Impact of Post-Transcriptional modifications on RNA structure and dynamics

Structure and Topology of RNA in Living Systems
Trento, 2023

Student: Valerio Piomponi

Supervisor: Giovanni Bussi

Molecular dynamics

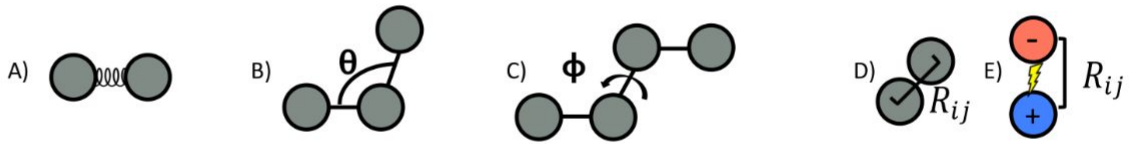
Powerful Computational tool to characterized molecular structural dynamics with "**unlimited**" resolution.

Empirical Force-fields:

Bonded

Non-bonded

$$E_{tot} = \sum_{bonds} K_r\,(r - r_{eq})^{\,2} + \sum_{angles} K_\theta\,(\theta - r\theta_{eq})^{\,2} + \sum_{dihedrals} \frac{V_n}{2}\,[1 + \cos(n\phi - \gamma)] + \sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}}\right]$$



Parameters derived from:

- Quantum Calculations (**only** small fragments)
- Experiments (**mostly** small fragments)

- No Polarization
- No chemical reactivity
- Quality of the force-fields is the bottleneck - Always integrate with experiments!

Combining MD with experiment

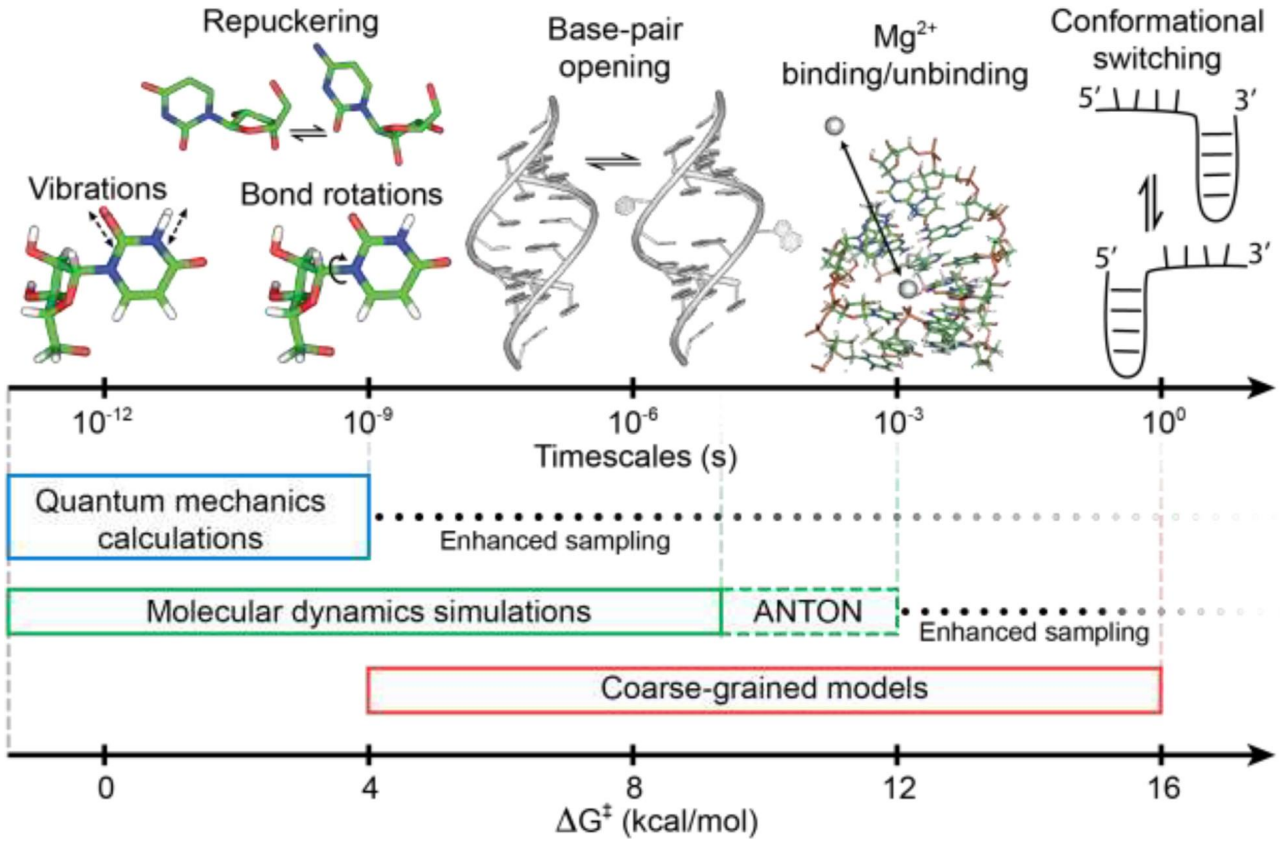
Transferable Force-Field Fitting

Ensemble refinement
(i.e Max Entropy, Minimum Parsimony)

RNA time-scales

Enhanced sampling techniques:

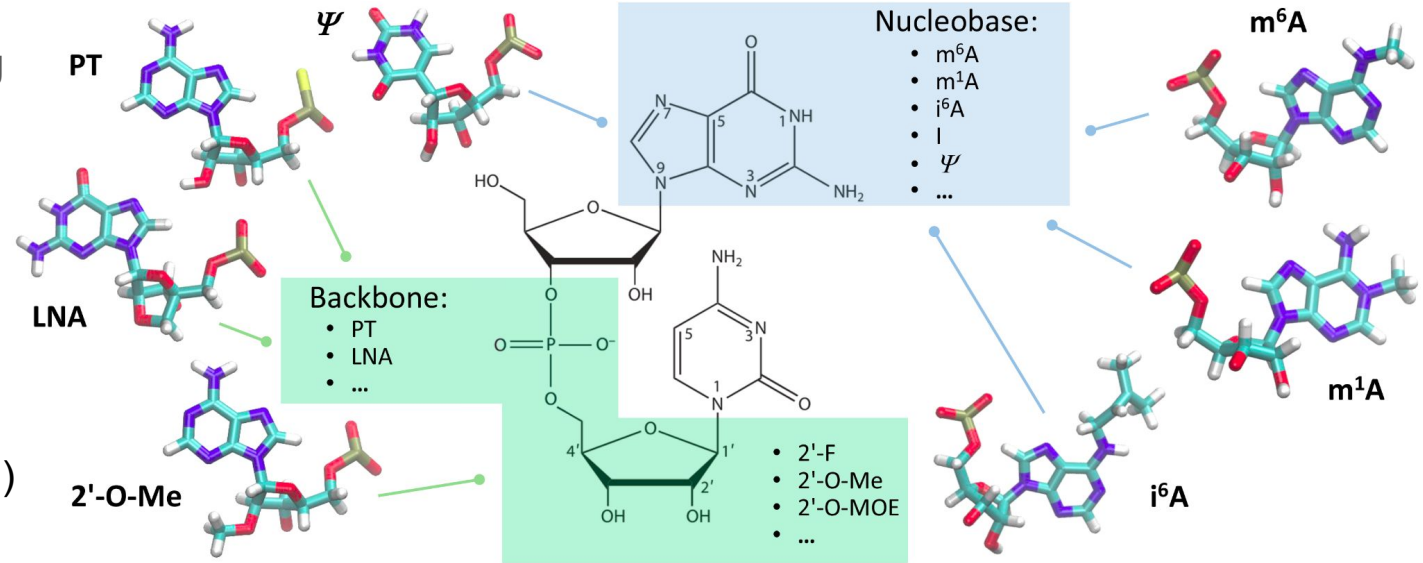
- Heat the system (e.g parallel tempering)
- **Bias** along a **reaction coordinate** (e.g Umbrella samplings; Metadynamics) More efficent, but *a priori* knowledge is needed



Sponer et al, **Chem. Rev.** (2018)

RNA modifications

- Artificial (mainly acting in the backbone to improve rigidity)
- Natural occurring (more than 100 modifications known)



Widespread (mRNA, tRNA, rRNA etc.) main roles:

- Affects RNA folding
- Affects target specificity of RNA interactions

Piomponi et al, **arXiv**, 2022

Accepted for publication in Springer book 2023 ***RNA Structure and Function***

MD to investigate impact of modification on RNA structural dynamics

Problem:

