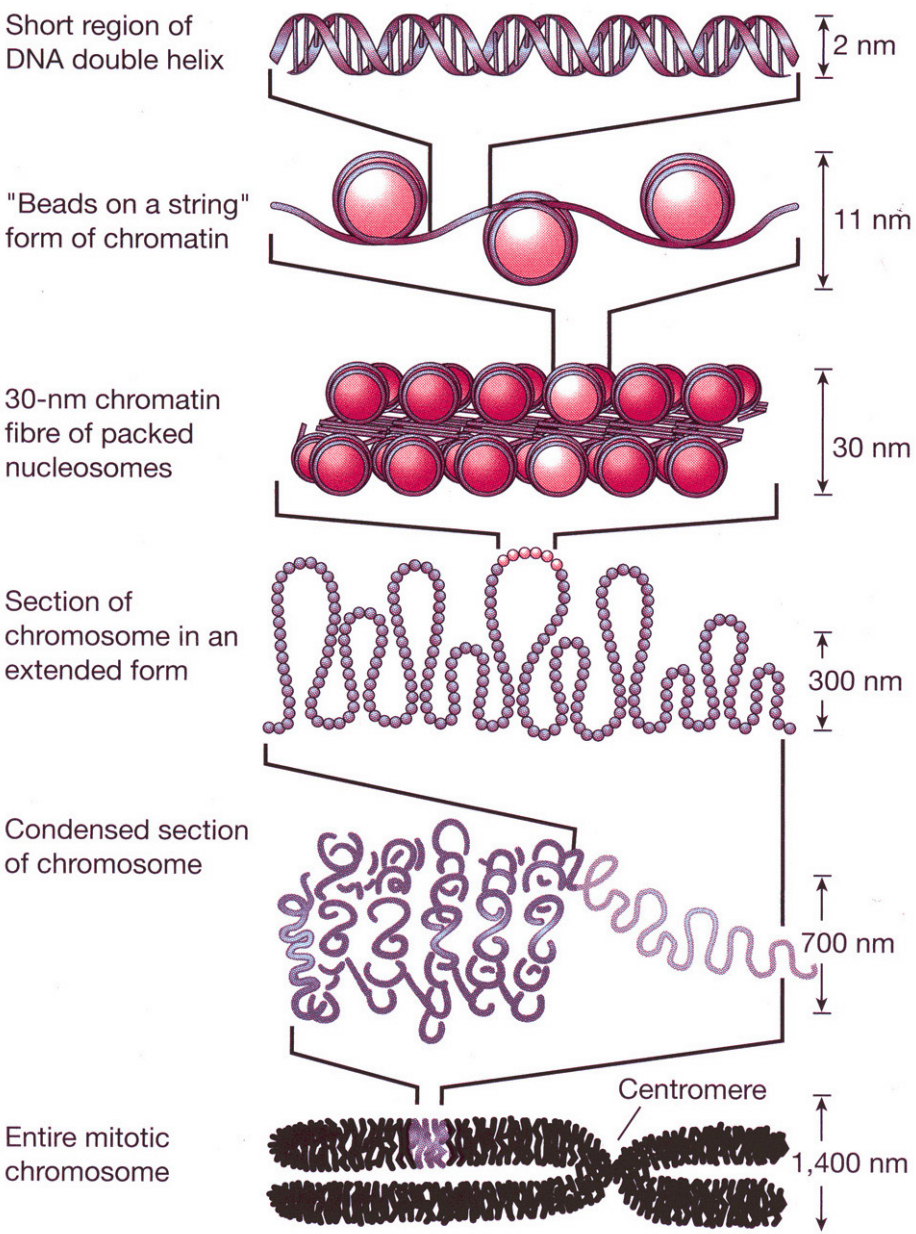
A genomic code for nucleosome positioning

Hierarchical DNA folding in eukaryotic chromosomes

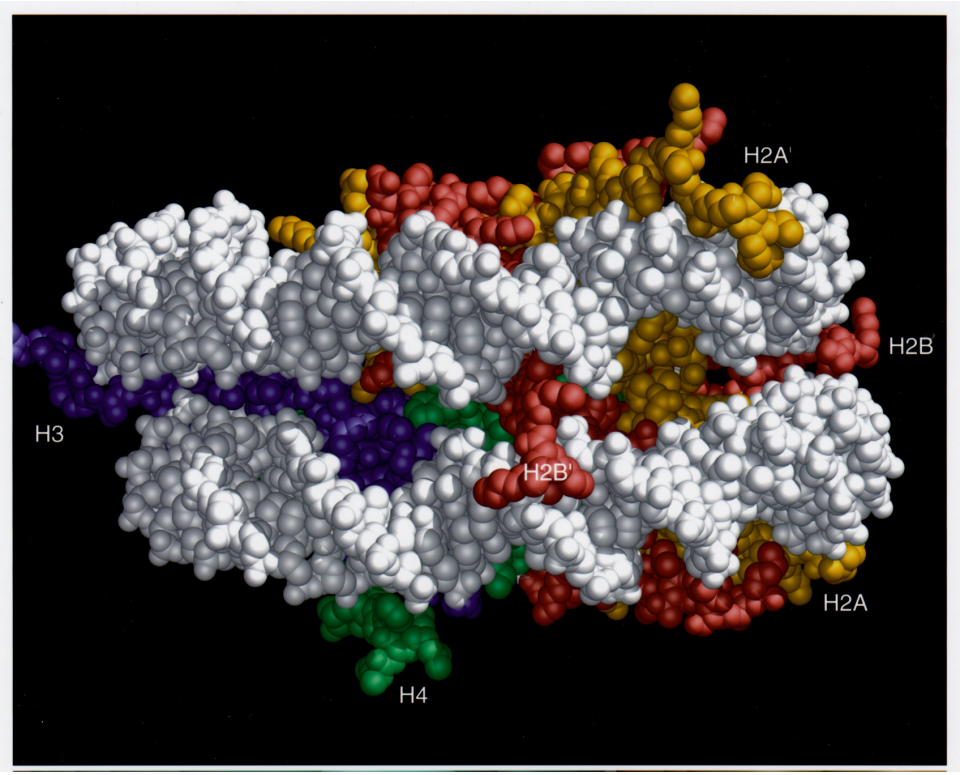


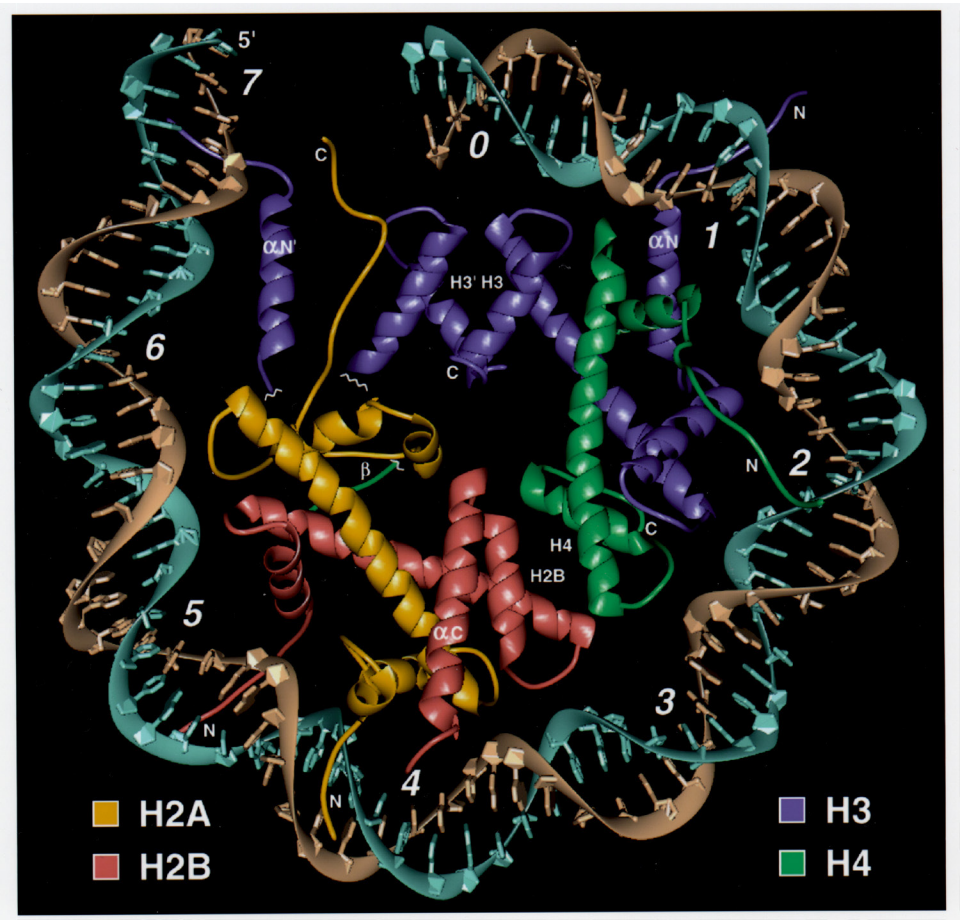
DNA double helix

Nucleosomes

Felsenfeld & Groudine,,
Nature 421: 448-453 (2003)

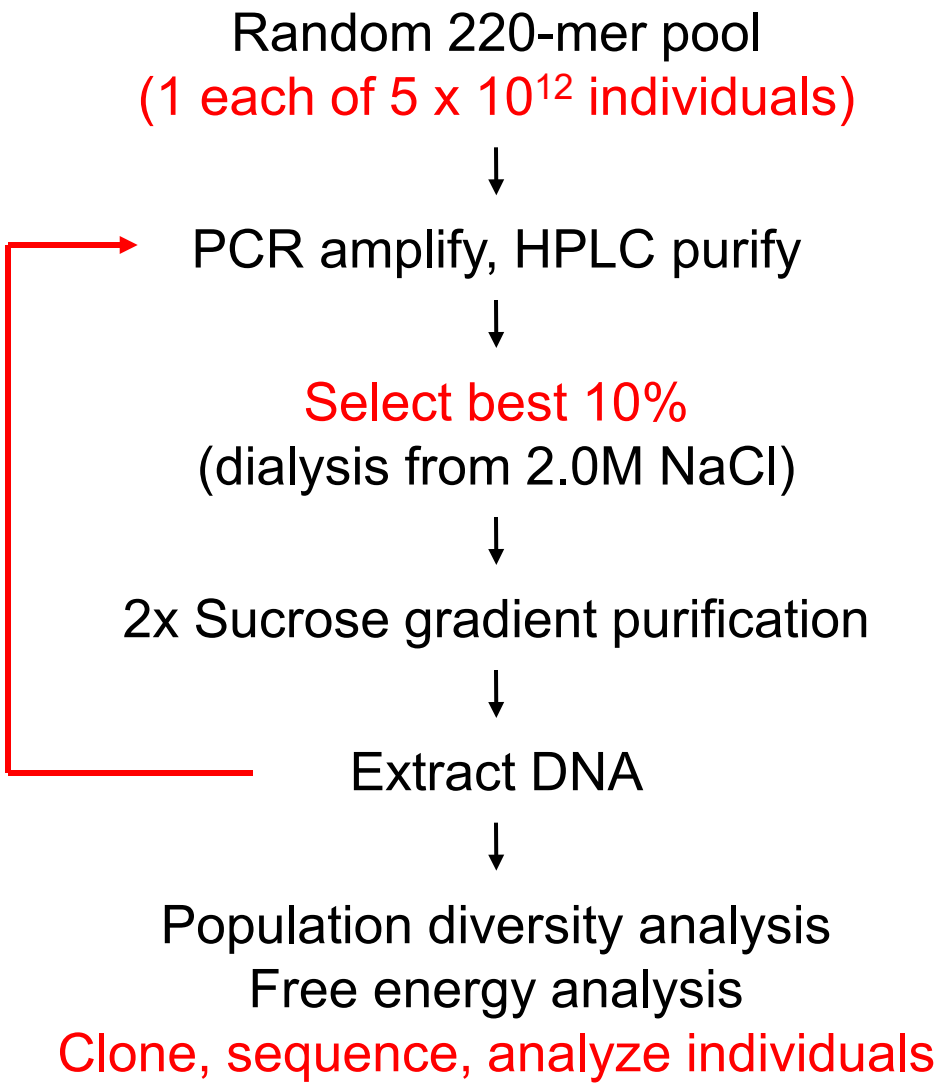
Most eukaryotic DNA is sharply looped





Luger et al., 1997

Physical selection for stable nucleosome formation on chemically synthetic random DNAs

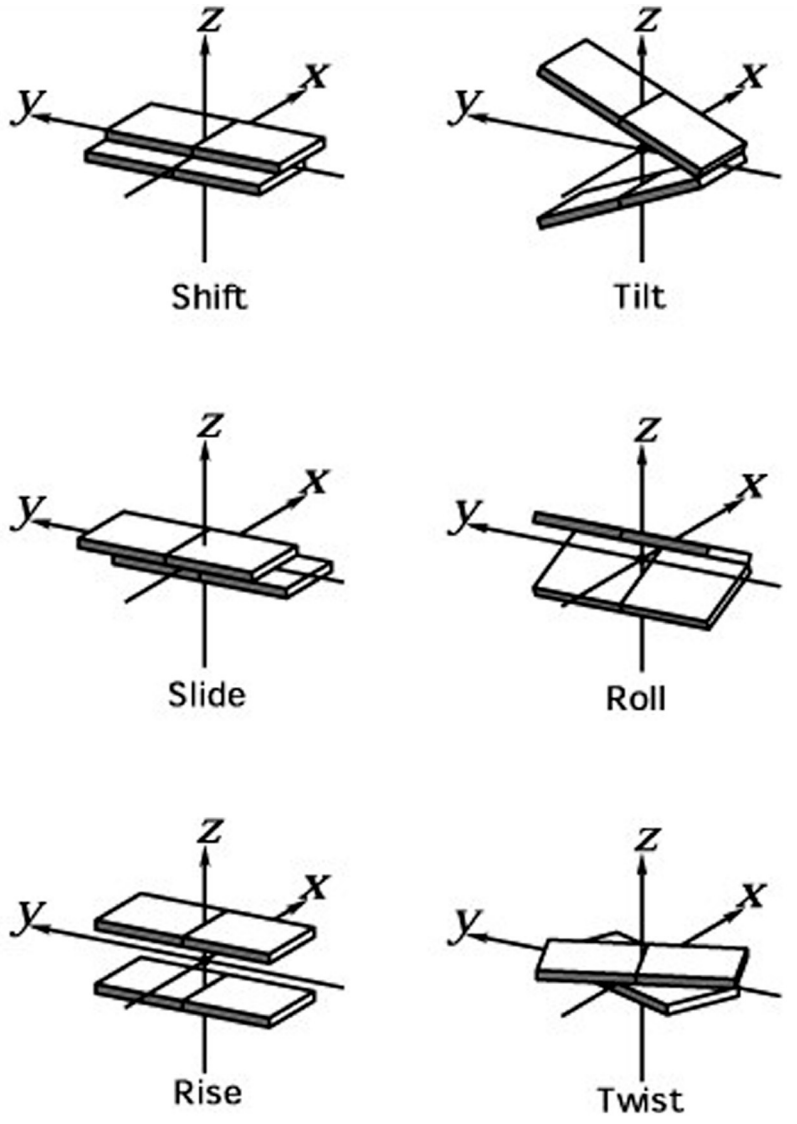


Lowary & Widom, 1998

- Differing DNA sequences exhibit a > 5,000-fold range of affinities for nucleosome formation

