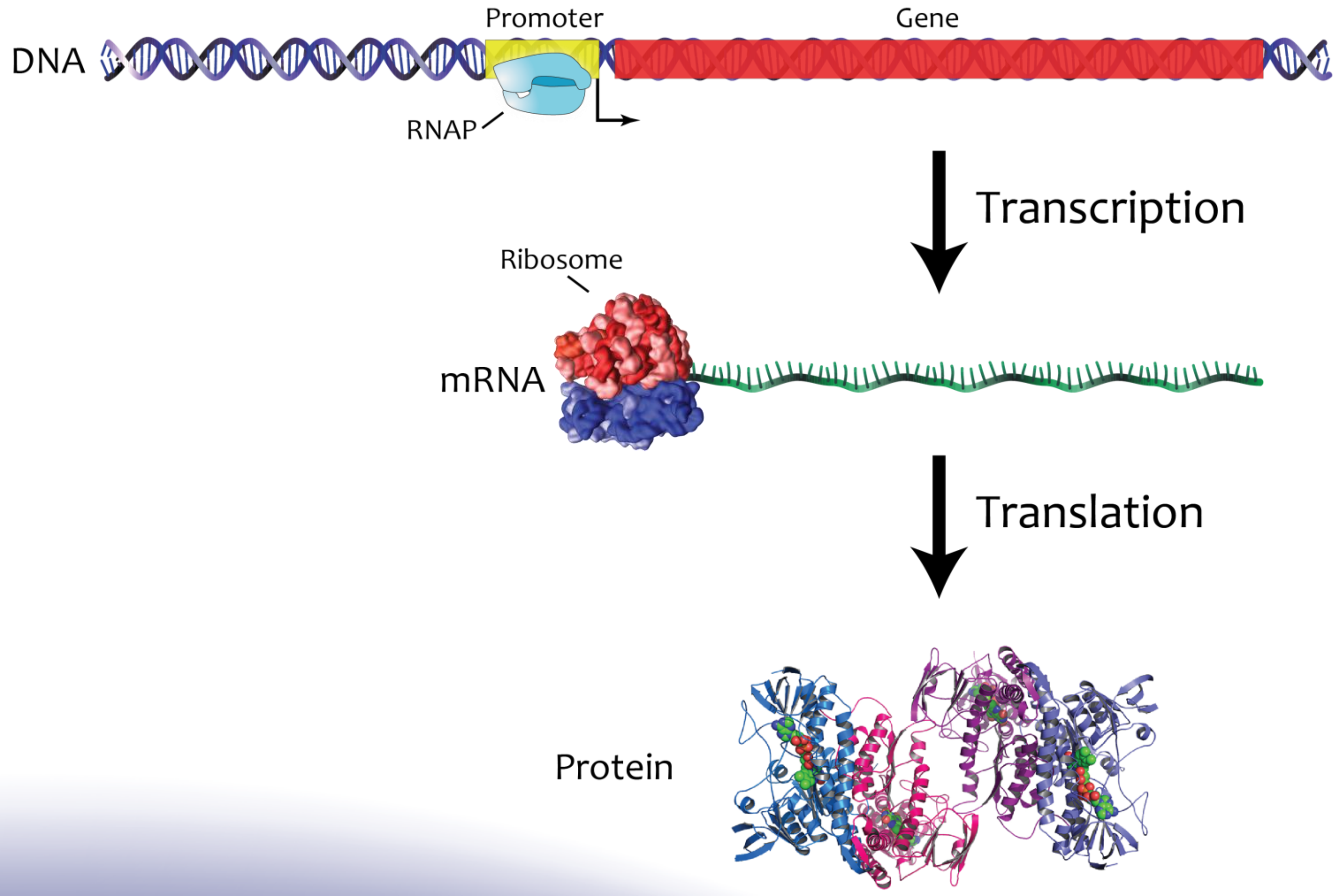
Enhancer Modeling by Monte-Carlo Simulations

Yaroslav Pollak

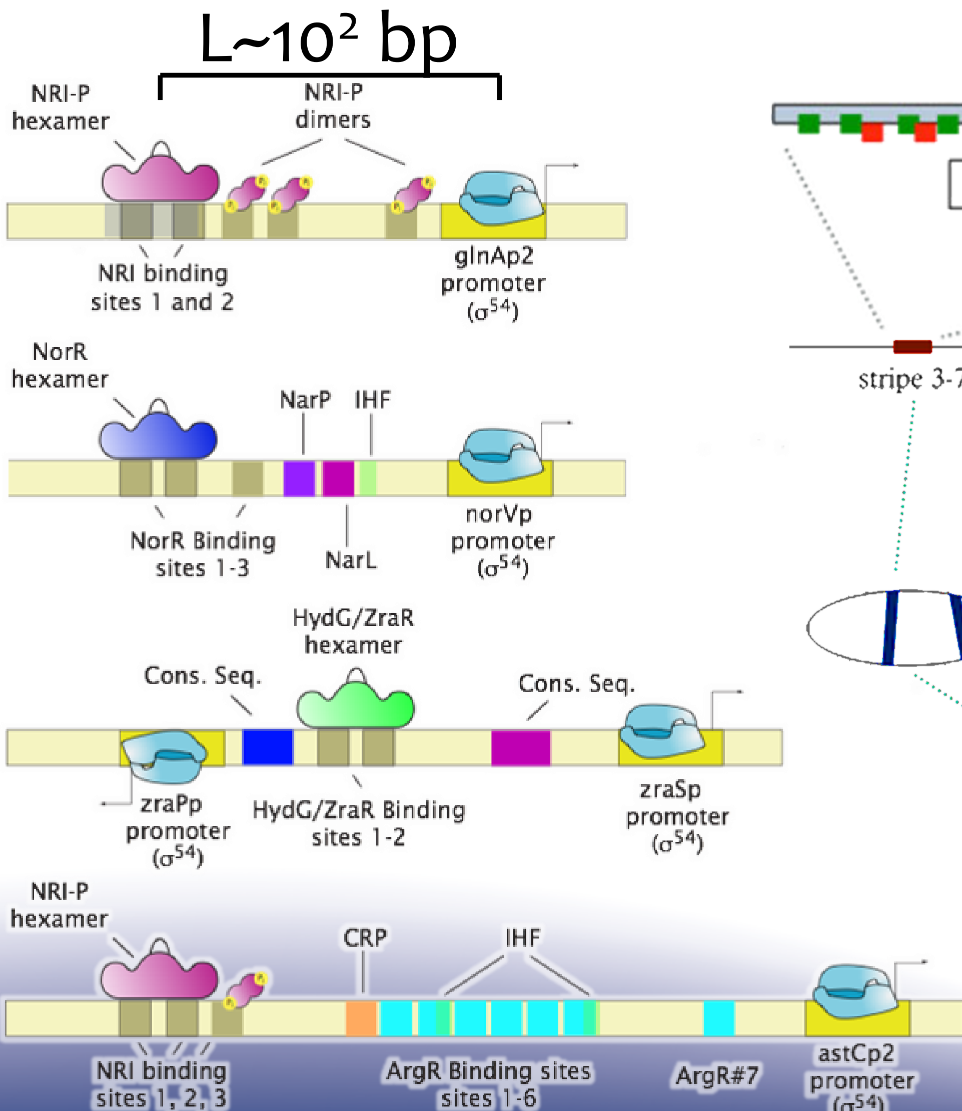
PhD seminar
November, 2016

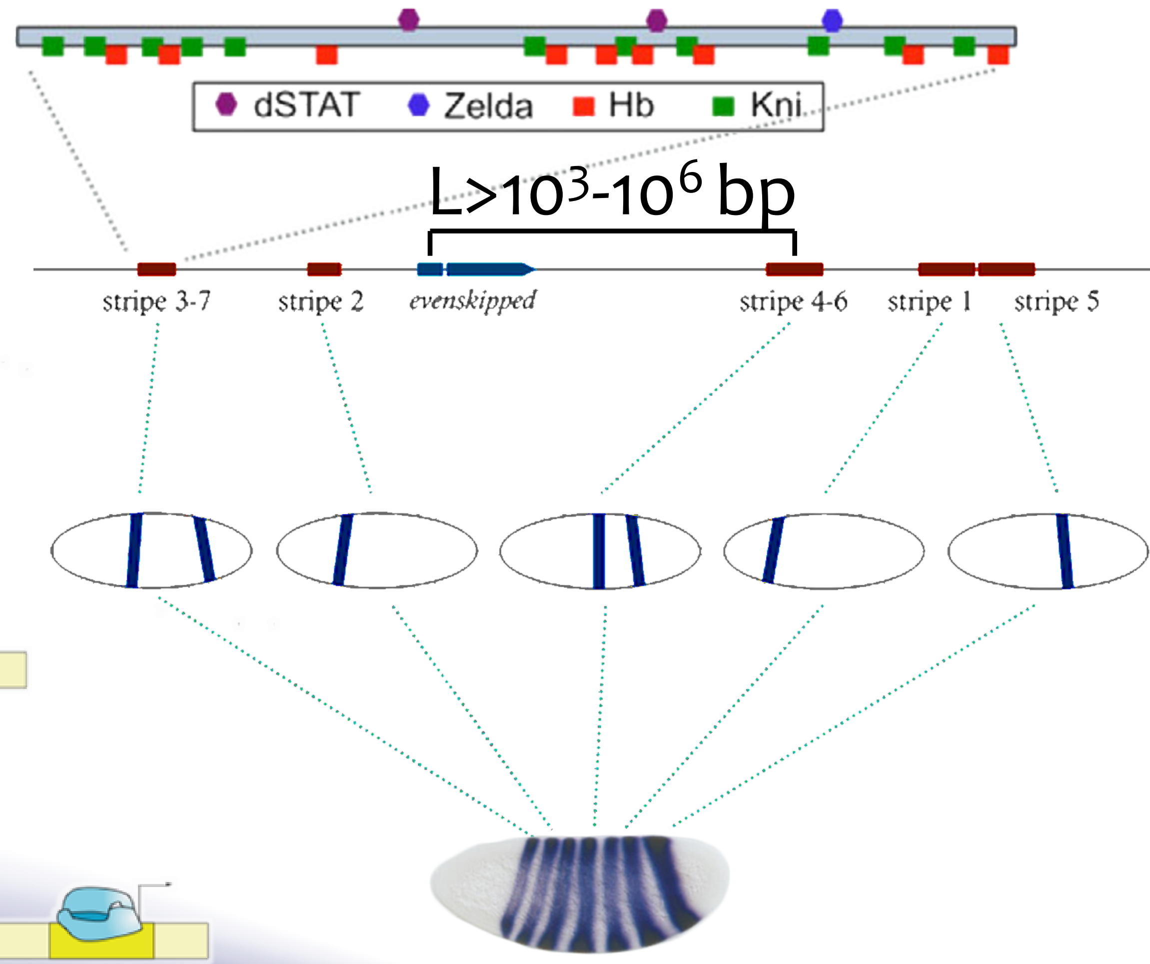
Introduction

Central Dogma of Biology



Enhancers

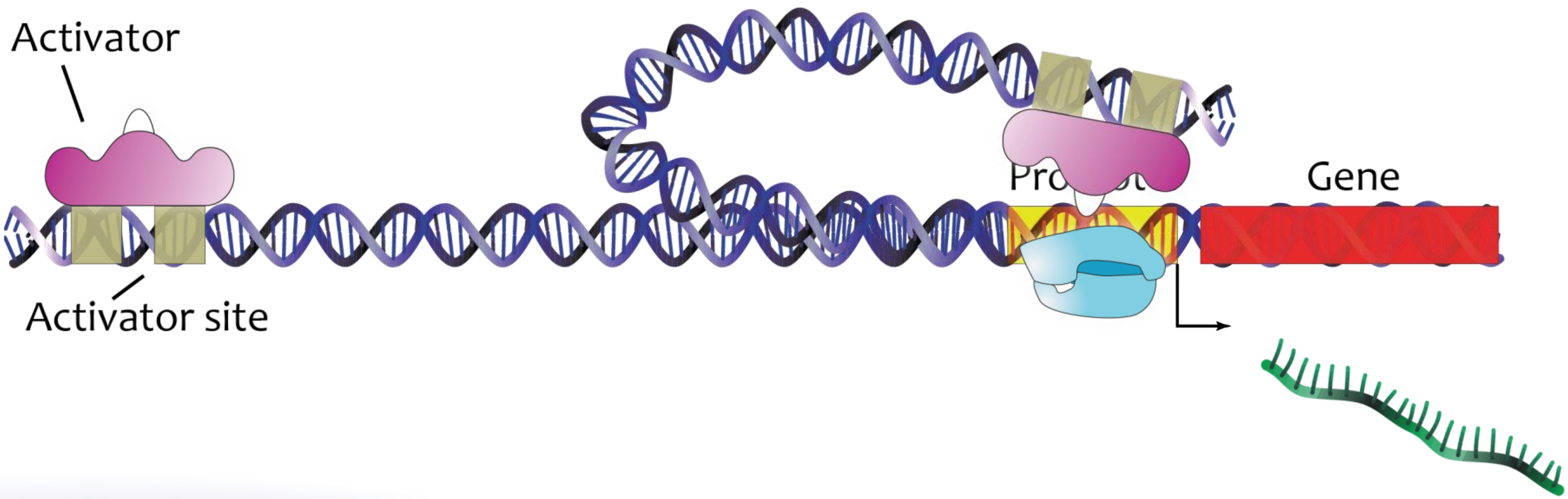




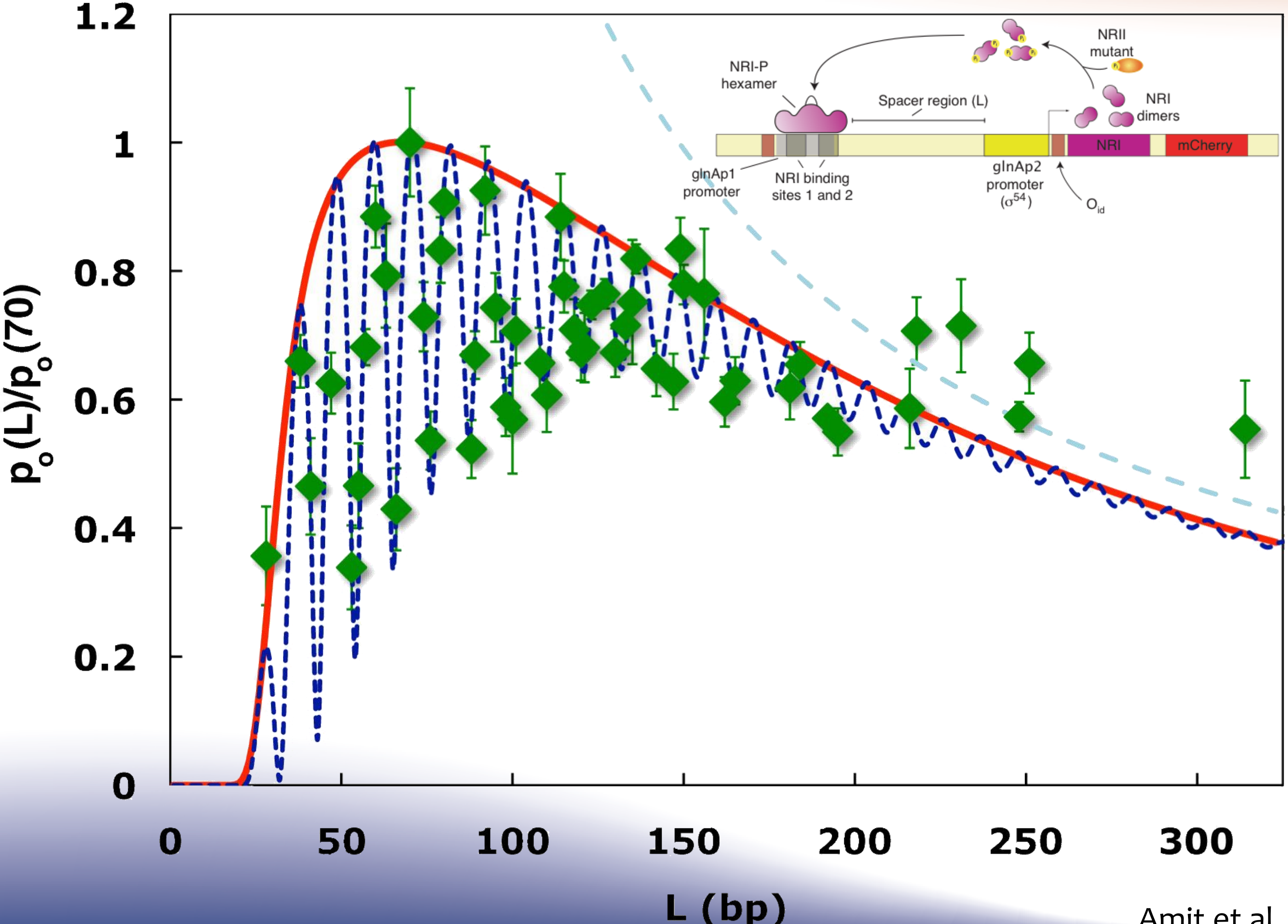
Struffi et al., Development, 2007.

Looping-based Transcription

- DNA is constantly changing conformations
- Every once in a while it loops (stochastic process)
- Looping probability $\Longrightarrow$ Transcriptional activity



