Simulating dynamic processes involving DNA

Krystyna Zakrzewska

Laboratoire de Biologie et Chimie des Protein

UMR 5086 CNRS - Université Lyon 1) - http://www.ibcp.fr

k.zakrzewska@ibcp.fr

Nice, october 09

I will be talking about

- Influence of Conformational Dynamics on the Exciton states of DNA
- Studies of DNA Conformation : ABC

Influence of Conformational Dynamics on the Exciton States of DNA

Motivation

- **Photoreactions upon absorption of UV radiation by DNA can have lethal or mutagenic effects.**
- **Whether the excited singlet states are localised on one base or extend over certain number**
- **Standard viewpoint: excitation is localised since exciton splitting is not observed**
- **The progress in experimental and theoretical methods**

Exciton formalism

$$H = H_0 + V$$

$$H_0 = \sum_{molecules,k}^{n} H_k \quad \text{and} \quad \left|\Phi_m^i\right\rangle = \Psi_m^i \prod_{molecules,k\neq m}^{n} \Psi_k^0$$

$$V = \iint \frac{\rho^{(k)}(\vec{r}^{(k)})\rho^{(l)}(\vec{r}^{(l)})}{\left|\vec{r}^{(k)} - \vec{r}^{(l)}\right|} d\vec{r}^{(k)}\, d\vec{r}^{(l)}$$

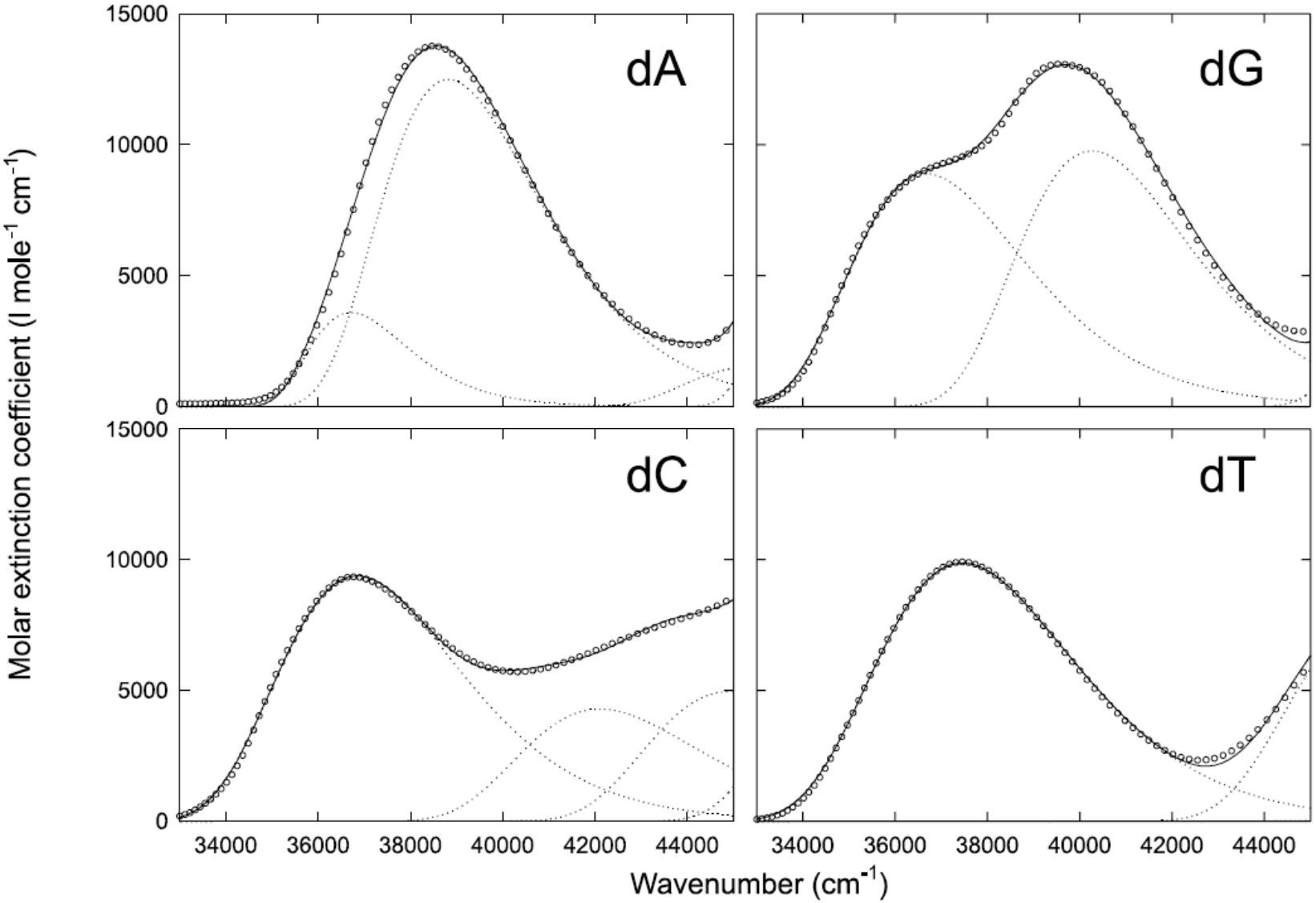
Longuet-Higgins

$$\left\langle \Psi_n^i \Psi_m^j \middle| V \middle| \Psi_n^k \Psi_m^l \right\rangle = \frac{\left\langle \Psi_n^i \middle| \rho^{(n)}\vec{r} \middle| \Psi_n^k \right\rangle \left\langle \Psi_m^j \middle| \rho^{(m)}\vec{r} \middle| \Psi_m^l \right\rangle}{\left|\vec{r}^{(n)} - \vec{r}^{(m)}\right|}$$

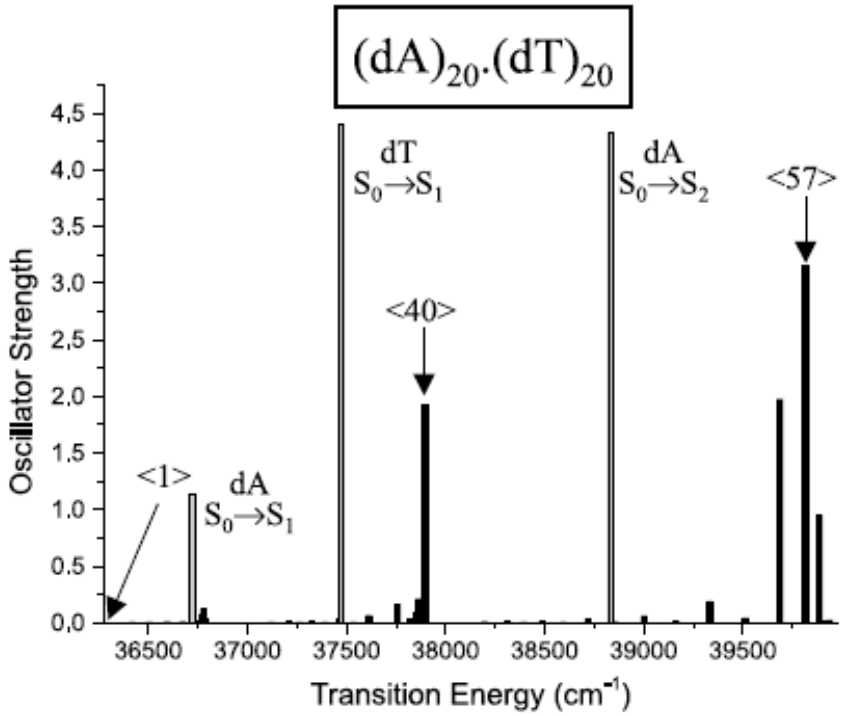
Interaction between transition charges

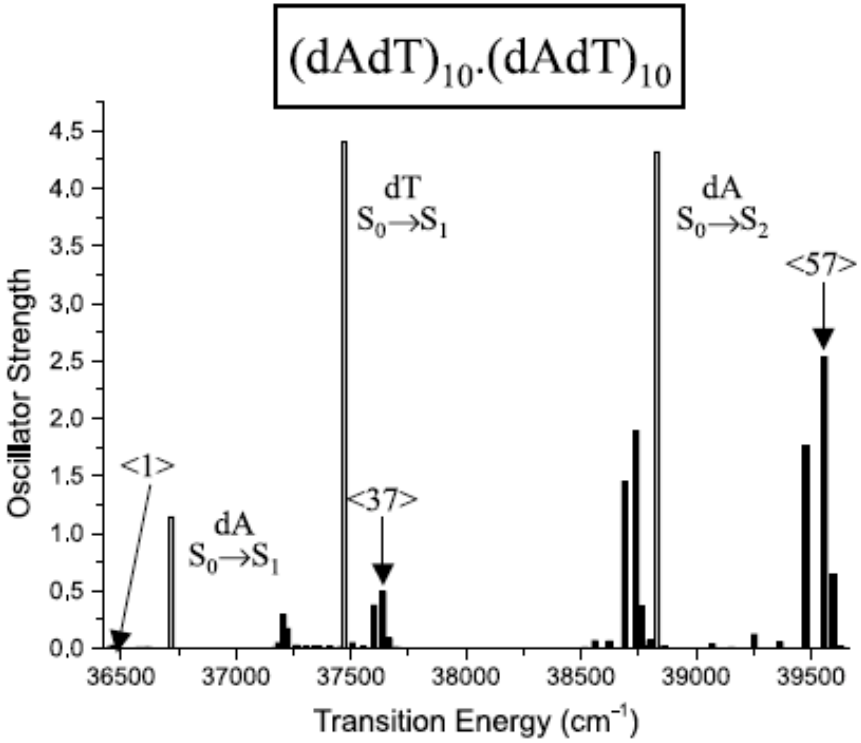
$$\left\langle \Phi_m^i \middle| H \middle| \Phi_m^i \right\rangle = E_0 + \underbrace{\left(\varepsilon_m^i - \varepsilon_m^0\right)}_{I} + \underbrace{\left(\sum_{n\neq m} E_{m,n}^{i,0} + \sum_{n\neq m}\sum_{p<n,p\neq m} E_{n,p}^{0,0}\right)}_{II} - \underbrace{\left(\sum_{n\neq m} E_{m,n}^{0,0} + \sum_{n\neq m}\sum_{p<n,p\neq m} E_{n,p}^{0,0}\right)}_{III}$$

Experimental absorption spectra of bases



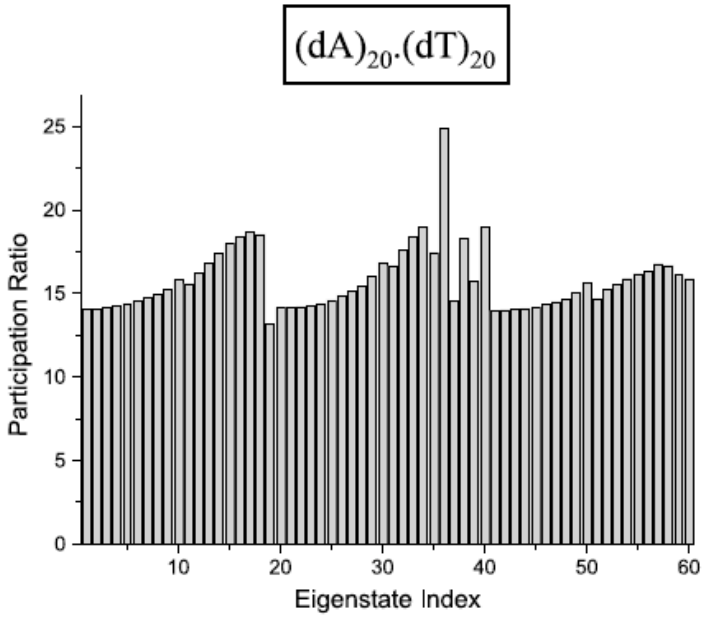
Regular B-DNA conformation

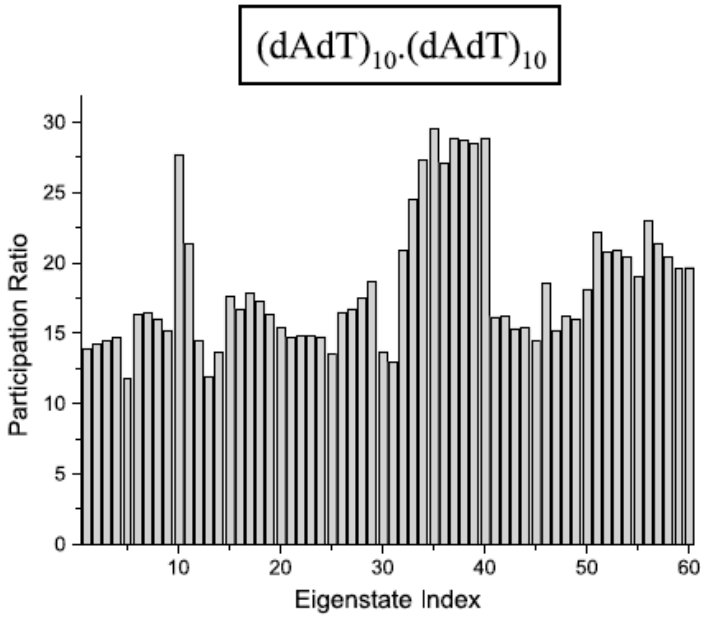




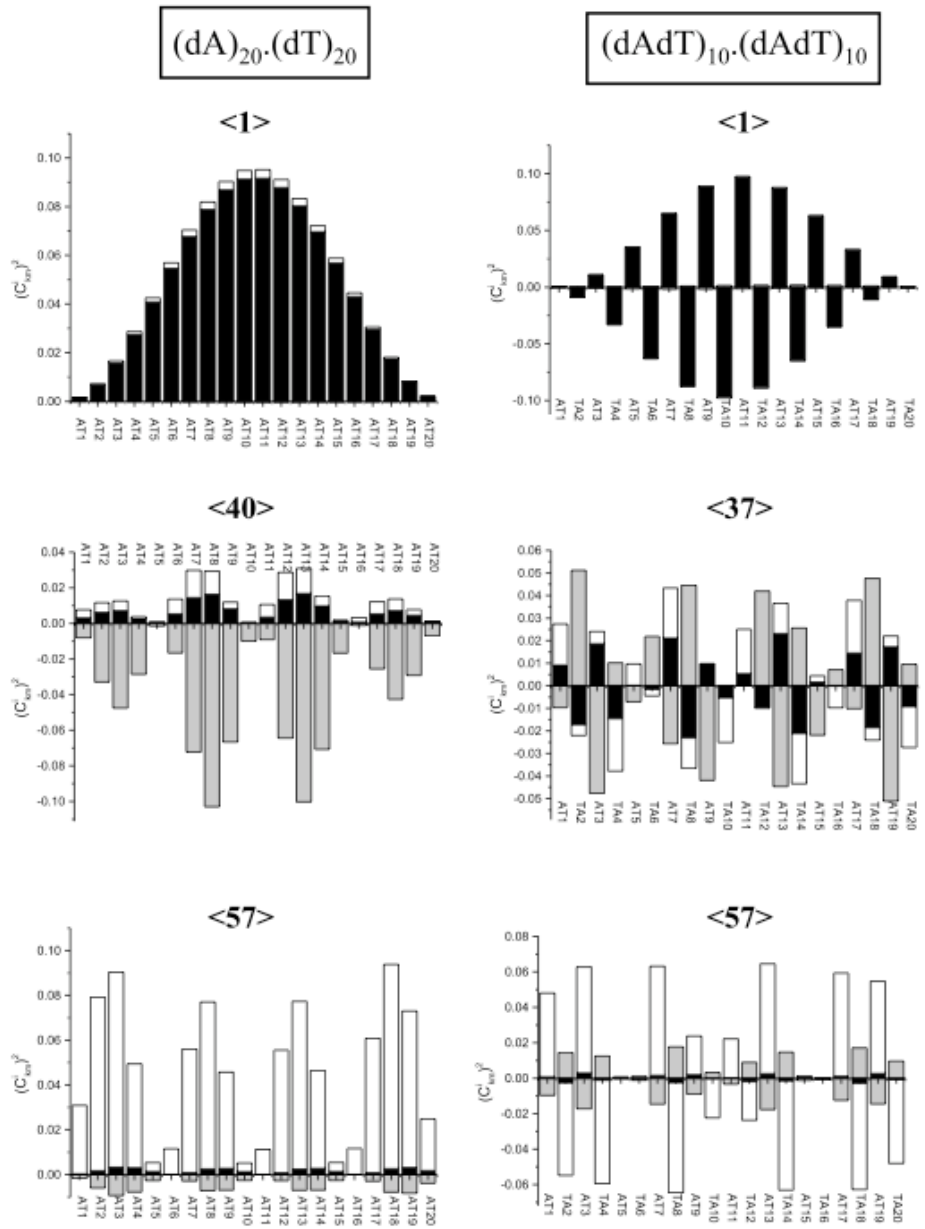
PR=1/L_K

$$L_k = \sum_{\text{molecules } m} \left[\sum_{\text{states } i} (C^i_{k,m})^2 \right]^2$$





