

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

Protein Review

Lec'03'slides

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

1 Protein structure

Why do we care?

Primary Structure & Amino Acids

Secondary Structure

Tertiary Structures & Domains

Quaternary Structure

2 Summary

3 Reading for next class

Protein structure

Why do we care?

Primary Structure & Amino Acids

Secondary Structure

Tertiary Structures & Domains

Quaternary Structure

Summary

Reading for next class

1 Protein structure

Why do we care?

Primary Structure & Amino Acids

Secondary Structure

Tertiary Structures & Domains

Quaternary Structure

2 Summary

3 Reading for next class

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

1 Protein structure

Why do we care?

Primary Structure & Amino Acids

Secondary Structure

Tertiary Structures & Domains

Quaternary Structure

2 Summary

3 Reading for next class

We want to predict 3D structures

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

Given AA sequence:

... lysine—arginine—glycine ...

Predict: 3D Protein Structure

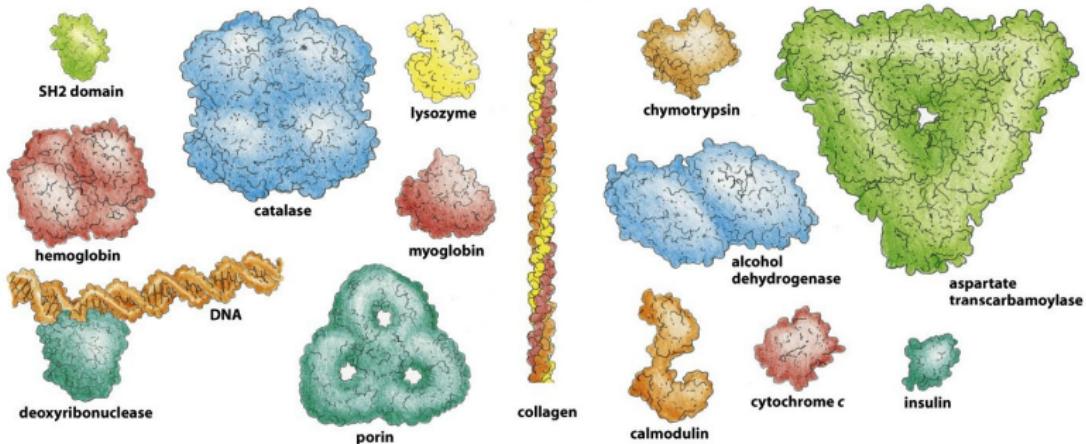


Figure 3-23 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Why do we care about protein shape? You're designing a new drug

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

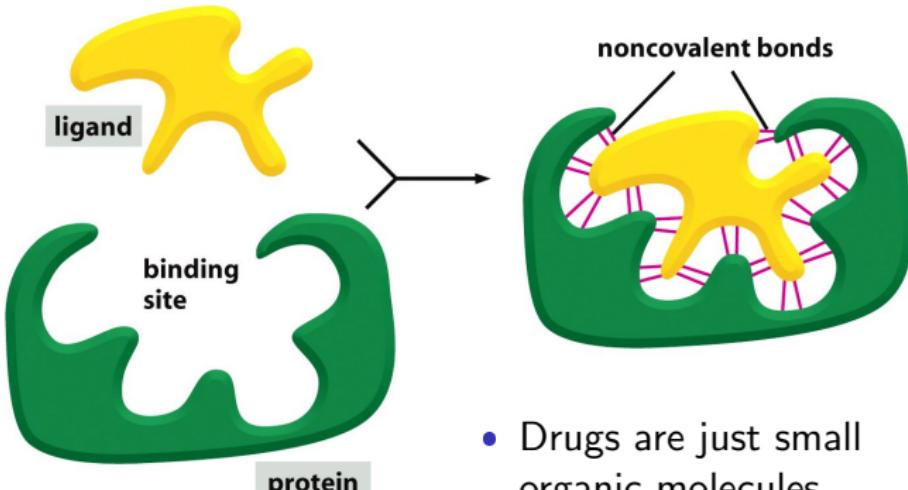


Figure 3-36 Molecular Biology of the Cell 5/e (© Garland Science 2008)

- Drugs are just small organic molecules
- Drugs mesh with proteins
- Chemists design shapes of molecules

Why do we care about protein shape?

Recall: protein-protein interactions

Protein structure

Why do we care?

