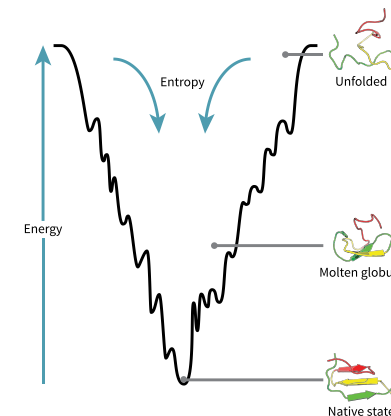
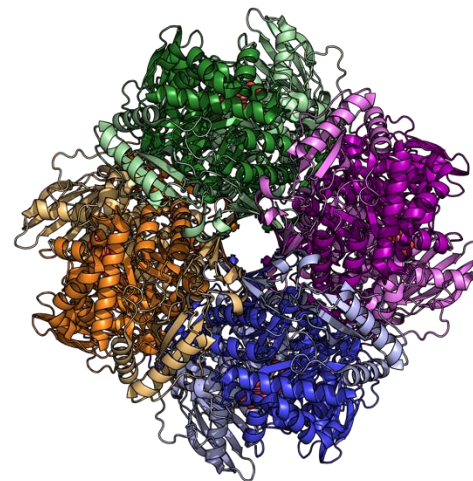


Protein Structure & Thermodynamics

Workshop

10th February 2025



Dr. Amanda Souza Câmara, IPK
Dr. Victor H. Rabesquine Nogueira, IPK
Prof. Dr. Antônio José da Costa Filho, University of São Paulo

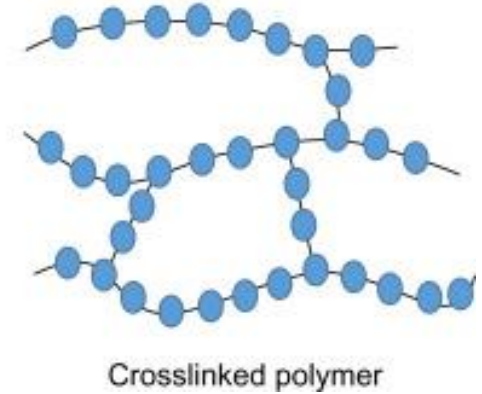
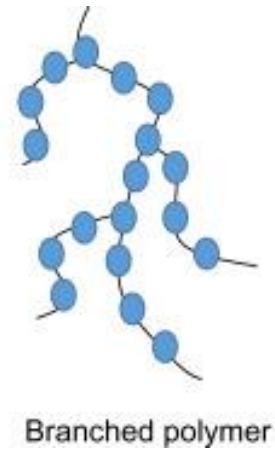
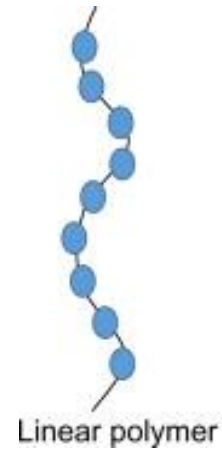
Programme

Monday 10th	13:00 – 17:00 – Welcome to the workshop, physics in biology, introduction to protein structure with atomic models
Tuesday 11th	08:30 – 12:00 – Introduction to protein thermodynamics
Wednesday 12th	08:30 – 12:00 – Continuing protein thermodynamics 13:00 – 14:00 - Vavilov Seminar by Prof. Dr. Antônio José da Costa Filho
Thursday 13th	08:30 – 12:00 – <i>In silico</i> analyses of protein structure with PyMOL, AlphaFold3 prediction and docking
Friday 14th	08:30 – 12:00 – Molecular dynamics simulation

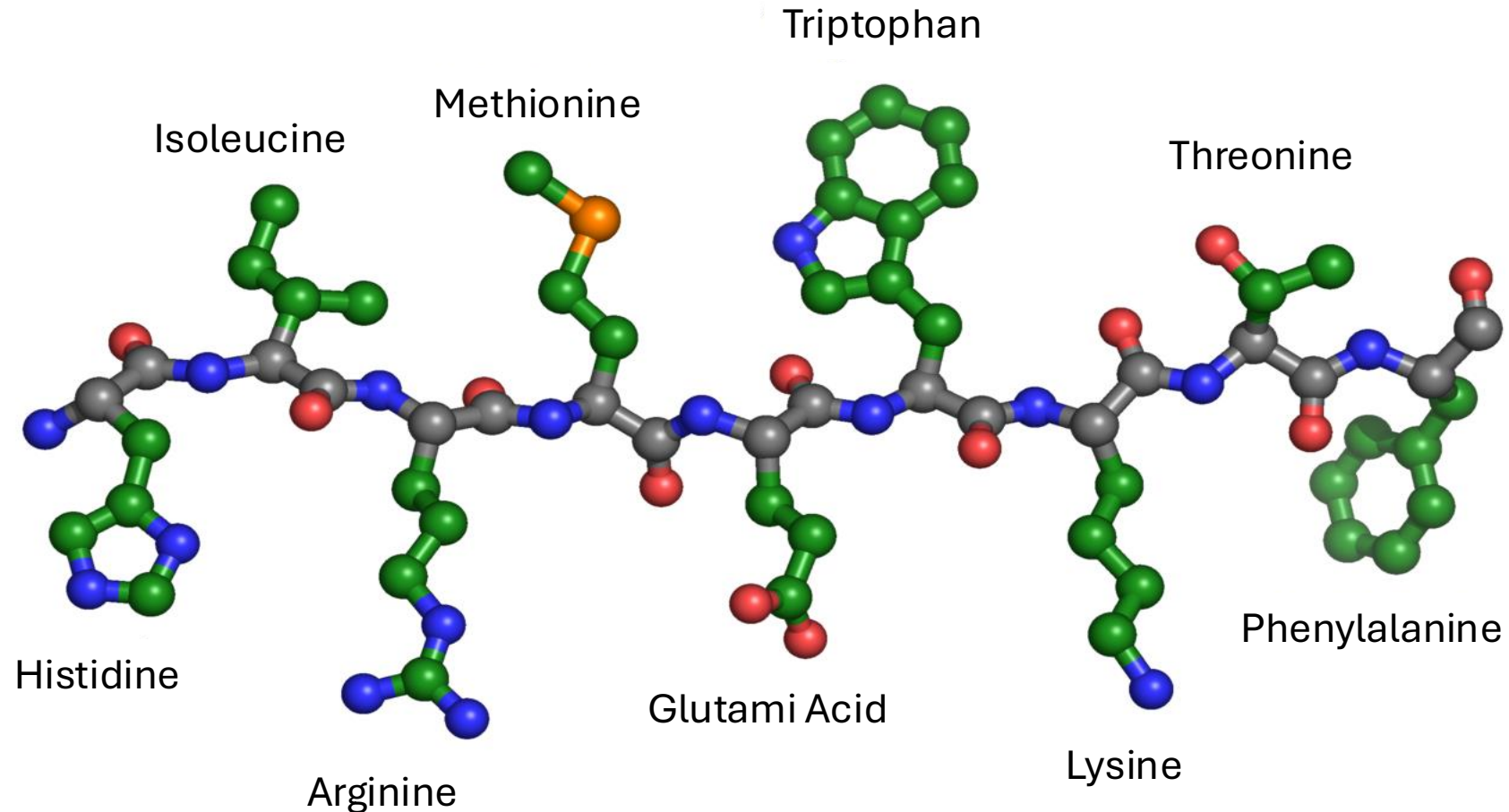
Introduction round

- Who are you?
- What is your background?
- Why are you interested in protein structure and thermodynamics?

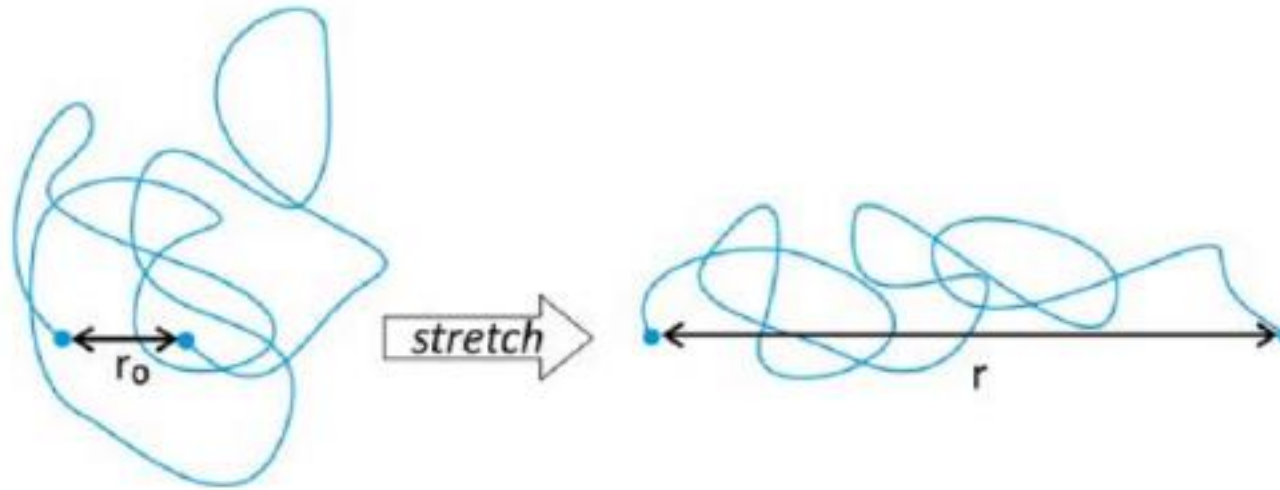
A protein is a linear polymer



A protein is a linear polymer



Conformations of a linear polymer

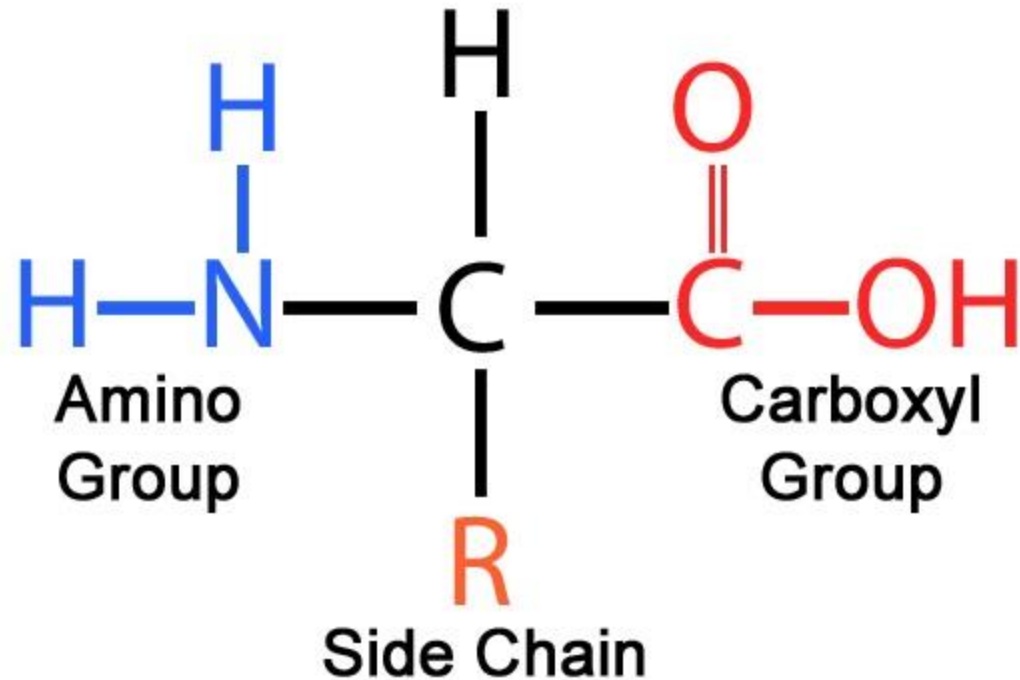


The completely stretched conformation is just one among countless possibilities.
But these conformations have restrictions!

