Zjišťování struktury DNA pomocí NMR a molekulových simulací

Jiří Kozelka
kozelka.jiri@gmail.com

4.10.2013

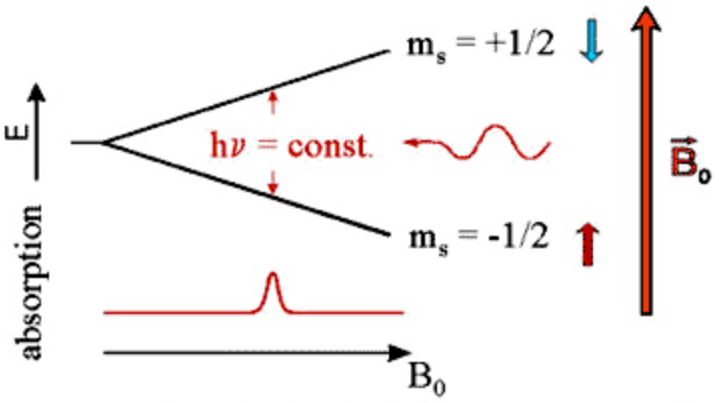
Použitý materiál:
http://chemistry.osu.edu/~foster.281/biochem766/download/PDF_files/766_nmrlecture.ppt.pdf
http://www.chem.ufl.edu/~nmr/lon/2D-1.pdf
http://www.bioc.aecom.yu.edu/labs/girvlab/nmr/course/COURSE_2012/Lecture_NucleicAcids_2012.pdf

Doporučené speciální kurzy:
C8950 NMR - Strukturní analýza, prof. RNDr. Radek Marek, Ph.D.

Doporučená literatura:

Wüthrich, K, 1984 "NMR of Proteins and Nucleic Acids" J. Wiley & Sons, NY.

NMR Spektroskopie



Využívá jaderného spinu nuklidů (^{1}H, ^{13}C, ^{31}P, ^{15}N)

Ve stacionárním magnetickém poli B_0 preferují spiny energeticky výhodnou polohu paralelní s B_0

Dodáním potřebné energie ΔE přejde spin do excitovaného stavu (poloha antiparalelní s B_0):

$$\Delta E = h\gamma B_0/2\pi$$

h: Planckova konstanta, : magnetogyrický faktor (charakteristický pro daný nuklid)

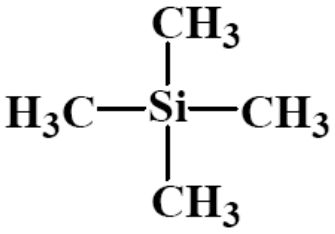
ΔE dodáme v podobě elektromagnetického záření o frekvenci ν:

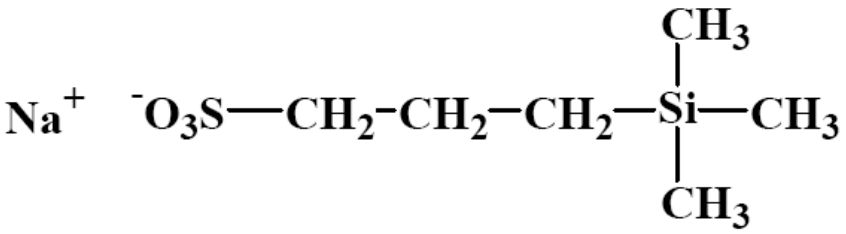
$$h\nu = h\gamma B_0/2\pi$$

$$\nu = \gamma B_0/2\pi$$

V molekule je magnetické pole B_0 modulováno magnetickými momenty elektronů: $B_0 \rightarrow B_L$ (lokální)=$(1-\sigma)\ B_0$
σ: „stínící konstanta“, charakteristizuje okolí daného jádra

frekvence, při které nastává rezonance: $\nu_L = \gamma B_L/2\pi$
ν_L je charakteristická při daném B_0 pro jádro v molekule
Jako standard používáme pro ^{1}H NMR např.:





Tetrametylsilan (TMS)
rozpustný v org. rozpouštědlech

4,4-Dimetyl-4-silapentansulfonan sodny (DSS)
rozpustný ve vodě

Rezonanční frekvenci udáváme zpravidla „chemickým posunem“ δ:

$\delta = 10^6 . (\nu_L - \nu_{ref})/\nu_{ref}$ [ppm] $= 10^6 . (B_L - B_{ref})/B_{ref}$ ($\nu_{ref} = \nu_{TMS}$ nebo ν_{DSS})

$B_L = (1-\sigma_L)\ B_0$ $B_{ref} = (1-\sigma_{ref})\ B_0$

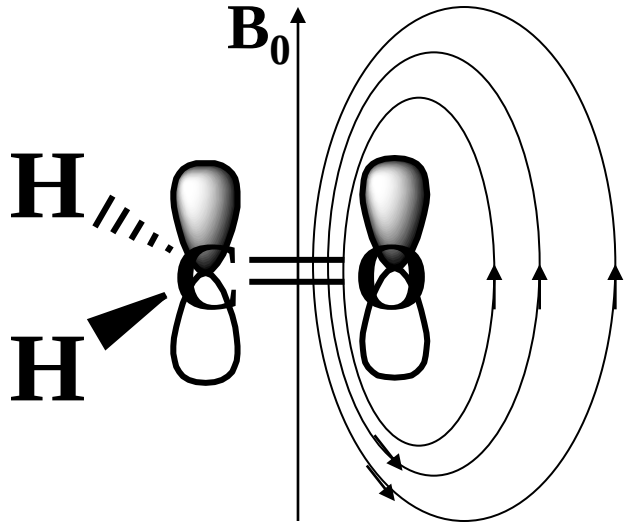
$$\delta = \frac{(1-\sigma_L)B_0 - (1-\sigma_{ref})B_0}{(1-\sigma_{ref})B_0} \cdot 10^6 = \frac{\sigma_{ref} - \sigma_L}{1-\sigma_{ref}} 10^6 \approx (\sigma_{ref} - \sigma_L) \cdot 10^6 [ppm]$$

Faktory ovlivňující δ:

1. **Indukční efekt elektronegativních atomů**

	δ(1H) [ppm]
CH_3J	2.16
CH_3Br	2.68
CH_3Cl	3.05
CH_3F	4.26

2. **Magnetická anisotropie dvojných vazeb**



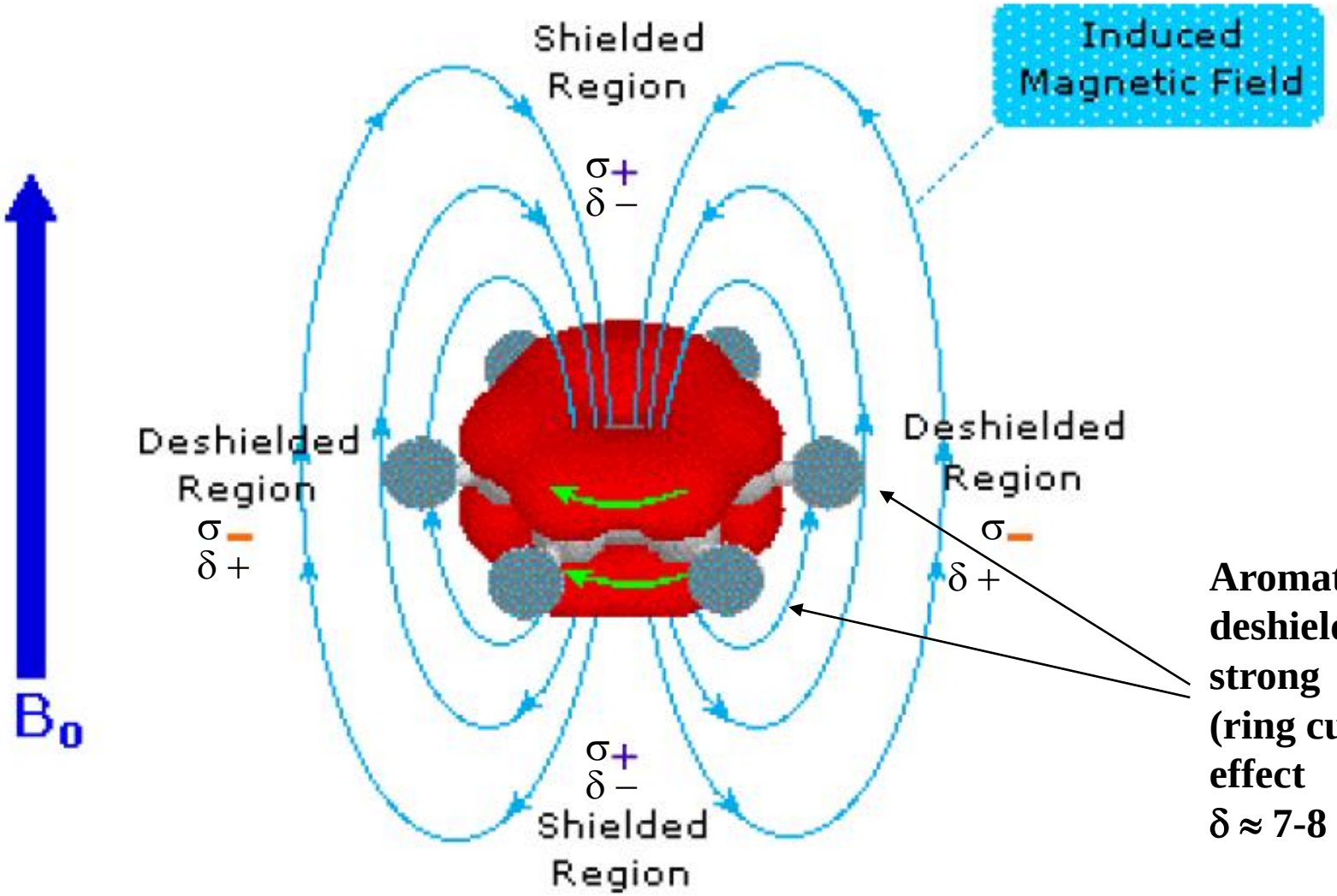
Aldehydic protons deshielded by a combination of inductive and anisotropy effects δ ≈ 10 ppm

In the molecular plane: Induced field of π-electrons parallel to B_0: deshielding

Below and above the molecular plane: Induced field of π-electrons antiparallel to B_0: shielding

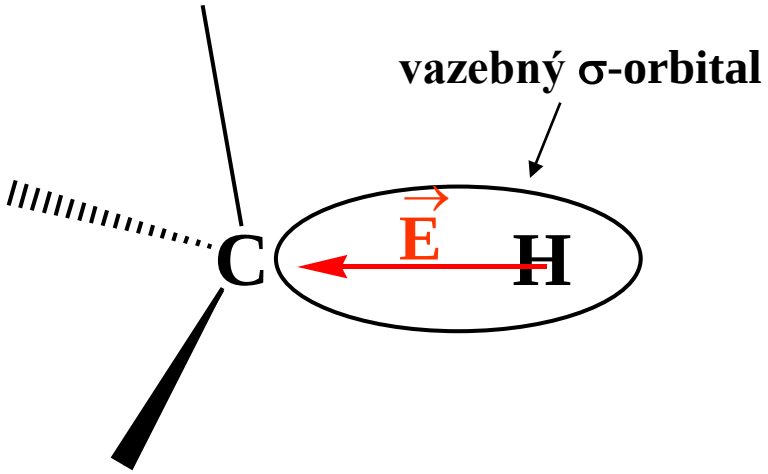
Faktory ovlivňující δ (pokračování):

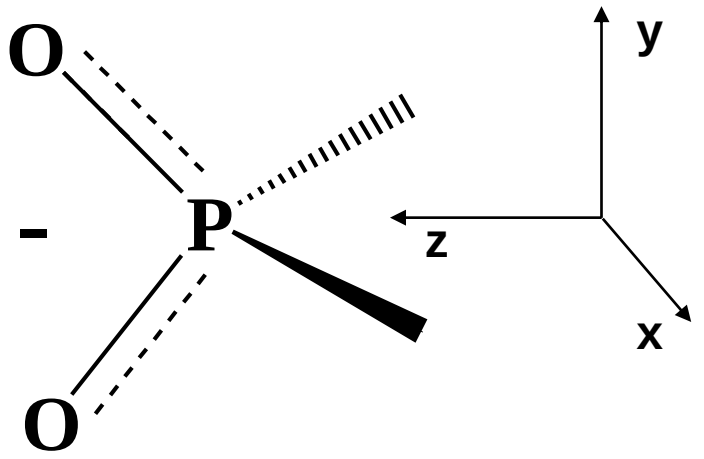
3. Kruhové proudy aromatických jader



Aromatic protons deshielded by a strong anisotropy (ring current) effect
δ ≈ 7-8 ppm

4. Polarizační efekt nabitých atomů





$$\delta_{\text{electric}} = A_Z E_Z + B E^2$$

stínění nebo odstínění podle polohy nabité skupiny (ve znázorněném případě odstínění)

vždy odstínění (deshieldng, Δδ>0)

1D ^{1}H NMR spectrum of DNA

The assignment problem:
Which peak corresponds to which atom in the molecule?

