1. THE DNA DOUBLE HELIX

Desoxyribo-nucleic acid (DNA) is the basic building template for life. It codes for the synthesis of proteins inside living cells by controlling the amino acid sequences that form the genes. DNA is formed of backbone phosphates, linked to desoxyribose sugars and side group amine bases. The charged phosphate groups are hydrophilic, the desoxyribose sugar groups are mostly hydrophobic, and the amine bases contain both hydrophobic and hydrophilic groups. DNA forms a helical structure which is stable because of the stacking of the amine bases and of the hydrogen bonding between them. The helical structure is effective at “hiding” the hydrophobic sugar groups from contact with water. The helix phase melts into disordered coils under various conditions including heating.

Figure 1: Schematic representation of the DNA helix and coil conformations

DNA macromolecules form helical structures in their active form and melt to a random coil phase in their denatured form. The denaturation transition consists in a helix-to-coil transition that can be promoted either using denaturing agents or through heating. Here, the simpler heating route is discussed using two characterization methods: the UV absorption spectroscopy and SANS.
2. UV ABSORPTION SPECTROSCOPY

Ultra-violet (UV) light absorption spectroscopy is sensitive to the stacking of \( \pi \)-bonded groups such as the amine bases in DNA. It is an effective method to monitor the helix-to-coil transition.

![UV Absorption Spectroscopy](image)

Figure 2: Typical UV absorption spectrum from DNA showing a characteristic peak at a wavelength of 260 nm.

The monitoring of the intensity at the peak position (260 nm) with increasing temperature yields a sigmoid function centered at the helix-to-coil transition temperature. The helix-to-coil transition is mediated by the un-stacking of the amine bases and the breaking of hydrogen bonds between these bases.

A sample containing 4 % salmon DNA (molecular weight of \( 1.3 \times 10^6 \) g/mol) in water is characterized by a helix-to-coil transition temperature of 94 °C. Since this temperature is too close to the boiling temperature for water (100 °C), another solvent was considered as well. A 4 % DNA in ethylene glycol is characterized by a more convenient helix-to-coil transition temperature of 38 °C. The characteristic sigmoid shape function was obtained in each case. The inflection point corresponds to the helix-to-coil melting temperature. In order to avoid saturation of the UV absorbance, 50 \( \mu \)m thin samples were measured.
Since the same DNA samples were investigated by UV absorption spectroscopy and SANS, deuterated solvents were used with both characterization methods. Moreover a 100 mM NaCl salt content was added throughout in order to screen the charges on the DNA phosphate groups.

![Graph showing UV absorption spectroscopy peak intensities at 260 nm with increasing temperature for 4% DNA in d-water or in d-ethylene glycol. The helix-to-coil transitions can be observed clearly.]

**3. HELIX-TO-COIL TRANSITION IN MIXED SOLVENTS**

The UV absorption spectroscopy provides an effective way to monitor the helix-to-coil transition in d-water/d-ethylene glycol mixed solvents. The same 4% salmon DNA weight fraction and 100 mM NaCl salt content were used.
Figure 4: Variation of the helix-to-coil transition temperature for 4 % DNA in mixed d-water/d-ethylene glycol mixed solvents.

The monotonic linear variation is attributed to the fact that the transition was approached from the helix side whereby solvents mix randomly (ideal solvent mixing behavior). The fact that the melting temperature decreases with d-ethylene glycol fraction points to the conclusion that the hydrophobic groups CD$_2$ in d-ethylene glycol play an important role in the melting transition. They help solvent molecules cross the hydrophobic zone of the desoxyribose sugar groups thereby loosening the helical structure. This argument helps understand the micellar nature of the DNA macromolecules in terms of a hydrophobic sugar region and hydrophilic phosphate and amine base regions. The ability of the solvent to cross the hydrophobic region controls the stability of the helix phase.
4. **THE HELIX-TO-COIL TRANSITION BY SANS**

