Collagen

7.88J Protein Folding

Prof. David Gossard

October 20, 2003

PDB Acknowledgements

The **Protein Data Bank** (PDB - http://www.pdb.org/) is the single worldwide repository for the processing and distribution of 3-D biological macromolecular structure data.

Berman, H. M., J. Westbrook, Z. Feng, G.Gilliland, T. N.Bhat, H.Weissig, I. N.Shindyalov, and P. E.Bourne. "The Protein Data Bank." *Nucleic Acids Research* 28 (2000): 235-242.

(PDB Advisory Notice on using materials available in the archive: http://www.rcsb.org/pdb/advisory.html)

PDB molecules and citations used in the "Collagen" Lecture Notes for 7.88J - Protein Folding

PDB ID: 1CGD

JRNL reference: Bella, J., M. Eaton, B. Brodsky, and H. M. Berman. "Crystal and molecular structure of a collagen-like peptide at 1.9A resolution." *Science* 266 (October 7, 1994):75-8.

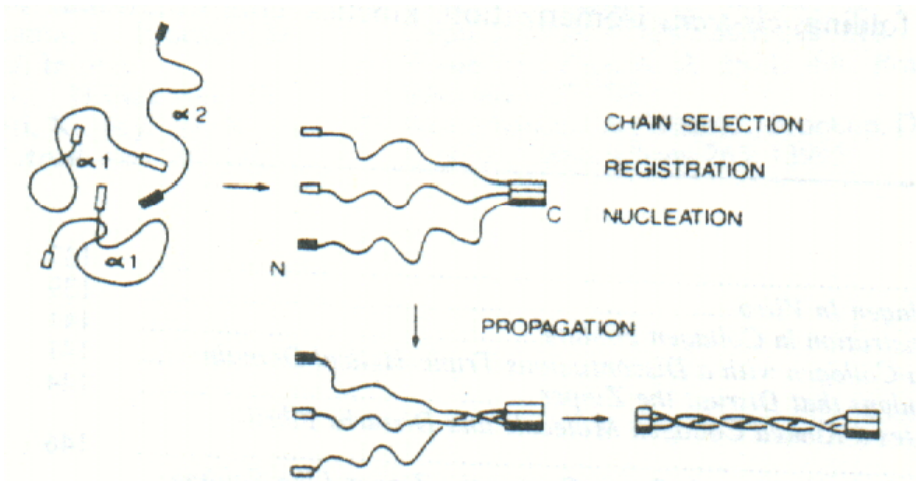
Pages: 16-17 ("Solved Structure – 1CGD")

Collagen

- ~20% of all proteins in human body are collagen
- Extracellular matrix protein family
  - At least 21 different types of collagen
- Structural protein
  - Bone, tendon, cartilage, cornea, etc.
- Mutations in collagen responsible for
  - Osteogenesis imperfecta
  - Hereditary aortic aneurysm