Primary Structure & Amino Acids

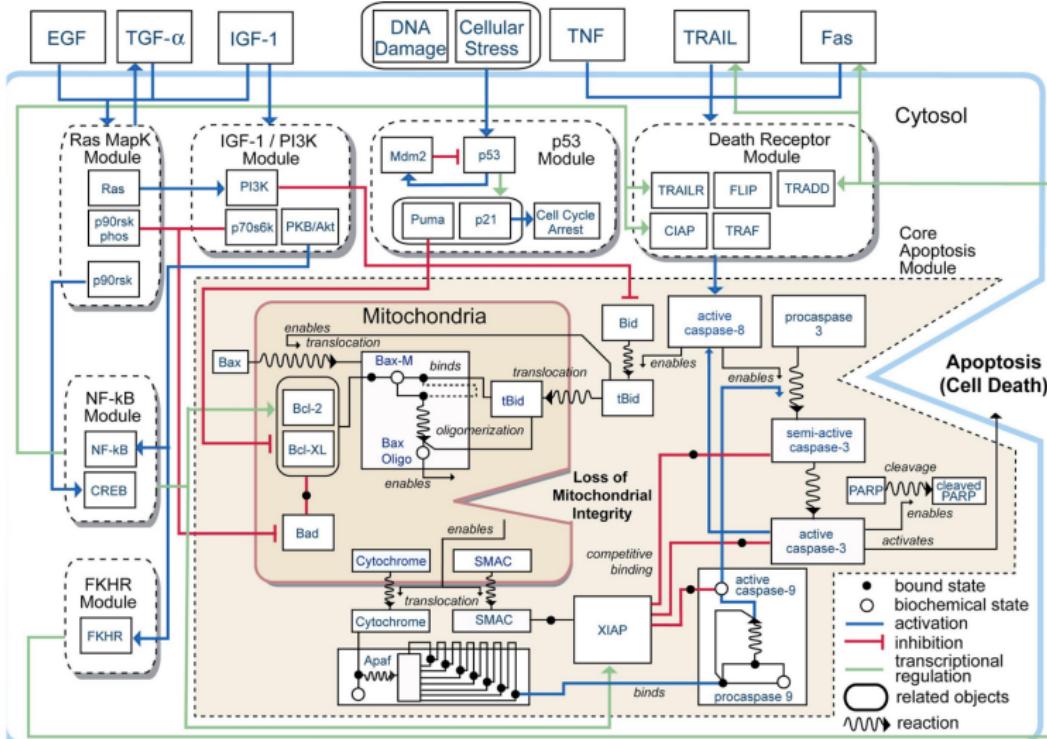
Secondary Structure

Tertiary Structures & Domains

Quaternary Structure

Summary

Reading for next class



X-ray diffraction is traditional solution

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

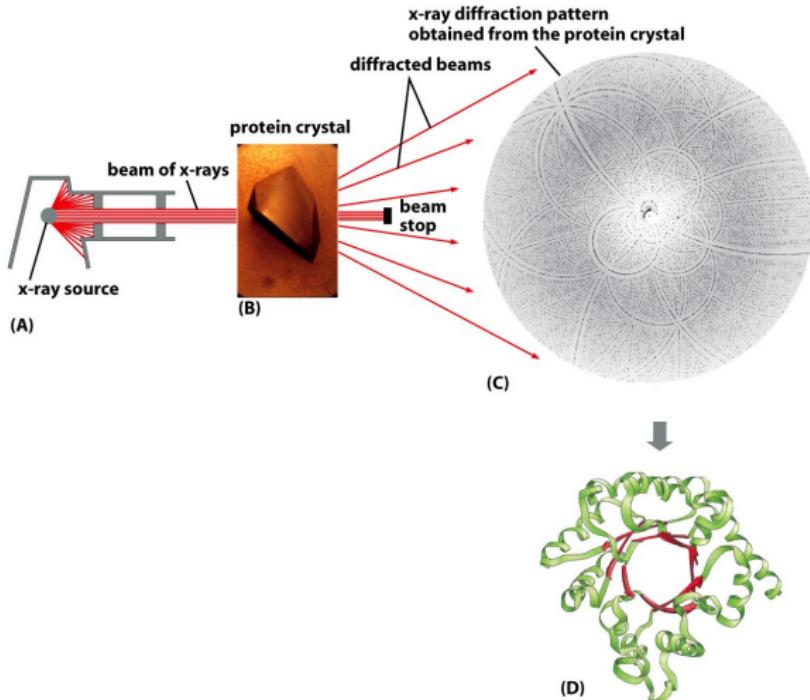


Figure 8-28 Molecular Biology of the Cell 5/e (© Garland Science 2008)

But X-ray diffraction has problems

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

- Protein crystals are squishy
 - Proteins crystals must retain ~40-60% H₂O
 - Not well ordered like hard, dry NaCl crystals.
 - So: **Imperfect ordering within crystal.**
- Post-translational glycosylation is inconsistent:
heterogeneous population → structure internally inconsistent.
- Limited resolution ~1.5 to 3 Å
 - 1 Å = 10^{-10} meter = 100 pm = 0.1 nm
 - X-ray wavelength/resolution: ~1.5 Å
 - Typical covalent bond length: 1.1 to 1.5 Å
 - Diameter of an H atom: 1 Å

Many important proteins do not crystallize

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

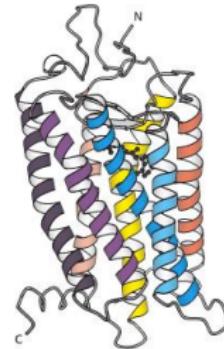
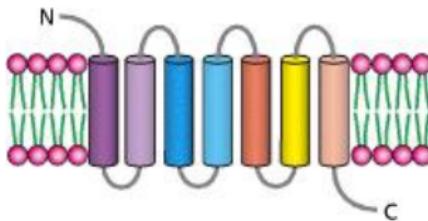
Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class



- No crystal → no X-Ray crystallography
- Example: G-protein coupled receptors (GPCRs)
 - Transmembrane proteins tricky to crystallize
 - **GPCRs are targets of 40% of current pharmaceuticals**

Why?

Nuclear Magnetic Resonance

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

- Nuclear magnetic resonance (NMR) can elucidate the structure of non-crystallized proteins.
- Proteins for NMR can exist in solution.
- Traditionally limited to ~20 kDaltons, recently improved to ~150kd.
- Poorer resolution than X-ray crystallography.

The fundamental problem

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

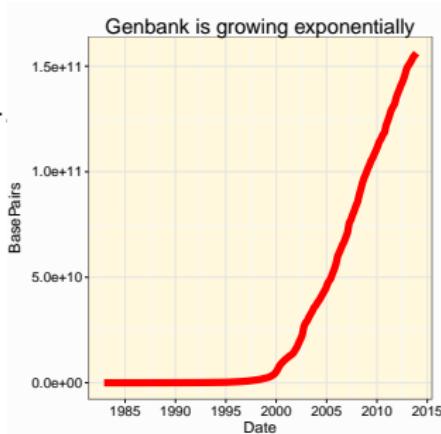
Summary

Reading for
next class

"Over the past five years, gene sequences have been discovered at a rate approximately 65 times faster than experimentalists have been able to determine the structure of the proteins they represent."

That quote is from 2004.
It's much worse now.

