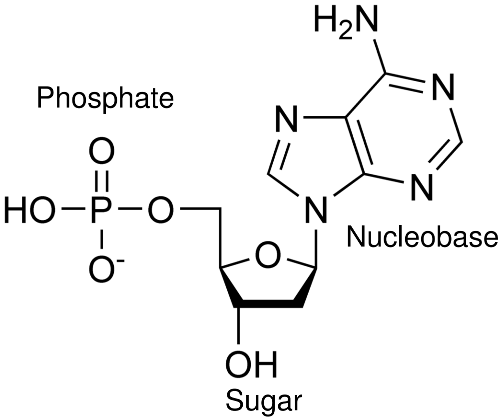
Hybrid particle-field model for DNA

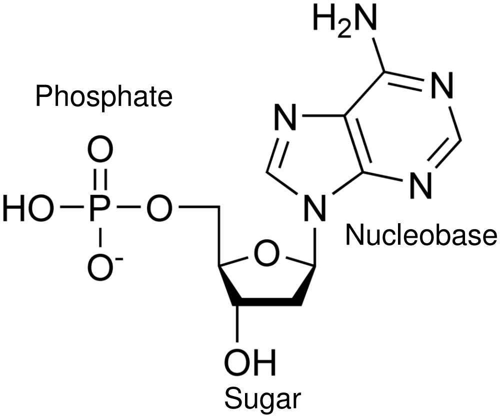
Sigbjørn Løland Bore
Weekley Hylleraas seminar

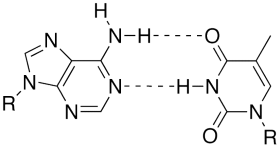
10.05.2019

Primer on DNA (1)



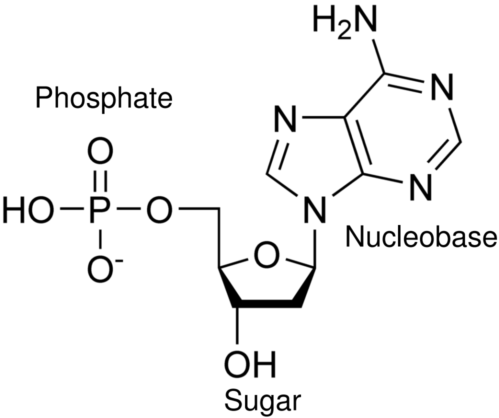
Primer on DNA (1)

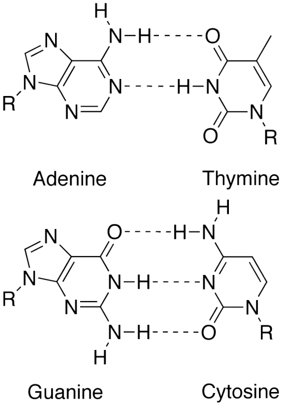




Adenine Thymine

Primer on DNA (1)



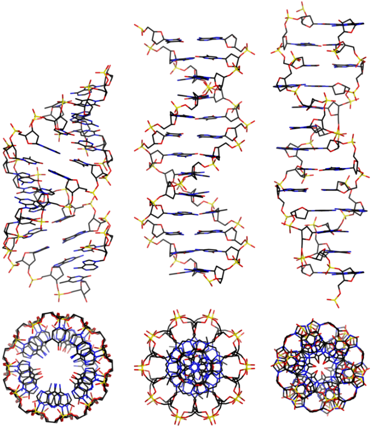


Primer on DNA (2)

- Watson and Crick pairing
- Double helix formation

Primer on DNA (2)

- Watson and Crick pairing
- Double helix formation
- Three types of helicies:
  - A-DNA
  - B-DNA
  - Z-DNA

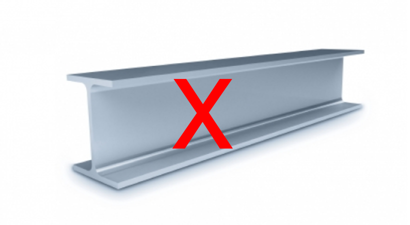


A-DNA B-DNA Z-DNA

Primer on DNA (3)