Pathway

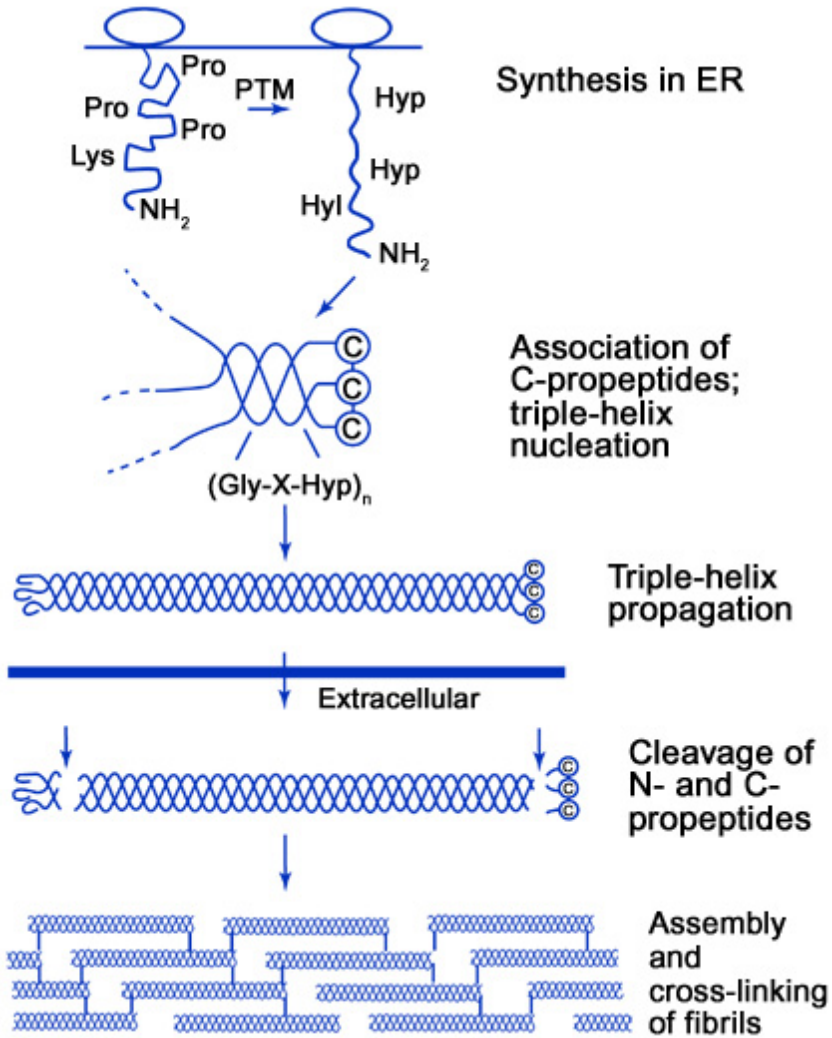
- Synthesized as longer precursors (procollagens) with globular extensions at both ends (propeptides)
- Propeptides form inter-chain disulfide bonds that align the chains prior to triple helix formation



Engel , J., and D. Prockop. "The zipper-like folding of collagen triple helices and the effects of mutations that distrupt the zipper", *Annu. Rev. Biophys. Biphys. Chem* 1991, **20**: 137-52.

Pathway

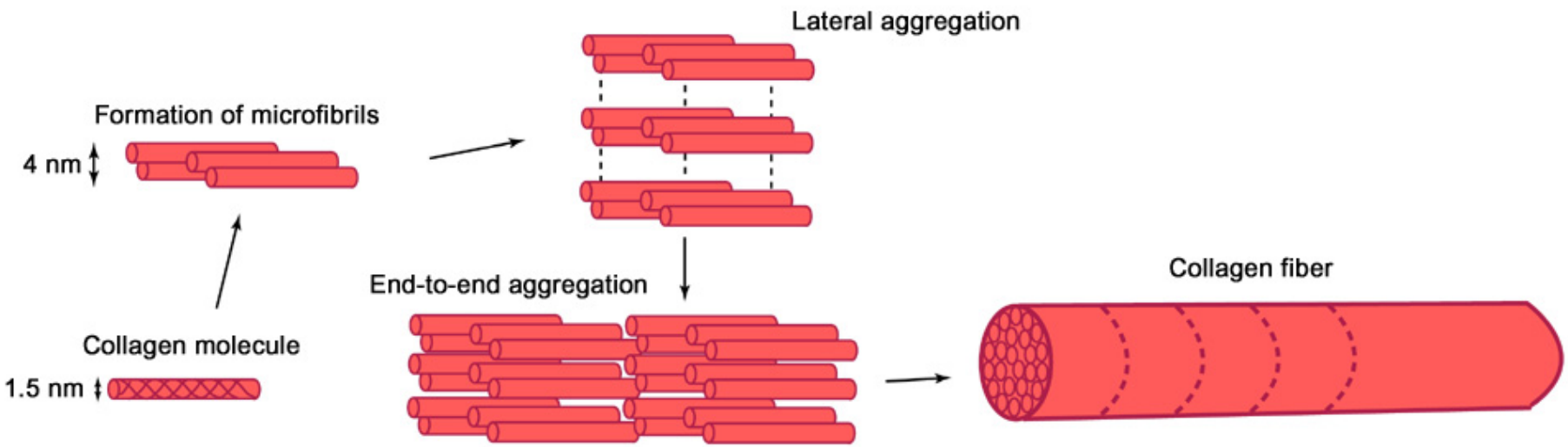
- Following exocytosis:
  - Propeptides cleaved off by extracellular enzymes
  - Triple-helix molecules polymerize into fibrils 50-200 nm long
  - Fibrils pack into fibers (stronger than steel of same size)
- When denatured, forms gelatin (missing propeptides lead to unordered cross-linking)



