Problem 28. Nucleic acids

Genetic information is encoded in a sequence of nucleobases which are bonded to a sugar–phosphate backbone. Deoxyribonucleic acid (DNA) contains adenine (A), cytosine (C), guanine (G), and thymine (T), whereas ribonucleic acid (RNA) contains uracil (U) instead of thymine.

The most common structures of nucleobases are shown in Figure 1, but these are not the only possible ones. Since nucleobases contain a number of double bonds, they may occur in several different tautomeric forms. Note that even zwitterionic tautomers are possible in principle, but the tasks below deal only with uncharged molecular structures.

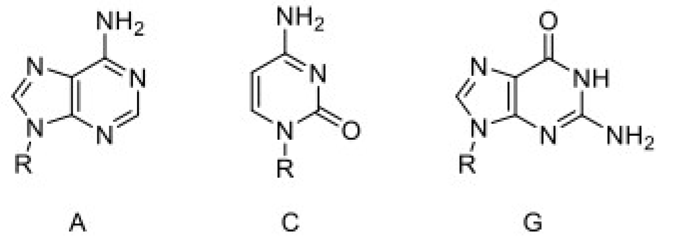
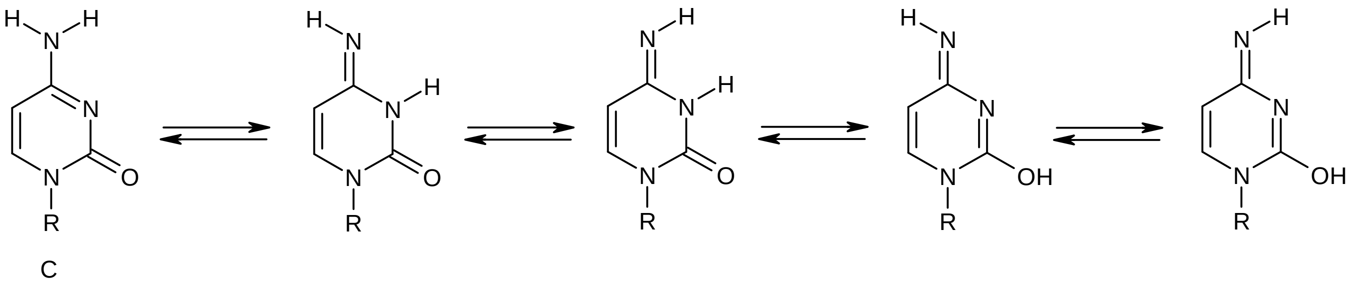


Figure 1. Structural formulae of nucleobases A, C, G, T, and U bonded to sugar-phosphate backbone (R).

28.1 Draw the structural formulae of all non-charged tautomers of cytosine. Assume the nucleobase is bonded to the sugar–phosphate backbone. Consider any pair of imino *E*/*Z* isomers as two different tautomers.



DNA undergoes so-called hybridization, in which two DNA strands form a complex in a helical shape. Hydrogen bonds between the nucleobases contribute to the correct pairing of two complementary strands of double-stranded DNA (dsDNA). Cytosine pairs with guanine, and adenine pairs with thymine (Figure 2). However, the presence of a rare tautomer in one of the DNA strands opens the possibility for non-standard pairing of nucleobases.

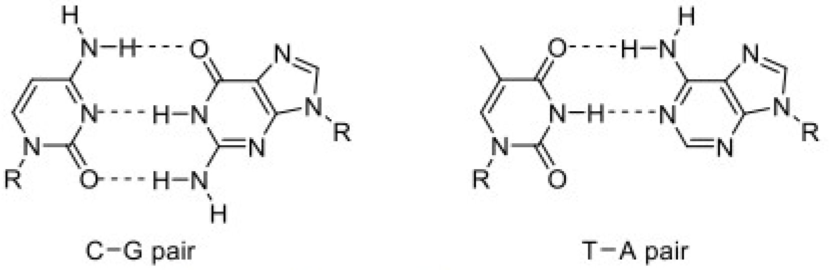
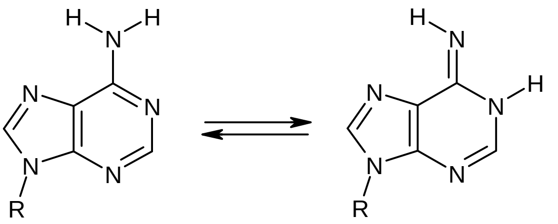


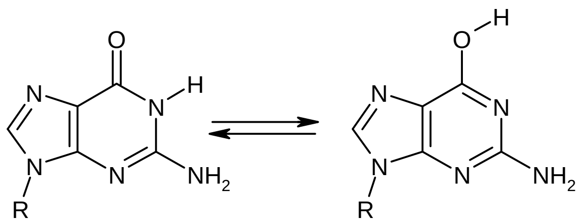
Figure 2. Standard DNA base pairs.