Force fields for modified nucleotides - not sufficient validations against experiments yet

Overview

- Published Work:

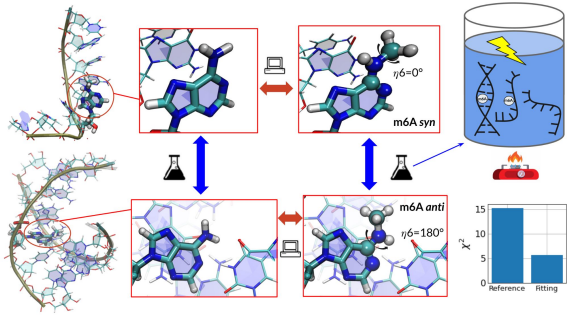
ACS central science

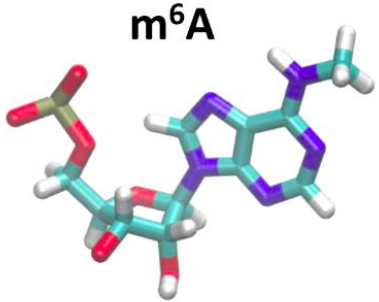
http://pubs.acs.org/journal/acscii

Research Article

Molecular Simulations Matching Denaturation Experiments for N^6-Methyladenosine

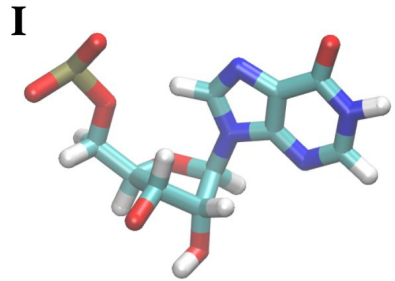
Valerio Piomponi, Thorben Fröhlking, Mattia Bernetti, and Giovanni Bussi*





Transferable Force-Field Fitting

- Ongoing: reconstruct ensemble of structures for a RNA helix containing **Inosines**, Combining MD and NMR experiments (**Maximum Entropy)**



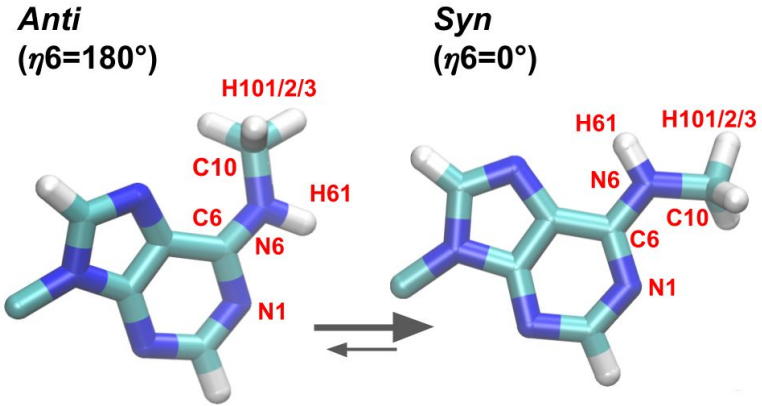
Ensemble Refinement

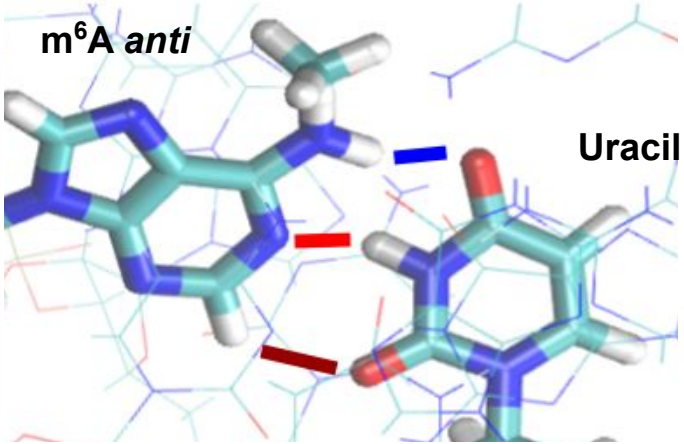
N6-methyladenosine (m^6A)

- Most common internal modification in eukaryotic RNAs (on average 1-2% of transcriptome)

- Two possible conformations: *syn* most favored in unpaired m^6A (10:1)
  anti most favored in WC paired m^6A (1:100) **[*]**

- Only available **force field** compatible with AMBER for m^6A (Aduri **[**]**) is not able to reproduce *syn/anti* populations, and other experimental evidences





*Liu et al, **Nat. Commun.** (2021)
Aduri et al, **JCTC (2007)

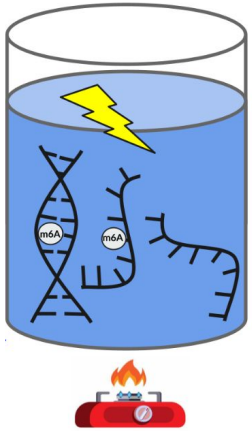
Fitting N6-methyladenosine (m^6A) force field against experiments

	System	Exp ΔΔG (kJ/mol)
A1	m6A $\Delta G_{syn/anti}$	6.3
A2	UACG6CUG AUGCUGAC	1.7 ± 0.9
A3	CGAU6GGU GCUAUCCA	7.1 ± 0.9
A4	6CGC GCG	-2.5 ± 1.2
A5	GCG6 CGC	-1.7 ± 0.9
B1	GUC6CUG CAGUGAC	2.5 ± 2.1
B2	ACU6UAGU UGAU6UCA	2.1 ± 1.3
B3	AGUU6ACU UCA6UUGA	5.4 ± 1.3
B4	CGGUG6UCG GCU6GUGGC	8.6 ± 0.8
B5	ACUUA6GU UG6AUUCA	1.7 ± 1.0

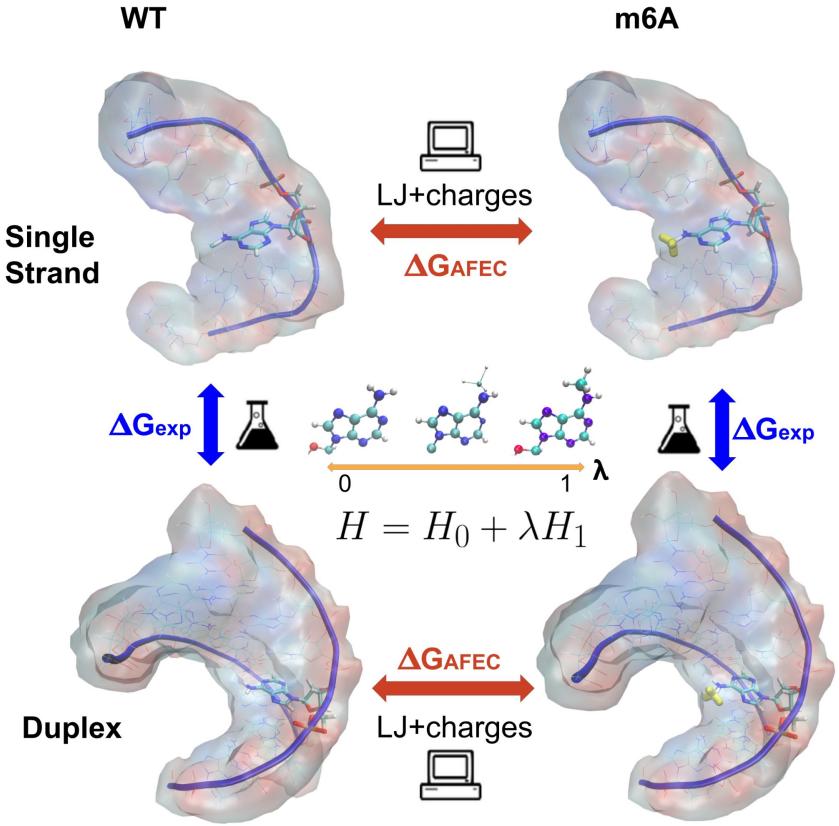
Roost et al, **JACS** (2015)

Kierzek et al, **Nat. Commun.** (2022)

Denaturation Experiments



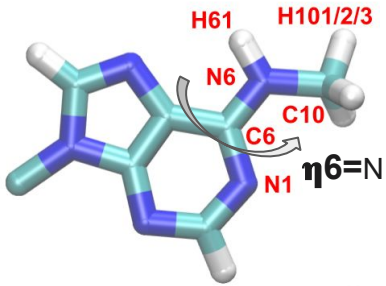
METHODS - Alchemical Free Energy Calculations (AFEC)