Recall:



Bioinformatics tools visualize protein structures

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

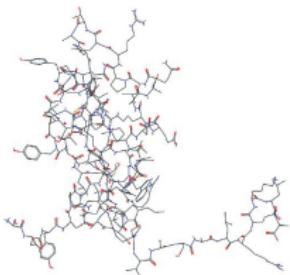
Secondary
Structure

Tertiary
Structures &
Domains

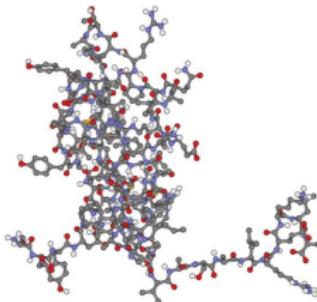
Quaternary
Structure

Summary

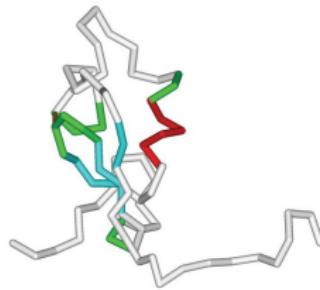
Reading for
next class



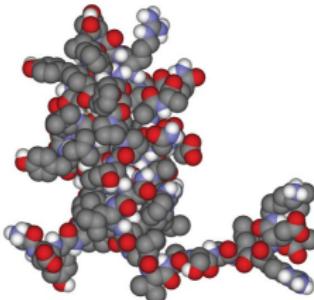
wire-frame



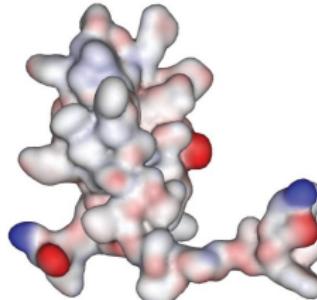
ball and stick



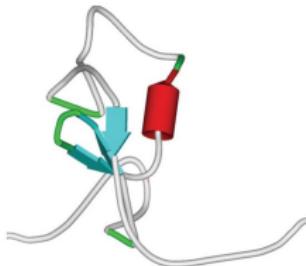
C_α representation



space-filling



surface



α/β schematic

Four main levels of protein structure

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

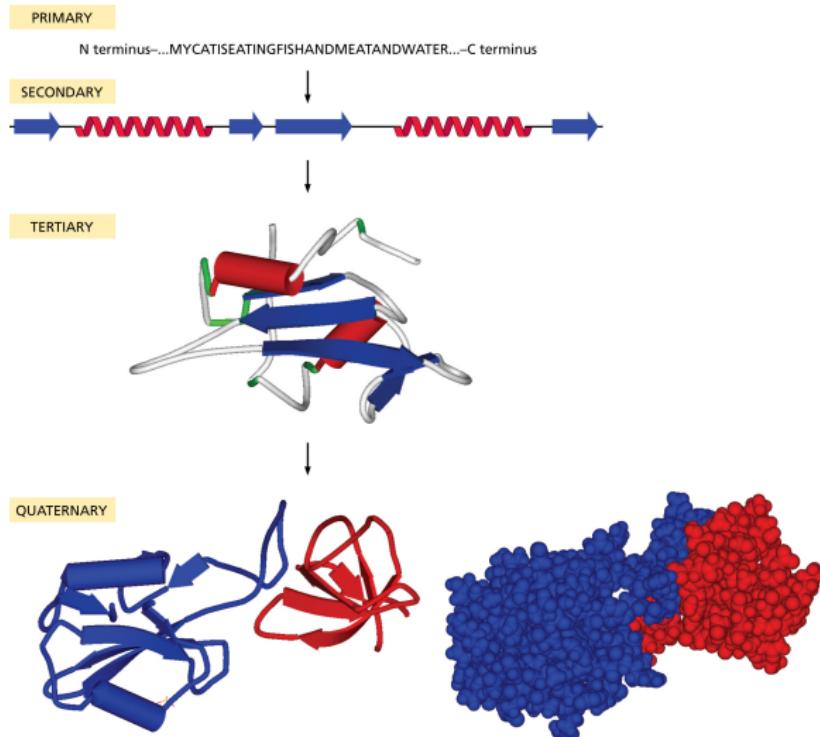
Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class



Another view of protein structure levels

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

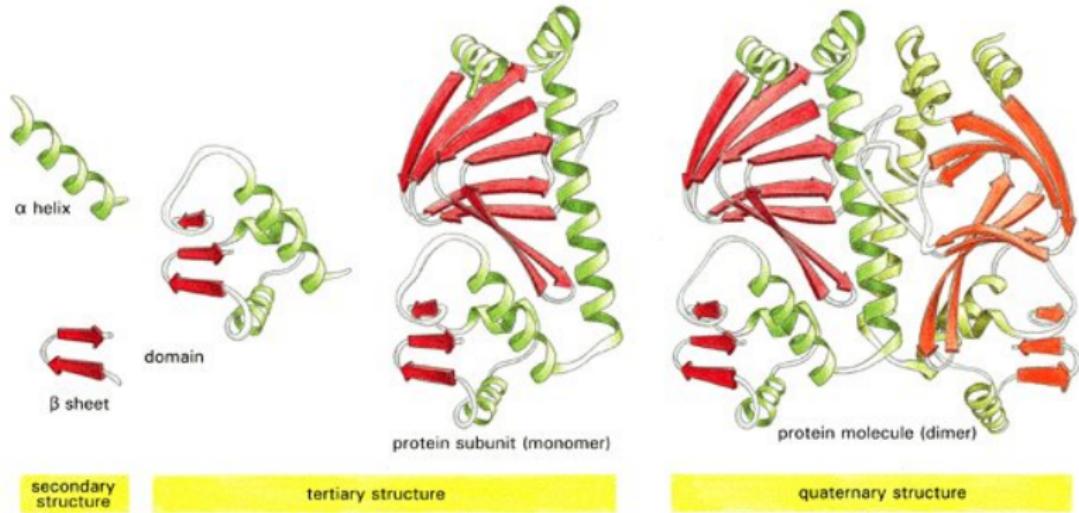
Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class



Note domain structure

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

1 Protein structure

Why do we care?