Higher-Order Structure

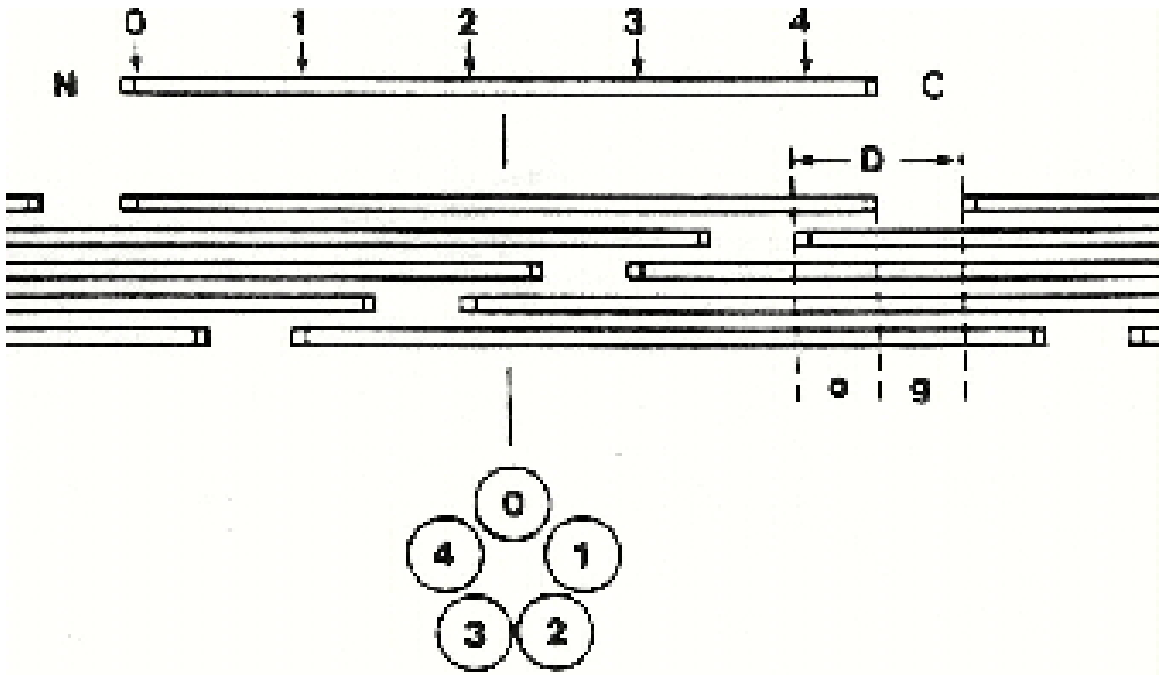
- Levels: molecule => microfibril => fiber

- "Stagger" of molecules gives rise to "banding" observed in fiber



Packing of Collagen Molecules

- 5 molecules/microfibril
- D/2 offset
- Gaps: locus of mineral deposition



Molecular Structure

- Triple helix
  - Rod-like bundle
  - Right-handed supercoil: 100A repeat
  - Length: 2800 A (~1000 aa)
- Each chain is extended, left-handed helix
  - 3.3 residues per turn (3.6 in α-helix)
  - 2.9 A rise per residue (1.5 in α-helix)
  - 9.6 A rise per turn (5.4 in α-helix)
- Helices do not form in isolation
- Chains are staggered by one residue

