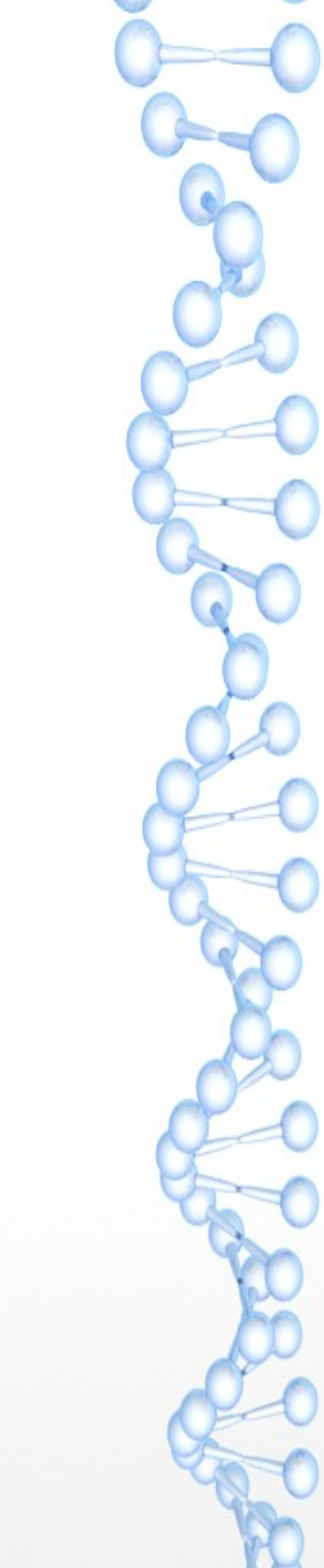
# DNA local conformers and their automatic assignment

