Chapter 2: Protein composition and structure

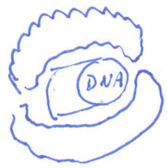
Chapter 2 : protein Composition & Structure

Introduction

1. protein are biopolymer build up by 20 amino Acids and fold up into 3D structure
2. Contains a variety of functional groups (alcohol : T, thiols : S, carboxylic acid : E)
3. interacts each other to form a complex assemblies

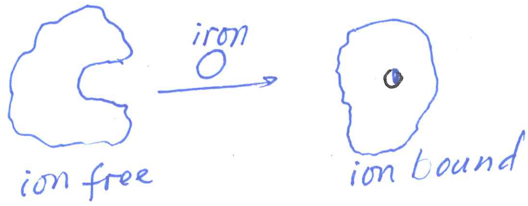
   ex replication of DNA

   signal transduction



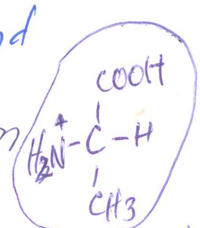
4. Some are rigid, Some are flixible

   (cytoskeleton), lactoferrin + Fe ⟶ conformational change



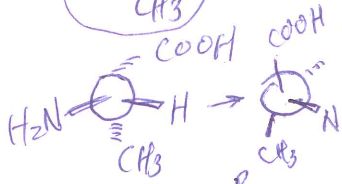
20 Building Blocks

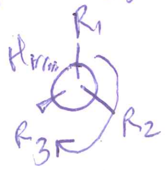
1. Fischer projection, a chiral center/carbon



2. R (rectus, right) } configuration

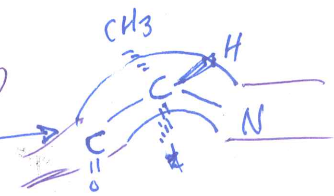
   S (sinister, left)



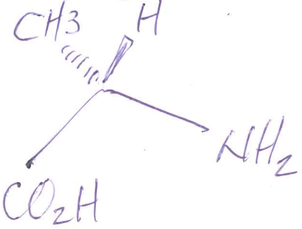


3. D (Dectro-lavorotary) } Form

   L (Levo-lavorotary)



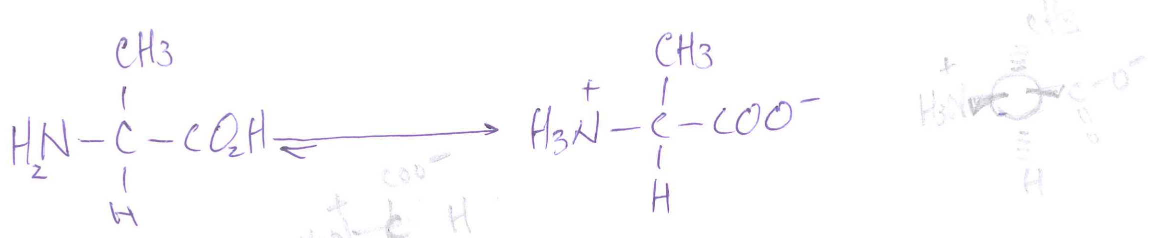
4. Carbon



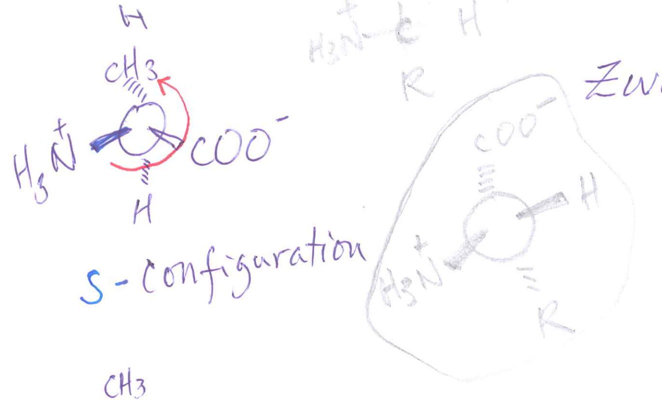
L-form.