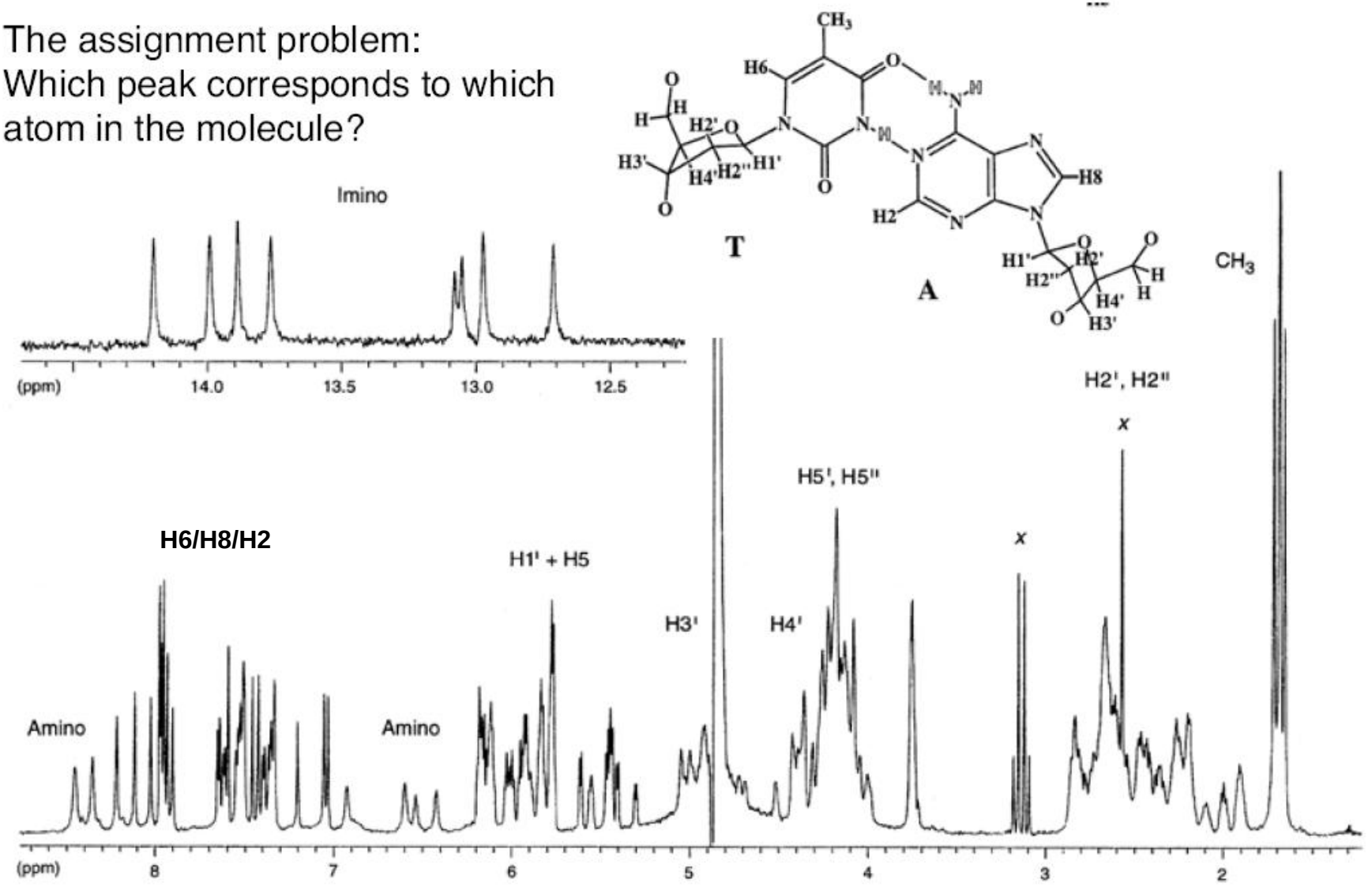
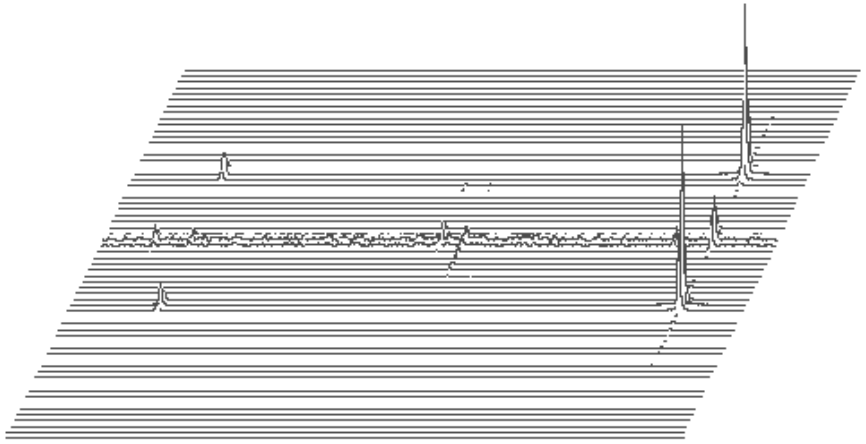


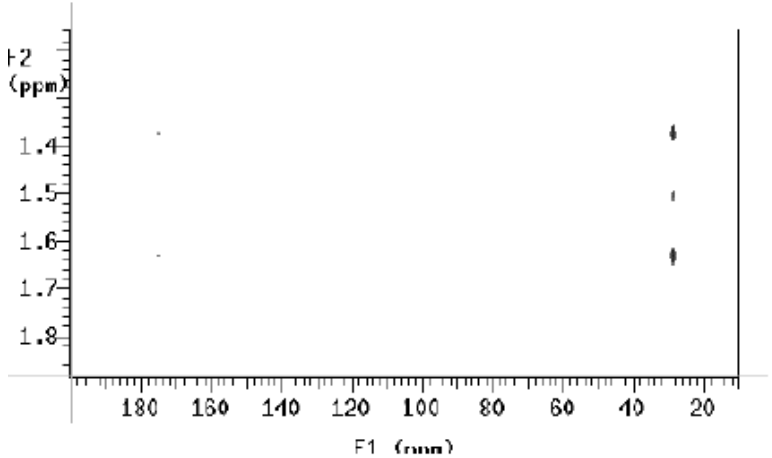
Figure 5-5
The one-dimensional 1D spectrum of the DNA oligomer d(CGCAAAAGGC)·d(GCCTTTTGCG) is shown (20°C in H_2O solution, solvent suppressed by presaturation), with the regions of the spectrum labeled. The imino and amino resonances of the terminal base pairs are missing due to exchange with solvent. Peaks marked with an *x* are small molecule contaminants.

2D NMR Spectroscopy
allows magnetic interactions (correlations) between spins to be detected and depicted in a 2-dimensional diagram.

2 D spectra are 3D objects



Cross-peaks reveal a correlation between the frequencies on the two frequency axes.



Correlations:

- Through bonds - based on scalar coupling.
- Through space - based on dipolar relaxation (nOe's). NOE = Nuclear Overhauser Effect

Sequential Resonance Assignments in ^{1}H NMR Spectra of Oligonucleotides by Two-Dimensional NMR Spectroscopy†

R. M. Scheek, R. Boelens, N. Russo,‡ J. H. van Boom,§ and R. Kaptein*

265 citations by 2.10.2013

2D NOESY spectrum of d(TGAGCGG)-d(CCGCTCA)

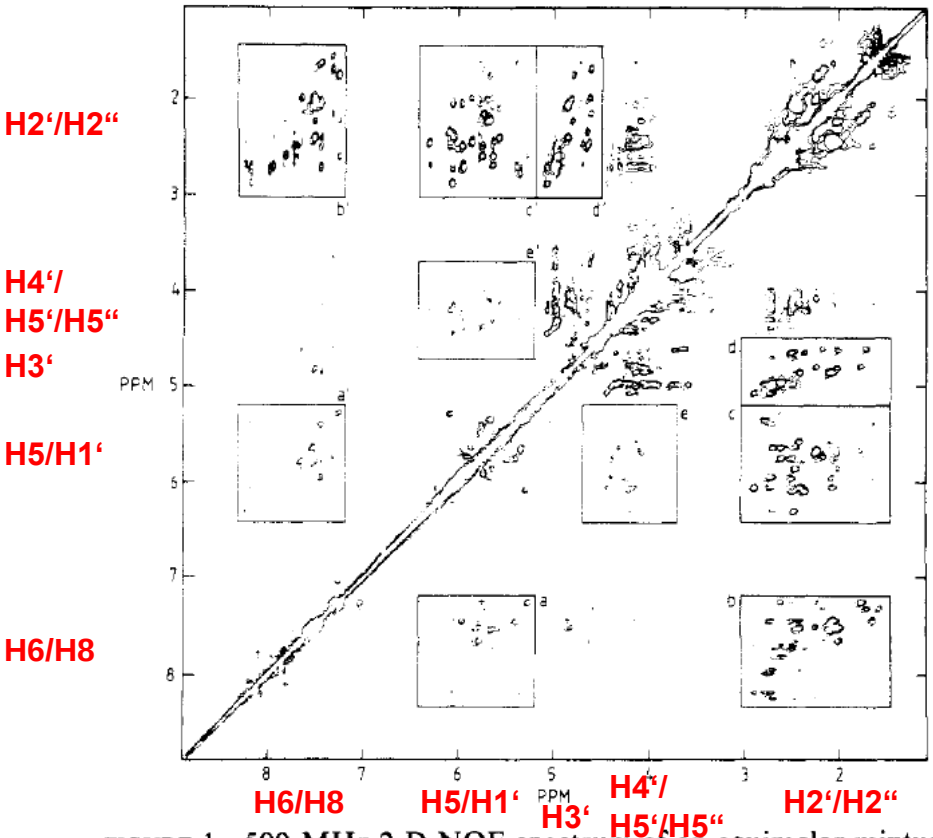
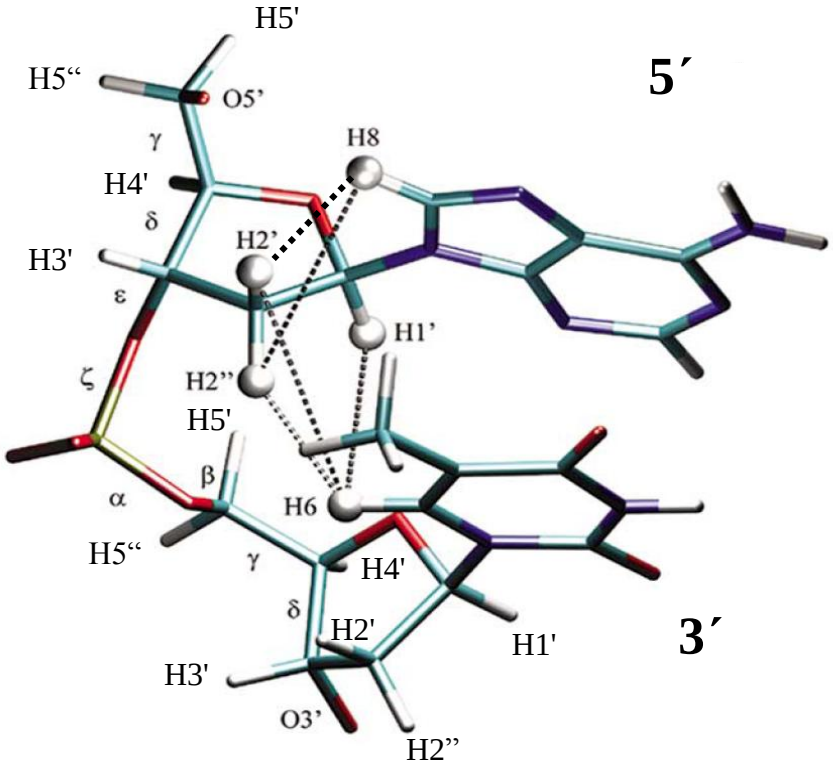


FIGURE 1: 500-MHz 2-D NOE spectrum of an equimolar mixture of d(TGAGCGG) and d(CCGCTCA). The spectrum is a sum of three 2-D NOE spectra with mixing times τ_m = 80, 120, and 160 ms that were recorded to obtain NOE buildup curves as shown in Figure 9.



Canonical 5‘-GT-3‘ fragment of B-DNA (serves as guideline for distances, this is not a fragment of the studied DNA duplex)

Sequential Resonance Assignments in ^{1}H NMR Spectra of Oligonucleotides by Two-Dimensional NMR Spectroscopy†

R. M. Scheek, R. Boelens, N. Russo,‡ J. H. van Boom,§ and R. Kaptein*

2D COSY spectrum of d(TGAGCGG)-d(CCGCTCA)

