Refinement of Nucleic Acid Structures with NtCs
(using dinucleotide conformers at dnatco.datmos.org)

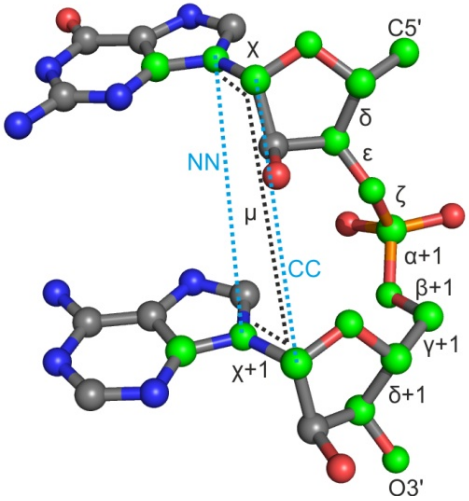
Jiří Černý

Laboratory of Structural Bioinformatics of Proteins
Institute of Biotechnology, Czech Academy of Sciences

jiri.cerny@ibt.cas.cz
ORCID, RG

3DBioInfo community

NtCs (dinucleotide conformers)



- 12D representation
- mostly backbone torsions
- 96 universal DNA/RNA conformers

https://dnatco.datmos.org

Structural alphabets (NtC and CANA) for conformational analysis of nucleic acids (DNA and RNA).

- *Annotation*
- *Validation*
- **Model building**
- **Modeling**
- Enhanced sampling MD
- **Refinement**

Refinement of experimental structures

initial 3D model (sequence → homology/*de novo* model)

X-ray: molecular replacement / experimental phasing
cryo-EM: (flexible) fitting of a model to the map

“few” cycles of refinement and model building using restraints (significant weights unless having very high resolution data)

- ideal bond lengths and bond angles (CDL)
- sugar-phosphate backbone conformation (torsions)
- pairing / stacking
- other experimental restraints

programs are using different functional forms and “blackbox” parameters

validation (and if necessary modify the model and continue)

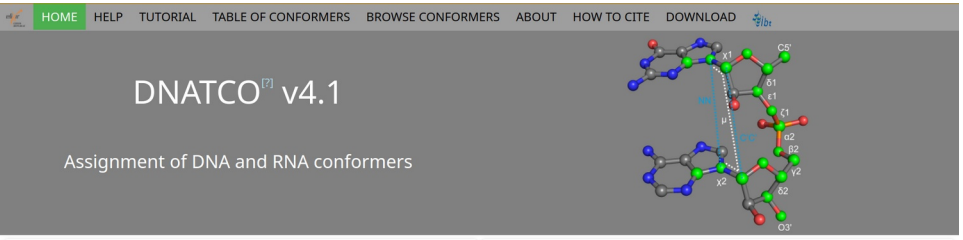
Refinement of experimental structures: tools

Coot → REFMAC/Phenix/Buster → Coot → → → → → → → PDB OneDep

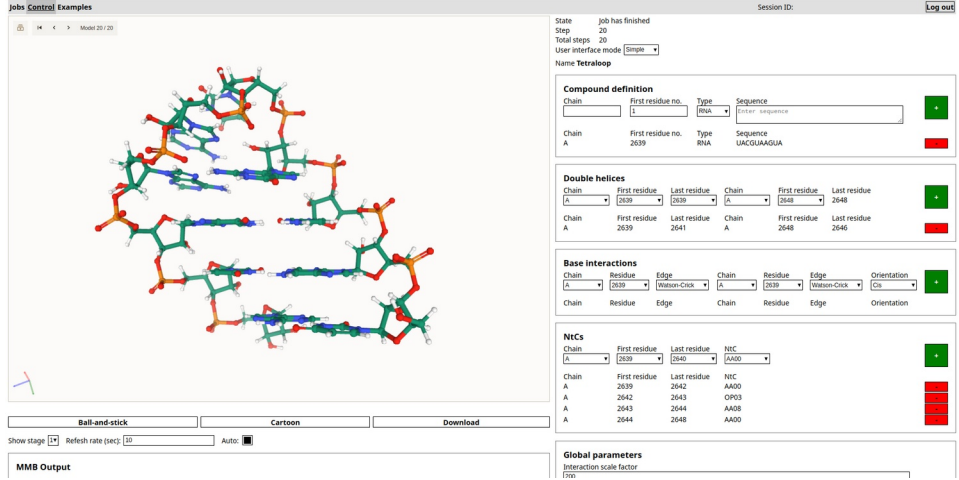
NtC restraints generated by dnatco.datmos.org are sometimes not enough to “force” the target conformation (when target is too far from initial value).