28.2 Draw the structural formulae of the non-standard pairs T–G*, T*–G, A–C* and A*–C, where any minor uncharged tautomer is marked with an asterisk. Keep the relative orientation of the sugar–phosphate backbone the same as in the standard pairs and maximize the number of hydrogen bonds between the nucleobases.

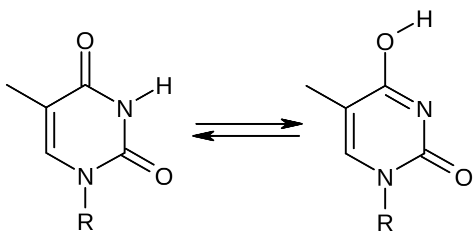
Per le quattro basi azotate i tautomeri più significativi sono i seguenti.



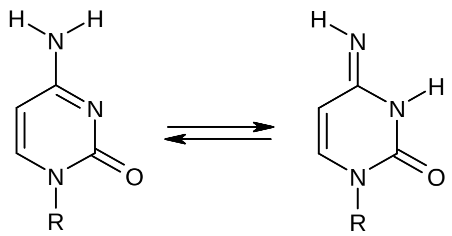
A A*



G G*

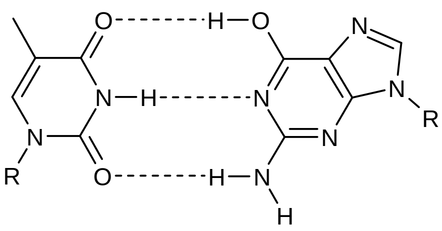


T T*

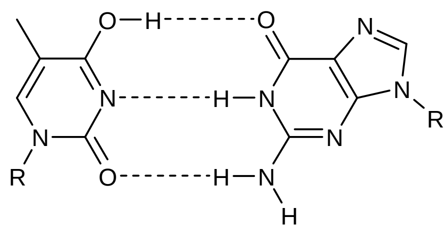


C C*

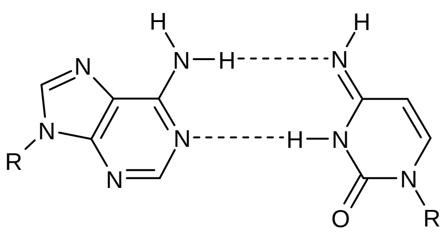
Gli accoppiamenti non standard sono quindi:



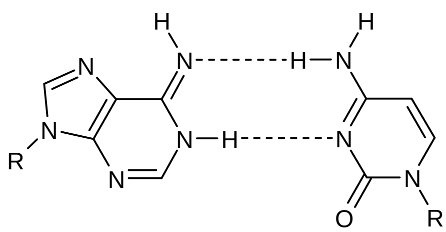
T G*



T* G



A C*



A* C

Spectrophotometry is an experimental technique that is particularly useful for investigating nucleic acids. Being aromatic, nucleobases absorb electromagnetic radiation in the UV range. At 260 nm, sample **1** of a nucleic acid with an unknown concentration of adenine transmits 11% UV light. A standard solution in which the concentration of adenine amounts to 27 μmol dm^{-3}, absorbs 57% UV light at the same wavelength.

28.3 Calculate the unknown concentration of adenine in sample **1**. Neglect any absorption at 260 nm by the other nucleobases and assume that both measurements were performed under identical experimental conditions (cuvette length, buffer composition, temperature, etc.).

L'assorbanza A è legata alla trasmittanza T dalla relazione $A = -\log T$

L'assorbanza A è legata alla concentrazione C dalla legge di Beer $A = \varepsilon\, l\, C$

Se la sostanza analizzata, la temperatura e la cella restano uguali, ε ed l restano uguali e quindi $C/A = k$

Quindi si può scrivere $\frac{C}{A} = \frac{C_o}{A_o}$ dove C e A si riferiscono alla soluzione incognita, mentre C_o e A_o si riferiscono alla soluzione nota.

