I. Protein Structure

- A. Primary
  1. Peptide Bond
  2. Determination
- B. Secondary
  1. α-Helix
  2. β-sheets/strands
  3. Connections between β-strands
  4. Super secondary structure
- C. Tertiary
  1. Picturing and classifications
  2. Topology
  3. Domains
  4. Intrinsically disordered
  5. Stability
- D. Quaternary

- Reading: Ch 4; 118-120
  Ch 1; 25-27
  Ch 5; 147-148, 150-152, 157
  Ch 6 ; 177, 209-210
- Homework #12

NEXT

- Reading: Ch 6; 178-179, 194-195, 179-181, 184-185, 186-188

II. Determination of Conformational Structure

- A. Quaternary structure
  1. How determined; Gel filtration & SDS-PAGE, Ultracentrifugation
- B. Tertiary structure
  1. X-ray diffraction/crystallography
  2. NMR spectroscopy
  3. Comparison: NMR *versus* X-ray crystallography
- C. Secondary structure
  1. Circular dichroism (CD)

III. Collagen

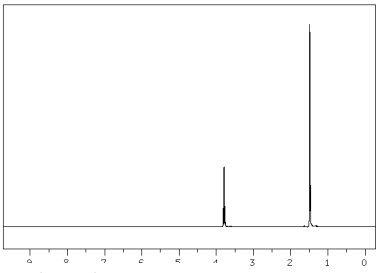
- A. Special Fibrous Protein:
- B. Clues to structure
- C. 4-S's
- D. Biosynthesis
- E. Disorders

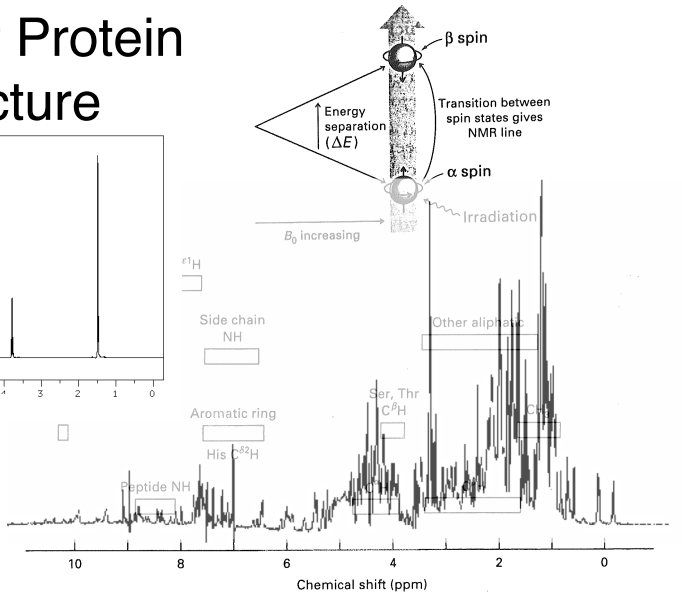
IV. ENZYMES: Binding & Catalysis (catalytic cycle, turnover number, binding reaction)

Lecture 12 (10/6/25)

Determination of the Conformational Structure of Proteins

NMR for Protein Structure





Making the assignments is the most difficult, challenging obstacle for protein-structure determination using NMR!

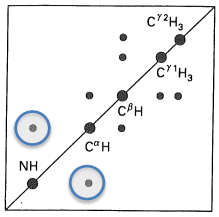
FIGURE 6.26
The range of ^{1}H-NMR chemical shifts observed for different hydrogen atoms of peptides in the random coil conformation.

Determination of the Conformational Structure of Proteins

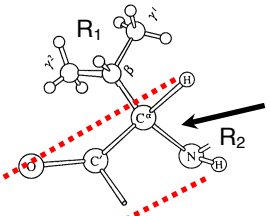
NMR for Protein Structure

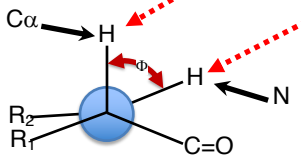
Correlation 2D spectroscopy (COSY)

J-coupling

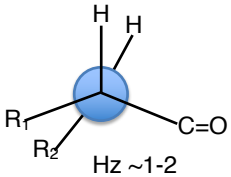


These intensities change depending on ϕ-angle





Hz ~4-5



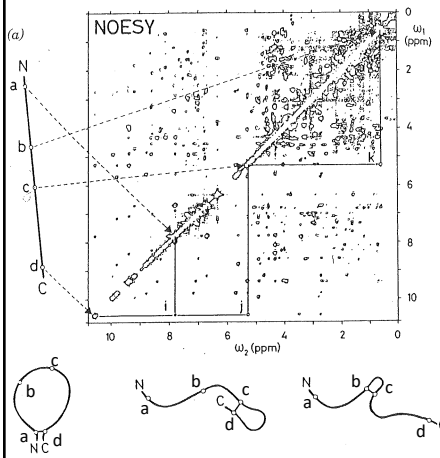
Hz ~1-2

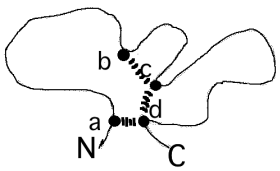
Can also get ϕ & ψ angles from secondary chemical shifts (observed from known random coil shifts)(TALOS & ANGLOR)