- Human DNA: 1-3 m
- Persistence length: 50 nm
- Very flexible

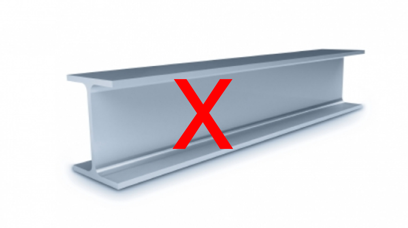
Primer on DNA (3)



- Human DNA: 1-3 m
- Persistence length: 50 nm
- Very flexible

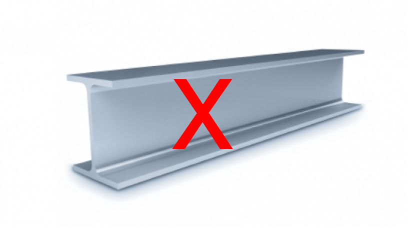
Primer on DNA (3)

- Human DNA: 1-3 m
- Persistence length: 50 nm
- Very flexible



Primer on DNA (3)

- Human DNA: 1-3 m
- Persistence length: 50 nm
- Very flexible
- Environmental effects
  - Temperature
  - Salt



The project:

- Reuse established CG-representations for DNA

The project:

- Reuse established CG-representations for DNA
- Model nonbonded interactions within hybrid particle-field framework

The project:

- Reuse established CG-representations for DNA
- Model nonbonded interactions within hybrid particle-field framework
- Parametrize the model

The project:

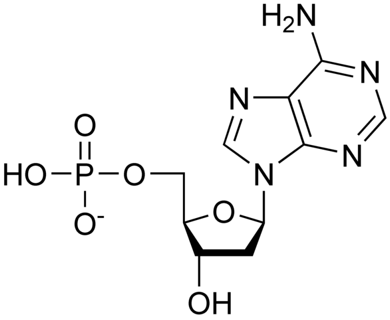
- Reuse established CG-representations for DNA
- Model nonbonded interactions within hybrid particle-field framework
- Parametrize the model
- Benchmark the model

The project:

- Reuse established CG-representations for DNA
- Model nonbonded interactions within hybrid particle-field framework
- Parametrize the model
- Benchmark the model
- No excuses, parallel implementation

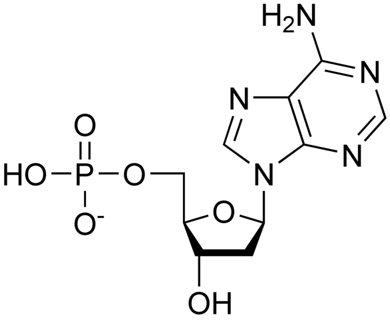
Coarse grain representation

- The coarse-grained representation should fulfill:



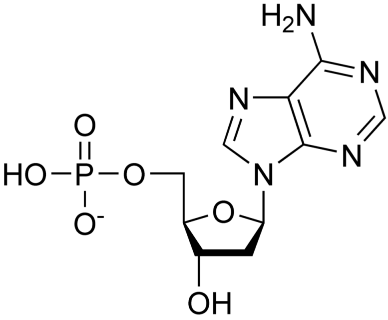
Coarse grain representation

- The coarse-grained representation should fulfill:
  - Represent the structural organization
  - 72 au per bead



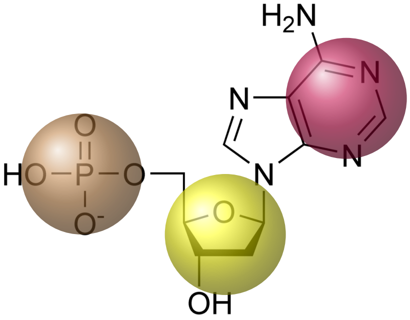
Coarse grain representation

- The coarse-grained representation should fulfill:
  - Represent the structural organization
  - 72 au per bead
- 3SPN-Model of Juan de Pablo
- Replace nonbonded interactions



Coarse grain representation

- The coarse-grained representation should fulfill:
  - Represent the structural organization
  - 72 au per bead
- 3SPN-Model of Juan de Pablo
- Replace nonbonded interactions