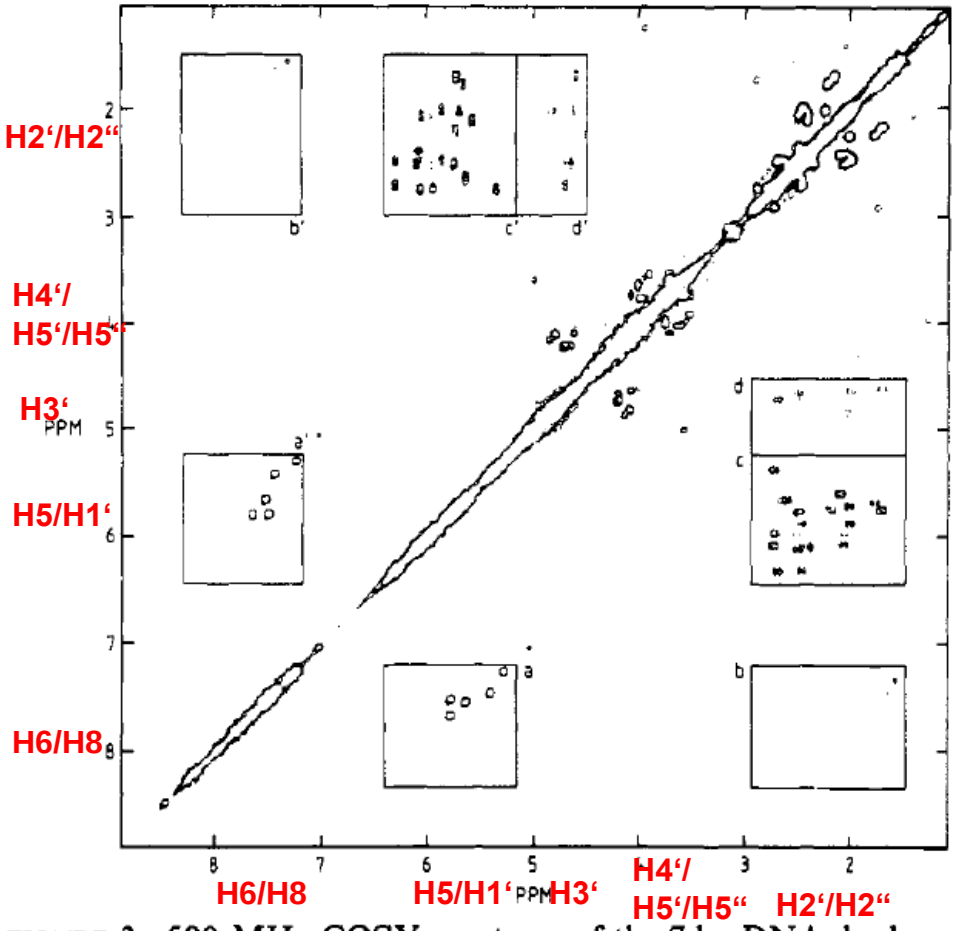
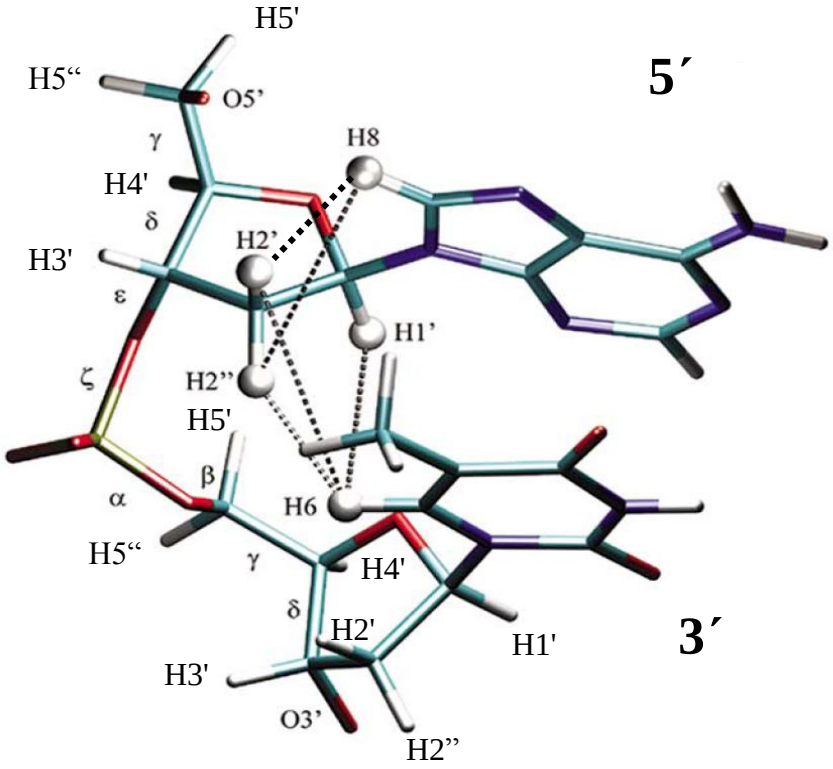


FIGURE 3: 500-MHz COSY spectrum of the 7 bp DNA duplex. A total of 64 scans was recorded for each FID. The boxed regions correspond to those in Figure 1.



Canonical 5'-GT-3' fragment of B-DNA (serves as guideline for distances, this is not a fragment of the studied DNA duplex)

Chemical shifts of the non-exchangeable protons in a double-stranded DNA heptamer

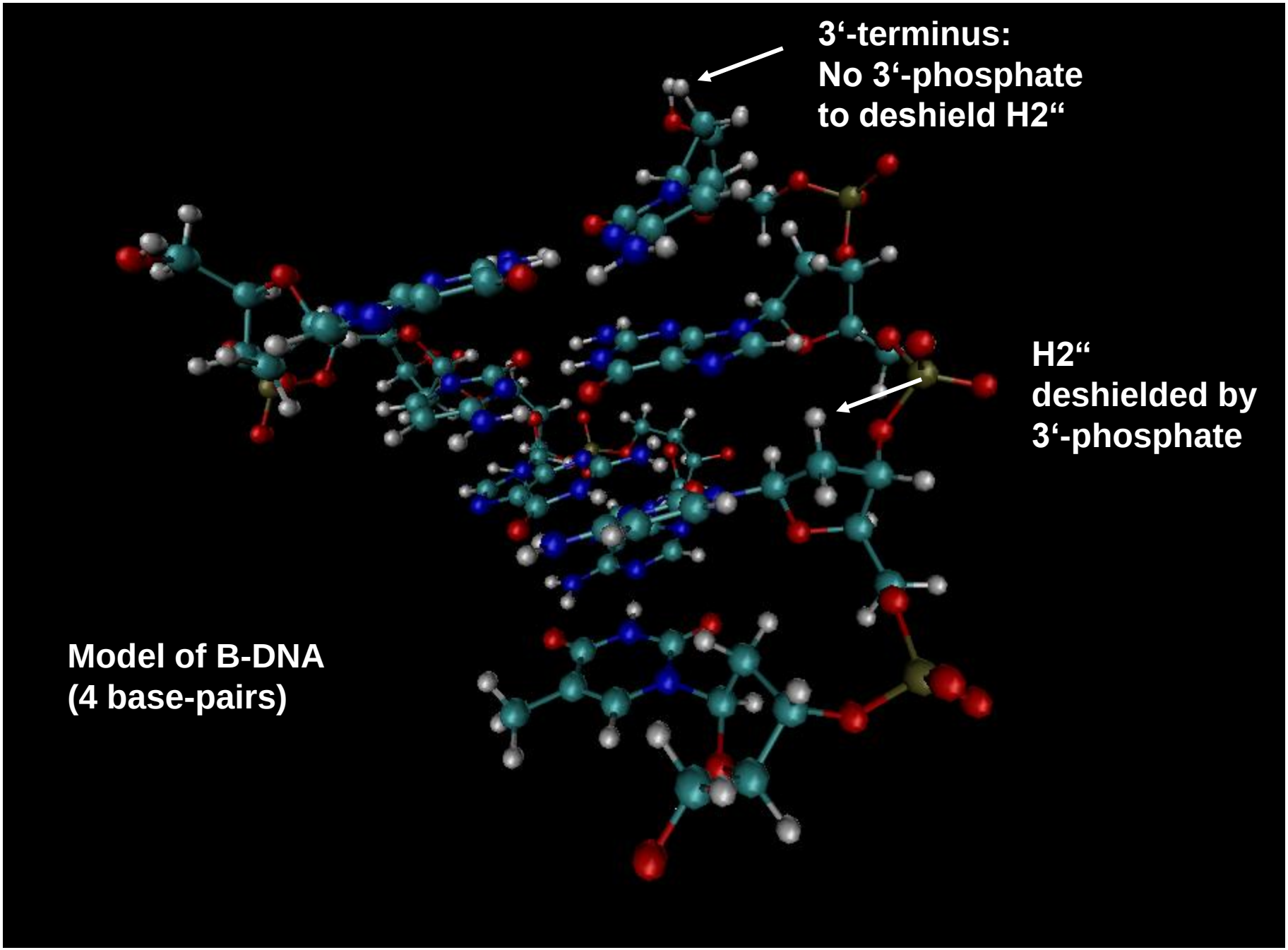
5'-TGAGCCG-3'
3'-ACTCGGC-5'

Table II: Chemical Shifts of Assigned ^{1}H Resonances (in ppm, Relative to DSS) in the Duplex of d(TGAGCGG) and d(CCGCTCA)

	H8 or H6	CH5 or TCH_3	H1′	H2′	H2″	H3′	H4′	H5′, H5″[a]
T1	7.31	1.58	5.74	1.73	2.18	4.62	4.02	NA
G2	7.99		5.36	2.74	2.80	4.98	4.31	4.04, 3.95
A3	8.18		6.06	2.74	2.90	5.08	4.46	4.22, 4.14
G4	7.69		5.75	2.50	2.62	4.98	4.39	4.21, 4.20
C5	7.25	5.28	5.67	1.79	2.25	4.79	4.12	NA
G6	7.82		5.65	2.62	2.71	4.95	4.33	4.07, 3.98
G7	7.73		6.10	2.50	2.37	4.65	4.22	NA
C1	7.65	5.78	5.86	2.02	2.47	4.63	4.10	NA
C2	7.52	5.64	5.58	2.12	2.43	4.86	4.15	NA
G3	7.94		5.95	2.71	2.76	5.02	4.41	4.14, 4.06
C4	7.44	5.40	5.96	2.08	2.52	4.74	4.25	NA
T5	7.43	1.65	6.05	2.08	2.42	4.86	4.16	NA
C6	7.50	5.78	5.70	2.03	2.25	4.82	4.09	NA
A7	8.24		6.30	2.70	2.48	4.71	4.21	NA

[a] NA, not assigned.

Scheek, R. M.; Boelens, R.; Russo, N.; van Boom, J. H.; Kaptein, R. Biochemistry 1984, 23, 1371-1376.


3‘-terminus:
No 3‘-phosphate
to deshield H2“
H2“
deshielded by
3‘-phosphate
Model of B-DNA
(4 base-pairs)

Chemical shifts of the non-exchangeable protons in a double-stranded DNA heptamer

5'-TGAGCCG-3'
3'-ACTCGGC-5'

Table II: Chemical Shifts of Assigned ^{1}H Resonances (in ppm, Relative to DSS) in the Duplex of d(TGAGCGG) and d(CCGCTCA)

	H8 or H6	CH5 or TCH_3	H1′	H2′	H2″	H3′	H4′	H5′, H5″[a]
T1	7.31	1.58	5.74	1.73	2.18	4.62	4.02	NA
G2	7.99		5.36	2.74	2.80	4.98	4.31	4.04, 3.95
A3	8.18		6.06	2.74	2.90	5.08	4.46	4.22, 4.14
G4	7.69		5.75	2.50	2.62	4.98	4.39	4.21, 4.20
C5	7.25	5.28	5.67	1.79	2.25	4.79	4.12	NA
G6	7.82		5.65	2.62	2.71	4.95	4.33	4.07, 3.98
G7	7.73		6.10	2.50	2.37	4.65	4.22	NA
C1	7.65	5.78	5.86	2.02	2.47	4.63	4.10	NA
C2	7.52	5.64	5.58	2.12	2.43	4.86	4.15	NA
G3	7.94		5.95	2.71	2.76	5.02	4.41	4.14, 4.06
C4	7.44	5.40	5.96	2.08	2.52	4.74	4.25	NA
T5	7.43	1.65	6.05	2.08	2.42	4.86	4.16	NA
C6	7.50	5.78	5.70	2.03	2.25	4.82	4.09	NA
A7	8.24		6.30	2.70	2.48	4.71	4.21	NA

[a] NA, not assigned.

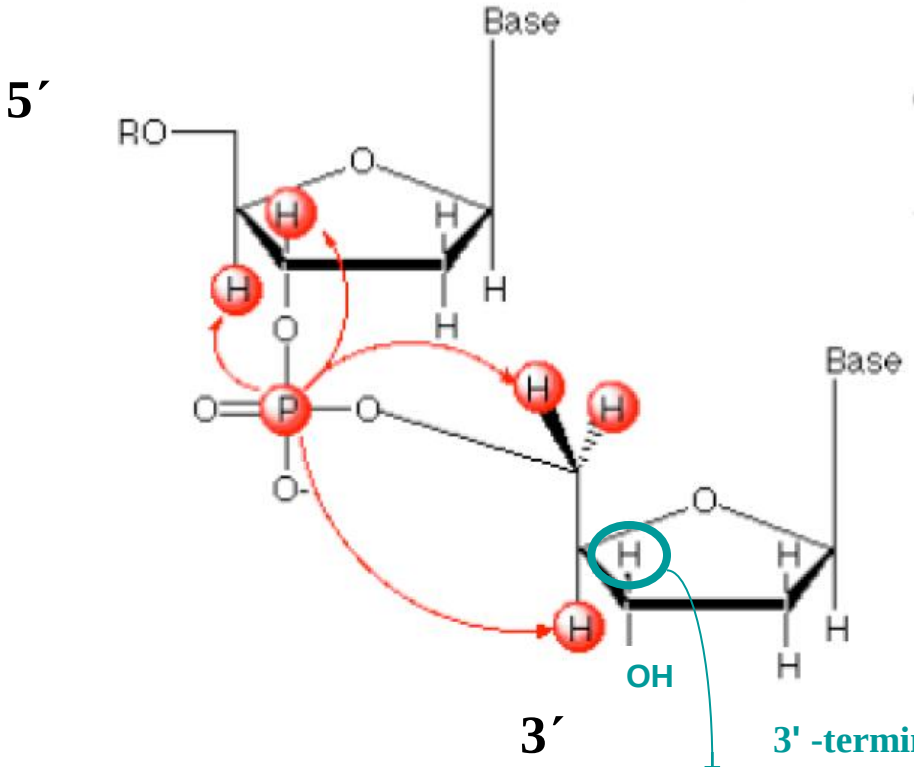
3‘-terminal nucleotide: no deshielding of H2“ by the 3‘-phosphate