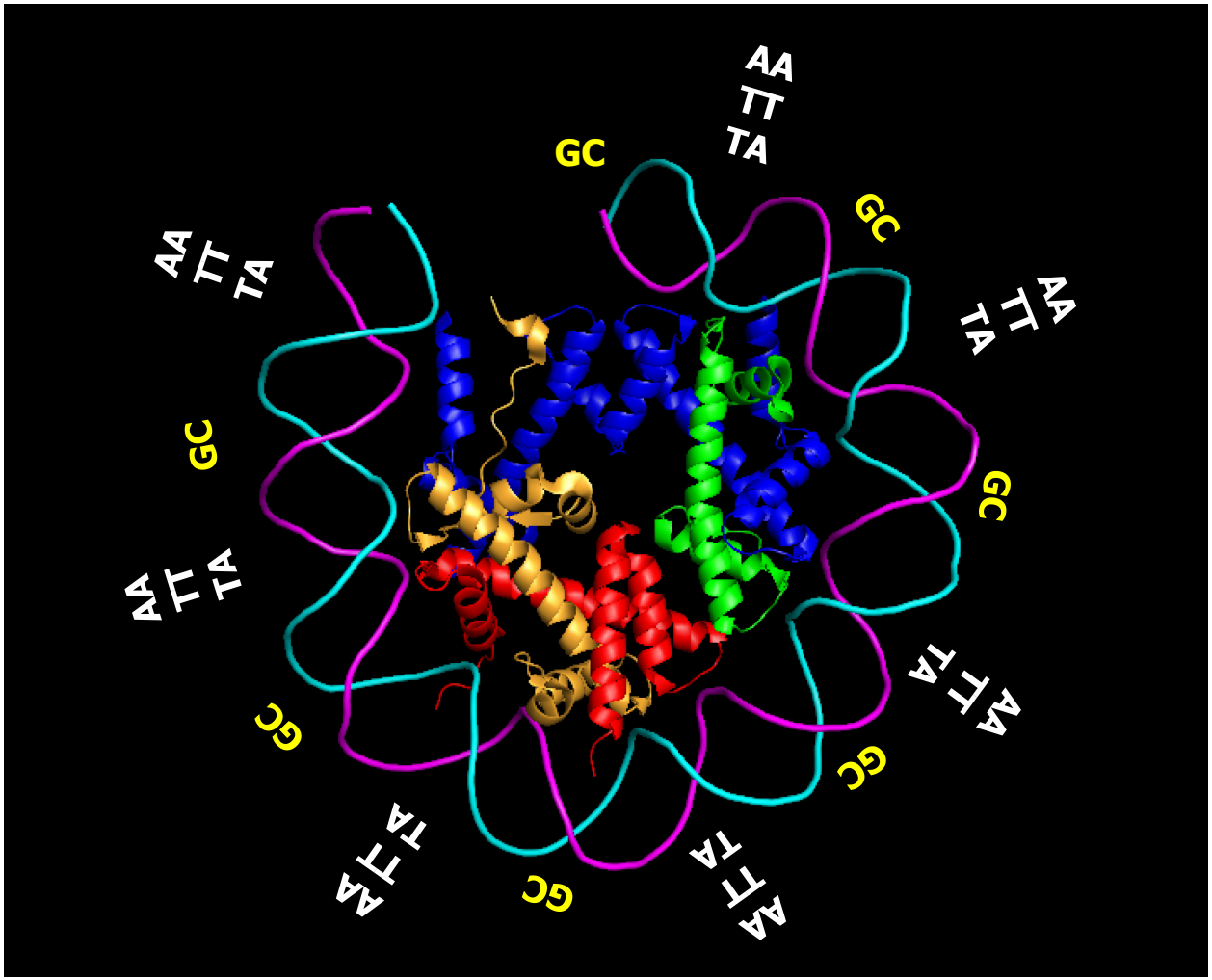
Lowary & Widom, 1998
Thåström et al., 1999
Widom, 2001
Thåström et al., 2004

Basepair steps as fundamental units of DNA mechanics



Zhurkin
Olson

DNA sequence motifs that stabilize nucleosomes and facilitate spontaneous sharp looping

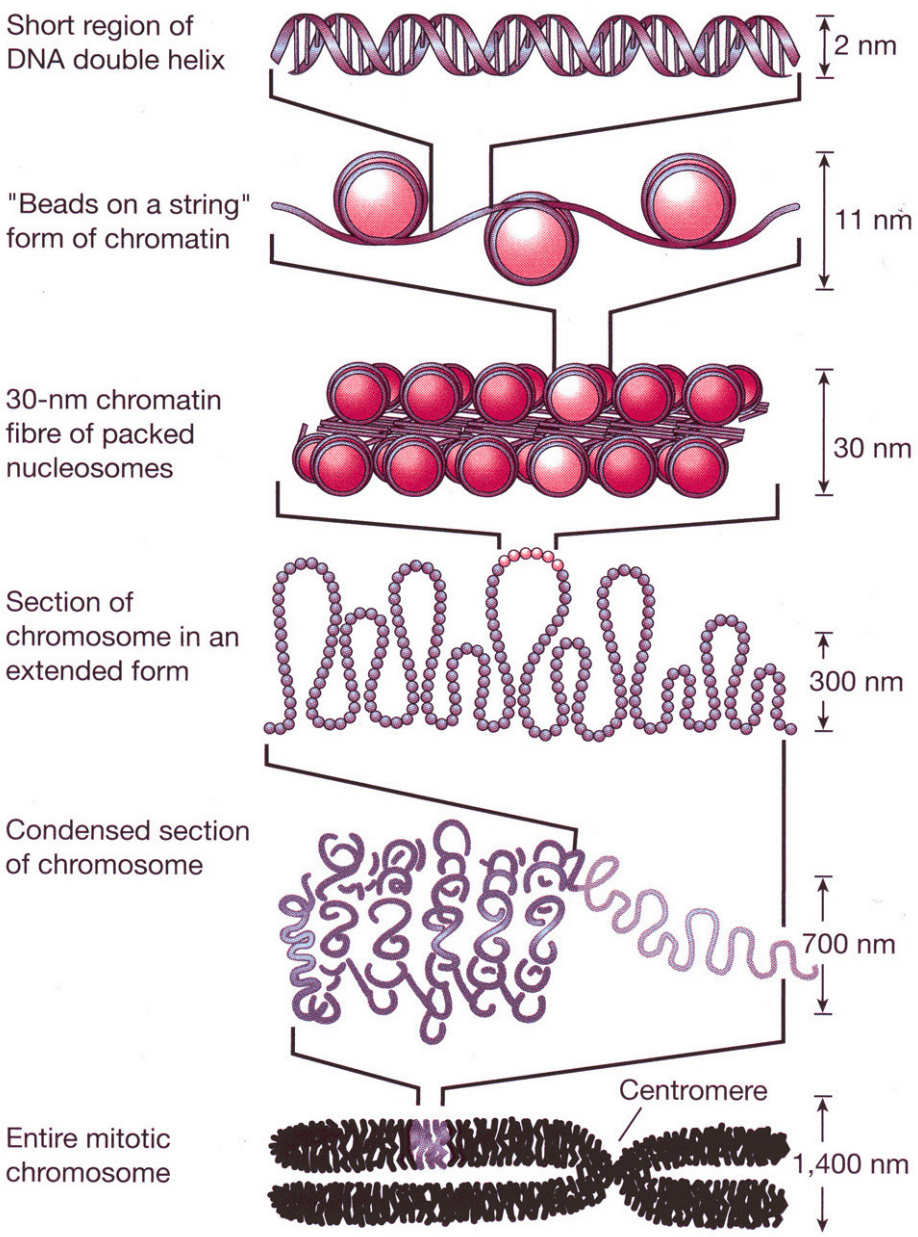


Segal et al., 2006

Why DNA some sequences have especially high affinity for histone octamer

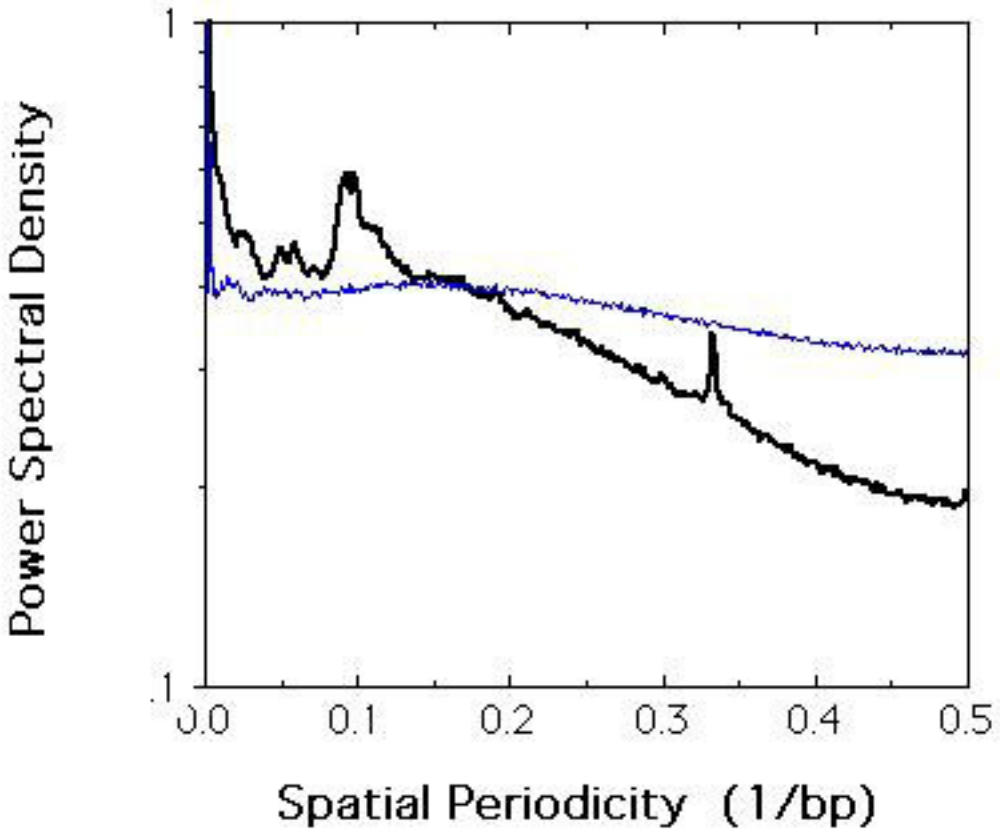
- More or better bonds
- Appropriately bent
- More easily bendable
- Appropriate twist
- More easily twistable

Widom, 2001
Cloutier & Widom, 2004



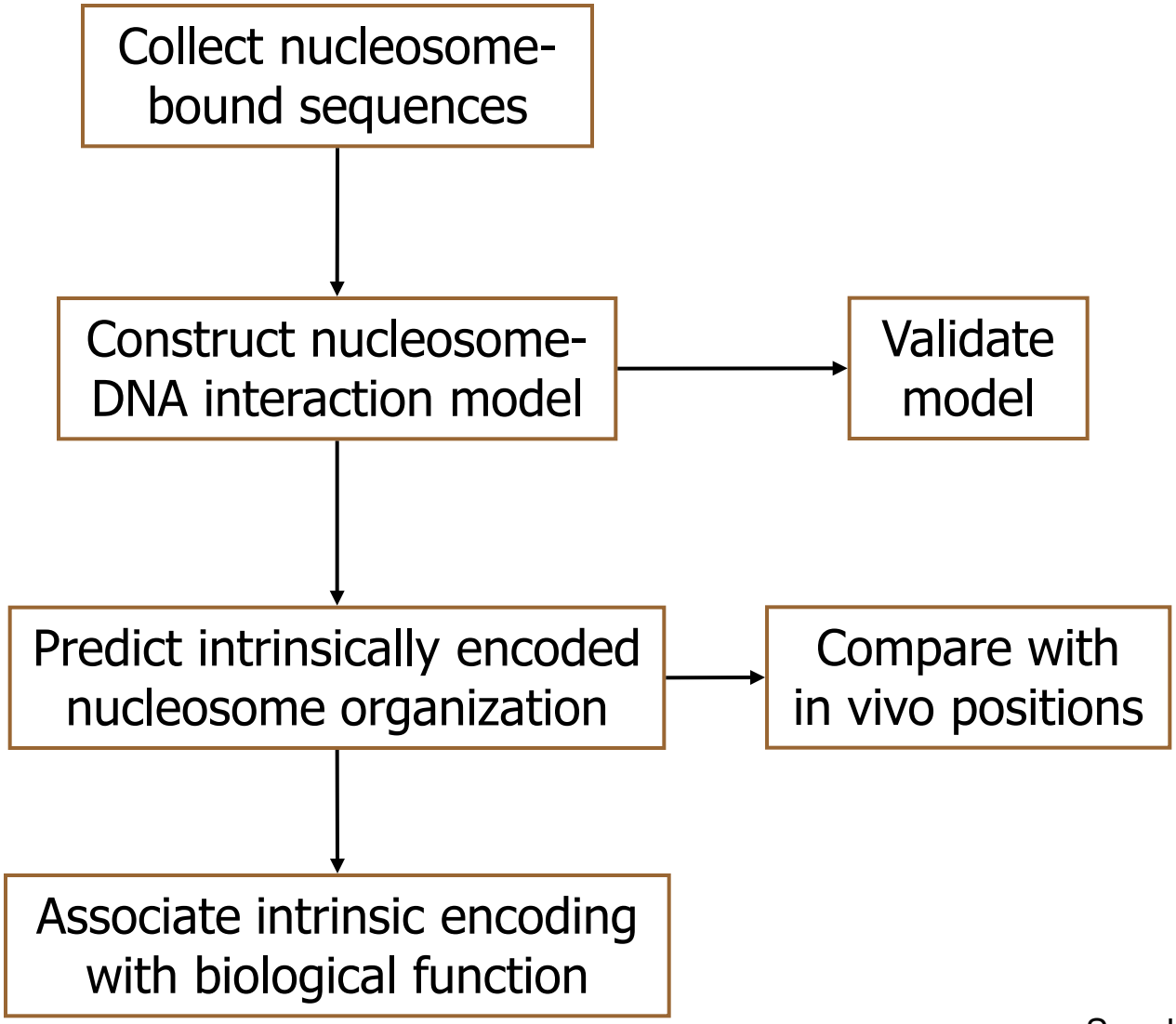
Felsenfeld, G. & Groudine, M. (2003), *Nature* 421: 448-453

