

I metodi della Biologia Strutturale

Laurea Magistrale in Biotecnologie Mediche
Curriculum Nanobiotecnologie

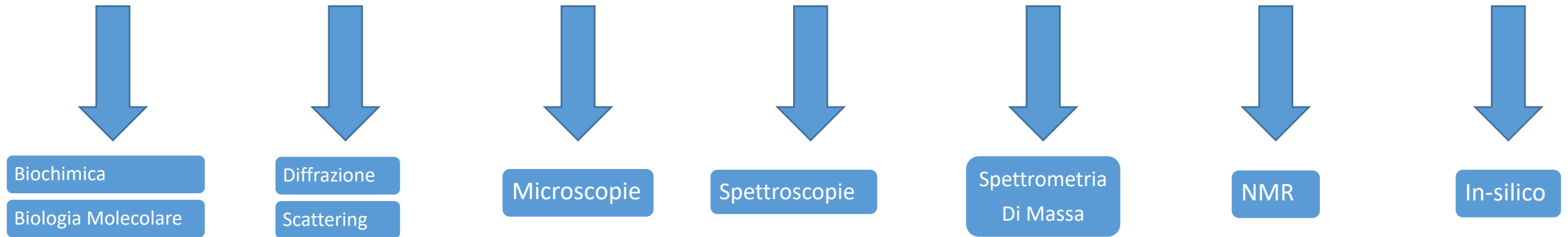
A.A. 2022-23

Toolbox del biologo strutturale

Il Toolbox del Biologo Strutturale

Il fine ultimo in uno studio di biologia strutturale è quello di acquisire le informazioni sulla struttura molecolare della macromolecola, in relazione al problema biologico in cui essa è coinvolta.

In questo senso il biologo strutturale ha a disposizione diversi strumenti per poter raggiungere lo scopo prefissato



Utilità dei vari metodi

Informazioni non-
strutturale

- Metodi Biochimici
- Biologia Molecolare
- Spettroscopie (assorbimento, Fluorescenza, FRET)
- Metodi Biofisici (Calorimetrie, SPR, Spettroscopia di Massa)
- NMR

Informazione
strutturale locale
(limitata)

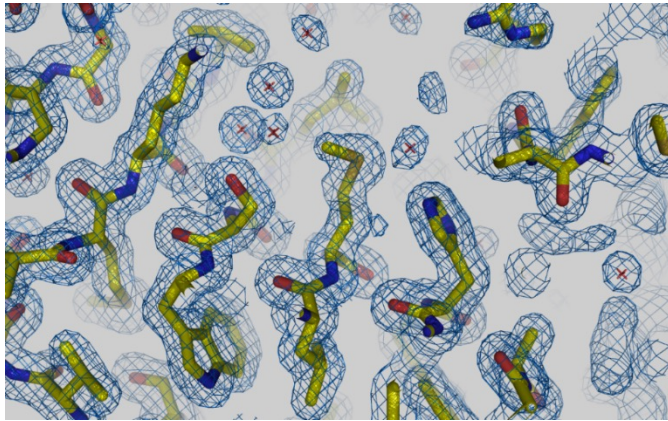
- Spettroscopie (EXAFS, Dicroismo circolare, FRET)
- Scattering (SAXS, Diffrazione da fibre)
- NMR
- Spettrometria di Massa (HDX, fingerprinting)

Informazione
Strutturale
Completa

- Diffrazione di raggi X (cristallo singolo)
- Cryo-Electron Microscopy
- NMR
- Metodi computazionali

Scegliere il metodo più opportuno

Abbiamo visto che esistono vari approcci allo studio della struttura delle macromolecole.
I metodi scelti devono essere i più adatti alle finalità dello studio intrapreso.