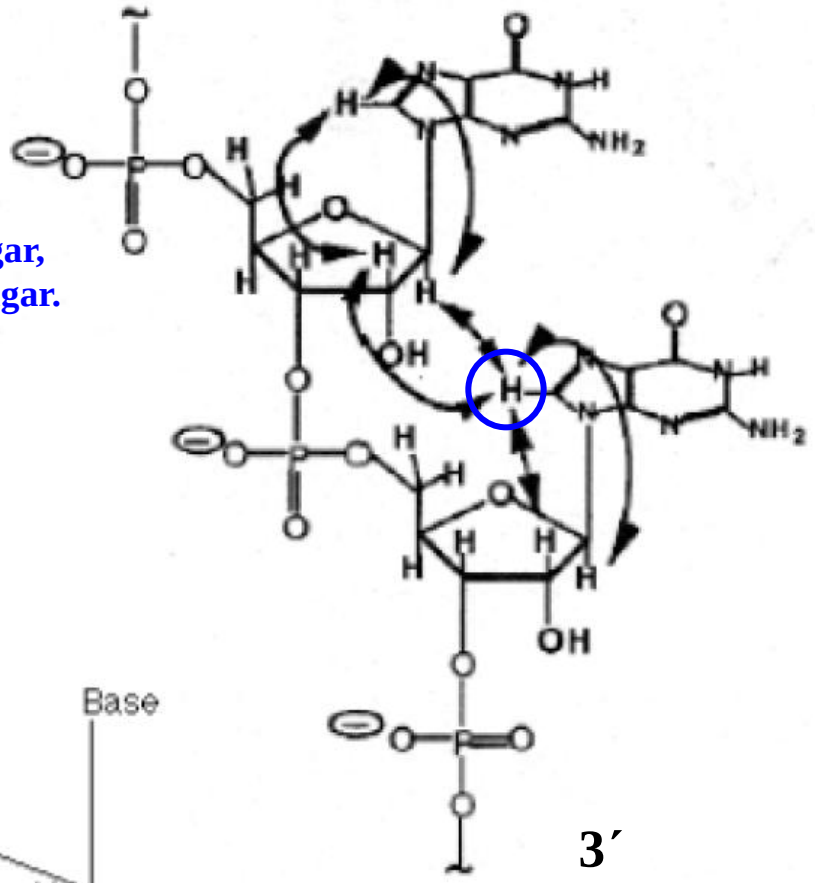
Scheek, R. M.; Boelens, R.; Russo, N.; van Boom, J. H.; Kaptein, R. Biochemistry 1984, 23, 1371-1376.

Assignments of non-exchangeable protons:

Sequential Assignments

The base proton has correlations with its own sugar, as well as with the 5'-sugar, but not with the 3'-sugar.



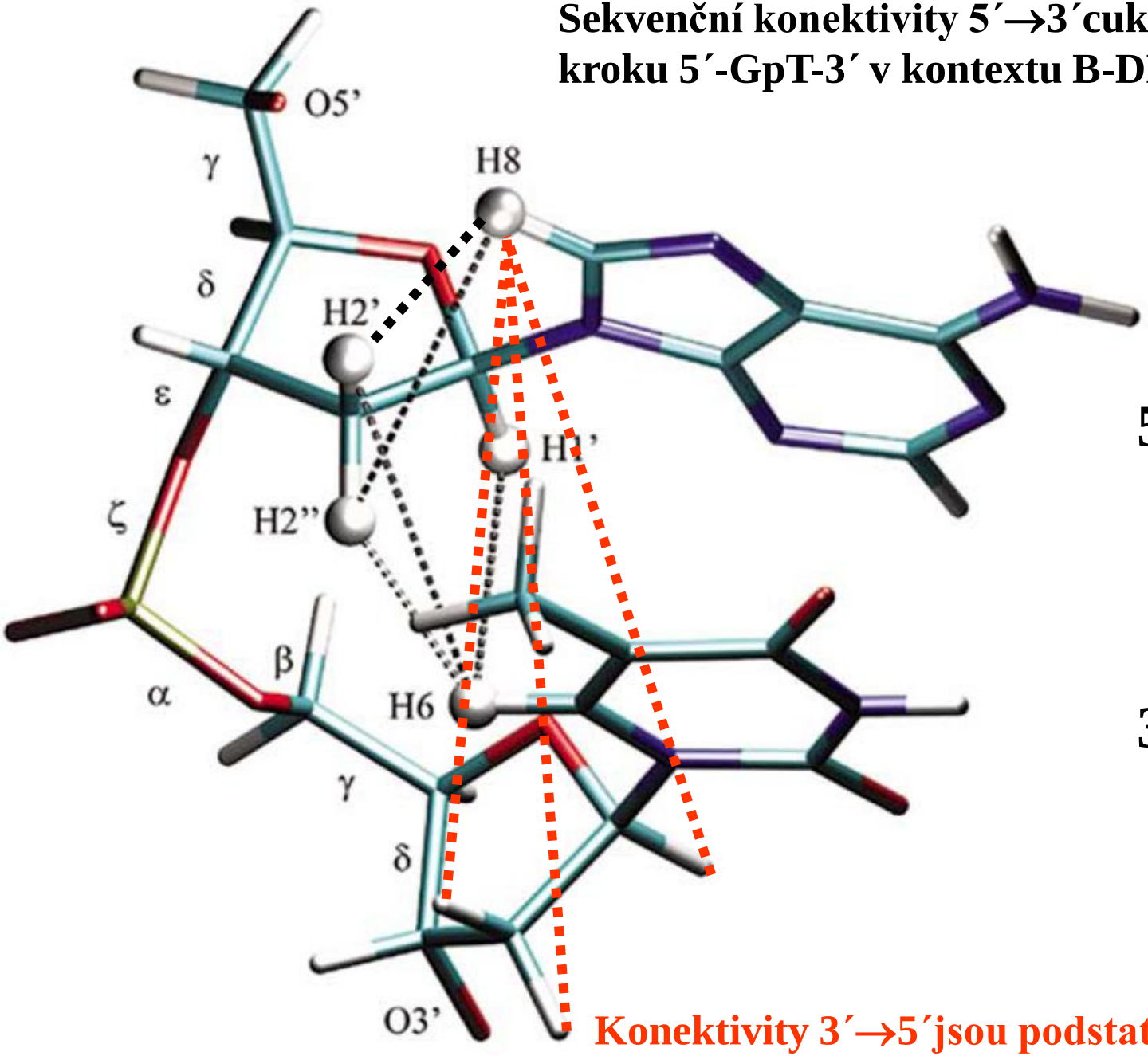


via ^{1}H-^{1}H NOEs

via ^{1}H-^{31}P scalar connections

3' -terminus: H3' has no P to couple with
⇒ identification of the 3' -nucleotide

Sekvenční konektivity 5´→3´cukr-báze kroku 5´-GpT-3´ v kontextu B-DNA



5´-G

3´-T

Konektivity 3´→5´jsou podstatně delší!

Sequential assignment via ^1H-^1H NOEs:

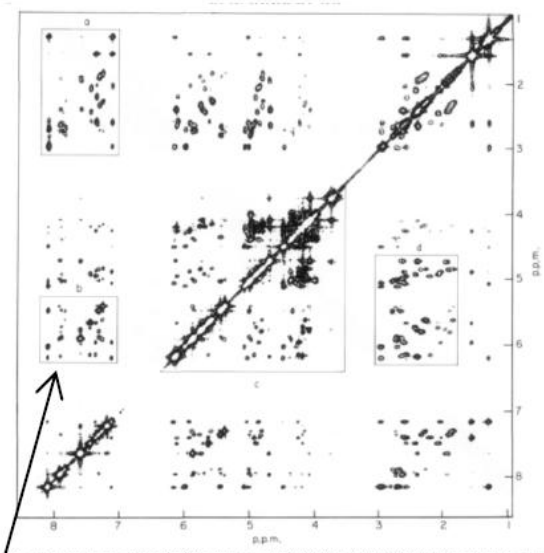
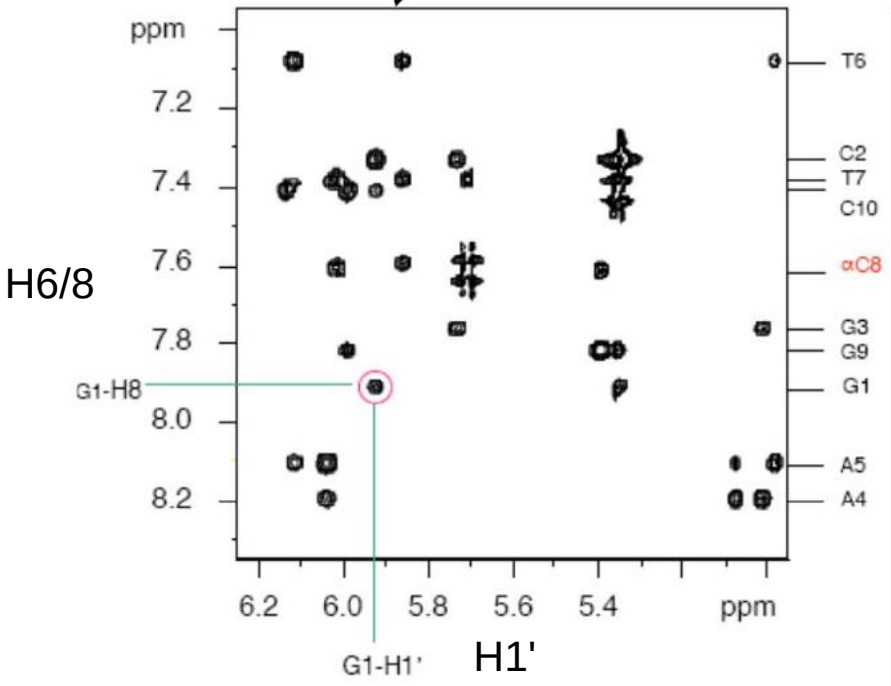
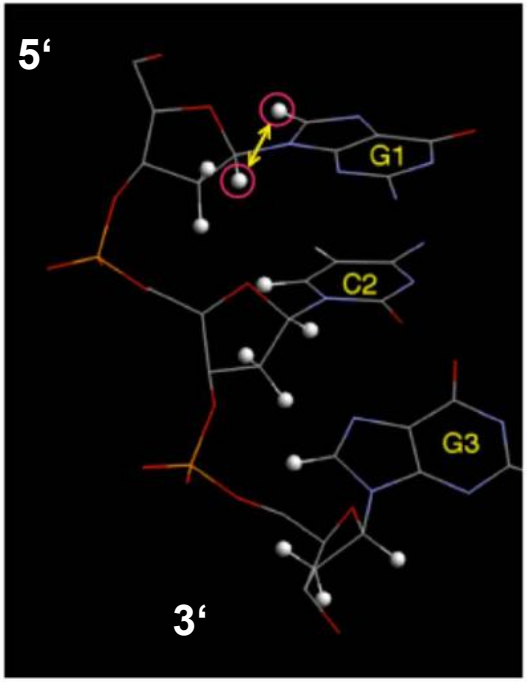
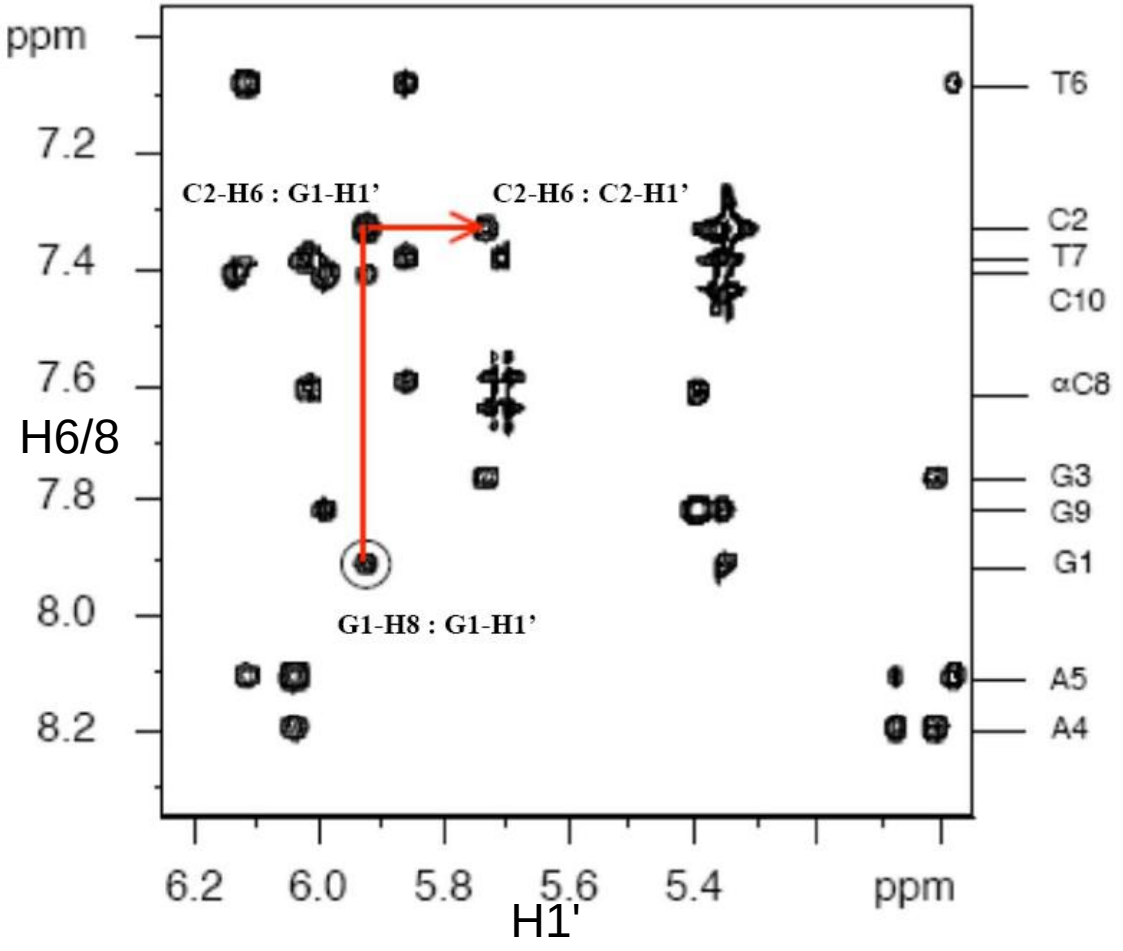