Determination of the Conformational Structure of Proteins

NMR for Protein Structure

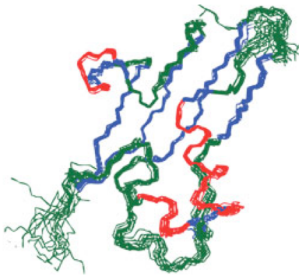
Nuclear-Overhauser Effect Spectroscopy (NOESY)



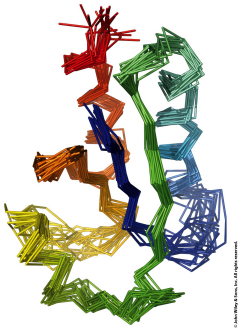


Determination of the Conformational Structure of Proteins

Typical NMR structure



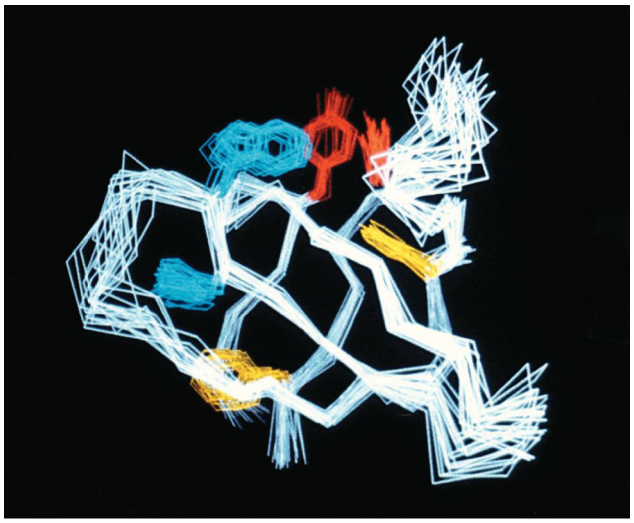
Blue = β-sheet, red = α-helix, green = loops without 2° structure.



Blue = N-term, red = C-term
"Rainbow coloring."

Determination of the Conformational Structure of Proteins

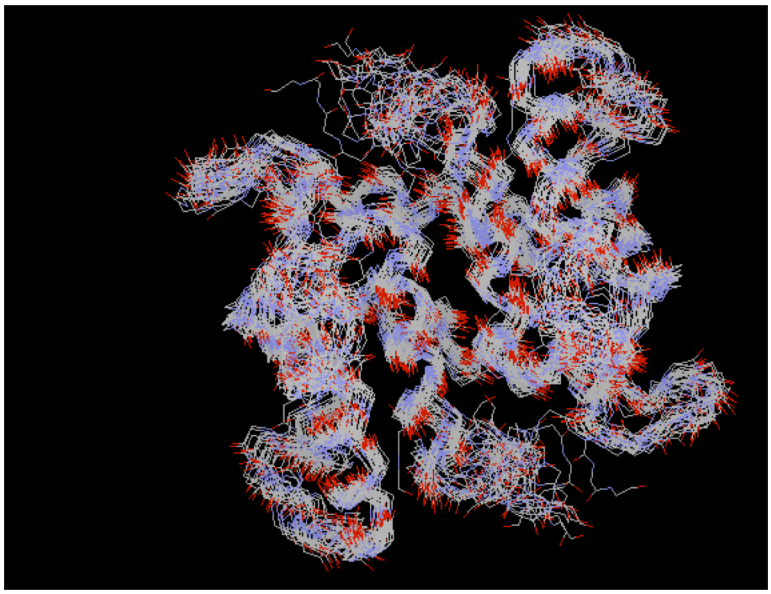
NMR Structure of a Protein



Courtesy Stuart Schrieber, Harvard University

Determination of the Conformational Structure of Proteins

Typical NMR structure



Notice there are many overlapping structures that all fit the NMR data. Where it is tight, you have higher resolution and where it is loose you have parts of the molecule that are more mobile

Determination of the Conformational Structure of Proteins

Tertiary Structure: Summary

Compare/Contrast X-ray crystallography and NMR:

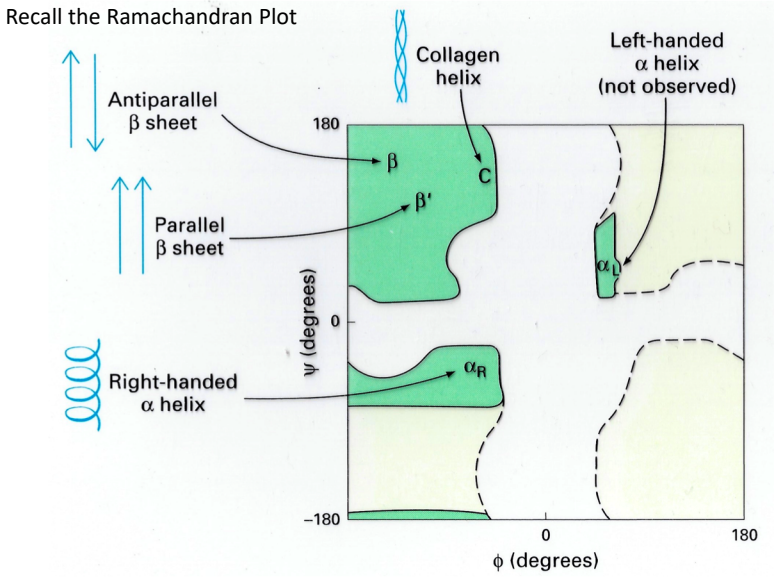
1) Crystal vs. solution structures the same; not significant crystal constraints
2) NMR not as high resolution
3) NMR better at predicting regions that are dynamic; X-ray uses “B-factors” or even does not show, i.e., “disordered”
4) X-ray cannot distinguish “rotomers” of Asn, Gln, Thr; NMR is unambiguous
5) X-ray much better at larger structures; NMR has assignment problem: generally useful for up to 30-40 kDa

Determination of the Conformational Structure of Proteins

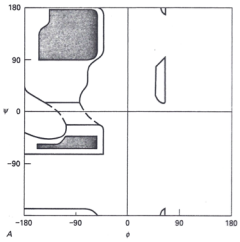
Secondary Structure

82

Determination of the Conformational Structure of Proteins

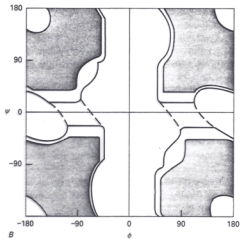


Determination of the Conformational Structure of Proteins

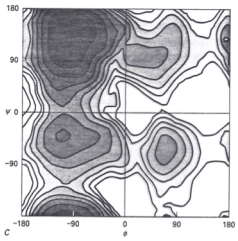


Ala-Ala

(assuming hard spheres)



Gly-Gly



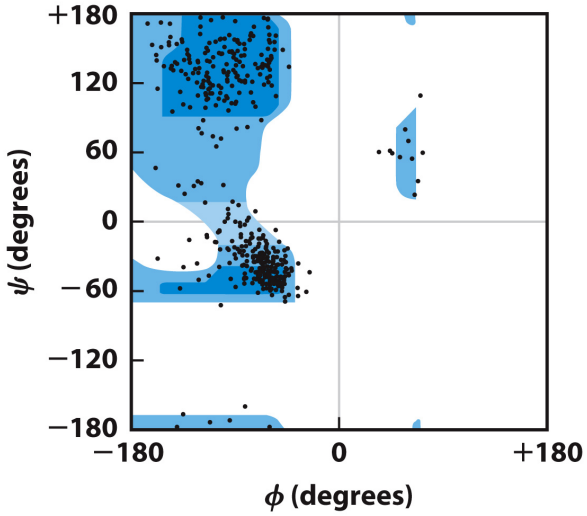
$NAc\text{-}Ala\text{-}Ala\text{-}NH_2$

(calculation with molecular dynamics including solvent (water))

FIGURE 5.2
Ramachandran plots of the permitted values of ϕ and ψ for different residues. Each two-dimensional plot is continuous at the edges, because a rotation of $-180°$ is the same as one of $+180°$. The original plots that considered only repulsions between hard-sphere atoms are shown in *A* and *B* for Ala and Gly residues, respectively. The fully allowed regions are shaded; the partially allowed regions are enclosed by a solid line. The connecting regions enclosed by the dashed lines are permissible with slight flexibility of bond angles. The much greater flexibility of the Gly residue compared with Ala is apparent, as is the symmetry of the plot for Gly residues resulting from the absence of a chiral side chain. (From G. N. Ramachandran & V. Saisekharan, *Adv. Protein Chem.* 23:283–437, 1968.) *C*: Ramachandran plot computed for the dipeptide *N*-acetyl-Ala-Ala-amide using molecular dynamics simulations and including water as the solvent. The apparent free energies for the various values of ϕ and ψ are given as contours of 2 kJ/mol (0.5 kcal/mol) relative to the lowest free energy in the upper left-hand corner that is shaded black. Note that the differences between allowed and disallowed regions are much less distinct than in *A* and *B*. (Figure kindly provided by J. Hermans.)

