Rotational and Translational Positioning Patterns in the Yeast Nucleosomes Mapped by Paired-end Sequencing

Feng Cui, Difei Wang, Hope A. Cole, David J. Clark

and Victor B. Zhurkin (NIH)
Left-handed helix

E.T. in Paris, 1989
Sequence patterns associated with rotational and translational positioning of nucleosomes

Rotational positioning
- Trifonov (1980) AA:TT
- Zhurkin (1983) YR, RY
- Travers (1986) A+T vs. G+C

Translational positioning
- Long A-tracts in the linkers (Struhl 1985; Rando et al. 2005)

What else? (GC-rich center?)
Periodic patterns of AT-containing fragments

Chicken NCPs

WW (AA+TT+AT+TA)

Satchwell et al. (1986) *J Mol. Biol.*

V.B. Zhurkin
I. Yeast (5.3 mil)  

Cole et al. (2011) Nucleic Acid Res.

II. Yeast (70,000)  


III. Fly (60,000)  

Mavrich et al. (2008) Nature

WW = AA + TT + AT + TA  

SS = GG + CC + GC + CG

V.B. Zhurkin
Yeast nucleosomes (~55,000) from F. Pugh lab

Mavrich et al., Genome Res 2008

Original Set

Realigned Set

Size unknown: The ends are NOT paired

V.B. Zhurkin

~30 nt
Paired-end sequences can be analyzed without additional re-alignment

(Pugh, 2008 (Fly); Segal, 2008 (Yeast); Clark, Henikoff labs, 2011)

Nucleosome sequences based on ‘single reads’ need re-alignment

(Segal, Pugh labs, 2006-2010)
Translational positioning of nucleosomes

“Not all animals kinks are born equal”

Distribution of WW dimers (AA:TT, AT, TA); Clark et al. 2011
Two histone motifs interacting with SHL ±0.5

A wide minor groove at SHL ±0.5. GC-rich sequences are favorable

Translational positioning signal #1 ?

V.B. Zhurkin
Yeast nucleosomes contain a novel pattern at SHL ±4.5

Translational positioning signal #2?

May be related to histone H2A sequence in yeast

In yeast nucleosomes, Adenines are more frequent in the leading strand; Thymines are more frequent in the complementary strand at SHL – 4.5.
The residues in histone H2A N-tail interacting with DNA at SHL ±4.5 are not conserved. The H2A N-tail in yeast is more hydrophobic than in other species. This increase in hydrophobicity may explain species-specific pattern in yeast nucleosomal DNA sequences (at SHL ±4.5).
TT : AA “wedges” may help DNA bending in this case
Paired-end yeast nucleosomal sequences can be analyzed without additional re-alignment

→ Rotational positioning patterns, WW vs SS.

→ Two Translational positioning signals:

#1. SHL ± 0.5: GC-rich center (universal – yeast, fly, chicken?)

#2: SHL ± 4.5: Adenines ‘inside’ vs Thymines ‘outside’
   (yeast specific signal related to histone H2 sequence?)
Part II: Role of flexible YR dimeric steps

• DNA - Histone interactions (Arg in the minor groove)

• ‘Kink-and-Slide’ deformations in nucleosome

• YYRR (e.g., TTAA, CCAA) most favorable for minor groove bending

• RYRY → Major groove bend

• Using these ‘rules’ we predicted NU positioning in vitro (rotational and translational)

IN VIVO: YYRR in Yeast nucleosomes
‘Kink-and-Slide’ deformations in nucleosome

Tolstorukov et al., JMB 2007

Slide ≈ 2.5 Å

Roll ≈ −20°

Blue → minor groove. Red → Major groove.