Topography of 3 typical eigenstates



black –adenine $\mathbf{S_1}$

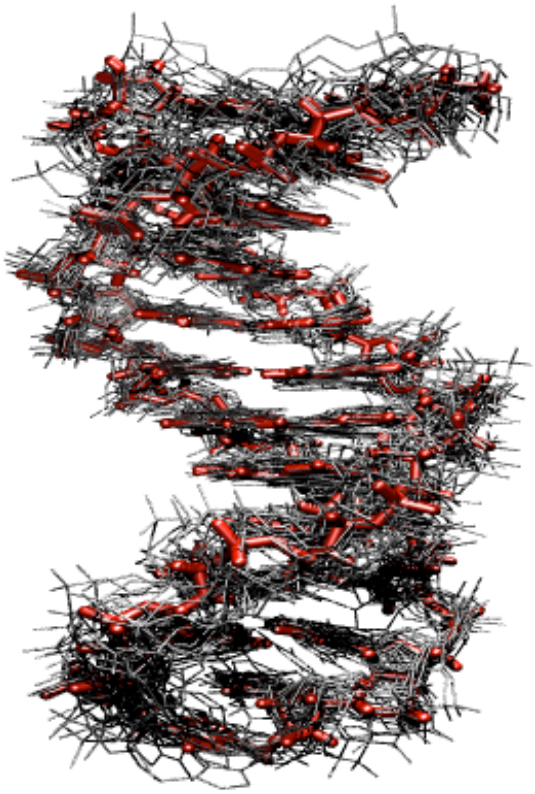
white – adenine $\mathbf{S_2}$

grey thymine S_1

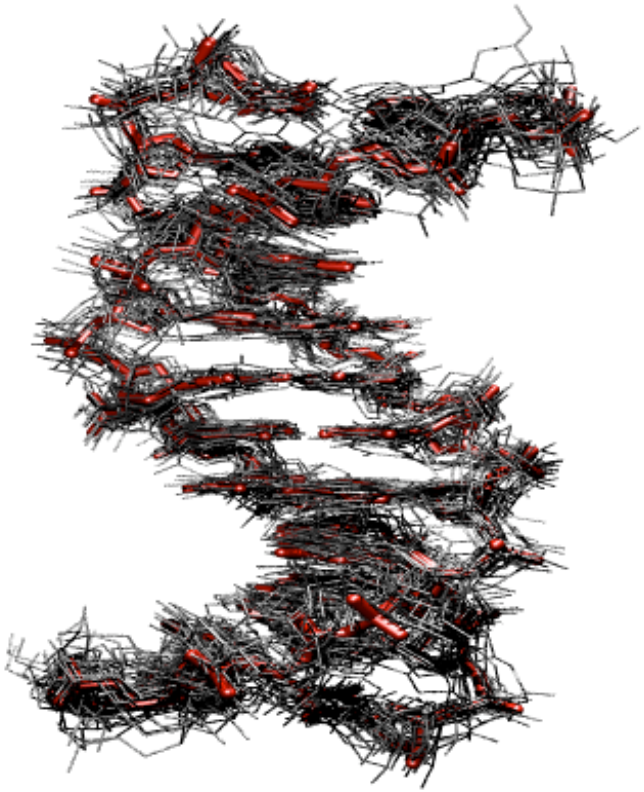
highest oscillator strength

DNA Dynamic Structure

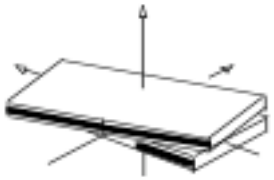
$(dA)_{12}.(dT)_{12}$



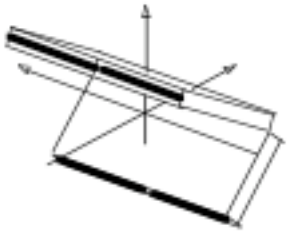
$(dAdT)_6.(dT)_6$



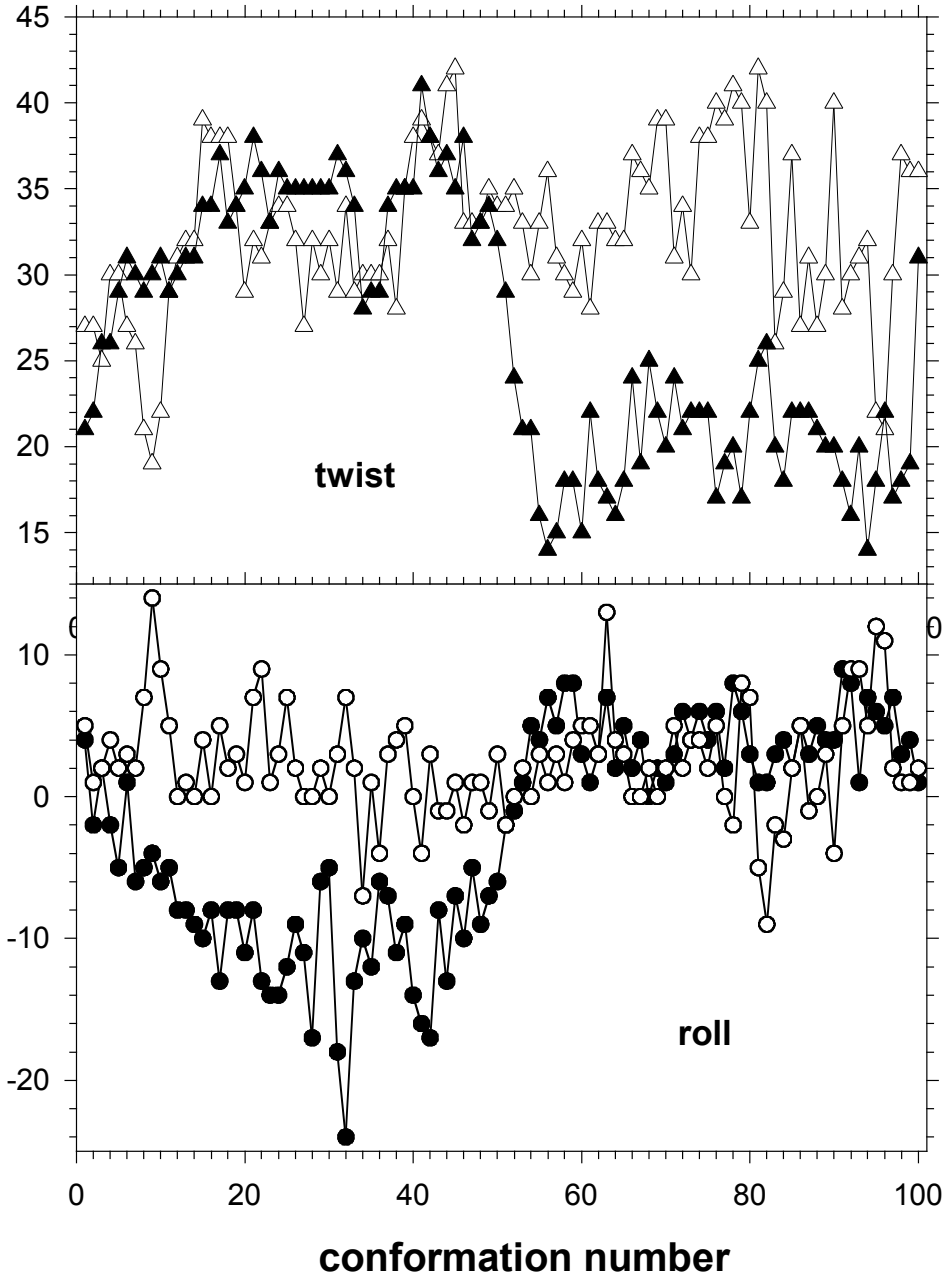
Structural parameters



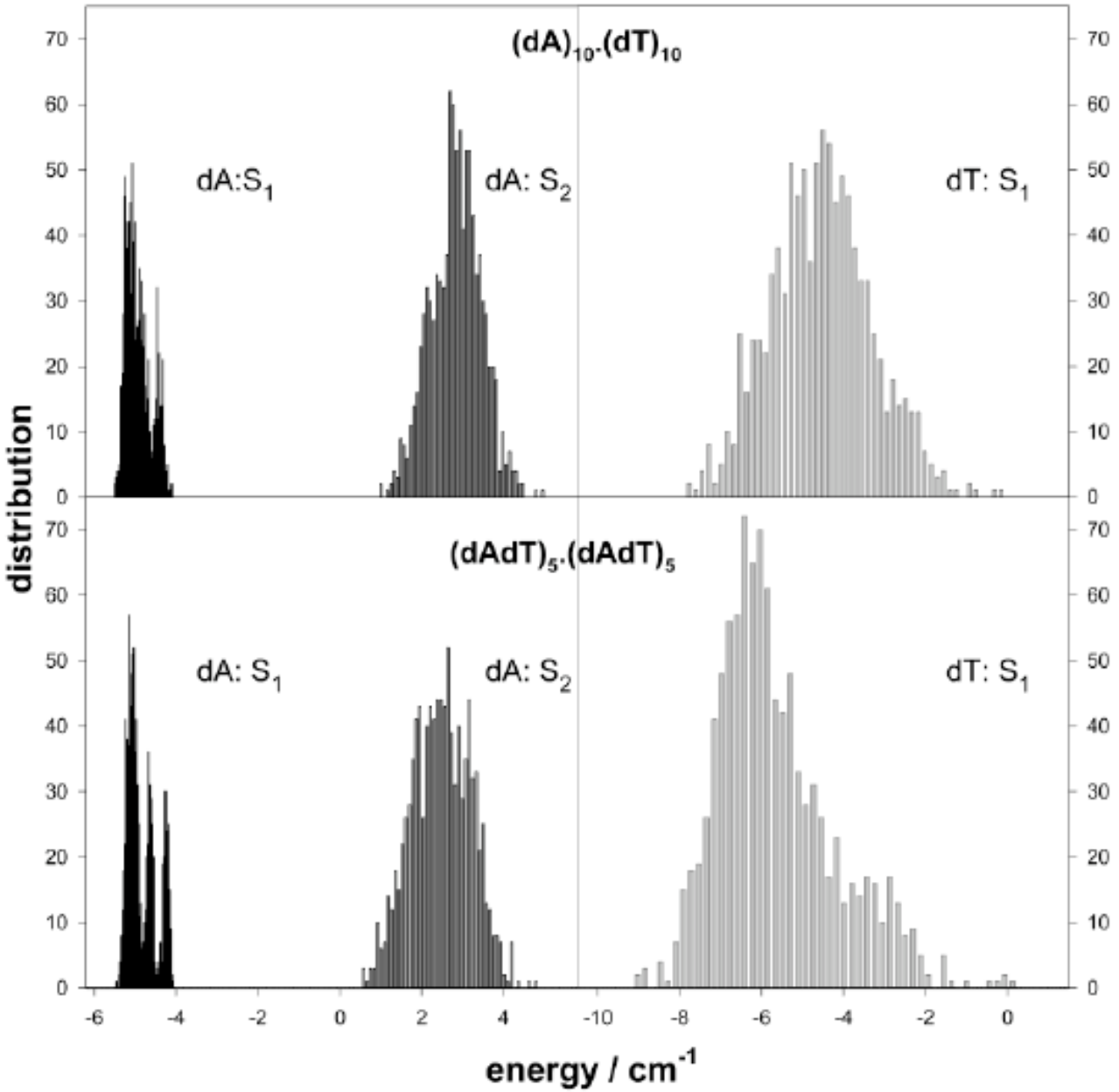
Twist



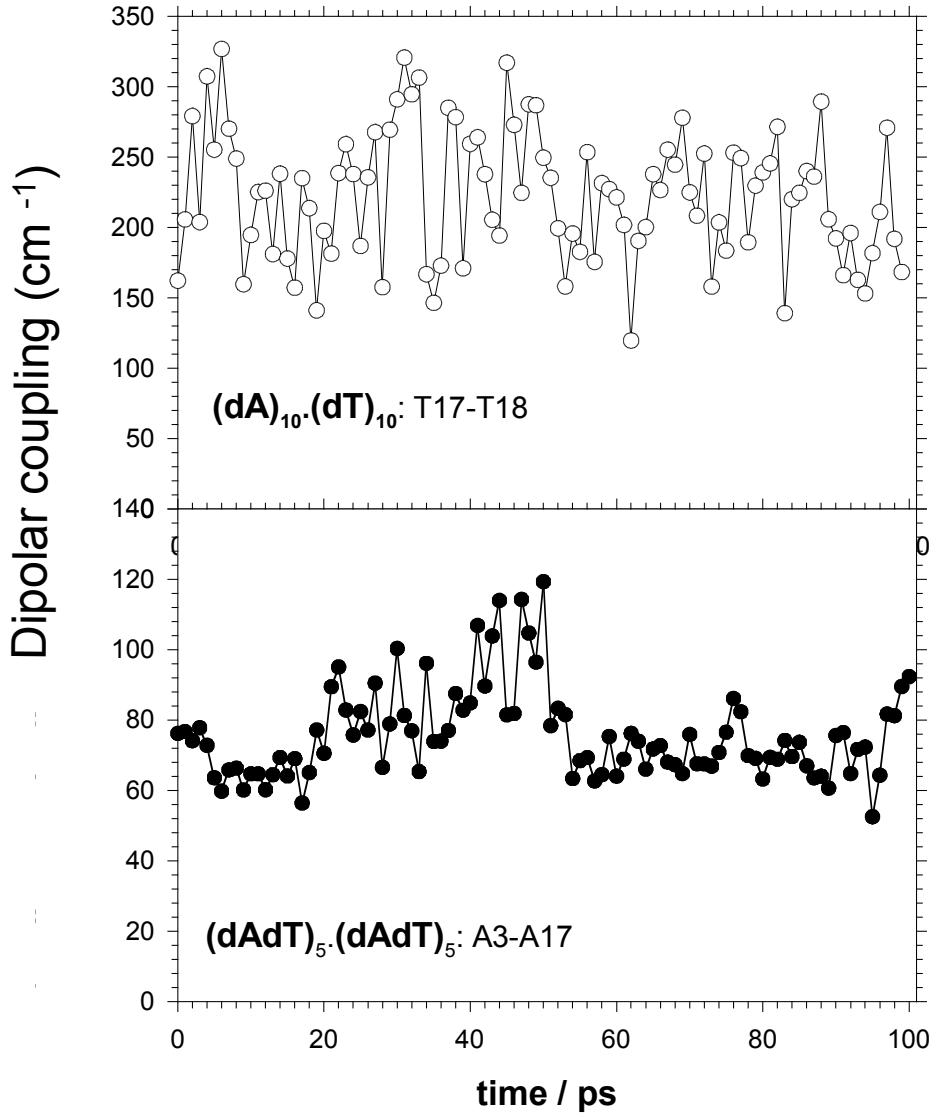