~10.2 bp periodicity of AA/TT steps in the *C. elegans* genome



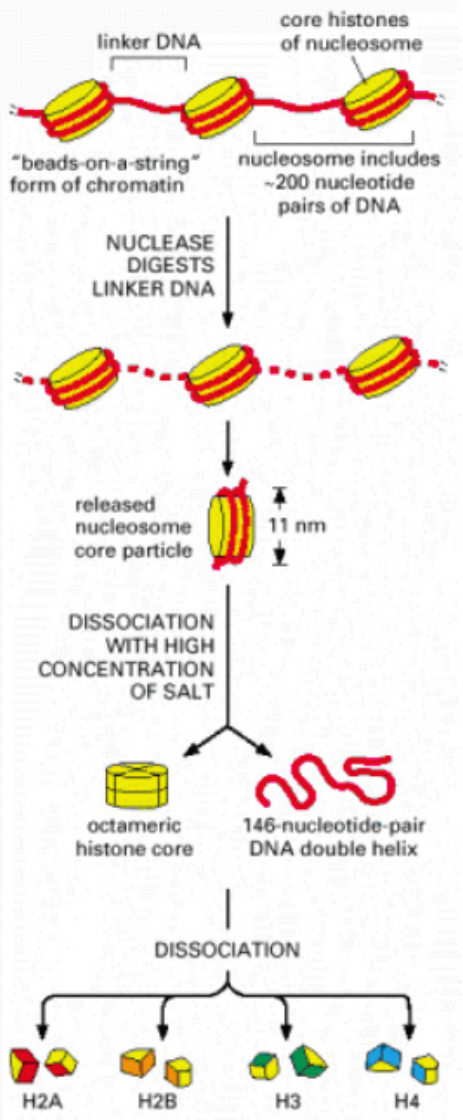
Widom, 1996

Understanding and predicting the genome's nucleosome-forming potential



Segal et al., 2006

Isolation of natural nucleosome core DNA



Alberts et al., 4th ed., Fig. 4–24 (2002)

Center alignment of yeast nucleosome DNAs

ACGTAGCTGTAGTGTACTGACGTACGTCGTC

ACTAGCTGATACGGAGACCCGCGCGATTTTGCGGTC

ACTGTTCGTCGTGTGTGTGTGCTGCTGTAGACTTGTGTG

TACTGTTTTTATTTTGCGGGCATGCTTGT

Center align

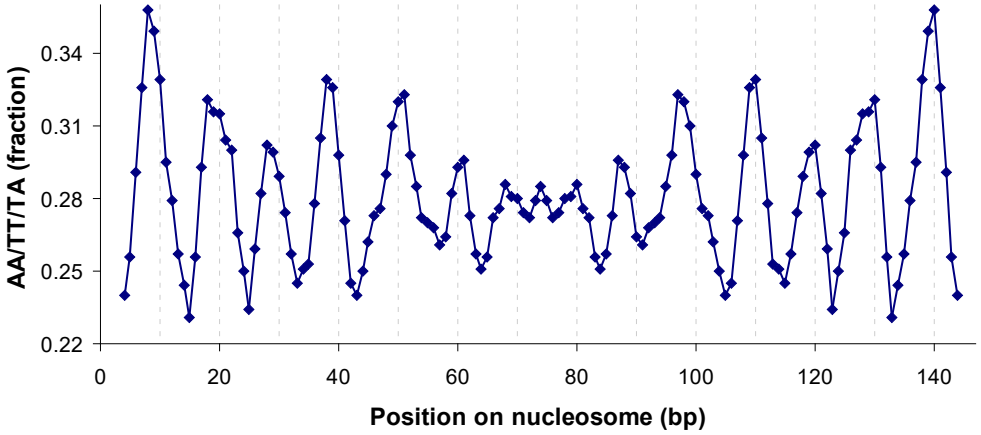
ACGTAGCTGTAGTGTACTGACGTACGTCGTC

ACTAGCTGATACGGAGACCCGCGCGATTTTGCGGTC

ACTGTTCGTCGTGTGTGTGTGCTGCTGTAGACTTGTGTG

TACTGTTTTTATTTTGCGGGCATGCTTGT

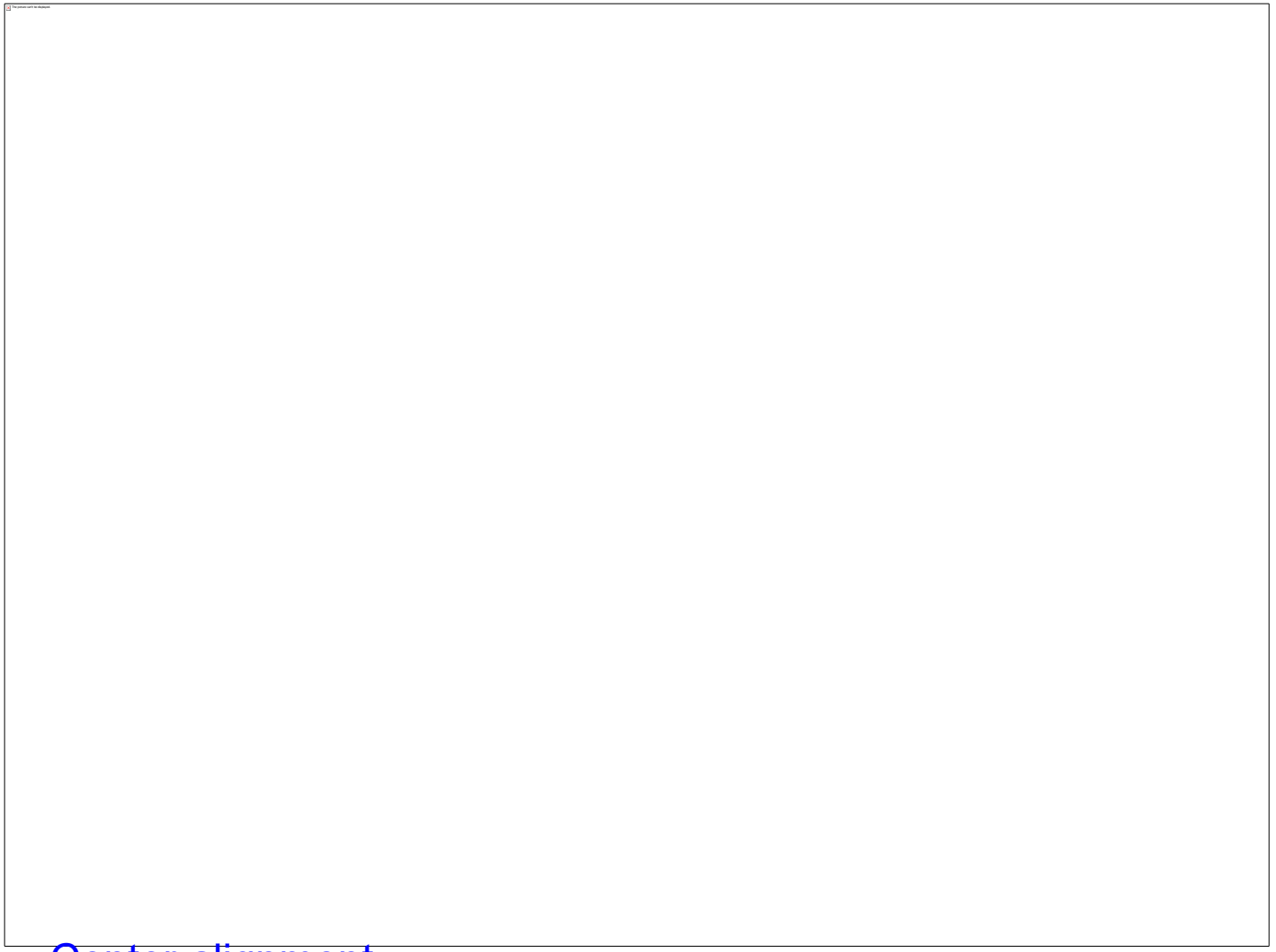
- ~10bp periodicity of AA/TT/TA
- Same period for GC, out of phase with AA/TT/TA
- Signals important for DNA bending
- NO signal from alignment of randomly chosen genomic regions



Yeast (in vivo nucleosomes)

Segal et al., 2006

Location mixture model alignment *vs* center alignment

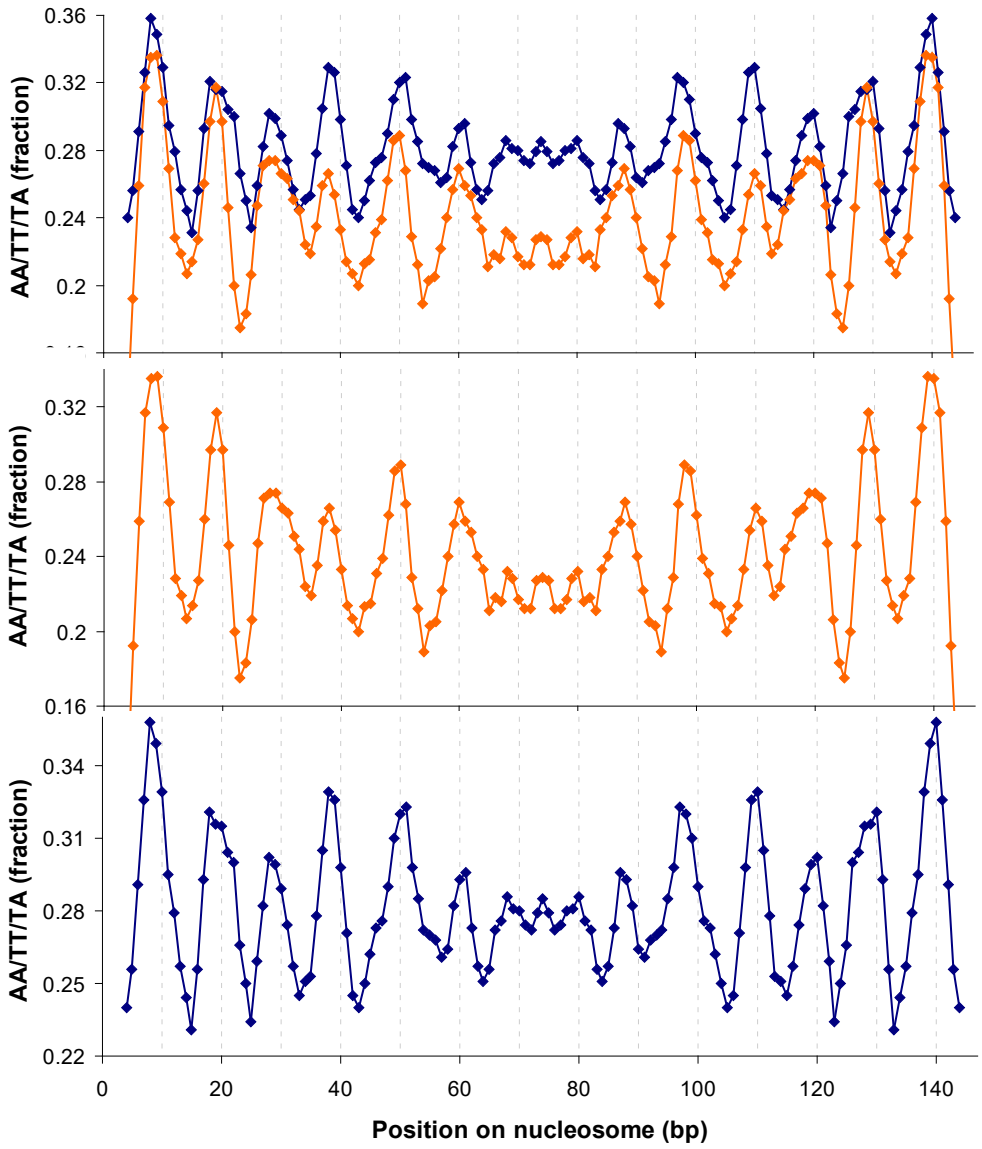


Center alignment

Location mixture model alignment

Wang & Widom, 2005

Center alignment of chicken nucleosome DNAs



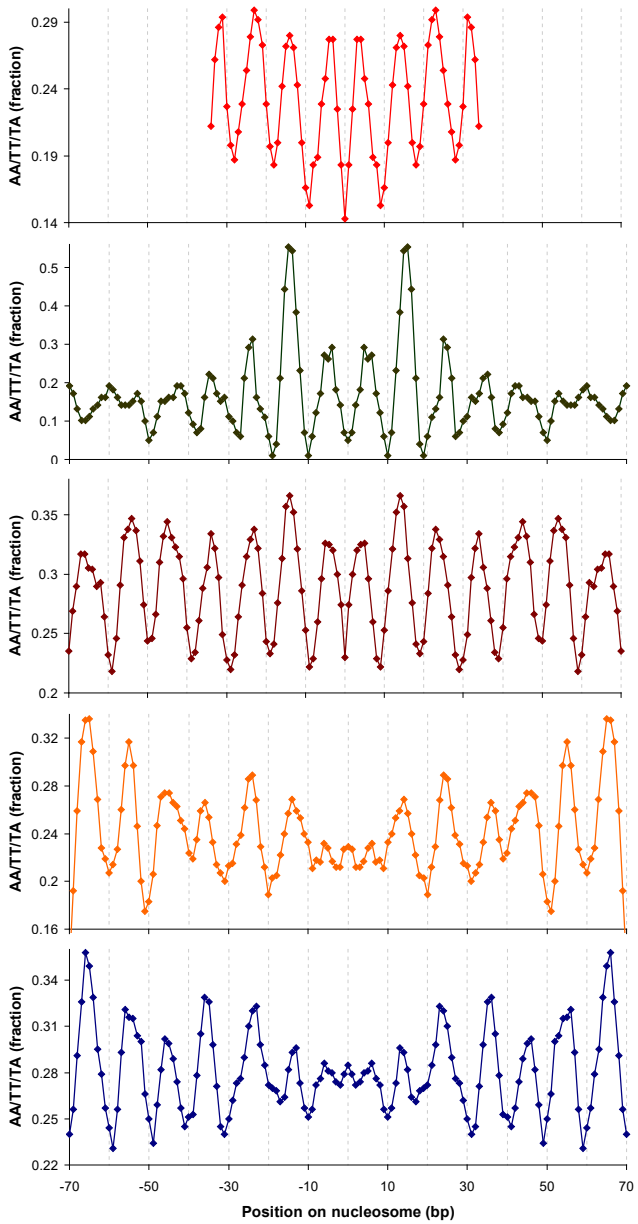
Chicken + Yeast merge

Chicken (in vivo) (Satchwell et al., 1986)

