Biological matter: physical description of biopolymers (1)

Jean-Louis Sikorav

EPFL October 2 2008

Goal

- Provide a general description of the structure of informational biopolymers (DNA, RNA, proteins).
- Use concepts of polymer physics to better understand their (static) properties, in vitro and in vivo.

Structure of biopolymers

- Primary structure: complete covalent structure. The sequence of monomers plus other covalent bonds [disulfide bridges, intra- or inter-chain, post-translational modifications of amino acids (hydroxylation, acetylation, glycosylation, phosphorylation, … or bases (methylation…)].
- Secondary structure: description of tridimensional structures of locally ordered monomers. Corresponds to helical structures (DNA: A, B, Z; proteins: a helix, b sheet…).

Peptide bond

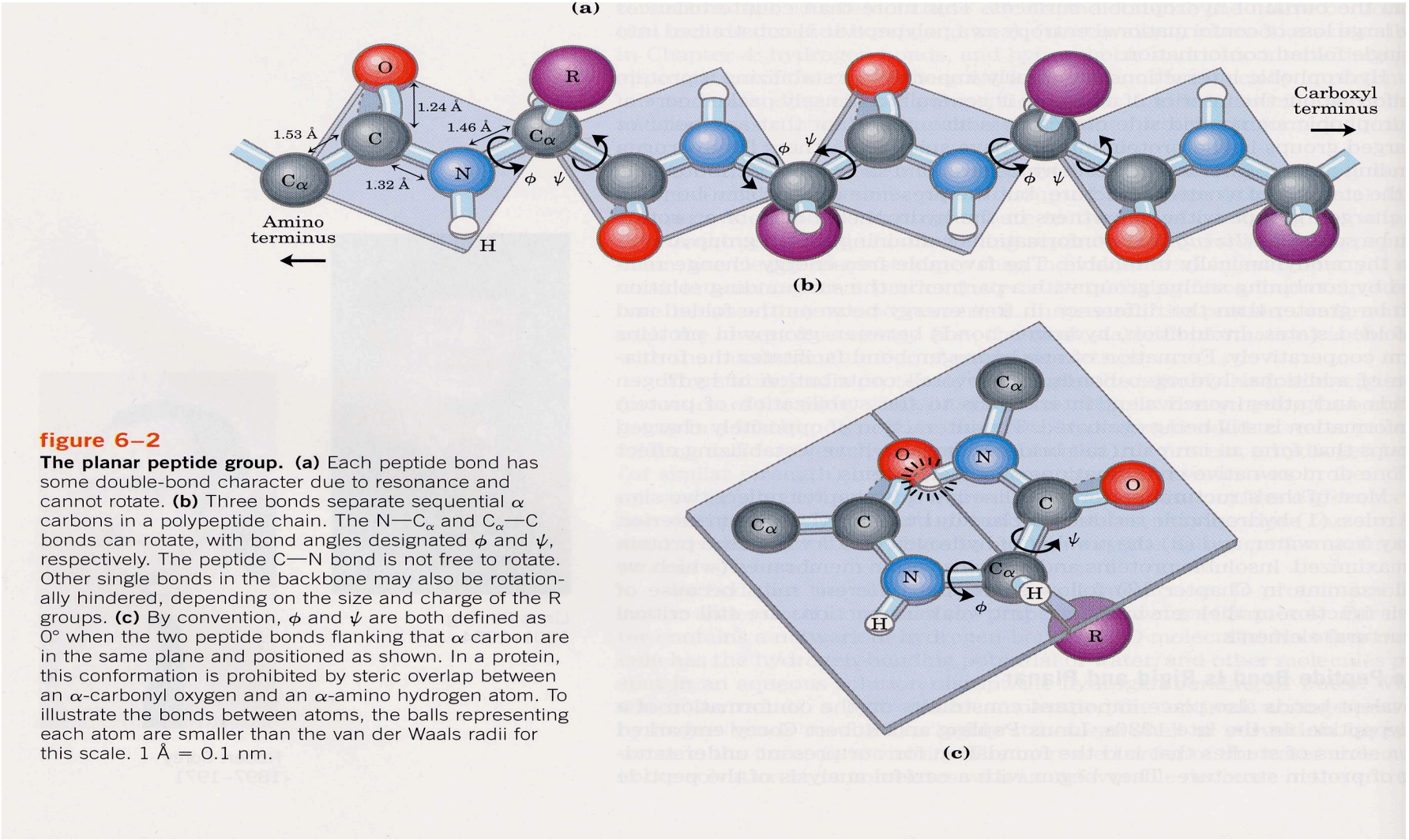


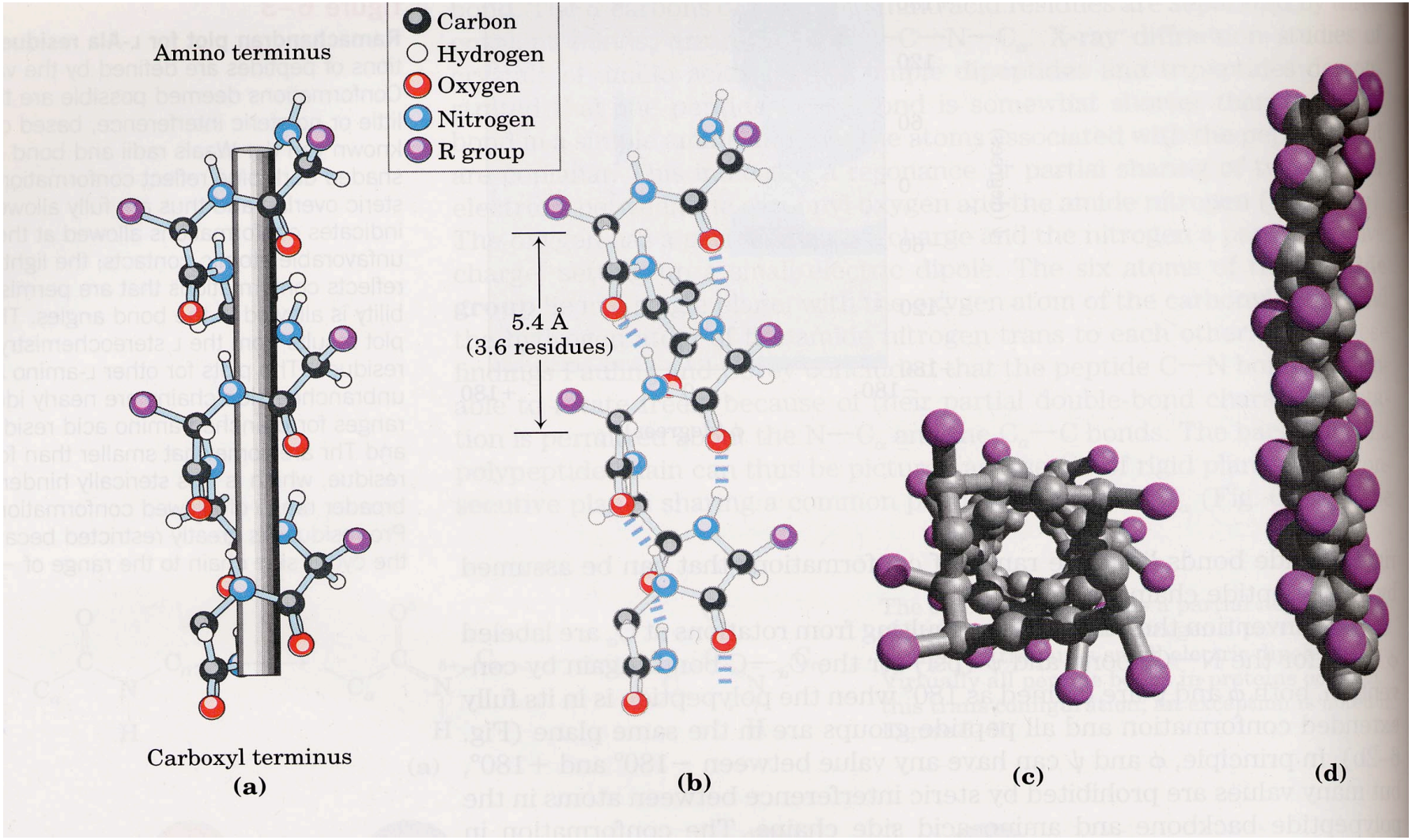
figure 6–2

The planar peptide group. (a) Each peptide bond has some double-bond character due to resonance and cannot rotate. **(b)** Three bonds separate sequential α carbons in a polypeptide chain. The N—C_α and C_α—C bonds can rotate, with bond angles designated ϕ and ψ, respectively. The peptide C—N bond is not free to rotate. Other single bonds in the backbone may also be rotationally hindered, depending on the size and charge of the R groups. **(c)** By convention, ϕ and ψ are both defined as 0° when the two peptide bonds flanking that α carbon are in the same plane and positioned as shown. In a protein, this conformation is prohibited by steric overlap between an α-carbonyl oxygen and an α-amino hydrogen atom. To illustrate the bonds between atoms, the balls representing each atom are smaller than the van der Waals radii for this scale. 1 Å = 0.1 nm.

α Helix

- Right-handed helix 3.6 amino acids (AA) per turn, (helical pitch 1.5 Å $\times$ 3.6 = 5.4 Å ; diameter (without side groups) 6 Å.
- Each carbonyl group is an acceptor for a hydrogen bond with the NH group located 4 residues farther.
- Side groups of the AA are outside and do not influence the helical structure. Proline is an exception: it is used to make a left-handed helix as found in collagen.