5. LSd

6. Zwitterion (dipolar ion)



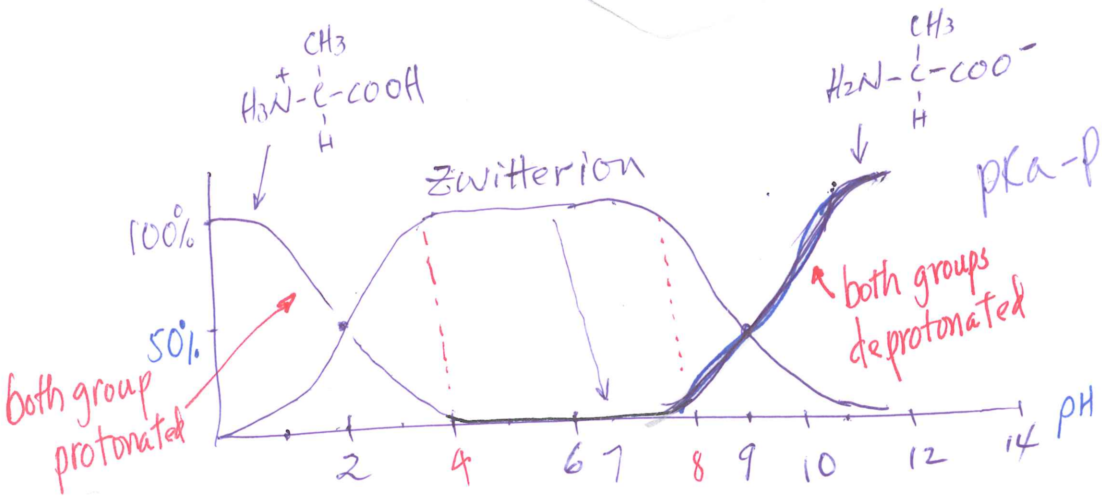
Zwitterion



S-configuration

$R\text{-}COOH \rightleftharpoons RCOO^- + H^+, \ pKa \doteq 2.8$

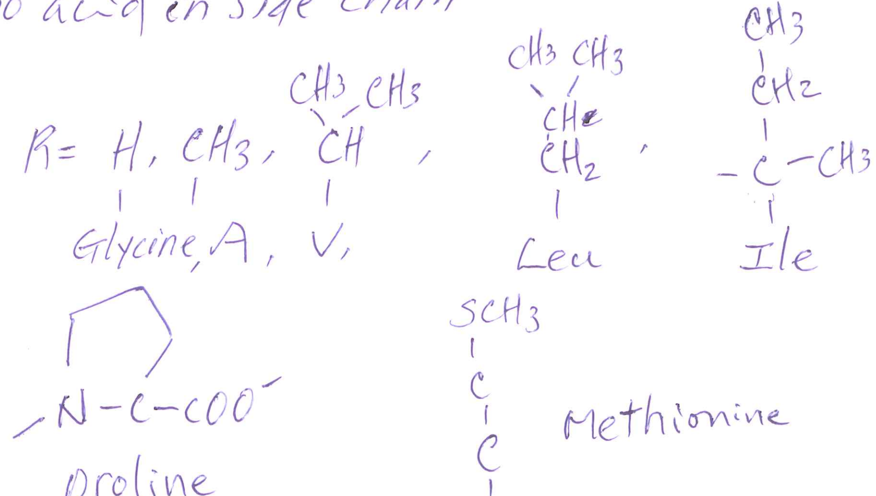
$R\text{-}\overset{+}{N}H_3 \rightleftharpoons R\text{-}NH_2 + H^+ \ pKa \doteq 9$



$$pKa - pH = \log \frac{[RCOOH]}{[RCOO^-]}$$

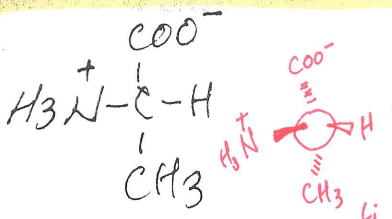
7. 4 family of amino acid in side chain

a. Non-polar : R = H, CH_3,



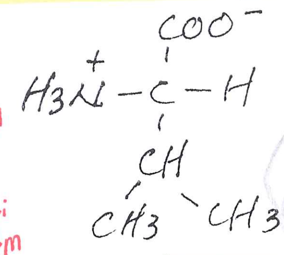
b. polar amino Acid

Amino Acid with non-polar side chain (9)

