

Problem 1

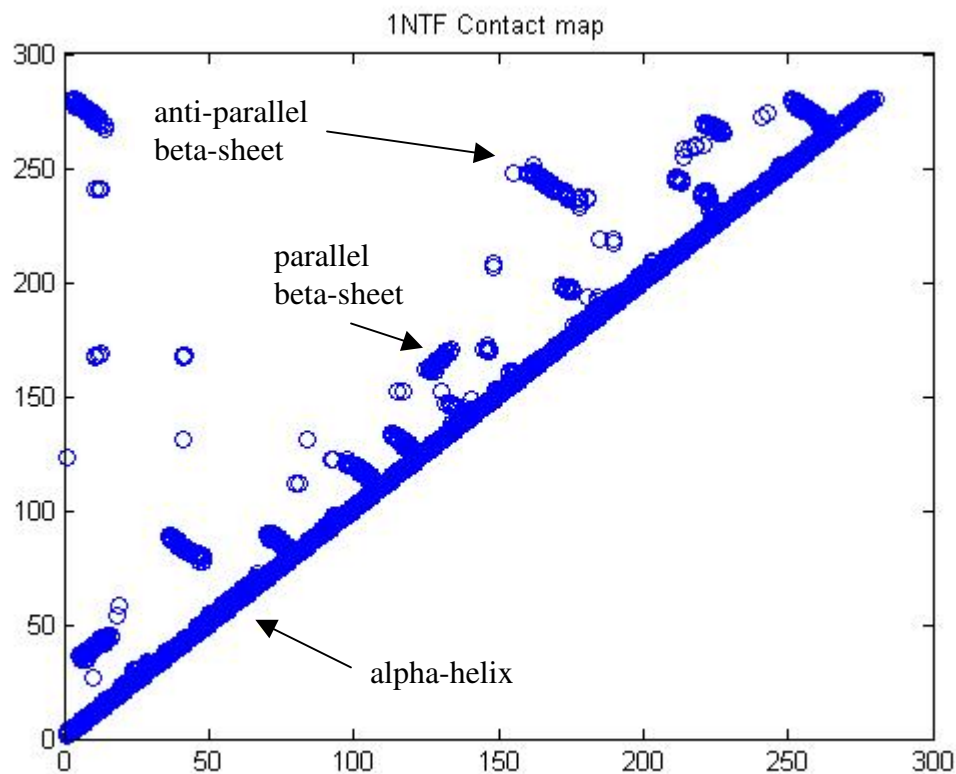
The file 1NTF.ca holds the coordinates of the C α atoms for the protein Cimex Nitrophorin, which is a protein in the saliva of the blood sucking bed bug (for real!).

- Compute and plot the contact map for this protein and comment about secondary structures you identify in the contact map.
- Go to the PDB website (<http://www.rcsb.org/pdb/>) and view the structure of this protein (the PDB ID for it is 1NTF). Compare your results from part a to the structure you see.

