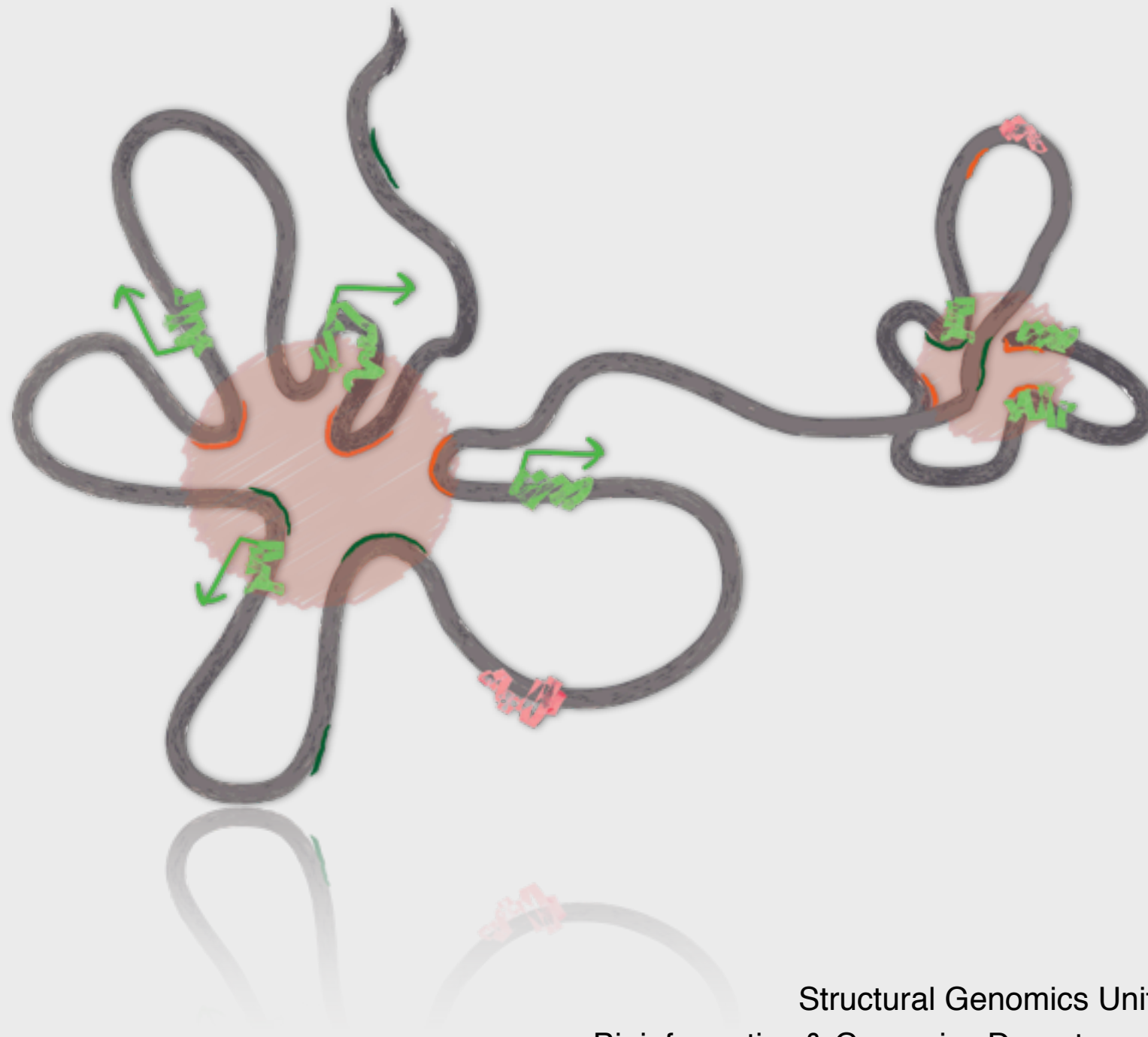


# 3D folding of chromosomal domains in relation to gene expression



Marc A. Marti-Renom

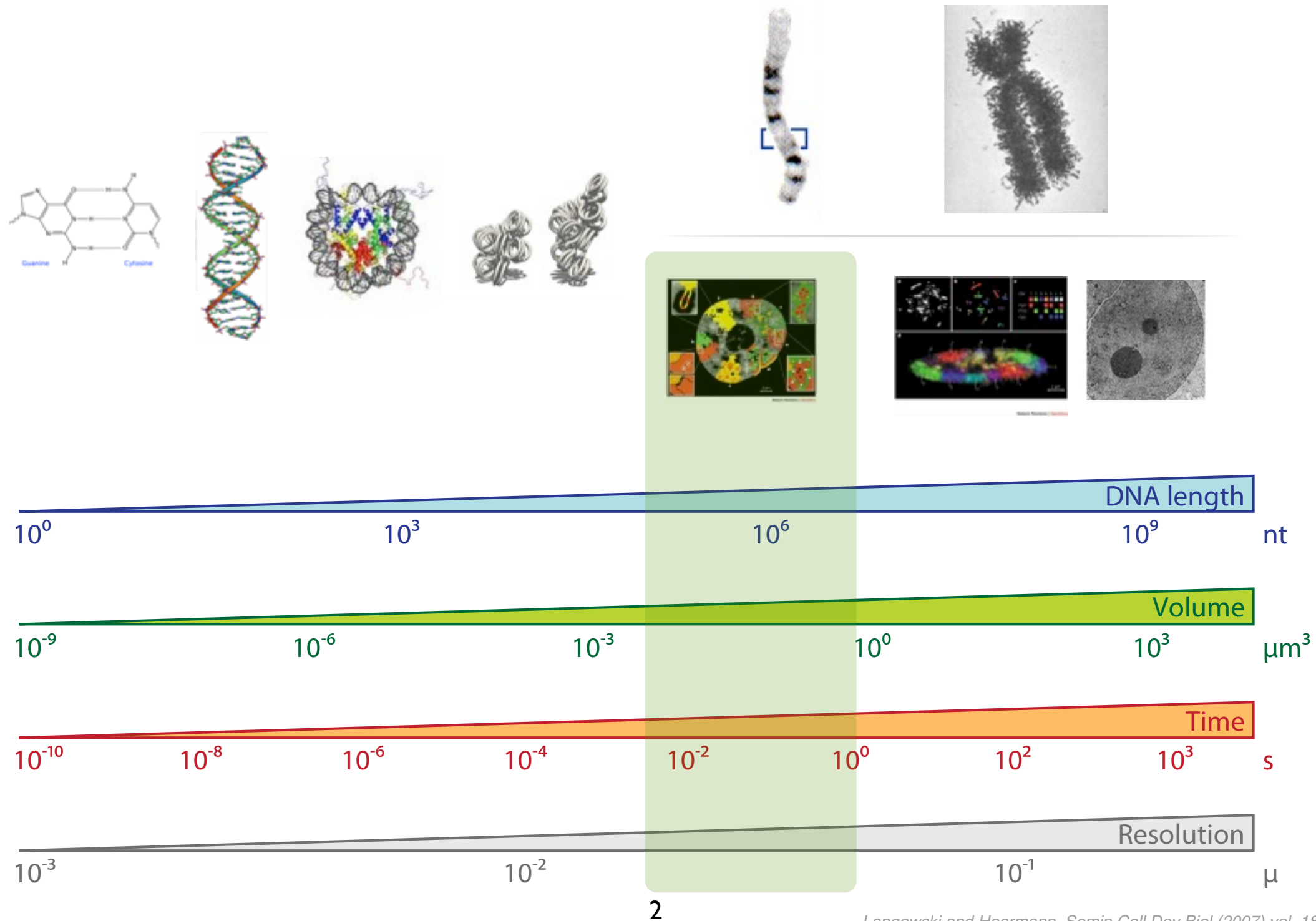
<http://sgu.bioinfo.cipf.es>



# Resolution

Limited knowledge...

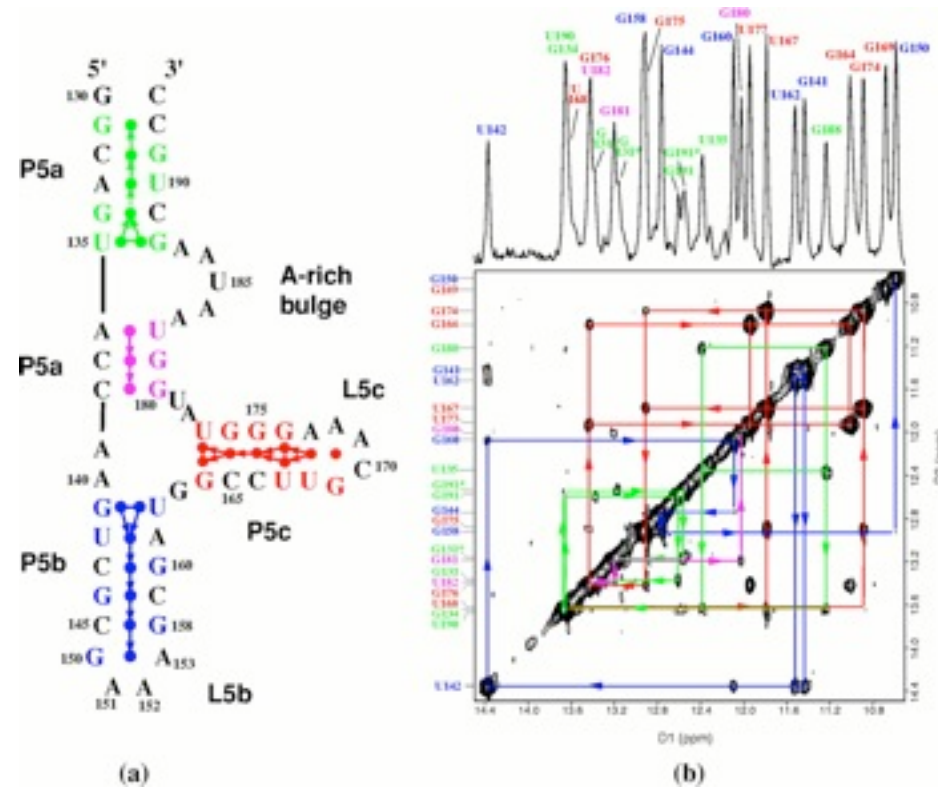
Knowledge



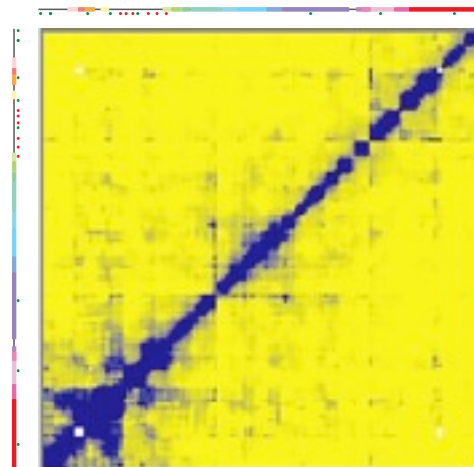
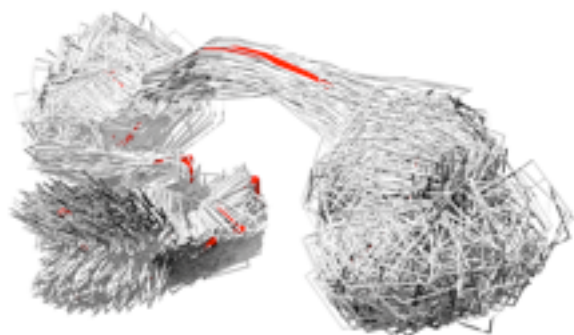
Adapted from:  
Langowski and Heermann. *Semin Cell Dev Biol* (2007) vol. 18 (5) pp. 659-67