V.B. Zhurkin
Arginines in the minor groove facilitate DNA Slide \( \approx 2-3 \, \text{Å} \) (R77, R43 from histone H2A)
Arginine in the minor groove and DNA Slide

Slide = 2-3 Å

Sequence dependence

5' YR
Favorable: no R-R clash

5' RY
Unfavorable: R-R clash

Arg

PDB: 1kx5
SHL -5.5

V.B. Zhurkin

Wang et al., JBSD 2010
Steric restraints mimicking Arg-DNA interactions
(Slide = 2.5 Å; Roll = −20°)

**Favorable for ‘Kink-and-Slide’**
YYRR: TTAA, CCAA:TTGG (± CCGG)
Explains data by Crothers and Widom:
TTAA in minor groove

**Unfavorable for ‘Kink-and-Slide’**
RRYY: AATT, AACC:GGTT, GGCC

RYRY → Major groove bend

Wang et al., JBSD 2010

New strategy: Looking for YYRR tetramers in the minor-groove bending positions

V.B. Zhurkin
Distribution of **DIMERS** and **TETRAMERS** in paired-end nucleosomes

Both **YR** and **RY** $\rightarrow$ max in **Major groove** bends (agrees with max **RYRY**)

**YYRR** $\rightarrow$ max in **minor groove** bends.

**YR** behavior depends on the context.

*Clark et al. 2011;  L= [147-152]*
Part II: Role of flexible YR dimeric steps

- DNA - Histone interactions (Arg in the minor groove)

- ‘Kink-and-Slide’ deformations in nucleosome

- YYRR (e.g., TTAA, CCAA) most favorable for minor groove bending

- RYRY → Major groove bend

- Using these ‘rules’ we predicted Nu positioning *in vitro* (rotational and translational)

*In vivo:* YYRR in Yeast nucleosomes
Part III: Role of Linker Histones in ...

Linker histone (LH) H5 / H1°
IV: Heterogeneous sizes of the yeast nucleosomes. Is this related to MNase cleavage?

MNase preference for A+T sequences:
Cleavage occurs at $W|WS$ sequences ($W=A+T$, $S=G+C$)
($A|TG$, $A|TC$, $T|TG$, etc.)

![Histogram showing the distribution of read lengths with a peak around 147-152 bp (~5.3 mil)]
173dimer-2 (MNase+Exo)
(num ~ 16 million)

Length of nucleosomal fragment, bp

% of total fragments

[145-147] (~ 3 mil)
Future: MNase + exonuclease “cocktails”?  

ChIP-exo:

$\lambda$-exonuclease, 5’-to-3’ (Rhee, Pugh 2011)

MNase

Exo III

Comparable to J. Widom’s ‘chemical mapping’

$\rightarrow$ Easily applicable to higher eukaryotes!!
173 dimer-2_MN+exo [145-147] (10-pile) (~1 mil)
173 dimer_2_MN+exo [145-147] (~3 mil)
CC [147-152] (~5 mil)
Translational positioning of nucleosomes

“Not all kinks are born equal”

X-ray:
The strongest DNA deformations observed in
SHL ± 5.5, 3.5, 1.5 (minor groove)
SHL ± 2 (Major groove)
Eduardo de Haifa at NIH  ca. 1995
Flexibility of YR dimers: $TA:TA \geq CA:TG >> CG:CG$
(tendency for Sliding + Twisting + Bending into the minor groove)
Acknowledgements:

T. Nikitina (LCB, NCI), S. Grigoryev (Penn. State):

experiments with MNase cleavage of nsm 601.
173 dimer-2_MN+exo [145-147] (10-pile) (~1 mil)
173 dimer_2_MN+exo [145-147] (~3 mil)
CC [147-152] (~5 mil)

V.B. Zhurkin
Micrococcal Nuclease, exo-cellular Ca^{2+} dependent nuclease

MNase cleavage of the P-O5’ bond: **A+T** specificity

View from the minor groove:
- phosphate at the catalytic center
- flipped base (A,T)
- phosphate “subsites”

