

LSM 3243

By the end of the course, you should be able to

1. Describe protein backbone and side-chain conformations in terms of dihedral angles and interactions
2. Understand the factors that determine protein conformations
3. Know common nucleic acid conformations and to understand the interactions that stabilize such ordered conformations
4. Know the composition of cell membrane and understand unique features of membrane proteins
5. Understand the structures of micelles and bilayers
6. Understand how small molecules are transported across membranes
7. Understand basic concepts of spectroscopic techniques: circular dichroism, fluorescence and nuclear magnetic resonance
8. Be knowledgeable to the application of these techniques to life sciences
9. State transition
10. Protein folding
11. Protein-ligand interaction

Week1 Protein Structure

Introduction

1. Only L-amino acids are found in natural proteins.

Look along the H—C_α bond, seeing CO-R-N in clockwise indicates L.

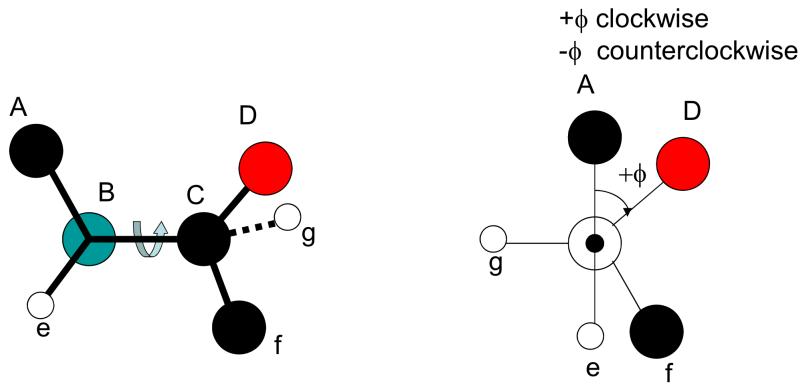
2. Nomenclature of main/side chain atoms: α, β, γ...

Polypeptide chain

peptide bonds

- N is partially positive (resonance?) while O is partially negative.
- The 6 atoms are coplanar!

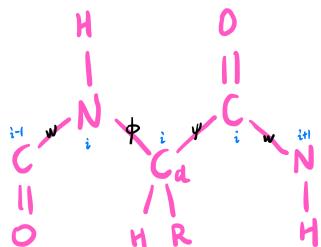
torsion angle



- Look through B-C, if B-A (the nearer) needs to rotate clockwise to reach superposition with C-D, the angle is positive.
- But in which direction we look through doesn't matter because the result is the same.

backbone dihedral angles

- ϕ : N-C_α
- ψ : C_O-C_α
- ω : C_O-N

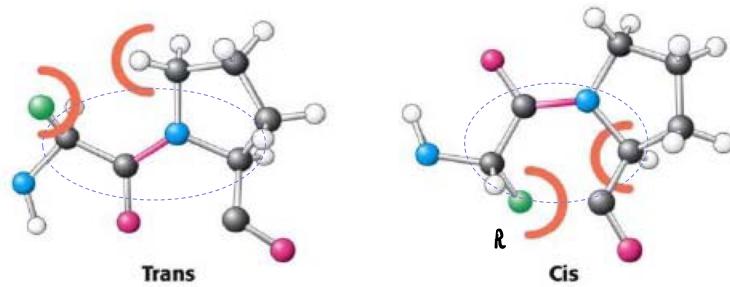


The protein structure is almost determined by these angles.

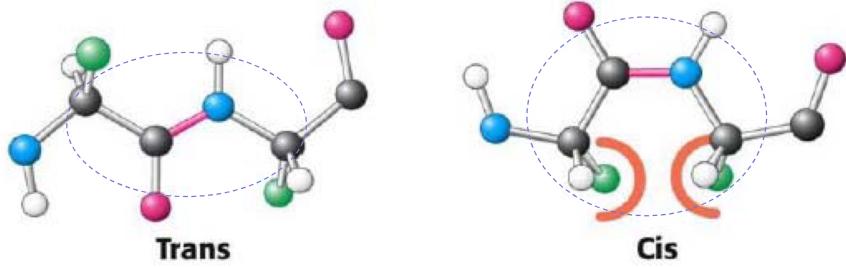
ω angle

It is usually fixed.

- For proline, groups in trans is less repulsive, but there's cis.



- For others, trans is almost always favored (8 kJ/mol).

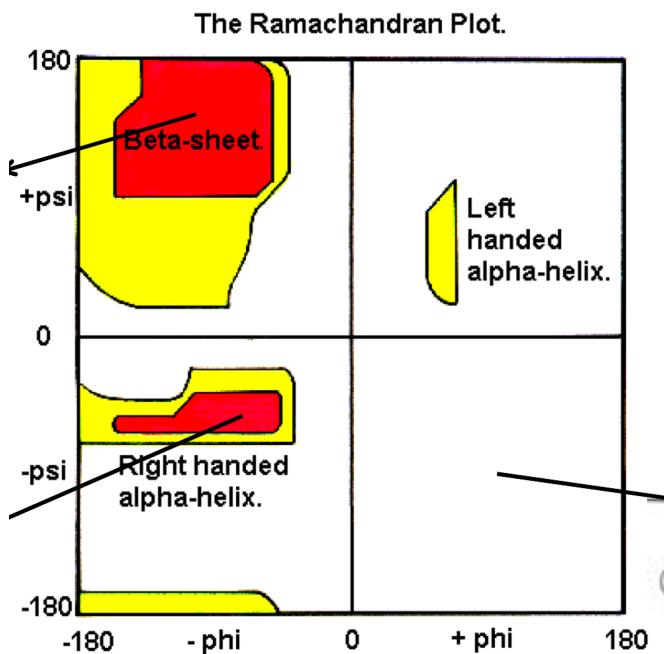


- However, some X-Pro structures are cis

ϕ and ψ angles, the Ramachandran Map/Plot

Assumption: No two atoms in a molecule can come any closer than the sum of their **van der Waals radii**. Those who disobeys it are called "disallowed conformations".

- The plot is similar for most amino acids.



- There are semi-allowed regions, where the distances are slightly less than the sum of vdw radii.
- For a certain kind of protein, we can draw the plot according to statistical data. Those who are out of allowed region might be in linkers, turns or loops.
- two special AAs
 - The plot of Gly has a larger area and is centralsymmetric. Both are due to the R group is simply a small H atom.
 - ϕ of Pro is fixed to about 60° because of the fixed cyclopentane. ψ is most favored at about -60° and 160°
 - Gly often interrupts α and β . Pro cannot exist in anti-parallel β sheet. Instead, they are often located in turns and loops.

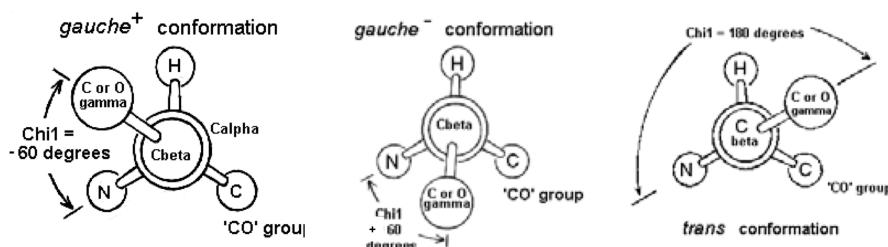
structure features of typical conformations

conformations	approximate ϕ and ψ ($^\circ$)	features	special structures
---------------	--	----------	--------------------

conformations	approximate ϕ and ψ ($^{\circ}$)	features	special structures
α helix	-57, -47	3.6 residues/turn, 100° /residue CO of i th and NH of $i+4$ th AA form H bond side chains protude radically outside helical wheel projection: projection along the axis of α helix. 18 AAs, 5 cycles. Maybe nonpolar AAs are located on the same side and polar AAs on the other side	3_{10} helix: at the termini of a regular α helix
anti-para β sheet	-139, 135	CO and NH on adjacent chains form H bond with each other to stabilize	parallel β sheet: a little more tortuous H bonds
β turn	--	$i+1$ is usually Pro and $i+2$ usually Gly at the surface of protein (reverse the direction) nucleation center of folding	

side chain dihedral angles

- $\chi_1 : N-C_{\alpha}-C_{\beta}-C_{\gamma}$, $\chi_2 : C_{\alpha}-C_{\beta}-C_{\gamma}-C_{\delta}$, then goes along the side chain.
- Possibilities of χ_1 . All are avoiding overlapping with the main chain.
Because C=O is the bulkiest group, gauche⁺ ($\chi_1 = -60^{\circ}$) is the most favored. (trans is 2nd where $\chi_1 = 180^{\circ}$)



- χ tend to adopt staggered conformations (交错式构象).
 - For aromatic residues, χ_2 tends to be $\pm 90^{\circ}$ to minimize close contact with main chain.
 - For residues who form H bond with the environment, the last χ (Asp, Asn: χ_2 ; Glu: χ_3) adopts a wide range since that bond always rotates.

For those like Tyr who have two ways on the chain, both sides have the same χ number.
This is not so meaningful.

In addition

1. tertiary structure: combination of secondary structures
2. quaternary structure: specific assembly of folded subunits (3rd struct)
3. examples
 - kertain: coiled coil (α)
 - silkworm silk fibers: β
 - collagen: left-handed triple helix (Pro)

Week2-1 Interactions in Proteins

Basic concepts

Internal energy (内能) involves:

- kinetic energy
- potential energy

Recall three thermodynamical laws.

Life is but a interplay of weak forces (non-bonding interactions).

bond	energy (kJ/mol)
ionic (COO^- and NH_3^+)	20~80
hydrogen	4~50, depending on distance and orientation
dipole (CO and CO)	5~10
vdw	0.5~3 per atom pair
hydrophobic	4~8 per non-polar groups

functions:

- additional strength to stabilize (high level structure)
- flexible, interact and perform functions

general electrostatic interactions

ion-ion

atoms carrying charge

- equation

$$V = \frac{Q_i Q_j}{4\pi \varepsilon_0 \varepsilon_r r_{ij}}$$

- ε_r : relative dielectric constant
- The force is weaker in water (larger $\varepsilon_r = 78.5$) than in hydrophobic core (though hard to measure)/other organic solvents.

- salt effect

$$I = \frac{1}{2} \sum c_i q_i^2$$

$$D = \sqrt{\frac{\varepsilon_0 \varepsilon_r k_B T}{2 N_A e^2 I}}$$

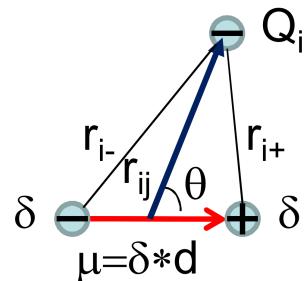
$$V = \frac{Q_i Q_j}{4\pi \varepsilon_0 \varepsilon_r} \frac{\exp(-r_{ij}/D)}{r_{ij}}$$

$I \uparrow, V \downarrow$. "Neutralizing charges"

eg: N and C terminal, side chain (salt bridge)

ion-dipole

- ion with groups carrying no formal charge
- dipole: **negative** → **positive** point charge
- dipolar molecule: those with dipole moment (H_2O , CO)

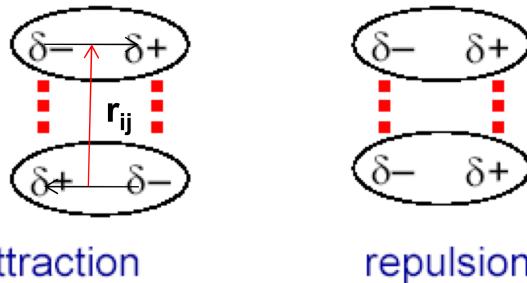


$$V = \frac{Q_i \delta}{4\pi \varepsilon_0 \varepsilon_r} \left(\frac{1}{r_{i-}} - \frac{1}{r_{i+}} \right) = \frac{Q_i \delta}{4\pi \varepsilon_0 \varepsilon_r} \cdot \frac{r_{i+} - r_{i-}}{r_{i-} r_{i+}} \approx \frac{Q_i \mu}{4\pi \varepsilon_0 \varepsilon_r r_{ij}^2}$$

eg: dissolve salt. $\Delta H = E_{lattice} - E_{hydration}$, hydration: ion with water.

entropy may increase.

dipole-dipole



$$V = \frac{1}{4\pi\epsilon_0\epsilon_r} \left[\frac{\mu_i \cdot \mu_j}{r_{ij}^3} - \frac{3(\mu_i \cdot r_{ij}) \cdot (\mu_j \cdot r_{ij})}{r_{ij}^5} \right]$$

The potential depends on the relative orientation. The system tends to lower the energy.

eg:

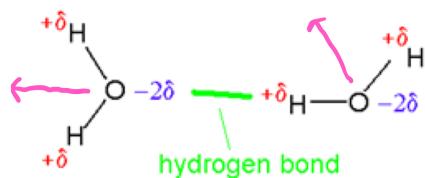
- d-d interaction: $\text{H}_2\text{O} > \text{HCl}$
- interaction between α helices (accumulation of peptide bond polarity), important for attracting charged molecules and enhances reactions.

van der Waals

here refers to dispersion forces

- features
 - temporary dipole, induced non-uniform e⁻ distribution
 - contact distance, L-J potential
- important for molecules both with and without permanent dipoles
- significant for large molecules

hydrogen bond



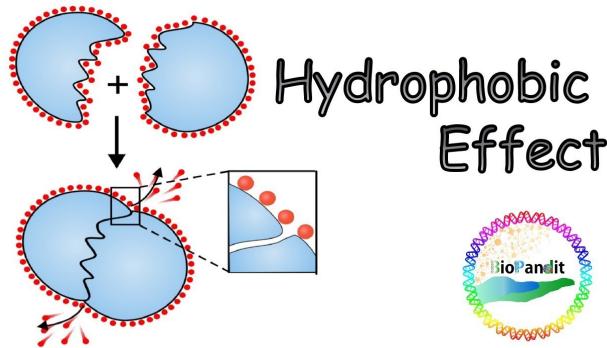
- features
 - short range: $r_{bond} \ll r < R_{vdw}$
 - directionality: 120~180°
- affects secondary structure & recognition, but not dominating/determining folding/assembly
- only intramolecular H bonds in the interior of a protein are favorable in presence of competition from water

actually vdW and H bond are both special cases of dipole-dipole interaction

hydrophobic interaction

Oil doesn't mix with water. The molecules attract themselves more than each other. Thus mixing water and oil is only a process where water molecules form a cage (more ordered, to strengthen interaction) rather than thorough mixing.

Their attraction increases ($\Delta H < 0$) but the deterministic factor is $\Delta S(H_2O) < 0$. To minimize the entropy decrease, hydrophobic parts tend to gather to make their volume lowest and water molecules to form a cage least.



Disulfide bonds

Fold first, then disulfide bonds!

