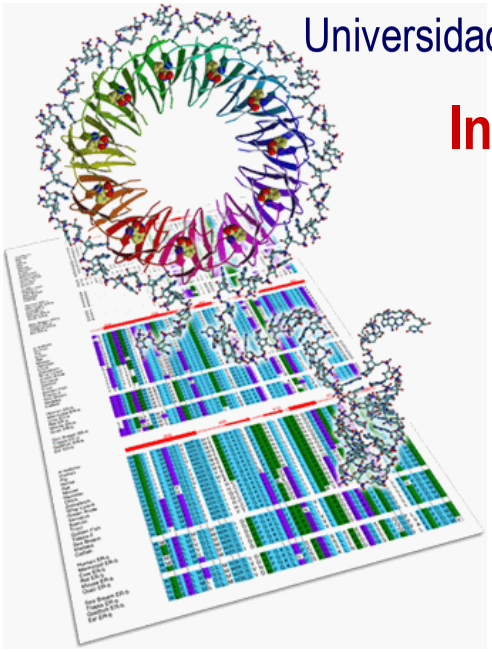
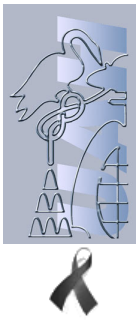
BioInfo 2004

Curso de Doctorado: BIOINFORMÁTICA
Universidad Autónoma de Madrid. Marzo-Abril 2004



Interacciones entre proteínas y moléculas pequeñas (I)

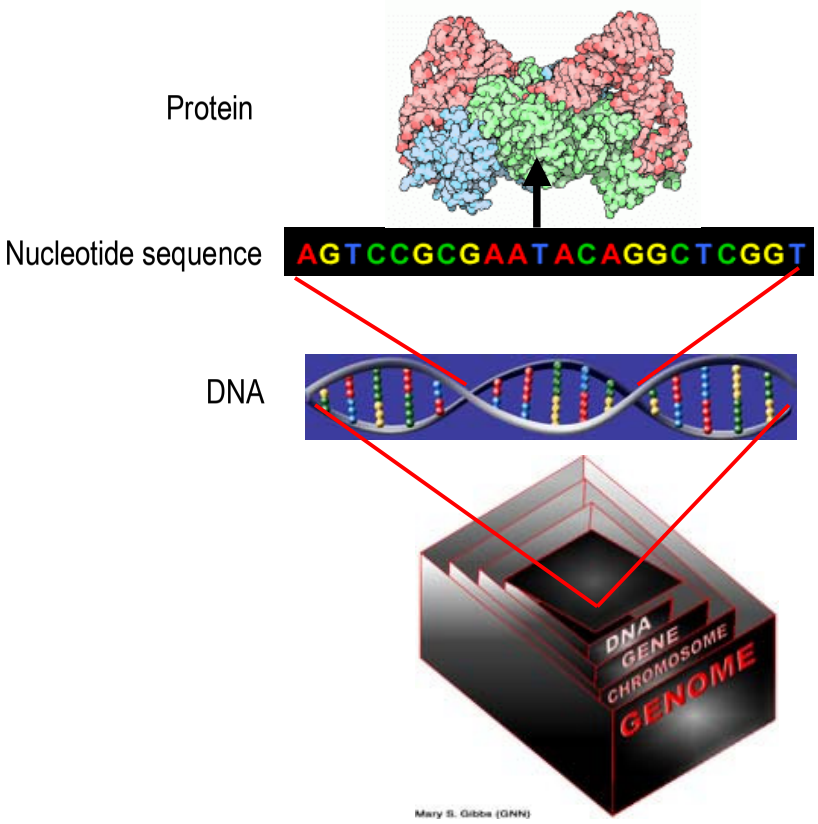
Federico Gago
Departamento de Farmacología
Universidad de Alcalá, Madrid



Interacciones entre proteínas y moléculas pequeñas

1. Funciones de las proteínas y movimientos asociados.
2. Concepto de ligando y sitio de unión. Ejemplos.
3. Bases de datos estructurales y programas asociados.
4. Caracterización estructural de moléculas pequeñas y sus complejos con proteínas.
5. Acoplamiento ligando-receptor (“docking”): algoritmos y programas.
6. Cribado virtual.
7. Relaciones estructura-actividad: QSAR y 3D-QSAR.
8. Diseño de nuevos ligandos.

Gene expression = Protein production



How Proteins Work

Proteins recognize and reversibly bind to other molecules: *cofactors, substrates, inhibitors*… Also *ions* and other *proteins*.

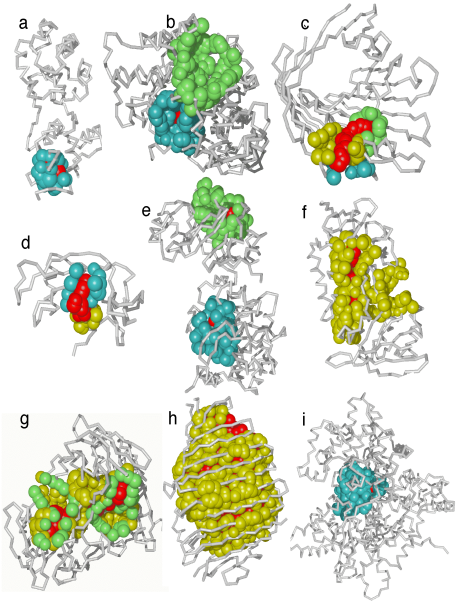
The bound molecule is called a **ligand**.

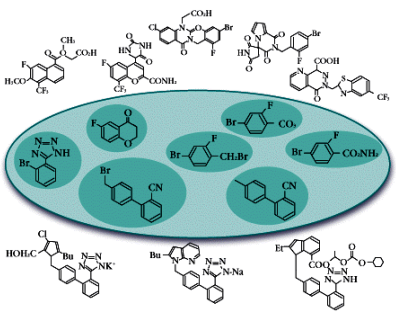
The region of a protein that associates with substrates and products is called the **active site**.

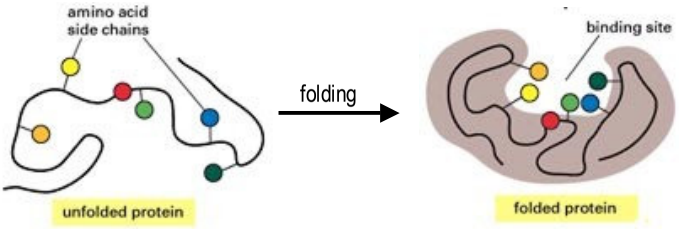
The region of a protein that associates with activator or inhibitor molecules is called an **allosteric site**.

Proteins can have > 1 binding site for different ligands

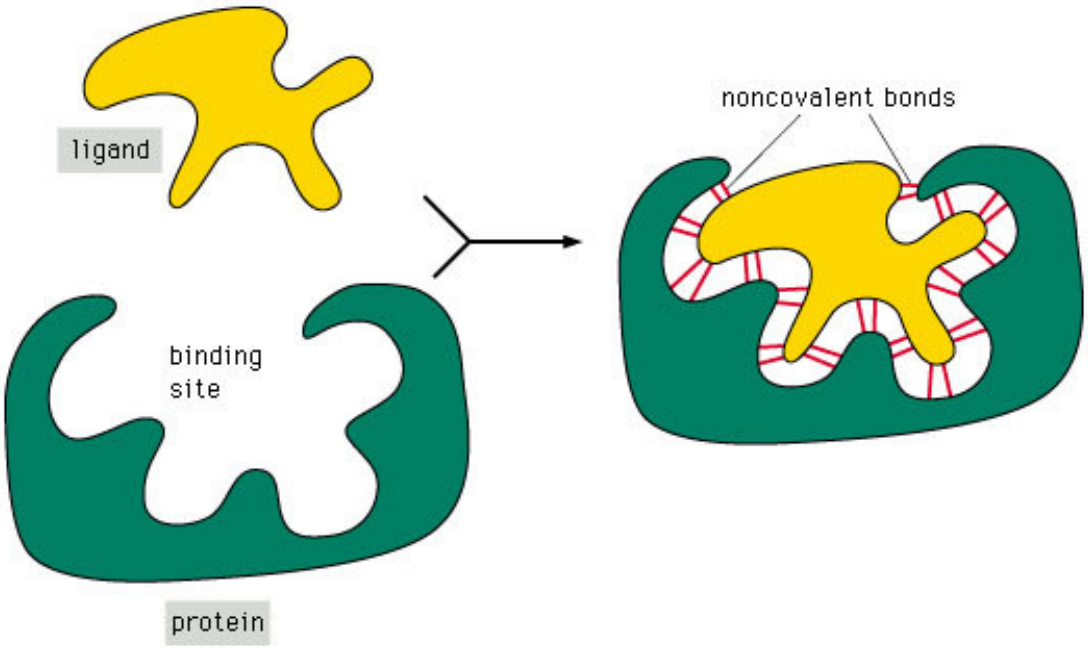
Proteins fold in such a way that they create specific sites that are the right *size*, *shape*, and *polarity* for their ligands.



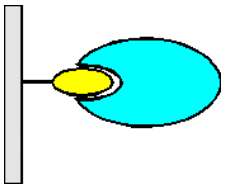




Ligand binding is highly selective

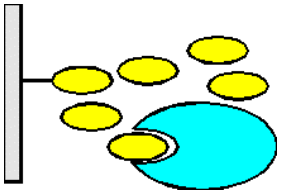


Affinity chromatography is a powerful purification method: the protein binds specifically to a **ligand**



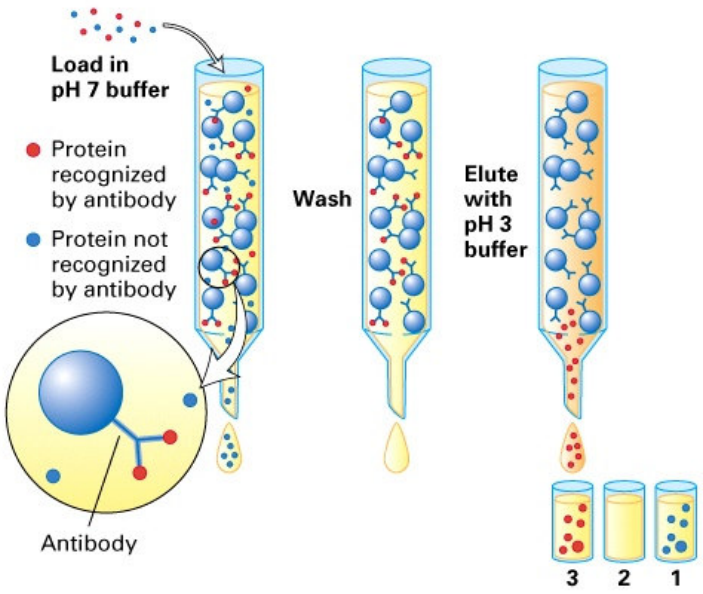
Ligand is covalently bound to the column

Protein specifically binds to a ligand for which it has a high affinity.