Informazione necessaria:

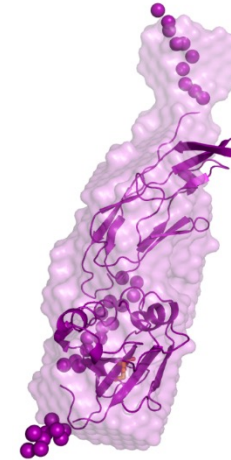
Struttura molecolare a livello atomico

Metodi Utili:

Cristallografia di raggi-X

NMR

CryoEM



Informazione necessaria:

Forma della molecola

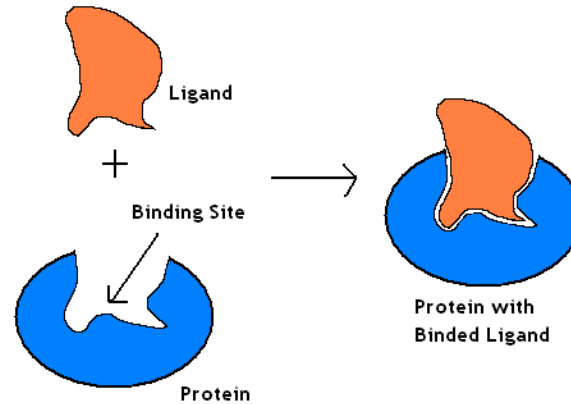
Metodi Utili:

SAXS

CryoEM (bassa risoluzione)

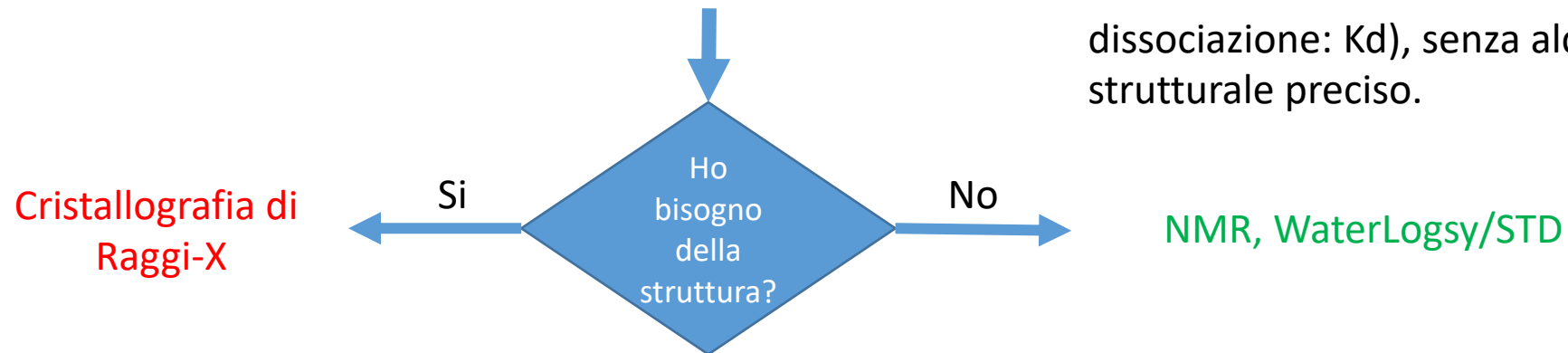
Esempio: Binding di un ligando

Lo studio dell'interazione tra una proteina e un suo ligando è uno degli studi più frequenti in biologia strutturale



Molto spesso è richiesta la determinazione della struttura molecolare del complesso a risoluzione atomica.

Altre volte basta l'informazione relativa ai soli residui coinvolti nell'interazione, o sull'efficacia dell'interazione (costante di dissociazione: K_d), senza alcun dettaglio strutturale preciso.



Metodi Strumentali

Spettroscopie:

- UV-Visibile
 - Assorbimento
 - fluorescenza
- Infrarosso
- Dicroismo Circolare
- EXAFS/XANES
- Fotoemissione
- **Risonanza Magnetica Nucleare (NMR)**

Spettrometrie di Massa

Metodi Calorimetrici

- ITC
- DSC

Scattering:

- **Diffrazione di raggi-X**
- Diffrazione di elettroni
- Diffrazione di neutroni
- Small Angle X-ray scattering
- Wide Angle X-ray scattering
- Dynamic Light Scattering