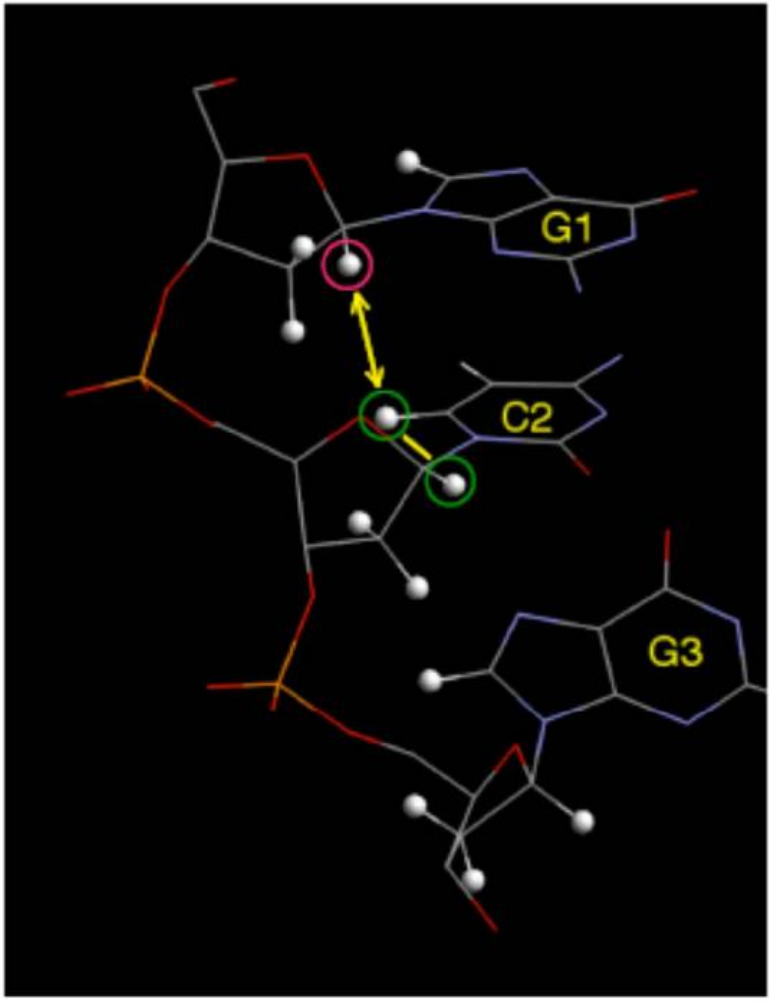
Fig. 3. Contour plot of the NOESY spectrum of d(CG-C-G-A-A-T-T-C-G-C-G). The boxed regions b, c and d correspond to expanded plots given in Figs 5, 6 and 7, respectively.

$d(GCGAATT\alpha CGC)_2$

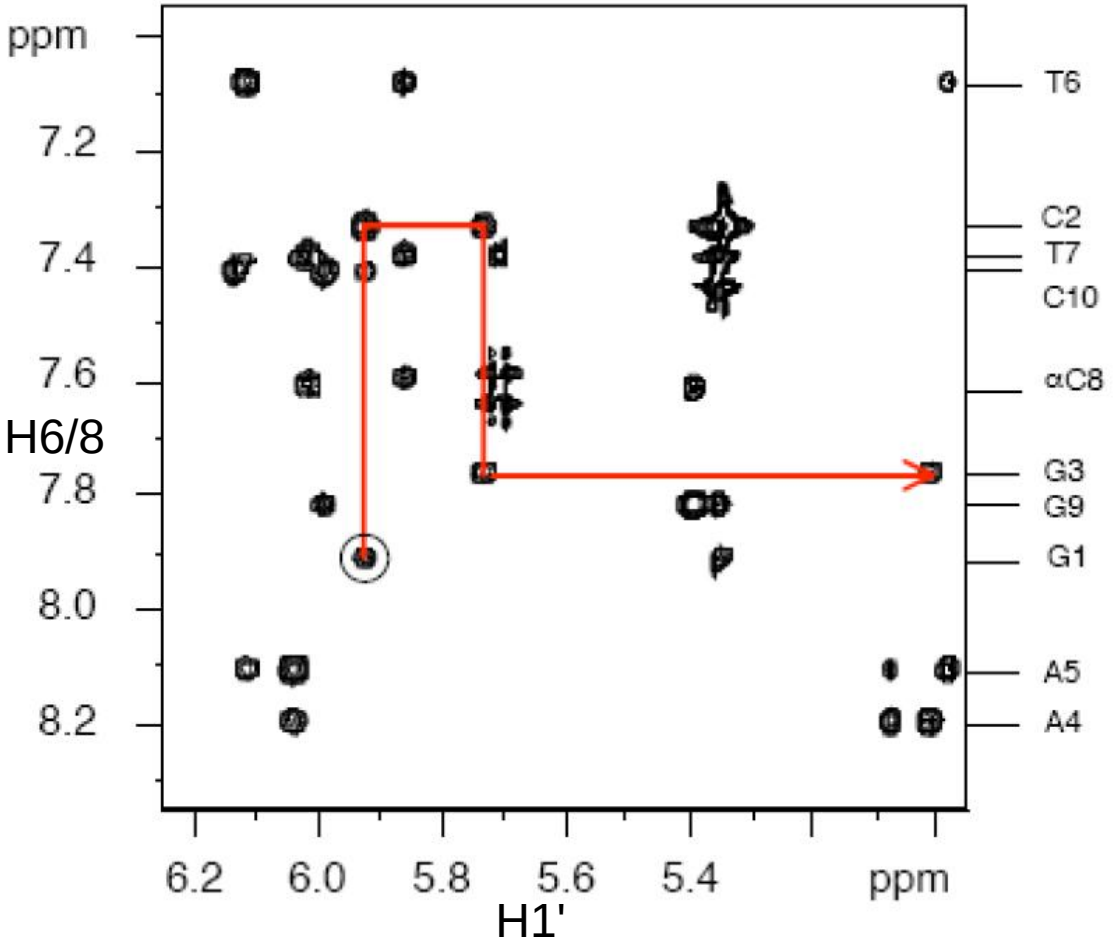


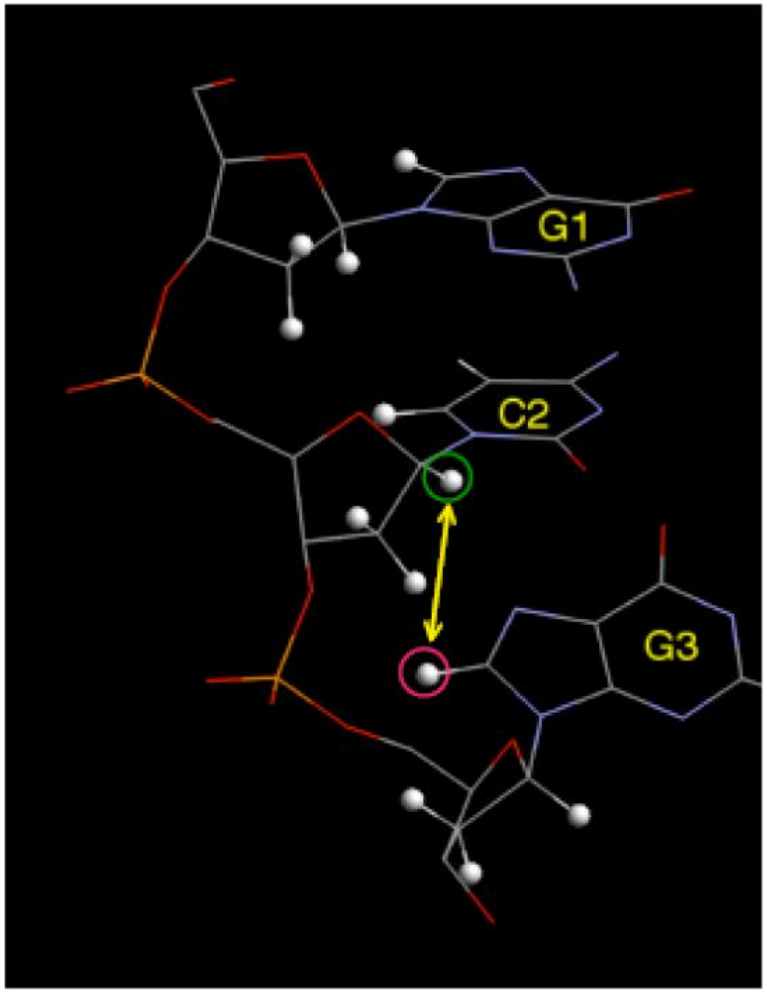






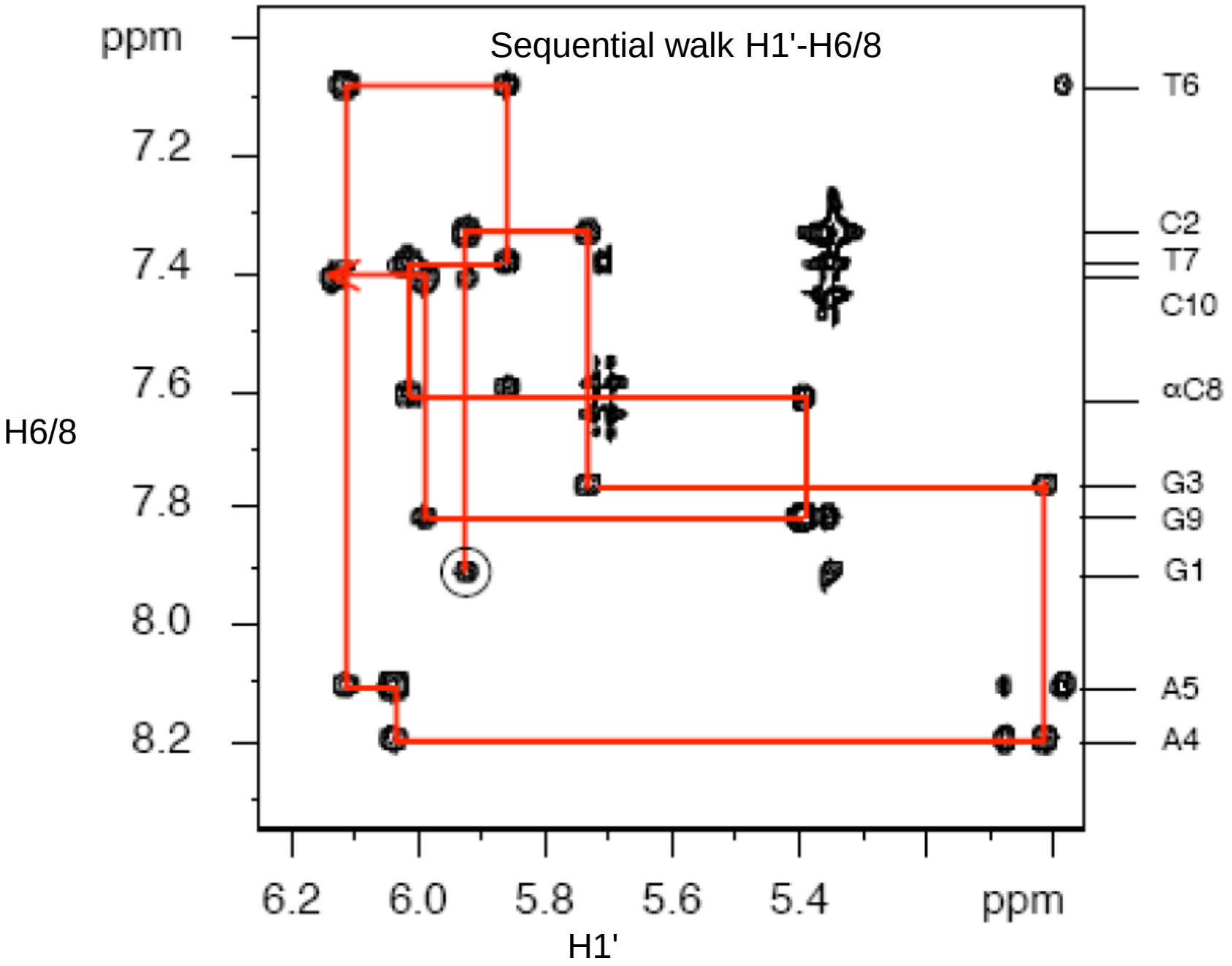
d(GCGAATTαCGC)$_2$





d(GCGAATTαCGC)$_2$

d(GCGAATTαCGC)$_2$



Strukturní informace získaná ze spekter NMR

1. **Odstupy mezi atomy ze spekter NOESY (Nuclear Overhauser Effect SpectroscopY): intenzita cross-peaku** $\propto r^{-6}$