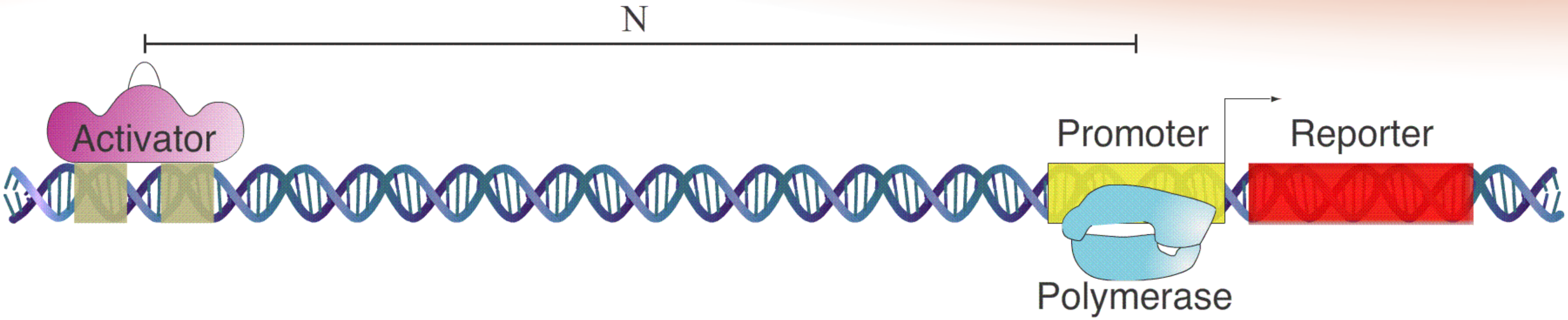
Looping Shown in Bacteria

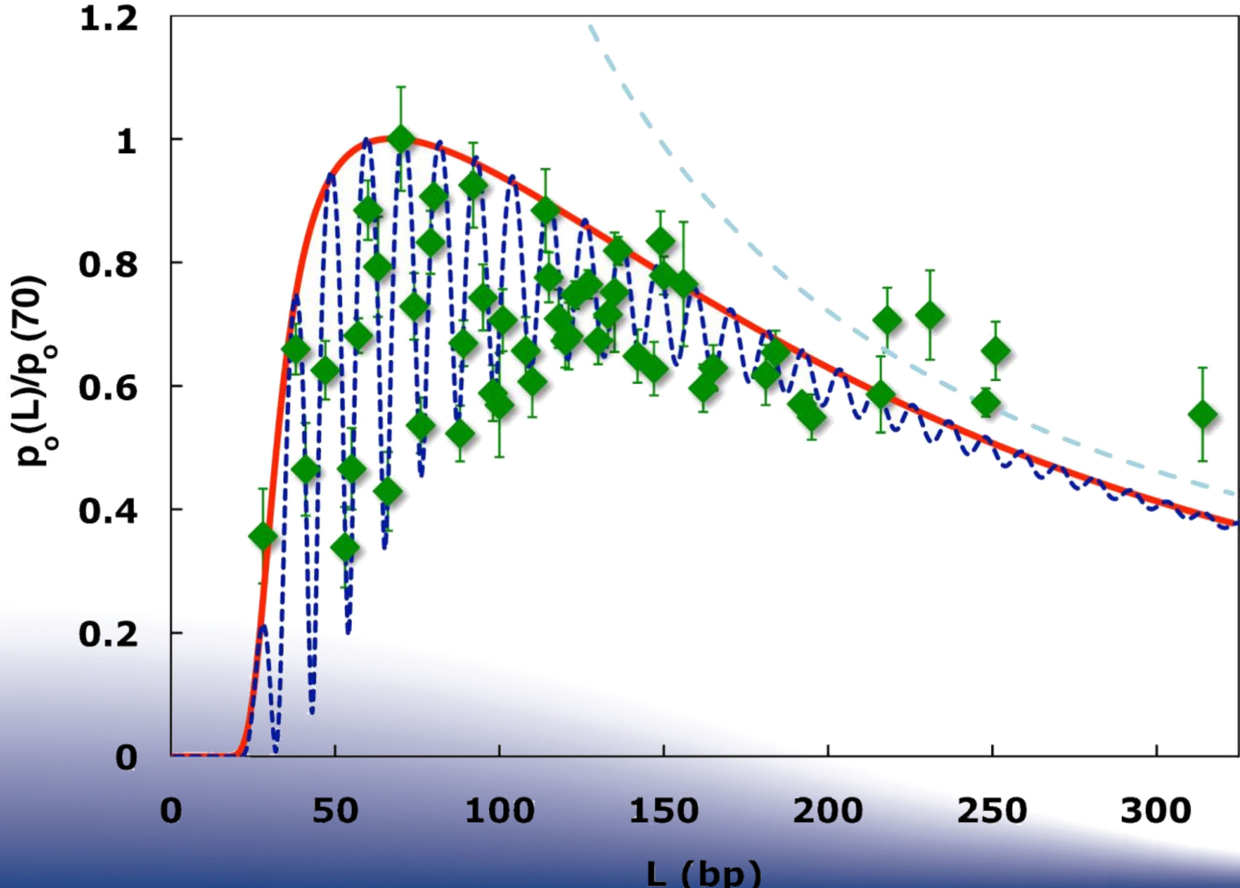


Amit et al., Cell 2011

Protein Binding to DNA

Oscillations

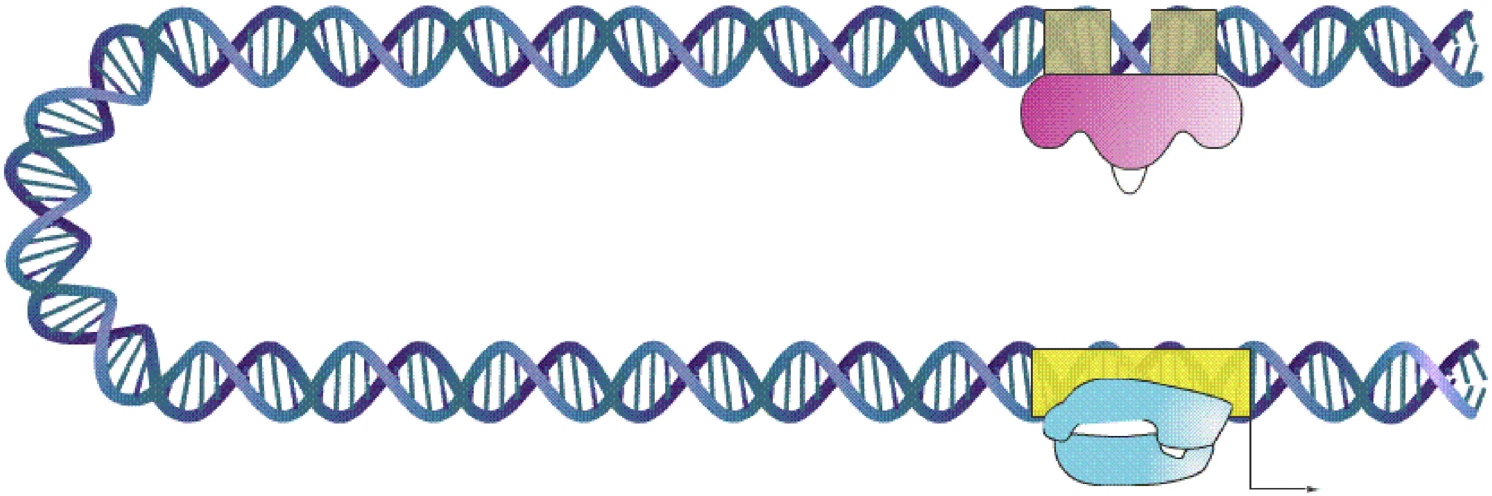




Amit et al., Cell 2011

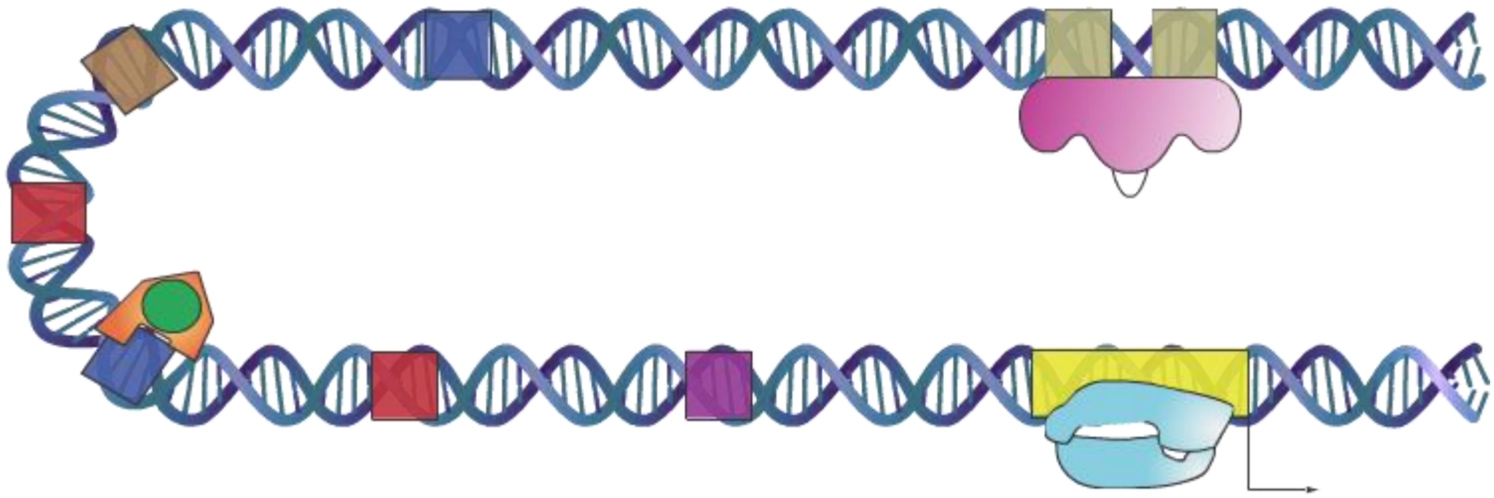
Transcription Regulation

- Regulation by Transcription Factors (TFs) binding DNA



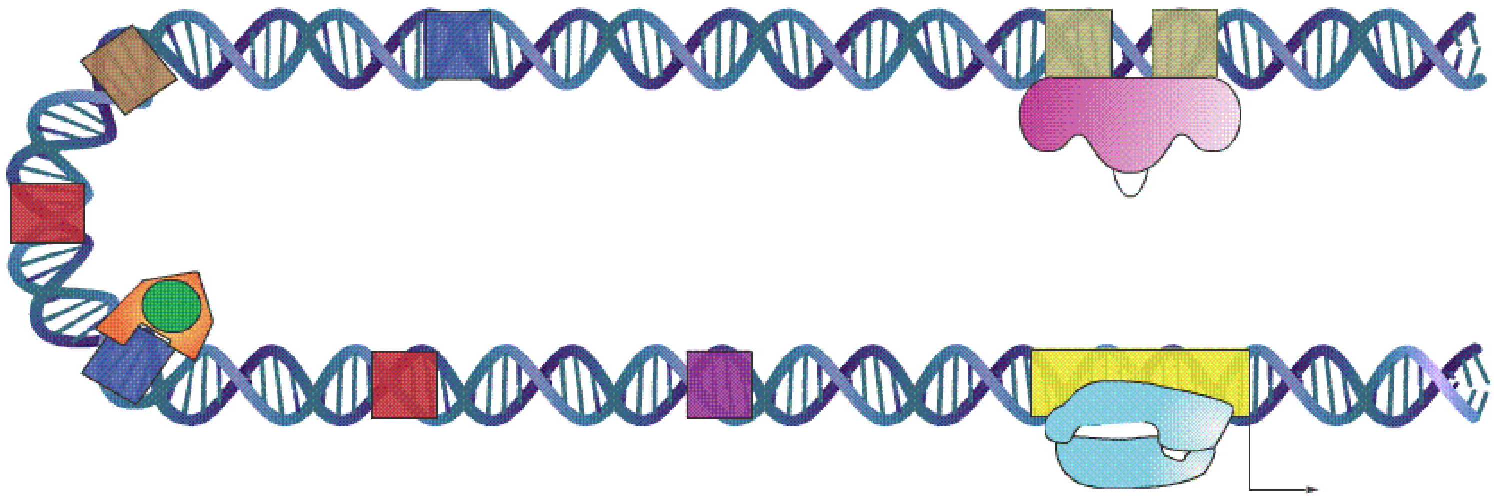
Transcription Regulation

- Regulation by Transcription Factors (TFs) binding DNA



Transcription Regulation

- Regulation by Transcription Factors (TFs) binding DNA

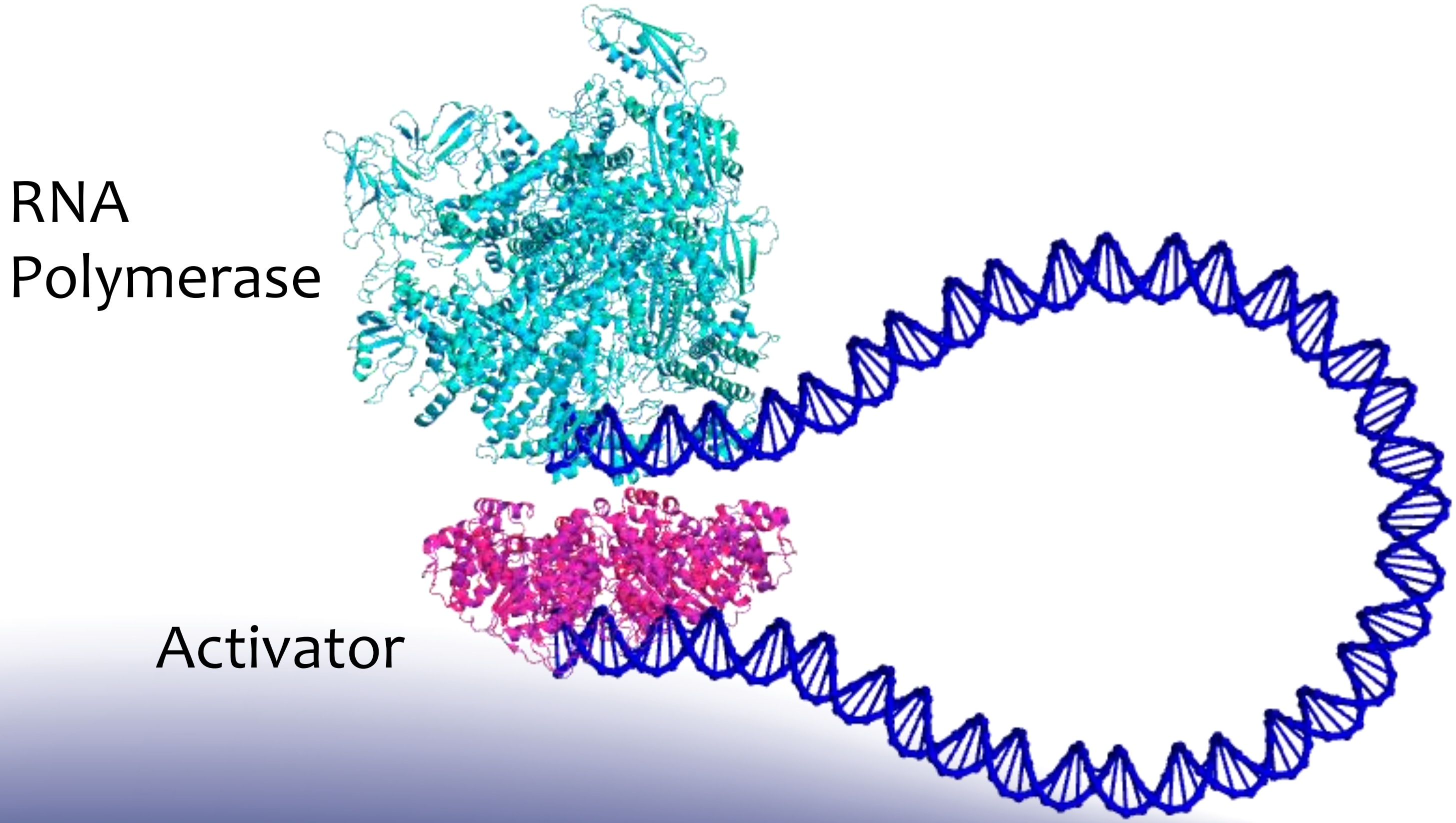


Physical Mechanism

- Known mechanisms: TF alters DNA locally
  - Bending – No long-range effect, few TFs bend DNA
  - Twisting – No long-range effect, few TFs twist DNA
  - Stiffening – No long-range effect

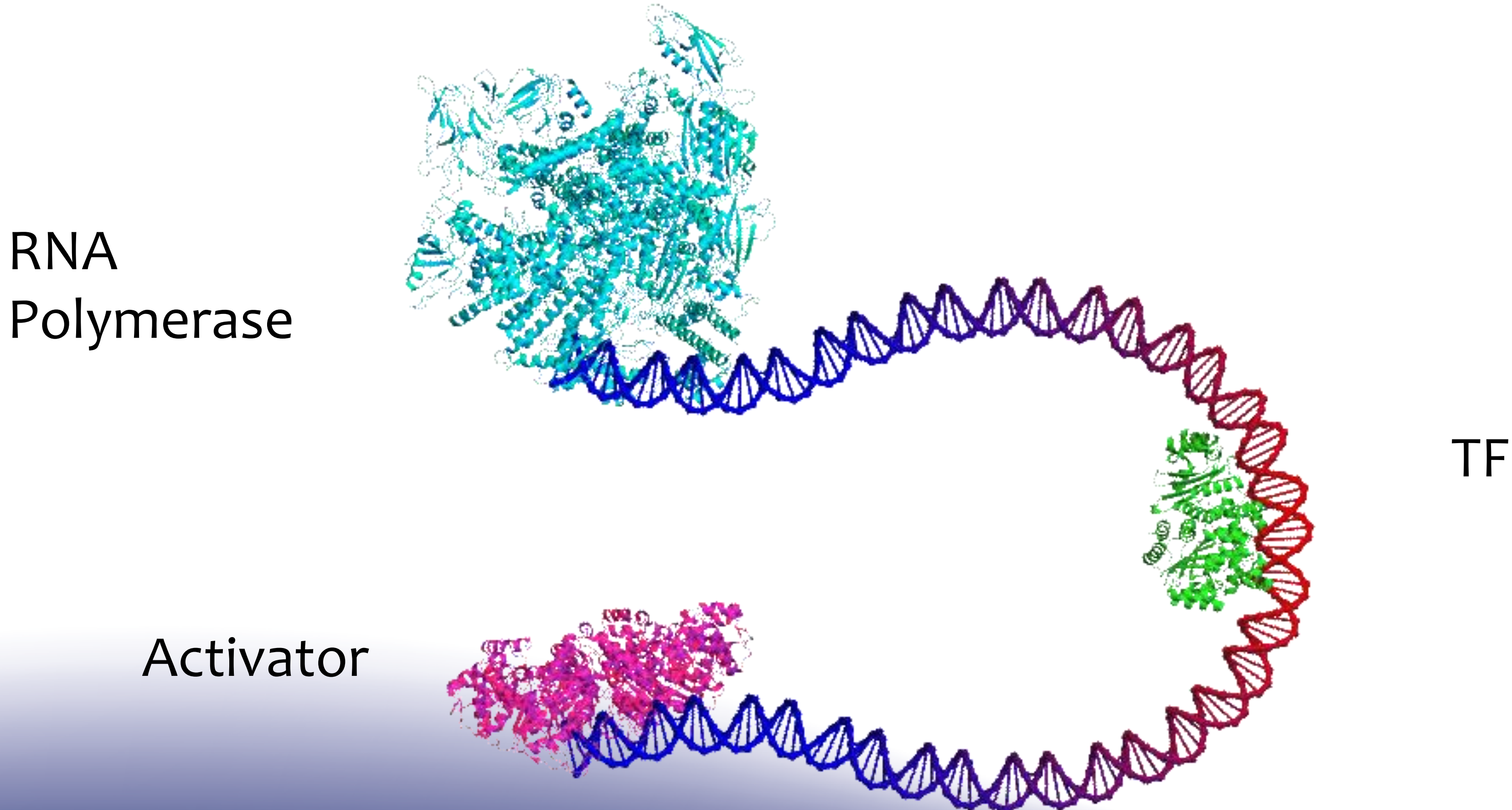
- Not addressing chromatin modification

Excluded Volume



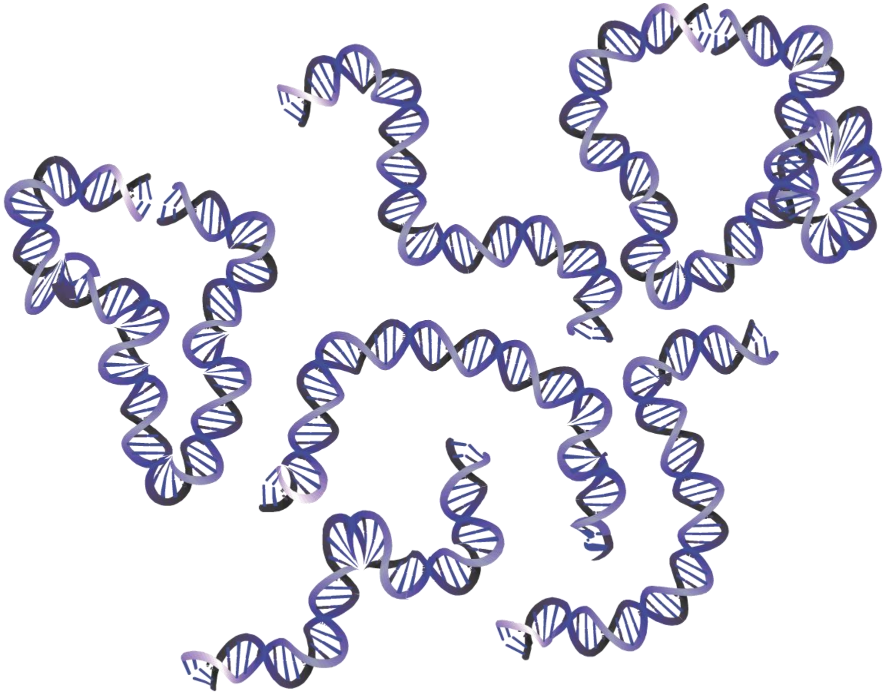
Excluded Volume

- Any* DNA binding protein can increase or decrease looping probability depending its orientation

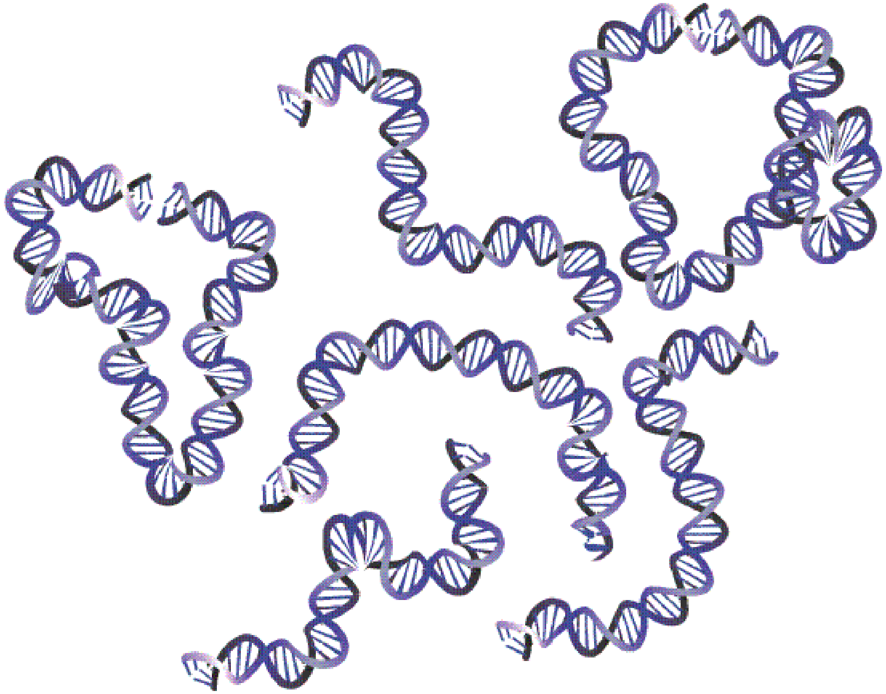


Modeling

Looping Probability



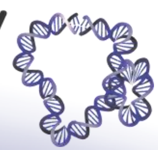
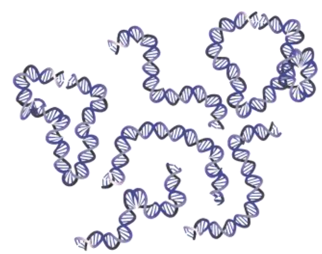
Looping Probability



Looping Probability

$$\frac{1}{k_b T}$$

$$P_{loop} = \frac{e^{-\beta E} + e^{-\beta E}}{e^{-\beta E} + e^{-\beta E} + e^{-\beta E} + e^{-\beta E} + e^{-\beta E} + e^{-\beta E}}$$

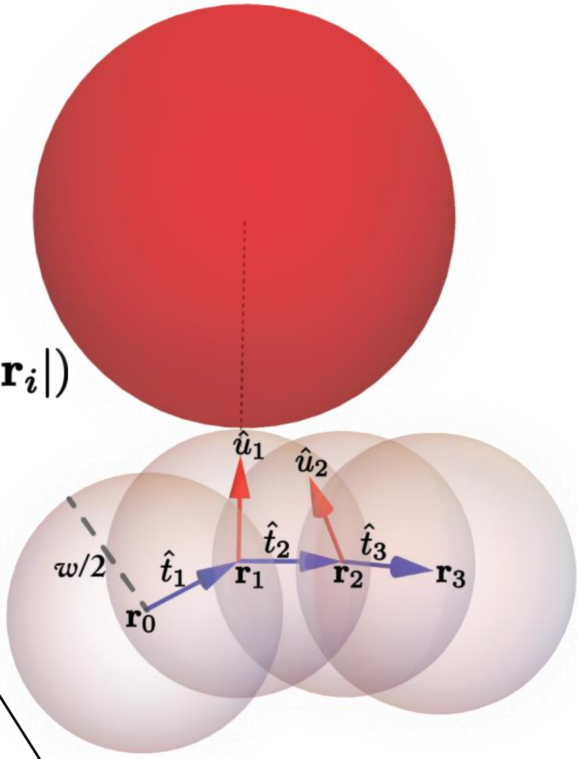
E  ? 