Determination of the Conformational Structure of Proteins

Ramachandran Plot

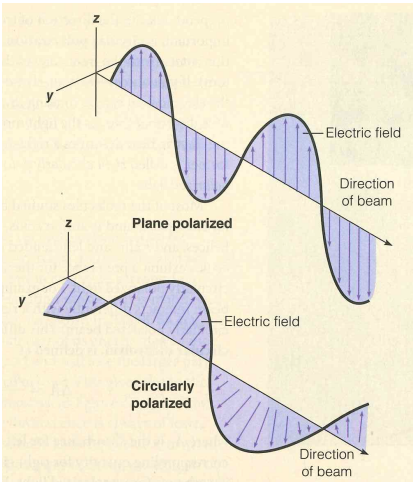


Pyruvate kinase

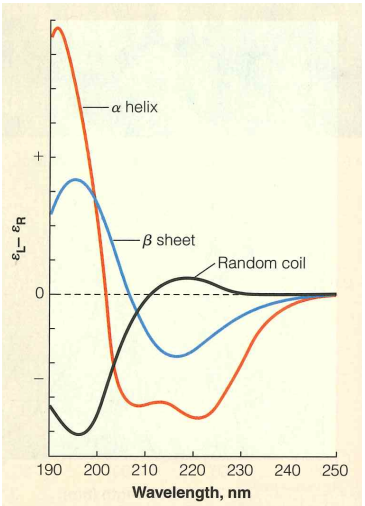
Figure 4-9b
Lehninger Principles of Biochemistry, Seventh Edition

Determination of the Conformational Structure of Proteins

Secondary Structure



(a) Polarization of light. Above: Plane polarized light, in which the amplitude of the electric field oscillates in a single plane. Below: In circularly polarized light, the oscillation of the electric field follows a helical path around the axis describing the direction of the beam.

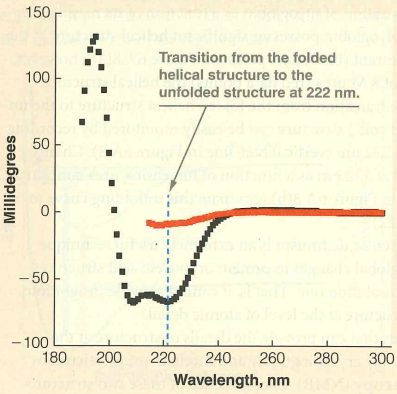


(b) Circular dichroism spectra for polypeptides in various conformations. Here the *y*-axis records differences in molar absorptivity (ε) between left- and right-circularly polarized light.

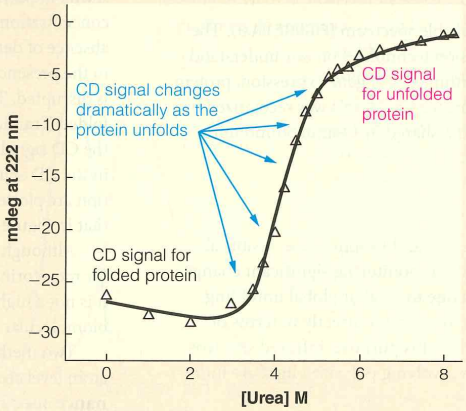
CD Demo:
http://cddemo.szialab.org/

Determination of the Conformational Structure of Proteins

Secondary Structure



(a) Denaturation of myoglobin as a function of increasing urea concentration. The black spectrum was obtained under native conditions. The red spectrum was obtained in the presence of 8 M urea.



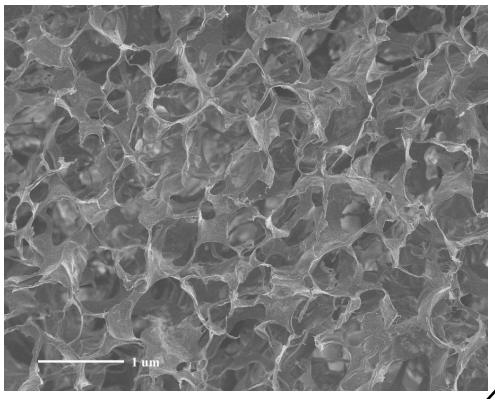
(b) Changes in the CD signal at 222 nm as a function of increasing urea concentration.