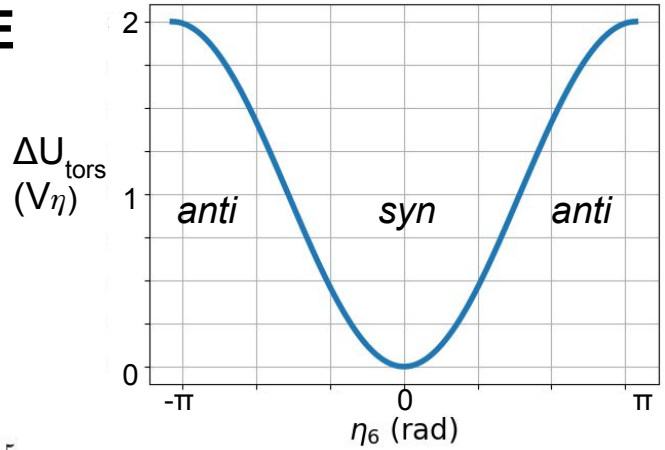


FITTING PROCEDURE

Fitting on:

- 6 partial charges
- **η6 torsional** potential to favor *syn* over *anti*





- Energy change associated to perturbed parameters:

$$\Delta U(x) = \sum_{i=5}^{5} K_i(x)\Delta Q_i + \sum_{i=1}^{5}\sum_{j=i}^{5} K_{ij}(x)\Delta Q_i \Delta Q_j + V_\eta[1+cos(\eta_6(x_i)-\pi)]$$

- ΔGs computed through a reweighting procedure

$$C = \chi^2 + \alpha \sum_{i=1}^{5} \Delta Q_i^2 + \beta V_\eta^2$$

- Fitting: minimization of a cost function:

$$\chi^2 = \sum_{i=1}^{N_{exp}} \frac{(\Delta\Delta G_{afec} - \Delta\Delta G_{exp})^2}{\sigma_i^2}$$

- Hyperparameters chosen in order to avoid overfitting (**Cross validation**) and ensure sufficient Statistical significance (**Kish Size**)

RESULTS

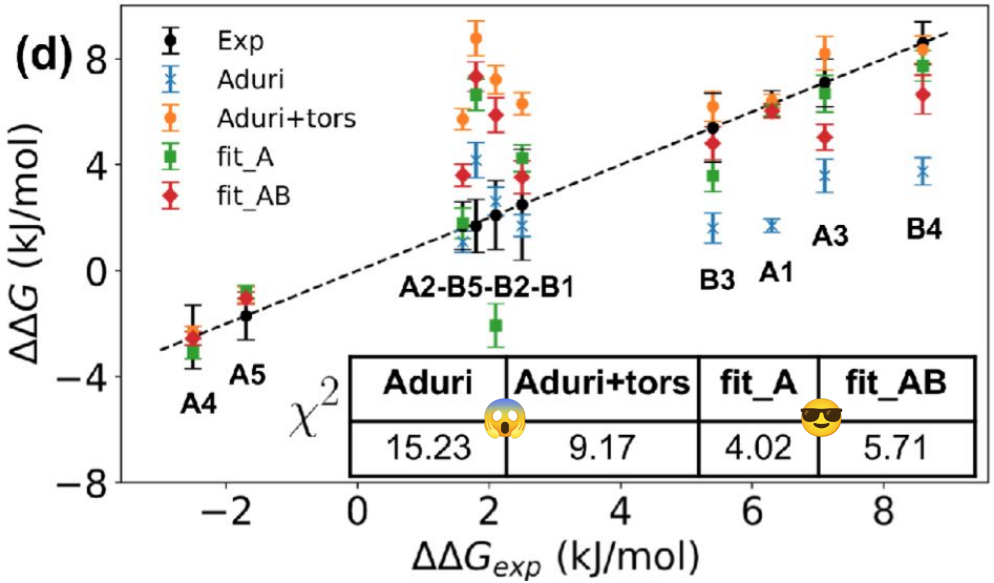
- 4 parametrization are compared:

  - Aduri (Reference)

  - Aduri+tors (tors. Potential to enforce A1)

  - Fit_A (fitting on A1-A5 with α=10 e^{-2} β=0)

  - Fit_AB (fitting on the entire data set with α=50 e^{-2} and β=0)



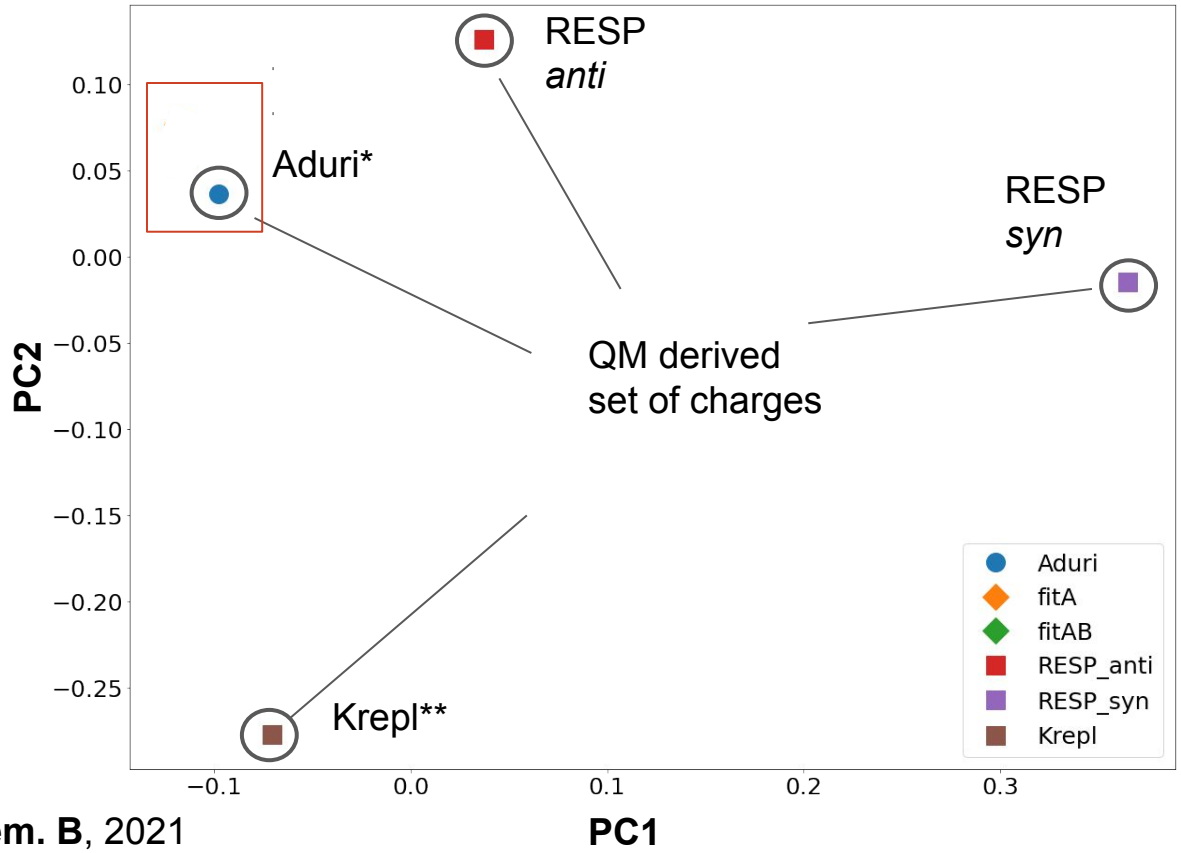
- Fitting is able to **improve agreement** with experiments with relatively "**small perturbation on the charges"**

ΔQ (e)	C6	N6	H61	N1	C100	H101
fit_A	0.019	0.077	0.099	-0.046	0.004	-0.051
fit_AB	0.009	0.049	0.067	-0.053	0.033	-0.035

	Aduri	+tors	fit_A	fit_AB
2*Vη (kJ/mol)	**0**	**4.70**	**4.92**	**4.98**

“Small perturbation on the charges”

Fitted charges are “closer” to reference compared to other QM derived set of charges