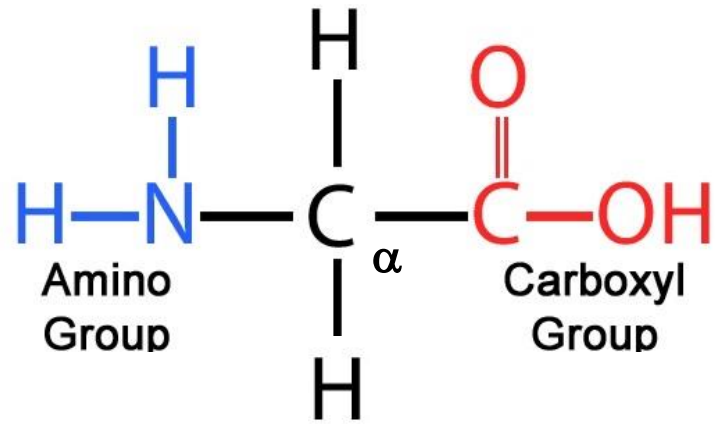
Our GitHub page

- <https://github.com/InsilicoGenebankProteomics/Protein-Structure-and-Thermodynamics-Workshop-2025>

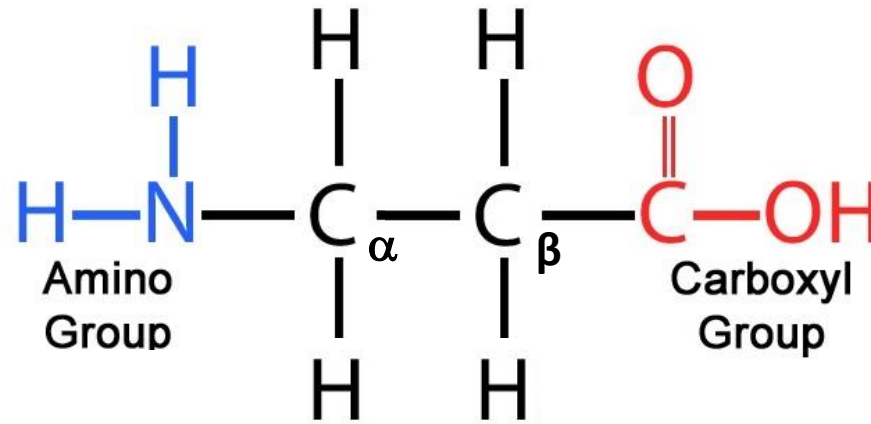
Building an amino acid



Building an α amino acid

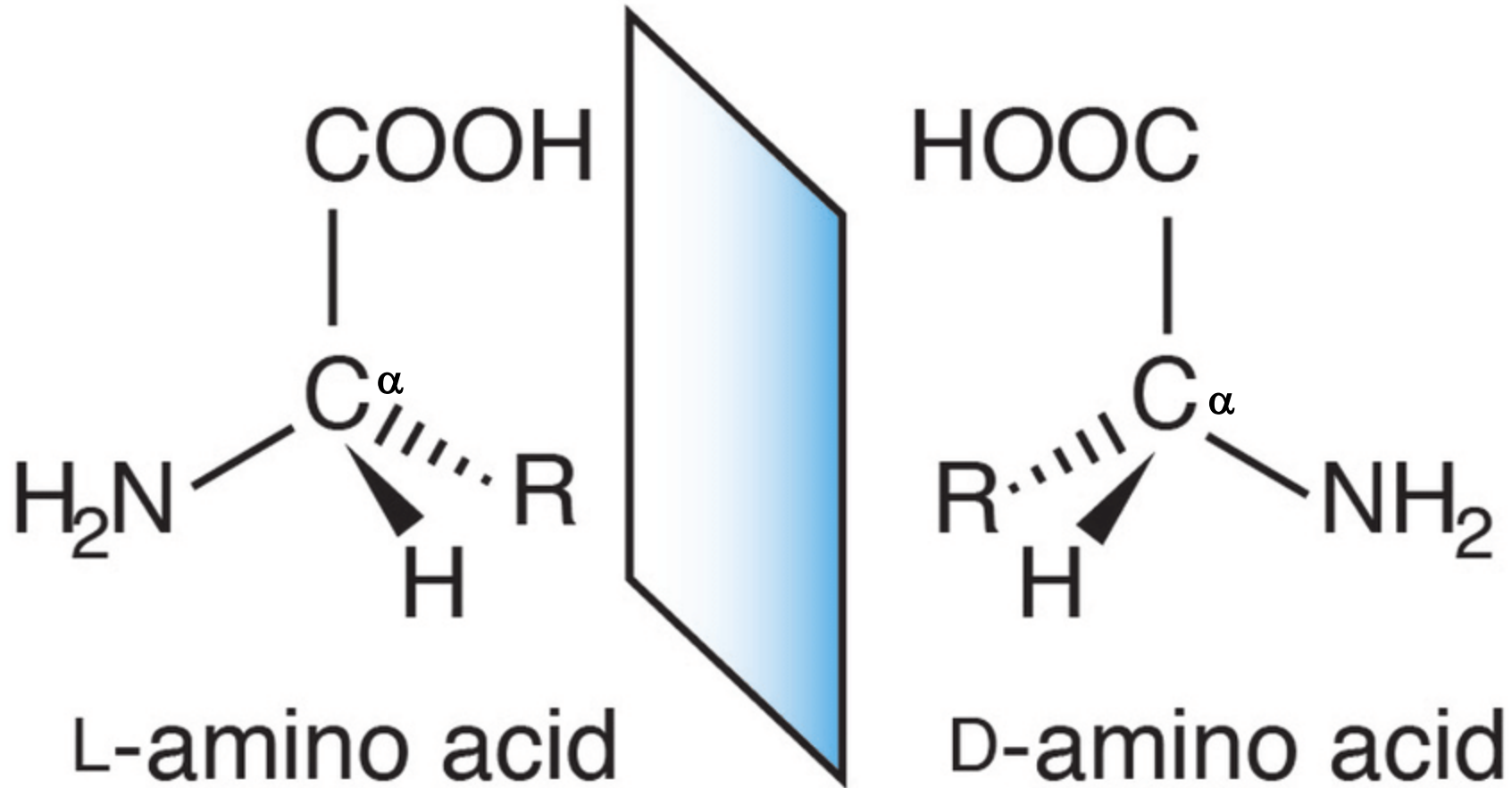


α amino acid



β amino acid

α carbons are chiral centers



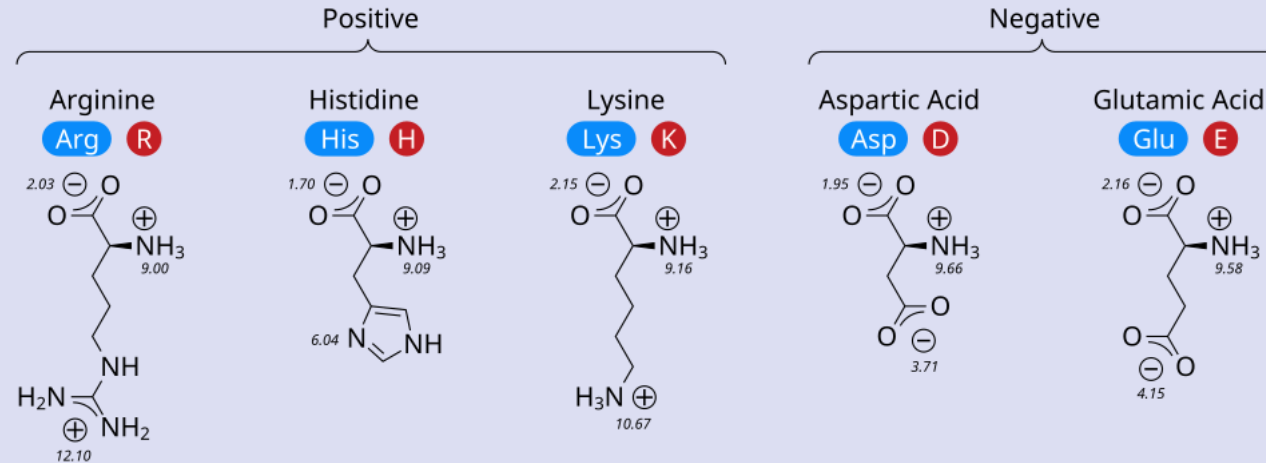
TWENTY-ONE PROTEINOGENIC α -AMINO ACIDS

Side chain charge
at physiological
pH 7.4

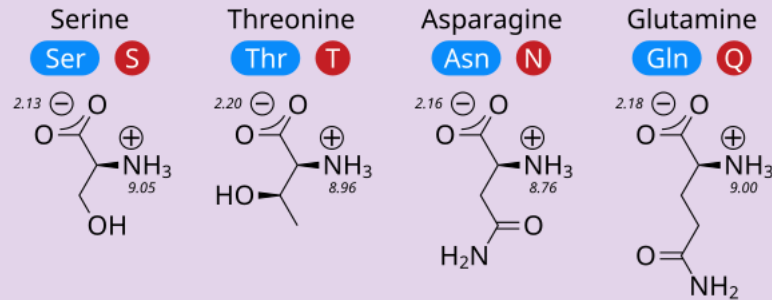
pK_a values shown
italicized

⊕ Positive
⊖ Negative

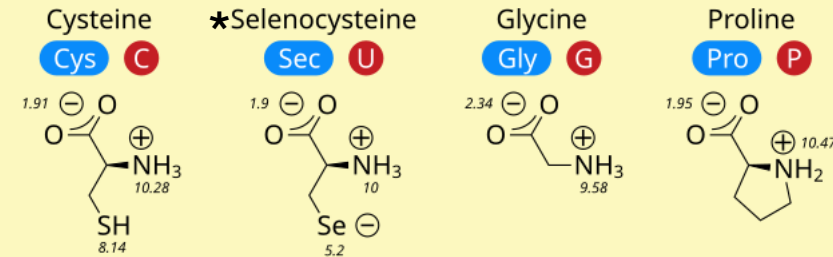
A. Amino Acids with Electrically Charged Side Chains



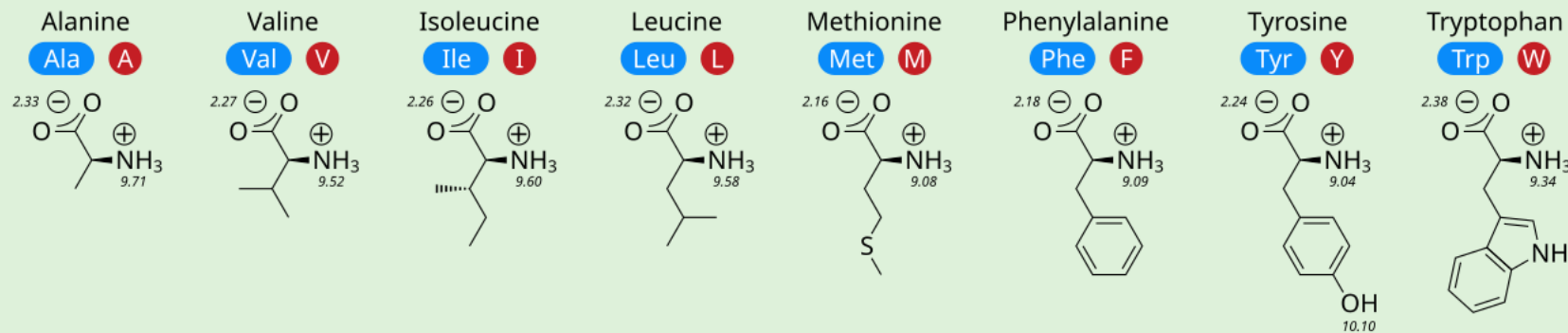
B. Amino Acids with Polar Uncharged Side Chains



C. Special Cases

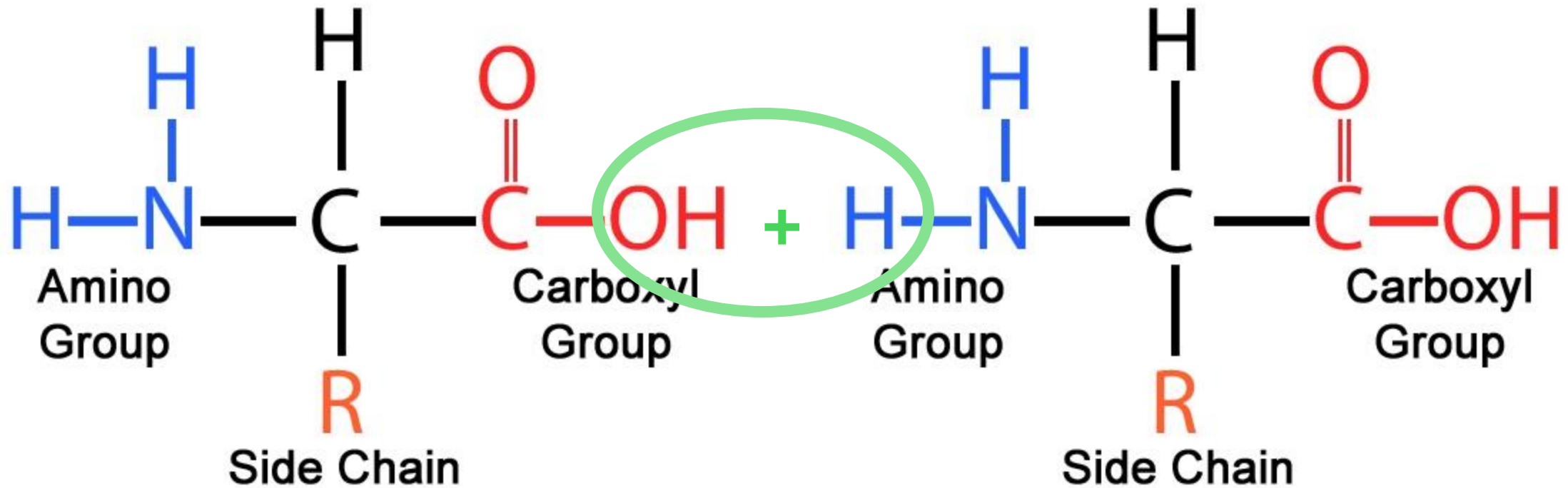


D. Amino Acids with Hydrophobic Side Chains

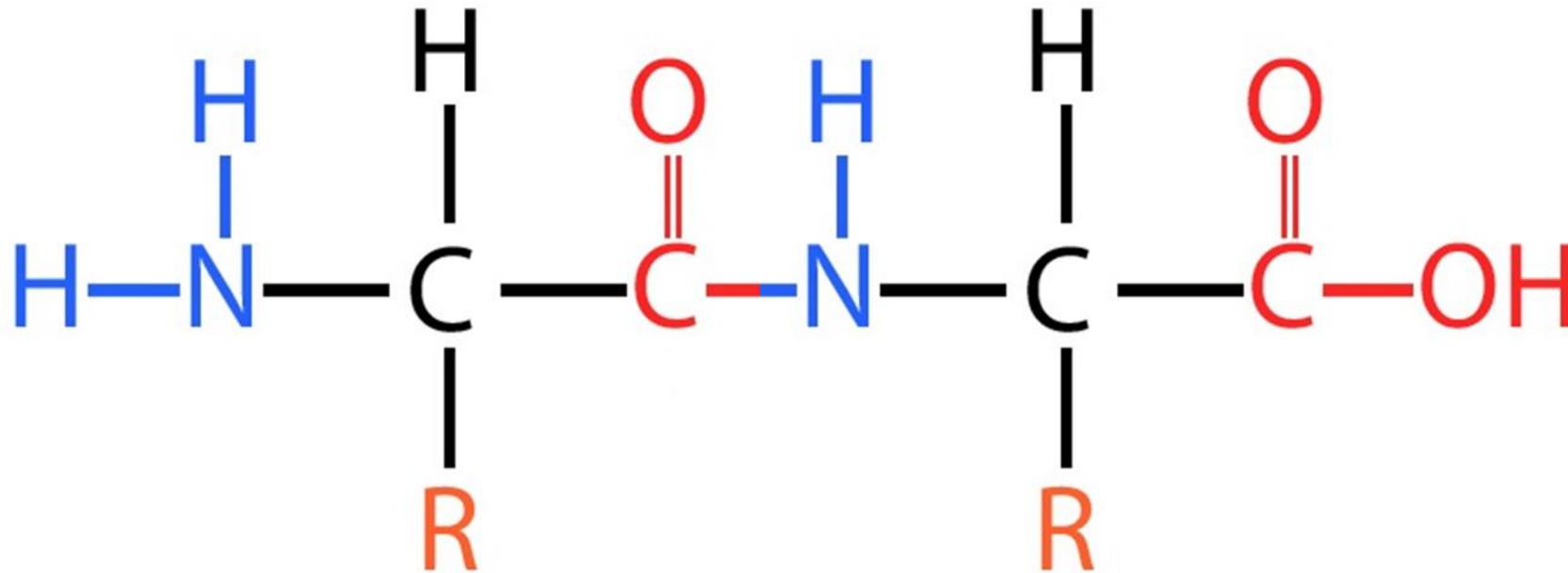


*Selenocysteine needs translational recoding.