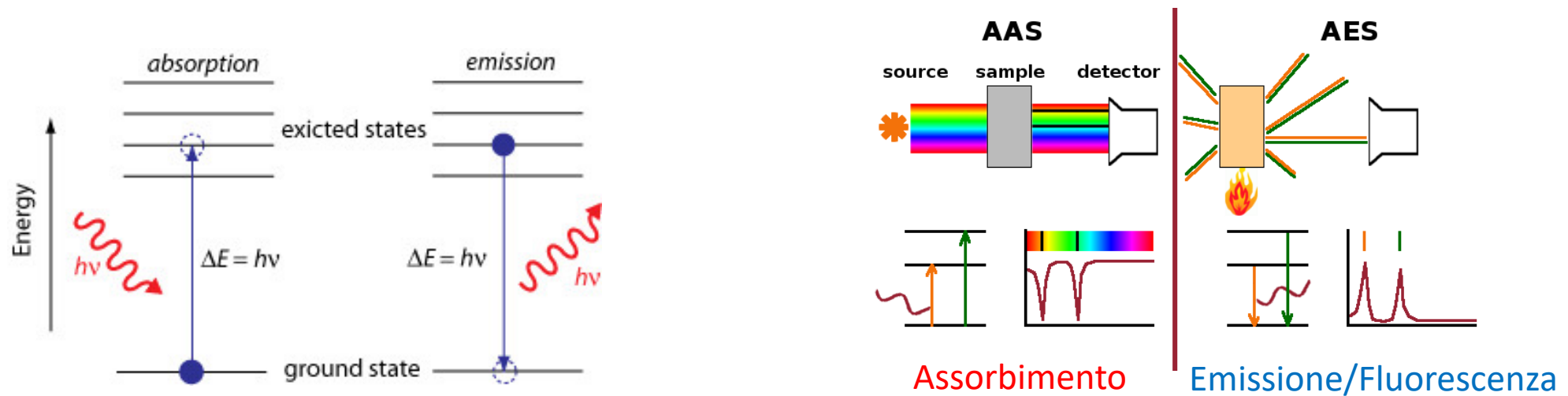
Microspie:

Cryo-Electron Microscopy (CryoEM)

Surface Plasmon Resonance (SPR)

Spettroscopie

Sono tecniche di indagine basate sull'assorbimento di una radiazione elettromagnetica da parte di molecole, atomi o nuclei in relazione ad un fenomeno fisico preciso (transizione da un livello energetico inferiore ad uno di energia superiore), dipendente dalla lunghezza d'onda λ della radiazione. La misura della quantità di radiazione assorbita (assorbimento) o emessa successivamente (emissione/fluorescenza) può fornire informazioni di natura strutturale sull'sistema studiato.



Il principio è assolutamente generale, e la radiazione elettromagnetica usata può andare dalle onde radio all'Infrarosso per arrivare ai raggi-X.

Spettroscopie in Biologia Strutturale

Esistono molte spettroscopie, tra loro diverse poiché basate su fenomeni fisici differenti.

Diverse spettroscopie possono essere utili in Biologia Strutturale: assorbimento nell'UV-Visibile, Infrarosso, Raman, Spettroscopia di fotoemissione, Dicroismo Circolare, EXAFS.

Queste tecniche coinvolgono transizioni elettroniche (o vibrazionali, come l'Infrarosso).

Altre tecniche coinvolgono invece transizioni di tipo diverso, come l'**NMR** (transizioni di spin nucleare).

Tra le tecniche più utilizzate :

- **Dicroismo Circolare**
- **FRET**
- **EXAFS**

Risonanza Magnetica Nucleare (NMR)

