

# A pinch of salt on your DNA ?

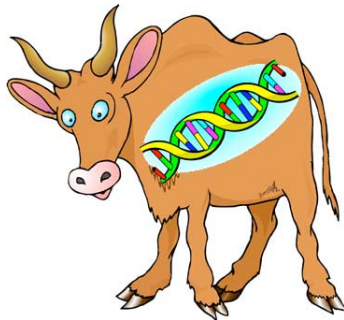
Sam Meyer

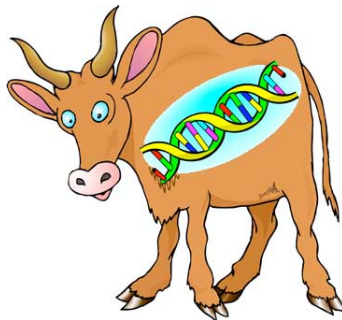
LIRIS, INSA Lyon

Journées du Centre Blaise Pascal  
28.11.2013









# Outline

## DNA transactions: a matter of electrostatics

- Electrostatics in DNA-protein interactions

- The problem of sequence recognition

## Analysis of MD simulation data

- Ions around a DNA oligomer

- Data analysis: convergence

- Data analysis: effect of sequence

## From simulation data to implicit solvent models

- General problem

- Benchmark

- Conclusion

# Outline

## DNA transactions: a matter of electrostatics

Electrostatics in DNA-protein interactions

The problem of sequence recognition

Analysis of MD simulation data

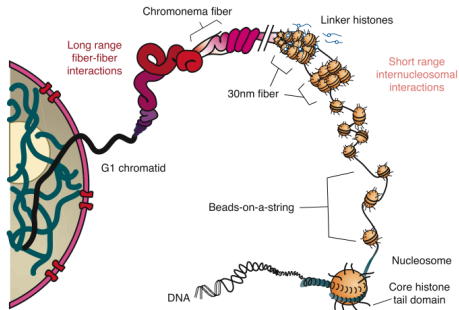
From simulation data to implicit solvent models

## DNA-protein interactions in the nucleus



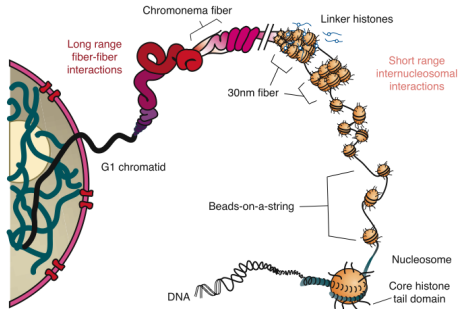
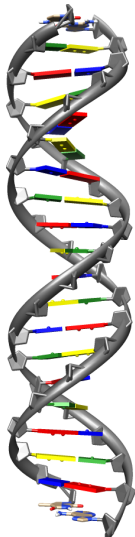


## DNA-protein interactions in the nucleus

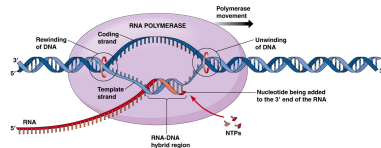


Source Horn 2002

## DNA-protein interactions in the nucleus



Source Horn 2002



## The role of salt in electrostatic interactions

### Ions in the solvent

- ▶  $\text{NaCl} \rightarrow \text{Na}^+, \text{Cl}^-$
- ▶  $\text{KCl} \rightarrow \text{K}^+, \text{Cl}^-$
- ▶ ...

Physiological concentration

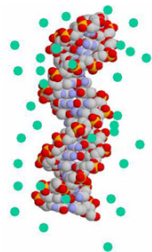
$\sim 0.1 \text{ mol/L}$

## The role of salt in electrostatic interactions

### Ions in the solvent

- ▶  $\text{NaCl} \rightarrow \text{Na}^+, \text{Cl}^-$
- ▶  $\text{KCl} \rightarrow \text{K}^+, \text{Cl}^-$
- ▶ ...

