



Modelado de Proteínas.



Modelado de Proteínas

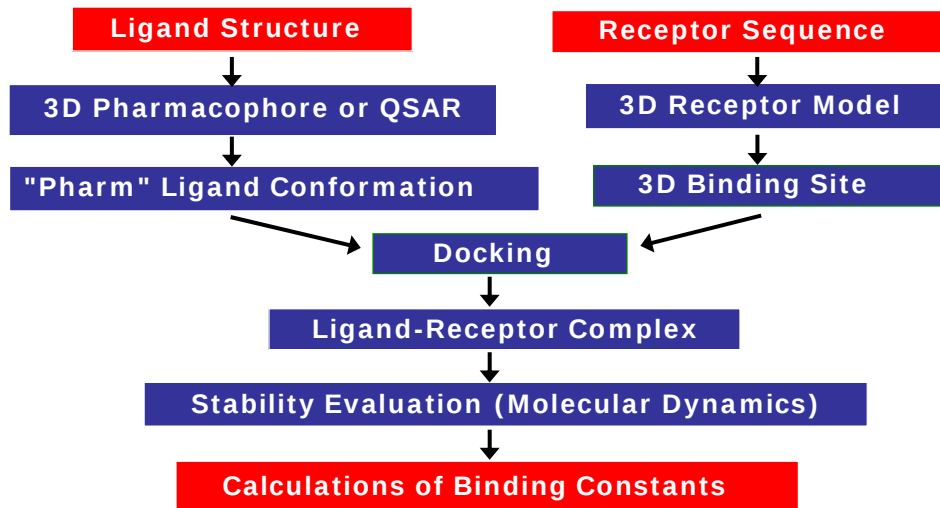
I- Introducción. Estructura de las proteínas.

II- Estructura experimental de proteínas.

III- Modelado de proteínas.

IV- Metodología del modelado de proteínas.

Computer Modelling of Drug-Receptor Interactions



ESTRUCTURA DE LAS PROTEÍNAS:

- Funciones biológicas de las proteínas
- Naturaleza química de las proteínas
- Diversidad estructural de las proteínas

Conexión entre secuencia – plegamiento – función

TEMA 7

I- Introducción.

II- Estructura experimental de proteínas:

1. Cristalografía de rayos-X.
2. Resonancia magnética nuclear.
3. Microscopía electrónica.
4. Otros.

III- Modelado de proteínas.

IV- Metodología del modelado de proteínas.

TEMA 7

I- Introducción.

II- Estructura experimental de proteínas.

III- Modelado de proteínas: Métodos de predicción de estructura.

1. Predicción de estructura secundaria.
2. Reconocimiento del plegamiento.
3. Modelado por homología.
4. Modelado por topología de proteínas de transmembrana.

IV- Metodología del modelado de proteínas.

III- Modelado de proteínas: Métodos.

Clasificaciones de métodos de predicción de estructura de proteínas

Nivel estructura proteínas	Secundaria	terciaria	cuaternaria	
Representación de la proteína	1D 	2D 	3D 	4D 
Uso de información extra				
<i>Ab Initio</i>	pred. str. secundaria	mutaciones correlacionadas	- dinámica molecular - minimización de energía	docking
<i>No Ab-Initio</i>	pred. str. secundaria	- modelado por homología - threading		docking con filtros

Secondary Structure Prediction

Three-state model: helix, strand, coil.

Given a protein sequence:
- NWVLSTAADMQGVVTDGMASGLDKD . . .
↓
Predict a secondary structure sequence:
- LLEEEELLLLHHHHHHHHHHHLLHHHL . . .

Methods:

- statistical
- stereochemical

Accuracy: 50-85%

Statistical Methods

Residue conformational preferences:

Glu, Ala, Leu, Met, Gln, Lys, Arg -	helix
Val, Ile, Tyr, Cys, Trp, Phe, Thr -	strand
Gly, Asn, Pro, Ser, Asp -	turn

Chou-Fasman algorithm:

Identification of helix and sheet "nuclei"
helix - 4 out of 6 residues with high helix propensity
sheet - 3 out of 5 residues with high sheet propensity
Propagation until termination criteria met

Modelado de Proteínas

I- Introducción.

II- Estructura experimental de proteínas.

III- Modelado de proteínas: Métodos de predicción de estructura.

1. Predicción de estructura secundaria.
2. Reconocimiento del plegamiento ("threading").
3. Modelado por homología.
4. Modelado por topología de proteínas de transmembrana.

IV- Metodología del modelado de proteínas.

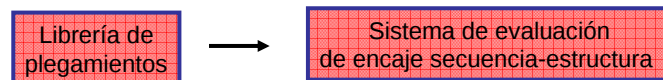
2. Reconocimiento del plegamiento ("threading"). Nace de la observación de que proteínas no relacionadas entre sí adoptan plegamientos similares.