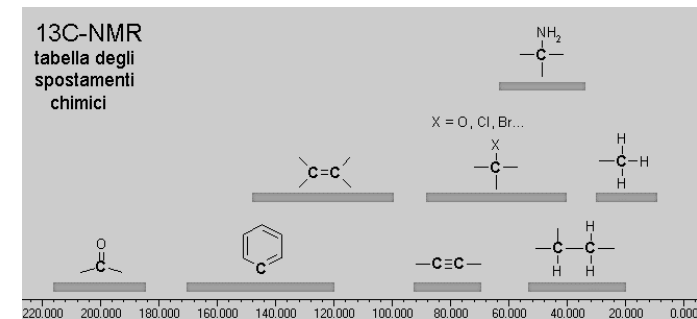
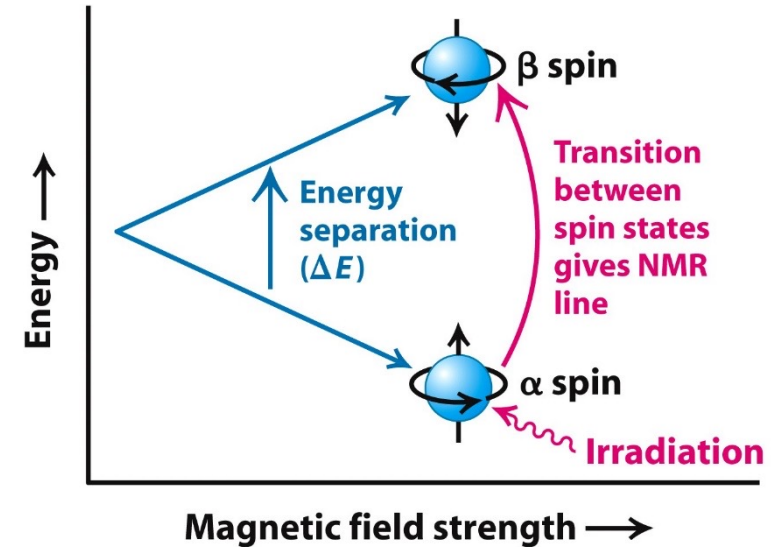
NMR

Nella spettroscopia NMR è coinvolto lo **spin nucleare** e la sua interazione con un **campo magnetico esterno**.

Il momento magnetico nucleare di nuclei con spin $\frac{1}{2}$ (^1H , ^{13}C , ^{15}N , ^{19}F , ^{31}P) si orienta in presenza di un campo magnetico esterno definendo così due stati a spin opposto di diversa energia.

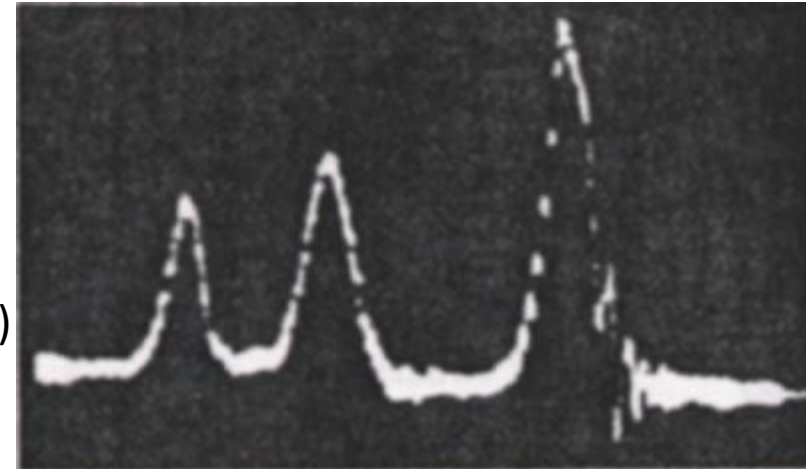
Applicando un campo elettromagnetico è possibile indurre la transizione dallo stato a più bassa energia verso lo stato a più alta energia, questa transizione è alla base del NMR.

I nuclei risentono del loro intorno chimico per cui nuclei di atomi appartenenti a gruppi chimicamente diversi (alifatici, aromatici...) avranno frequenze di assorbimento diverse.



Breve Storia dell'NMR

- Il fenomeno di Risonanza Magnetica Nucleare fu descritto per primo da Isidore Raabi nel 1938
- Felix Bloch e Edward Purcell, da ricerche indipendenti, dimostrarono il fenomeno dell'NMR nello stato condensato nel 1946
- Il primo spettro NMR (etanolo) fu acquisito nel 1951
- Nel 1952 Varian mette in commercio il primo spettrometro NMR (30 MHz)
- Nel 1953 Overhauser descrive l'effetto che porterà il suo nome
- Nel 1957 primi spettri della Bovine Pancreatic trypsin Inhibitor (BPTI)
- Nel 1976 Introduzione della tecnica 2D NMR
- Nel 1985 Kurt Wutrich pubblica la prima assegnazione completa della BPTI