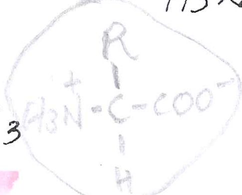


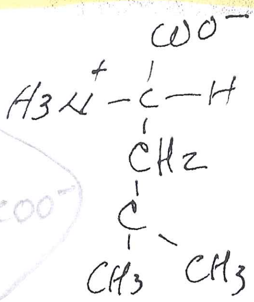
Alanine

S confi
L form

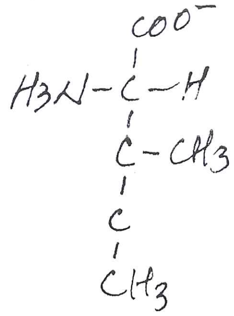


Valine

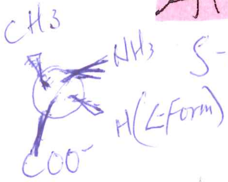




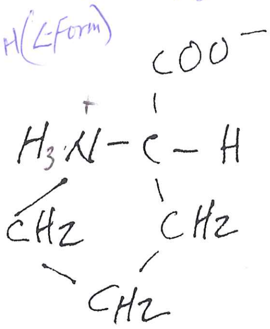
Leucine



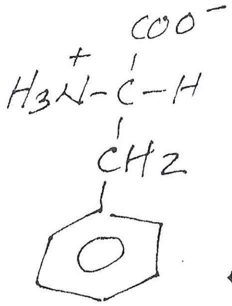
Ile (Isoleucine)



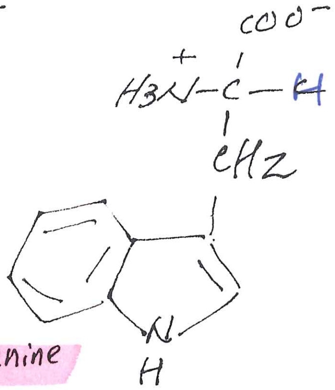
S-configuration!



proline

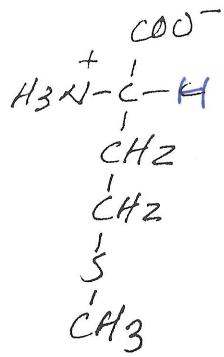


phenylalanine

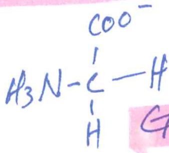


tryptophan

Trp

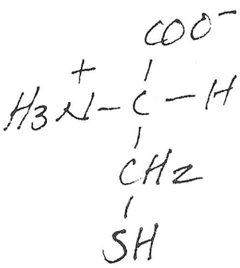


methionine

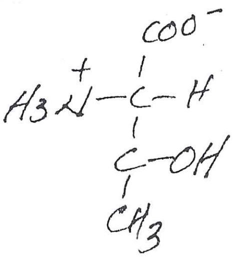


Glycine

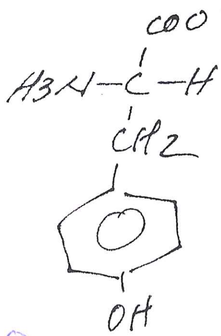
Amino Acid with polar side chain (electronically neutral) (6)



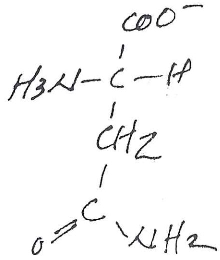
① cysteine
② serine



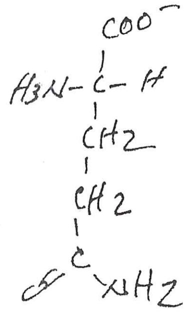
③ Threonine



④ tyrosine



Asn (N)
⑤ Asparagine



Gln (Q)
⑥ Glutamine

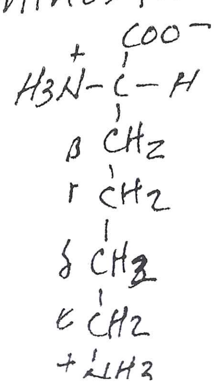
negative side-chain

Amino Acid with carboxyl Group (COO−) in side chain: Asp. Glu (2)

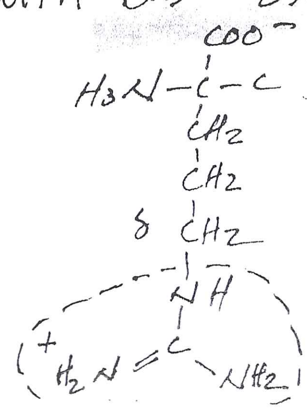
aspartic acid (D)