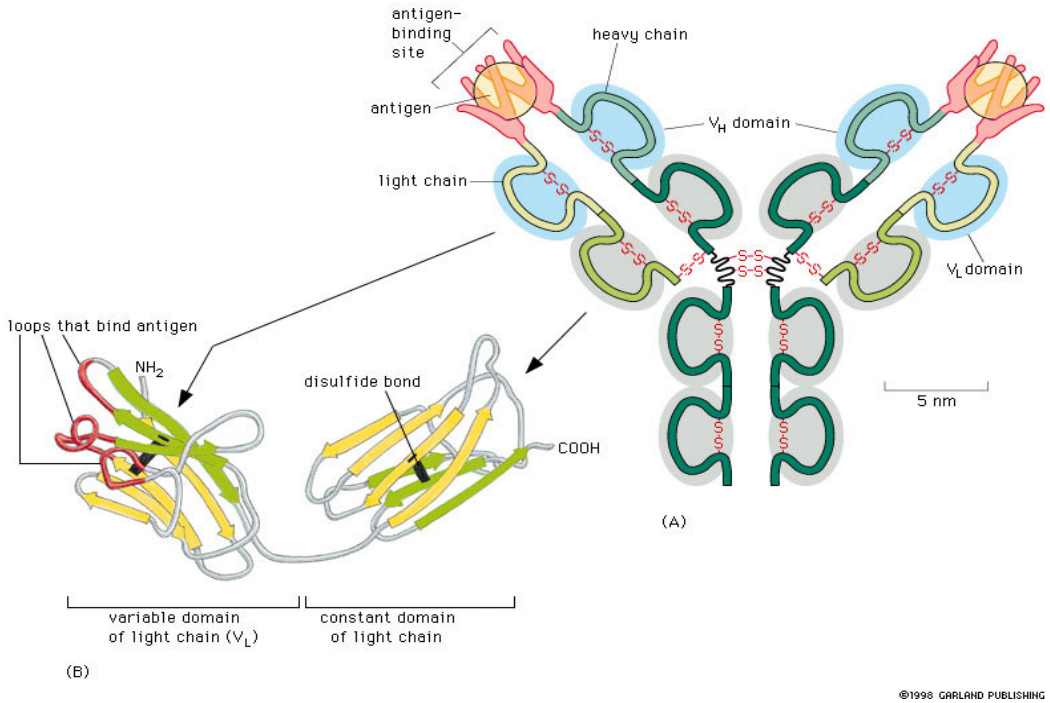


Elution of protein with unbound ligand

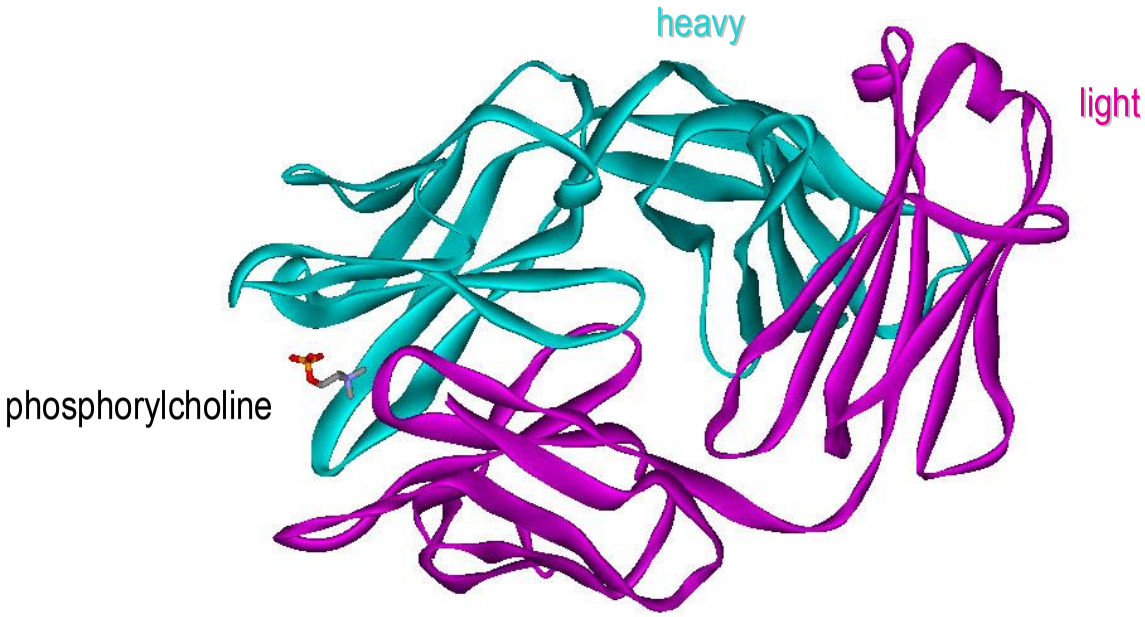
Separation of proteins by specific binding to another molecule: affinity chromatography



Antibodies selectively bind to antigens



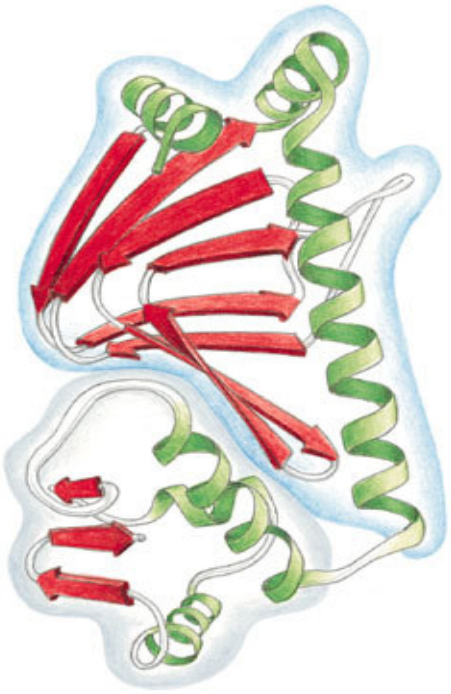
Immunoglobulin McPC603 Fab-Phosphocholine Complex (2mcp.pdb)



Protein Domains

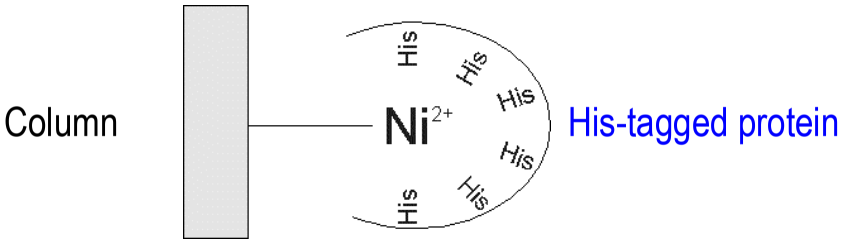
Different parts of a polypeptide chain can fold independently to form a stable structure called a domain.

The different domains of a protein often have different functions such as the DNA binding domain (small) and the cyclic AMP binding domain of the CAP protein shown.



protein molecule
made of two
different domains

Smallest ligand binding domain: His-His-His-His-His-His His-tag



Advantages:

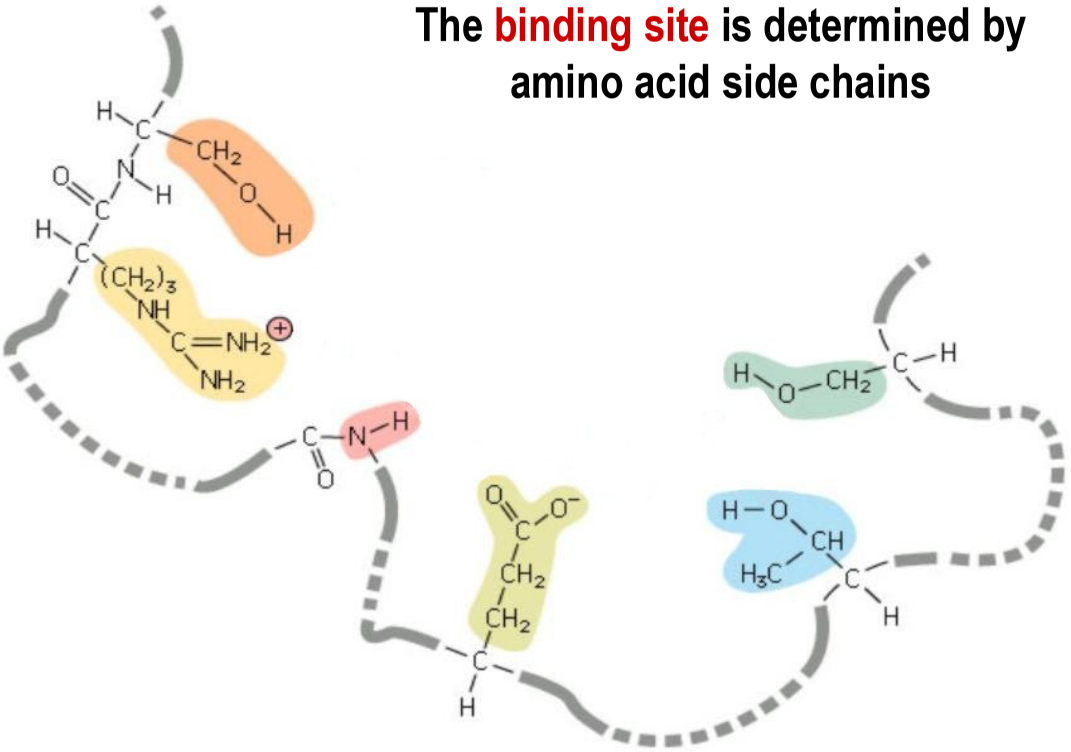
His-tags usually do not influence activity of protein, no need for removal

