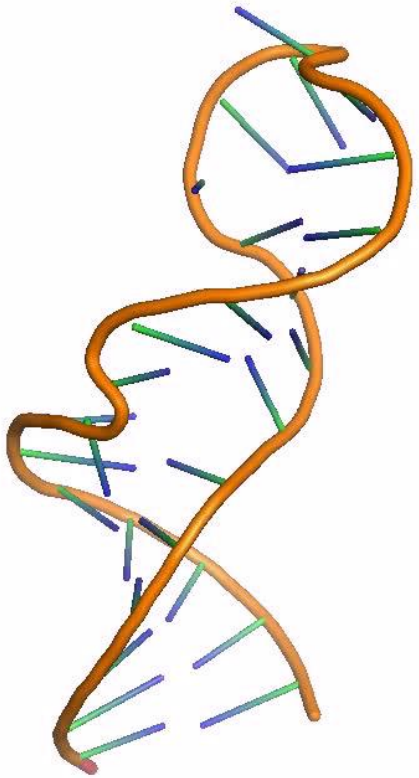
Uncovering Macromolecule Conformation Ensembles with X-ray Scattering Interferometry

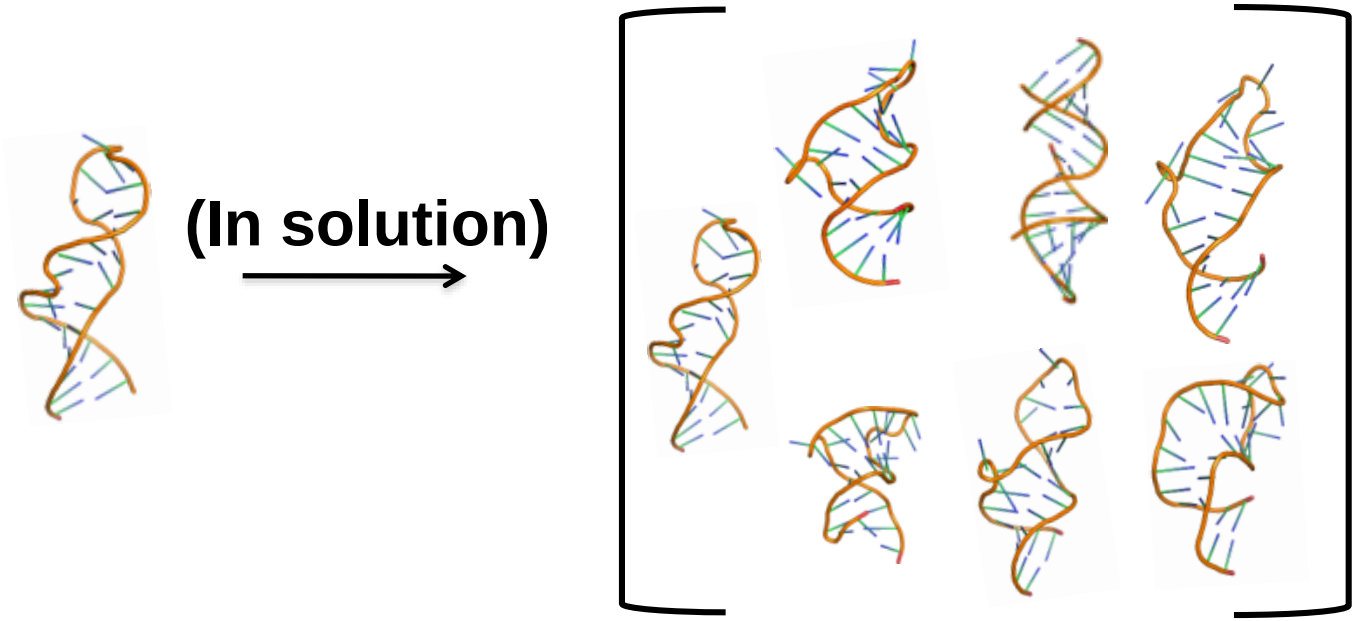
Xuesong Shi

Department of Biochemistry
Herschlag Lab
Stanford University

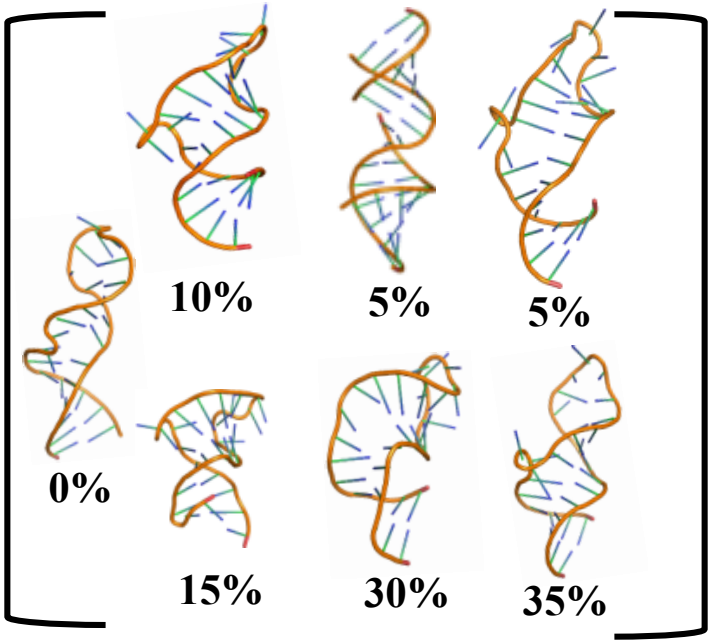
Macromolecules are dynamic in solution

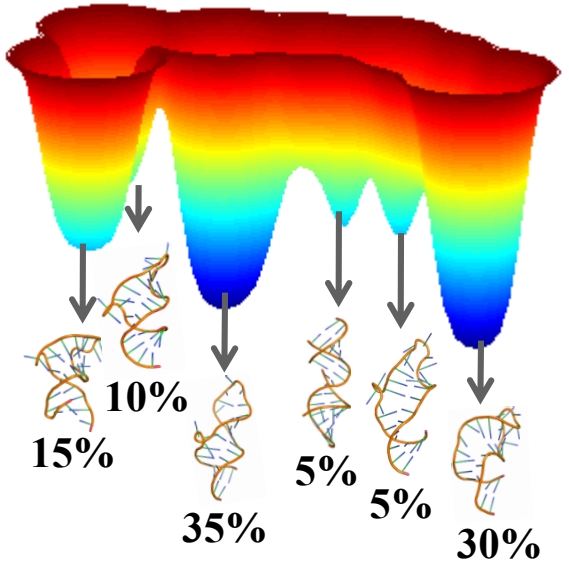


Macromolecules Exist as Conformational Ensembles



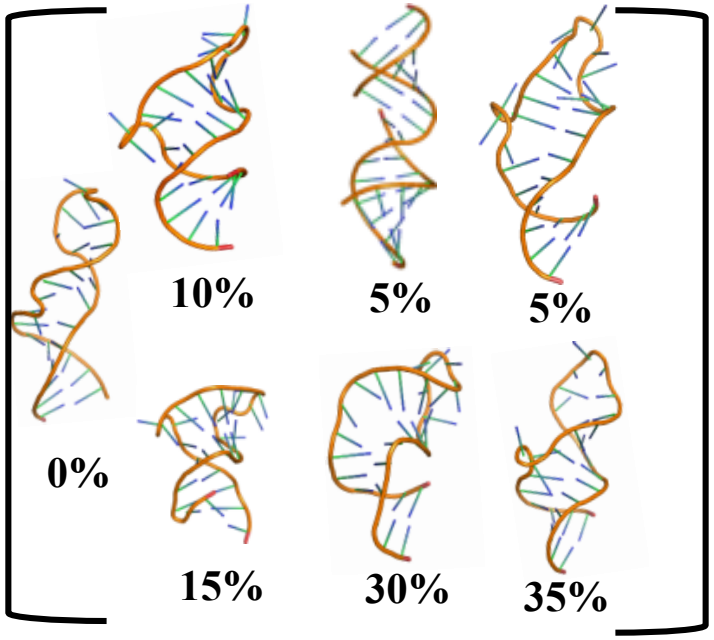
Macromolecules Exist as Conformational Ensembles

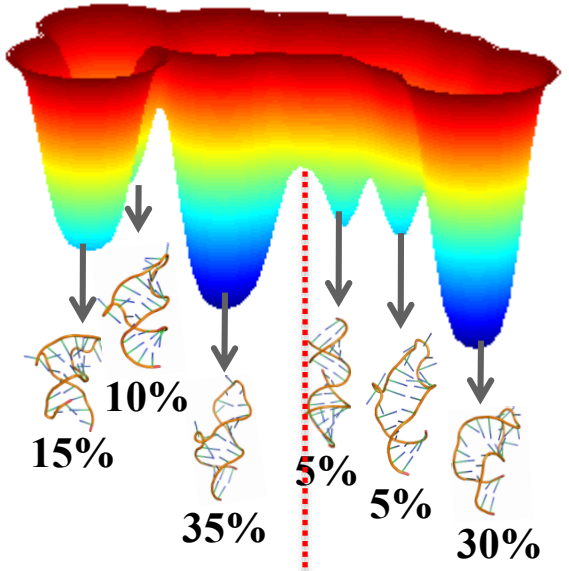


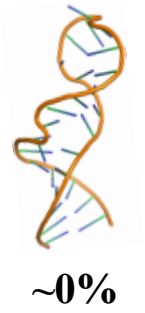


Conformational ensemble ⟺ **Energy landscape**

Macromolecules Exist as Conformational Ensembles

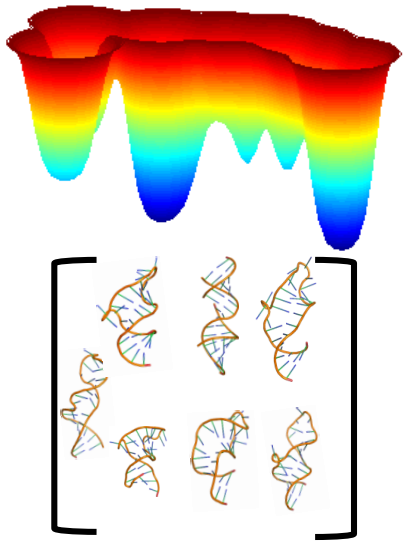






the average structure could be rarely populated in the actual ensemble

Why Care About Conformational Ensembles

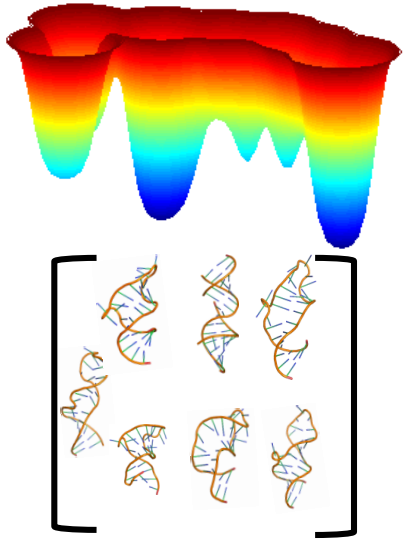


Why Care About Conformational Ensembles

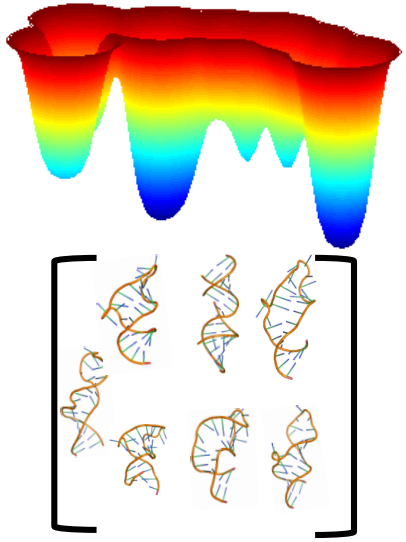
Solution structures

- **More accurate than an average structure**

~0%



Why Care About Conformational Ensembles



Solution structures

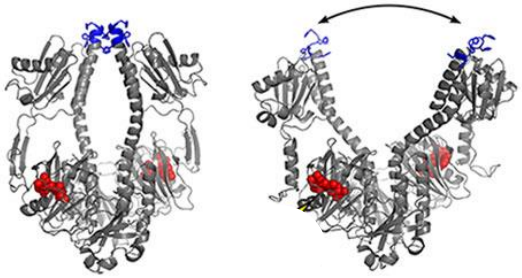
- **More accurate than an average structure**

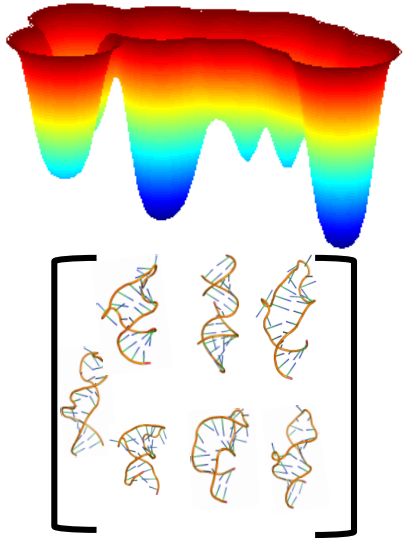
~0%

- **Too floppy to get an average structure (e.g. no crystals)**

Why Care About Conformational Ensembles

Conformational change and function





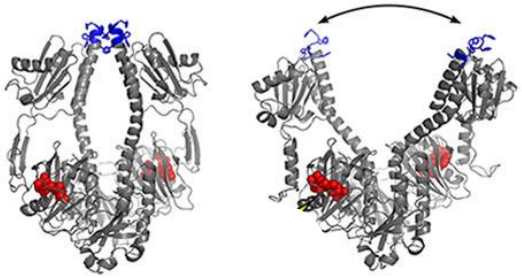
Solution structures

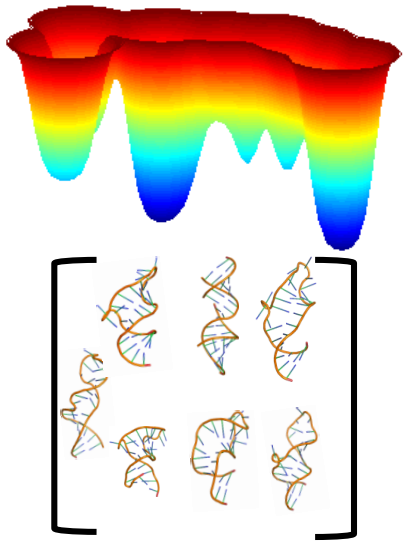
- **More accurate than an average structure**

- **Too floppy to get an average structure (e.g. no crystals)**

Why Care About Conformational Ensembles

Conformational change and function

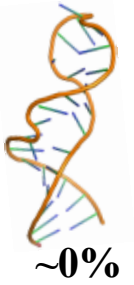




Folding and assembly

Solution structures

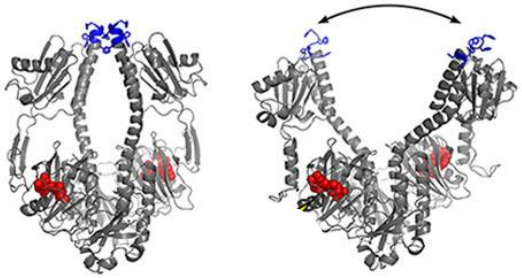
- **More accurate than an average structure**

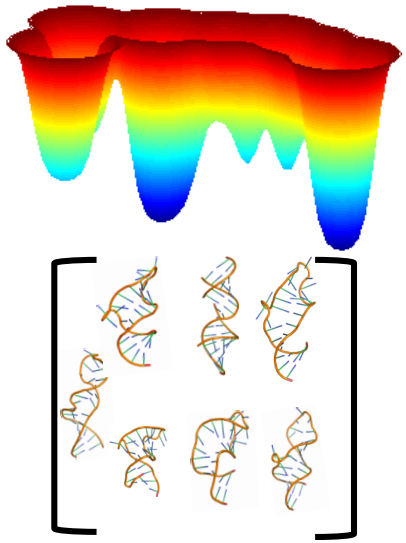


- **Too floppy to get an average structure (e.g. no crystals)**

Why Care About Conformational Ensembles

Conformational change and function



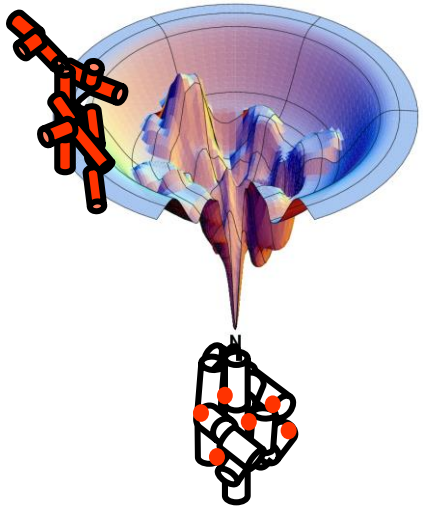


Solution structures

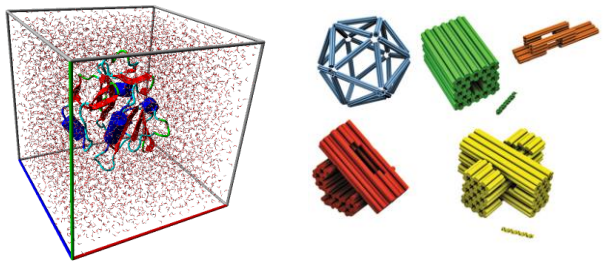
- **More accurate than an average structure**

- **Too floppy to get an average structure (e.g. no crystals)**

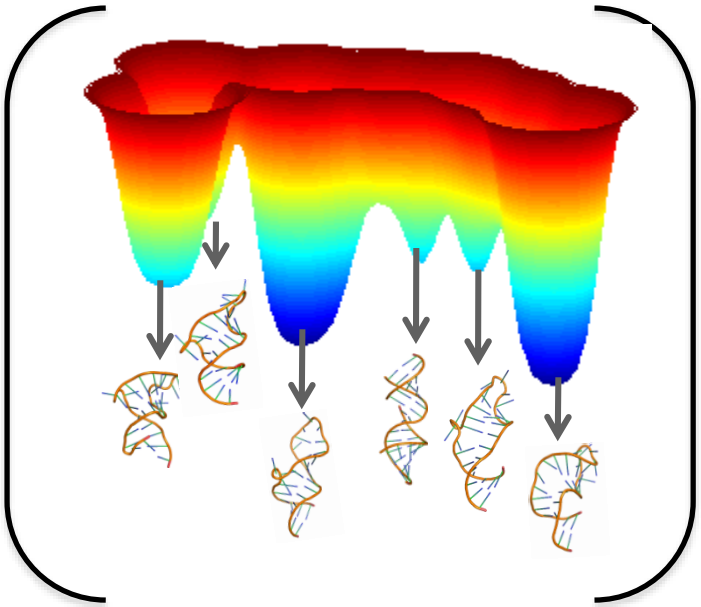
Folding and assembly



Better physical models for bio-computation and design



Conformational Ensembles Are Hard to Study

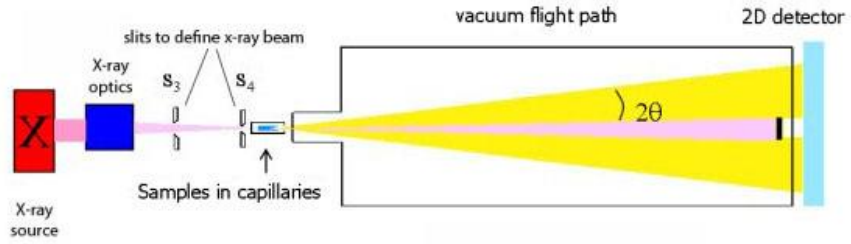


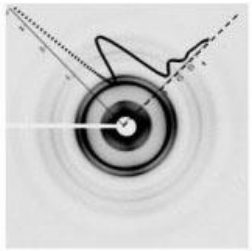
Averaged structural information

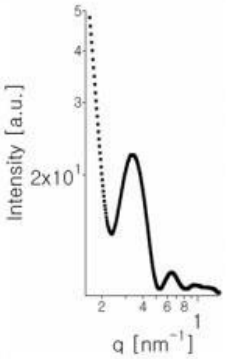
An average conformation

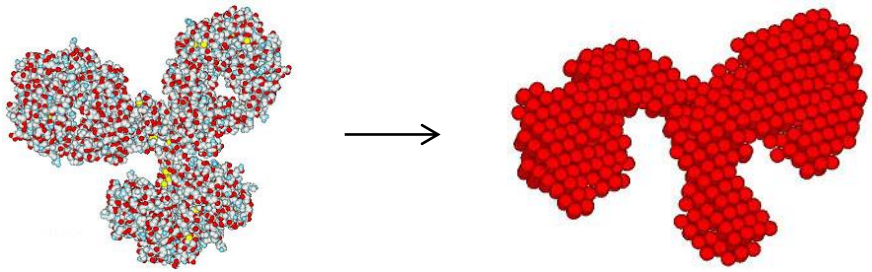
HOW TO PRECISELY MEASURE AN ENSEMBLE?

Small-Angle X-ray Scattering





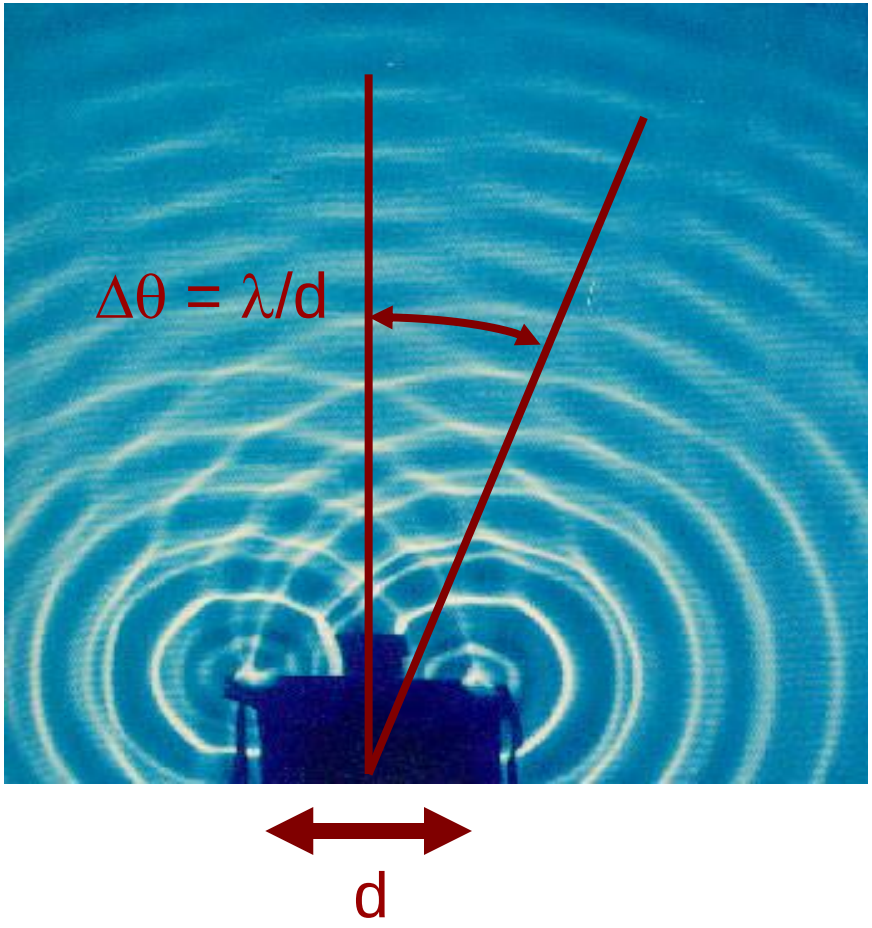




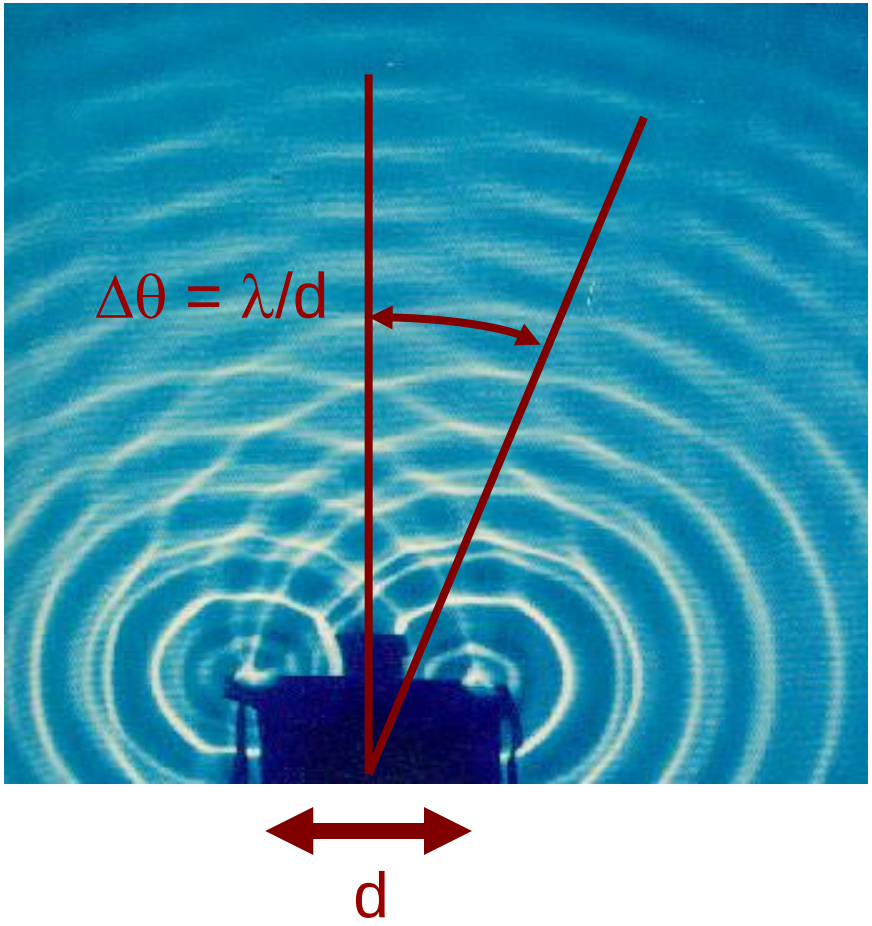
Strengths: solution phase, instantaneous, direct, gives absolute distances.