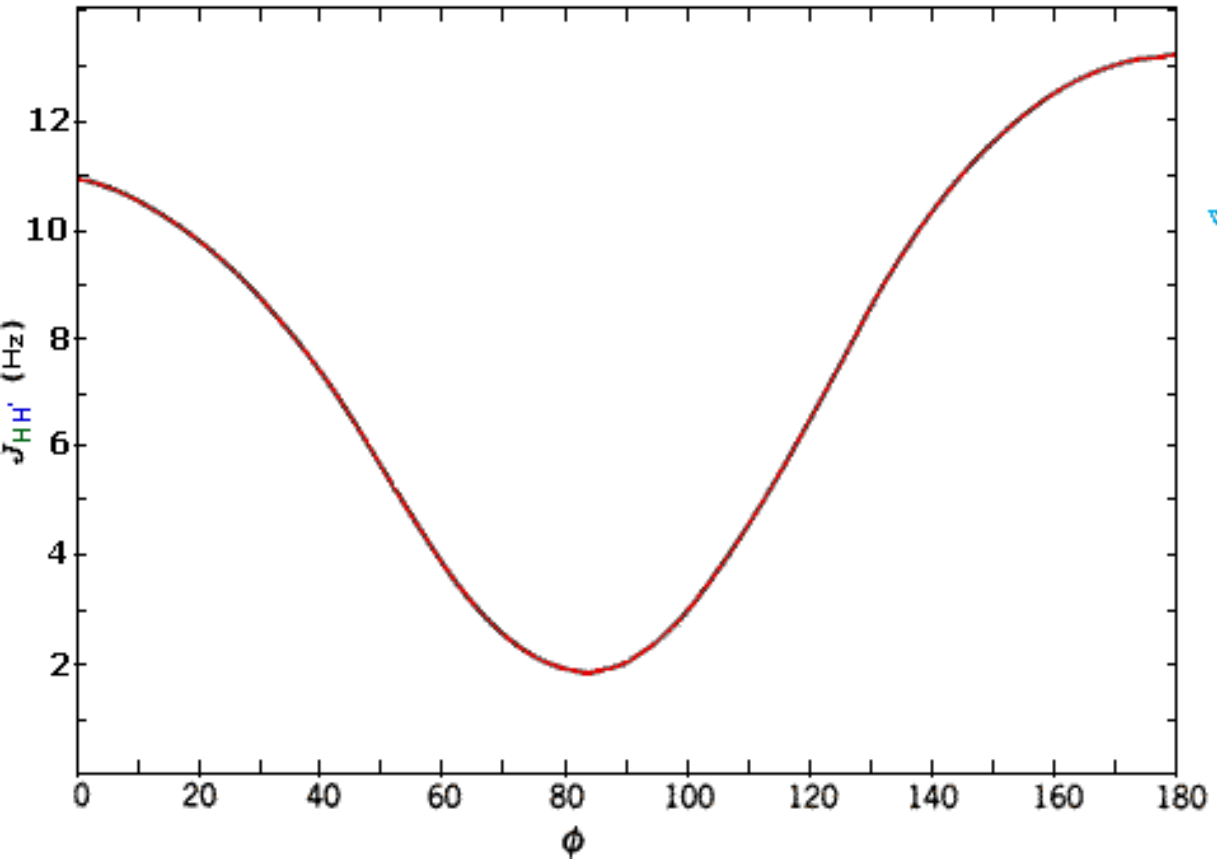
2. **Dihedrální úhly odvozené ze spin-spinových konstant získaných ze spekter COSY (COrrelated SpectroscopY) pomocí Karplusovy rovnice**

3. Informace získané z chemických posunů

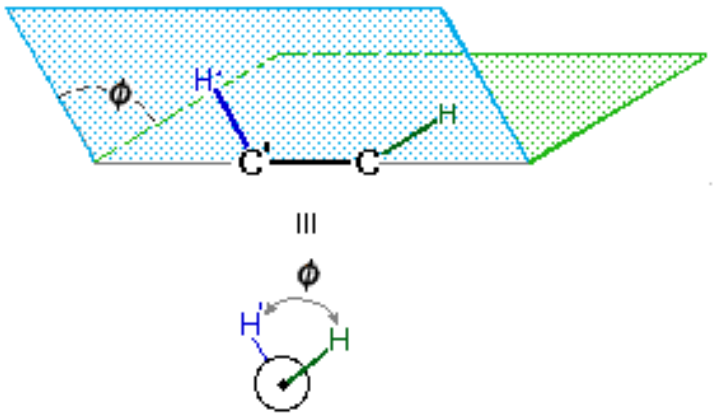
1., 2: Je možno zohlednit ve výpočtech molekulového modelování jako „restraints“

3.: Kvantifikace vlivu okolí na chemický posun je komplikovaná, proto jej těžko lze zohlednit automaticky při výpočtu struktury.

Velikost vicinálních spin-spin-interakčních konstant závisí na dihedrálním úhlu H-X-X-H. Tuto závislost popisuje tzv. Karplusova rovnice.



$$J = A - B\cos\phi + C\cos 2\phi$$



V jiných pramenech má Karplusova rovnice tvar uvedený dole. Ukažte, že obě rovnice jsou ekvivalentní a vypočtěte vztah mezi koeficienty A, B, C a P, Q, R

$$^3J_{HH} = P\cos^2\phi + Q\cos\phi + R$$

http://www2.chemistry.msu.edu/faculty/reusch/virttxtjml/spectrpy/nmr/nmr2.htm

Strukturní informace získaná ze spekter NMR

1. **Odstupy mezi atomy ze spekter NOESY (Nuclear Overhauser Effect SpectroscopY): intenzita cross-peaku** $\propto r^{-6}$

2. **Dihedrální úhly odvozené ze spin-spinových konstant získaných ze spekter COSY (COrrelated SpectroscopY) pomocí Karplusovy rovnice**

3. Informace získané z chemických posunů

1., 2: Je možno zohlednit ve výpočtech molekulového modelování jako „restraints“

3.: Kvantifikace vlivu okolí na chemický posun je komplikovaná, proto jej těžko lze zohlednit automaticky při výpočtu struktury.

Molecular modeling is an ensemble of methods for the study of molecular conformations, based on energy calculations. It comprises geometry optimizations by minimizing the energy and molecular dynamics simulations.

Empirical energy calculation:

$$\mathcal{V}(\mathbf{r}^N) = \sum_{\text{bonds}} \frac{k_i}{2}(l_i - l_{i,0})^2 + \sum_{\text{angles}} \frac{k_i}{2}(\theta_i - \theta_{i,0})^2$$

$$+ \sum_{\text{torsions}} \frac{V_n}{2}(1 + \cos(n\omega - \gamma))$$

$$+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right)$$

Structural variables $\{f(\mathbf{r}^N)\}$

Force-field parameters (fix)

$$\mathcal{V}(\mathbf{r}^N) = \sum_{\text{bonds}} \frac{k_i}{2}(l_i - l_{i,0})^2 + \sum_{\text{angles}} \frac{k_i}{2}(\theta_i - \theta_{i,0})^2$$

$$+ \sum_{\text{torsions}} \frac{V_n}{2}(1 + \cos(n\omega - \gamma))$$

$$+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right)$$

$$\mathcal{V}(\mathbf{r}^N) = \sum_{\text{bonds}} \frac{k_i}{2}(l_i - l_{i,0})^2 + \sum_{\text{angles}} \frac{k_i}{2}(\theta_i - \theta_{i,0})^2$$

$$+ \sum_{\text{torsions}} \frac{V_n}{2}(1 + \cos(n\omega - \gamma))$$

$$+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right)$$

Two main techniques are used to optimize the geometry of a molecular model:

1. Molecular mechanics calculations: Using an optimization algorithm, we move the atomic coordinates so as to minimize the empirical potential.

2. Molecular dynamics simulations: We assign random velocities to the atoms, whose mean square corresponds to a chosen temperature. Then, in infinitesimal steps, we let the atoms move in the empirical potential.
MD simulations allow us:
- to describe the dynamical behaviour of the molecule and to include solvent
- to overcome energy barriers
- to estimate the Gibbs free energy of the solvated molecule

$$\mathcal{V}(\mathbf{r}^N) = \sum_{\text{bonds}} \frac{k_i}{2}(l_i - l_{i,0})^2 + \sum_{\text{angles}} \frac{k_i}{2}(\theta_i - \theta_{i,0})^2$$

$$+ \sum_{\text{torsions}} \frac{V_n}{2}(1 + \cos(n\omega - \gamma))$$

$$+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right)$$

Principal problem of Molecular mechanics calculations:
For most of the molecules, many conformations close in energy exist, and our calculations are not accurate enought to define which conformation corresponds to the global minimum.

Therefore, we have to check the calculated model against experimental structural data.

This can be done by adding to the empirical energy a „penalty function“, accounting for discrepancies between experiment and the calculated model.

$$\mathcal{V}(\mathbf{r}^N) = \sum_{\text{bonds}} \frac{k_i}{2}(l_i - l_{i,0})^2 + \sum_{\text{angles}} \frac{k_i}{2}(\theta_i - \theta_{i,0})^2$$

$$+ \sum_{\text{torsions}} \frac{V_n}{2}(1 + \cos(n\omega - \gamma))$$

$$+ \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left(4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \right)$$

From NOESY spectra, we usually obtain upper and lower limits for a given interatomic distance d_j, d_j^{upper} and d_j^{lower}. A pseudo-energy term is then added for distances higner than the upper limit or lower than the lower limit:

$$E_{distance}^{upper} = \sum_{upper} \begin{cases} A_j(d_j - d_j^{upper})^2 & \text{if } d_j > d_j^{upper} \\ 0 & \text{otherwise.} \end{cases}$$

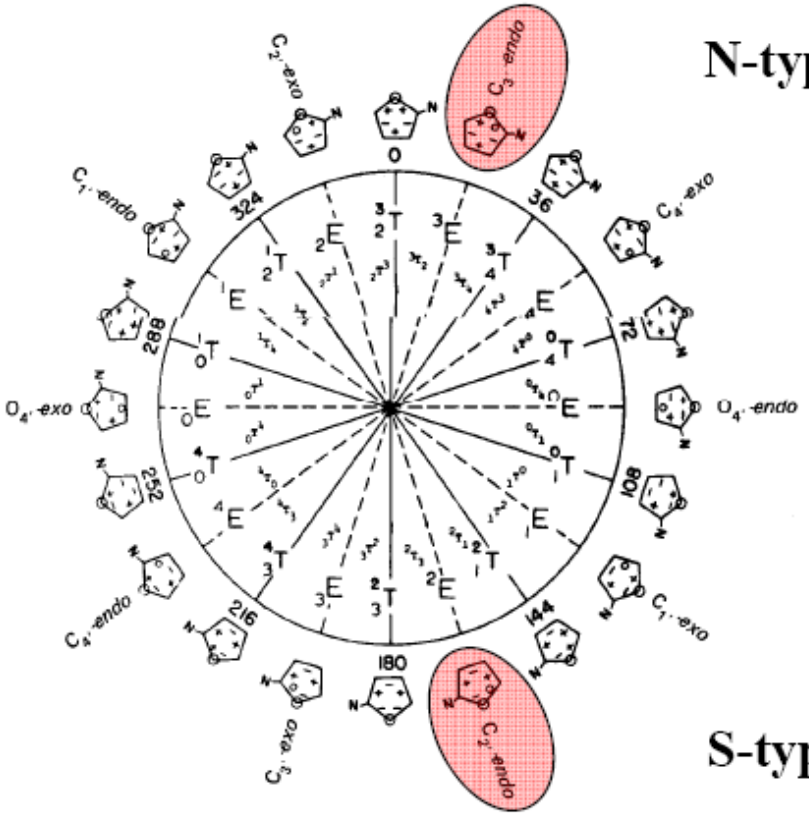
$$E_{distance}^{lower} = \sum_{lower} \begin{cases} A_j(d_j - d_j^{lower})^2 & \text{if } d_j < d_j^{lower} \\ 0 & \text{otherwise.} \end{cases}$$

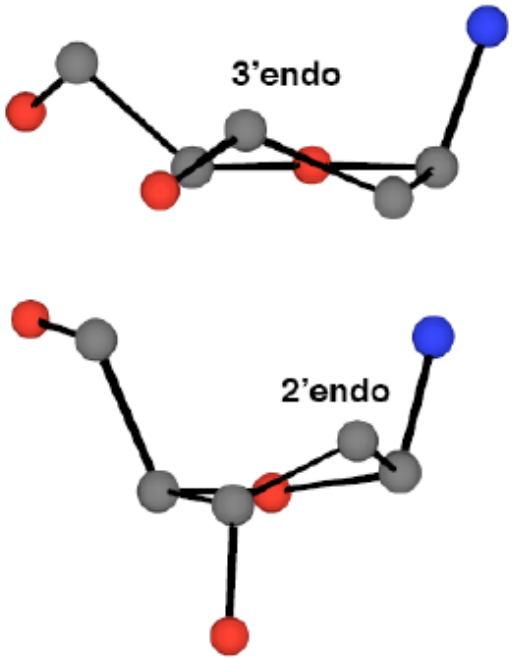
Similar penalty terms are added for deviations of dihedral angles from the experimental upper and lower limits.

Structural Parameter: The Sugar Pucker

The five membered furanose ring is not planar. It can be puckered in an envelope form (E) with 4 atoms in a plane or it can be in a twist form (T). The geometry is defined by two parameters: **the pseudorotation phase angle (P)** and the **pucker amplitude (Φ)**.

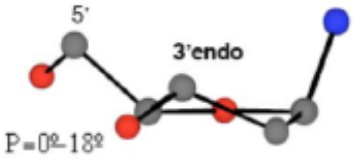
The pseudorotation cycle of the furanose ring shows the relationships among the phase angle of pseudoration (**P**) and the alternative **E/T** (envelope/twist) and ***endo/exo*** notations. Conformations corresponding to the northern half are designated as N-type and those corresponding to the southern half are designated as S-type. The P values commonly observed for the N and S conformations are represented by C3'-endo (^{3}E) and C2'-endo (^{2}E), respectively.