Branden and Tooze

Sequence

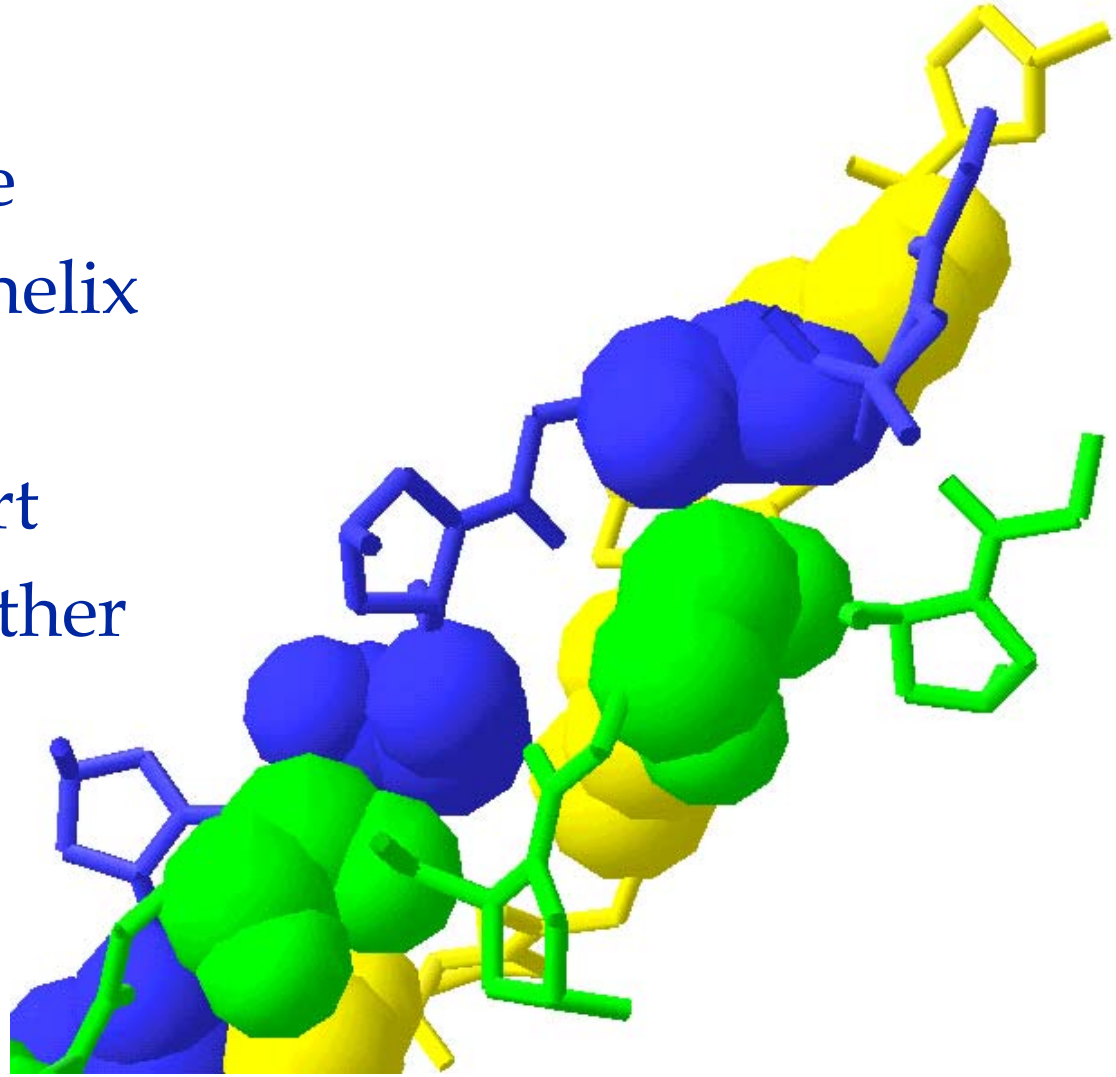
- Glycine at every third residue
  - $(Gly\ –X – Y)_n$
  - X often proline (Pro)
  - Y often hydroxyproline (Hyp)
    - In collagen Pro & Hyp constitute about 20% (40% ?) of all residues
- Glycine at center of triple helix
- Pro & Hyp side-chains fully exposed to solvent

The corresponding figure that illustrates these points may be found in:

Introduction to protein structure / Carl Branden, John Tooze. New York : Garland Pub., 1991.