Physiological concentration  
 $\sim 0.1 \text{ mol/L}$

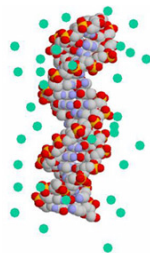


## The role of salt in electrostatic interactions

### Ions in the solvent

- ▶ NaCl  $\rightarrow$  Na<sup>+</sup>, Cl<sup>-</sup>
- ▶ KCl  $\rightarrow$  K<sup>+</sup>, Cl<sup>-</sup>
- ▶ ...

Physiological concentration  
 $\sim 0.1$  mol/L



### Interaction between charged particles

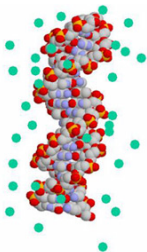
- ▶ No counterions:  $V(r) \propto 1/r$
- ▶ Counterions:  $V(r) \propto e^{-\kappa r}/r$

## The role of salt in electrostatic interactions

### Ions in the solvent

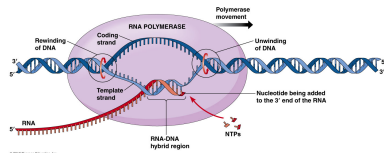
- ▶ NaCl  $\rightarrow$  Na<sup>+</sup>, Cl<sup>-</sup>
- ▶ KCl  $\rightarrow$  K<sup>+</sup>, Cl<sup>-</sup>
- ▶ ...

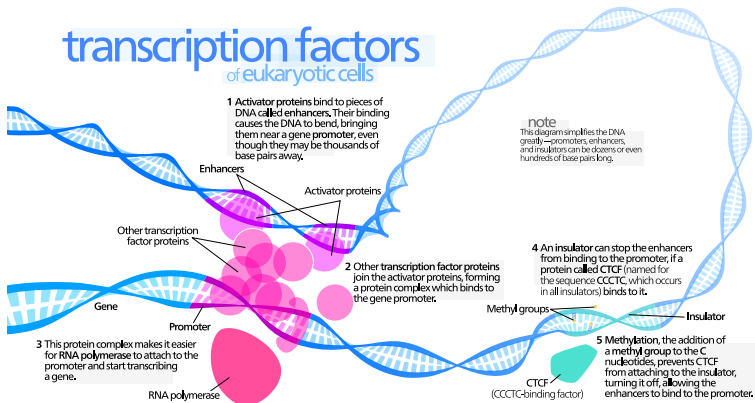
Physiological concentration  
 $\sim 0.1$  mol/L

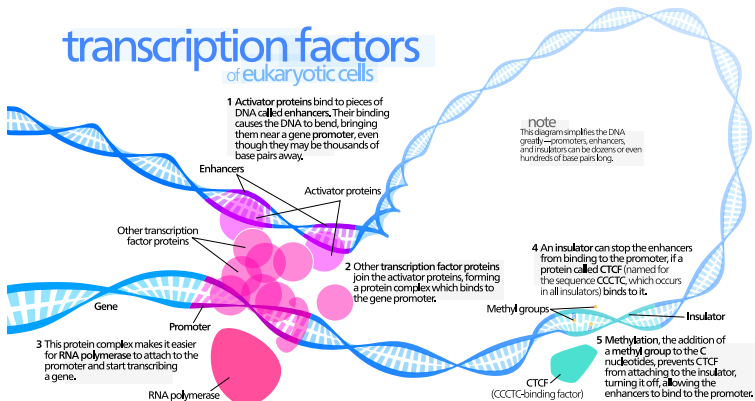


### Interaction between charged particles

- ▶ No counterions:  $V(r) \propto 1/r$
- ▶ Counterions:  $V(r) \propto e^{-\kappa r}/r$





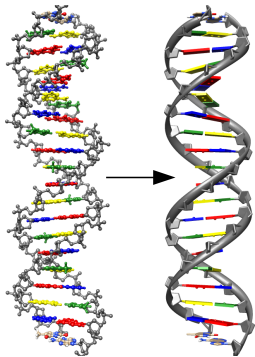


Mechanisms of indirect sequence recognition by proteins ?