The SANS technique is effective at determining macromolecular structures. A series of measurements were performed from a 4% DNA/d-ethylene glycol/100 mM NaCl sample at temperatures ranging from 10 °C to 80 °C (at 5 °C intervals). A figure shows a typical SANS spectrum at two temperatures; one below (25 °C) and the other one above (50 °C) the helix-to-coil transition temperature. This temperature is known to be 38 °C based on US absorption measurements (Hammouda-Worcester, 2006).
Figure 6: SANS from a 4 % mass fraction DNA/d-ethylene glycol/100 mM NaCl sample measured at temperatures below (25 °C) and above (50 °C) the helix-to-coil melting temperature.

The high-Q signal is distinctively different in the two cases. The data show an abrupt decrease in the high-Q intensity for the helix phase but a gradual decrease for the coil phase.

The SANS data were fit to the following empirical functional form that reproduces the main data features:

\[
I(Q) = \frac{A}{Q^n} + \frac{C}{1 + (Q\xi)^m} + B. \tag{1}
\]

The term \(A/Q^n\) represents the low-Q clustering (network) scattering part and the term \(C/[1+(Q\xi)^m]\) represents the high-Q solvation part. B represents a Q-independent (mostly incoherent) background. The low-Q part represents scattering from a large gel network structure. It does not change much across the melting transition. Our focus here is on the high-Q signal exclusively.
A figure shows the variation of the “solvation intensity” (the fitted quantity C) for increasing temperature. The intensity drop between 25 °C and 40 °C characterizes the helix melting transition. Lowering temperature shows that this transition is weakly reversible with substantial hysteresis. Further temperature increase beyond the melting transition increases the solvation intensity. This result is typical of water-soluble polymers which are characterized by a Lower Critical Solution Temperature (LCST).

![Figure 7: Variation of the SANS solvation intensity](image)

The correlation length $\xi$ also varies across the melting transition. This correlation length represents a weighted-average inter-distance between the hydrogen-containing (sugar-amine base) groups. It is around 8.5 Å in the helix phase and increases to 12.3 Å in the coil phase. In the helix phase the sugar-amine base groups are closer together than in the coil phase. This increase in $\xi$ is due to the opening of the tight helical structure into a loser coil configuration. This correlation length is not a measure of the DNA radius. Raising the temperature further in the coil phase increases the correlation length even further; this is a familiar trend for LCST systems.
Figure 8: Variation of the correlation length (the quantity $\xi$ in the empirical model) for increasing temperature. After melting, DNA coils swell with further temperature increase.

Finally the high-Q Porod exponent $m$ is seen to vary between values around 3.7 in the helix phase to values close to 1.7 in the coil phase. DNA helices are appearing like cylinders with fairly smooth surfaces (Porod exponents close to 4) and DNA coils behave like polymer chains in good solvent conditions or in a fully swollen chain configuration (Porod exponent of $5/3 = 1.67$).
Figure 9: Variation of the high-Q Porod exponent $m$ for increasing temperature. This exponent varies from 3.7 (cylinder) to 1.7 (swollen coil).

It is noted that the rod-like nature of DNA (Porod exponent around 1) has not been seen due to the clustering signal overwhelming the low-Q scattering. It is also noted that once the melting transition has taken place, DNA coils behave like typical water-soluble synthetic polymer chains.

5. A HELIX-TO-COIL TRANSITION MODEL

Helix-to-coil transition models have been published by many authors including Zimm (Zimm, 1959). The formulation from Flory’s book (Flory, 1969) will be followed closely here.