Yeast (in vivo)

Segal et al., 2006

Alignments of nucleosomes selected in vitro



Mouse (in vitro) (Widlund et al., 1997)

Random DNA (in vitro) (Lowary & Widom, 1998)

Yeast (in vitro)

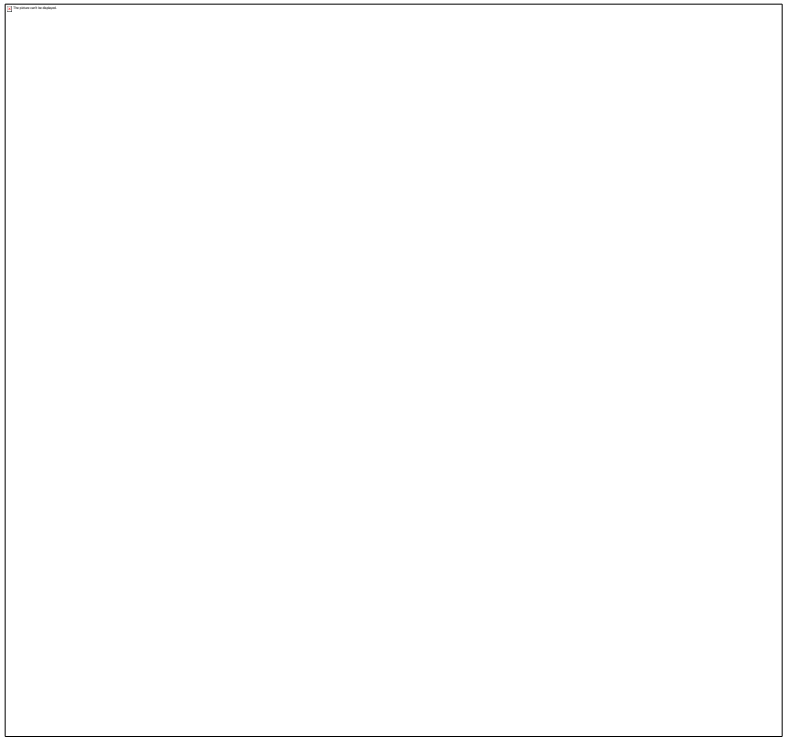
Chicken (in vivo) (Satchwell et al., 1986)

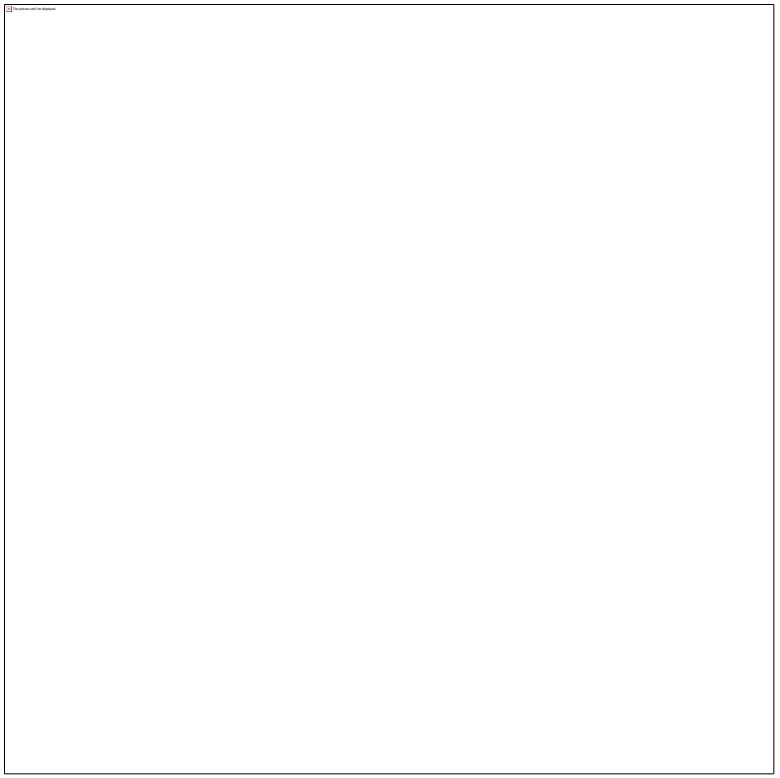
Yeast (in vivo)

Segal et al., 2006

In vitro experimental validation of histone-DNA interaction model

- Adding key motifs increases nucleosome affinity
- Deleting motifs or disrupting their spacing decreases affinity

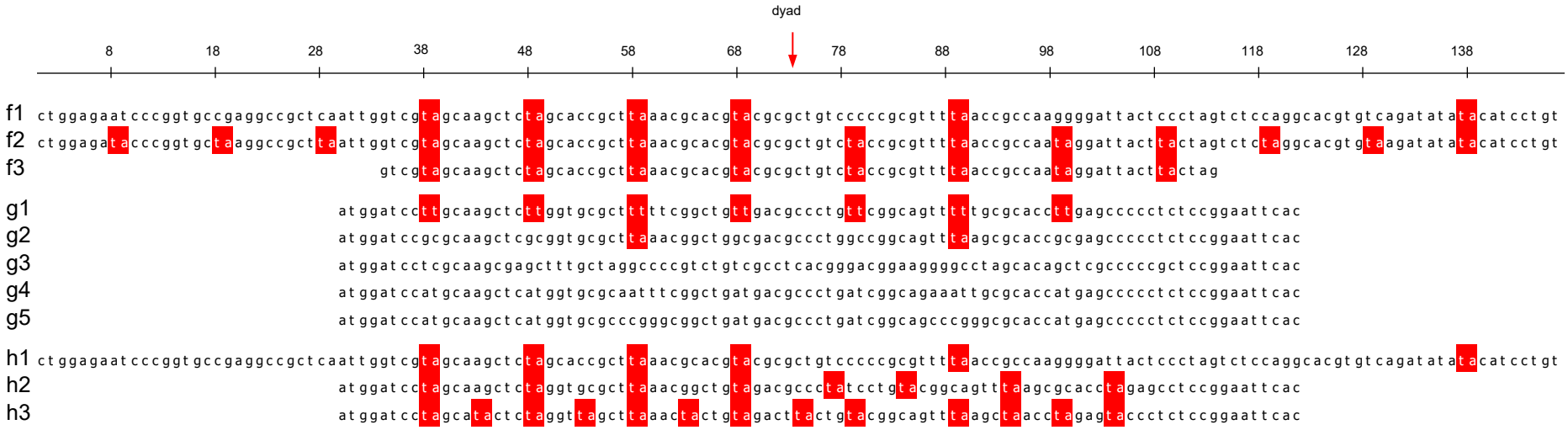


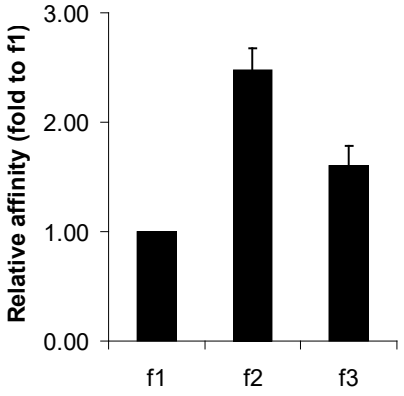


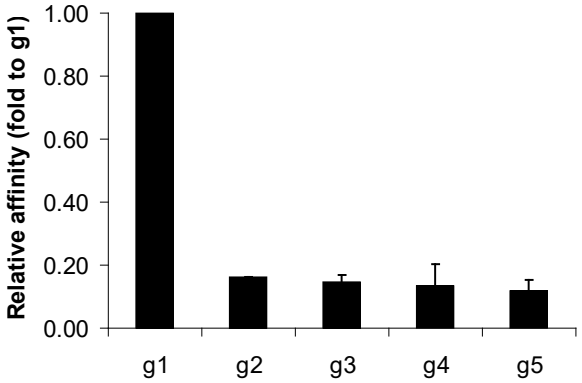
Segal et al., 2006

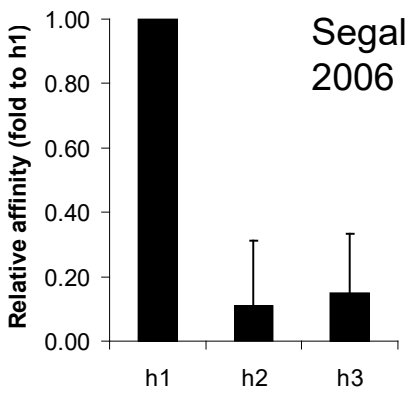
In vitro experimental validation of histone-DNA interaction model

- Adding key motifs increases nucleosome affinity
- Deleting motifs or disrupting their spacing decreases affinity









Segal et al., 2006

Summary

Differing DNA sequences exhibit a > 5,000-fold range of affinities for nucleosome formation

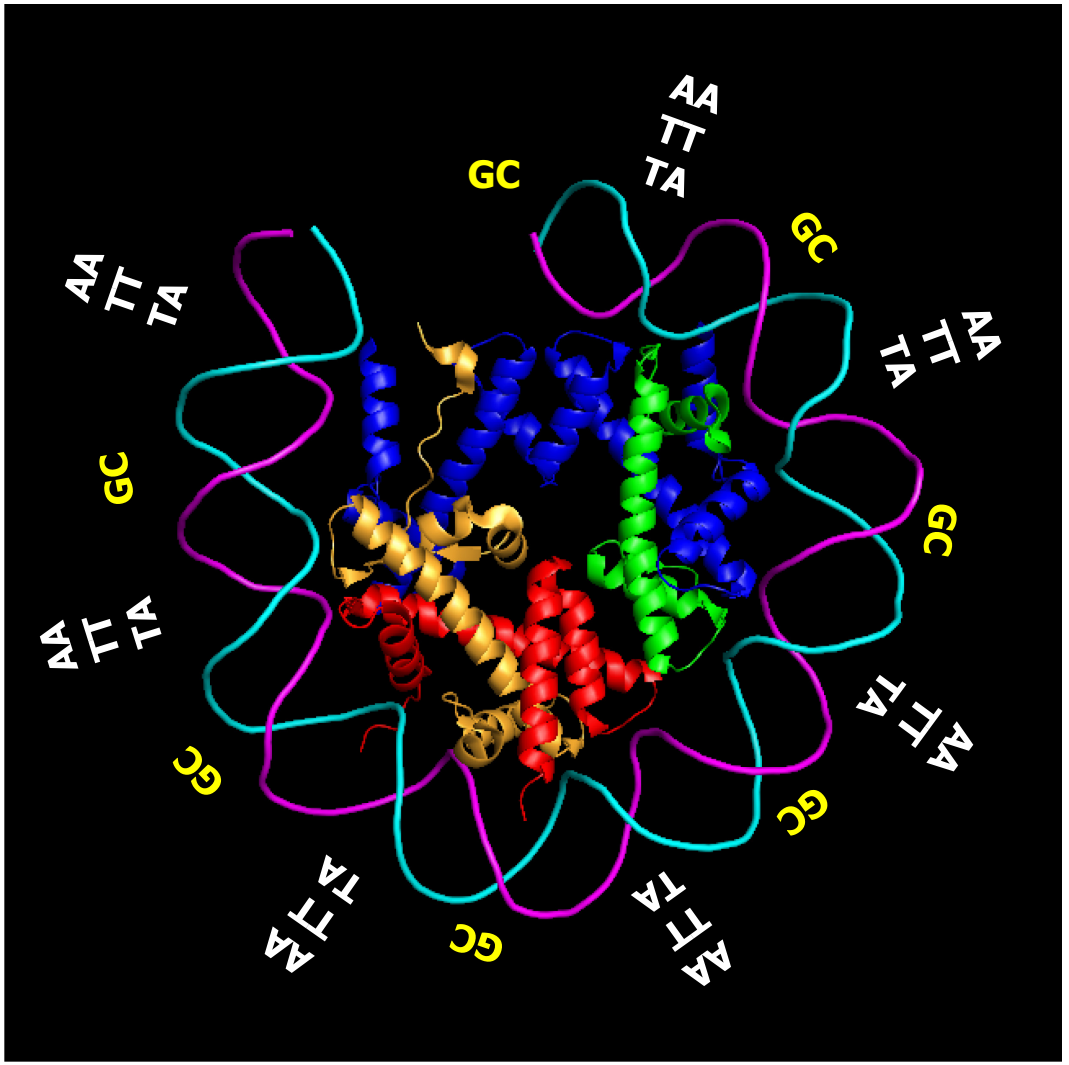
We have a predictive understanding of the DNA sequence motifs that are responsible

Nucleosome–DNA model

Differences from TF model

- Signal is weak
- Information in dinucleotides
- Two-fold symmetry axis

- Position specific dinucleotide distributions
- Add reverse complement of each sequence
- Local (3 bp) averaging



Segal et al., 2006

Summary

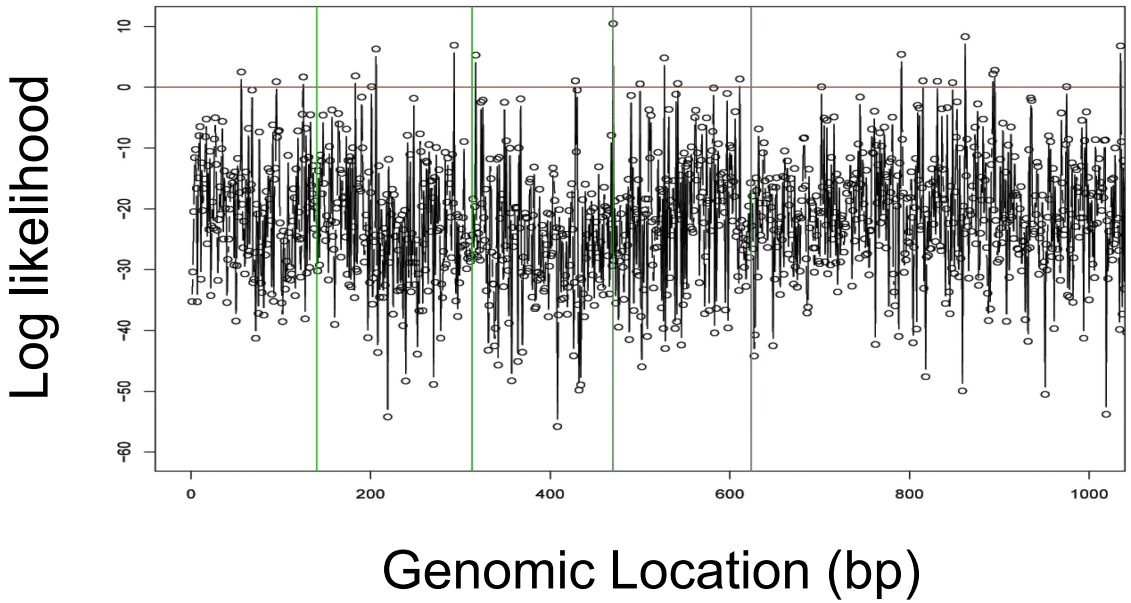
Differing DNA sequences exhibit a > 5,000-fold range of affinities for nucleosome formation

We have a predictive understanding of the DNA sequence motifs that are responsible

Sequences matching these motifs are abundant in eukaryotic genomes, and are occupied by nucleosomes in vivo

Placing nucleosomes on the genome

A free energy landscape, not just scores and a threshold !!