Glycines

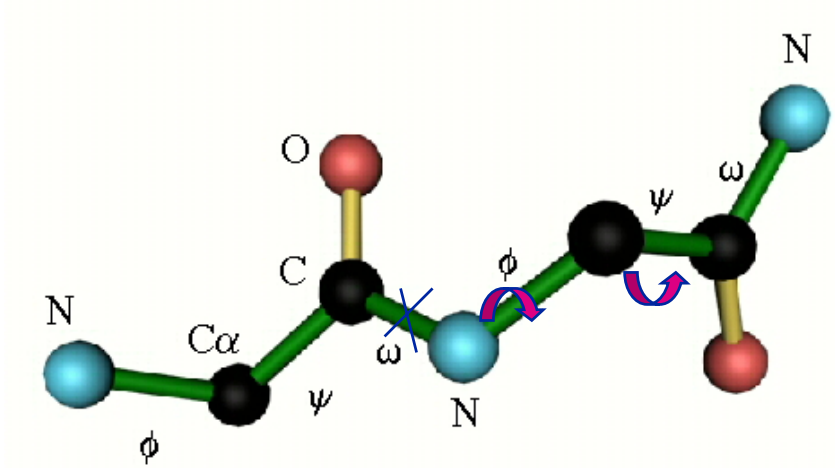
- On the inside of the triple helix
- Staggered wrt to Gly's on other chains

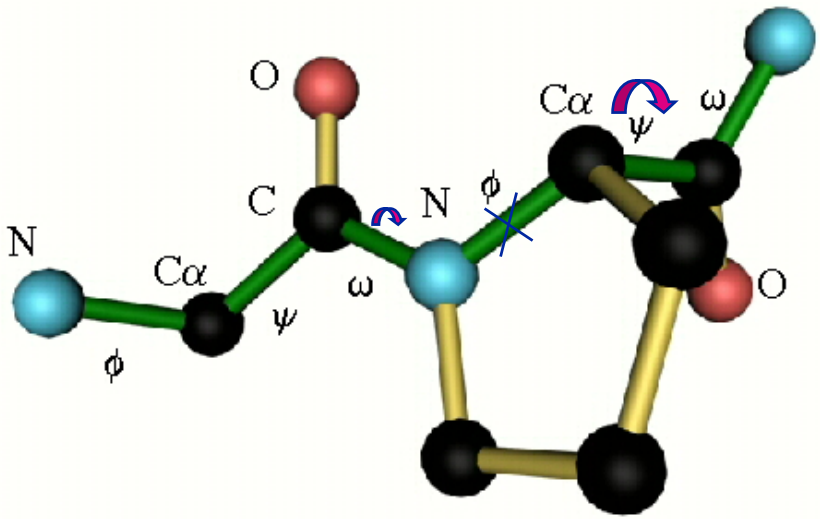


Proline (Pro)

- Increases "stiffness" of chain
  - Eliminates one rotational degree of freedom (ψ–angle)
  - Slightly increases another (ω-angle)