3. Modelado por homología: la proteína problema es similar a una de estructura 3D conocida.

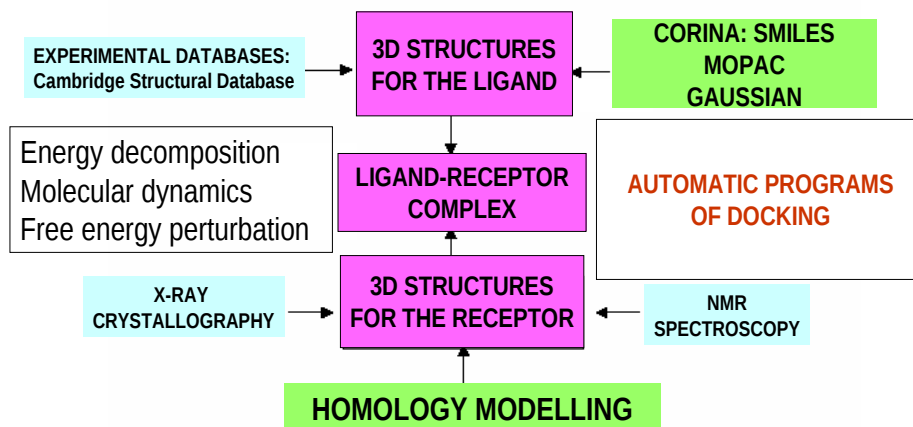
· Estos dos tipos de métodos usan información no contenida en la secuencia de aminoácidos y en ello se diferencian de los métodos *ab-initio*, que solo usan esa información pero que aún están en un grado de desarrollo que los hace inaplicables de manera general en la predicción de estructura.

· La información "extra" que usan ambos métodos proviene, básicamente, de otras proteínas cuya estructura se ha determinado experimentalmente.

· Ambos métodos se basan en la relación encontrada entre parecido estructural y parecido en secuencia.



¿Cuándo usamos el Modelado por Homología?



¿QUÉ ES EL MODELADO POR HOMOLOGÍA?

Es un método que permite predecir la **Estructura terciaria (3D)** de una proteína deseada (**Diana**) conociendo:

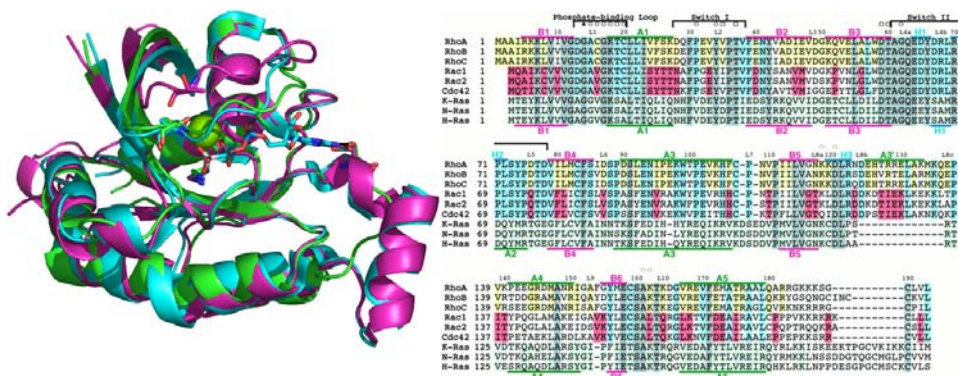
- Su secuencia de aminoácidos (**Estructura primaria**)
- La estructura terciaria de una **homóloga** (**templado/molde**) resuelta por rayos-X o NMR

¿EN QUÉ PRINCIPIOS SE BASA EL MODELADO POR HOMOLOGÍA?

- Secuencias similares adoptan estructuras similares
- En una familia, los residuos conservados en la secuencia, conservan la misma estructura
- La topología de los residuos del sitio activo también es conservada
- El proceso de evolución por selección natural "tolera" variaciones de la secuencia

¿EN QUÉ PRINCIPIOS SE BASA EL MODELADO POR HOMOLOGÍA?

- Secuencias similares adoptan estructuras similares
- En una familia, los residuos conservados → misma estructura
- La topología de los residuos del sitio activo también es conservada



PROCEDIMIENTO

- 1- Encontrar un molde adecuado
- 2- Alineamiento Diana-Molde (es el paso más importante).
ALINEAMIENTOS:
 - DE SECUENCIA (CLUSTALW, TCOFFEE)
 - ESTRUCTURALES (DALI, MAMMOTH)
 - 2.1 Identificar las regiones conservadas
 - 2.2 Identificar las regiones variables y predecir su estructura.
- 3- Construcción del modelo (existen diferentes aproximaciones)
Modelado de las cadenas laterales
- 4- Refinado y evaluación del modelo: PROCHECK, WHATIF

1- BÚSQUEDA DEL TEMPLADO O MOLDE

La primera etapa es determinar homólogos de la proteína problema.

Se intenta determinar a que “familia” pertenece la proteína. BLASTP, FASTA.

IMPORTANTE:

Conocer la proteína a la hora de seleccionar el molde (Swiss-Prot).

Sin templado NO hay modelo. Al menos una de las proteínas de la familia debe tener estructura conocida experimentalmente.