Allows purification in large quantities

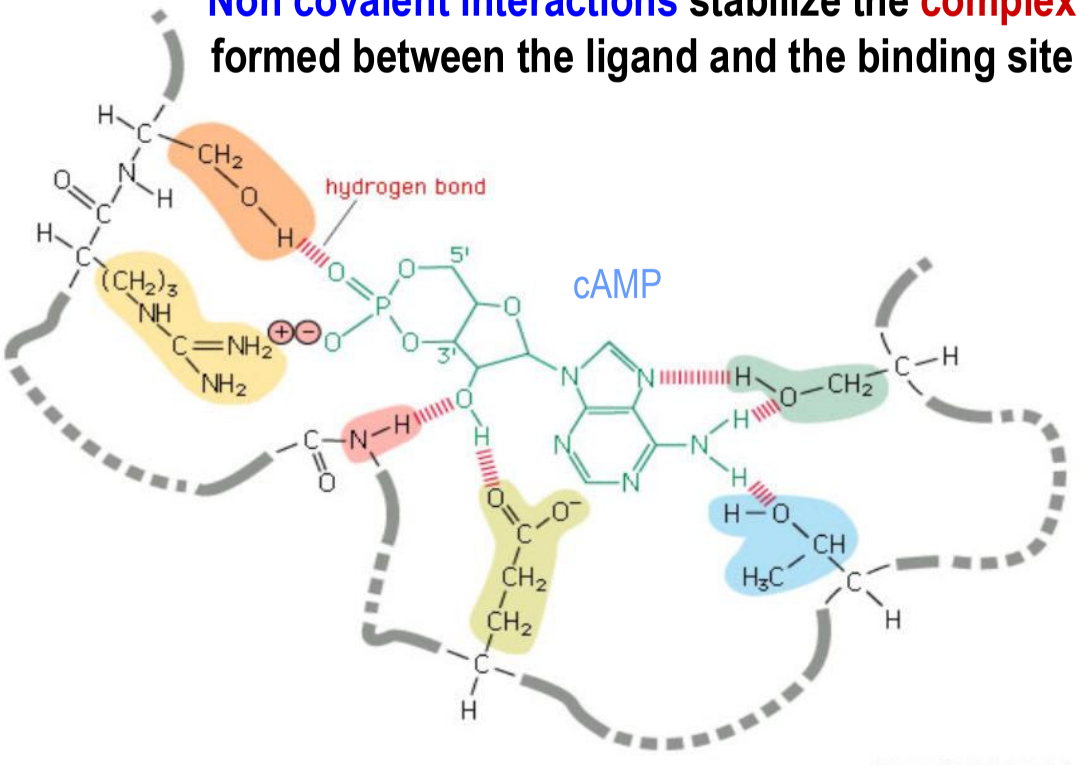
Simple to construct

Antibodies against the His-tag are available: detection of protein without need for specific antibodies

The **binding site** is determined by amino acid side chains



Non covalent interactions stabilize the **complex** formed between the ligand and the binding site



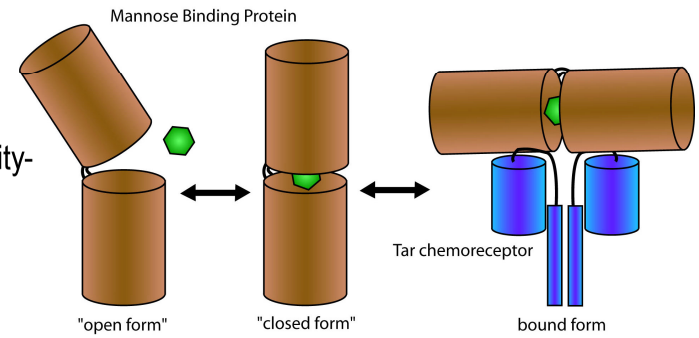
Proteins are:

Function	Example
Enzymes	DNA polymerase
Structural	collagen
Transporters	hemoglobin
Motors	myosin
Storage molecules	casein
Signalling molecules	insulin
Receptor molecules	rhodopsin
Regulatory molecules	lactose repressor
Speciality molecules	antifreeze
Defenses	antibodies

Proteins are flexible

- conformational changes can be **small**
  (molecular vibrations, small movements of amino acids);
  "breathing"
- or relatively **large**
  structural domains moving several nm

"Induced fit": the complementarity-enhancing structural adaptation that occurs between protein and ligand

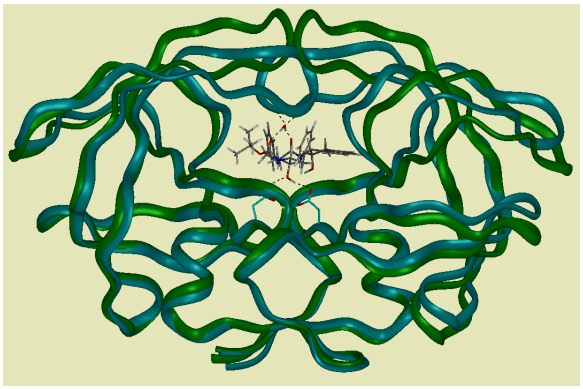


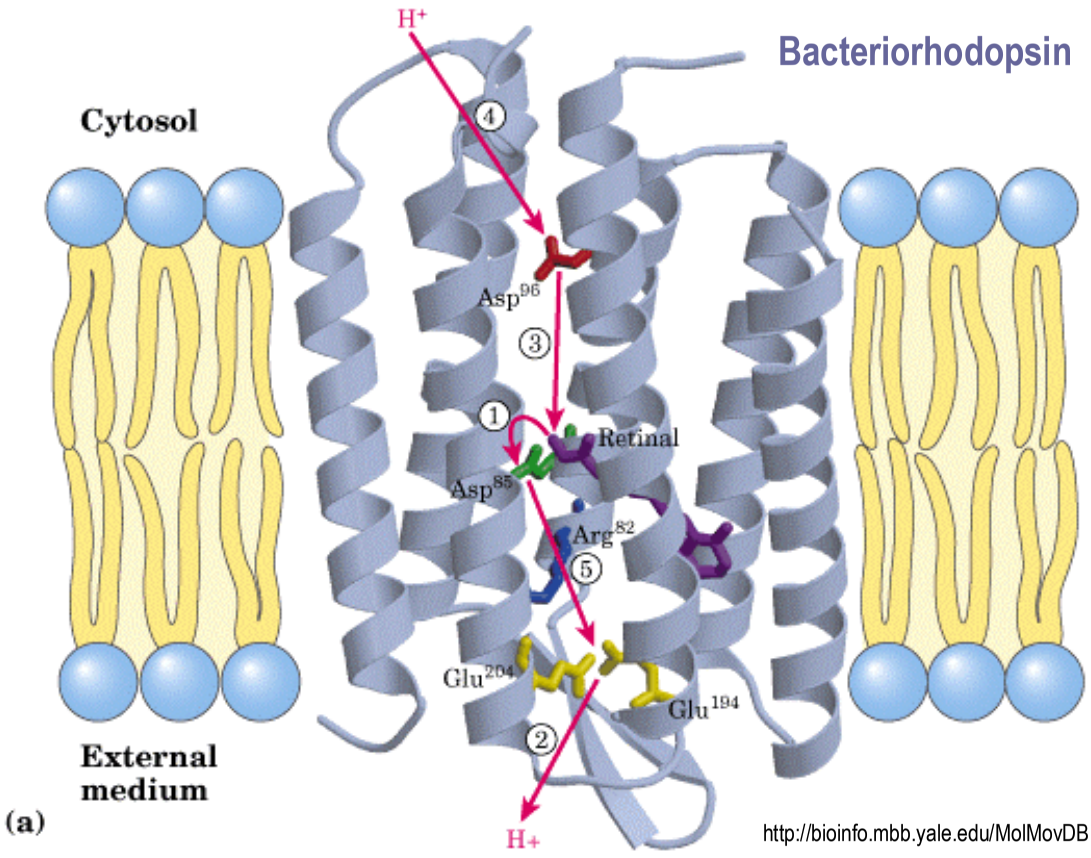
Conformational coupling: Ligand binding orients PBP receptor-binding face for signaling complex formation

I. Motions of Fragments Smaller than Domains

A. Motion is predominantly shear - Proteins for which two or more conformations are known: dihydrofolate reductase, insulin, thymidylate synthase, bacteriorhodopsin…

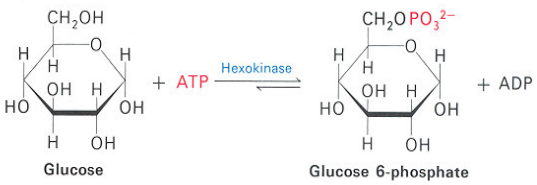
B. Motion is predominantly hinge - Proteins for which two or more conformations are known: annexin V (Trp motion), cystatin, enolase, HIV-1 protease, HhaI methyltransferase, immunoglobulin (CDR motion), isocitrate dehydrogenase, lactate dehydrogenase, lipase, malate dehydrogenase, seryl-tRNA synthetase, triglyceride lipase, triose phosphate isomerase, *Yersinia* protein tyrosine phosphatase, ras protein, recvinm ...

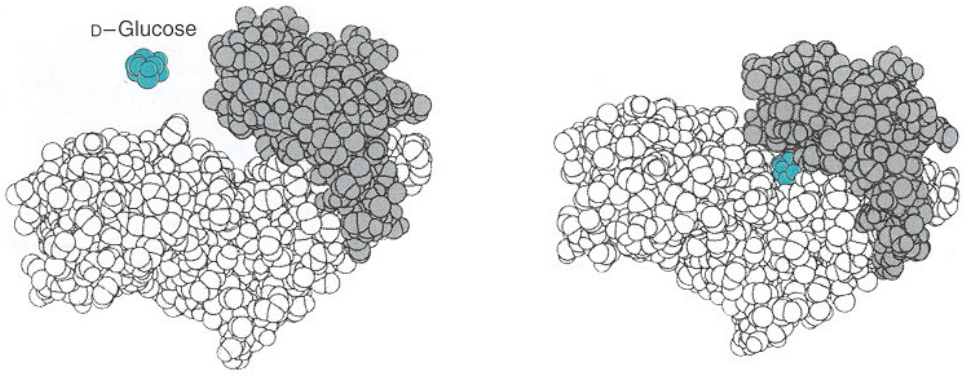




II. Domain Motions

A. Motion is predominantly shear - Proteins for which two or more conformations are known: alcohol dehydrogenase, aspartate amino transferase, citrate synthase, endothiapepsin, glyceraldehyde-3-phosphate dehydrogenase, glycerol kinase, hexokinase, human interleukin 5, phosphofructokinase (not allosteric transition), Trp repressor…





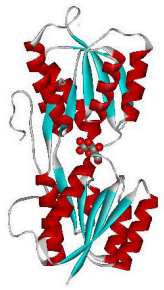
II. Domain Motions

B. Motion is predominantly hinge - Proteins for which two or more conformations are known: Acetylcholinesterase, adenylate kinase, annexin V (breathing motion), calbindin, calmodulin, canine lymphoma immunoglobulin (Fc-Fab hinge), catabolite gene activator protein (CAP), cell adhesion molecule CD2, DNA polymerase beta, diphtheria toxin, *E. coli*. periplasmic dipeptide binding protein, family-5 endoglucanase CelC, formate dehydrogenase, glutamate dehydrogenase, glutamine binding protein, GroEL domain, heat shock transcription factor, interferon-gamma, iron sulfur protein (bc1 complex), lactoferrin, Lysine/Arginine/Ornithine (LAO) binding protein, maltodextrin binding protein, phosphoglycerate kinase, recoverin, T4 lysozyme mutants (Ile3->Pro & Met6->Ile), TBSV coat protein, troponin-C, tryptophan synthase, c-Src tyrosine kinase, cAMP-dependent protein kinase (catalytic domain)…

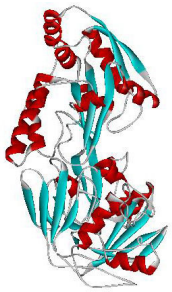
C. Motion involves partial refolding of tertiary structure - Proteins for which two or more conformations are known: Gα, HIV-1 reverse transcriptase, haemagglutinin, serpins…