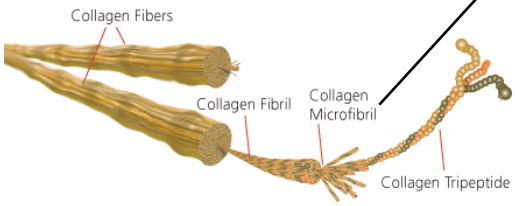
Collagen

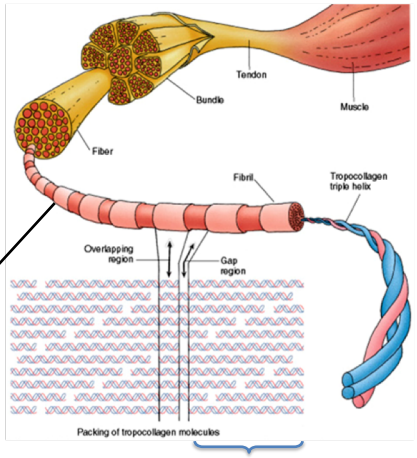
Protein Structure – Collagen

Most abundant protein in mammals

Extracellular: cartilage, tendons, bones, teeth, skin, vessels, lungs







3000 Å

Protein Structure – Collagen

Primary & Secondary structure

Primary structure

Amino Acid Composition of Fibrous Proteins

Amino Acid	α-Keratin (Wool)	Fibroin (Silk)	Collagen (Bovine Tendon)	Elastin (Pig Aorta)
Gly	8.1	44.6	32.7	32.3
Ala	5.0	29.4	12.0	23.0
Ser	10.2	12.2	3.4	1.3
Glu + Gln	12.1	1.0	7.7	2.1
Cys	11.2	0	0	—[e]
Pro	7.5	0.3	22.1[a]	10.7[c]
Arg	7.2	0.5	5.0	0.6
Leu	6.9	0.5	2.1	5.1
Thr	6.5	0.9	1.6	1.6
Asp + Asn	6.0	1.3	4.5	0.9
Val	5.1	2.2	1.8	12.1
Tyr	4.2	5.2	0.4	1.7
Ile	2.8	0.7	0.9	1.9
Phe	2.5	0.5	1.2	3.2
Lys	2.3	0.3	3.7[b]	3.6[d]
Trp	1.2	0.2	0	—[e]
His	0.7	0.2	0.3	—[e]
Met	0.5	0	0.7	—[e]

boxed,

Note: The three most abundant amino acids in each protein are ~~indicated in red~~. Values given are in mole percent.

[a]About 39% of this is hydroxyproline.
[b]About 14% of this is hydroxylysine.
[c]About 13% of this is hydroxyproline.
[d]Most (about 80%) is involved in cross-links.
[e]Essentially absent.

Amino Acid Sequence

$(Gly\text{-}Pro/Ala\text{-}X)_{330}$

X=hydroxyl-Pro*, Glx, Arg, Asx, hydroxyl-Lys (Cδ)*, Ser

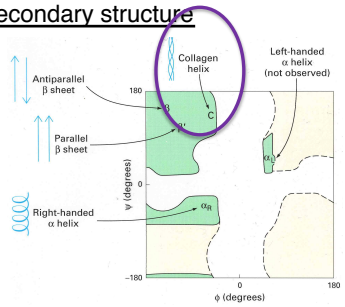
*Hydroxylation reactions catalyzed by proline- and lysine-hydroxylase; require vitamin C

Secondary structure

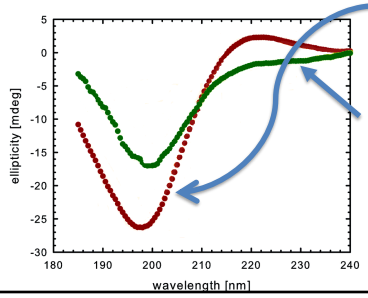
Ramachandran plot

ϕ = -65°

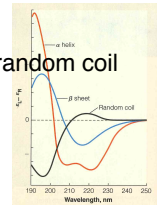
ψ = +130°



CD of collagen



random coil



Protein Structure – Collagen

Amino Acid	α-Keratin (Wool)	Fibroin (Silk)	Collagen (Bovine Tendon)	Elastin (Pig Aorta)
Gly	8.1	44.6	32.7	32.3
Ala	5.0	29.4	12.0	23.0
Ser	10.2	12.2	3.4	1.3
Glu + Gln	12.1	1.0	7.7	2.1
Cys	11.2	0	0	—[e]
Pro	7.5	0.3	22.1[a]	10.7[c]
Arg	7.2	0.5	5.0	0.6
Leu	6.9	0.5	2.1	5.1
Thr	6.5	0.9	1.6	1.6
Asp + Asn	6.0	1.3	4.5	0.9
Val	5.1	2.2	1.8	12.1
Tyr	4.2	5.2	0.4	1.7
Ile	2.8	0.7	0.9	1.9
Phe	2.5	0.5	1.2	3.2
Lys	2.3	0.3	3.7[b]	3.6[d]
Trp	1.2	0.2	0	—[e]
His	0.7	0.2	0.3	—[e]
Met	0.5	0	0.7	—[e]

boxed,

Note: The three most abundant amino acids in each protein are ~~indicated in red~~. Values given are in mole percent.