# Structure determination

Integrative Modeling Platform  
<http://www.integrativemodeling.org>



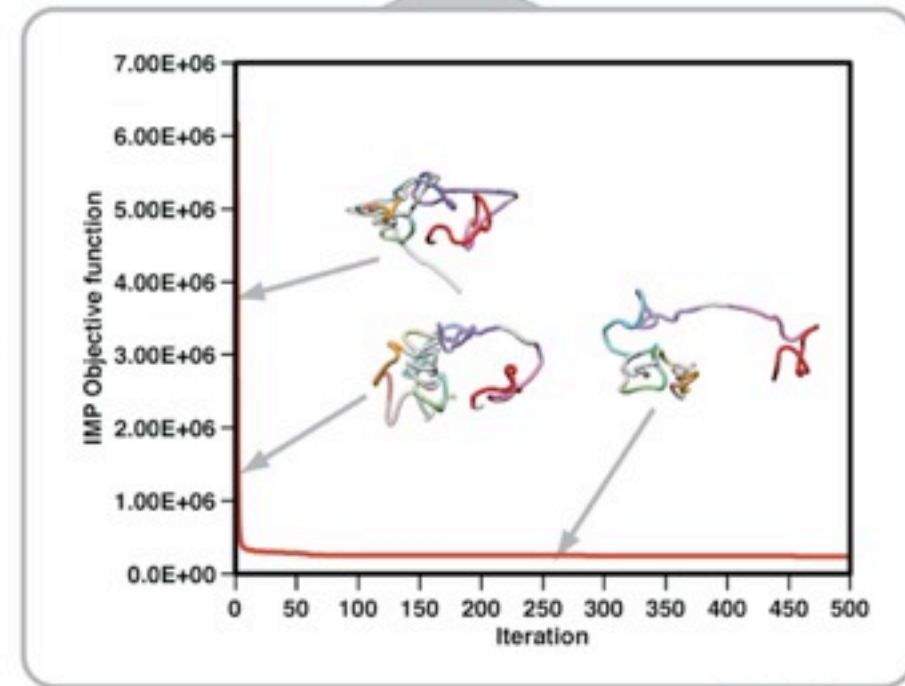
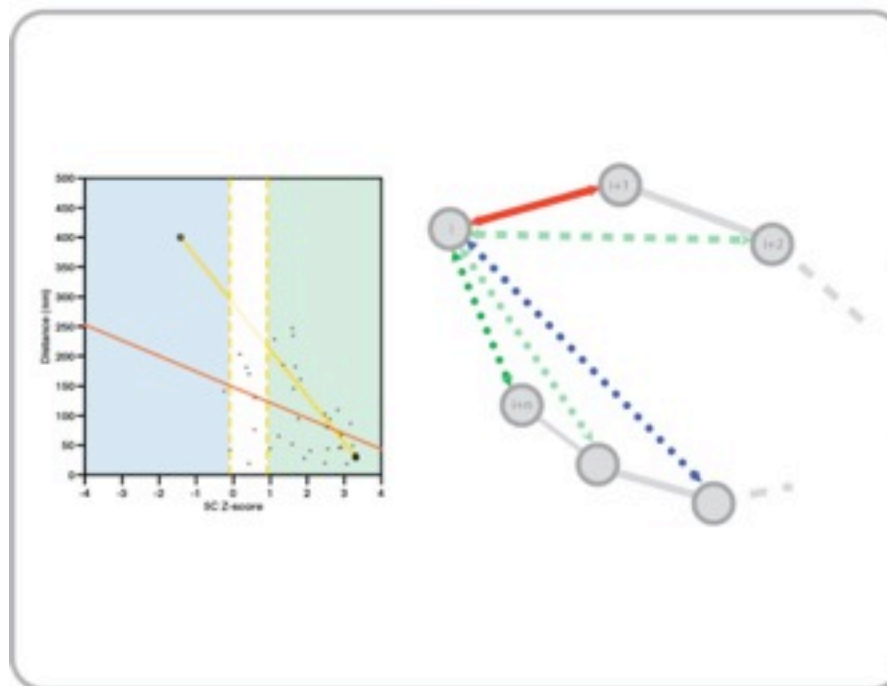
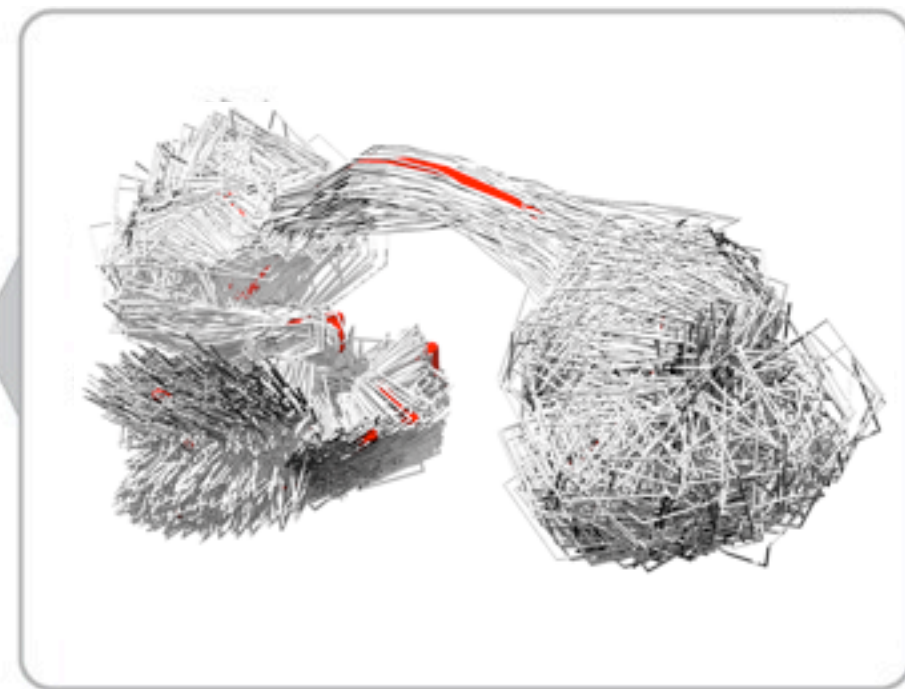
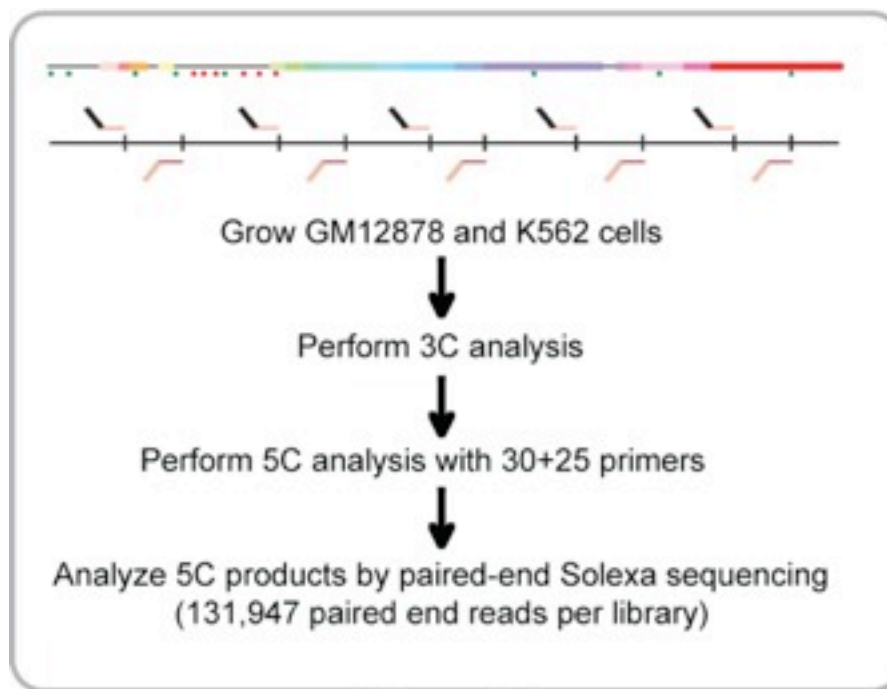
NMR structure determination  
*2D-NOESY data*



Chromosome structure determination  
*5C data*

# Integrative and iterative approach

## Experiments



Computation 



# Human $\alpha$ -globin domain

## ENm008 genomic structure and environment



- RNA expression data from the transcriptome group at Affymetrix and Cold Spring Harbor Laboratories.
- CTCF and histone modifications (H3K4me3) from the Broad Institute and the Bradley E. Bernstein Lab at the Massachusetts General Hospital/Harvard Medical School.
- DNaseI data from the Crawford Lab at Duke University and at the Collins Lab at NHGRI.

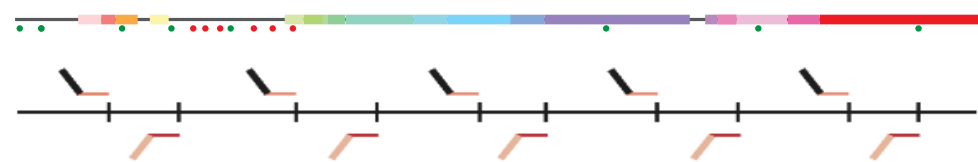
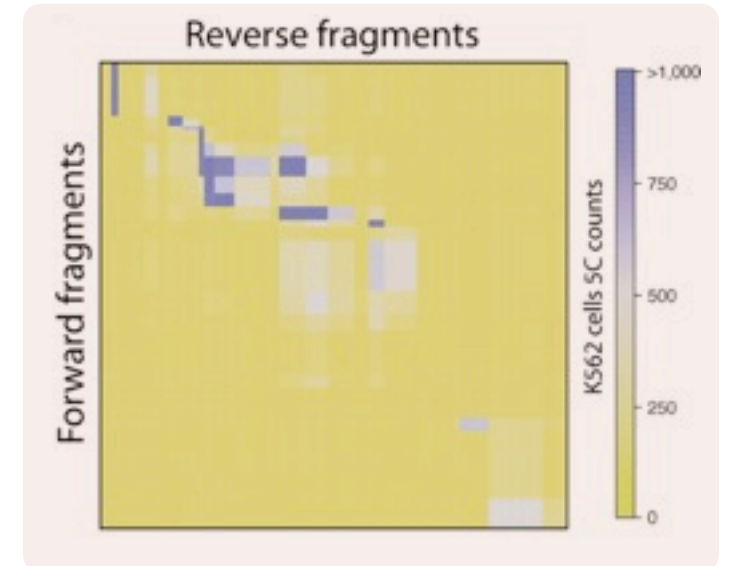
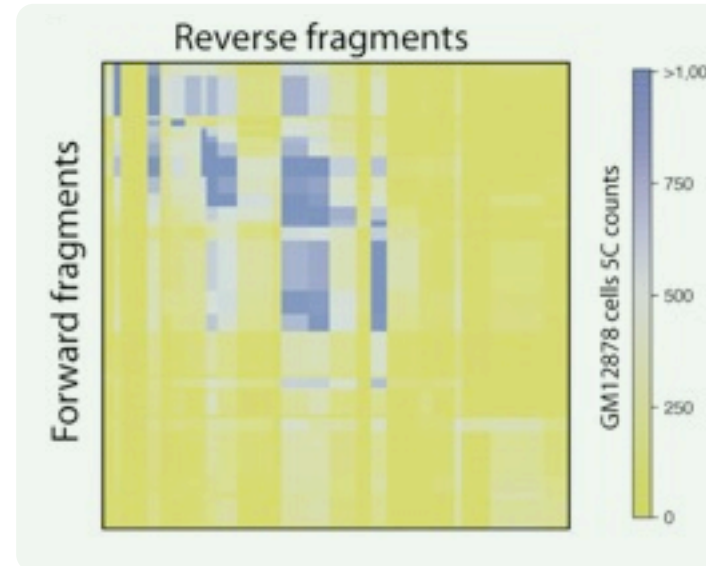
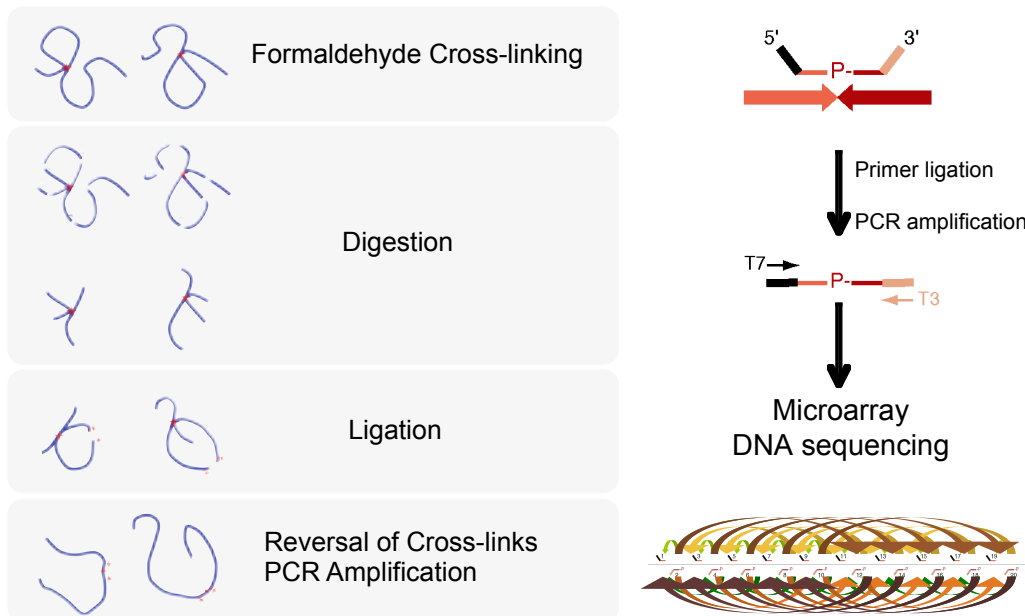
ENCODE Consortium. Nature (2007) vol. 447 (7146) pp. 799-816