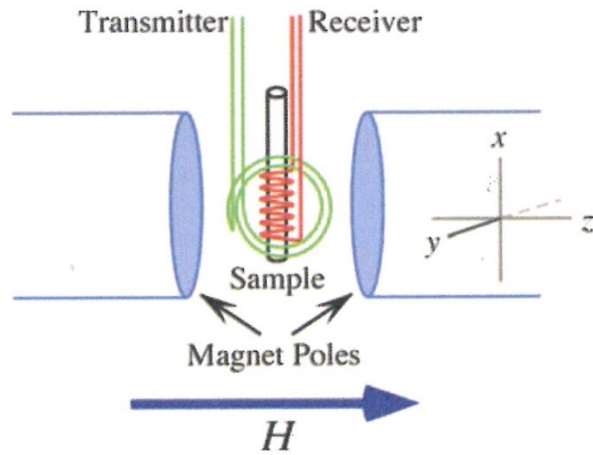
Condizioni sperimentali

- Concentrazione: 0.5 – 2 mM
- Volume: 100 – 500 μL
- Purezza: > 90%
- Stato di Aggregazione: monodisperso
- Buffers: Deuterati e concentrazione 10-20 mM
- pH: preferibile non > 7.5 (scambio protoni)

- Per spettri ^{13}C e ^{15}N è richiesta la marcatura con gli isotopi opportuni

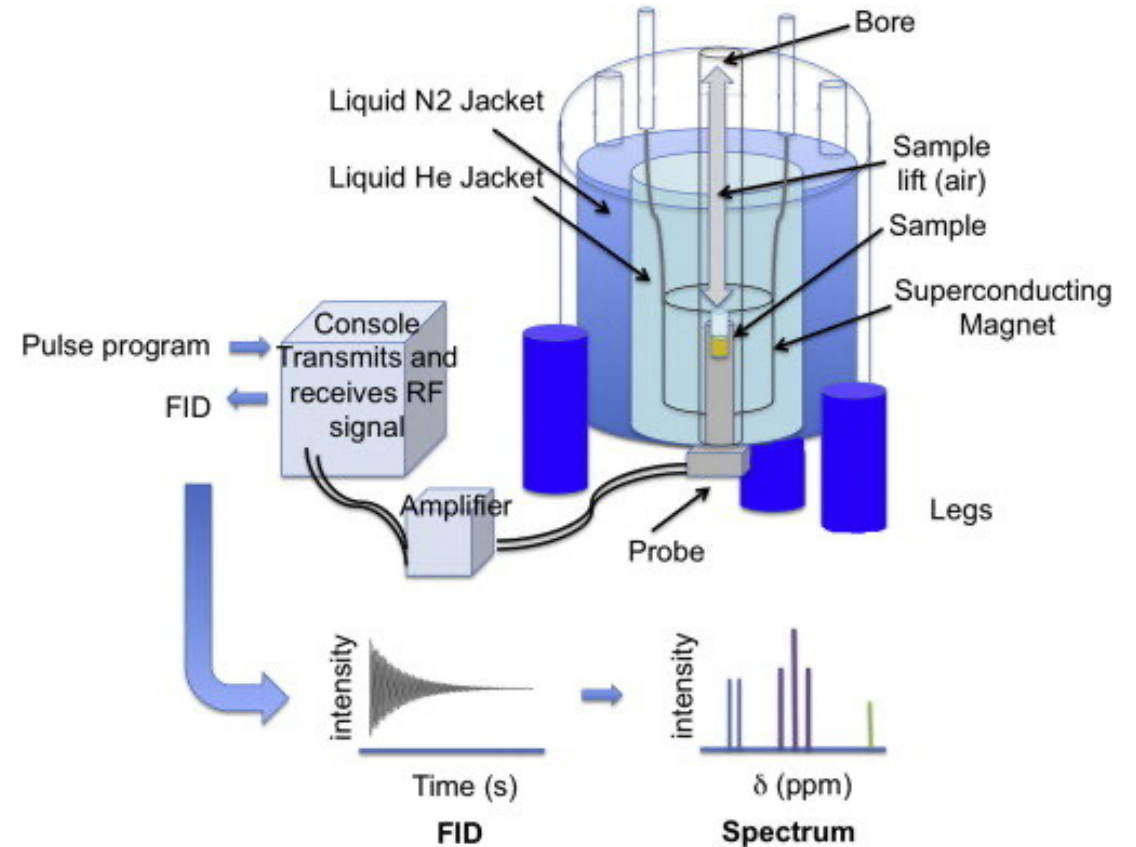
- MW: non maggiore di 50 kDa

Apparato Sperimentale



L'apparato sperimentale si compone di:

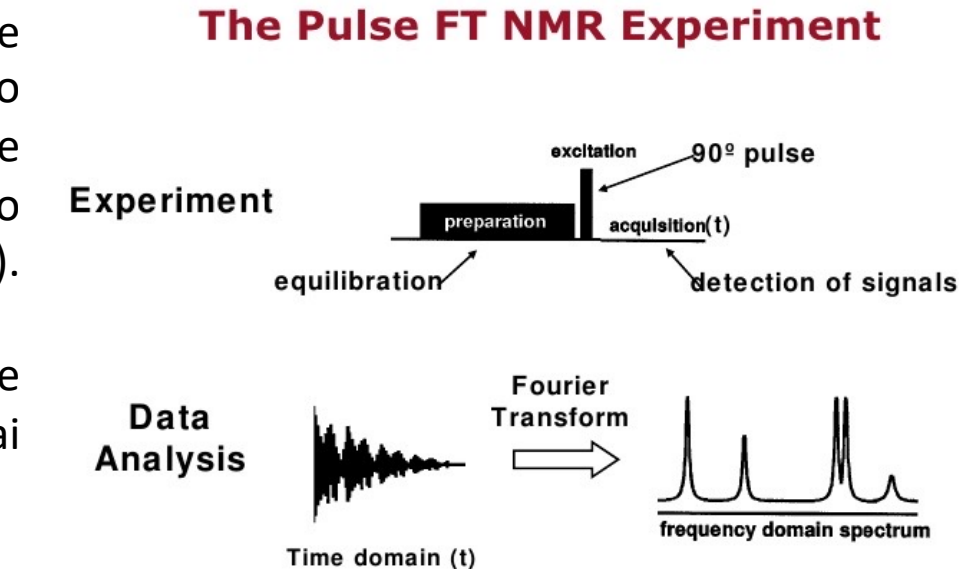
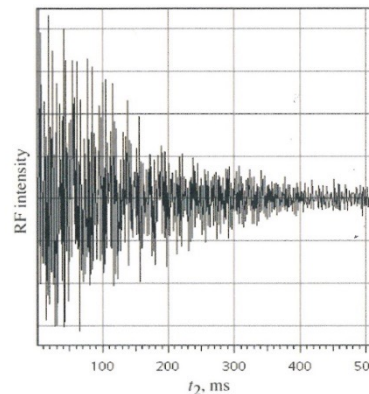
- Magnete superconduttore
- Trasmettitore (di impulsi)
- Ricevitore



L'esperimento NMR

L'esperimento, nella sua forma più semplice unidimensionale (1D) ^1H , consiste nell'inviare un impulso che ecciti tutti i nuclei, e quindi registrare il segnale (radiazione elettromagnetica) prodotto dal rilassamento dei nuclei eccitati (Free Induction Decay: FID). Rilassandosi i nuclei tornano allo stato iniziale. Utilizzando la Trasformata di Fourier è quindi possibile avere uno spettro con in frequenza con i segnali relativi ai singoli nuclei.