Model has to be modified to be closer to target values. Currently still mostly manual and very tedious work in Coot.

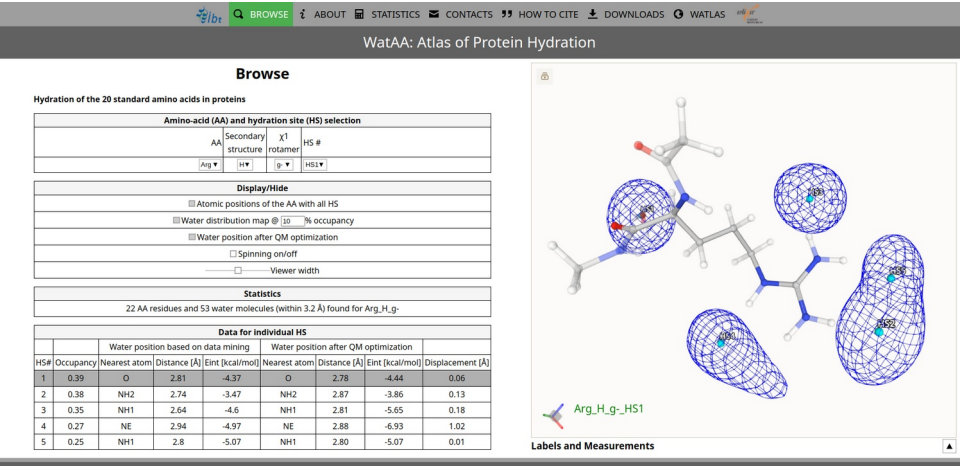
Alternatives provided by our extensions of the MMB program (MacroMoleculeBuilder, formerly RNAbuilder) including full support for modeling NtCs, mmCIF reading/writing, and CCP4/MRC map fitting. Recently we also introduced NtC modeling into Coot (but using Coot 1, so more development is needed on both sides).