[a]About 39% of this is hydroxyproline.
[b]About 14% of this is hydroxylysine.
[c]About 13% of this is hydroxyproline.
[d]Most (about 80%) is involved in cross-links.
[e]Essentially absent.

Protein Structure – Collagen

The 4 S's for Collagen:

- Size
- Shape
- Stability
- Solubility

Protein Structure – Collagen

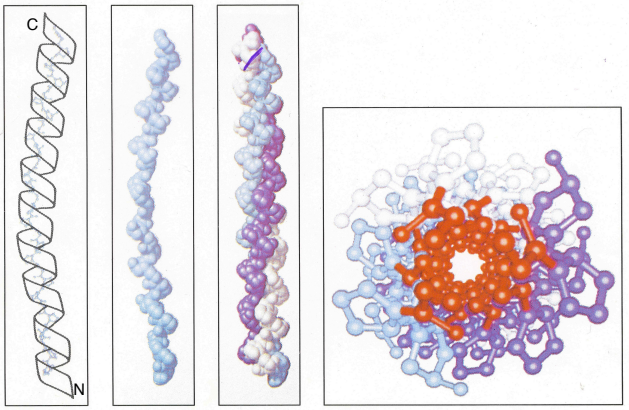
Structure	Φ (°)	Ψ (°)	Rise (Dist/residue) (Å)	Residues/ Repeat	Pitch (Distance/repeat) (Å)	Diameter (Å)
α-helix	-57	-47	1.5	3.6	5.4	5.0
Anti- ⇌ β-sheet	-139	+135	3.4	2	6.8	-
Parallel ⇒ β-sheet	-119	+113	3.2	2	6.4	-
β-turn-Type I				4	0	-
i + 1	-60	-30	-			
i + 2	-90	0	-			
β-turn-Type II				4	0	-
i + 1	-60	120	-			
i + 2	80	0	-			
Collagen	-65	+130	3	3	9	14(triple)

Size

MW = 285,000 Da
Long strands (3000 x 14 Å)
helix dimensions/parameters

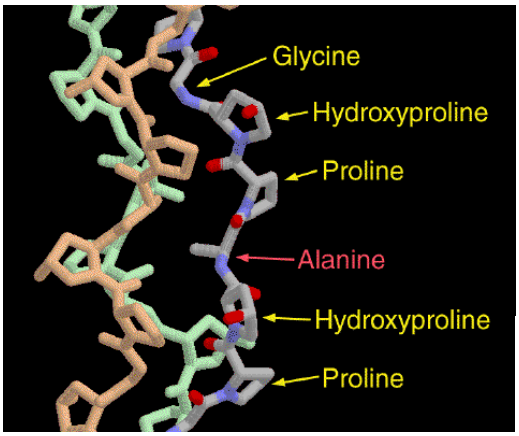
Protein Structure – Collagen

The Collagen Triple Helix

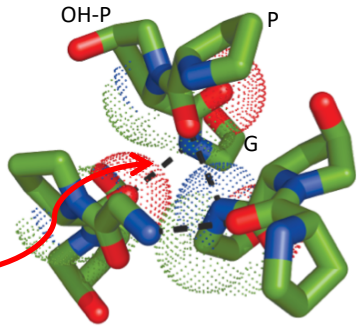


Shape; Collagen-triple helix
- left handed
- right handed twist to three helices
- role of glycine — Allows for packing of triple helix
- role of proline — Locks the helix

Protein Structure – Collagen



Stability
- Packing of Gly
- Inter-stand H-bonds



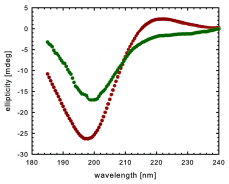
Collagen model peptide
PDBid 1CAG

Protein Structure – Collagen

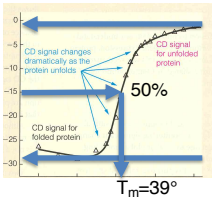
Stability

- Melting (viscosity or CD vs. temp)
- cooperativity = sigmoidal plot
- role of hydroxyl-proline- T_m

Melting of Collagen



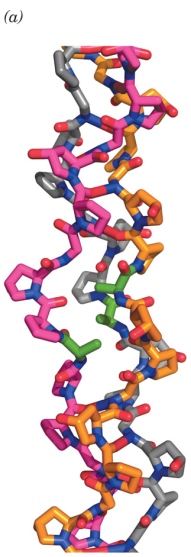
Species	Body Temperature	Collagen T_m
Calf	37	39
Shark, barracuda	26	29
Cod, deep sea redfish	14	16