# 5C experiments

<http://my5c.umassmed.edu>

GM12878

K562

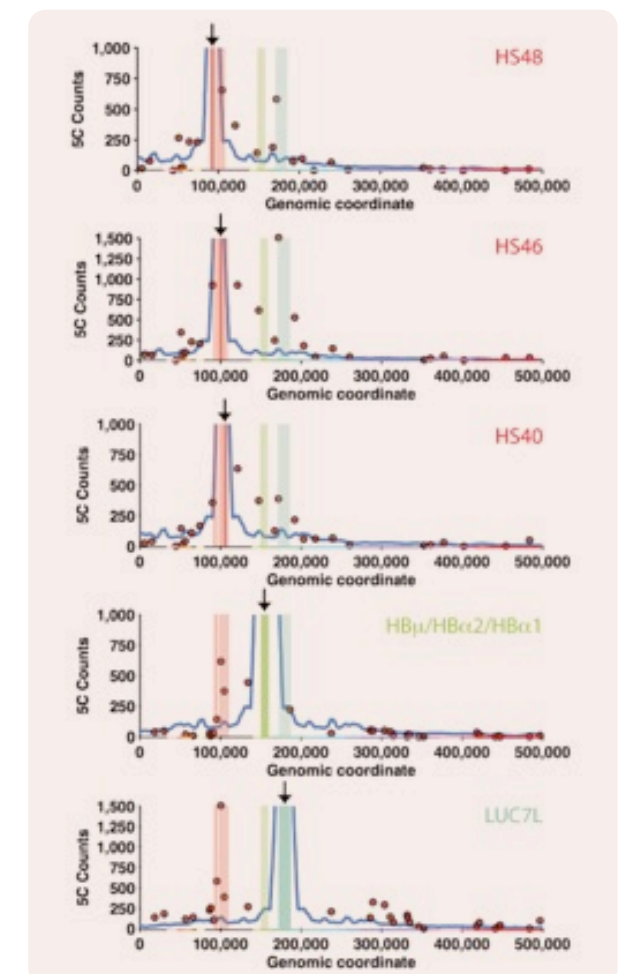
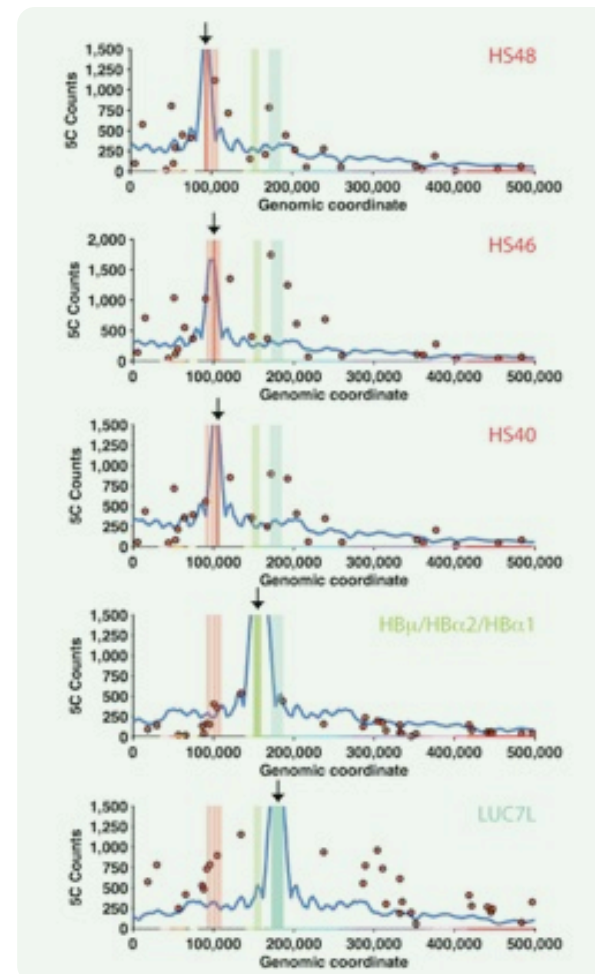


Grow GM12878 and K562 cells

Perform 3C analysis

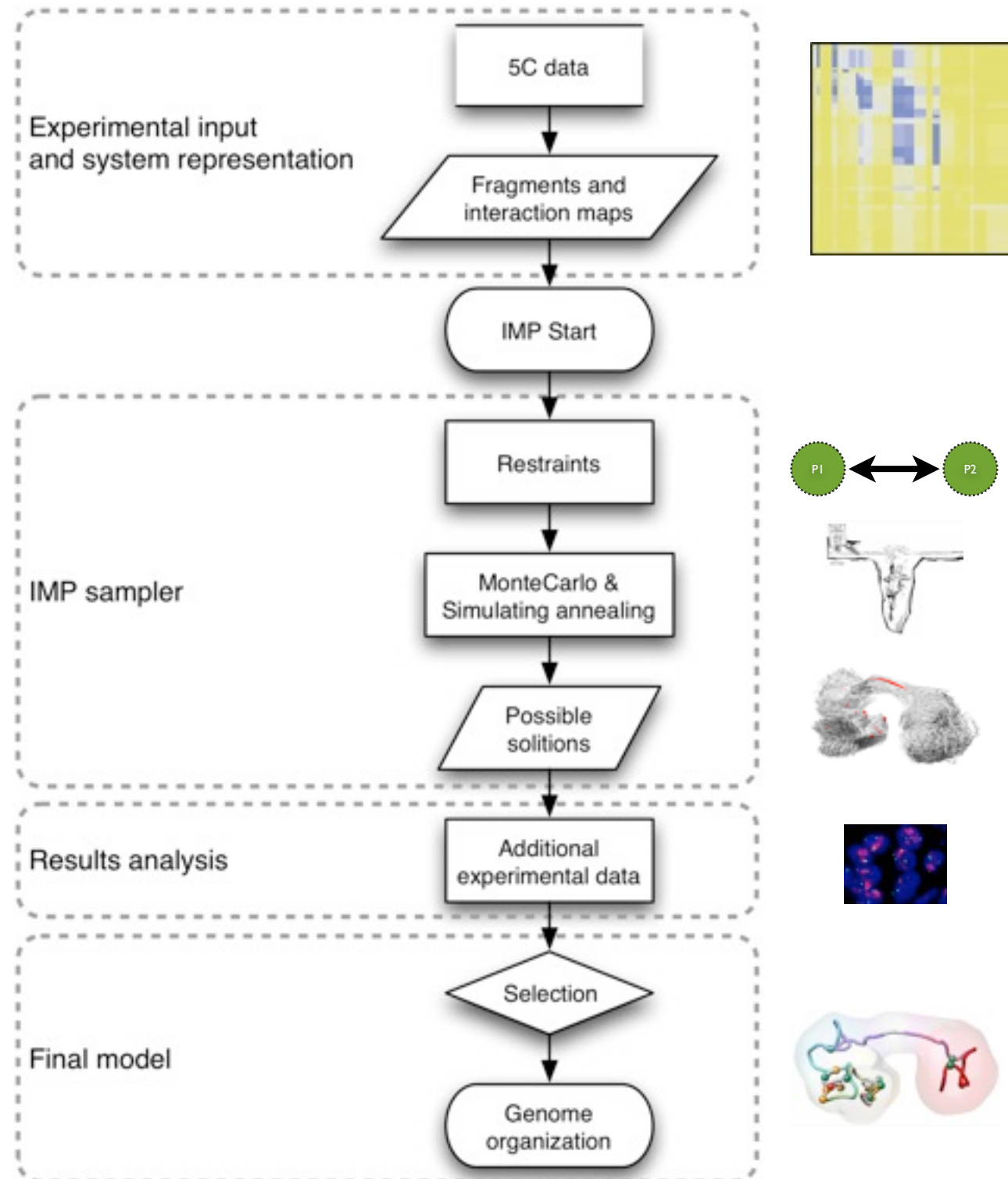
Perform 5C analysis with 30+25 primers

Analyze 5C products by paired-end Solexa sequencing  
(131,947 paired end reads per library)



# Integrative Modeling

<http://www.integrativemodeling.org>



# Representation

## Harmonic

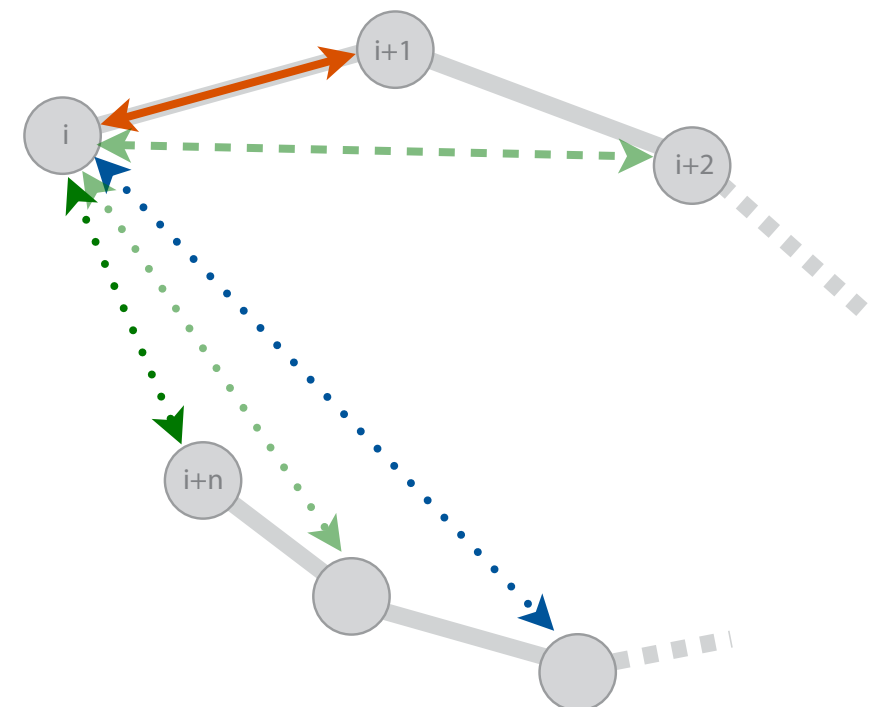
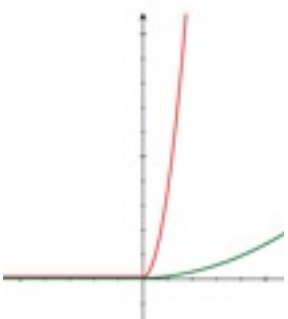
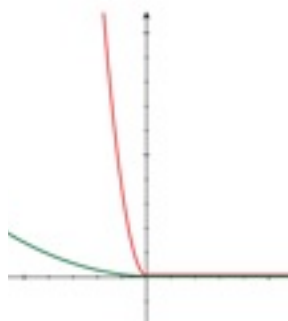
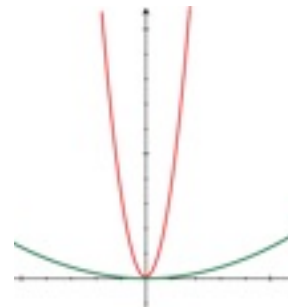
$$H_{i,j} = k(d_{i,j} - d_{i,j}^0)^2$$

## Harmonic Lower Bound

$$\begin{cases} \text{if } d_{i,j} \leq d_{i,j}^0; & lbH_{i,j} = k(d_{i,j} - d_{i,j}^0)^2 \\ \text{if } d_{i,j} > d_{i,j}^0; & lbH_{i,j} = 0 \end{cases}$$

## Harmonic Upper Bound

$$\begin{cases} \text{if } d_{i,j} \geq d_{i,j}^0; & ubH_{i,j} = k(d_{i,j} - d_{i,j}^0)^2 \\ \text{if } d_{i,j} < d_{i,j}^0; & ubH_{i,j} = 0 \end{cases}$$

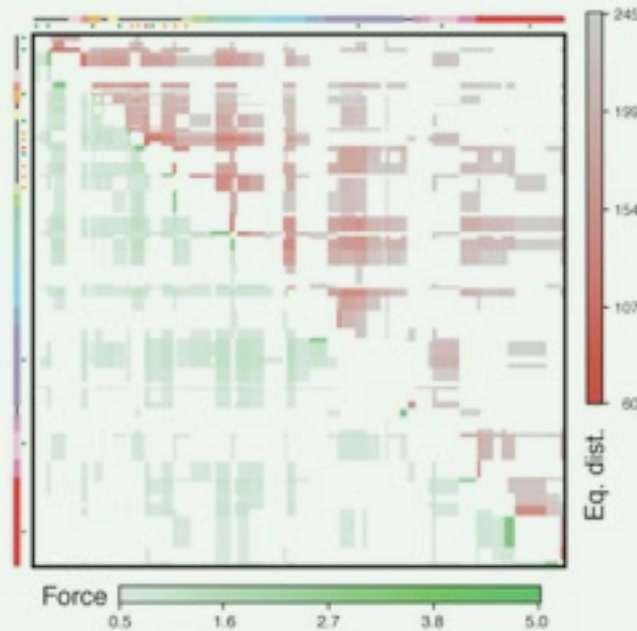
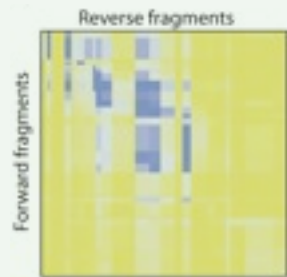