Consider a macromolecule consisting of $N$ units (think residues) comprising $v$ helical sequences. There is a total of $N_H$ helical units and $N_C = N - N_H$ coil units. Define the partition function for the melting of one helical unit as $s = \exp(\Delta H_m/RT)$ where $\Delta H_m$ is
the enthalpy needed, \( R \) is the molar gas constant (related to the Boltzman constant \( k_B \) through Avogadro’s number \( N_A \) as \( R = k_B N_A \)) and \( T \) is the temperature in absolute units. Assume that it takes no enthalpy to form a coil so that the partition function for a coil unit is equal to 1. Define the partition function for the removal of one helical sequence as \( \sigma \). The partition function for the helix-to-coil melting process is:

\[
Z = \sum \prod s^{NH} \sigma^\nu.
\]

(2)

The product \( \prod \) is taken over all helical units \( N_H \) and all helical sequences \( \nu \) and the summation \( \sum \) is taken over all possible configurations (i.e., over all possible unit arrangements to form the macromolecule with \( N \) units).

Figure 10: Schematic representation of the helical and coil sequences.

The partition function can be expressed in matrix notation as:

\[
Z = J^* U^N J.
\]

(3)

With:

\[
J^* = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & s\sigma \\ 1 & s \end{bmatrix}, \quad J = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.
\]

(4)

\( J^* \) means that the macromolecule starts with a coil unit and \( J \) means that it finishes with either a coil or a helical unit. In order to perform the \( U^N \) product, the configuration matrix
U is diagonalized into the form $U = \Lambda \Lambda^{-1}$ where $\Lambda$ is a diagonal matrix. $U^N$ simplifies as $U^N = \Lambda \Lambda^N \Lambda^{-1}$. The eigenvalues $\lambda_1$ and $\lambda_2$ (diagonal elements of matrix $\Lambda$) are given by:

$$
\lambda_1 = \frac{(1 + s) + \sqrt{(1 - s)^2 + 4\sigma s}}{2} \quad \lambda_2 = \frac{(1 + s) - \sqrt{(1 - s)^2 + 4\sigma s}}{2}.
$$

(5)

The partition function can then be summed up to become:

$$
Z = \frac{(1 - \lambda_2)}{(\lambda_1 - \lambda_2)} \lambda_1^{N+1} + \frac{(\lambda_1 - 1)}{(\lambda_1 - \lambda_2)} \lambda_2^{N+1}.
$$

(6)

The fraction of units in the helical state is given by:

$$
p_H = \frac{1}{N} \frac{\partial \ln(Z)}{\partial \ln(s)}.
$$

(7)

The fraction of units in the coil state is $p_C = 1 - p_H$. In the case of long macromolecules ($N >> 1/2 \sqrt{\sigma}$), $p_H$ simplifies to:

$$
p_H = \frac{\partial \ln(\lambda_1)}{\partial \ln(s)} = \frac{(\lambda_1 - 1)}{(\lambda_1 - \lambda_2)}.
$$

(8)

The relative number of helical sequences is given by:

$$
p_\sigma = \frac{\partial \ln(\lambda_1)}{\partial \ln(\sigma)} = \frac{(\lambda_1 - 1)(1 - \lambda_2)}{\lambda_1(\lambda_1 - \lambda_2)}.
$$

(9)

The average number of helical units per helical sequence $y_H$ is given by the ratio:

$$
y_H = \frac{p_H}{p_\sigma} = \frac{\lambda_1}{1 - \lambda_2}.
$$

(10)

In the notation used here, the total number of helical units is $N_H = N p_H$ and the number of helical sequences is $\nu = N p_\sigma$.

The meaning of the various parameters is discussed here. First what is the meaning of parameter $s$? The helix-to-coil transition is driven by heating of the sample, i.e., by the enthalpy needed to melt one unit $\Delta H_m$. From the definition of $s$, one can express the deviation from the melting temperature $T_m$ as:

$$
T_m - T = \frac{RTT_m}{\Delta H_m} (s-1).
$$

(11)
This simple relation is obtained by expanding the exponential in the definition of s. The dimensionless variable s can be used instead of T. How to understand the meaning of parameter $\sigma$? Note that at $s = 1$ (middle of the helix-to-coil melting transition corresponding to temperature $T_m$) the preceding results simplify to:

$$p_H = \frac{1}{2}, \quad p_\sigma \cong \frac{\sqrt{\sigma}}{2}, \quad y_H \cong \frac{1}{\sqrt{\sigma}}. \quad (12)$$