Primary Structure & Amino Acids

Secondary Structure

Tertiary Structures & Domains

Quaternary Structure

2 Summary

3 Reading for next class

Protein primary structure

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

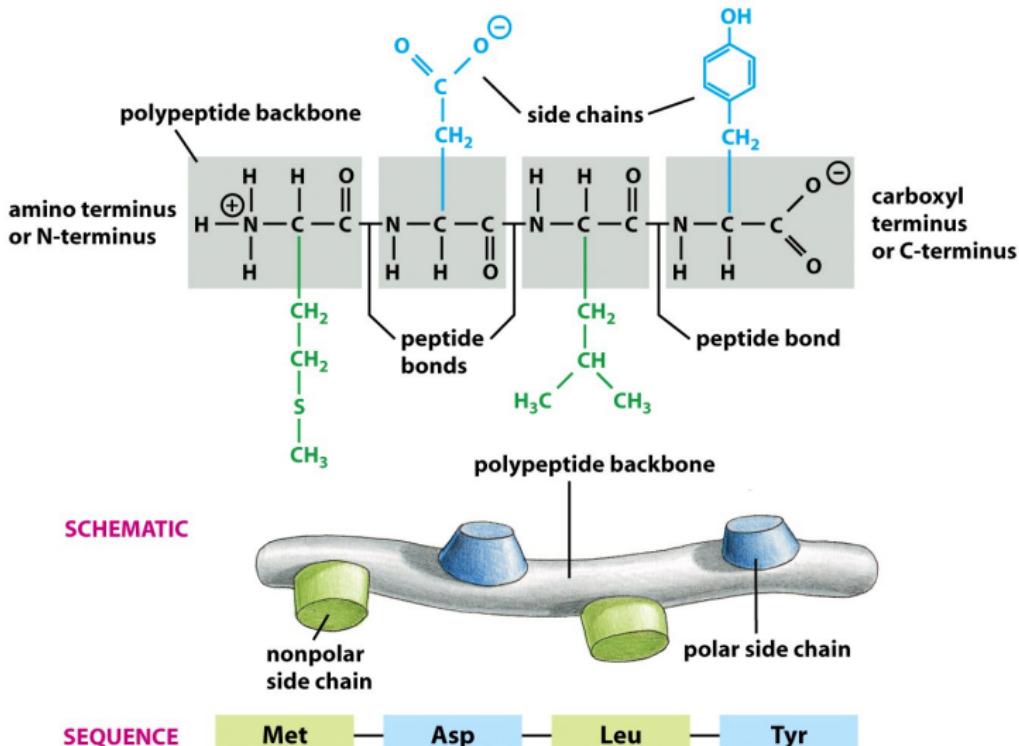
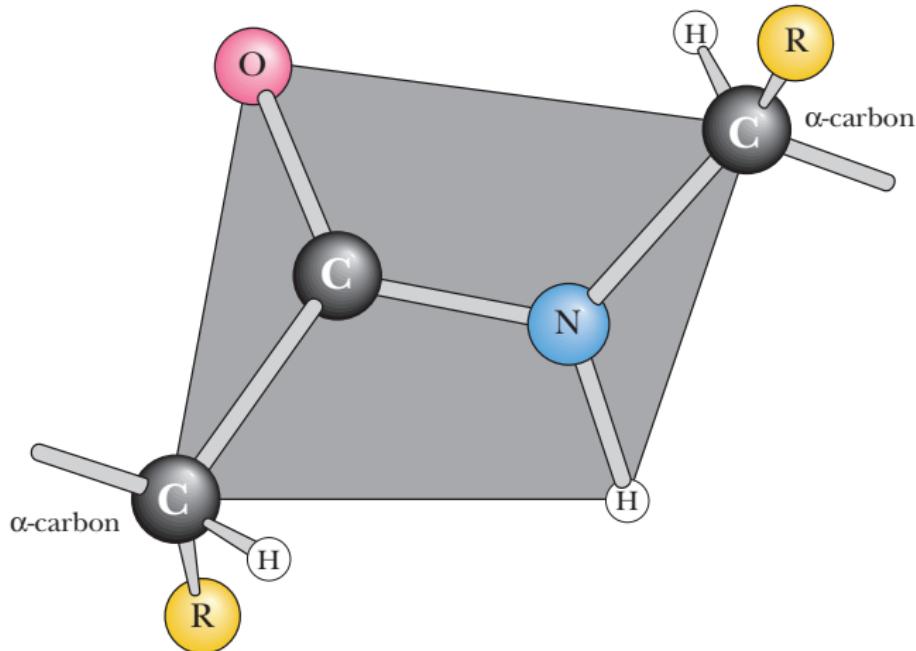


Figure 3-1 part 2 of 2 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Peptide bond is planar



Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

Amino acids come in four flavors

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

AMINO ACID		SIDE CHAIN		AMINO ACID		SIDE CHAIN	
Aspartic acid	Asp	D	negative	Alanine	Ala	A	nonpolar
Glutamic acid	Glu	E	negative	Glycine	Gly	G	nonpolar
Arginine	Arg	R	positive	Valine	Val	V	nonpolar
Lysine	Lys	K	positive	Leucine	Leu	L	nonpolar
Histidine	His	H	positive	Isoleucine	Ile	I	nonpolar
Asparagine	Asn	N	uncharged polar	Proline	Pro	P	nonpolar
Glutamine	Gln	Q	uncharged polar	Phenylalanine	Phe	F	nonpolar
Serine	Ser	S	uncharged polar	Methionine	Met	M	nonpolar
Threonine	Thr	T	uncharged polar	Tryptophan	Trp	W	nonpolar
Tyrosine	Tyr	Y	uncharged polar	Cysteine	Cys	C	nonpolar

POLAR AMINO ACIDS

Figure 3-2 Molecular Biology of the Cell 5/e (© Garland Science 2008)

NONPOLAR AMINO ACIDS

Chemical details of amino acids

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

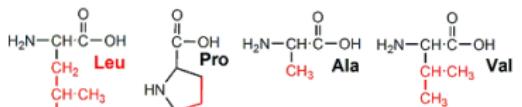
Secondary
Structure

Tertiary
Structures &
Domains

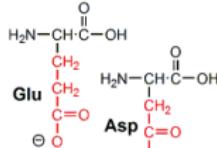
Quaternary
Structure

Summary

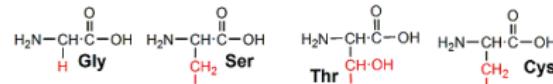
Reading for
next class



nonpolar



polar, charged (-)



polar, charged (+)



polar uncharged

Amino
Backbone Side
Chain

How might evolution substitute between these?

Deep question: Why is the code arranged this way?