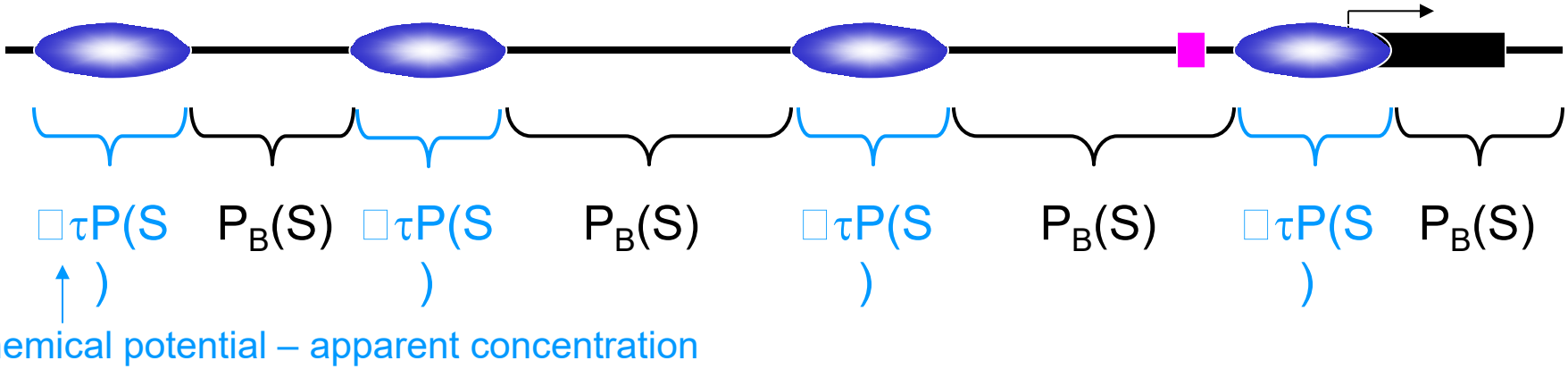


- Nucleosomes occupy 147 bp and exclude 157 bp

Segal et al., 2006

Equilibrium configurations of nucleosomes on the genome

- One of *very many* possible configurations



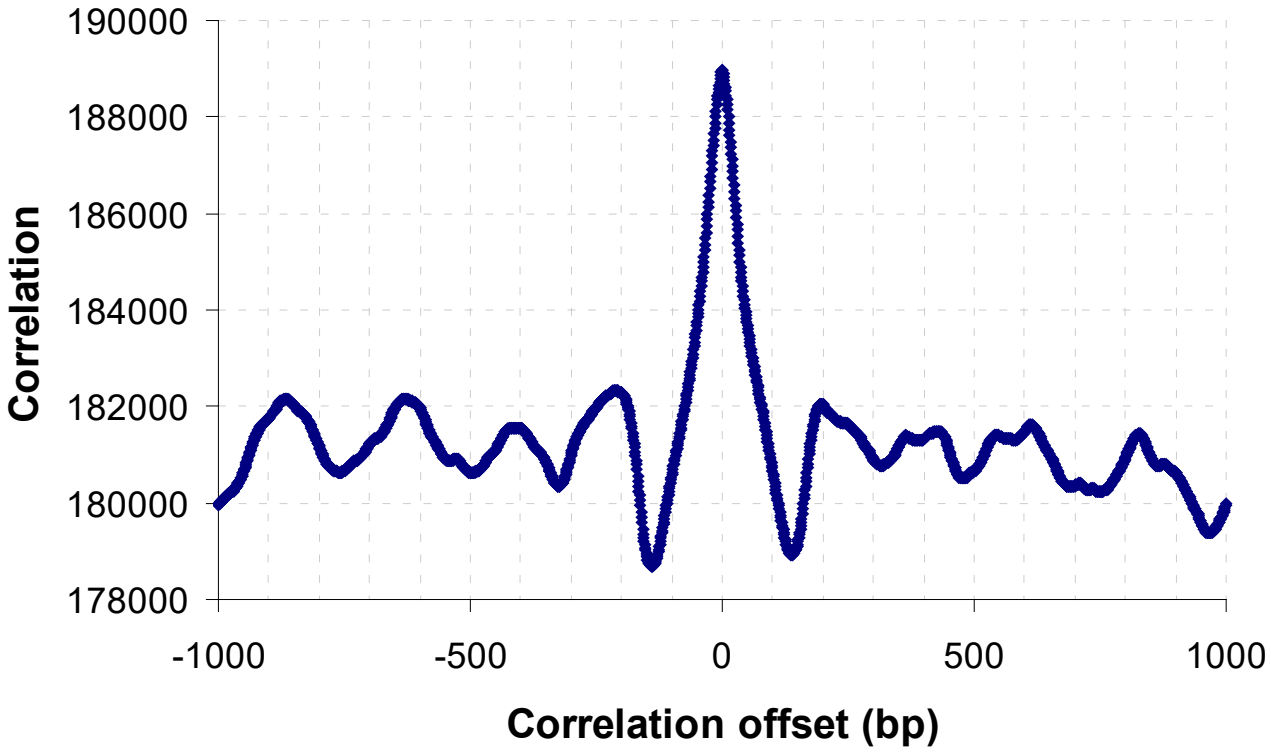
Probability of placing a nucleosome starting at each allowed basepair *i* of S

Probability of *any* nucleosome covering position *i* (≡ average occupancy)

Locations *i* with high probability for starting a nucleosome (≡ stable nucleosomes)

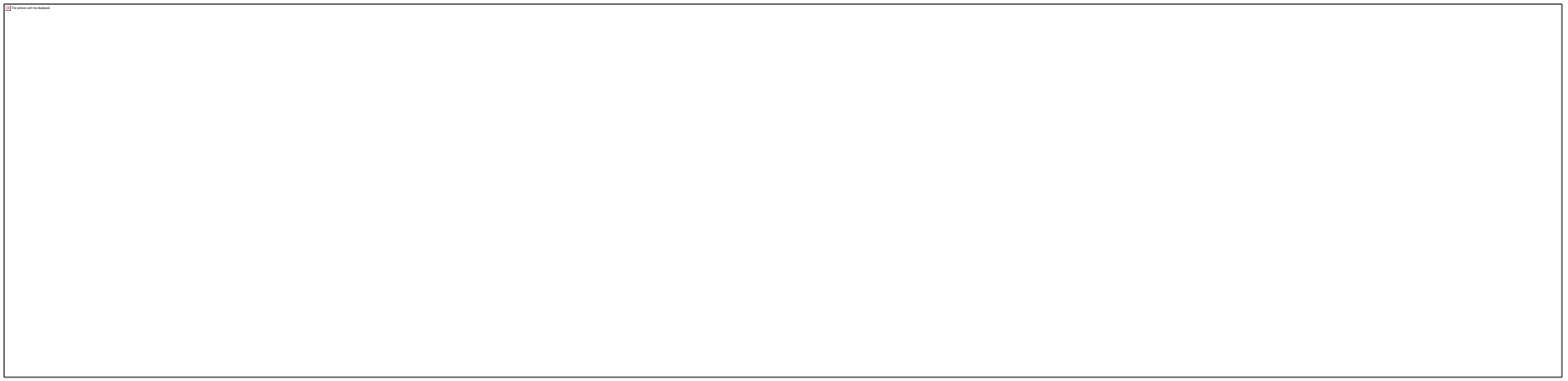
Segal et al., 2006

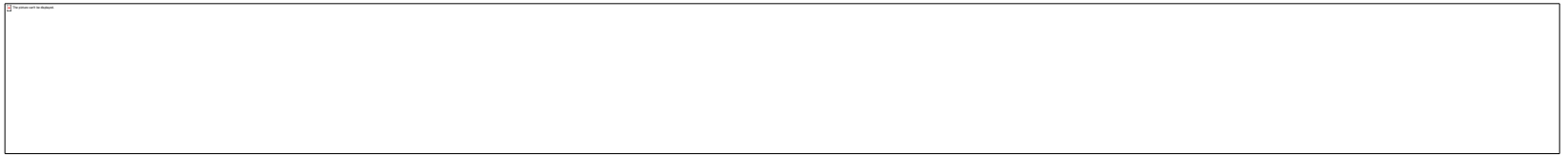
Cross-correlation of average occupancies predicted using yeast and chicken models



Segal et al., 2006

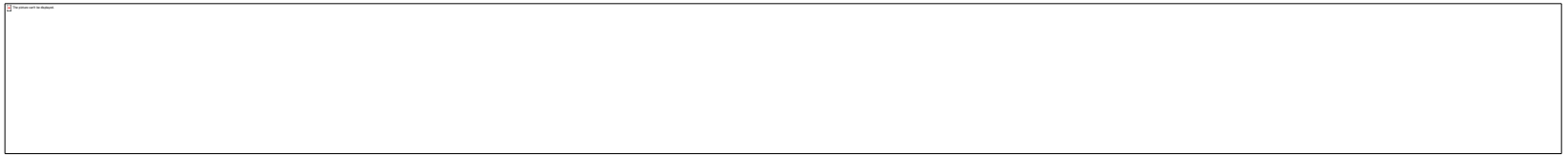
Nucleosome coding potential at the *GAL1–10* locus: predicted distribution compared to experimental





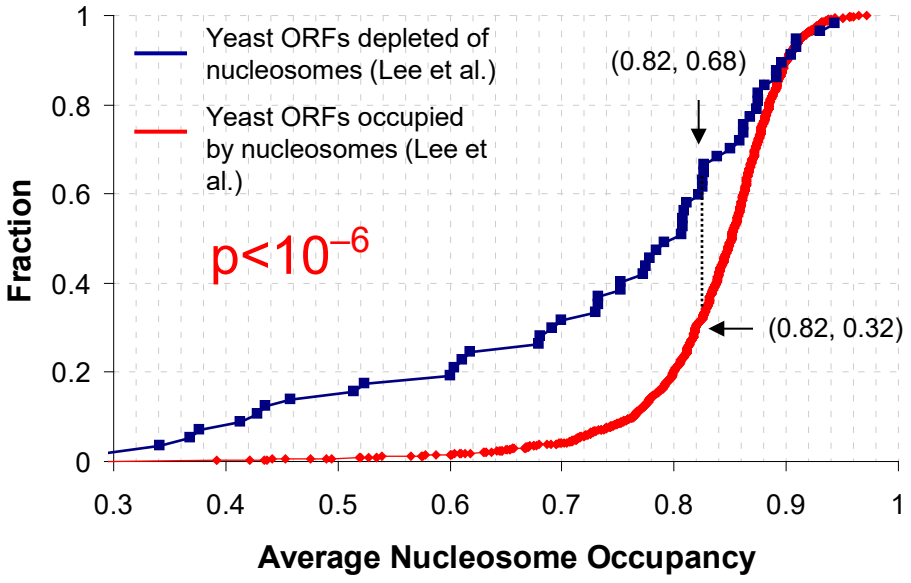
Segal et al., 2006

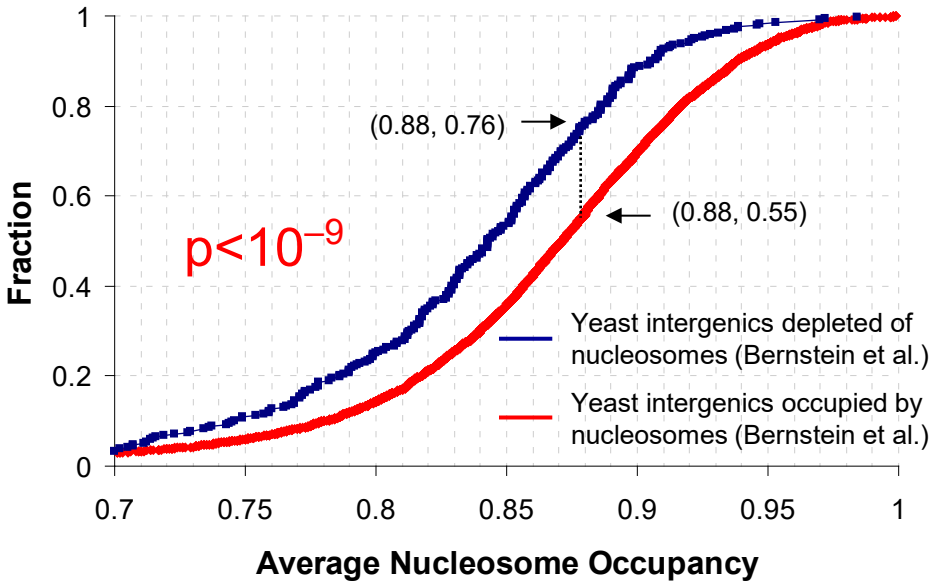
Nucleosome coding potential at the *CHA1* locus



