BE.342/442 Thursday, November 10, 2005
Topic: Molecular Structure and Self-Assembly of DNA and RNA

Administrative:

We had no class yesterday – hope everyone enjoyed the movie.

A special surprise on the course website: a **signed** paper by Francis Crick!

Next Tuesday: guest lecturer.

Next Thursday: lecture on DNA as a molecular machine (and the end of this lecture). Mid-terms due!

DNA as a Material

Is DNA a structural material? Using DNA as a material is almost like burning books to keep warm in the winter: in biology, DNA is used for information storage.

Ned Seeman, a former postdoc here at MIT along with Alexander Rich, spent his time at NYU "tinkering" with DNA as if it were a toy. Last year, he wrote a cover story for *Scientific American* on "DNA Oragami" – methods of linking DNA to form materials, just as we use proteins, lipids, and carbohydrates.

The technology for manipulating DNA as a material has advanced far faster than for proteins, lipids, and carbohydrates – why? After all, DNA is far less abundant in nature, and at one time was difficult to obtain. When Prof. Zhang was doing research, DNA cost $25 per base. Today, it cost close to 10 cents! That's a 250x drop in price! That's because DNA is designed to be replicated, and can be replicated in a machine.

Meanwhile, the biotech sector has exploded -- AmGene, founded in 1978, is now worth 30 billion dollars!

DNA History

The Central Dogma of Biology: DNA replication, transcription into RNA, and translation of RNA into protein. Although reverse transcriptase can create DNA from RNA, today there exists no way to "reverse translate" a protein.

In the 1950's there were only a handful of institutions studying molecular biology in the United States. One of the few researchers was Max Delbrück. Some major players in the development of molecular biology were:

> Francis Crick
> Leslie Orgel
> Alexander Rich
> James Watson
> Sydney Brenner
> Max Perutz

From yesterday's movie, we learned how the structure of DNA was "guessed" from the suggestion of a helical structure with a rise of 0.32 nm gleaned from Rosalind Franklin's x-ray patterns. The first atomic resolution of nucleotide base pairings was discovered through an electron density map produced by Alexander Rich and colleagues in 1974.

Base Pairing

DNA and RNA differ only by a single –OH group on the second carbon of the 5-membered ring forming the ribose sugar.

Purines: Adenine (A), Guanine (G)
Purines contain one six-membered ring linked along a side to a five-membered ring.

Pyrimidines: Cytosine (C), Thymine (T), Uracil (U)
Pyrimidines contain a single six-membered rings.

See figures in the textbook to understand the numbering system for purines and purimidines.

Both of these bases can assume a keto form and an enol form. In the movie, we saw that Jerry Donahue pointed out to James Watson that the bases are more stable in the keto form. This clue was tremendous, because the tautomeric form determines the spacing between the base pairs, so, as we saw, this helped Watson and Crick solve the structure.

The A-T pair forms 2 H-bonds: one donor comes from adenine, and the other comes from thymine.

The C-G pair forms 3 H-bonds, with guanine donating one H-bond and cytosine donating 2 H-bonds.

There are a number of possible hydrogen bonds in the base pairing structure: you will need to map all of them as part of your take-home midterm.

Triplets of base pairing are possible such as T-A-T Hoogsteen type base pairing, which use the mechanism of Watson-Crick pairing.

 "Non-Watson-Crick" base pairings include G-U, the pairing in RNA where U replaces T, and several "unusual" base pairings. These include: G-U parallel pairing (as opposed to the antiparallel pairing seen in DNA), parallel A-T pairing, triplets of G-G-C, A-A-U, and G-A pairing. Although these are rarely seen in biology, these pairings are useful for applications of DNA as a material.

Other Structural Features

Sugar puckering is an important feature that affects phosphodiester bridge orientation and glycosidic bonding in DNA. The $C_{2'}$-*endo* puckering is found in B-DNA, whereas $C_{3'}$-*endo* puckering is found in A-DNA. (Recall the 2 different DNA structures Rosalind Franklin found in

the movie.) B-DNA is predominant in the genome, because it is more stable. In the B-form, all the bases occur in the center, and the 2 strands are helically aligned with a major groove and a minor groove.
Major groove: position 5, 6, 7, 8 for purines; position 4, 5, 6 for pyrimidines.
Minor groove: position 1, 2, 3, 4 for purines; position 1, 2, 3 for pyrimidines

DNA concentration measurements are based on UV absorption at 260 nm, although the UV absorption spectra differ slightly for each base. The peaks of UV absorption for the bases occur at:

Base	Peak absorpiton
CMP	250 nm
AMP	258 nm
UMP	260 nm
CMP	275 nm

Syn vs *anti* configuration:
In the *syn* configuration, the base faces toward the sugar ring, causing crowding that is less stable than in the *anti* confirguration.

Structural comparison of A-, B-, and Z-DNA:

B-DNA is the "tightest" and most stable structure, with a helix diameter of 23.7Å, a rise of 3.4 Å per base pair, a pitch of 35.4 Å per turn, a tilt of 1 degree, and 10.4 base pairs per turn. All bases assume the *anti* configuration. The major groove is wide and deep, and the minor drove is narrow and shallow.

A-DNA has a helix diameter of 25.5Å, a rise of 2.3 Å per base pair, a pitch of 25.3 Å per turn, a tilt of 19 degrees, and 11 base pairs per turn. All bases assume the *anti* configuration. The major groove is narrow and deep, and the minor drove is broad and shallow.

Z-DNA is an unusual structure, only found in the presence of specific bases. Its structure was solved in 1979. It has a helix diameter of 18.4Å, a rise of 3.8 Å per base pair, a pitch of 45.6 Å per turn, a tilt of 9 degrees, and 12 base pairs per turn. C and T assume the *anti* configuration, but G assumes the *syn* configuration. The major groove is flat, and the minor groove is narrow and deep.

The **G4-Quartet** structure was not solved until 1992 (by Alexander Rich). In this structure, four guanines form complementary hydrogen bonds, serving as "bookends" for chromosomes called telomeres.

The **B-Z Junction** was recently solved, published in *Nature* on October 20, 2005. The structure flips out, and all the neighboring bases must accommodate the flipping. This junction is important in gene promoters, which contain $(CG)_n$ and $(CATG)_n$ repeating sequences. As example of a gene with such a promoter is IGF: Insulin-like Growth Factor, necessary for every cell. The flipping of a single base pair at the junction was never before understood!

3-stranded **H-DNA** is another unusual DNA structure that only forms at low pH or negative supercoiling conditions. It can be used to block DNA synthesis and gene transcription, and has interesting materials properties that have been used to modify DNA nanostructure.

Additionally, DNA and RNA can fold on itself, forming **stem-loops**, **hairpins**, and **pseudoknots**. Pseudoknots are used to stop protein translation, and are being used in therapeutic applications.

DNA Packing

Every cell (except sperm and eggs) contains 2 meters of DNA!
DNA tomoisomers are identical in size and sequences, but have different folding.

Bacteria and viruses fit in all that DNA through negative supercoiling.
We do it through histones, which won't be covered here.

Supercoiling is much like rolling a thread to twist the strands together tightly. (Demonstration:) Try twisting a rubber band while stretching it out. Then relax the rubber band. What happens?

DNA can assume Type I, II, and III coiled confirmations. Type II includes supercoiled structure, as well as strained circles and unstrained circles.

DNA Recombination

The Holliday Junction for a 4-stranded DNA structure shows how two DNA double-helices come together. This structure was not solved by Alexander Rich… but by his postdoc!

(This lecture will be continued next Thursday.)