Measuring the Sugar Pucker by NMR

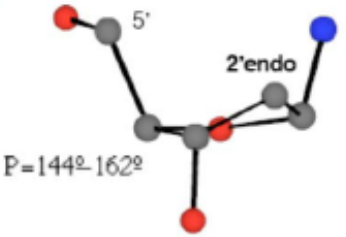
Very weak $J_{H1'-H2'}$ and strong $J_{H3'-H4'}$ cross-peaks correspond to *pure N-type conformation* (preferred conformation in RNA). Strong $J_{H1'-H2'}$ and weak $J_{H3'-H4'}$ cross-peaks correspond to *pure S-type conformation*. $J_{H2'-H3'}$ is similar in both states. Intermediate intensities indicate an equilibrium between N and S states.



P=0º-18º

Ribose: $^3J_{H1'-H2'} \approx 1$ Hz

Deoxyribose: $^3J_{H1'-H2'} \approx 1.8$ Hz



P=144º-162º

Ribose: $^3J_{H1'-H2'} \approx 7.9$ Hz

Deoxyribose: $^3J_{H1'-H2'} \approx 10$ Hz

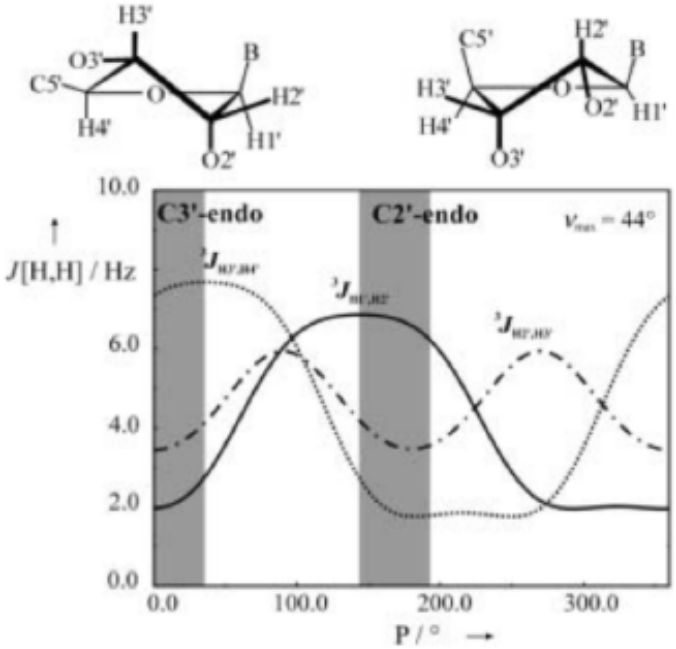
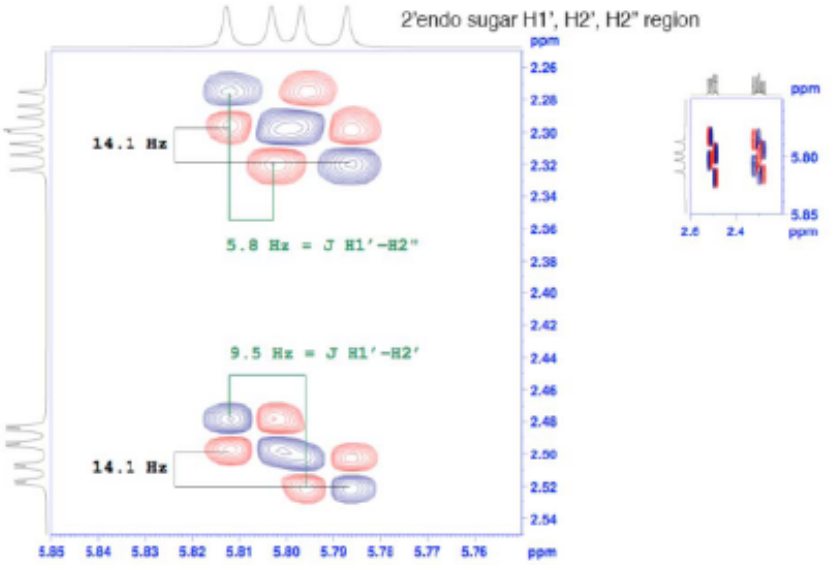


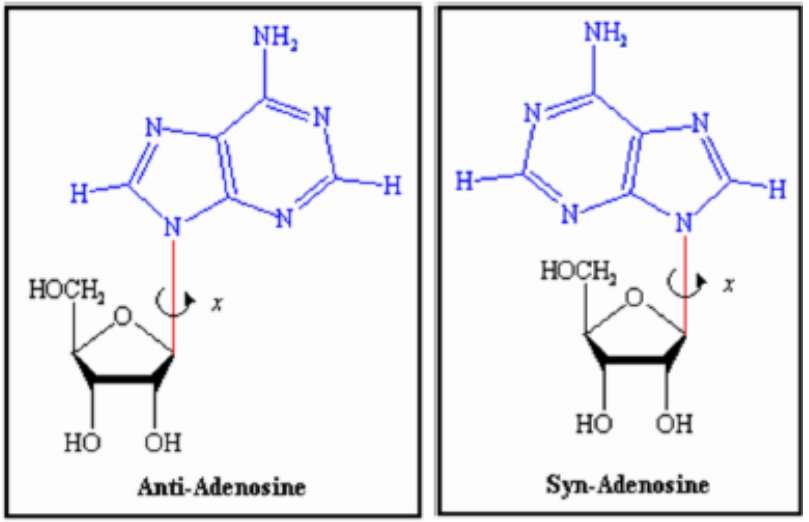
Figure 27. Karplus relation of ^3J(H1',H2'), ^3J(H2',H3'), and ^3J(H3',H4') coupling constants depending on the pseudorotation phase P at a pseudorotation amplitude ν_{max} of 44°.



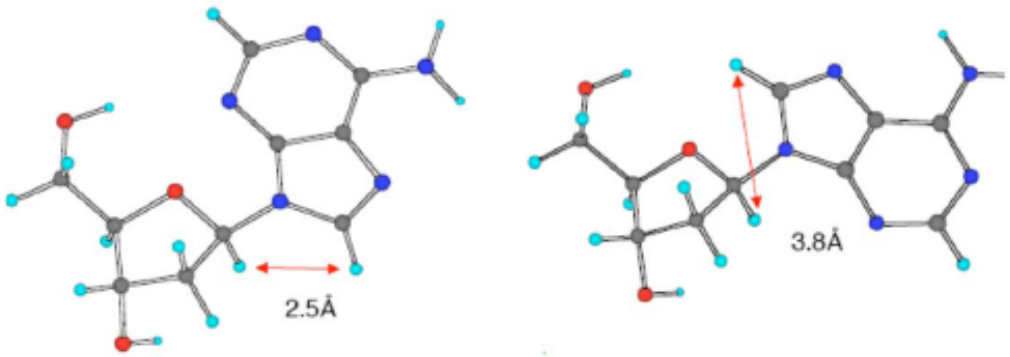
Structural Parameter: The Glycosidic Torsion Angle Measuring Using Distances

The base can exist in 2 distinct orientations about the N-glycosidic bond. These conformations are identified as, ***syn*** and ***anti***.

The *anti* conformation predominates.



Distance information (H8 - H1') determines the glycosidic torsion angle:



syn ***anti***

Furthermore, although H8 resonances from nucleotides in the *anti* glycosidic conformation are in close contact to several sugar proton resonances, in the *syn* conformation no close contact to sugar protons is expected (except to the anomeric resonances).

Measuring the Glycosidic Torsion Angle Using Scalar Couplings

Base-to-sugar three-bond C6/C8-Hl' and Cl'- H6/H8 coupling constants can also be used to define the glycosidic angle χ:

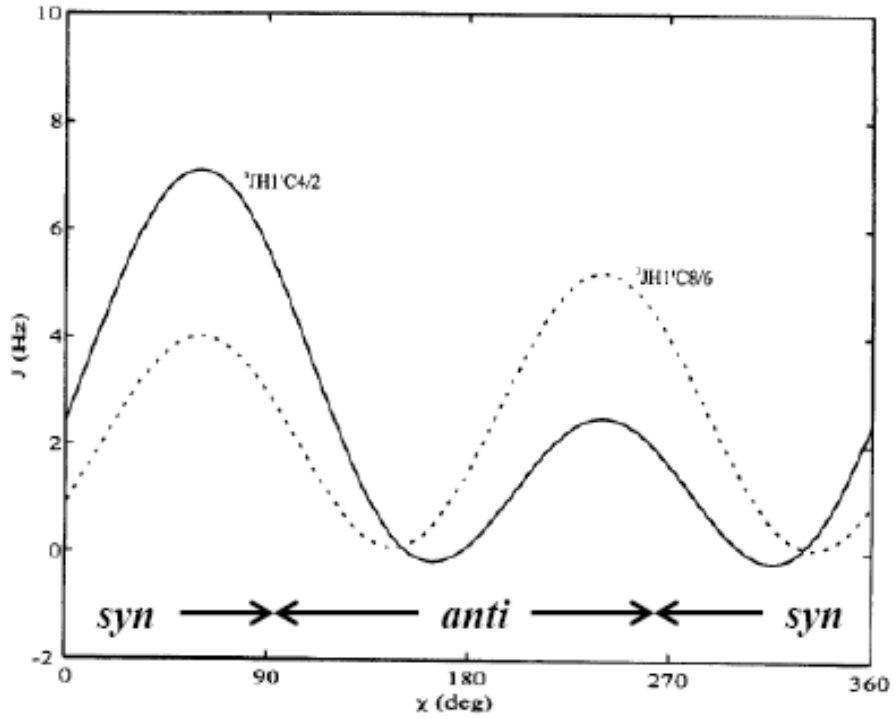


Fig. 8. The $^3J_{H1'C8}$- and $^3J_{H1'C4}$-coupling constants calculated as a function of the torsion angle χ on the basis of their Karplus relations (see text).

χ

$^3J_{H1'-C6}$ (U,C,T)

$^3J_{H1'-C2}$ (U,C,T)

$^3J_{H1'-C8}$ (A,G)

$^3J_{H1'-C4}$ (A,G)

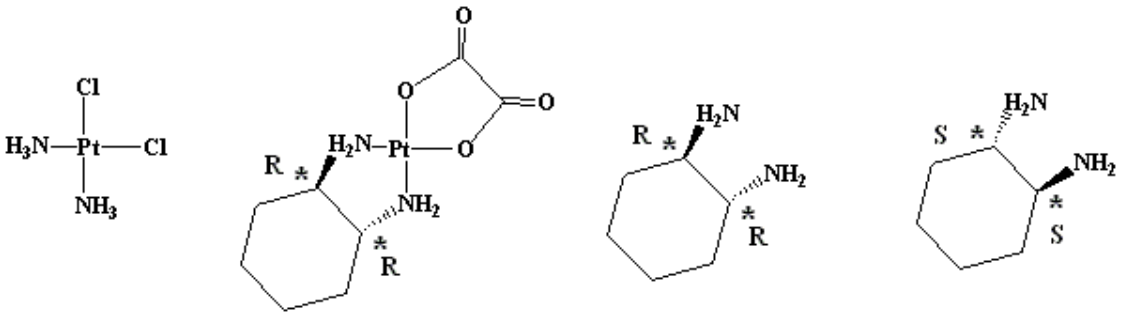
Příklad využití technik NMR pro strukturní studii

Unexpected intrastrand-to-interstrand rearrangement of Pt-GG crosslinks formed between an analogue of the antitumor drug cisplatin and a DNA duplex: Evidence for kinetic instability of Pt-N bonds

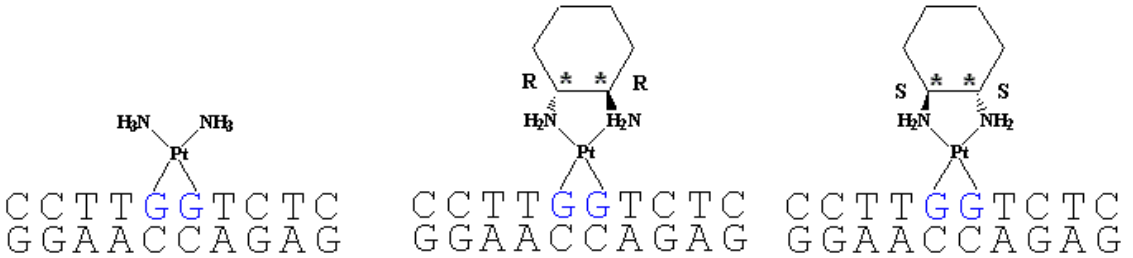
Kubicek, K.; Monnet, J.; Scintilla, S.; Kopecna, J.; Arnesano, F.; Trantirek, L.; Chopard, C.; Natile, G.; Kozelka, J. *Chem. Asian J*. 2010, **5**, 244-247

Context

Cisplatin and Oxaliplatin are efficient antitumor drugs with different cytotoxic properties.[1] The differences in their biological effects are believed to be related to different structural perturbations they cause to their cellular target, DNA. Oxaliplatin contains a chiral diamine ligand, and the question arises how the chirality affects the adduct structure.