α Helix


Carbon
Hydrogen
Oxygen
Nitrogen
R group
Amino terminus
5.4 Å
(3.6 residues)
Carboxyl terminus
(a)
(b)
(c)
(d)

β Sheet

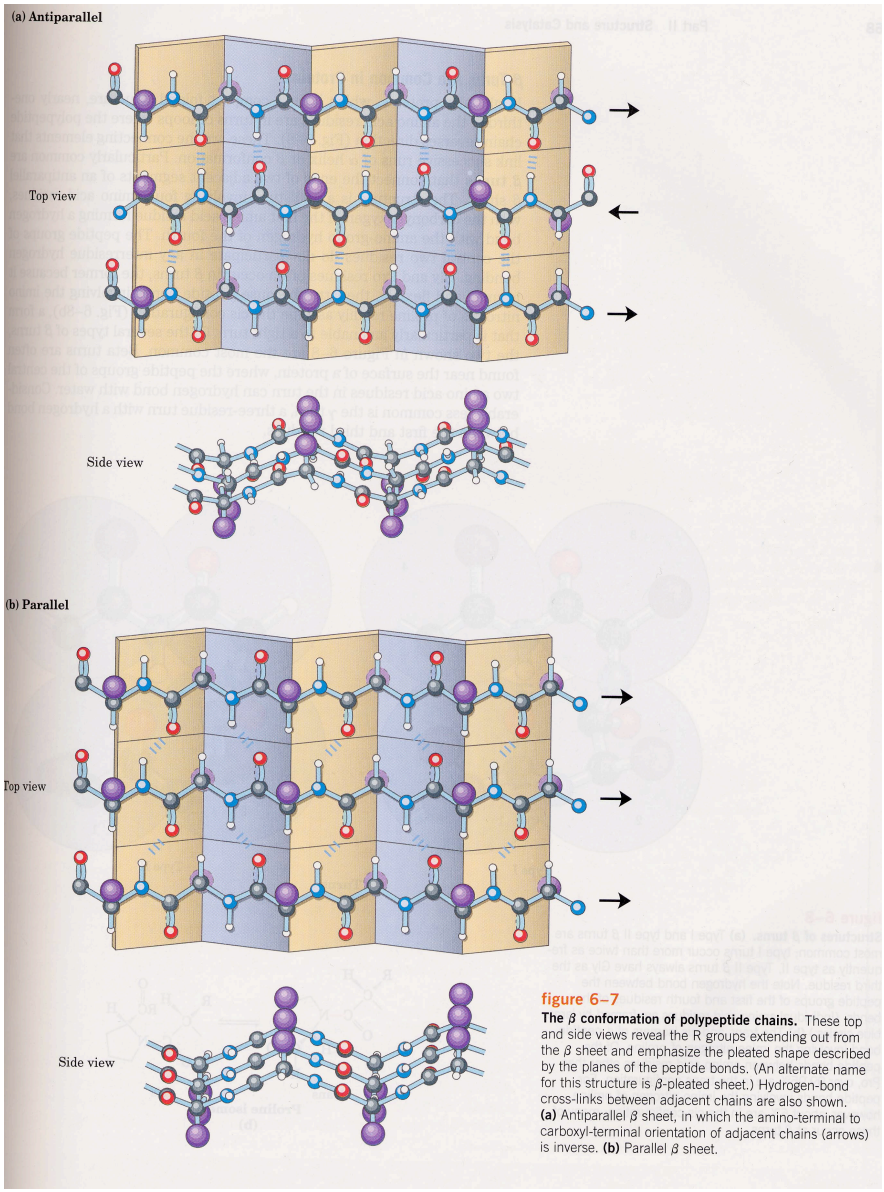


figure 6–7
The β conformation of polypeptide chains. These top and side views reveal the R groups extending out from the β sheet and emphasize the pleated shape described by the planes of the peptide bonds. (An alternate name for this structure is β-pleated sheet.) Hydrogen-bond cross-links between adjacent chains are also shown. **(a)** Antiparallel β sheet, in which the amino-terminal to carboxyl-terminal orientation of adjacent chains (arrows) is inverse. **(b)** Parallel β sheet.

β Turn

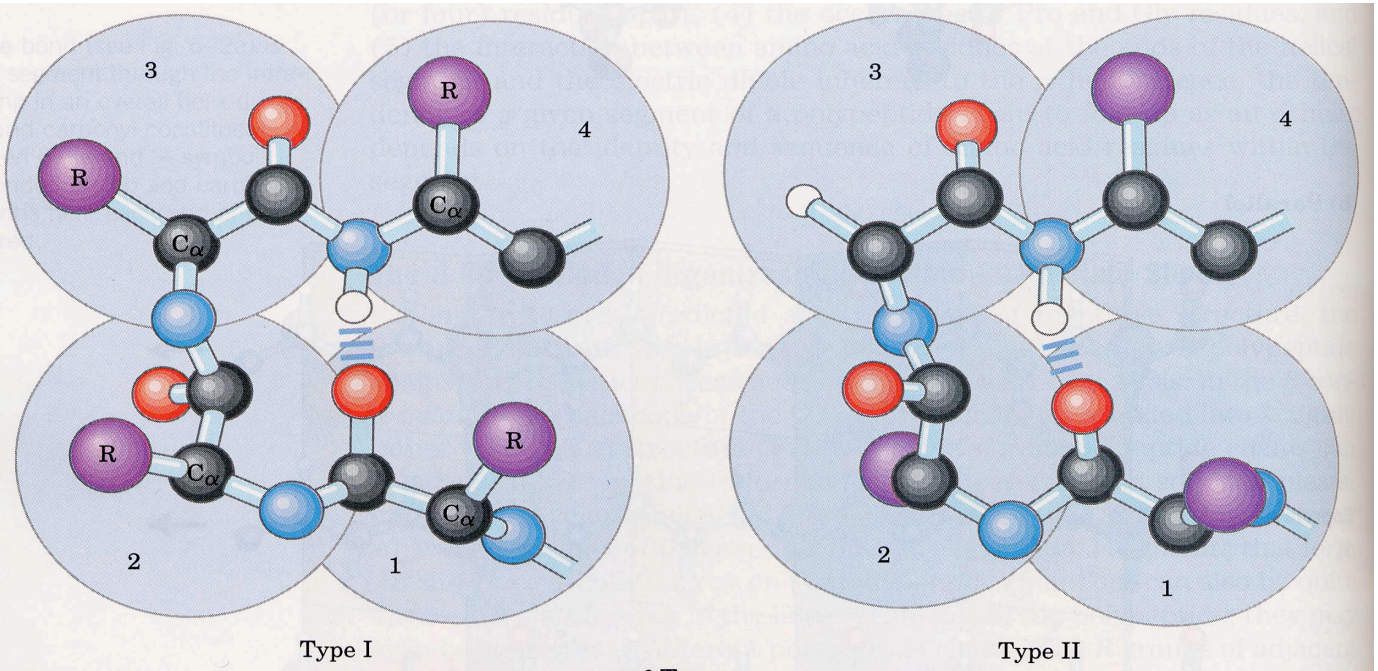
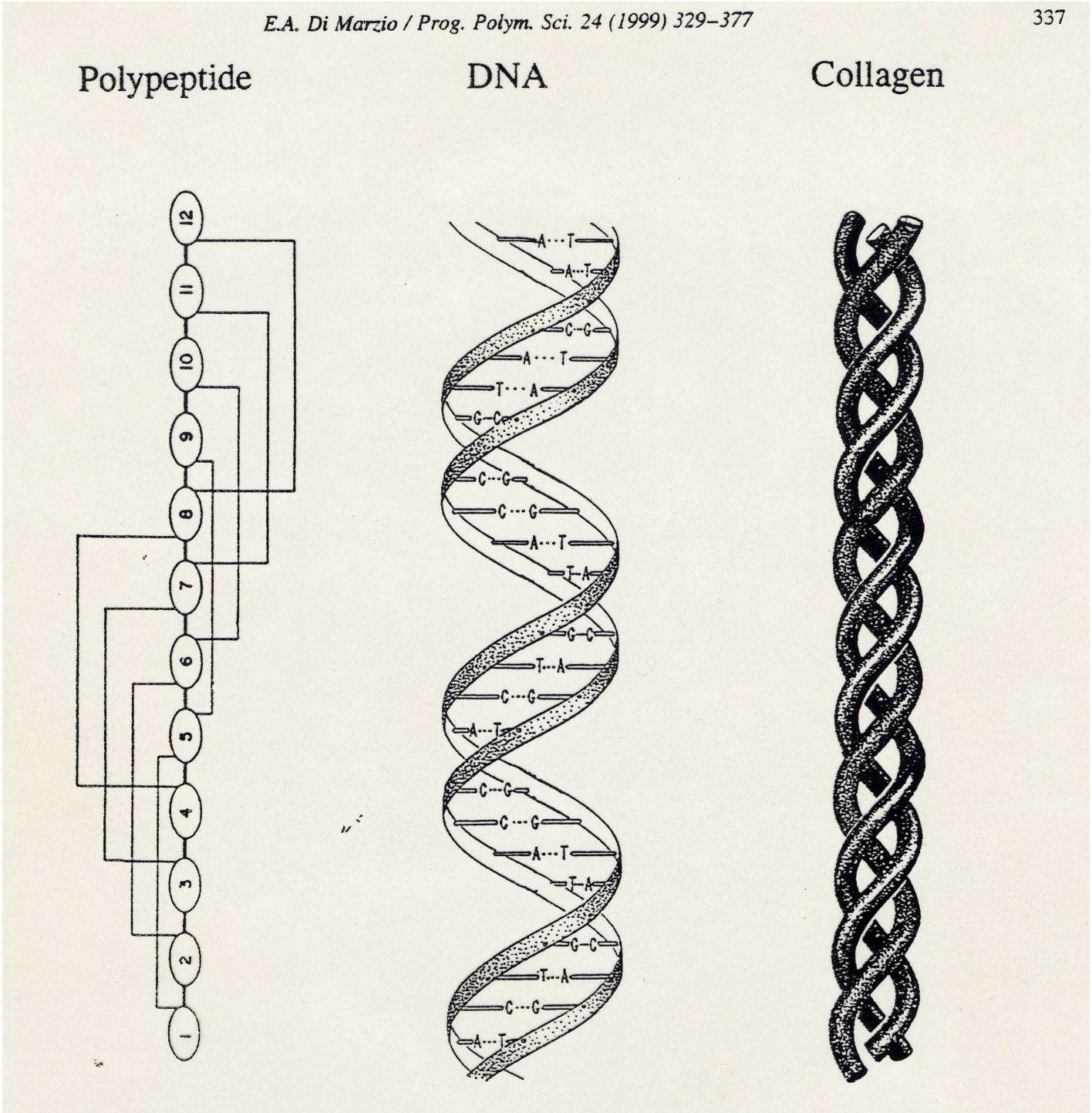


