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 (In solution)
Macromolecules Exist as Conformational Ensembles

Conformational ensemble \(\leftrightarrow\) 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
Why Care About Conformational Ensembles

Solution structures
- More accurate than an average structure
- Too floppy to get an average structure (e.g. no crystals)

Why Care About Conformational Ensembles

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)
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
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

\[ \Delta \theta = \frac{\lambda}{d} \]
Scattering interference between two scatters

\[ \Delta \theta = \frac{\lambda}{d} \]
X-ray Scattering Interferometry

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

Pehr Harbury
Rebecca Mathew-Fenn

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

Mathew-Fenn et al. Science 2008
Shi et al. PNAS 2013
**X-ray Scattering Interferometry**

![Diagram showing Au-Au scattering pattern and distance distribution](Image)

**Same to SAXS:**
- Solution phase
- Instantaneous (time independent)

**Additional advantages:**
- Precise absolute distance (model independent)
- A distribution

*Mathew-Fenn et al. Science 2008*
*Shi et al. PNAS 2013*
Proof of principles

Precise absolute distance (sub-Å level reproducibility)

single base pair discrimination

Mathew-Fenn et al. plos one 2008
Proof of principles

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: $\sim 4^\circ$

Bending persistence length:
$B \sim 100 \text{ nm}$

Olson (1998) PNAS
Single molecule studies of DNA polymer

~50k bp

Mechanical properties of long DNA (> 1k bp)

Bending persistence length: \( B = 50 \text{ nm} \) (150 bp)

Average bending fluctuation per base step: \( \sim 6^\circ \)
Discrepancy

( macroscopic)  
Single molecule  
Stretching of long DNA (> 1k bp)

( microscopic)  
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

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

Base-step 10s of bps kbs, genome

4° or 6°? (to be determined by XSI)
Basic information from an Au-Au distance distribution

Parameters:
Mean
Variance (sd²)

Probabilty

Variance^0.5

Mean

Mean

Au-Au distance (Å)
XSI determined B-DNA helical-structure

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 persistence length: 55 ± 10 nm

Shi et al. PNAS 2013
From microscopic to macroscopic behavior

Base-step

10s of bps

kbs, genome

6° (XSI)
DNA SEQUENCE DEPENDENT VARIATION
DNA sequence dependence is important

CREB

...ATGACGTCAT...

p53

...GAGCATGCTCA...

Nucleosome
DNA A-tract, e.g. NNNAAAAAAANNN

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
DNA ENSEMBLES

Shi et al. PNAS 2014
Model System:
DNA Duplex with a single-stranded Bulge

Gohlke, P.N.A.S. (1994)
Wozniak, P.N.A.S. (2008)
A single Au-Au distribution

No bulge 3A-bulge

Shi et al. PNAS 2014
Degeneracy in a single distribution

$d_1 = d_2$

Can not distinguish
conformation 1 and 2

Conformation 1

Conformation 2
Reduce degeneracy with multiple Au pairs

Conformation 1

Au pair 1

\[ d_1 \neq d_2 \]

Can distinguish conformation 1 and 2

Conformation 2

Au pair 1

\[ d_1 = d_2 \]

Can not distinguish conformation 1 and 2

Au pair 2

\[ d_1 \neq d_2 \]

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)
6 Au pairs

50000 basis-conformations (topologically allowed)

XSI determined ensemble

6 Au-Au distributions

Shi et al. PNAS 2014
Molecular Dynamics Versus XSI Data

MD ensemble

XSI ensemble

MD optimum state (stacking)

XSI optimum state (H-bond)

Shi et al. PNAS 2014
Test of MD vs XSI Model Predictions

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

5′-GGC[A]AAAGAT
3′-CCG[......]CTA

Shi et al. PNAS 2014
Test of MD vs XSI Model Predictions

(XSI)  (MD)

5′-GGC AAAA GAT 3′-CCG ....... CTA

Relative Fluorescence Intensity

<table>
<thead>
<tr>
<th></th>
<th>Single-stranded</th>
<th>Bulged duplex</th>
</tr>
</thead>
<tbody>
<tr>
<td>5′-GGCA AAGAT</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>5′-GGCA AAGAT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3′-CCG ....... CTA</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

Yellow A is replaced by 2-AP

Shi et al. PNAS 2014
Test of MD vs XSI Model Predictions

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.
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

Start from individual elements

Junction  Structural motif

Too complex
Complex RNAs consist of similar building blocks

Junction

Structural motif

Too complex

Start from individual elements
Fundamental questions

Structural motif

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

- A single rigid conformation or a dynamic ensemble in solution?
- Does binding of a structure protein limit it to a single conformation?
Fundamental questions

- 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

- Consensus kink-turn

Kink-turns in blue
KT•L7Ae is a reoccurring RNA•Protein motif:

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

The kink-turn motif

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

• To what extent are RNA solution structures represented by crystal structures?
Kink-turns investigated

KTA (Kt7)

KtB (Box C/D)

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

3nt-bulge:
5′…NN $\text{NNN}$ NN NN…
3′…NN $\text{NNN}$…

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

How to detect an ensemble

An ensemble broadens the Au-Au distance distribution (e.g. lower peak height with normalized area)
KtA (Kt7 in ribosome) exists as an equilibrium of kinked and unkinked conformations.

Au-Au distance distributions for KtA

Salt conditions

<table>
<thead>
<tr>
<th>Salt conditions</th>
<th>Mg$^2+$ (mM)</th>
<th>Na$^+$ (mM)</th>
<th>K$^+$ (mM)</th>
<th>Tris (mM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>0</td>
<td>20</td>
<td>0</td>
<td>30</td>
</tr>
<tr>
<td>2.</td>
<td>0</td>
<td>60</td>
<td>0</td>
<td>70</td>
</tr>
<tr>
<td>3.</td>
<td>1</td>
<td>10</td>
<td>150</td>
<td>70</td>
</tr>
<tr>
<td>4.</td>
<td>0</td>
<td>510</td>
<td>0</td>
<td>70</td>
</tr>
<tr>
<td>5.</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>70</td>
</tr>
</tbody>
</table>
Isolate the kinked-state

Kinked distributions
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

The ensemble likely consists of mainly two types of kinked conformation

Questions

The kink-turn motif

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

• To what extent are RNA solution structures represented by crystal structures?
Effects of L7Ae protein binding

L7Ae binding stabilizes kinked conformations, consistent with FRET results

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

Fitted KtA•L7Ae

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

**KtA•L7Ae**
- The mean changes with salt, multiple solution structures
- Broader distribution from high to low salt (5 to 3 to 2).

**Consistent with an ensemble**

**KtB•L7Ae**
- 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

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

• To what extent are RNA solution structures
  represented by crystal structures?

RNA alone

RNA•Protein complex
Comparison with the KtA•L7Ae crystal structure

<table>
<thead>
<tr>
<th>Probability</th>
<th>Low salt</th>
<th>Mid salt</th>
<th>High salt</th>
</tr>
</thead>
</table>

![Comparison diagram showing the comparison of the KtA•L7Ae crystal structure with different salt concentrations.](image)

L7Ae (4BW0)  
KtA
Comparison with the KtB•L7Ae crystal structure

- 2 low salt
- 3 mid salt
- 5 high salt

Mean Au-Au distance (Å)
Questions

The kink-turn motif

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

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

More dimensions

Junction

More complex

Structural motif

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!

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