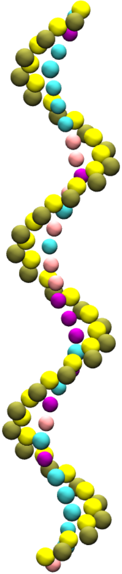


Bonded interactions

$$H_0(\{\mathbf{r}\}) = \sum_i^{N_\text{atom}} \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 + \sum_i^{N_\text{bond}} \frac{1}{2} k_r (r_i - r_{i0})^2$$
$$+ \sum_i^{N_\text{bend}} \frac{1}{2} k_\theta (\theta_i - \theta_{i0})^2 - \sum_i^{N_\text{tor}} k_\phi \exp\left[-\frac{(\phi_i - \phi_{0i})^2}{2\sigma_\phi^2} \right],$$

Bonded interactions

Equilibrium structure



$$H_0(\{\mathbf{r}\}) = \sum_i^{N_{\text{atom}}} \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 + \sum_i^{N_{\text{bond}}} \frac{1}{2} k_r (r_i - r_{i0})^2$$
$$+ \sum_i^{N_{\text{bend}}} \frac{1}{2} k_\theta (\theta_i - \theta_{i0})^2 - \sum_i^{N_{\text{tor}}} k_\phi \exp\left[-\frac{(\phi_i - \phi_{0i})^2}{2\sigma_\phi^2}\right],$$

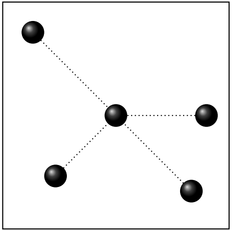
Bond	$r_{i0}\,/\,\mathrm{nm}$	Bend	$\theta_{i0}\,/\,\mathrm{deg}$	Torsional	$\phi_{i0}\,/\,\mathrm{deg}$
S-P	0.3899	S-P-S	94.49	P-S-P-S	-154.8
P-S	0.3559	P-S-P	120.15	S-P-S-P	-179.2
S-A	0.4670	A-S-P	112.07	A-S-P-S	-32.8
S-T	0.4189	P-S-A	103.53	S-P-S-A	54.8
S-G	0.4829	T-S-P	116.68	T-S-P-S	-44.8
S-C	0.3844	P-S-T	92.06	S-P-S-T	58.0
		G-S-P	110.12	G-S-P-S	-29.1
		P-S-G	107.40	S-P-S-G	53.9
		C-S-P	110.33	C-S-P-S	-34.1
		P-S-C	103.79	S-P-S-C	57.0

Hybrid particle field method

Mesoscale potentials in molecular dynamics:

$$V_{\text{ext},i} = \frac{1}{\tilde{\phi}_0}\left(k_{\text{b}} T \sum_j \chi_{ij}\phi_j(\mathbf{r}) + \frac{1}{\kappa}\left(\sum_j \phi_j(\mathbf{r}) - \tilde{\phi}_0\right)\right)$$

χ_{ij}: Flory-Huggins parameter. κ: compressibility. $\tilde{\phi}_0$: system density.



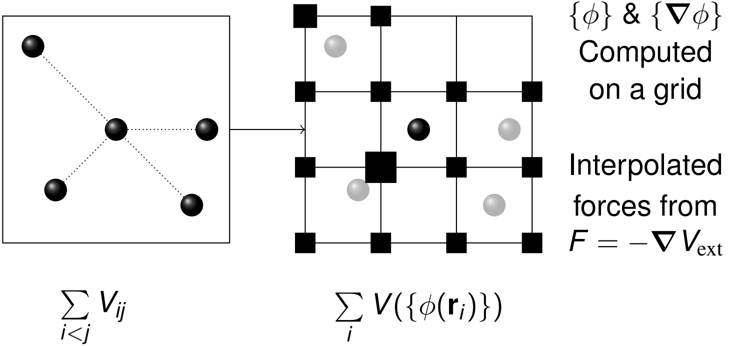
$$\sum_{i<j} V_{ij}$$

Hybrid particle field method

Mesoscale potentials in molecular dynamics:

$$V_{\text{ext},i} = \frac{1}{\tilde{\phi}_0}\left(k_b T \sum_j \chi_{ij}\phi_j(\mathbf{r}) + \frac{1}{\kappa}\left(\sum_j \phi_j(\mathbf{r}) - \tilde{\phi}_0\right)\right)$$