figure 6–8

Structures of β turns. (a) Type I and type II β turns are most common; type I turns occur more than twice as frequently as type II. Type II β turns always have Gly as the third residue. Note the hydrogen bond between the peptide groups of the first and fourth residues of the bends. (Individual amino acid residues are framed by large blue circles.) **(b)** The trans and cis isomers of a peptide bond involving the imino nitrogen of proline. Of the peptide bonds between amino acid residues other than Pro, over 99.95% are in the trans configuration. For peptide bonds involving the imino nitrogen of proline, however, about 6% are in the cis configuration; many of these occur at β turns.

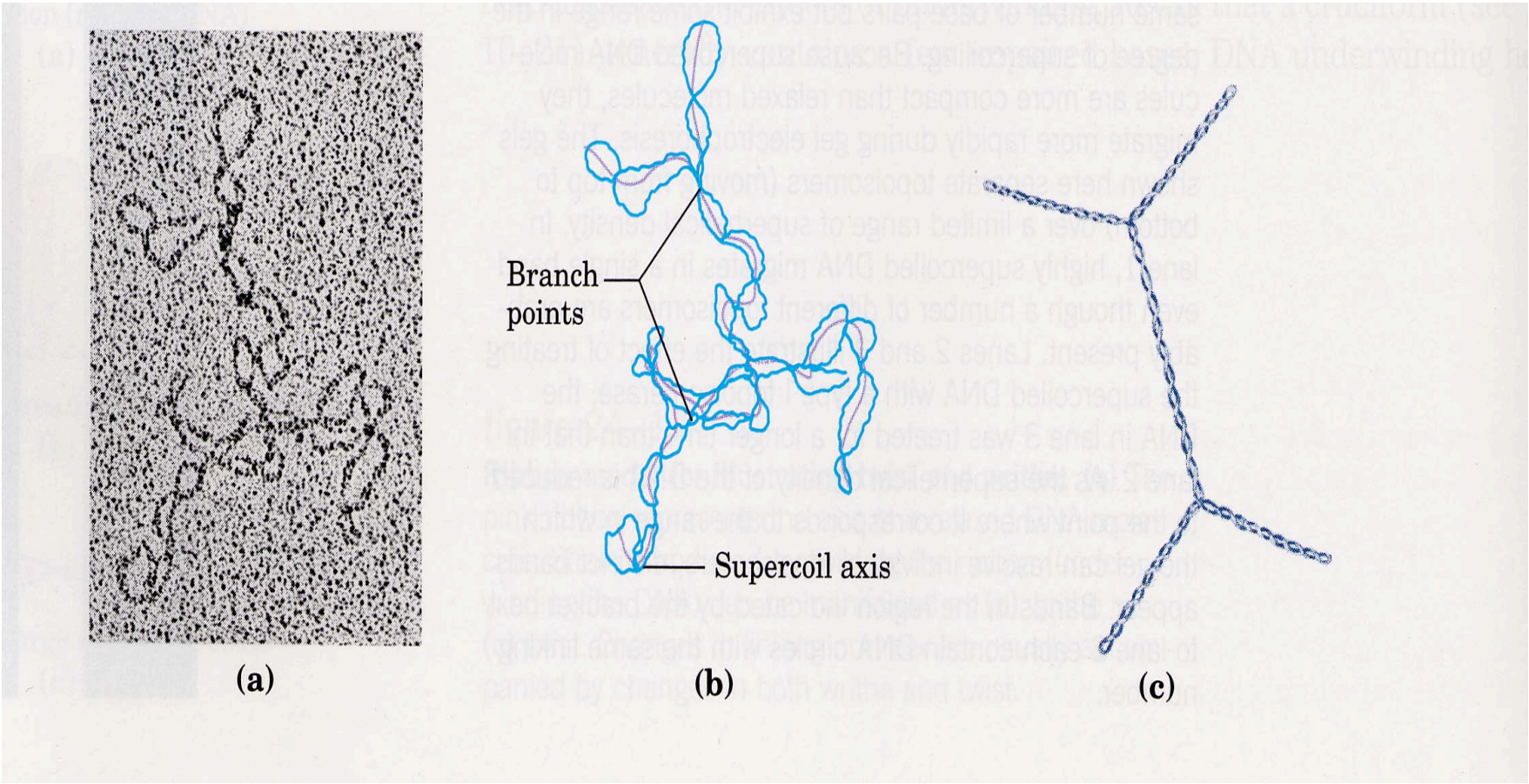
Helices with 1, 2 or 3 strands



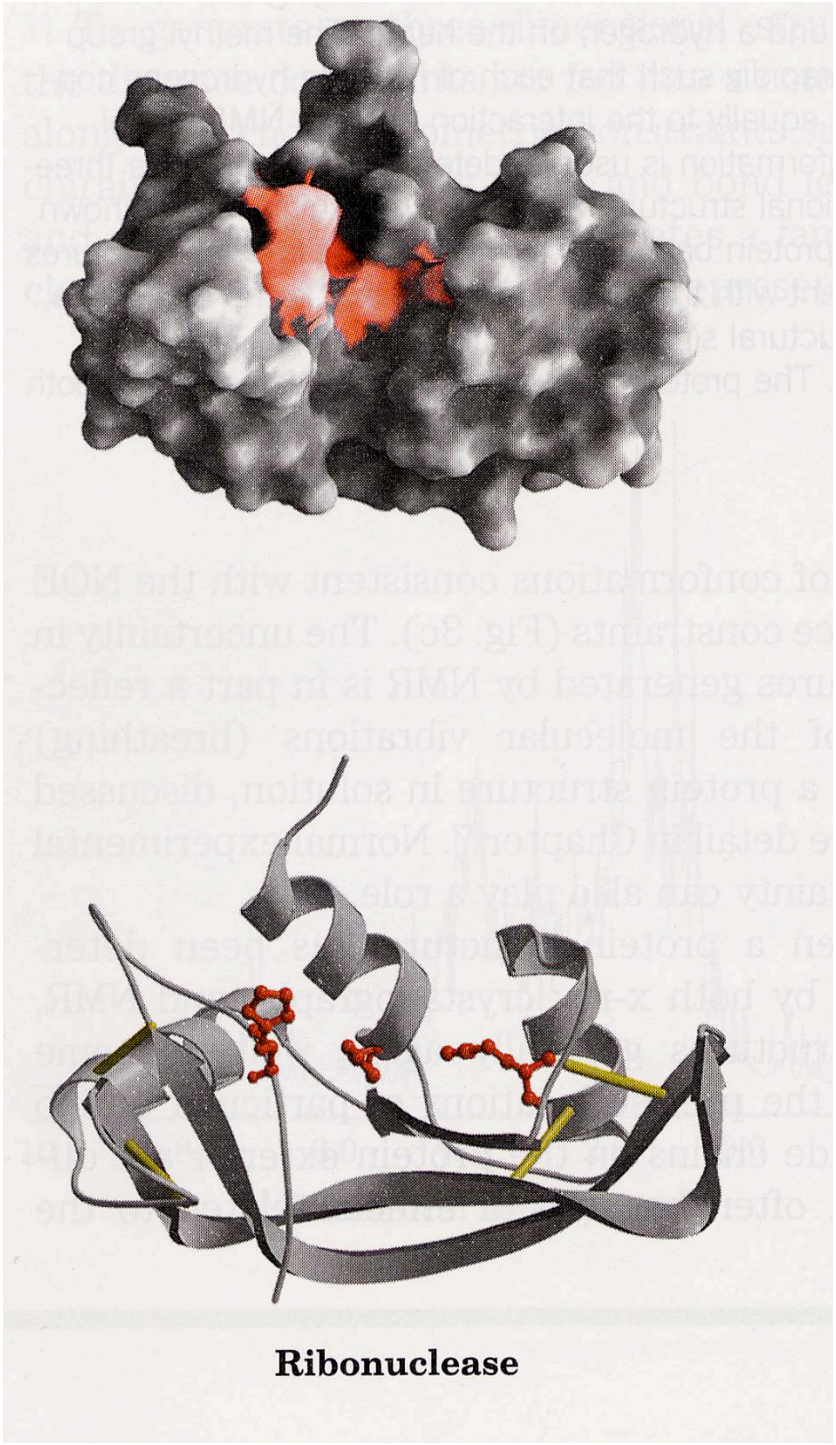
Structure of biopolymers

- Tertiary structure: the complete tridimensional structure of all the monomers of a chain. This includes a description of the secondary structure (examples: circular DNA, globular protein). Concept of folding.
- Quaternary structure: made by non-covalent associations of tertiary structures. Examples: multimeric proteins, nucleic-acid protein complexes (ribosome, filamentous virus capsid viral).