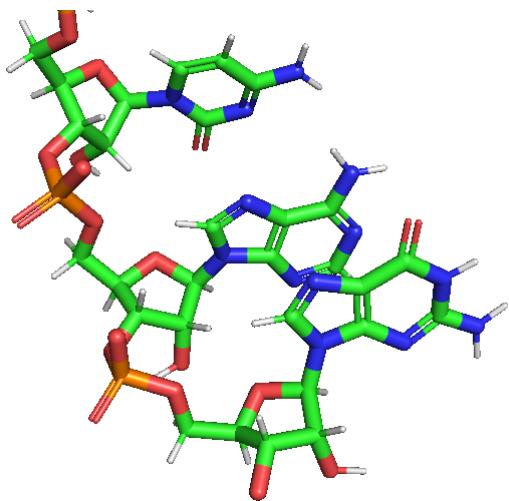
Week2-2 Nucleic Acid Structure

component

- pentose (difference)
- base
- phosphate (phosphoester bond)

conformation

- bases point inwards
- plots of torsion angle
 - main chain: extended (trans, 180°)
 - χ : anti (0°)

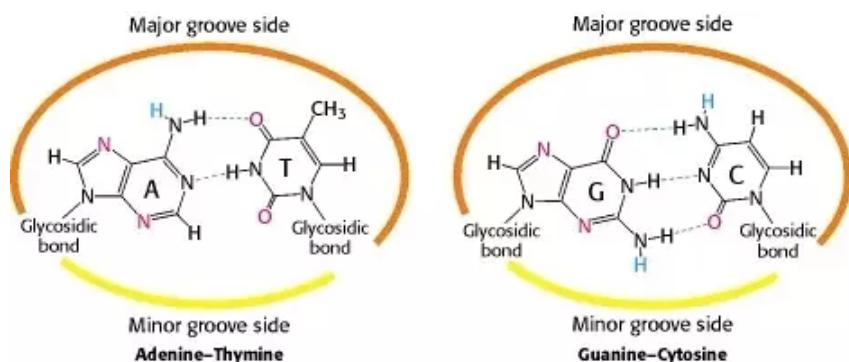


secondary structure

double helix

helix	(B) DNA	(α) protein
units per turn	10.5 (10)	3.6
degree per unit	36°	100°
rise per unit	0.34 nm	
diameter	22 nm	

grooves



grooves	major	minor
part of base	more	less
actual H bonds	fewer	more
bind	proteins	drugs

Most proteins perform non-sequence-specific binding, switching on/off gene expression.

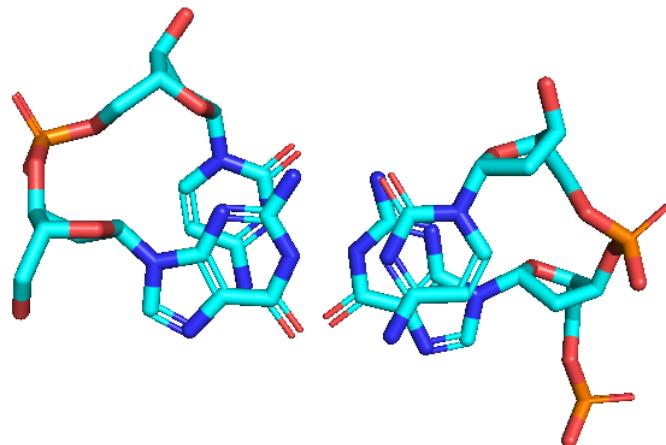
forces to stabilize

hydrogen bond

Watson-Crick base pairs vs unusual pairs

fraction of GC → melting temperature. AT: local unwinding

stacking interaction



GC superimposition

is a combination of

- vdW, where 0.34 nm is the optimal distance
- dipole-dipole, same direction ⇔ repel

bases on the next level rotate 36° as a balance

electrostatic interaction

negative charged (-1 per bp) phosphate backbone: repulsion, controls distance between two strands

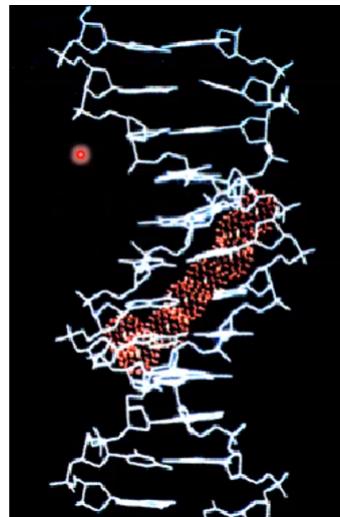
- low-salt: (phosphates) try to be trans; bases farther, prefer single strand (PCR)
- high-salt: screen (掩蔽) them to form double strand DNA

DNA prefers to take helix to 1) form stacking 2) avoid repulsion

with water

Water forms H bond with bases, pentoses and phosphate groups. The major and minor grooves potentially form same amount of H bonds. But due to larger space, the major groove needs more water molecules. This is the cause of **entropy loss**.

As a result, the minor groove forms **more H bonds** with water, i.e. more enthalpy gain. Thus, it can trap more water molecules (a cluster that may not form H bonds) and compensate the entropy loss. In contrast, the major groove only maintains one shallow layer of water.



spine of hydration

other

- tightness: Z form (left-handed, seq-specific)>A form (seq-specific/RNA, high salt)>B form (normal)
- RNA structure

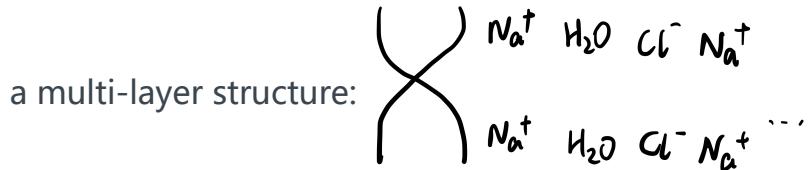
Week3-1 HW

1. interactions

	protein	nucleic acid
interactions	covalent bond (disulfide) electrostatic (ion and dipole) vdW H bond hydrophobic	covalent bond H bond stacking (vdW, hydrophobic) electrostatic
secondary	H bond	H bond (stacking)
tertiary	H bond, hydrophobic	H bond, stacking

2. heat and salt?

salt stabilizes the backbone, pushing the phosphate groups nearer and thus harder to be denatured by heat.



Week3-2 Structure of Cell Membrane

functions of cell membranes

- compartmentalization
- selective permeable
- communication, signal transduction
- organize activities, energy production

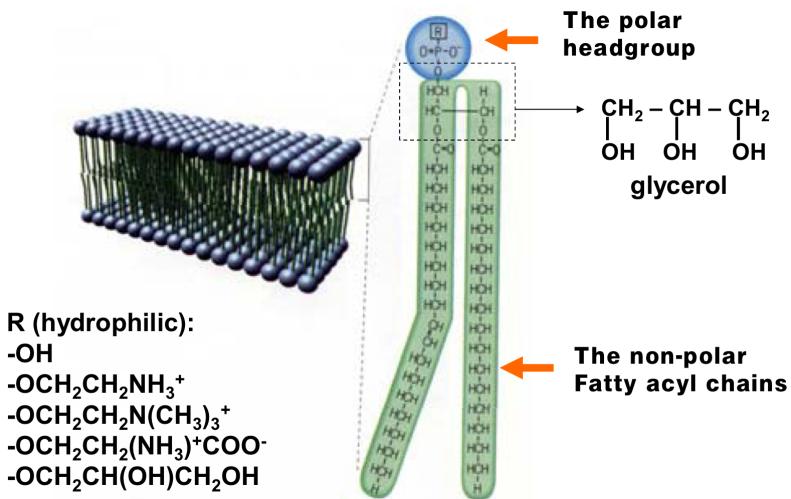
lipids

types

- glycerophospholipids
- sphingolipids
- cholesterol

features

- **amphiphilic**
- different content in different species or organs



lipids on the membrane

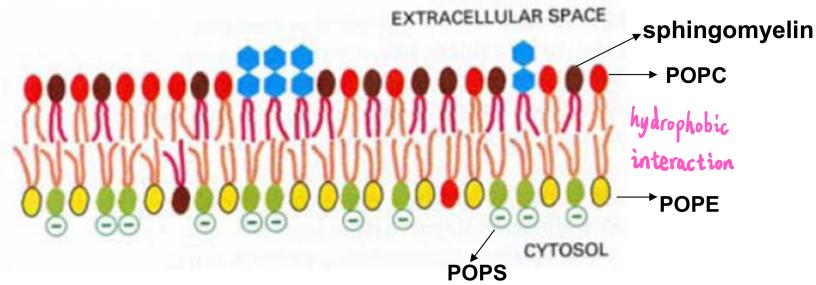
1. assymetric distribution (in two layers)

- negative charged phospholipids tend to point at cytosol while the neutral ones point at extracellular space

when cell undergoes apoptosis, POPS goes to the outer membrane because entropy increases.

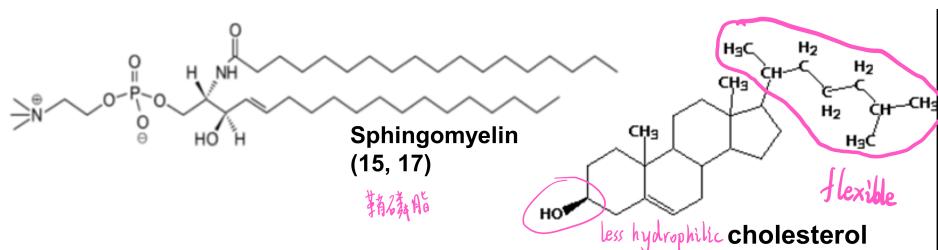
2. random distribution (in the same layer)

- lipids are randomly distributed, and swim.. (bulk phase)
- vdW forces are not strong enough to hold lipids together**
- hydrophobic effect drives the lipid to assemble, so there must be water



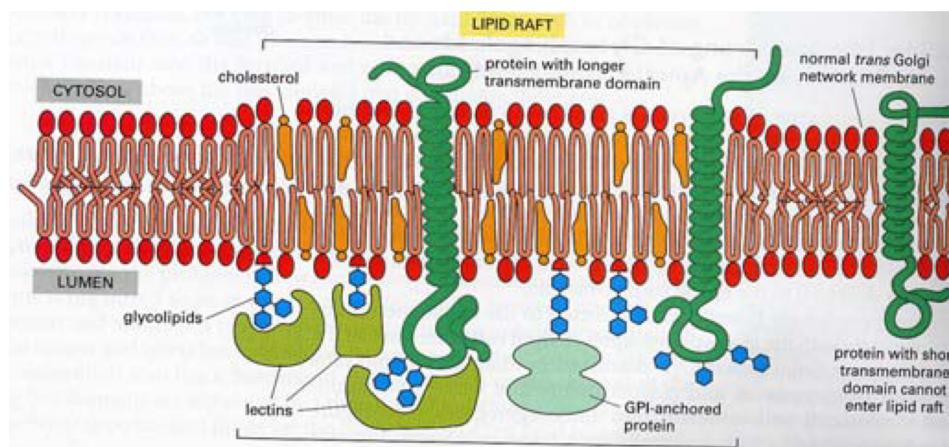
lipid rafts

- sphingolipids have longer and straighter fatty acids thus have bigger vdW forces to hold themselves together (transiently)
 - glycerophospholipids: about 13?
- cholesterol can hold more lipid molecules
 - it's also very "hydrophobic"
 - it interacts with fatty acids in both sides of the plane, and has bigger surface area.
 - its ring is a rigid structure that makes lipid rafts stable.



They assembly into lipid rafts, which

- are thicker (longer) than other parts of the membrane;
- are more resistant to detergents
- accommodate and gather proteins for specific functions like singaling



proteins

ratio ↑, function ↑

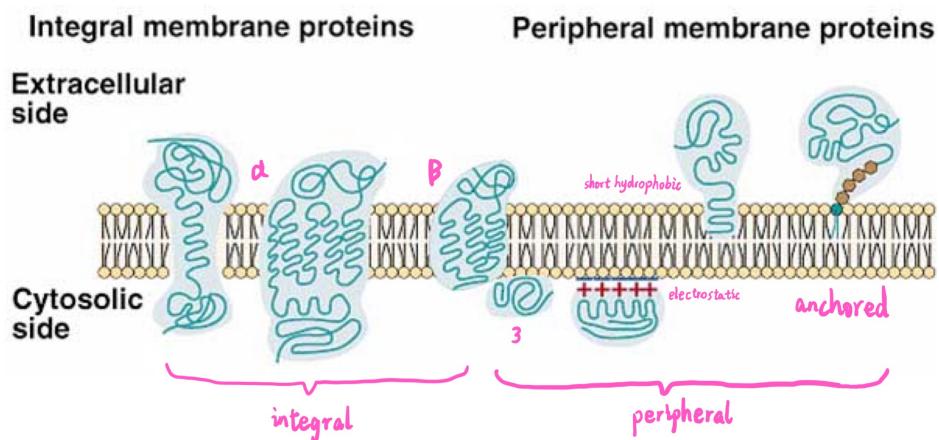
types:

https://en.wikipedia.org/wiki/Membrane_protein for figures

<https://www.expasy.org/resources/protscale> hydrophobicity plot

hydrophobic R chains point outwards (maybe hydrophilic inwards)

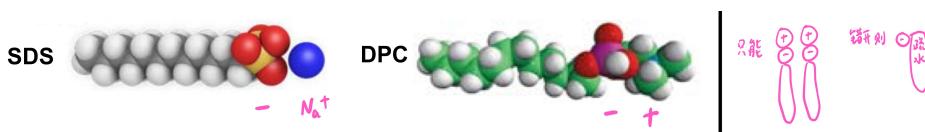
- integral (整合)
 - one or multiple **transmembrane** segments
 - α helix or β barrel
 - some are not trans-membrane
- peripheral (外周)
 - no covalent bonds, only **non-covalent** bonds
 - electrostatically---with polar head
 - terminal hydrophobic group---with bilayer core
 - or bound to an integral protein
 - either outside and inside the membrane
- anchored (锚定)
 - **covalently** bond to the lipids
 - GPI-anchored proteins (G: glycosylphosphatidylinositol-linked)
 - protein---phosphoethanolamine---tetrasaccharide---inositol---
 - where lipase C functions
 - phosphate---diacylglycerol, which is a part of bilayer
 - function: enzyme, antigen, adhesion



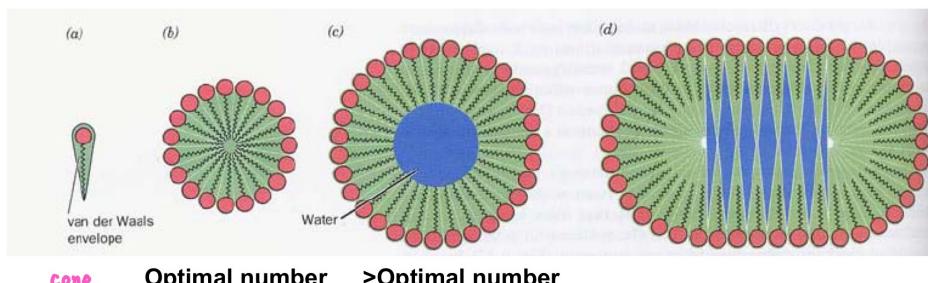
micelles

https://en.wikibooks.org/wiki/Structural_Biochemistry/Lipids/Micelles

- as the concentration of detergent/lipid molecules rises (over Critical Micelle Concentration, CMC), they are no more a layer on the surface, but forms micelles (胶束).
- structure
 - SDS can be seen as a cone (圆锥) model
 - which is due to electrostatic repulsion on the heads
 - driven by hydrophobic interaction, they form a sphere
 - when concentration get higher, the sphere get bigger and water enter inside, thus hydrophobic interaction is disrupted.
 - DPC is a zwitterion but has to be parallel (repulsion direction) to avoid contacting hydrophobic groups.
 - triglyceride may form a cylinder, which favours the formation of bilayer.



structure of molecules



structure of SDS micelles

lipid mobility

- the more lipids are ordered/the closer lipids are placed, fluidity ↓, melting point ↑
- Bilayer has two phases. gel phase: solid; fluid phase: liquid.
- saturated FAs adopt all-trans to achieve the closest contact, but all bonds are freely rotatable, especially those which are near the center of bilayer.
 - gauche--trans--gauche makes a kink
 - cis-double bond makes a bend

which reduces the packing density

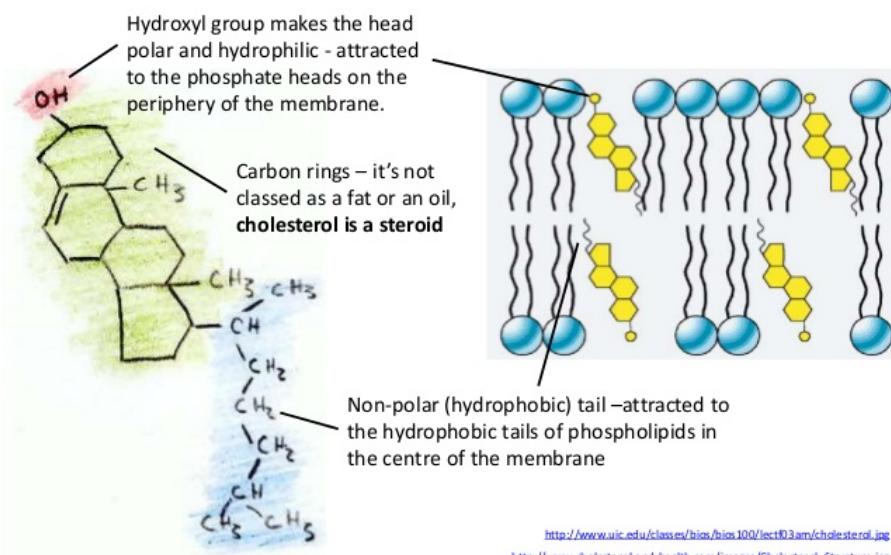


- cholesterol
 - block motions
 - disrupt ordered structure

which makes a balance, modulating the fluidity.