## Role of sequence-dependent DNA mechanics

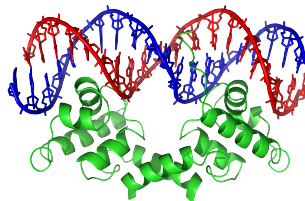
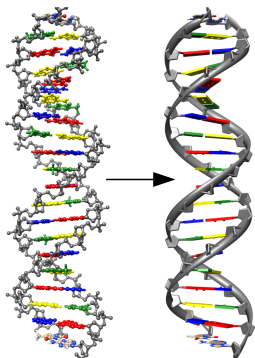
Nanoscale mechanical models of  
DNA from all-atomic simulation



*S. Meyer et al. Biophysical J. 2013*

## Role of sequence-dependent DNA mechanics

Nanoscale mechanical models of  
DNA from all-atomic simulation

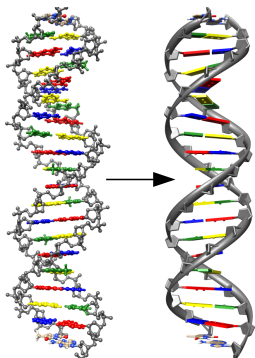


*N. Becker et al. Nucl. Ac. Res. 2006*

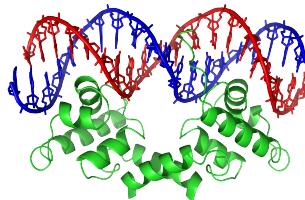
*S. Meyer et al. Biophysical J. 2013*

## Role of sequence-dependent DNA mechanics

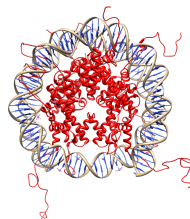
Nanoscale mechanical models of  
DNA from all-atomic simulation



*S. Meyer et al. Biophysical J. 2013*



*N. Becker et al. Nucl. Ac. Res. 2006*



*S. Meyer et al. In preparation*

Role of the solvent:  
sequence-dependent distribution of counter-ions ?

# Outline

DNA transactions: a matter of electrostatics

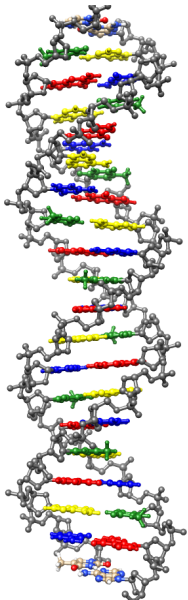
Analysis of MD simulation data

Ions around a DNA oligomer

Data analysis: convergence

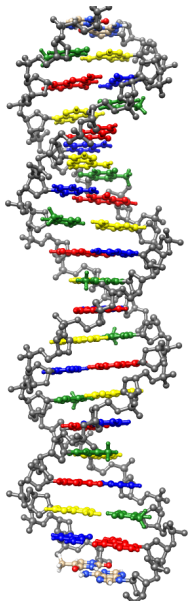
Data analysis: effect of sequence

From simulation data to implicit solvent models



## Source of data ?

- ▶ Scarce experimental data
- ▶ Molecular Mechanics model. Empirical force fields for DNA [Cornell et al. 1995] and water (SPC/E)
- ▶ Simulation method: Molecular Dynamics (Newton's laws): "Numerical experiment"
- ▶ In the present study, DNA is kept rigid.

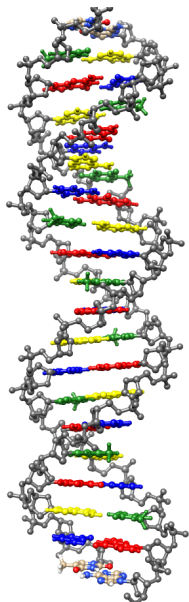


## Source of data ?

- ▶ Scarce experimental data
- ▶ Molecular Mechanics model. Empirical force fields for DNA [Cornell et al. 1995] and water (SPC/E)
- ▶ Simulation method: Molecular Dynamics (Newton's laws): "Numerical experiment"
- ▶ In the present study, DNA is kept rigid.