Se requiere que toda la secuencia del “target” éste representada en el templado o, al menos las secuencias de las regiones conservadas.

Se pueden utilizar más de un templado.

2- ALINEAMIENTO: PASO CRÍTICO EN HM

Lo óptimo es una identidad > 40%

Entre 25-40% zona gris

El alineamiento debe contener el menor número de GAPS
y que éstos se encuentren a ser posible en regiones variables.

Alineamientos múltiples mejores que los de a pares

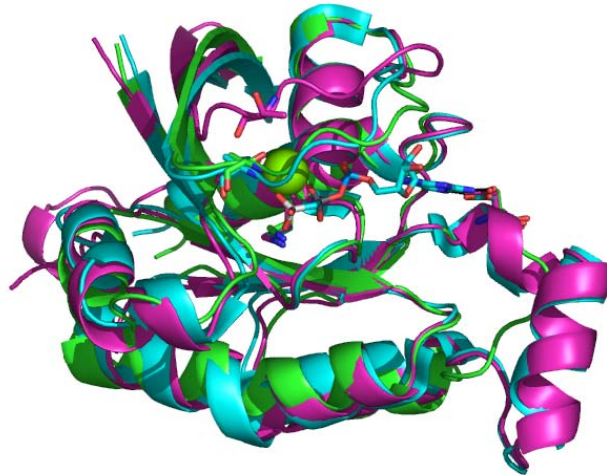
Pequeños errores en alineamiento grandes errores en el modelo

A veces incluir datos de alineamientos estructurales puede ser útil

EJEMPLO: Proteínas de la familia GTPasa

			Phosphate-binding Loop	Switch I		Switch II	
			B1 10 L1 20 A1 30 L2 40 B2 L3 50 B3 60 L4a H1 L4b 70				
RhoA	1	MAAIRKKLVIVGDGACGKTCLLVFSKDOFPEVYVPTVFENYVADIEVDGKQVELALWDTAGQEDYDRLR					
RhoB	1	MAAIRKKLVVVGDGACGKTCLLVFSKDEFFPEVYVPTVFENYVADIEVDGKQVELALWDTAGQEDYDRLR					
RhoC	1	MAAIRKKLVIVGDGACGKTCLLVFSKDOFPEVYVPTVFENYVADIEVDGKQVELALWDTAGQEDYDRLR					
Rac1	1	MQAIKCVVVGDAVGKTCLLISYTTNAPFCEYIPTVFENYVADIEVDGKQVELALWDTAGQEDYDRLR					
Rac2	1	MQAIKCVVVGDAVGKTCLLISYTTNAPFCEYIPTVFENYVADIEVDGKQVELALWDTAGQEDYDRLR					
Cdc42	1	MQTIKCVVVGDAVGKTCLLISYTTNAPFCEYIPTVFENYVADIEVDGKQVELALWDTAGQEDYDRLR					
K-Ras	1	MTEYKLVVVGAGGVGKSALTITQLQNHFFVDEYDPTIEDSYRKQVVIDGETCLLDILDLAGQEEYSAMR					
N-Ras	1	MTEYKLVVVGAGGVGKSALTITQLQNHFFVDEYDPTIEDSYRKQVVIDGETCLLDILDLAGQEEYSAMR					
H-Ras	1	MTEYKLVVVGAGGVGKSALTITQLQNHFFVDEYDPTIEDSYRKQVVIDGETCLLDILDLAGQEEYSAMR					
			B1 A1 B2 B3 H1				
			H2 L5 80 B4 L6 90 A3 100 L7 110 B5 L8a 120 H3 L8b A3' 130 L8c				
RhoA	71	PLSYPDTDVILMCFSDSPDSLENIPEKWTPEVKHFC-P-NVPIILVGNKDLRNDHETRELAKMKQEP					
RhoB	71	PLSYPDTDVILMCFSDSPDSLENIPEKWTPEVKHFC-P-NVPIILVGNKDLRNDHETRELAKMKQEP					
RhoC	71	PLSYPDTDVILMCFSDSPDSLENIPEKWTPEVKHFC-P-NVPIILVGNKDLRNDHETRELAKMKQEP					
Rac1	69	PLSYPDTDVILMCFSDSPDSLENIPEKWTPEVKHFC-P-NVPIILVGNKDLRNDHETRELAKMKQEP					
Rac2	69	PLSYPDTDVILMCFSDSPDSLENIPEKWTPEVKHFC-P-NVPIILVGNKDLRNDHETRELAKMKQEP					
Cdc42	69	PLSYPDTDVILMCFSDSPDSLENIPEKWTPEVKHFC-P-NVPIILVGNKDLRNDHETRELAKMKQEP					
K-Ras	69	DQYMRGTGEGFLCVFAINNTKSFEDIH-HYREQIKRVKSDSDVPMVLVGNKCDLPS-----RT					
N-Ras	69	DQYMRGTGEGFLCVFAINNTKSFEDIH-HYREQIKRVKSDSDVPMVLVGNKCDLPS-----RT					
H-Ras	69	DQYMRGTGEGFLCVFAINNTKSFEDIH-HYREQIKRVKSDSDVPMVLVGNKCDLPS-----RT					
			A2 B4 A3 B5				
			A4 140 150 L9 B6 160 L10 A5 170 180 190				
RhoA	139	VKPEEGRDMANRIGAFGYMECSAKTKDGVREVFEMATRAALQARRGKKKSG-----CLVL					
RhoB	139	VRTDDGRAMAVRIQAYDYLECSAKTKDGVREVFEMATRAALQARRGKKKSG-----CLVL					
RhoC	139	VSEEGRDMANRIGAFGYMECSAKTKDGVREVFEMATRAALQARRGKKKSG-----CLVL					
Rac1	137	ITYPQGLAMAKEIGAVKYLECSALTQGLKTVFDEAIRAVLCPPPVKKRRK-----CLIL					
Rac2	137	ITYPQGLAMAKEIGAVKYLECSALTQGLKTVFDEAIRAVLCPPPVKKRRK-----CLIL					
Cdc42	137	ITYPQGLAMAKEIGAVKYLECSALTQGLKTVFDEAIRAVLCPPPVKKRRK-----CLIL					
K-Ras	125	VDTKQAQDLARSYGI-PFIETSAKTRQGVDAFYTLVREIRQYRLKKISKKEKTPGCVKIKKCIIM					
N-Ras	125	VDTKQAQDLARSYGI-PFIETSAKTRQGVDAFYTLVREIRQYRLKKISKKEKTPGCVKIKKCIIM					
H-Ras	125	VESRQAQDLARSYGI-PFIETSAKTRQGVDAFYTLVREIRQYRLKKISKKEKTPGCVKIKKCIIM					
			A4 B6 A5				