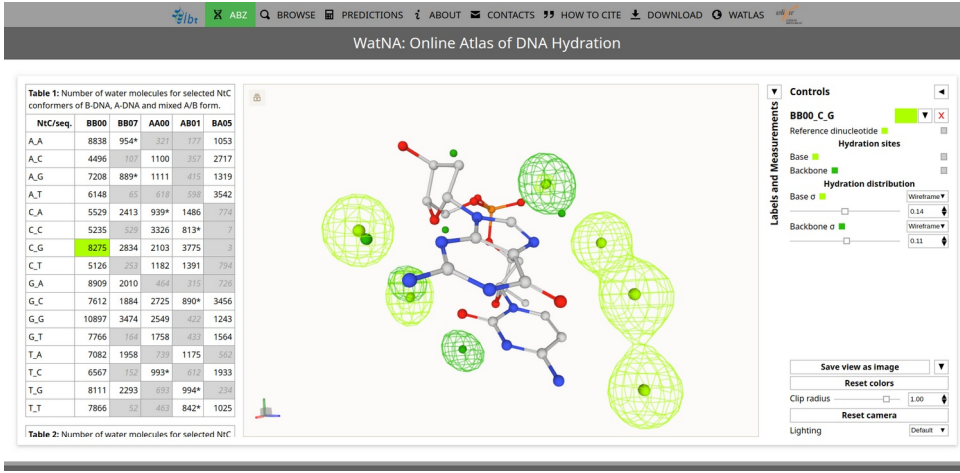
https://dnatco.datmos.org



https://webmmb.datmos.org



https://wataa.datmos.org



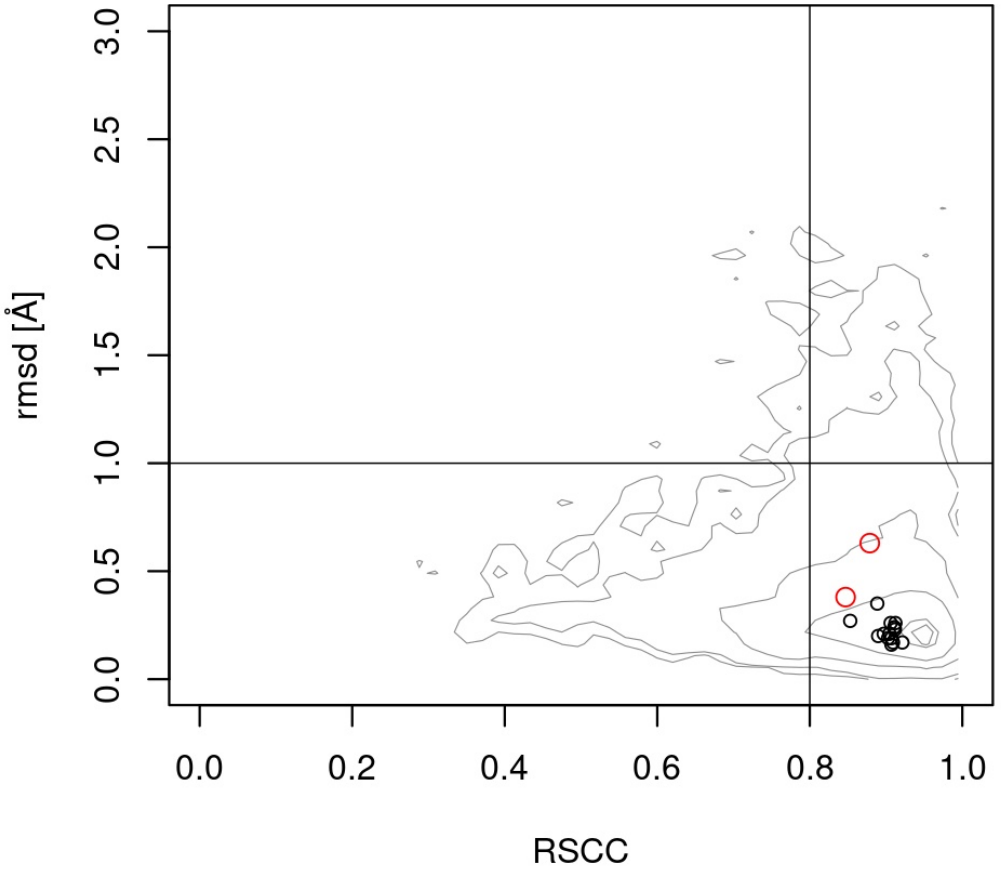
https://watna.datmos.org

DNATCO and webMMB demo

REFMAC restraints

```
# dinucleotide conformers (NtC) restraints generated by https://dnatco.datmos.org/v4.1/
# contains restraints for steps within 0.50 Å rmsd of selected NtC.
# sigma values for step 1q93_A_G_2_G_3 are scaled by 2.00
external torsion first chain A residue 2 atom C5' next chain A residue 2 atom C4' next chain A residue 2 atom C3' next chain A residue 2 atom O3' value 82.08 sigma 15.47 period 1
external torsion first chain A residue 2 atom C4' next chain A residue 2 atom C3' next chain A residue 2 atom O3' next chain A residue 3 atom P value 206.27 sigma 23.51 period 1
external torsion first chain A residue 2 atom C3' next chain A residue 2 atom O3' next chain A residue 3 atom P next chain A residue 3 atom O5' value 287.91 sigma 19.61 period 1
external torsion first chain A residue 2 atom O3' next chain A residue 3 atom P next chain A residue 3 atom O5' next chain A residue 3 atom C5' value 293.46 sigma 18.69 period 1
external torsion first chain A residue 3 atom P next chain A residue 3 atom O5' next chain A residue 3 atom C5' next chain A residue 3 atom C4' value 172.56 sigma 20.26 period 1
external torsion first chain A residue 3 atom O5' next chain A residue 3 atom C5' next chain A residue 3 atom C4' next chain A residue 3 atom C3' value 54.93 sigma 19.20 period 1
external torsion first chain A residue 3 atom C5' next chain A residue 3 atom C4' next chain A residue 3 atom C3' next chain A residue 3 atom O3' value 81.86 sigma 13.91 period 1
external torsion first chain A residue 2 atom O4' next chain A residue 2 atom C1' next chain A residue 2 atom N9 next chain A residue 2 atom C4 value 198.67 sigma 32.72 period 1
external torsion first chain A residue 3 atom O4' next chain A residue 3 atom C1' next chain A residue 3 atom N9 next chain A residue 3 atom C4 value 200.44 sigma 30.99 period 1
external distance first chain A residue 2 atom N9 second chain A residue 3 atom N9 value 4.767 sigma 0.750 type 1
external distance first chain A residue 2 atom C1'  second chain A residue 3 atom C1'  value 5.45 sigma 0.682 type 1
external torsion first chain A residue 2 atom N9 next chain A residue 2 atom C1' next chain A residue 3 atom C1' next chain A residue 3 atom N9 value 18.24 sigma 16.11 period 1