Periplasmic Binding Proteins

a structurally conserved family of bilobate, soluble receptor proteins

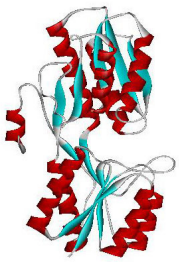


Glucose/Galactose Binding Protein

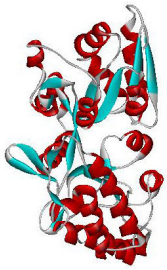


Dipeptide Binding Protein

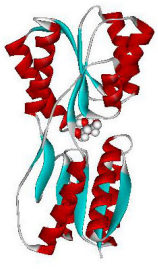
Leu, Ile, Val Binding Protein



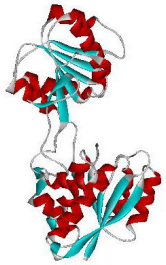
Arabinose Binding Protein



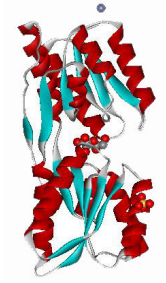
Maltose Binding Protein



Ribose Binding Protein



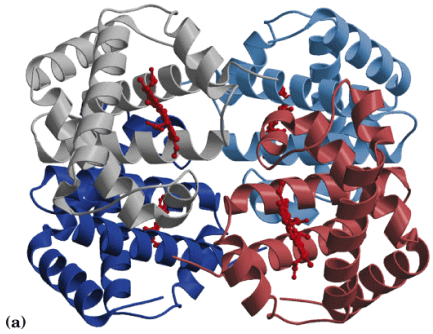
Leucine Binding Protein



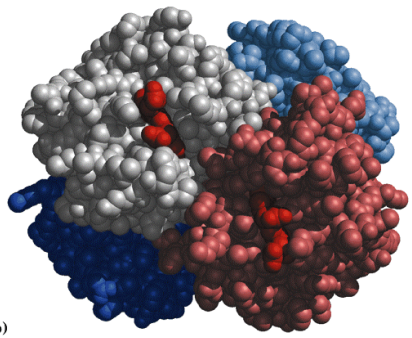
D-Allose Binding Protein

III. Larger Movements than Domain Movements involving the Motion of Subunits

A. Motion involves an allosteric transition - Proteins for which two or more conformations are known: aspartate transcarbamoylase, fructose-1,6-biphosphatase, glycogen phosphorylase, hemoglobin, Lac repressor core (allosteric motion), Lac repressor upon binding DNA (subunit motion via tetramerization domain), phosphofructokinase…



(a)



(b)

B. Motion does not involve an allosteric transition - Proteins for which two or more conformations are known: aspartate receptor, Bam HI endonuclease, immunoglobulin (VL-VH movement), *S. cerevisiae* PPR1 Zn-finger DNA recognition protein, erythropoietin receptor, F1-ATPase, polymerase processivity factor PCNA…

Dynamics and Relaxation

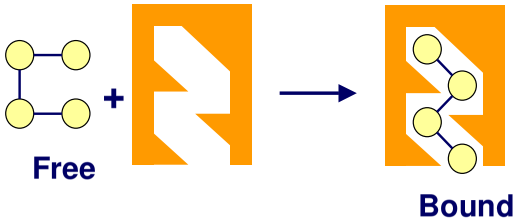
- **Time scales and molecular motions**

Atomic fluctuations, vibrations	10^{-15} to 10^{-12} s	< 1Å
Group motions (covalently linked units)	10^{-12} – 10^{-3} s	< 1 Å – 50 Å
Molecular rotation, reorientation	10^{-12} – 10^{-9} s	
Molecular translation, diffusion		
Rotation of methyl groups	10^{-12} – 10^{-9} s	
Flips of aromatic rings	10^{-9} – 10^{-6} s	
Domain motions	10^{-8} – 10^{-3} s	
Proline isomerization	> 10^{-3} s	

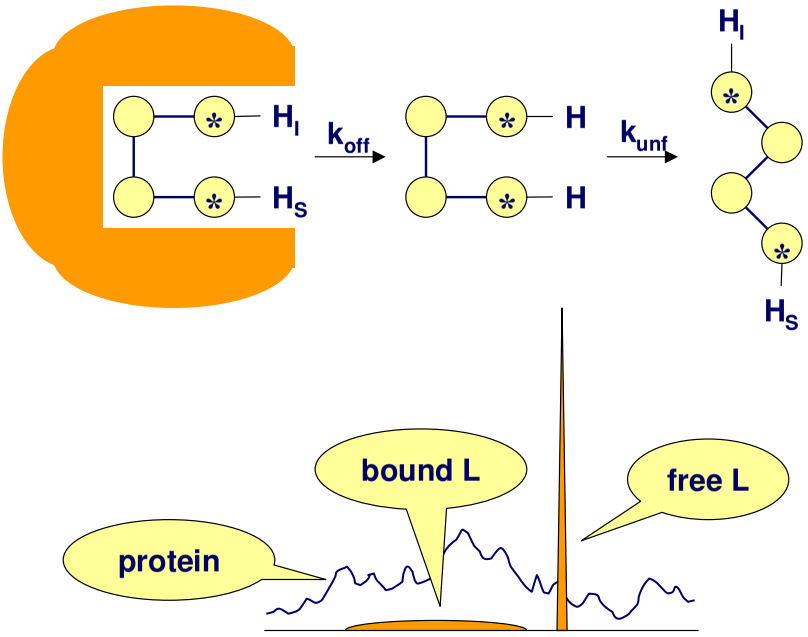
Chemical exchange (*e.g.* two protein conformations)
Amide exchange
Ligand binding

- **Time scales and molecular motions**

Atomic fluctuations, vibrations	Influences bond length measurements
Group motions (covalently linked units)	
Molecular rotation, reorientation	Relaxation, linewidths, correlation times
Molecular translation, diffusion	DOSY NMR
Rotation of methyl groups	^{2}H NMR
Flips of aromatic rings	^{2}H NMR
Domain motions	^{2}H NMR
Chemical exchange, Pro isomerization	Chemical shifts
Amide exchange	^{15}N-^{1}H HSQC
Ligand binding	Transferred NOE measurements



Transferred NOEs



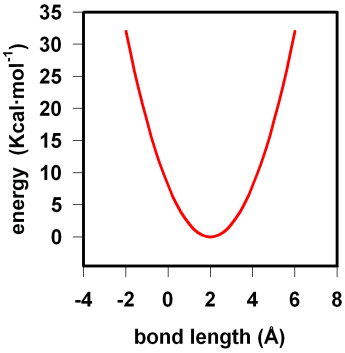
Molecular Mechanics

$$E_{pot} = E_{bonded} + E_{non\text{-}bonded}$$

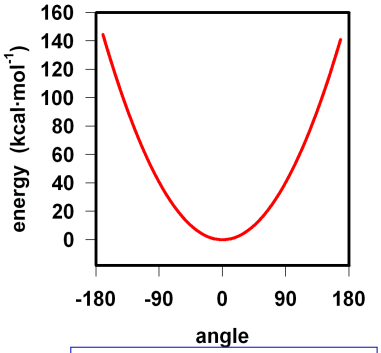
$$E_{bonded} = \Sigma_i \; E_{bond} + \Sigma_i \, E_{angle} + \Sigma_i \, E_{dihedral}$$

$$E_{non\text{-}bonded} = \Sigma_i \, E_{electrostatic} + \Sigma_i \, E_{van\ der\ Waals}$$

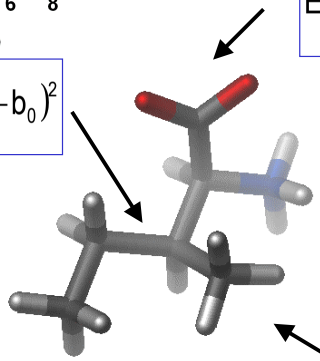
BONDING TERMS

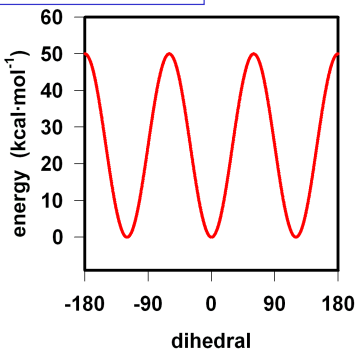


$$E_{bonds} = \sum_{bonds} \frac{1}{2} k_b (b - b_0)^2$$



$$E_{angle} = \sum_{angles} \frac{1}{2} k_\theta (\theta - \theta_0)^2$$



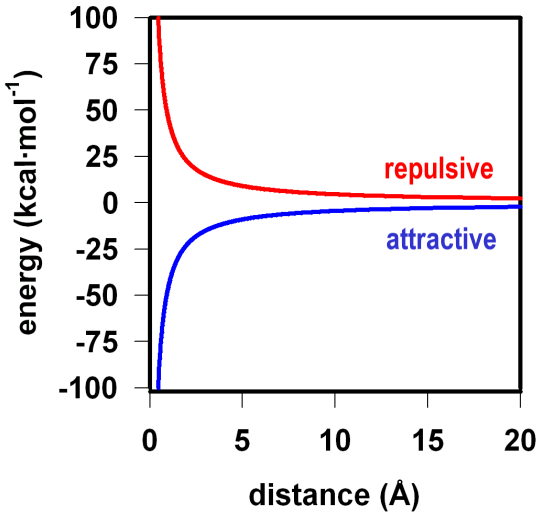


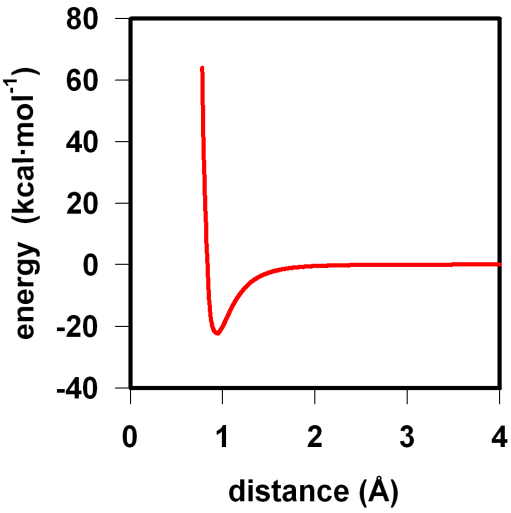
$$E_{dihedral} = \sum_{dihedrals} \frac{1}{2} k_d [1 + \cos(\varphi - \varphi_0)]$$

NON-BONDING TERMS

$$E_{electrostatic} = \frac{1}{4\pi\pi_0\varepsilon} \sum_{ij} \frac{q_i q_j}{r_{ij}}$$

$$E_{Lennard-Jones} = \sum_{ij} \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}$$