The peptide bond

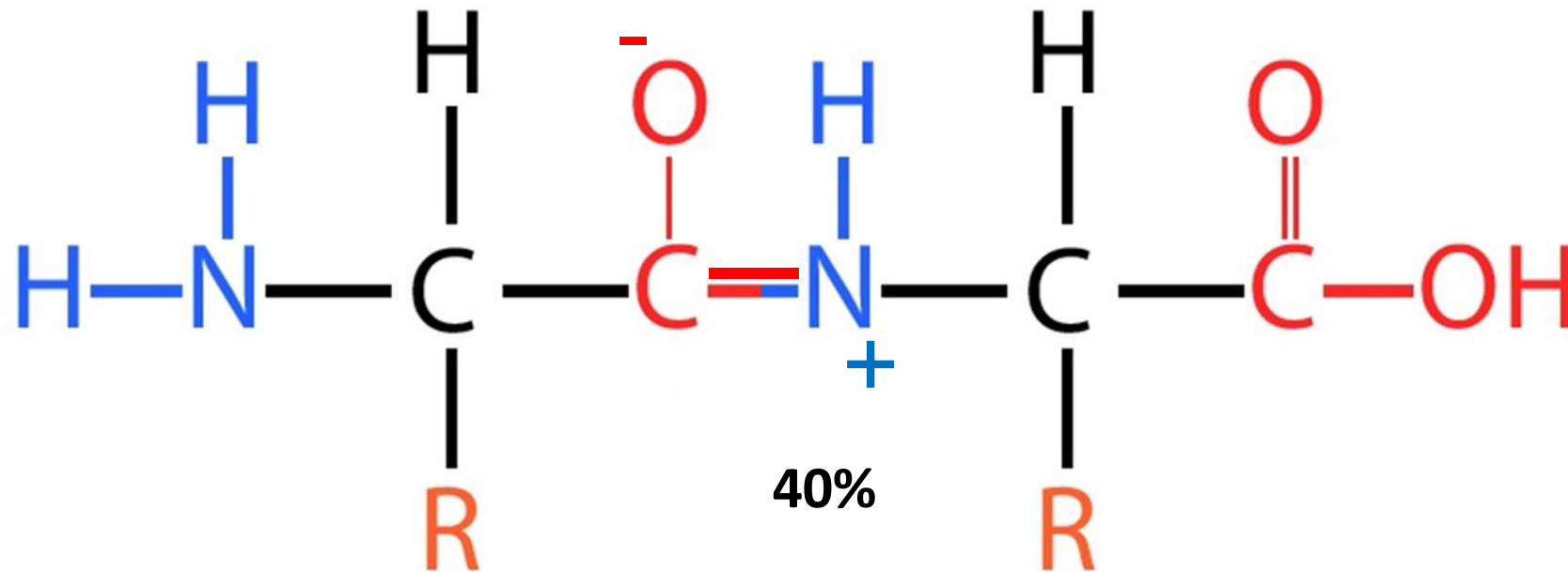


The peptide bond

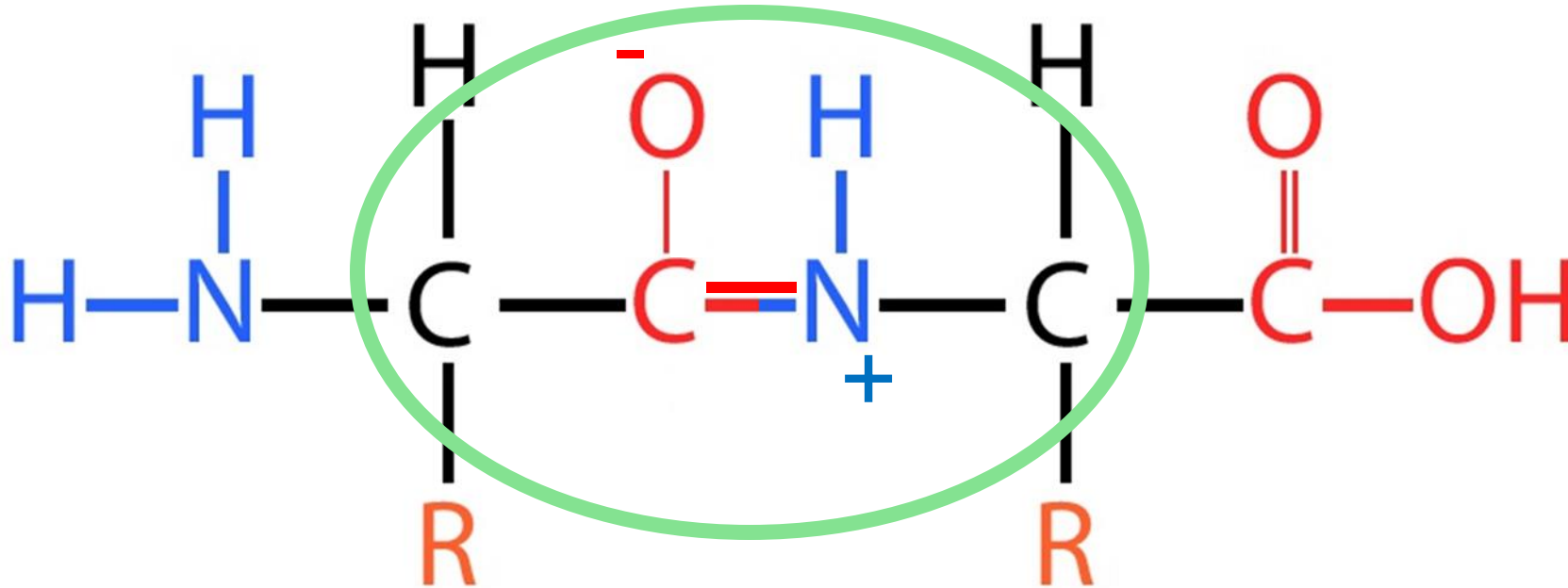


From the amino-terminal end to the carboxyl-terminal end

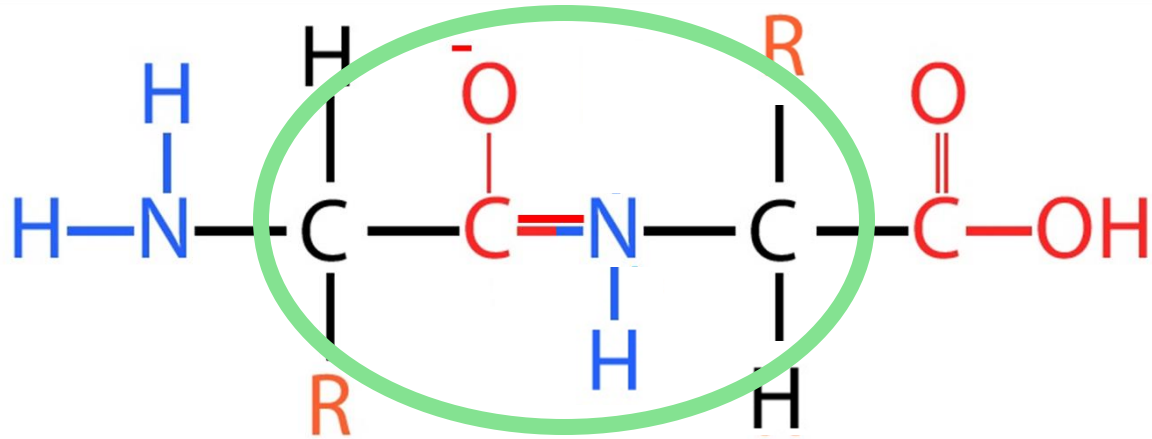
The peptide bond has two resonance structures