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

TTT	Phe	TCT	Ser	TAT	Tyr	TGT	Cys
TTC	Phe	TCC	Ser	TAC	Tyr	TGC	Cys
TTA	Leu	TCA	Ser	TAA	stop	TGA	stop
TTG	Leu	TCG	Ser	TAG	stop	TGG	Trp

CTT	Leu	CCT	Pro	CAT	His	CGT	Arg
CTC	Leu	CCC	Pro	CAC	His	CGC	Arg
CTA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
CTG	Leu	CCG	Pro	CAG	Gln	CGG	Arg

ATT	Ile	ACT	Thr	AAT	Asn	AGT	Ser
ATC	Ile	ACC	Thr	AAC	Asn	AGC	Ser
ATA	Ile	ACA	Thr	AAA	Lys	AGA	Arg
ATG	Met	ACG	Thr	AAG	Lys	AGG	Arg

GTT	Val	GCT	Ala	GAT	Asp	GGT	Gly
GTC	Val	GCC	Ala	GAC	Asp	GGC	Gly
GTA	Val	GCA	Ala	GAA	Glu	GGA	Gly
GTG	Val	GCG	Ala	GAG	Glu	GGG	Gly

Key:

polar

nonpolar

acid

base

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains
Quaternary
Structure

Summary

Reading for
next class

1 Protein structure

Why do we care?

Primary Structure & Amino Acids

Secondary Structure

Tertiary Structures & Domains

Quaternary Structure

2 Summary

3 Reading for next class

α -helix stabilized by hydrogen bonds

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

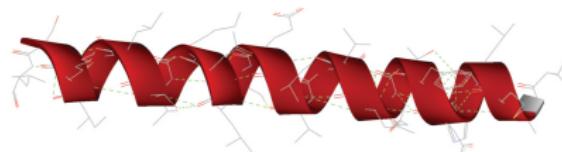
Tertiary
Structures &
Domains

Quaternary
Structure

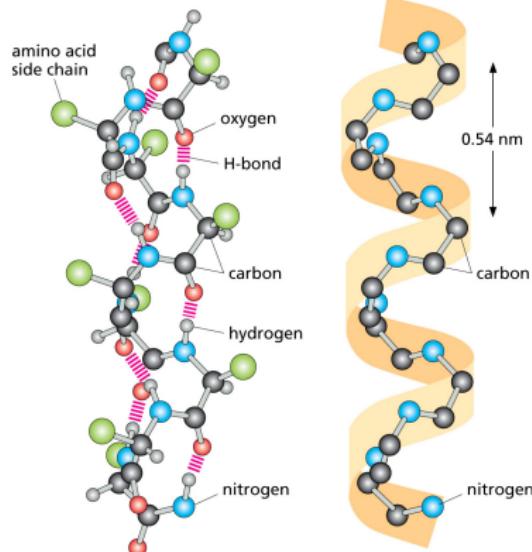
Summary

Reading for
next class

(A)



(B)



β -sheet stabilized by hydrogen bonds

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

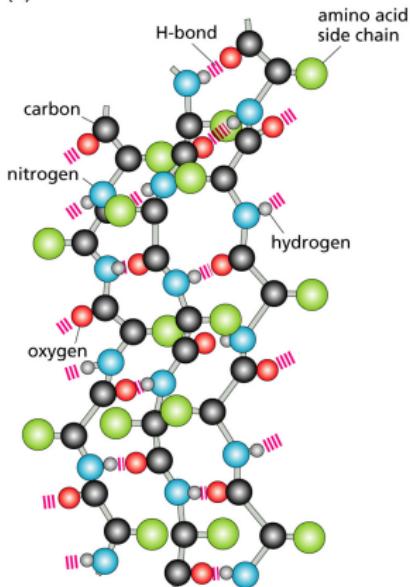
Tertiary
Structures &
Domains

Quaternary
Structure

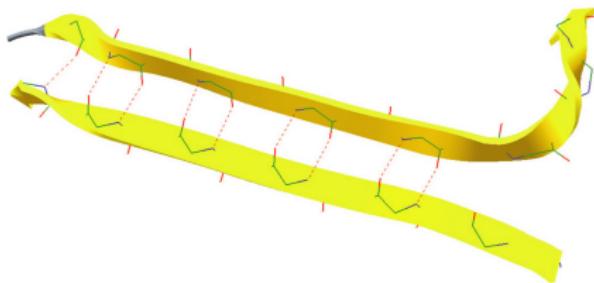
Summary

Reading for
next class

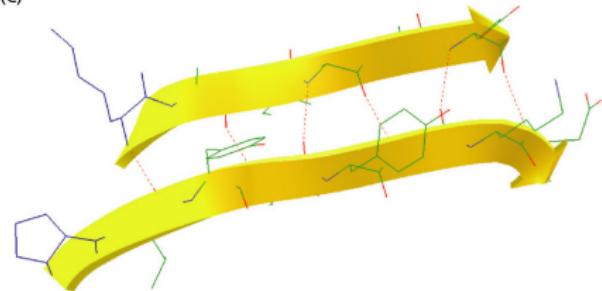
(A)



(B)



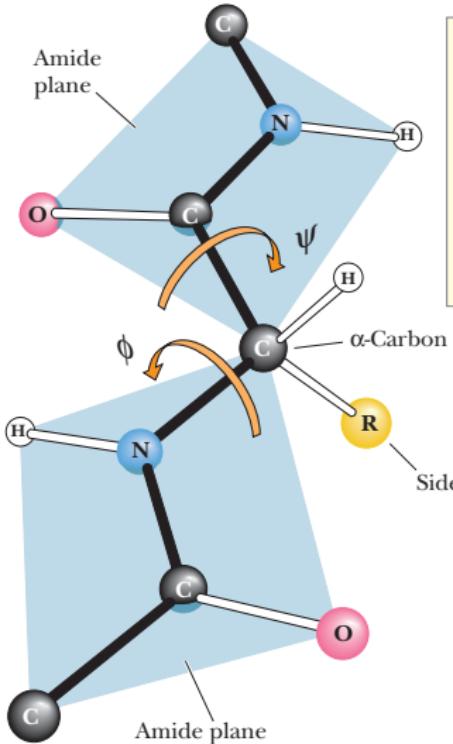
(C)