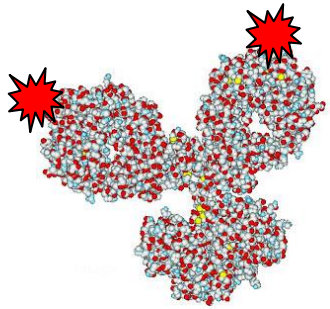
Weaknesses: summation of all intramolecular distances, admits degenerate models.

Scattering interference between two scatters

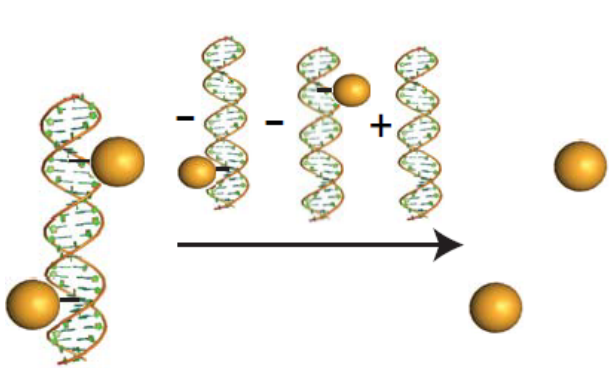


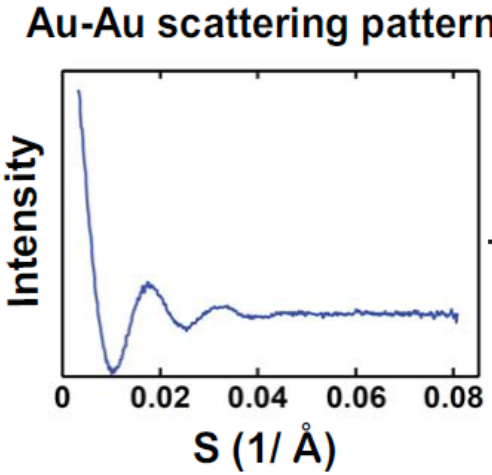
Scattering interference between two scatters

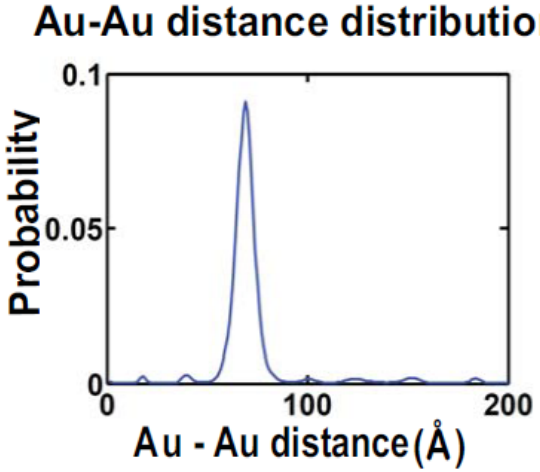




X-ray Scattering Interferometry





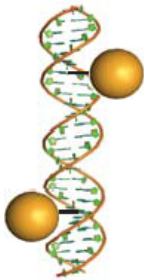


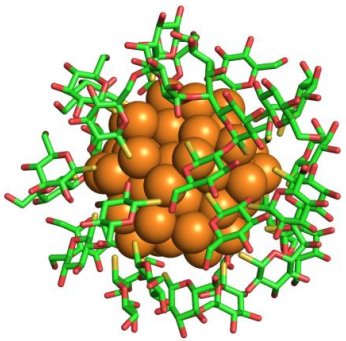
Pehr Harbury

Rebecca Mathew-Fenn

Mathew-Fenn et al. Science 2008
Shi et al. PNAS 2013

X-ray Scattering Interferometry



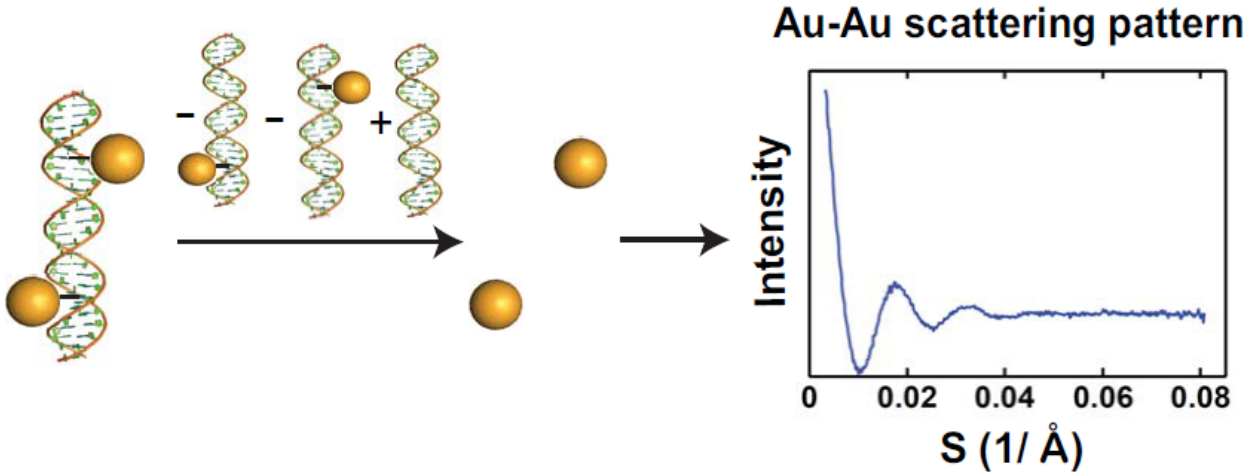


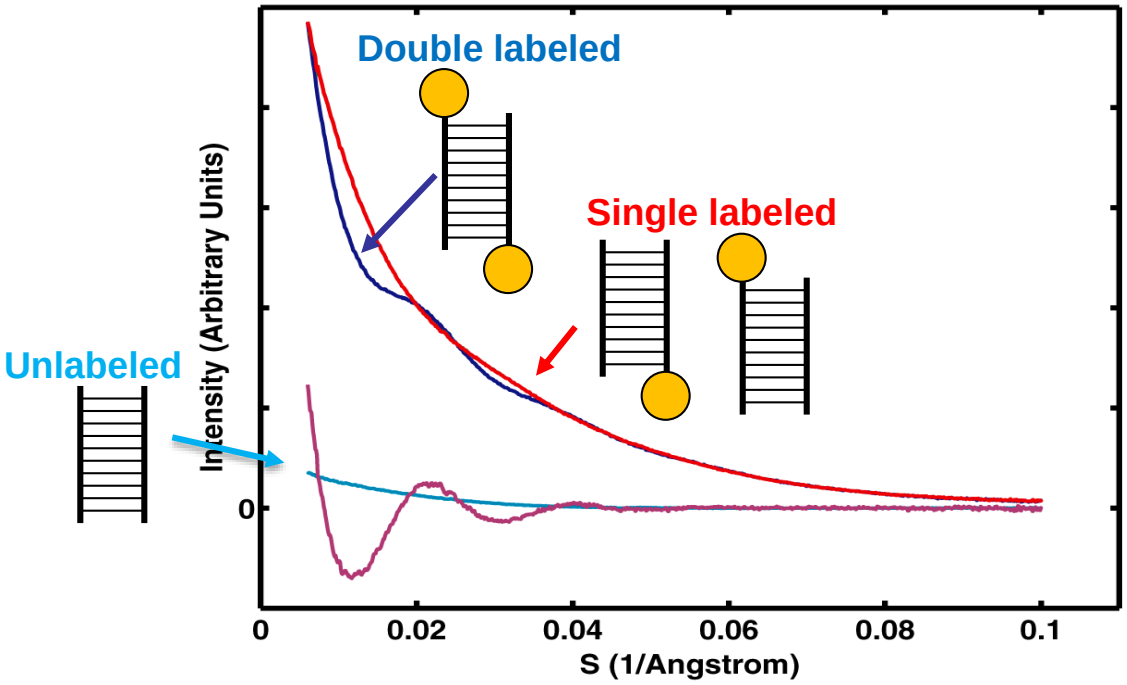
1.2 nm

2 nm

Mathew-Fenn et al. Science 2008
Shi et al. PNAS 2013

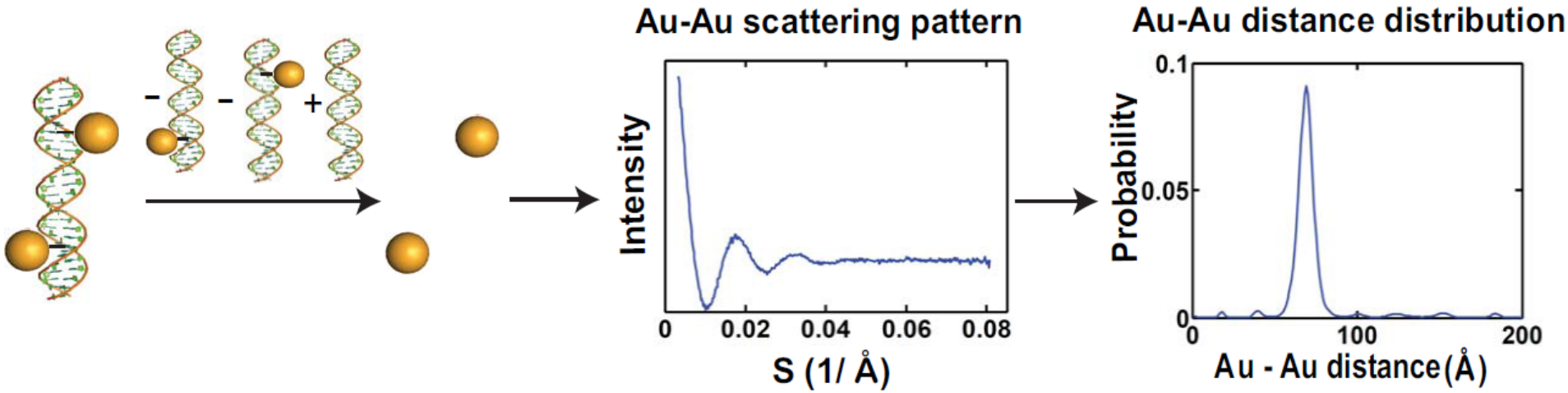
X-ray Scattering Interferometry





Mathew-Fenn et al. Science 2008
Shi et al. PNAS 2013

X-ray Scattering Interferometry



Same to SAXS:

- **Solution phase**
- **Instantaneous (time independent)**

Pehr Harbury

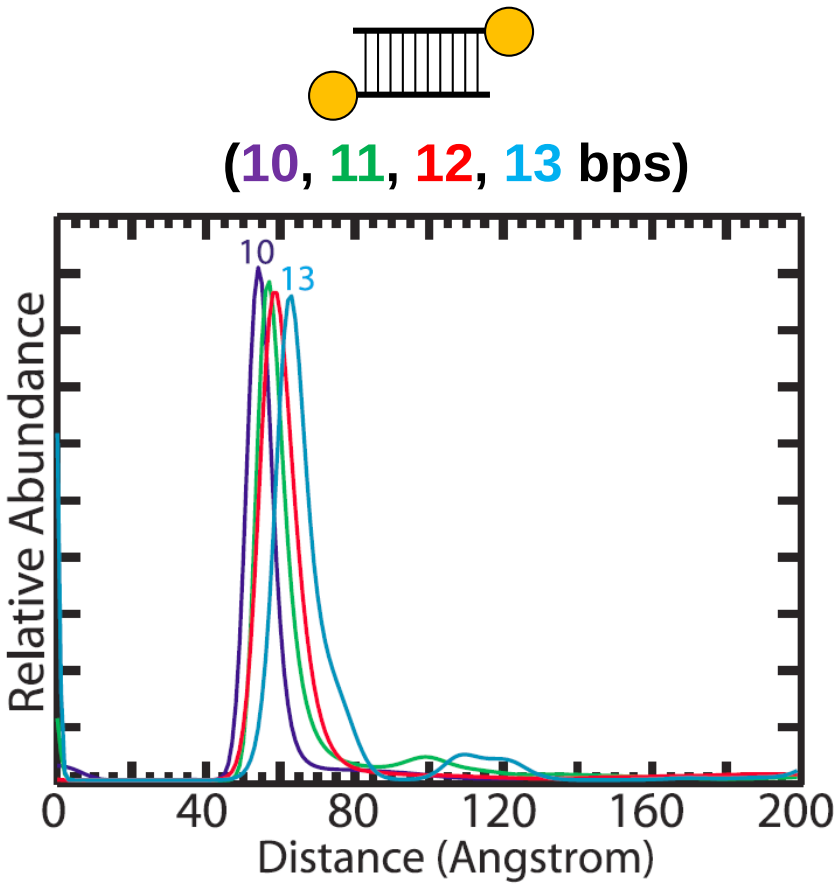
Rebecca Mathew-Fenn

Additional advantages:

- **Precise absolute distance (model independent)**
- **A distribution**

Mathew-Fenn et al. Science 2008
Shi et al. PNAS 2013

Proof of principles

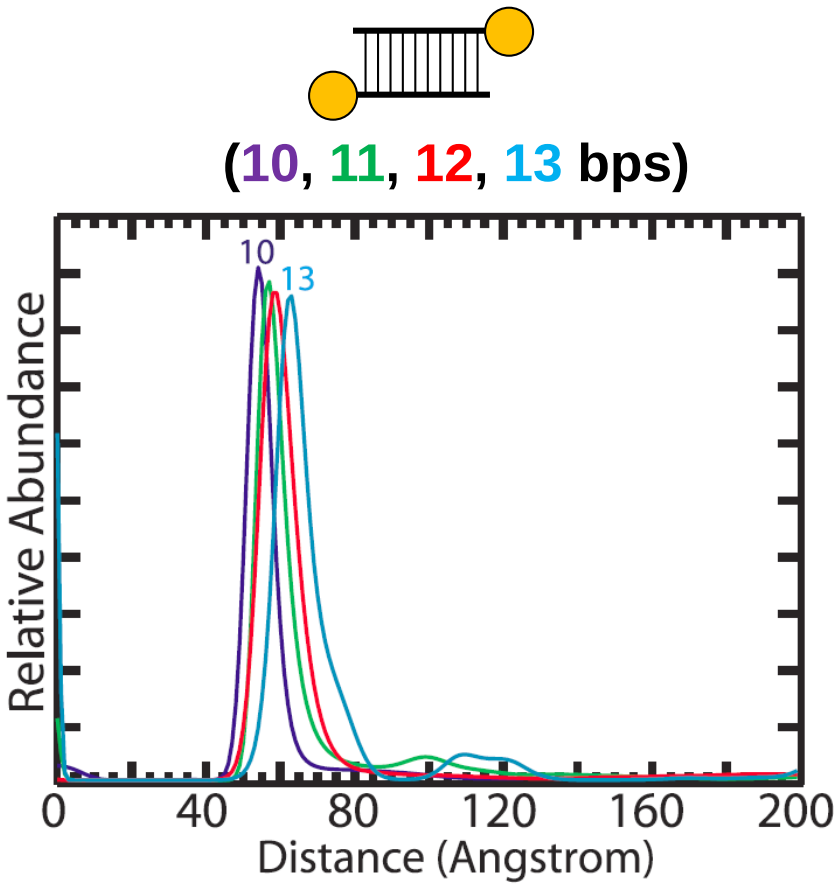


single base pair discrimination

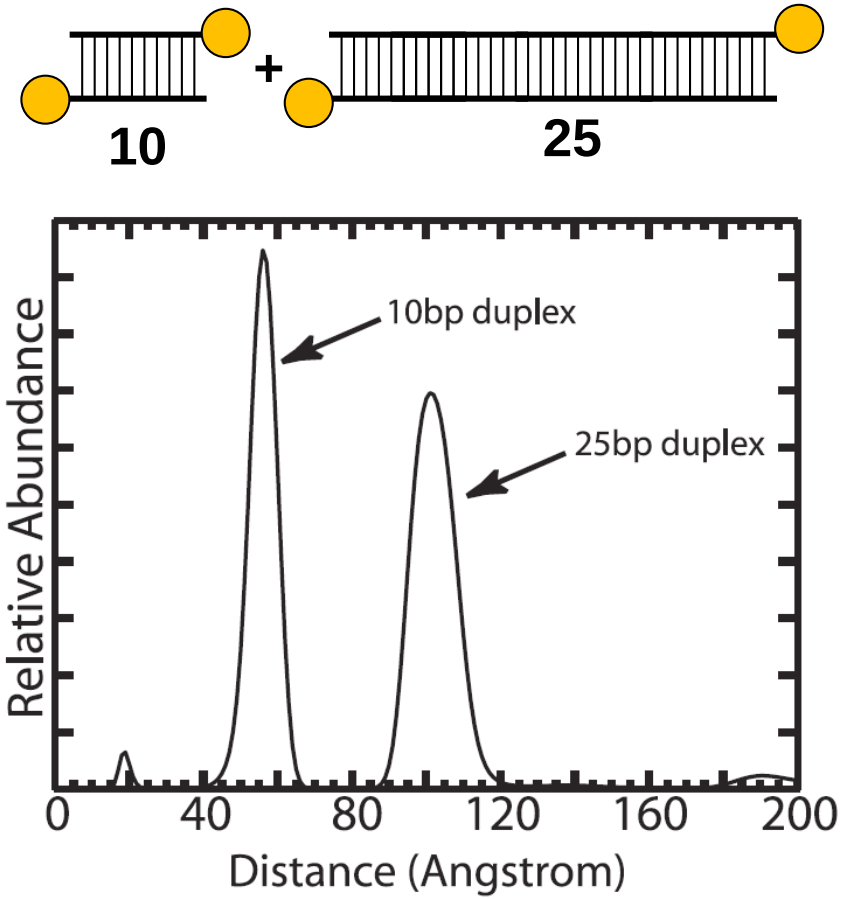
Precise absolute distance (sub-Å level reproducibility)

Mathew-Fenn et al. plos one 2008

Proof of principles



single base pair discrimination



population mixtures

Mathew-Fenn et al. plos one 2008

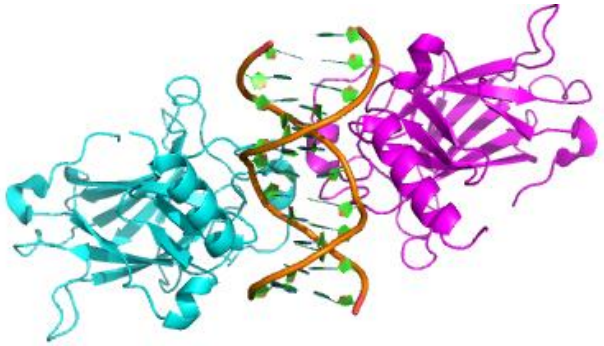
DNA (outline)

- **Elasticity and helical structure**
- **Sequence dependent variation**
- **Ensemble**
- **Atomic and physical models**

ELASTICITY OF DNA

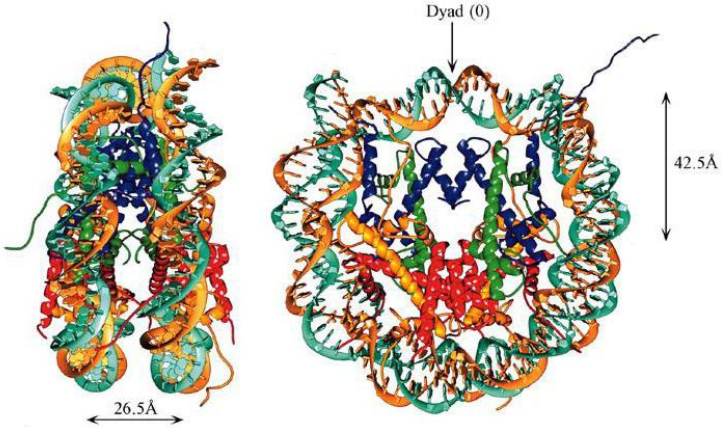
Shi et al. PNAS 2013

Why Does DNA Elasticity Matter?