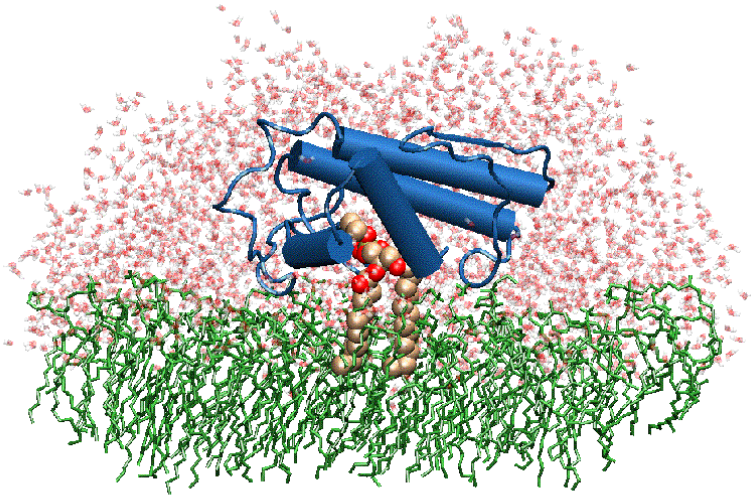




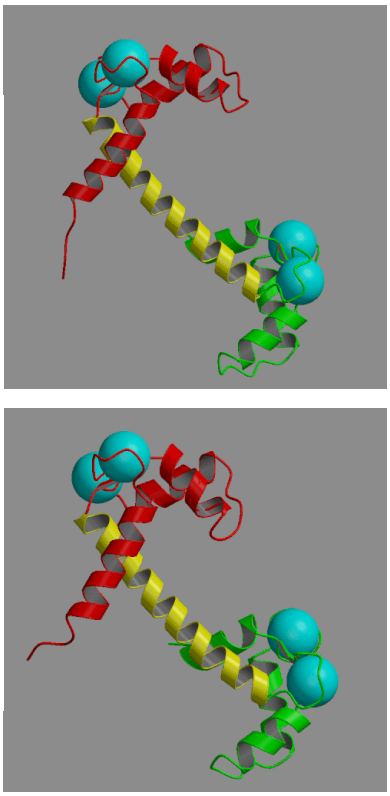
Molecular Dynamics

Newton's second law

$$-\frac{dV}{dx} = F = m \cdot a = m \cdot \frac{v}{t} = m \cdot \frac{d^2x}{dt^2}$$

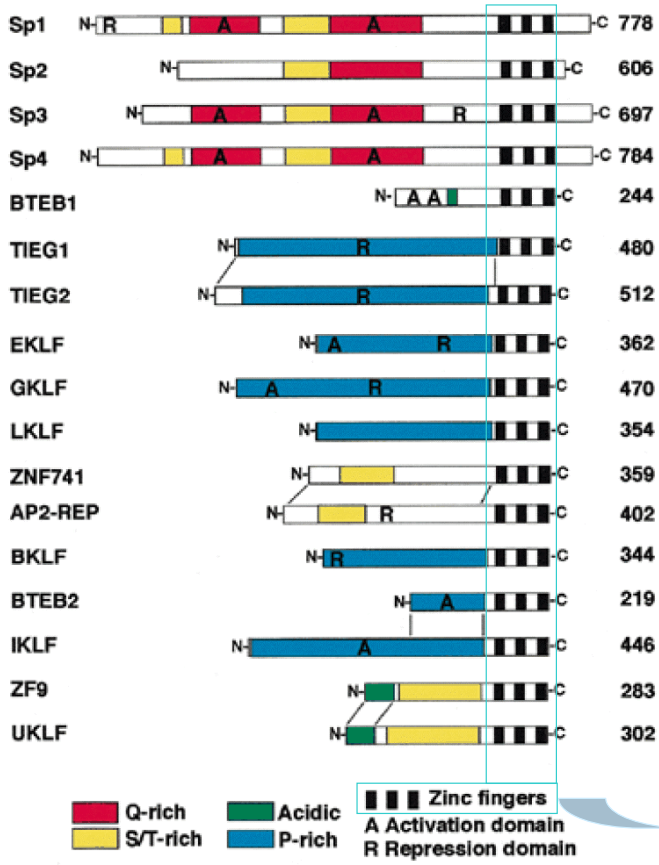


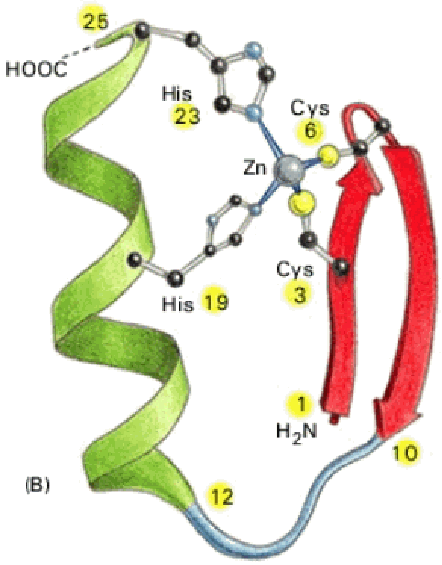
F. Zhou & K. Schulten: Molecular Dynamics Study of Phospholipase A_2 on a Membrane Surface. Proteins: Structure, Function, and Genetics, 25:12-27 (1996)