Paulína Božíková

Charles university in Prague, Faculty of Science

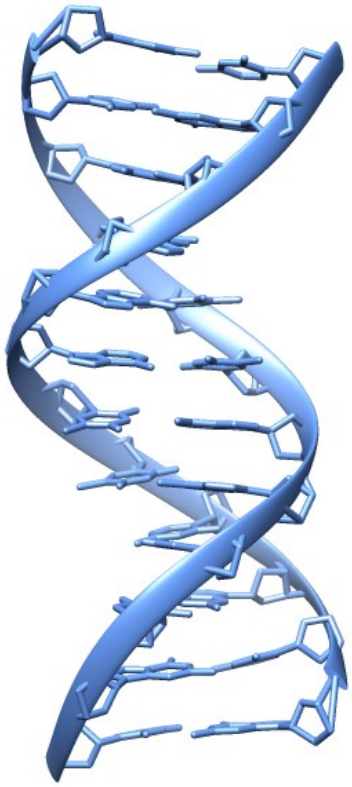


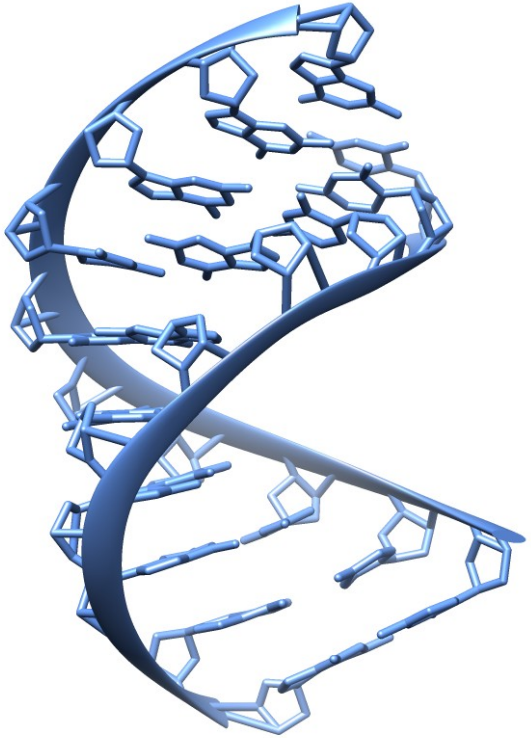


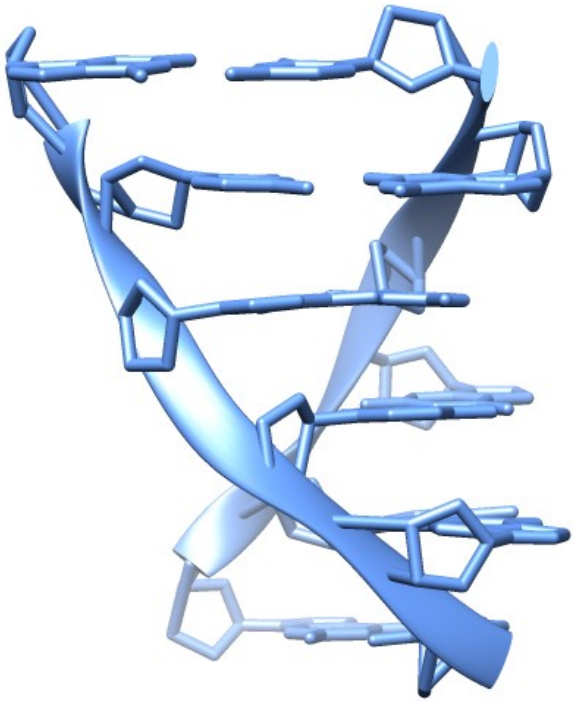
# Introduction

# DNA

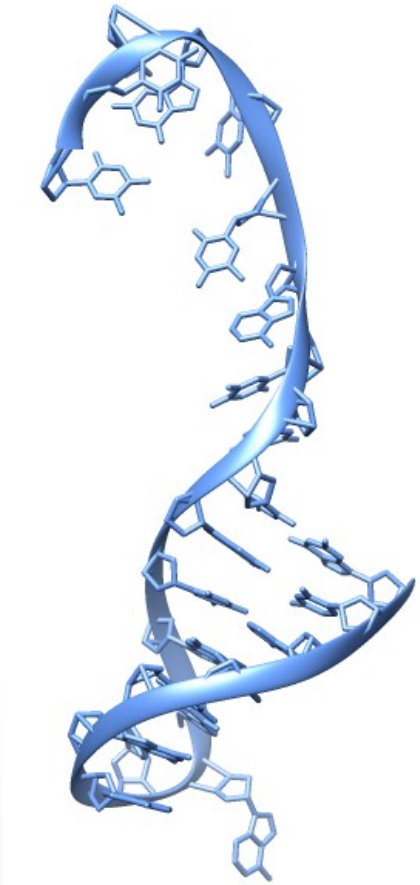
- 3 known forms (A-, B-, Z-DNA)

