Mathematical Modelling of DNA

Prof. John Maddocks
Session 7

Version Spring 2021

A. Patelli, R. Singh

Note: Last week's exercise was long and this Series is relatively short, but you can do various bits at whatever time you like. In the exercises below it's probably easy to use cgDNAweb+ for visualising shapes, but for computation related to stiffness matrices its's probably easiest to use the cgDNA+ matlab package.

1 Understanding groundstate and Stiffness

This exercise repeats in part the points illustrated in Qu 5 of Session 6 (in particular the sequence in Qu 5.1 serie 6 and in Qu 1.1 of this serie is not a palindrome, while the sequence in Qu 5.2 serie 6 and in Qu 1.2 of this serie a palindrome) but now after the lectures on the Crick-Watson transformation between strands, and for very short sequences where it is easier to use spy to view the nonzero entries in the entire stiffness matrix.

1) Compute the ground state and stiffness matrix for the tetramers AAAA and TTTT (a non palindromic complementary pair of sequences). Visualise the two ground states, and check that you understand the relations between the two ground states. Use spy with a tolerance (see Qu 5.1 in corr 6) to understand the location of the zero entries in the corresponding stiffness matrices.

2) Now compute the ground state and stiffness matrix for the palindromic tetramer AATT, understand the symmetries in its ground state, and use spy to observe the set of zero entries in its stiffness matrix (put a tolerance of 10^{-10} while using spy).

3) Plot the ground state for the uniform sequences A_{20} and G_{20}. Observe that despite the sequences being uniform, the ground state of each is not uniform because of end effects. For each sequence the stiffness blocks are only different for the first and last junctions, as are the sigma shape parameters. Nevertheless the ground state only approaches an interior uniform value five or so base pairs from each end. Now use sequence $A_{10}G_{10}$ and note the nonlocal disruption to the groundstate at the interior boundary between the polyA and polyG segments.

2 Effect of a point mutation on the Shape and the Stiffness

In this exercise we will show how a single point mutation (change of a letter) in a given sequence will lead to a change on the reconstruction of the shape coordinates. For that consider the two following sequences:

S_1 = GTGAAAAAAAAAGC,

S_2 = GTGAAAATAAAAGC.

Using cgDNAweb+ construct the shapes for S_1 and S_2. And to construct stiffness matrices you can use cgDNA+ matlab package. Then do the following:

1) Compare the 3D structure of the groundstates. What can you say?

2) For which base–pair the change of Stagger is biggest when passing from S_1 to S_2?

3) For which base–pair is the change of Propeller is biggest when passing from S_1 to S_2?

4) Of all the junctions, where the value of Roll changes by more than 1% when passing from S_1 to S_2, which junction is closest to the end of the oligomer?

5) What do you expect to be the changes in the stiffness matrices?

3 Using cgDNAweb+ with interesting sequences

For this exercise only use of cgDNAweb+ (`http://cgdnaweb.epfl.ch`) is required. Read details in section getting started on cgDNAweb+ and understand the modelling differences between cgDNA model (parameter set PS4) and cgDNA+ model (parameter set PS+1).

3.1 Variation in shape with different parameter sets

In this exercise we will see how two different parameter sets (PS4 and PS+1) for a given sequence will lead to a change of the shape coordinates. For that consider the following sequence (phased A-tracts):

$$S = (A_5G_5T_5C_5G)_2A_5G_5T_5C_5G(A_5G_5T_5C_5G)_2.$$

Using cgDNAweb+ construct the shapes for S using parameters sets (PS4 and PS+1) and then do the following:

1) Compare the 3D structure of the ground-states. What can you say (observe strong intrinsic bends)?

3) What can you say for internal coordinates? Use 2D plot in cgDNAweb+ and observe large negative propeller for this sequence.

3.2 Visualising left-handed and right-handed superhelices

In this exercise we will see example sequences for which groundstate form superhelices that are left-handed and right-handed. For that consider the two following sequences:

$S_1 = (A_5CACG_2)_{11}$,

$S_2 = (A_7GAG_2)_{10}$.

Using cgDNAweb+ construct the shapes for S_1 and S_2. Then compare the 3D structure for both sequences and observe that groundstate corresponding to sequence S_1 is left-handed helix and groundstate corresponding to sequence S_2 form right-handed helix.

4 cgDNA+ reconstruction of the 6 distinct poly–dinucleotide

First set $D = \{AT, GC, AA, AC, AG, GG\}$ a set of 6 independent dimers, using the dimer steps in D we can define all 6 distinct poly–dinucleotide sequences (denoted by $\text{poly}(\alpha\beta)_N$) that are the sequences obtained by N repetitions of the dimer $\alpha\beta \in D$.

Using the cgDNA+ matlab package reconstruct the groundstate of the 6 $\text{poly}(\alpha\beta)_{150}$, $\alpha\beta \in D$. Then:

1. Visualize the entries of the groundstates coordinates and compare them. Note that the ground-state intra vary very considerably between the six sequences.

2. Visualize the 3D reconstruction of all the groundstates and compare (in the eye–ball metric) the radius and the pitch of the resulting helical structures. Here you can actually just plot the xyz coordinates of all the base-pair positions.

5 Downloading and compiling the cgDNApmc code

Download cgDNApmc code from `https://lcvmwww.epfl.ch/teaching/modelling_dna/protected_files/codes_exercises/cgDNApmc_code.zip` and compile it on your machine. Follow the on–line documentation (`https://lcvmwww.epfl.ch/software/cgDNAmc/doc/index.html`) and report anything unclear. Next week we will do Monte Carlo computations with the cgDNA+ model which will need installed version of cgDNApmc.