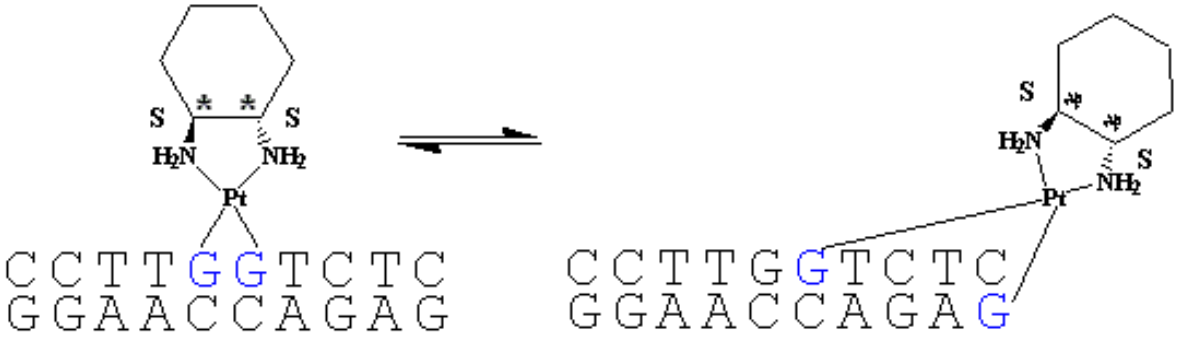
We are therefore investigating DNA oligonucleotides bearing adducts of *cis*-Pt$(NH_3)_2^{2+}$, Pt(*R,R*-DACH)$^{2+}$, and Pt(*S,S*-DACH)$^{2+}$, using NMR spectroscopy and molecular modeling.



[1] E. Raymond, S. Faivre, S. Chaney, J. Woynarowski, E. Cvitkovic, Mol.Cancer Ther. 2002, 1, 227-235.

An unexpected result

In this work, we have prepared and purified the adduct with Pt(*S,S*-DACH)$^{2+}$. During the NMR analysis, we observed that the intrastrand crosslink reversibly rearranged to a new species that we identified as an interstrand crosslink with the 5'-terminal G.



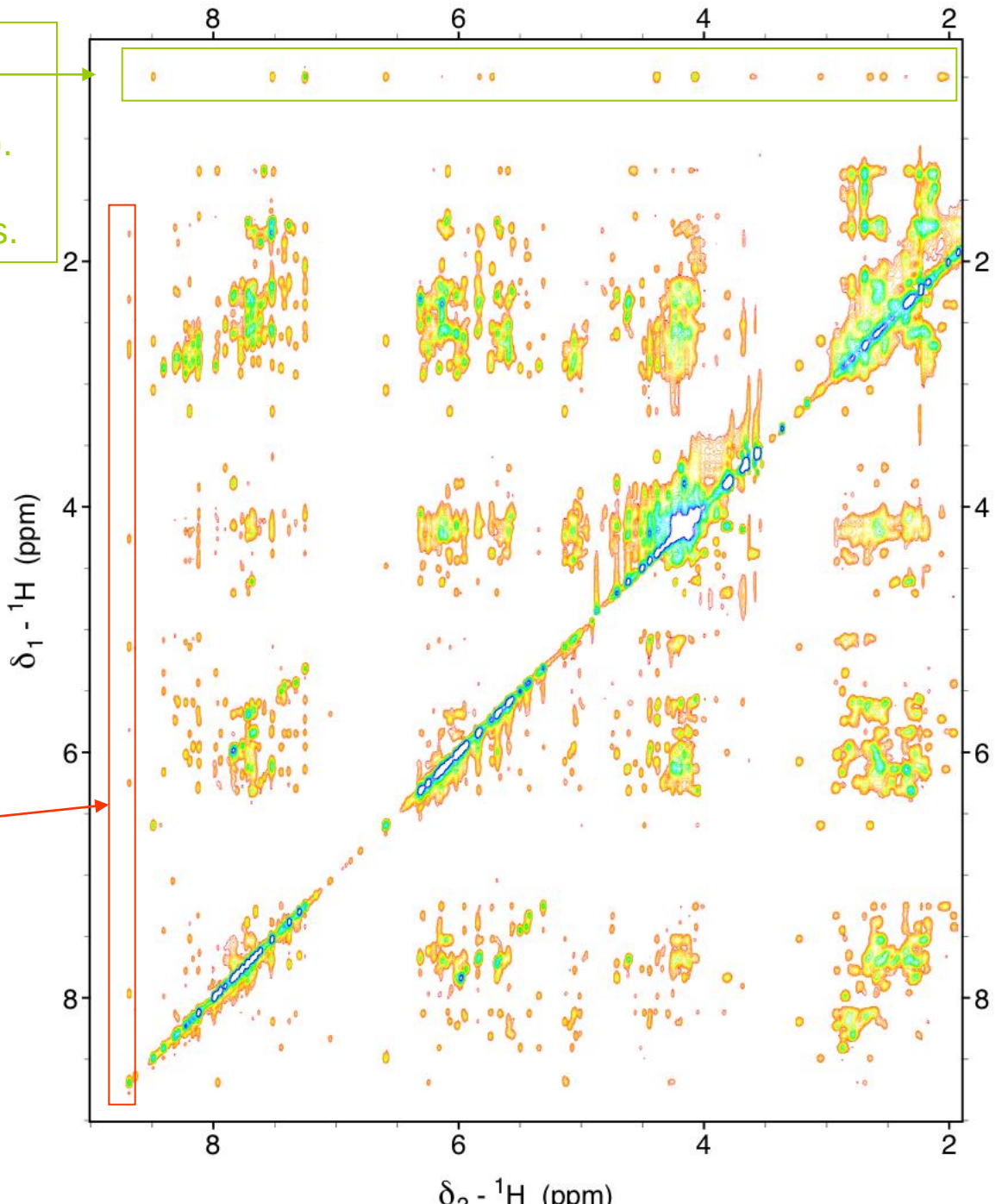
A simple method to investigate the interstrand crosslink separately: heating to 40 °C

At room temperature, the intrastrand and interstrand crosslinks form an approximately 1:1 equilibrium. Above 15 °C, the intrastrand crosslink starts to melt and at 40 °C, its signal coalesce and disappear from the NOESY spectra. At this temperature, it was thus possible to study the interstrand crosslink separately.

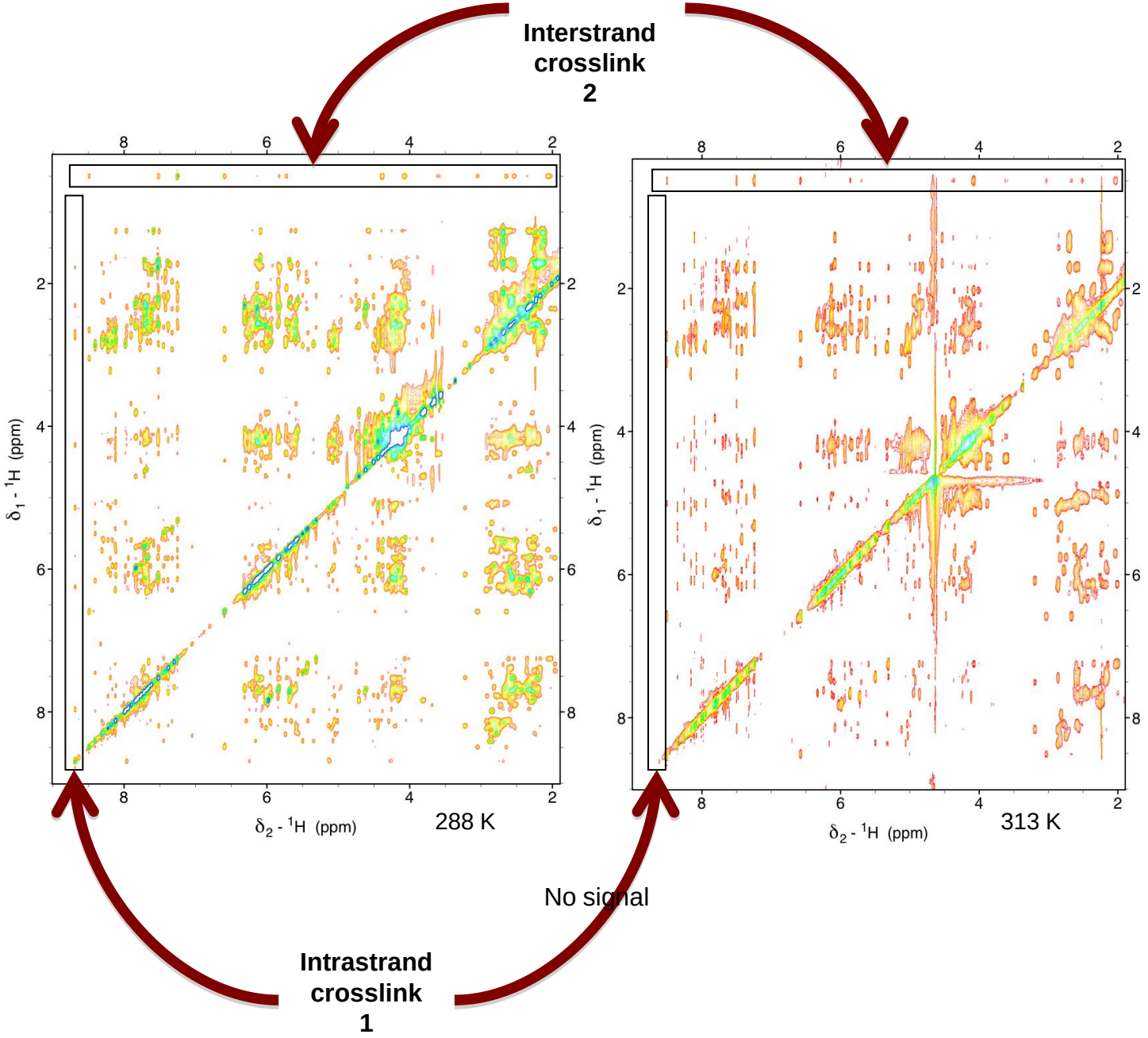
Peaks of newly formed species (turned out to be a DNA duplex containing an interstrand crosslink). They correspond to a strongly shielded proton with many contacts.

500 ms NOESY spectra of a ~2.6 mM solution pH 7.2 (20 mM phosphate) containing 1 and 2

peaks of initial species (Pt-crosslinked DNA duplex)



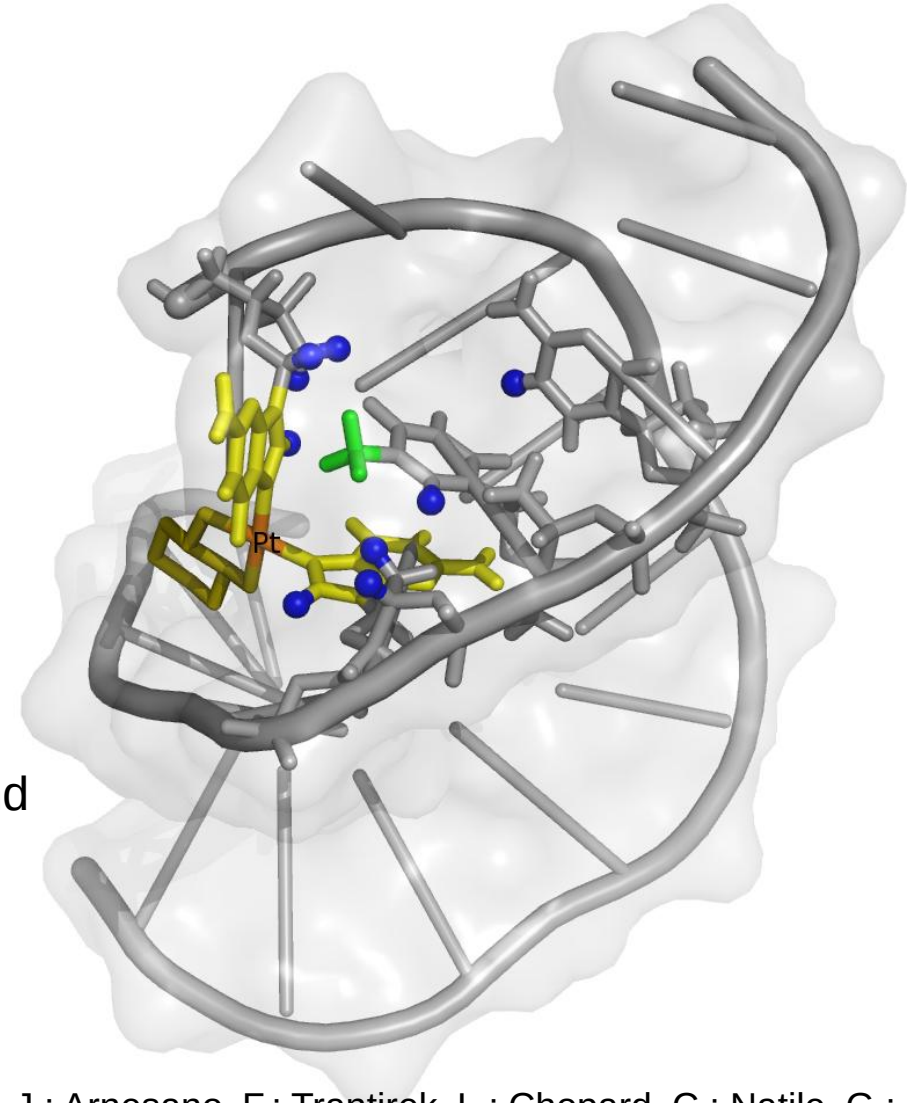
500 ms NOESY spectra of a ~2.6 mM solution pH 7.2 (20 mM phosphate) containing 1 and 2



3D model of the interstrand crosslink shows a very unusual DNA structure

The methyl group of T7 (green) is intercalated between the two platinated guanines (yellow), with many contacts to other protons (blue). This intercalation resembles to that seen in recognition complexes between cellular proteins and kinked DNA.

The intercalation of thymine CH_3 between two guanines explains the strong shielding of this methyl group!



Conclusion

Pt-N bonds are kinetically unstable and can break at physiologically relevant rate.

Kubicek, K.; Monnet, J.; Scintilla, S.; Kopecna, J.; Arnesano, F.; Trantirek, L.; Chopard, C.; Natile, G.; Kozelka, J. *Chem. Asian J*. 2010, **5**, 244-247