Bonds rotate around α -carbon

Protein structure
Why do we care?
Primary Structure & Amino Acids
Secondary Structure
Tertiary Structures & Domains
Quaternary Structure

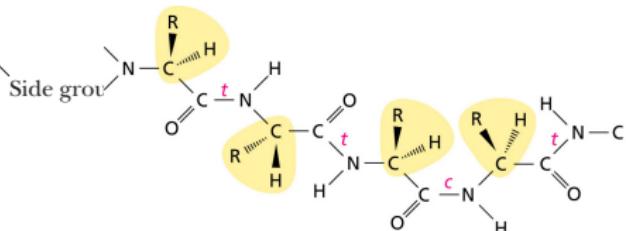
Summary
Reading for next class



Key idea:

If you know every Φ and Ψ ...

then you know the overall 3D shape of the protein



$$\phi = 180^\circ, \psi = 180^\circ$$

Steric hindrance favors trans over cis

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

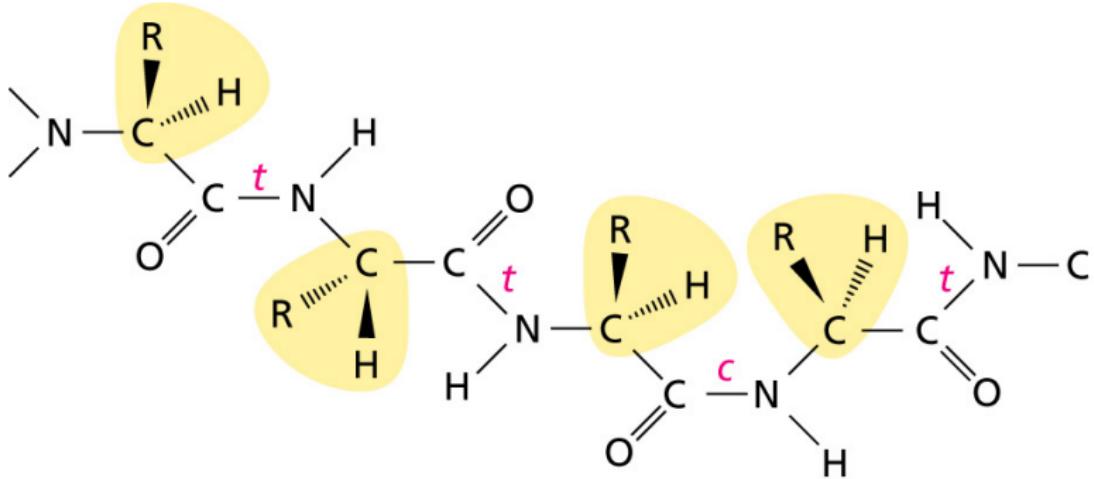
Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class



How can a computer predicting protein folding use this fact?

Ramachandran plot shows Φ and Ψ of α -helices and β -sheets

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

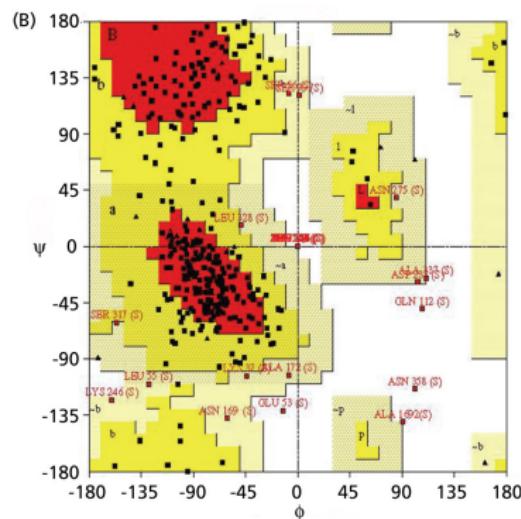
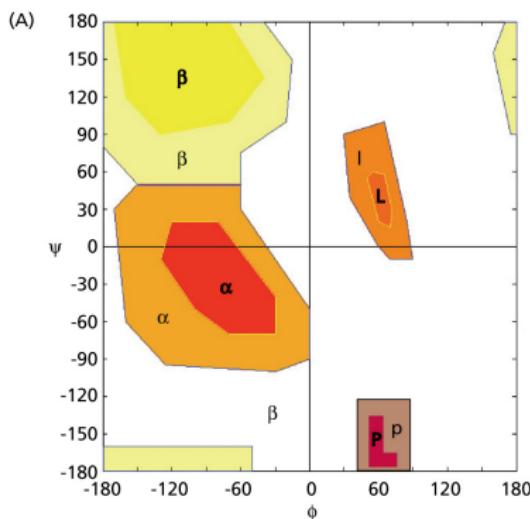
Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class



Cool animation of steric hindrance of varying Φ and Ψ : http://www.biochem.arizona.edu/classes/bioc462/462a/NOTES/Protein_Structure/Rama_animation.htm

www.biochem.arizona.edu/classes/bioc462/462a/NOTES/Protein_Structure/Rama_animation.htm

Hydrogen bonds stabilize turns

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

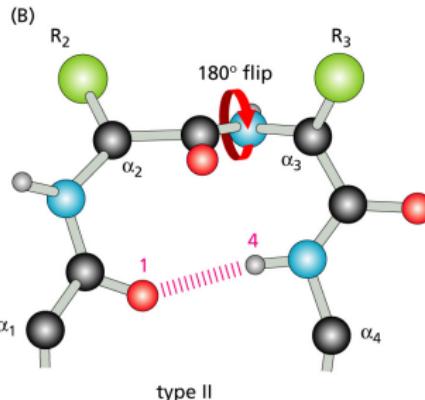
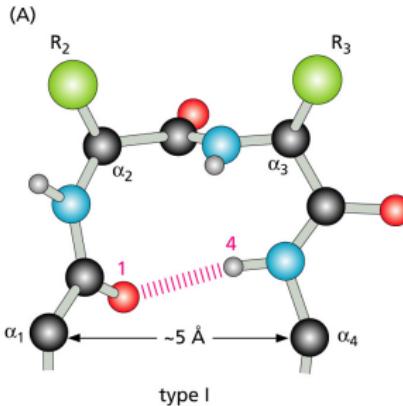
Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class



Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

1 Protein structure

Why do we care?