C. Anfinsen, 1971 (Review)
X-ray structure of MNase + THP + Ca$^{2+}$

Model: MNase + DNA + Ca$^{2+}$

F. Cotton et al., 1979

D. Wang (unpublished)

V.B. Zhurkin
His46 from the “finger” loop may “poke” the base out of the DNA duplex (A or T)

V.B. Zhurkin
MNase cleaves DNA (yellow strand) and the linker/core junction

Yellow: 145 bp core DNA (nsm 601)

V.B. Zhurkin
**Sequence ‘601’**

(Lowary & Widom, 1998)

5’…CGCCC|TGGAGAAT-----CCTGTGCA|ATGTGG…3’ wt-601. 149-bp core DNA (MNase cleavage)

---

**MNase cleavage of 601-AT mononucleosomes**

<table>
<thead>
<tr>
<th>601-wt</th>
<th>5’-CTATACGCGGCGGCCC</th>
<th>TGGAGAATCCCCGG</th>
<th>“left” linker</th>
</tr>
</thead>
<tbody>
<tr>
<td>601-AT (5)</td>
<td>5’-CTATACGCGGCGGCCC</td>
<td>GGCATGATCCCCGG</td>
<td>... ATGGTGCGTAGACAGCT</td>
</tr>
<tr>
<td>601-AT (3)</td>
<td>5’-CTATACGCGGCGGCCC</td>
<td>CATGGAATCCCCGG</td>
<td></td>
</tr>
<tr>
<td>601-AT (1)</td>
<td>5’-CTATACGCGGCGGCGCA</td>
<td>TGGAGAATCCCCGG</td>
<td></td>
</tr>
<tr>
<td>601-AT (-2)</td>
<td>5’-CTATACGCGGCGCCATGC</td>
<td>GGGAGAATCCCCGG</td>
<td></td>
</tr>
<tr>
<td>601-AT (-5)</td>
<td>5’-CTATACGCCATGGCCC</td>
<td>GGGAGAATCCCCGG</td>
<td></td>
</tr>
</tbody>
</table>

ACGTGTAGATATATACATCCTTGCAAAGTGTGGATCCGAAT-3’

“right” linker
601-AT (5) CTATACGGC\textsubscript{5}CGC\textsubscript{10}GCCGCCGCTCAATTGGTCGTAGACAG\textsubscript{43} CT
601-AT (3) CTATACGGC\textsubscript{3}CGCCCGCTCAATTGGTCGTAGACAG\textsubscript{43} CT
601-AT (1) CTATACGGC\textsubscript{1}CGCCCGCTCAATTGGTCGTAGACAG\textsubscript{43} CT
601-AT (-2) CTATACGGCG\textsubscript{2}CGGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAG\textsubscript{43} CT
601-AT (-5) CTATACGCC\textsubscript{5}GCGCCCGCGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAG\textsubscript{43} CT

V.B. Zhurkin
<table>
<thead>
<tr>
<th>ExoIII</th>
<th>AT(5)</th>
<th>AT(5)</th>
<th>AT(1)</th>
<th>AT(1)</th>
<th>AT(-5)</th>
<th>AT(-5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

37

MNase + exoIII

V.B. Zhurkin
601-AT (5)  CTATA CGCGGCGCCCGCCGCGCATGATCCCGGTGCCGAGGCGCTCAATTGGTCGTAGACAGCT

601-AT (3)  CTATA CGCGGCGCCCGCCGCGCATGGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCT

601-AT (1)  CTATA CGCGGCGCCCGCCGCGCATGAGGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCT

601-AT (-2)  CTATA CGCGGCGCCGCGCGAGGGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCT

601-AT (-5)  CTATA CGCGCGCATGCCCGGGAGAATCCCGGTGCCGAGGCCGCTCAATTGGTCGTAGACAGCT

V.B. Zhurkin
**MNase cleavage of mononucleosomes**

- **Length of read, bp**
  - 120
  - 130
  - 140
  - 150
  - 160
  - 170
  - 180

- **% of total reads**
  - 0
  - 1
  - 2
  - 3
  - 4
  - 5
  - 6
  - 7

- **MNase cleavage of mononucleosomes**
  - ~10 bp
  - 2-3 bp
  - 140 bp
  - 160-165 bp