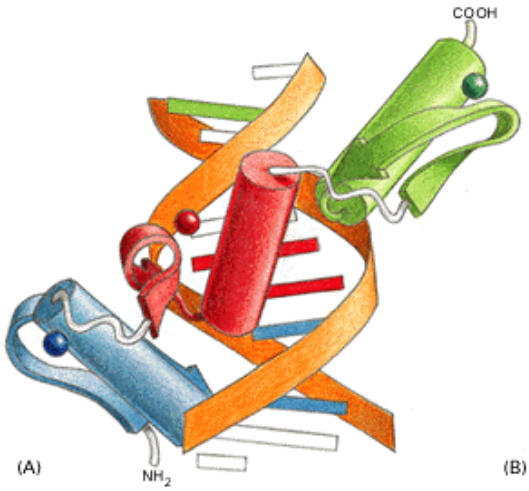


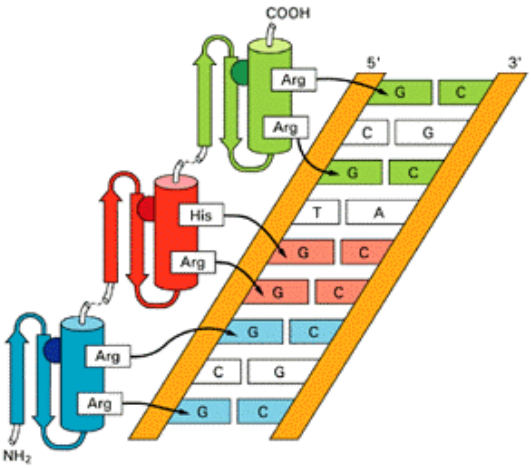
CALMODULIN

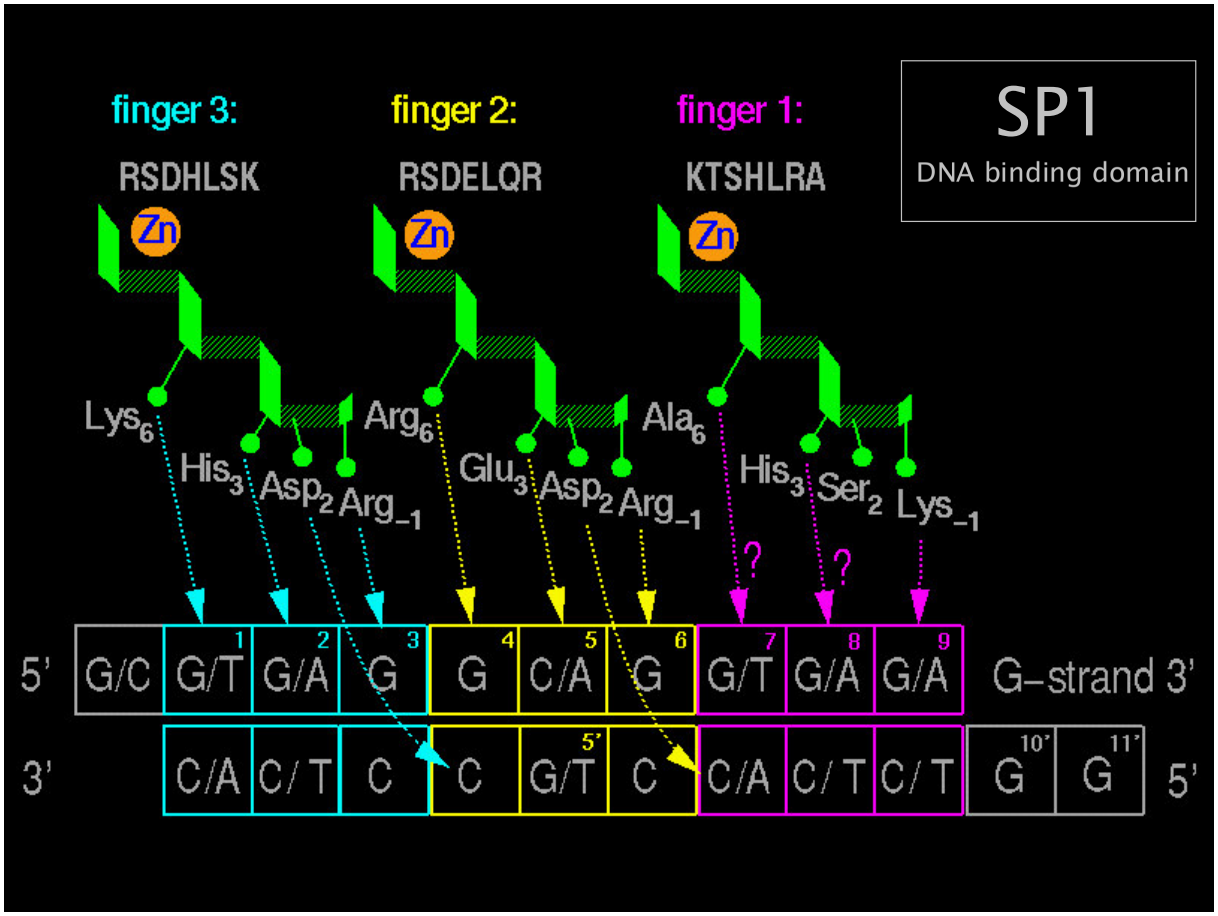
ZINC FINGERS

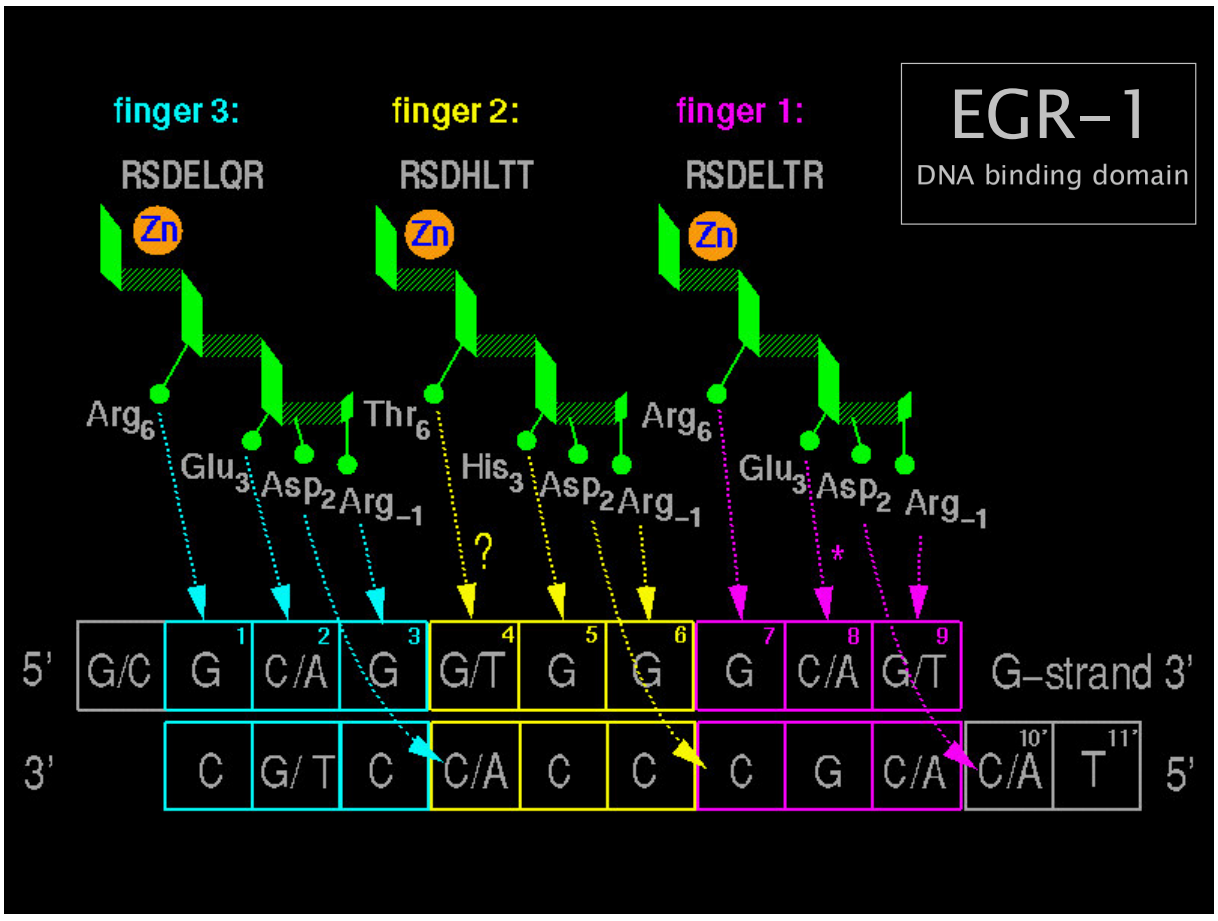




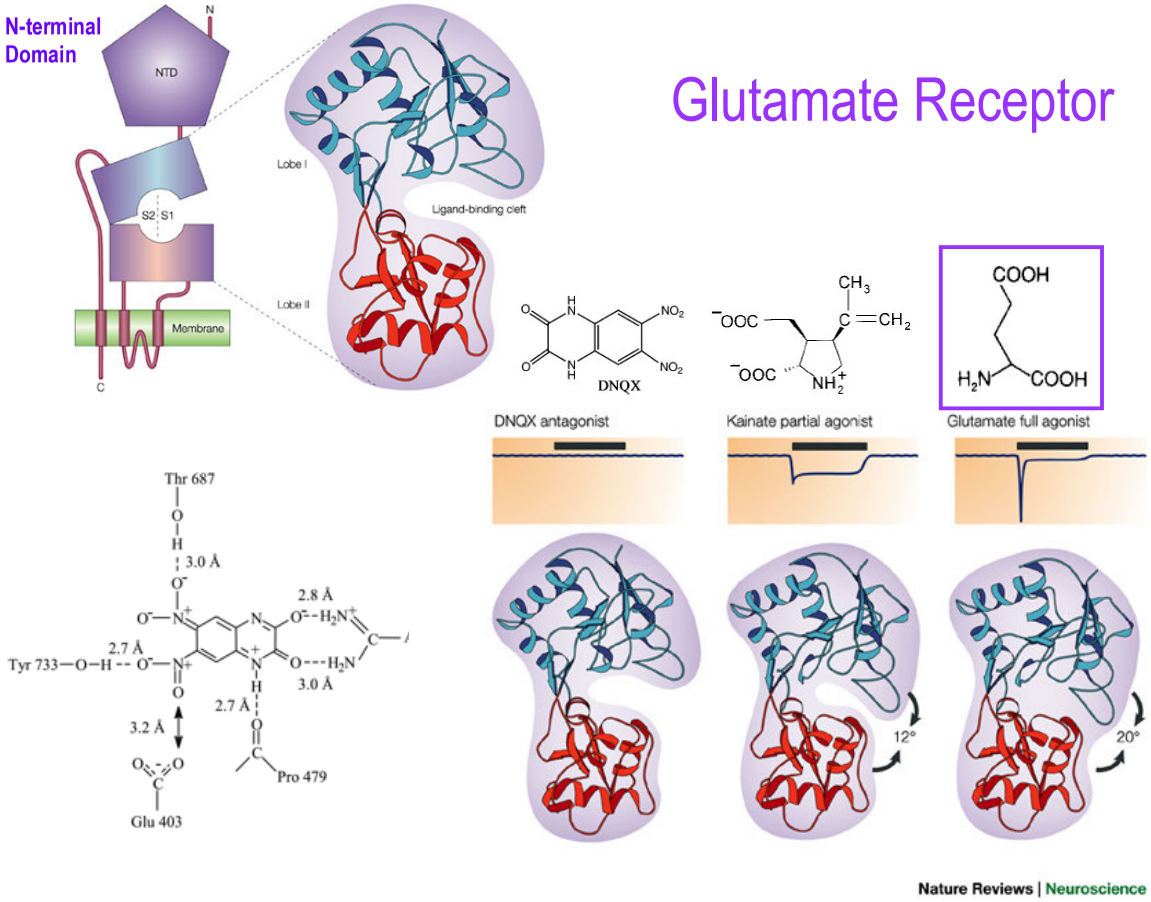





SP1
DNA binding domain
finger 3:
RSDHLSK
finger 2:
RSDELQR
finger 1:
KTSHLRA
Zn
Zn
Zn
Lys6
His3
Asp2
Arg−1
Arg6
Glu3
Asp2
Arg−1
Ala6
His3
Ser2
Lys−1
?
?
5'
G/C
G/T
G/A
G
G
C/A
G
G/T
G/A
G/A
G–strand 3'
3'
C/A
C/T
C
C
G/T
C
C/A
C/T
C/T
G
G
5'
1
2
3
4
5
6
7
8
9
10'
11'
5'


EGR-1
DNA binding domain
finger 3:
RSDELQR
finger 2:
RSDHLTT
finger 1:
RSDELTR
Zn
Zn
Zn
Arg6
Glu3
Asp2
Arg−1
Thr6
His3
Asp2
Arg−1
Arg6
Glu3
Asp2
Arg−1
?
*
5'
G/C
G
C/A
G
G/T
G
G
G
C/A
G/T
G–strand 3'
3'
C
G/T
C
C/A
C
C
C
G
C/A
C/A
T
5'
1
2
3
4
5
6
7
8
9
10'
11'

Glutamate Receptor

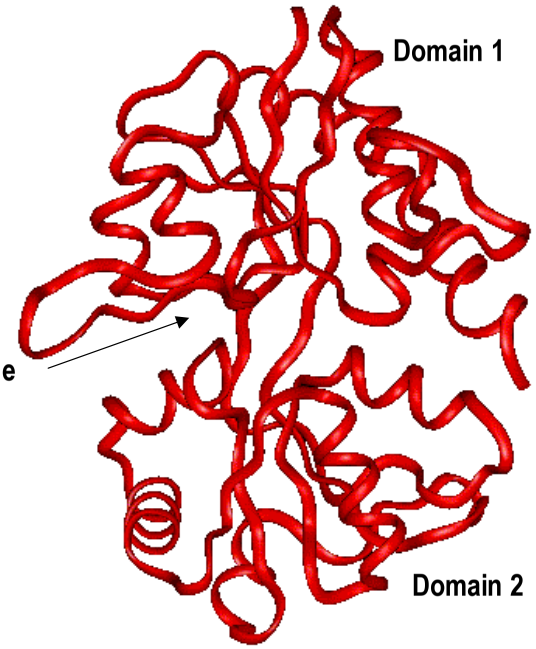


Nature Reviews | Neuroscience

Structural homology between PBP and S1S2 Glu R2

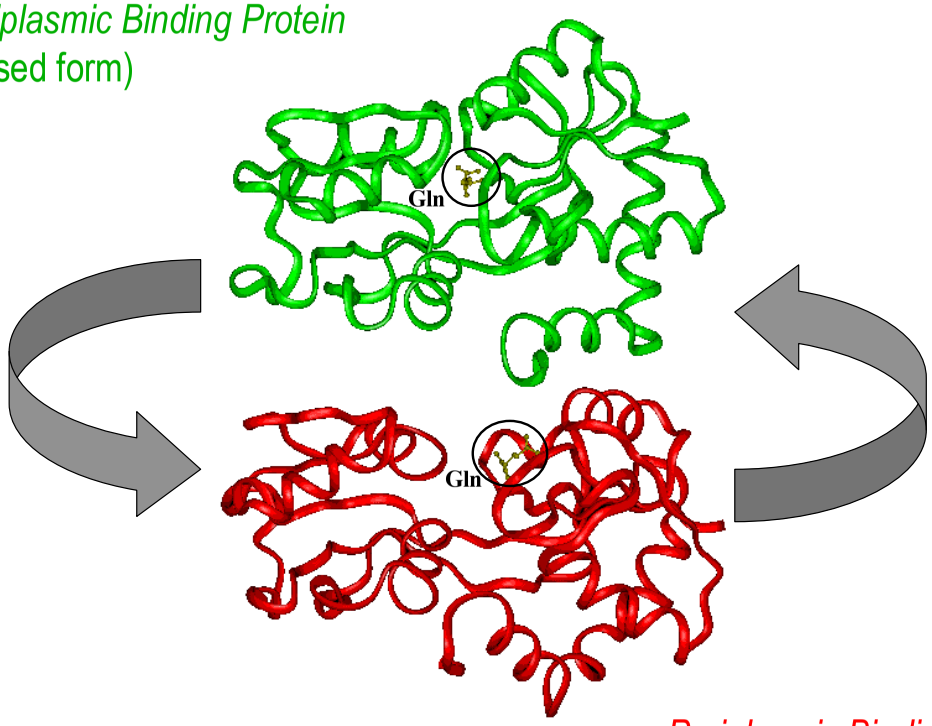
Periplasmic Binding Protein

Glutamate Receptor 2 (+kainate)