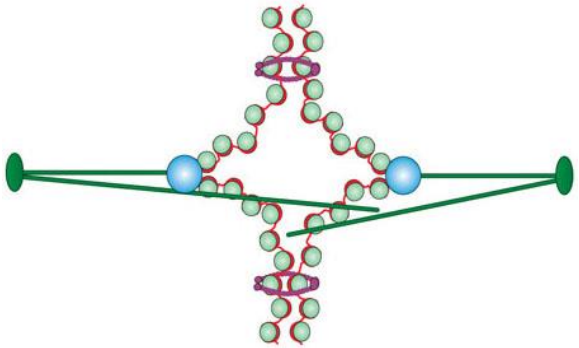


DNA-protein interactions

transcription and replication
twisting

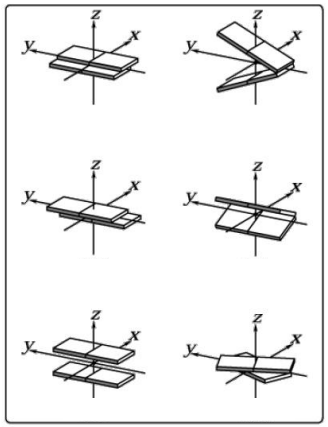


DNA packing and regulation
bending

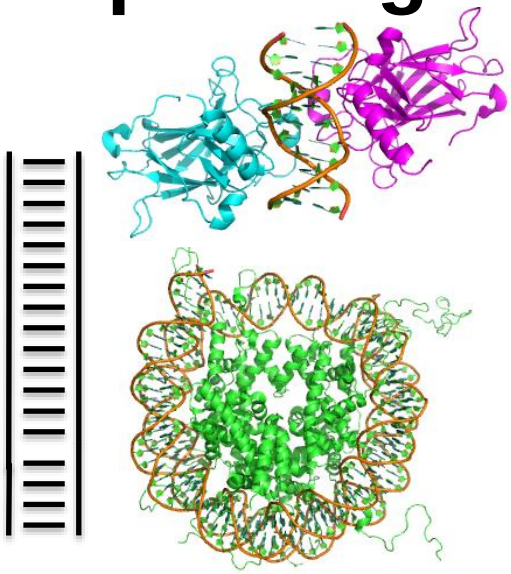


e.g. chromosome
segregation
stretching

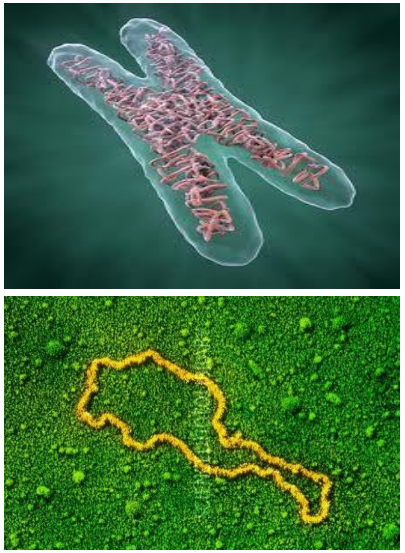
Understand DNA from macroscopic to microscopic length scale



Base-step

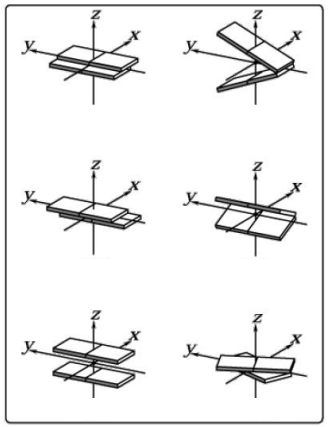


10s of bps

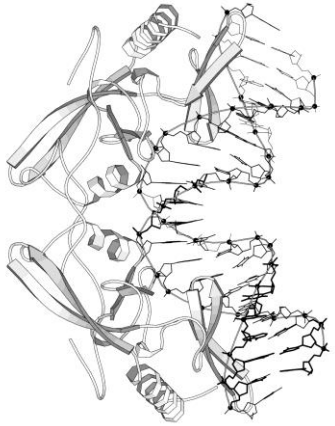


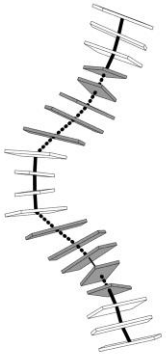
kbs, genome

“Knowledge-Based” Microscopic DNA Model

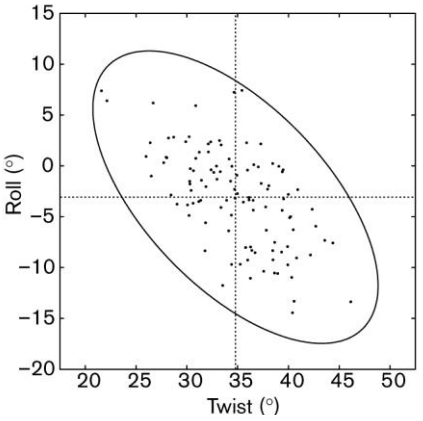


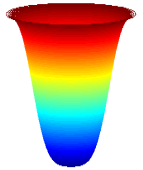
Base-step



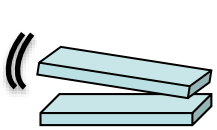


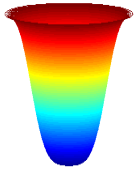
Crystal-structure database of DNA-protein complexes





Average bending fluctuation per base step: ~4°

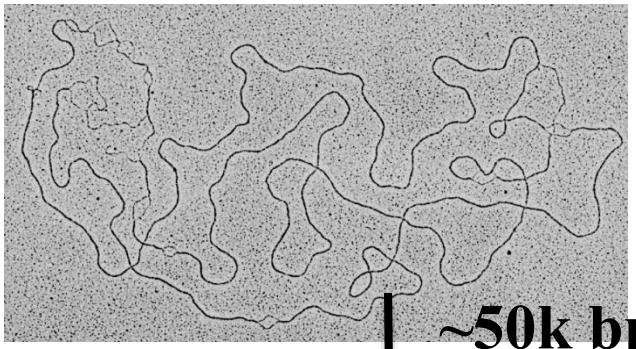




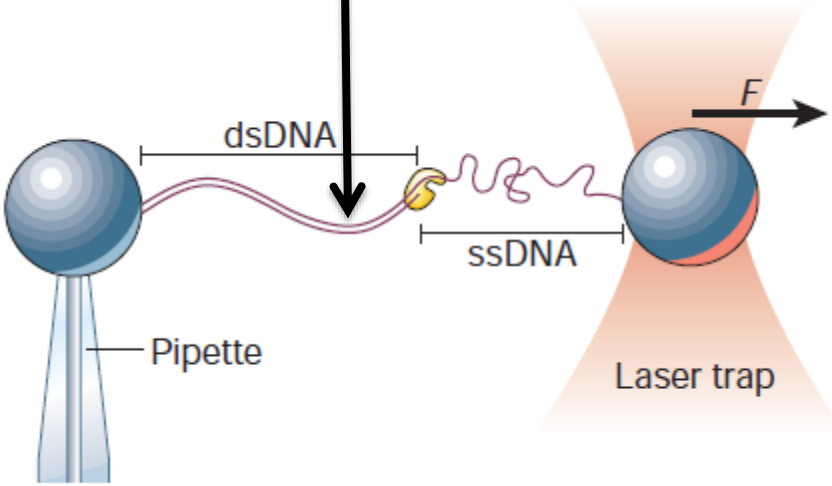
Bending persistence length:
B ~ 100 nm

Olson (1998) PNAS

Single molecule studies of DNA polymer



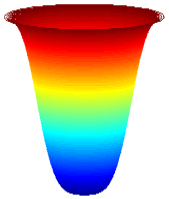
~50k bp



Mechanical properties of long DNA (> 1k bp)

Bending persistence length:
B = 50 nm (150 bp)

Average bending fluctuation per base step: ~6°



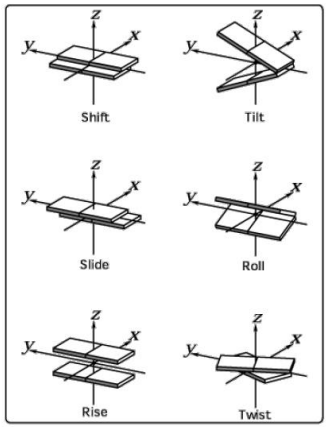
Discrepancy

	(macroscopic)	(microscopic)
	Single molecule Stretching of long DNA (> 1k bp)	**Survey DNA/protein crystal-structures**
Bending persistence length:	**50 nm**	**100 nm**
Average bending fluctuation per base step:	**~6°**	**~4°**

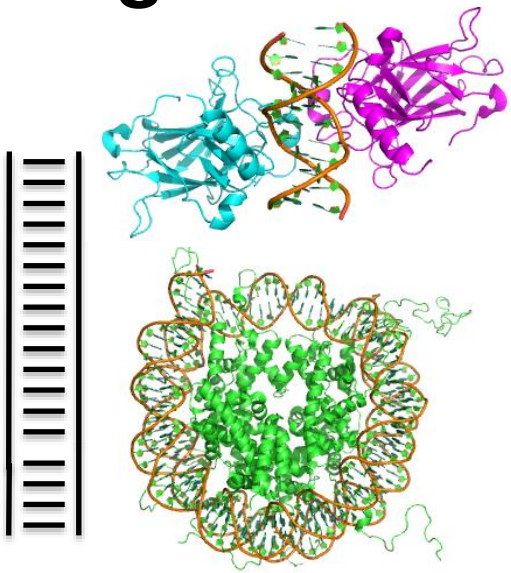
- Macroscopic elasticity does not reflect microscopic elasticity?
  (e.g. Macroscopic elasticity is dominated by infrequent DNA kinks)

- DNA/protein crystal structures do not reflect free DNA in solution?

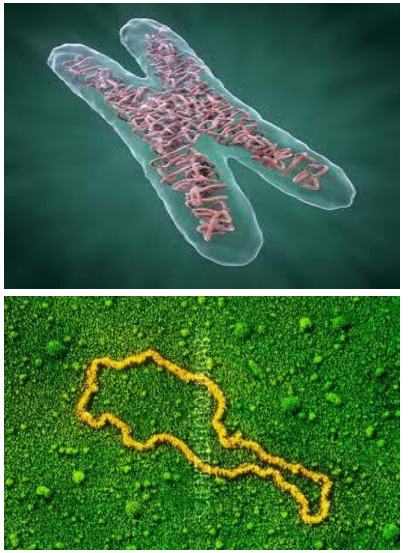
From microscopic to macroscopic length scale



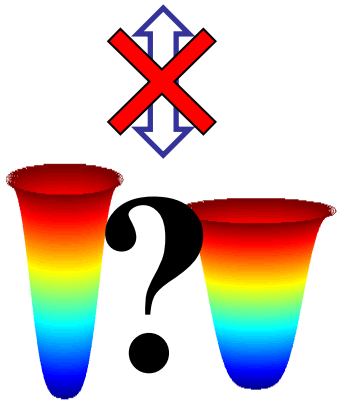
Base-step



10s of bps

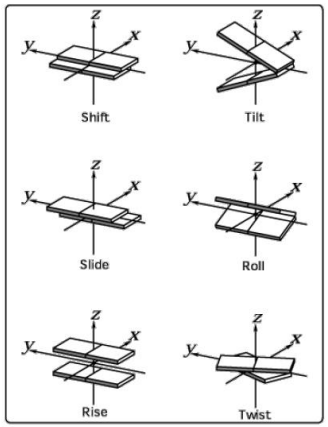


kbs, genome

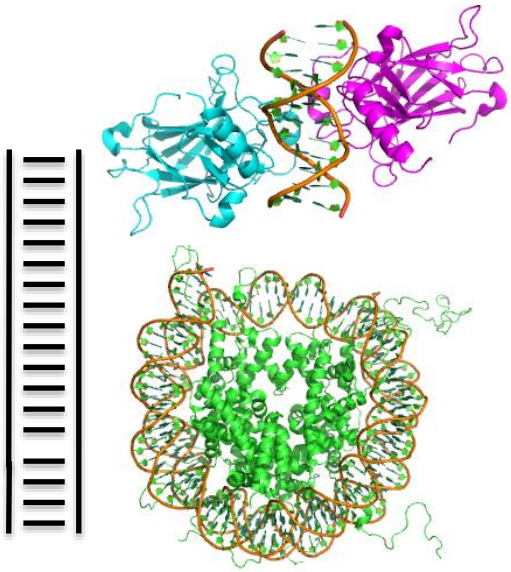


Need to more directly probe DNA elasticity on a microscopic length scale

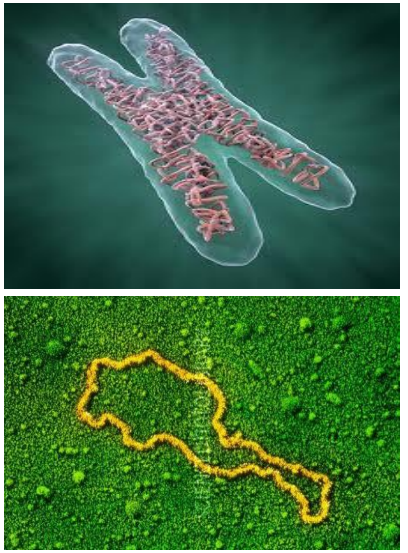
From microscopic to macroscopic behavior



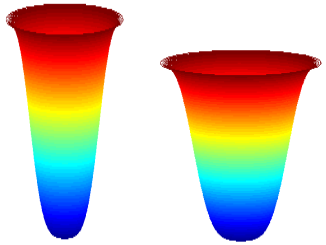
Base-step



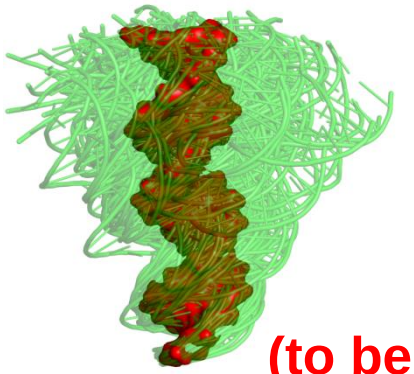
10s of bps



kbs, genome

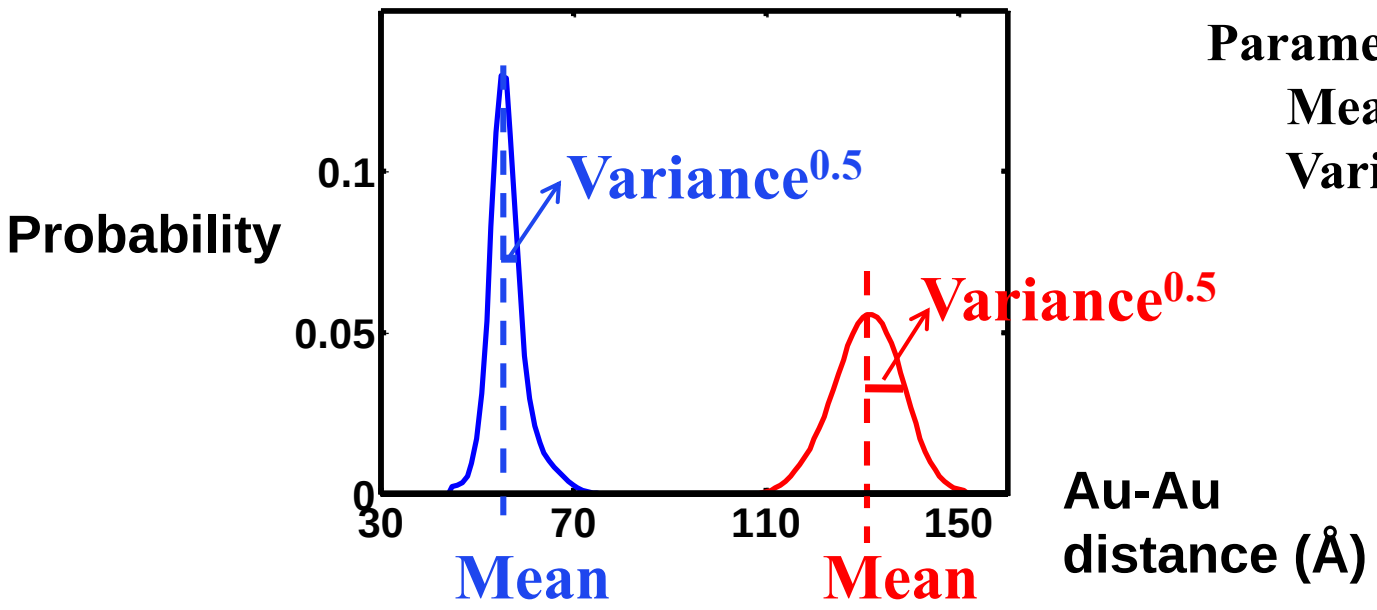


4º or 6º ?

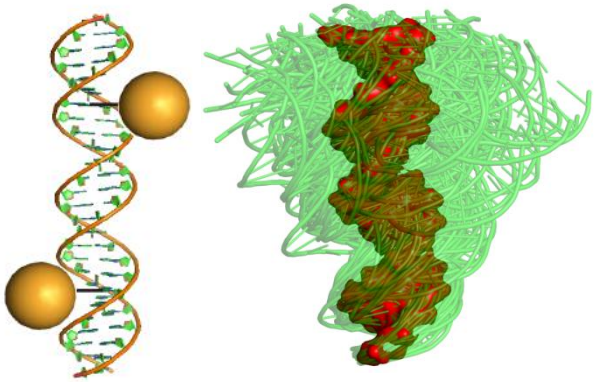


(to be determined by XSI)

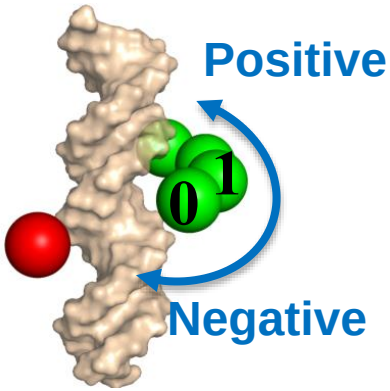
Basic information from an Au-Au distance distribution



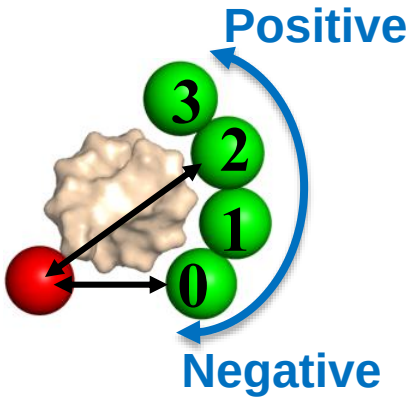
Parameters:
Mean
Variance (sd^2)



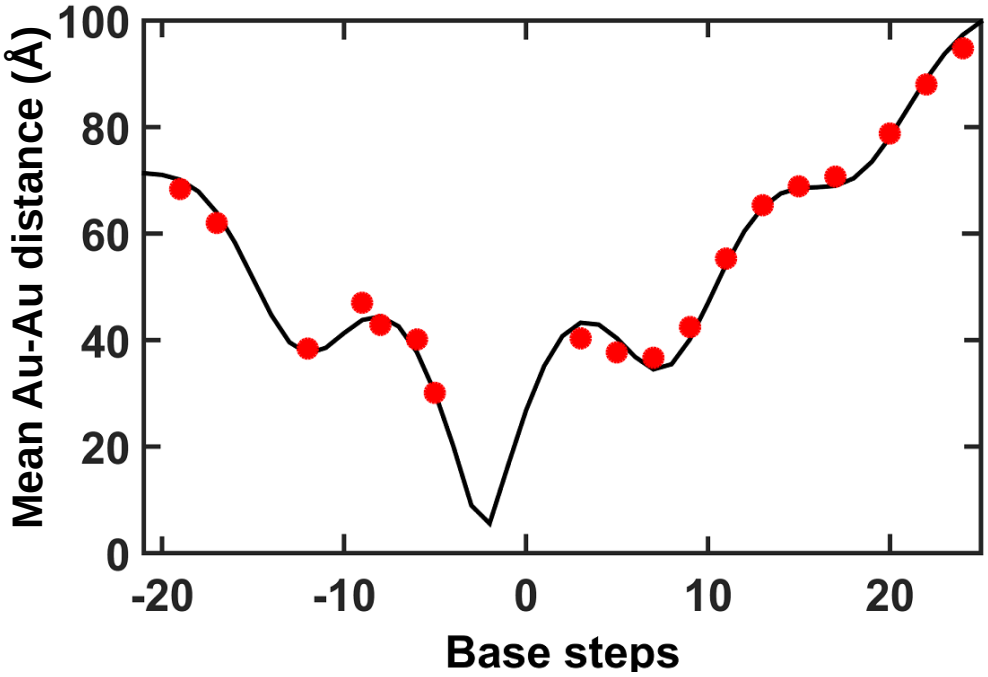
XSI determined B-DNA helical-structure



(side view)



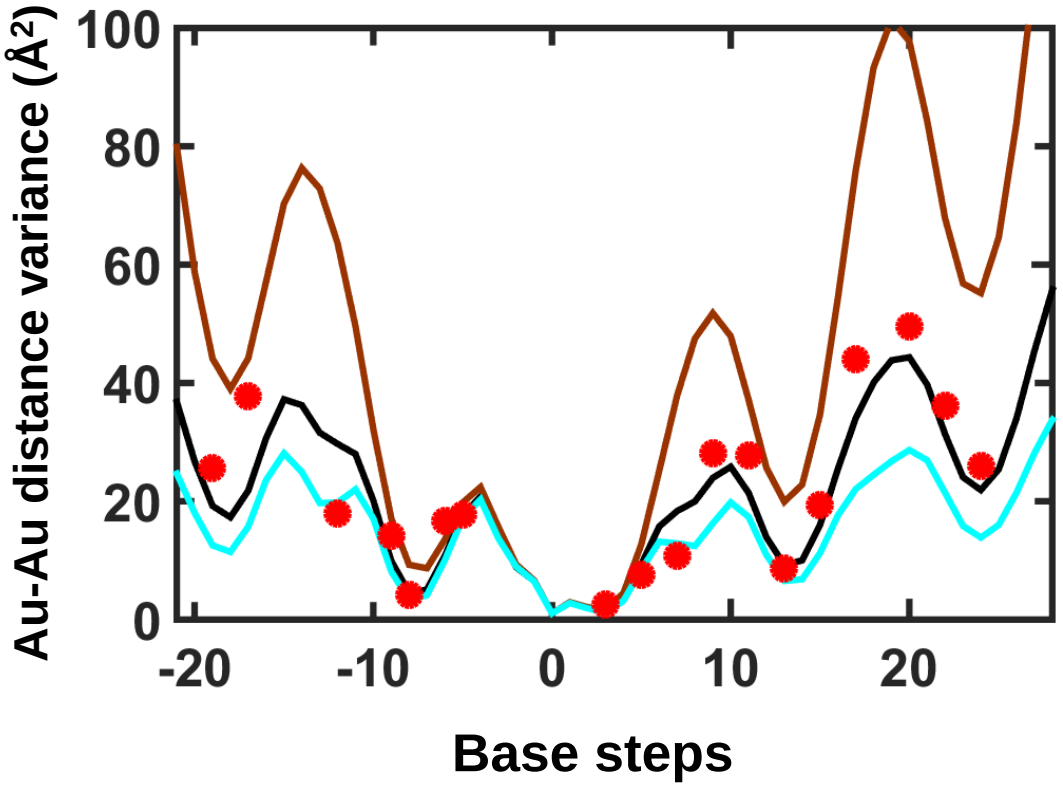
(top view)



3.36 Å in rise per base-step
10.5 bases per helical-turn

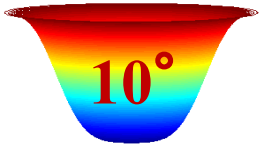
Shi et al. PNAS 2013

XSI determined B-DNA elasticity

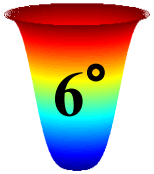


(bending fluctuation per base-step)

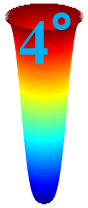
20 nm



55 nm



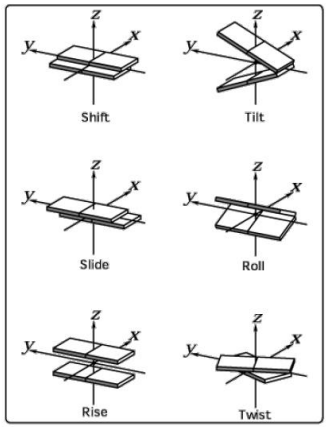
100 nm



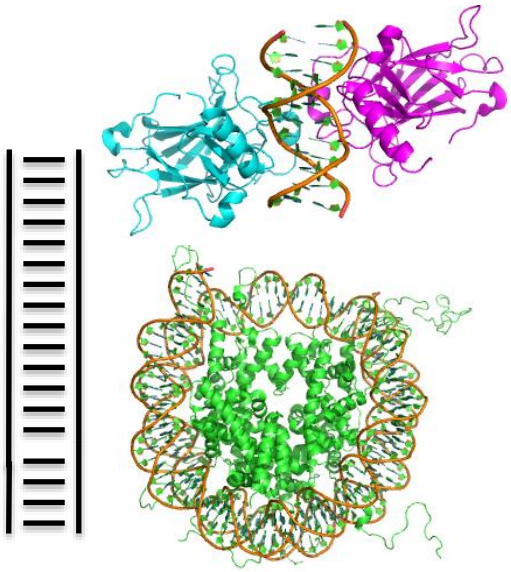
Bending persistence length: 55 ± 10 nm

Shi et al. PNAS 2013

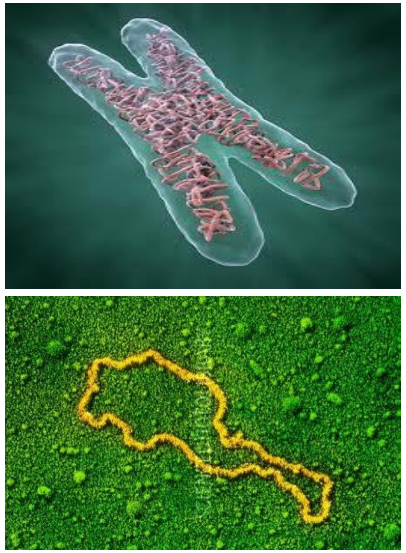
From microscopic to macroscopic behavior