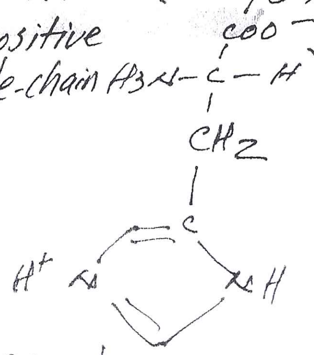
Glutamic acid (E)

Amino Acid with basic side chain: LYS, ARG, HIS



positive side-chain





guanidino

Histidine
Arginine
Lysine
(3)

positively charged: K, R, H

-COOH

pH=3

$-COO^-$

pH=14

$pH = -\log[H^+]$

Negatively charged: D, E

Asp Glu

$RCOOH \rightleftharpoons RCOO^- + H^+$

$$K_a = \frac{[RCOO^-][H^+]}{[RCOOH]}$$

$pK_a = pH +$

$$\log K_a = \log[H^+] + \log\frac{[RCOO^-]}{[RCOOH]}$$

$$pK_a = pH + \log\frac{[RCOOH]}{[RCOO^-]}$$

$$pK_a - pH = \log\frac{[RCOOH]}{[RCOO^-]}$$

(当 pH=7 时 $4 = \log\frac{[RCOOH]}{[RCOO^-]}$

$= 10^4$

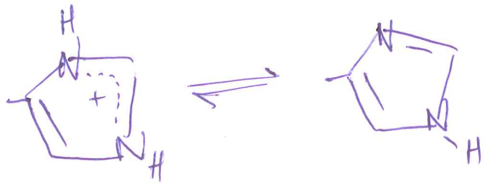
Typical pKa

C-Terminal $-R-COOH \longrightarrow R-COO^-$ $pK_a = 3.1$ (~3)

side-chain D, E $\overset{COOH}{\underset{C-COOH}{CH_2}} \longrightarrow \overset{COO^-}{\underset{R-COO^-}{CH_2}}$ $pK_a = 4.1$

V-D-H-C-Y-K

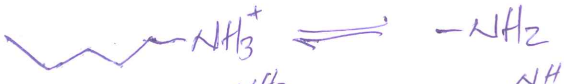
at pH=7

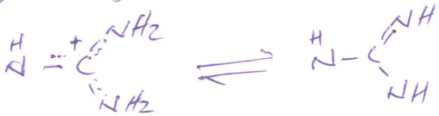
Histidine  $pK_a = 6$

N-Terminal $-NH_3^+ \rightleftharpoons -NH_2$ $pK_a = 8$

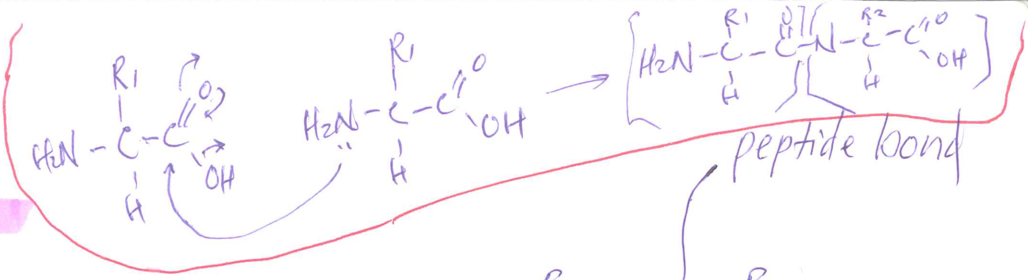
Cysteine $-SH \rightleftharpoons -S^-$ $pK_a = 8.3$