1.3.U3 Cholesterol is a component of animal cell membranes.

Cholesterol



High conc cholesterol may abolish phase transition, keeping the membrane of eg. heart cells in cold blood animals functioning at a low tempearture.

- other motions
 - lateral (侧向): much faster in fluid phase
 - flip-flop: very slow

Week4-1 Membrane Equilibrium

terms

- extensive property: depend on the size or amount. additive
 - intensive property: e.g. density.
- for intensive properties, define partial molar quantity: $\bar{Y}_i = \left(\frac{\partial Y}{\partial n_i} \right)_{T,P,n_j}$
 - then we can apply the additive rule.
- chemical potential: PMQ of Gibbs free energy

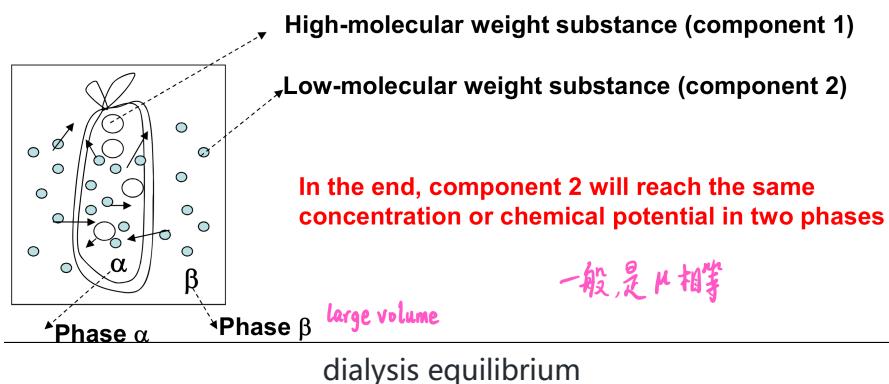
$$\mu_i = \mu_i^0 + RT \ln a_i$$

chemical potential equilibrium

- for an open system (where n may vary but μ doesn't vary under constant T, P)

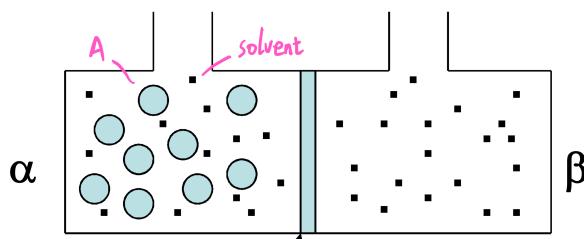
$$dG = \sum_i \mu_i dn_i$$

- for a multiphase system in equilibrium, which is separated by semi-permeable membrane
 - for each component μ_i should be the same and constant in all systems
 - only when this component is permeable. ($dn_i \neq 0$)
 - when not in equilibrium, μ may change first. But we only calculate about the equilibrium state



- when equilibrium is broken

If only one side (α) has solute A of molar concentration C that cannot pass through semi-permeable membrane. On the other side (β) is pure water with the same T and P (pressure of atmosphere on the liquid).



- equilibrium can never be reached. **Osmotic pressure** is generated.
- Imagine we change the pressure to achieve equilibrium, with approximation, we can get

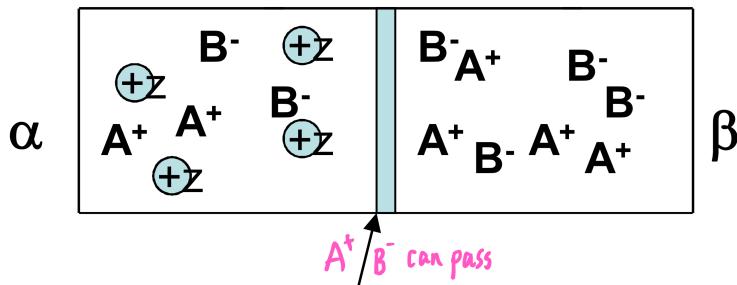
$$\pi = P^\alpha - P^\beta = RTC$$

- the solute can be either macromolecules or small molecules

Donnan effect

the ion may start transferring, but have to keep electrical neutrality.

Donnan effect: when one side contains impermeable charged molecule, concentrations of ions A,B in different side are different.



Write the chemical potential

$$\mu_1 = \mu_A^0 + \mu_B^0 + RT \ln C_A^\alpha + RT \ln C_B^\alpha$$

$$\mu_2 = \mu_A^0 + \mu_B^0 + RT \ln C_A^\beta + RT \ln C_B^\beta$$

At equilibrium, $\mu_1 = \mu_2$

$$C_A^\alpha \cdot C_B^\alpha = C_A^\beta \cdot C_B^\beta$$

Note z is the charge and C_M is concentration of the macromolecule.

According to electrical neutrality, let

$$C_A^\alpha + zC_M = C_B^\alpha = a$$

$$C_A^\beta = C_B^\beta = b$$

$$r = \frac{zC_M}{2C_A^\beta}$$

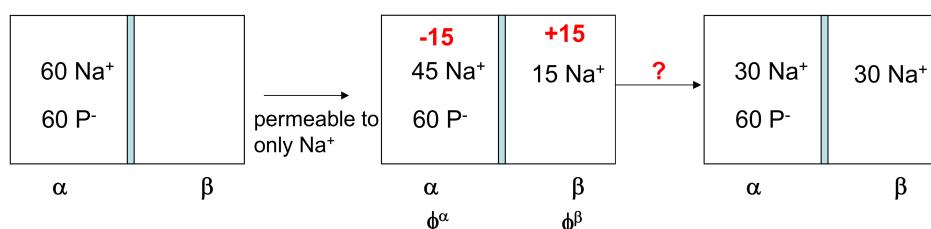
then solve it

$$r_D = \frac{C_A^\beta}{C_A^\alpha} = \frac{C_B^\alpha}{C_B^\beta} = r + (r^2 + 1)^{0.5}$$

If A is H^+ , then $pH^\alpha - pH^\beta = \log r_D$

combined effect: membrane potential

Consider this situation. Na^+ won't pass through because of the electric potential. Chemical potential equilibrium can never be reached, as in the example about osmotic pressure. This system looks like the cell where no ions can freely pass through.



Part two of this section refers to the effect of different concentration, while part three refers to the effect of charged molecules.

Define electrochemical potential as their combination

$$\mu' = \mu + Zf\phi$$

where Z is charge, f is Faraday constant, ϕ is the relative electric potential.

At equilibrium, for single ion A on one side:

$$\Delta\varphi = \varphi^{in} - \varphi^{out} = -\frac{RT}{Zf} \ln \frac{c^{in}}{c^{out}}$$

Ionic concentration difference and membrane potential are generated simultaneously. Maintaining a concentration difference, we get membrane potential. Still, those which cannot pass through the membrane don't contribute.

For multiple ions ([reference](#)):

$$\Delta\varphi = \varphi^{in} - \varphi^{out} = -\frac{RT}{Zf} \ln \frac{\sum_{positive} p_i c_i^{in} + \sum_{negative} p_j c_j^{out}}{\sum_{positive} p_i c_i^{out} + \sum_{negative} p_j c_j^{in}}$$

where p_i is permeability. The cell mainly regulates permeability and sometimes concentration.

summary:

By putting charged macromolecules and selectively letting ions pass through, the cell sets up the difference of concentration of ions as well as the electrical potential, acting as the foundation of activities like neural signaling.

in addition

inference of osmotic pressure

$$\begin{aligned}\mu_1^\alpha &= \mu_1^0(T, P^\alpha) + RT \ln c_1^\alpha y_1^\alpha \\ \mu_2^\beta &= \mu_2^0(T, P^\beta) + RT \ln c_2^\beta y_2^\beta\end{aligned}$$

where c is the molar concentration of water. Here it's approximately the molar ratio (?)

Due to $G = U + PV - TS$, $\mu = V_0 P$, V_0 is molar volume of water.

When equilibrium, $\mu_1^\alpha = \mu_2^\beta$. Assume $y = 1$ So

$$\mu_1^0(T, P^\alpha) - \mu_2^0(T, P^\beta) = V_0(P^\alpha - P^\beta) = -RT \ln \frac{c_1^\alpha}{c_2^\beta}$$

Week4-2 HW

1. Why do most transmembrane helices contain more than the minimum number of residues?

Prof: Usually helices form other H bonds with water except for the C=O...H–N. But now the terminal few residues have to protrude out of the bilayer and form H bond with water to stabilize if the head group can't.

other reasons might be:

- the helices are not perpendicular to the bilayer...
- a few residues are used to anchor in the bilayer...

2. function of cholesterol:

- the formation of lipid rafts
- modulate the fluidity

3. calculation

try to calculate transfer

Week5-1 Membrane Transport

ionic composition

	main cation	main anion
intracellular	K	organic phosphates and acids
intercellular	Na	Cl
blood plasma	Na	Cl

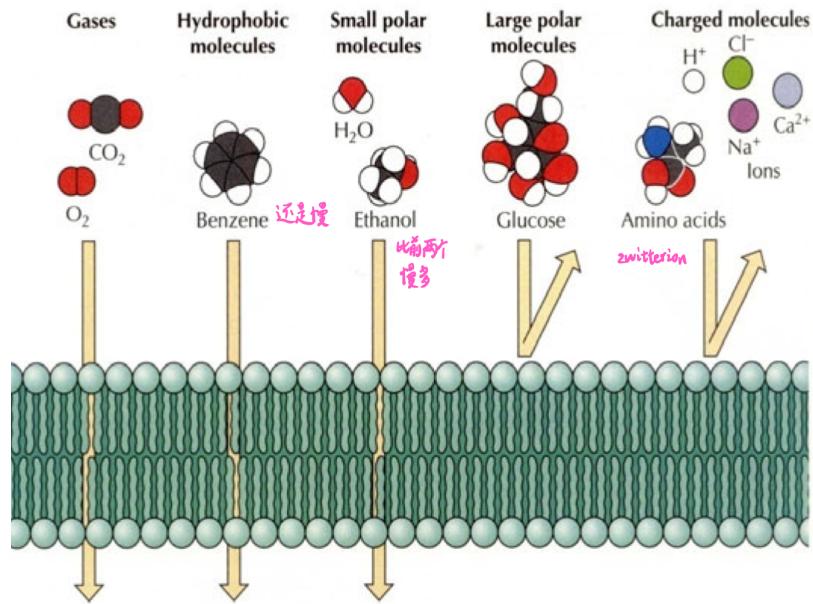
They are related to diseases. [reference here](#)

Principle

- mostly decided by its **soluability in oil**
- also highly depends on the size

diffusion through a protein-free lipid bilayer

reference here



- can diffuse:
 - gas: freely pass
 - hydrophobic: able to diffuse
 - small polar: much slower
- cannot diffuse: ions (charged) and large polar molecules (like zwitterions)
 - ions: hydrated, making it less soluble
 - also: may be attracted by the head group

classification

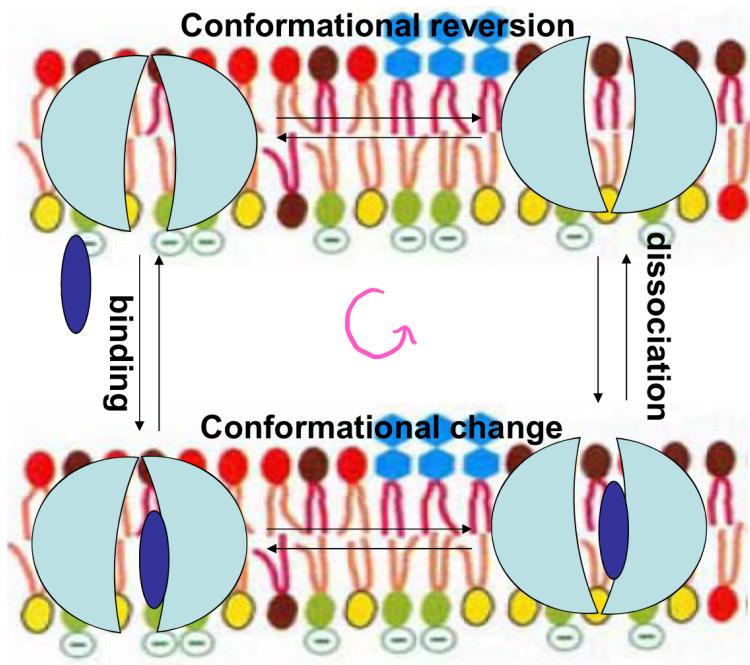
- mediated and non-mediated (gas, steriods...)
- passive and active

mediated transport

ionophores (离子载体): small molecules dissolved in the bilayer

carrier

action mechanisms

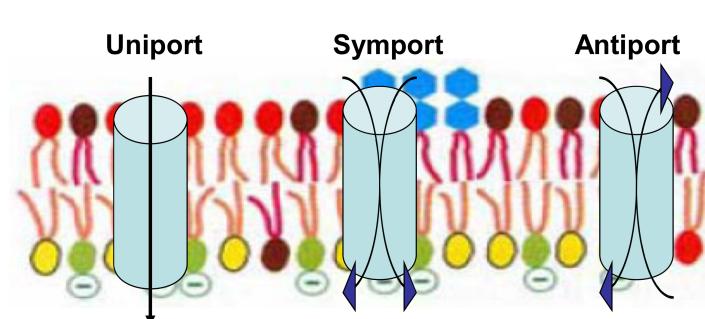
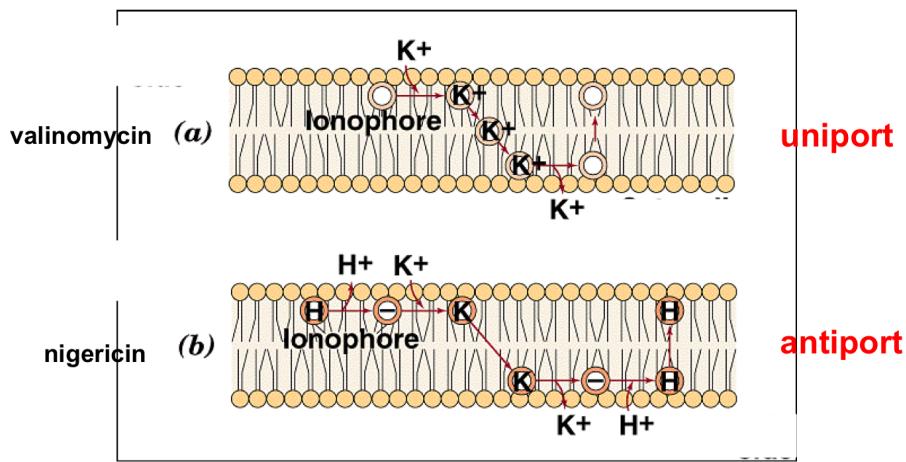