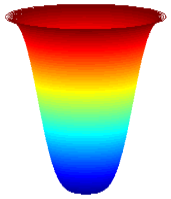
Base-step



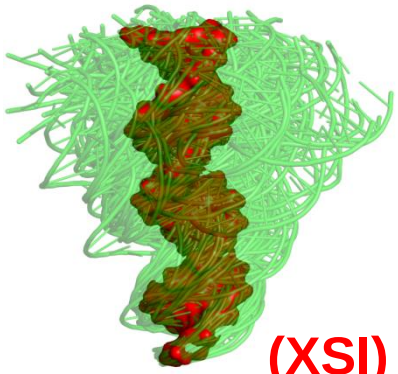
10s of bps



kbs, genome



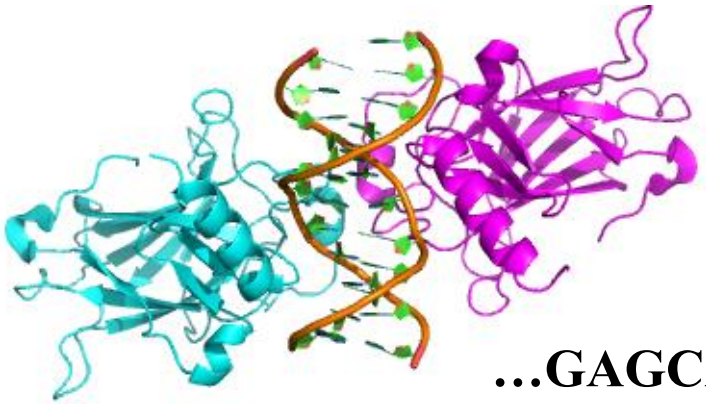
6º



(XSI)

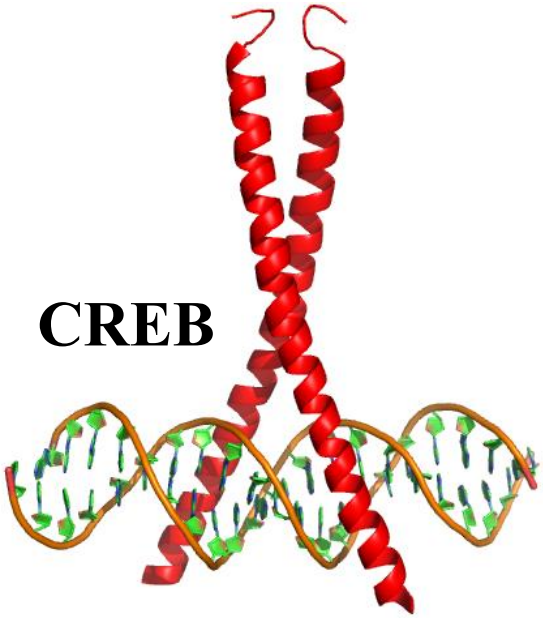
DNA SEQUENCE DEPENDENT VARIATION

DNA sequence dependence is important



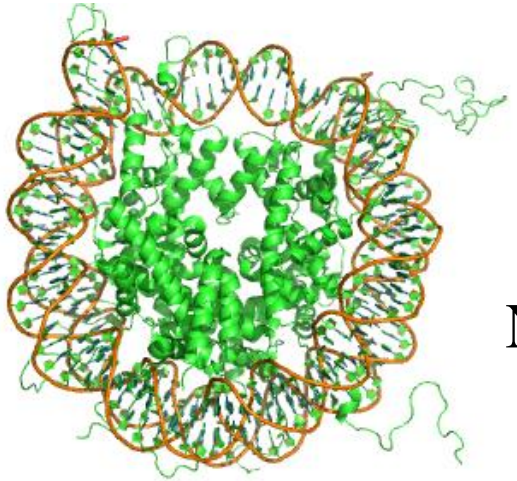
p53

…GAGCATGCTCA…



CREB

…ATGACGTCAT…

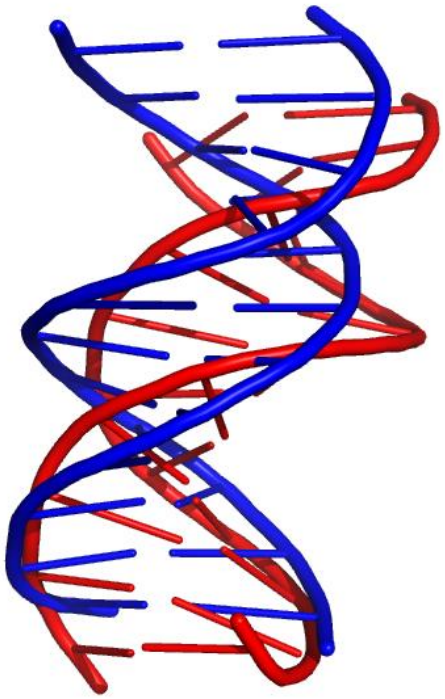


Nucleosome

DNA A-tract, e.g. NNNAAAAAANNN

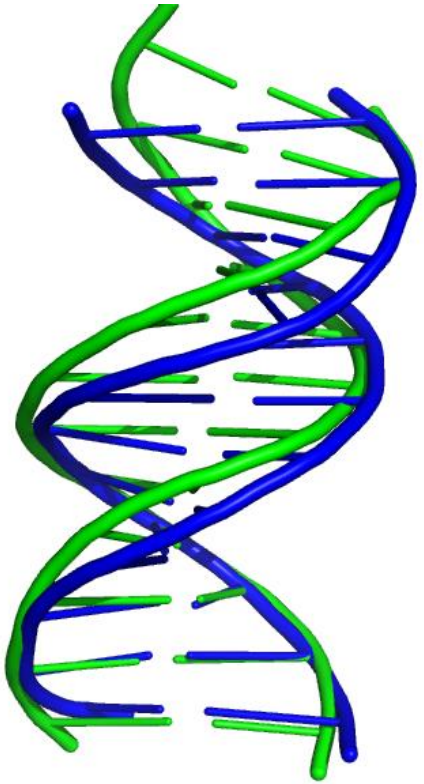
Standard B-DNA

NMR



Crystal structure I

NMR



Crystal structure II

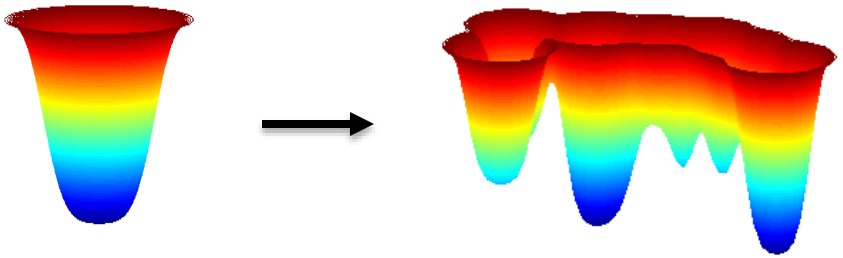
NMR

1FZX, *MacDonald et al., 2001 JMB;* 1D89 *DiGabriele and Steitz 1993 JMB*; 1D98 Nelson et al. 1987 Nature

DNA ENSEMBLES

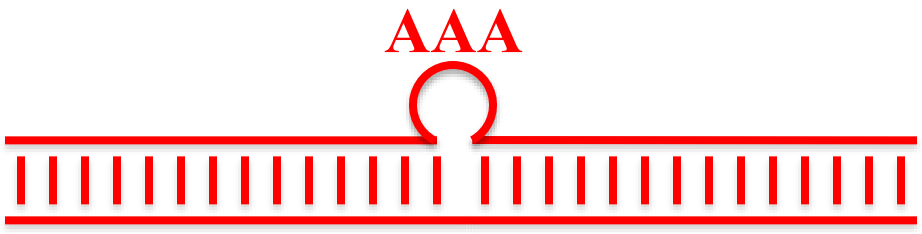
Shi et al. PNAS 2014

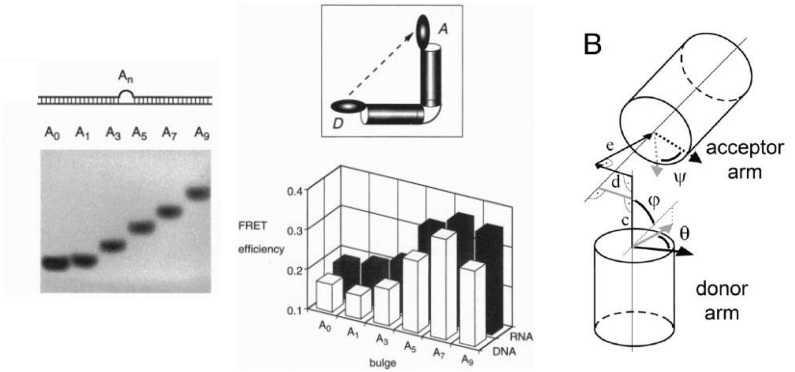
DNA ENSEMBLES



Shi et al. PNAS 2014

Model System:
DNA Duplex with a single-stranded Bulge

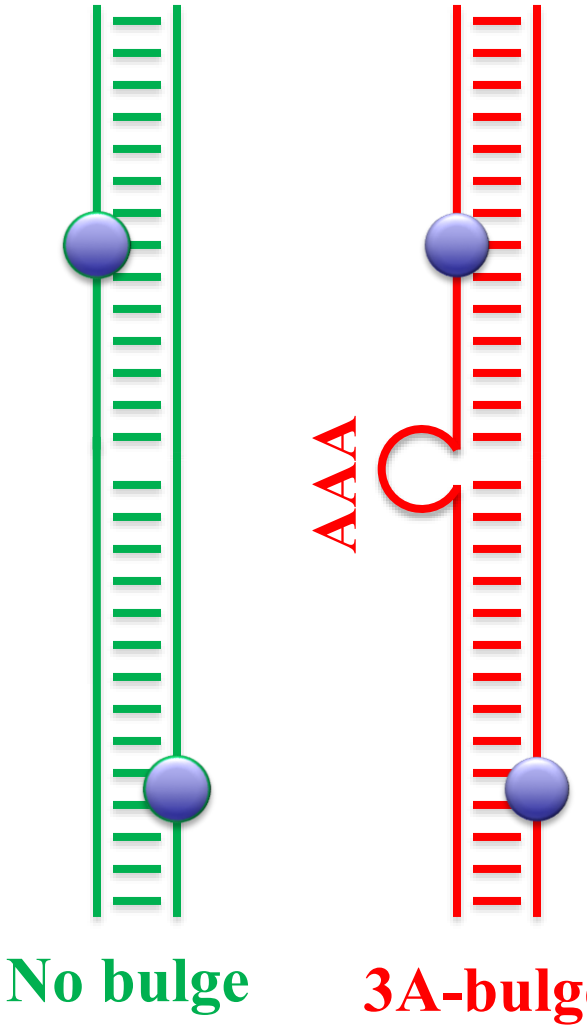


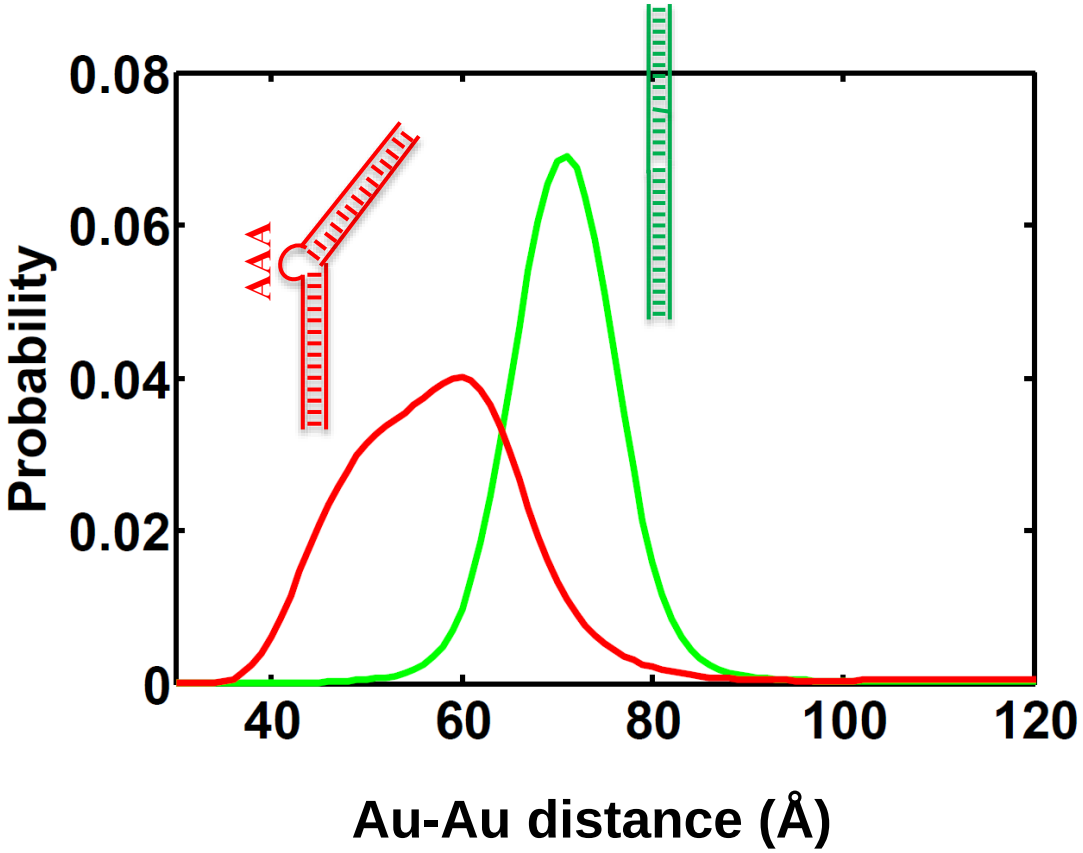


Gel FRET sm-FRET

Gohlke, *P.N.A.S.* (1994)
Lilley, *P.N.A.S.* (1995)
Wozniak, *P.N.A.S.* (2008)

A single Au-Au distribution

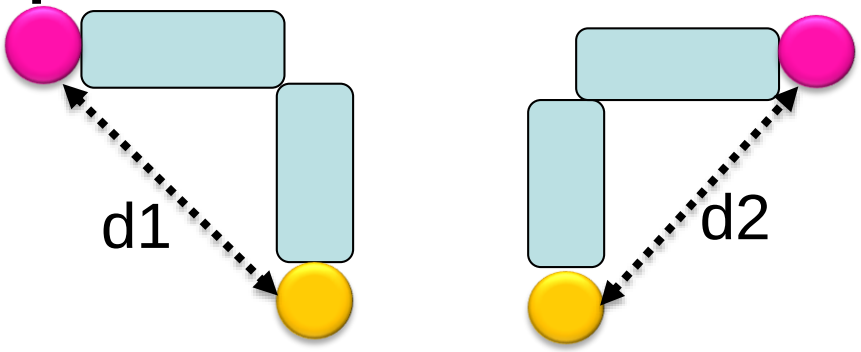




Shi et al. PNAS 2014

Degeneracy in a single distribution

Au pair 1



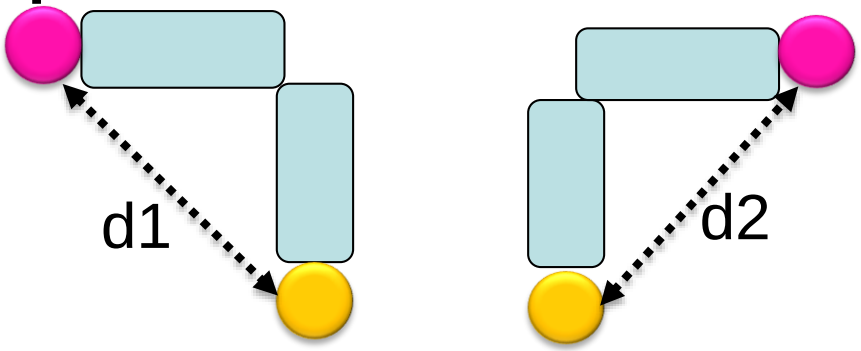
Conformation 1 Conformation 2

d1 = d2

Can not distinguish conformation 1 and 2

Reduce degeneracy with multiple Au pairs

Au pair 1

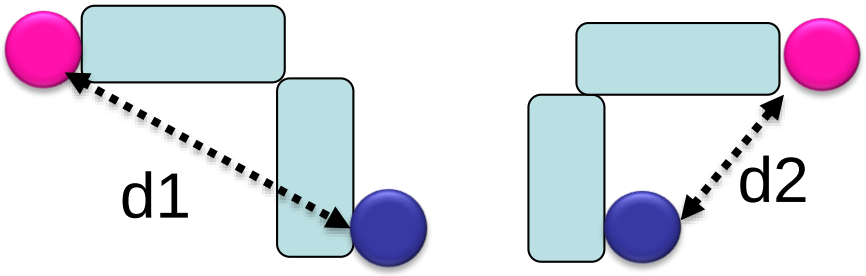


Conformation 1 Conformation 2

d1 = d2

Can not distinguish conformation 1 and 2

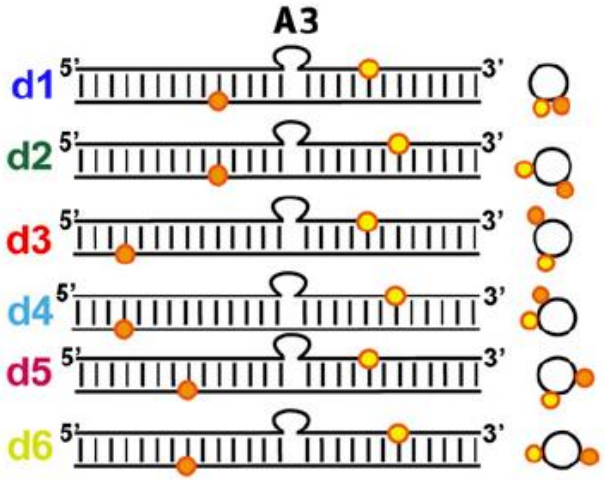
Au pair 2



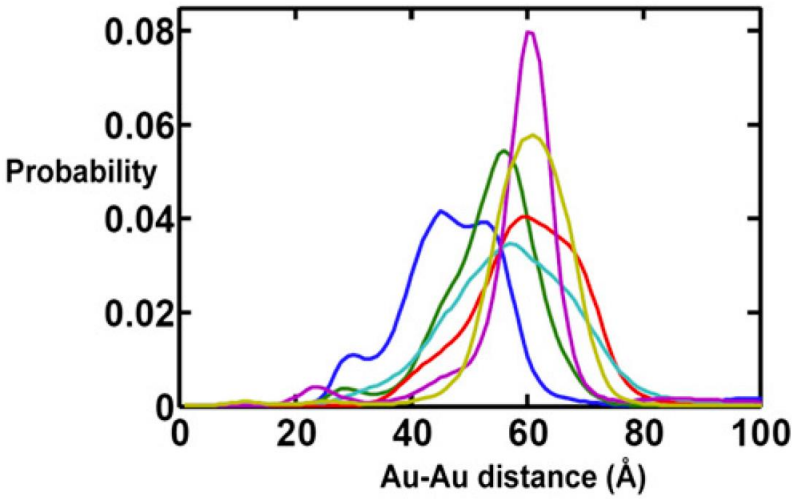
d1 ≠ d2

Can distinguish conformation 1 and 2

6 Au pairs

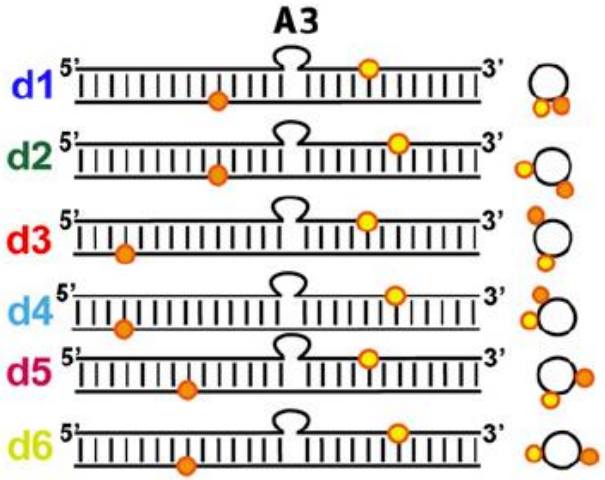


6 Au-Au distributions

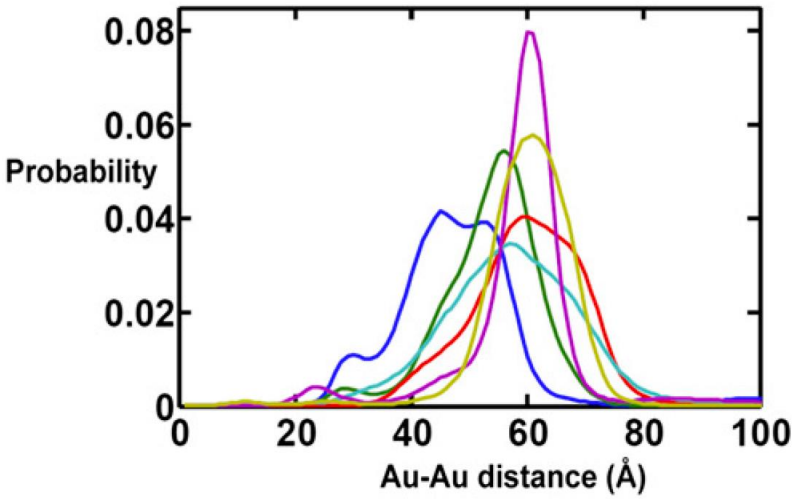


Shi et al. PNAS 2014

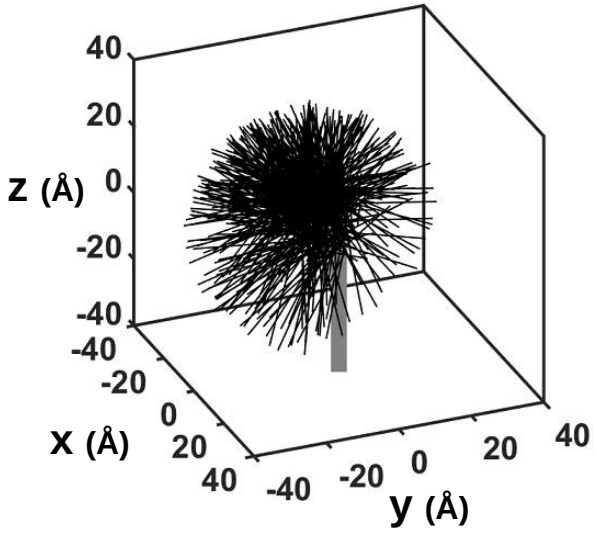
6 Au pairs



6 Au-Au distributions

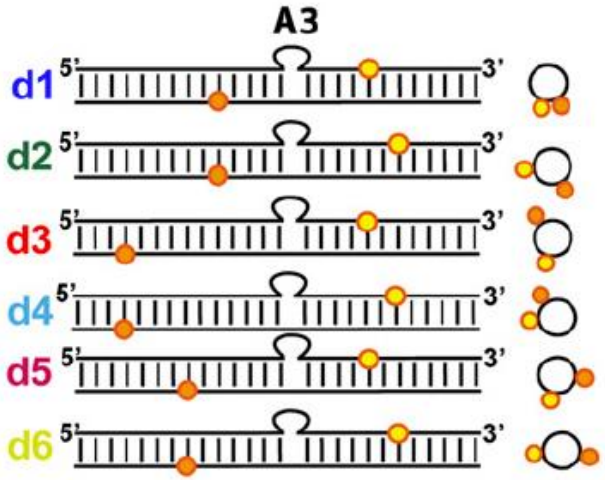


50000 basis-conformations (topologically allowed)