DNA Model

DNA

Discrete chain of straight links. For link i:

- Bending energy: $\beta E_i^{bend} = \frac{a}{2}\left|\hat{t}_i - \hat{t}_{i-1}\right|^2$
- Twisting energy: $\beta E_i^{twist} = c\left(\Omega_i - \Omega_0\right)^2$
- Chain volume: $\beta E_i^{hw} = \infty \sum_{j \neq i} \Theta\left(w - \left|\mathbf{r}_j - \mathbf{r}_i\right|\right)$



Regulator TF

- Local stiffening: $a \rightarrow a'$
- Local bending: $\hat{t}_i \rightarrow \hat{t}_i'$

Effective width > DNA width

Wang et al., Macromolecules 2011

TFs & RNAP

- Spherical volume

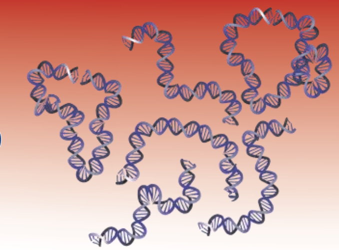
Pollak et al ,.PRE 2014

Brunwasser-Meirom, Pollak et al., Nat. Comm 2016

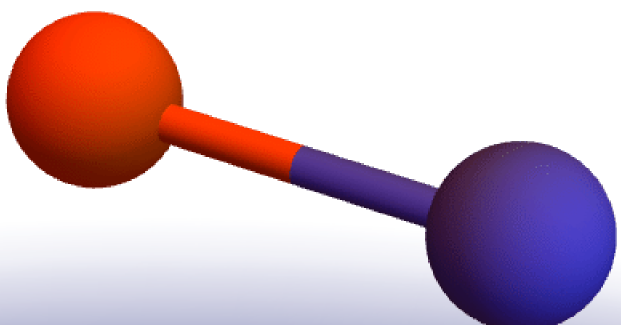
Monte-Carlo Simulations

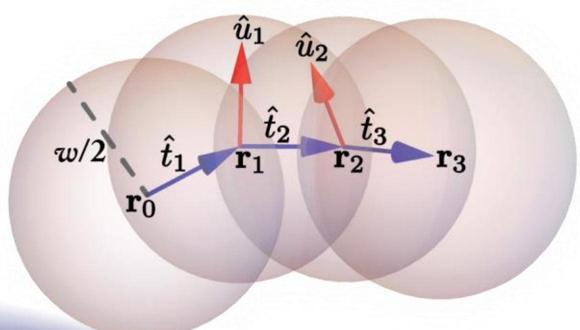
- Chain & TF volumes make considering all possible DNA conformations very difficult
- Static Monte-Carlo (Sequential Importance Sampling)

Monte-Carlo Simulations



- Static Monte-Carlo (Sequential Importance Sampling)
- Chains generated from scratch, link after link
- Samples of ~10^9 chains with & without TFs
- Comparing looping probability with & without TFs determines regulatory effect.





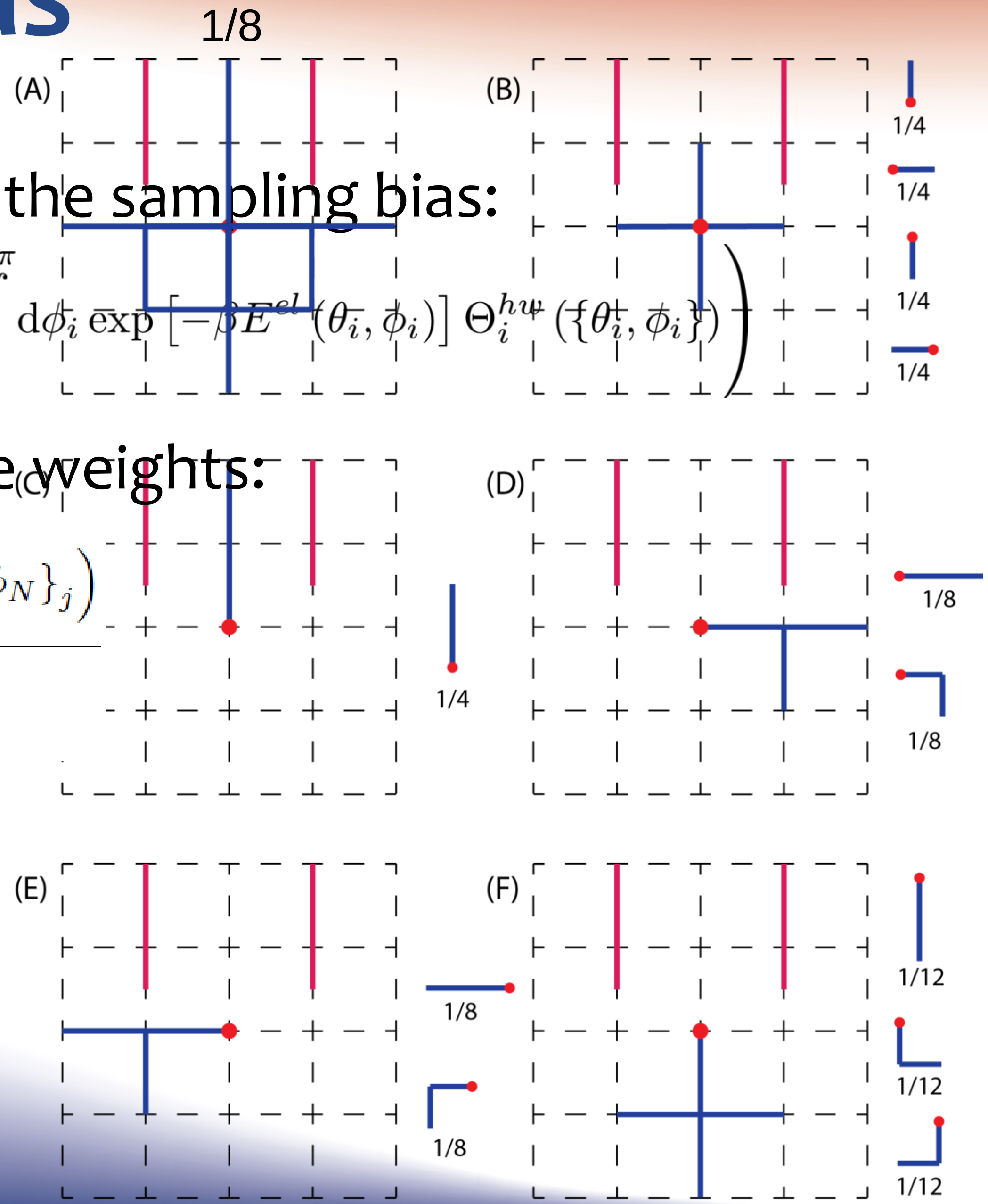
Sampling Bias

- Need weights to offset the sampling bias:

$$W\left(\{\theta_N,\phi_N\}\right)=\prod_{i=2}^{N}\left(\int_{-1}^{1}\mathrm{d}\cos\theta_i\int_{0}^{2\pi}\mathrm{d}\phi_i\exp\left[-\beta E^{el}(\theta_i,\phi_i)\right]\Theta_i^{hw}\left(\{\theta_i,\phi_i\}\right)\right)$$

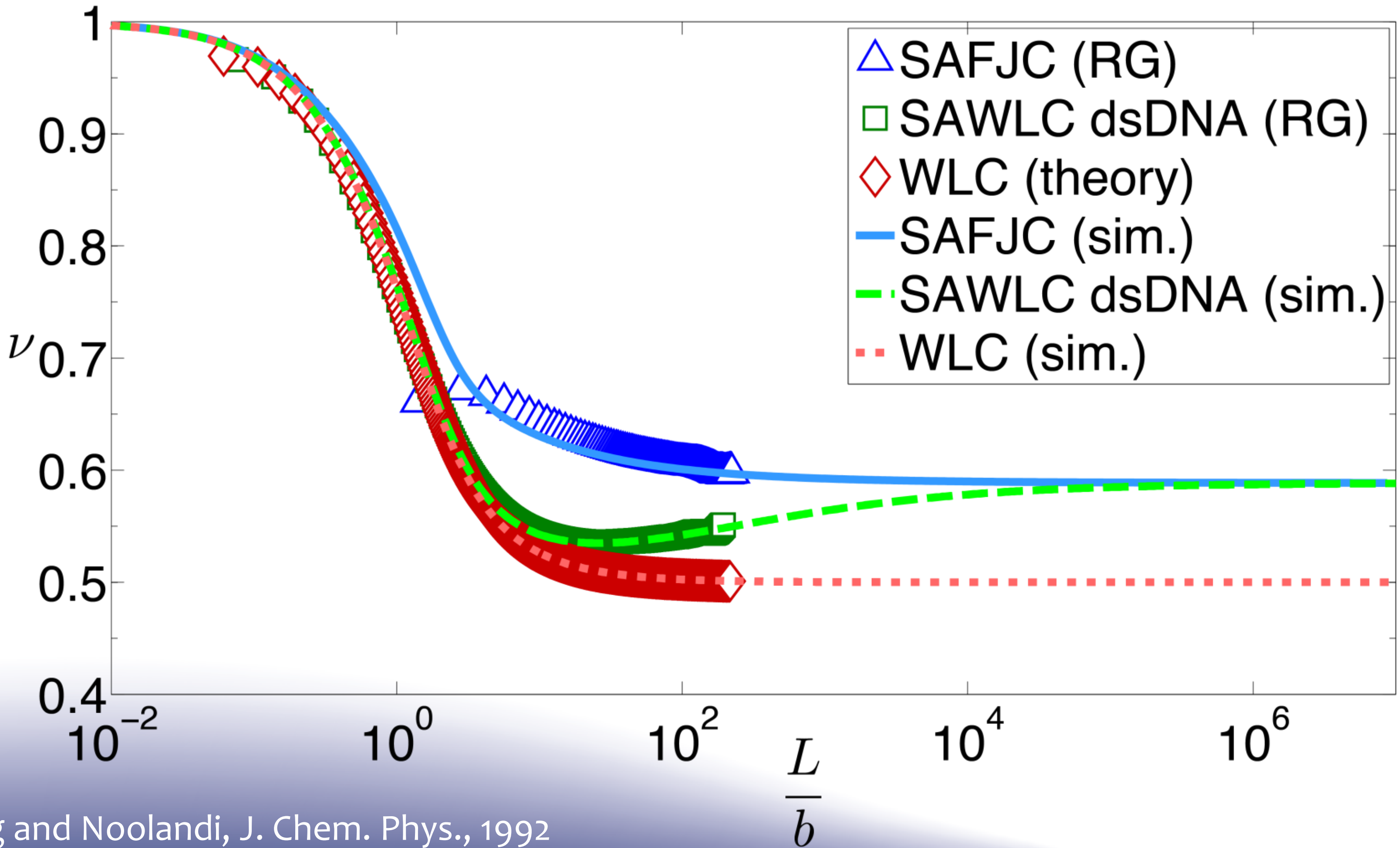
- Ensemble estimates use weights:

$$\langle f\rangle=\frac{\sum_{j=1}^{N_c} f\left(\{\theta_N,\phi_N\}_j\right)W\left(\{\theta_N,\phi_N\}_j\right)}{\sum_{j=1}^{N_c} W\left(\{\theta_N,\phi_N\}_j\right)}$$



DNA Model Verification

$$\sqrt{\langle R^2 \rangle} \propto \left(\frac{L}{b}\right)^{\nu}$$

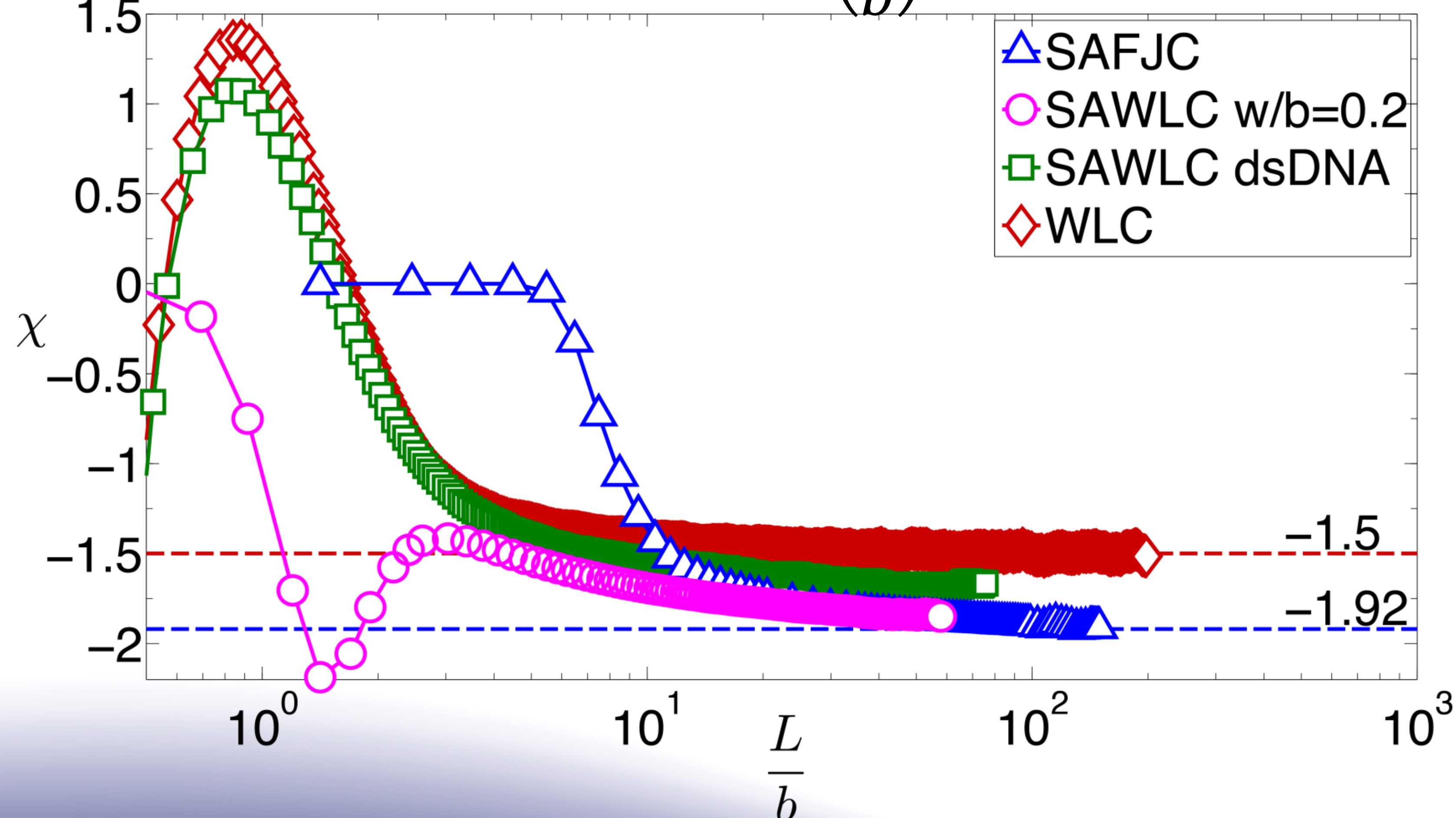


Zheng and Noolandi, J. Chem. Phys., 1992

Pollak et al., PRE 2014

DNA Model Verification

$$P_{looping} \propto \left(\frac{L}{b}\right)^{\chi}$$



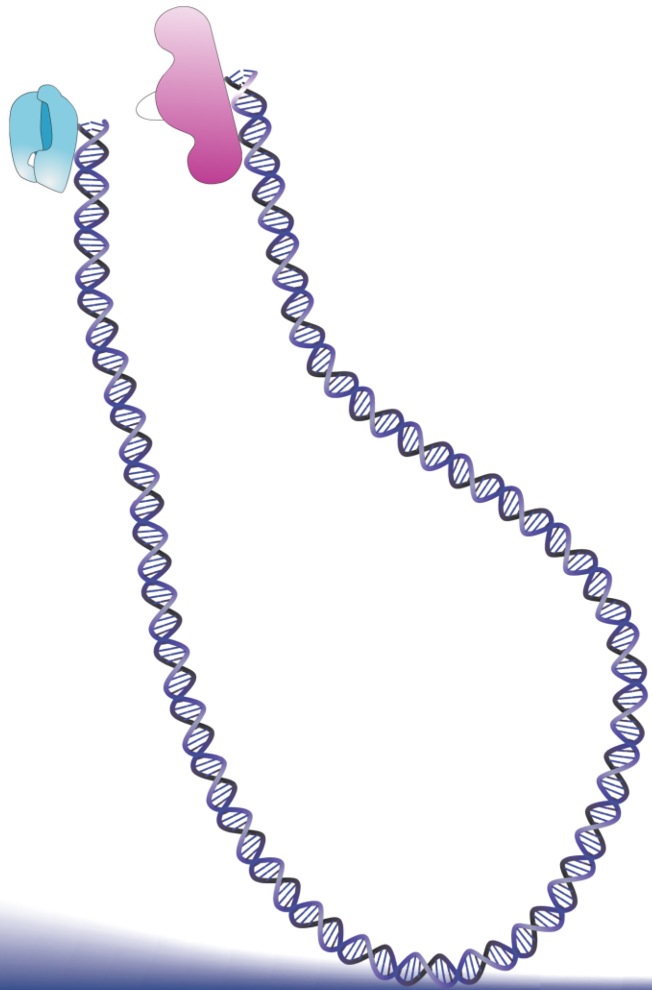
Gennes, Cornell U. Press, 1979, Sinclair et al., JACS, 1985

Pollak et al., PRE 2014

Short-Range
Looping Results

Simulation Looping Condition

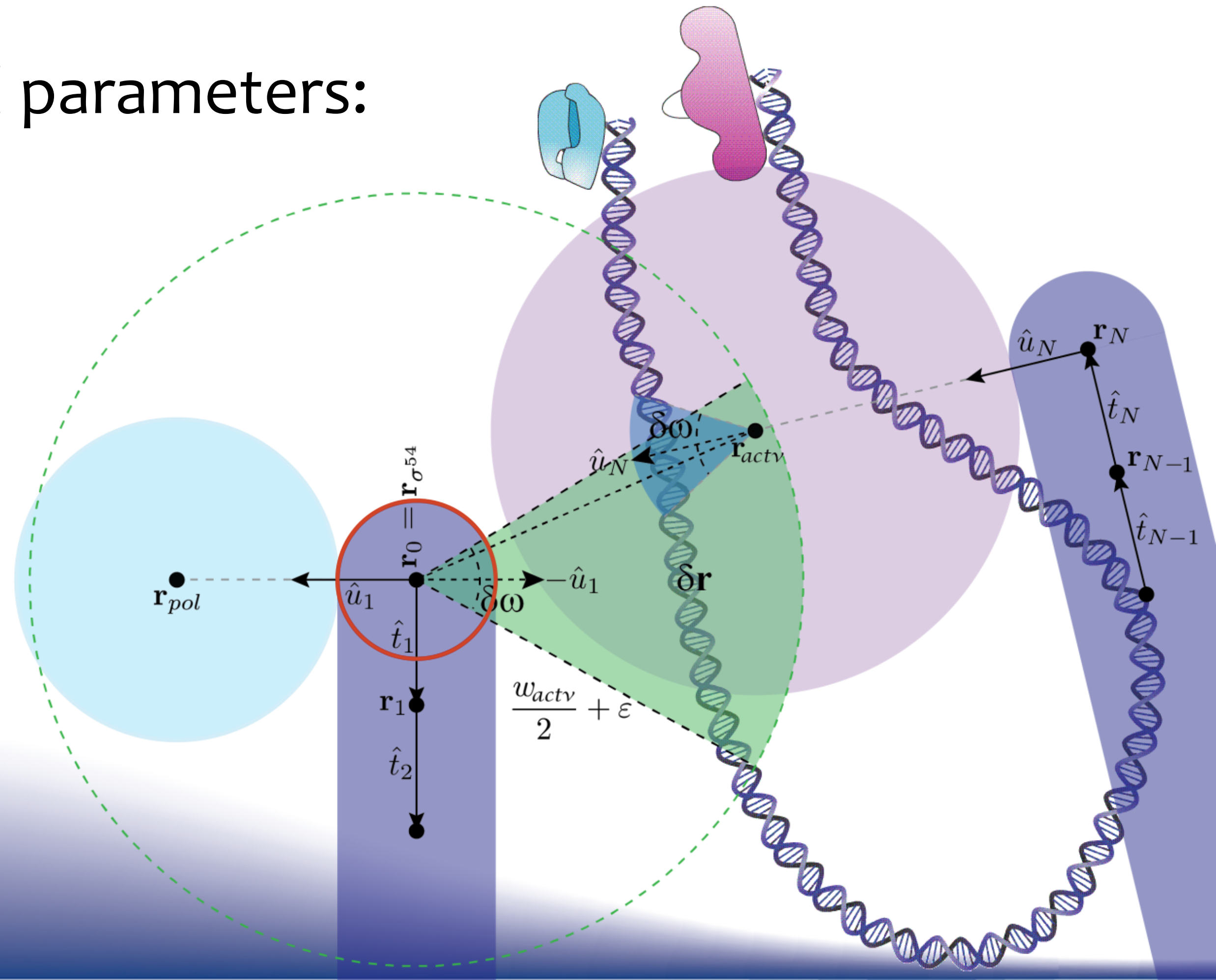
- Looping conditions mimicking bacterial σ^{54} promoters