Seq Id: **Ras-H/Rho-A**: 28%



Esther Marco y Federico Gago

BAJA IDENTIDAD DE SECUENCIA NO IMPLICA
NECESARIAMENTE BAJA HOMOLOGÍA
ESTRUCTURAL.

GFP/RFP: 25% ID

```
GFP  MGKGEELFTGVVFIIVELDGDVNGHKFSV
RFP  MRSSKNVKEFMRFKVRMEGTVNGHEFEI

GFP  SGEGEGDATYGLTLNFICTT.GKLPVPW
RFP  EGEGEGRPYEGHNTVVKLVTKGGPLPFAM

GFP  PTLVTTFSYGVQCFSRYPDHMKRHDFFKS
RFP  DILSPQFQVGSKVTVVHPADI..FDYKKL

GFP  AMPEGYVQERTIFFKDDGNYKTRAEVKFE
RFP  SFPEGFWKRVNMFEDGGVTVTQDSSLQ

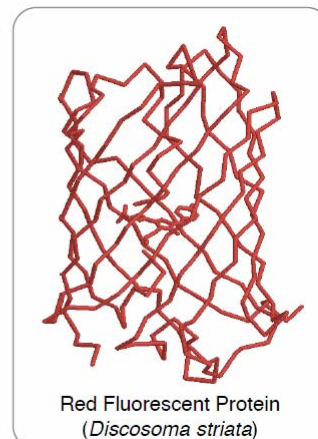
GFP  GDTLVNRIELKGIDFKEDGNILGHK.LEY
RFP  DGCIFYKVKFIGNVFPDGPVMQKTMGSH

GFP  NYNSHNVMADFKNGIKVNFKIRHNIE
RFP  EASTERLVPRDGLKGEIHKALKLK....

GFP  DGSVQLADHYQQNTFIDGGFVLLPDNHYL
RFP  DGGHYLVEFKSIY..MAKKFVQLFGYITV

GFP  STQSALSQDENKRDHMLLEFVTAAGIT
RFP  DSKLDITSH....NEDYTVIEQYERTEGR

GFP  HGMDELY
RFP  HHLF
```



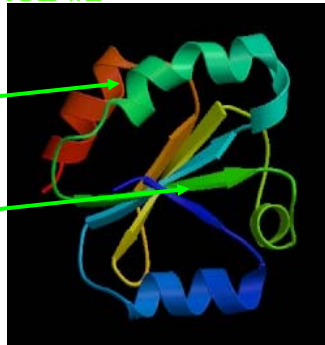
Búsqueda de regiones conservadas estructuralmente (RCE)

Target	ACDEF GH IKLMNPQRST--FGHQWERT---TYREWYEG
Tmpl.1	ASDEYAHLRILDPQRSTVAYAYE--KSFAPPGSFKWEYEA
Tmpl.2	MCDEYAHLRLMNP ER STVAGGHQWERT---GSFKEWYAA
	HHHHHHHHHHHHHHCCCCCCCCCCCCCCCCCCCCBBBBBBBBBB

RCE #1 RCE #2

Hélice RCE #1

Hoja B RCE #2



•Corresponden a las regiones estructuralmente mas estables de la proteína y a elementos de estructura secundaria.

•Altamente conservadas, y por lo general, sin GAPs.