Plectonemic configuration of supercoiled DNA

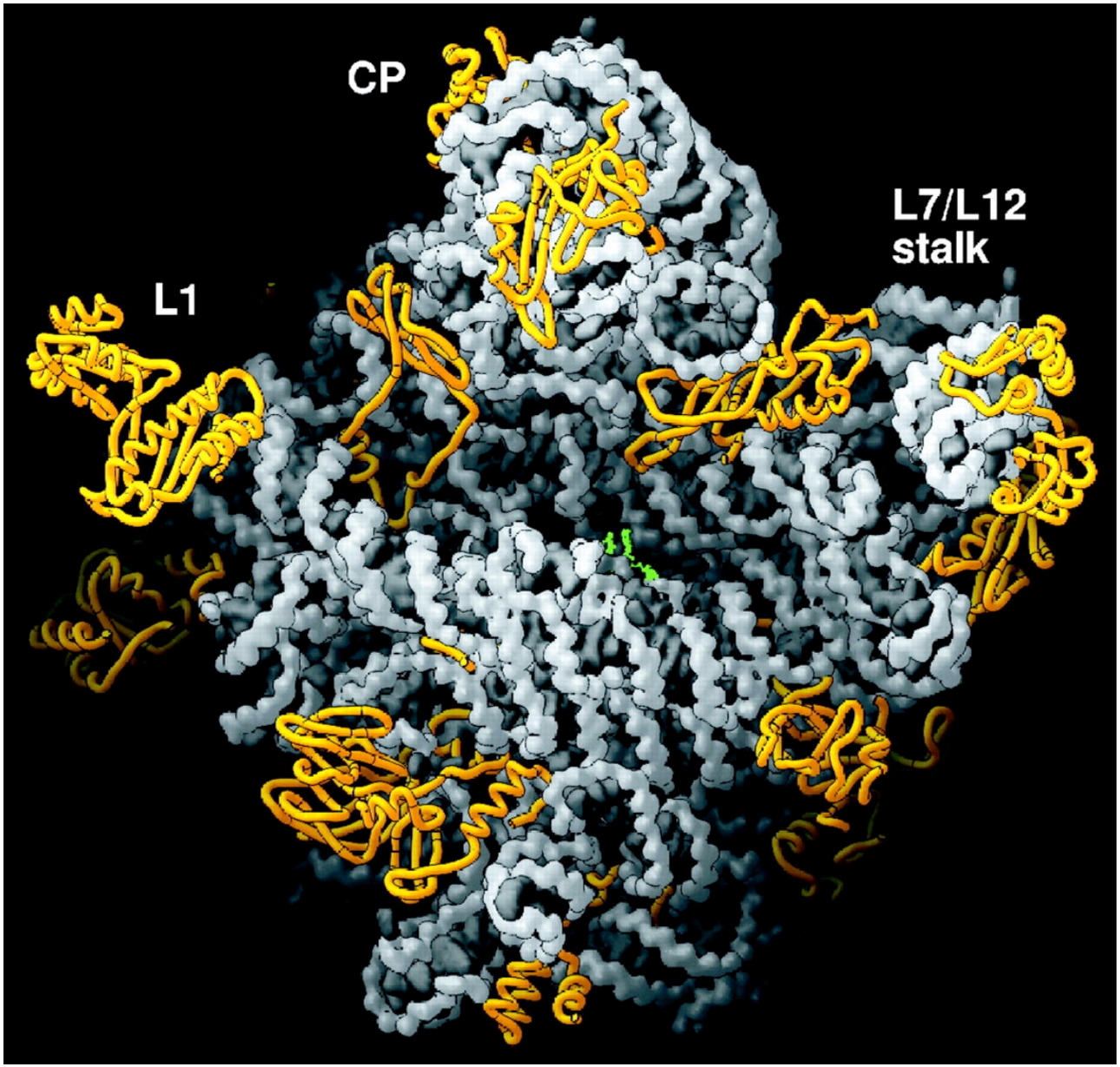


Ribonuclease



Ribonuclease

Ribosome



Symmetry and quaternary structures

- A relation between symmetry properties and assembly processes (equivalence or quasi-equivalence, helices and icoshedra).

- A relation between symmetry and allostery.

Physics of biopolymers

Phase transitions in biopolymers

Helix-coil, coil-globule, coil-stretch, adsorption, polymerization, threading through a pore.

Sol-gel, aggregation, liquid-crystalline mesophases.

Flory J. Pol. Sci. (1961), DiMarzio (1999)

Energy and molecular biology

- Thermal energy k_BT (4×10^{-21} J à 25°C)
- Energy of a covalent bond:
  100-200 k_BT (C-C : 144 k_BT)

Energy released by ATP hydrolysis:

ATP $\rightarrow$ AMP + PP_i environ 20 k_BT

Here we focus on conformational changes

Forces and bonds stabilizing biopolymers (1)

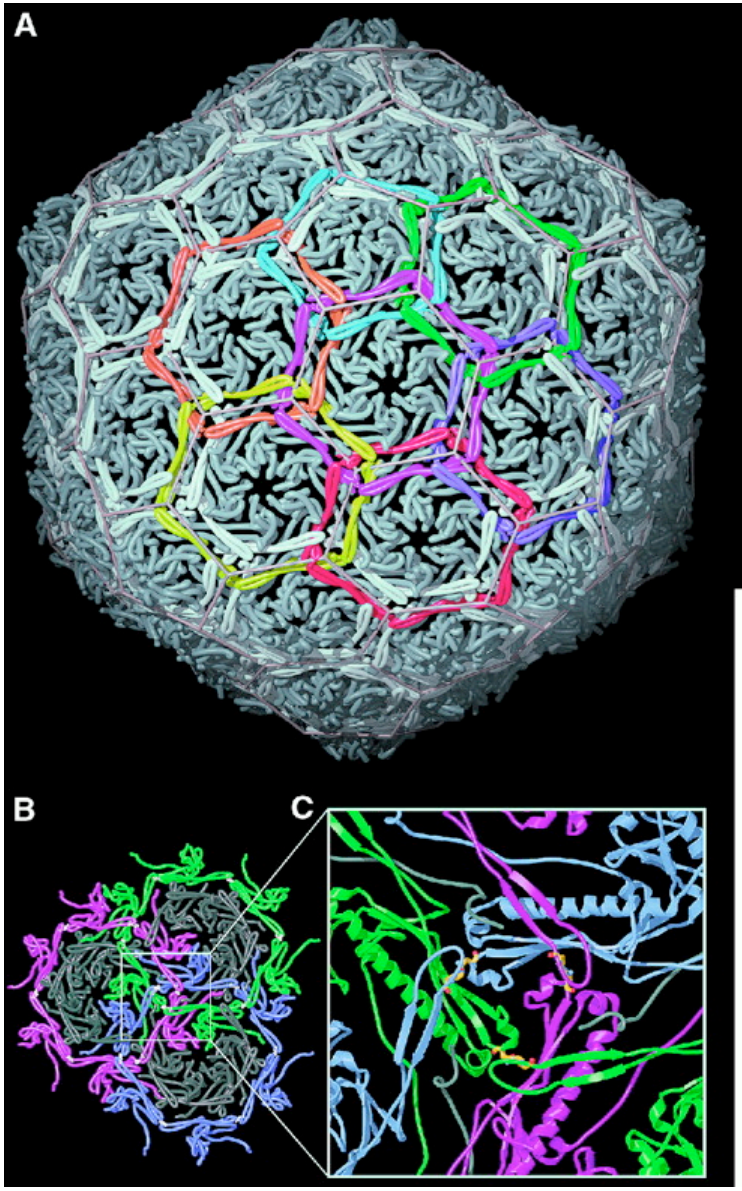
- Covalent bonds
- Interactions between a charge and a permanent dipole (charge/charge, charge/ permanent dipole)
- van der Waals interactions (decay as r^{-6}) (permanent dipole/permanent dipole, charge/ molecule, induced dipole /induced dipole)
- Hard-core repulsive forces (steric)
- Hydrogen bond
- Hydrophobic bond