χ_{ij}: *Flory-Huggins parameter.* κ: *compressibility.* $\tilde{\phi}_0$: *system density.*



Nonbonded interactions

$$W_{\text{elec}}[\rho] = \int d\mathbf{r}\ V_{\text{Coul}}(\mathbf{r})\rho(\mathbf{r})$$

$$W_{\text{non-elec}}[\{\phi\}] = \frac{1}{\tilde{\phi}_0}\int d\mathbf{r}\left[\frac{k_b T}{2}\sum_{k,\ell}\chi_{k\ell}\phi_k(\mathbf{r})\phi_\ell(\mathbf{r}) + \frac{1}{2\kappa}\left(\sum_k \phi_k(\mathbf{r}) - \tilde{\phi}_0\right)^2\right]$$

	P	S	A	T	C	G	W
P	χ_{PP}	0	0	0	0	0	χ_{PW}
S	0	0	0	0	0	0	0
A	0	0	0	χ_{NN}	0	0	χ_{NW}
T	0	0	χ_{NN}	0	0	0	χ_{NW}
C	0	0	0	0	0	χ_{NN}	χ_{NW}
G	0	0	0	0	χ_{NN}	0	χ_{NW}
W	χ_{PW}	0	χ_{NW}	χ_{NW}	χ_{NW}	χ_{NW}	0

Parametrization

- Parameters of the model:
  - k_r, k_θ, k_ϕ, χ_{NW}, χ_{NN}, χ_{PP}, χ_{PW}

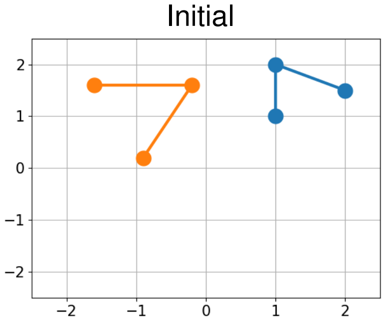
Parametrization

- Parameters of the model:
  - k_r, k_θ, k_ϕ, χ_{NW}, χ_{NN}, χ_{PP}, χ_{PW}
- Goals:
  - Reproduces well the strcuture of B-DNA
  - Reproduce the persistence length of SS- and DS-DNA

Optimization procedure(1): Fitness parameter

$$\eta = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(\mathbf{r}_{i,1} - \mathbf{r}_{i,2})^2}$$

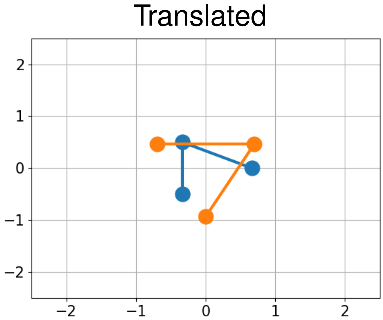
- Kabsch algorithm



Optimization procedure(1): Fitness parameter

$$\eta = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(\mathbf{r}_{i,1} - \mathbf{r}_{i,2})^2}$$

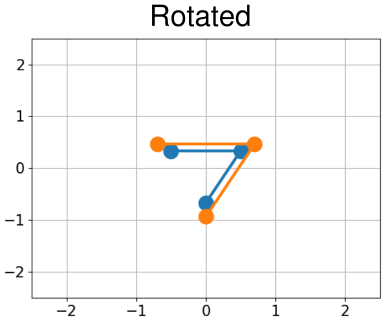
- Kabsch algorithm



Optimization procedure(1): Fitness parameter

$$\eta = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(\mathbf{r}_{i,1} - \mathbf{r}_{i,2})^2}$$

- Kabsch algorithm



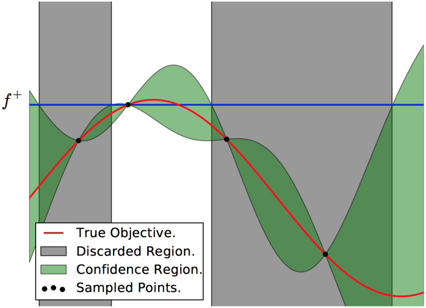
Optimization procedure(2): Optimization method