Tyrosin  $pK_a = 10.9$

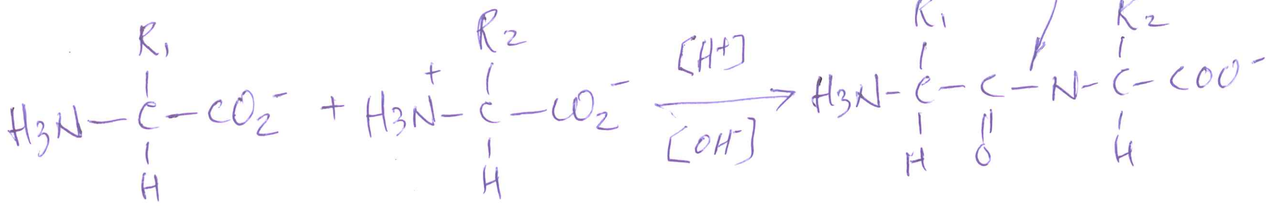
K  $pK_a = 10.8$

R  $pK_a = 12.5$

peptide bond



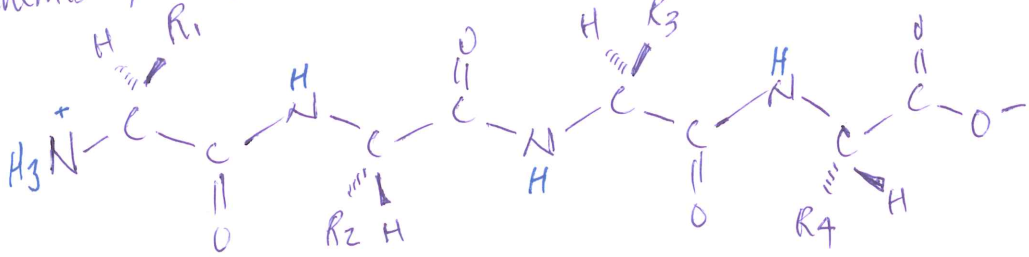
peptide bond



Tyr-Gly-Gly-phe-Leu

N-ter ⟶ C-terminal

Stereo-Chemistry of peptides



Sequence 的重要性

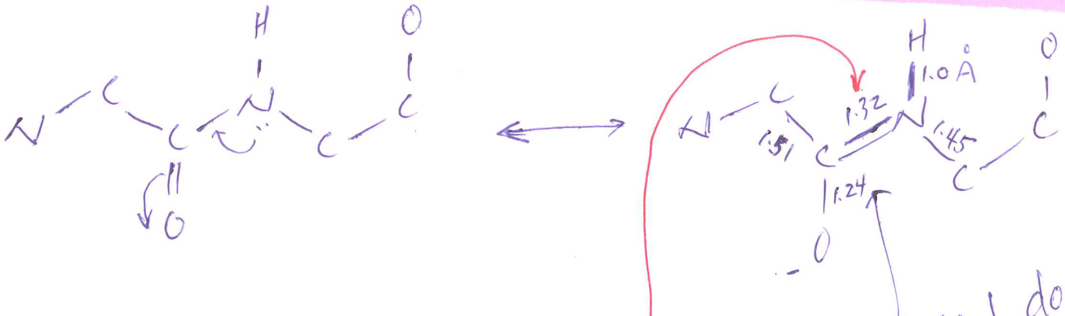
1. catalytic mechanism
2. 3D structure
3. molecular pathology
4. evolutionary History

In 1953. Sanger determine Insulin 的 aa sequence

[1]G — C C — C — C N[21] ⟵ —— 上海

[1]F — C —— C G[20] —— A[30] —— 北京

polypetide chain are flexible yet conformationlly Restricted



partial double bond

C—N bond length 1.49 Å ⟵ 1.32 Å

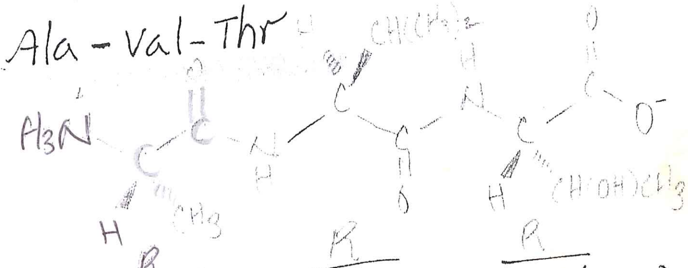
C=N " " 1.27 Å

Chapter 8: 3D Structure of protein

X-ray: pepsin structure (1934) –
α helix & β-sheet of protein (1951) — pauling (+1962)
DNA double helix (1953) — Watson & crick
Myoglobin (1958) — John Kendrew
Hemoglobin (1968) — Max perutz
30,000 structure in PDB (2006)

primary structure: Amino Acid Sequence
Secondary structure: backbone
Tertiary structure: backbone + side chain
Quanternary structure: spatial arrangement of subunit