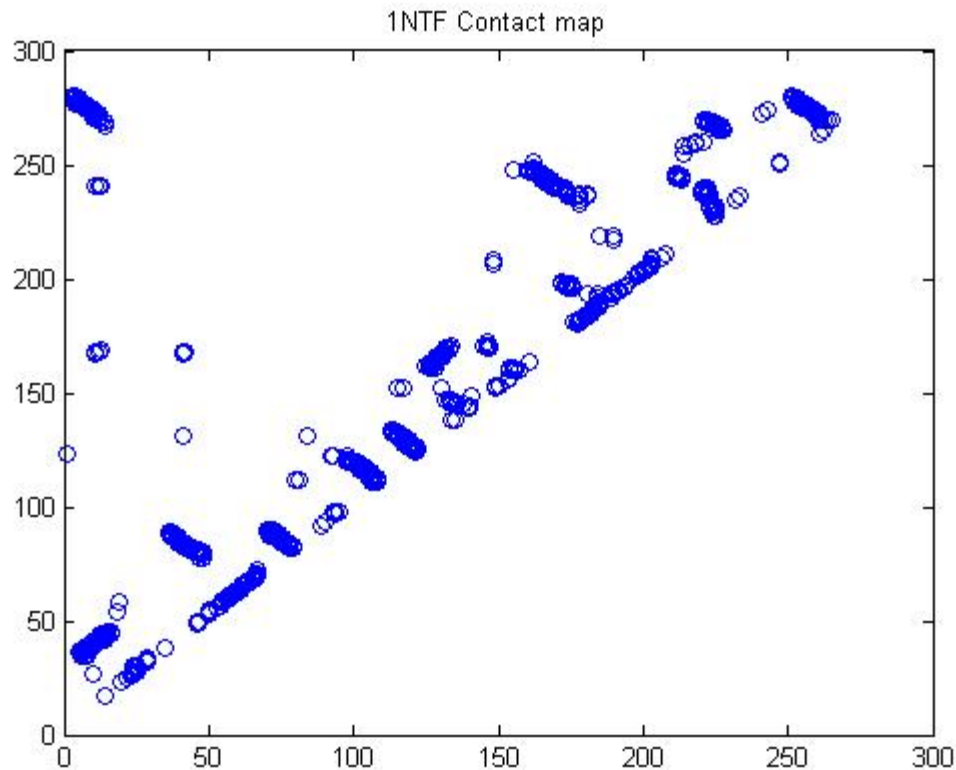
Solution 1

- Below is a plot of the contact map.

A representative of each type of secondary structure is noted. A diagonal parallel to the main diagonal indicates a parallel beta-sheet; A diagonal perpendicular to the main diagonal indicates an anti-parallel beta-sheet; A thickening of the main diagonal indicates an alpha-helix (see detailed picture).



To accentuate the helices below is a plot without the main diagonal and the two diagonals adjacent to it. We can recognize about 3 main helices and maybe another 'weak' one. We can also easily recognize all 11 strands of beta-sheets. There are also other contacts in the map, but there is little we can say about those. It pretty much agrees with the material on PDB, except for the first helix we got at about residues 10 to 40 (the lowest diagonal in the contact map) that doesn't appear on PDB.



Problem 2

In CASP5 held at December 2002 many groups tried to predict the structure of the protein Cimex Nitrophorin.

The files T0142TS464.ca T0142TS419.ca T0142TS427.ca hold the coordinates of the C α atoms of 3 of those predictions.

- Compare each of these predictions with the experimental structure from Problem 1, using RMS distance.
- Order these 3 predicted structures from best to worse.

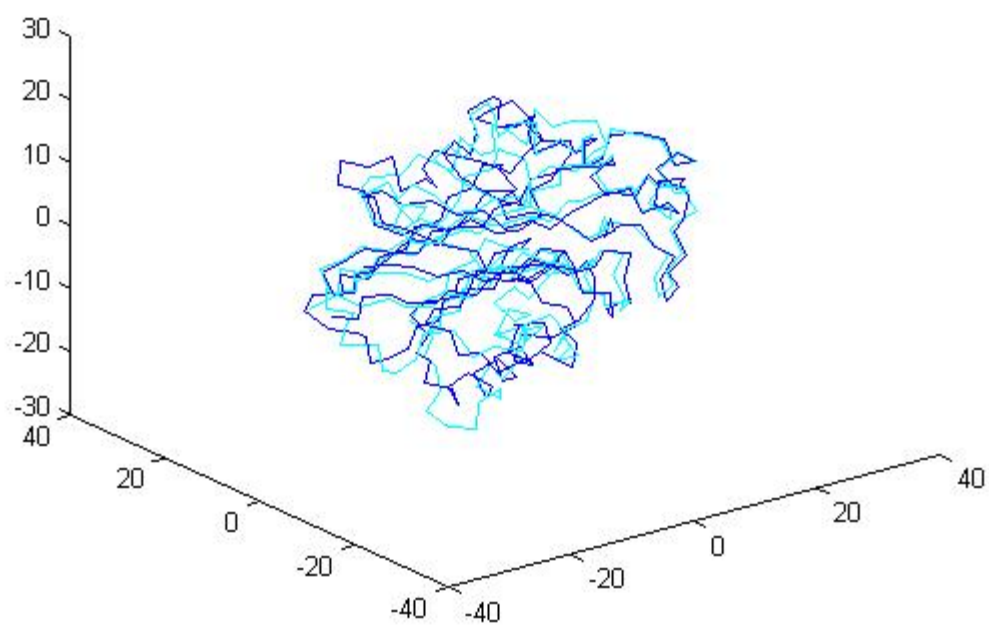
Solution 1

a) Structures T0142TS419_1 and 1NTF yielded an RMSD of 12.39 Angstrom;
 Structures T0142TS427_1 and 1NTF yielded an RMSD of 3.44 Angstrom;
 Structures T0142TS464_1 and 1NTF yielded an RMSD of 26.31 Angstrom;

b) T0142TS427_1 is clearly the best prediction to 1NTF. T0142TS419_1 is a pretty bad prediction (RMSD > 10) of 1NTF, bad enough to be considered meaningless. T0142TS464_1 is clearly the worst prediction of the three, with a score so bad that a random protein of the same length should get on average a better RMSD!

Below are simple plots of the superposition of the best and worst predictions (cyan) over 1NTF (dark blue).

1NTF vs. T0142TS427₁ rms=3.4403



1NTF vs. T0142TS464₁ rms=26.318