Roll



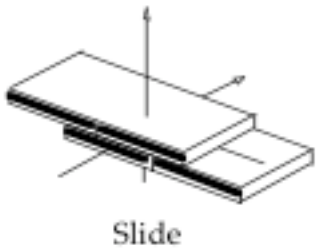
Diagonal term distribution

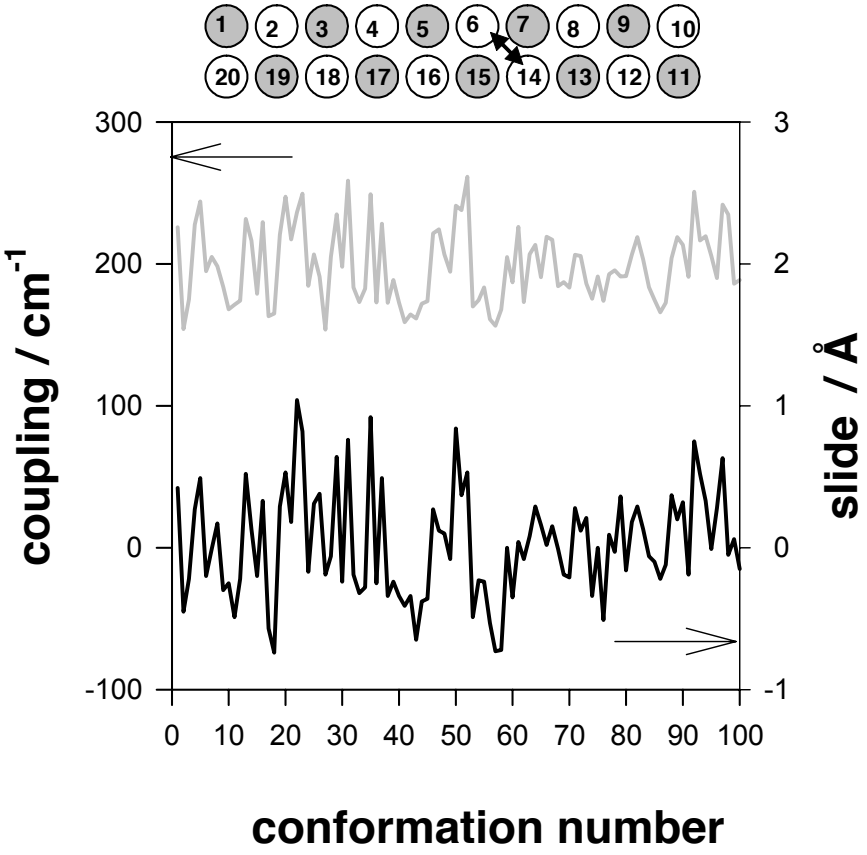


Dipolar coupling for 100 conformations

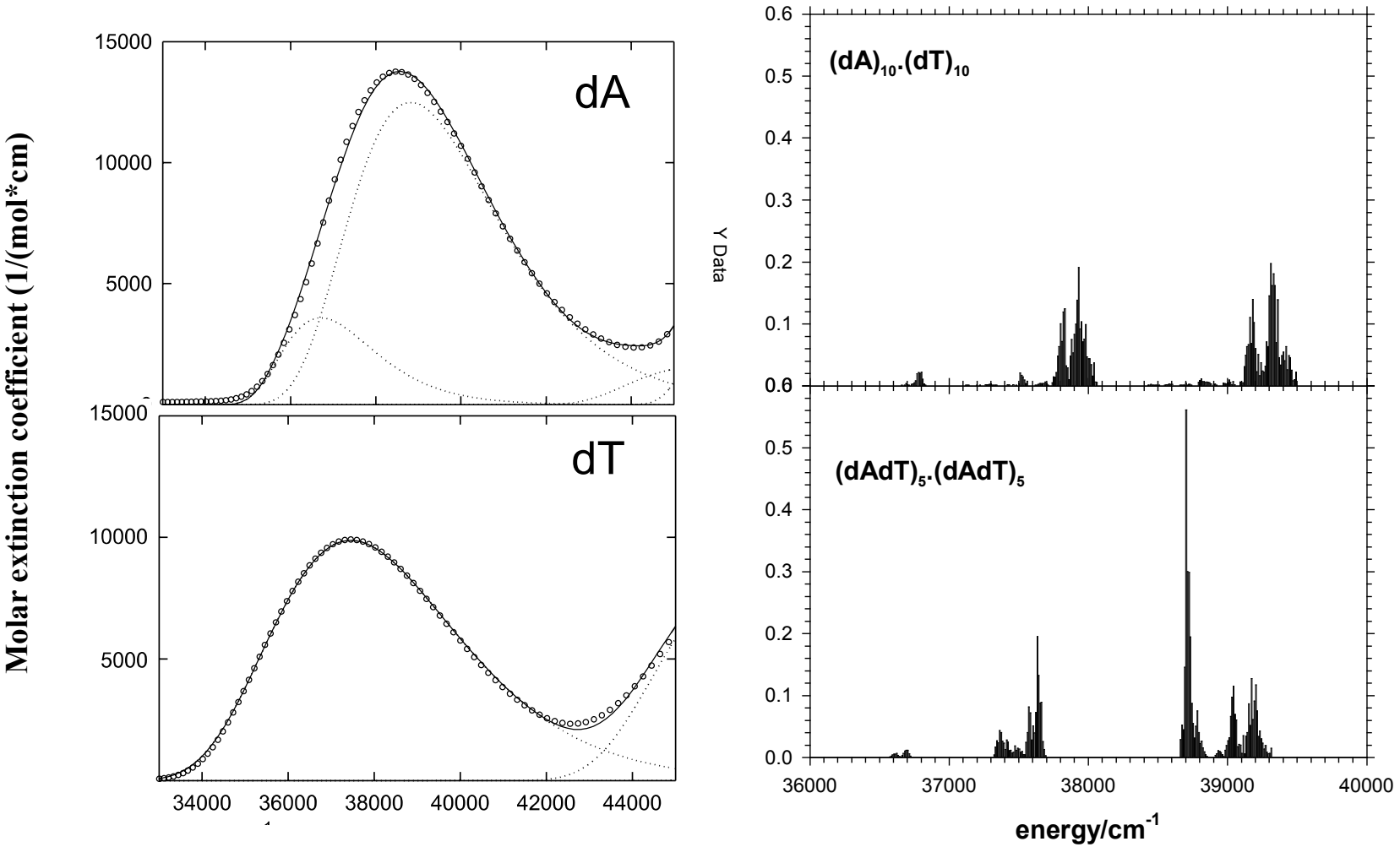


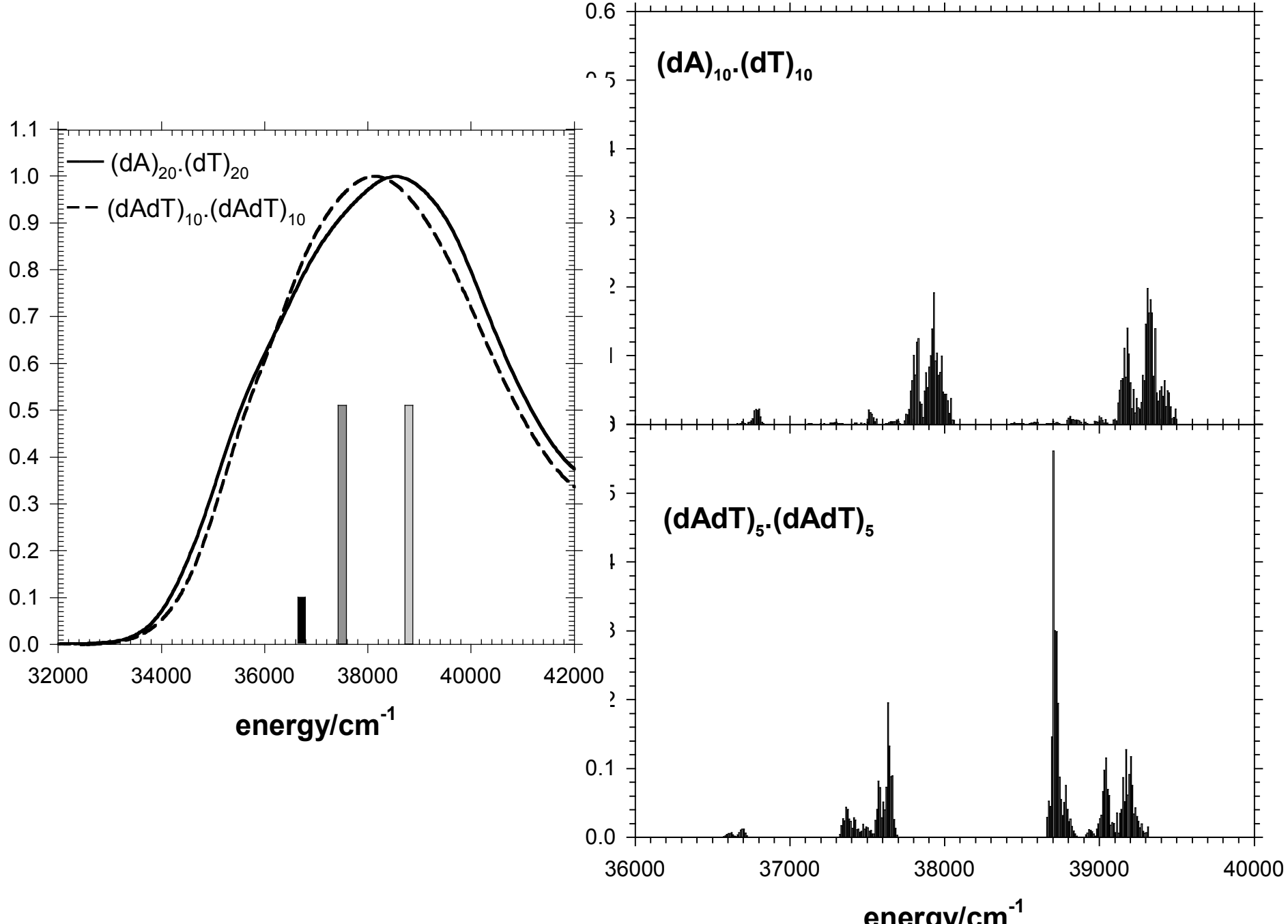
Correlation between dipolar coupling and helicoidal parameters



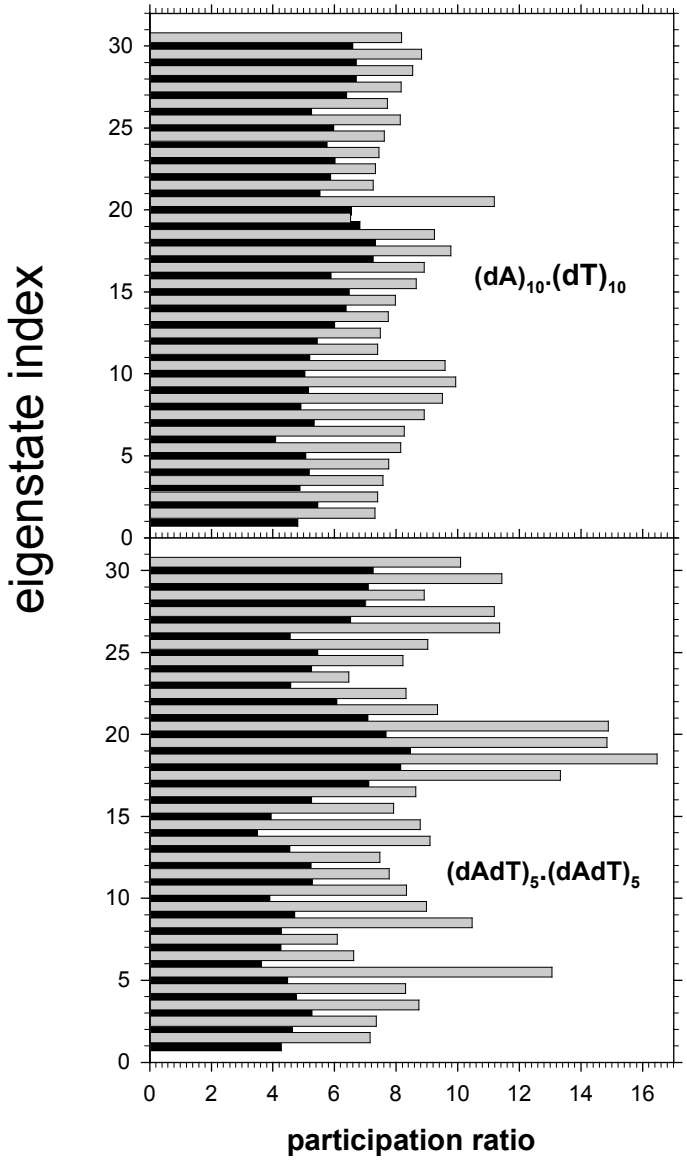


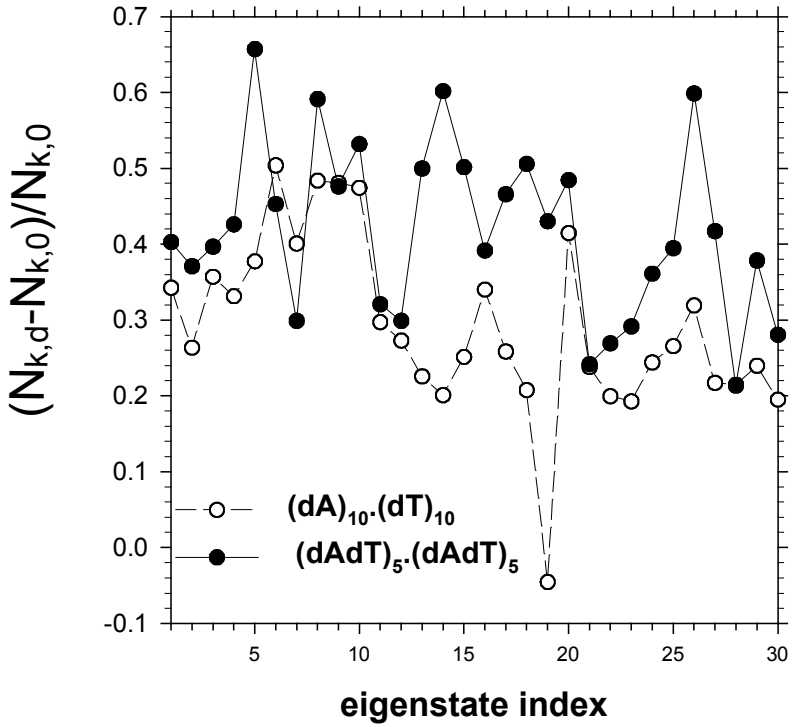
Comparison of the absorption spectra




(dA)20.(dT)20
(dAdT)10.(dAdT)10
energy/cm-1
(dA)10.(dT)10
(dAdT)5.(dAdT)5
energy/cm-1