- reversible conformation change.
- the direction is determined by the direction of gradient
- **must just fit the site!**
- may need energy (light; ATP; other "downhill" potential)

classes

- uniport: only one, driven by gradient
- symport: one substrate driven by the gradient of another
- antiport: another substrate in reverse direction

Proposed action mechanism for valinomycin & nigericin-mediated transport of K⁺



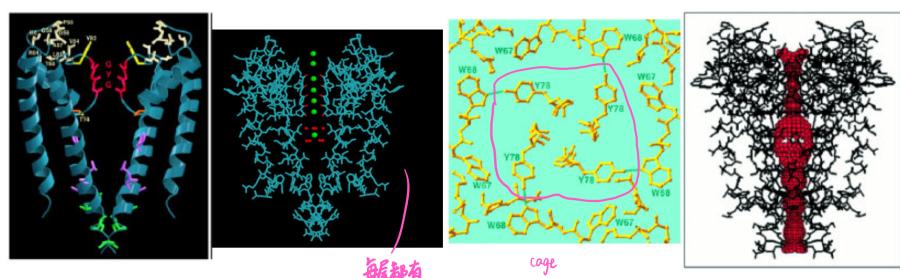
examples

- valinomycin (缬氨霉素) carrying K^+
 - takes the ion and move to the other side (size-specific)
 - can only bind to K^+ rather than Na^+
- lactose permease

channel

examples

- gramicidin (短杆菌肽): its dimer helices allows K^+ and Na^+ to pass through
- K^+ channel:



- acidic residues: attracting K^+ and exclude anions
- $C=O$ groups: coordinate with K^+ and replace its water molecules (cannot for Na^+)
- the first one pushes the second out
- aquaporins
 - about 2.8 angstroms the narrowest, only allowing water molecules to passage

Spectroscopic Tools

categories

- electromagnetic wave
- scatter
 - static
 - dynamic
- absorption
 - UV/Vis
 - circular dichroism
 - IR
- fluorescence
 - excitation
 - emission
 - anisotropy
- nucleus
 - NMR

Week7-1 Circular Dichroism

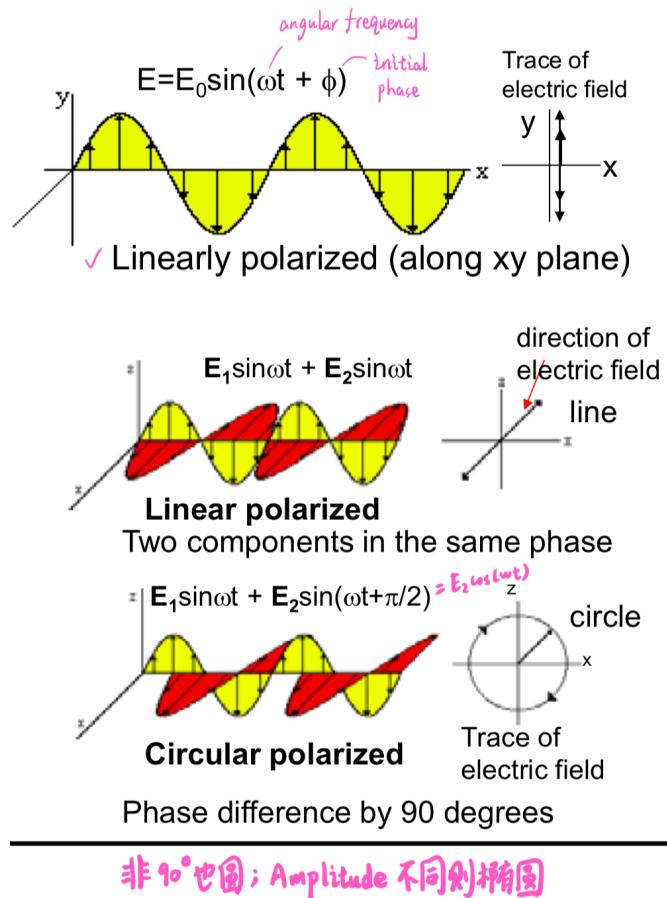
- Circular Dichroism (CD) spectroscopy - YouTube

Light and Polarization

is a kind of electromagnetic wave with mutually perpendicular oscillating electric and magnetic field.

Natural light is un polarized, i.e. oscillating at all directions.

Two linearly polarized lights combine into linear/circular/elliptical polarized light.



Principle of CD

Linear polarized light can be seen as a combination of two circular polarized lights with different direction but same amplitude and initial phase.

Optically active substances (containing chiral center) has different absorbance to right/left-handed light (change A and ϕ). The emission light would be elliptically polarized light.

Ellipticity (椭圆度)

$$\theta = \arctan \frac{A_{min}}{A_{max}}$$

where A_{min} and A_{max} is the minimum and maximum amplitude of **the difference** (also elliptical) of two emission lights. Unit is degree.

Molar Ellipticity (deg · cm²/dmol)

$$[\theta] = 100\theta/cL$$

where c (mol/L) is concentration and L (cm) is light path length.

notes on experiments:

- usually for 0.1 mM DNA/protein in 1cm cuvette, θ is 0.01~0.1 degree. usually use 0.2~0.5 mg/mL protein
- volume: 20 μ L~1 mL; buffer: 10mM phosphate (lower interference)
- Cl⁻ has strong absorbance at wavelength < 200 nm

features:

- CD occurs only when $\epsilon \neq 0$ (molar extinction coefficient).
- CD can be either positive or negative
- More averaged structure (AGCT; denatured) causes less intense signal; more asymmetry more intense

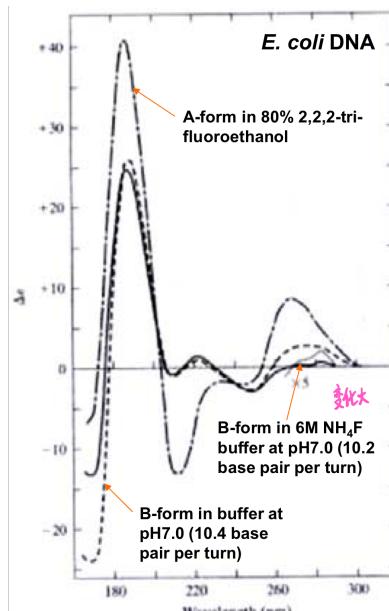
Application

- determine protein/DNA overall secondary structure, study their conformational change as well as kinetics
- study the effect of ligand binding, mutation of protein. study the folding and denaturation

CD of DNA

features

- high salt B form: almost 0 at 280 nm
- no water A form: intense positive at 190 nm



study conformational change: change temperature or water/organic solvent, plot spectrums together.

DNA SS is not so interesting as that of RNA/protein.

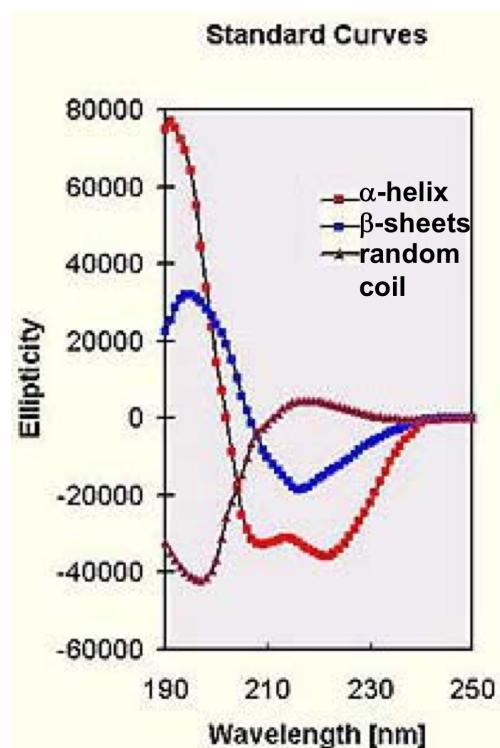
CD of protein

regions

- far-UV: 180~250 nm, 2° structures
- near-UV: 250(230?)~300 nm, 3° structure, if Phe/Tyr/Trp is in asymmetric environment (often folded structure)

standard curve

eg. poly-Lys in different environment



SS type	maximum	minimum
α-helix	193 nm	222,208 nm
β-sheets	198 nm	213 nm
random	almost 0 > 220	negative 190~200

application in protein study

determine SS fraction

For any protein, use

$$\Theta = f_\alpha \Theta_\alpha + f_\beta \Theta_\beta + f_u \Theta_u$$

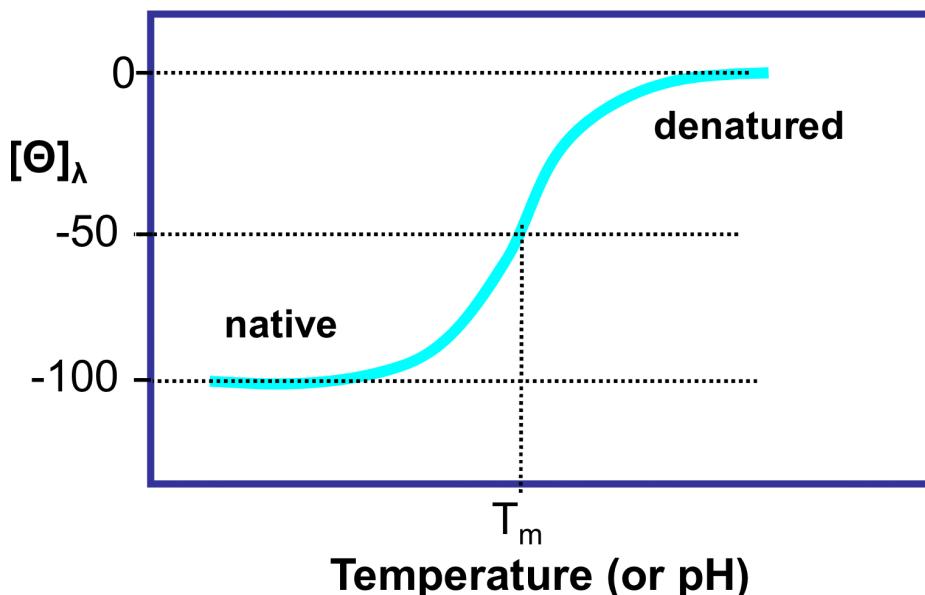
to fit the $\Theta - \lambda$ curve to find the parameters, i.e. the content of all SS.

- results are good for α -helix and tolerable for the other two (only qualitatively)
- may also use combination of other proteins with known content as basis (extract standard curve from them)

measure structural properties under various conditions

- change temperature/ligand concentration/pH/organic solvent/substrate
- record the change of a typical wavelength
- fit the curve to determine T_m , K_d , etc.

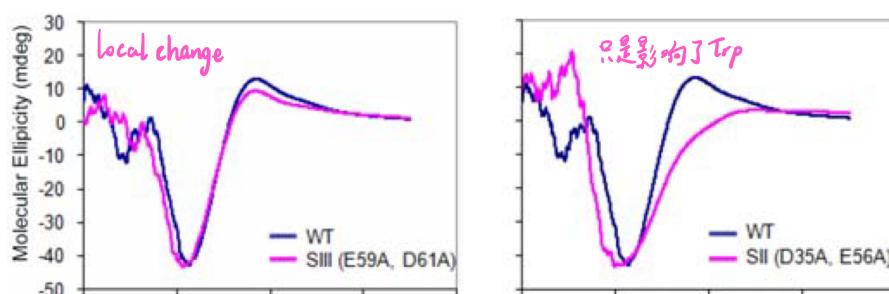
- usually the one where the biggest difference between native and final state occurred.
- For α -helix, can pick 222 nm, often avoid 200 nm



determine the effect of mutation

just compare the CD of WT and mutant

- if there's small change, but protein's function is affected, then this residue is responsible for the function
- if there's large change, then this residue affects the overall structure. We know it's critical for the structure but cannot determine if the residue is really responsible for the function (or at active site)



study protein folding dynamics/kinetics

mix protein and denaturant, CD spectrum is recorded quickly. can get a $[\theta]$ -time curve (at picked wavelength). may fit with exponential equation

summary of plots

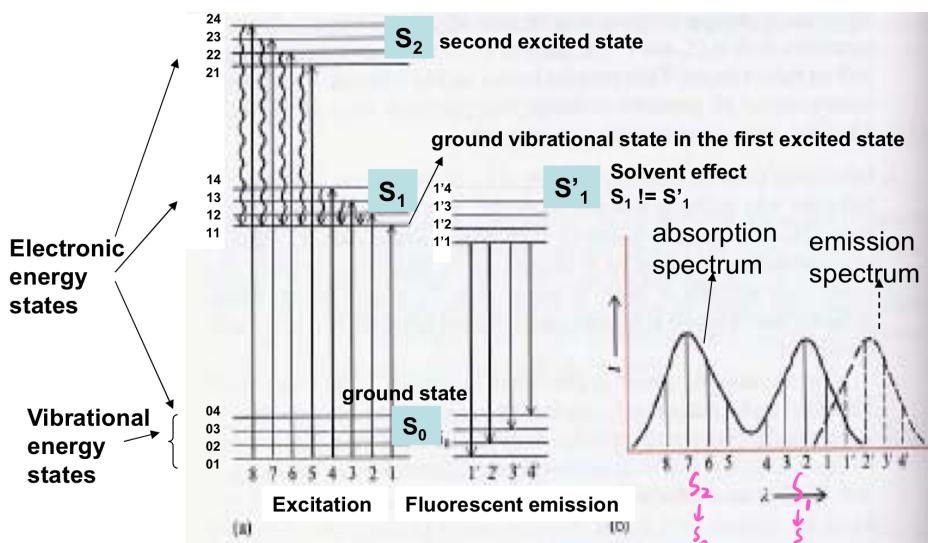
- CD spectra ($\theta-\lambda$), SS and mutant
- $[\theta]$ -experiment condition
- $[\theta]$ -time

Week7-2 Fluorescence Spectroscopy

- [Fluorescence Polarization - YouTube](#)
- [Fluorescence Spectroscopy Tutorial - Typical Applications - YouTube](#)
- slides in USTC

Principle

generation



- electron has multiple energy levels
 - only when the input energy equals energy difference can an electron be excited
- vibration makes each energy state split
 - at equilibrium distance the energy is lowest
 - there is zero point energy
- fluorescence
 - only consider singlet states
 - range of spectrums

- emission: from ground vibrational state in the first excited state (actually S_1') to (any) ground state
- absorption: from ground vibrational state in the ground state to any higher electronic energy state
 - thus absorption energy might be higher but little difference
- features of the spectra
 - Stokes shift
 - energy of the first excited state might be lowered because of **the solvent**, etc., which makes the difference in absorption and emission spectrum.
 - shape of the emission spectrum is independent of absorption
 - they may intercept a little
 - and form kind of mirror symmetry
 - but the intensity is corresponding to the absorption
 - we'd choose the most absorbed wave length (intrinsic property)

properties of fluorescence

intensity is dependent on

- molar extinction coefficient (摩尔消光系数)
- fluorescence quantum yield
- concentration
- optical pathlength
- collect efficiency

considerations

- usually use relative intensity
- absorbance might turn nonlinear when concentration increases

fluorophores

- intrinsic: Trp >> Tyr
- nucleic acids' fluorescence is so weak
- small molecule to help: ANS, fluorescein, EB

characteristics: conjugated C=C, complex rings

fluorescence polarization

Situation: adding probes into the macromolecule. Probes:

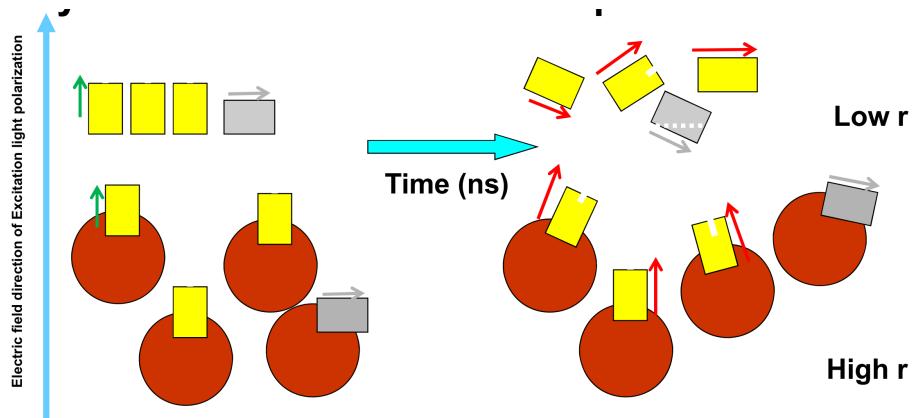
- without target or in aqueous solution, their fluorescence is quenched
- their fluorescence is enhanced in non-polar or rigid environment (eg DNA)

Use horizontally polarized light to excite, and detect horizontal and vertical signals (all are perpendicular to the exciting light)

$$r = \frac{I_H - I_V}{I_H + 2I_V}$$

This r actually measures **anisotropy** of the system (containing ligand).