Segal et al., 2006

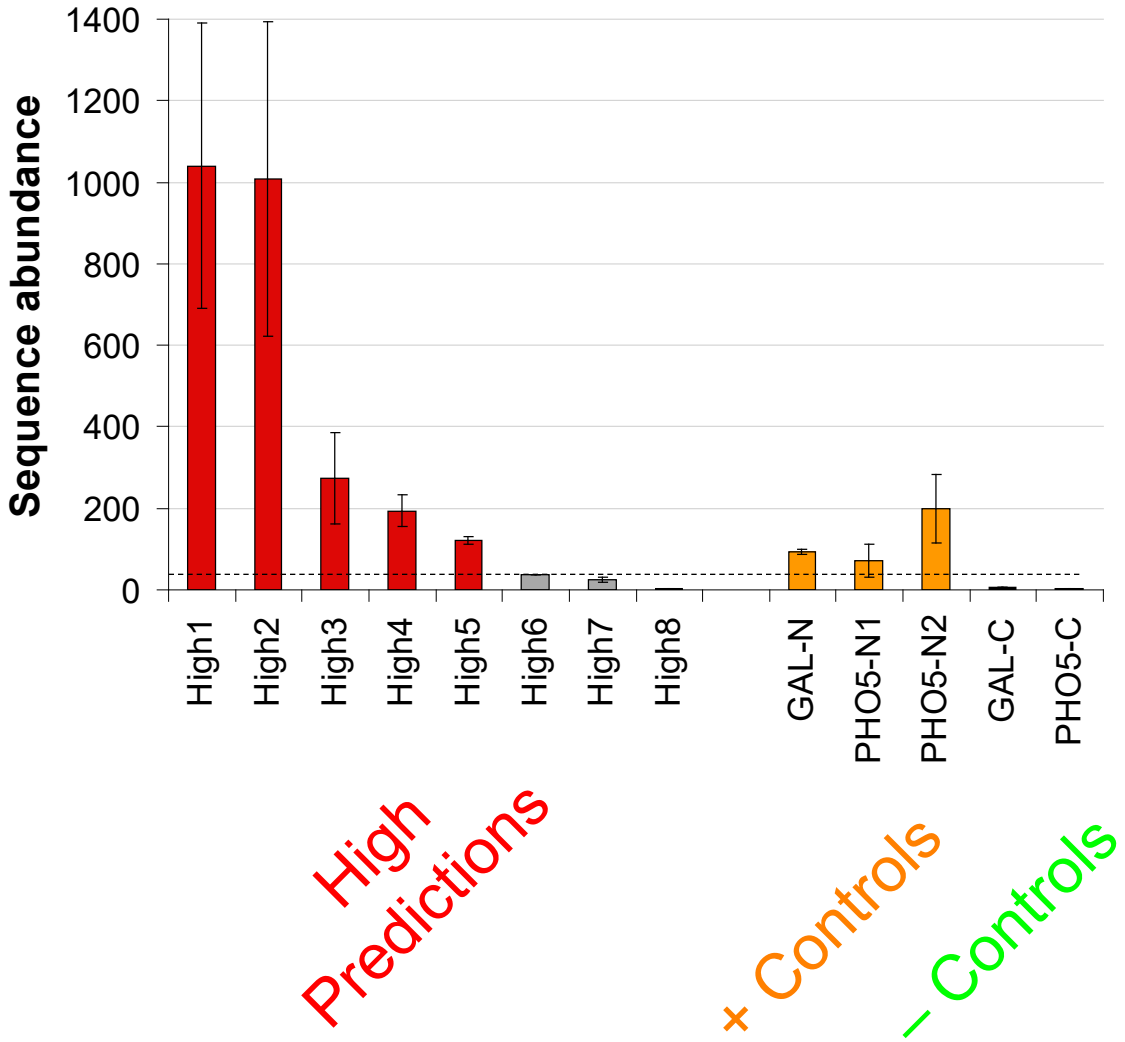
Reliable classification of nucleosome occupancies on depleted *vs* occupied regions





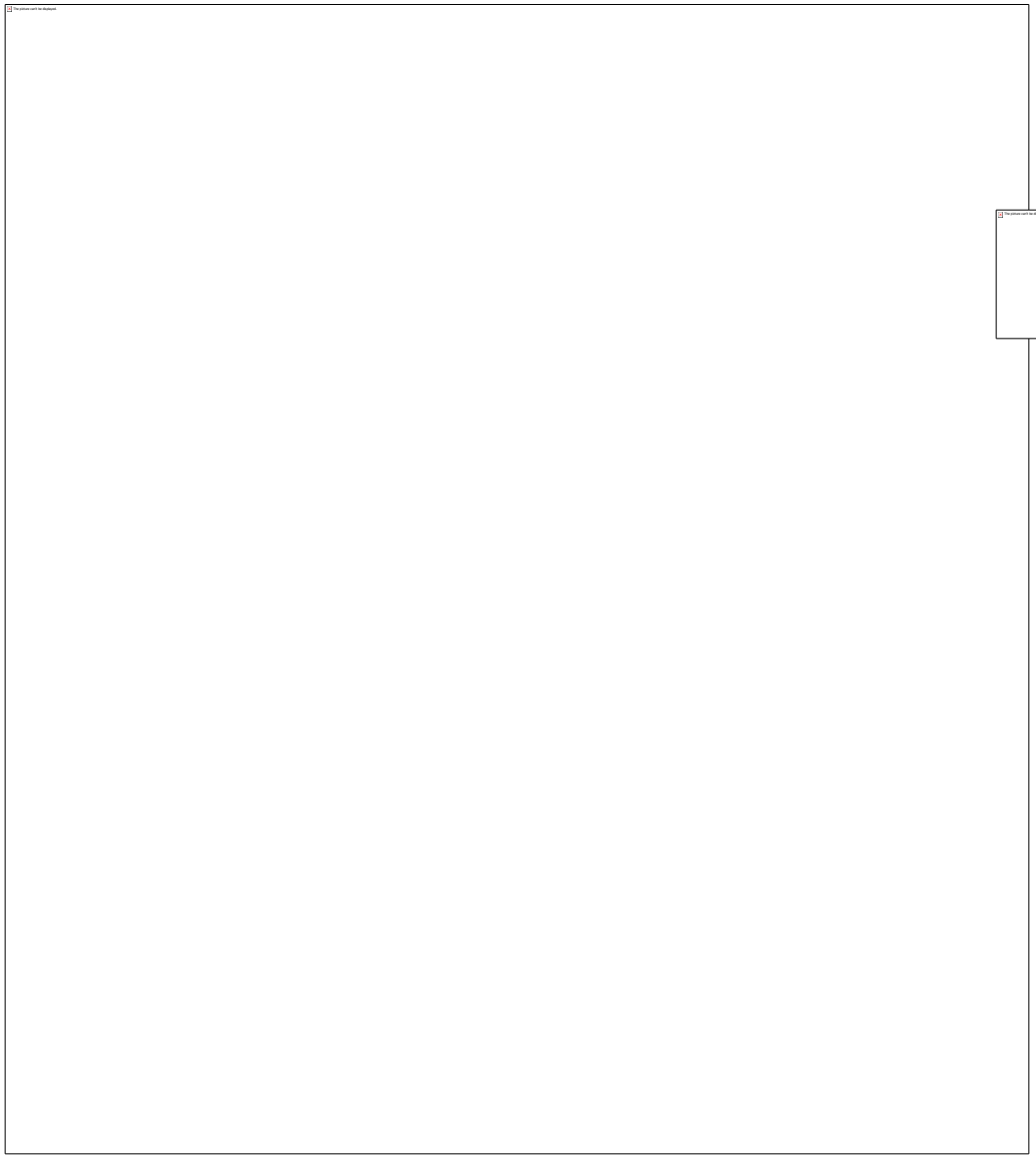
Segal et al., 2006

In vivo nucleosome occupancies at predicted high-occupancy locations



Segal et al., 2006

In vivo nucleosome occupancies at predicted *low*-occupancy locations



Segal et al., 2006

Summary

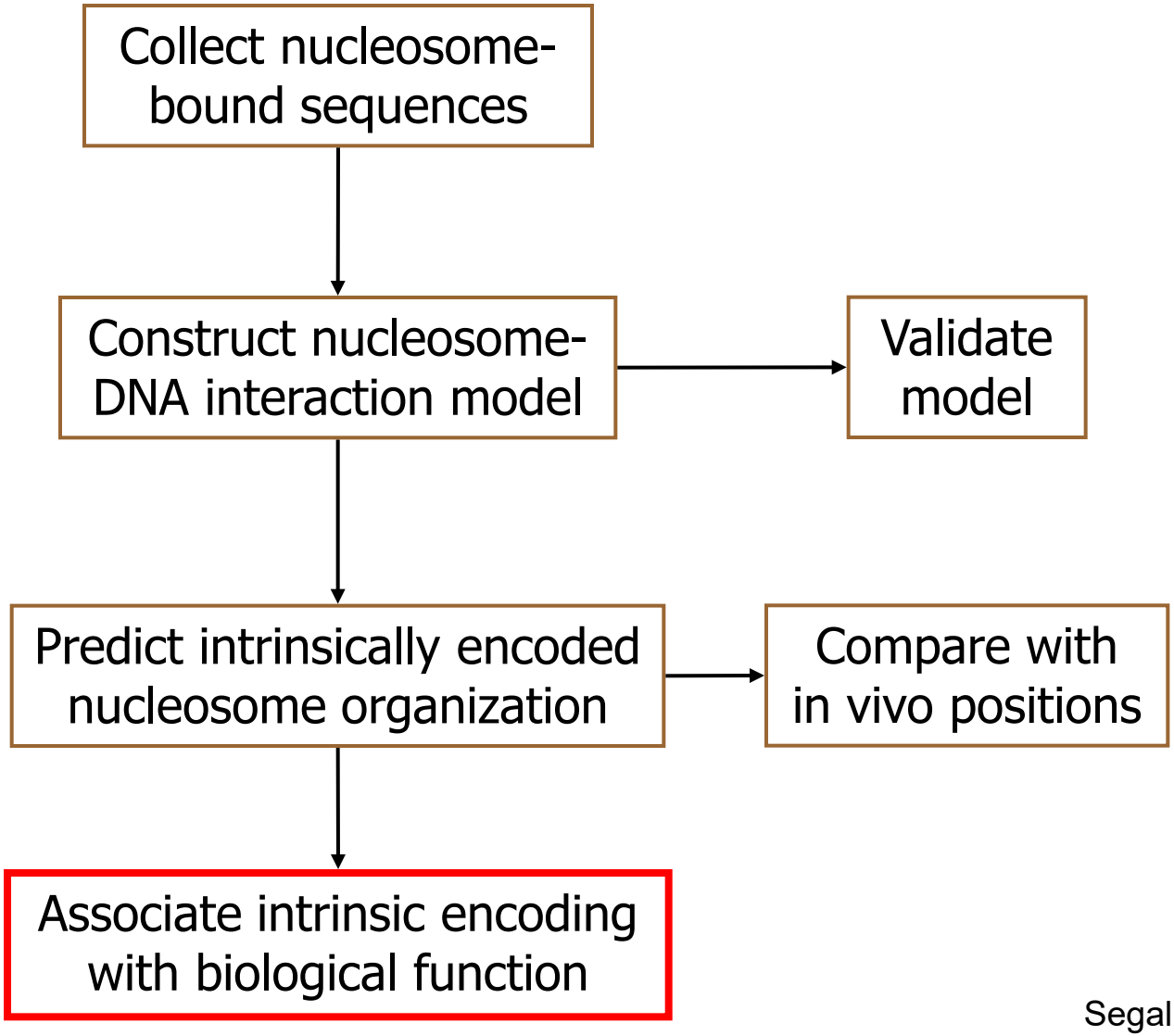
Differing DNA sequences exhibit a > 5,000-fold range of affinities for nucleosome formation

We have a predictive understanding of the DNA sequence motifs that are responsible

Sequences matching these motifs are abundant in eukaryotic genomes, and are occupied by nucleosomes in vivo

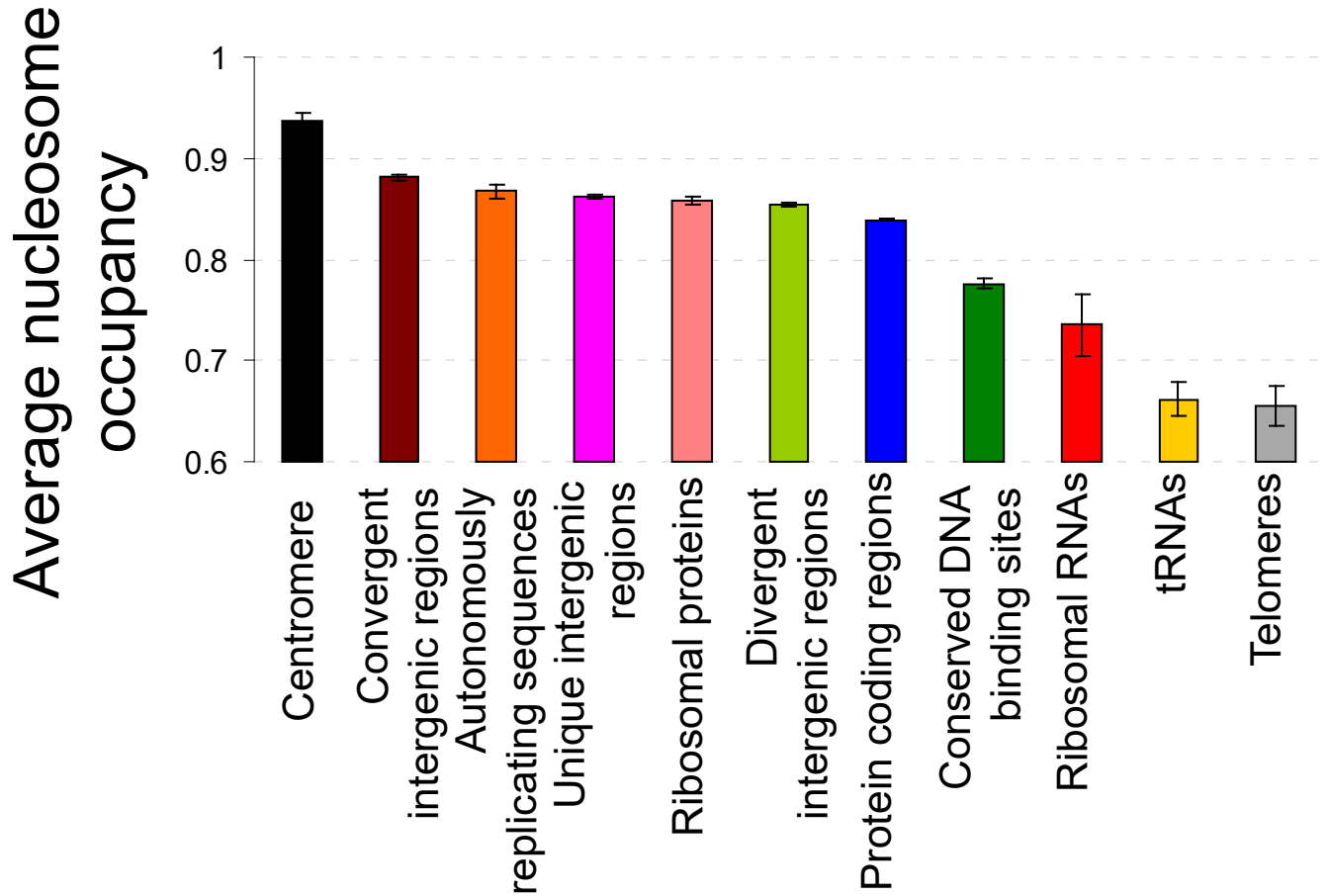
A model based only on these DNA sequence motifs and nucleosome-nucleosome exclusion explains ~50% of in vivo nucleosome positions