- **G+C**

- **140-160 bp**

- **H2AZ ?**

- **V.B. Zhurkin**
Nucleosome Positioning
Jonathan Widom (NWU)
The nucleosome signature at higher resolution

- **Chemical map**
- **MNase map, 147 bp only**

**Graph Details:**
- Y-axis: AA/TT/TA/AT Frequency
- X-axis: Distance (bp)
- Horizontal lines at 0.3 and 0.4
II. Translational positioning of nucleosomes

“Not all animals kinks are born equal”

X-ray:
The strongest DNA deformations observed in

SHL ± 5.5, 3.5, 1.5 (minor groove)

SHL ± 2 (Major groove)
This idea apparently ‘works’ *in vitro*.

Calculate the Scores for various patterns:
- YYRR, WW (A+T) → minor groove
- RYRY, SS (G+C) → major-groove

Incorrectly predicted *in vitro* nucleosome positioning

Correctly predicted *in vitro* nucleosome positioning

Error = 0

‘601’

In *vitro* (20 positions)

Cui & Zhurkin, *JBSD* 2010
Segal; Clark; Henikoff

Clark [147-152] (~1.3 million)

YYRR vs RYRY

V.B. Zhurkin
Yeast $5 \times 10^6$ nucleosomes, 147-152 bp (D. Clark et al.)

Distribution of WW dimers (AA:TT, AT, TA)
**TTAA versus ATAT: X-ray data**

(free DNA & protein+DNA better than 2.5 Å)

<table>
<thead>
<tr>
<th></th>
<th>Twist (°)</th>
<th>Roll (°)</th>
<th>Slide (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>45 structures</td>
<td>TTAA</td>
<td>36 ± 6</td>
<td>5.0 ± 6</td>
</tr>
<tr>
<td>49 structures</td>
<td>ATAT</td>
<td>41 ± 4</td>
<td>-0.7 ± 4</td>
</tr>
</tbody>
</table>

Yanagy, Privet & Dickerson (1991): YCAR vs XCAX (e.g., RCAY)

YCAR: High Twist Profile (Twist ~45; Roll <0)

Buckle

**TTAA**
- 0.5 ± 6
- 0.5 ± 6
- -2.5 ± 7

Buckle

**ATAT**
- 0.5 ± 6
- 5.0 ± 6

V.B. Zhurkin
Henikoff-yeast (~95 million)

% of total reads

Length, bp

[147-152] (~12 mil)

% of total reads

Length, bp

Henikoff-yeast (~95 million)  PNAS USA  2011

V.B. Zhurkin
IV. Nematode (44 mil)

V. Human (865 million)

Size unknown: The ends are NOT paired

~30 nt
Size unknown: The ends are NOT paired

~30 nt  ~30 nt
II. Translational positioning of nucleosomes

“Not all animals kinks are born equal”
Genome-wide nucleosome maps of yeast

**Paired-end sequencing**
- MNase treatment of yeast chromatin
- Both ends of the digested fragments sequenced and mapped to genome

(Hope, Howard & Clark, *NAR* 2011)
(Henikoff, *PNAS USA* 2011)

most prevalent length: ~150 bp

(Fields,… & Segal: ~160 bp)

\[
\begin{array}{c}
\text{[147-152] (\sim 5.3 \text{ mil})} \\
\end{array}
\]

\[
\begin{array}{c}
\text{[147-152] (\sim 70,000)} \\
\end{array}
\]
AT-containing fragments are depleted in nucleosome center: A translational positioning signal?

**Yeast** (Cole *et al.*)

**Yeast** (Field *et al.*)

**Fly**

SHL -0.5

V.B. Zhurkin