external torsion first chain A residue 2 atom C4' next chain A residue 2 atom O4' next chain A residue 2 atom C1' next chain A residue 2 atom C2' value 1.05 sigma 19.67 period 1
external torsion first chain A residue 2 atom O4' next chain A residue 2 atom C1' next chain A residue 2 atom C2' next chain A residue 2 atom C3' value 334.06 sigma 18.83 period 1
external torsion first chain A residue 2 atom C1' next chain A residue 2 atom C2' next chain A residue 2 atom C3' next chain A residue 2 atom C4' value 39.44 sigma 15.62 period 1
external torsion first chain A residue 2 atom C2' next chain A residue 2 atom C3' next chain A residue 2 atom C4' next chain A residue 2 atom O4' value 320.3 sigma 13.93 period 1
external torsion first chain A residue 2 atom C3' next chain A residue 2 atom C4' next chain A residue 2 atom O4' next chain A residue 2 atom C1' value 24.38 sigma 16.82 period 1
external torsion first chain A residue 3 atom C4' next chain A residue 3 atom O4' next chain A residue 3 atom C1' next chain A residue 3 atom C2' value 0.13 sigma 16.75 period 1
external torsion first chain A residue 3 atom O4' next chain A residue 3 atom C1' next chain A residue 3 atom C2' next chain A residue 3 atom C3' value 334.75 sigma 18.40 period 1
external torsion first chain A residue 3 atom C1' next chain A residue 3 atom C2' next chain A residue 3 atom C3' next chain A residue 3 atom C4' value 39.17 sigma 15.99 period 1
external torsion first chain A residue 3 atom C2' next chain A residue 3 atom C3' next chain A residue 3 atom C4' next chain A residue 3 atom O4' value 319.98 sigma 12.19 period 1
external torsion first chain A residue 3 atom C3' next chain A residue 3 atom C4' next chain A residue 3 atom O4' next chain A residue 3 atom C1' value 25.17 sigma 12.60 period 1
```

Additional independent validation



Conclusions

NtC restraints generated by dnatco.datmos.org used with REFMAC or Phenix provide better quality structures – lower rmsd and higher RSCC.

The MMB program now fully supports modeling of NtCs, mmCIF reading/writing, and CCP4/MRC map fitting. (next to the already available all base pairing patterns)

Soon also Coot version with NtC support will be available.

More automatic version of the workflow is under development.

Use mmCIF format in your tools/services, many existing tools can be updated easily with help of the gemmi library.