Simulation Looping Condition

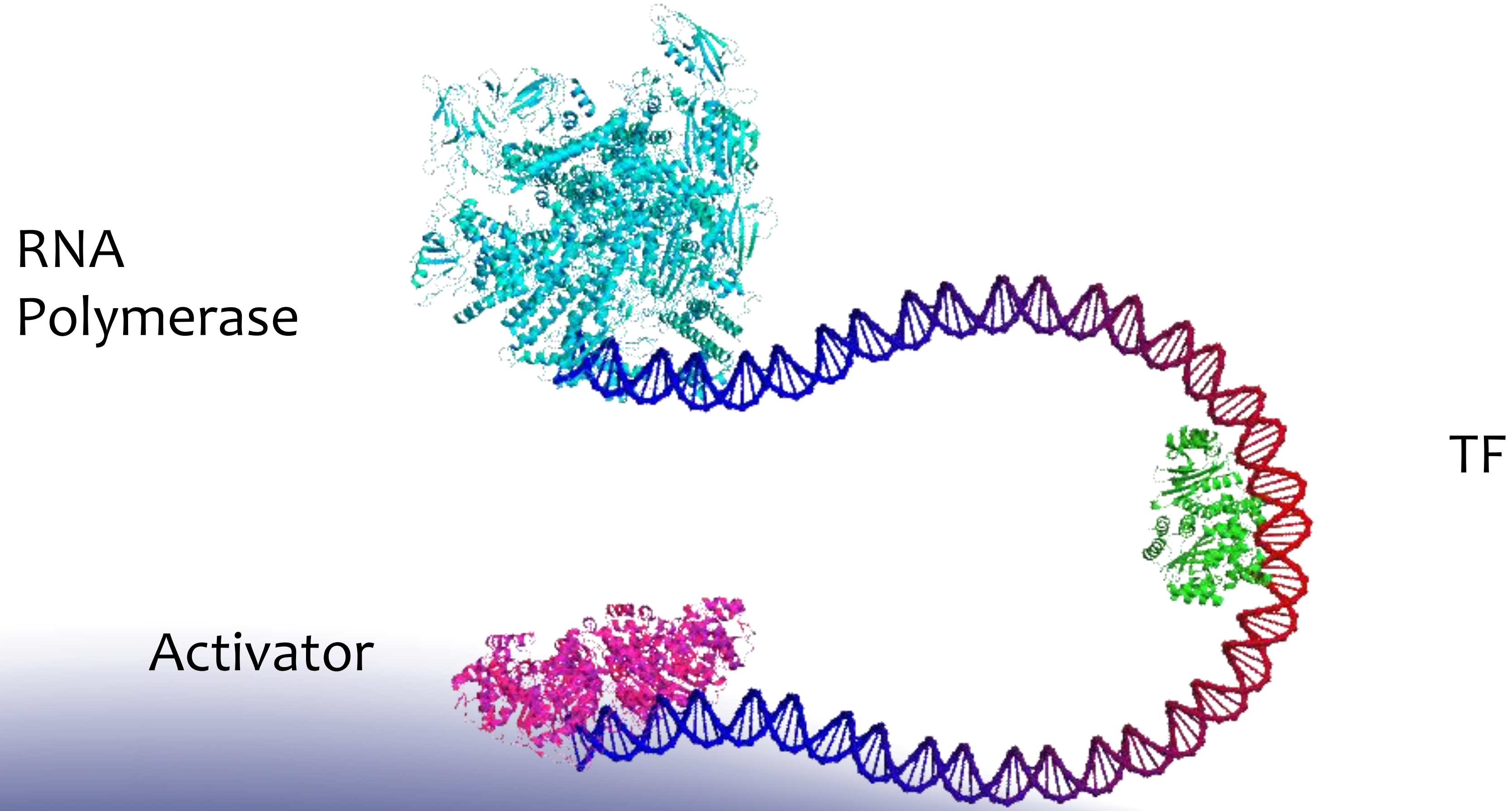
- Looping conditions mimicking bacterial σ^{54} promoters

- Additional model parameters:
  - $\delta\mathbf{r}\ (\delta\omega, \varepsilon)$
  - $\delta\omega'$
  - $\mathbf{r}_{pol}, \mathbf{r}_{actv}$



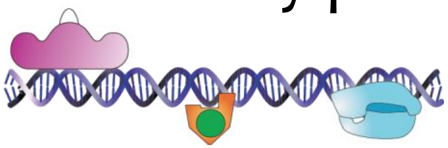
Excluded Volume - Reminder

- Any* DNA binding protein can increase or decrease looping probability depending its orientation

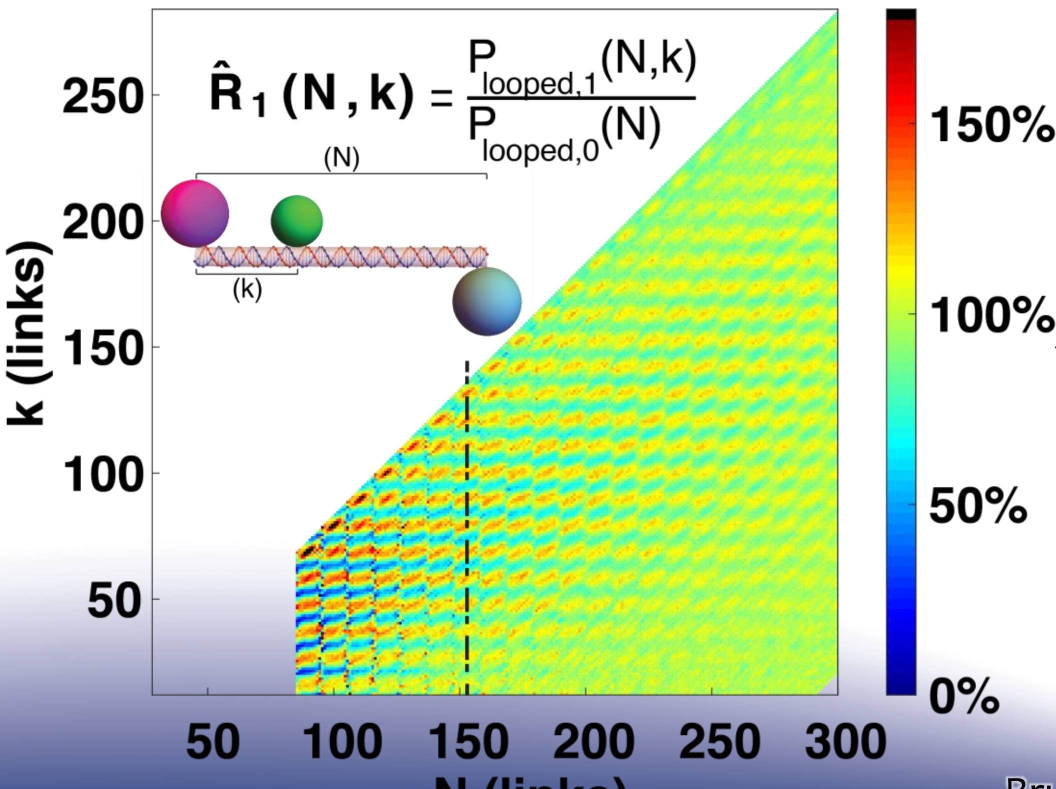


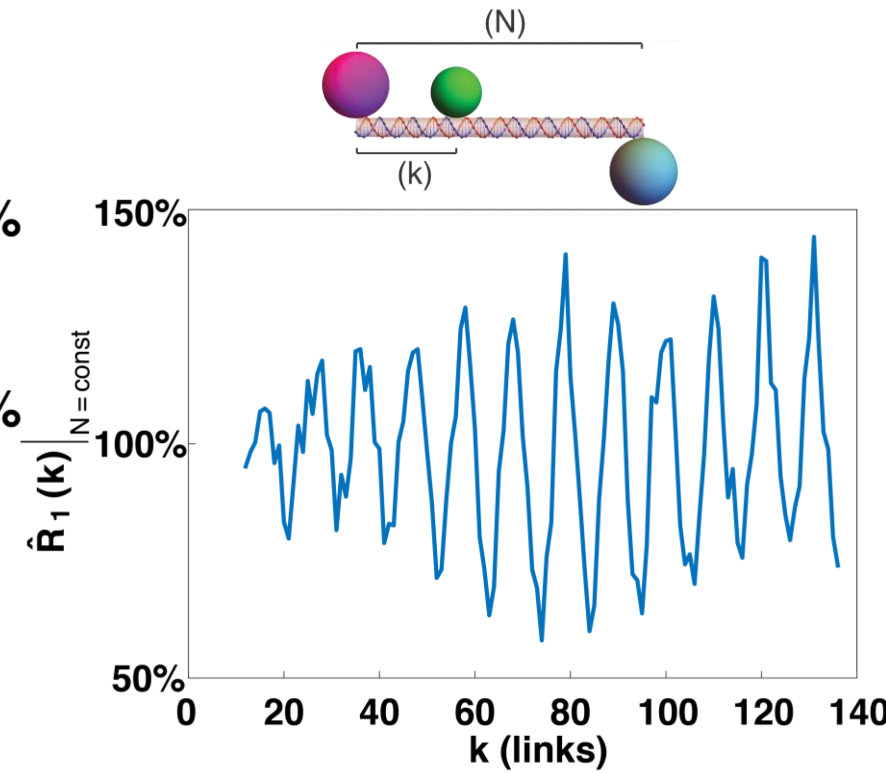
Excluded Volume

- 11 bp oscillatory pattern of up & down-regulation

Up: 

down: 

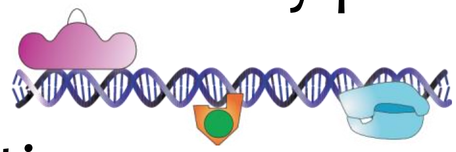
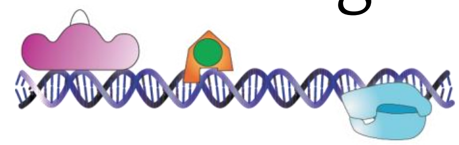




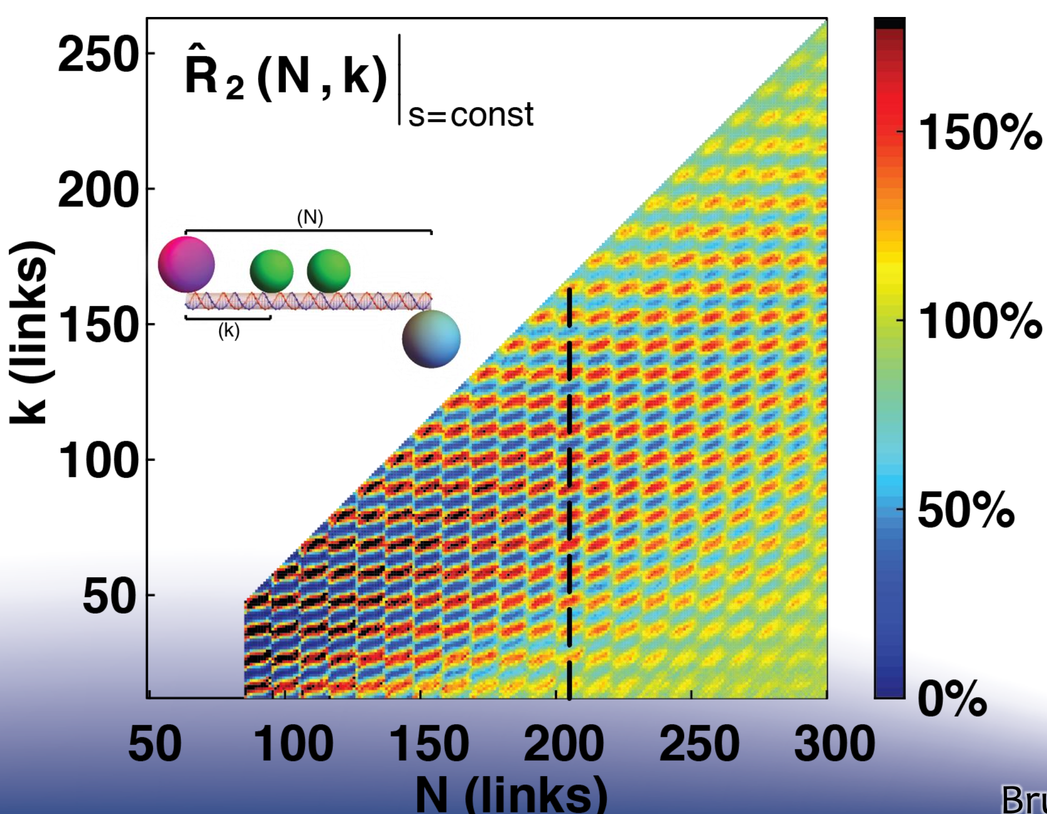
Brunwasser-Meirom, Pollak et al., Nat. Comm 2016

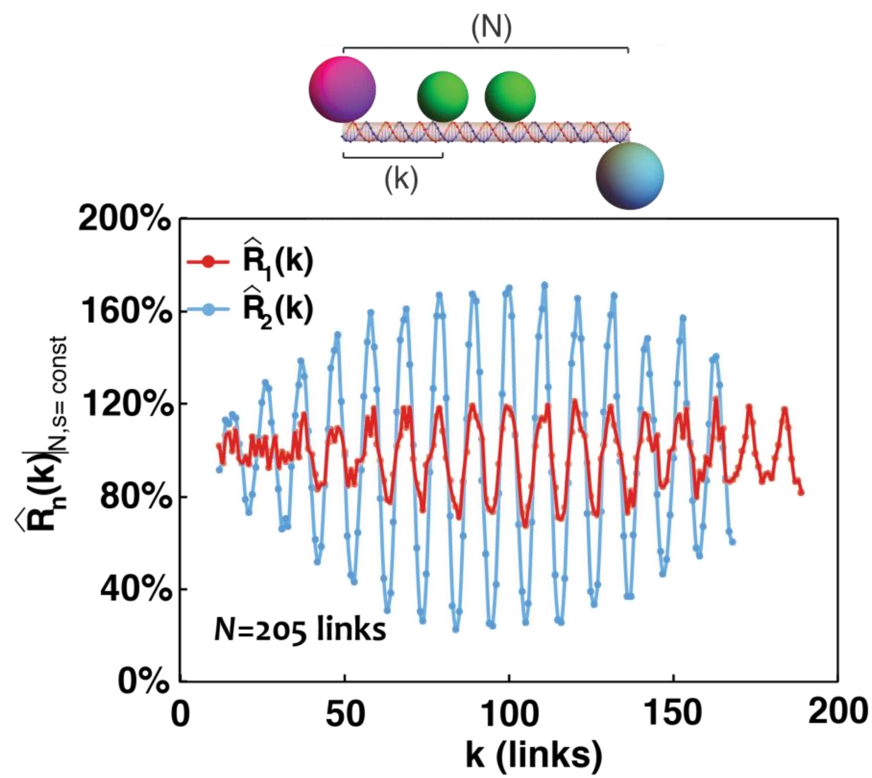
Excluded Volume

- 11 bp oscillatory pattern of up & down-regulation

Up:  down: 

- Additive

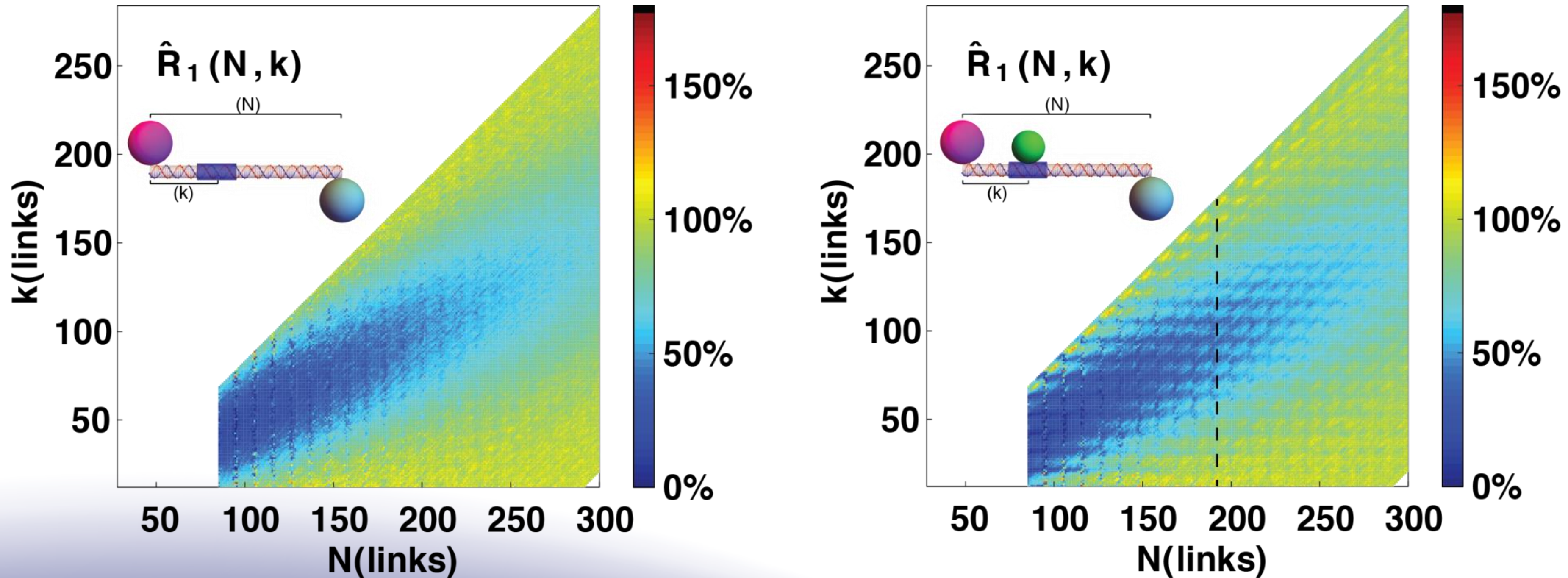




Brunwasser-Meirom, Pollak et al., Nat. Comm 2016

Stiffening

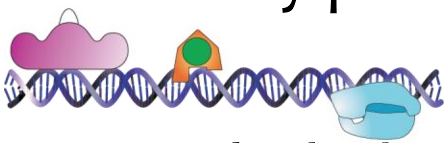
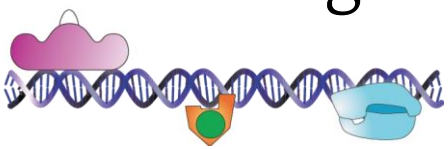
- Down-regulation, no oscillatory pattern
- Additive with the volume effect



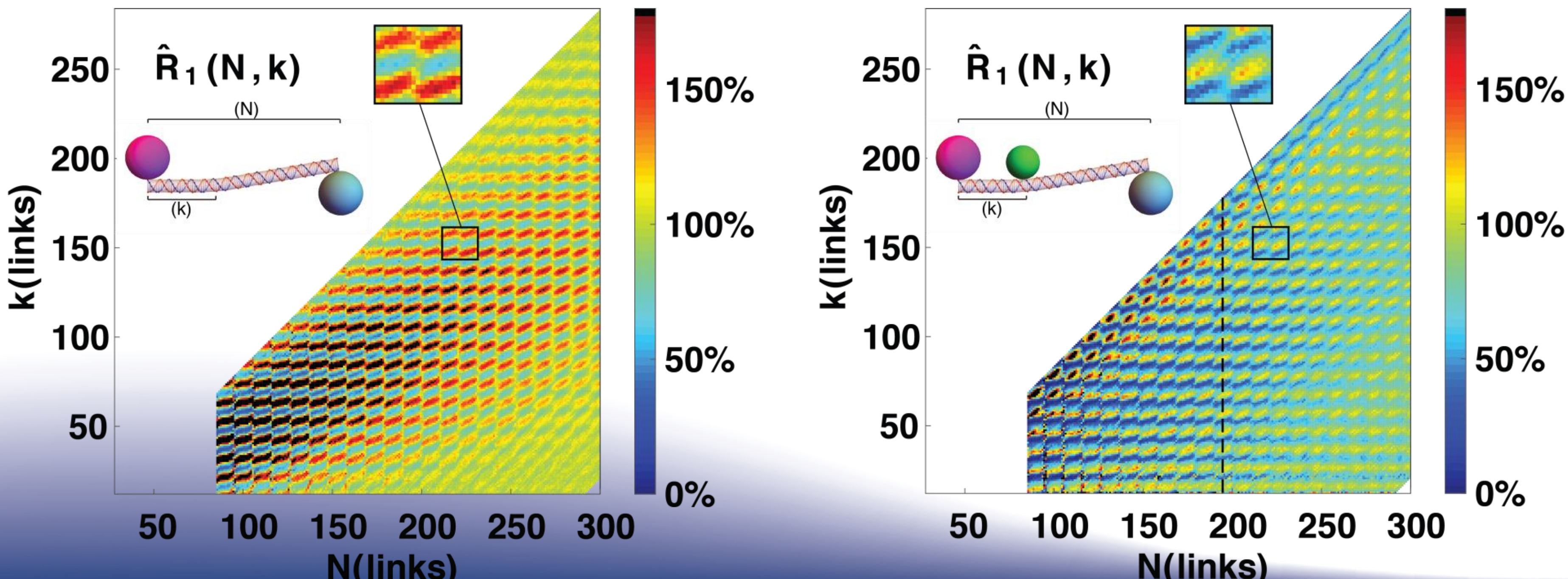
Brunwasser-Meirom, Pollak et al., Nat. Comm 2016

Bending

- 11 bp oscillatory pattern of up & down-regulation

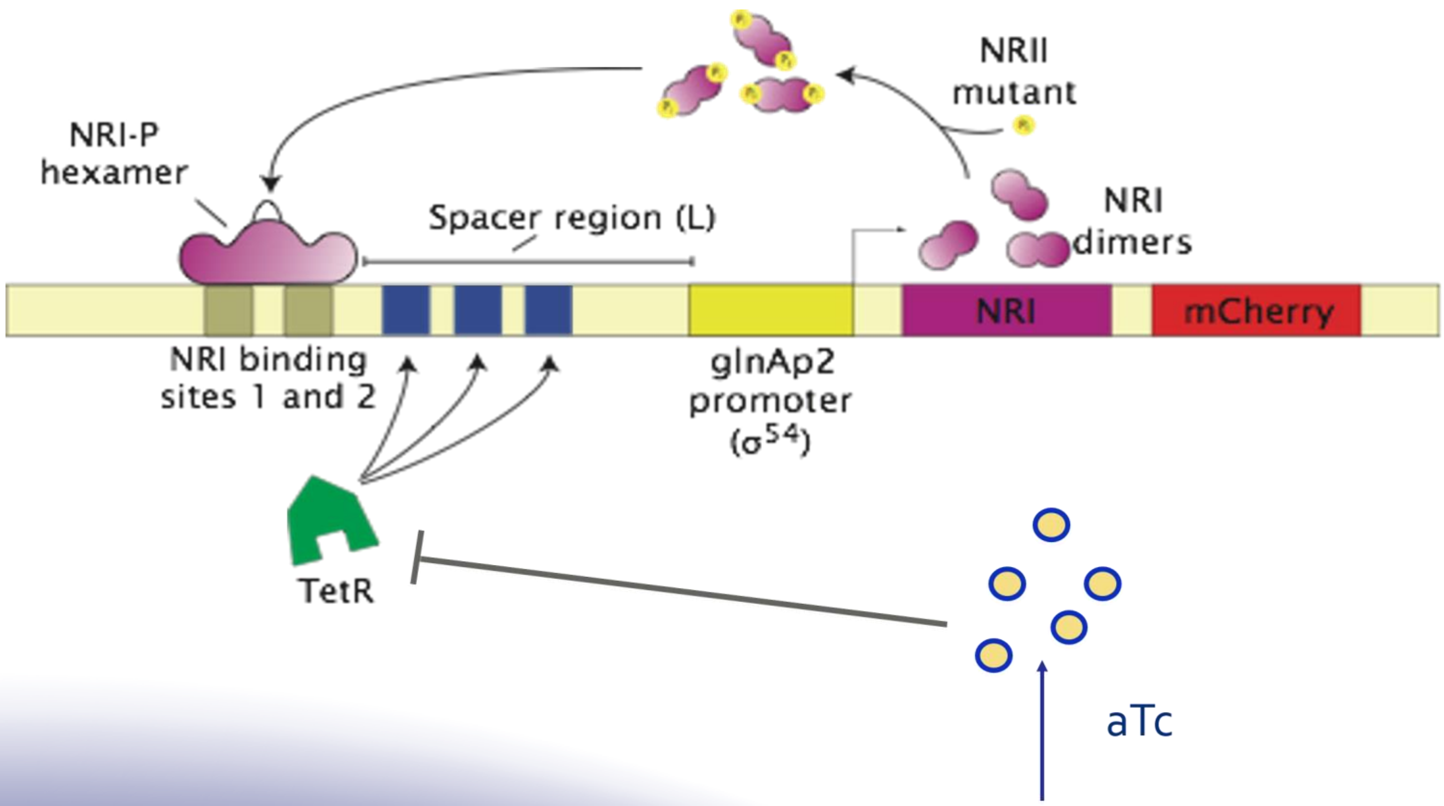
Up:  down: 