Primary Structure & Amino Acids

Secondary Structure

Tertiary Structures & Domains

Quaternary Structure

2 Summary

3 Reading for next class

Tertiary structure

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

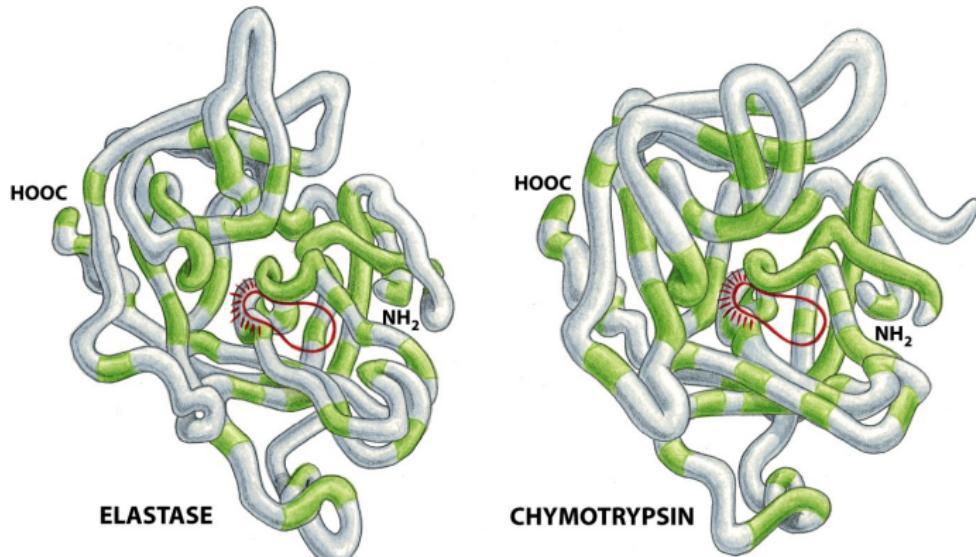


Figure 3-12 Molecular Biology of the Cell 5/e (© Garland Science 2008)

These proteins' tertiary structures look similar.
How can that happen?

Key idea: Globular proteins have hydrophobic cores and hydrophilic exteriors

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

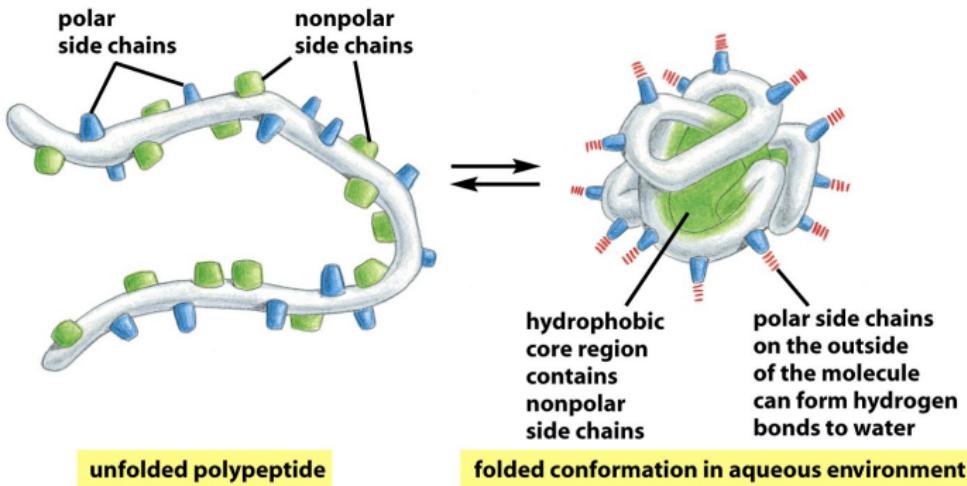


Figure 3-5 Molecular Biology of the Cell 5/e (© Garland Science 2008)

- 1) Why hydrophilic on the outside, hydrophobic on the inside?
- 2) How does this help computers predict the structure?

Domains are independent self-folding units

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

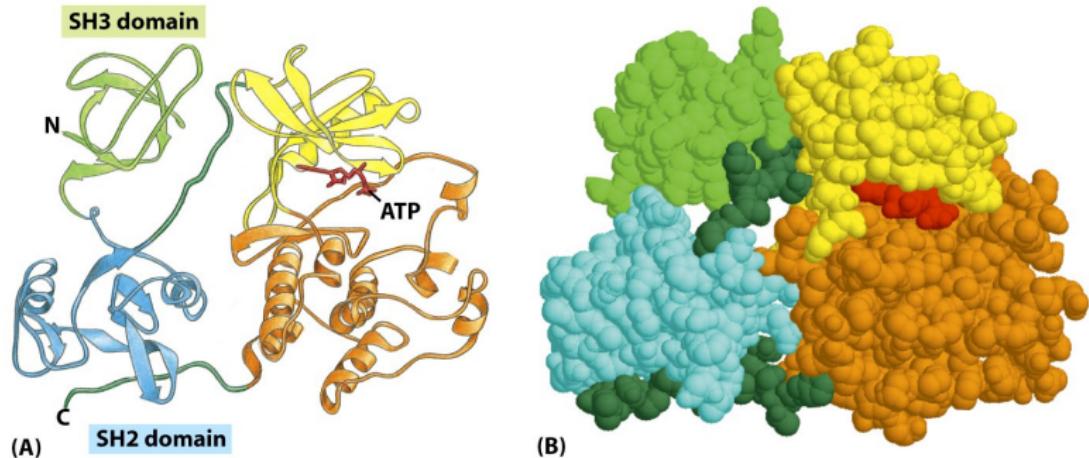


Figure 3-10 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Domains have hydrophobic cores →