## Computational aspects

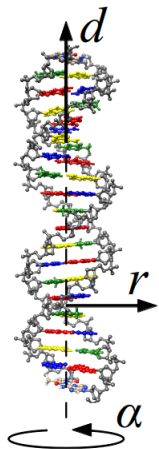
- ▶ 18-mer of DNA, with 3x repeating AGCT sequence
- ▶ ~ 30 000 atoms in the simulation box (solvent 90%)
- ▶ Timestep: 2 fs. Trajectory: 400 ns = 20 million steps. Computation time: ~ 6 weeks on 32 CPUs (PSMN)



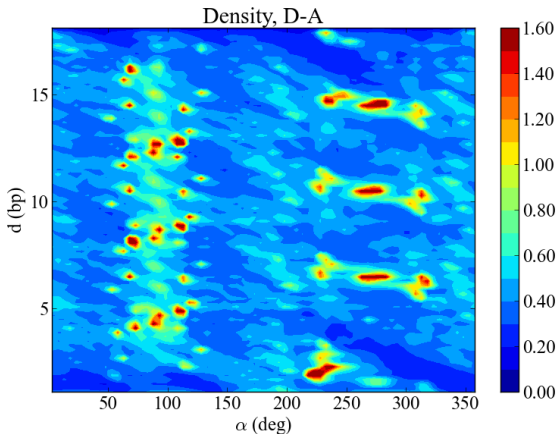


## Analysis: helical coordinates following DNA geometry

- ▶ Ion distribution wrt the DNA structure [*R. Lavery et al 2009*]
- ▶ Convergence? Ion distribution reflects the sequence

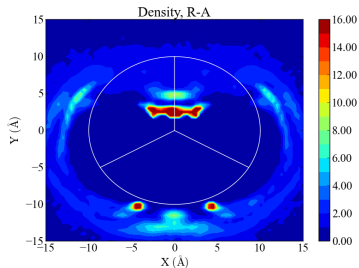
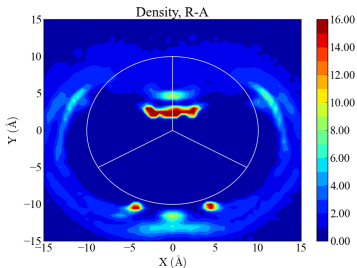


C  
G  
T  
C  
G  
A  
T  
C  
G  
A  
T  
C  
C  
G

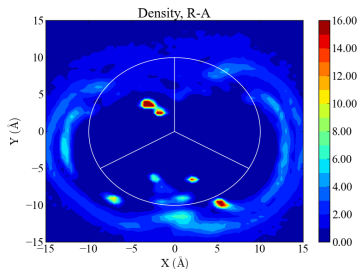
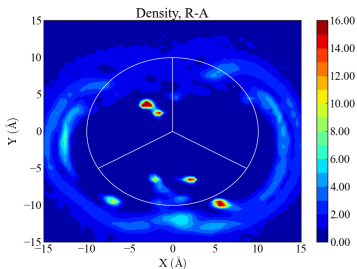


## Effect of sequence on ion distribution

GC

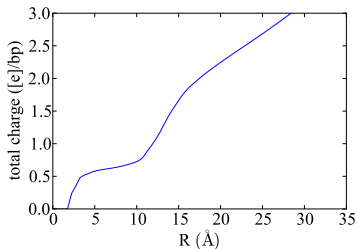


AG

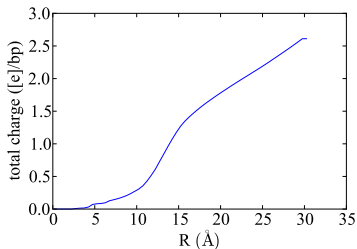


## Integrated radial cation density

GC

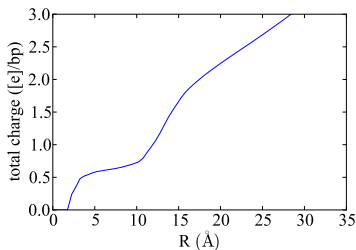


AG

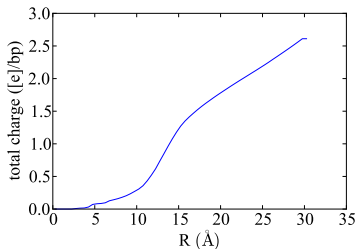


## Integrated radial cation density

GC



AG

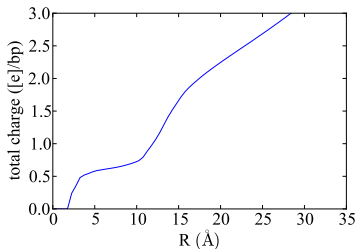


Mechanism of indirect sequence recognition by DNA proteins ?