- Requirements:
  - No gradients
  - Handle noisy fitness
  - Few function calls

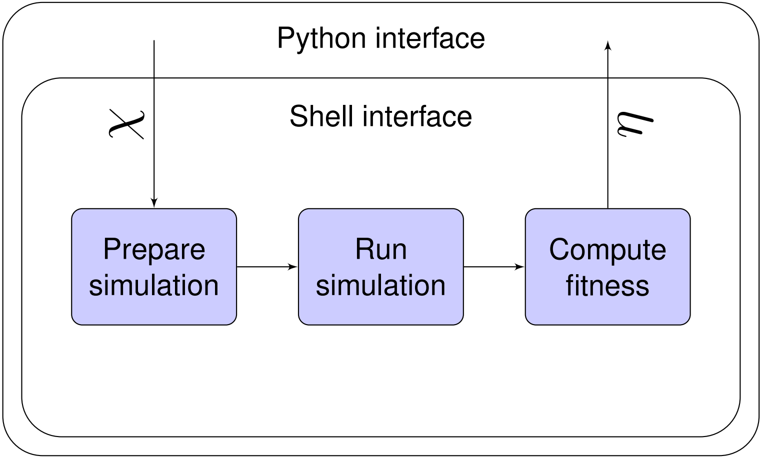
Optimization procedure(2): Optimization method