Búsqueda de regiones variables estructuralmente (RVE)

Target	ACDEF GH IKLMNPQRST--FGHQWERT---TYREWYEG
Tmpl.1	ASDEYAHLRILDPQRSTVAYAYE--KSFAPPGSFKWEYEA
Tmpl.2	MCDEYAHLRLMNP ER STVAGGHQWERT---GSFKEWYAA
	HHHHHHHHHHHHHHCCCCCCCCCCCCCCCCCCCCBBBBBBBBBB

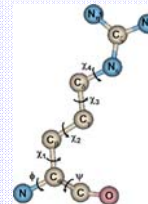
RVE (loop)

RVE (loop)



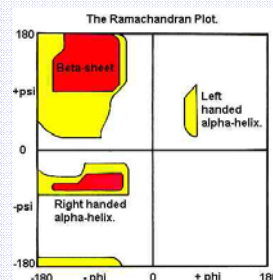
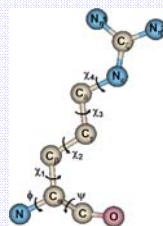
3- GENERAR COORDENADAS DEL MODELO (x,y,z)

- Usando las RCE como anclaje:
- Aminoácidos idénticos: transferir todas las coordenadas del mismo
- Aminoácidos similares: transferir "backbone", reemplazar cadena lateral respetando las torsiones (χ)
- Aminoácidos diferentes: transferir sólo el "backbone"
- Agregar loops de las RVE (bibliotecas de loops, ModLoop <http://alto.compbio.ucsf.edu/modloop/>)
- Modelado de las cadenas laterales MaxSprout <http://www.ebi.ac.uk/maxsprout/>



4- REFINADO Y EVALUACIÓN DEL MODELO

- Refinado del modelo: optimización de la geometría
- Evaluación del modelo
Los programas de evaluación verifican si el modelo tiene:
"Todo aquello que una buena estructura debe tener"
Una buena estructura posee (I):
 - Mínimo número de ángulos de torsión no permitidos (Ramachandran plot)



4- REFINADO Y EVALUACIÓN DEL MODELO

- Una buena estructura posee (y II):
 - Maximizado el empaquetamiento de residuos hidrofóbicos
 - Minimizada la superficie hidrofóbica expuesta a solvente
 - Maximizada la superficie hidrofílica expuesta a solvente
 - Maximizado número de enlaces de hidrógeno
 - Correcta estereoquímica
- La verificación última es ver si el modelo justifica los datos experimentales y si tiene capacidad predictiva

**UN MODELO NO ES CORRECTO O ERRÓNEO,
ES ÚTIL O INÚTIL**

Servidores de modelado por homología:

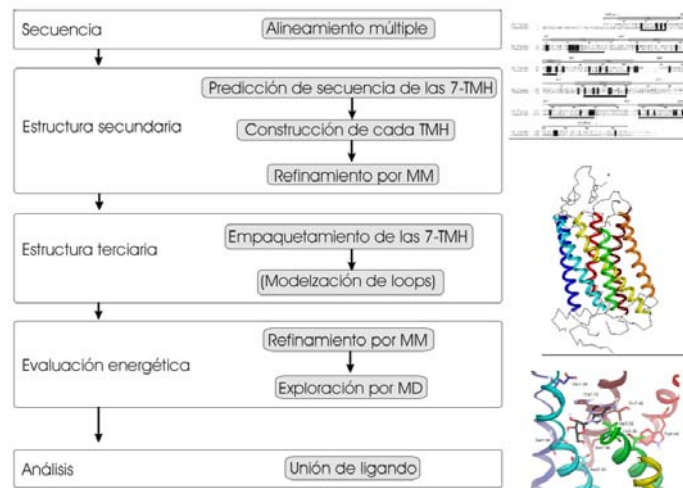
Swiss-Model	SWISS-MODEL. An Automated Comparative Protein Modelling Server (ExPASy, CH)
TITO/Modeller	Modeller at the CBS Bioinformatic Tools Server (Montpellier, France)
CPHmodels	The CPHmodels Server using comparative (homology) modelling (BioCentrum, Denmark)
SDSC1	"SDSC1" - SDSC Protein Structure Homology Modeling Server (San Diego, USA)
3D-JIGSAW	Automated system to build three-dimensional models for proteins. (Cancer Research UK)
Loops Database	A table of five protein loop classes. (Cancer Research UK).

Servidores de evaluación de modelos:

AQUA	AQUA - Alignment Quality Assessment
BIOTECH Validation Suite	Biotech Validation Suite for Protein Structures (PROCHECK, PROVE & WHAT IF)
WHATIF	WHAT IF Web interface: homology modelling, drug docking, electrostatics calculations, structure validation and visualisation.
3D Evaluation	Verify3D & PROSA II at the CBS Bioinformatic Tools Server (Montpellier, France)

III- Modelado de proteínas:

4. Modelado por topología de proteínas de transmembrana.



Modelado de Proteínas

I- Introducción.