Reduction of spatial extent by dynamics

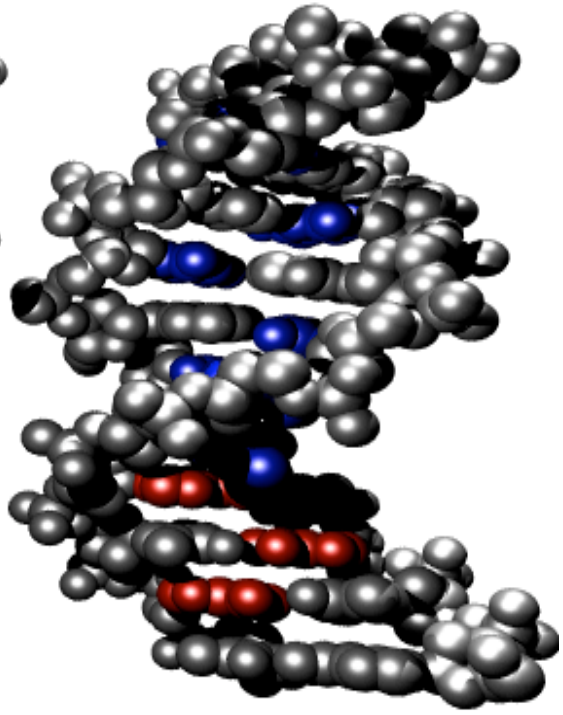




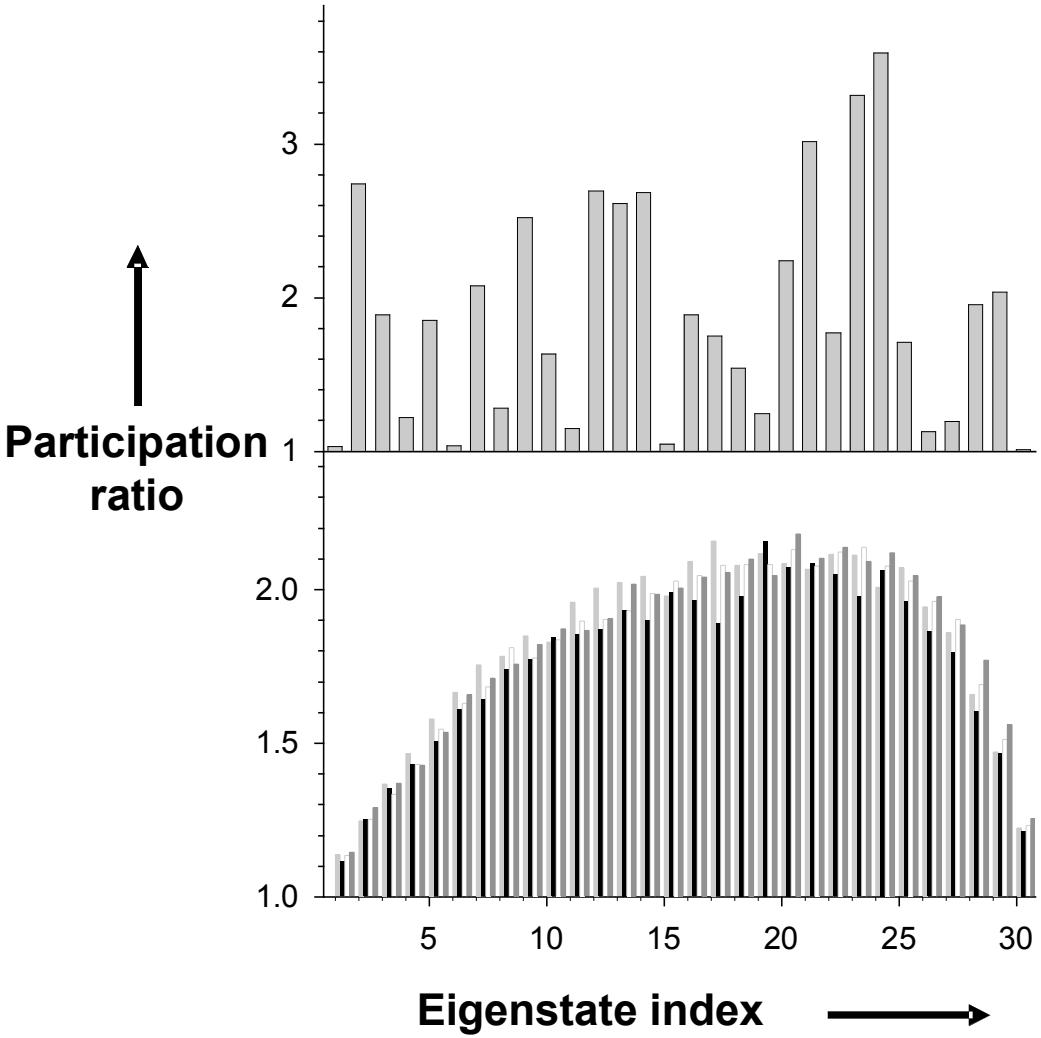
Topography of the eigenstates <1> in red, <11> in blue