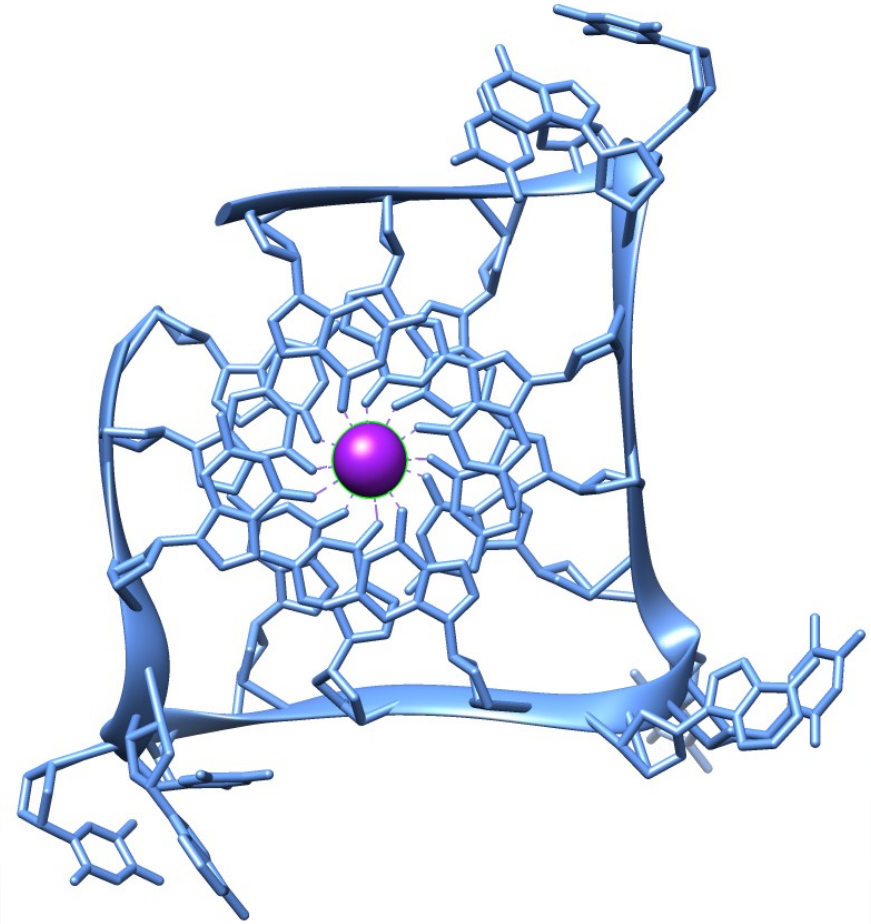




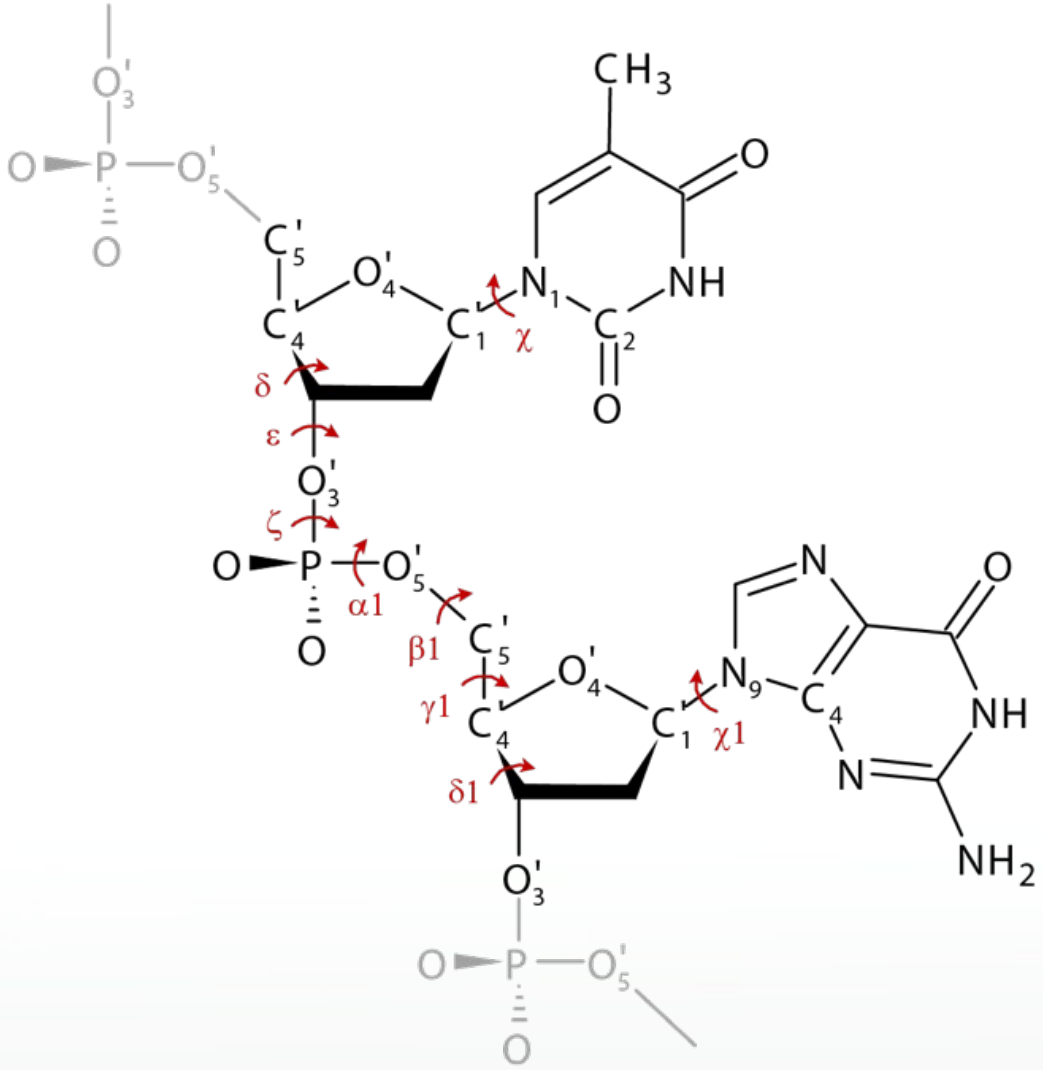


# However ...



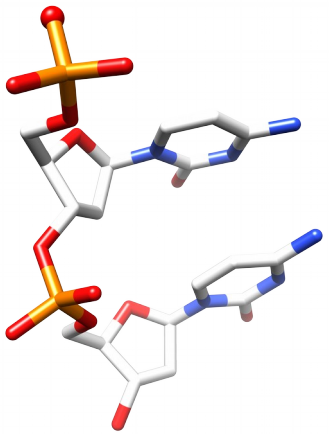


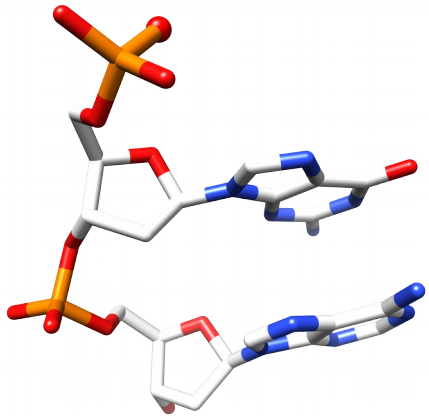
# Dinucleotide conformer definition

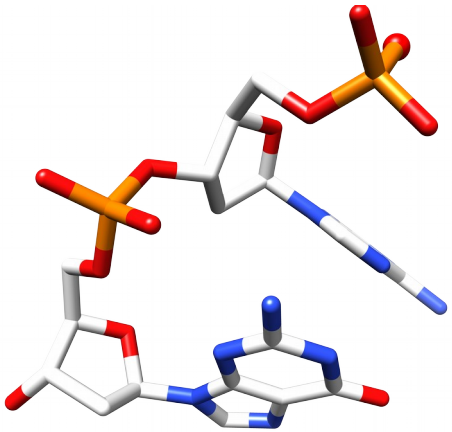


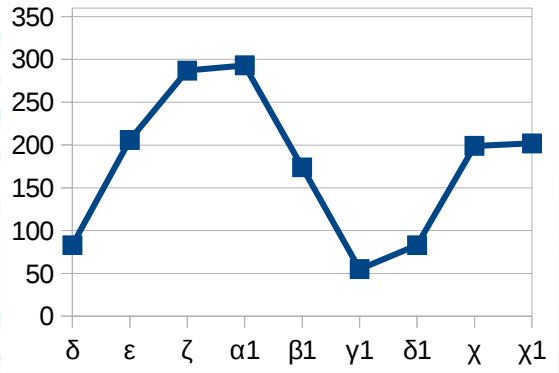
# Dinucleotide conformers

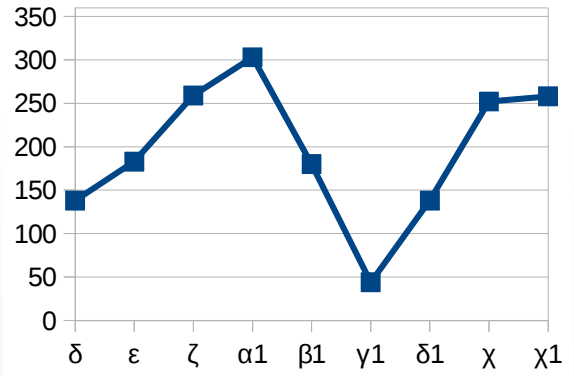
- A-DNA B-DNA quadruplex

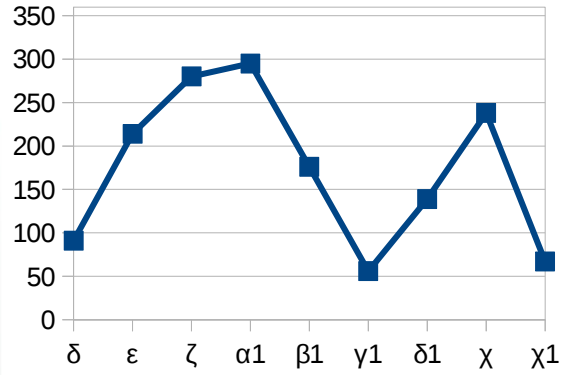








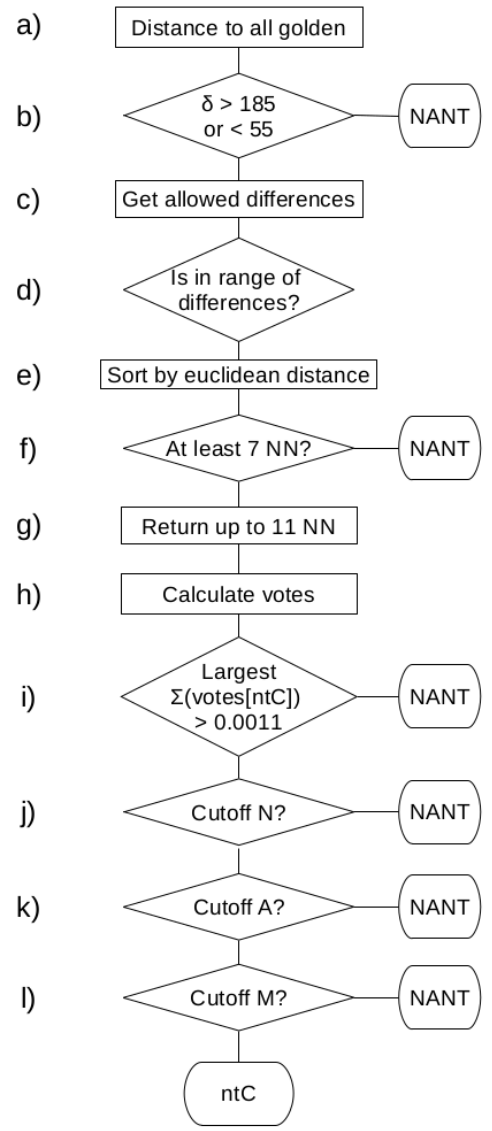
# Methods

# Selection of the structures

- Protein Data Bank (PDB)
- Crystallographic resolution 3.0 Å
- 2406 protein/DNA complexes, 882 naked DNA
- Excluding sequentially redundant structures
- Further analyzed 58,039 dinucleotides