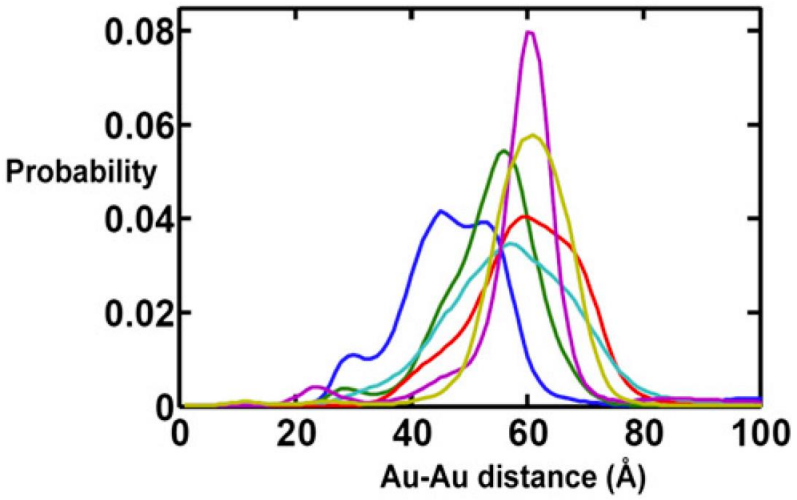


Shi et al. PNAS 2014

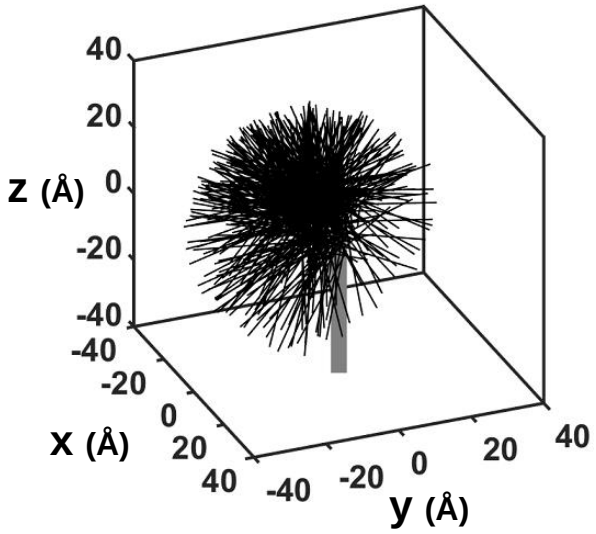
6 Au pairs



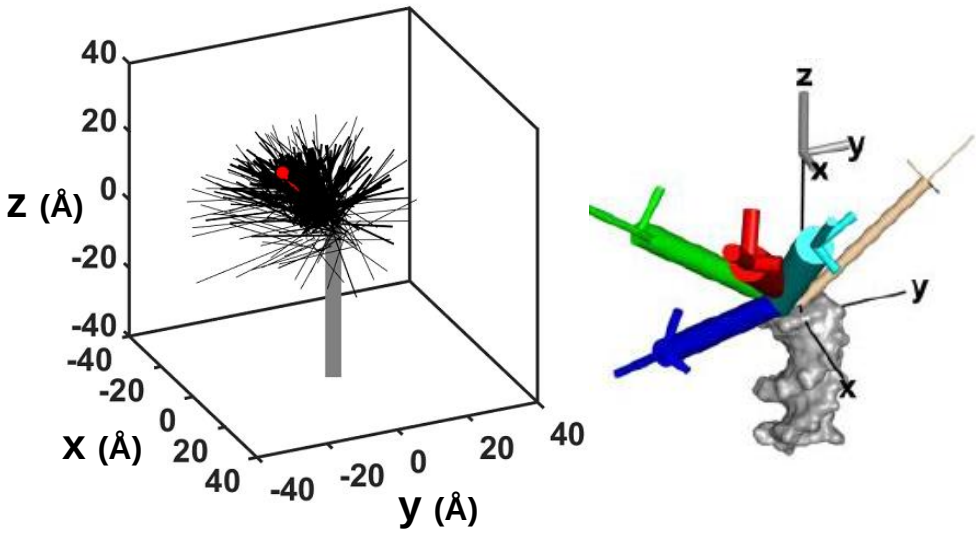
6 Au-Au distributions



50000 basis-conformations (topologically allowed)



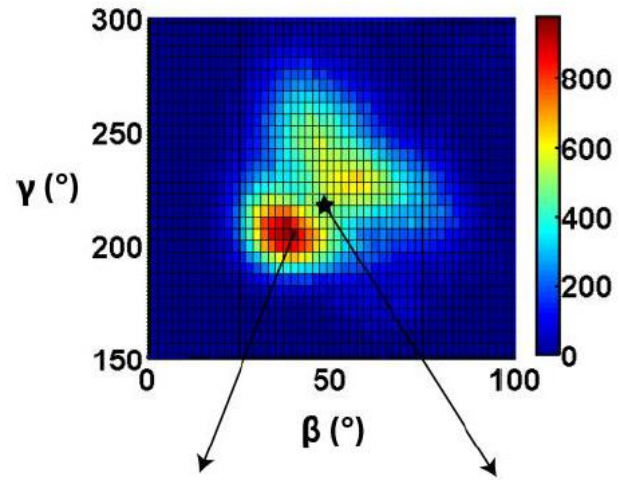
XSI determined ensemble



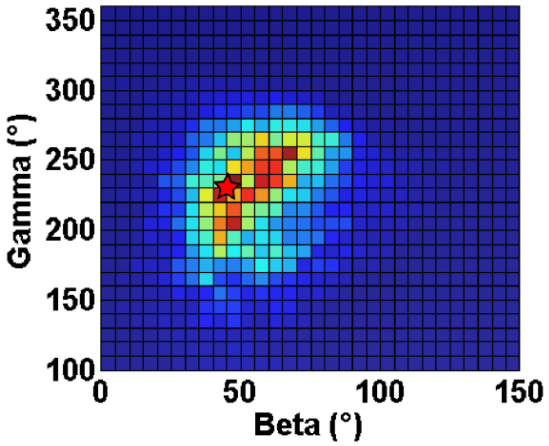
Shi et al. PNAS 2014

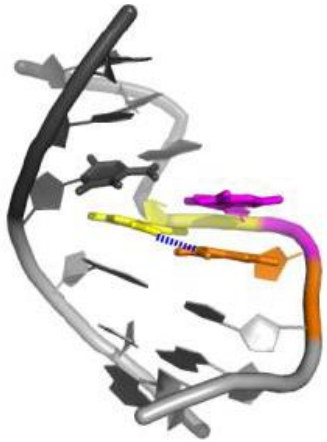
Molecular Dynamics Versus XSI Data

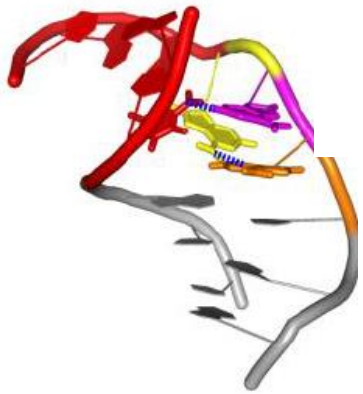
MD ensemble



XSI ensemble







MD optimum state

(stacking)

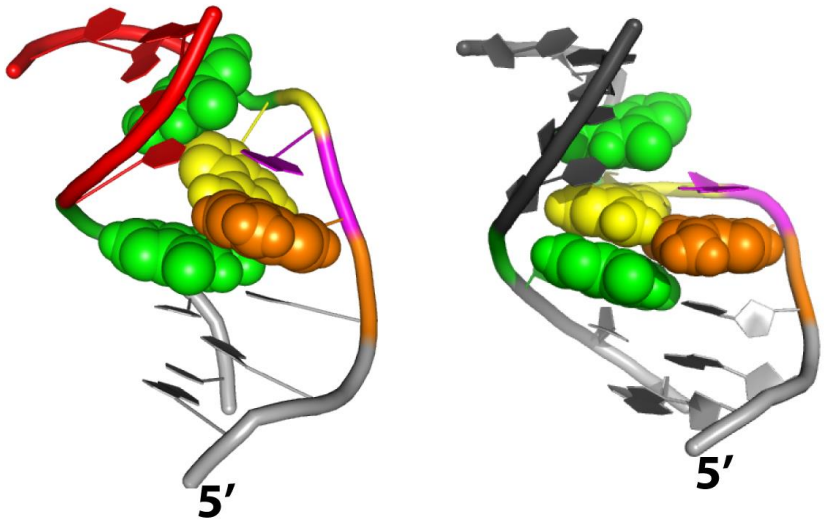
XSI optimum state (I)

(H-bond)

Shi et al. PNAS 2014

Test of MD vs XSI Model Predictions

(XSI) (MD)



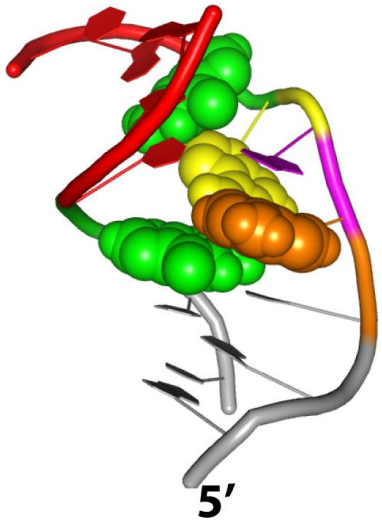
5′-GGCAAAGAT
3′-CCG…….CTA

To be tested by replacing a bulge A with a 2-AP, a fluorescence A-analog.

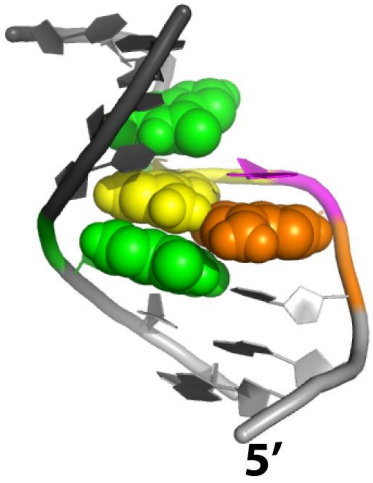
Shi et al. PNAS 2014

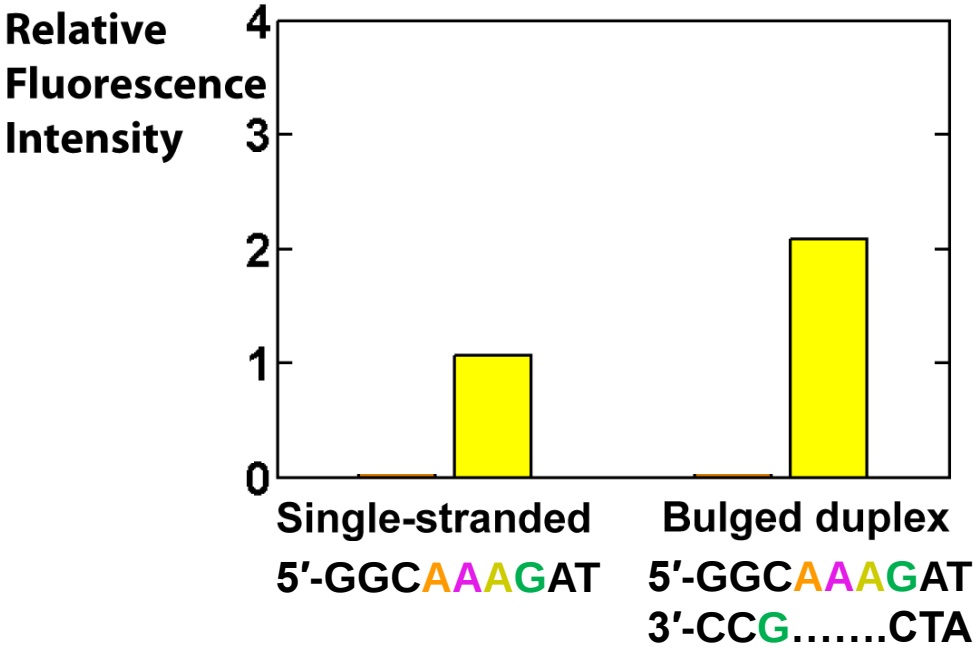
Test of MD vs XSI Model Predictions

(XSI)



(MD)



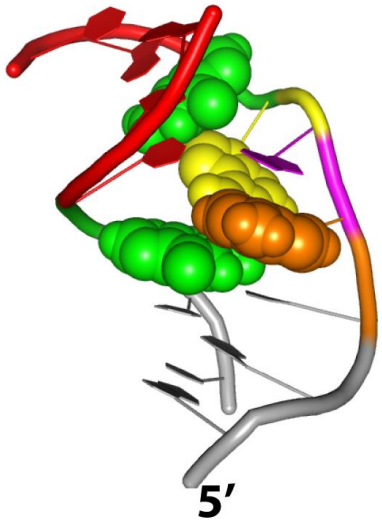


Yellow A is replaced by 2-AP

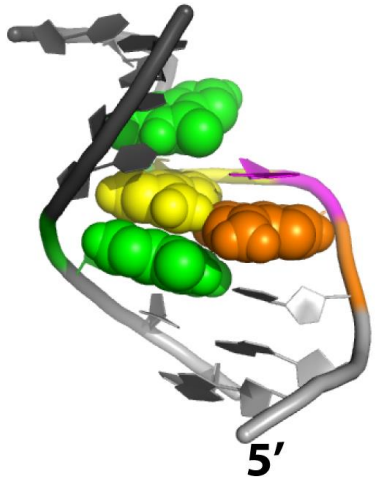
Shi et al. PNAS 2014

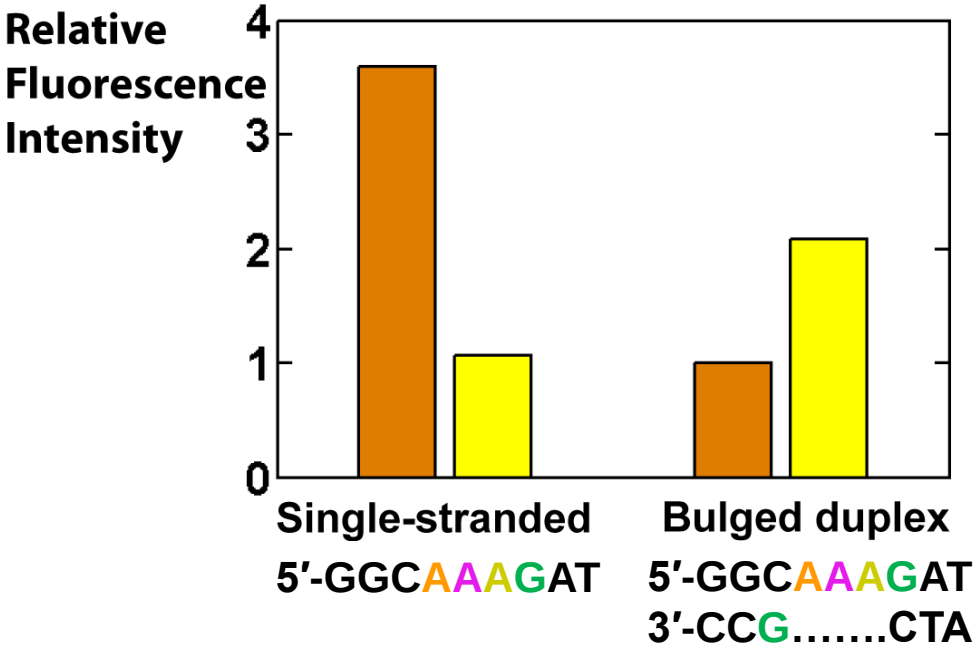
Test of MD vs XSI Model Predictions

(XSI)



(MD)



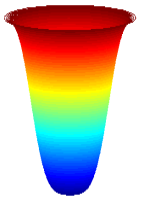


In a separate experiment,
the orange A is replaced by 2-AP

Shi et al. PNAS 2014

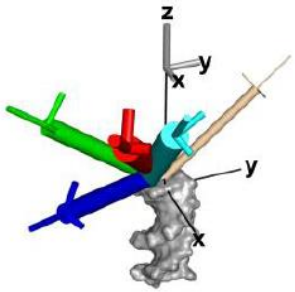
DNA

- **Elasticity and helical structure**

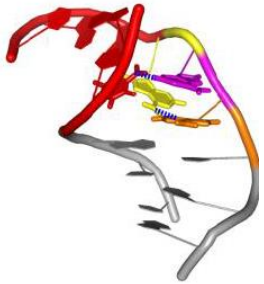


- **Sequence dependent variation**

- **Ensemble**



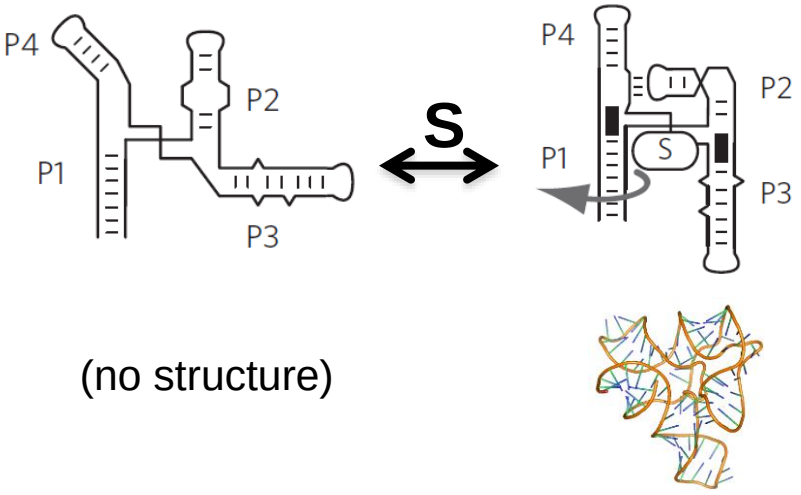
- **Atomic and physical models**



FUNCTIONAL RNAs

Functional RNAs are molecular machines that can undergo conformational changes

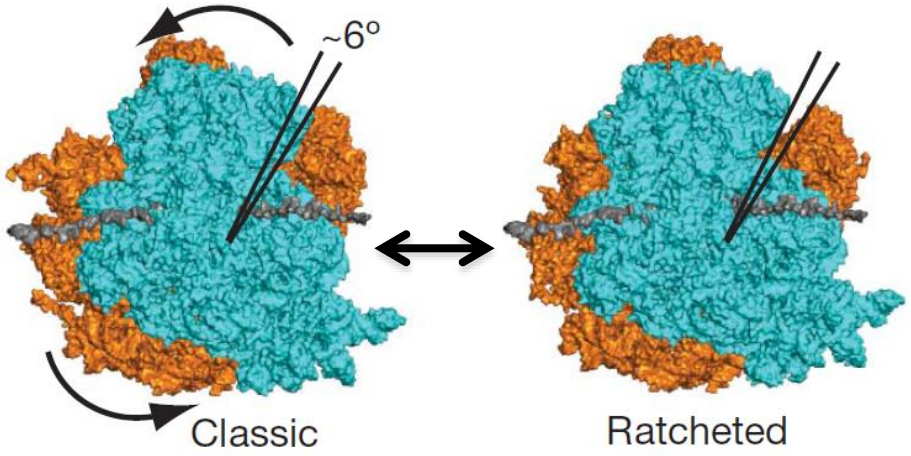
Riboswitch (SAM-I)



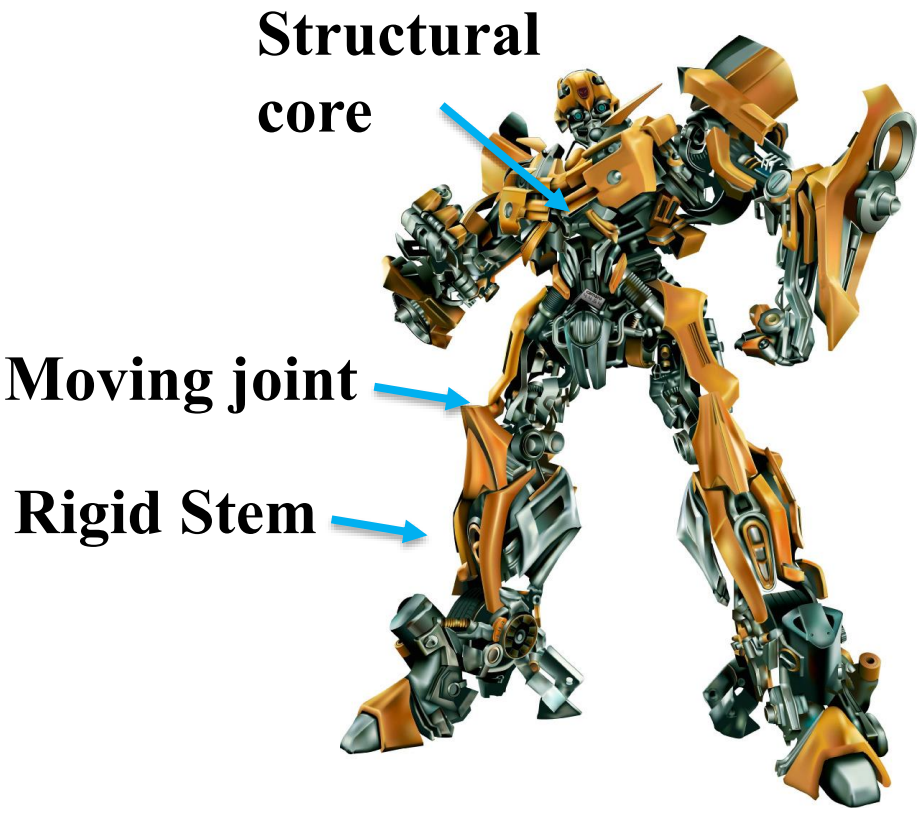
(no structure)