L'assorbanza A della soluzione incognita si ottiene dalla $A = -\log T$ $\quad A = \log\frac{1}{T} = \log\frac{1}{0,11}$ $\quad A = 0,9586$

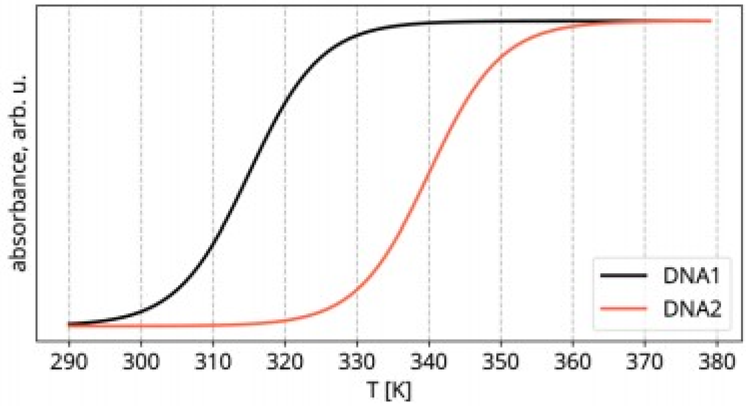
La trasmittanza % della soluzione nota è 100 - 57 = 43%, quindi la trasmittanza è 0,43

L'assorbanza A_o della soluzione nota si ottiene dalla $A_o = \log\frac{1}{T_o}$ $\quad A_o = \log\frac{1}{0,43}$ $\quad A_o = 0,3665$

Quindi $C = \frac{C_o}{A_o} A$ $\quad C = \frac{0,9586}{0,3665} 27 = 70,62\ \mu mol\ L^{-1}$

La concentrazione incognita è quindi di 71 $\mu mol\ L^{-1}$.

Spectrophotometry in the near-UV region is a useful tool to monitor the hybridization of DNA as the temperature changes. Melting temperature T_m is defined as the temperature at which 50% of the original amount of DNA double helices are dissociated into separated strands. Nucleobases within dsDNA absorb less strongly than those in ssDNA, thus the dissociation of dsDNA manifests itself by an increase of absorbance. The plot below shows the absorbance at 260 nm as a function of temperature for two different DNA species (DNA1 and DNA2). Assume that both DNA species have equal molar absorption coefficients and that all the measurements were performed under otherwise identical conditions using identical equipment (initial concentrations, buffers, cuvette, etc.).



28.4 Considering the plot shown above, decide whether the following statements are true or false or whether that cannot be answered based only on the plot.

a) At 320 K, the concentration of dsDNA1 is lower than the concentration of dsDNA2.

True False Cannot be answered

b) The melting temperature T_m of DNA1 is higher than the melting temperature of DNA2.

True False Cannot be answered

c) dsDNA of the species DNA1 is more thermodynamically stable than that of the DNA2 with respect to their single-stranded forms.

True False Cannot be answered

d) dsDNA1 is composed of a larger number of nucleobase pairs than dsDNA2

True False Cannot be answered

----

a) True A 320 K il DNA2 è ancora a doppia elica (basso assorbimento), mentre il DNA1 è quasi completamente dissociato (assorbimento maggiore) quindi a 320 K c'è poco DNA1 a doppia elica
b) False La Tm del DNA1 è di circa 315 K, mentre quella del DNA2 e di circa 340 K
c) False Se il DNA1 si dissocia a T minore, significa che è termodinamicamente meno stabile
d) False Il DNA1 si dissocia a T minore, quindi è più ricco di coppie di basi A-T che sono meno fortemente legate (2 legami idrogeno) di quelle G-C (3 legami idrogeno). Le basi A-T assorbono di più a 260 nm di quelle G-C, quindi dato che a temperatura ambiente l'assorbanza è la stessa, il DNA1 deve essere un po' più corto.

The Rous sarcoma virus is a retrovirus. Its genetic information is stored in a single strand of RNA rather than in dsDNA; recall that RNA contains uracil instead of thymine (Figure 1). The virus uses an enzyme, reverse transcriptase, to synthesize its complementary DNA (cDNA) strand, which is then transcribed to messenger RNA (mRNA). Finally, the mRNA is translated to a polypeptide strand in the ribosome of the infected cell.

The following fragment of 8 nucleotides was identified in the RNA of the virus: 5′-CCCCAGGU-3′.

28.5 Write the sequences of cDNA and mRNA corresponding to the octanucleotide. Mind the orientation of the molecule, and identify the 5′- and 3′-termini.

5'-CCCCAGGU-3' RNA virale
↓
3'-GGGGTCCA-5'
5'-CCCCAGGT-3' cDNA
↓
5'-CCCCAGGU-3' mRNA

28.6 How many possible single-stranded RNA octanucleotides do exist?

Dato che la testa e la coda dell'RNA sono distinguibili (3' e 5'), in un oligonucleotide composto di n nucleotidi si possono avere 4^n filamenti diversi, in un ottanucleotide si possono avere 4^8 (65536) diverse catene.

Soluzione proposta da Mauro Tonellato - Padova