# Conformational assignment of a dinucleotide

- Original golden standard - 24 ntC classes (Čech et al., BMC
- Bioinformatics, 2013)
- Physically possible delta
- Differences 5 ESD
- Nearest neighbours
- 1/(Euclidean distance)^2
- Cutoffs 28°, ±60º



# New conformers

- Original set – 24 conformers
- Not assigned dinucleotides
  - Clusters – R package
    - Hclust
    - Circular distance
    - Heatmap

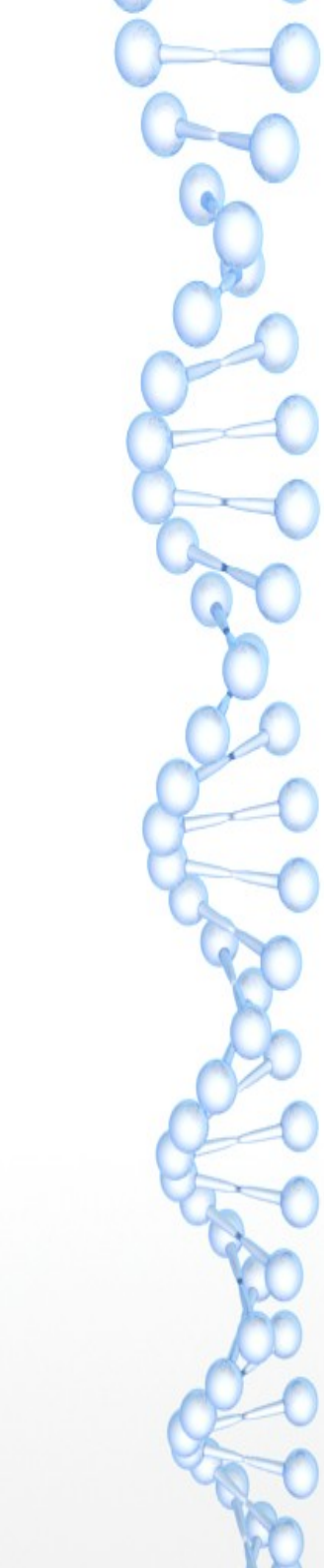
# Self-consistent golden set

- Used the previously described method
- High-resolution structures
- One dinucleotide compared to the rest
- If not assigned to the same class excluded from the golden set

# Results

# Basic

- Described 57 dinucleotide conformers
- To simplify – 12 ntA classes comprising similar ntC