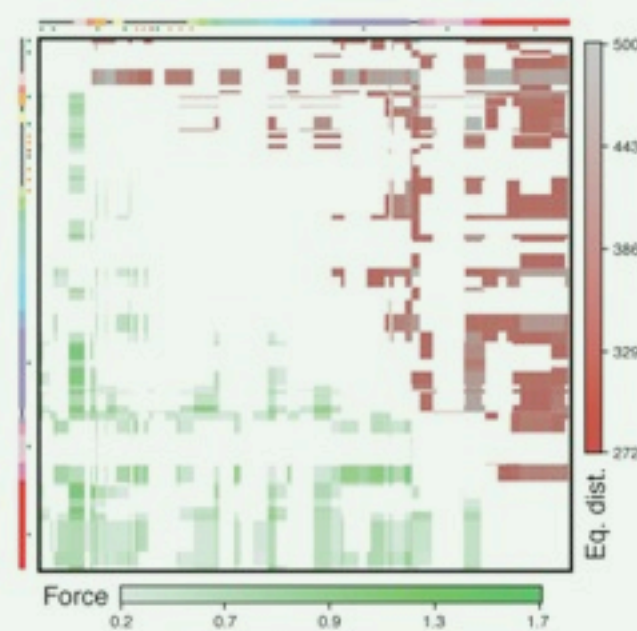
# Scoring

GM12878

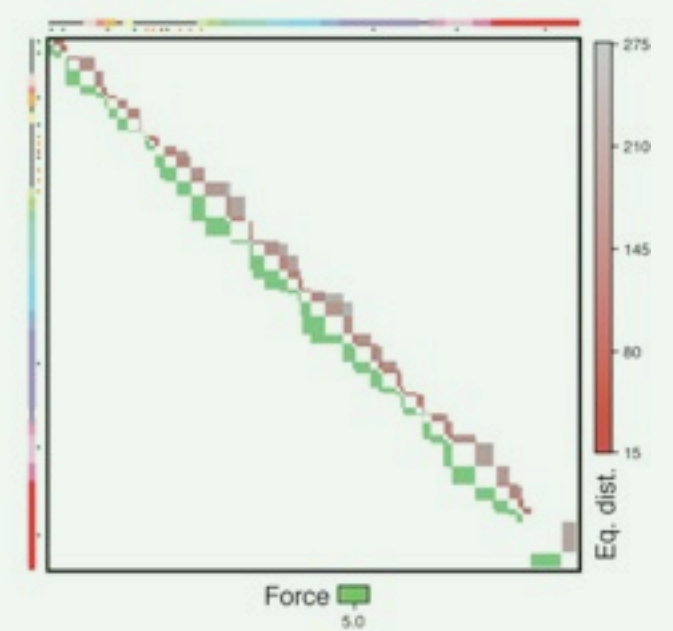
70 fragments  
1,520 restraints



Harmonic



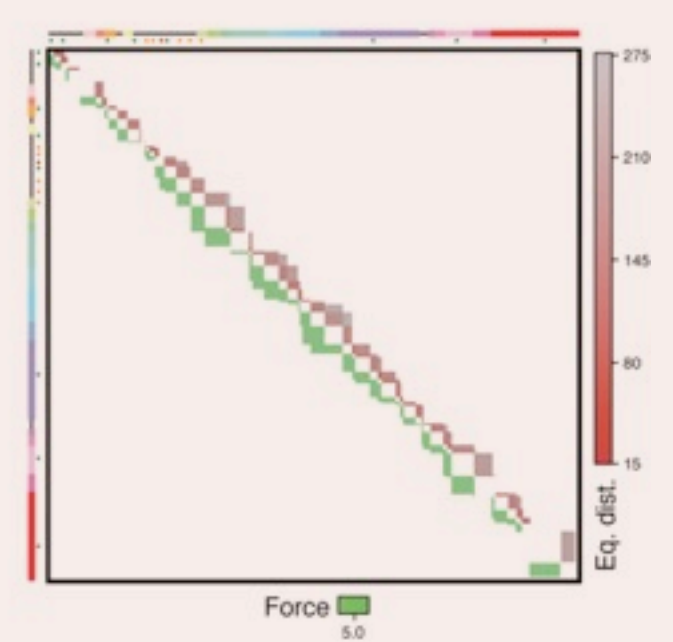
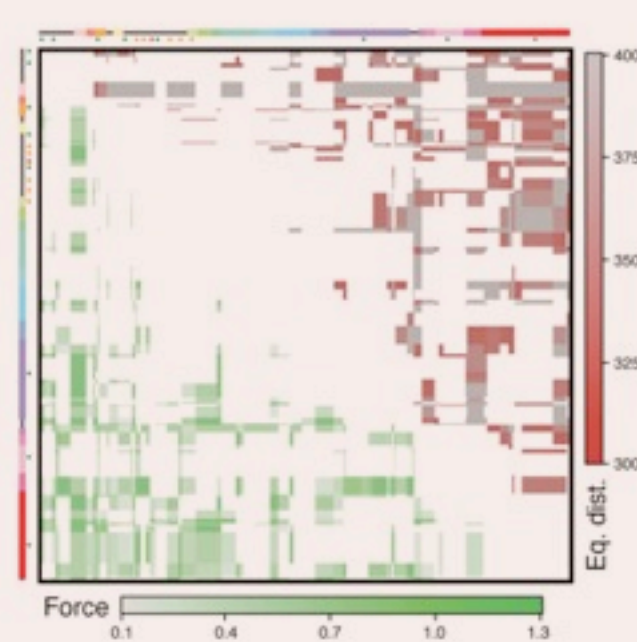
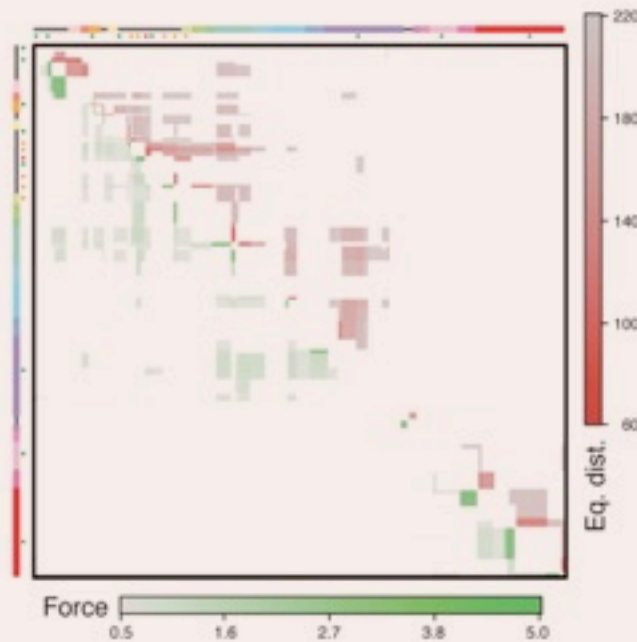
Harmonic Lower Bound



Harmonic Upper Bound

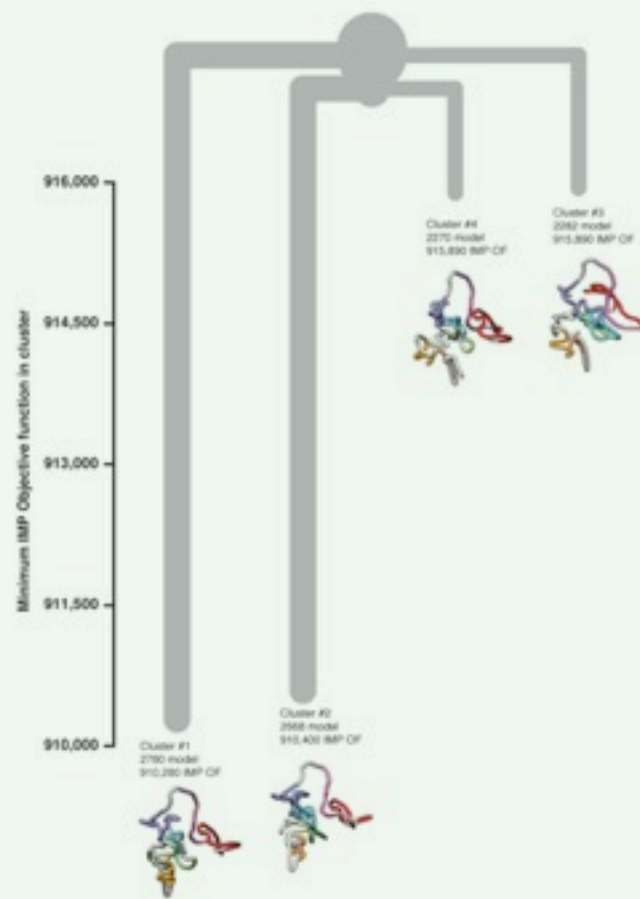
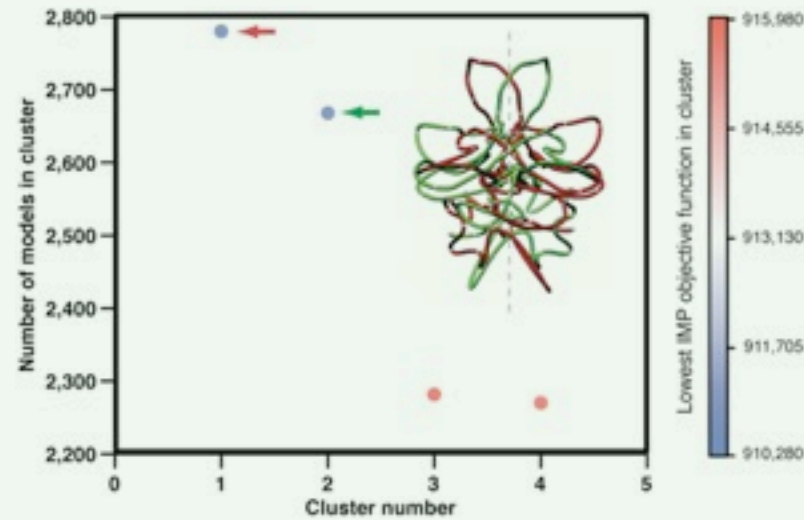
K562

70 fragments  
1,049 restraints

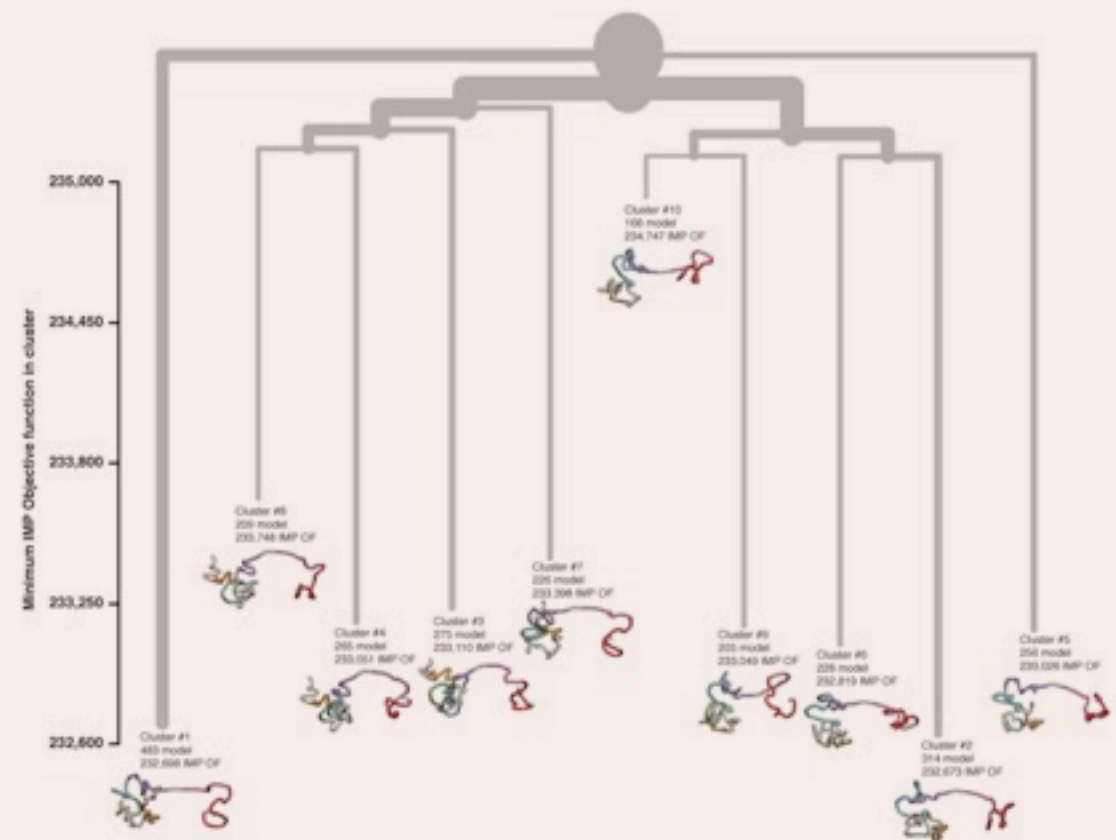
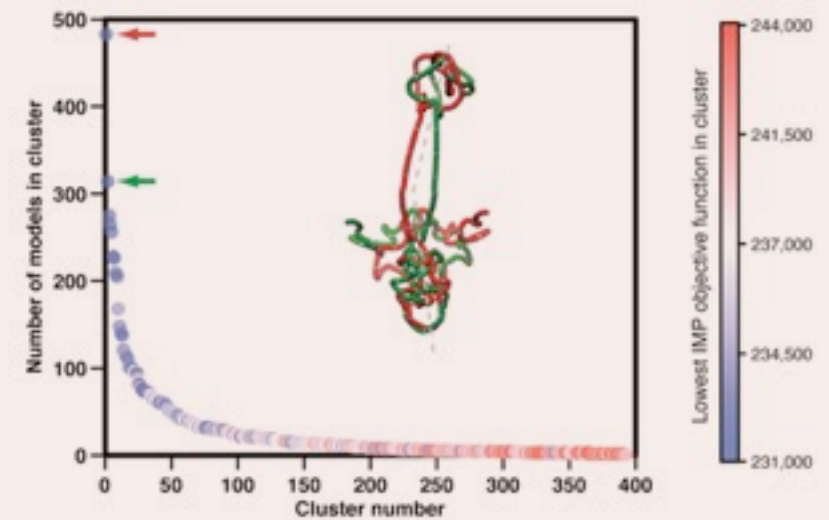