- Promotes an extended (not globular) conformation

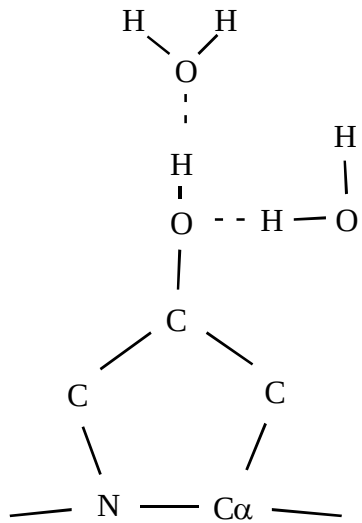




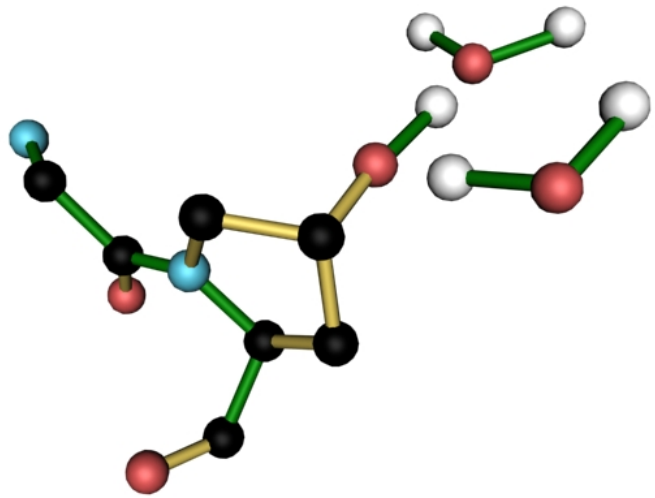
Key Features of Collagen

- High content of Hyp residues
- Unique interaction with water

Hydroxyproline (Hyp)



- Produced from Pro by post-translational modification
- Is unusual, always found in triple helix domains in animal proteins, rarely in other proteins
- Provides binding sites for water molecules
  - Important to stability



Hydrogen Bonding Pattern

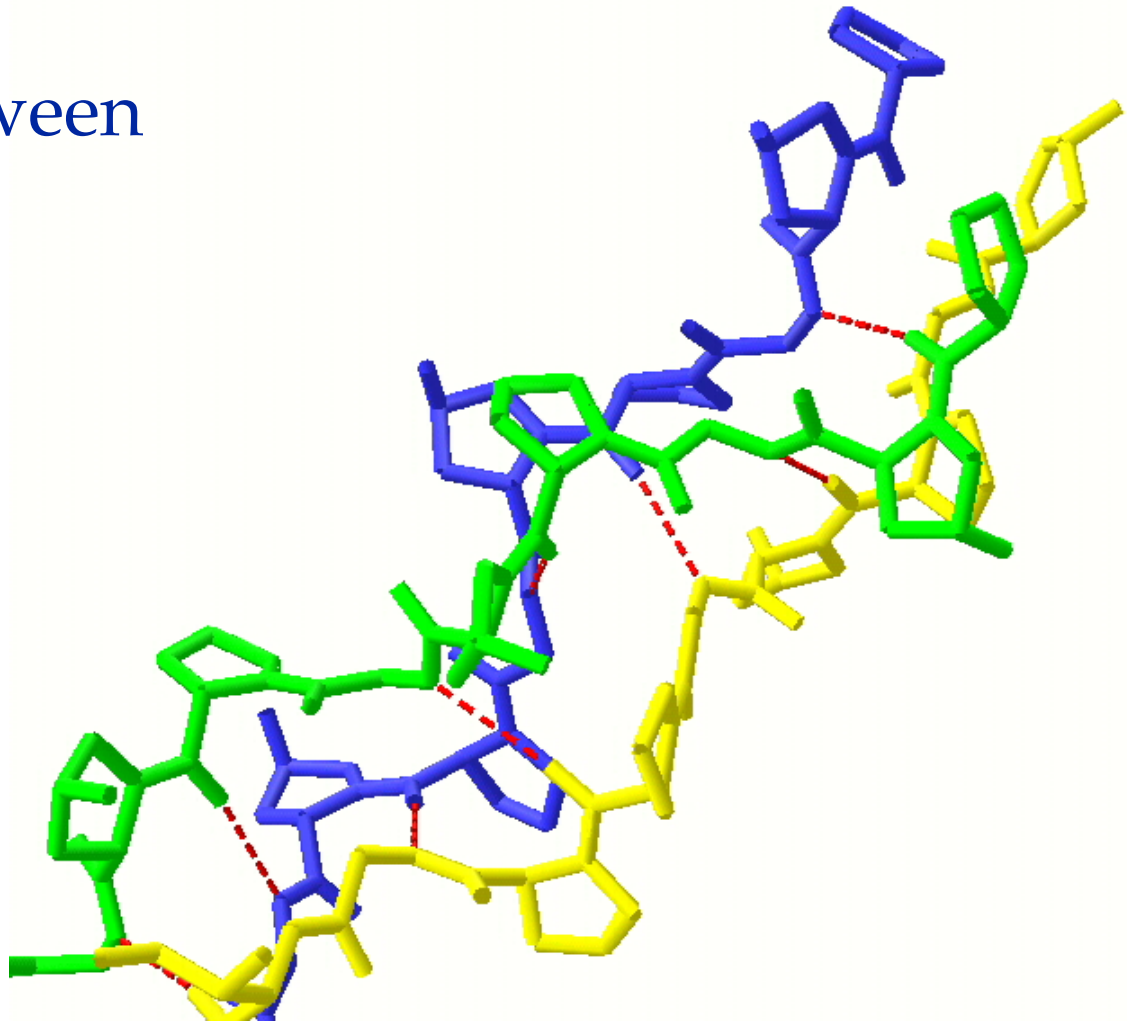
- H-bonds between
  - Gly N - H
  - Pro C=O