| ntA letter | all steps | | protein/ DNA | | naked DNA | |
|---|---|---|---|---|---|---|
| | Number | Percent | Number | Percent | Number | Percent |
| AAA | 3057 | 5.3 | 2002 | 3.9 | 1055 | 14.5 |
| A-B | 2770 | 4.8 | 2400 | 4.7 | 370 | 5.1 |
| B1A | 3991 | 6.9 | 3397 | 6.7 | 594 | 8.1 |
| B2A | 321 | 0.6 | 284 | 0.6 | 37 | 0.5 |
| BBB | 19675 | 33.9 | 18075 | 35.6 | 1600 | 21.9 |
| 2B1 | 3742 | 6.4 | 3224 | 6.4 | 518 | 7.1 |
| 3B1 | 2938 | 5.1 | 2884 | 5.7 | 54 | 0.7 |
| B12 | 3631 | 6.3 | 3417 | 6.7 | 214 | 2.9 |
| BB2 | 3229 | 5.6 | 2617 | 5.2 | 612 | 8.4 |
| miB | 2973 | 5.1 | 2758 | 5.4 | 215 | 2.9 |
| SQX | 501 | 0.9 | 214 | 0.4 | 287 | 3.9 |
| ZZZ | 342 | 0.6 | 158 | 0.3 | 184 | 2.5 |
| NAN | 10869 | 18.7 | 9314 | 18.4 | 1555 | 21.3 |