Understanding and predicting the genome's nucleosome-forming potential



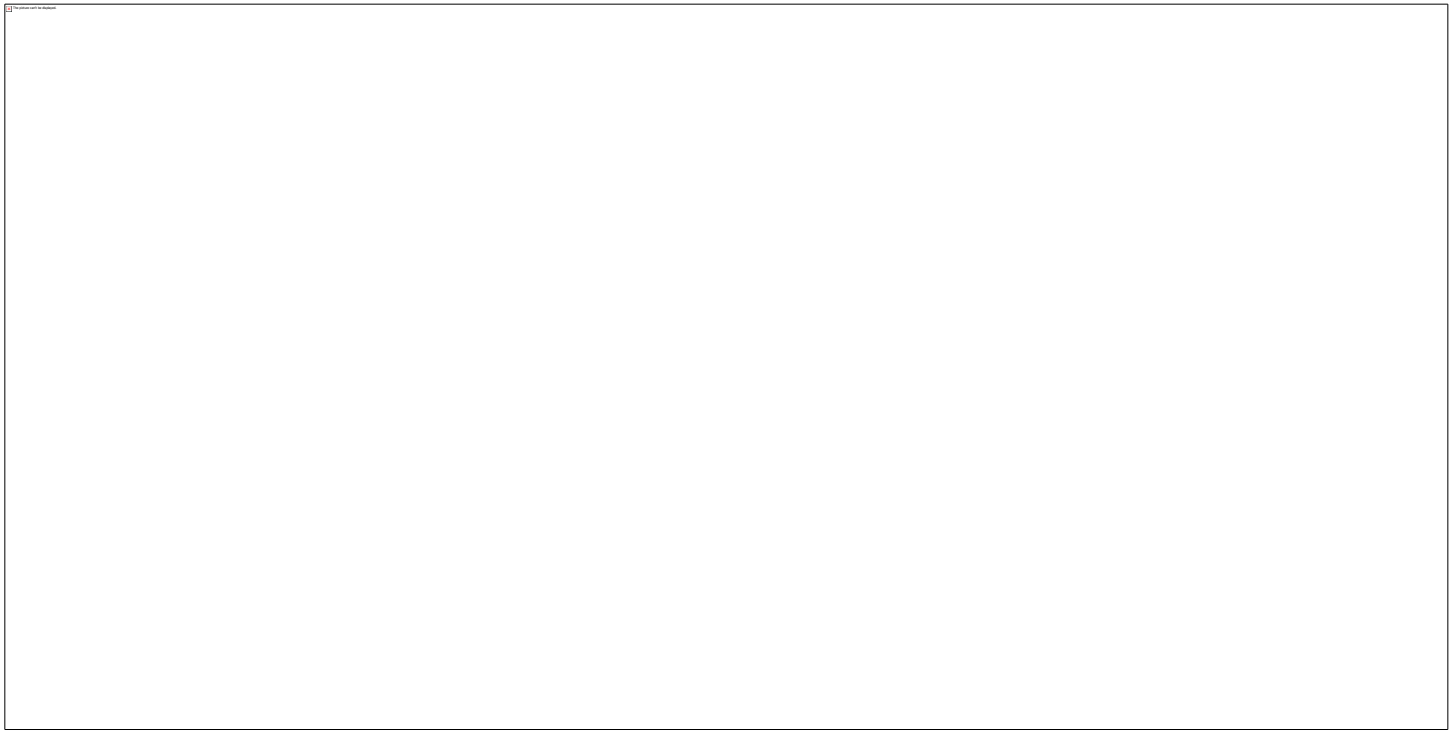
Segal et al., 2006

Nucleosome occupancy varies with chromosome region type



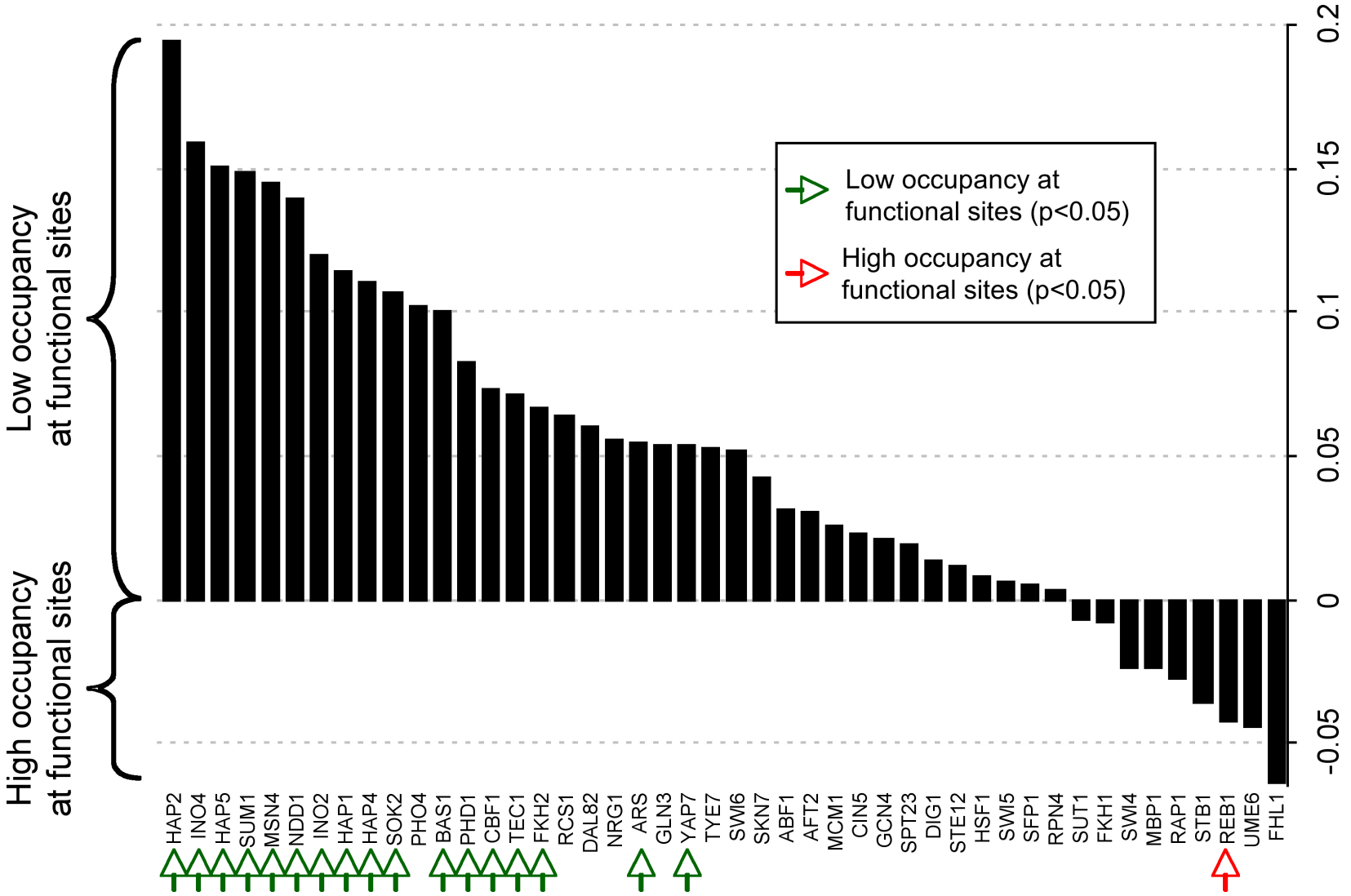
Segal et al., 2006

Does the genome's intrinsic nucleosome organization facilitate occupancy of functional binding sites?