- Requirements:
  - No gradients
  - Handle noisy fitness
  - Few function calls

$\Longrightarrow$ Bayesian Optimization

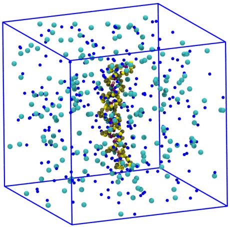


Optimization procedure(3): Implementation



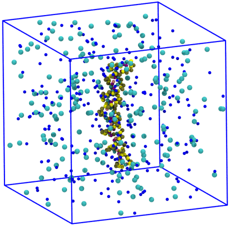
Application of optimization

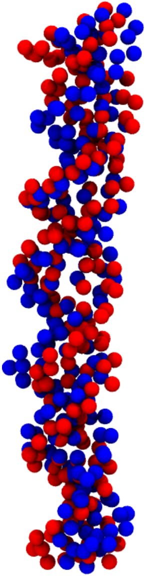
- 32 bp DNA, 100 mM salt
- 120ns



Application of optimization

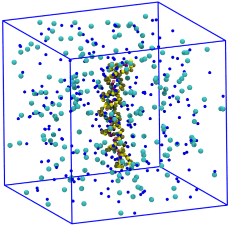
- 32 bp DNA, 100 mM salt
- 120ns

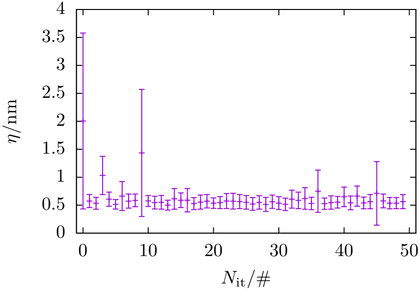


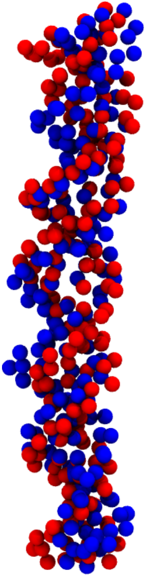


Application of optimization

- 32 bp DNA, 100 mM salt
- 120ns





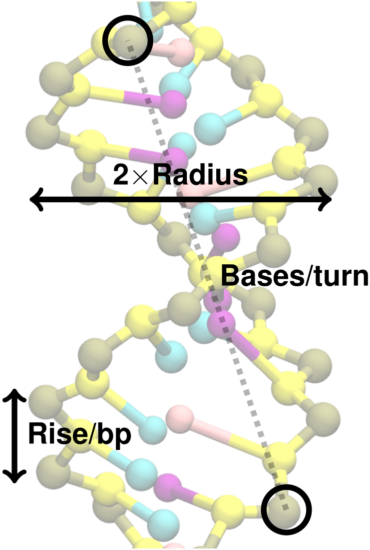


Applications: Structural properties

- Best set:
  - $\chi_{NW} = 19.0\,\mathrm{kJ\,mol^{-1}}$
  - $\chi_{NN} = -12.7\,\mathrm{kJ\,mol^{-1}}$
  - $\chi_{PW} = -7.2\,\mathrm{kJ\,mol^{-1}}$
  - $\chi_{PP} = -4.2\,\mathrm{kJ\,mol^{-1}}$

Applications: Structural properties

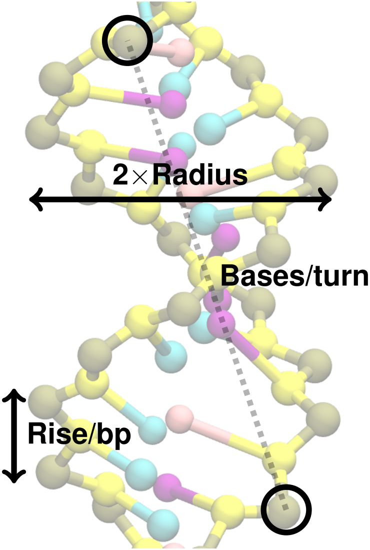
- Best set:
  - $\chi_{NW} = 19.0\,\mathrm{kJ\,mol^{-1}}$
  - $\chi_{NN} = -12.7\,\mathrm{kJ\,mol^{-1}}$
  - $\chi_{PW} = -7.2\,\mathrm{kJ\,mol^{-1}}$
  - $\chi_{PP} = -4.2\,\mathrm{kJ\,mol^{-1}}$



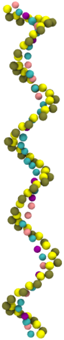
Applications: Structural properties

- Best set:
  - $\chi_{NW} = 19.0\,\mathrm{kJ\,mol^{-1}}$
  - $\chi_{NN} = -12.7\,\mathrm{kJ\,mol^{-1}}$
  - $\chi_{PW} = -7.2\,\mathrm{kJ\,mol^{-1}}$
  - $\chi_{PP} = -4.2\,\mathrm{kJ\,mol^{-1}}$

Property	simulation	expt.
Bases per turn	9.6 ± 0.3	10
Rise pr bp/nm	$0.34(5) \pm 0.01$	0.34
Radius/nm	0.88 ± 0.04	0.94

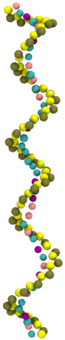


Applications: Hairpin-formation

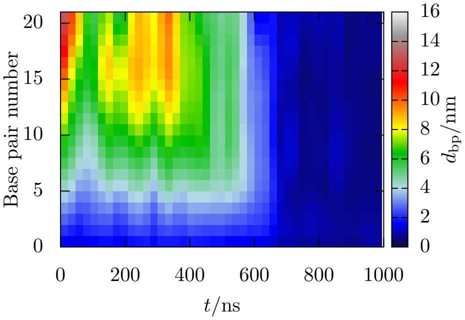


Initial

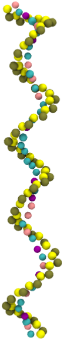
Applications: Hairpin-formation



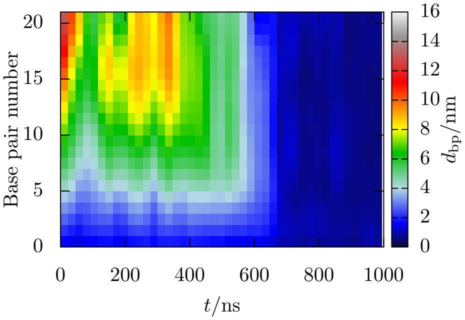
Initial

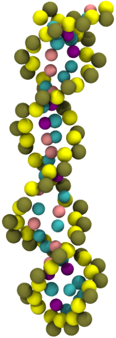


Applications: Hairpin-formation



Initial





Final

Persistence length DS-DNA

$$\left\langle \hat{\mathbf{t}} \cdot \hat{\mathbf{t}}_l \right\rangle = e^{-l/l_\mathrm{p}}, \quad \mathbf{t} \equiv \mathbf{r}_{\mathrm{P},i+10} - \mathbf{r}_{\mathrm{P},i}$$