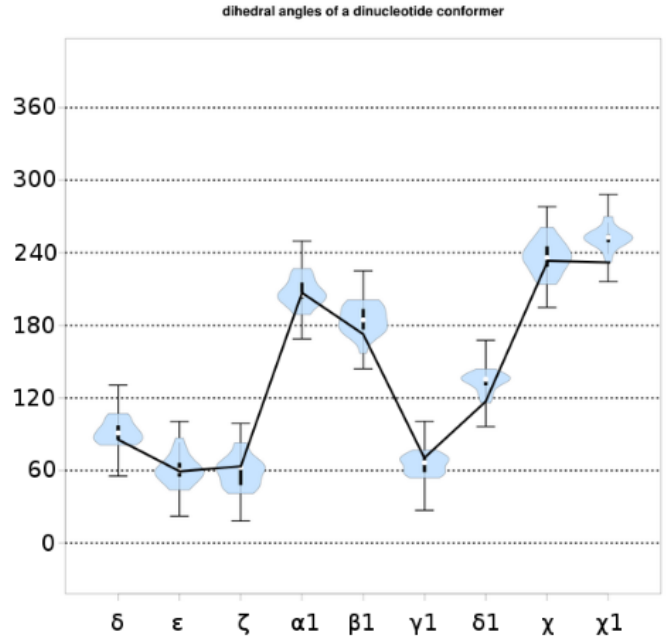
# Annotation of the dinucleotide conformers

- ntC AB02
- Extremely low

  ε near 60°
- Compensated by

  a value of ζ
- Double helices,

  hairpins, unpaired strand ends

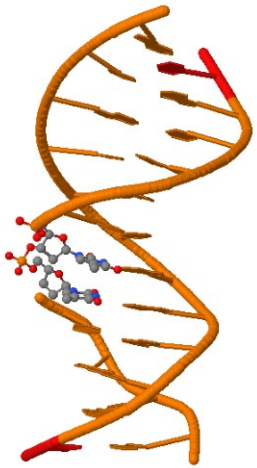



| | δ | ε | ζ | α1 | β1 | γ1 | δ1 | χ | χ1 |
|---|---|---|---|---|---|---|---|---|---|
| step_torsions | 85.8 | 59.3 | 63.5 | 206.9 | 172.7 | 70.4 | 116.9 | 233.4 | 232.0 |
| ntC_average | 92.4 | 60.7 | 59.2 | 208.0 | 184.7 | 64.2 | 133.3 | 236.7 | 252.5 |
| Δ torsions | -6.6 | -1.4 | 4.3 | -1.1 | -12.0 | 6.2 | -16.4 | -3.3 | -20.5 |

# Dickerson-Drew dodecamer

Home

Center view on ◉ step or ○ molecule



PDB ID 1bna already assigned, checking structure ...
Downloading PDB ID 1bna ... structure didn't change.

**Results of the assignment of 22 detected steps, can be also downloaded as csv file.**

**Click a row in table for analysis of results.**

| Step name | ntA | ntC |
|---|---|---|
| 1bna_A_DC1_DG2 | B12 | BB04 |
| 1bna_A_DG2_DC3 | B1A | BA06 |
| 1bna_A_DC3_DG4 | A-B | AB01 |
| 1bna_A_DG4_DA5 | B12 | BB04 |
| 1bna_A_DA5_DA6 | 2B1 | BB01 |
| 1bna_A_DA6_DT7 | B1A | BA06 |
| 1bna_A_DT7_DT8 | B1A | BA06 |
| 1bna_A_DT8_DC9 | 2B1 | BB01 |
| 1bna_A_DC9_DG10 | BBB | BB00 |
| 1bna_A_DG10_DC11 | BB2 | BB07 |
| 1bna_A_DC11_DG12 | B1A | BA06 |
| 1bna_B_DC13_DG14 | BBB | BB00 |
| 1bna_B_DG14_DC15 | B1A | BA06 |
| 1bna_B_DC15_DG16 | A-B | AB01 |
| 1bna_B_DG16_DA17 | BBB | BB00 |
| 1bna_B_DA17_DA18 | BBB | BB00 |
| 1bna_B_DA18_DT19 | B1A | BA06 |
| 1bna_B_DT19_DT20 | 2B1 | BB01 |
| 1bna_B_DT20_DC21 | 2B1 | BB01 |
| 1bna_B_DC21_DG22 | BBB | BB00 |
| 1bna_B_DG22_DC23 | B2A | BA17 |
| 1bna_B_DC23_DG24 | B1A | BA06 |

Steps with non-standard or missing atoms have not been assigned, description of conformers is defined in the table.