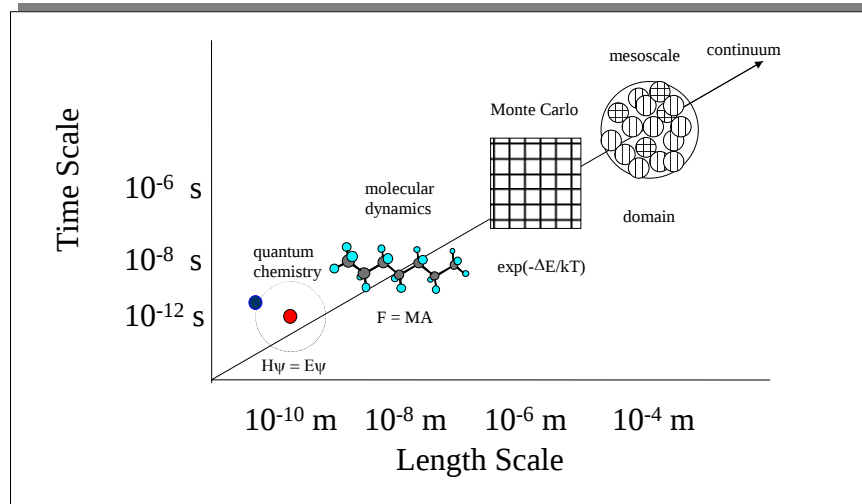
II- Estructura experimental de proteínas.

III- Modelado de proteínas.

IV- Metodología del modelado de proteínas:

1. Dificultades.
2. Química Cuántica.
3. Mecánica Molecular.
4. Dinámica Molecular.
5. Métodos híbridos (QM/MM).
5. Potenciales de Interacción Molecular.

What and Where: Scales in Simulation



IV- Metodología del modelado de proteínas:

1. Dificultades:

-Cálculo de la energía → derivar propiedades

-Optimización de la geometría:

Métodos basados en el gradiente de energía (derivadas de la E con respecto a los grados de libertad geométricos):

- Steepest descent
- Conjugated gradients
- Newton-Raphson

-Simulación de la variación a lo largo del tiempo: Dinámica Molecular.

Modelado de Proteínas

I- Introducción.

II- Estructura experimental de proteínas.

III- Modelado de proteínas.

IV- Metodología del modelado de proteínas:

1. Dificultades.
2. Química Cuántica.
- 3. Mecánica Molecular.**
4. Dinámica Molecular.
5. Métodos híbridos (QM/MM).
5. Potenciales de Interacción Molecular.

3. Mecánica Molecular.

The "mechanical" molecular model was developed out of a need to describe **molecular structures** and **properties** in as practical a manner as possible.

Range of applicability:

- Molecules containing thousands of atoms.
- Organics, oligonucleotides, peptides, and saccharides.
- Vacuum, implicit, or explicit solvent environments.
- Ground state only.
- Thermodynamic and kinetic (via molecular dynamics) properties.

Molecular mechanics methods are based on the following principles:

- Nuclei and electrons are lumped into atom-like particles.
- Atom-like particles are spherical (radii obtained from measurements or theory) and have a net charge (obtained from theory).
- Interactions are based on springs and classical potentials.
- Interactions must be preassigned to specific sets of atoms.
- Interactions determine the spatial distribution of atom-like particles and their energies.

Note how these principles differ from those of Quantum Mechanics!!

Quantum Mechanics methods are based on the following principles:

- Nuclei and electrons are distinguished from each other.
- Electron-electron (usually averaged) and electron-nuclear interactions are explicit.
- Interactions are governed by nuclear and electron charges (i.e. potential energy) and electron motions.
- Interactions determine the spatial distribution of nuclei and electrons and their energies.

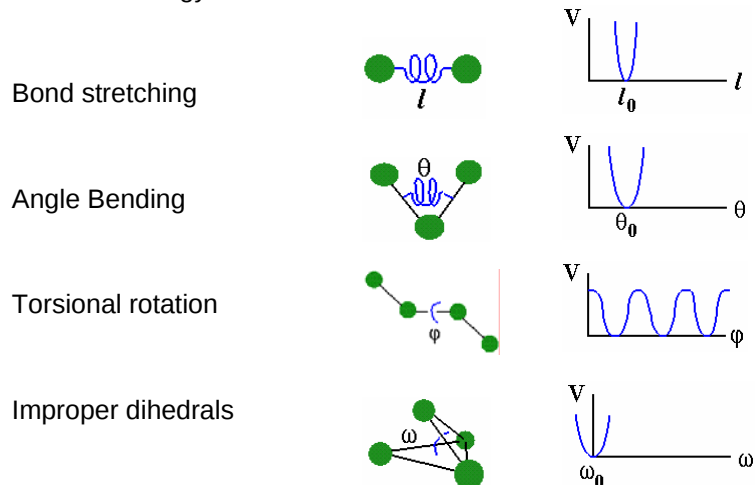
The Anatomy of a Molecular Mechanics Force-Field

- The mechanical molecular model: atoms as spheres, and bonds as springs.
- The mathematics of spring deformation can be used to describe the ability of bonds to stretch, bend, and twist.
- Non-bonded atoms (greater than two bonds apart) interact through van der Waals attraction, steric repulsion, and electrostatic attraction/repulsion.
- These properties are easiest to describe mathematically when atoms are considered as spheres of characteristic radii.
- MM Energies: no meaning as absolute quantities.
Differences in energy between two or more conformations: meaning.
- A simple molecular mechanics energy equation is given by:
Energy = Stretching E + Bending E + Torsion E + Non-Bonded Interaction E

These equations together with the data (parameters) required to describe the behaviour of different kinds of atoms and bonds, is called a **force-field**.