- Opposes excluded-volume



Synthetic Biology Experiments

Michal Meirom

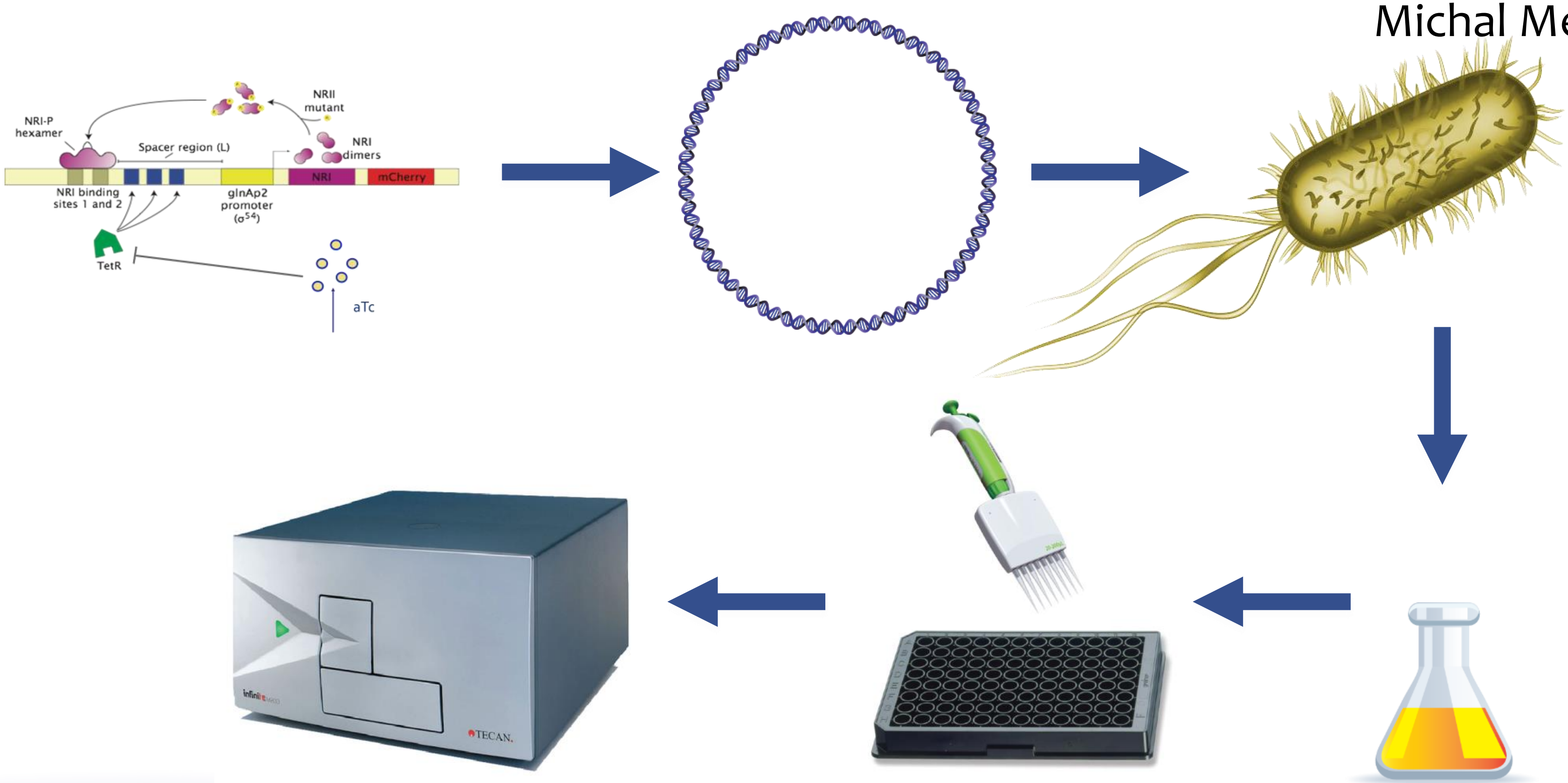


Amit et al., Cell 2011

Brunwasser-Meirom, Pollak et al., Nat. Comm 2016

Synthetic Biology Experiments

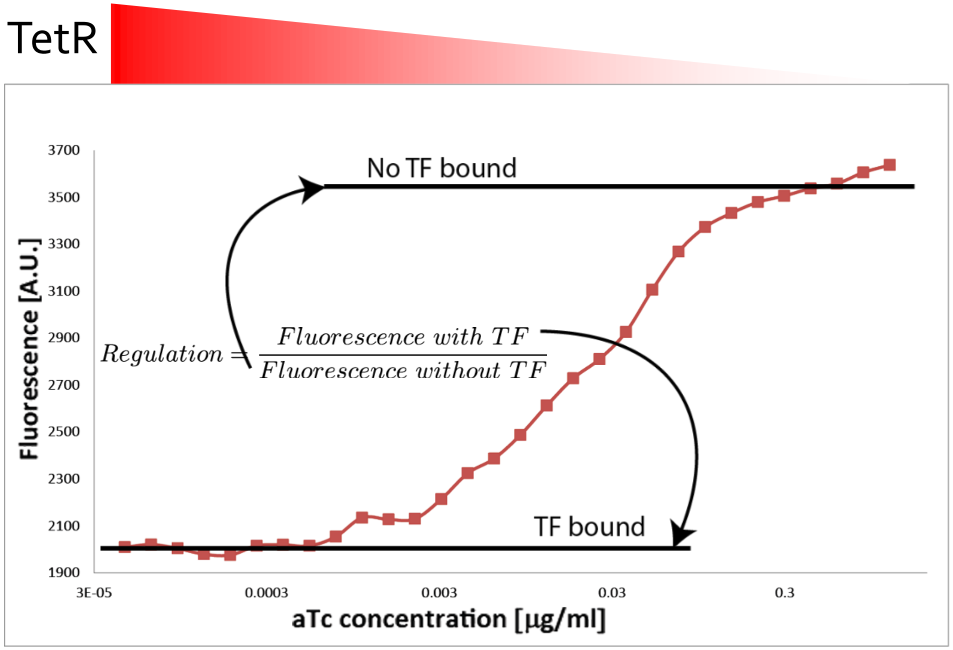
Michal Meirom

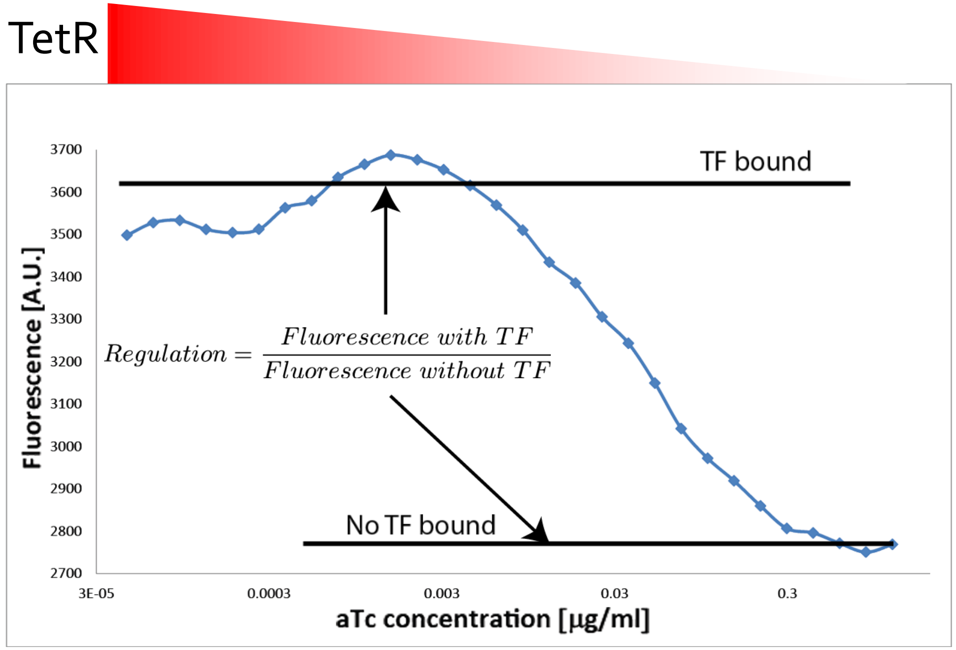


Brunwasser-Meirom, Pollak et al., Nat. Comm 2016

Measurements

- OD normalized fluorescence
- Fluorescence ratio ≈ looping probability ratio





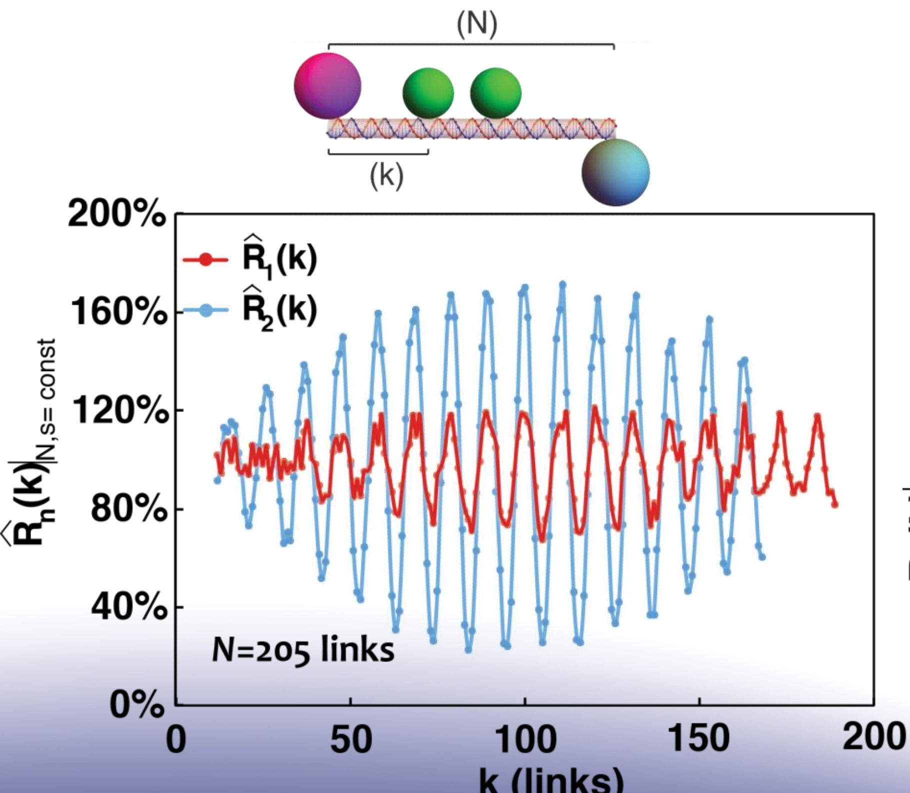
TetR, k=64 bp | Δk=25bp | TetR, k=89 bp

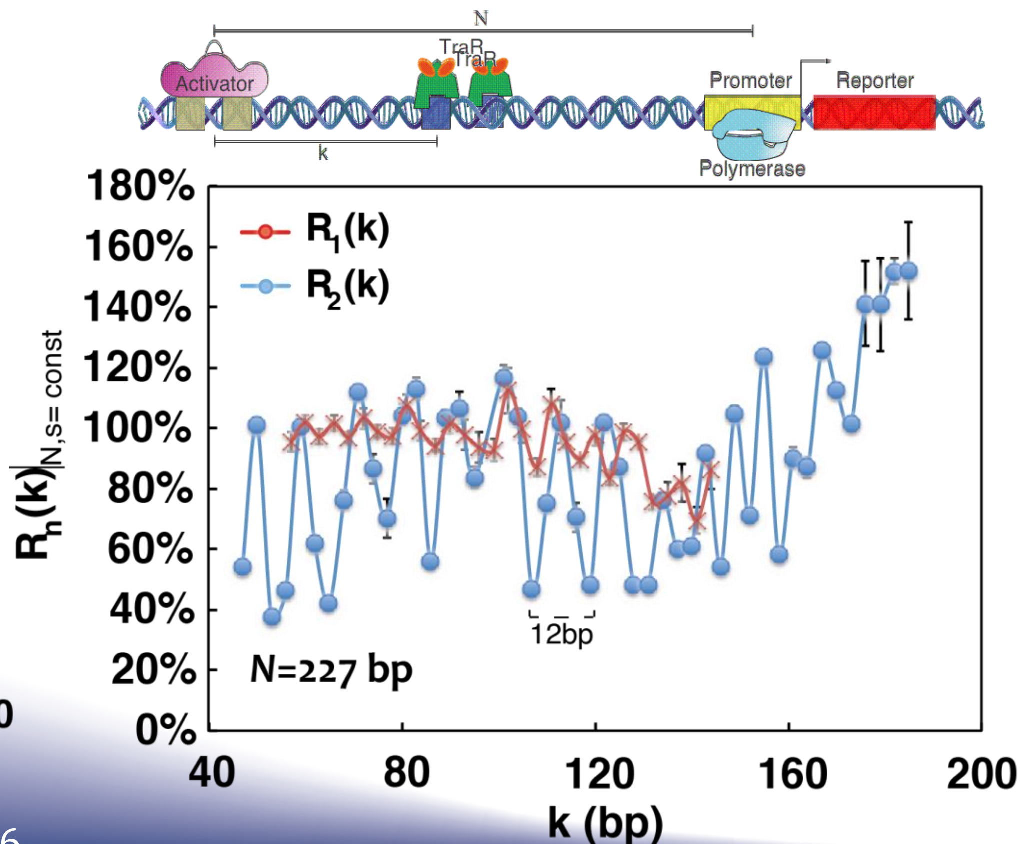
Down-regulation | Up-regulation

Brunwasser-Meirom, Pollak et al., Nat. Comm 2016

TraR Experimental Results

- 11 bp oscillatory pattern of up & down-regulation
- Additive effect

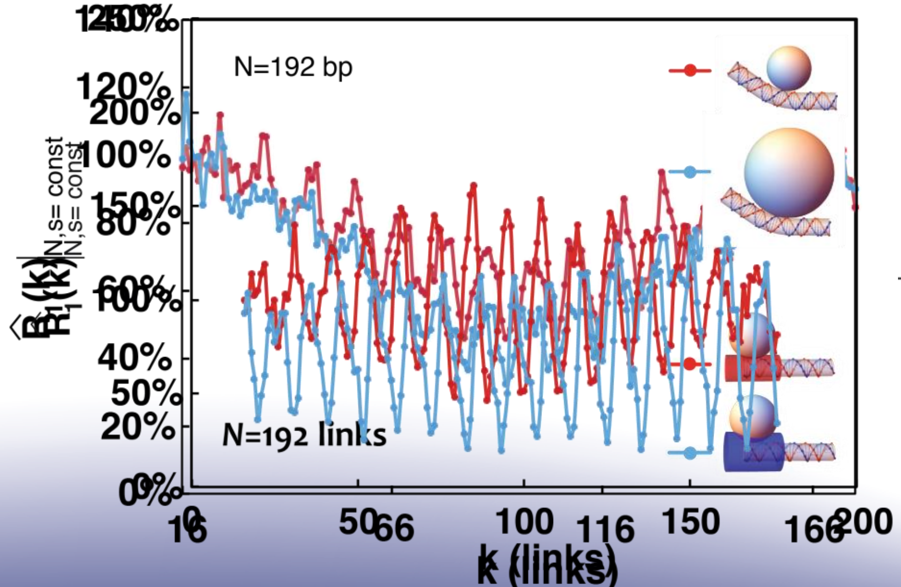


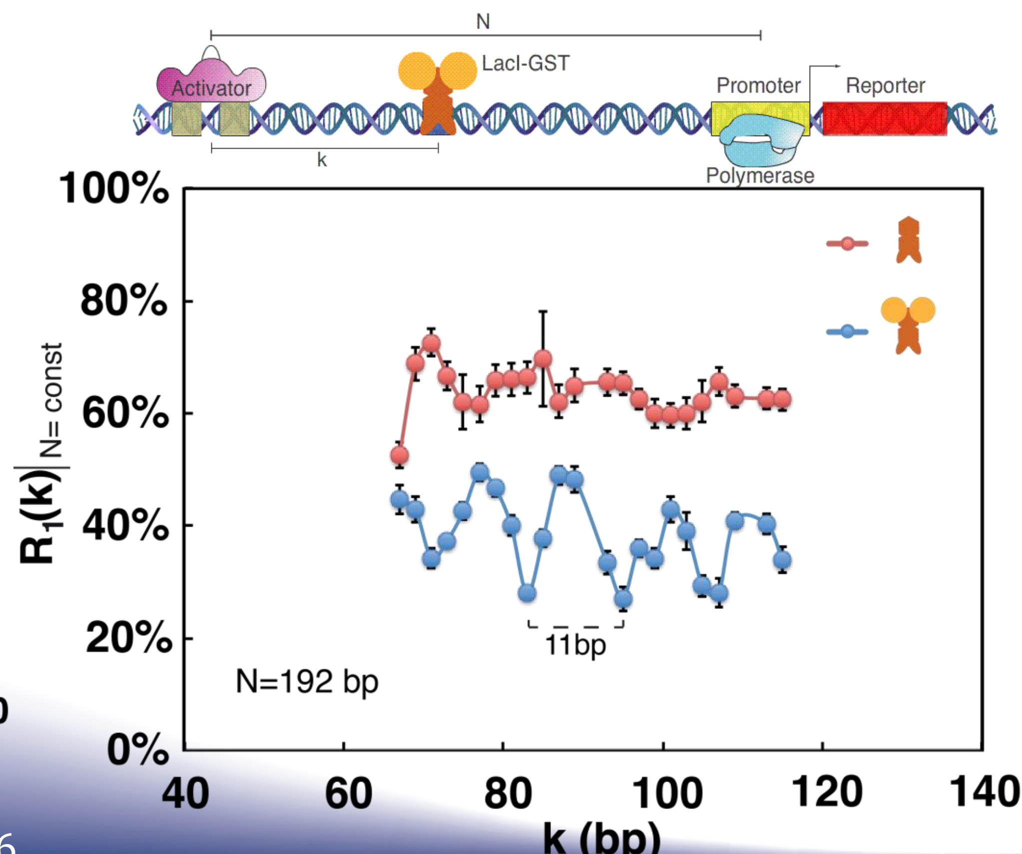


Brunwasser-Meirom, Pollak et al., Nat. Comm 2016

LacI Experimental Results

- Larger volume ⟹ larger amplitude
- Oscillations phase flip
- Shift towards down regulation

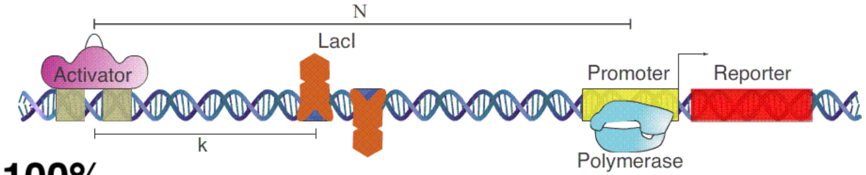


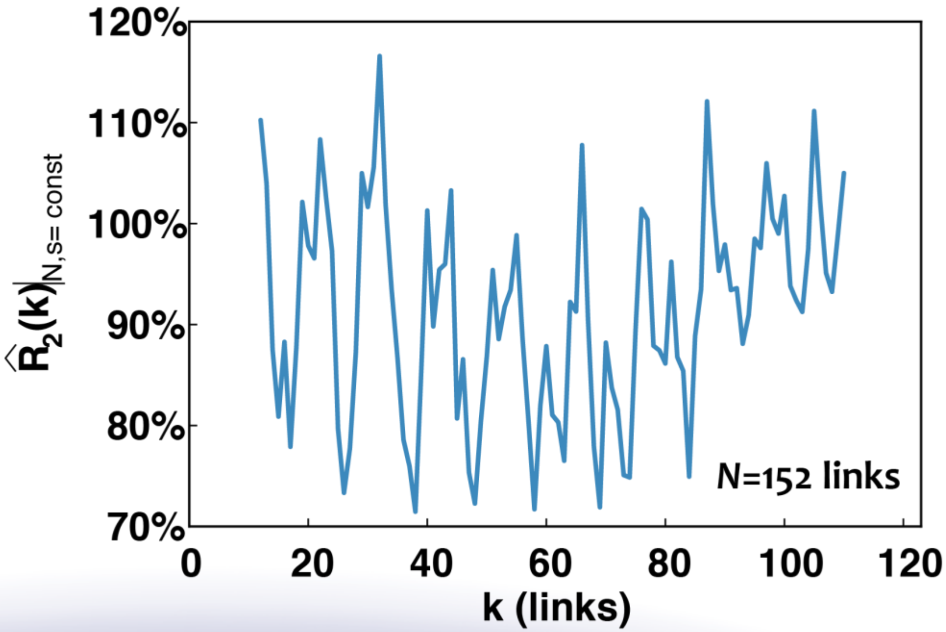


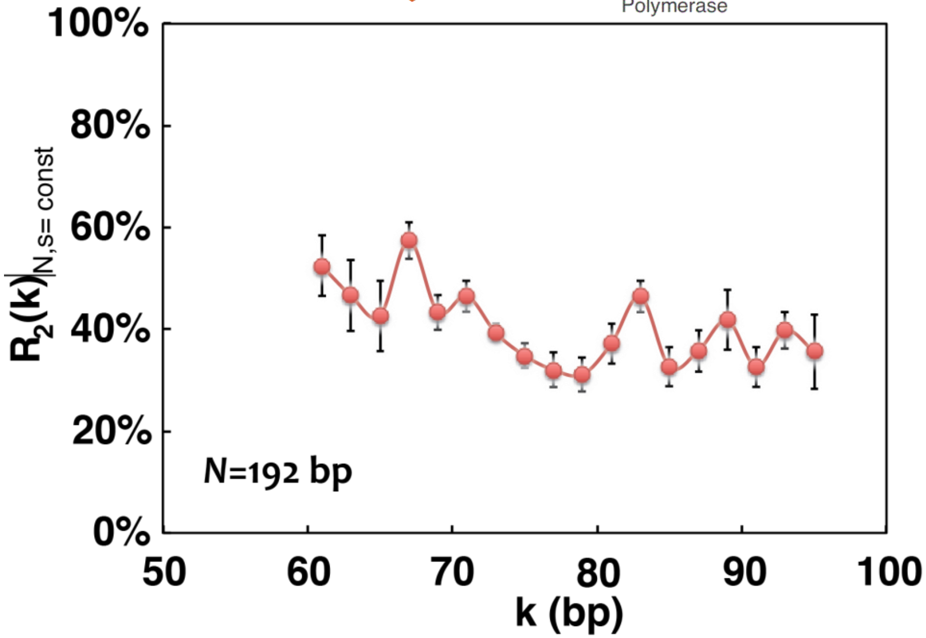
Brunwasser-Meirom, Pollak et al., Nat. Comm 2016

Multiple TFs – Out of Phase

- Deletory, weak down-regulation with ½ DNA helical repeat periodicity.