- only molecules on certain orientation will be excited.
- at excited state (~nanoseconds) the molecules move randomly, thus the emission light is on all directions ($I_H \approx I_V$), making r low (for single small molecule dye)
- the rotation is slowed (large r) if
 - the probe binds to a big molecule
 - solution with high viscosity



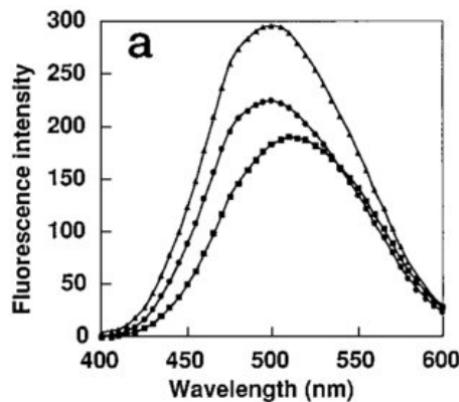
Application examples

In summary, the key points are

- r value is affected by molecular weight
- from hydrophobic to exposed environment, red shift occurs and intensity probably decreases
- figures
 - intensity/ r – λ curve
 - intensity/ r –condition (mole,...but **usually no temperature**)

protein-protein/peptide/ligand

- One of the two has a fluorophore and the binding will effect that site.
- $r - \lambda$ curve
 - bind to **cover Trp** into hydrophobic site → blue shift
 - hard to predict intensity



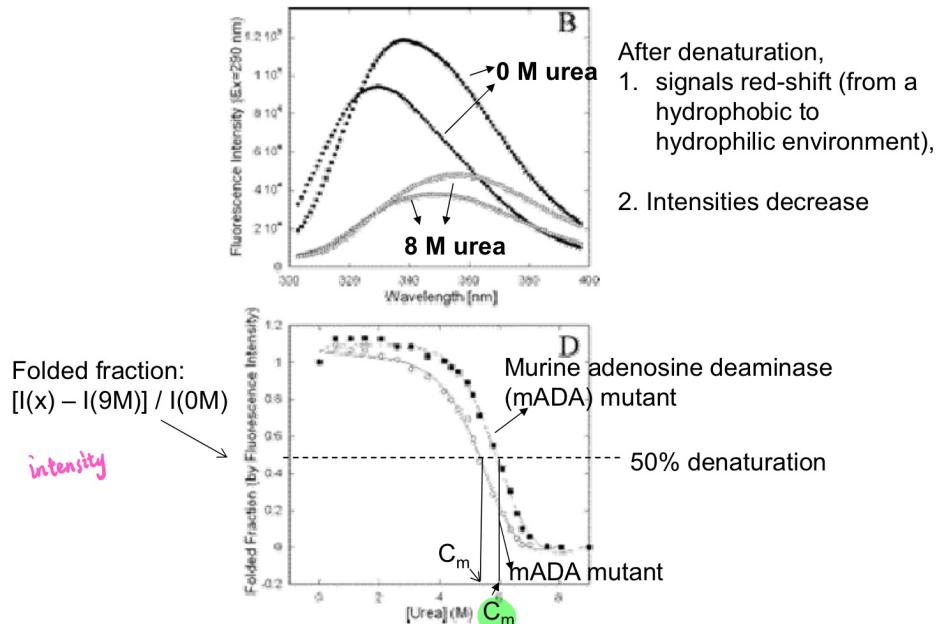
- r —molar ratio
 - is binding? see shape
 - how many? see ratio (there must be apparent mw change)
 - binding mode (close to Trp)? see r
 - affinity??

conformational change

eg: add urea to protein (with Trp!)

- slight red shift
- intensity significantly reduced

make a plot of fold fraction (calculated from intensity) and urea concentration to study denaturation/folding/Kd...



introduce to living cell

Design these proteins to detect the content of small molecules, like FICRrh/cAMP.

As a technique to study the change/visualize processes ...

Fluorescence Resonance Energy Transfer

principle

FRET is a kind of dipole-dipole interaction between excited states. The donor is excited (eg by light) and then excites the acceptor. eg: CFP and YFP

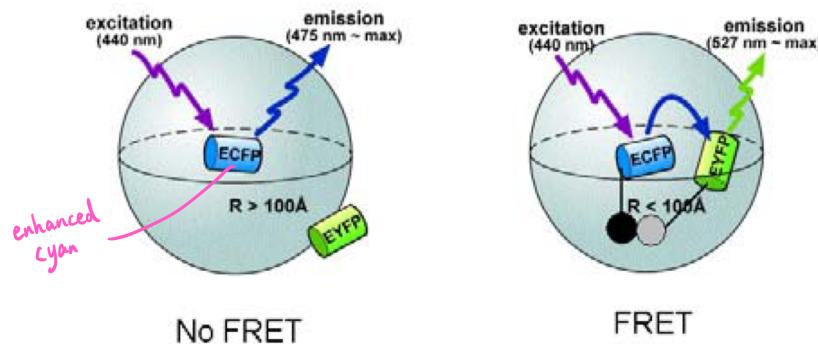
The efficiency of energy transfer is $\frac{R_0^6}{R_0^6 + r^6}$, which decreases quickly through distance.

Primary conditions:

- donor and acceptor are close contact ($10\text{~}\sim\text{~}100\text{ \AA}$)
- the absorption spectrum of acceptor **overlaps** with the emission spectrum of the donor (but separable).
- their dipole orientation must be approximately parallel

application

Strength: just fusion expression of a fluorescent protein, easier than chemical dye in processing and maintaining biological activity (stability).



Can quantify:

- Protein-Protein/DNA Interaction
 - if they can interact, you'll see the emission spectrum of the acceptor (predominant) and donor, instead of only the donor's.
 - *in vivo* or *vitro*
- protein conformational changes
 - in a specifically designed conformation, the donor and acceptor can interact; condition changes, spectrum changes
 - maybe observe signal to change with time, if slow

<https://www.biomart.cn/mobile/experiment/article/25123.htm>

Quencher

Molecules which can interact with others (fluorophore) and reduce the intensity of fluorescence.

eg: small molecules as quenchers. like oxygen, acrylamide, $\text{Cs}^+ \cdot 4 \text{H}_2\text{O}$

Put (different concentration of) quencher together with fluorescent protein (marcomolecule) to measure the **accessibility** for the quencher into the aromatic residues. Just measure F_0/F_c , the intensity before and after adding quencher.

Week8-2 NMR

- [Relaxation \(NMR\) - Wikipedia](#)
- [How MRI Works - Part 1 - NMR Basics - YouTube](#) detailed
 - [NMR spectroscopy visualized - YouTube](#)
 - [eLearning - NMR Relaxation - YouTube](#)

Principles

- static magnetic field: make the energy difference of nucleus
 - spin energy levels split, because spin has direction
- inducing electromagnetic wave: to excite the nucleus when the energy matches

$$\Delta E = h\nu = h\gamma B_0$$

- ν is the frequency of inducing electromagnetic wave
- γ is only related to atom type
- B_0 is the strength of static megnetic field
- the absorption and emission energy is the same, so we often measure the latter
- for ${}^1\text{H}$, $42.58 \text{ MHz} \sim 1 \text{ Tesla}$

Phenomenons

Chemical shift

Definition

ring current of electron → induced magnetic field → B on the nucleus changes

$$B_{eff} = B_0(1 - \sigma)$$

so

$$\nu = \gamma B_{eff}$$
$$\delta = \frac{\nu - \nu_{ref}}{\nu_{ref}} \times 10^6$$

ν_{ref} is from a standard substance.

Factors influencing δ

The induced magnetic field is related to

- Electron density: more electron (more shield), higher current, strong field, lower δ
- the direction (spatial orientation) of other atoms

so the factors include:

1. inductive effect (electronegativity)

connecting to such an atom lowers the electron density (OH, C=O, C⁺ ...)

thus farther from TMS i.e. large δ

roughly: $\delta \propto$ Electronegativity of connected atom

2. hybridization effect (anisotropy)

usually vinyl Hs have big δ

for biomolecules, those on the aryl (F, W, Y)

3. other effects (structure)

- H bond
- dihedral angle, secondary structure
- temperature
 - random coil is more variable
- solvent

J coupling (spin-spin coupling)

For two protons on two connected C, when one of them is on ground state/excited spin state, it generates different induced magnetic field (only differ in directions). It functions on the other and makes the ΔE change. The magnetic field in one direction strengthens B_0 while the other weakens, so ΔE splits into two whose energy difference is J . This effect is mutual.

- often < distance of 3 bonds.
 - transmitted via electrons in the intervening bonds
- NMR intensity: number of Hs; split: number of adjacent Hs
 - just add up the effect of all Hs, 2 choices each
 - get probability distribution/strength: Pascal triangle

Relaxation

Relaxation time T_2 describes how fast the NMR signal decays in time-domain spectroscopy.

In frequency-domain spectroscopy, full width at half maximum (FWHM, linewidth) is π/T_2 (The AUC is constant). Broad peak/large FWHM means

- small T_2 /slower decay, lower height
- molecule size is big, or this group is rigid (like -Ph)

Labile hydrogens

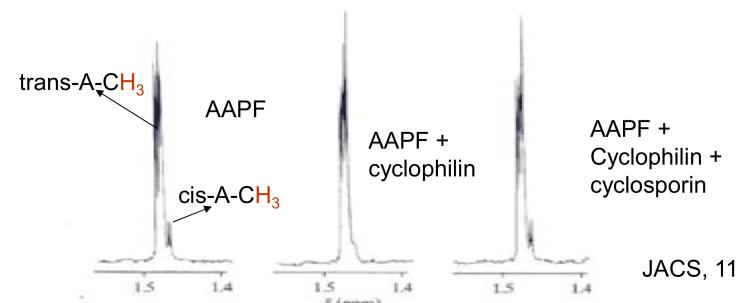
H on O,N,-COOH are shielded by hydrogen-deuterium exchange (add D₂O)

Applications

only 1D ¹H NMR

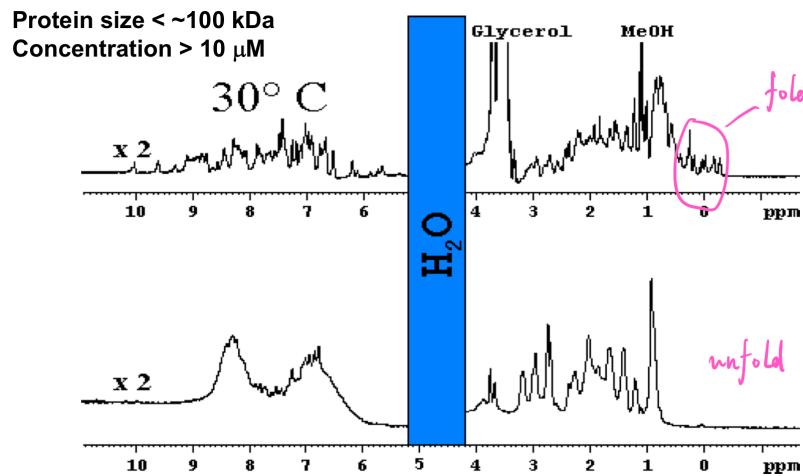
identify inhibitors

eg: Cyclophilin catalyzes isomerization of Pro in the substrate polypeptide from cis to trans. cis exists in single substrate, disappears when adding Cyclophilin, reappears when further adding its inhibitor.



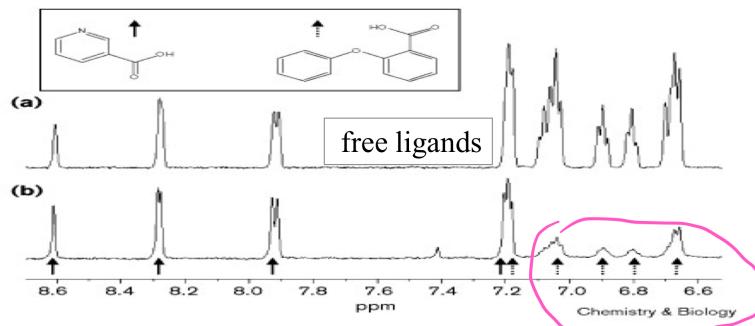
study protein folding

- signals of unfolded protein are similar and overlap
- that of folded protein have dispersion, and may have signals <0.5 (alkyl) or >9 (peptide bond) ppm.



drug discovery

Compare signals of small molecules. The signals may look like the target after binding. eg: They move together, so peaks of ligands become broader



In the presence of P38 MAP Kinase

metabonomics

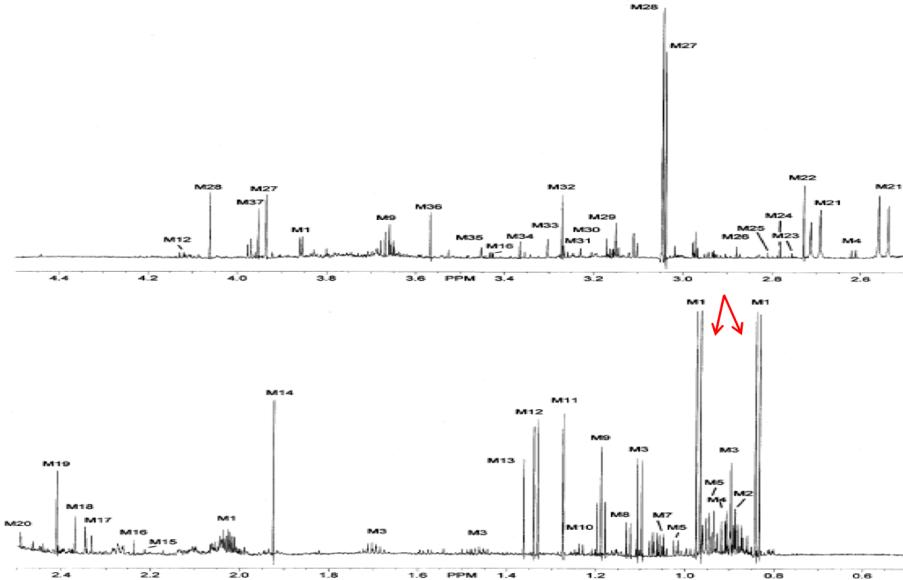
First we'd know the characteristic peak of compounds.

biomarker identification

Compare the spectra of samples (like urine) from patients and healthy ones.
Find the abnormal metabolite corresponding to abnormal signals.

toxicology

Compare the spectra of samples from rat treated with a drug/toxin and control. Identify which is up/downregulated.