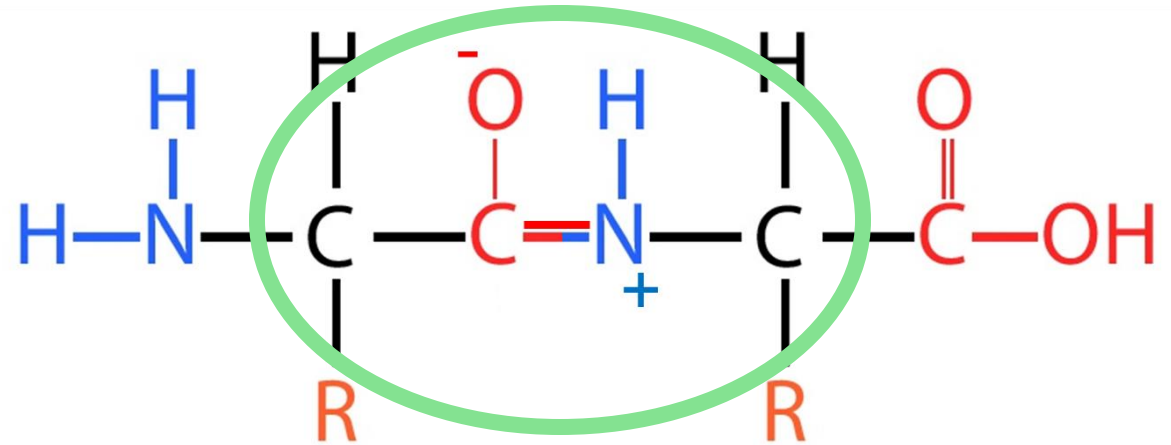
The peptide bond makes a plane



The peptide bond, cis or trans conformations

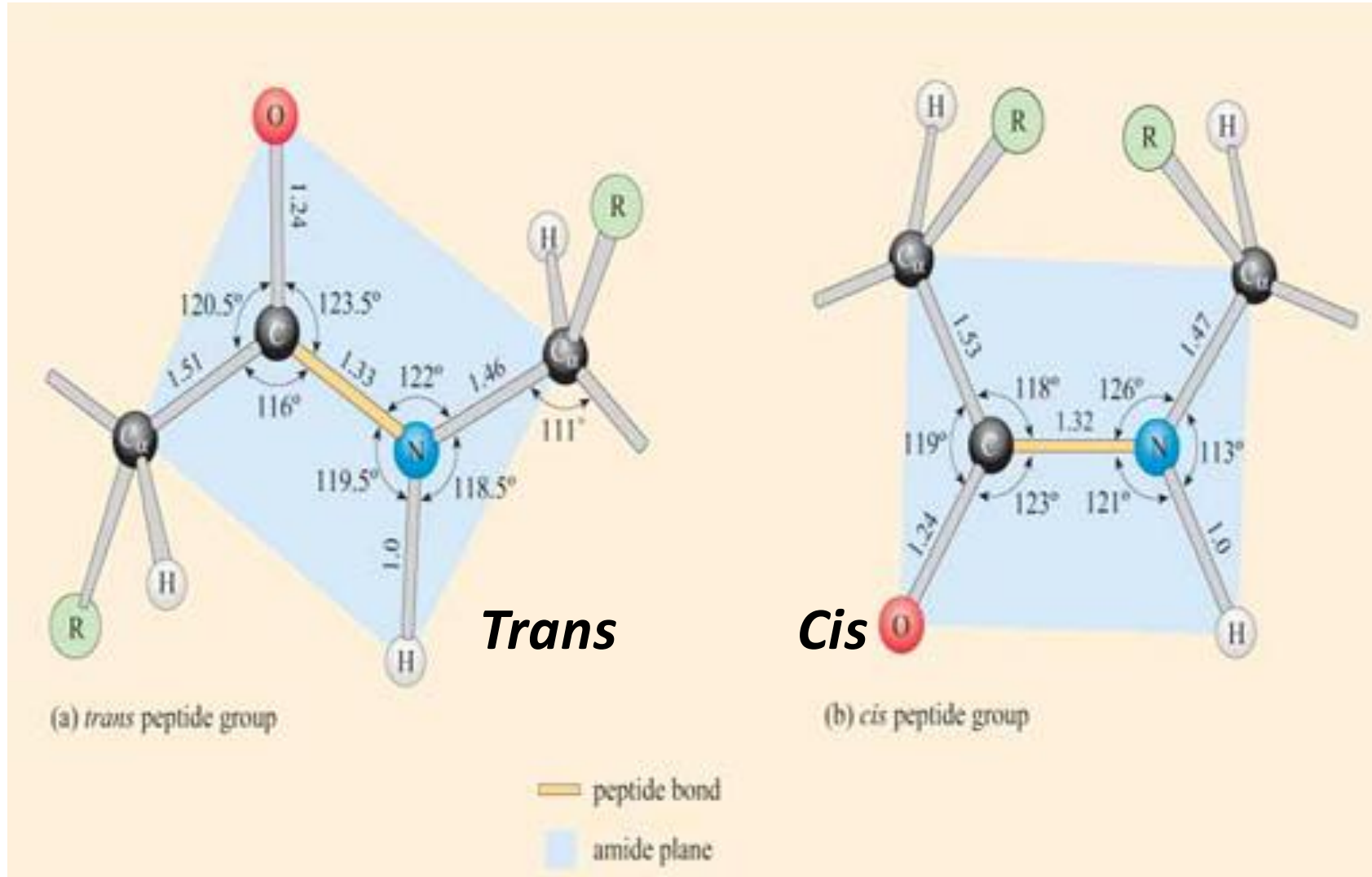


trans

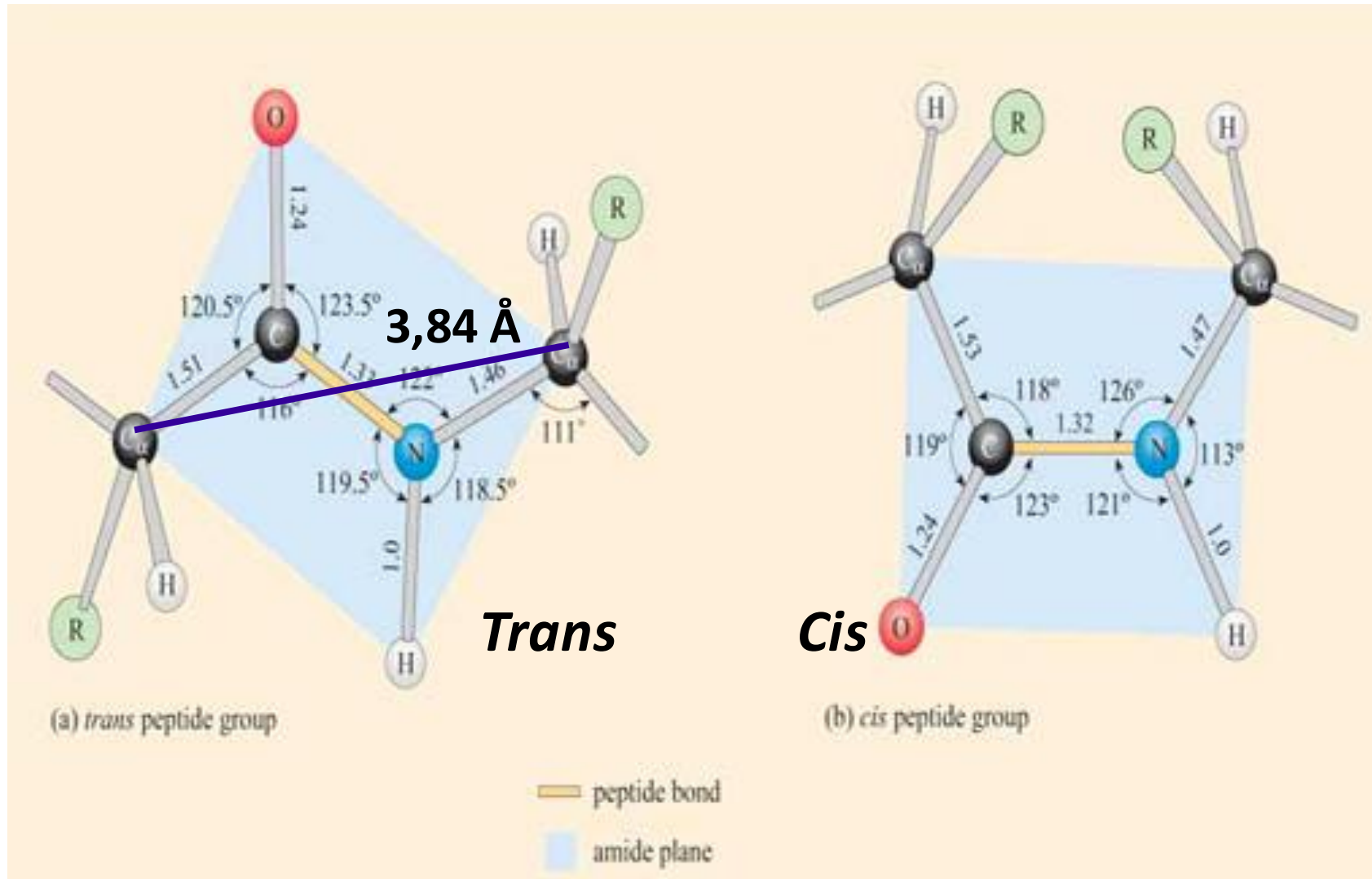


cis

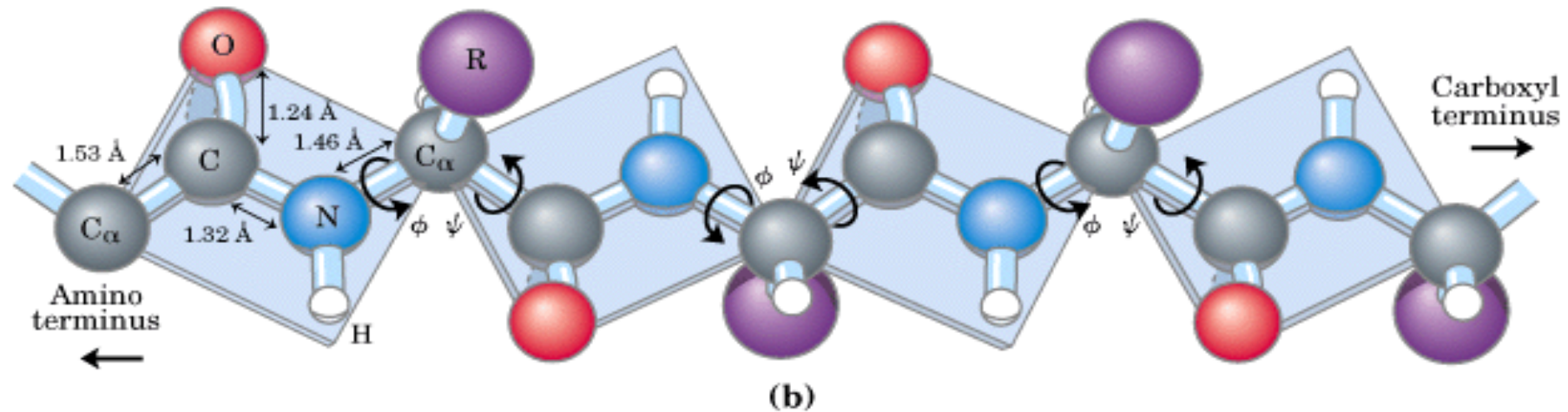
The peptide bond, cis or trans conformations



The peptide bond, cis or trans conformations



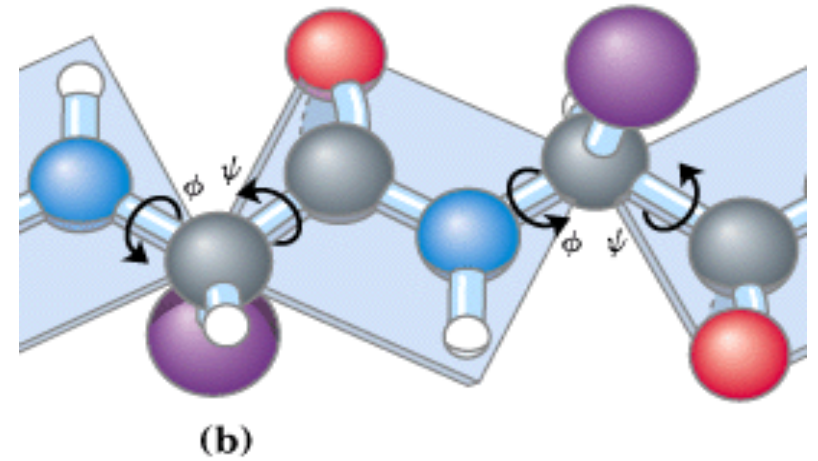
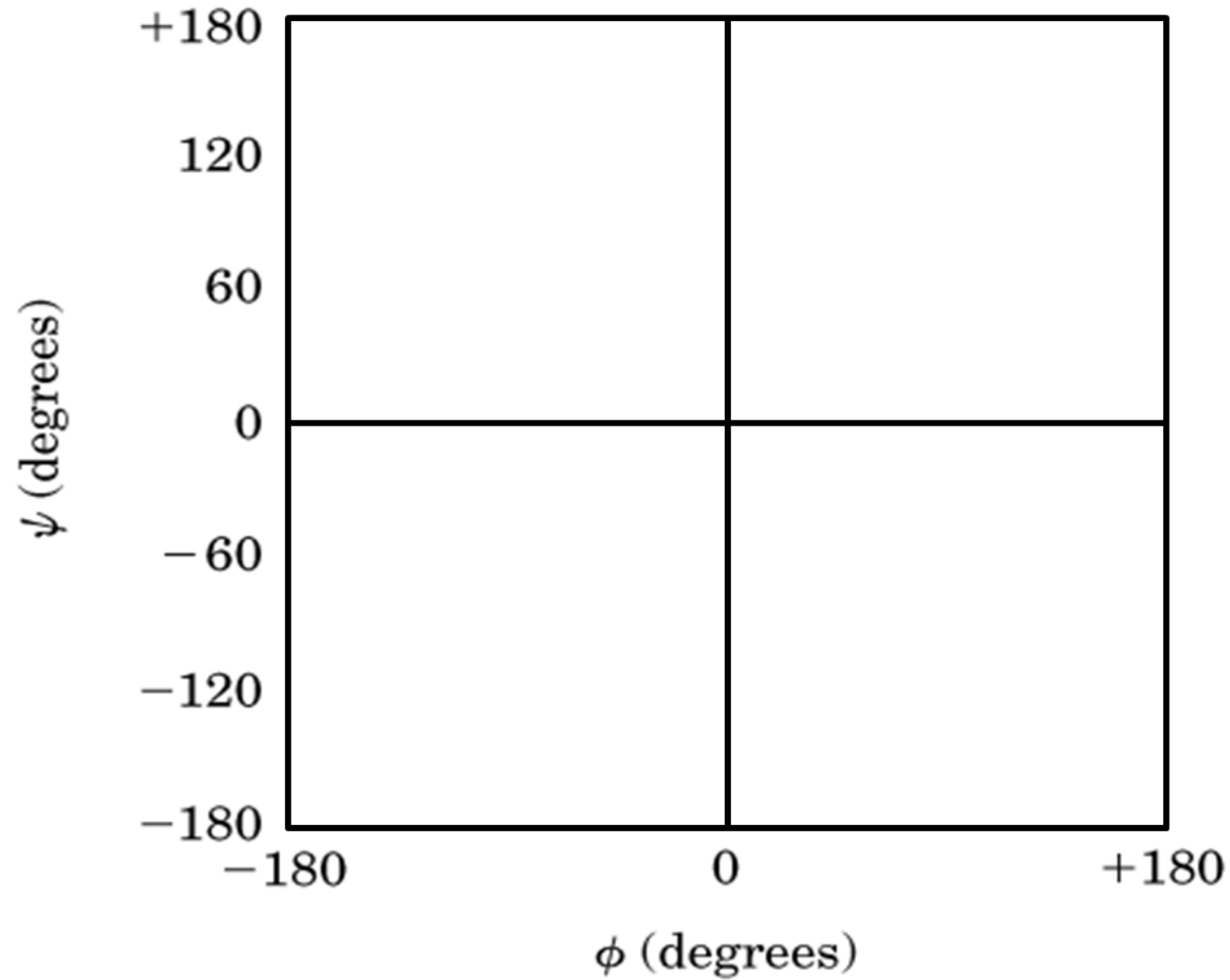
Between peptide planes are two torsion angles