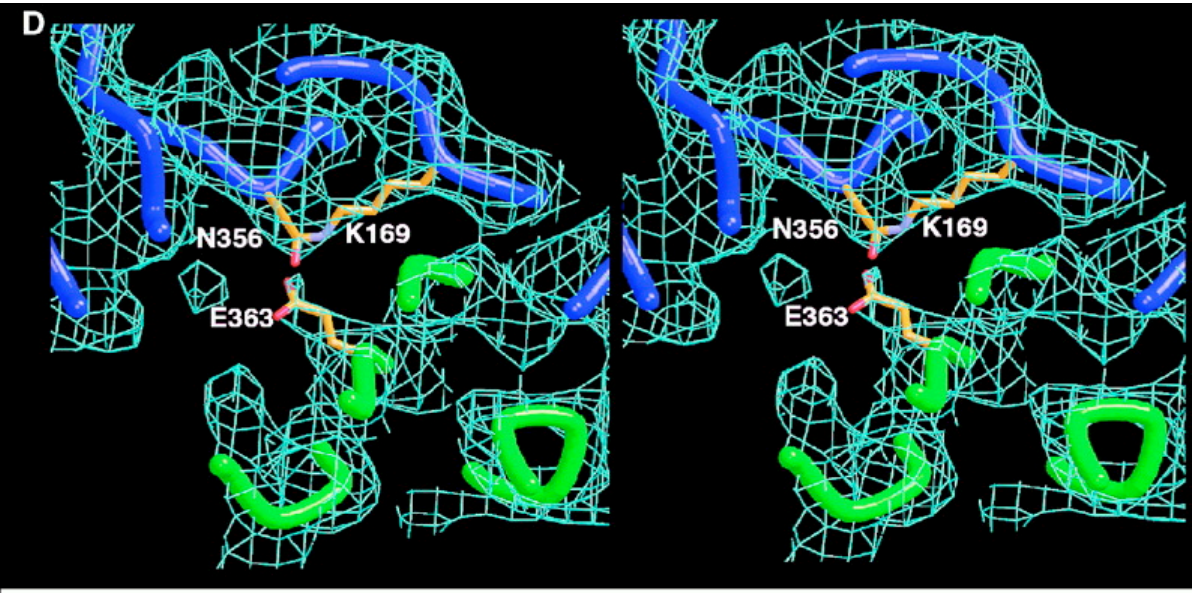
See for instance Daune (Biophysique moléculaire) and Israelachvili (Intermolecular & Surface Forces)

Forces and bonds stabilizing biopolymers (2)

Topological bonds:

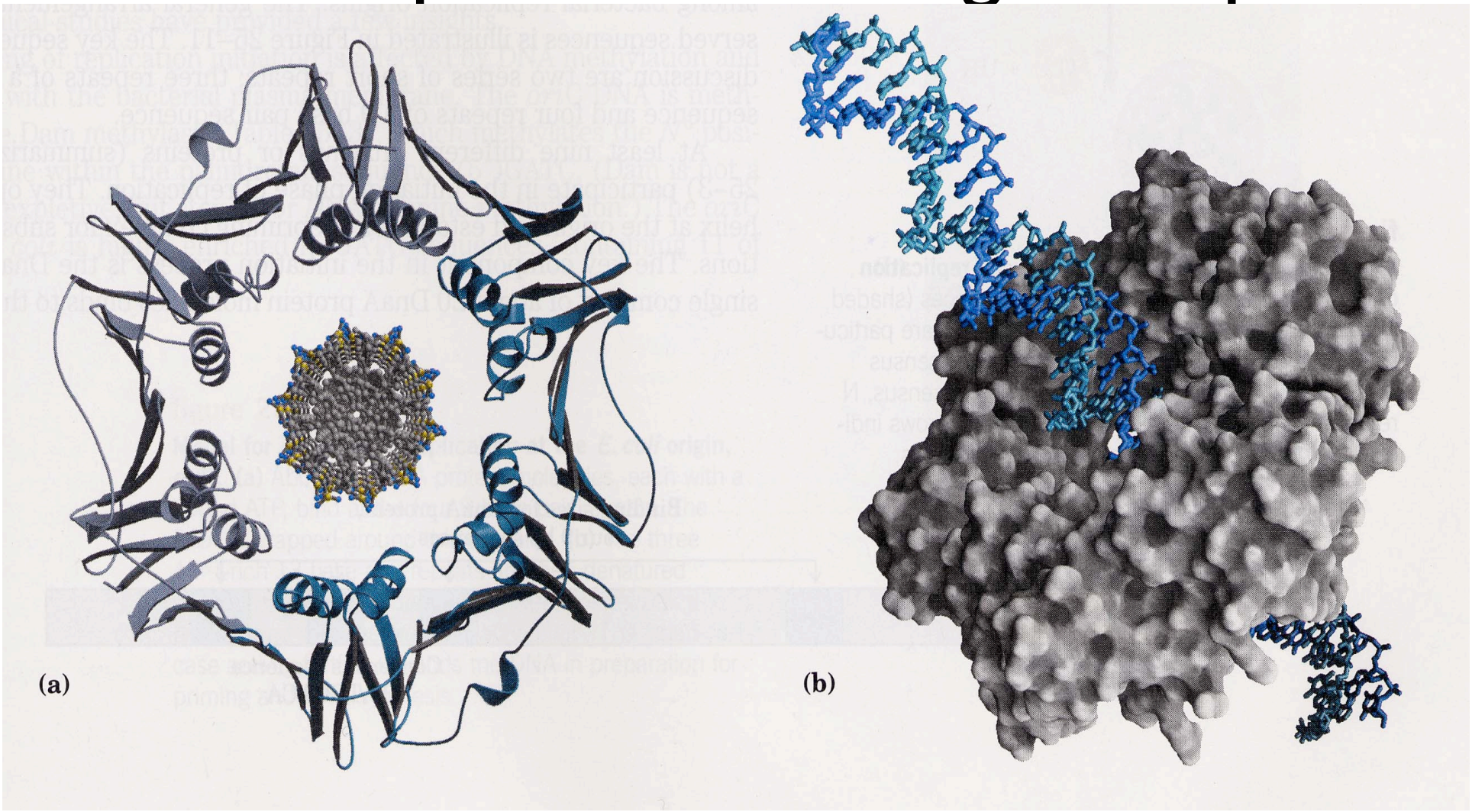
- Circular DNA (knots and catenanes)
- Coat of mail of viral caspids
- DNA clamps





Coat of mail HK97
Science (2000)
289:2129-33

β Sub-unit of *Escherichia coli* DNA pol III: a sliding clamp

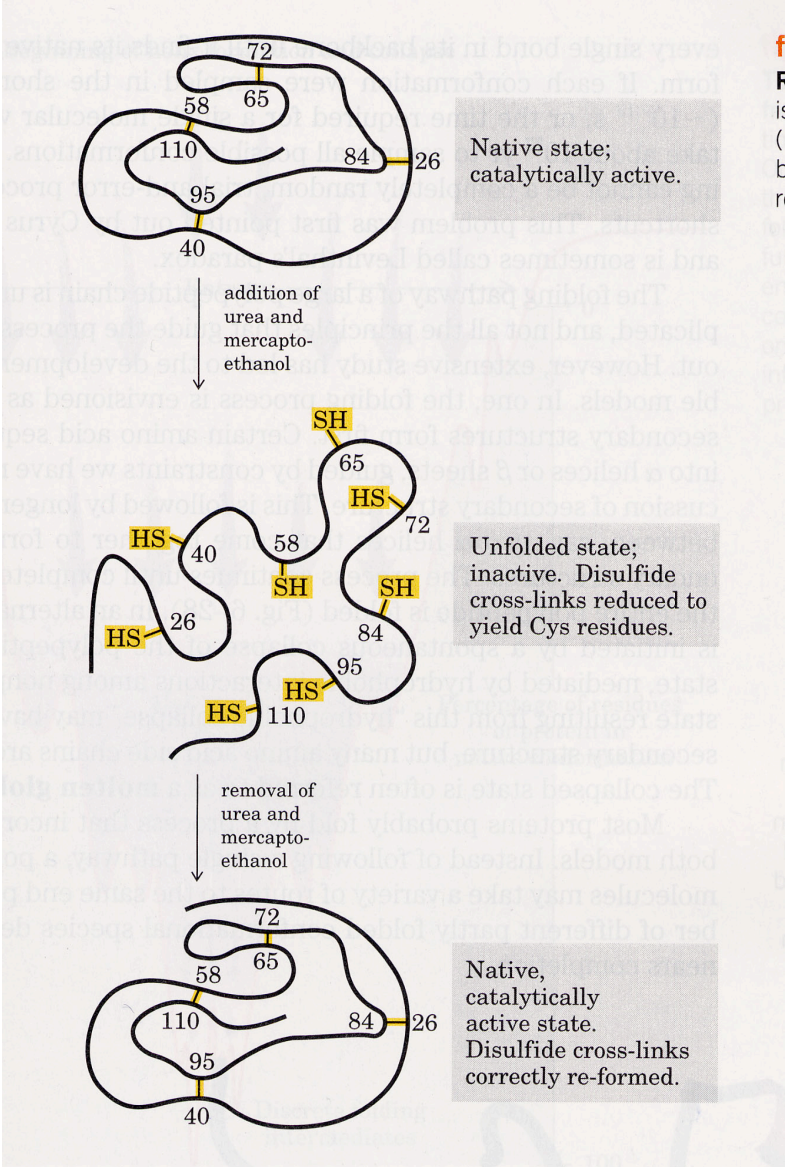


Static properties of biopolymers

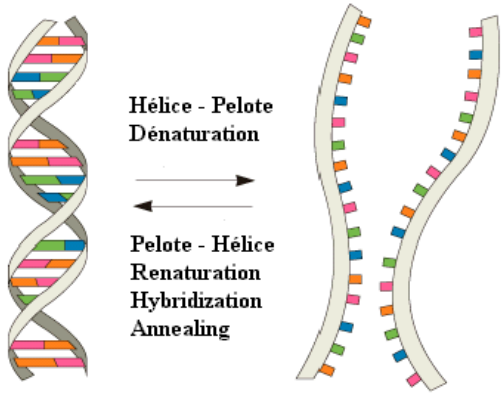
- Approach: study of simples models (homopolymer) of the secondary structure
- Concept of persistence length, a consequence of helical structure
- Transition between helical and coil conformations (locally disordered)