$(dA)_{12}.(dT)_{12}$

$(dAdT)_6.(dAdT)_6$



Participation ratio for 30 eigenstates of $(dA)_{10}(dT)_{10}$



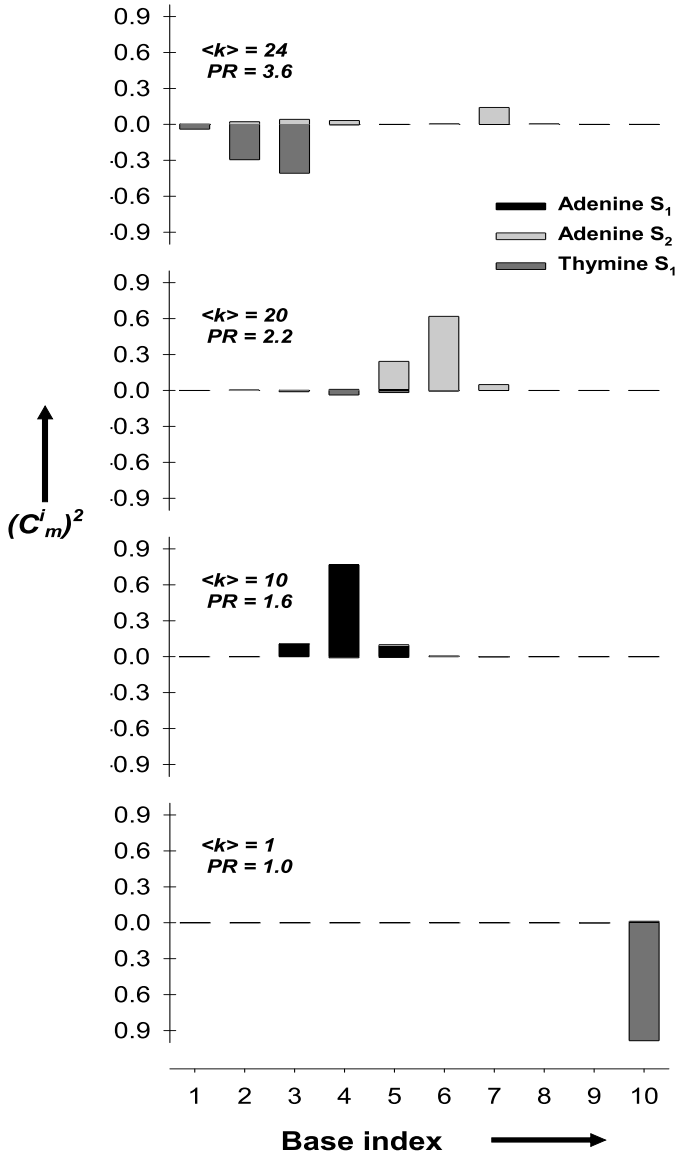
Single conformation

Single distribution of monomer transition energies

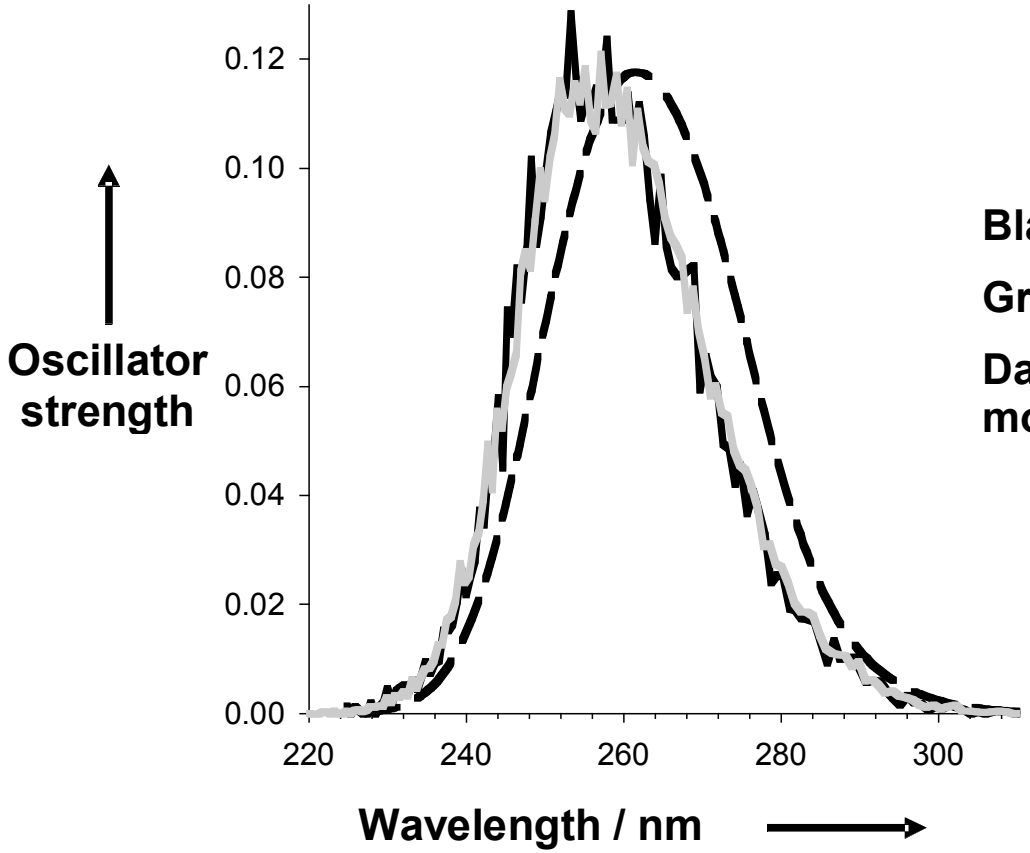
Four conformations

500 distributions of monomer transition energies

Topography of a few eigenstates, $(AT)_{10}\,(AT)_{10}$



Calculated absorption spectra for $(AT)_{10}$ $(AT)_{10}$

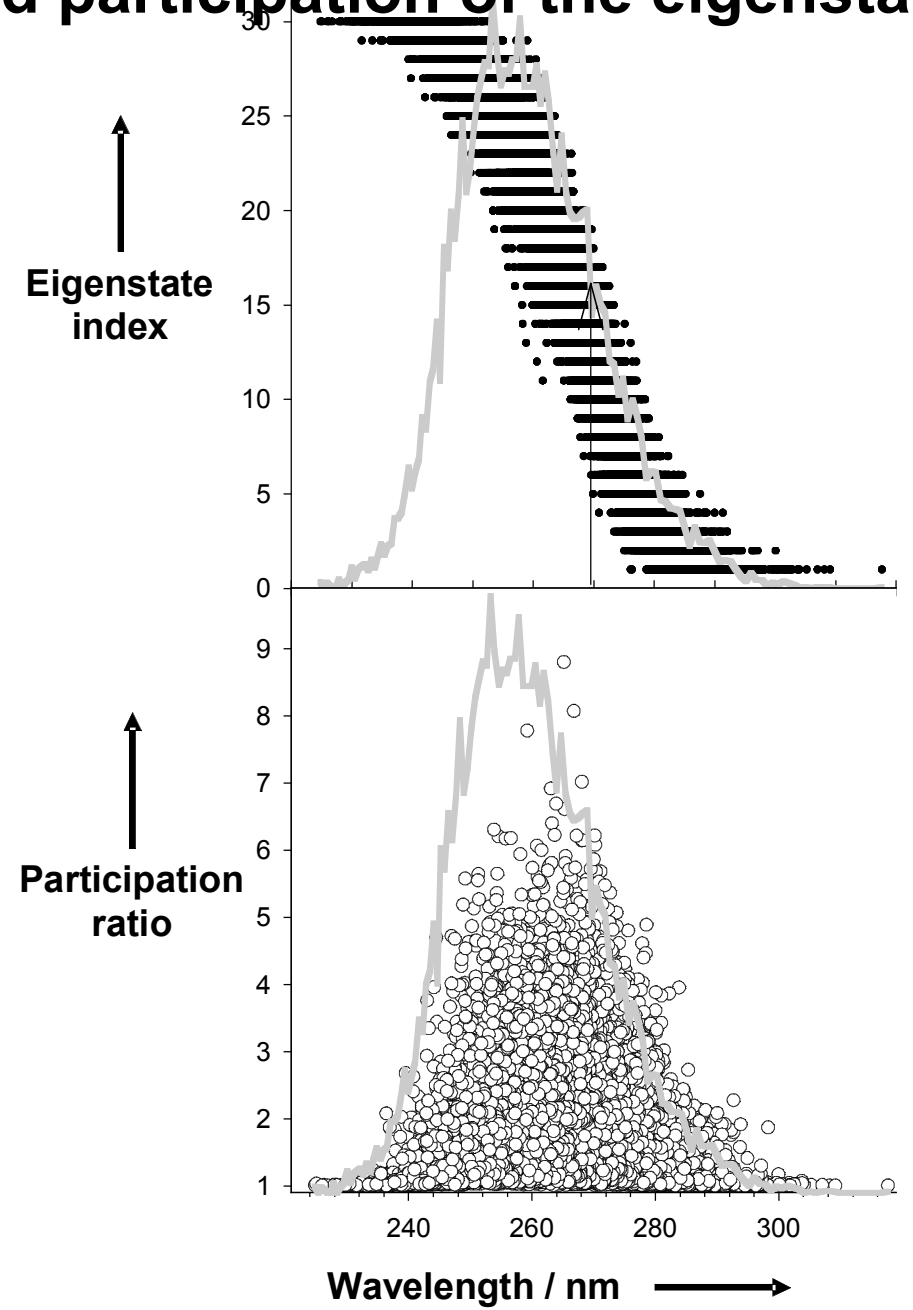


Black- single conformation

Grey -4 conformations

Dashed –non interacting monomers

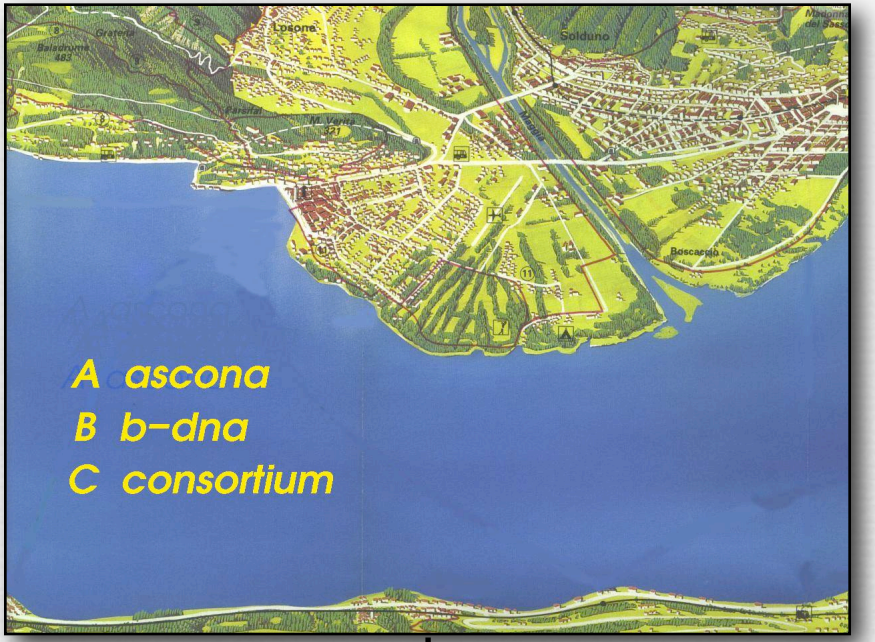
Position and participation of the eigenstates



Conclusions

- Oligomer absorption is only slightly shiftd to higher energies
- Off-diagonal terms correlated with helicoidal parameter fluctuations
- Fluctuations reduce the delocalisation
- Influence of off-diagonal disorder smaller then diagonal disorder

DNA Conformational studies : ABC



D. Beveridge, Wesleyan U., USA

T. Bishop, Tulane U., USA

D. Case, Rutgers U., USA

T. Cheatham, U. Utah, USA

B. Jayaram, IIT New Delhi, India

F. Lankas, Prague, Czech Rep.

R. Lavery, IBCP, France

C. Laughton, Nottingham, UK

J. Maddocks, EPFL, Switzerland

M. Orozco, Barcelona, Spain

R. Osman, Mount Sinai, NY, USA

A. Pérez, Barcelona, Spain

H. Sklenar, MDC, Germany

N. Spackova, Brno, Czech Rep.

K. Sponar, Brno, Czech Rep.

M. Young, Berkeley, USA

K. Zakrzewska, IBCP, France

Beveridge et al. *Biophys. J.* 87 (2004) 3799

Dixit et al. *Biophys. J.* 89 (2006) 3721

Lavery et al. *Nucleic Acids Res.* (2009) submitted

Molecular Dynamics

Molecular dynamics simulation consists of the numerical, step-by-step, solution of the classical equations of motion, which for a simple atomic system may be written:

$$F_i = m_i a_i$$

$$F_i = -\nabla V$$

$$-\frac{dV}{dr_i} = m_i \frac{d^2 r_i}{dt^2}$$

Integration algorithms

Velocity Verlet algorithm

$$r(t+dt)=r(t)+v(t)dt+\tfrac{1}{2}a(t)dt^2$$

$$v(t+dt)=v(t)+\tfrac{1}{2}[a(t)+a(t+dt)]dt$$

Leap-frog algorithm

$$r(t+dt)=r(t)+v(t+\tfrac{1}{2}dt)dt$$
$$v(t+\tfrac{1}{2}dt)=v(t-\tfrac{1}{2}dt)+a(t)dt$$

Force field used

$$\sum_{valence} K(\theta - \theta_{eq})^2 + \sum_{dihedrals} V_n / 2[1 + \cos(n\phi - \gamma)]$$

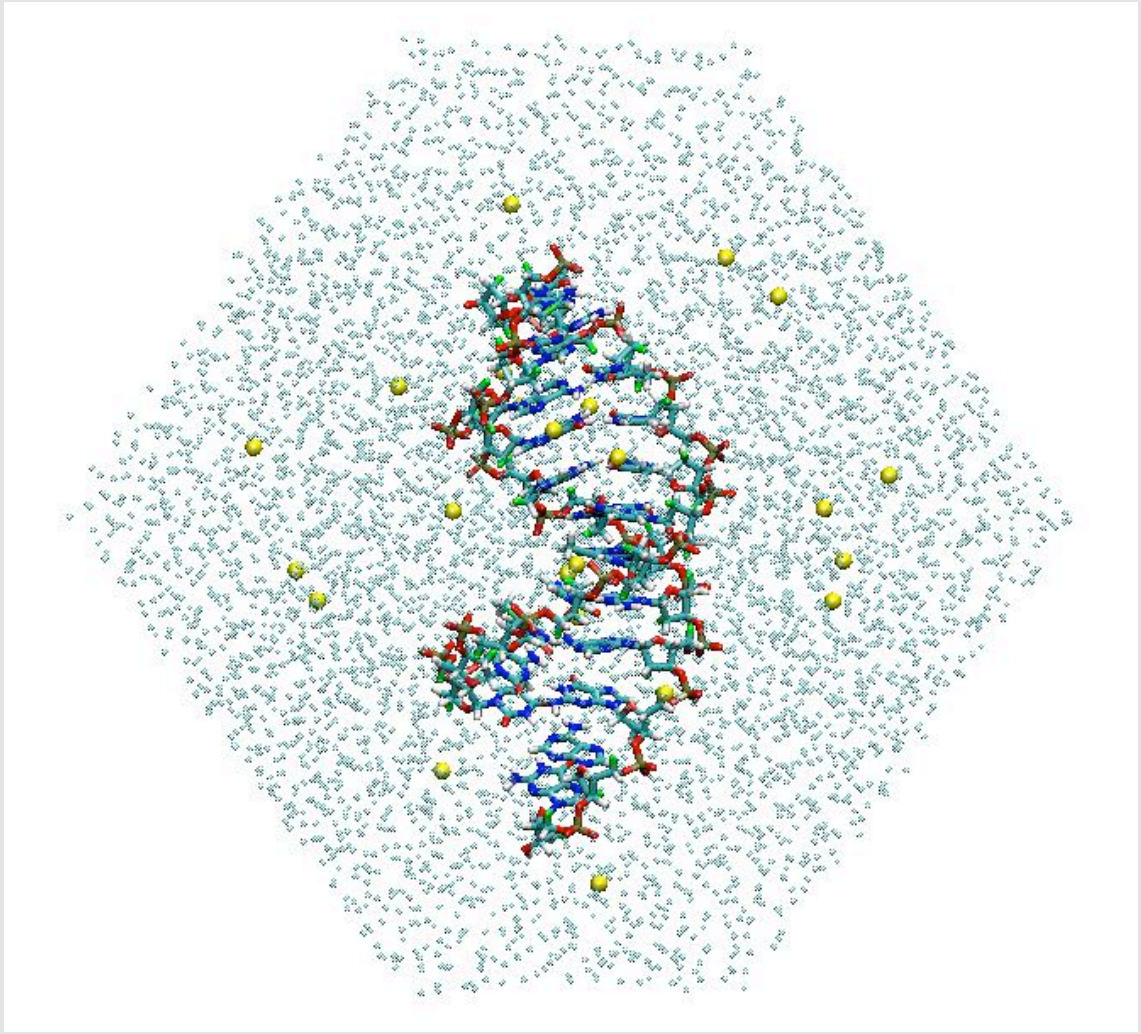
Bonded energy

$$+ \sum_i \sum_{j=i+4} (\varepsilon_{ij} \sigma_{ij}^{12} / r_{ij}^{12} - 2\varepsilon_{ij} \sigma_{ij}^{6} / r_{ij}^{6})$$

Non-bonded energy

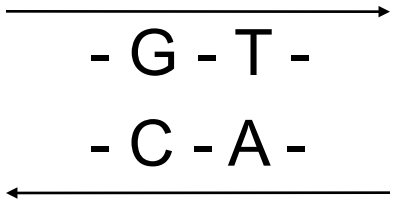
$$+ \; 332 \sum_{i=1,n} \sum_{j=i+1,n} q_i q_j / r_{ij}$$

Periodic boundary conditions



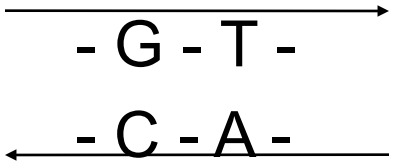
16 possible dinucleotides

GG	G	A	C	T
G	GG	GA	GC	GT
A	AG	AA	AC	AT
C	CG	CA	CC	CT
T	TG	TA	TC	TT



10 unique dinucleotides

GG	G	A	C	T
G	GG	GA	GC	GT
A	AG	AA		AT
C	CG	CA		
T		TA		



136 unique tetramers

AAAA AAAC AAAG AAAT AAGA AAGC AAGG AAGT
AATA AATC AATG AATT ACGA ACGC ACGG ACGT
AGAA AGAC AGAG AGAT AGCA AGCC AGCG AGCT
AGGA AGGC AGGG AGGT AGTA AGTC AGTG AGTT
ATAA ATAC ATAG ATAT ATGA ATGC ATGG ATGT
CAAA CAAC CAAG CAAT CAGA CAGC CAGG CAGT
CATA CATG CCGA CCGG CGAA CGAC CGAG CGAT
CGCA CGCG CGGA CGGC CGGG CGGT CGTA CGTC
CGTG CGTT CTAA CTAG CTGA CTGC CTGG CTGT
GAAA GAAC GAAG GAAT GAGA GAGC GAGG GAGT
GATA GATC GATG GCGA GCGC GCGG GGAA GGAC
GGAG GGAT GGCA GGCC GGCG GGGA GGGC GGGG
GGGT GGTA GGTC GGTG GGTT GTAA GTAC GTAG
GTGA GTGC GTGG GTGT TAAA TAAC TAAG TAAT
TAGA TAGC TAGG TAGT TATA TCGA TGAA TGAC
TGAG TGAT TGCA TGGA TGGC TGGG TGGT TGTA
TGTC TGTG TGTT TTAA TTGA TTGC TTGG TTGT

39 oligomer database

GGGGGGGGGGGGG	AAAAAAAAAAAAA	CGCGCGCGCGCGC
TATATATATATAT	AGAGAGAGAGAGA	TGTGTGTGTGTGT
AGGGAGGGAGGGA	CGGGCGGGCGGGC	TGGGTGGGTGGGT
GAAAGAAAGAAAG	CAAACAAACAAAC	TAAATAAATAAAT
CGGCCGGCCGGCC	AGGAAGGAAGGAA	TGGTTGGTTGGTT
TAATTAATTAATT	CGGACGGACGGAC	AGGCAGGCAGGCA
AGGTAGGTAGGTA	TGGATGGATGGAT	CGGTCGGTCGGTC
TGGCTGGCTGGCT	CAAGCAAGCAAGC	GAACGAACGAACG
TAACTAACTAACT	CAATCAATCAATC	TAAGTAAGTAAGT
GAATGAATGAATG	TGAGTGAGTGAGT	CGAGCGAGCGAGC
TGCGTGCGTGCGT	TAGATAGATAGAT	GACAGACAGACAG
TACATACATACAT	AGCTAGCTAGCTA	TGCATGCATGCAT
CGATCGATCGATC	TGACTGACTGACT	CGTACGTACGTAC