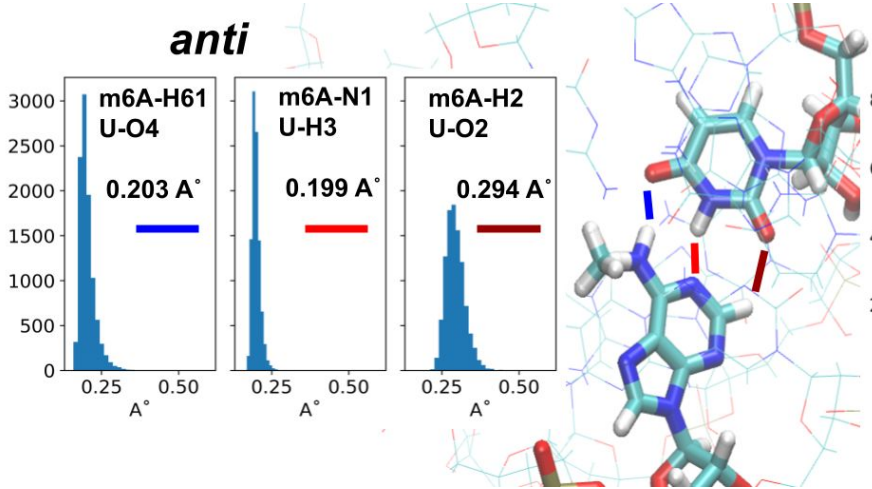
* *Aduri et al*, **JCTC**, 2007

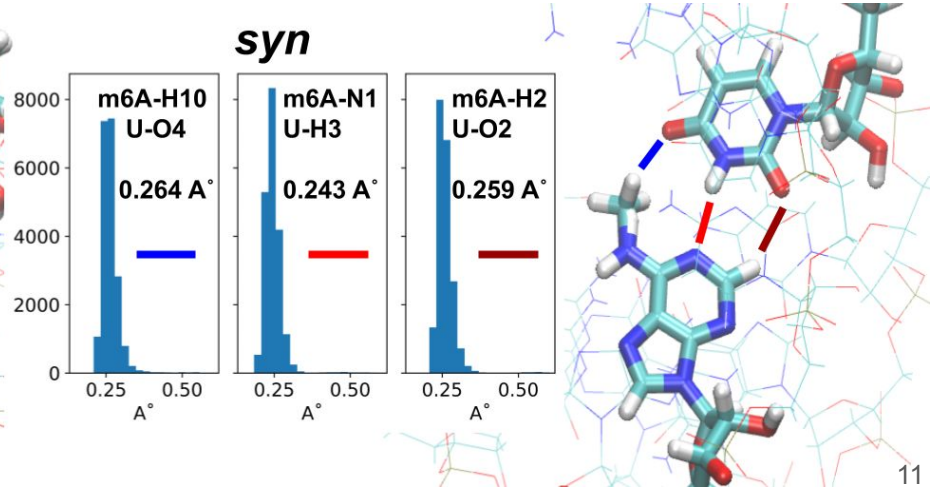
** *Krepl et al,* **J. Phys. Chem. B**, 2021

RESULTS

The fitting is **transferable**: **ΔG***syn/anti* in duplexes A2-A3 was not included in the fitting, but it is better reproduced by the fitted parameters

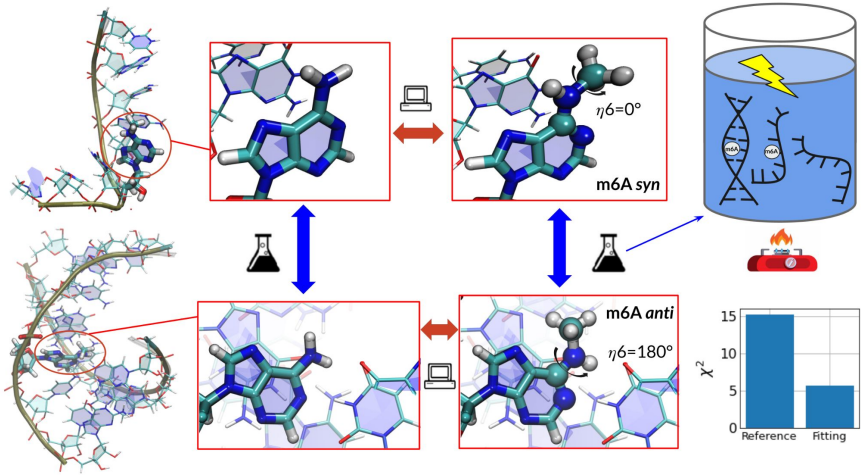
ΔG*syn/anti*	**Aduri**	**Aduri+tors**	**fit_A**	**fit_AB**	**Exp**
A1 (kJ/mol)	1.71 ± 0.25 😱	6.33 ± 0.25 😎	6.07 ± 0.21 😎	6.04 ± 0.26 😎	6.3
A2 (kJ/mol)	-7.7 ± 0.5 😎	-3.1 ± 0.4 😱	-10.4 ± 0.5 😎	-7.8 ± 0.4 😎	~ -11
A3 (kJ/mol)	-5.4 ± 0.5 😎	-0.8 ± 0.4 😱	-4.9 ± 0.4 😎	-5.8 ± 0.5 😎	<< 0
2*Vη (kJ/mol)	**0**	**4.70**	**4.92**	**4.98**	





1st part - CONCLUSIONS

- **First** attempt to tune partial charges of a biomolecular force field based on experiments performed on macromolecular complexes.
- Methodological Contribution: Fitting Strategy that allows **AFEC** to be use as a reference
- The fitting allows *syn*/*anti* balance and optical melting experiments to be reproduced with a very small perturbation on the charges.
- The fitting is transferable and opens the way to the use of MD to quantitatively **investigate the effects of N6 methylations on RNA structural dynamics and recognition** (ongoing)



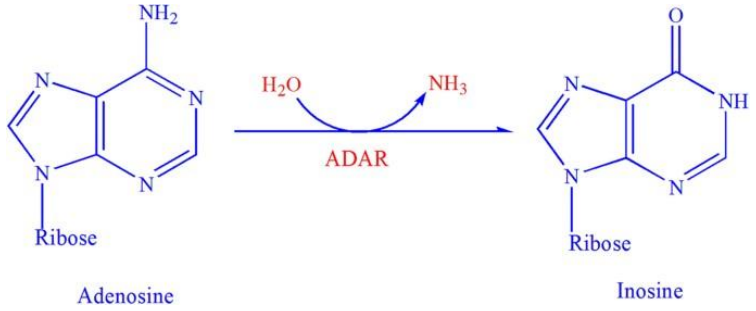
Piomponi et al, **ACS**, 2022

Combining MD and experiments to investigate the structures of a inosine-rich 20-bp RNA helix

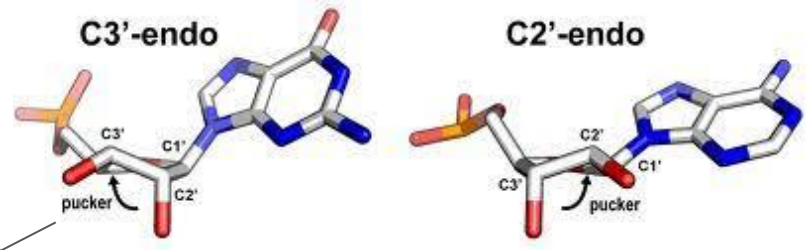
A - U
C - G
U - A
G - C
G - C
A - U
C - G
A - U
I - U
I - U
U - I
I - U
C - G
U - A
C - G
C - G
G - C
A - U
G - C
G - C

- A-to-I editing regulation affects immune response
- I-U bp introduce **flexibility** on dsRNA

Eisenberg et al, **Nat. Rev. Gen**. 2018



NMR J-couplings data (Sattler group in Munich) indicate high C2'-endo populations in the center part of the helix



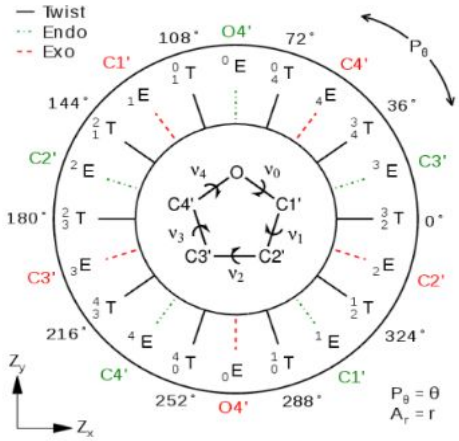
Common configuration in RNA A-helix

AIM: Use MD to construct an ensemble of structures compatible with NMR data

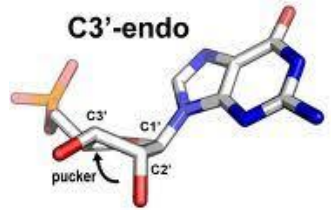
METHODS - enhanced sampling

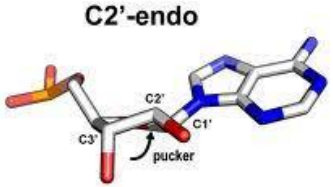
Independent well-tempered MetaDynamics along Z.x variable for 24 nucleotides