The corresponding figure that illustrates these points may be found in:

Introduction to protein structure / Carl Branden, John Tooze. New York : Garland Pub., 1991.

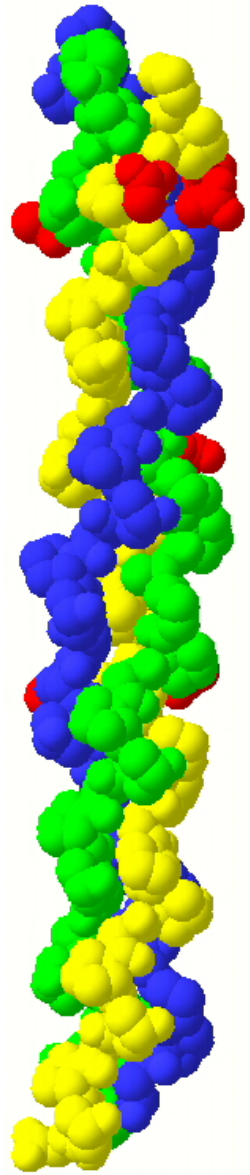
3D Hydrogen Bonding Pattern

- H-bonds between
  - Gly N - H
  - Pro C=O



Principles of icosahedral virus structure published by Caspar and Klug, Cold Spring Harbor Laboratory Press, Symp. Quant. Biol. vol 27, 1962.

Solved Structure –1CGD



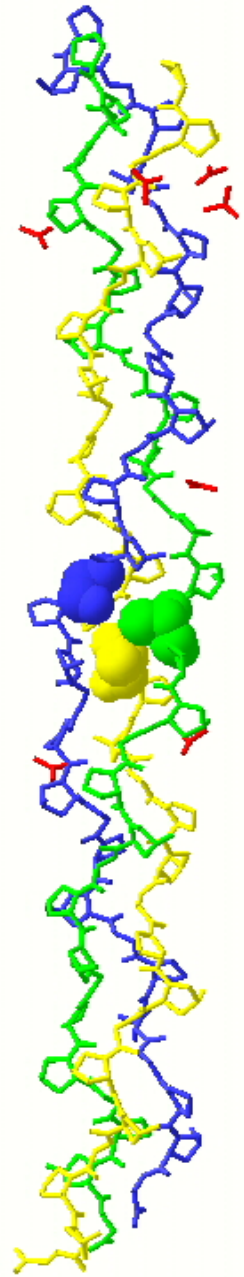
- Gly -> Ala
  (single aa substitution, all 3 chains)
- Formed crystals instead of fibrils
- 3 chains – triple helix
- 30 residues (not 1000 !)
- Tm: 62 °C -> 29 °C !!!