Right at the transition point (i.e., at $T = T_m$), $\nu = N \sqrt{\sigma}/2$ represents the number of helical sequences. One can think of $\sigma$ as a helical sequence “nucleation” parameter. The helix-to-coil transition happens through either a few or many helix-to-coil sequences (think “nucleation centers”) depending on the temperature conditions. This transition is similar to the melting transition of crystalline materials that happens through nucleation centers.

Note that this simple model applies to the simplest form of helix-to-coil transition involving single-strand helical structures and “isolated” macromolecules. This would apply to very dilute polypeptide sequences. The case of double-strand helical structures (as in DNA) would involve larger configuration matrices U with different enthalpies for the various processes (hydrogen bonding and base stacking for AT or GC pairs). The results described here are approximate but still give useful predictions for the DNA case. They are applied to some of our UV absorption data.

6. APPLICATION OF THE MODEL TO UV ABSORPTION DATA

Consider the helix-to-coil transition UV absorption data obtained for 4 % DNA/100 mM NaCl in d-ethylene glycol. In order to apply the simple model described above, two rescalings of the UV data have are performed: (1) modification of the horizontal temperature axis into the variable s axis using the relationship between T and s given in the previous section and (2) rescaling of the UV data vertical axis to a variation between 0 and 1. Moreover the values $R = 1.989$ cal/mol.K and 1 cal = 4.18 J are used. The model described here is not sensitive enough to let both $\Delta H_m$ and $\sigma$ float. The melting temperature ($T = T_m = 38^{\circ}C$) corresponds to $s = 1$. A reasonable value for the enthalpy of melting $\Delta H_m$ is taken to be $\Delta H_m = -6$ kcal/mole. The data are plotted along with model best fit for the predicted $p_C$ using this value and the best fit value of $\sigma = 0.0037$. This simple model reproduces the sigmoid shape of the UV absorption data well.
Figure 11: The UV absorption data across the helix-to-coil transition is compared to model prediction for $p_C = 1 - p_H$ with the best fit parameter $\sigma = 0.0037$. $\Delta H_m = -6$ kcal/mole has been used.

The relative number of helical sequences $p_{o}$ increases, reaches a maximum at $s=1$ (or $T=T_m$) then decreases. Note that at the melting transition (i.e., for $T=T_m$), $N_H = N/2$ and $p_o = 0.029$ which means that there are 29 helical sequences per 1000 base units.

REFERENCES


QUESTIONS

1. DNA is formed of what units?
2. What drives the formation of the helical structure of DNA?
3. What is the analytical measurement method of choice for observing the helix-to-coil transition?
4. What is the typical helix-to-coil transition temperature for DNA/water? How about for DNA/ethylene glycol?
5. What is the SANS Porod exponent for the helical structure? How about for the coil structure? What do these exponents mean?
6. What is the activation enthalpy for the melting of a helical unit?

ANSWERS

1. DNA is formed of nucleotides.
2. The helical structure of DNA is driven by the stacking of the amine bases and the hydrogen-bonding between them.
3. UV absorption spectroscopy is the analytical measurement method of choice for observing the helix-to-coil transition.
4. A typical helix-to-coil transition temperature of 94 °C characterizes DNA/water. That temperature is around 38 °C for DNA/ethylene glycol.
5. A SANS Porod exponent close to 4 characterizes the helical structure and an exponent close to 5/3 characterizes the coil structure. A Porod exponent of 4 is for a cylinder with smooth surface whereas an exponent of 5/3 is for a fully swollen coil.
6. The melting of a helical unit is characterized by an activation enthalpy of -6 kcal/mol.