ABC oligomers – construction

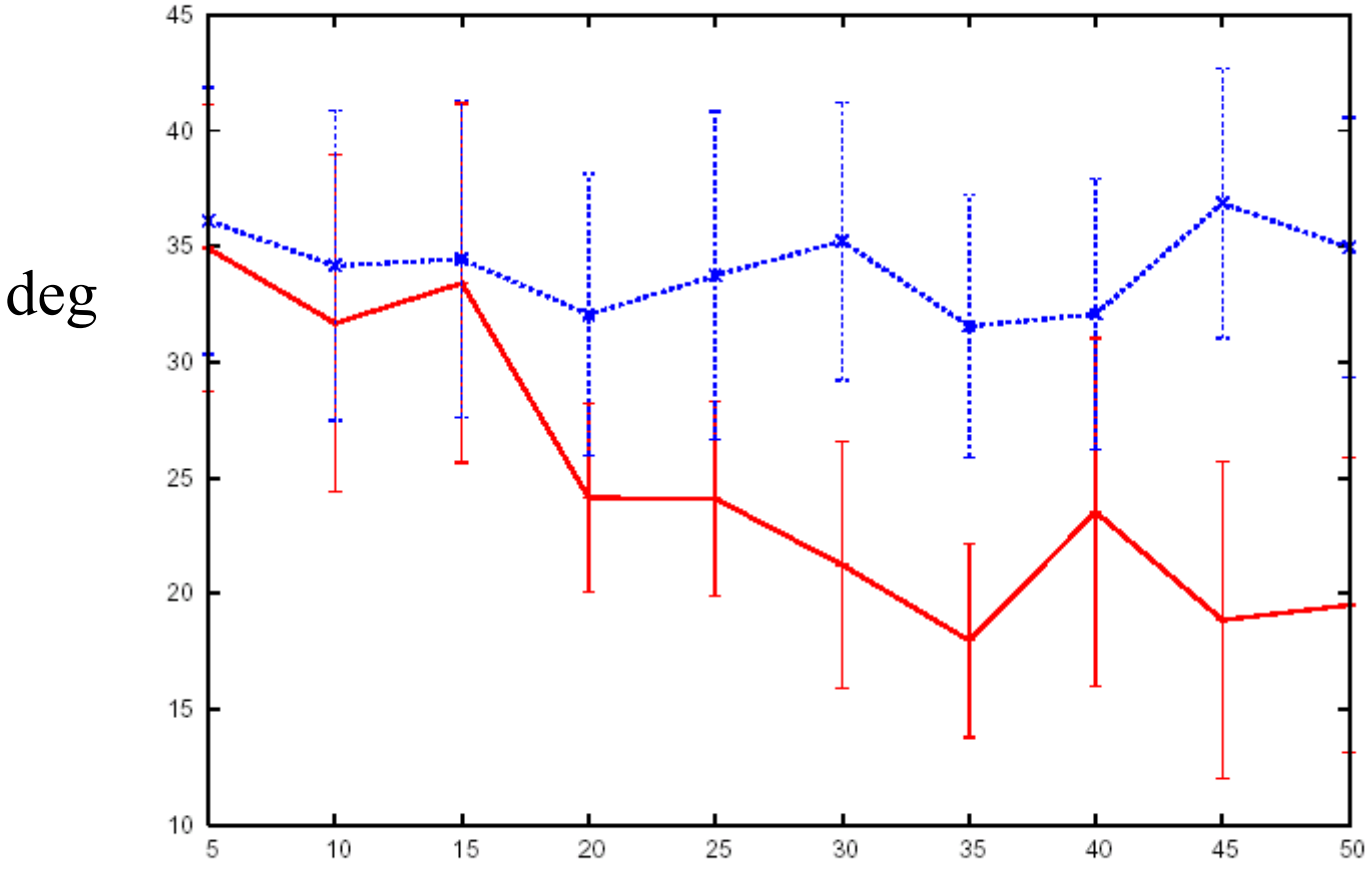
G-D-ABCD-ABCD-ABCD-G

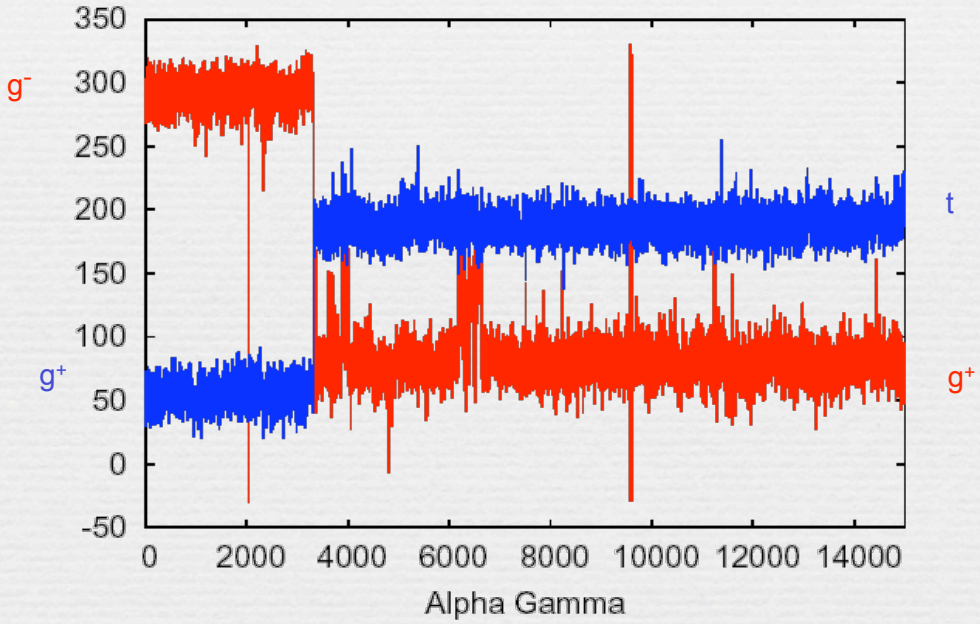
- 15 base pairs
- Central tetranucleotide repeats
- GC terminal base pairs for stability
- No sampling for i<3 or i>13
- Two copies of each tetranucleotide

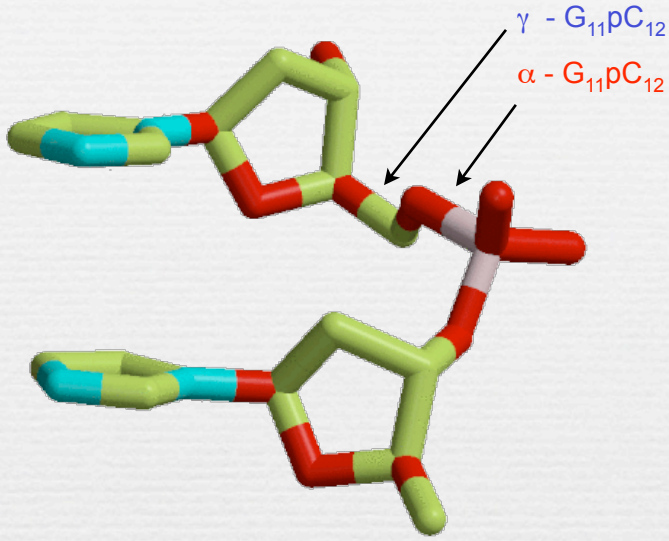
Simulation protocol

- AMBER program
- PARM94 parameters
- Truncated octahedral box (~7600 waters)
- Neutralising K+ counterions
- Particle mesh Ewald electrostatics
- 2 fs timestep (SHAKE on X-H)
- Careful equilibration, NVT→NPT
- Save configuration every 1ps
- 15 ns trajectories (1st cycle)

Mtf binding site – twist_8
Na^+ red, K^+ blue





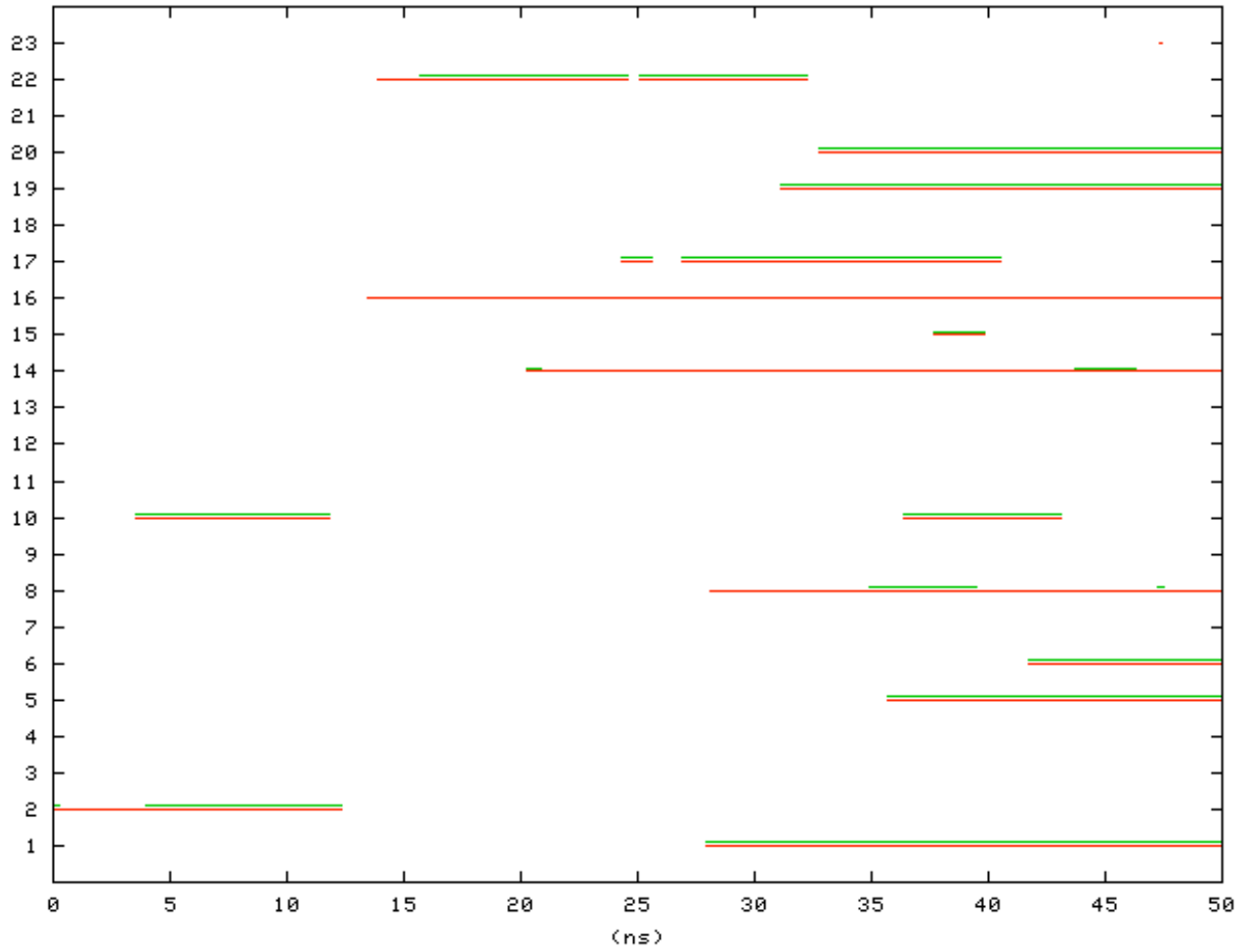


αγ *transition*

α/γ transitions in Na^+ dynamics

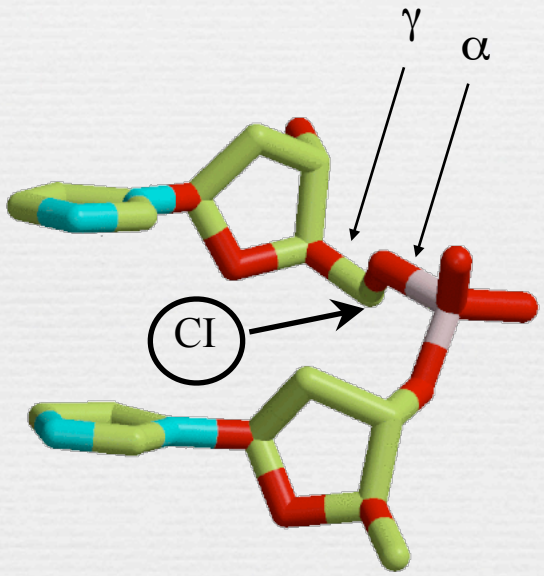
γ

α



Refinement of the amber force field for nucleic acids simulations: PARMBSC0

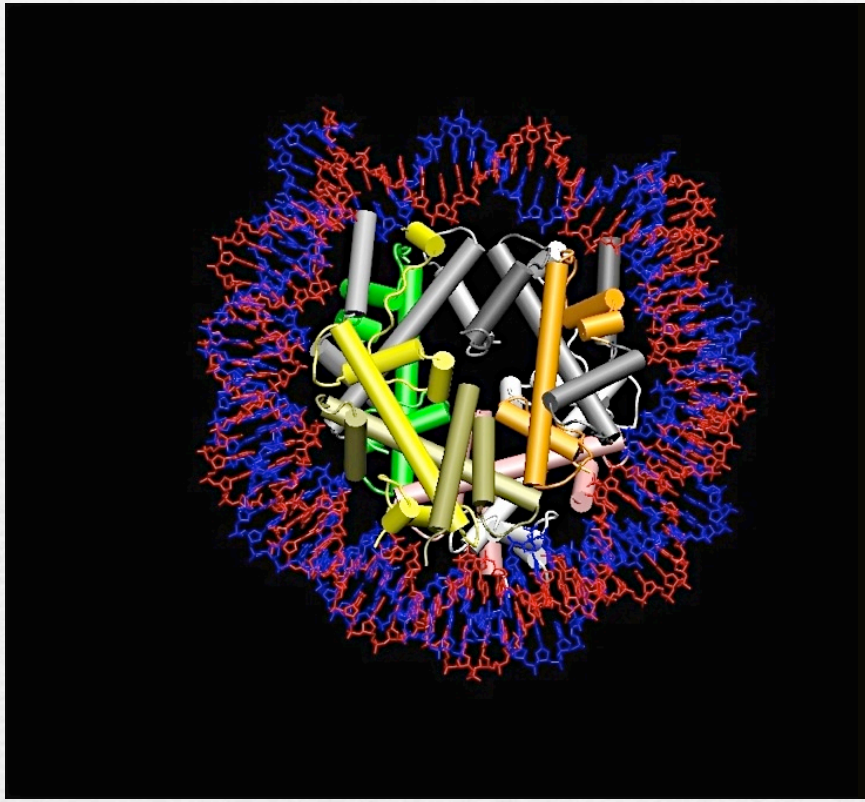
```
OS-P -OS-CI     1        0.185181    31.79508    -1.0   alfa
OS-P -OS-CI     1        1.256531   351.95960    -2.0   alfa
OS-P -OS-CI     1        0.354858   357.24748     3.0   alfa
OH-P -OS-CI     1        0.185181    31.79508    -1.0   alfa
OH-P -OS-CI     1        1.256531   351.95960    -2.0   alfa
OH-P -OS-CI     1        0.354858   357.24748     3.0   alfa
CT-CT-CI-OS     1        1.178040   190.97653    -1.0   gamma
CT-CT-CI-OS     1        0.092102   295.63279    -2.0   gamma
CT-CT-CI-OS     1        0.962830   348.09535     3.0   gamma
CT-CT-CI-OH     1        1.178040   190.97653    -1.0   gamma
CT-CT-CI-OH     1        0.092102   295.63279    -2.0   gamma
CT-CT-CI-OH     1        0.962830   348.09535     3.0   gamma
```