fold independently

Example: tumor suppressor p53 has several domains

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

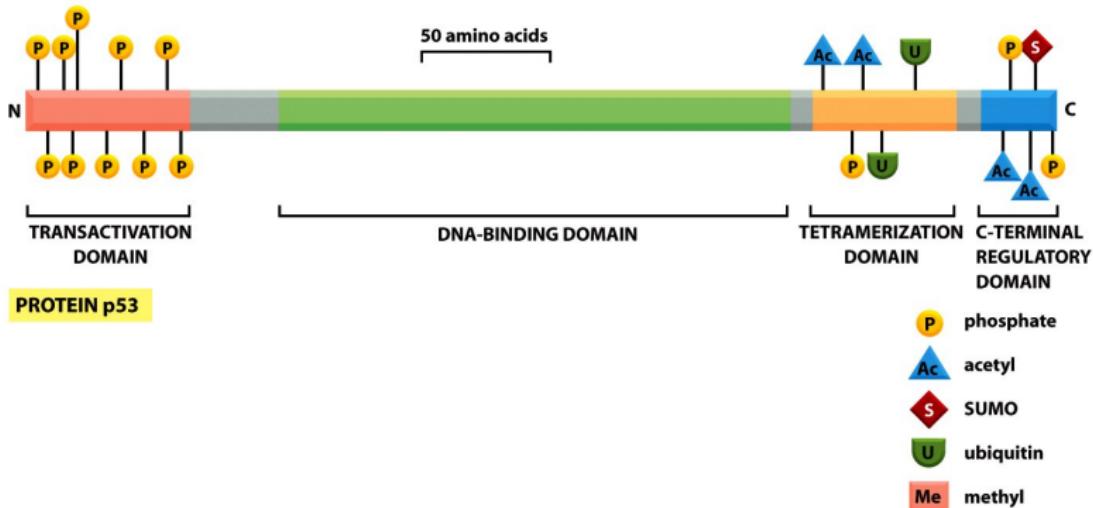


Figure 3-81a Molecular Biology of the Cell 5/e (© Garland Science 2008)

Domains are modular

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

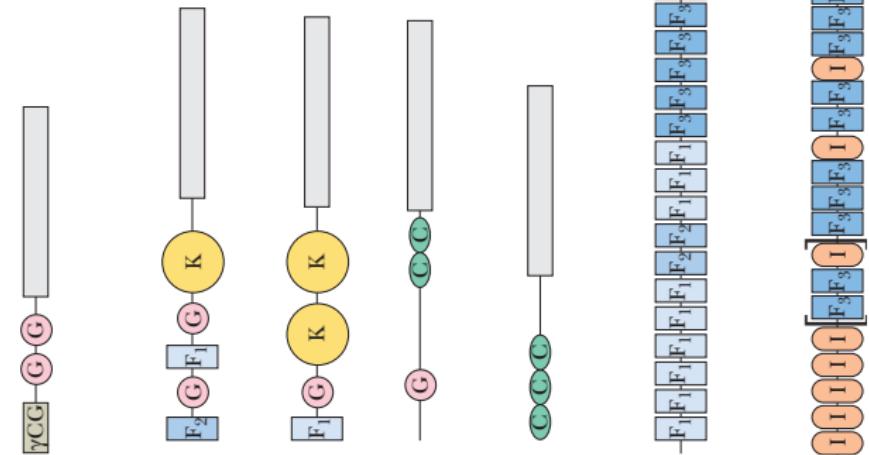
Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

How did molecular evolution do this????



Factor VII,
IX, X and
protein C

Factor XII

tPA

Clr,Cls

C2,
factor B

Fibronectin

Twitchi

Hint: How do self-folding domains relate to evolution?

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

1 Protein structure

Why do we care?

Primary Structure & Amino Acids

Secondary Structure

Tertiary Structures & Domains

Quaternary Structure

2 Summary

3 Reading for next class

Quaternary structure of hemoglobin heterotetramer

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

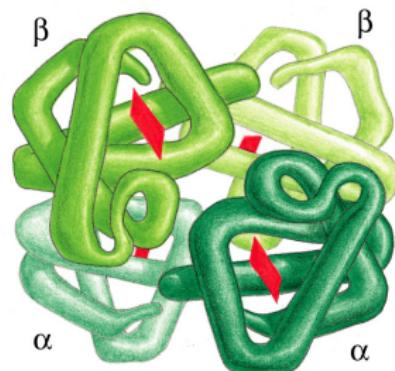
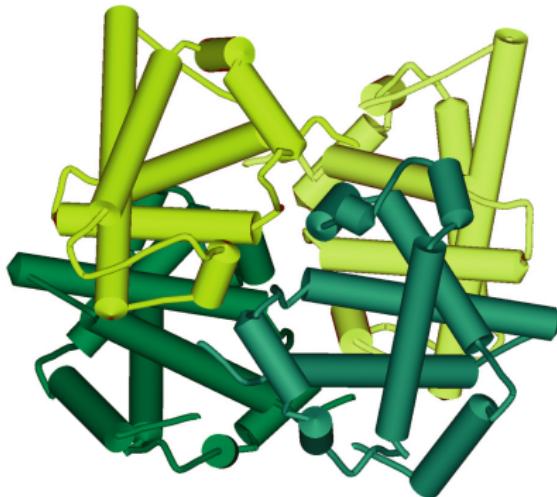
Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class



- Hemoglobin is a hetero-tetramer
- Dissect the word *hetero-tetramer*
- Do these genes have to be neighbors on the chromosome?

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

1 Protein structure

Why do we care?

Primary Structure & Amino Acids

Secondary Structure

Tertiary Structures & Domains

Quaternary Structure

2 Summary

3 Reading for next class

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

1 Protein structure

Why do we care?

Primary Structure & Amino Acids

Secondary Structure

Tertiary Structures & Domains

Quaternary Structure

2 Summary

3 Reading for next class

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

1 Protein structure

Why do we care?

Primary Structure & Amino Acids

Secondary Structure

Tertiary Structures & Domains

Quaternary Structure

2 Summary

3 Reading for next class

Reading for next class

Protein
structure

Why do we care?

Primary
Structure &
Amino Acids

Secondary
Structure

Tertiary
Structures &
Domains

Quaternary
Structure

Summary

Reading for
next class

Chapter 2

Section “Introduction to Biological Databases”

to

Section “Command-line Access to Data at NCBI”

Pages	Notes
19–42	Read