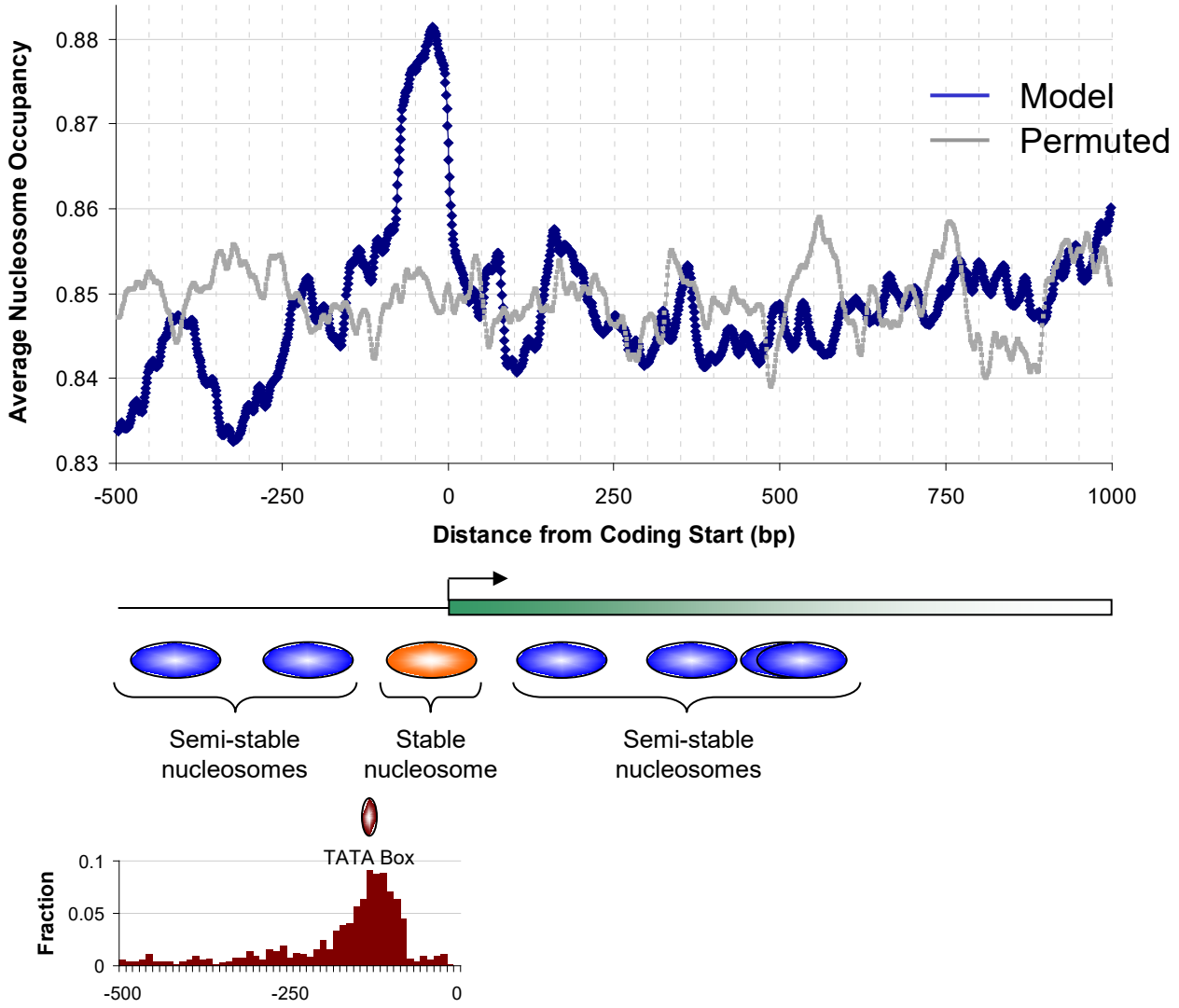
Segal et al., 2006

The yeast genome encodes low nucleosome occupancy over functional binding sites



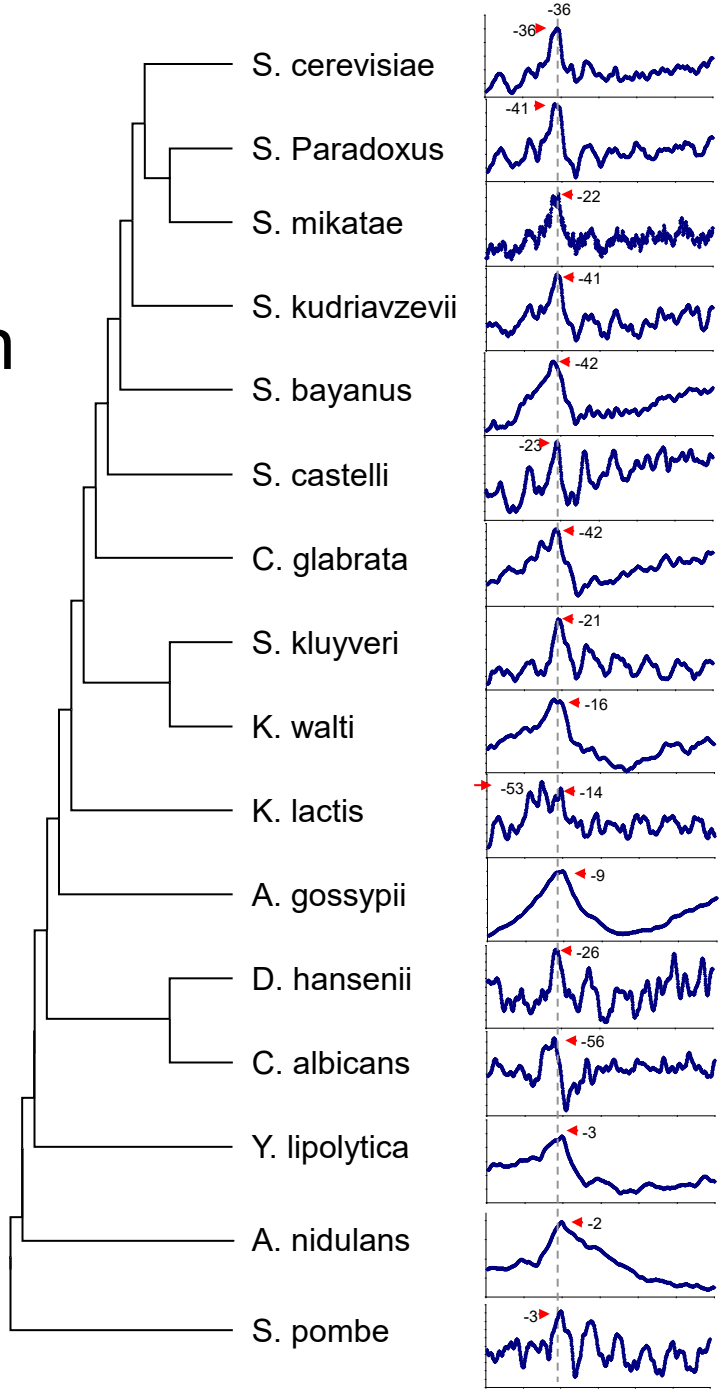
Segal et al., 2006

Distinctive nucleosome occupancy adjacent to TATA elements at yeast promoters



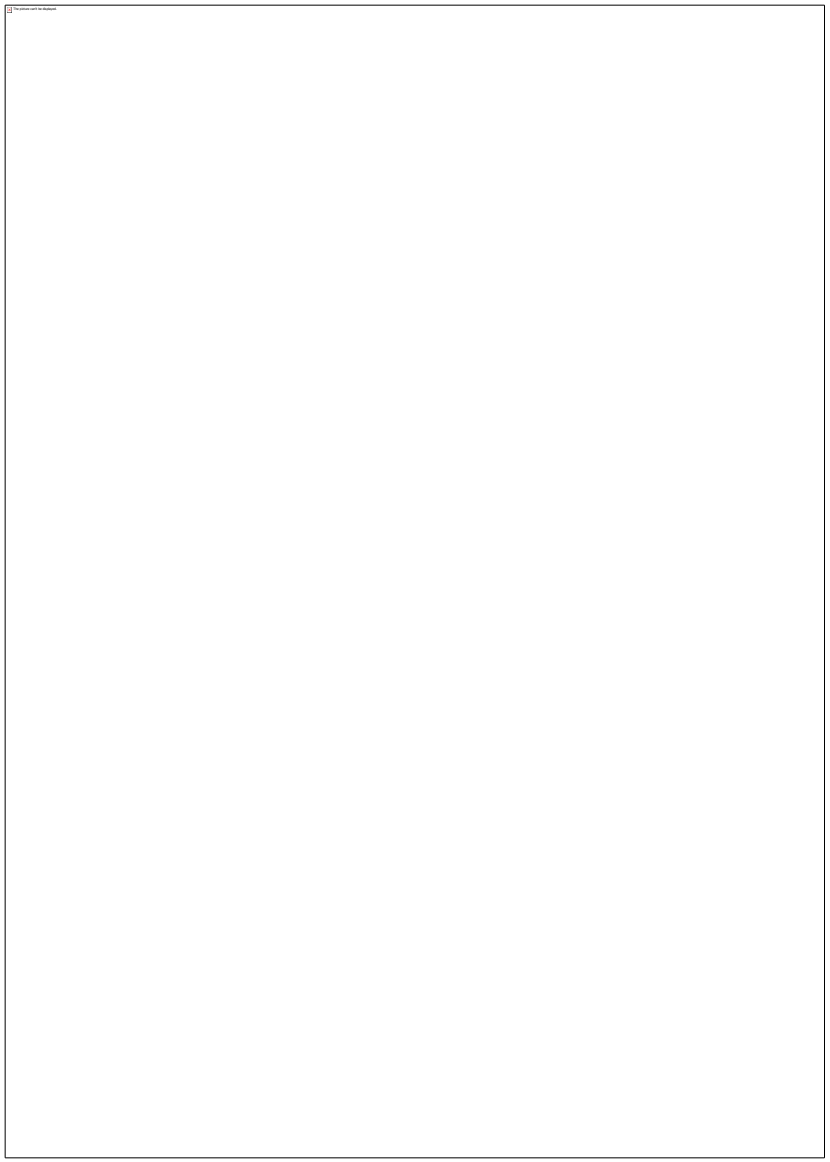
Segal et al., 2006

Nucleosome organization near 5’ ends of genes is conserved through evolution



Segal et al., 2006

Predicted nucleosome organization near 5’ ends of genes – comparison to experiment



Segal et al., 2006

Summary

Differing DNA sequences exhibit a > 5,000-fold range of affinities for nucleosome formation

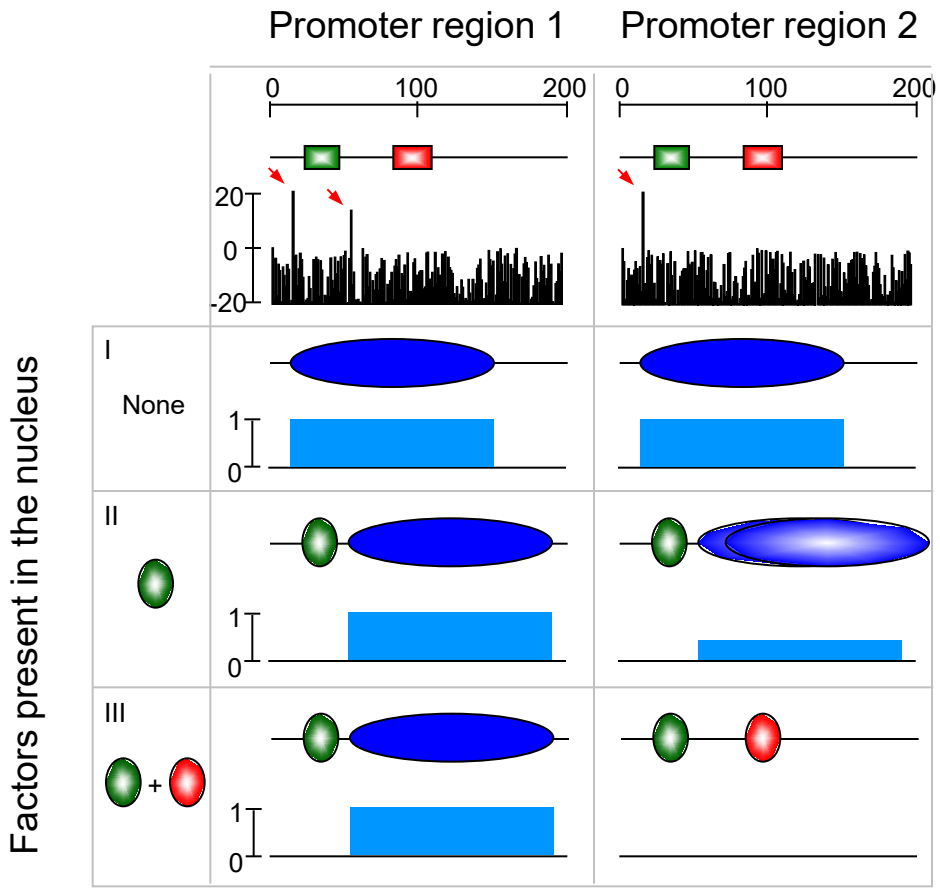
We have a predictive understanding of the DNA sequence motifs that are responsible

Sequences matching these motifs are abundant in eukaryotic genomes, and are occupied by nucleosomes in vivo

A model based only on these DNA sequence motifs and nucleosome-nucleosome exclusion explains ~50% of in vivo nucleosome positions

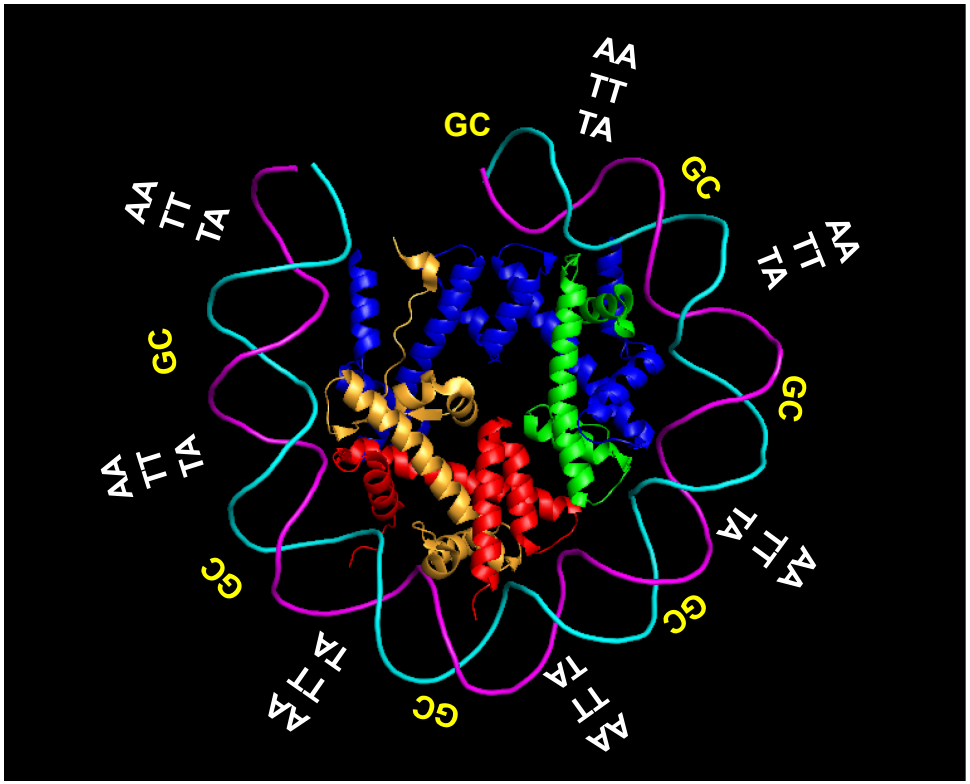
These intrinsically encoded nucleosome positions are correlated with, and may facilitate, essential aspects of chromosome structure and function

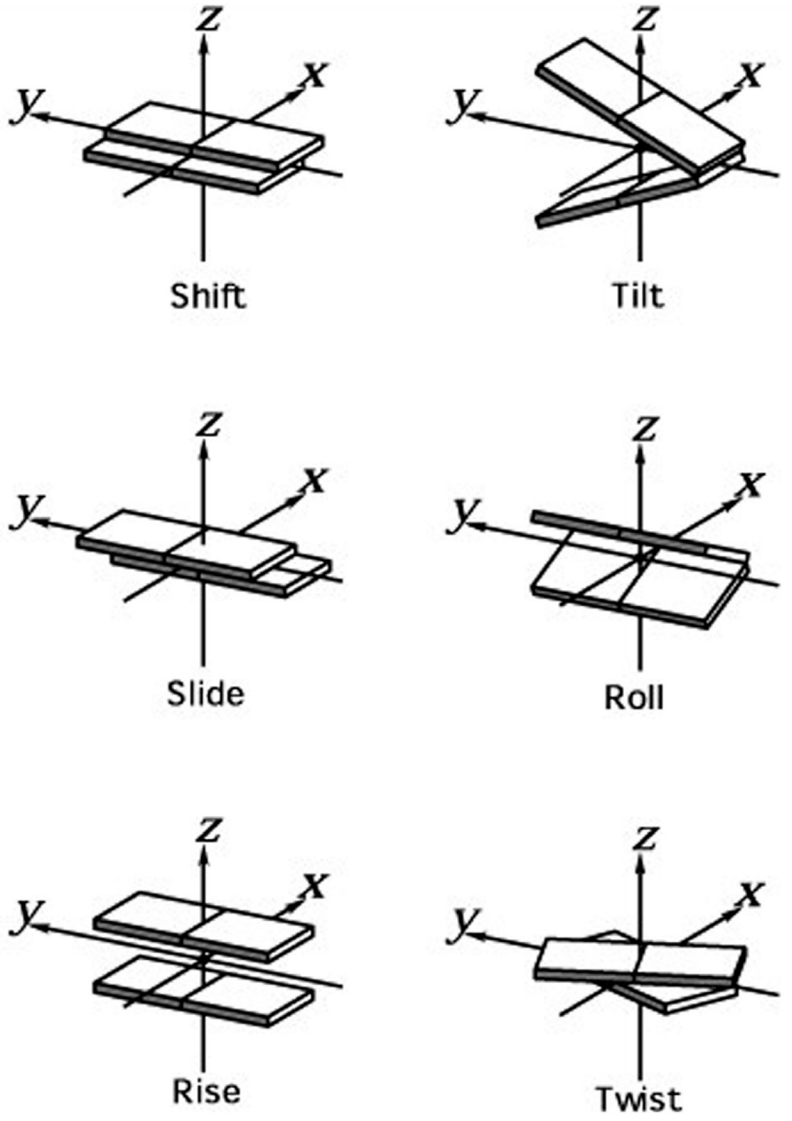
Genome sequence influences the competition between nucleosomes and DNA binding proteins



Segal et al., 2006

Toward a proper free energy model for the sequence-dependent cost of DNA wrapping





Morozov, Segal, Widom, & Siggia

Acknowledgements

Nucleosome positioning in vivo

Eran Segal (Weizmann Inst.)
Yvonne Fondufe-Mittendorf
Lingyi Chen
Annchristine Thåström
Yair Field (Weizmann Inst.)
Irene Moore
Jiping Wang (Northwestern U. Statistics)

Alexandre Morozov (Rockefeller U.)
Eric Siggia (Rockefeller U.)