750 MHz 1H NMR spectrum of the urine from a MSUD patient
M1, M11: 2-hydroxy-isovalerate (sweet smell), M13: 3-amino-

Summary on biophysical techniques

- CD tells about SS, but don't know if it's properly folded
- fluorescence/NMR can tell if it's folded, but cannot tell exactly what SS is

having α helix does not mean it's properly folded/packed!

Week10-1 Conformational Transition

Fundamental Representations

For the transition of a macromolecule in 2 conformations $A \rightleftharpoons B$, the equilibrium constant is $K = \frac{[B]}{[A]}$, the probability of observing B is

$$P_B = \frac{[B]}{[A] + [B]} = \frac{K}{1 + K}$$

Assume there are many states ($i = 1, 2, \dots$) in equilibrium, and a supposed reference state ($i = 0$, only for the sake of mathematics), define **weight** as the equilibrium constant between states (S):

$$\omega_i = \frac{[S_i]}{[S_0]} = \exp(-\Delta G_{i,0}/RT)$$

so ω_i is only related to energy level. We combine those in the same energy level, the number of states is called **degeneracy** g_i . So we redefine: i does not refer to a single state but the combination of states in the same energy level, so

$$\omega_i = g_i \exp(-\Delta G_{i,0}/RT)$$

We also define **partition function** as

$$Q = \sum_{i=1}^{N_{tot}} \omega_i$$

and the probability will be

$$P_i = \frac{\omega_i}{Q}$$

Sum of concentrations on this energy level divided by the total concentration. All divided by $[S_0]$ and sum up states in the same energy level to get ω_i .

Conformational Transition Models

Notation:

- For a DNA/protein of length N , each unit has a state (helix or coil, etc.), represented as a,b.
- Suppose ΔG of changing a single unit from a to b is s .
 - so $\omega_i = g_i s^i$
- Probability of the state with i units in b state is named P_i .
- Fraction of a,b in all units of all states is P_a, P_b, \dots
 - if define fraction of b in state i as $f_i = \frac{i}{N}$,
 - then $P_b = \frac{f_i \omega_i}{Q}$

all-or-none (highly-cooperative) model

Only two states: all is a, and all is b.

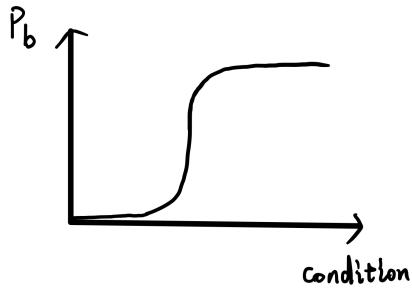
$$K = s^N$$

$$P_0 = P_a = \frac{1}{1 + s^N}$$

$$P_N = P_b = \frac{s^N}{1 + s^N}$$

Interpretation:

- If $s = 1$ (no energy difference), $P_0 = P_N = 1/2$
- When N is large,
 - as condition (pH, temp) change, s changes
 - as long as $s > 1$, P_b is close to 1; if $s < 1$, P_b is close to 0
 - which means P_b will not stay at the middle but either close 0 or 1
 - the graph looks like \downarrow , where it seems the two states change suddenly



non-cooperative model

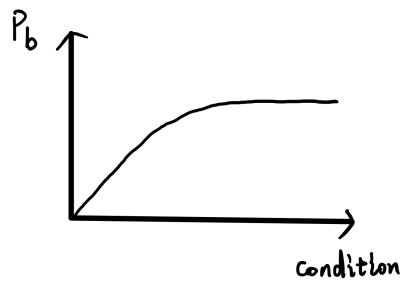
All units change independently. So there will be $N + 1$ states and

$$g_i = \binom{i}{N} = \frac{N!}{i!(N-i)!}$$

$$Q = \sum \binom{i}{N} s^i = (1+s)^N$$

Interpretation

- if $s = 1$, P_i will be a binomial distribution; otherwise just modified by s^N
- $P_b = \dots = \frac{s}{s+1}$, and this might be a long dependence



- $P_b/P_a = s$ which is independent of N . The fraction of b in all states is the same as single a and b.

Zipper Model

A classical model to describe realistic transitions.

- Assume it's impossible for two sites to transition simultaneously.
- Assume the first step is more difficult (**nucleation**, the one to break original complete structure, see examples later), which needs extra energy.

$$\sigma = \Delta G_{extra} > 0$$
$$\omega_1 = \exp[-(s + \sigma)/RT]$$

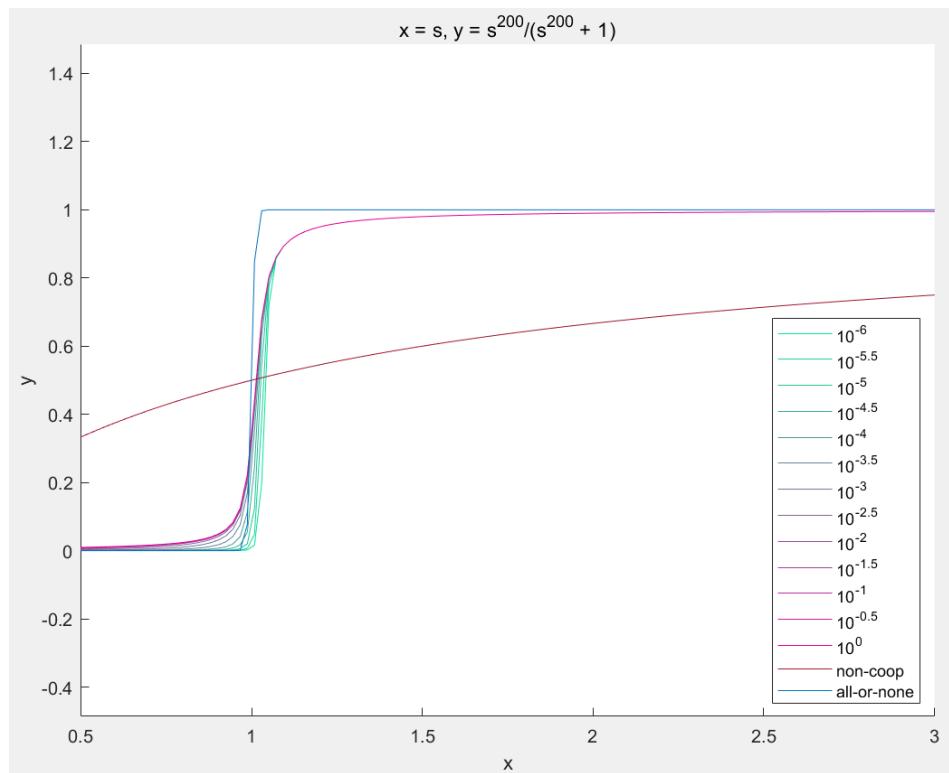
- Also, only those next to the nucleation center will perform transition.
State i will have continuous b units of number i .

So

$$g_i = \begin{cases} 1 & i = 0 \\ N - i + 1 & i \geq 1 \end{cases}$$
$$\omega_i = \begin{cases} 1 & i = 0 \\ g_i \sigma s^i & i \geq 1 \end{cases}$$
$$P_b = \frac{\sum_{i=1}^N \frac{i}{N} (N - i + 1) \sigma s^i}{1 + \sum_{i=1}^N (N - i + 1) \sigma s^i}$$

Results (try to simulate yourself!):

- it will look like highly-cooperative model if
 - $\sigma \downarrow (\ll s)$, difficult for the first to occur.
 - $N \uparrow$, σ is more distributed on the sequenceso σ is called "cooperativity coefficient"
- (not like in slides) it will not look like non-cooperative model because definition of degeneracy has implied cooperativity.



Applications

in proteins

coil-helix transition

From coil to helix, the first 4 (/3/5) events are difficult. Then just need to move one residue if the neighboring three are ready.

- $s \approx 1$ because enthalpy gain from H bonding is counteracted by entropy loss (helix is more fixed)
 - sequence dependence: AA type and its environment
- $\sigma \approx 2 \times 10^{-4}$ for poly- γ -benzyl-L-Glu
 - Accounting for entropy loss from 3 more residues (N just minus 3). Thus, $\sigma_{3-10} > \sigma_\alpha$ ($i-i+3$)
 - sequence dependence: Lys $\sigma \approx 10^{-3}$

helix melting (the reverse process)

Melting one residue may start from the end or middle. The latter situation will break H bonds at both ends. Maybe break 5 H bonds and gain entropy once? A larger σ ???

Still using the original s which > 1 .

- from the end: $\omega_{N-1} = 2\sigma s^{N-1}$
- from the middle: $\omega_{N-1} = (N-2)\sigma^2 s^{N-1}$
 - the first term is degeneracy. 2 ends, others are middle

- if you consider the reversed numbering, $\omega_1 = (N - 2)\sigma^2 \frac{1}{s}$, divided by s^N

So when the probabilities equal,

$$N = \frac{2}{\sigma} + 2$$

N is independent of s but only decided by σ

Larger N distributes the nucleation loss and more likely to start in the middle.

But more start from the end. eg: poly- γ -benzyl-L-Glu will melt in the middle if $N > 10002\dots$

in DNA

It follows the same equations to determine whether to start from the end or middle. More probable in the middle? $\sigma \approx 10^{-7}\dots$

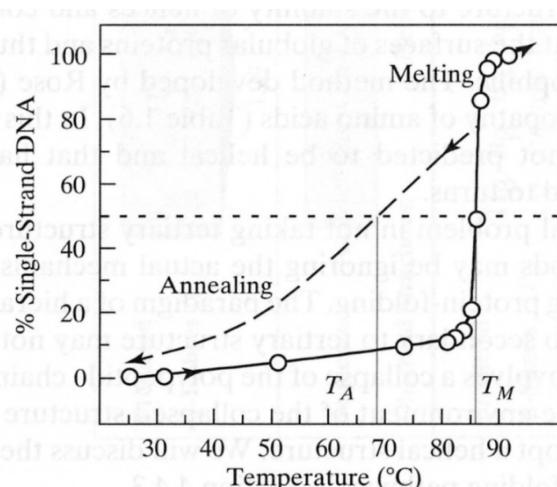
DNA melting

From dsDNA to ssDNA,

- s is composed of energy to break (2/3) H bonds and a pair of base stacking (main force in DNA secondary structure!).
- σ means additional energy to break base stacking on the other side.

DNA annealing

- dependent on the concentration of ssDNA, so $\omega_1 = \kappa\sigma s$
- sequence dependence
- in annealing it needs finding proper combination of base pairs (eg. a primer searches for complementary sequence), so cooperativity is lower than melting. Also $T_m \gg T_a$. I think it's combining $\kappa\sigma$ as σ and this is just amending non-quasi-static kinetic error.



DNA helical transition

In ethanol (remove water), B-DNA \rightarrow A-DNA.

- think of DNA becoming a little bit tighter.
- s : the difference between average energy per base pair in two forms
- σ : A-B junction, it's a little distorted when an A form base connect with a B
 - $5\sim 8 \text{ kJ/mol}$, $\sigma = 0.14 \sim 0.04$

Rise salt conc/alcohol conc/temperature, B-DNA \rightarrow Z-DNA

- σ turns to the energy used to turn right-handed helix into left-handed helix
- σ is larger if conditions suitable for Z-DNA (above) is satisfied.

DNA supercoil state transition

For double-strand circular DNA

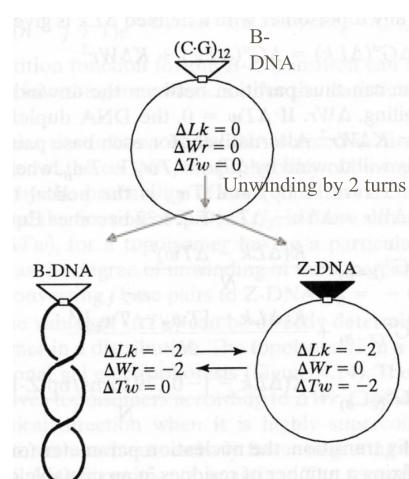
- linking number: how many times one strand comes around the other.
 - Might change only when double strand is broken, like using topoisomerase
- twisting number: how many "turns" the two strands rotate along central axis.
 - total number of base pairs in a DNA molecule divided by the number of bases per turn
- writhing number: the number of supercoiling

$$L = T + W$$

$$\Delta L = \Delta T + \Delta W$$

situations

1. In DNA replication, helicase functions. L remains. T behind decreases, T in front increases.
2. An enzyme unwinds DNA by 2 turns through cutting it off. $\Delta L = -2$
 - maintaining circular shape: $\Delta W = 0$
 - relieve all stress: $\Delta T = 0$



Transition occurs in a local region of alternating C and G

why supercoil?

- save space
- for replication and transcription (Z-DNA)

In addition for DNA melting: how does s change with condition?

- at T_m , melting and renaturation reach a equilibrium, $s_{T_m} = \frac{[\text{melted}]}{[\text{unmelted}]} = 1$
- according to [Van 't Hoff equation](#), $s = -\frac{\Delta H^\Theta}{R} \left(\frac{1}{T} - \frac{1}{T_m} \right)$
- when $T > T_m$, $s > 1$ thus can melt; when $T < T_m$, $s < 1$

Week10-2 Protein Folding

Basics

Why study protein folding?

1. sequence-structure relationship
2. related to diseases
3. biotechniques eg. purify

some facts

- Protein folds in milliseconds
- Protein folding must follow certain pathways
 - instead of randomly search in the huge conformation space.
- Protein refolds automatically upon dilution of denaturant.
- Unfolded states are heterogenous (all kinds) while native states are narrow.
- Typical denaturant: 8M urea, 6M GdnCl

concepts

1. denatured state (D)
 - an ensemble of dynamic inter-converting species
 - with few nonlocal interactions; only retain few local interactions
(interactions between atoms that are close in sequence)
2. intermediate (I)

- occur during refolding of large proteins
- on-pathway intermediate has native-like interactions that makes conformation searching easier
- trapped intermediate slows down folding, which needs chaperons, otherwise aggregates/precipitates (even lower energy than native)

3. transition state (TS)

- highest energy, corresponding to rate-limiting step
- backbone: forms secondary structure
- side chain: specific rotamers, a bridge from disordered conformation to native structure

Folding Models

1. framework model

- quickly form secondary structure first, leading to tertiary structure
- segments of native-like structures in TS, some residues have high ϕ value
- preferred in bigger protein
- subvisions: "diffusion collision" and "propagation"

2. nucleation model

- form a compact nucleus first, then further condensation
- partially formed interactions; residues have equalized ϕ value
- preferred in smaller protein

Equilibrium Studies

principles

states: D, N, Intermediates (know how many!)

choose a property y (eg. Fluo/CD under a certain λ that maximizes the difference).

fit the fraction f of each state.

$$y = f_N y_N + f_D y_D + \sum_i f_i y_i$$

$$f_N + f_D + \sum_i f_i = 1$$

define the normalized property (distance from denatured)

$$d_i = \frac{y_i - y_N}{y_D - y_N} \in [0, 1]$$

and several ratios

$$K_i = \frac{f_i}{f_N}$$

$$K_D = \frac{f_D}{f_N}$$

and apparent (表观) fractional denaturation/native

$$f_{app} = f_D + \sum_i f_i d_i$$

$$1 - f_{app} = f_N + \sum_i f_i (1 - d_i)$$

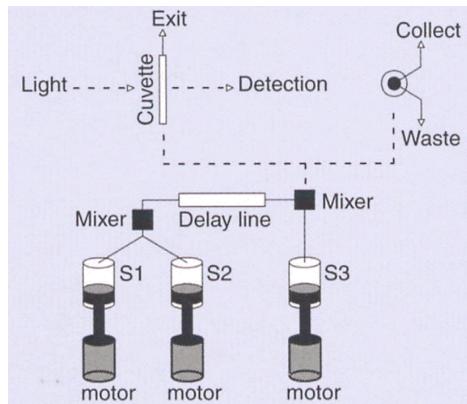
so the apparent equilibrium constants (between apparent denatured and native)

$$K_{app} = \frac{f_{app}}{1 - f_{app}} = \frac{K_D + \sum_i K_i d_i}{1 + \sum_i K_i (1 - d_i)}$$

- if there's no intermediates (2 states), $K_i = 0$
condition: $K_{app} = K_D = -\exp(\Delta G/RT)$
or is constant in different property experiments
- otherwise $K_{app} \neq K_D$, or varies in experiments since d_i varies

experiment and fitting

add denaturant/refolding buffer; get signals under different conditions



denaturant (GdnCl/urea, Fluo/CD)

(Assume $y_D > y_N$) 2 states, $f_D = \frac{[D]}{[D] + [N]} = \frac{K_D}{K_D + 1}$.

