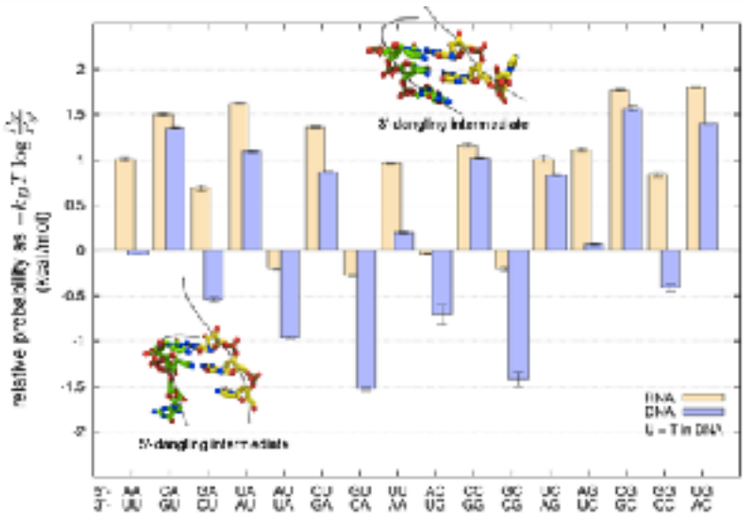
Processivity is higher when displaced strand is pyrimidine rich

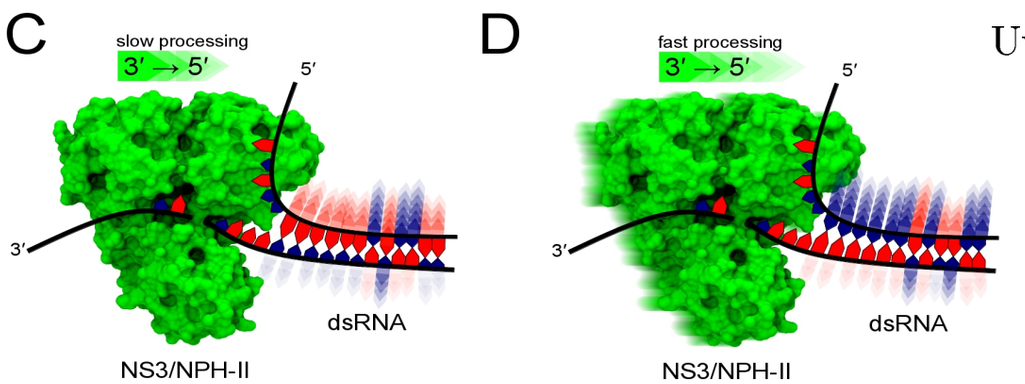
Valid on the 3 tested helicases!
1/8 prob to get this by chance :-)

Conditions (n=3 repetitions):
50nM dsDNA
500nM helicase
1mM MgCl2
0.1mg/mL BSA
Addition of 1mM ATP

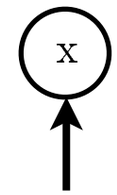
Colizzi et al, PNAS (2019)

Speculation: helicase directionality





Family	processing DNA 5′→3′	processing DNA 3′→5′	processing RNA 5′→3′	processing RNA 3′→5′	Class
DEAH/RHA		x		x	SF2
NS3/NPH-II		x		x	SF2
XPD/RAD3	x				SF2
Ski2-like		x		x	SF2
Suv3[(a)]				x	SF2
RIG1-like		x		x	SF2
RecQ-like		x			SF2
RecG-like		x			SF2
UvrD/Rep/PcrA		x			SF1
Pif1-like	x				SF1
Upf1-like	x				SF1



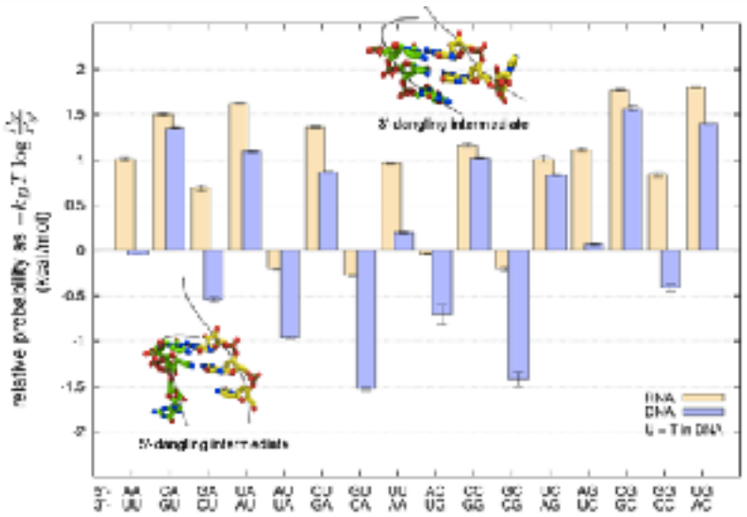
exception
(requires cofactors)