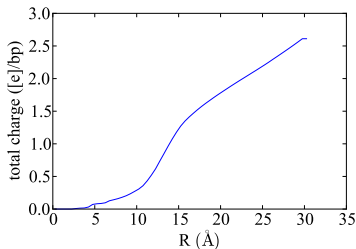
- ▶ DNA negative charges are differently screened by the ions

## Integrated radial cation density

GC



AG



## Mechanism of indirect sequence recognition by DNA proteins ?

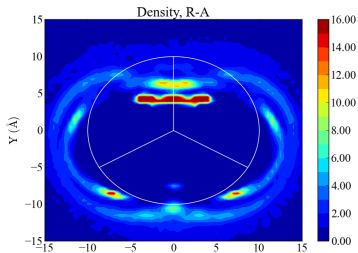
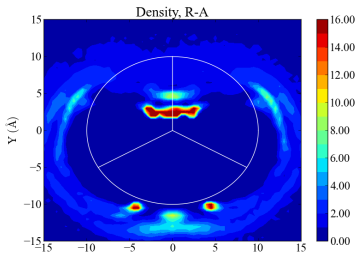
- ▶ DNA negative charges are differently screened by the ions
- ▶ Charged proteins can experience the sequence-dependent electrostatic potential, even without direct contact with the bases

## Mechanism: DNA sequence-dependent structure vs. chemistry

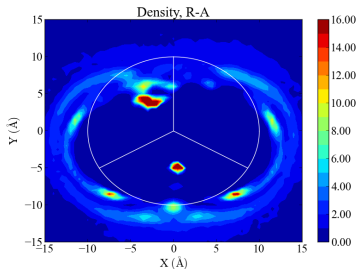
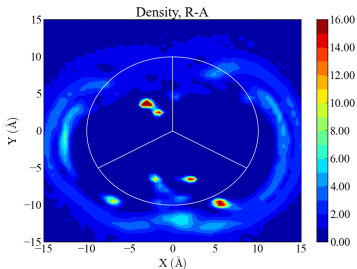
Seq. dependent structure

B-DNA

GC



AG



## Conclusion of data analysis

- ▶ Requires a specific analysis adapted to the physico-chemical properties of the system
- ▶ Sequence-dependent charge distribution in the solvent
- ▶ Possible mechanism of solvent-mediated sequence recognition

# Outline

DNA transactions: a matter of electrostatics

Analysis of MD simulation data

From simulation data to implicit solvent models

General problem

Benchmark

Conclusion



## Objective: modeling of protein-DNA interactions

- ▶ Larger systems
- ▶ Computation of binding free energies

## Objective: modeling of protein-DNA interactions

- ▶ Larger systems
- ▶ Computation of binding free energies

## Solvent models

- ▶ Explicit MD

## Objective: modeling of protein-DNA interactions

- ▶ Larger systems
- ▶ Computation of binding free energies

## Solvent models

- ▶ Explicit MD
- ▶ Implicit solvent: continuous description of lower resolution (coarse-grained)

## Objective: modeling of protein-DNA interactions

- ▶ Larger systems
- ▶ Computation of binding free energies

## Solvent models

- ▶ Explicit MD
- ▶ Implicit solvent: continuous description of lower resolution (coarse-grained)

## Poisson-Boltzmann theory and derivatives

- ▶ Mean-field continuous description of the ion positions.