For $N \rightleftharpoons D$, ΔG has approximately linear relationship with conc of denaturant c . Thus $K_D = \exp[-(\Delta G_0 + mc)/RT]$, where ΔG_0 is when $c = 0$.

Plug in $y = y_N + (y_D - y_N) \cdot f_D$ (or reverse D/N), get

$$y = y_N + (y_D - y_N) \cdot \frac{\exp[-(\Delta G_0 + mc)/RT]}{\exp[-(\Delta G_0 + mc)/RT] + 1}$$

Fit the y - c curve to get m and ΔG_0 . Either set y_D, y_N as parameters or not.

temperature (CD)

2 states, still $y = y_N + (y_D - y_N) \cdot \frac{K_D}{K_D + 1}$, but K_D is related to T .

Fit curve to get

- ΔH_m : at T_m
- ΔC_p : of reaction, assume as constant
- T_m : D and N reach equilibrium
- initial protein concentration

Kinetics Studies

time-scale experiment

For 2-state model $N \xrightleftharpoons[k_{-1}]{k_1} D$, $y = y_\infty + A \exp(-k_u t)$. Fit $\ln(y - y_\infty) - t$

to get k_u . Reversely get k_f .

for 3-state model $N \xrightleftharpoons[k_{-1}]{k_1} X \xrightleftharpoons[k_{-2}]{k_2} D$,

$y - y_0 = A_1 \exp(-\lambda_1 t) + A_2 \exp(-\lambda_2 t)$. Maybe fit each subsection.

Also, intermediate of folding and refolding may not be the same.

get rate for 2-state model

equations

For $N \rightleftharpoons D$, do unfolding/refolding separately.

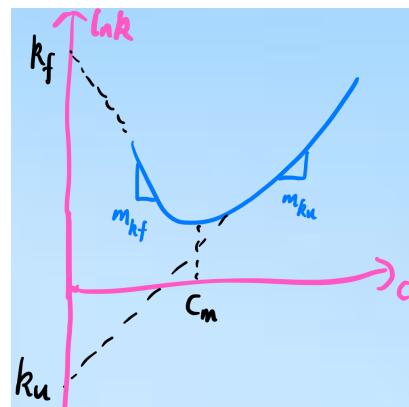
- Add denaturant into folded protein
 - **when conc is high, it's only unfolding**, discard refolding reaction (vice versa)
- Repeat this 2-state experiment with different concentration. Get a $k - c$ curve.
 - Actually regard this as $N + \text{urea} \xrightarrow{k_1} D + \text{urea}$.
- Fit $\ln k_f = \ln k_{H_2O} - m_{k_f} c$.
 - The linearity indicates if it's 2-state.

Do the same reversely and combine the plots, the two arms.

Or just fit all data with this to get Chavron plot (assume folding rate in high concentration is much slower than unfolding and vice versa).

$$k = k_{f,H_2O} \cdot \exp(-m_{k_f} c) + k_{u,H_2O} \cdot \exp(m_{k_u} c)$$

Interpretation



- All parameters > 0 .
 - k_{H_2O} is the rate when $c = 0$
 - m_k is an important factor to indicate how denaturant changes the rates
- The intersect of two lines indicates the concentration when half the protein is unfolded, indicating thermodynamical stability.
 - $c_m = \frac{k_{f,H_2O} - k_{u,H_2O}}{m_{k_f} + m_{k_u}}$
- simple compare: kinetically more stable if
 - m_{k_f} equals and m_{k_u} smaller;
 - m_{k_u} equals and m_{k_f} bigger;
- $\ln k_f$ and ΔG_F are both linear with c , and the coefficient should be the same.
 - But if we use only one m is equilibrium study, i.e treat it as 2-states, add up two ΔG , then the total $m = m_{k_f} + m_{k_u}$.
 - Larger m leads to steeper equilibrium curve (like all-or-none), indicating more stable besides the above two criteria.

Local State Studies

NMR

Collect the signal of all $-NH-$ groups. All residues in HSQC have been assigned.

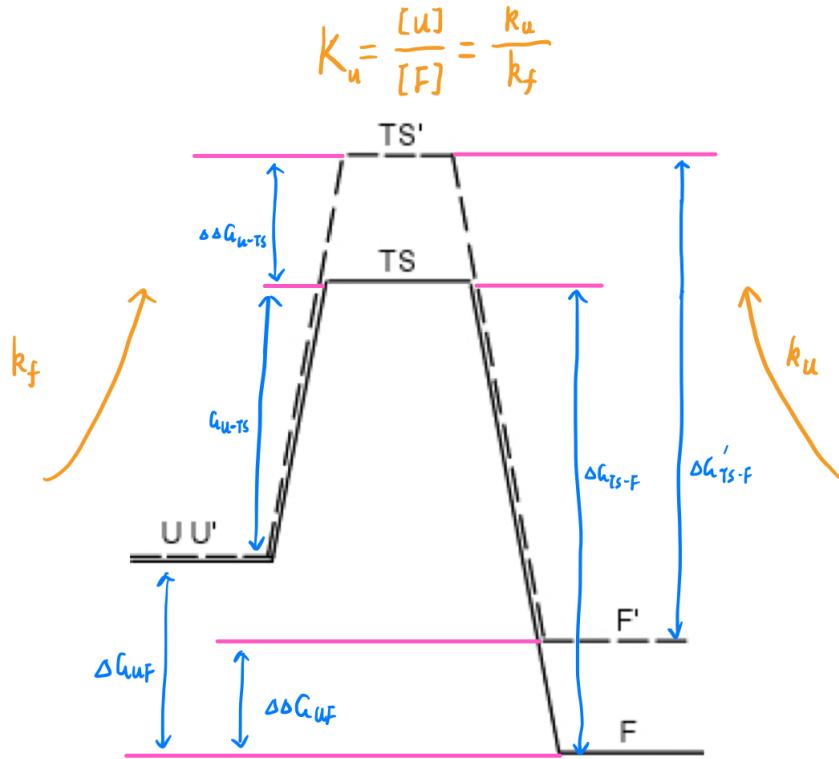
- Compare the 1H chemical shift of (half?) denatured protein and random coil
 - whose chemical shift is lack of dispersion (8~8.5 ppm). If some residues are out of this (7.5~9.5 ppm), it may maintain folded structure.
- Compare the $H-N-C_\alpha-H_\alpha$ J-coupling constant, which is related to ϕ angle.
- Check the NOEs.

???

transition state (ϕ value analysis)

For $N \rightleftharpoons TS \rightleftharpoons D$, compare WT and one-residue mutant.

Experimentally get k_f and k_u (I guess under appropriate denaturant concentration, if we assume m does not change. But they do change...).



Think of activation energy: $k_f = k_{f,0} \exp(-\Delta G_{U-TS}/RT)$. Also for f_u .

Intuition: larger k_f , smaller $|\Delta G_{U-TS}|$.

Strange definition:

$$\Delta\Delta G_{U-TS} = \Delta G_{U-TS} - \Delta G_{U-TS}^* = -\ln \frac{k_f}{k_f^*}$$

$$\Delta\Delta G_{TS-N} = \Delta G_{TS-N}^* - \Delta G_{TS-N} = -\ln \frac{k_u^*}{k_u}$$

then

$$\phi_f = \Delta\Delta G_{TS-N} / \Delta\Delta G_{U-N}$$

$$\phi_u = \Delta\Delta G_{U-TS} / \Delta\Delta G_{U-N}$$

- for 2-state model, $\phi_f + \phi_u = 1$
- if $\phi_f \approx 1$, no difference in TS and U, so this residue is absent in TS formation; if $\phi_u \approx 1$, no difference in ΔG_{TS-N} , so this residue is almost done in TS
- if most residues show 0 or 1, the protein is completing local structure first: **Framework model**
- if most residues show fractional ϕ , the protein is forming partial structure first overall: **Nucleation model**

Notes

- k_f is on the "unfold" side.
- ϕ_f is on the "fold" side.
- $K_u = \frac{k_u}{k_f}$ for $N \rightleftharpoons U$, because $r_u \propto k_u[N]$

pulsed hydrogen labeling

- unfold the protein in D_2O , thus all $-NH-$ carry D
- dilute with D_2O buffer to make it partially refolded (for a while)
- dilute with excessive H_2O (for a while) at high pH (not that many H)
 - those which are unfolded carry H.
 - those which folded before maintains D: analyze by 2D NMR
- finish folding by lowering pH

Week12-1 Protein-Protein/Ligand Interaction

Basics

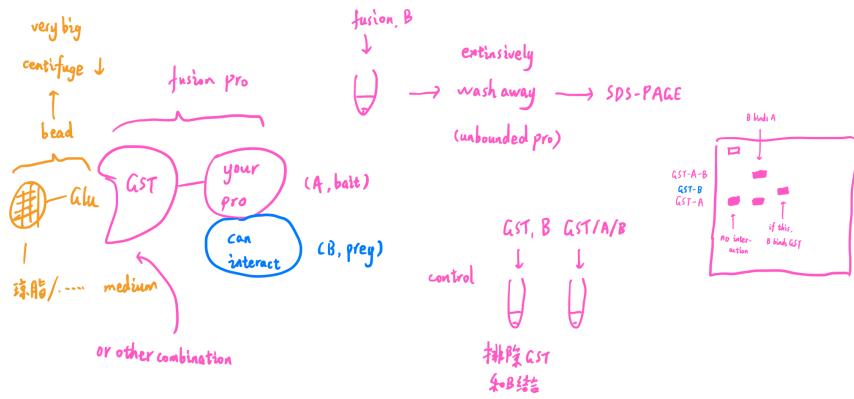
- non-covalent interactions: hydrophobic, ionic, H bonds...
- binding sites: inside? narrow?
 - small ligand: pockets in/outside
 - larger: surface (peptide) or deeply inside (heme)
 - huge: protein in grooves of DNA
- structure
 - lock-and-key model: complementarity
 - structural changes to minimize energy (flexible)

Qualitative Study

"pull-down" assay

[Pull-Down Assays | Thermo Fisher Scientific - SG](#)

- *in vitro*
- fusion protein
- other combination

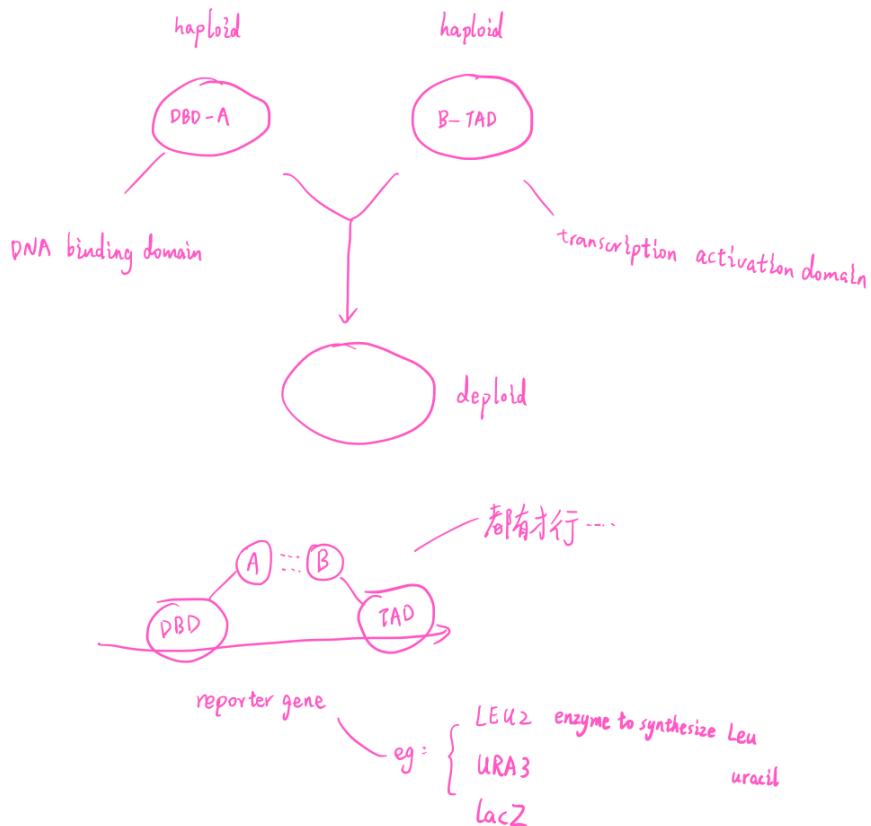


NMR CSP

- in vitro
- residues that shift a lot are important

Δppm

yeast two hybrid



Applications

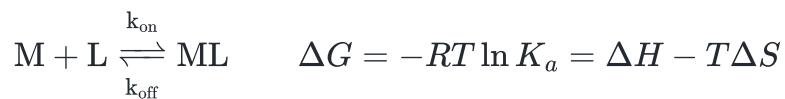
- screening cDNA library to assign functions; identify drug target
- mutation study (residues involved in interaction/binding)
- drug discovery/find regulator compounds

Maybe we can compare the interaction strength by observing expression level (increase/decrease all matter)...