Fig. 6-1 Ala - Val - Thr

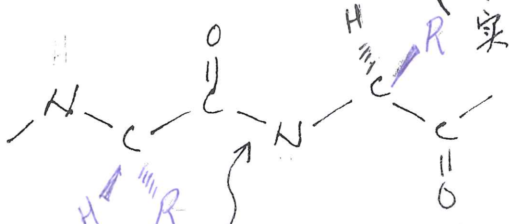


extended form

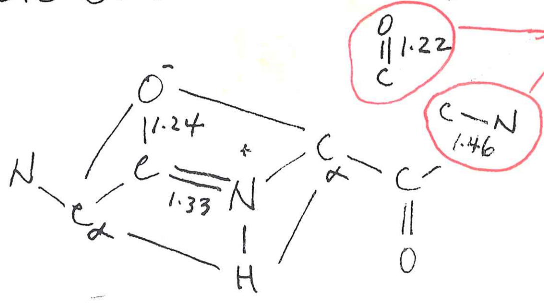
Secondary structure

peptide bond (~40% double bond character)

支鍵: (上右, 下左 虛) 实



peptide bond ⇅



正常

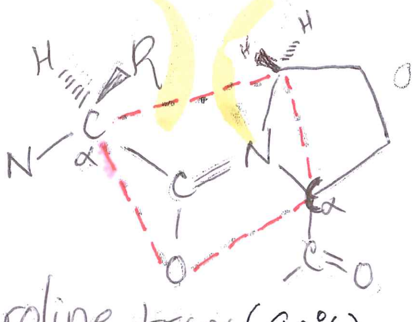
~40% of double bond

2方 Cα 在 trans.

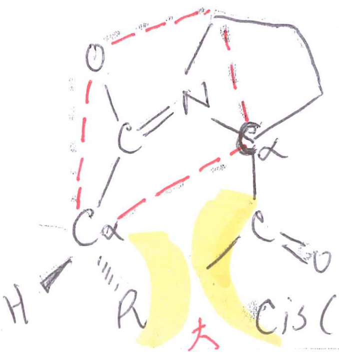
~~E LL cis conformatn, ~2 Kcal/mol~~

$E_{cis} = E_{trans} + 2\ Kcal/mol$

小



proline trans (90%)

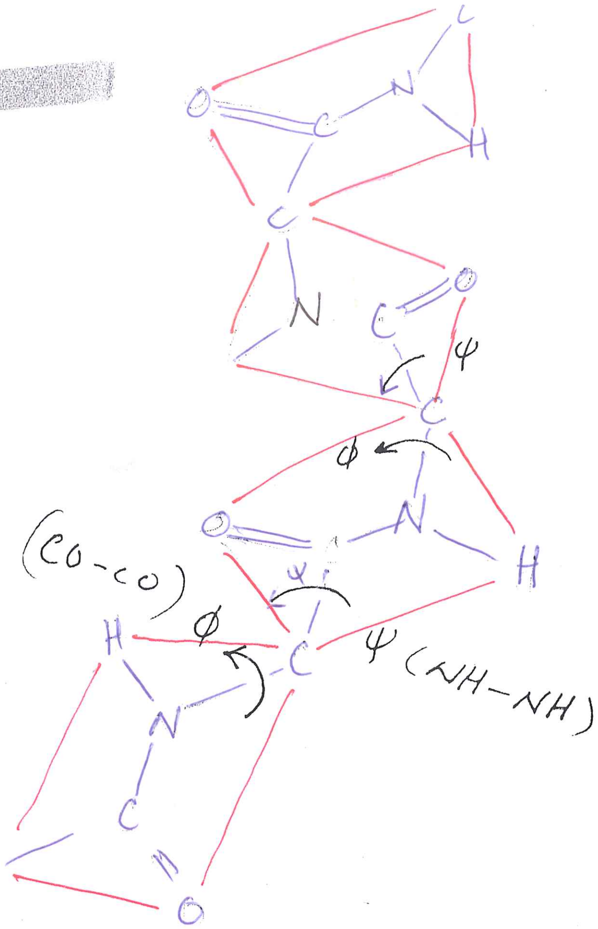


proline cis (10%)