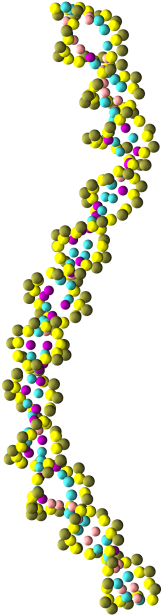
Persistence length DS-DNA

$$\left\langle \hat{\mathbf{t}} \cdot \hat{\mathbf{t}}_l \right\rangle = e^{-l/l_\mathrm{p}}, \quad \mathbf{t} \equiv \mathbf{r}_{\mathrm{P},i+10} - \mathbf{r}_{\mathrm{P},i}$$

- Experimental: l_P =40-60 nm

Persistence length DS-DNA

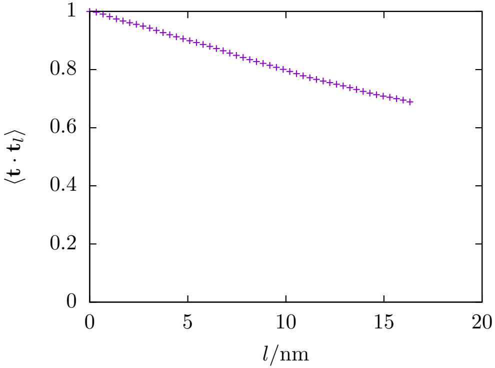
$$\left\langle \hat{\mathbf{t}} \cdot \hat{\mathbf{t}}_l \right\rangle = e^{-l/l_\mathrm{p}}, \quad \mathbf{t} \equiv \mathbf{r}_{\mathrm{P},i+10} - \mathbf{r}_{\mathrm{P},i}$$

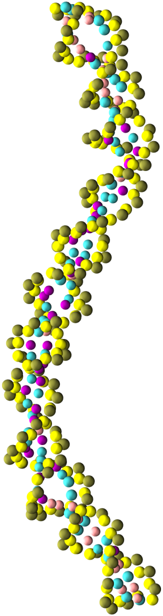


- Experimental: l_P =40-60 nm

Persistence length DS-DNA

$$\left\langle \hat{\mathbf{t}} \cdot \hat{\mathbf{t}}_l \right\rangle = e^{-l/l_\mathrm{P}}, \quad \mathbf{t} \equiv \mathbf{r}_{\mathrm{P},i+10} - \mathbf{r}_{\mathrm{P},i}$$

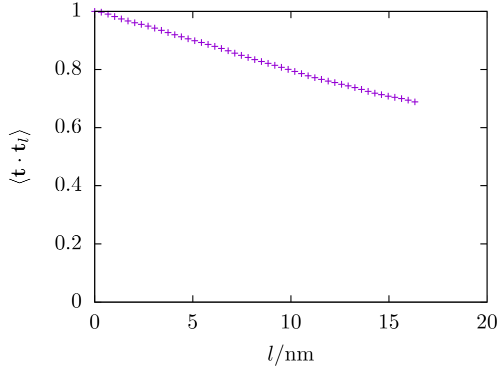


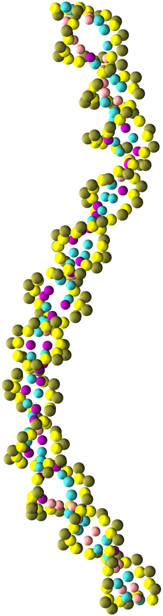


- Experimental: l_P =40-60 nm

Persistence length DS-DNA

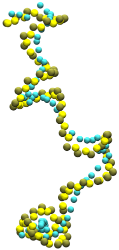
$$\left\langle \hat{\mathbf{t}} \cdot \hat{\mathbf{t}}_l \right\rangle = e^{-l/l_\mathrm{p}}, \quad \mathbf{t} \equiv \mathbf{r}_{\mathrm{P},i+10} - \mathbf{r}_{\mathrm{P},i}$$