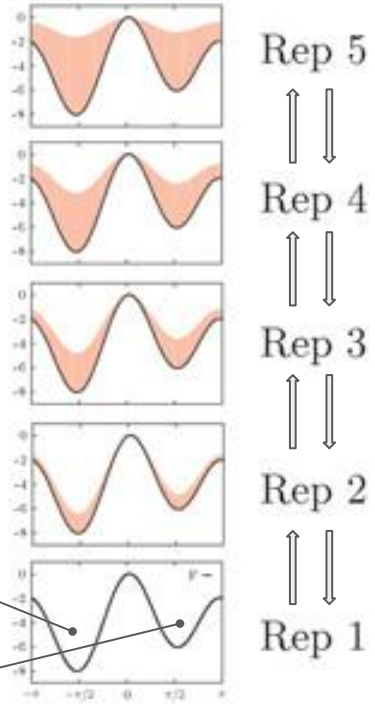
$$Z_x = \frac{\nu_1 + \nu_3}{2\cos(4\pi/5)}$$



- Concurrent metadynamics integrated in HREX.
- Different replicas have different strengths of the bias potential
- Replica 1 is unbiased







Gil-Ley et al, **JCTC**, 2015

METHODS - Maximum Entropy

Find Prob distr. As close as possible to the prior. among those compatible with **experimental averages**:

$$U(x_j) = U_0(x_j) + K_B T \sum_i^m \lambda_i F_i(x_j)$$

Lagrangian Multipliers λ found **minimizing Γ** function

$$\Gamma = \log(Z(\lambda)) + \sum_i^m \lambda_i F_i^{EXP} + \frac{1}{2} K \vec{\lambda}^2$$

A **regularization term** is added to avoid overfitting

$$Z(\lambda) = \sum_j^N w_j^0 \exp[-\sum_i^m \lambda_i F_i(\mathbf{x}_j)]$$

Forward model: Karplus Equations

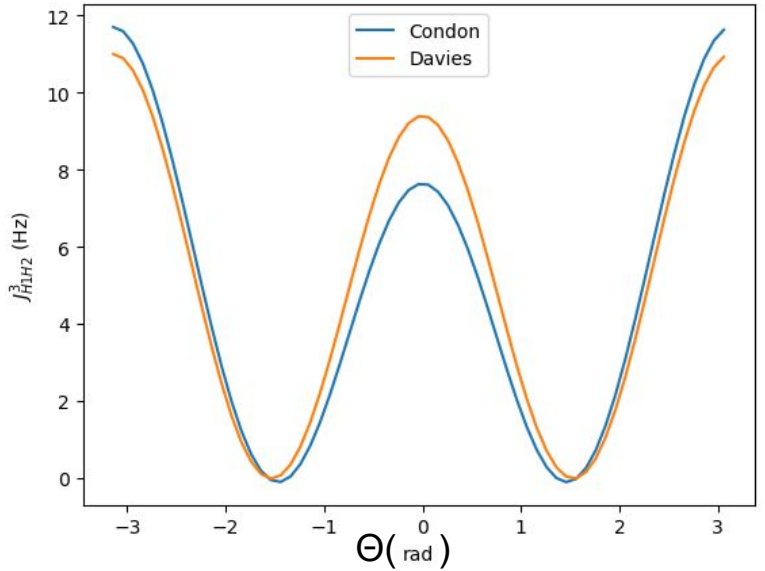
$$J_{H1H2} = A\cos^2(\theta) + B\cos(\theta)$$

Condon:
A=9.67 Hz ; B= -2.03 Hz

Condon et al, **JCTC**, 2015

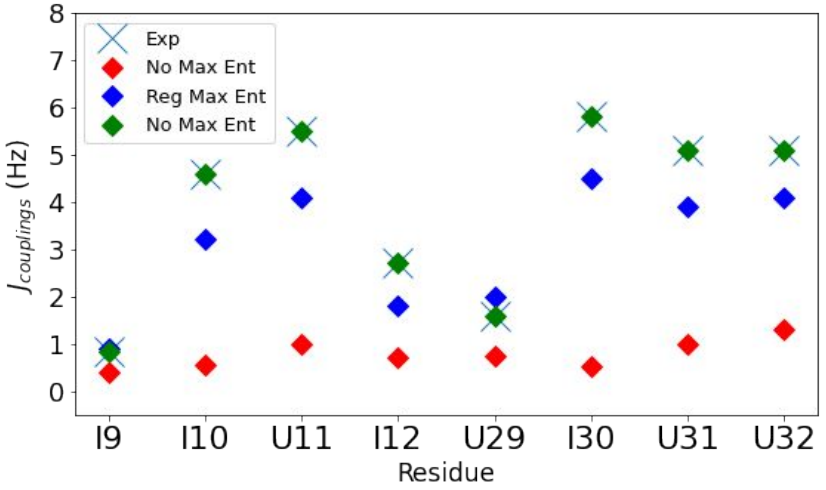
Davies:
A=10.2 Hx ; B= -0.8 Hz

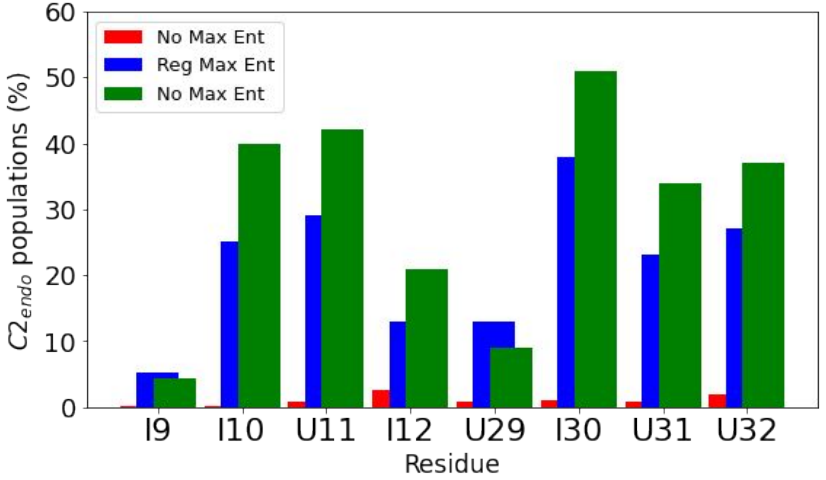
Davies et al, **Prog. NMR Spec.** 1977



Preliminary Results

- MD predicts very low populations of the C2’-endo conformation (~ 1/2 % for central nucleotides)
- Max Ent is able to perfectly enforce
  Experimental J couplings by increasing the weights of the C2’-endo structures





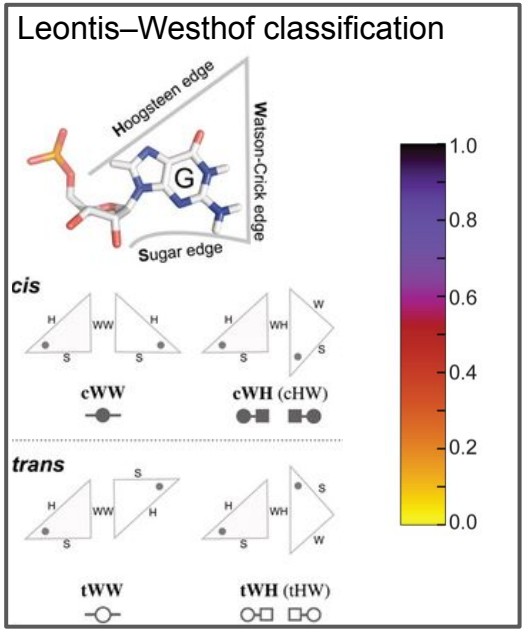
Clustering based on C2endo populations

Dynamics secondary structure *Bottaro et al*, RNA, 2019

C2’-endo

C3’-endo

Leontis–Westhof classification

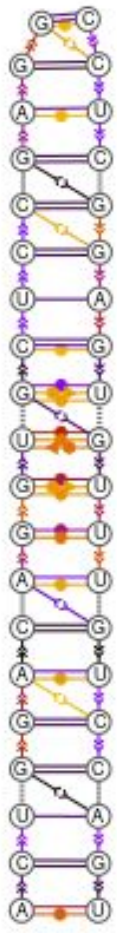


Populations:

Condon Reg. ME 33.9%

Davies Reg. ME 28.1%

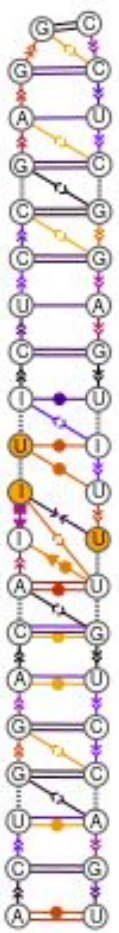
Davies Not Reg. ME 8.6%



2.8%

5.1%

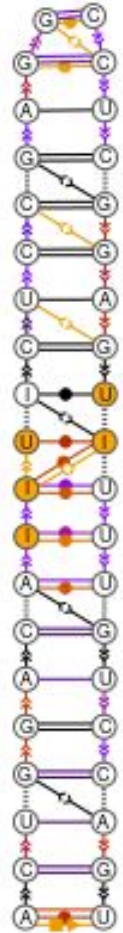
8.6%



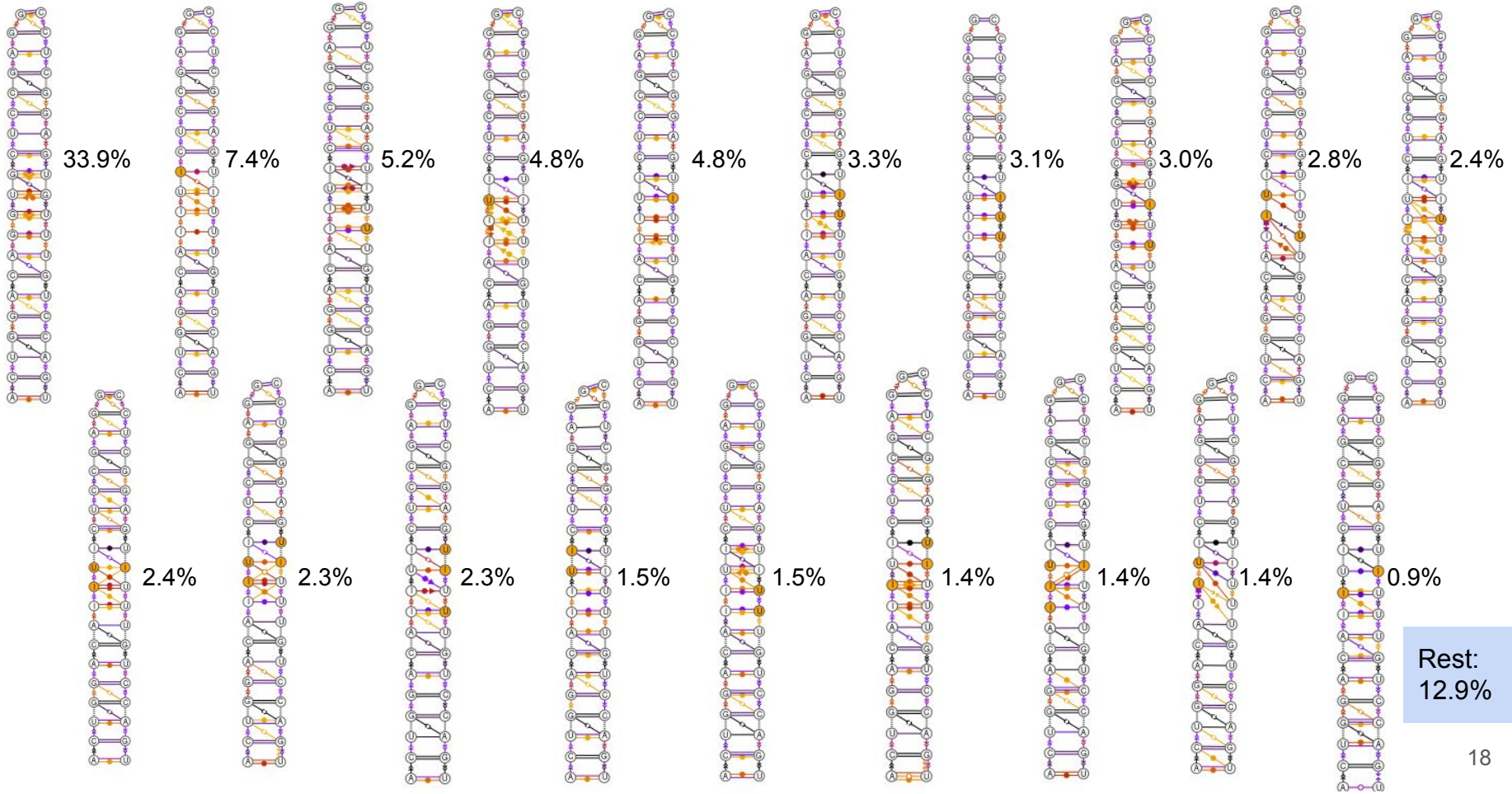
1.0%

1.3%

2.1%

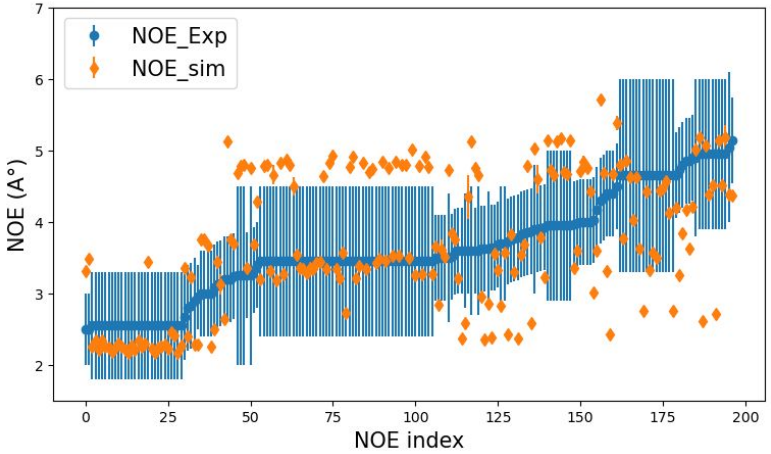


Clusters for Reg. Max. Ent. using Condon parameters



Validation on other solution experiments

- NOEs (NMR)
  Signal related to
  Protons distances



- No Max Ent - 60% agreement
- Reg Max Ent - 69 % agreement

- SAXS
  (related to Radius of Gyration)

	Adenosine helix	Inosine helix	Ino. helix Reg Max Ent	Ino. helix Max Ent
<rG> (Å)	1.81	1.79	1.84	1.87
Std (Å)	0.06	0.08	0.11	0.13

Acknowledgments

Christoph Müller-Hermes

Prof. Dr. Michael Sattler

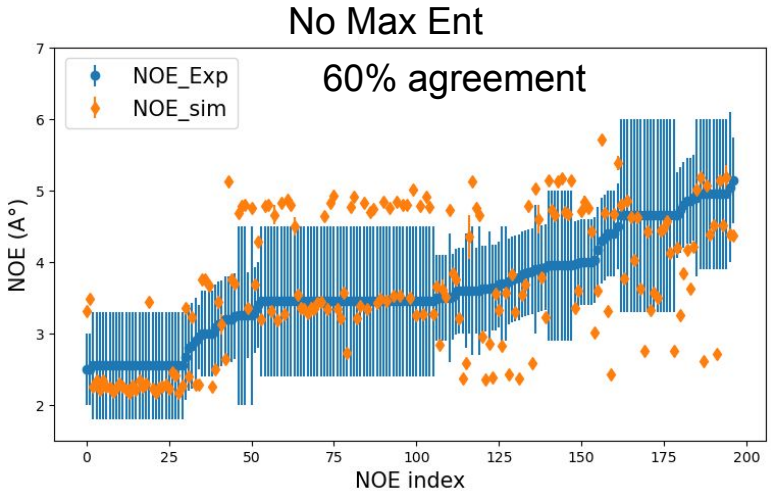
HELMHOLTZ MUNICH

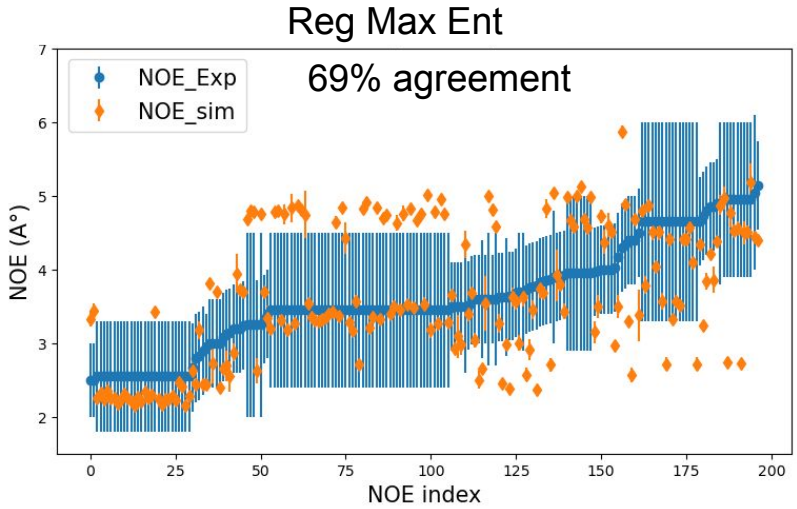
Technical University of Munich

Thanks for your Attention!