Quantitative Study

thermodynamics

For 1:1 binding



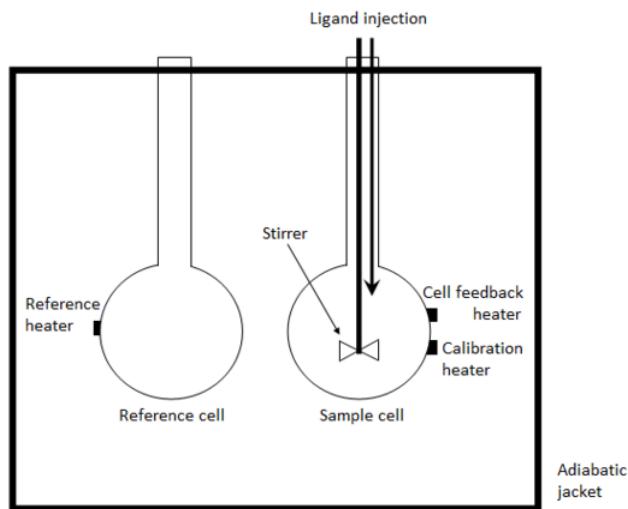
where $K_d = 1/K_a$. ($[ML] = K_a[M][L]$; $[M][L] = K_d[ML]$)

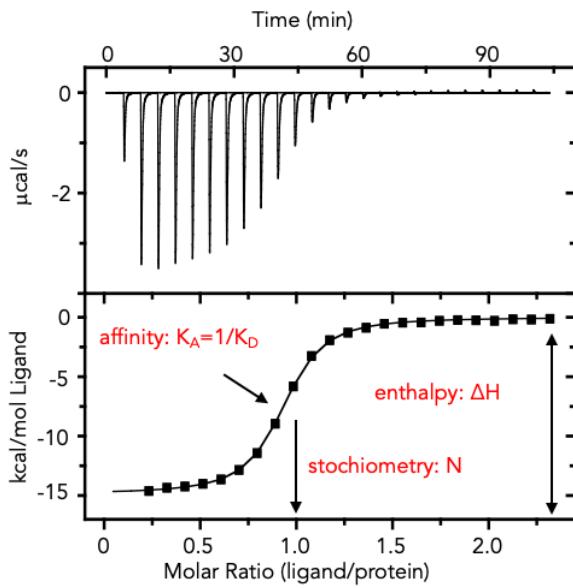
We wish low K_d and negative ΔG

ITC

[Isothermal Titration Calorimetry \(ITC\) | Center for Macromolecular Interactions \(harvard.edu\)](#)

apparatus:





Fit the curve, get ΔH , ΔS , K_a , n (stoichiometry, 化学计量)

equilibrium and saturation binding

single site

define fraction of M that has L on the site:

$$\nu = \frac{[ML]}{[ML] + [M]} = \frac{K_a[L]}{K_a[L] + 1} = \frac{[L]}{K_d + [L]}$$

assuming $[L] \gg [M]$ (L not consumed in binding, use $[L]_{total}$)

This will be a Michaelis-Menten-like curve.

multiple equivalent sites, no interaction

Just multiply by n because they are identical and independent. Also, the effect is proportional to experimental measure. We also define binding fraction of all sites.

$$\begin{aligned}\nu^* &= \frac{nK_a[L]}{K_a[L] + 1} \in [0, n) \\ Y &= \frac{\nu^*}{n} = \frac{K_a[L]}{K_a[L] + 1} \in [0, 1)\end{aligned}$$

equivalently,

$$\frac{\nu^*}{[L]} = K_a n - K_a \nu^*$$

this is called **Scatchard equation**. We can get K_a (slope) and n (x intercept).

multiple non-equivalent sites, no interaction

Due to independence, we can add up the results of all kinds of sites.

$$\nu^* = \sum_i \frac{n_i K_{a,i}[L]}{K_{a,i}[L] + 1}$$

But the results given:

- there's n separate lines, the slope of which is $K_{a,i}$
- We can still know $\sum_i n_i$ from x intercept.

multiple non-equivalent sites, interacting

The Scatchard plot becomes a curve. Modify:

$$\nu^* = \frac{n K_\pi [L]^{n_h}}{K_\pi [L]^{n_h} + 1}$$

$$K_\pi = \prod_i K_i$$

still $Y = \frac{\nu^*}{n}$, which can come from simple Scatchard plot...

define $\theta = Y/(1 - Y) = K_\pi [L]^{n_h}$, then got Hill equation:

$$\ln \theta = \ln K_\pi + n_h \ln [L]$$

Interpretation:

- fit the equation to get K_π (exp y-intercept) and n_h (slope)
- if no interaction, $\ln \theta - \ln [L]$ curve will be straight, or $n_h = 1$
- if $n_h < 1$, negative cooperativity, affinity \downarrow
- if interaction, the curve looks like



At begining, too few, no cooperativity; in the end, no more sites. Thus slope=1?

extrapolate the first and last part of the line and get x-intercept:

- $-\ln K_{a,1}/n_{h,1} = -\ln K_{a,1}$, also $-\ln K_{a,n}$
 - if $n_h > 1$, it looks like the left figure, $K_{a,1} < K_{a,n}$
 - if $n_h < 1$, it looks like the right figure, $K_{a,1} > K_{a,n}$

CD and fluorescence

change $[L]$ and get different signal

- fluorescence: strength of maximum difference. 280 nm for Trp

- fluorescence polarization: use r as ν^*

- rotation relaxation time: $p = \frac{3nV}{RT}$

where n is viscosity and V is molecular volume

- $r = \frac{I_H - I_V}{I_H + 2I_V}$, and according to Perrin equation

larger V , longer p , larger r

- advantages

- independent on light intensity, fluorophore conc, etc, not measuring intensity but the ratio between components
- less affected by instrumental change

- CD: both far (2°) and near UV (3°)

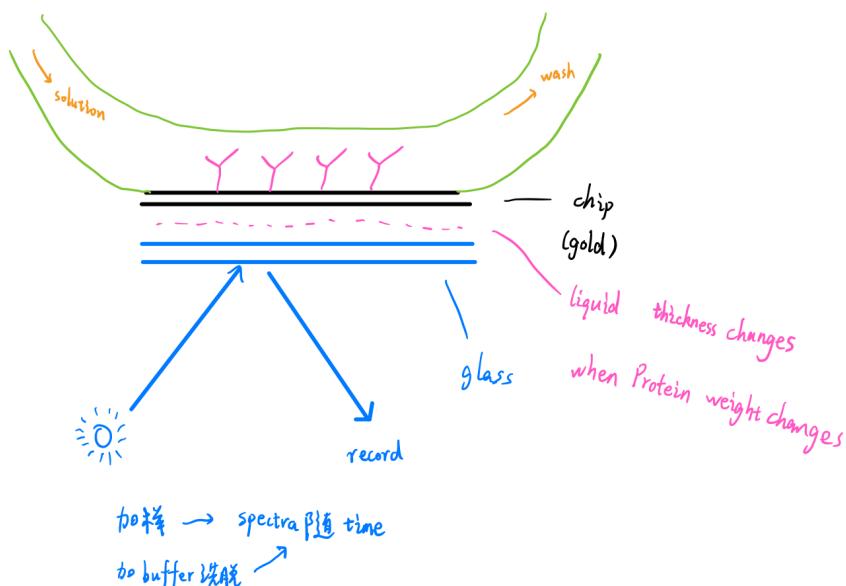
- eg: single site saturation. fit $Y = \frac{B_{max}[L]}{K_d + [L]}$

maybe: a linear term can be added to account for non-specific background binding

Kinetics of Binding

experiments

surface plasmon resonance (SPR)



steps

- immobilization: fix protein/ligand to the surface covalently
- association: add ligand/protein
- equilibrium: measure association
- dissociation: measure
- regeneration: break the remaining complex by acid/base

biolayer interferometry

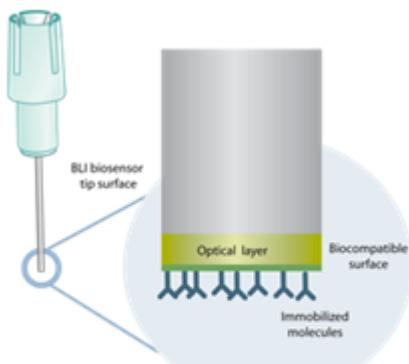


FIGURE 2

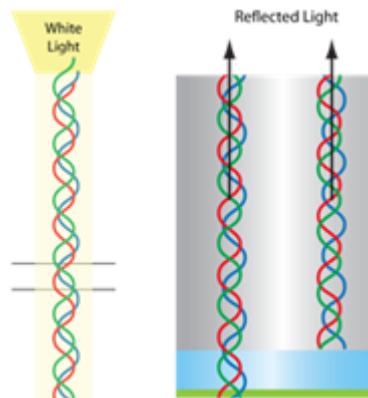


FIGURE 3

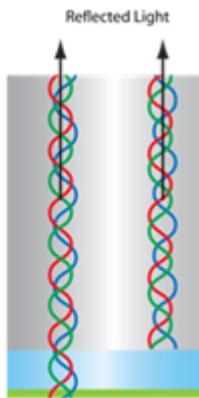
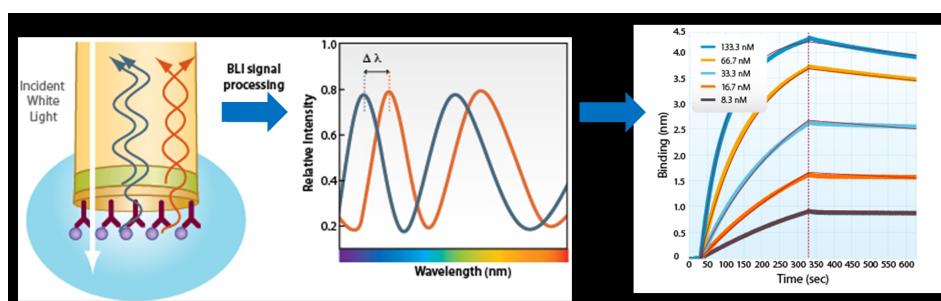
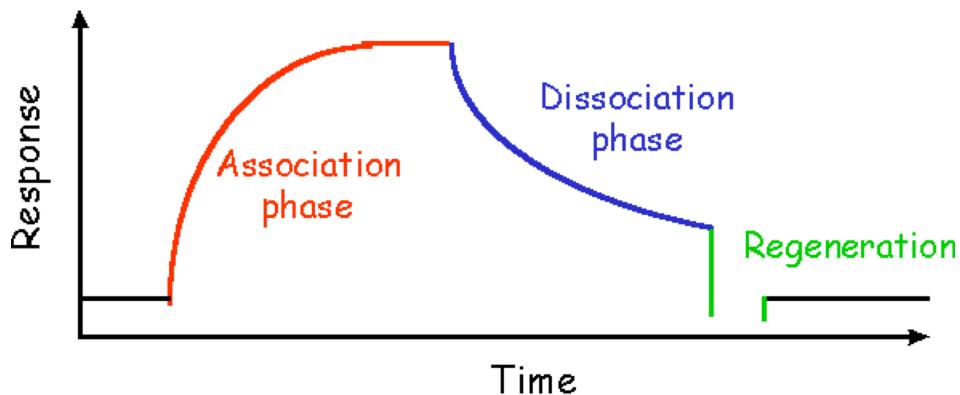


FIGURE 4



- adding ligand changes the thickness of biolayer, causing a spectrum shift. (interference, thus intensity changes)
- record the shift in different wavelengths.
- extract an index from it as "response". (like the peak shift)
- collect this "response" as a function of time
- operation: just dip the tip into ligand/buffer solutions

sensograms



k_{on} and k_{off} are fitted with $y = y_\infty + A \exp(-kt)$

LSM3243 outline

1. protein structure

- CORN
- peptide planar
- torsion angle definition
- ω : trans, Pro
- ϕ and ψ : N–C _{α} , C_O–C _{α}
- ramachandran plot
 - vdw radii, 3 regions
 - special: Gly, Pro. interrupt α , β
- structure features of typical conformations
 - helix, sheet, turn, coil
- side chain
 - definition, χ_1 , χ_2 , other, last χ .

2. interactions

- electrostatic
 - effect of solvent/environment, salt,
 - which parts carry charge
 - dipole: no charge. approximation. solvent with solute/solvent
-

3. nucleic acid

-

4. structure of membrane

- lipids
 - features

- amphiphilic (phospholipid, cholesterol)
 - differ in composition
 - asymmetric (two)
 - random (same)
- types
 - glycerophospholipids: most, 13 C
 - sphingolipids: long and straight chain, stronger vdW to hold
 - cholesterol: hold more (rigid, two faces, hydrophobic)
- lipid rafts
 - thicker
 - resistant
 - functional proteins
- protein
 - intergral
 - peripheral
 - anchored
- micelles
 - CMC, single layer--micelles--disrupted
 - DPC, triglyceride
-

5. membrane potential

6. membrane transport

7. CD

- principle
- applications
 - DNA conformation
 - protein secondary structure
 - protein folding under various conditions
(folding: does it lose normal structure)
 - protein-ligand
 - protein mutation
 - folding kinetics

8. fluorescence

- principle
 - basic
 - polarization
- applications
 - interactions and binding parameters
 - protein folding under various conditions
 - visualization
- FRET
 - conditions
 - study interactions (conformation)
- quenching: study accessibility into aromatic

9. NMR

- principle
 - basic, chemical shift, J coupling, T2
- applications
 - inhibitor/drug identification
 - protein folding
 - metabonomics (biomarker)

summary

10. conformation transition

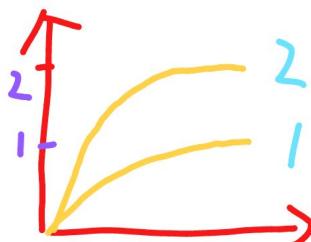
- model
 - basic and simple models
 - zipper model

11. protein folding

12. protein interaction

final exam

1. a protein with 10 beta (barrel), 2 helix, no Cys/Trp
 - a. possible interactions in helices? interaction that mainly hold the barrel?
possible interactions between helix and helix-cap?
 - b. reasons why 5% SDS makes unfolded
 - c. design a protocol to verify it's unfolded
 - d. how to determine if it interact with a 310-AA protein? if so, how to determine binding ratio?
2. a G₁₂ and C₁₂ dsDNA
 - a. 60°C, no salt, what form? why?
 - b. 1M NaCl, what form, why?
 - c. what technique to use?
3. a POPS vesicle, liposome-like
 - a. glucose, alanine, methanol. who can retain in the vesicle?
i might forgot "polar"
 - b. add 1 mM NaCl and 0.2 mM protein with 10 charge (Cl⁻ as counter ion) outside, calculate equilibrium concentration of all
4. two plots of v and [L]. method 1: measure the concentration of protein that bound on the ligand. method 2: measure the concentration of ligand that bound on the protein. data also given



- a. how many binding sites on the protein? why?

- b. K_d value?
- c. $[L]$ at 5 mM, the fraction of the protein bounded?
- d. 0.9 of the protein bounded, $[L]$?
- e. draw and label the scatchard plot of method 2.
how to determine n and K_a ? is the sites equivalent?
- f. suggest technique for each method and explain.

beyond course

Differential Scanning Fluorimetry (DSF)

- 差示扫描量热法 - 维基百科, 自由的百科全书 ([wikipedia.org](https://zh.wikipedia.org))
- Differential Scanning Fluorimetry (DSF) | Center for Macromolecular Interactions ([harvard.edu](https://chemistry.harvard.edu/research/groups/center-macromolecular-interactions/differential-scanning-fluorimetry-dsf))

more on NMR

radiofrequency pulse

free induction decay

when the nucleus magnet re-aligns, its direction changes and induces a chaning electric field.

HSQC spectra

- Wikipedia - Heteronuclear single quantum coherence spectroscopy
- <https://www.protein-nmr.org.uk/solution-nmr/spectrum-descriptions/1h-15n-hsqc/>

directly bonded atoms

protein-ligand

<https://www.originlab.com/index.aspx?go=Solutions/CaseStudies&pid=1701>

Note

I'll use quotation box to highlight important content, grey text to give additional notes, other colored text to denote other environments, so that it's clear.

Unimportant or very short part will be entitled "in addition" or inside a code box and a simple list. Do not spend much time on them.

so that it's clear.

The note is to help with reviewing rather than for publishing.

Do think and remember more than typing!!!

The colors: https://blog.csdn.net/qq_40862304/article/details/109893291

平常有时间做一点，就为考试周减轻压力。有时间还可多看书。

blubook主题可能为了图表的断页而压缩列表空间和字体。。

化生81资料编写小组

2020年11月23日