JSmol

**Download**

Results (as csv)
Definition of the conformers as torsion averages (csv) and their esd values (csv) or as their Cartesian coordinates of the conformers

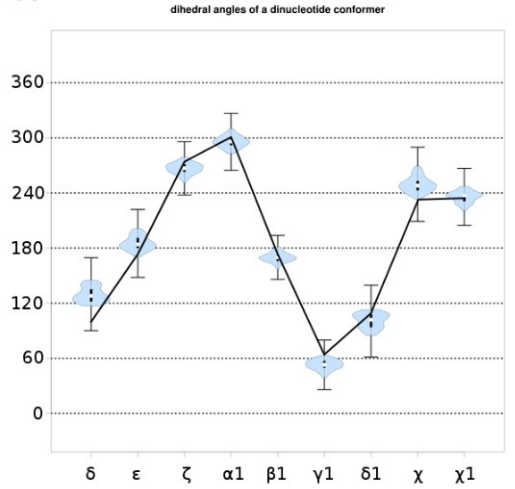
**Download the papers:**

definition of conformers:
Svozil et al., Nucleic Acids Research, 36, 3690 (2008).
description of the assignment process:
Čech et. al, BMC Bioinformatics, 14:205 (2013).

Use left click and drag to rotate, wheel to zoom, and right click to show the menu. See the J(S)mol wiki for detailed description.

© Jiří Černý & Bohdan Schneider

BI-to-A complex conformer with δ1 from C3' to O4'-endo, χ/χ1 are B-like



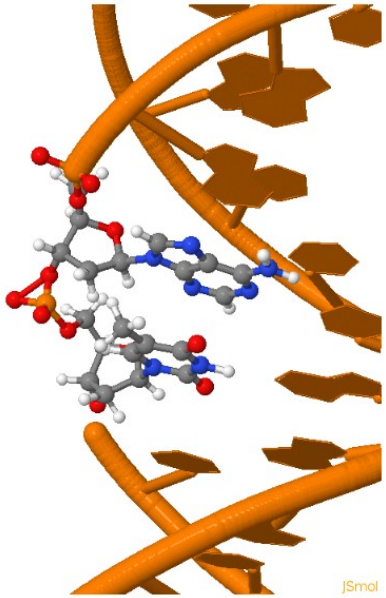

| | δ | ε | ζ | α1 | β1 | γ1 | δ1 | χ | χ1 |
|---|---|---|---|---|---|---|---|---|---|
| step_torsions | 98.9 | 173.6 | 274.1 | 300.8 | 173.4 | 64.1 | 108.9 | 232.7 | 234.3 |
| ntC_average | 129.0 | 185.1 | 266.8 | 295.7 | 169.9 | 52.9 | 100.9 | 248.8 | 235.7 |
| Δ torsions | -30.1 | -11.5 | 7.3 | 5.1 | 3.5 | 11.2 | 8.0 | -16.1 | -1.4 |

www.dnatco.org (Černý et al., NAR, 2016)

# “Average B-DNA”

Home

Center view on (•) step or ( ) molecule




preprocessing pdb file ...
measuring torsions ...
assigning conformers (24 residues, it might take some time) ...
... DONE in 8 seconds

**Results of the assignment of 22 detected steps, can be also downloaded as csv file.**

**Click a row in table for analysis of results.**

| Step name | ntA | ntC |
|---|---|---|
| average-bdna_A_DC1_DG2 | NAN | NANT |
| average-bdna_A_DG2_DC3 | NAN | NANT |
| average-bdna_A_DC3_DG4 | NAN | NANT |
| average-bdna_A_DG4_DA5 | NAN | NANT |
| average-bdna_A_DA5_DA6 | miB | BB15 |
| average-bdna_A_DA6_DT7 | miB | BB15 |
| average-bdna_A_DT7_DT8 | miB | BB15 |
| average-bdna_A_DT8_DC9 | NAN | NANT |
| average-bdna_A_DC9_DG10 | NAN | NANT |
| average-bdna_A_DG10_DC11 | NAN | NANT |
| average-bdna_A_DC11_DG12 | NAN | NANT |
| average-bdna_B_DC1_DG2 | NAN | NANT |
| average-bdna_B_DG2_DC3 | NAN | NANT |
| average-bdna_B_DC3_DG4 | NAN | NANT |
| average-bdna_B_DG4_DA5 | NAN | NANT |
| average-bdna_B_DA5_DA6 | miB | BB15 |
| average-bdna_B_DA6_DT7 | miB | BB15 |
| average-bdna_B_DT7_DT8 | NAN | NANT |
| average-bdna_B_DT8_DC9 | NAN | NANT |
| average-bdna_B_DC9_DG10 | NAN | NANT |
| average-bdna_B_DG10_DC11 | NAN | NANT |
| average-bdna_B_DC11_DG12 | NAN | NANT |

Steps with non-standard or missing atoms have not been assigned, description of conformers is defined in the table.