Bella, J., M. Eaton, B. Brodsky, and H. M. Berman. "Crystal and molecular structure of a collagen-like peptide at 1.9A resolution." *Science* **266 (**October 7, 1994) :75-81.

Principles of icosahedral virus structure published by Caspar and Klug, Cold Spring Harbor Laboratory Press, Symp. Quant. Biol. vol 27, 1962.

Solved Structure – 1CGD

	Chain 1				**Chain 2**				**Chain 3**		
1	Pro	Hyp	Gly	31	Pro	Hyp	Gly	61	Pro	Hyp	Gly
4	Pro	Hyp	Gly	34	Pro	Hyp	Gly	64	Pro	Hyp	Gly
7	Pro	Hyp	Gly	37	Pro	Hyp	Gly	67	Pro	Hyp	Gly
10	Pro	Hyp	Gly	40	Pro	Hyp	Gly	70	Pro	Hyp	Gly
13	Pro	Hyp	**Ala**	43	Pro	Hyp	**Ala**	73	Pro	Hyp	**Ala**
16	Pro	Hyp	Gly	46	Pro	Hyp	Gly	76	Pro	Hyp	Gly
19	Pro	Hyp	Gly	49	Pro	Hyp	Gly	79	Pro	Hyp	Gly
22	Pro	Hyp	Gly	52	Pro	Hyp	Gly	82	Pro	Hyp	Gly
25	Pro	Hyp	Gly	55	Pro	Hyp	Gly	85	Pro	Hyp	Gly
28	Pro	Hyp	Gly	58	Pro	Hyp	Gly	88	Pro	Hyp	Gly



Principles of icosahedral virus structure published by Caspar and Klug, Cold Spring Harbor Laboratory Press, Symp. Quant. Biol. vol 27, 1962.

Hydrogen Bonding Pattern - 1CGD

- At Gly -> Ala substitution site
  - triple helix "unwinds" slightly
  - H-bonds broken
  - 4 Water molecules establish bridges between the groups with the broken H-bonds

The corresponding figure that illustrates these points may be found in:

Introduction to protein structure / Carl Branden, John Tooze. New York : Garland Pub., 1991.

Types of Water Bridges

Table 2 may be found on page 896 of:

Bella, J., B. Brodsky, and H. M. Berman. “Hydration structure of a collagen peptide.” *Structure* **3 (**September 15, 1995**)**: 893-906.

Location of Water Bridges

Figure 2 may be found on page 897 of:

Bella, J., B. Brodsky, and H. M. Berman. “Hydration structure of a collagen peptide.” *Structure* **3 (**September 15, 1995**)**: 893-906.

Interchain & Intrachain Water Bridges

- A: interchain,
  1 water molecule
- B: interchain,
  2 water molecules
- C: Intrachain,
  3 water molecules
- D: Intrachain,
  network of water

Figures may be found in:
Bella, J., M. Eaton, B. Brodsky, and H. M. Berman. “Crystal and molecular structure of a collagen-like peptide at 1.9A resolution.” *Science* **266 (**October 7, 1994**)**: 75-81.

Water Bridges

Images may be found in:

Bella, J., B. Brodsky, and H. M. Berman. “Hydration structure of a collagen peptide.” *Structure* **3 (**September 15, 1995**)**: 893-906.

Hydration Shell

Image may be found in:

Bella, J., Brodsky, B. and Berman, H.M., “Hydration structure of a collagen peptide”, Structure, **3**:893-906, September 15, 1995.

Packing of Triple Helices

- Hexameric
- Anti-parallel
- Separation distance between axes (14 A) too large for direct contact
- Water matrix

Corresponding image may be found in:
Bella, J., M. Eaton, B. Brodsky, and H. M. Berman. “Crystal and molecular structure of a collagen-like peptide at 1.9A resolution.” *Science* **266 (**October 7, 1994**)**: 75-81.

Molecular Surface