# Not just *the* solution

GM12878

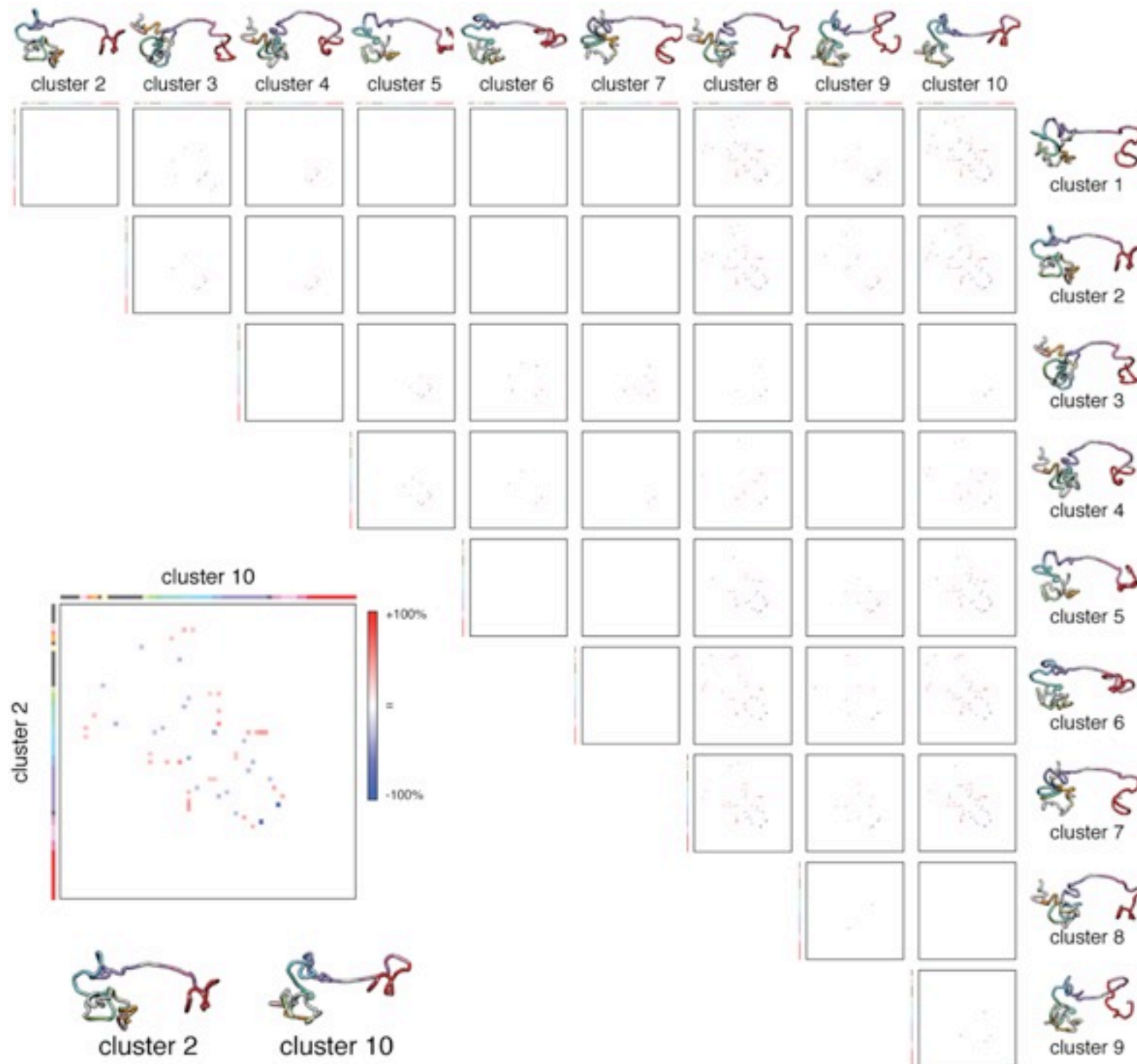


K562



# Not just *the* solution

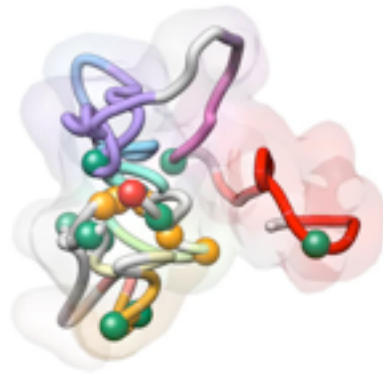
and we can de-convolute them!



# Consistency

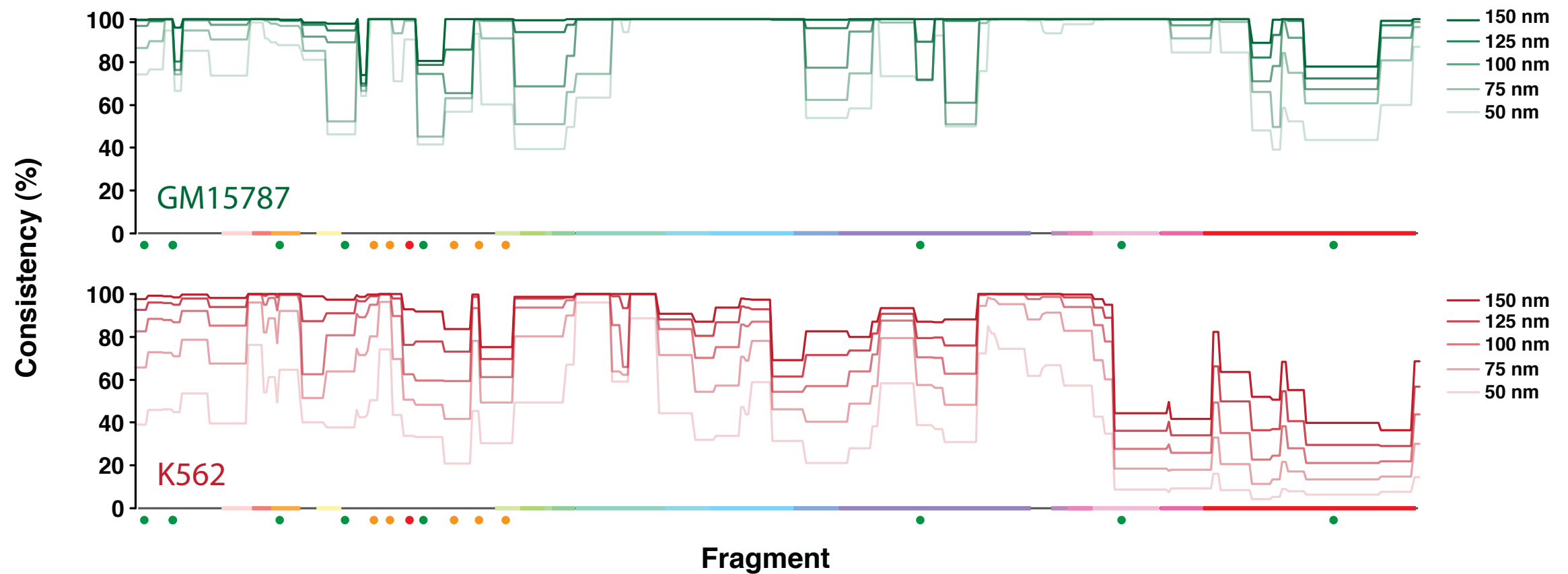
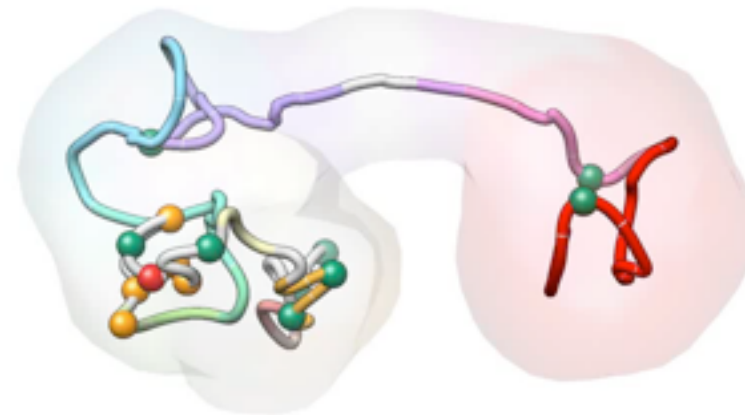
## GM12878

Cluster #1  
2780 model  
910,280 IMP OF



## K562

Cluster #2  
314 model  
232,673 IMP OF

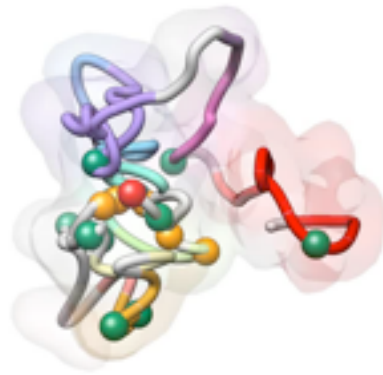




# Compactness

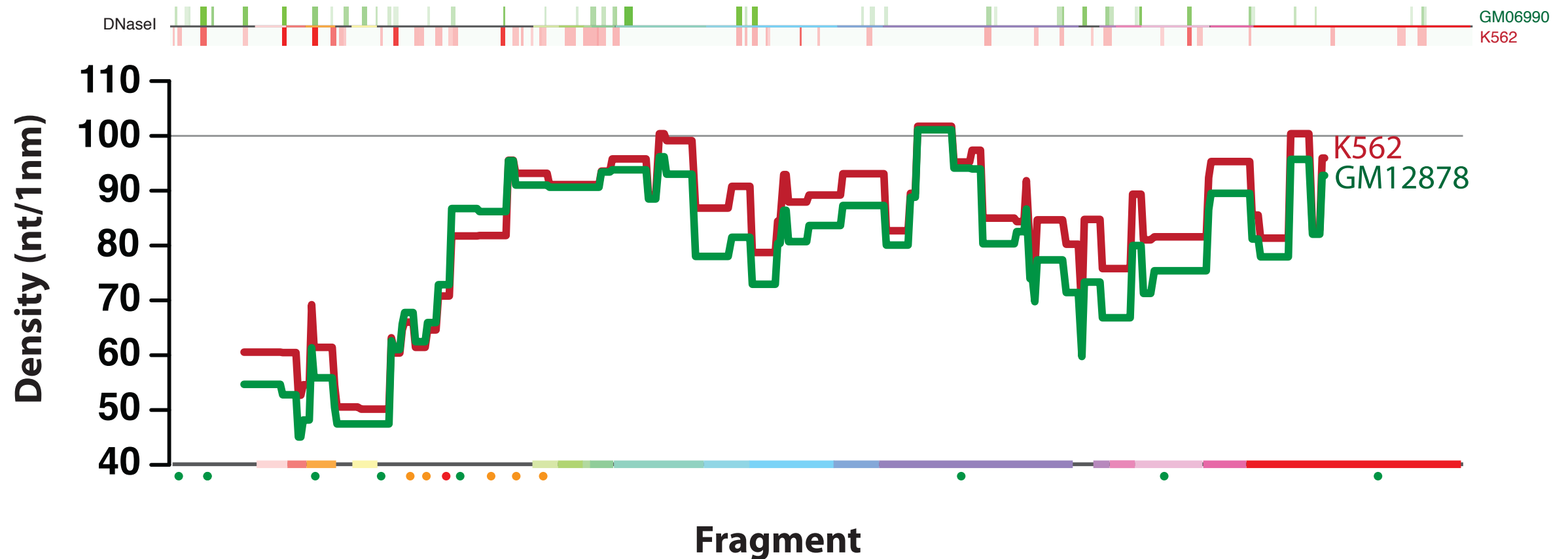
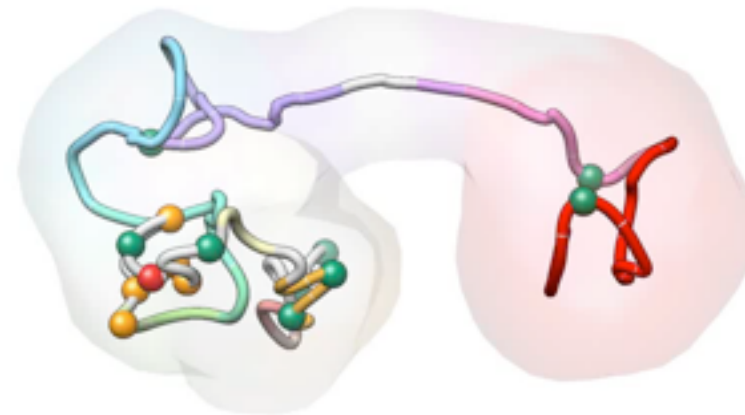
**GM12878**