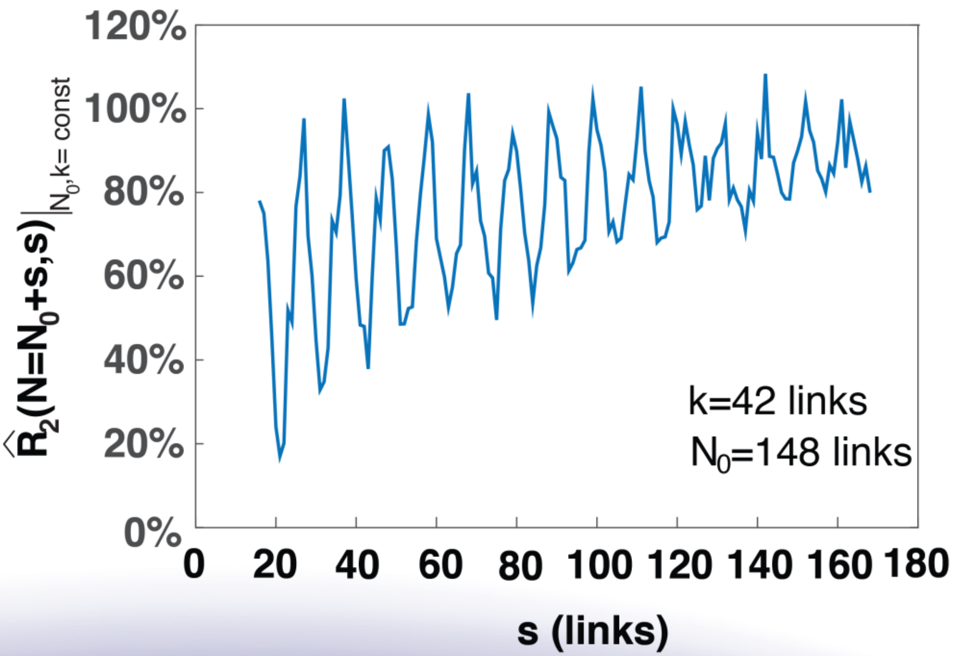


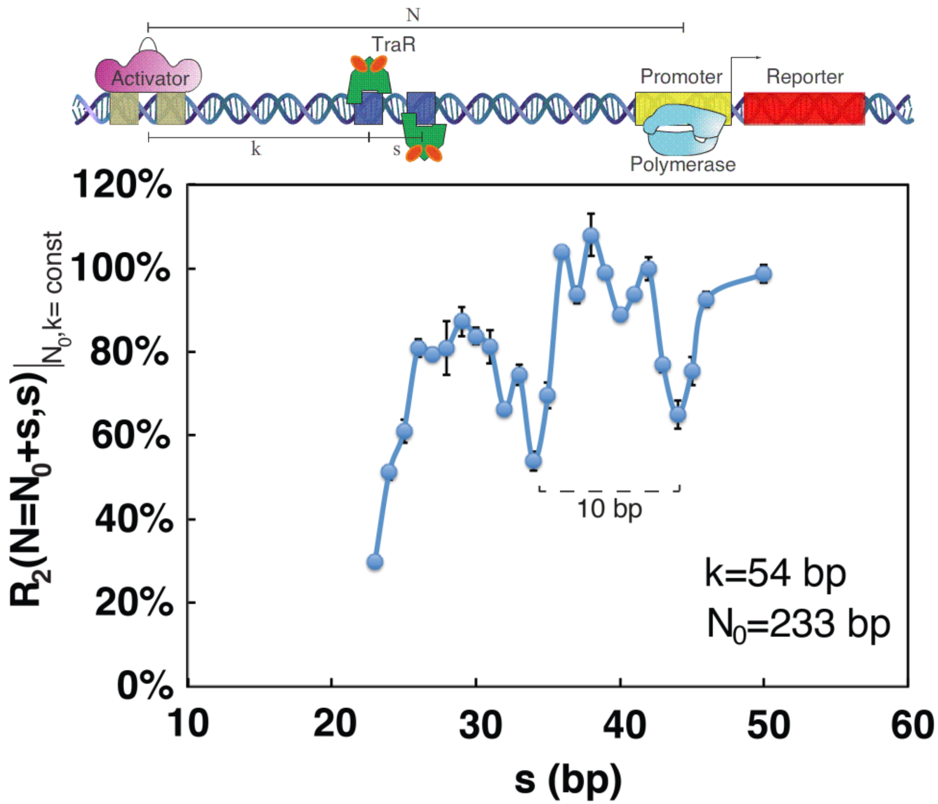




Brunwasser-Meirom, Pollak et al., Nat. Comm 2016

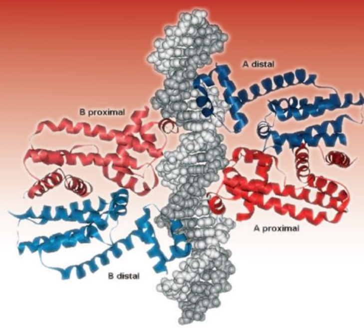
Multiple TFs



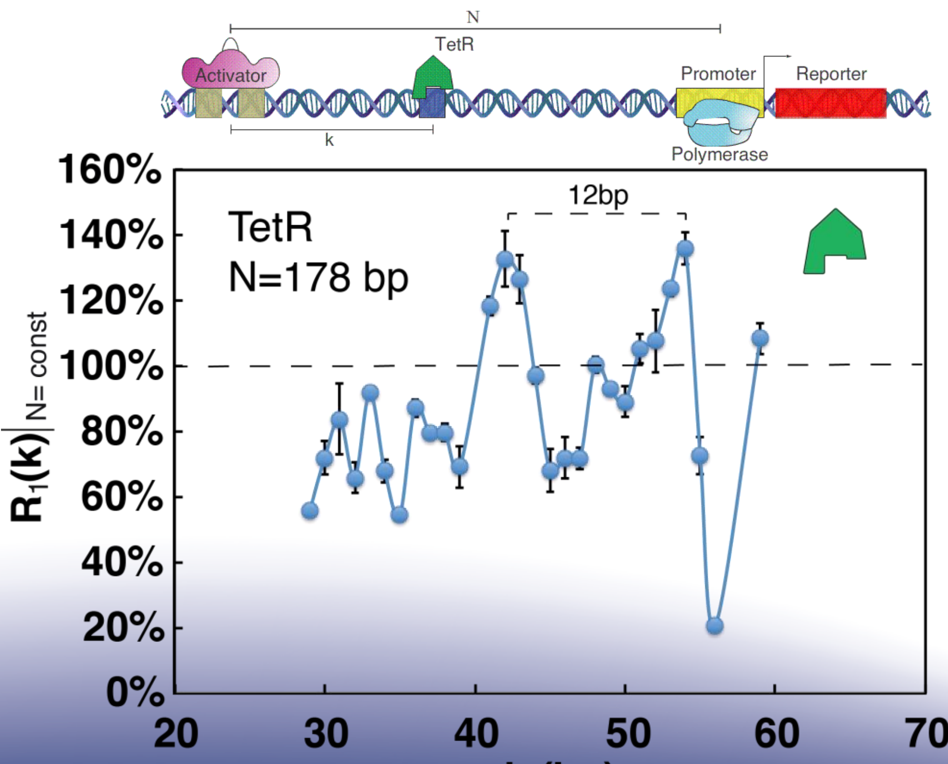


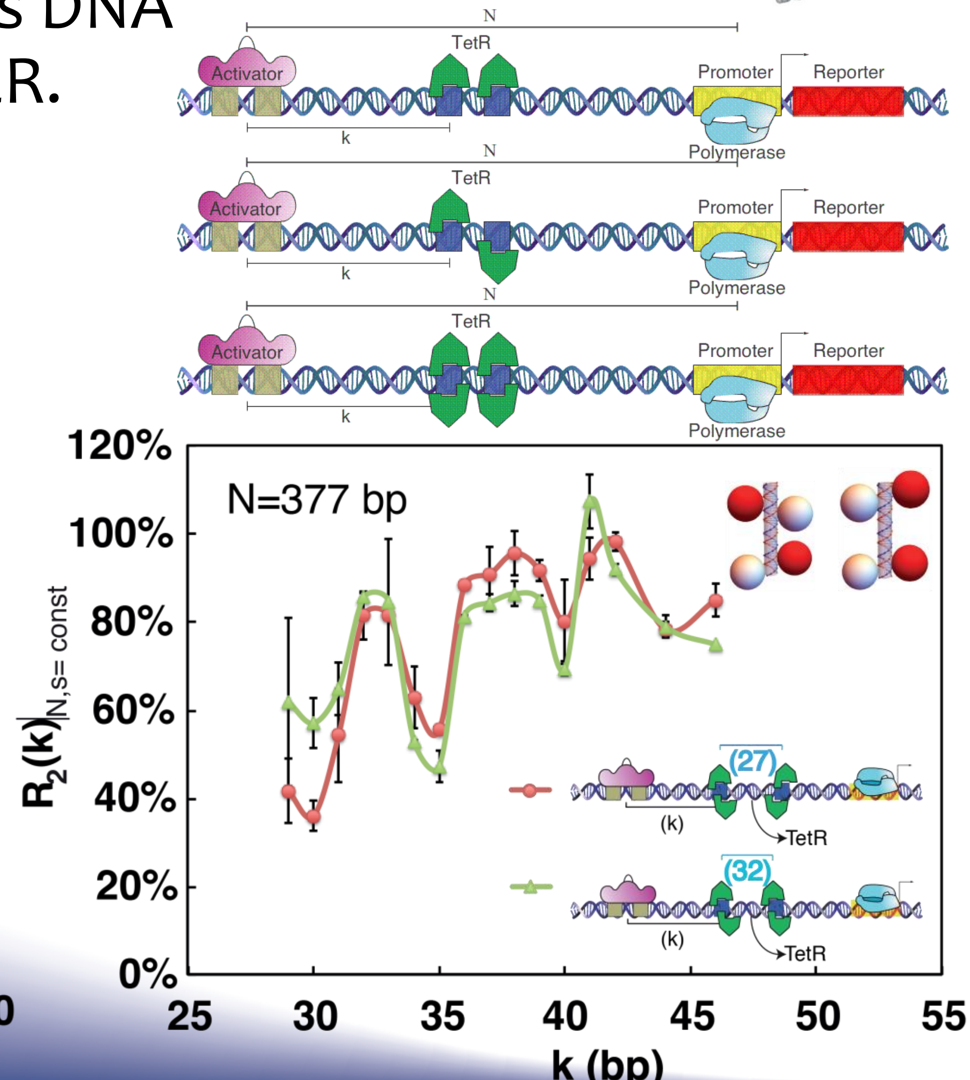
Brunwasser-Meirom, Pollak et al., Nat. Comm 2016

Structural Insights



- Results suggest TetR binds DNA similar to its homolog QacR.



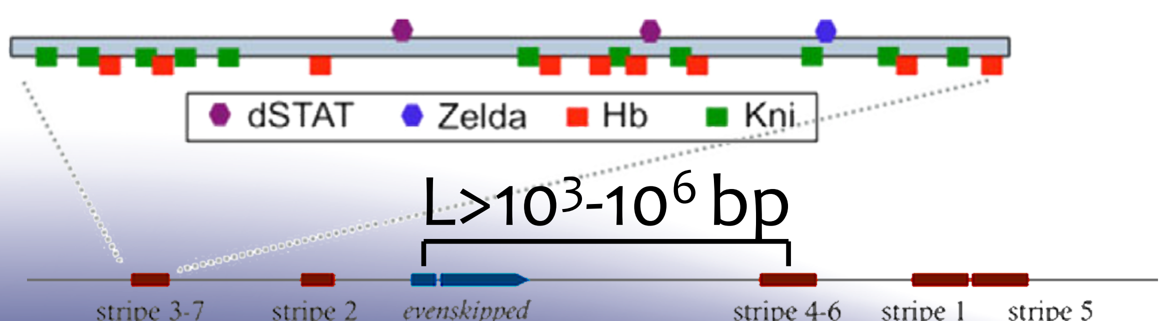


Brunwasser-Meirom, Pollak et al., Nat. Comm 2016

Long-Range
Looping Results

Short vs. Long range

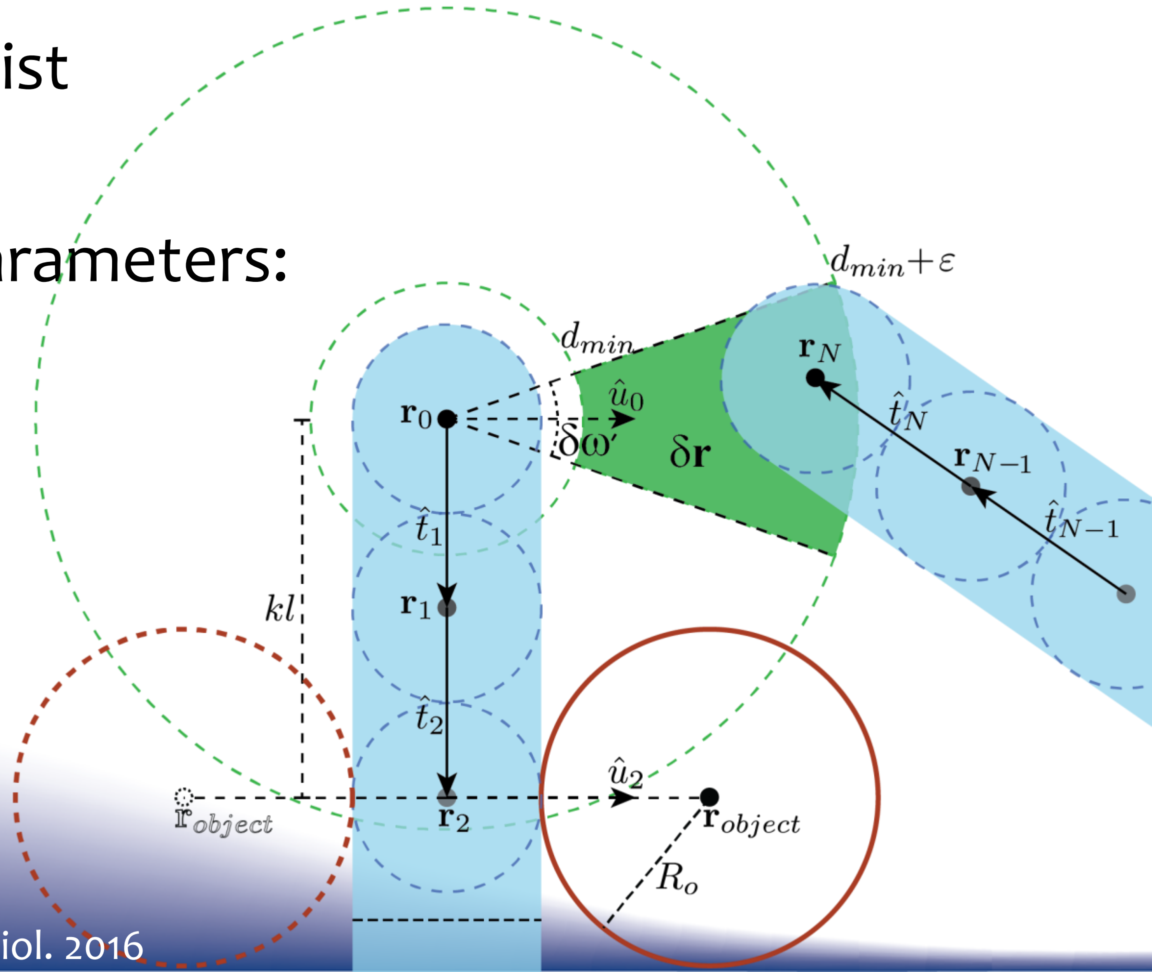
- Short range (elastic regime) (~10^2bp)
  - Bending
  - Stiffening
  - Excluded volume
- Long range (entropic regime) (>10^3bp)
  - ~~Bending~~
  - ~~Stiffening~~
  - Excluded volume



Simulation Looping Condition

- No activator & RNA polymerase
- Terminating end must reside in an off-axis cone
- Neglecting chain twist

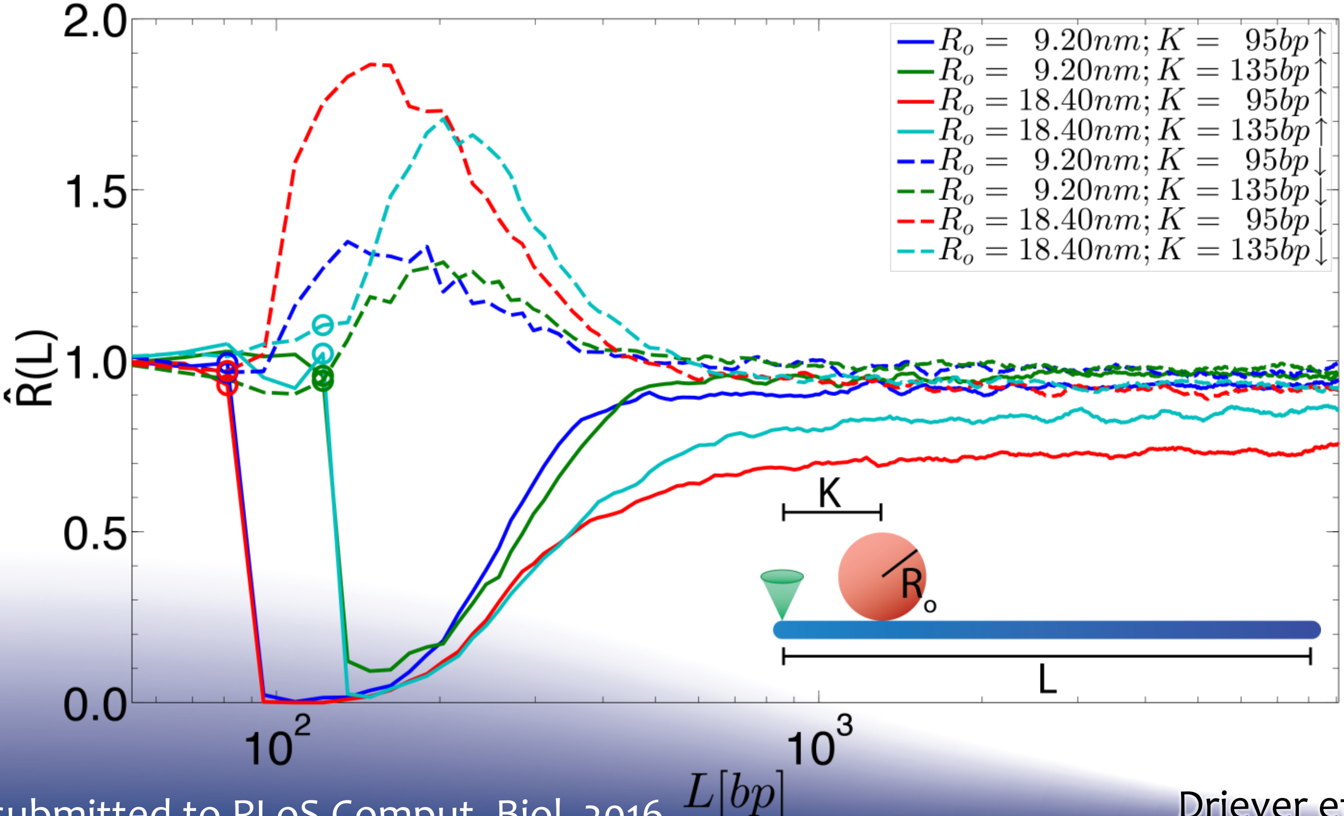
- Additional model parameters:
  - $\delta \mathbf{r}\,(\delta\omega, \varepsilon, d_{min})$
  - ~~$\delta\omega'$~~
  - ~~$\mathbf{r}_{pol}, \mathbf{r}_{actv}$~~