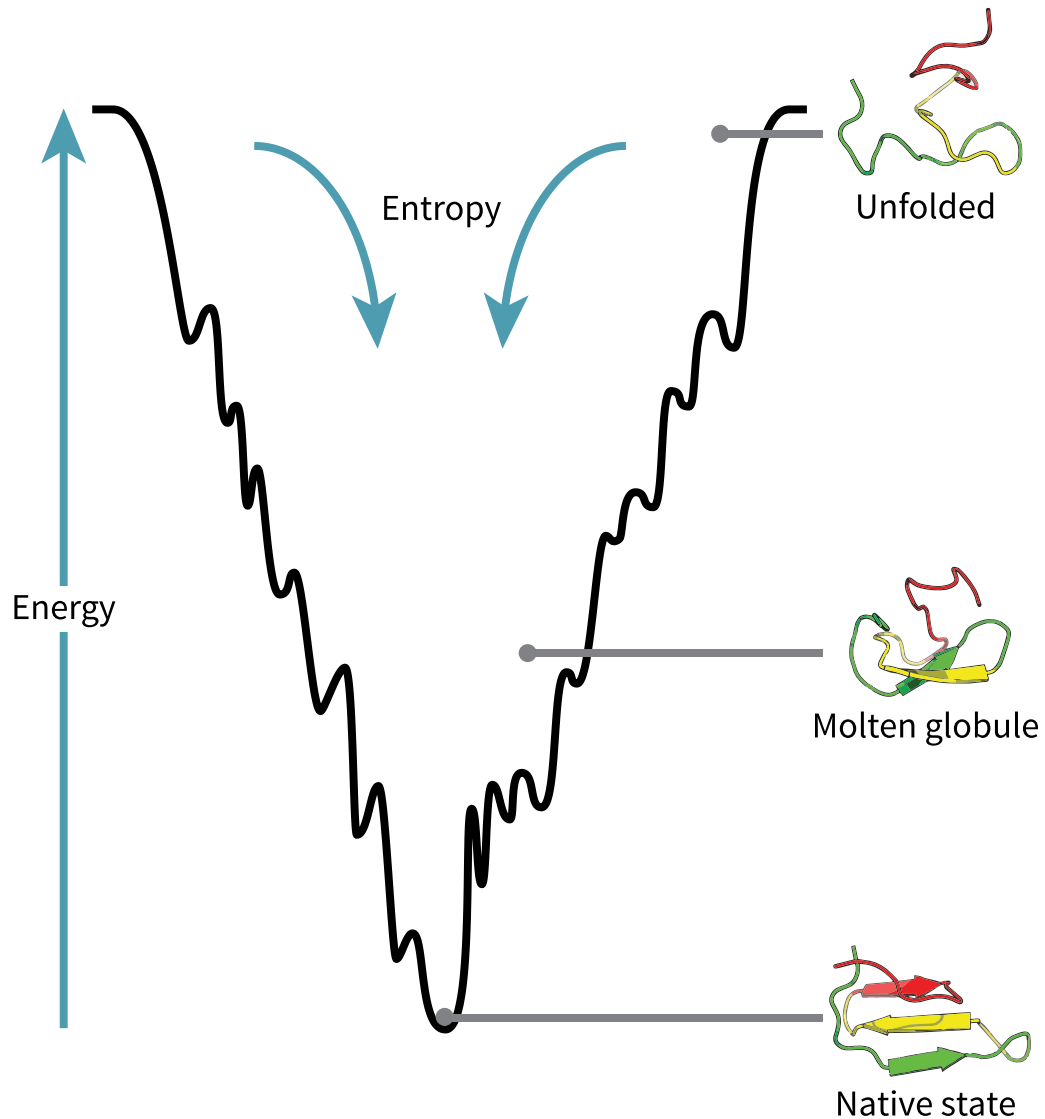
©1999 GARLAND PUBLISHING INC.
A member of the Taylor & Francis Group

Only the ϕ and ψ angles are free to rotate
And the positions of the main-chain atoms can be entirely defined by them.

The Ramachandran plot



©1999 GARLAND PUBLISHING INC.
A member of the Taylor & Francis Group



Extended conformations have low probability and high energy

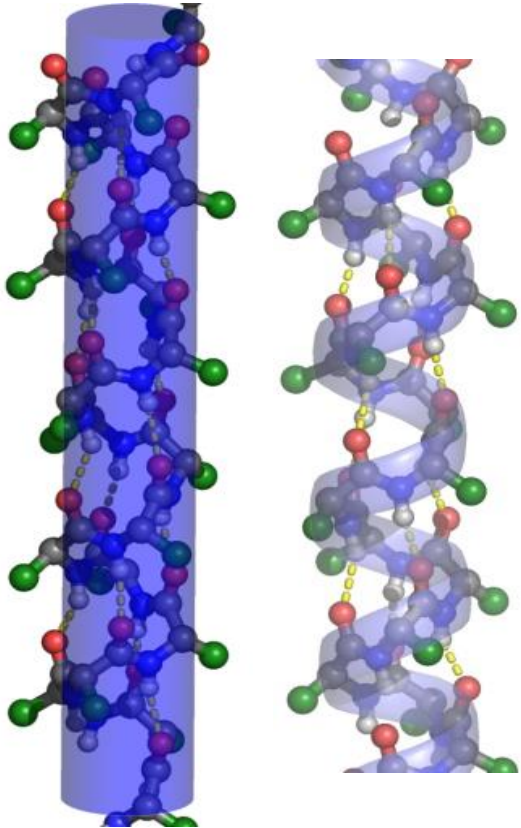
Globular conformations have a hydrophobic core

Lacking water, polar groups of the main chain make hydrogen bonds with each other, forming Secondary Structures (helices and sheets)

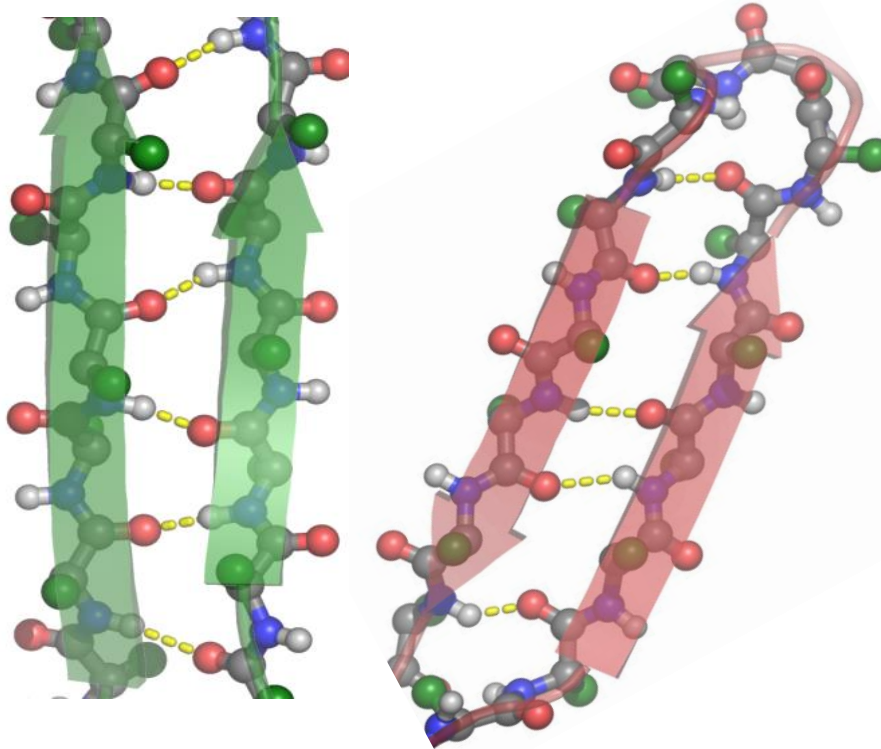
Secondary Structures arrange into motives and domains, which arrange into Tertiary Structures

Other protein chains arrange to form Quaternary Structures

Secondary structure

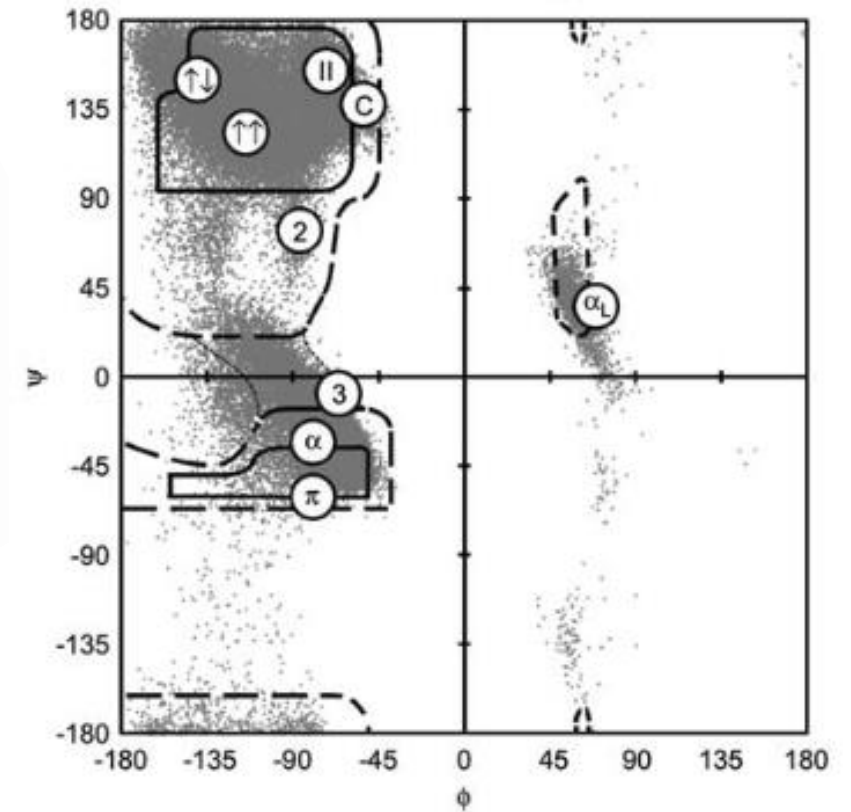


α -helices

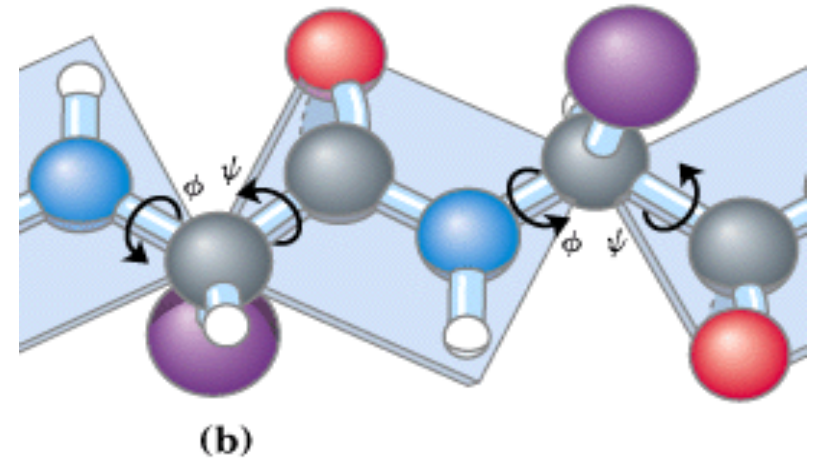
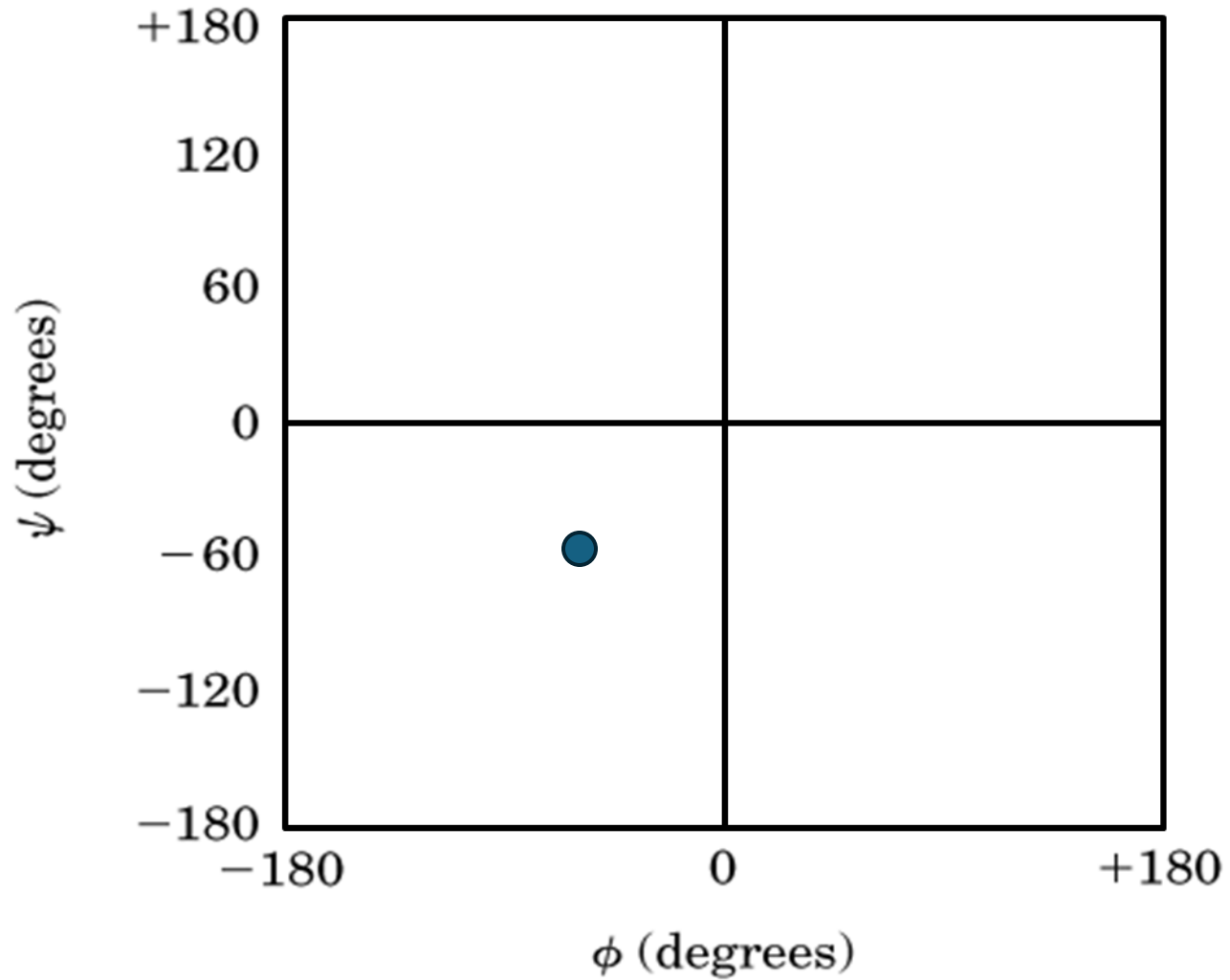