Functional RNAs are molecular machines that can undergo conformational changes

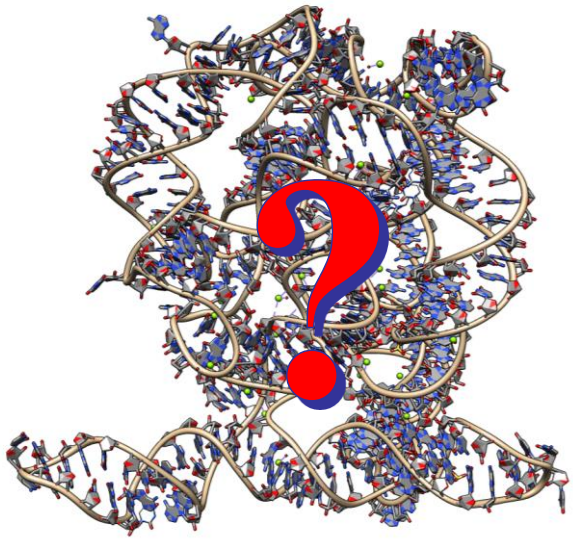
Ribosome

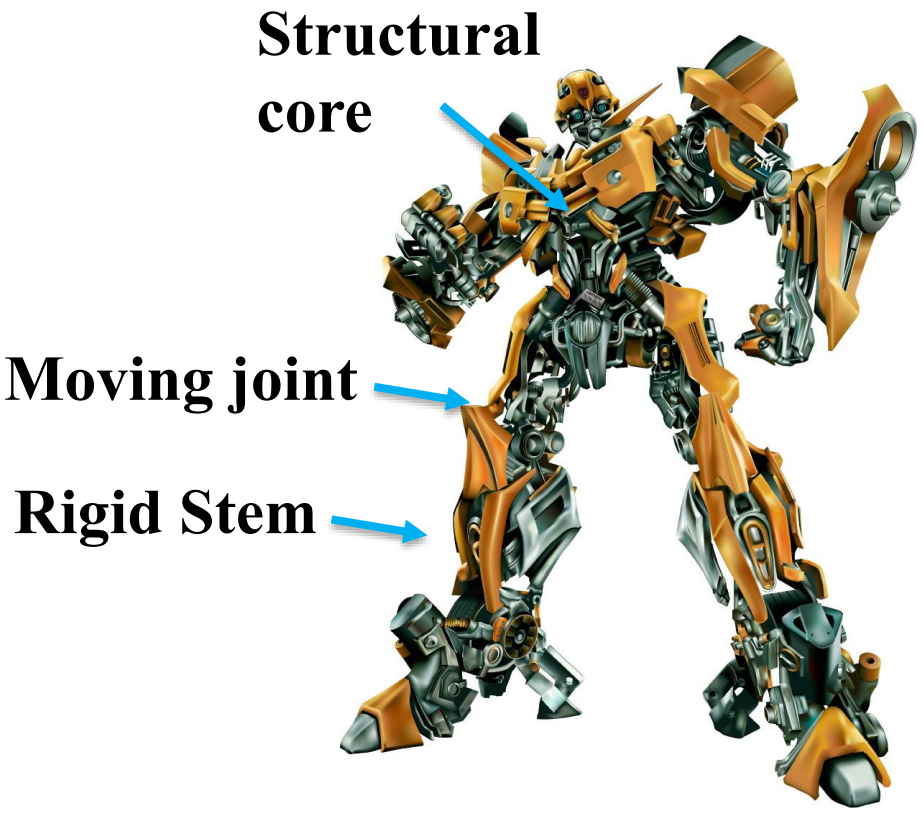


While it is easy for a robot,

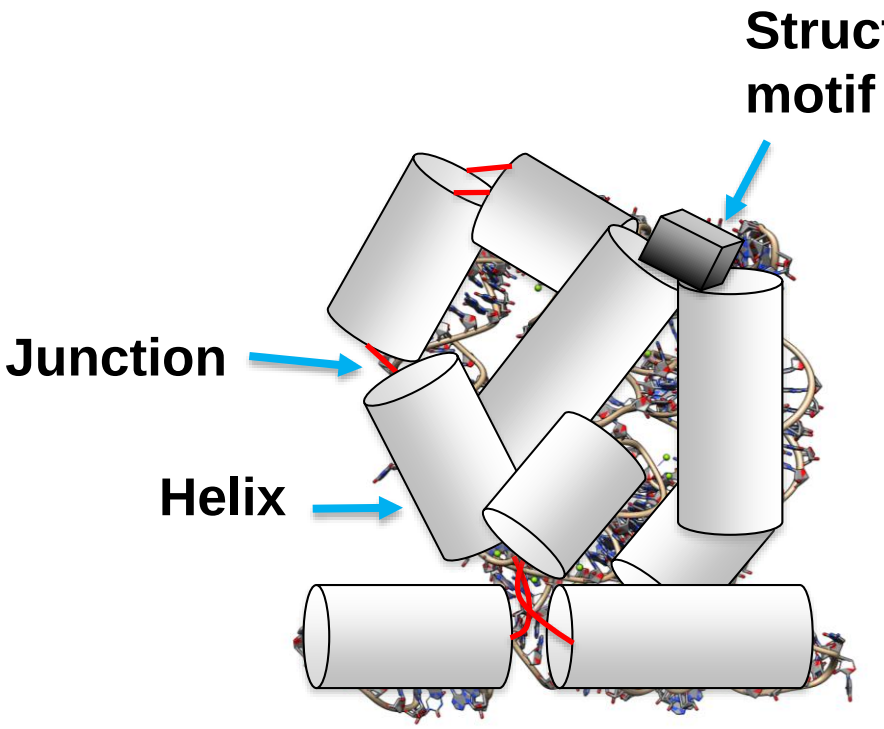


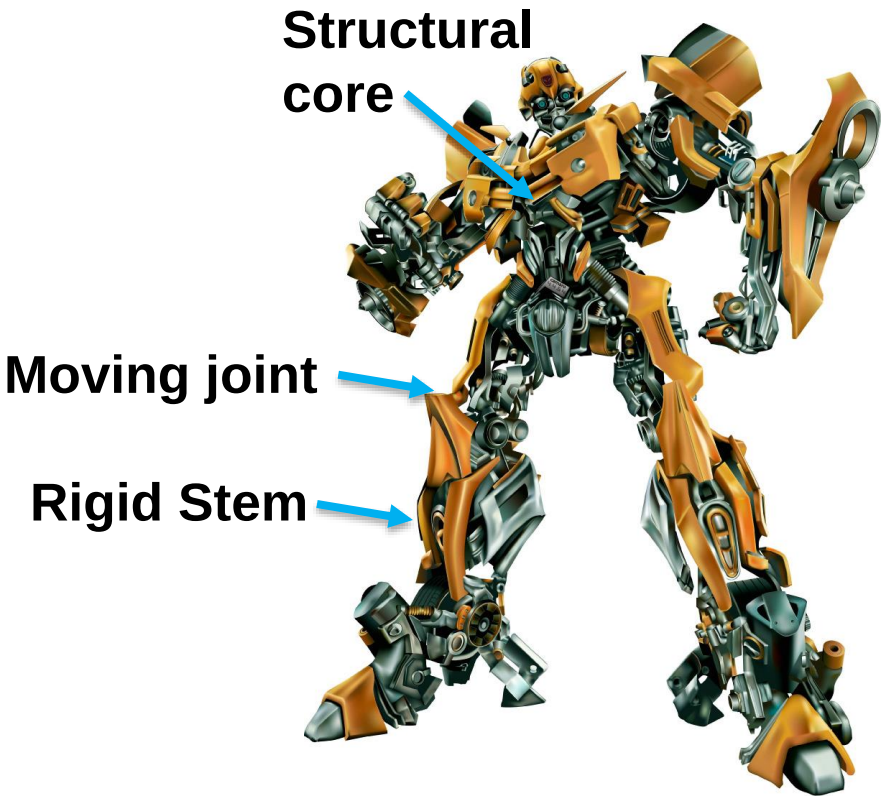
While it is easy for a robot, how do structured RNA and RNA•Protein complexes achieve these conformational changes?



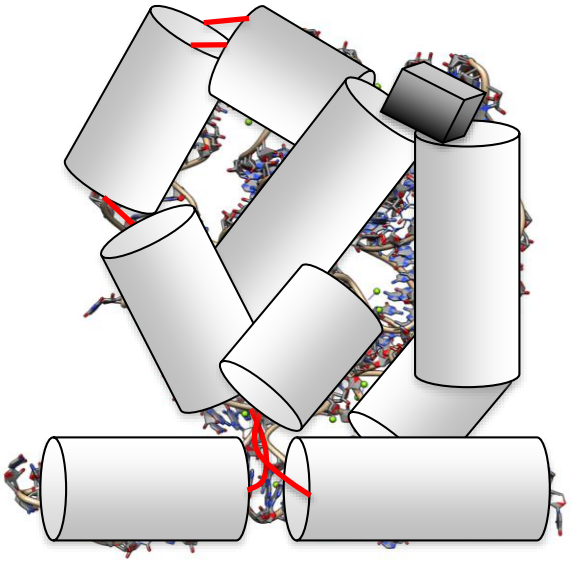


Complex RNAs consist of a similar sets of building blocks

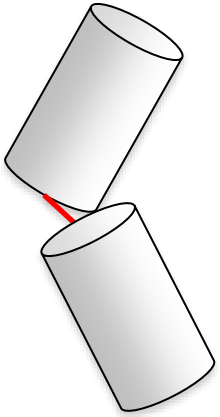




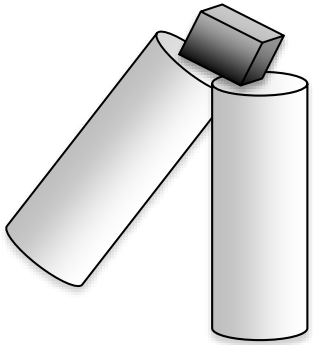
Complex RNAs consist of similar building blocks



Too complex



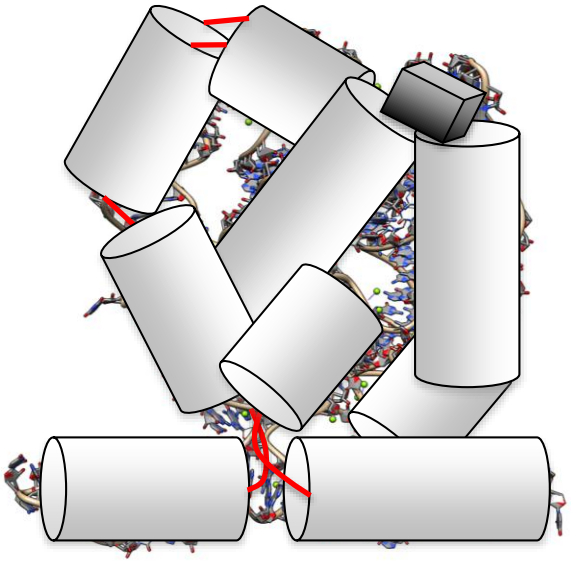
Junction



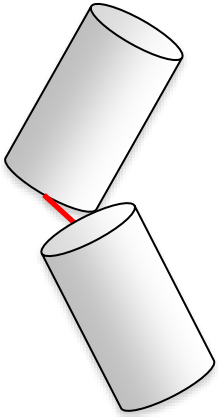
Structural motif

Start from individual elements

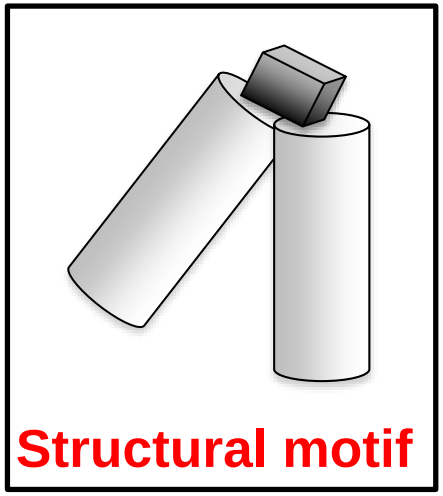
Complex RNAs consist of similar building blocks



Too complex



Junction

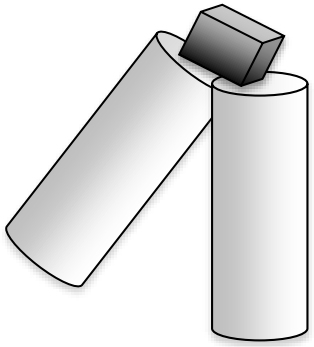


Structural motif

Start from individual elements

Fundamental questions

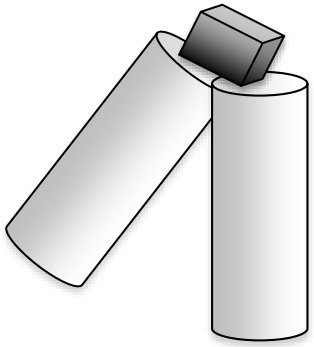
Structural motif



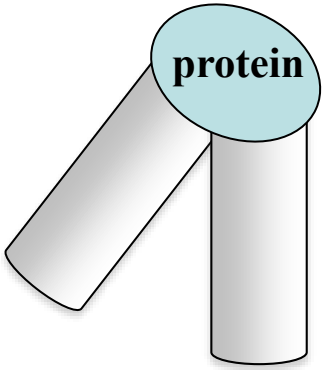
- **A single rigid conformation or a dynamic ensemble in solution?**

Fundamental questions

Structural motif



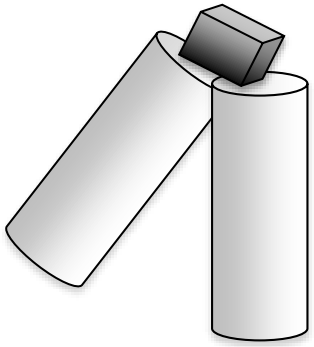
- **A single rigid conformation or a dynamic ensemble in solution?**



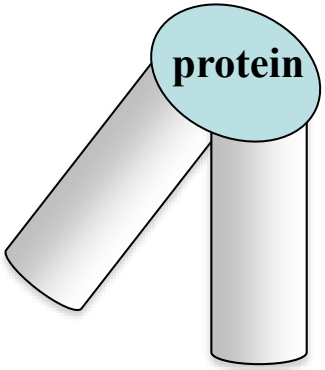
- **Does binding of a structure protein limit it to a single conformation?**

Fundamental questions

Structural motif



- **A single rigid conformation or a dynamic ensemble in solution?**



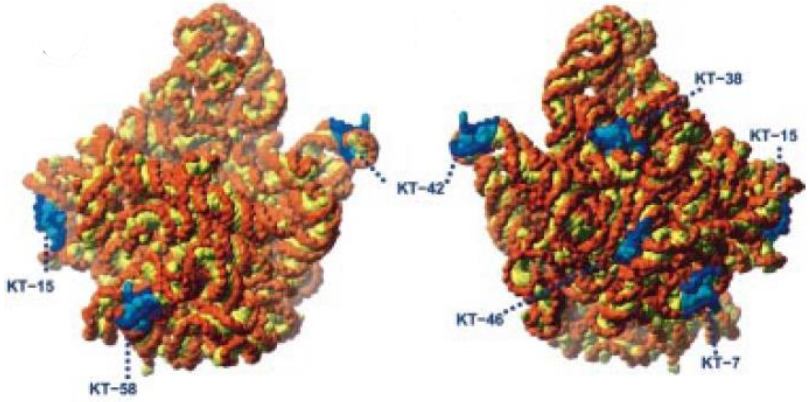
- **Does binding of a structure protein limit it to a single conformation?**

- **crystal structure vs solution structure**

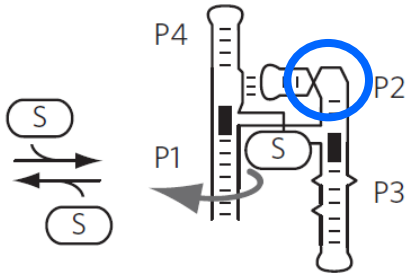
RNA kink-turn motif (KT)

- **Kink-turn is widespread and in all major classes of structured RNA**

e.g. Ribosome

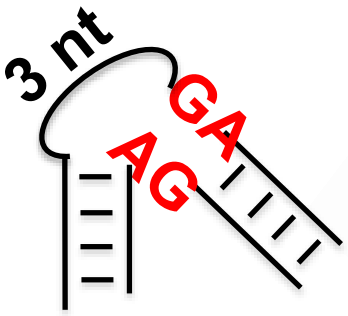


e.g. SAM-I Riboswitch

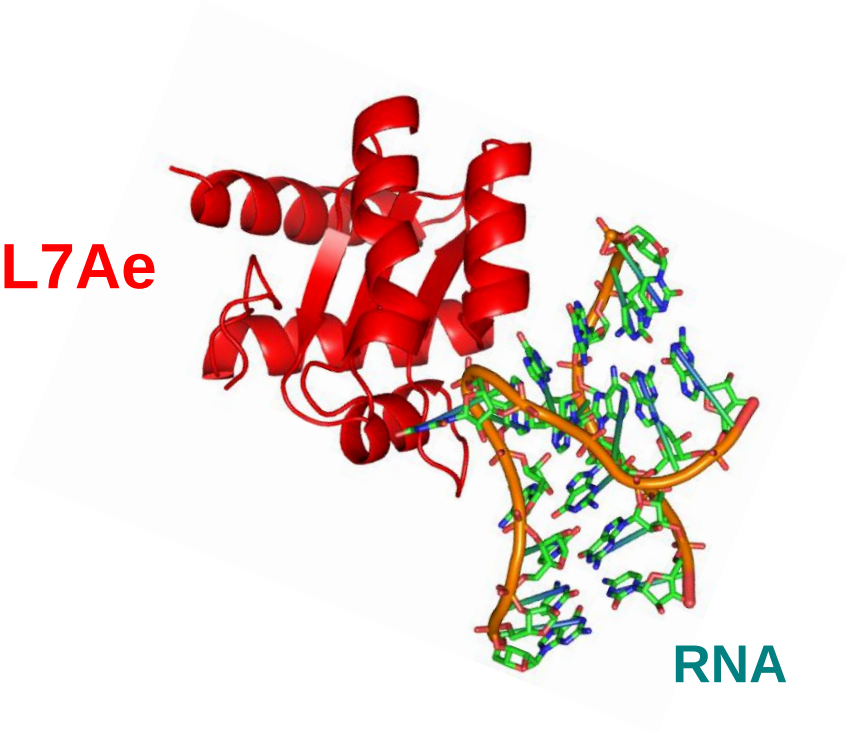


Kink-turns in blue

- **Consensus kink-turn**



RNA•Protein motif: KT•L7Ae

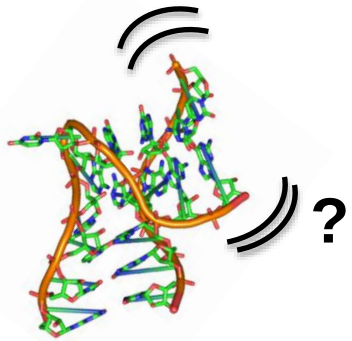


KT•L7Ae is a reoccurring RNA•Protein motif:

- **Ribosome**
- **Box C/D s(no)RNP**
- **Spliceosome**
- **RNase P**

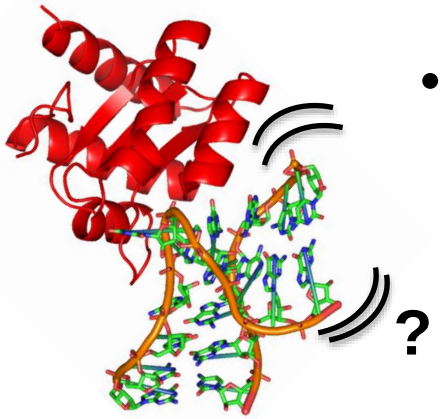
Questions

The kink-turn motif



RNA alone

- **Does folded**
  KT
  KT•L7Ae
  exist as a single conformation
  or as a dynamic ensemble?

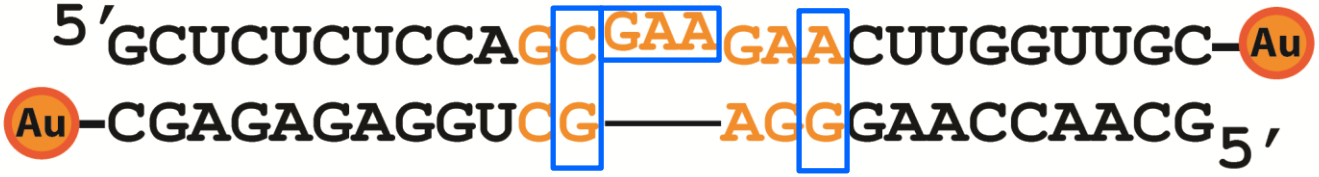


RNA•Protein complex

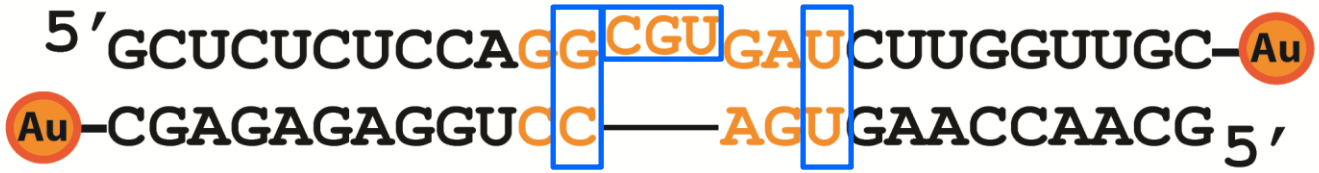
- **To what extent are RNA solution structures represented by crystal structures?**

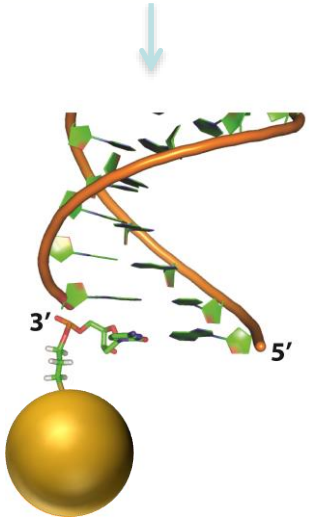
Kink-turns investigated

KTA (Kt7)



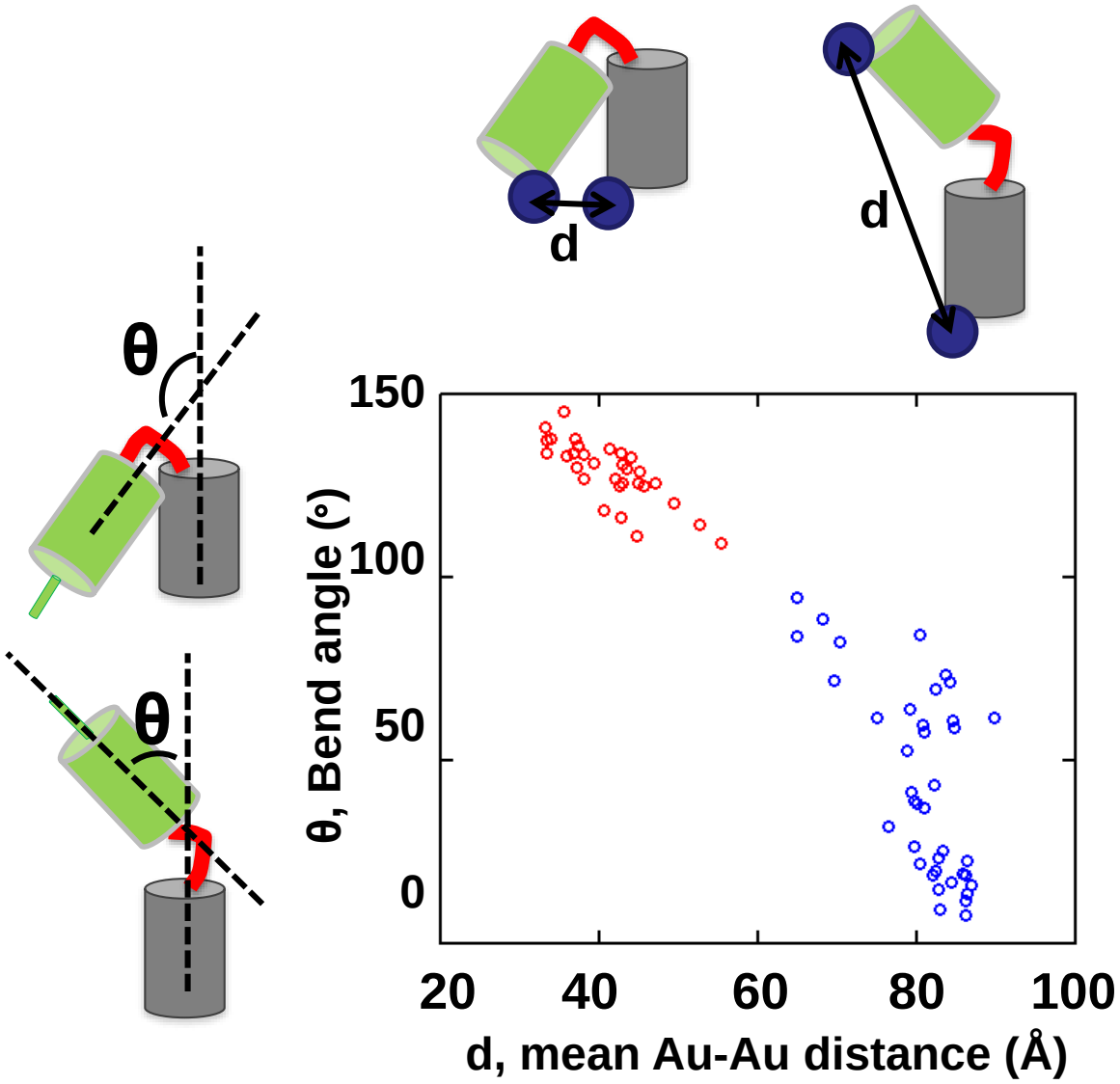
KtB (Box C/D)





Shi et al. Nat. Chem. Biol., 2016

Basis-set conformations for RNA 0-3 junction obtained from crystal structure database