Cluster #1  
2780 model  
910,280 IMP OF



**K562**

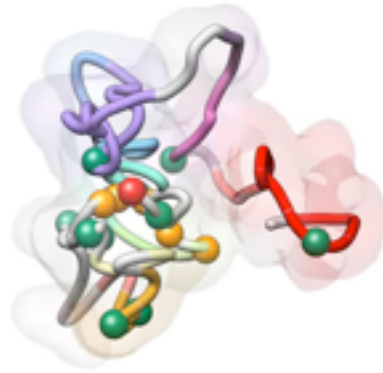
Cluster #2  
314 model  
232,673 IMP OF



# Regulatory elements

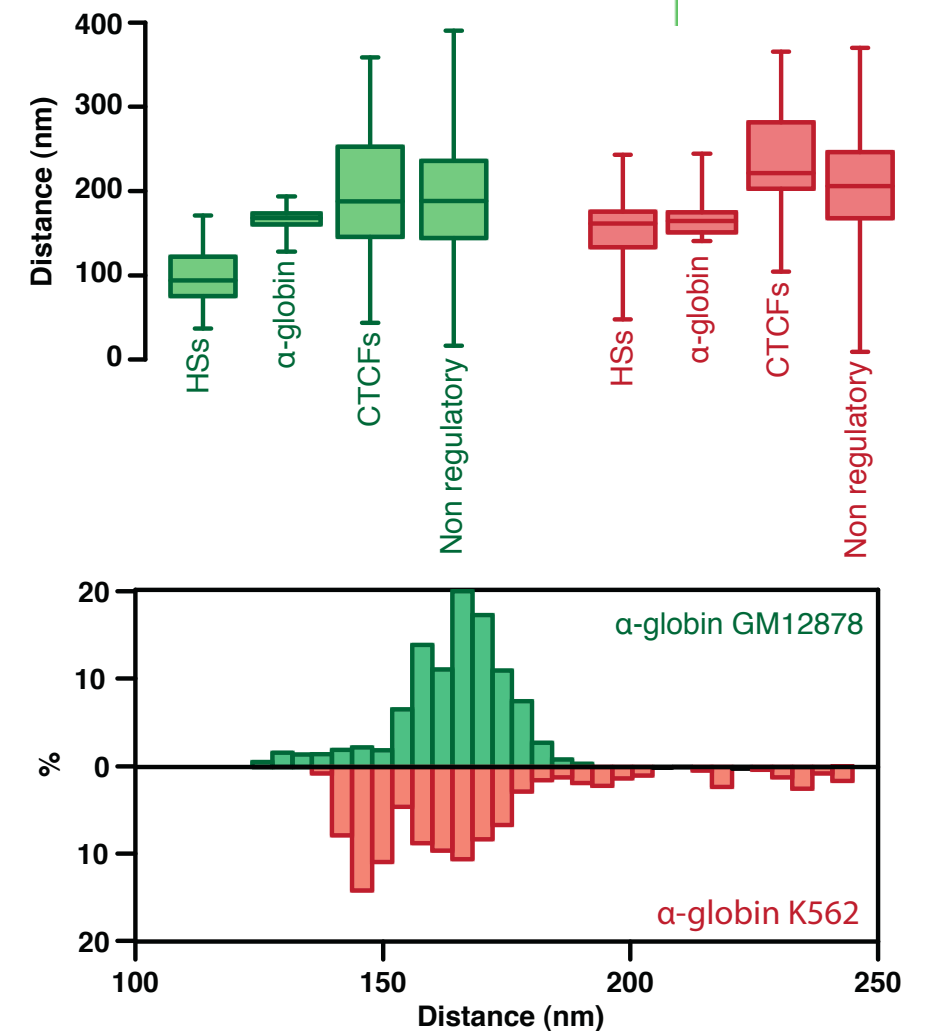
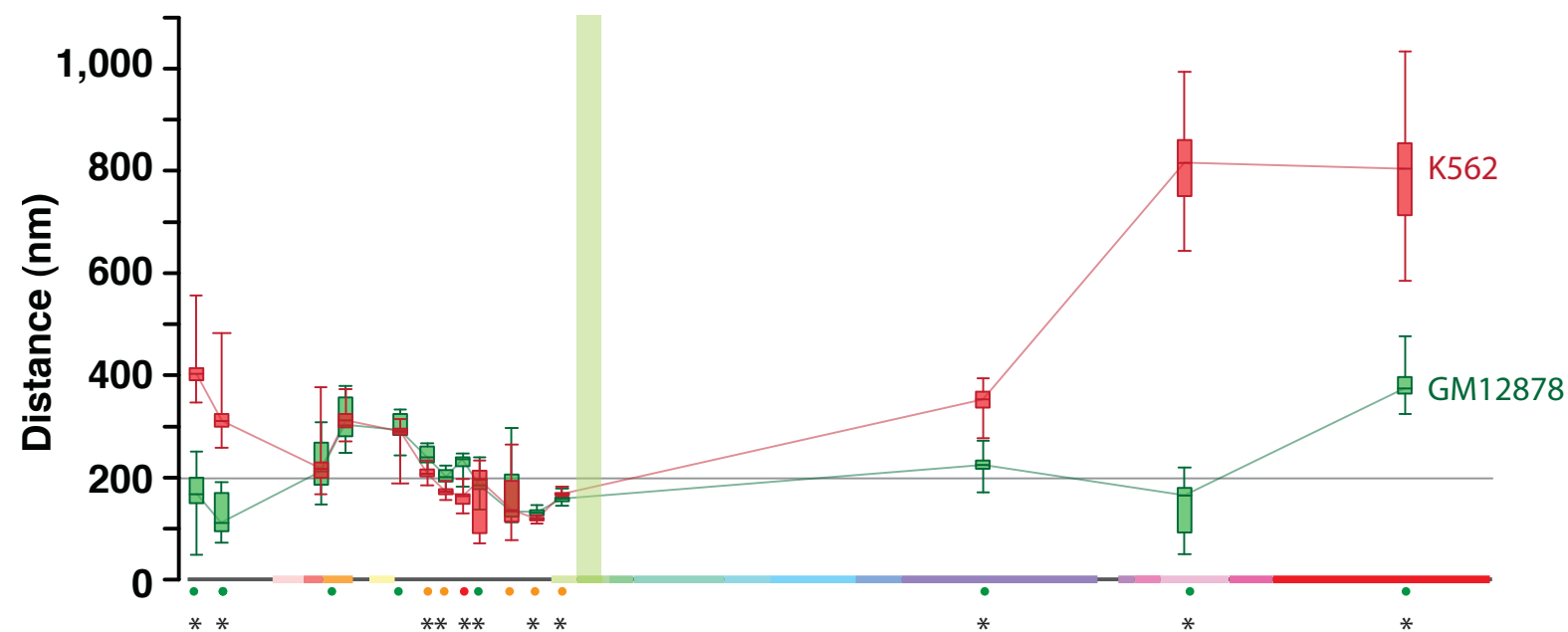
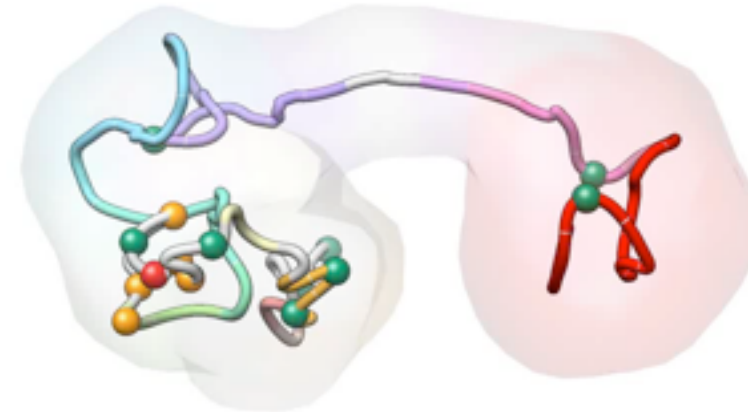
**GM12878**

Cluster #1  
2780 model  
910,280 IMP OF



**K562**

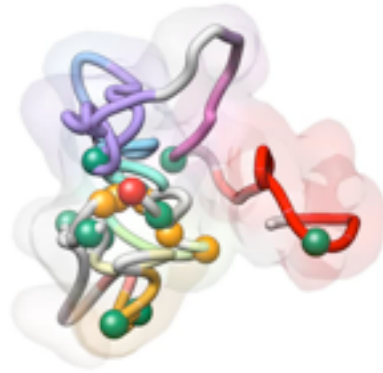
Cluster #2  
314 model  
232,673 IMP OF



# Multi-loops

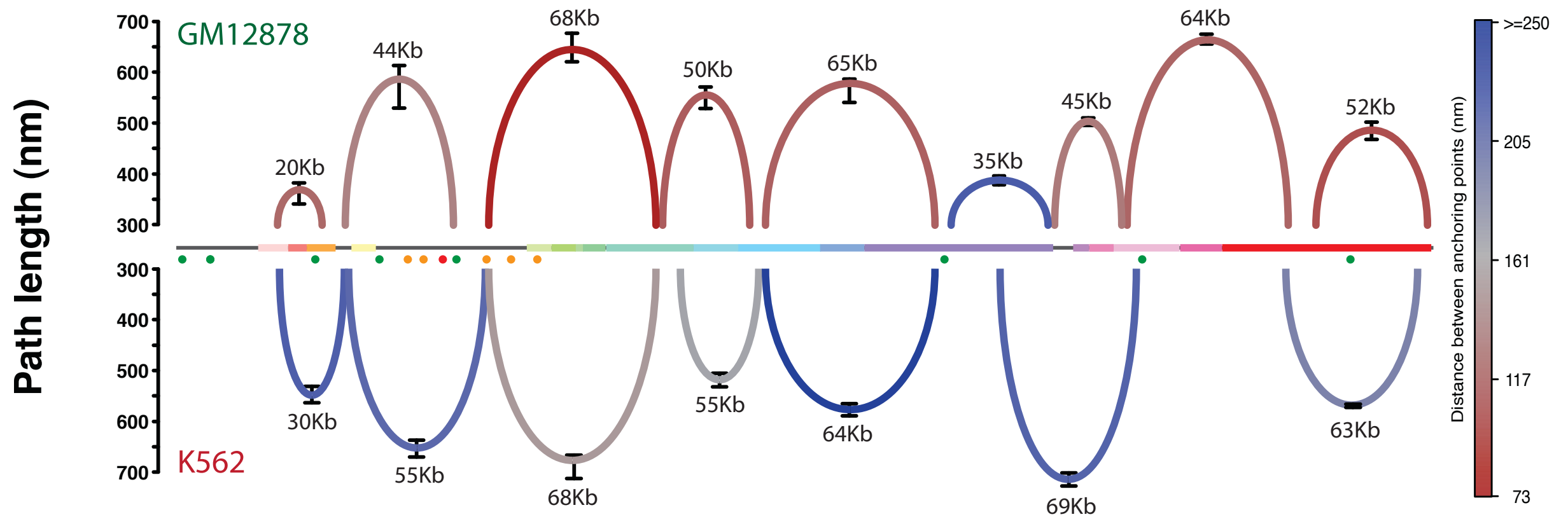
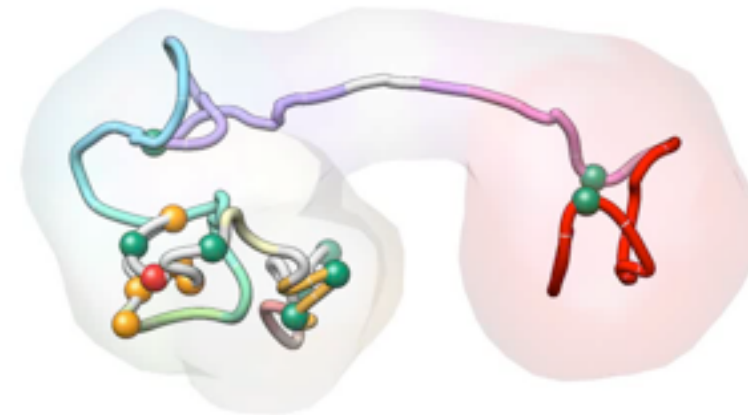
**GM12878**

