Summary of works

Biological events are controlled through flow of information in a structural manner. DNA, the information store house, is used as a template to produce messages in the form of messenger RNA that is translated into proteins that control both the structural as well as enzymatic pathways in living organisms.

It has been established that there is a strong structure-function correlation and that it depends on the sequence specificity to a large extent. Many biological processes are initiated by the supercoiling energy and the negative superhelicity often leads to structural transitions in which melting plays a crucial role. Our aim is (i) to find out the sequence dependent structural flexibility of various forms of DNA and its causes, and (ii) to understand the melting transition of supercoiled DNA.

This thesis comprises of four chapters. In the introductory chapter (Ch. 1) the structural polymorphism of DNA has been discussed in general to provide the background knowledge for the problem in hand. It is now accepted that DNA sequence significantly affects the structure and flexibility of various forms of DNA through control of energy cost of its conformational change.

In chapter 2, the DNA has been considered as a worm like chain and the persistence length (P) and torsional rigidity (C) (two measuring parameters of the flexibility) have been calculated by analysing crystal structure database using 'NUPARM', developed by Prof. Manju Bansal's group. Following this, the flexibilities of various homo- and co-polymeric sequences of various forms of DNA have been discussed. The value of the parameters have been found to be in good agreement with those estimated by others. Persistence lengths for the homopolymeric sequences, namely poly(dA).poly(dT) and poly(dG).poly(dC), are significantly large compared to mixed B-DNA sequence resulting in the inability of these sequences to form nucleosome under normal conditions. The heteropolymeric sequences poly(dA-dC).poly(dG-dT) and poly(dG-dC).poly(dG-dC), on the other hand have smaller persistence lengths. It has been shown that genetic disease forming triplet repeat sequences (d(CTG).d(CAG) and d(CGG).d(CCG)) are more flexible than normal B-DNA sequence and hence, wrap around the histone octamer efficiently and modulate the gene expression. We have carried out model building studies using NUCGEN and HBFIND programme to find out the possibility of any cross
strand hydrogen bonds. The rigidity of any given DNA sequence is often controlled by its ability to form cross strand bifurcated hydrogen bonds between the successive base pairs. We have also carried out molecular orbital calculation using MOPAC to determine the conformational energy parameters. The results of P and C values of non B-form DNA show that they are less sequence dependent than B-form.

Nucleosome is now considered to be the structural framework within which replication and transcriptional machinery operates most effectively to regulate gene duplication and expression. The histone octamers can be associated with an immense variety of DNA sequences which very often adopt well defined locations with respect to the octamers. In chapter 3, the persistence length data has been used to determine the nucleosomal positioning pattern. Our theoretical prediction has been compared with the experimental data available. The robustness of the algorithm developed has been shown to be useful for the analysis of nucleosomal positioning. This methodology has been further extended to show that flexibility profiles of mouse and Sachharomyces are closer to Methanosarcina barkeri and Methanococcus jannaschii than the two prokaryotes - E.coli and Paracoccus. This persistence length data has been calculated over long stretches of nearly 5000 base pairs in the samples used. Archaea was found to resemble eukarya more closely than bacteria and hence reinforces the view that at the evolutionary level from protogene, bacteria branched off earliest.

DNA in living organisms almost always exists in supercoiled form and the excess free energy of supercoiled DNA facilitates several biochemical processes, like transcription, replication, recombination and the binding of unwinding pattern. The calculation of partition function of a supercoiled DNA undergoing conformational transition has been the standard method for analysing the structural basis for major biological events like replication, recombination, etc. In 4th Chapter, a theoretical investigation of the denaturation characteristics of a supercoiled DNA has been presented employing a Metropolis Monte Carlo algorithm to examine the overall melting profiles of a supercoiled DNA at various temperature and pH. This simpler method can explain almost all the overall denaturation characteristics and it also correctly calculates the detailed denaturation probabilities of each base pair at various degrees of supercoiling. The interesting result is the plasmid pBR322 has very pronounced peaks between 4150 and
4350 for native levels of supercoiling. These are the terminator and promoter regions of the beta-lactamase gene. Similar analysis with \( \phi X174 \) showed multiple regions of high peaks. These overlap all the gene encoded in the DNA sequence. It would be interesting to note that our method is sensitive enough to identify the transcribing regions from any segment of DNA through a melting probability calculation. Our method, using the Monte Carlo algorithm, can qualitatively reproduce the denaturation profiles; however the agreement with experiment in the case of alkali induced denaturation is not as good as in the case of thermal denaturation. The alkali denatured supercoiled DNA is an interesting form which has been shown to resemble structure involved in replication initiation complex. Sources of discrepancy between theory and experiments have also been discussed.

DNA double helical form is mainly stabilised by stacking energy of the successive base pair. The number of data points of each of the ensemble of local doublet parameters of each of the ten dinucleotides are not enough. We have calculated how stacking energies change with the changes of the values of the local doublet parameters. Since energy of a state reflects the population of that point, one can generate the histogram of the population of a particular doublet for a local doublet parameter. We have also compared these values with that derived from X-ray crystallographic database. All the previous calculations discussed in Chapter 2 are based on the assumptions that the number of dinucleotides reduces from sixteen to ten because AA, AC, CA, CT, TC, CC are equivalent to TT, GT, TG, AG, GA, GG respectively. Considering all of the sixteen dinucleotides separately, the results have been discussed. We have also calculated the correlation of the local doublet parameter of each of the ten dinucleotides. It has been shown that this correlation is also sequence specific.
Manuscript Acceptance Form

Journal: Biophysical Chemistry (BIOCHE)
Format: Paper
Type of Paper: Normal
Full Title: Denaturation of superboilded DNA: A Monte Carlo Study
Accepting Editor: Professor G. Schwarz
Editor's Ms. No. GS 382
Corresponding Author: Professor Ashoke Ranjan Thakur
Departement of Biophysics, Molecular Biology and Genetica
University of Calcutta, 92 Acharya Prafulla Chandra Road
Calcutta 700 009, India
Tel: ++91/350-8386/6387/6396/1397 351-0359
Fax: ++91-33-351-0360
E-mail: arthakur@cubmb.ernet.ind
Received Date: 03.03.1998
Revised Rec. Date: 26.08.1998
Accepted Date: 10.09.1998