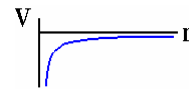
How do we describe the potential energy $V(\mathbf{x})$ for a molecule?

Potential Energy includes terms for

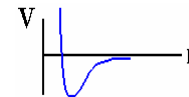
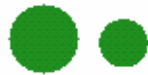


Potential energy includes terms for (contd.)

Electrostatic Interactions



van der Waals Interactions



Equation for covalent terms in P.E.

$$V_{bonded}(\vec{R}) = \sum_{bonds} k_l (l - l_0)^2 + \sum_{angles} k_\theta (\theta - \theta_0)^2 + \sum_{impropers} k_\omega (\omega - \omega_0)^2 + \sum_{torsions} A_n [1 + \cos(n\phi - \phi_0)]$$

Equation for non-bonded terms in P.E.

$$V_{nonbonded}(\vec{R}) = \sum_{i < j} \left(\epsilon_{ij} \left[\left(\frac{r_{ij}^{min}}{r_{ij}} \right)^{12} - 2 \left(\frac{r_{ij}^{min}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_r \epsilon_0 r_{ij}} \right)$$

van der Waals term
(potencial de Lennard-Jones)

Electrostatic term
(ley de Coulomb)

Modelado de Proteínas

I- Introducción.

II- Estructura experimental de proteínas.

III- Modelado de proteínas.

IV- Metodología del modelado de proteínas:

1. Dificultades.
2. Química Cuántica.
3. Mecánica Molecular.
- 4. Dinámica Molecular.**
5. Métodos híbridos (QM/MM).
5. Potenciales de Interacción Molecular.

Molecular Dynamics

- At physiological conditions, the biomolecules undergo several movements and changes
- The time-scales of the motions are diverse, ranging from few femtoseconds to few seconds
- These motions are crucial for the function of the biomolecules

Newton's second law of motion

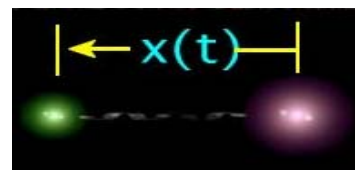
$$-\frac{dV}{dx} = F = ma = m \frac{d^2x}{dt^2}$$

Potential energy

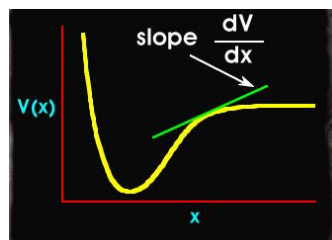
Position

We need to know

The motion of the
atoms in a molecule, $x(t)$



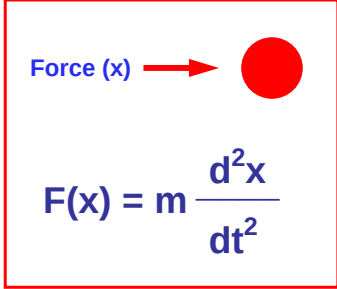
and therefore,



the potential energy, $V(x)$

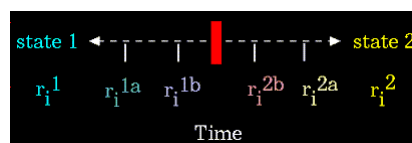
- Each of these interactions exerts a force onto a given atom of the molecule
- The total resulting force on each atom is calculated using the PE function

Knowing the force on an atom, its movement due to the force is then calculated:



$$F(x) = m \frac{d^2x}{dt^2}$$

For the "i" atoms of the system:



Newton's equation is used in the molecular dynamics formalism to simulate atomic motion:

$$\text{force} = \text{mass} \times \text{acceleration} (F_i = m_i a_i)$$

The force on an atom can be calculated from the change in energy between its current position and its position a small distance away

$$-\frac{dE}{dr_i} = F_i$$

Trajectory: the resulting series of snapshots of structural changes over time.

$$-\frac{dE}{dr_i} = m_i \frac{d^2 r_i}{dt^2}$$

In practice, trajectories are not directly obtained from Newton's equation. First, the atomic accelerations are computed from the forces and masses. The velocities are next calculated from the accelerations based on:

$$a_i = \frac{dv_i}{dt}$$

Lastly, the positions are calculated from the velocities:

$$v_i = \frac{dr_i}{dt}$$

Trajectory between two states: subdivided into a series of sub-states separated by a small time step, "Δt" (e.g. 1 femtosecond):