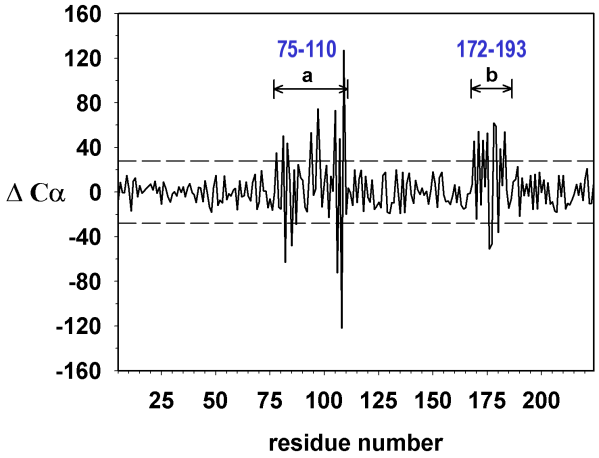
hinge

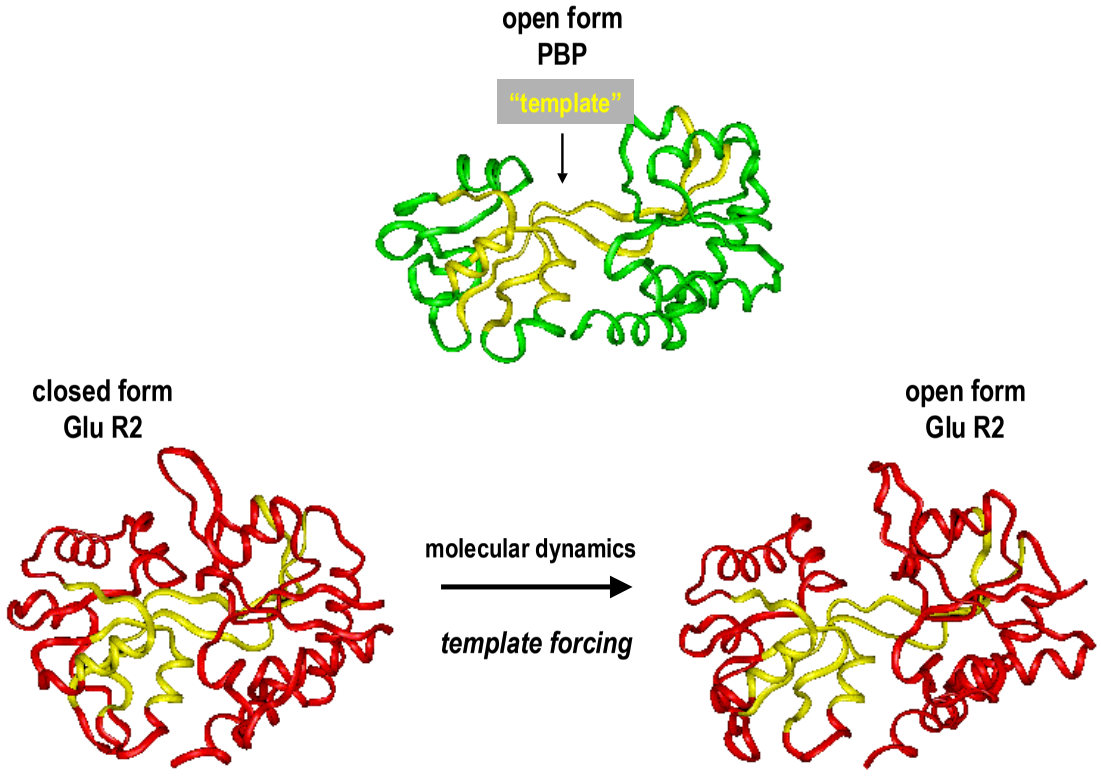
Periplasmic Binding Protein
(closed form)

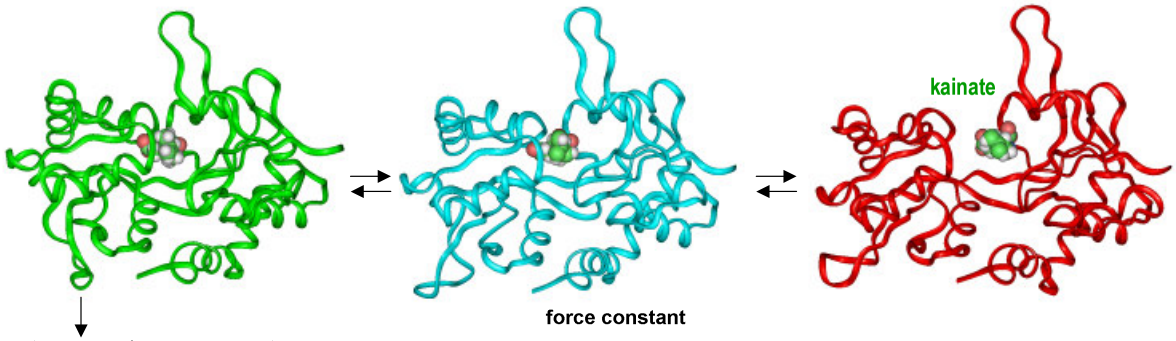


Periplasmic Binding Protein
(open form)

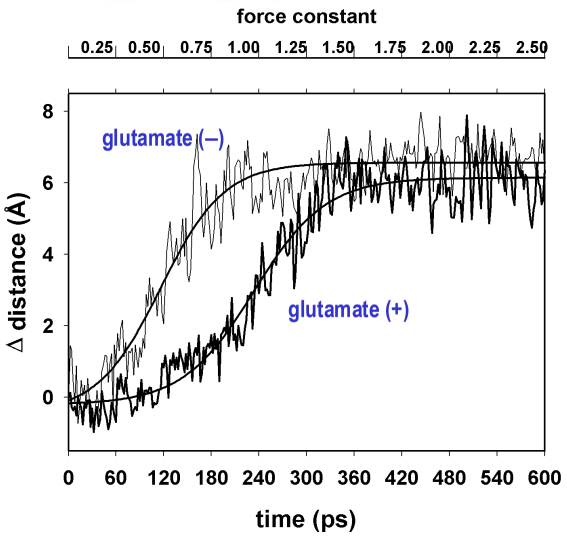
Characterization of the hinge region in periplasmic glutamine-binding protein (QBP)





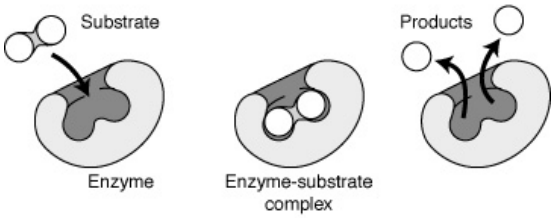


transmembrane segment insertion loop.



distance between Cα carbons of Ser-158 and Arg-108

Enzymes: A special case of protein function

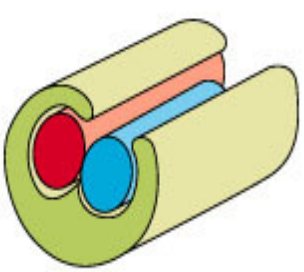


Enzymes bind and assist in the chemical transformation of other molecules

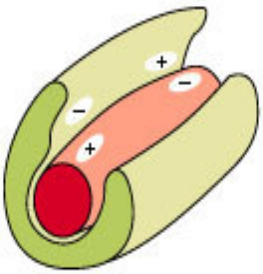
Substrate: molecule acted upon by an enzyme (analogous to ligand)

Catalytic site: substrate-binding site (analogous to ligand-binding site)

How do enzymes catalyze reactions?



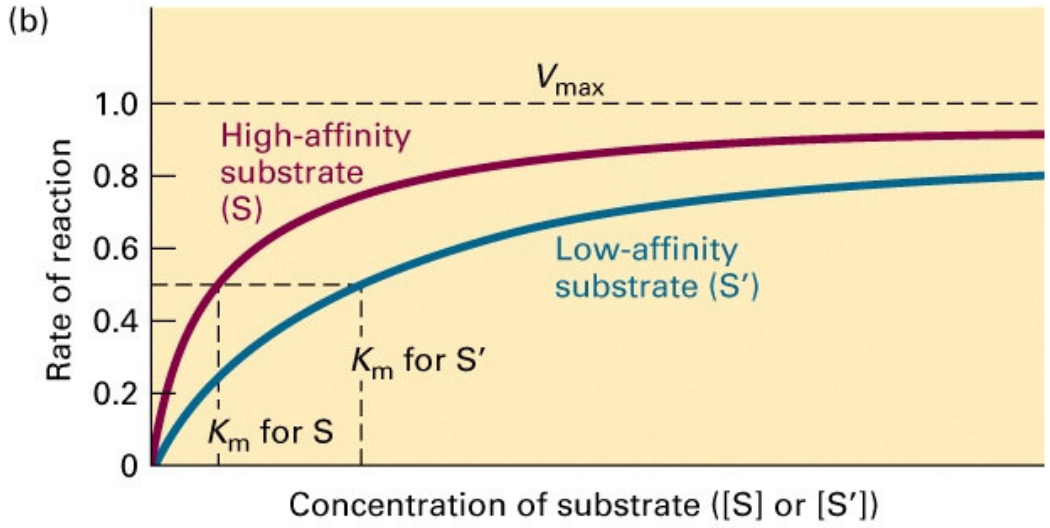
(A) enzyme binds to two substrate molecules and orients them precisely to encourage a reaction to occur between them



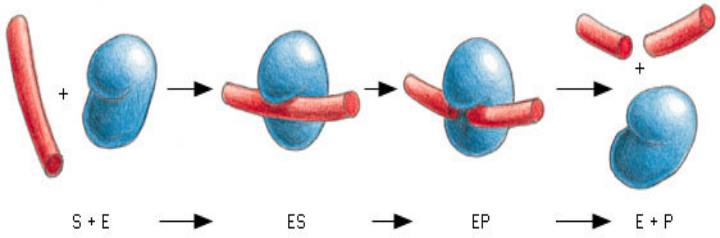
(B) binding of substrate to enzyme rearranges electrons in the substrate, creating partial negative and positive charges that favor a reaction

(C) enzyme strains the bound substrate molecule, forcing it toward a transition state to favor a reaction

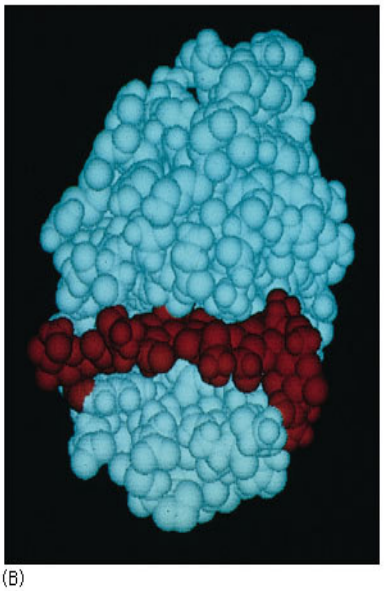
Kinetics of an enzymatic reaction are described by V_{max} and K_m