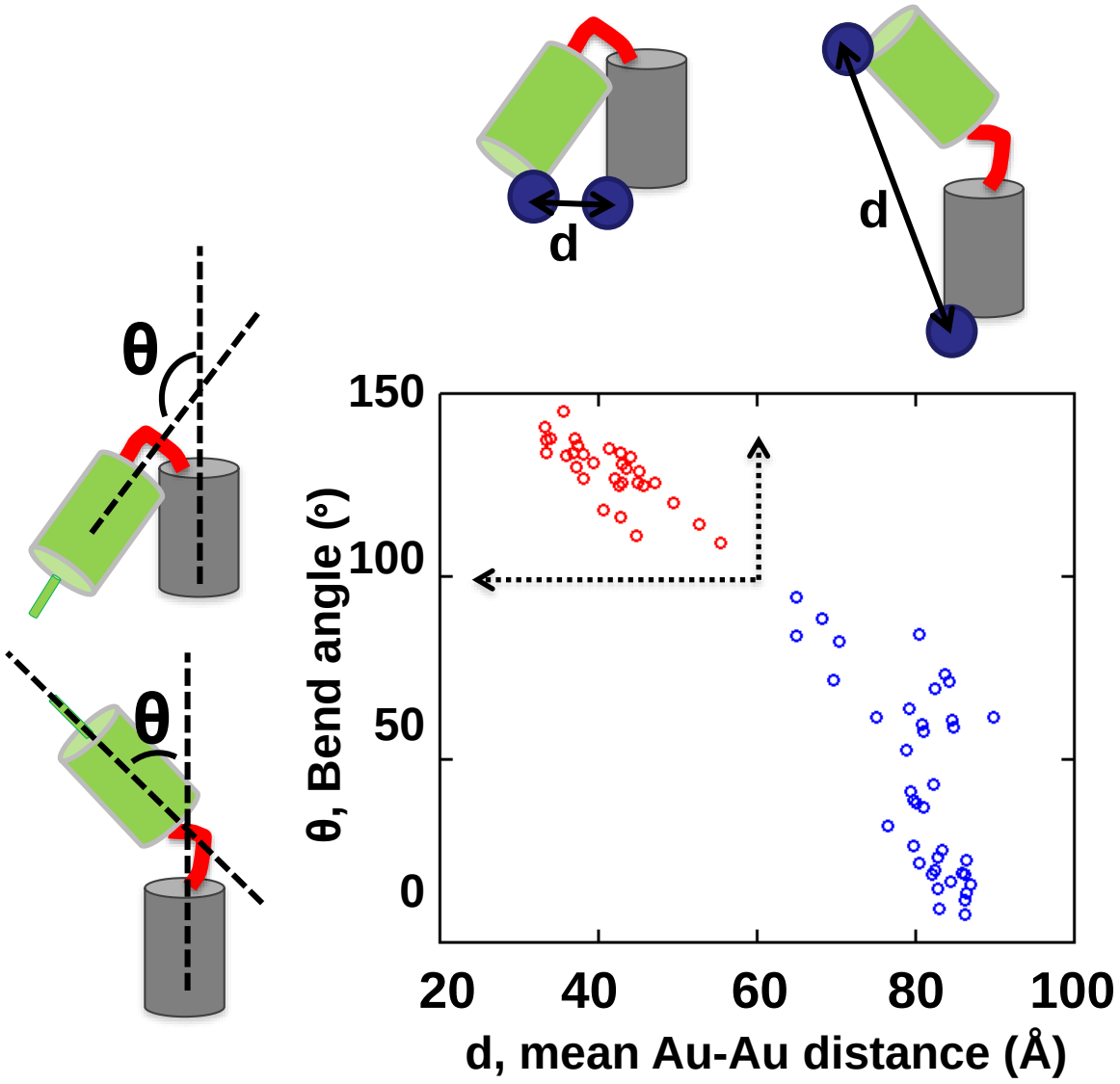
3nt-bulge:

5′-…NN NNN NNN…

3′-…NN —— NNN…

Shi et al. Nat. Chem. Biol., 2016

Basis-set conformations for RNA 0-3 junction obtained from crystal structure database



Kink-turn:

5′-…NN^{NNN}**GA**N…

3′-…NN ——**AN**N…

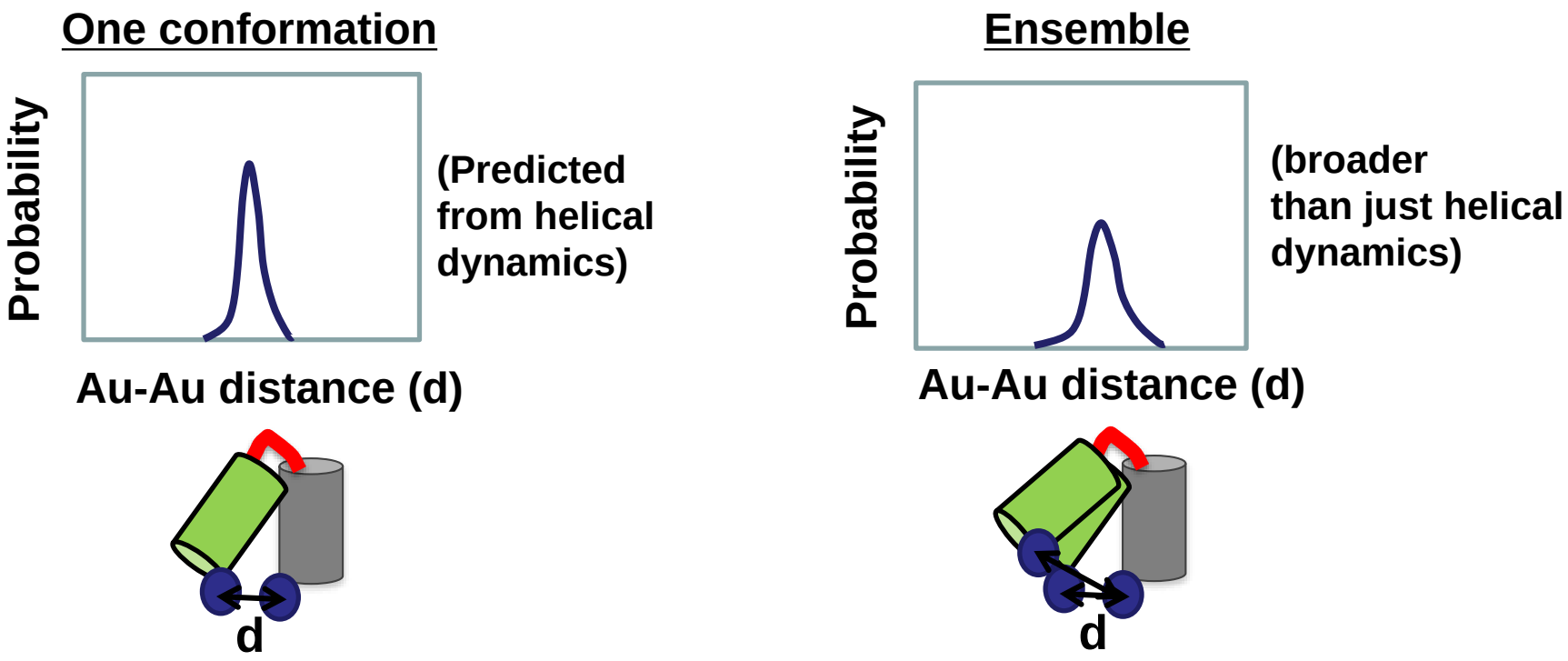
3nt-bulge:

5′-…NN^{NNN}**NN**N…

3′-…NN—— **NN**N…

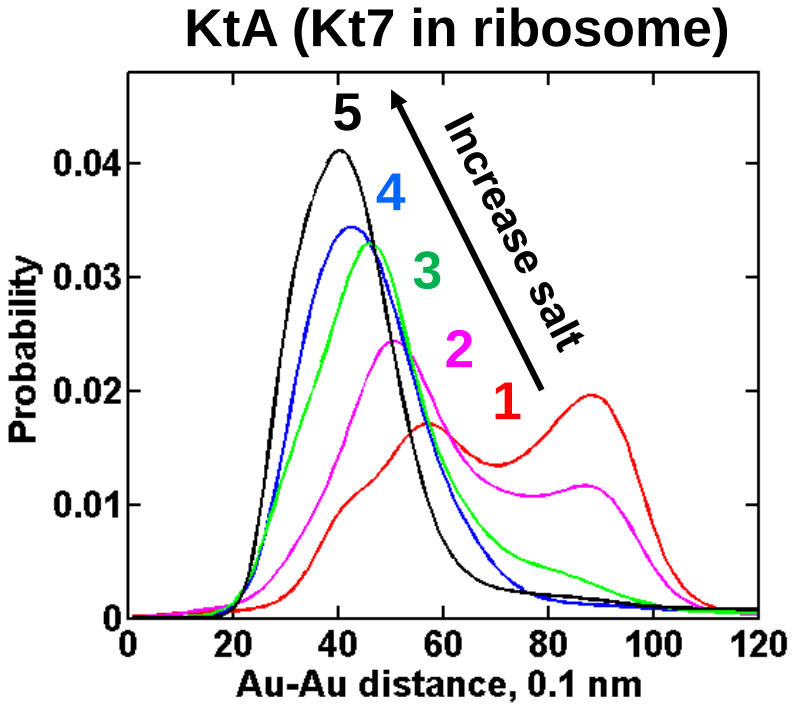
Shi et al. Nat. Chem. Biol., 2016

How to detect an ensemble

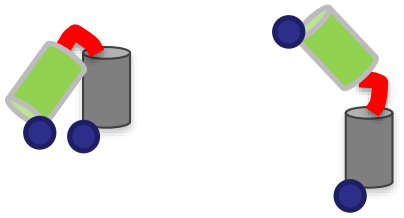


An ensemble broadens the Au-Au distance distribution (e.g. lower peak height with normalized area)

Au-Au distance distributions for KtA



Salt conditions	Mg^{2+}	Na^{+}	K^{+}	Tris	(mM)
1.	0	20	0	30	
2.	0	60	0	70	
3.	1	10	150	70	
4.	0	510	0	70	
5.	10	10	0	70	

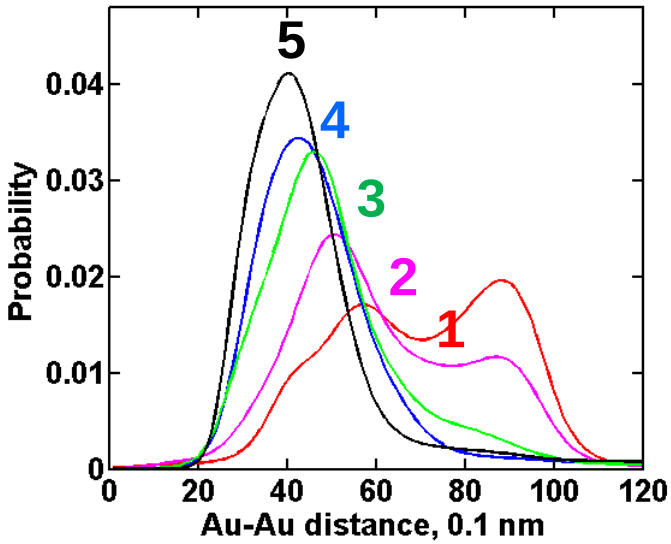


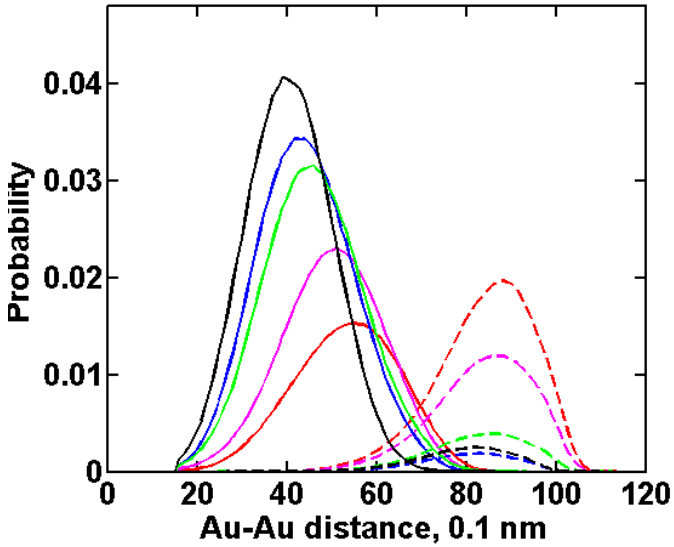
(Kinked) (unkinked)

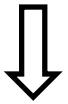
KtA exists as an equilibrium of kinked and unkinked conformations

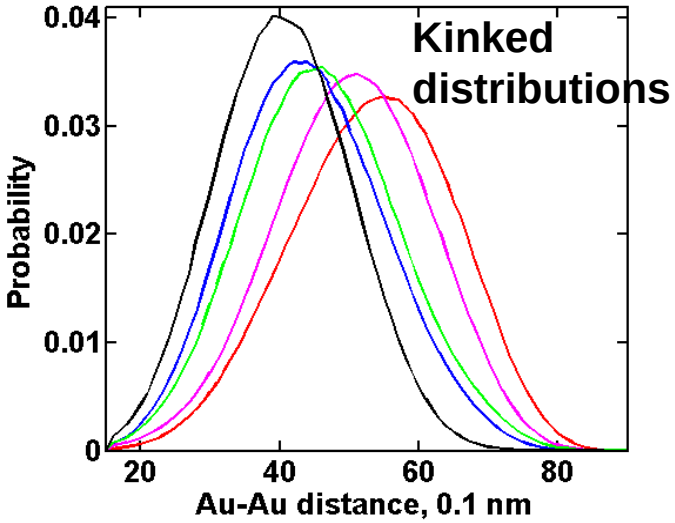
Shi et al. Nat. Chem. Biol., 2016

Isolate the kinked-state

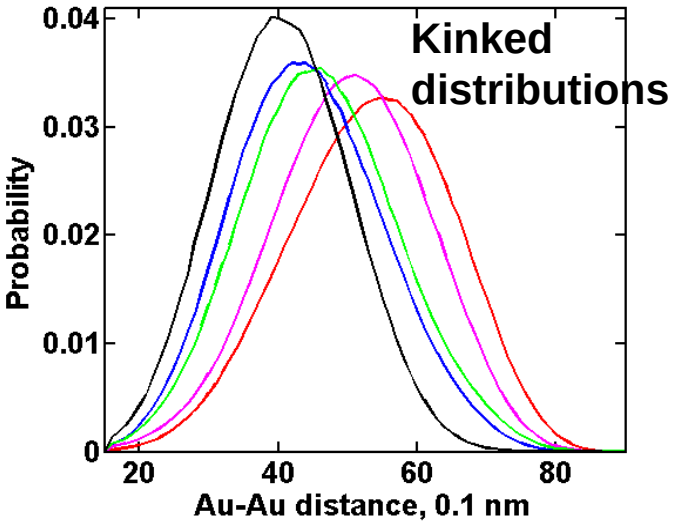








Is the kinked state an ensemble?

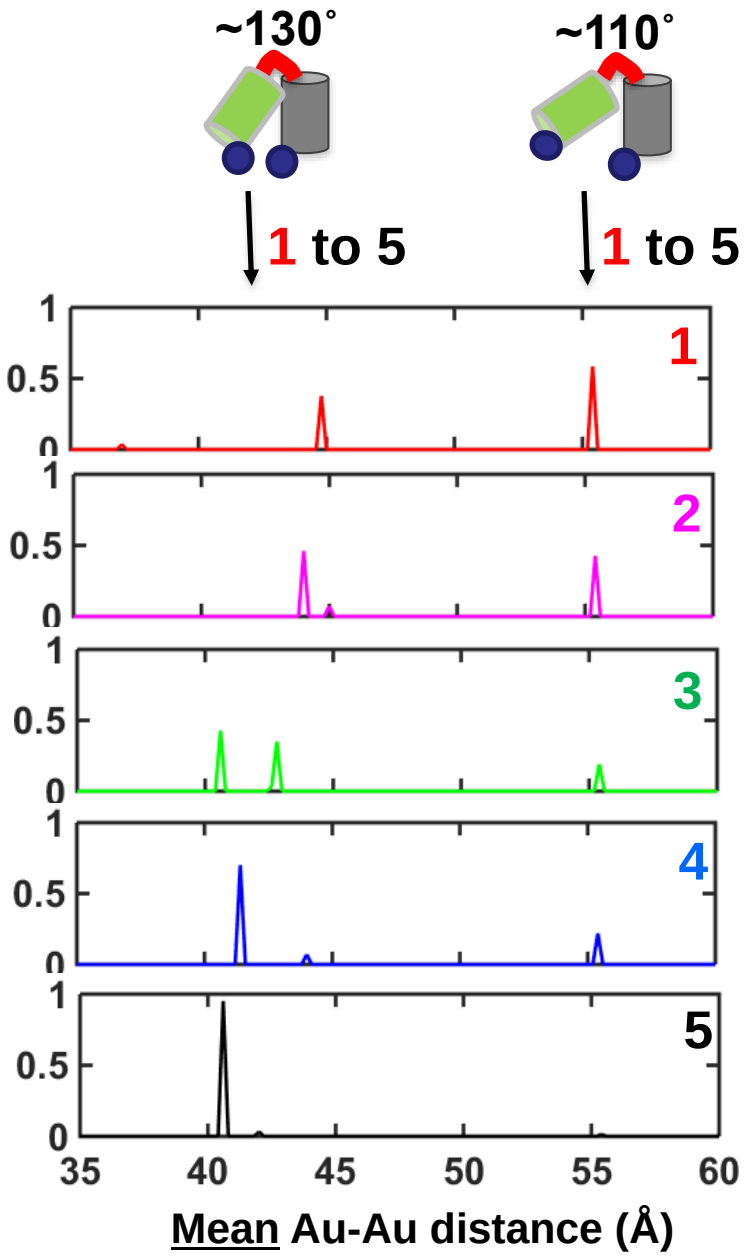


Two supports for a varied ensemble of >1 structure

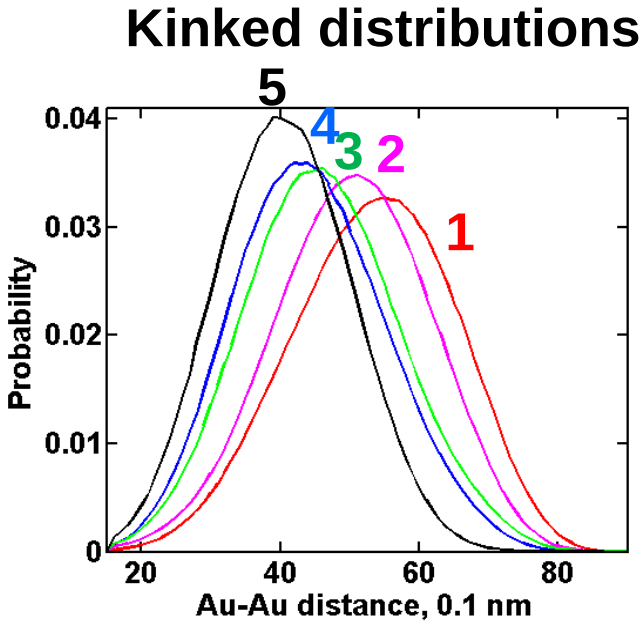
1. **The mean changes with salt, suggesting multiple kinked structures in solution.**

2. **The distributions become broader from 5 (black) to 1 (red).**

An ensemble model for the kinked KtA



Best fit

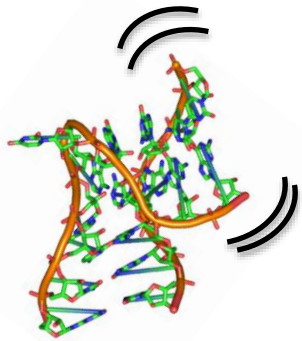


The ensemble likely consists of mainly two types of kinked conformation

Shi et al. Nat. Chem. Biol., 2016

Questions

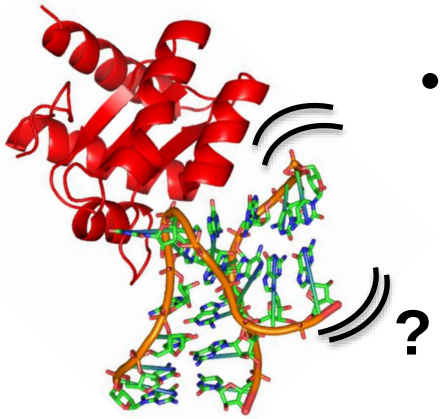
The kink-turn motif



RNA alone

- **Does folded**
  KT
  KT•L7Ae
  exist as a single conformation
  or as a dynamic ensemble?

Yes, an ensemble

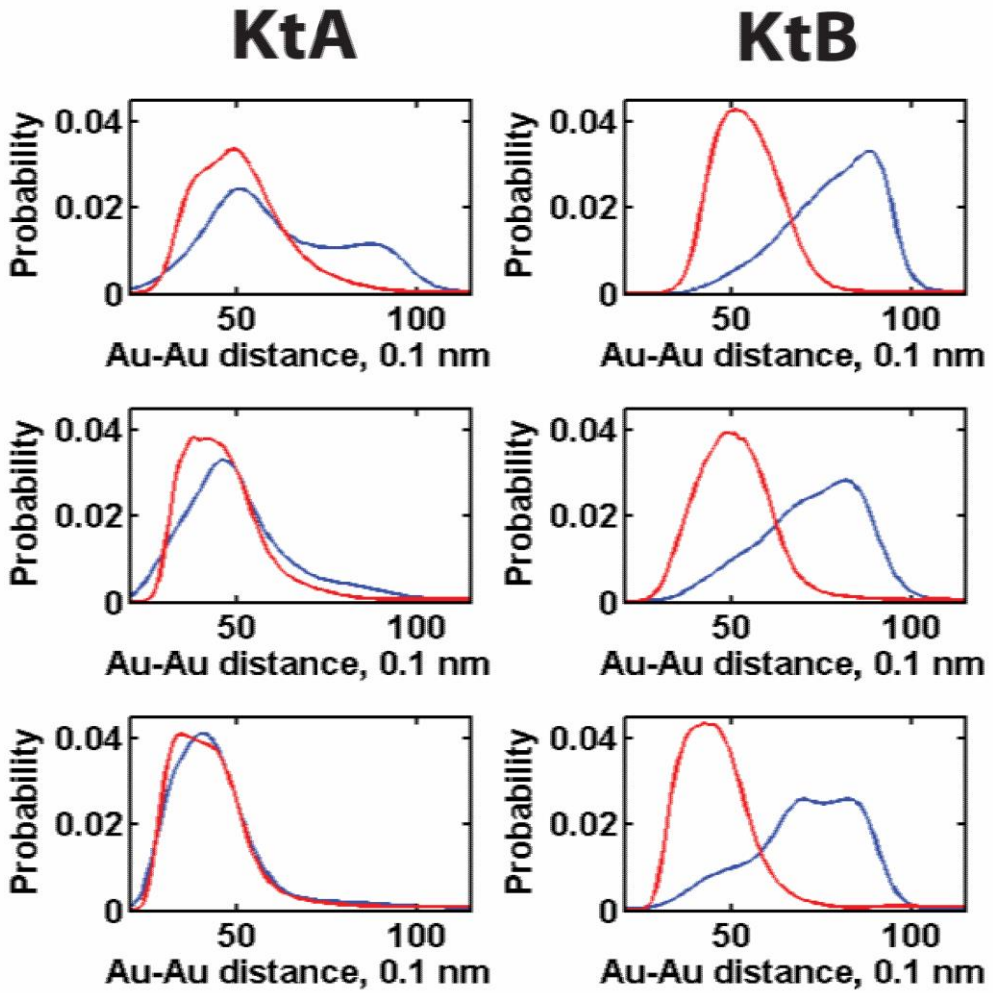


?