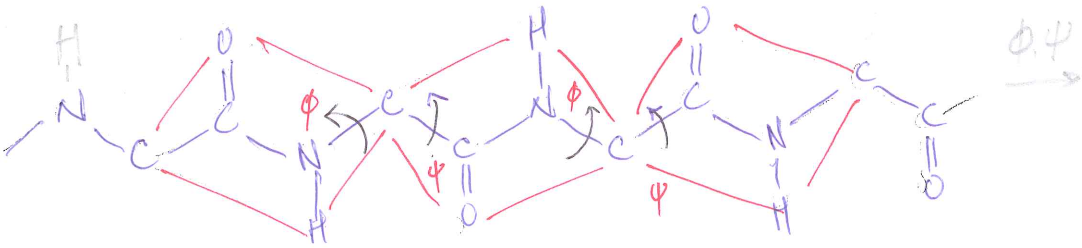
大

cis (10%)

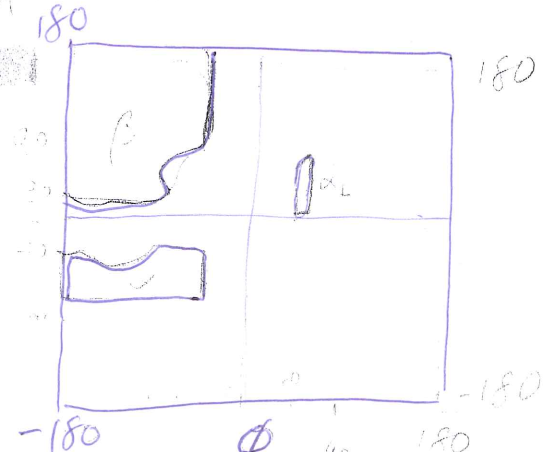
φ,ψ angle



φ,ψ angle of a polypeptide chain

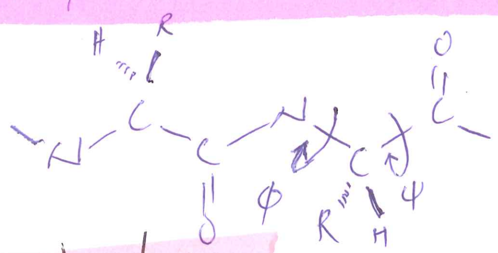


Ramachandran Diagram



proline cis/trans isomer

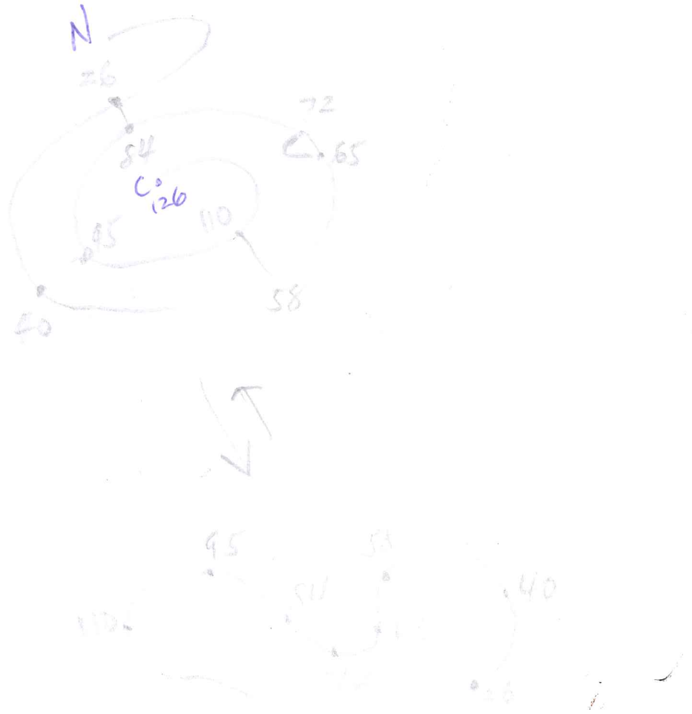
ψ & ϕ angle / Ramachandra plot



primary structure

Secondary structure

α-helix
β-sheet } 見右



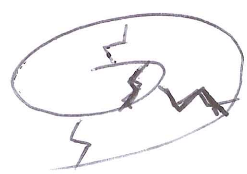
Tetiary Structure

side-chain & 3D

Quaternary Structure

Subunit.
Hemoglobin (tetramer)
Cro protein (Dimer)

Motif (helix-turn-helix)



Domain (CD4 consists 4 similar Domain)

Christian Anfinsen

bovin ribonuclease

β-mercapto-ethanol

scramble
$C(8,2)=105$
$\frac{1}{7}\times\frac{1}{5}\times\frac{1}{3}\times\frac{1}{1}$