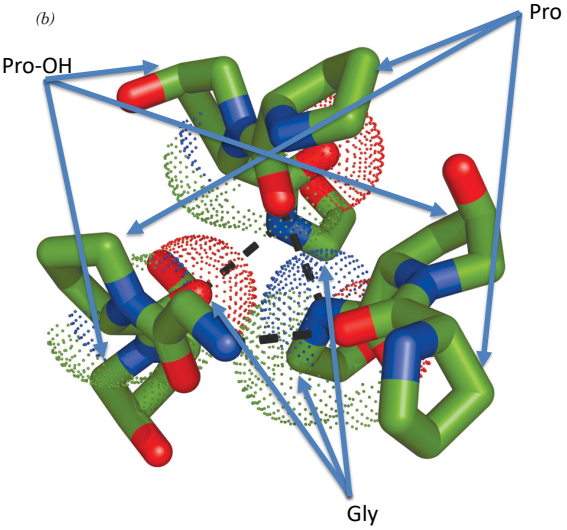


CD *vs.* temp

Higher T_m correlates with higher OH-Pro/Pro

Protein Structure – Collagen





$(Gly\text{-}Pro/Ala\text{-}X)_{330}$ $(Gly\text{-}Pro\text{-}Pro^{hydroxy})$

X=hydroxyl-Pro, Glx, Arg, Asx, hydroxyl-Lys (Cδ), Ser

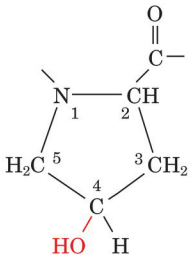
Protein Structure – Collagen

The 4 S's for Collagen:

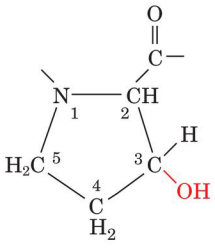
- ✓ Size
- ✓ Shape
- ✓ Stability
- Solubility - not

Protein Structure – Collagen

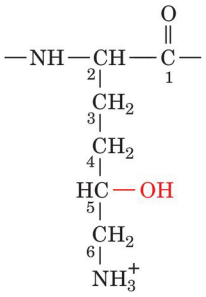
Modified Residues in Collagen & Elastin



4-Hydroxyprolyl residue (Hyp)



3-Hydroxyprolyl residue



5-Hydroxylysyl residue (Hyl)

1.4-9%

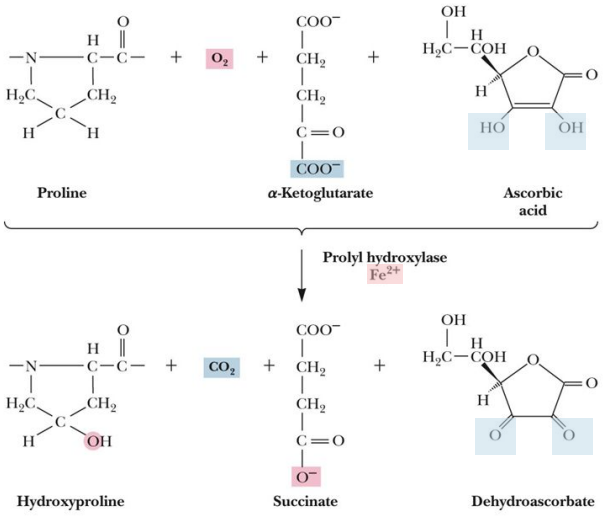
Higher in collagen

0.5-3%

Higher in elastin

Protein Structure – Collagen

Prolylhydroxylase



Substrates get:

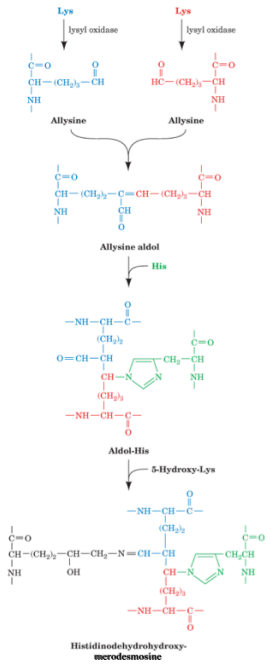
ox

red

Lysylhydroxylase (uses a similar mechnism)

Scurvy: Vitamin C Deficiency

Protein Structure – Collagen

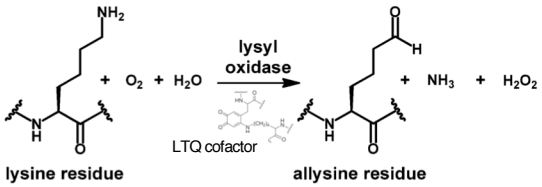


Cross-linking of Collagen

Solubility:

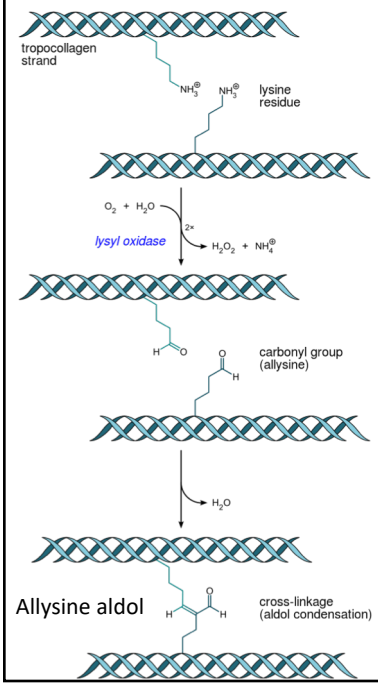
insoluble due to cross-linking between FOUR triple helices using 3 Lys and 1 His residues

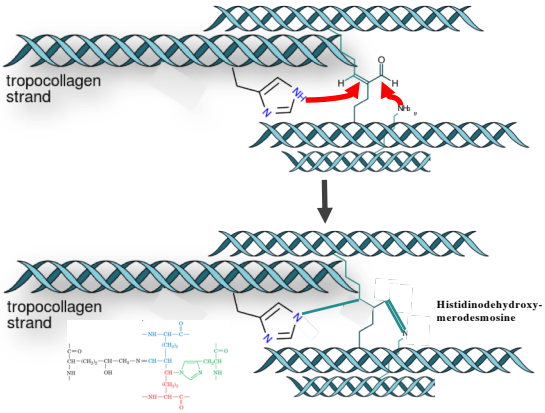
lysyl oxidase → allysine



Protein Structure – Collagen

Cross-linking of Collagen





Modified from Tropokollagen_Quervernetzung_Lysyl_Oxidase_(DE).svg: Sponk (talk)translated by Sponk (talk) - Tropokollagen_Quervernetzung_Lysyl_Oxidase_(DE).svg, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=11183362

Protein Structure – Collagen

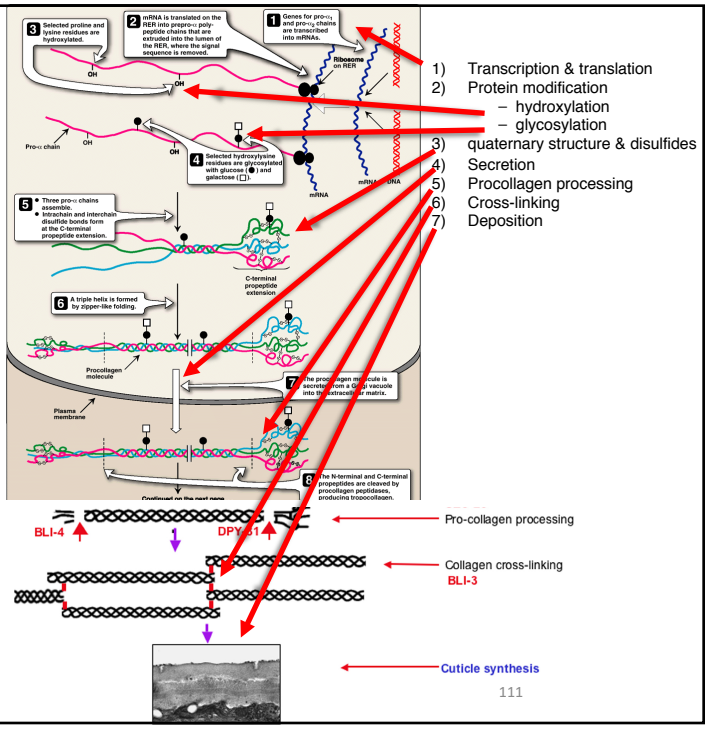
The 4 S's for Collagen:

- ✓ Size: Long strands (3000 x 14 Å; ~300 Kda)
- ✓ Shape: left-handed **triple** helix
- ✓ Stability: controlled by degree of Pro-OH
- ✓ Solubility: - not

How and where is collagen produced? biosynthesis

Protein Structure – Collagen

Collagen Biosynthesis



111

Protein Structure – Collagen

Collagen Disorders

One thousand mutations have been identified in twelve out of more than twenty types of collagen. These mutations can lead to various diseases at the tissue level

Diseases

Scurvy –	caused by lack of Vitamin C needed for hydroxylating enzyme that makes Pro-OH (Hyp) & Lys-OH (Hyl)
Osteogenesis imperfecta –	caused by a mutation in **type 1 collagen** weak bones and irregular connective tissue
Ehlers-Danlos Syndrome –	caused by a mutation in **type 3 collagen (EDS, type 4)** Ten different types of this disorder that lead to deformities in connective tissue; e.g., Rubber man.
Osteoporosis –	Not inherited genetically, brought on with age reduced levels of collagen in the skin and bone
Knobloch syndrome –	Caused by a mutation in the **collagen XVIII gene** protrusion of the brain tissue and degeneration of the retina