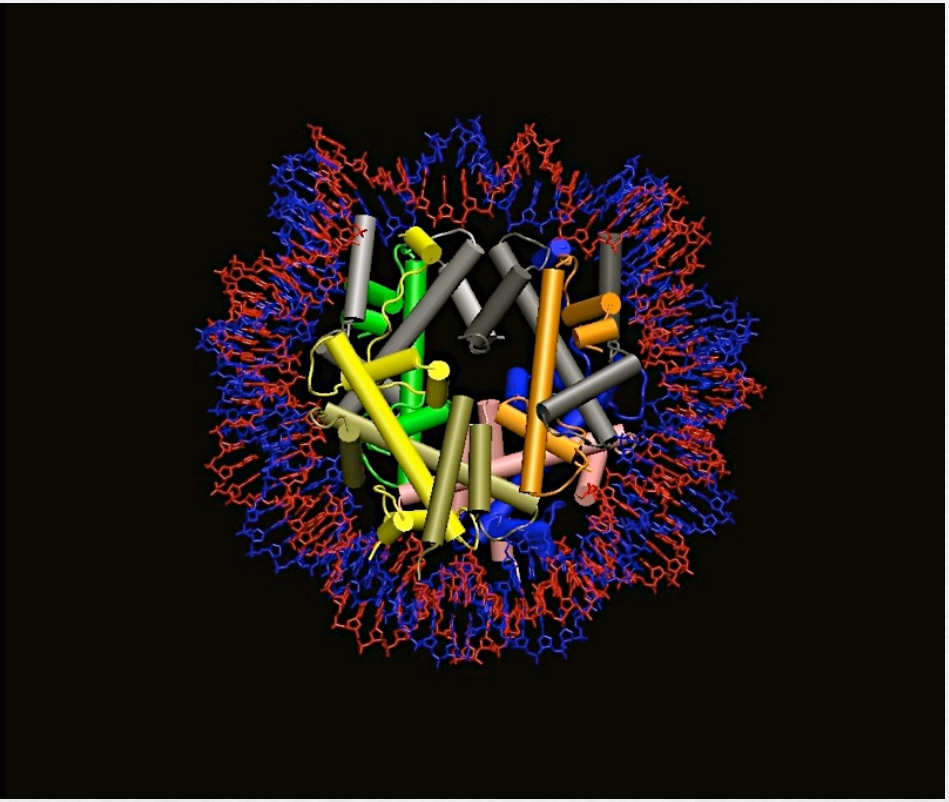


M. Orozco: http://mmb.pcb.ub.es/PARMBSC0/

Pérez et al. Biophys. J. 92 (2007) 3817

AMBER PARM BSC0





X-ray

Minimized

BSC0: Pérez et al. *Biophys. J.* 92 (2007) 3817

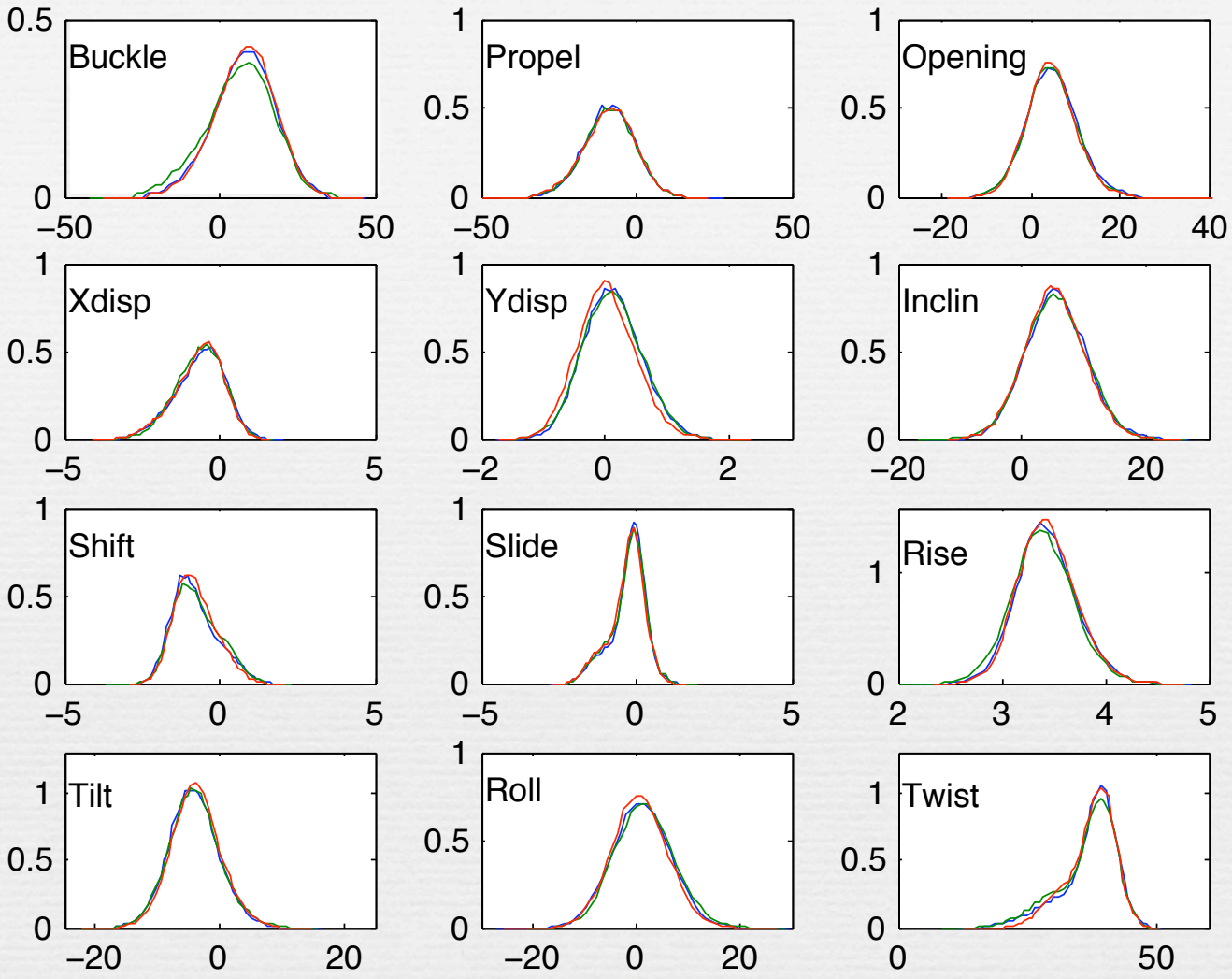
ABC - round 2 - 39 x 18-mers

GCAAAAAAAAAAAAAAGC	GCATCAATCAATCAATGC	GCGGGGGGGGGGGGGGGC
GCACAAACAAACAAACGC	GCCGCGCGCGCGCGCGGC	GCGTGGGTGGGTGGGTGC
GCATAAATAAATAAATGC	GCGACGGACGGACGGAGC	GCACTAACTAACTAACGC
GCAGAGAGAGAGAGAGGC	GCGCCGGCCGGCCGGCGC	GCAGTAAGTAAGTAAGGC
GCCGAGCGAGCGAGCGGC	GCGTCGGTCGGTCGGTGC	GCATTAATTAATTAATGC
GCCTAGCTAGCTAGCTGC	GCTACGTACGTACGTAGC	GCTATATATATATATAGC
GCGAAGGAAGGAAGGAGC	GCTGCGTGCGTGCGTGGC	GCGATCGATCGATCGAGC
GCGCAGGCAGGCAGGCGC	GCAAGAAAGAAAGAAAGC	GCGATGGATGGATGGAGC
GCGTAGGTAGGTAGGTGC	GCACGAACGAACGAACGC	GCGCTGGCTGGCTGGCGC
GCTCAGTCAGTCAGTCGC	GCATGAATGAATGAATGC	GCGTTGGTTGGTTGGTGC
GCTGAGTGAGTGAGTGGC	GCTAGATAGATAGATAGC	GCTATGTATGTATGTAGC
GCGCATGCATGCATGCGC	GCGAGGGAGGGAGGGAGC	GCTCTGTCTGTCTGTCGC
GCAGCAAGCAAGCAAGGC	GCGCGGGCGGGCGGGCGC	GCTGTGTGTGTGTGTGGC

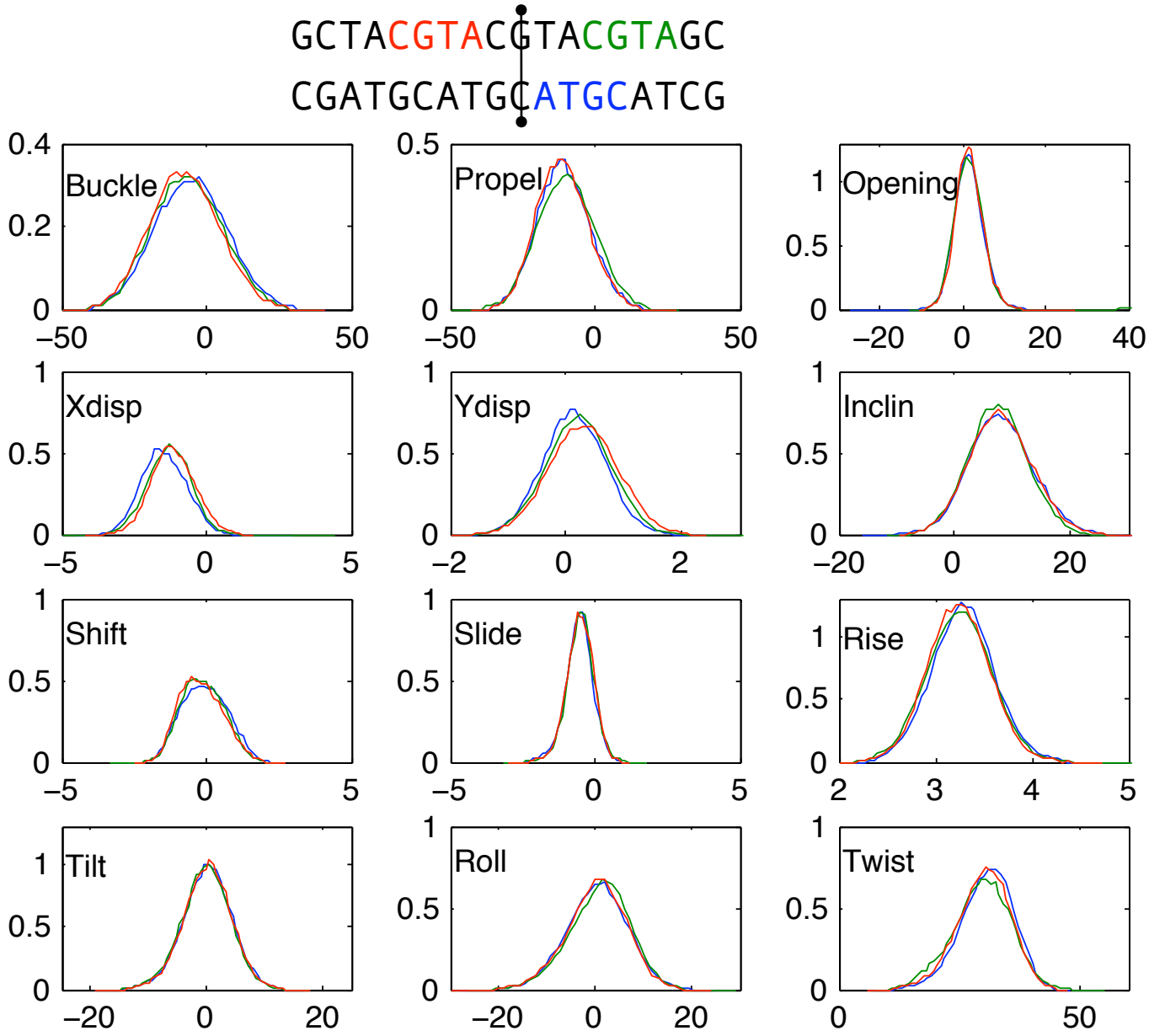
Amber parmbsc0
150 mM KCl
11,500 water molecules
37,000-47,000 atoms

2.75 μs of trajectories
50-100 ns / oligomer
3,000,000 snapshots

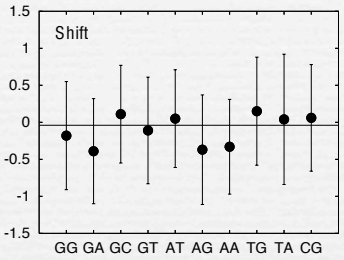
- data on all base pairs with all possible nearest neighbours (32 cases + 4 nn neighbours)
- data on all inter-base pair steps with all possible nearest neighbours (136 cases)

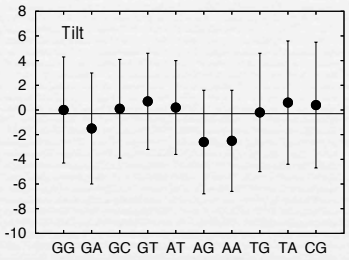


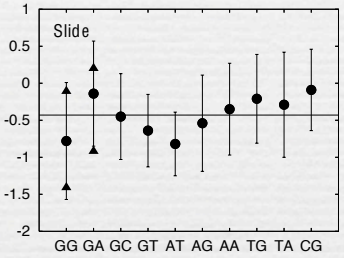
Distributions of helical parameters for A_9 and A_9G_{10} in the AGTC oligomer:
0-50 ns with SPC/E water (—)
50-100 ns with SPC/E water (—)
0-50 ns with TIP4PEW water (—)

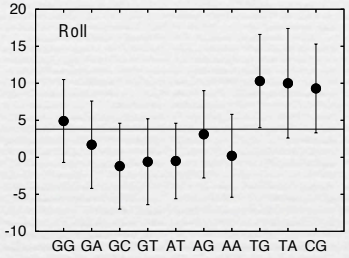


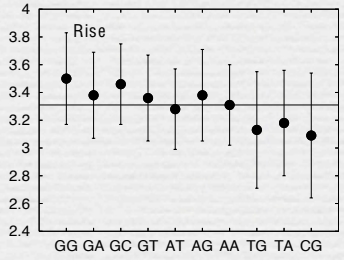
Distributions of helical parameters for the central G base pair and central GT step in three CGTA tetranucleotide fragments in positions 5→8 (—), 11→14 (—) and 13→16 (—)