Validation on other solution experiments

- NOEs (NMR)



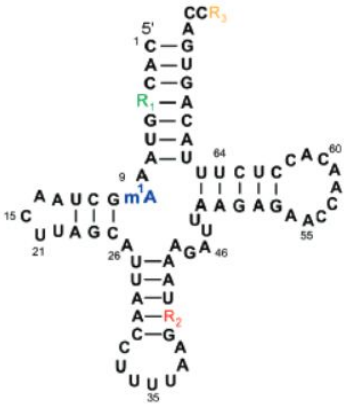


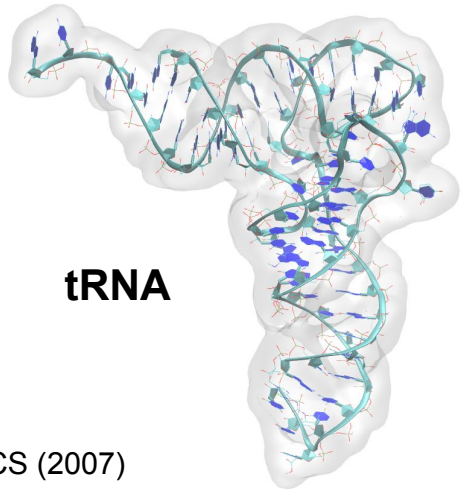
- SAXS (related to Radius of Gyration)

	Adenosine helix	Inosine helix	Ino. helix Reg Max Ent	Ino. helix Max Ent
<rG> (Å)	1.81	1.79	1.84	1.87
Std (Å)	0.06	0.08	0.11	0.13

RNA structure and function

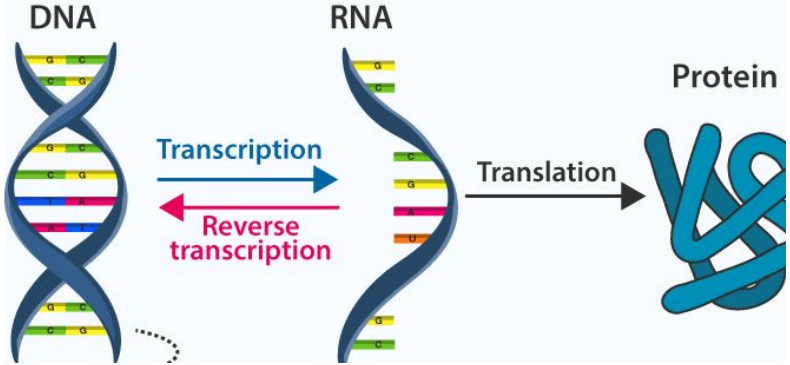
- Polymeric molecule essential for **coding** (mRNA)
- Many other fundamental roles in the cell (non-coding RNA: tRNA, rRNA, etc.)
- **structural dynamics** is important for function

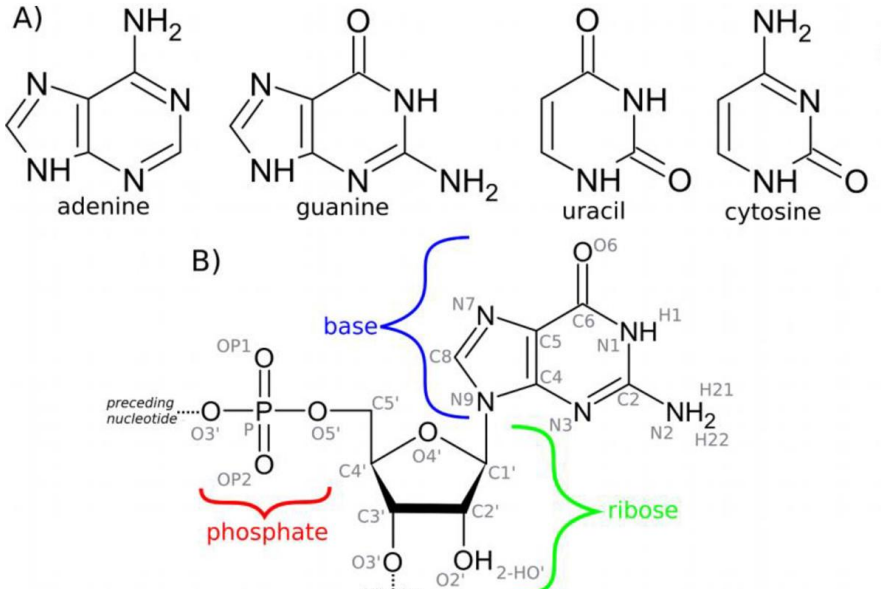




tRNA

Voigts-Hoffman et al, JACS (2007)

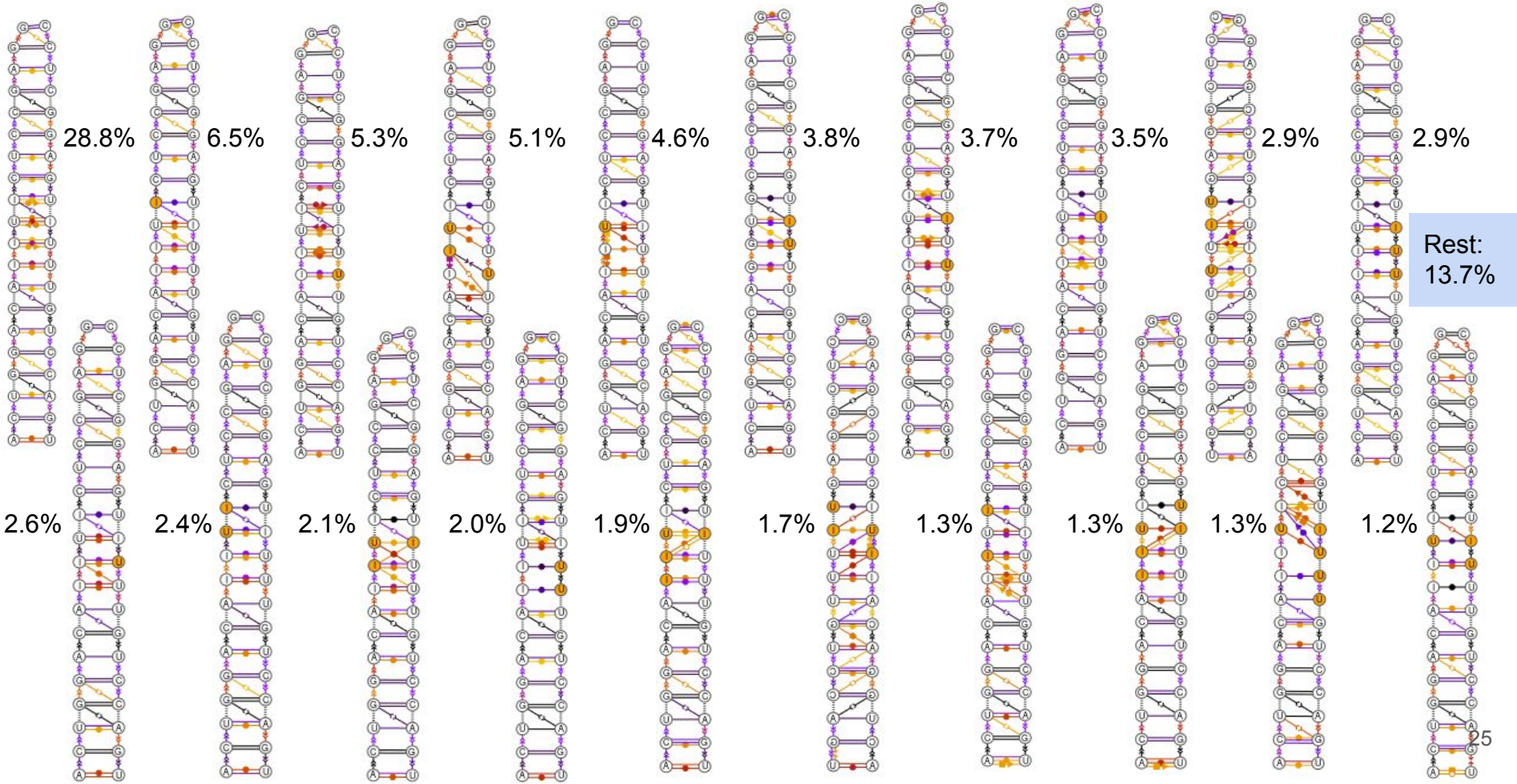




Sponer, Bussi, et al, Chem Rev (2018)

Thanks for your Attention!