β -sheets

Ramachandran plot



α -helices



©1999 GARLAND PUBLISHING INC.
A member of the Taylor & Francis Group

α -helices

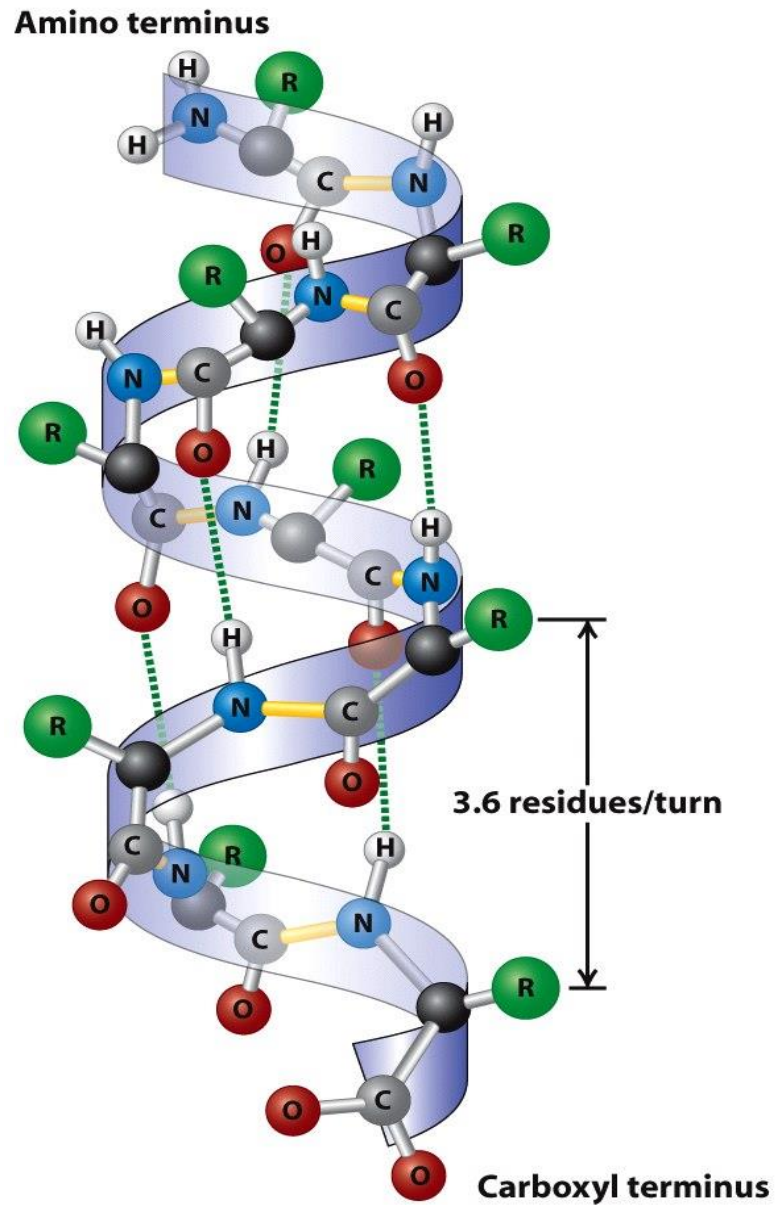
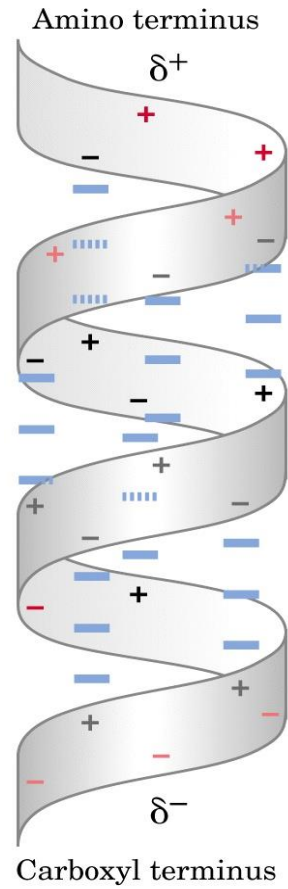
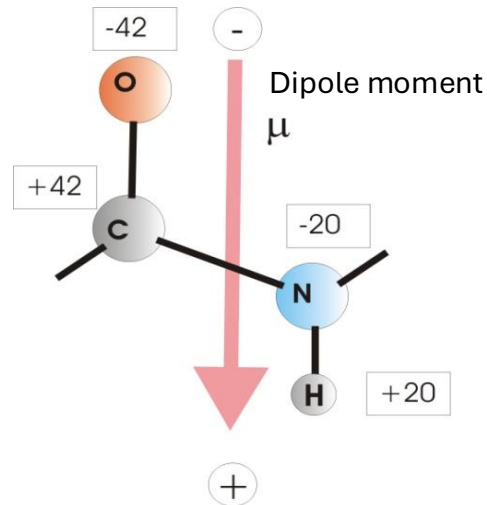


Figure 3-4
Molecular Cell Biology, Sixth Edition
© 2008 W. H. Freeman and Company

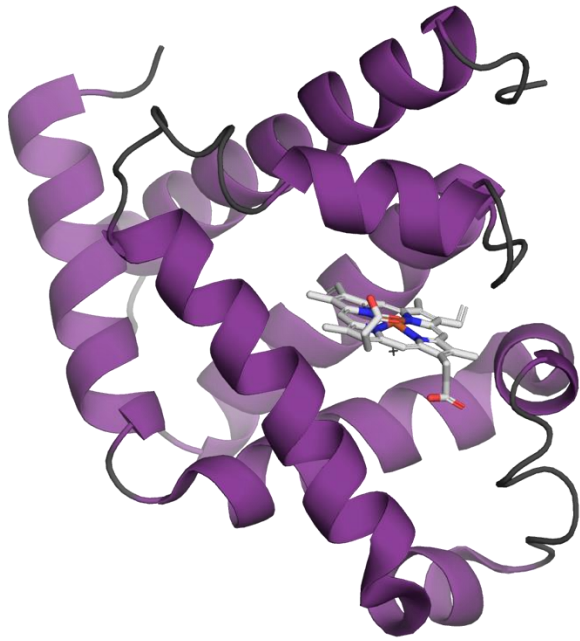
Electrical dipole



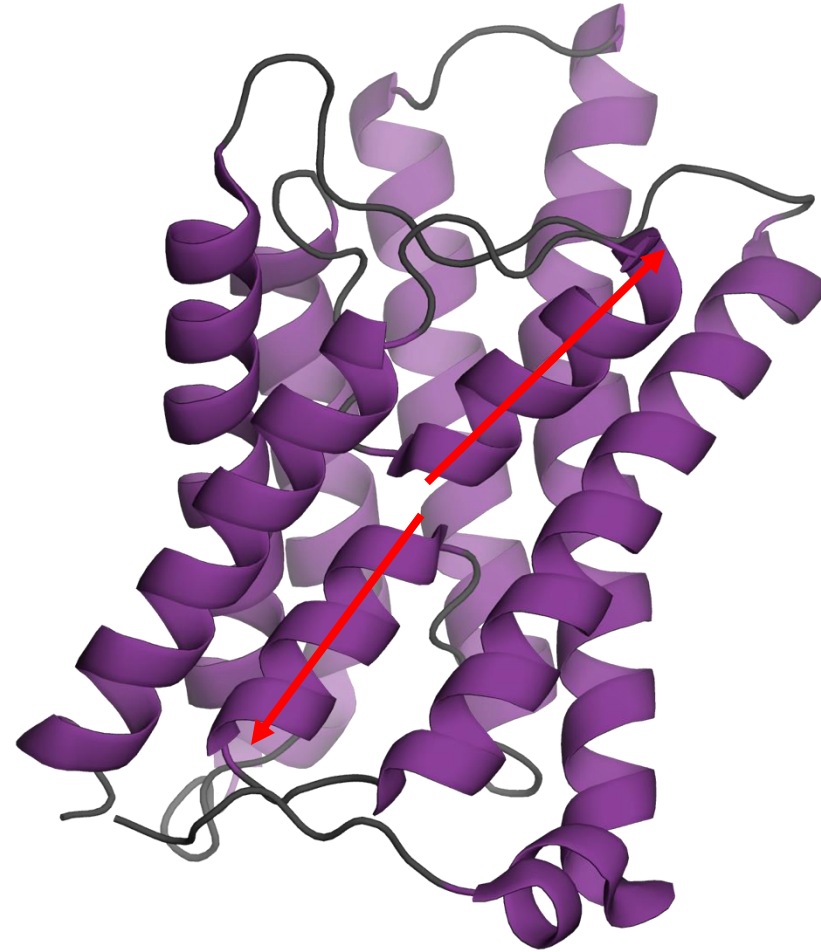
Biological Function

- Help in binding charged cofactors
- Long range attraction (K^+ and Cl^- channels)
- Change the nucleophilic properties of neighbouring residues for catalysis

All alpha protein example

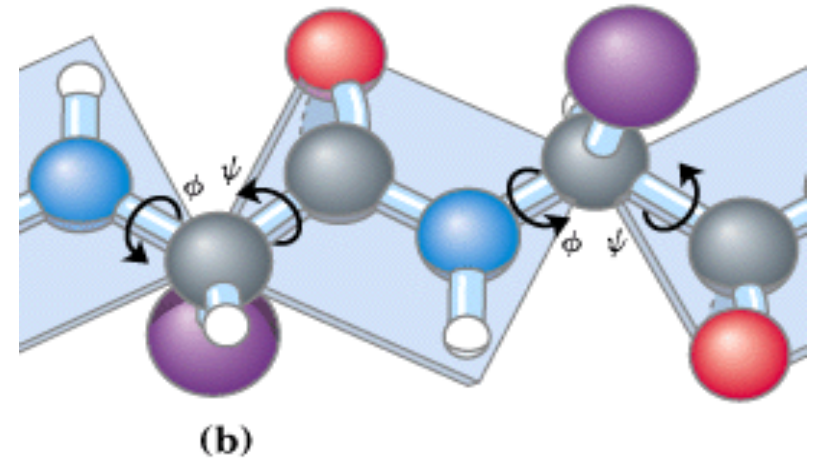
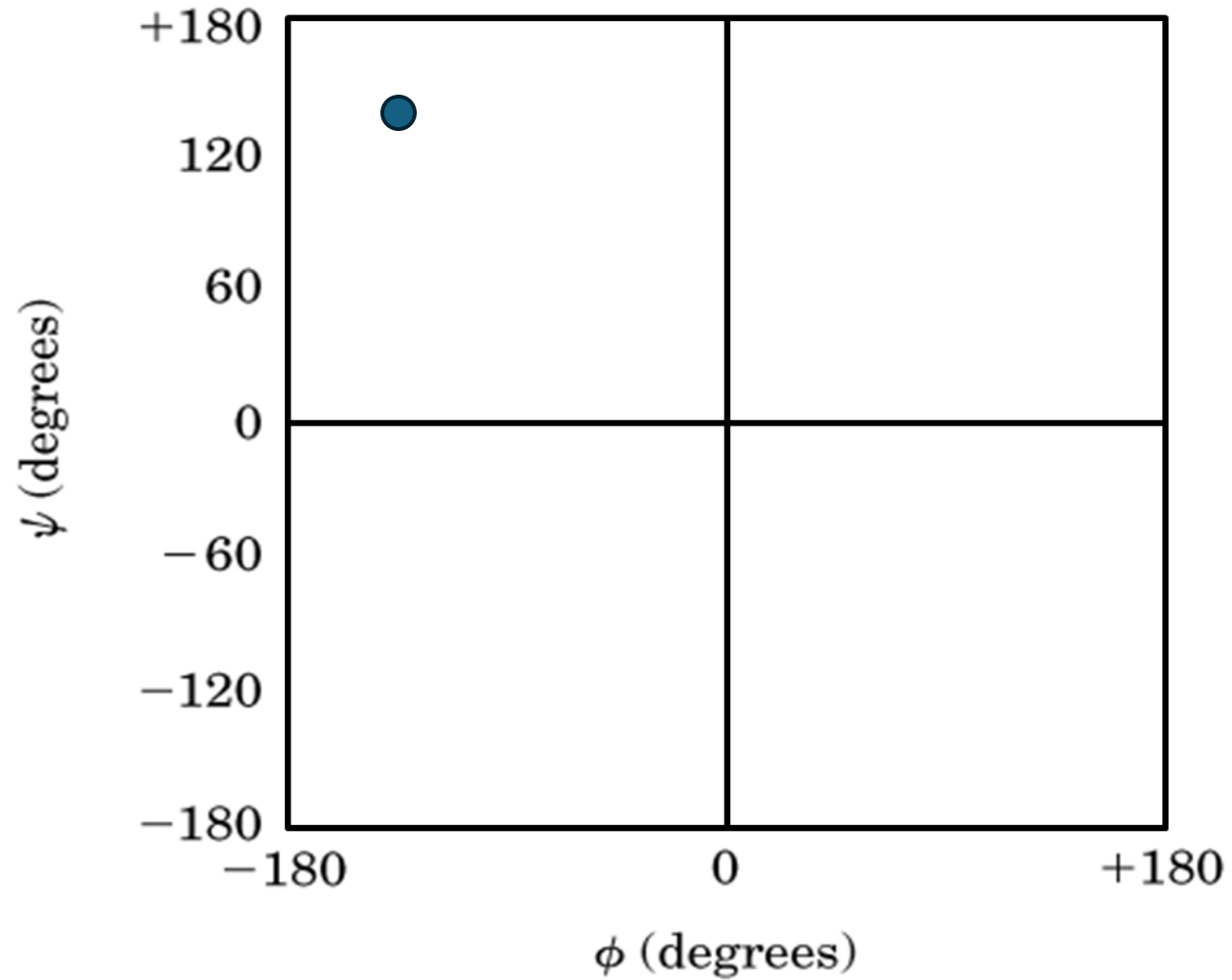