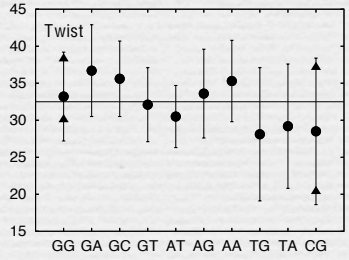


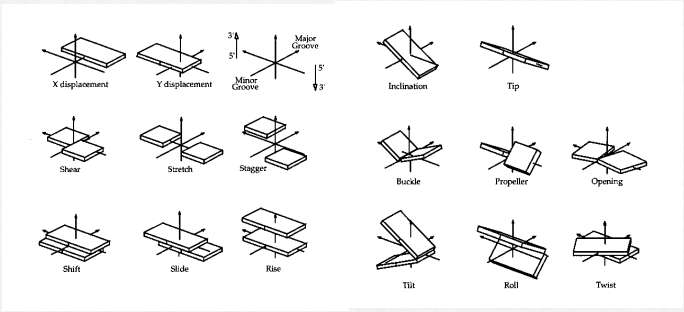


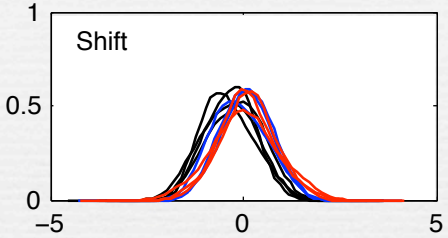


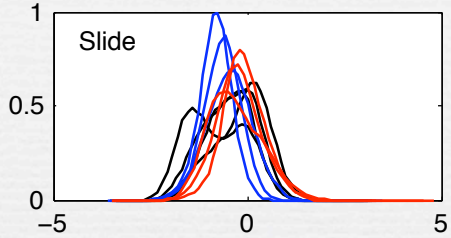


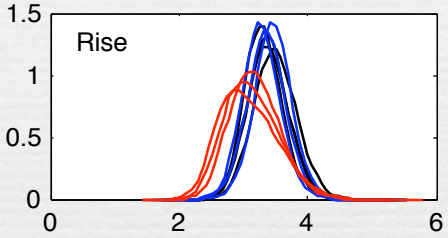


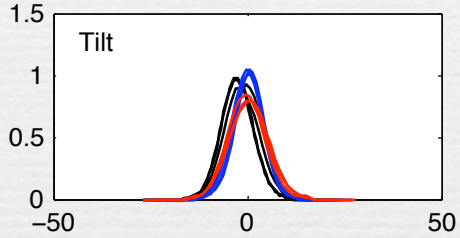


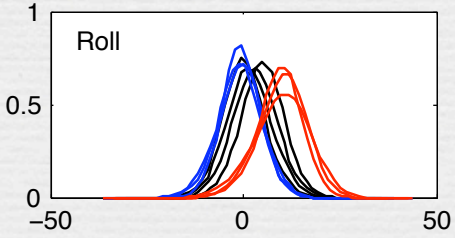


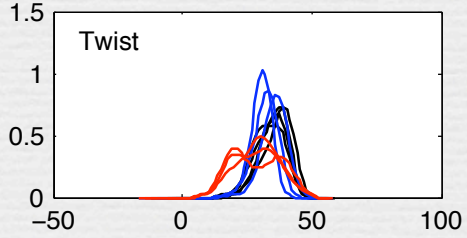






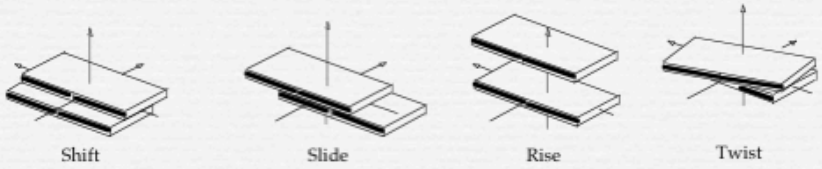


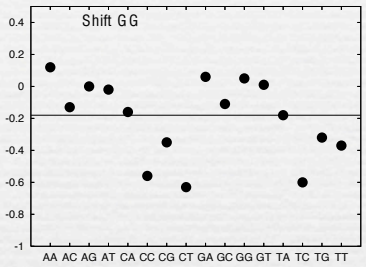


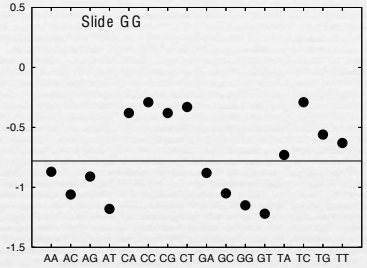


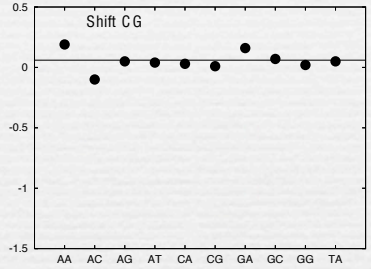
RR
YR
RY

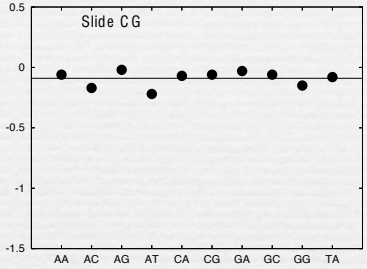
Average base pair step parameters and distributions





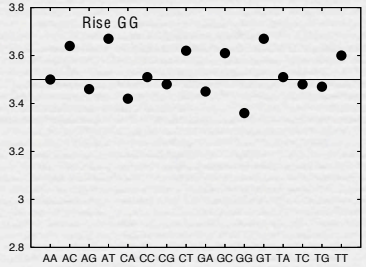


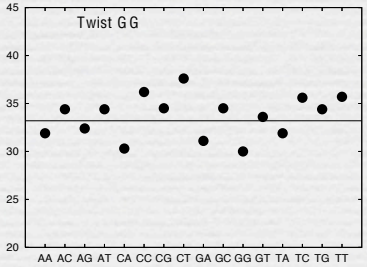


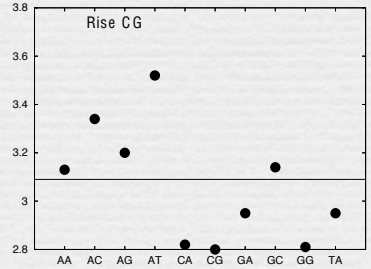


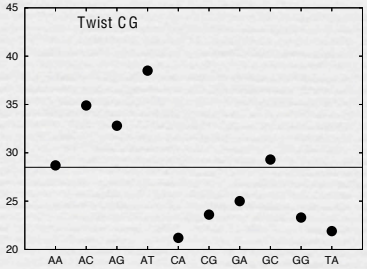
GG

CG

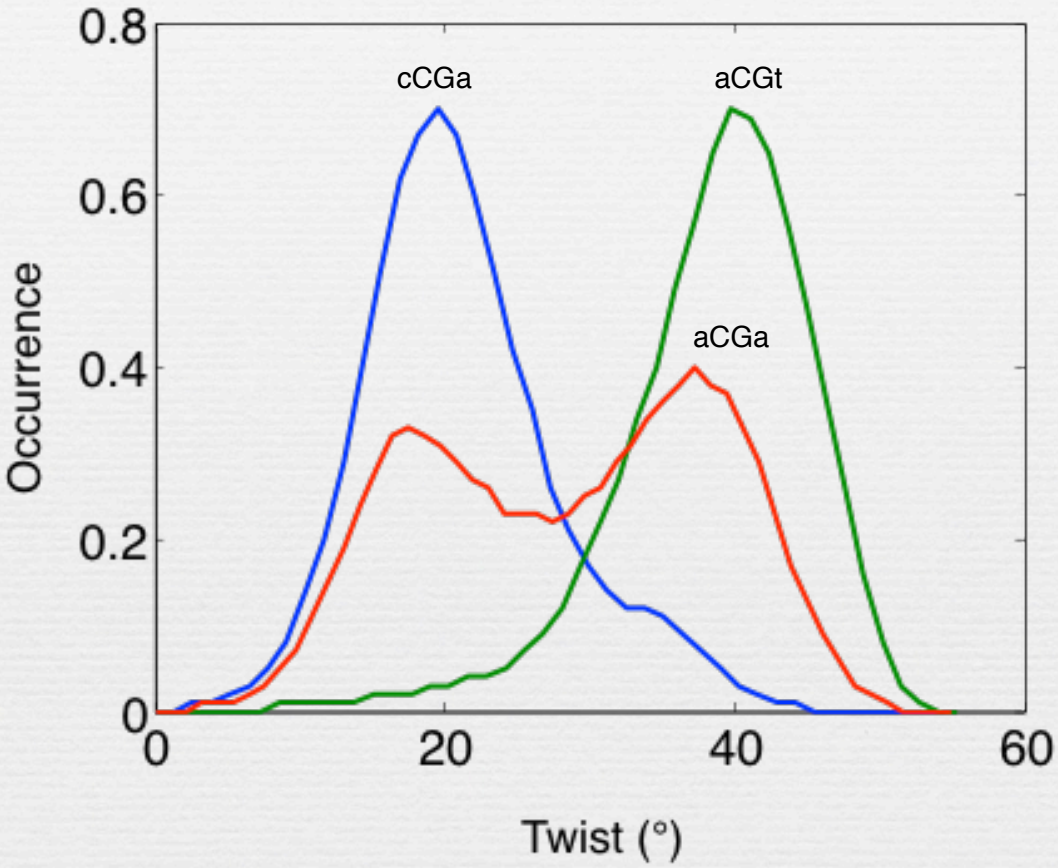




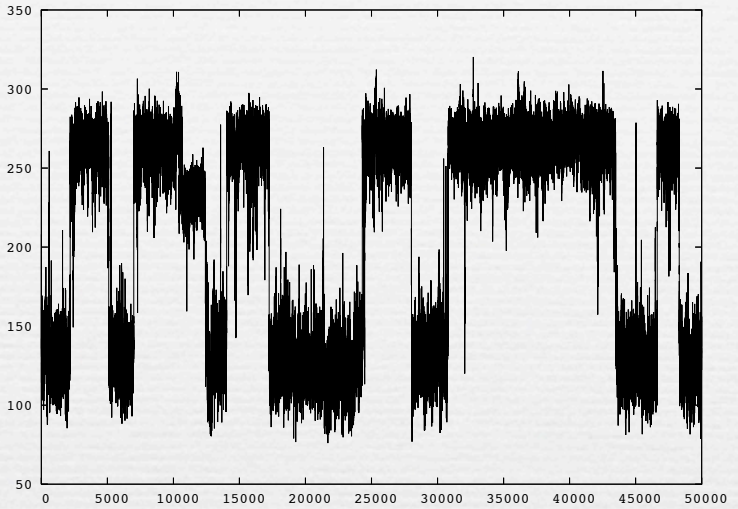




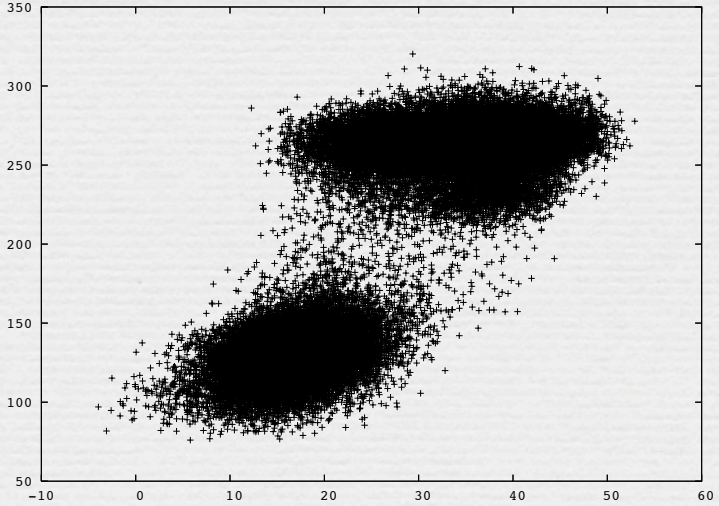
Nearest-neighbour effects on base pair step parameters



ζ (O3'-P)



aCGa twist



Next-nearest neighbour effects on base pair parameters ...

Sequence	Xdisp	Inclination	Propeller	Buckle
A**AAA**A	**-1 Å**	**0°**	-16°	**3°**
C**AAA**C	**-1.6 Å**	**7.5°**	-18°	**8°**
G**CAT**G	**-1.8 Å**	**10°**	-12°	-3°
T**CAT**T	**-1.3 Å**	**4°**	-10°	-3°
G**GGG**G	**-3.1 Å**	11°	**-4°**	-2°
C**GGG**C	**-1.6 Å**	7°	**-14°**	-6°
A**CGT**A	-1.3 Å	**8°**	-11°	**-6°**
T**CGT**T	-0.9 Å	**4°**	-8°	**1°**

And so onwards to hexanucleotides .....

GC-Y-XABCDY-XABCDY-X-GC

GCAAACGTAA**ACGT**AAACGTAAGC
GCCAACGTCA**ACGT**CAACGTCAGC
GCGAACGTGA**ACGT**GAACGTGAGC
GCTAACGTTA**ACGT**TAACGTTAGC
GCACACGTAC**ACGT**ACACGTACGC
GCGCACGTGC**ACGT**GCACGTGCGC
GCAGACGTAG**ACGT**AGACGTAGGC
GCCGACGTCG**ACGT**CGACGTCGGC
GCGGACGTGG**ACGT**GGACGTGGGC
GCATACGTAT**ACGT**ATACGTATGC

GCAAACGAAA**ACGA**AAACGAAAGC
GCCAACGACA**ACGA**CAACGACAGC
GCGAACGAGA**ACGA**GAACGAGAGC
GCTAACGATA**ACGA**TAACGATAGC
GCACACGAAC**ACGA**ACACGAACGC
GCGCACGAGC**ACGA**GCACGAGCGC
GCAGACGAAG**ACGA**AGACGAAGGC
GCCGACGACG**ACGA**CGACGACGGC
GCGGACGAGG**ACGA**GGACGAGGGC
GCATACGAAT**ACGA**ATACGAATGC
GCTTACGATT**ACGA**TTTCGATTGC
GCTGACGATG**ACGA**TGACGATGGC
GCTCACGATC**ACGA**TCACGATCGC
GCGTACGAGT**ACGA**GTACGAGTGC
GCCTACGACT**ACGA**CTACGACTGC
GCCCACGACC**ACGA**CCACGACCGC

..... 2080 unique hexanucleotides (≈ 350 oligomers)

Acknowledgements

Peter Varnai

Benjamin Bouvier

Jean-Pierre Dognon

Emanuela Emanuele

Dimitra Markovitsi

Philippe Millié

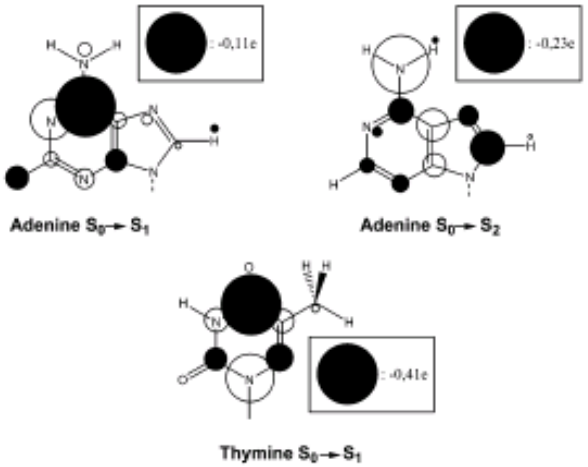
Delphine Onidas

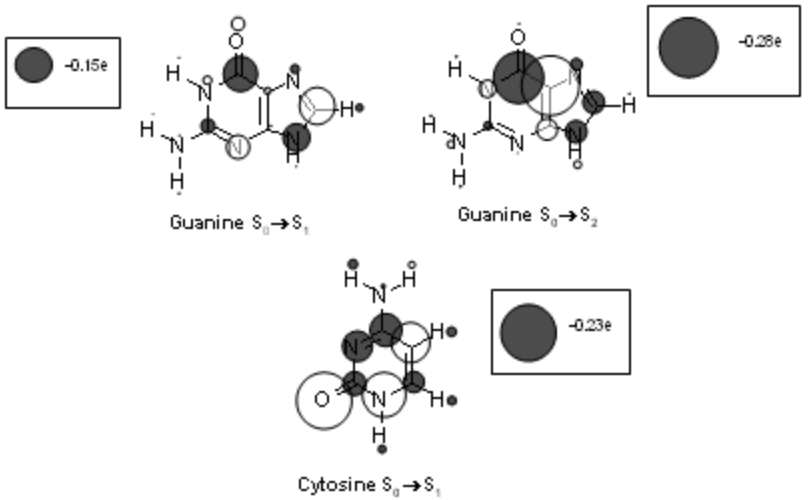
Laboratoire Francis Perrin

CEA - CNRS

ABC Consortium

Charges


-0,11e
-0,23e
-0,41e
Adenine $S_0 \rightarrow S_1$
Adenine $S_0 \rightarrow S_2$
Thymine $S_0 \rightarrow S_1$


-0.15e
-0.28e
-0.23e
Guanine $S_0 \rightarrow S_1$
Guanine $S_0 \rightarrow S_2$
Cytosine $S_0 \rightarrow S_1$