RNA helicases always process 3’⇛5’

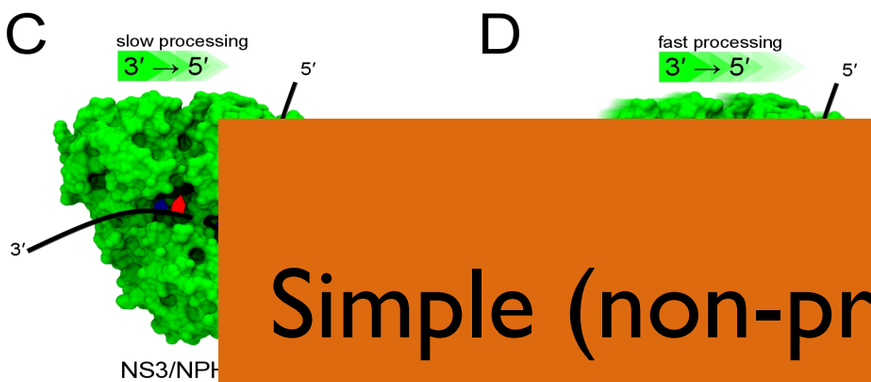
DNA helicases exist for both directions

Colizzi et al, PNAS (2019)

Speculation: helicase directionality



Family	processing DNA 5′→3′	processing DNA 3′→5′	processing RNA 5′→3′	processing RNA 3′→5′	Class
DEAH/RHA		x		x	SF2
NS3/NPH-II		x		x	SF2
XPD/RAD3	x				SF2
Ski2-like		x		x	SF2
Suv3[(a)]				x	SF2
RIG1-like		x		x	SF2
RecQ-like		x			SF2
RecG-like		x			SF2
UvrD/Rep/PcrA		x			SF1
Pif1-like	x				SF1



Summary

Simple (non-predictive?) simulations + intuition can lead to experimentally testable hypotheses

Intrinsic helix dynamics has an impact on the dynamics of nucleic acids processing machineries

RNA helicases always process 3’⇛5’

DNA helicases exist for both directions

Colizzi et al, PNAS (2019)

Agenda

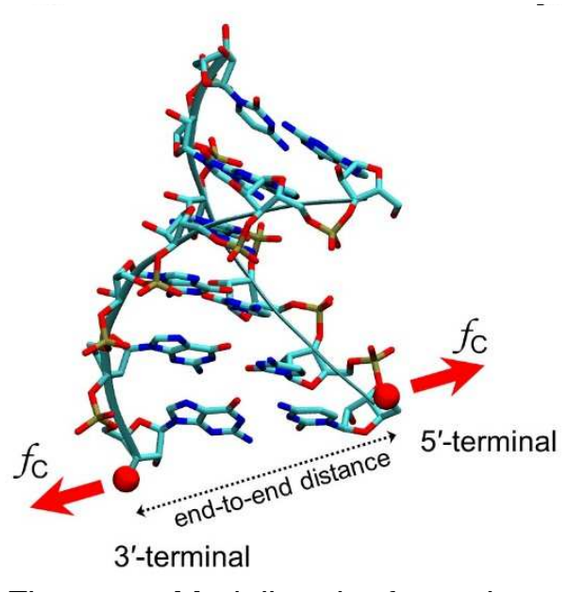
<u>Combine experiment (NMR) and MD</u>

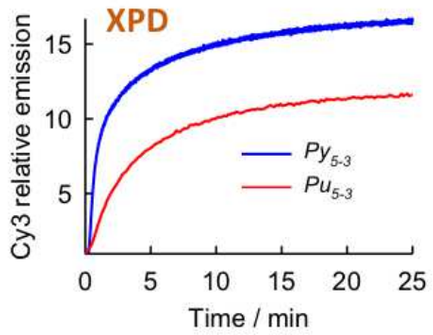
- enforce averages*
- dynamics of a RNA hairpin%

$$S[P] = -\sum_x P(x) \log \frac{P(x)}{Q(x)}$$

<u>RNA unzipping dynamics&</u>

- Go model
- Helicase mediated unwinding
- Experimental validation





*Cesari, Reisser, and Bussi, Computation (2018)
%Podbevsek et al, Sci. Rep. (2018); Reisser et al, NAR (2020)
&Colizzi et al, PNAS (2019)

Acknowledgements

Mattia Bernetti
Nicola Calonaci
Thorben Frohelking
Valerio Piomponi

Sabine Reißer
(now at MDC, Berlin)
Francesco Colizzi
(now at IRB, Barcelona)

Carlos Penedo
C. Perez-Gonzalez
Remi Fritzen
Malcom F White
(St Andrews, UK)

Koby Levy
(Weizmann, Israel)

Silvia Zucchelli (SISSA+Uni Novara)†
Stefano Gustincich (IIT Genova)