Cluster #1  
2780 model  
910,280 IMP OF



**K562**

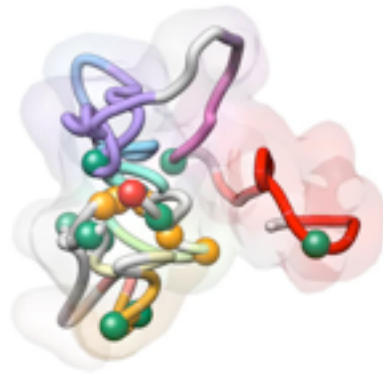
Cluster #2  
314 model  
232,673 IMP OF



# Expression

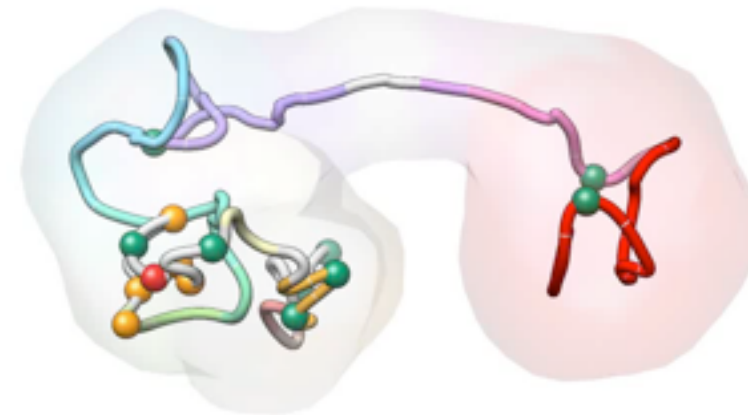
## GM12878

Cluster #1  
2780 model  
910,280 IMP OF



## K562

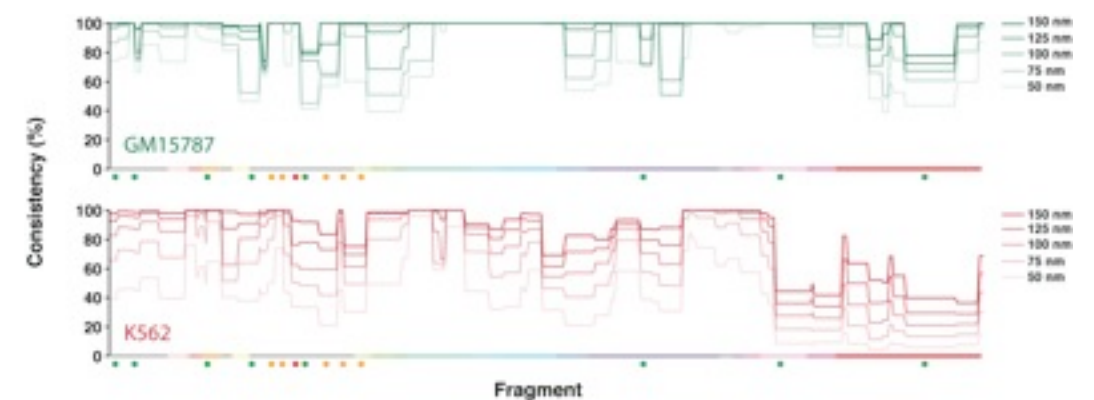
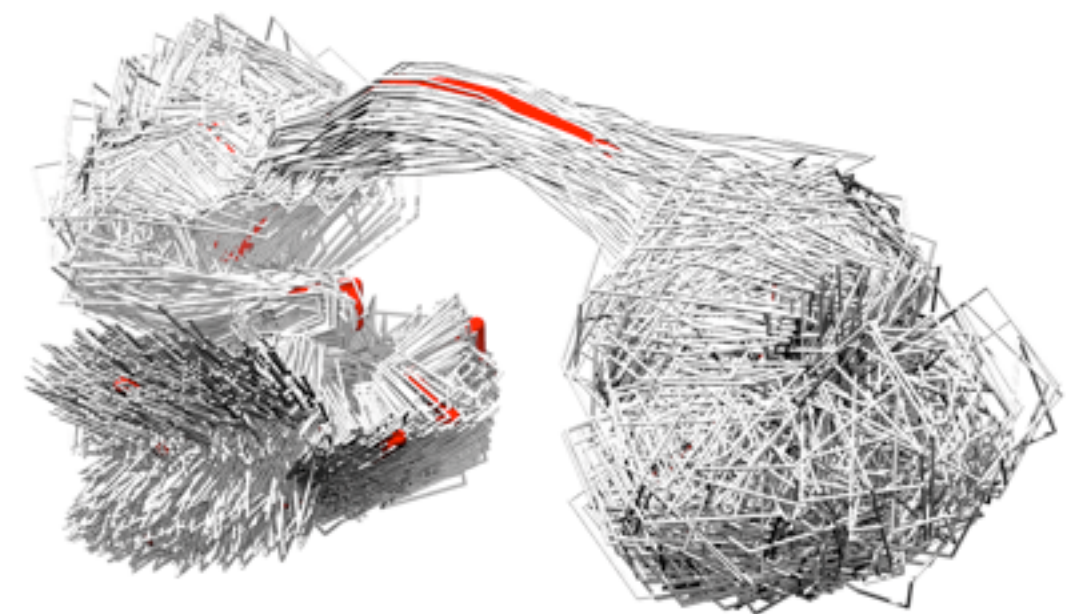
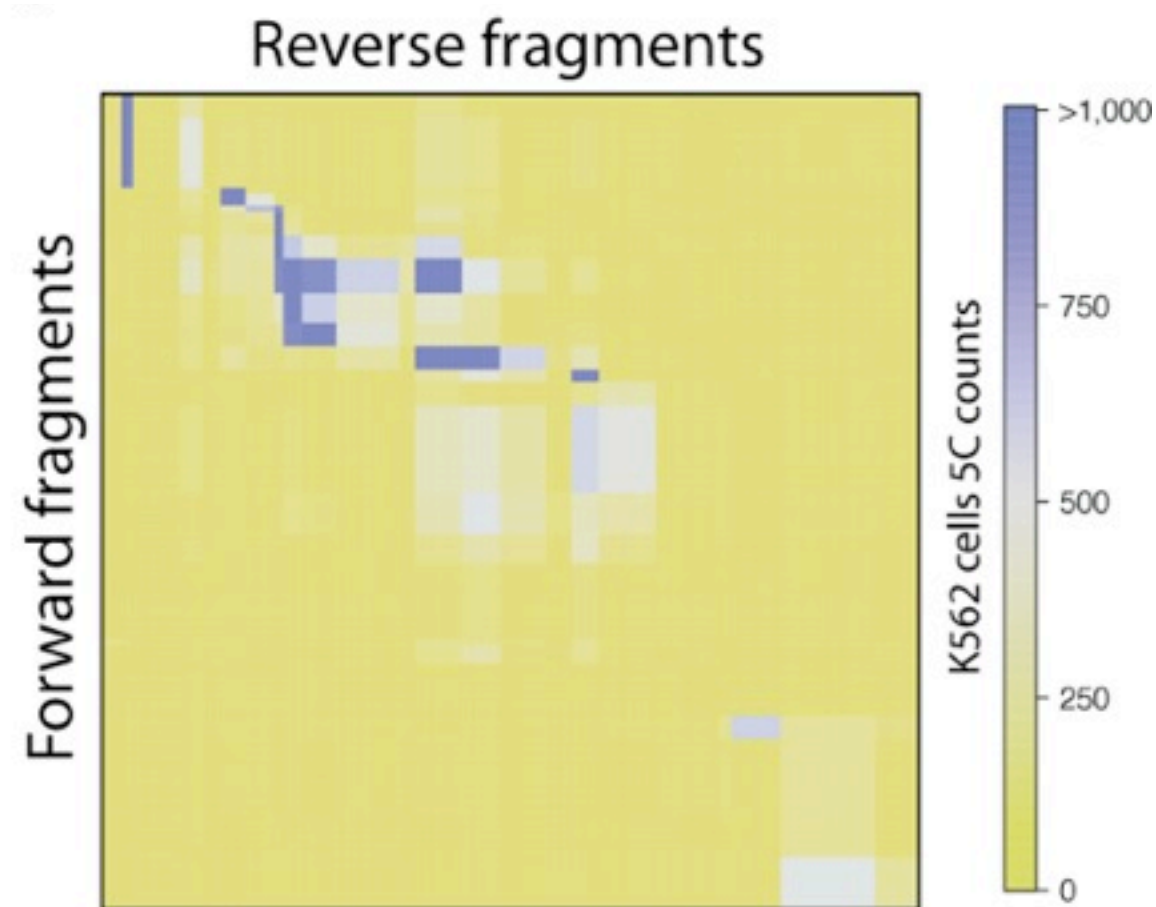
Cluster #2  
314 model  
232,673 IMP OF





# Summary

5C data results in comprehensive interaction matrices to build a consistent 3D model



# Summary

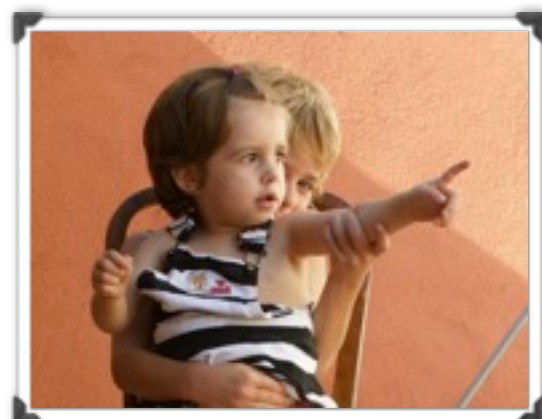
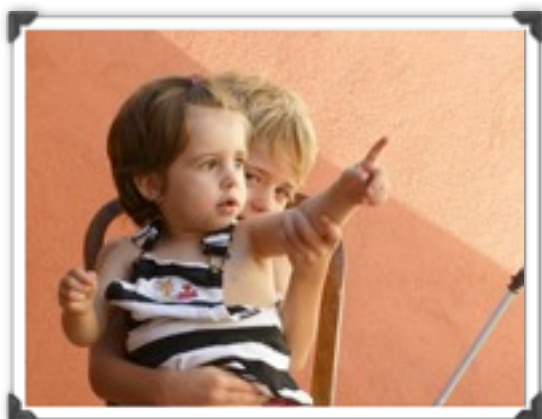
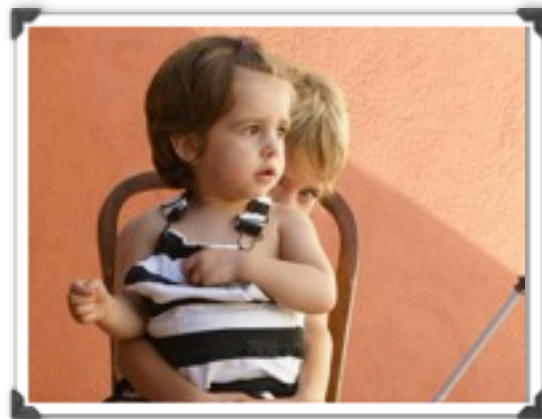
Models allow for 5C data de-convolution





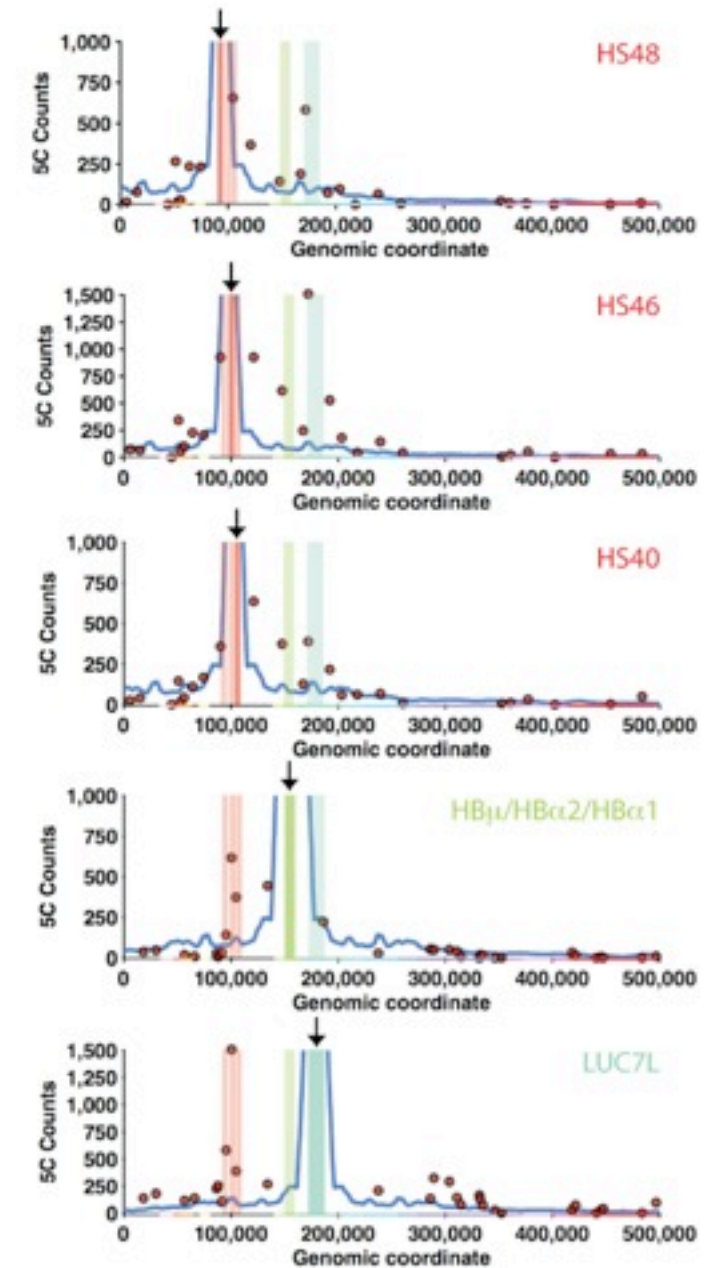
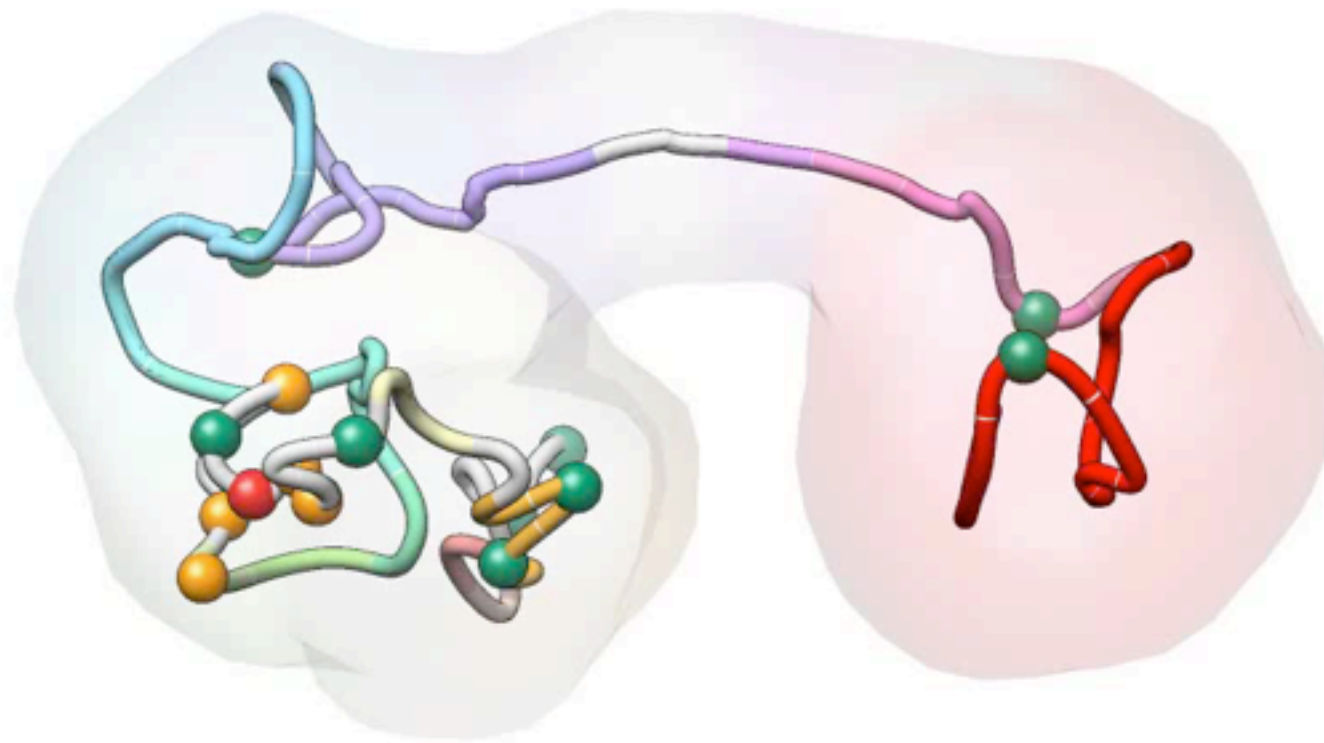
# Summary