RNA•Protein complex

- **To what extent are RNA solution structures represented by crystal structures?**

Effects of L7Ae protein binding



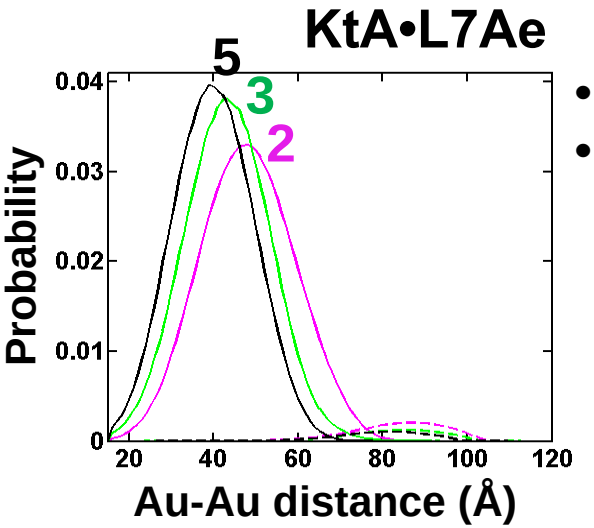
Blue: - protein

Red: + protein

L7Ae binding stabilizes kinked conformations, consistent with FRET results

Shi et al. Nat. Chem. Biol., 2016

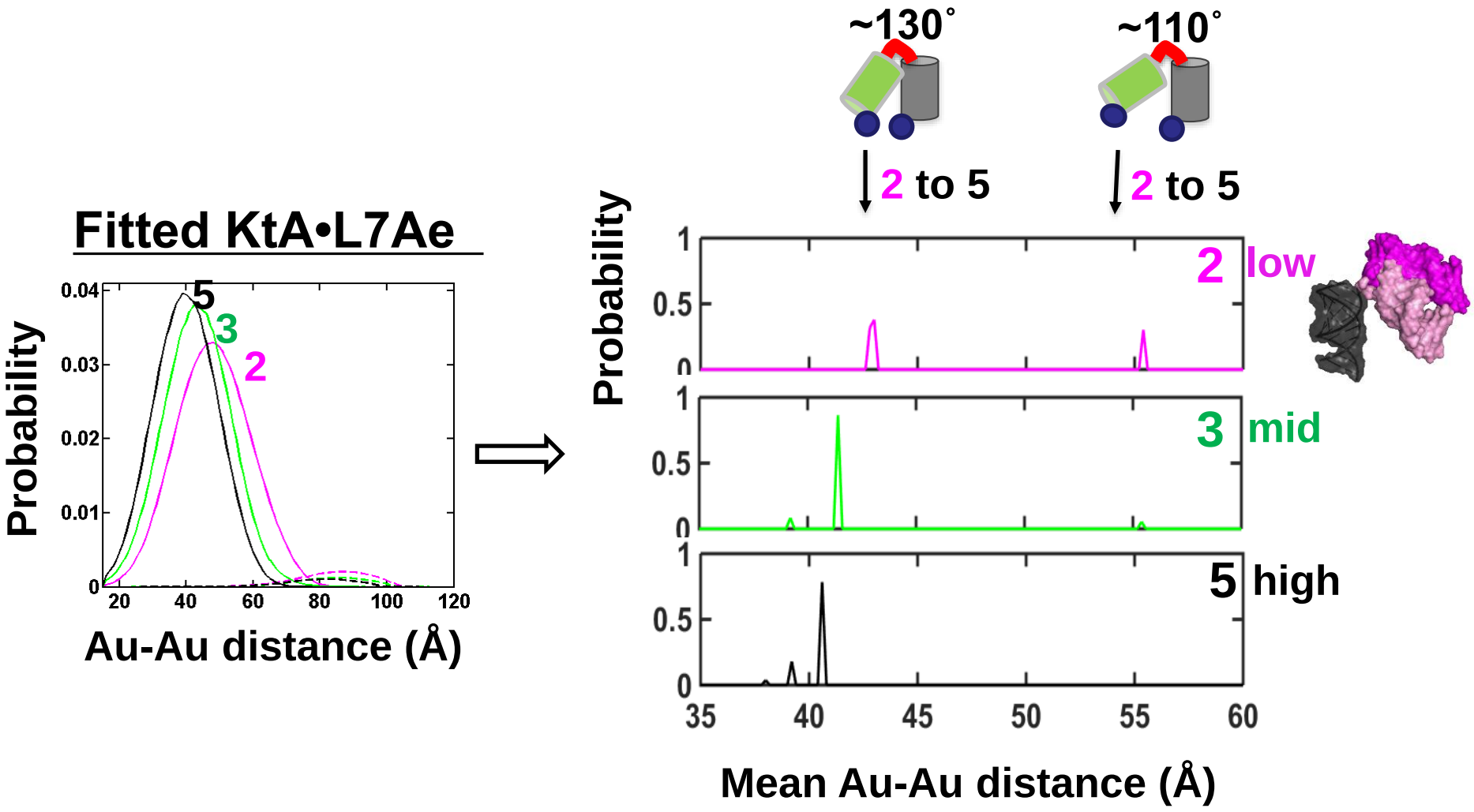
Does the kinked population remain an ensemble or is it reduced to one state with bound protein?



- The mean changes with salt, multiple solution structures
- Broader distribution from high to low salt (5 to 3 to 2).

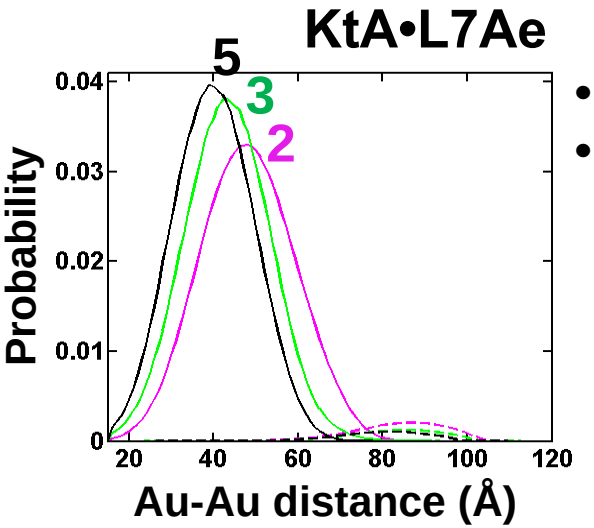
Consistent with an ensemble

An ensemble model for KtA•L7Ae



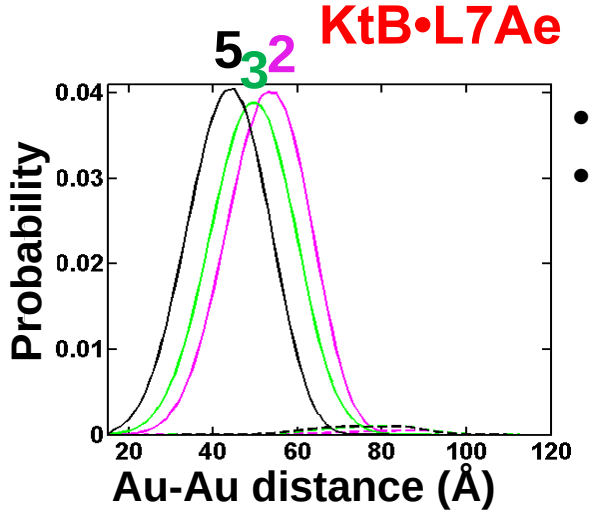
Similar trends to KtA alone

Does the kinked population remain an ensemble or is it reduced to one state with bound protein?



- The mean changes with salt, multiple solution structures
- Broader distribution from high to low salt (5 to 3 to 2).

Consistent with an ensemble

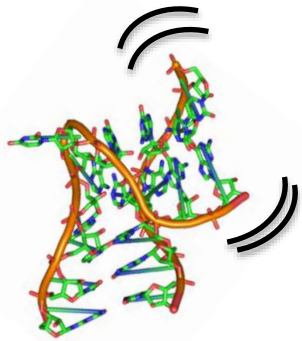


- The mean changes with salt, multiple solution structures
- Broader distribution at salt 3 compared to 2 and 5

Consistent with multiple conformations

Questions

The kink-turn motif



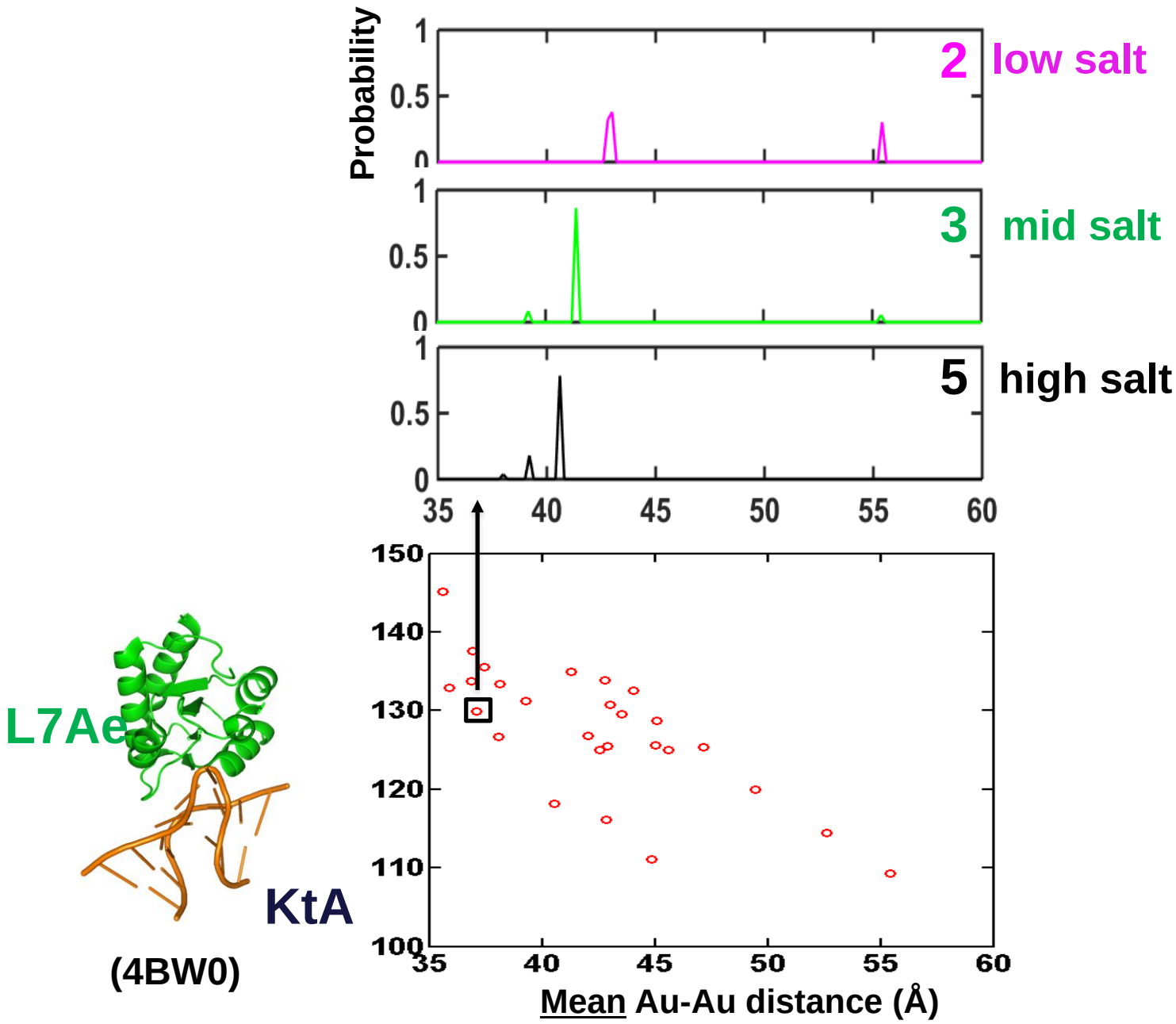
RNA alone

- **Does folded**
  KT **Yes, an ensemble**
  KT•L7Ae **Yes, an ensemble**
  exist as a single conformation or as a dynamic ensemble?

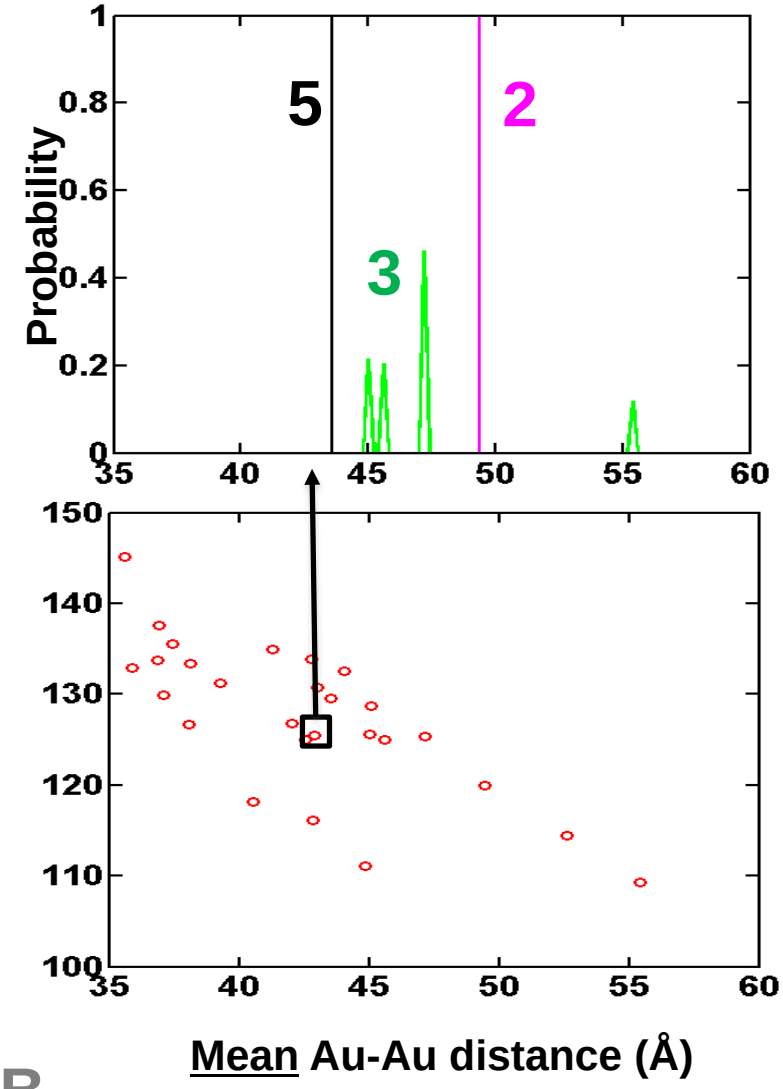
RNA•Protein complex

- **To what extent are RNA solution structures represented by crystal structures?**

Comparison with the KtA•L7Ae crystal structure



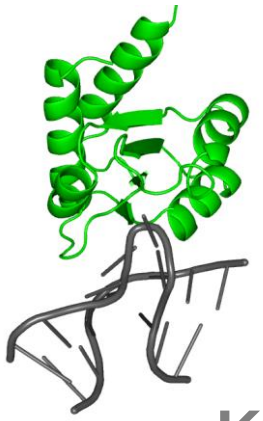
Comparison with the KtB•L7Ae crystal structure



2 low salt

3 mid salt

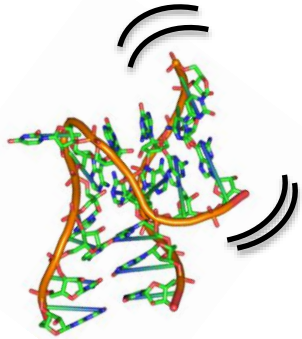
5 high salt



(1RLG)

Questions

The kink-turn motif



RNA alone

- **Does folded**
  KT Yes, an ensemble
  KT•L7Ae Yes, an ensemble
  exist as a single conformation or as a dynamic ensemble?

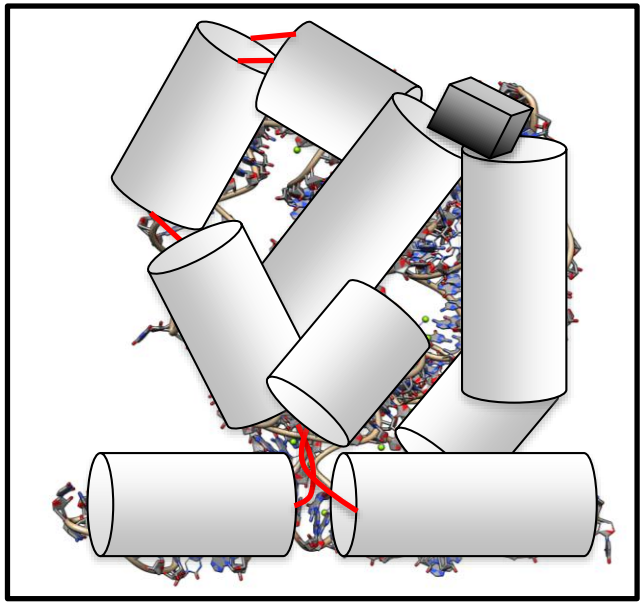
RNA•Protein complex

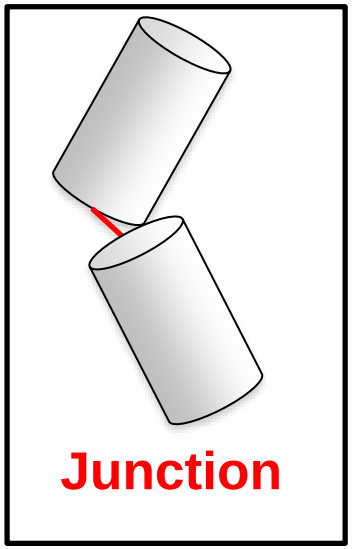
- **To what extent are RNA solution structures represented by crystal structures?**
  (1) Solution structure changes with salt and cannot be represented by a single crystal structure
  (2) Crystal structures are closer to high salt solution structure

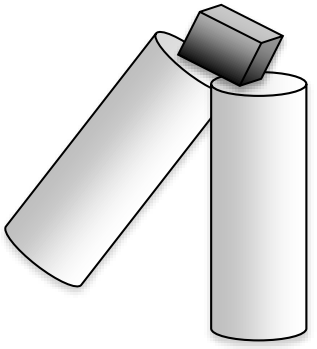
To be addressed …

- **Is the kink-turn behavior generalizable to other motifs?**
- **Is the kink-turn dynamics regulated in biology?**

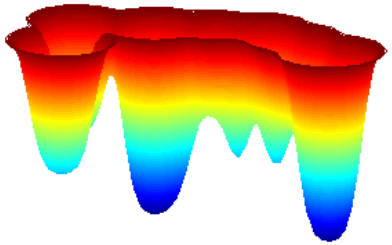
Building blocks of complex RNA



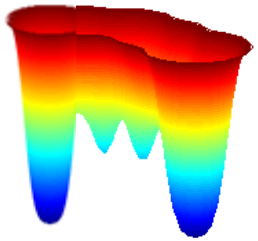




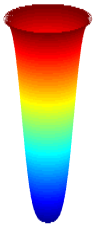
Structural motif



More dimensions



More complex

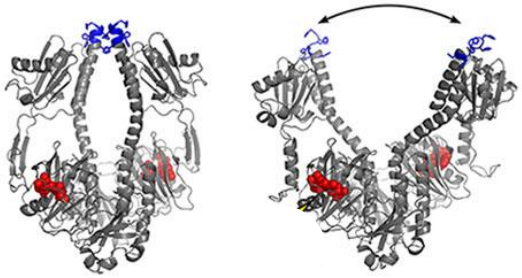


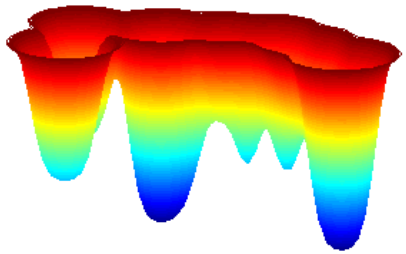
Restricted

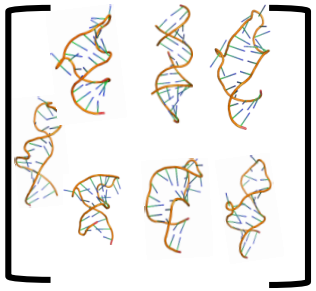
FUTURE PERSPECTIVE

Why Care About Conformational Ensembles

Conformational change and function





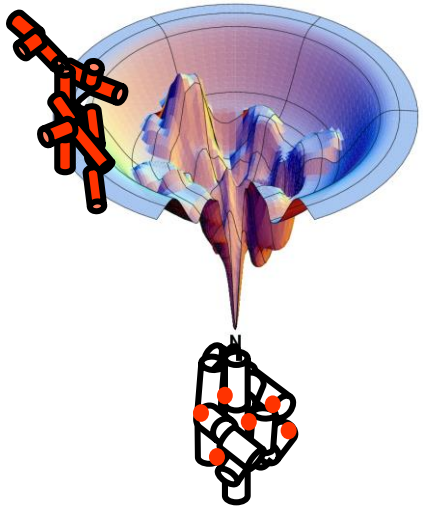


Solution structures

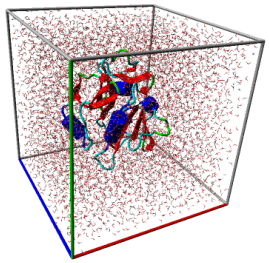
- **More accurate than an average structure**

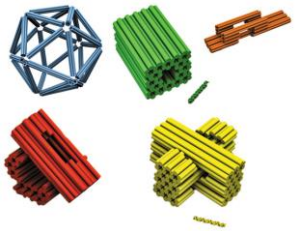
- **Too floppy to get an average structure (e.g. no crystals)**

Folding and assembly



Better physical models for bio-computation and design

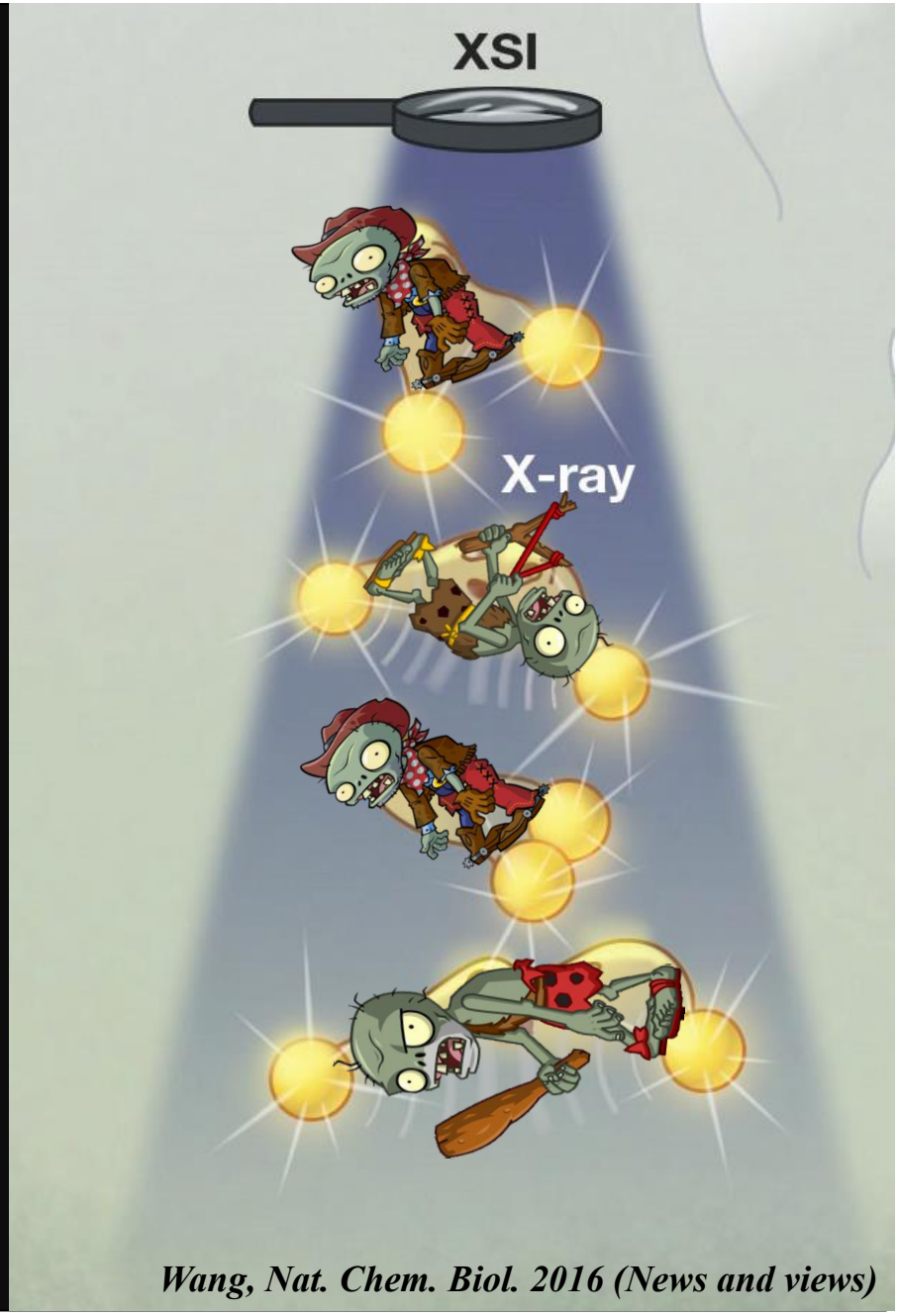




“I sense a horde, but can’t see their heads!”

Walkers in the dark

Now, go get them!



Wang, Nat. Chem. Biol. 2016 (News and views)

Acknowledgement

Daniel Herschlag

Members of the Herschlag lab

Collaborators:

Pehr B. Harbury
Rebecca Mathew
Rhiju Das
David Lilley (Dundee)
Lin Huang (Dundee)
Kyle Beauchamp
Peter Walker

SSRL beamline 4-2:

Hiro Tsuruta
Thomas Weiss
Tsutomu Matsui
Ivan Rajkovic