Lysozyme catalyzes the cutting of a polysaccharide chain

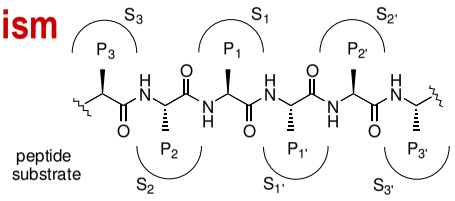


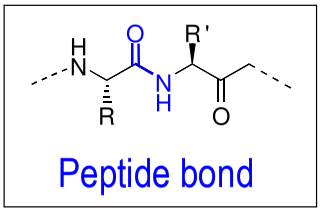
(A)

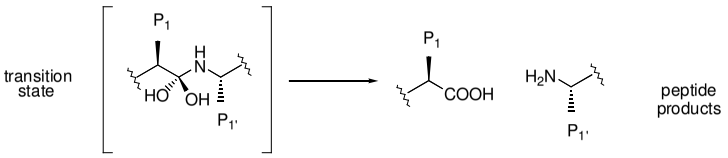


(B)

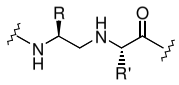
Hydrolytic mechanism in HIV-1 protease



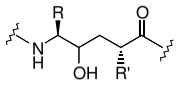




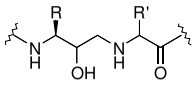
Examples of non-hydrolyzable isosteres of the peptide bond cleaved by HIV-1 protease



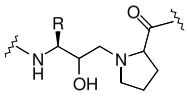
Reduced amide



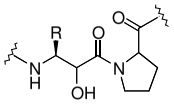
Hydroxyethylene



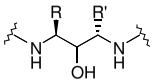
Hydroxyethylamine (acyclic R')



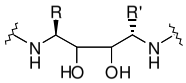
Hydroxyethylamine (cyclic R')



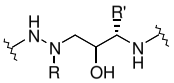
Norstatine



Mono-ol

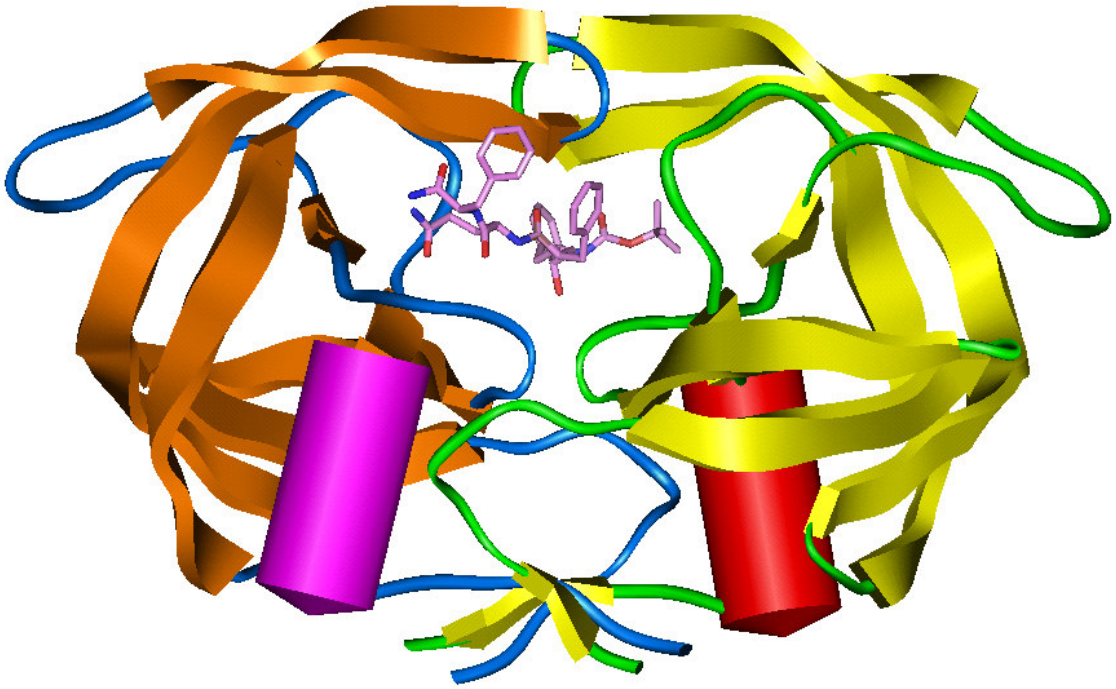


Diol

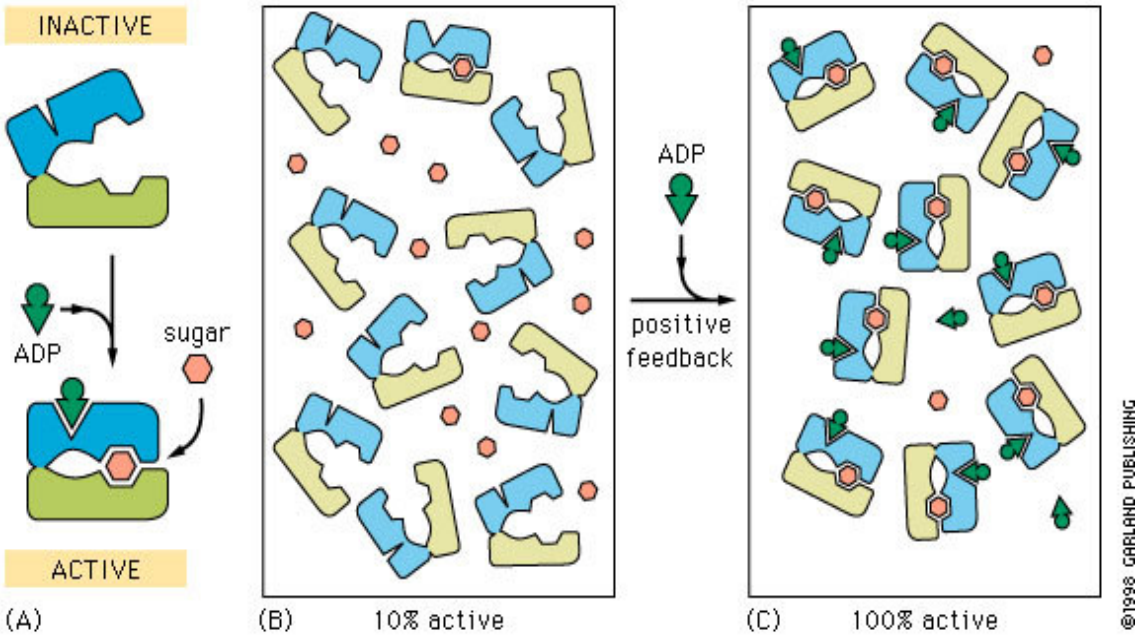


Azapeptide

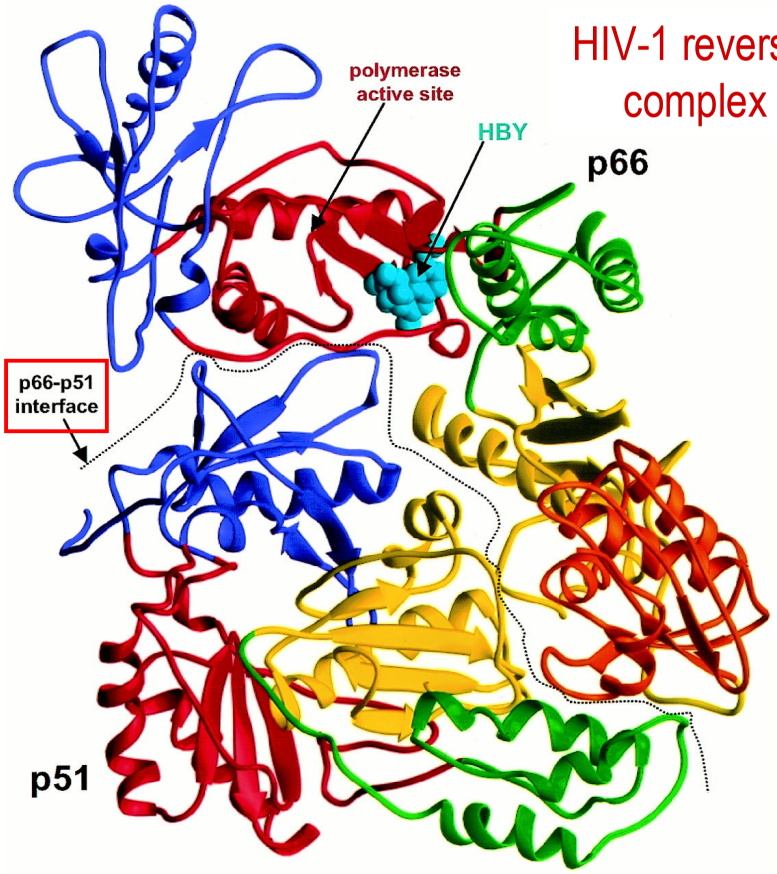
HIV-1 protease in complex with inhibitor QF-34

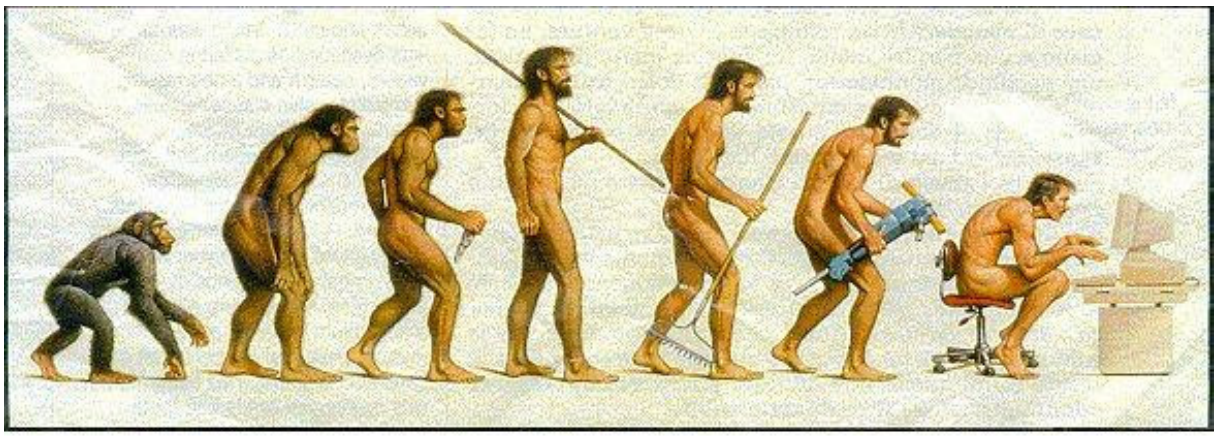


Enzyme activation caused by an allosteric change



HIV-1 reverse transcriptase in complex with inhibitor HBY





Somewhere, something went terribly wrong

QUESTIONS WELCOME

PREGUNTAS, POR FAVOR

E-mail: federico.gago@uah.es