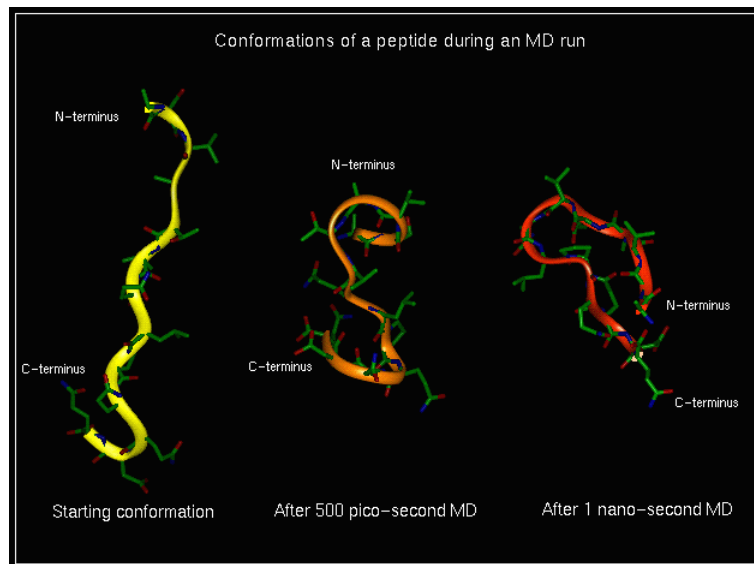


Steps summarized:

1	solve for a_i at t using:	$-\frac{dE}{dr_i} = F_i = m_i a_i(t)$
2	update v_i at $t + \Delta t/2$ using:	$v_i(t + \Delta t/2) = v_i(t - \Delta t/2) + a_i(t) \Delta t$
3	update r_i at $t + \Delta t$ using:	$r_i(t + \Delta t) = r_i(t) + v_i(t + \Delta t/2) \Delta t$

**Generalizing these ideas,
the trajectories for all the atoms of a molecule can
be calculated.**

How a molecule changes during MD



The Necessary Ingredients

- Description of the structure: atoms and connectivity
- Initial structure: geometry of the system
- Potential Energy Function: force field
 - AMBER
 - CVFF
 - CFF95
 - Universal

Protein-specific Applications of MD

- Calculation of thermodynamic properties such as internal energy, free energy
- Studying the protein folding / unfolding process
- Studying conformational properties and transitions due to environmental conditions
- Studying conformational distributions in molecular system.

An overview of various motions in proteins (I)

Motion	Spatial extent (nm)	Log ₁₀ of characteristic time (s)
Relative vibration of bonded atoms	0.2 to 0.5	-14 to -13
Elastic vibration of globular region	1 to 2	-12 to -11
Rotation of side chains at surface	0.5 to 1	-11 to -10
Torsional libration of buried groups	0.5 to 1	-14 to -13

An overview of various motions in proteins (II)

Motion	Spatial Extent (nm)	Log ₁₀ of characteristic time (s)
Relative motion of different globular regions (hinge bending)	1 to 2	-11 to -7
Rotation of medium-sized side chains in interior	0.5	-4 to 0
Allosteric transitions	0.5 to 4	-5 to 0
Local denaturation	0.5 to 1	-5 to 1
Protein folding	???	-5 to 2

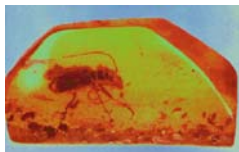
A typical MD simulation protocol

- Initial random structure generation
- Initial energy minimization
- Equilibration
- Dynamics run – with capture of conformations at regular intervals
- Energy minimization of each captured conformation

Essential Parameters for MD (to be set by user)

- Temperature
- Pressure
- Time step
- Dielectric constant
- Force field
- Durations of equilibration and MD run
- pH effect (addition of ions)

WHAT IS AMBER?



(Ugly bug from an earlier version)

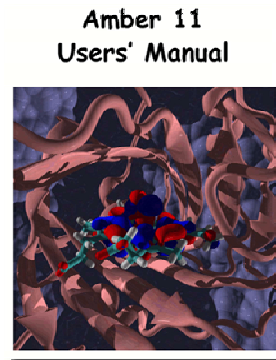


"The bugs have magically disappeared!"

- AMBER (Assisted Model Building with Energy Refinement).
- Allows users to carry out molecular dynamics simulations
- Updated forcefield for proteins and nucleic acids
- Parallelized dynamics codes
- Ewald sum periodicity
- New graphical and text-based tools for building molecules
- Powerful tools for NMR spectral simulations
- New dynamics and free energy program

WHY AMBER?

- Most widely used program: approximately 5000 users world over.
- Over 1000 research papers have been published using AMBER.
- Program available at a nominal price for academic users.
- Complete source code available with the package.
- Available for most machine configurations.
- Developed by Prof. Peter Kollman
Department of Pharmaceutical Chemistry
University of California San Francisco



BASIC INFORMATION FLOW IN AMBER