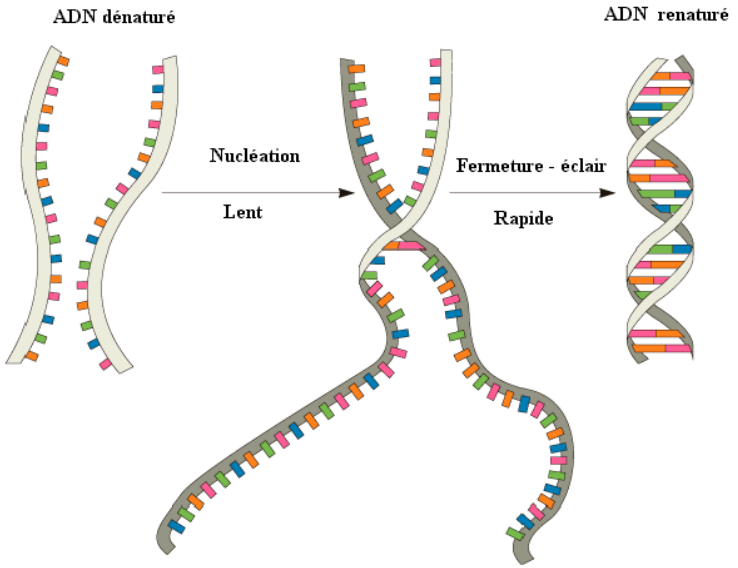
  A simple model for the denaturation-renaturation of proteins and nucleic acids

Reversible denaturation-renaturation of ribonuclease (Anfisen)



Reversible denaturation-renaturation of DNA (Marmur)





Models of non-interacting polymeric chains

Ideal chain (Gaussian, freely-jointed, random-walk, random coil)

Excluded volume effect (swollen coil, collapse)

Freely rotating:

$$< \vec{a}_i . \vec{a}_{i+1} >= a^2 \cos\theta$$

Persistence length l_p: length over which the chain persists in the initial direction:

$$l_p = \lim_{N \to \infty} < \frac{\vec{a}_1}{a} . \vec{R}_N >= \frac{a}{1 - \cos\theta}$$

Semi-flexible Chain (1)

Wormlike chain, persistent chain

Limit of a freely rotating chain with a contour length L and $l_p = a/(1\text{-}\cos\theta)$ kept constant, divided into N segments of length $a = L/N$

One finds $(\cos\theta)^N \approx \exp\text{-}L/l_p$

Semi-flexible Chain (2)

Correlation length for:

$$< \vec{u}_1 . \vec{u}_2 >= \exp - \frac{|s_2 - s_1|}{l_p}$$

Yielding:

$$< \vec{R}^2 >= 2 l_p L \left[1 - \frac{l_p}{L} (1 - \exp - \frac{L}{l_p}) \right]$$

$$L \to \infty \quad < \vec{R}^2 >= 2 l_p L$$

$$L \to 0 \quad < \vec{R}^2 >= L^2 (1 - 3 \frac{L}{l_p} + \frac{1}{12} \left(\frac{L}{l_p} \right)^2 + \ldots)$$

Examples

- DNA: $l_p \approx 50$ nm or 150 base pairs de (polyelectrolyte, l_p depends on salt)
- α Helix: 20 Å (poly-L-alanine), 220 Å (poly-L-proline)
- Filamentous bacteriophage F_c $l_p \approx 2{,}2$ mm ($L = 0{,}895$ mm)
- Actin filament $l_p \approx 10$ mm
- Microtubule $l_p \approx 1$ mm (hollow cylinder)

Relation between l_p et chain elasticity

Bending fluctuations of long molecules (Landau & Lifshitz)

- Bending of a semi-flexible chain:

$$\frac{dF}{ds} = 1/2K\left(\frac{d\vartheta}{ds}\right)^2$$

- *K* characterizes the resistance to bending

$$l_p = \frac{K}{k_B T}$$

Helix to coil transition

- Examples: α helix $\rightarrow$ random coil
  dsDNA $\rightarrow$ 2 complementary strands
  Collagen $\rightarrow$ 3 strands.
- Similar transitions:
  Helix $\rightarrow$helix helix $\rightarrow$ β sheet

  Goal: describe quantitatively the transition for a homopolymer then for genomic dsDNA

Denaturation of poly-g-benzyl-L-glutamate

Cantor & Schimmel

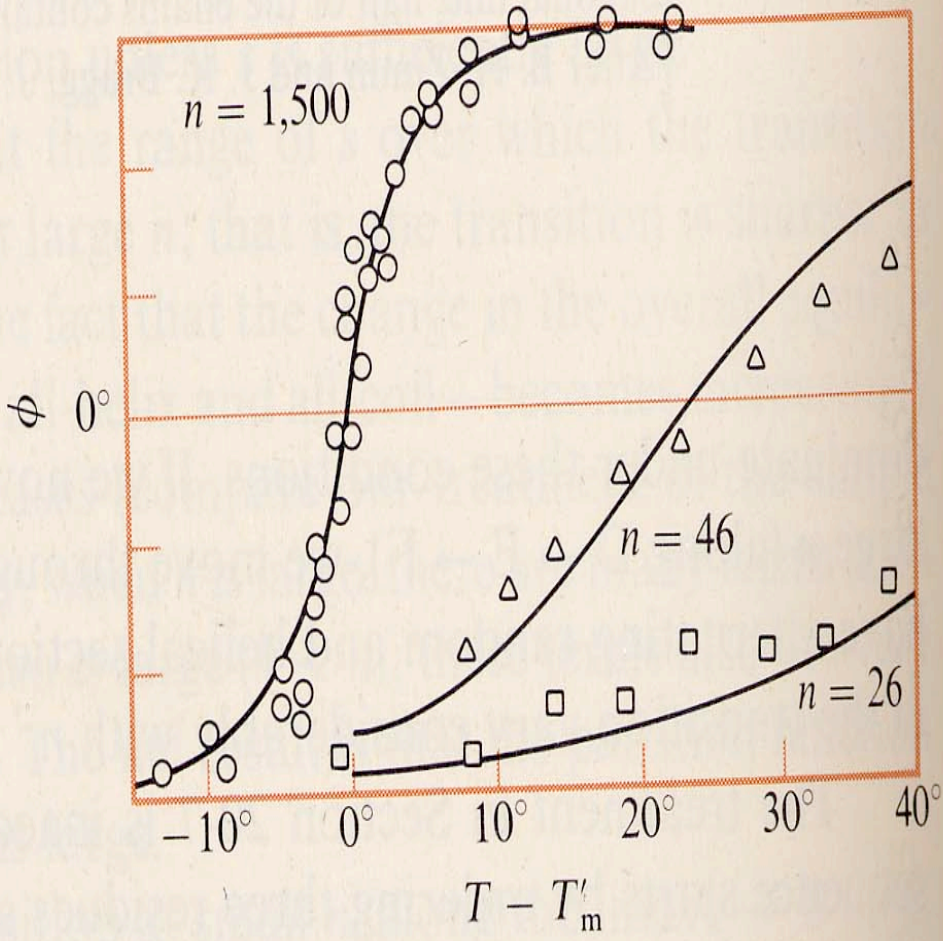


Figure 20-13

Theoretical curves and experimental results. The curves show optical rotation calculated from predicted fractional helicity as a function of $T - T'_m$ for poly-γ-benzyl-L-glutamate chains of three different lengths. The points represent experimental optical-rotation data. [After B. H. Zimm et al., *Proc. Natl. Acad. Sci. USA* 45:1601 (1959).]

Partition function

Partition function Z for an ensemble of N particles enclosed in a volume V at a temperature T. Let n_i be the number or particles with an energy U_i:

$$Z = \Sigma_i \exp(-U_i/k_B T)$$

The probability p_i of finding n_i particles with energy U_i is:

$$p_i = n_i/N = (1/Z)\exp(-U_i/k_B T)$$

Average value of a parameter $X = \Sigma_i p_i X_i$

(for instance fraction θ of monomers in a helical state)

Matrix method for the computation of the partition function of the chain

- We assume that each monomer can exist under 2 possible states, *h* et *c*. 2^N possible states for a chain of N monomers
- The transition from *c* to *h* corresponds to a change in free energy ΔG.
- $s = \exp(-\Delta G/k_B T)$ is the associated equilibrium constant. More accurately *s* is the equilibrium constant for the addition of a helical unit at the end of a helical sequence (propagation step).