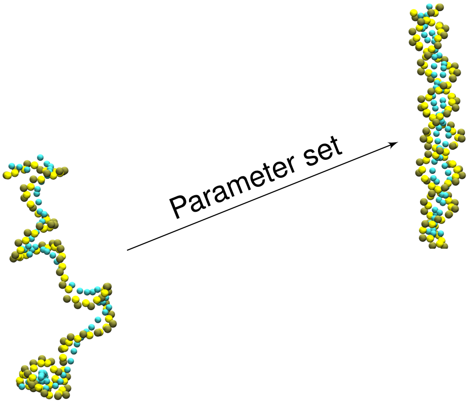


- Experimental: l_P =40-60 nm
- Simulation: l_P =43 nm

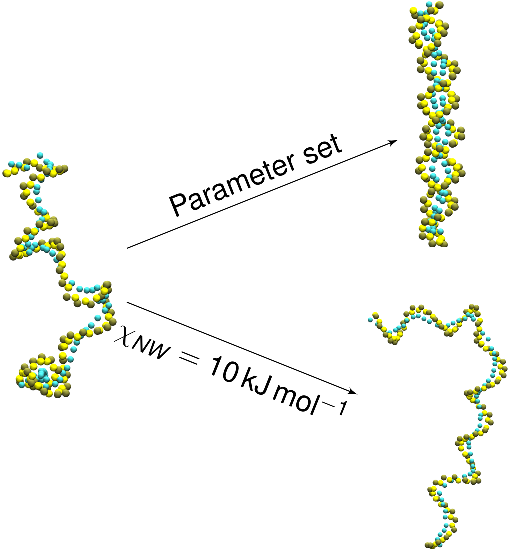
Problem with SS-DNA



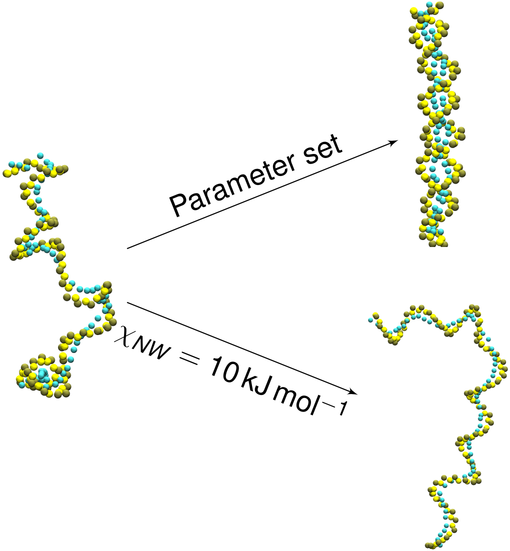
Problem with SS-DNA



Problem with SS-DNA



Problem with SS-DNA



Both are too stiff!

Outlook

- Redo optimization to get better SS-strand behaviour
  - Less stiff k_ϕ
  - Limit $\chi_{NW} \leq 10\,\mathrm{kJ\,mol^{-1}}$

Outlook

- Redo optimization to get better SS-strand behaviour
  - Less stiff k_ϕ
  - Limit $\chi_{NW} \leq 10\,\mathrm{kJ\,mol^{-1}}$
- Applications on longer doublestranded DNA

Outlook

- Redo optimization to get better SS-strand behaviour
  - Less stiff k_ϕ
  - Limit $\chi_{NW} \leq 10\,\mathrm{kJ\,mol^{-1}}$
- Applications on longer doublestranded DNA
- Investigate the effect of salt on persistence length

Outlook

- Redo optimization to get better SS-strand behaviour
  - Less stiff k_ϕ
  - Limit $\chi_{NW} \leq 10\,\mathrm{kJ\,mol^{-1}}$
- Applications on longer doublestranded DNA
- Investigate the effect of salt on persistence length
- Plans for applying optimization on other systems

Acknowledgements

Morten Ledum
Michele Cascella