Native $\xrightarrow[\beta\text{-mercapto-ethanol}]{8M\ urea}$ Denature & reduced $\xrightarrow{[O]}$ scrambled $\xrightarrow[\text{ethanol } [O]]{\text{mercapto}}$ Native

Structure prediction

α-Helix propensity

β-sheet "

Turn propensity

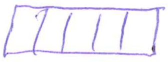
Protein Misfolding

, Mad cow disease (1990s, UK)

bovin spongiform encephalopathy (BSE)

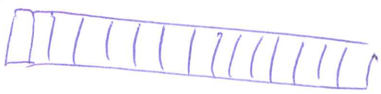
CJD; Creutzfeldt-Jakob Disease

Stanley Prusiner, 1997 (prions protein)



$\underbrace{\text{Normal}}_{}$ cellular protein

Normal PrP^c $\xrightarrow[\text{aggregate Nucleus}]{PrP^{sc}\ \text{Nucleus}}$ PrP^{sc}

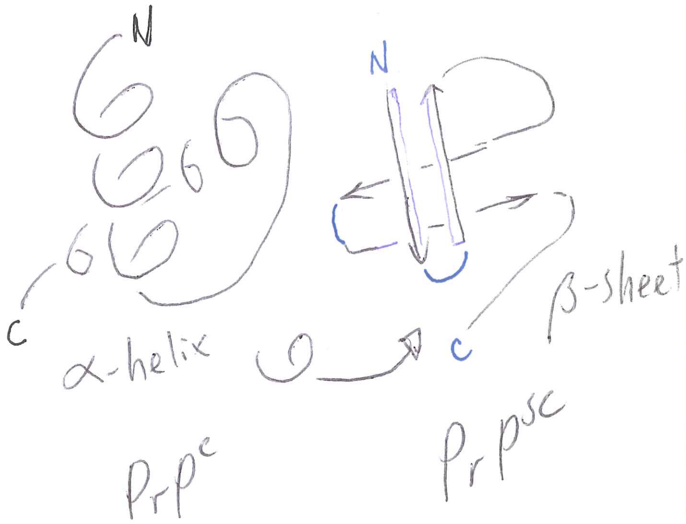


Alzheimer

Parkinson Disease

amyloid plaques (fribil)

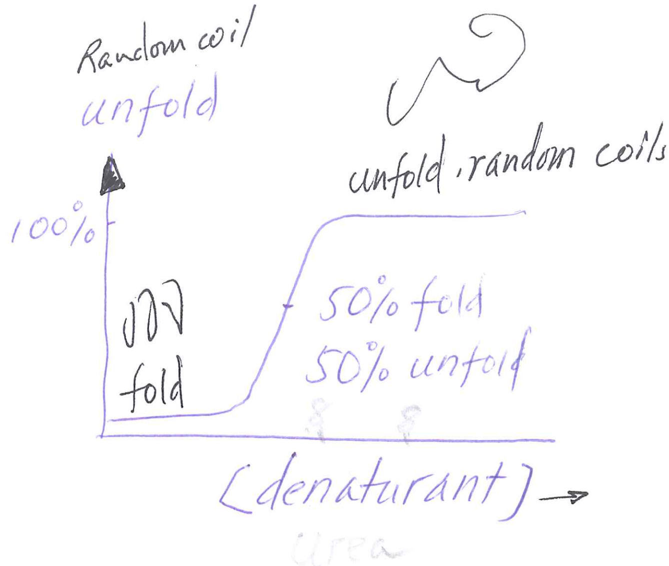
斑塊



scrapie

protein folding

Co-operativity : all or none



A statement

apparently valids
but leads to
logically unacceptable

john only cut the hair who don't cut hair themself

Cyrus Levinthal 似是而非的 paradox (MIT)

100 a.a protein

$3^{100} = 5 \times 10^{47}$

$\frac{5 \times 10^{47}}{10^{13}} = 5 \times 10^{34}$ sec $= 1.6 \times 10^{27}$ years

比地球的壽命都長 4.5×10^{9}

However, in cell all protein are folded in minutes, if not in seconds.

Cumulative Selection

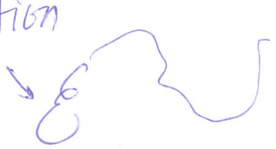
Monky & typewriter

Nucleation - Condensation

Here are the rules:
Ignore all rules.

The 2nd sentence is false,
The 1st sentence is true

Nucleation



Denature → Native