Il rilassamento (passaggio dallo stato eccitato a quello a più bassa energia) dei nuclei avviene sulla scala dei tempi dei ms



Free Induction Decay (FID)

Il chemical shift

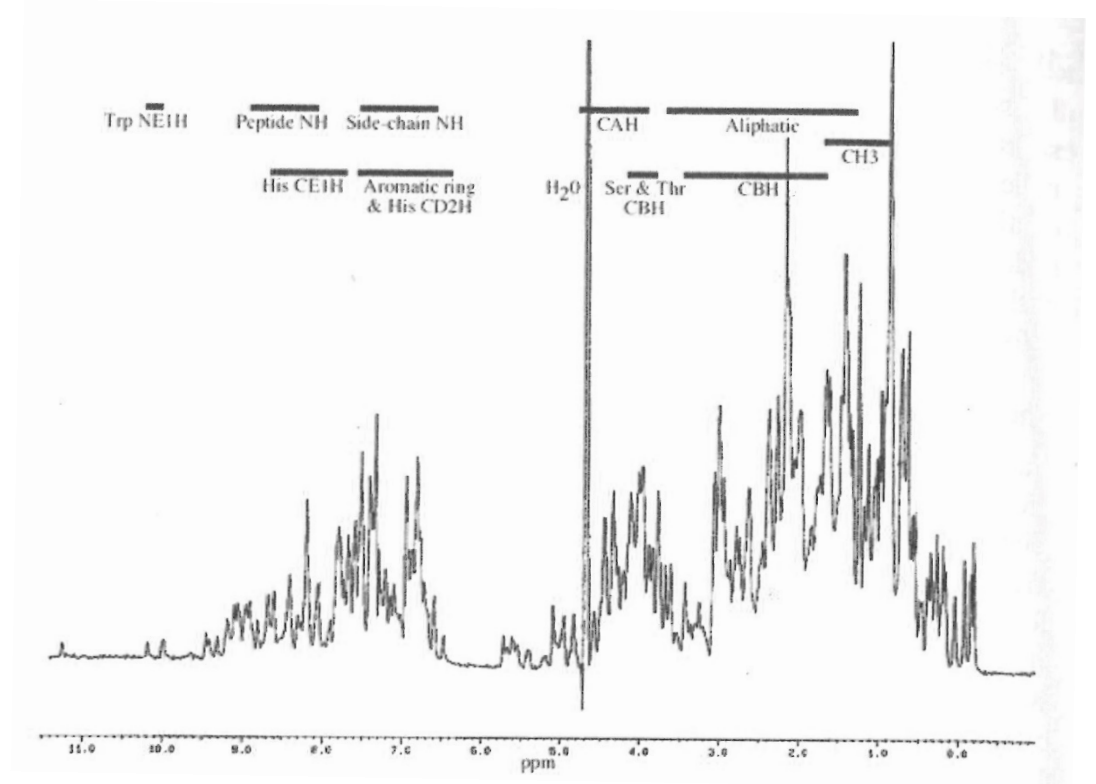
Nuclei diversi risentono dell'intorno chimico e di conseguenza *risuoneranno* (assorbiranno) a frequenze diverse

chemical shift
$$\delta(ppm) = 10^6 \frac{\nu - \nu_0}{\nu_0}$$

ν_0 è la frequenza di assorbimento (in Hz) di uno standard di riferimento quale il tetrametilsilano (TMS) per le molecole organiche e il Trimetilsililpropan solfonato (DSS) per le proteine)

L'uso di δ rende il **chemical shift** indipendente dall'intensità del campo magnetico applicato

Nelle proteine, i diversi protoni e atomi di carbonio hanno chemical shifts dipendenti dal tipo di residuo.



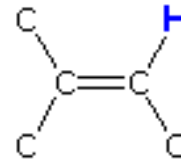
Accoppiamento scalare (J-coupling)

Il *J-coupling* o *accoppiamento scalare* è una forma di interazione che si trasmette attraverso i legami chimici (elettroni), tra nuclei 'vicini' (non più di 4 legami).

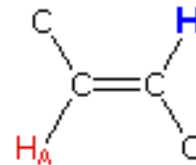
Questo accoppiamento causa uno 'sdoppiamento' dei picchi con chemical shift apparentemente simili.

L'effetto è tanto più forte quanto più i nuclei sono vicini.