Helix initiation in a polypeptide

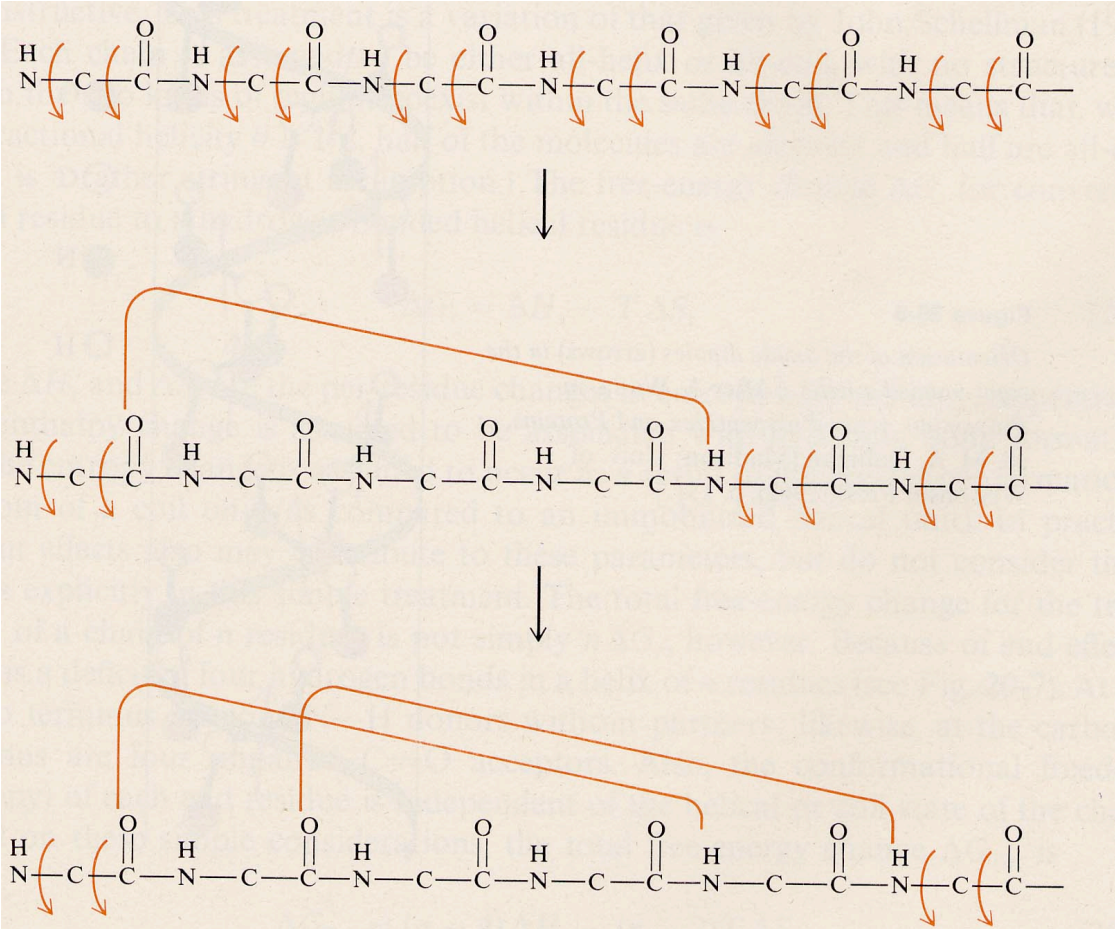


Figure 20-8

Helix initiation. This schematic illustration shows that six rotation angles are immobilized in making the first hydrogen bond, but only two angles are immobilized in making a bond adjacent to an existing one.

Nucleation

- It is more difficult to initiate a helical unit in a disordered (random coil) section
- One defines a nucleation parameter *s* such that the equilibrium constant pour the nucleation is σs.

  (one expects $\sigma << 1$)

Transfer matrix (1)

- For a given residue, there are 4 possible states taking into account the preceding residue (**nearest neighbor model**) :

residue *i*	residue *i*-1	statistical weight
c	*c*	1
c	*h*	1
h	*c*	σs
h	*h*	*s*

Transfer matrix (2)

- The partition function up to residue *i*-1 Z_{i-1} can be written:

  $Z_{i-1} = A_{i-1} + B_{i-1}$ (residue *i*-1 in the state *c* + residue *i*-1 in state *h*)

  $Z_i = A_i + B_i = (1+ \sigma s)A_{i-1} + (1+ s)\, B_{i-1}$

  Matrix representation:

$$\begin{pmatrix} A_i \\ B_i \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \sigma s & s \end{pmatrix} \begin{pmatrix} A_{i-1} \\ B_{i-1} \end{pmatrix}$$

$$Z_{\mathrm{N}} = \begin{pmatrix} 1 & 1 \end{pmatrix} M^{\mathrm{N}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Transfer matrix (3)

- Diagonalization of M : eigen values λ_1 and λ_2 ($\lambda_1 > \lambda_2$)
- For large N $Z_N \approx \lambda_1^N$

- This yields the fraction q of monomers in a helical state :

$$\theta = \frac{1}{N}\frac{\partial \ln Z_N}{\partial \ln s} \approx \frac{\partial \ln \lambda_1}{\partial \ln s}$$

For large N θ does not depend on N (not a phase transition for $N \rightarrow \infty$)

Transfer matrix (4)

- Mean number *H* of helical segments in a chain:

$$H = \frac{\partial \ln Z_N}{\partial \ln \sigma} = N \frac{\partial \ln \lambda_1}{\partial \ln \sigma}$$

Different cases (1)

σ = 1 (no cooperativity), $\lambda_1 = 1 + s$

and $\theta = s/(1 + s)$

σ = 0 (infinite nucleation energy)

$$\lambda_1 = \frac{1}{2}\left(1 + s + |1 - s|\right)$$

$$\lambda_1 = 1 \text{ and } \theta = 0 \text{ if } s < 1$$

$$\lambda_1 = s \text{ and } \theta = 1 \text{ if } s > 1$$

Different cases (2)

σ << 1 :

- Width of the transition:

$$\left(\frac{d\theta}{ds}\right)_{s=1} \approx \frac{1}{4\sigma^{1/2}}$$

Different approach :

$$s = \exp -\frac{\Delta G}{k_B T} \text{ and } \frac{d\ln s}{dT} = \frac{\Delta H}{k_B T^2}$$

$$\ln s = -\frac{\Delta H}{k_B}\left(\frac{1}{T} - \frac{1}{T_m}\right)$$

Width of the transition (2)

$$\Delta T = \left(\frac{d\theta}{\mathrm{dT}}\right)^{-1}_{T=T_m} = \frac{d\theta}{\mathrm{ds}}\frac{ds}{\mathrm{dT}}$$

$$= 4\sigma^{1/2} k_B T_m^2 / \Delta H$$

- Allows us to compare with experimental curves. One finds:

  ΔH = 990 cal.mol et $\sigma \approx 2\times 10^{-4}$

Denaturation of poly-γ-benzyl-L-glutamate

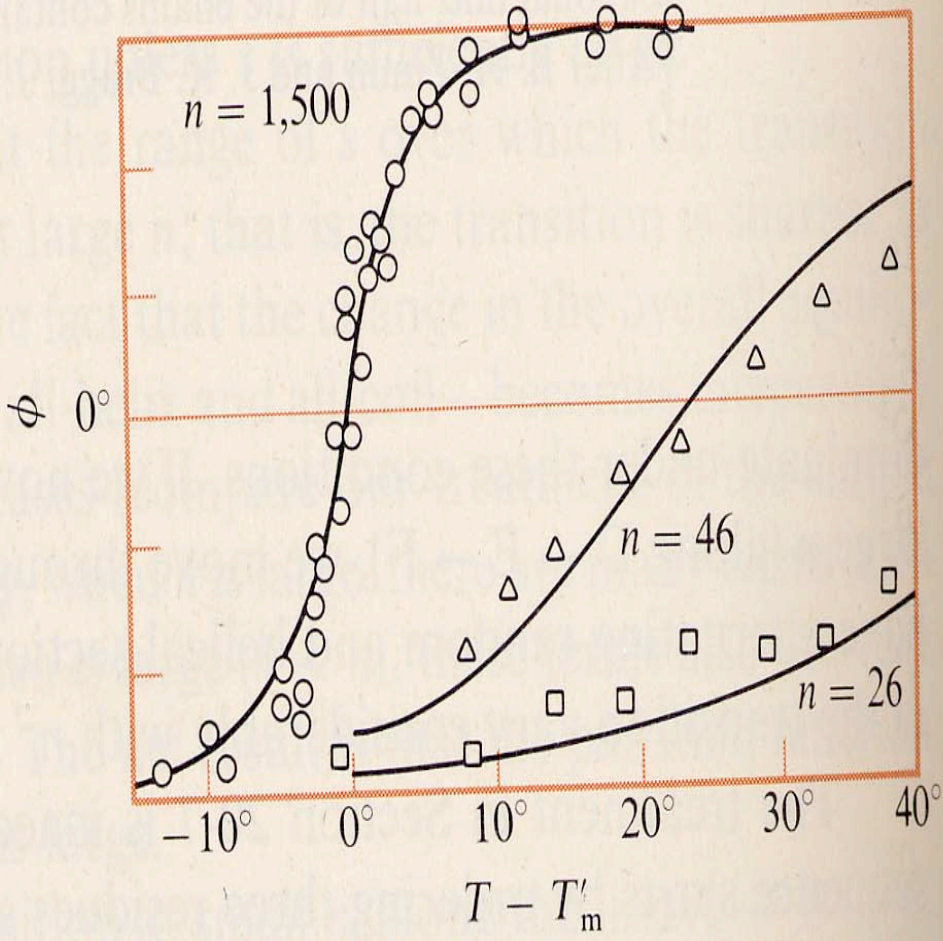


Figure 20-13

Theoretical curves and experimental results. The curves show optical rotation calculated from predicted fractional helicity as a function of $T - T'_m$ for poly-γ-benzyl-L-glutamate chains of three different lengths. The points represent experimental optical-rotation data. [After B. H. Zimm et al., *Proc. Natl. Acad. Sci. USA* 45:1601 (1959).]