Pollak et al., submitted to PLoS Comput. Biol. 2016

Excluded Volume Effect

- The effect is relatively constant at long range
- Always down-regulation
- Only sizable for TFs in-phase with the looping volume

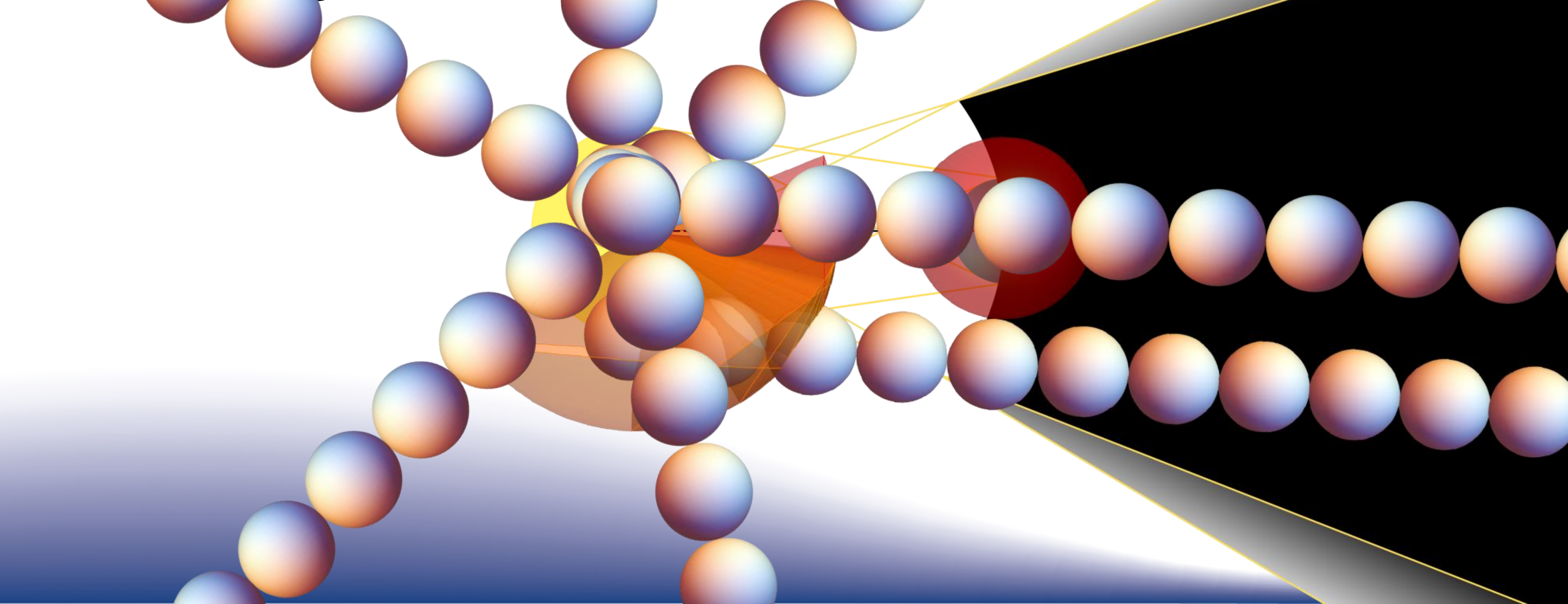


Pollak et al., submitted to PLoS Comput. Biol. 2016

Driever et al., Nature, 1989

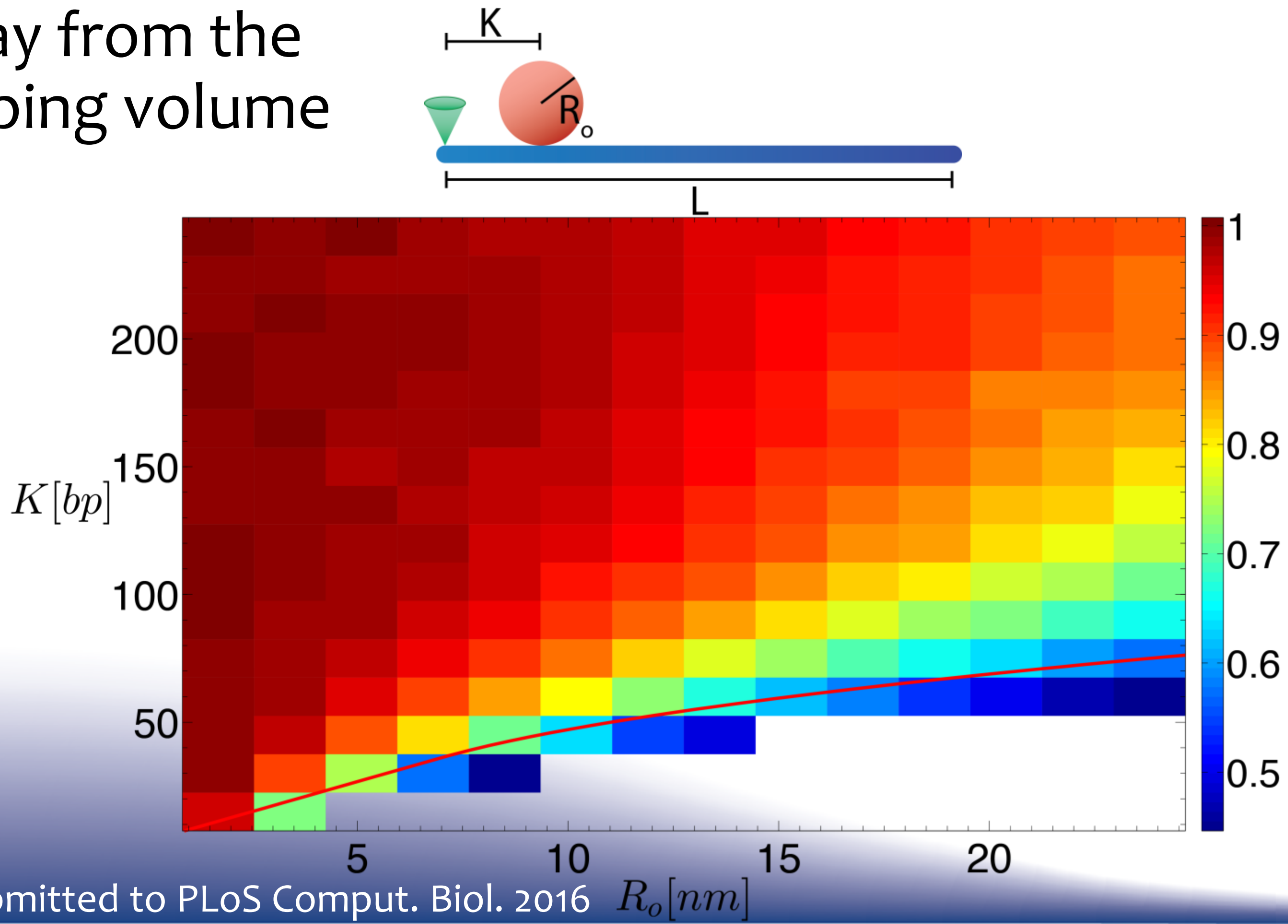
Mechanism - Eclipsing

- Chains emit light in all directions uniformly
- [illegible] depending on [illegible] angle at the [illegible]ping volume



Eclipsing Effect

- The effect is stronger with larger TF
- The effect dies out fast as the protrusion is moved away from the looping volume

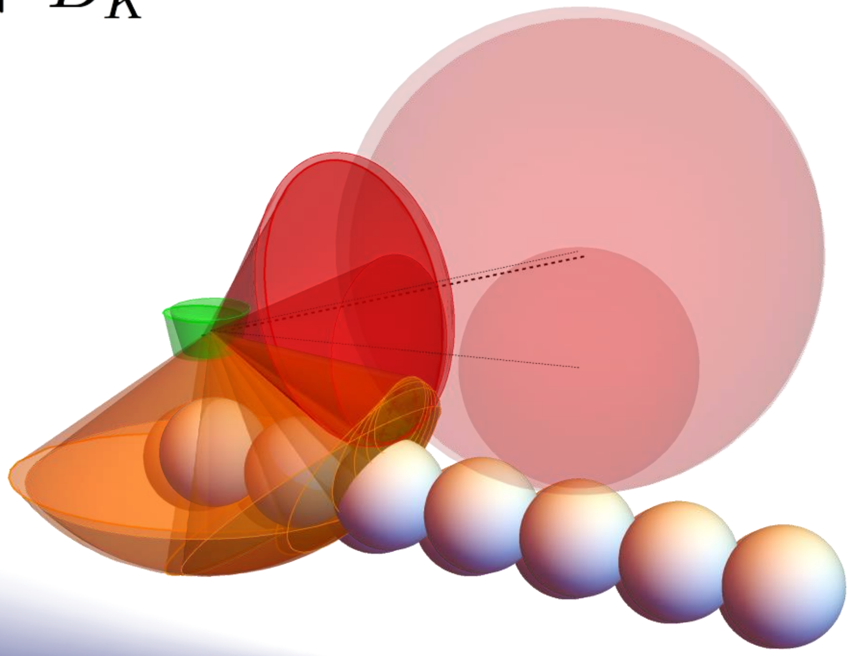


Pollak et al., submitted to PLoS Comput. Biol. 2016

Mechanism contd.

- Have to take the chain solid angle into account
- Distance is not straightforward
- $\hat{R}$ can be approximated for TF constant location

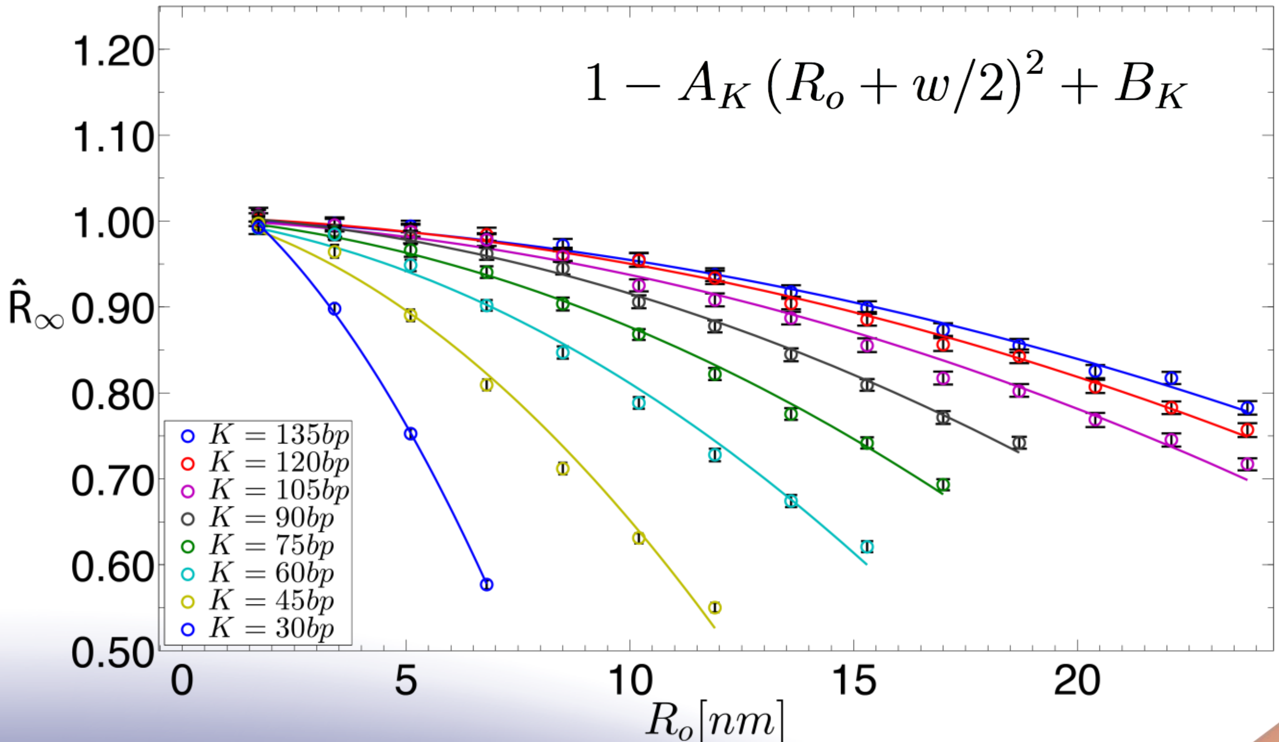
$$\hat{\mathsf{R}} \approx 1 - A_K \left(R_o + w/2\right)^2 + B_K$$

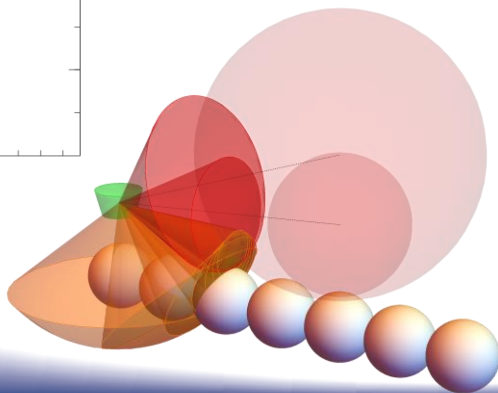


Pollak et al., submitted to PLoS Comput. Biol. 2016

Mechanism contd.

- $R^2 > 0.99$ for fit to quadratic functions in R_o.

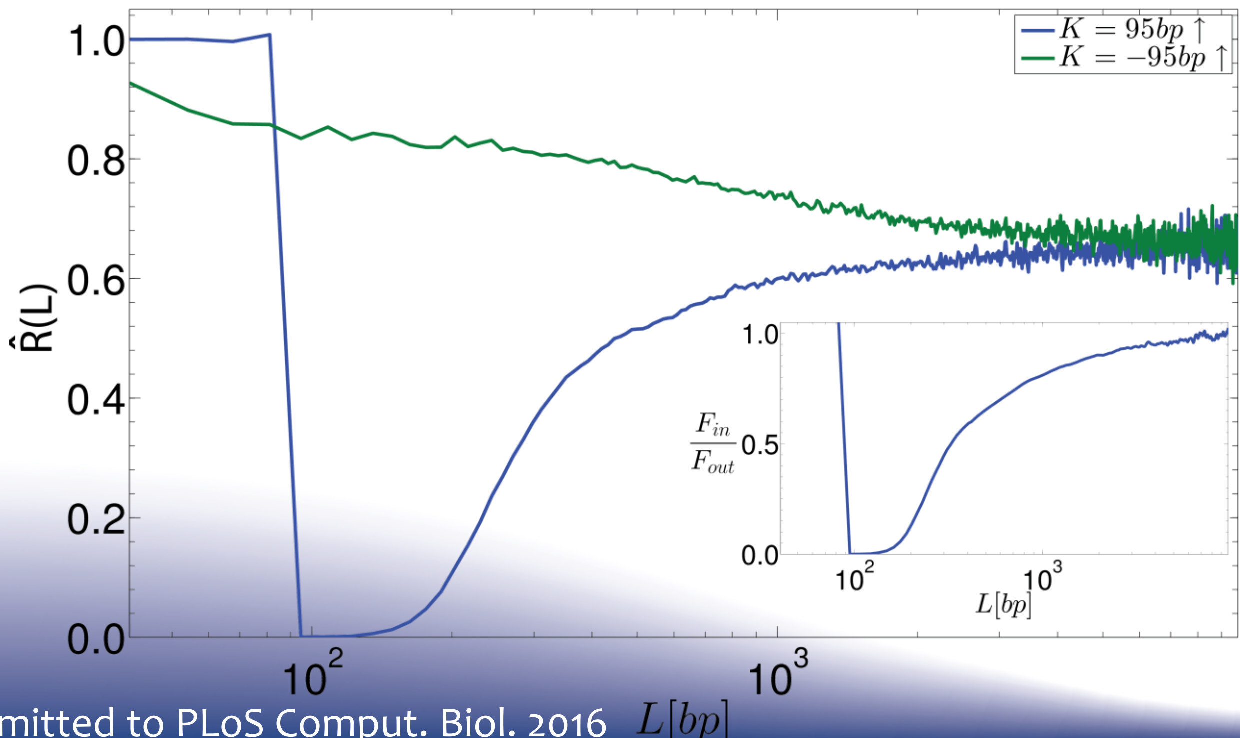




Pollak et al., submitted to PLoS Comput. Biol. 2016

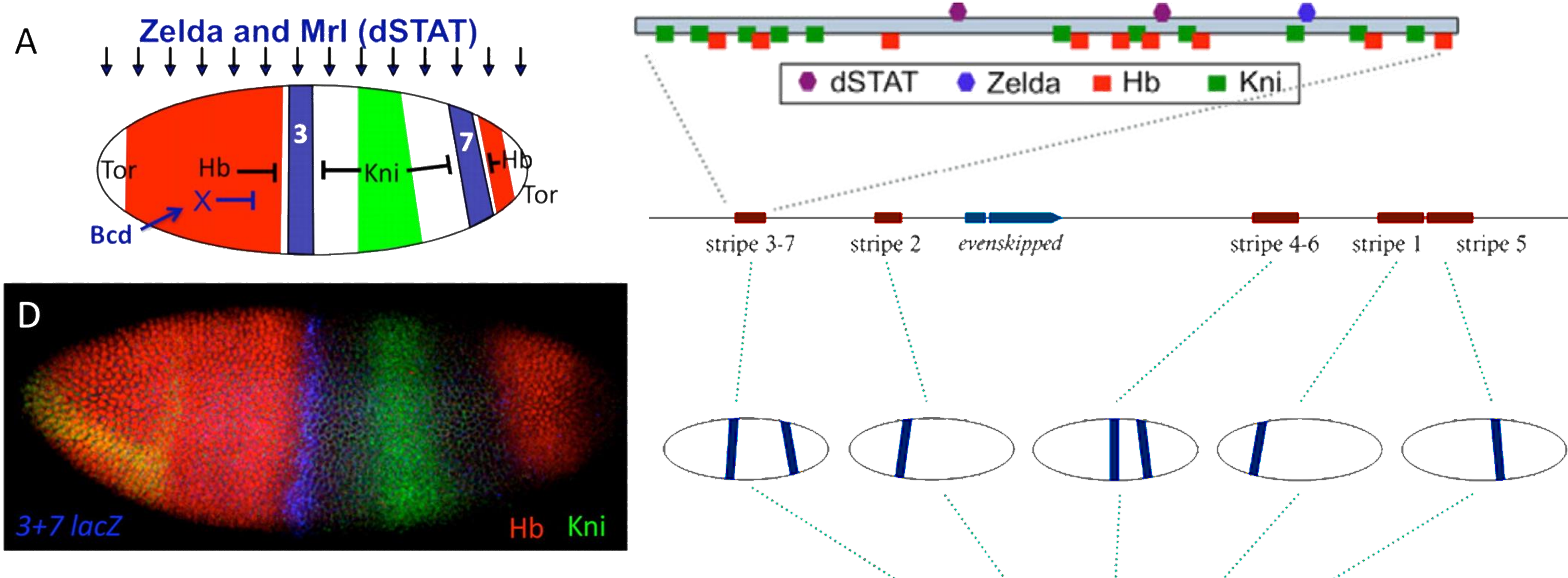
Upstream / Downstream

- The effect is independent of whether the TF is located upstream or downstream from the driver

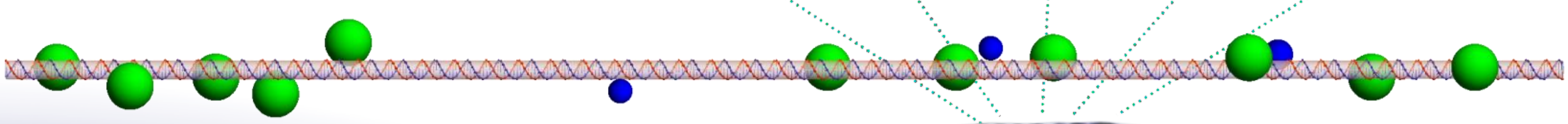


Pollak et al., submitted to PLoS Comput. Biol. 2016

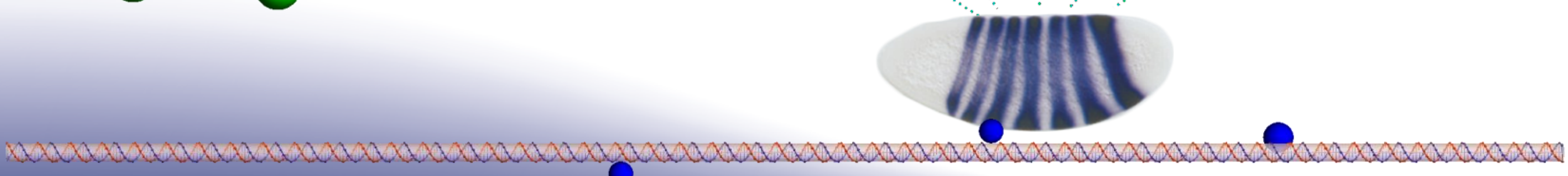
eve 3/7



- State 1:



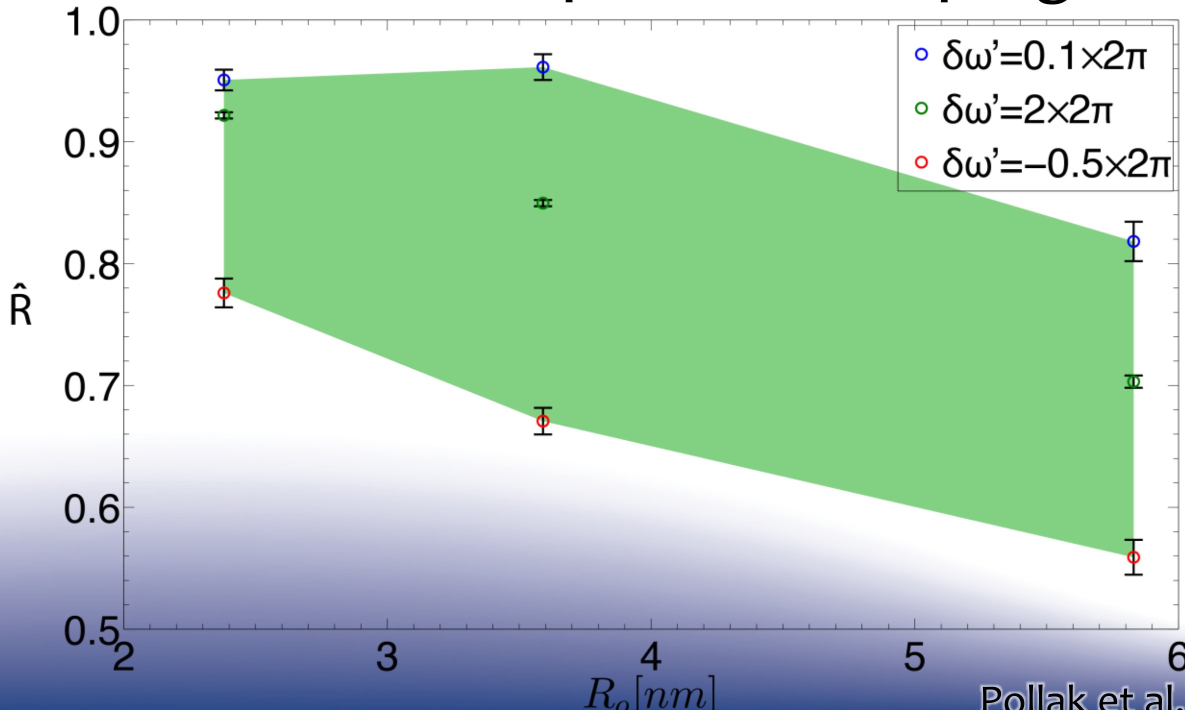
- State 2:

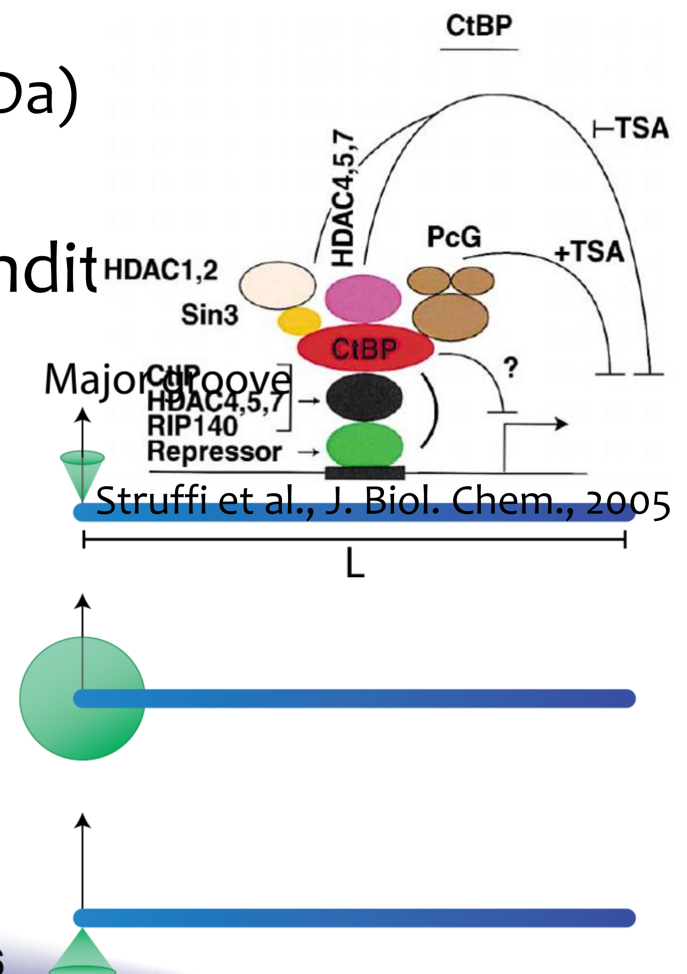


Struffi et al., Development, 2007.

eve 3/7 contd.

- Tested three possible repressor complex sizes:
  - Knirps alone (46 kDa)
  - Knirps bound to CtBP dimers (130 kDa)
  - Full putative 450 kDa complex
- Tested three possible looping condit





Struffi et al., J. Biol. Chem., 2005

Pollak et al., submitted to PLoS Comput. Biol. 2016

Conclusions

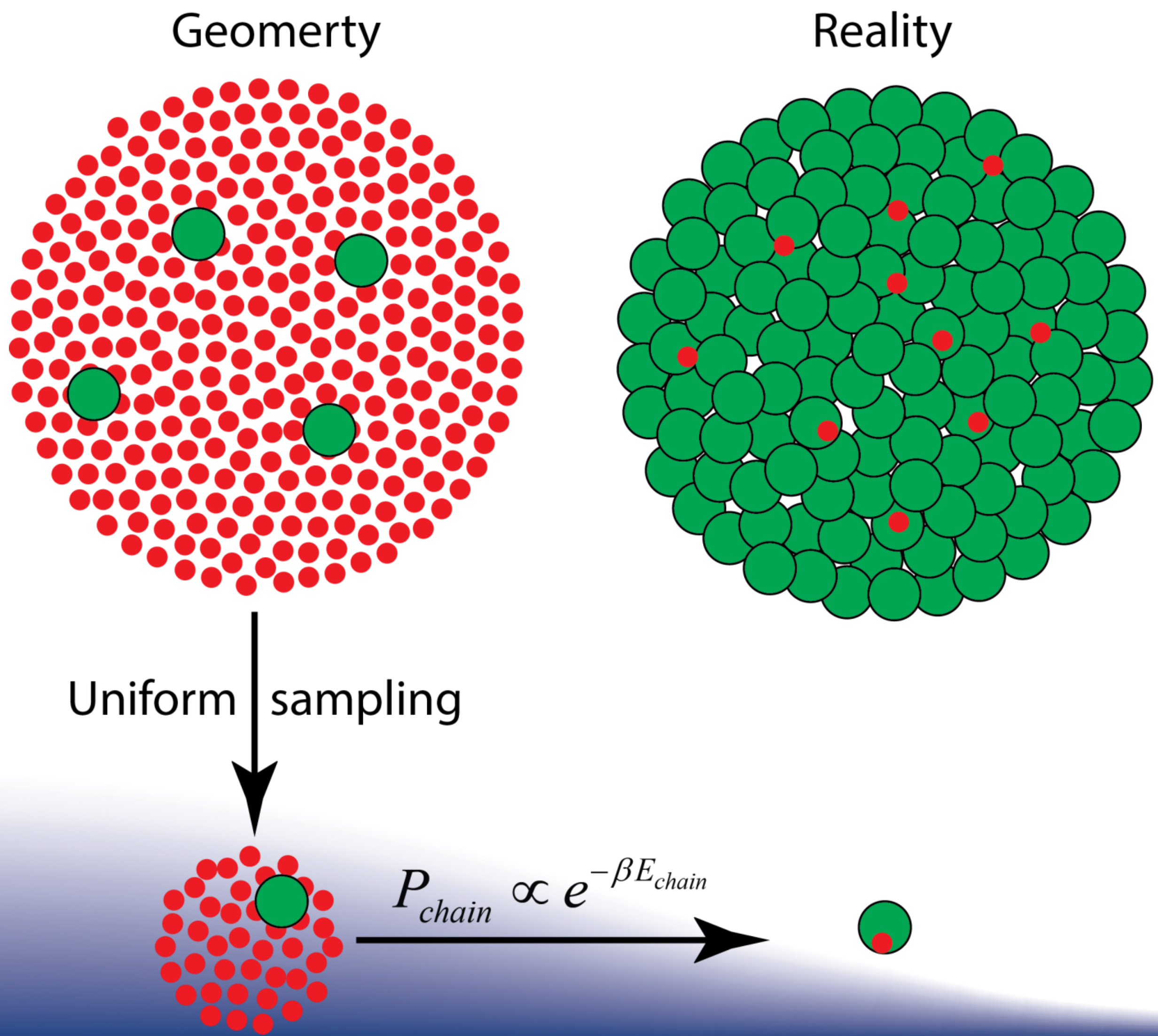
- Self-avoiding wormlike chain model can explain looping-based regulation
- Excluded volume can generate long-range repression
- Model predicts reduction in probability of looping for Knirps fully occupied eve 3/7 enhancer.

Thanks

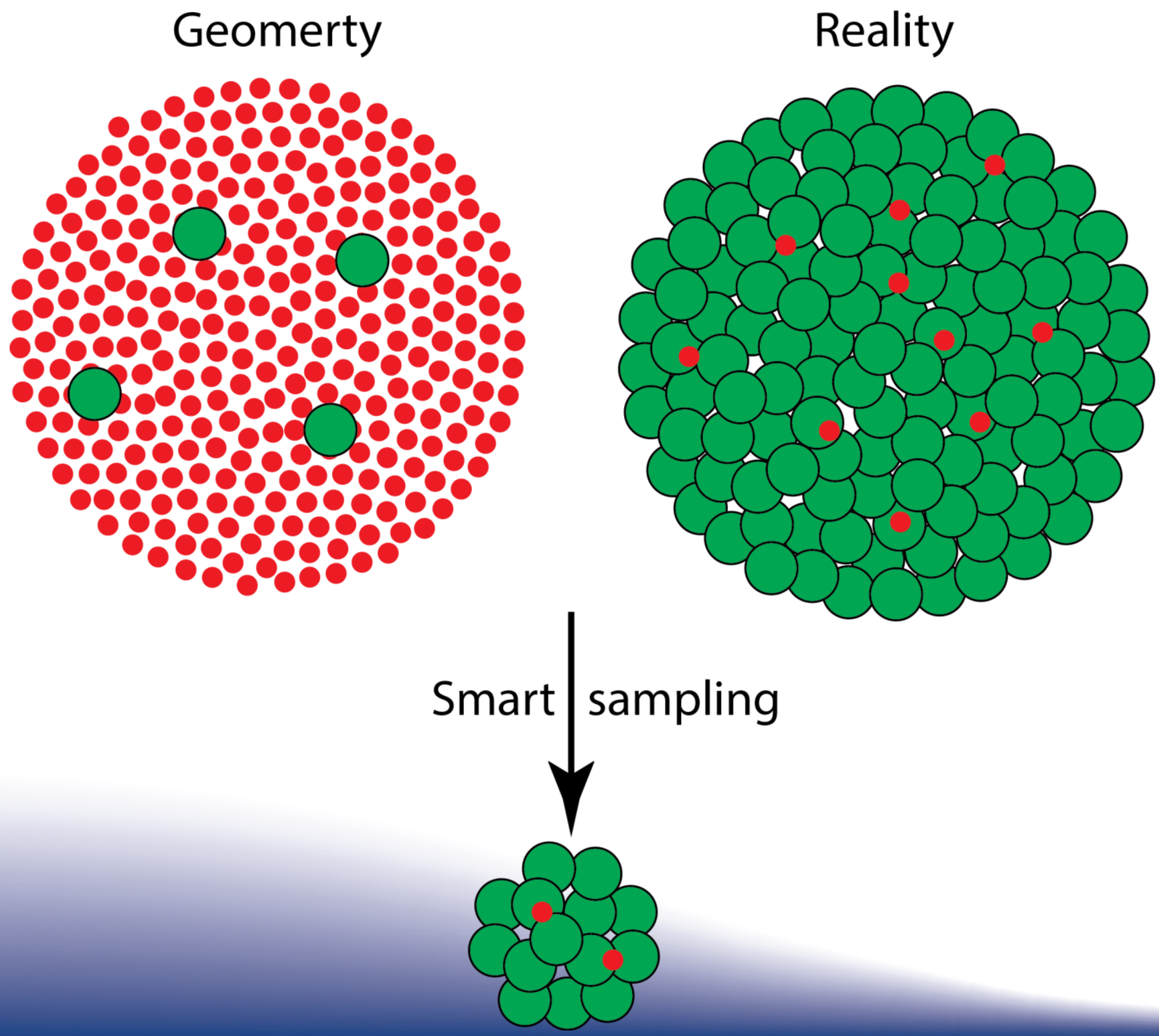
- Roee Amit
- Michal Meirom
- Sarah Goldberg
- Lior Levy
- Orna Atar
- Everybody else
- My Wife
- Our Parents

- RBNI – Daniel
- RBNI

Uniform Sampling



Importance Sampling



Sampling Bias

Would like to know: $$\langle v \rangle = \int v(\mathbf{x})\, \pi(\mathbf{x})\, \mathrm{d}\mathbf{x} = \frac{\int v(\mathbf{x}) \exp(-\beta E(\mathbf{x}))\, \mathrm{d}\mathbf{x}}{Z}$$

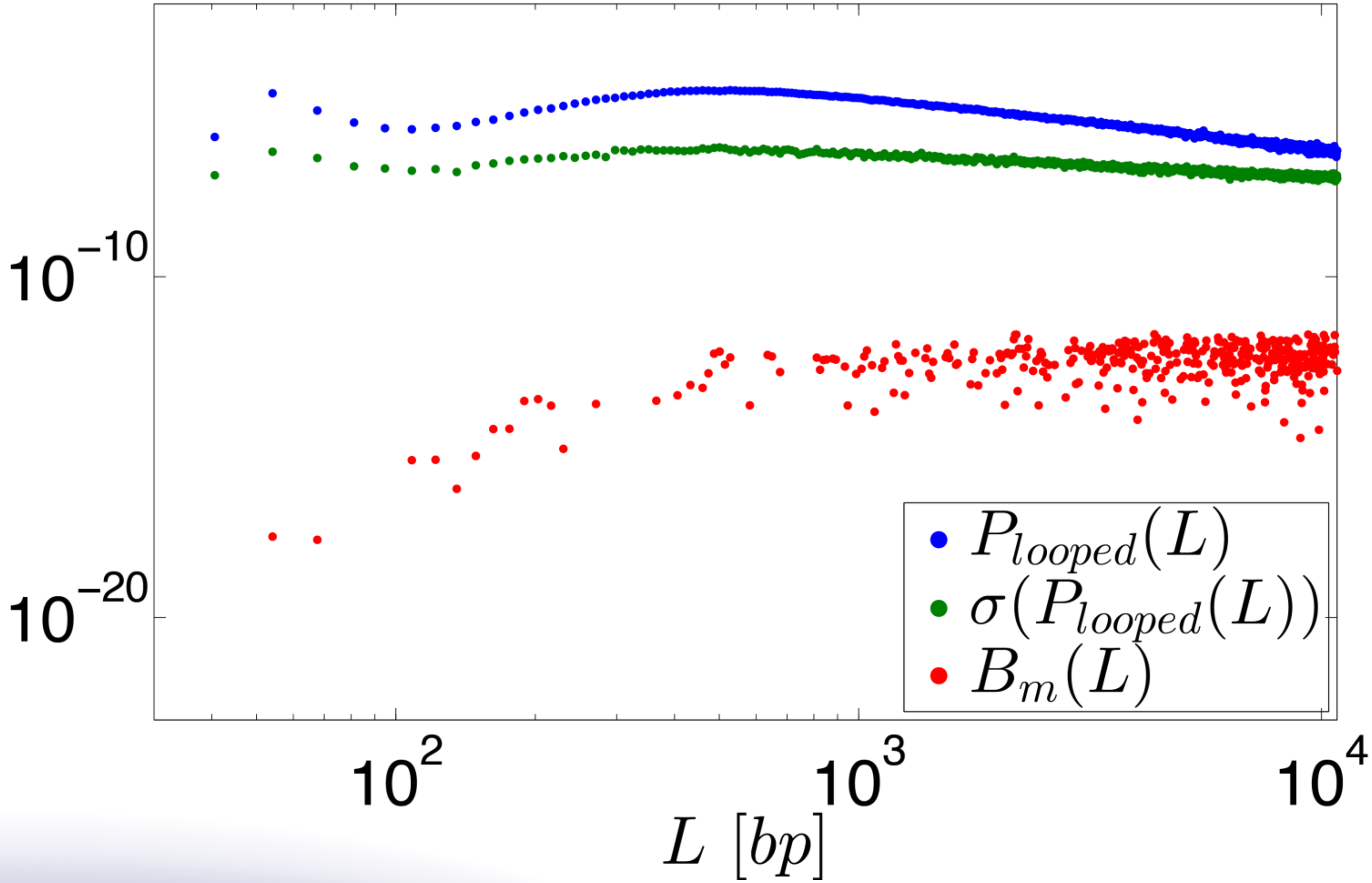
Estimate from ensemble: $$\langle \hat{v} \rangle = \frac{\sum_{i=1}^{m} w\left(\mathbf{x}^{(i)}\right) v\left(\mathbf{x}^{(i)}\right)}{\sum_{i=1}^{m} w\left(\mathbf{x}^{(i)}\right)}$$

Should have used: $$\langle \tilde{v} \rangle = \frac{1}{Z} \frac{1}{m} \sum_{i=1}^{m} w\left(\mathbf{x}^{(i)}\right) v\left(\mathbf{x}^{(i)}\right)$$

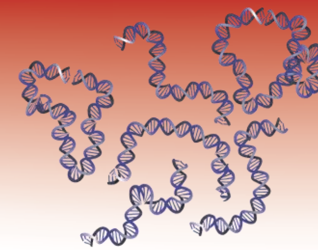
Sampling bias: $$B_m = E\left[\langle \hat{v} \rangle\right] - \langle v \rangle = -\frac{1}{mZ^2} \left(\int \left(v(\mathbf{x}) - \langle v \rangle\right) w(\mathbf{x}) \exp(-\beta E(\mathbf{x}))\, \mathrm{d}\mathbf{x} \right)$$

Practical bias estimator: $$B_m = E\left[\langle \hat{v} \rangle\right] - \langle v \rangle = -\frac{covar\left(\langle \hat{v} \rangle \hat{Z}\right)}{Z}$$

Sampling Bias



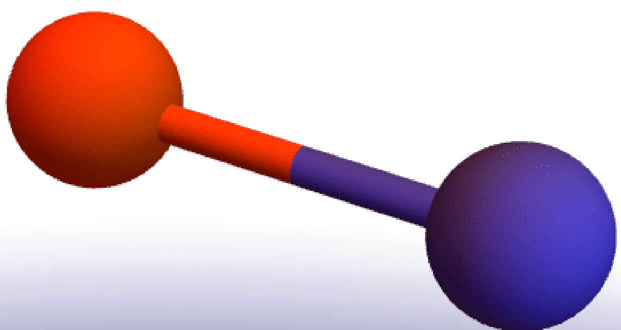
Monte-Carlo Simulations

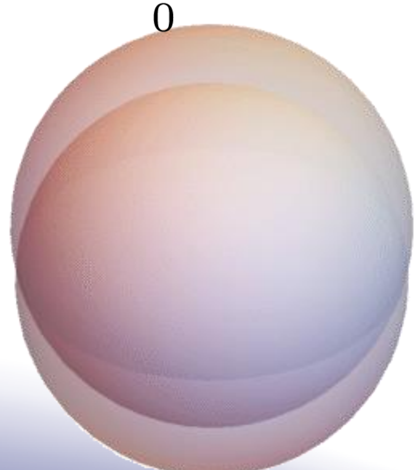


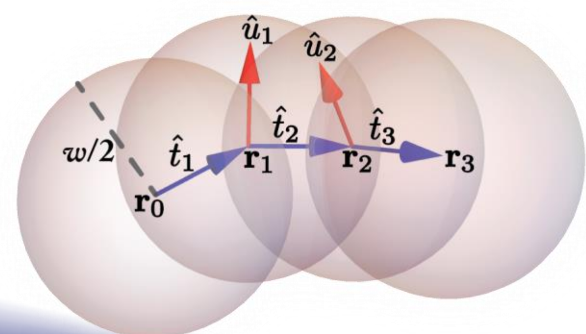
- Chains generated from scratch, link after link
- Links orientations mirror probabilities
- Samples of ~10^9 chains with & without TFs
- Comparing looping probability with & without TFs determines regulatory effect.

$$p_i\left(\{\theta_i,\phi_i\}\right) \propto \exp\left[-\beta E^{el+hw}\left(\{\theta_i,\phi_i\}\right)\right] \Longrightarrow P_{tot}\left(\{\theta_N,\phi_N\}\right) \not\propto \exp\left[-\beta E^{el+hw}_{chain}\right]$$

$$W\left(\{\theta_N,\phi_N\}\right) = \prod_{i=2}^{N}\left(\int_{-1}^{1} \mathrm{d}\cos\theta_i \int_{0}^{2\pi} \mathrm{d}\phi_i \exp\left[-\beta E^{el}\left(\theta_i,\phi_i\right)\right]\Theta_i^{hw}\left(\{\theta_i,\phi_i\}\right)\right)$$

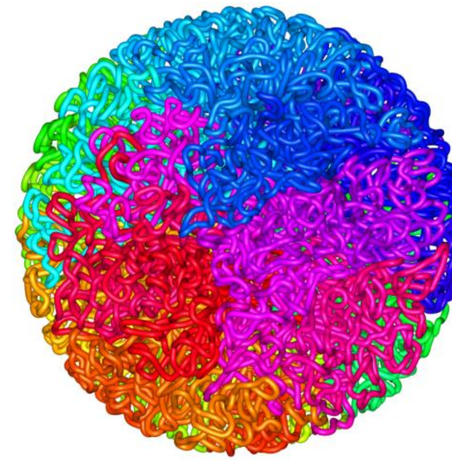




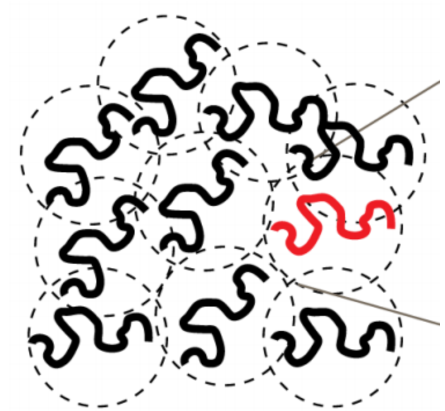


Model Applicability

- Cellular DNA:
  - In condensed globular state
  - Divided into autonomic domains – “blobs”
  - Blob size varies 300-10^3 nm



- Linearized active enhancer-promoter regions in organisms with low chromatin volume fractions (yeast, D. melanogaster) can explore the volume of the blob without inter-chromatin interactions



Confined DNA

- Confinement affects looping probability, not the ratio
- Enhancers/promoters are linearized
- Intermediate chain structure not important (ALA > *b*)
- Model applicable for small chromatin volume fractions