Myoglobin, 1976



Aquaporin

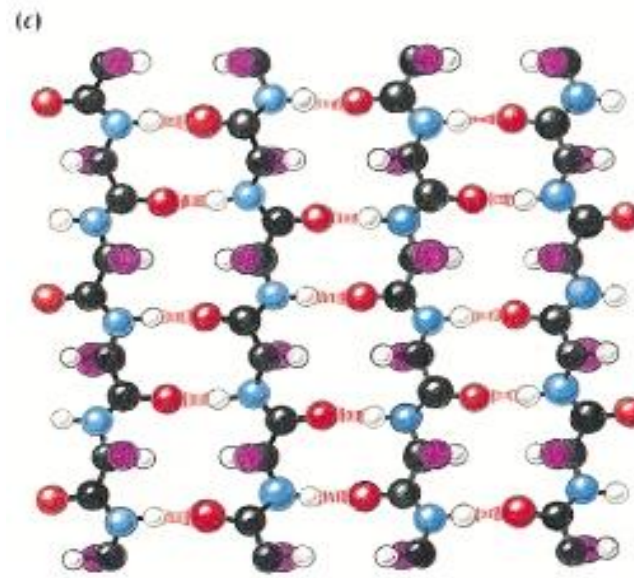
β -strands



©1999 GARLAND PUBLISHING INC.
A member of the Taylor & Francis Group

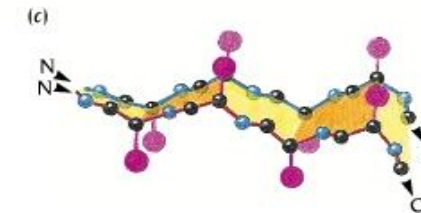
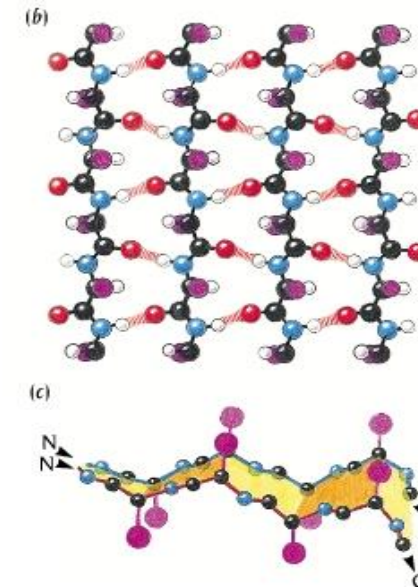
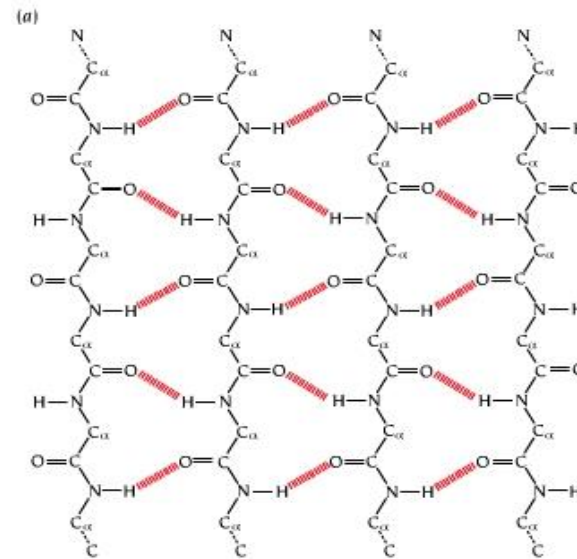
β -sheets

Anti-parallel β -sheet



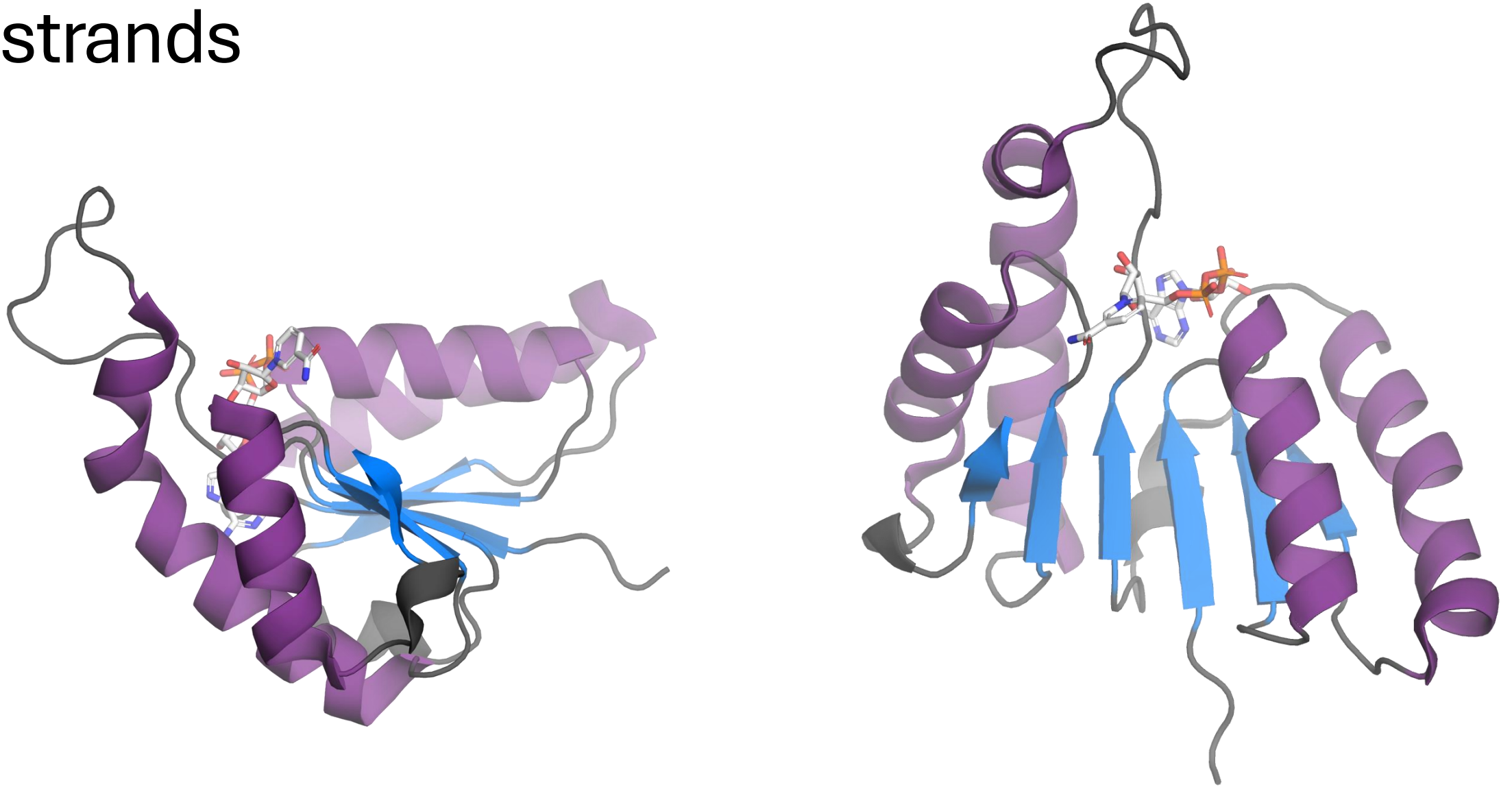
©1999 GARLAND PUBLISHING INC.
A member of the Taylor & Francis Group

Parallel β -sheet

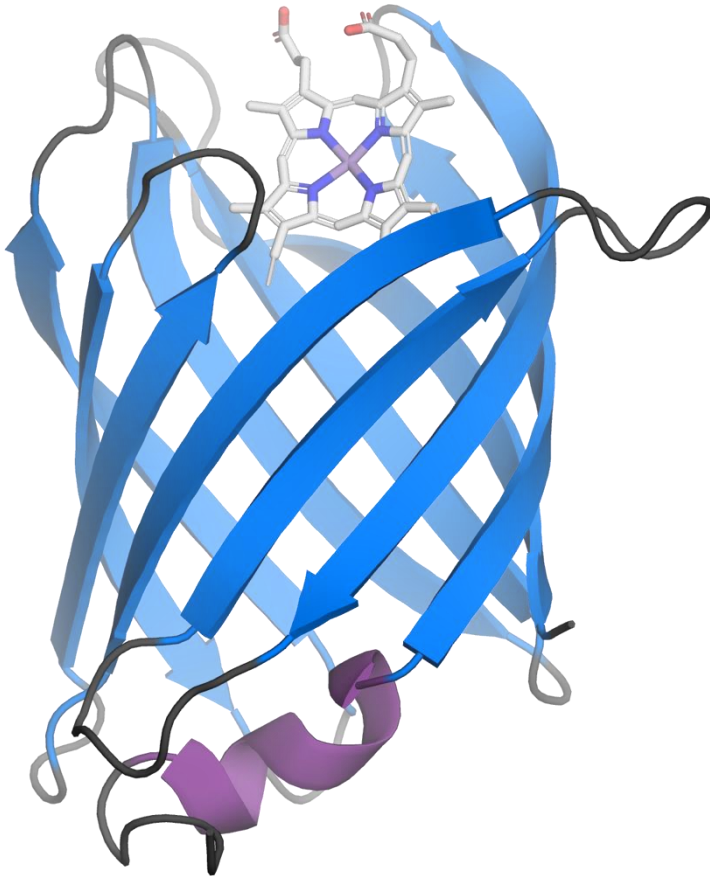


©1999 GARLAND PUBLISHING INC.
A member of the Taylor & Francis Group

Example of beta sheet with twisted beta strands

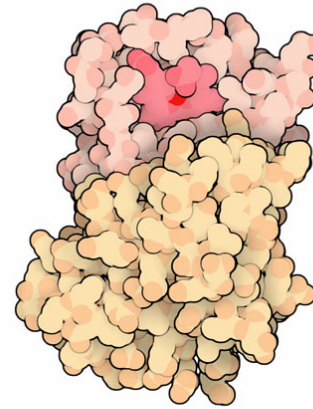


Example of a beta sheet with twisted beat strands



Nitrobindin

December 2010



Proteins perform most of the nanoscale tasks inside of cells, but occasionally, they need help from more exotic molecules. For instance, very small molecules like oxygen are difficult to capture, and proteins like hemoglobin use a heme to trap them. Heme is used in many other capacities as well, including the management of electrons and the capture of other gas molecules such as nitric oxide. So, when researchers at [CESG](#) discovered a new heme-containing protein in the plant *Arabidopsis*, they were faced with an exciting challenge: what is the heme doing?

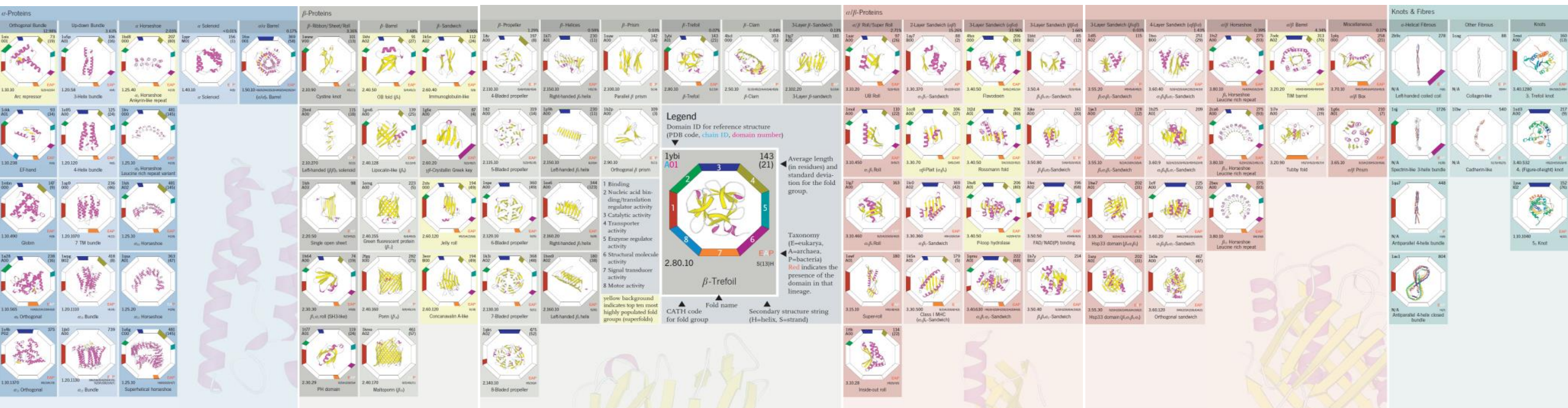
Heme Exposed

The heme in nitrobindin (PDB entry [3emm](#)) is unusual in that the iron atom is rather exposed to solvent. In many heme proteins, the heme is buried deep within the protein, with perfectly-placed amino acids guarding access to the iron atom. For instance, globins have a histidine on one side of the heme, which positions the iron in the proper place, and a histidine or glutamine on the other side, leaving just enough room for oxygen to bind. Nitrobindin, on the other hand, has a similar histidine coordinated directly to the iron, but the other side of the iron is free to interact with water. This has an unusual consequence: in the presence of oxygen, the iron atom is rapidly oxidized and shows only a weak interaction with oxygen.

Managing Nitric Oxide

Testing revealed, however, that the reduced form of the protein binds to nitric oxide (NO) with substantial affinity. This has posed a mystery about the function of the protein. Nitric oxide, in spite of its significant toxicity, is widely used in animal cells as a hormone, in particular, in the local control of blood flow. It plays a similar role in plant cells as part of a complex signaling network that decides what to do when cells are infected or wounded. One clue to the function of nitrobindin is provided by the similar NO-binding protein nitrophorin. Nitrophorin is made by blood-sucking insects and used to deliver NO to their victims, where it dilates the blood vessels and provides more blood for the insect. Nitrobindin may play a similar role in plants, providing a way to store NO safely until it is needed.