Clustering based on C2endo pop using Davies parameters



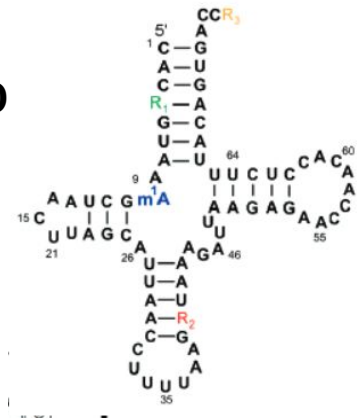
RNA structure and function

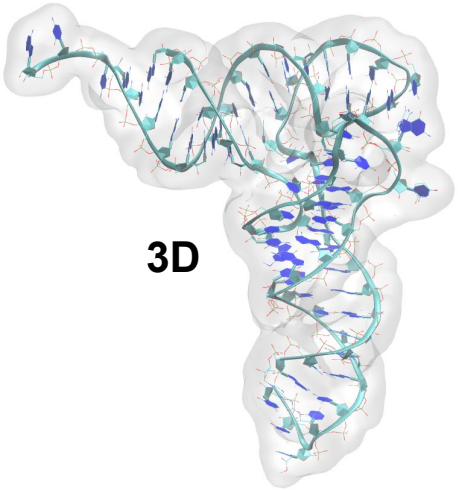
- Polymeric molecule essential for **coding** (mRNA)

- Many other fundamental roles in the cell (tRNA, rRNA, ncRNA) - 2D and 3D **structural dynamics** is important for function

1D ‘3-CCAGUGACAUUUCUCCACAACCAAGAG..-5’

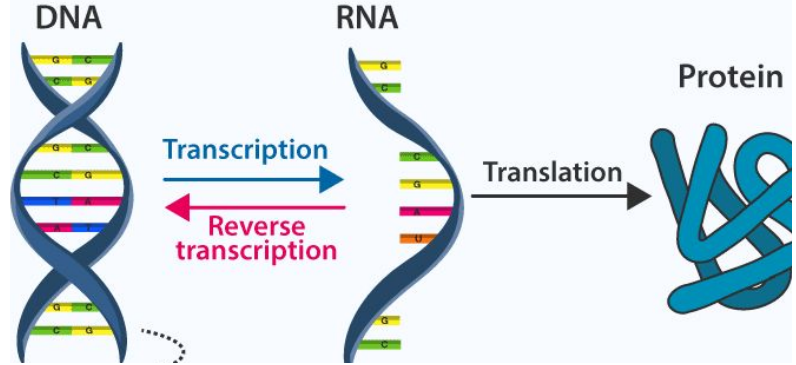
2D

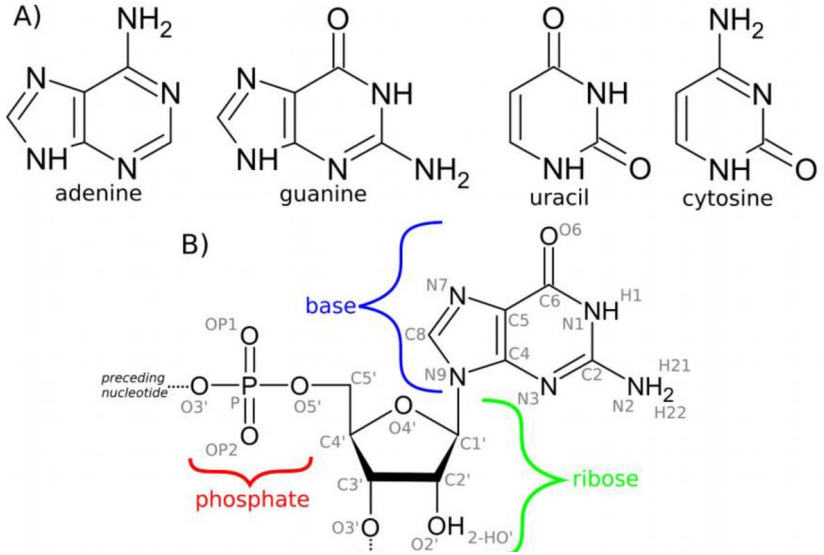




3D

Voigts-Hoffman et al, JACS (2007)





Sponer, Bussi, et al, Chem Rev (2010)

Preliminary Results

		No Max Ent		Reg Max Ent (Davies)		Max Ent (Davies)	
	Exp JC	MD JC	C2endo pop.	MD JC	C2endo pop.	MD JC	C2endo pop.
I9	0.84 Hz	0.40 Hz	0.025 %	0.90 Hz	5.2 %	0.84 Hz	4.4 %
I10	4.6 Hz	0.55 Hz	0.13 %	3.2 Hz	25 %	4.6 Hz	40 %
U11	5.5 Hz	1.0 Hz	0.80 %	4.1 Hz	29 %	5.5 Hz	42 %
I12	2.7 Hz	0.72 Hz	2.6 %	1.8 Hz	13 %	2.7 Hz	21%
U29	1.6 Hz	0.76 Hz	0.74 %	2.0 Hz	13 %	1.6 Hz	8.9 %
I30	5.8 Hz	0.54 Hz	0.92 %	4.5 Hz	38 %	5.8 Hz	51 %
U31	5.1 Hz	1.0 Hz	0.71 %	3.9 Hz	23 %	5.1 Hz	34 %
U32	5.1 Hz	1.3 Hz	2.0 %	4.1 Hz	27 %	5.1 Hz	37 %

Molecular dynamics

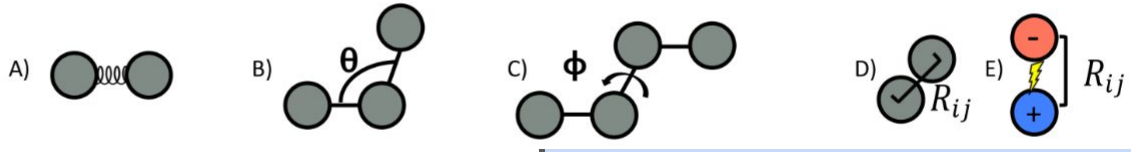
Powerful Computational tool to characterized molecular structural dynamics with “**unlimited”** resolution.

Empirical Force-fields:

Bonded

Non-bonded

$$E_{tot} = \sum_{bonds} K_r\,(r - r_{eq})^{\,2} + \sum_{angles} K_\theta\,(\theta - r\theta_{eq})^{\,2} + \sum_{dihedrals} \frac{V_n}{2}\left[1 + \cos(n\phi - \gamma)\right] + \sum_{i<j}\left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}}\right]$$



- No Polarization
- No chemical reactivity
- Quality of the force-fields is the bottleneck - Always integrate with experiments!

experiment

Field Fitting

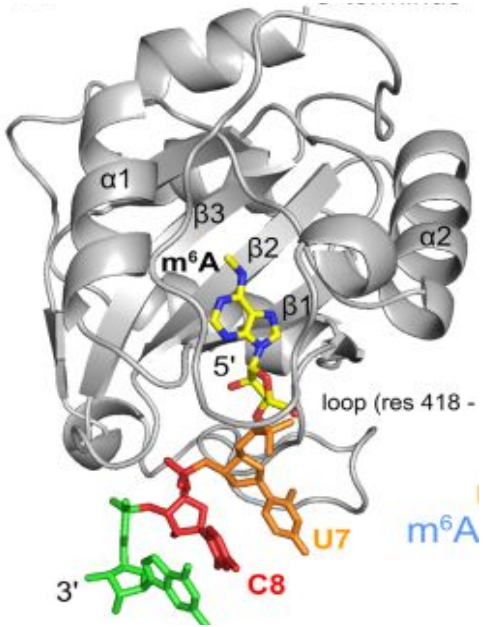
Ensemble refinement
(i.e Max Entropy, Minimum Parismony)

Future Project: Investigating the impact of m6A methylation on Free Energy of Bindings

2 months Exchange at Sponer Lab in Brno (Czech Republic).

- Use the developed AFEC methods and the refined force field to investigate the **impact of m6A on structure** and on **binding affinity** with m6A readers

- Combine the Alchemical Transformation with MetaD which enhances **water exchange** inside/outside the binding pocket



Krepl, **Jou. Chem. Phy. B**, 2021