Models allow for 5C data de-convolution



# Summary

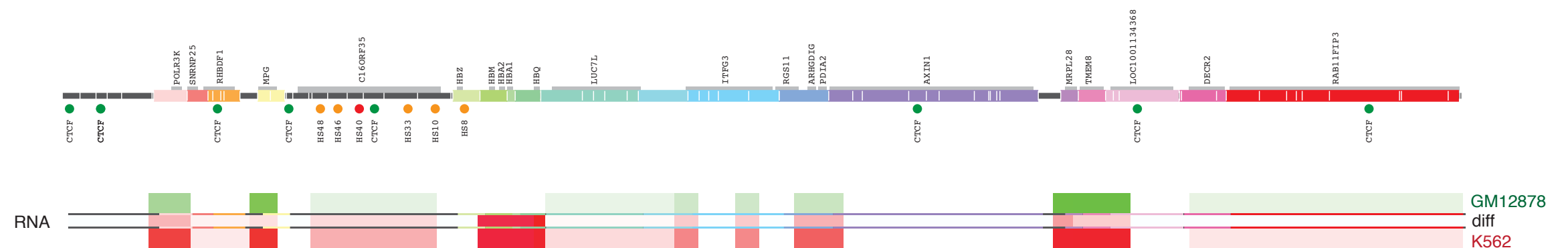
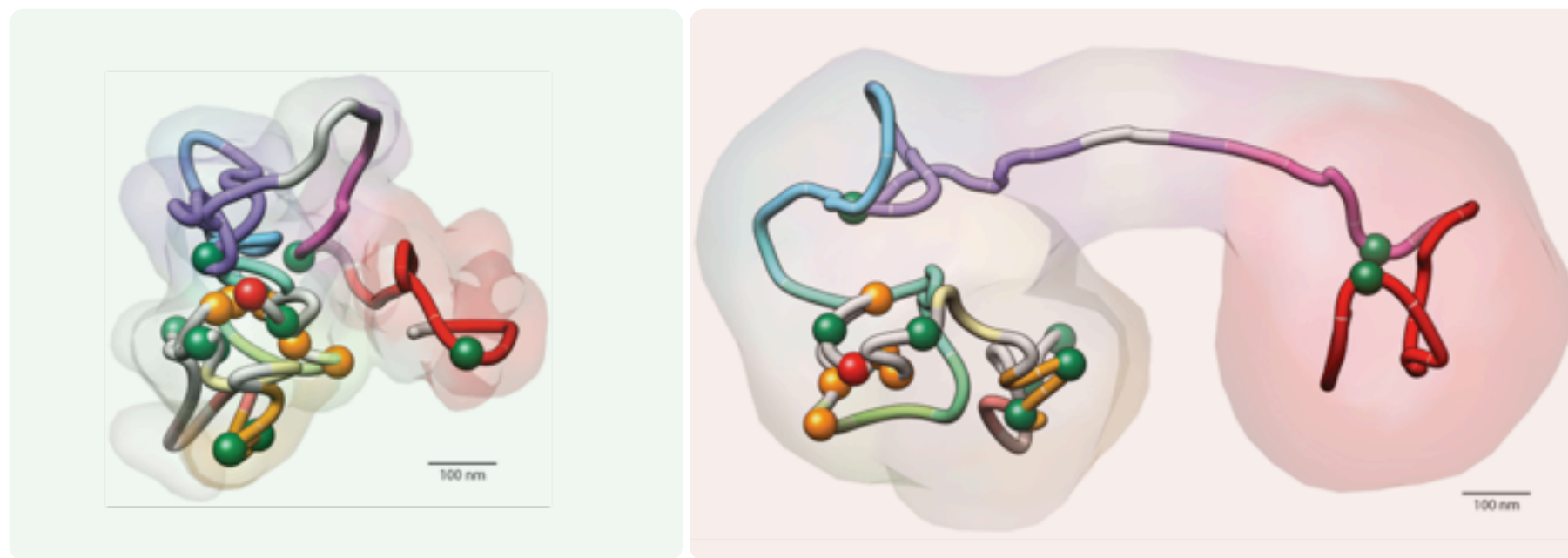
Selected models reproduce known (**and new**) interactions





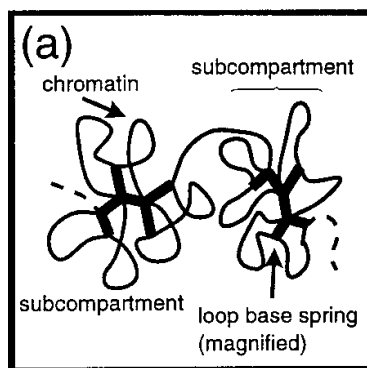
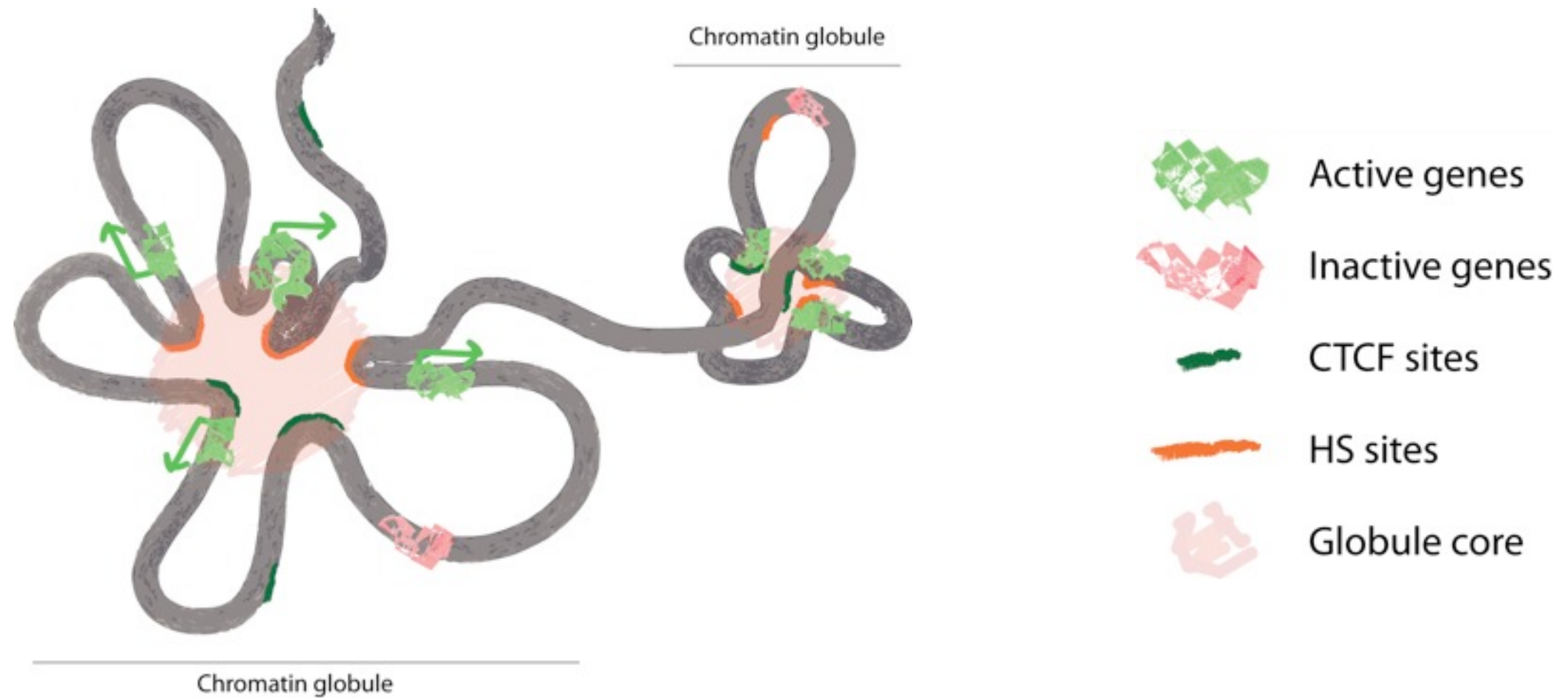
# Summary

Large-scale changes in conformation correlate with gene expression of resident genes

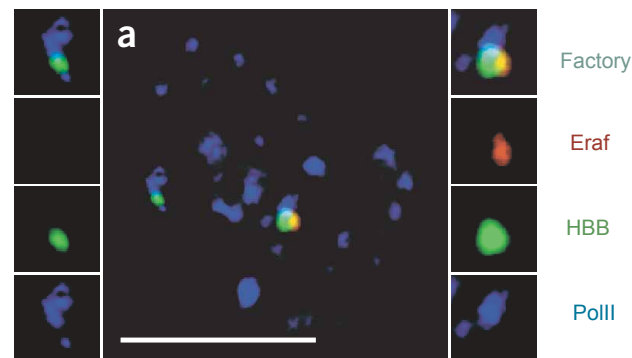


# Summary

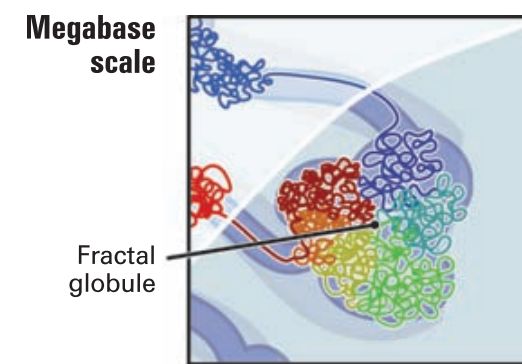
## “Chromatin Globule” model



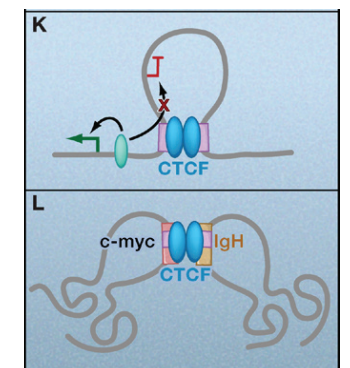
Münkel et al. JMB (1999)



Osborne et al. Nat Genet (2004)



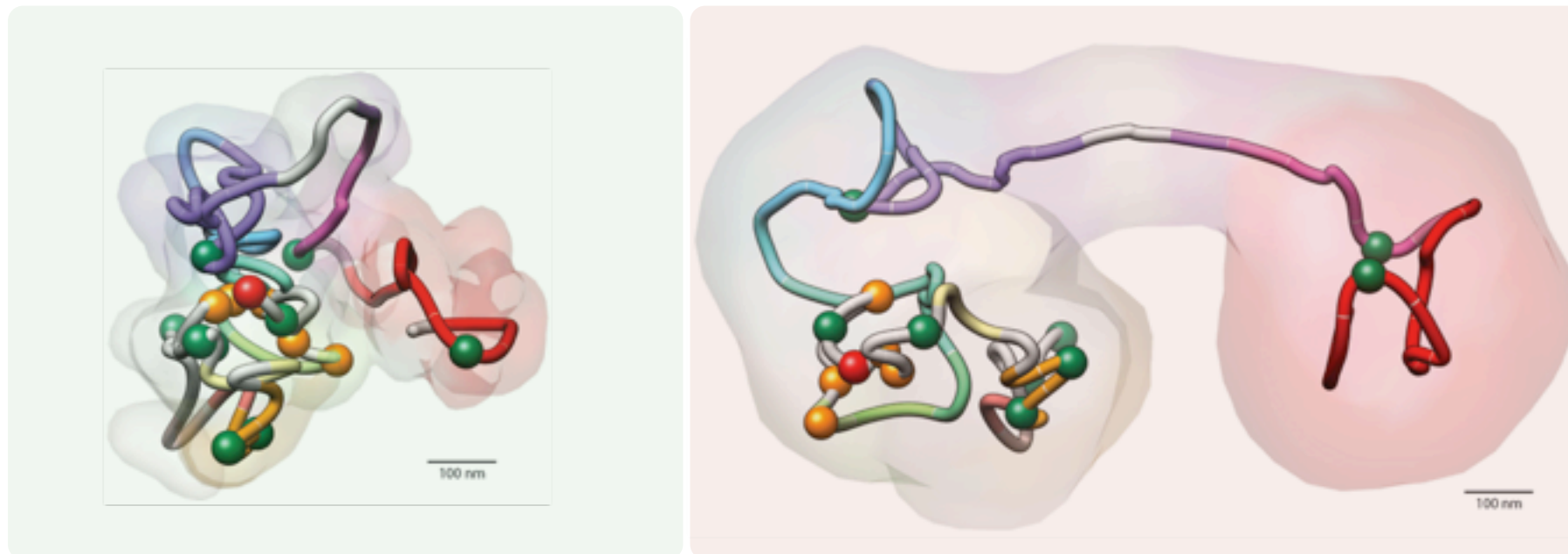
Lieberman-Aiden et al. Science (2009)



Phillips and Corces. Cell (2009)

# Open questions...

- Can we further **assess** the **accuracy** of our models?
- What **mechanisms** drive long-range interactions?
- Are **all** long-range interactions “functional”?
- What are **dynamics** of higher-order chromatin?



# Acknowledgments

## Job Dekker

Program in Gene Function and Expression  
Department of Biochemistry and Molecular Pharmacology  
University of Massachusetts Medical School  
Worcester, MA, USA



## Marc A. Marti-Renom

Structural Genomics Unit  
Bioinformatics and Genomics Department  
Centro de Investigación Príncipe Felipe  
Valencia, Spain



## Amartya Sanyal

Postdoctoral Fellow  
Dekker Lab



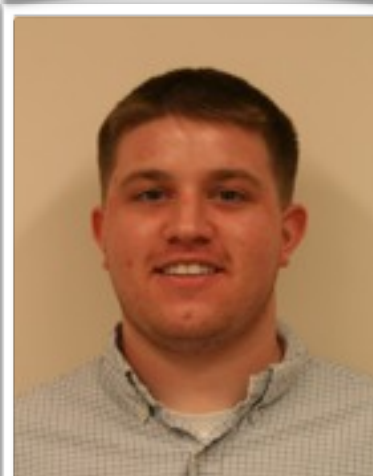
## Davide Baù

Postdoctoral fellow  
Structural Genomics Unit



## Bryan Lajoie

Bioinformatician  
Dekker Lab



## Emidio Capriotti

Postdoctoral fellow  
Structural Genomics Unit