## Objective: modeling of protein-DNA interactions

- ▶ Larger systems
- ▶ Computation of binding free energies

## Solvent models

- ▶ Explicit MD
- ▶ Implicit solvent: continuous description of lower resolution (coarse-grained)

## Poisson-Boltzmann theory and derivatives

- ▶ Mean-field continuous description of the ion positions.
- ▶ Commonly used in biomolecular modeling

## Objective: modeling of protein-DNA interactions

- ▶ Larger systems
- ▶ Computation of binding free energies

## Solvent models

- ▶ Explicit MD
- ▶ Implicit solvent: continuous description of lower resolution (coarse-grained)

## Poisson-Boltzmann theory and derivatives

- ▶ Mean-field continuous description of the ion positions.
- ▶ Commonly used in biomolecular modeling
- ▶ Recent improvements improve the description of the finite ion size [*P. Koehl et al. 2009*]

## Objective: modeling of protein-DNA interactions

- ▶ Larger systems
- ▶ Computation of binding free energies

## Solvent models

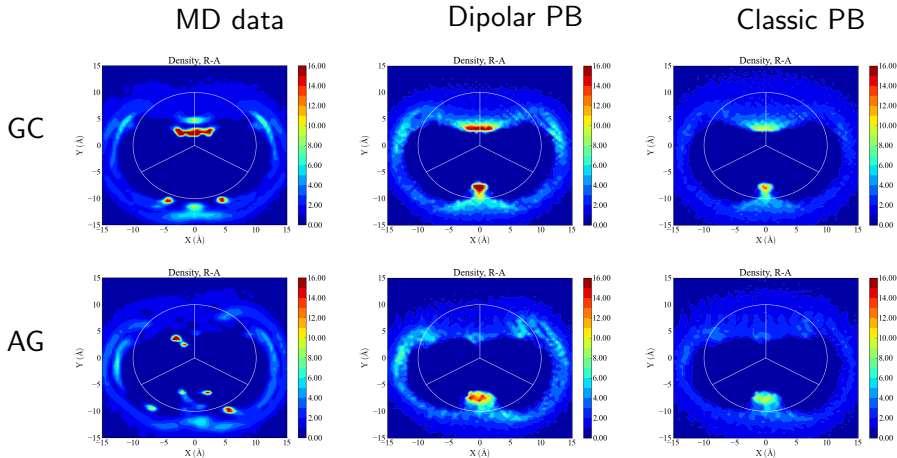
- ▶ Explicit MD
- ▶ Implicit solvent: continuous description of lower resolution (coarse-grained)

## Poisson-Boltzmann theory and derivatives

- ▶ Mean-field continuous description of the ion positions.
- ▶ Commonly used in biomolecular modeling
- ▶ Recent improvements improve the description of the finite ion size [*P. Koehl et al. 2009*]

Use simulation data as a benchmark for the coarse-grained models

## Comparison of coarse-grained models with MD data





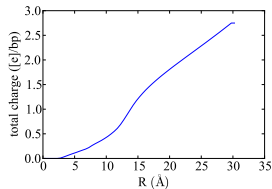
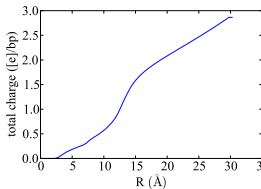
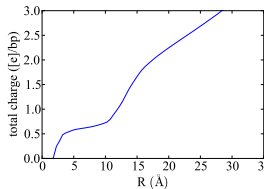
## Integrated radial cation density

MD data

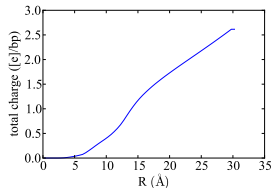
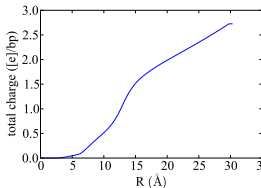
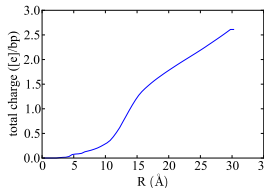
Dipolar PB

Classic PB

GC



AG



## First results

- ▶ Qualitative agreement
- ▶ At the quantitative level, coarse-grained models miss the strongly bound ions inside DNA

## First results

- ▶ Qualitative agreement
- ▶ At the quantitative level, coarse-grained models miss the strongly bound ions inside DNA

## Construction of hybrid models

- ▶ Impose the position of the strongly bound ions for the coarse-grained calculations
- ▶ Toward multi-scale models ?