A summary of everything



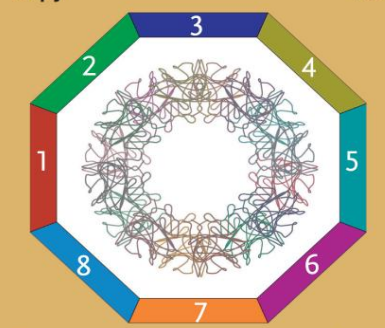
Legend

PDB code for
reference structure
▼

Number
of subunits
▼

1qtj

16



D₈

822

SAP from horseshoe crab

▲ Point group
symmetry
(Schoenflies
nomenclature)

Protein name

▲ Point group
symmetry
(International
nomenclature)

- 1 Multiple sites, cross-linking,
membrane association
- 2 Cooperativity/allostereism
- 3 Cavities, channels and pores
- 4 Functional (active) site formation
- 5 Size and stability
- 6 Economy of genetic material
- 7 "Rulers" (exact separation
between binding sites)
- 8 Multiple functions (in hetero-
oligomers)

Yellow background indicates
dihedral symmetry

Gray background indicates cubic/
icosahedral symmetry

 **WILEY-VCH**

Oligomeric Proteins

Highest Order Rotation Axis

1	2	3	4	5	6	7	> 7
<p>1aqd 2</p>  <p>C₁ 1</p> <p>HLA class II</p>	<p>1wrp 2</p>  <p>C₂ 2</p> <p>Trp repressor</p>	<p>1rtm 3</p>  <p>C₃ 3</p> <p>C-type mannose binding protein</p>	<p>1bi8 4</p>  <p>C₄ 4</p> <p>Potassium channel</p>	<p>1lts 5</p>  <p>C₅ 5</p> <p>Heat labile enterotoxin (B subunit)</p>	<p>1do0 6</p>  <p>C₆ 6</p> <p>HslU ATPase</p>	<p>1i8f 7</p>  <p>C₇ 7</p> <p>SmAP</p>	<p>1lgh 16</p>  <p>C₈ 8</p> <p>LHC II Rhodospirillum rubrum</p>
<p>3hvt 2</p>  <p>C₁ 1</p> <p>HIV reverse transcriptase</p>	<p>2pol 2</p>  <p>C₂ 2</p> <p>Bacterial polymerase III beta subunit (E. coli)</p>	<p>1cd5 6</p>  <p>D₃ 32</p> <p>Glucosamine 6-P deaminase</p>	<p>1cuk 4</p>  <p>C₄ 4</p> <p>RUVB (DNA recombination protein)</p>	<p>1msl 5</p>  <p>C₅ 5</p> <p>Mechanosensitive channel</p>	<p>1g8y 6</p>  <p>C₆ 6</p> <p>Helicase RepA of plasmid RSF 1010</p>	<p>7ahl 7</p>  <p>C₇ 7</p> <p>Alpha-hemolysin</p>	<p>1nkz 18</p>  <p>C₈ 9</p> <p>LHC II Rhodospirillum rubrum</p>
<p>1hzh 4</p>  <p>C₁ 1</p> <p>IgG</p>	<p>3phv 2</p>  <p>C₂ 2</p> <p>HIV protease</p>	<p>1raa 12</p>  <p>D₃ 32</p> <p>Aspartate transcarbamoylase</p>	<p>1dhn 8</p>  <p>D₄ 422</p> <p>7,8-Dihydropteridine aldolase</p>	<p>1gtp 10</p>  <p>D₅ 52</p> <p>GTP cyclohydrolase I</p>	<p>1y12 6</p>  <p>C₆ 6</p> <p>Protein secretion apparatus (HCP1)</p>	<p>1grl 14</p>  <p>D₇ 72</p> <p>GroEL</p>	<p>1wap 11</p>  <p>C₁₁ 11</p> <p>TRP RNA binding attenuation protein (TRAP)</p>
<p>2aai 2</p>  <p>C₂ 1</p> <p>Ricin</p>	<p>4hbb 4</p>  <p>C₄ 2</p> <p>Human haemoglobin</p>	<p>1dps 12</p>  <p>T 23</p> <p>DNA binding protein Dps</p>	<p>1a6d 16</p>  <p>D₄ 422</p> <p>Thermosome</p>	<p>1rvw 60</p>  <p>I 532</p> <p>Icosahedral lumazine synthase (B. subtilis)</p>	<p>1f52 12</p>  <p>D₆ 622</p> <p>Glutamine synthase</p>	<p>1pma 28</p>  <p>D₇ 72</p> <p>Proteasome</p>	<p>1qtj 16</p>  <p>D₈ 822</p> <p>SAP from horseshoe crab</p>
	<p>4pfk 4</p>  <p>D₂ 222</p> <p>Phosphofructokinase</p>	<p>3pcg 24</p>  <p>T 23</p> <p>Protocatechuate 3,4-dioxygenase</p>	<p>2fha 24</p>  <p>O 432</p> <p>Ferritin</p>	<p>1stm 60</p>  <p>I 532</p> <p>Satellite panicum mosaic virus</p>	<p>1g3k 24</p>  <p>D₄ 622</p> <p>HslUV ATP dependent protease</p>		

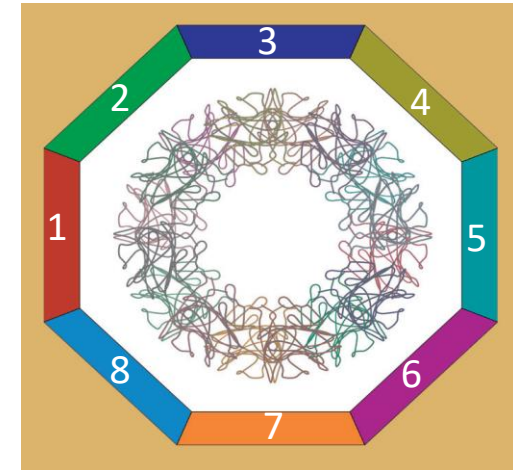
ISBN 978-3-527-31963-3



9 783527 319633

Why be oligomeric?

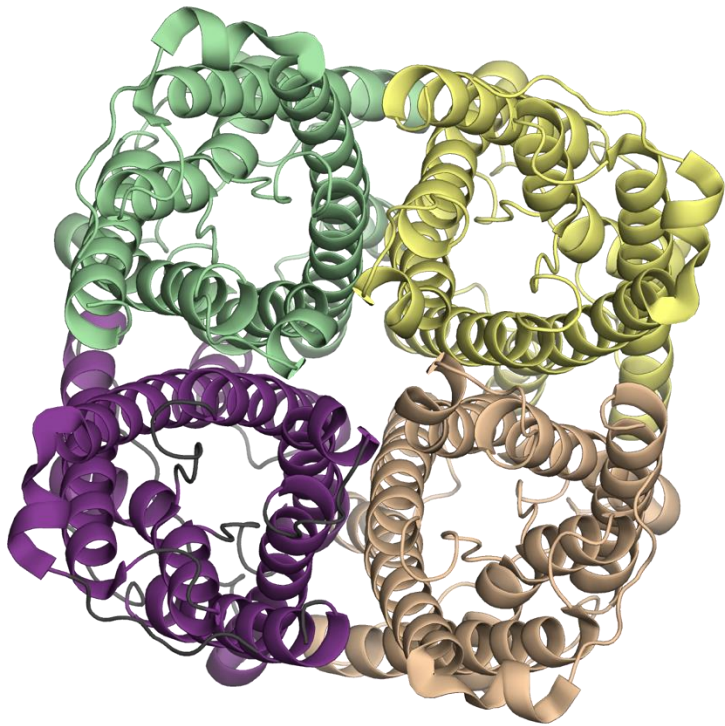
- 1) Multiple sites, cross-links, membrane association
- 2) Cooperativity/Allostery
- 3) Cavities, channels and pores
- 4) Formation of functional sites (active sites)
- 5) Size and Stability
- 6) Economy of genetic material
- 7) “Rulers” - exact operation between active sites
- 8) Multiple functions (in the case of hetero-oligomers)



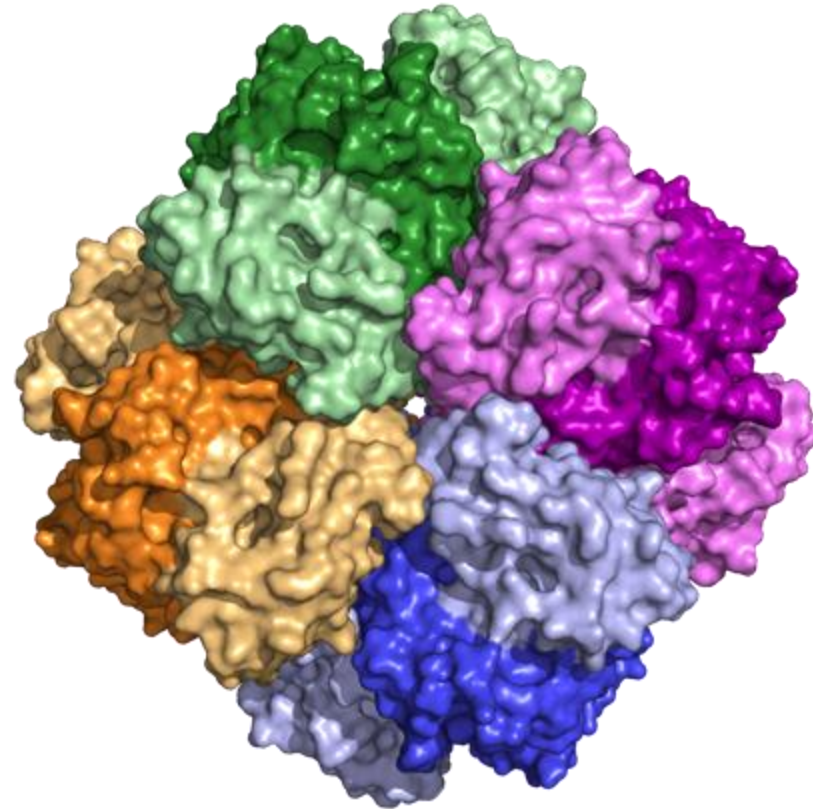
Oligomers distribution in *E. coli*

# subunits	% proteins
1	19.4
2	38.2
3	5.4
4	21.0
5	0.1
6	5.6
7	0.1
8	2.4
9	0.0
10	0.0
11	0.0
12	1.6
>12	2.2
Polimers	2.7

Oligomers



Aquaporin

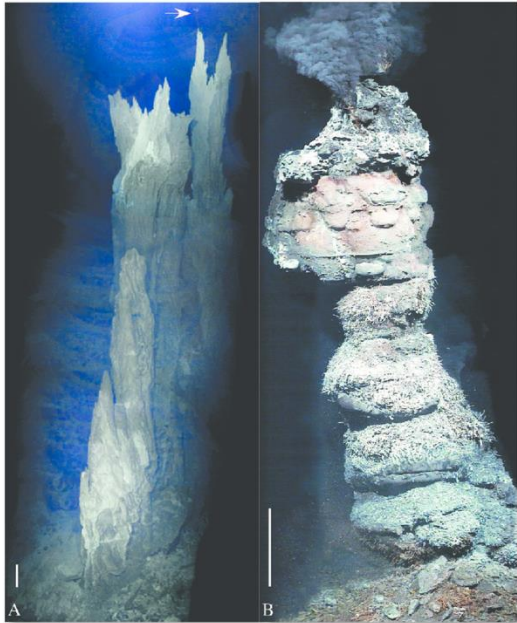


Rubisco

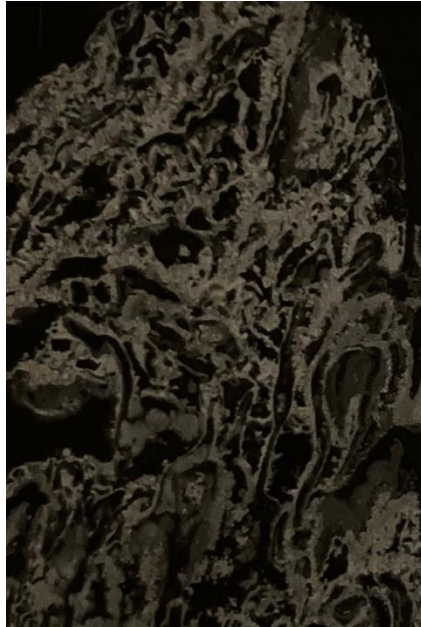
Why exactly these 20 amino acids?

Hypothetical codons of 2 base pairs: $4 \times 4 = 15$ possible amino acids + stop

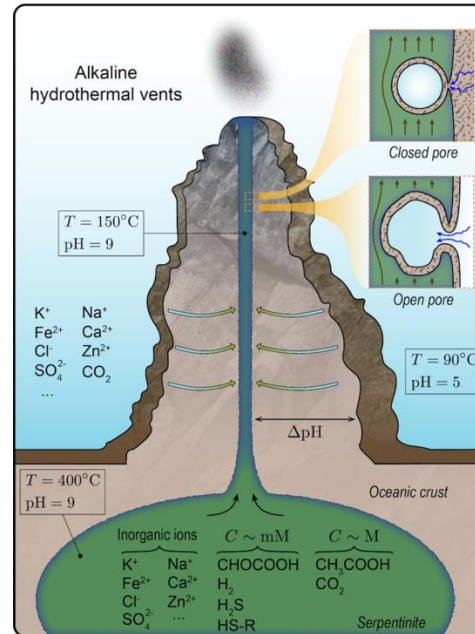
Codons of 3 base pairs: $4 \times 4 \times 4 = 20$ possible amino acids + stop and redundancy



Lane, 2015



Lane, Life ascending



Akbari and Palsso, 2023

1st letter – shared precursor
2nd letter – degree of solubility
3rd letter - degeneracy