**Download**

Results (as csv)
Definition of the conformers as torsion averages (csv) and their esd values (csv) or as their Cartesian coordinates of the conformers

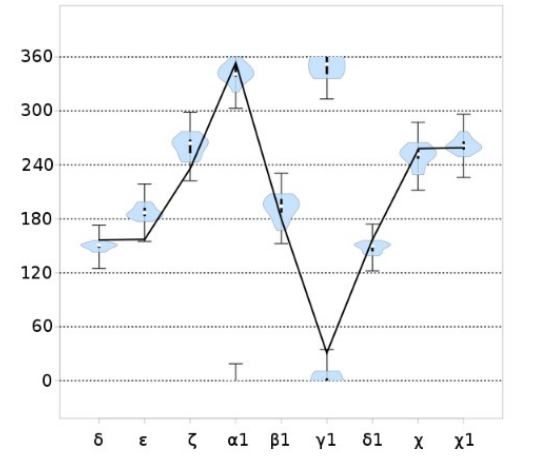
**Download the papers:**

definition of conformers:
Svozil et al., Nucleic Acids Research, 36, 3690 (2008).
description of the assignment process:
Čech et. al, BMC Bioinformatics, 14:205 (2013).



Use left click and drag to rotate, wheel to zoom, and right click to show the menu. See the J(S)mol wiki for detailed description.

BI with high α1 and γ1 near 0




| | δ | ε | ζ | α1 | β1 | γ1 | δ1 | χ | χ1 |
|---|---|---|---|---|---|---|---|---|---|
| step_torsions | 156.5 | 157.2 | 236.1 | 353.6 | 180.0 | 30.9 | 156.6 | 258.1 | 259.1 |
| ntC_average | 148.9 | 186.8 | 261.1 | 341.0 | 192.3 | 354.0 | 148.0 | 250.0 | 260.9 |
| Δ torsions | 7.6 | -29.6 | -25.0 | 12.6 | -12.3 | 36.9 | 8.6 | 8.1 | -1.8 |

# Summary

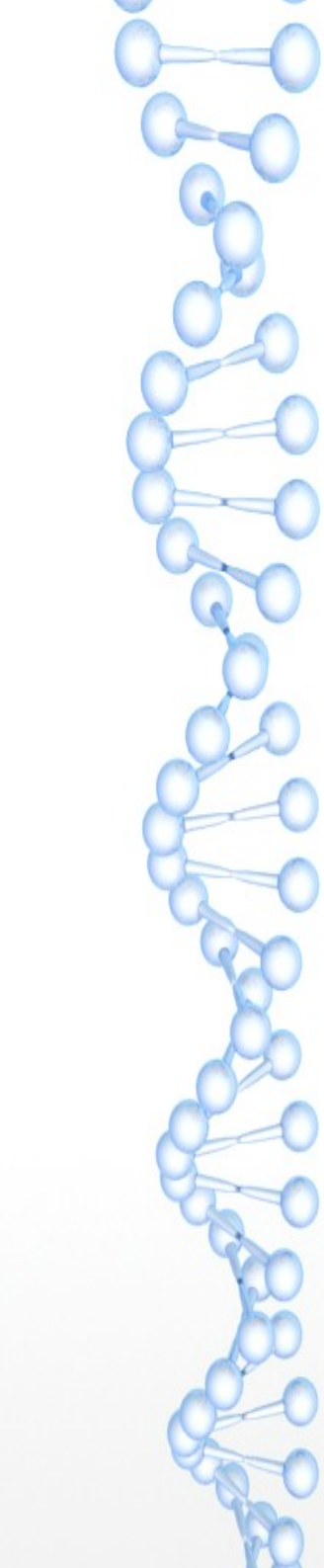
- New dinucleotide conformers
- Self-consistent golden set
- Assignment used to evaluate commonly used structures
- X-ray and NMR structures refinement (qualitative)

# Acknowledgement


- Bohdan Schneider (Laboratoř biomolekulárního rozpoznávání, Biotechnologický ústav)
- Jiří Černý (Strukturní bioinformatika proteinů, Biocev)

Collaboration

- Daniel Svozil (Vysoká škola chemicko-technologická)
- Petr Čech (Vysoká škola chemicko-technologická)



# Thank you for your attention