Relation with protein folding

- Two major additional difficulties

  1) The sequence must be taken into account.

  2) The nearest neighbor approximation is invalid. This leads to a large increase of the algorithmic complexity of the problem.

Helix-coil transition in dsDNA:

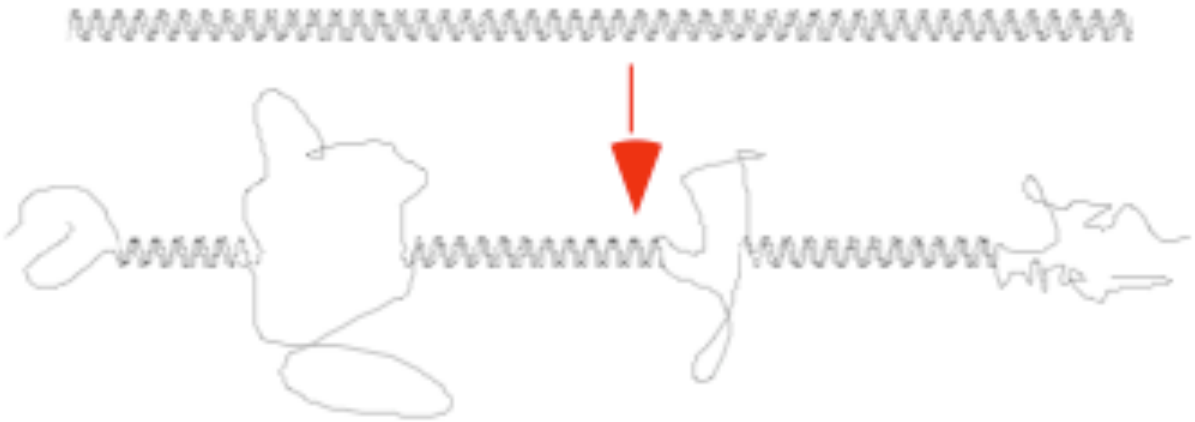
Genes and the physics of DNA Yeramian (2000)

Goal: describe the helix-coil transition in a linear dsDNA chain for a specific (genomic) sequence:

1) A sequence-dependent approach

2) The helix-coil transition beyond the nearest neighbor approximation

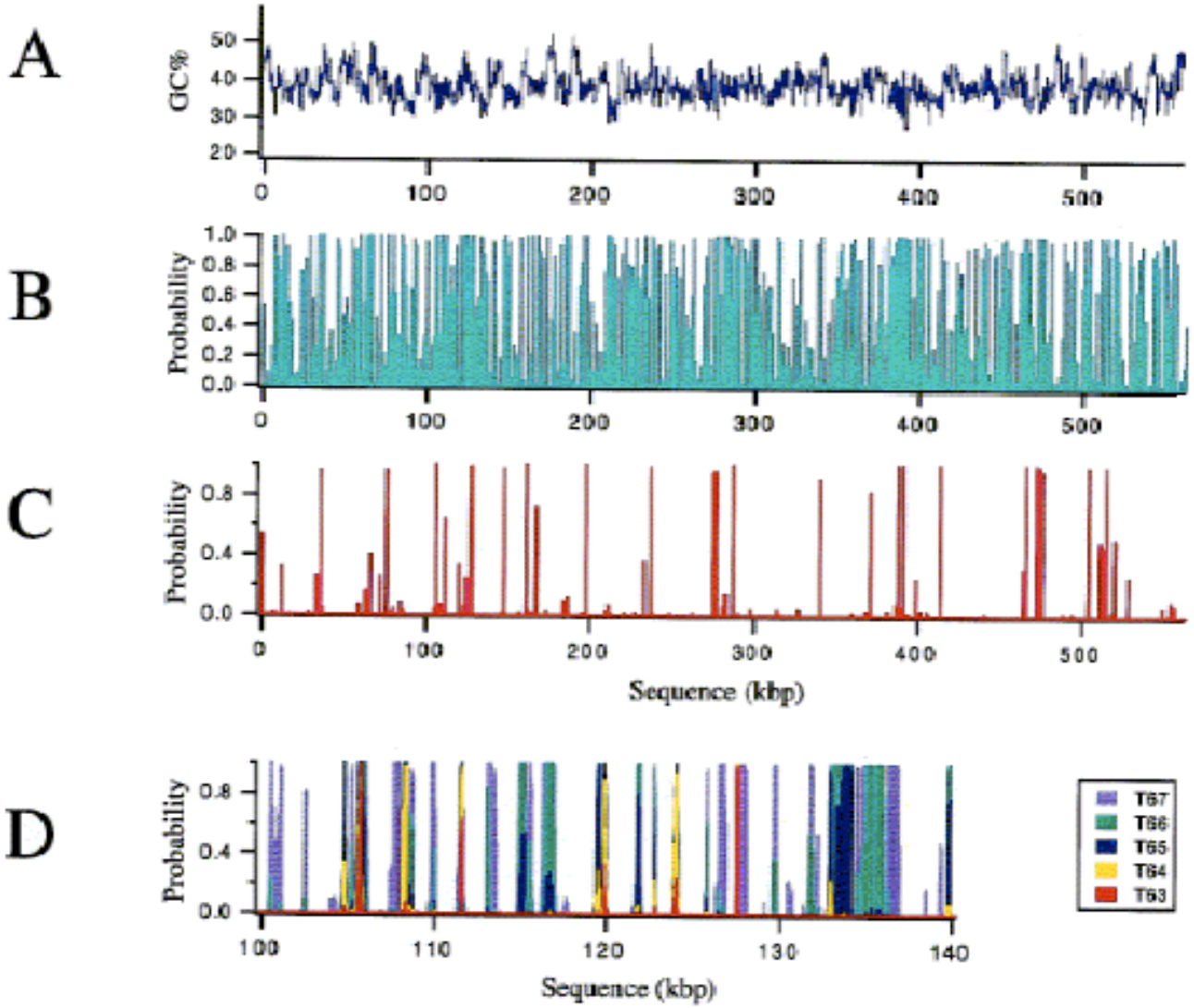
A long-range effect effect due to denaturation loops



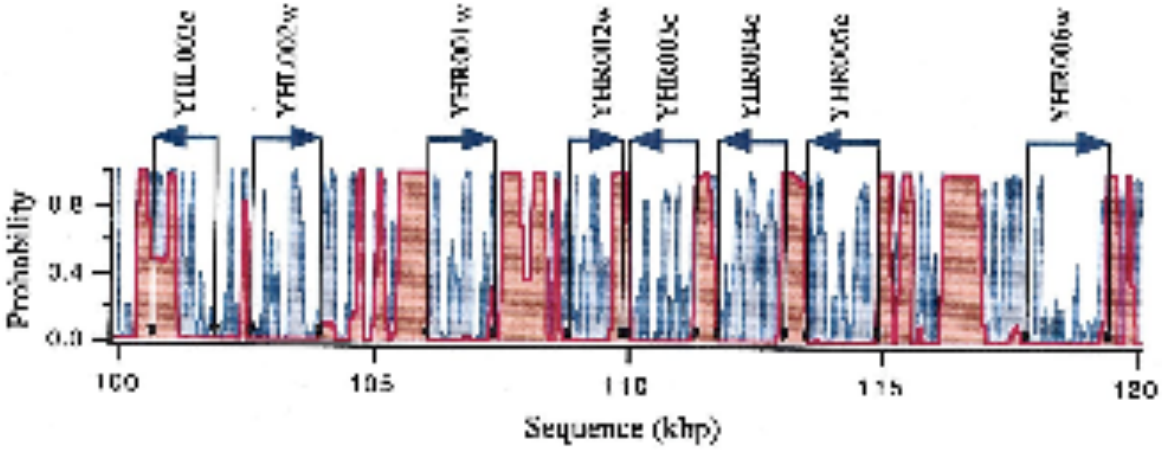
A statistical (entropic) weight associated with a loop of *j* monomers (assuming a random coil state for the monomers in the loop):

$$\omega(j) = \sigma j^{-3/2}$$

Yeast Chromosome VIII



Stability and genetic maps



Polymer (bio)physics and genetics

- A correlation between the (physical) stability map and the genetic map:

  The genes are regions of higher stability

  Not always true, numerous questions to be addressed (split genes with intron/ exon structures…)

Conclusions

- General description of the structure of biopolymers
- Concept of persistence length: a characteristic length of helical biopolymers.
- Elementary statistical mechanics of biopolymers:
- The transfer matrix applied to the helix-coil transition
- A relation between physical stability and genetic sequences in genomic DNA