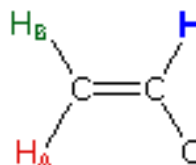
No Coupled Hydrogens



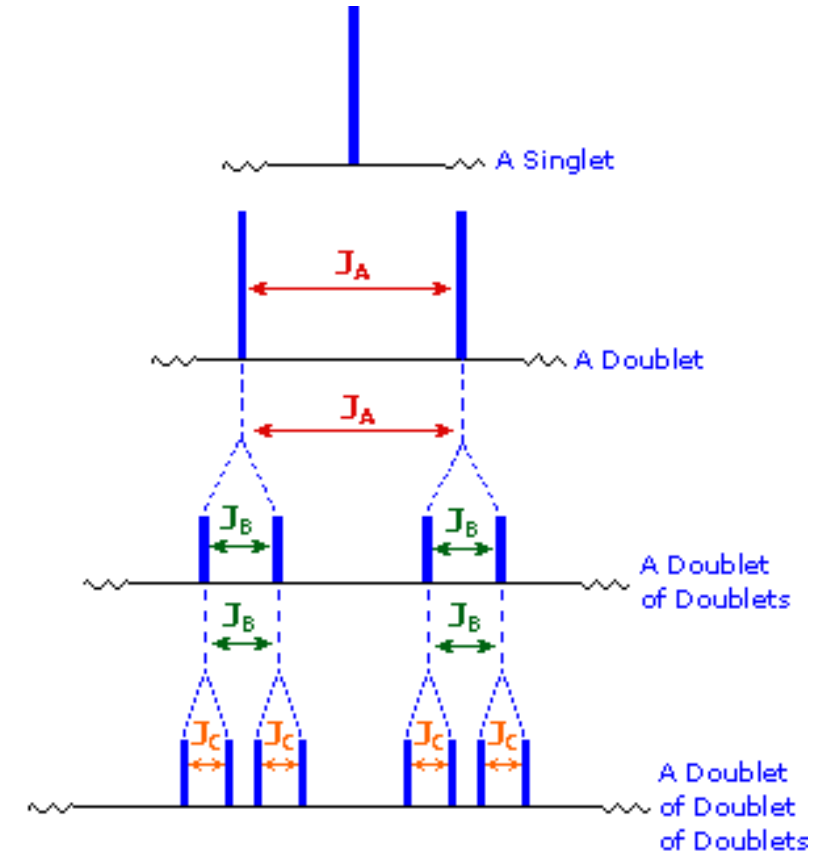
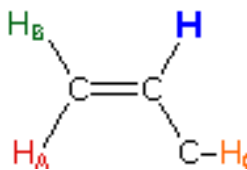
One Coupled Hydrogen



Two Coupled Hydrogens

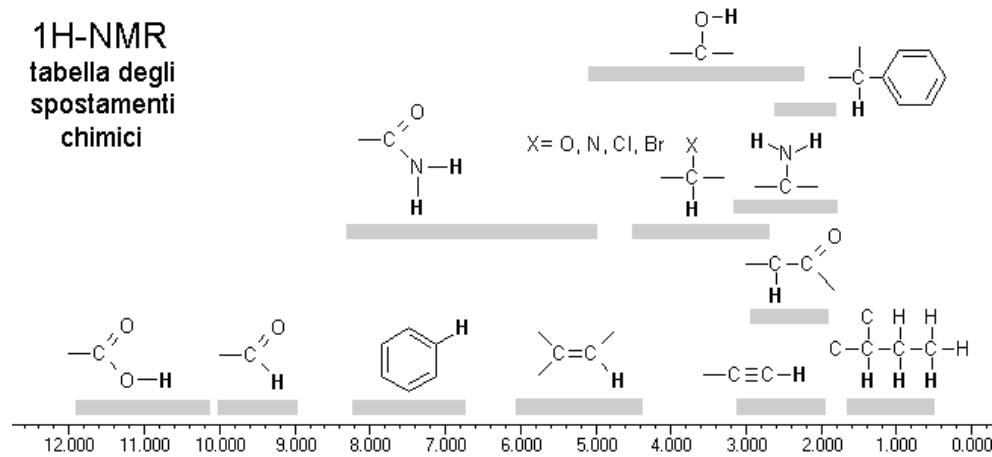


Three Coupled Hydrogens

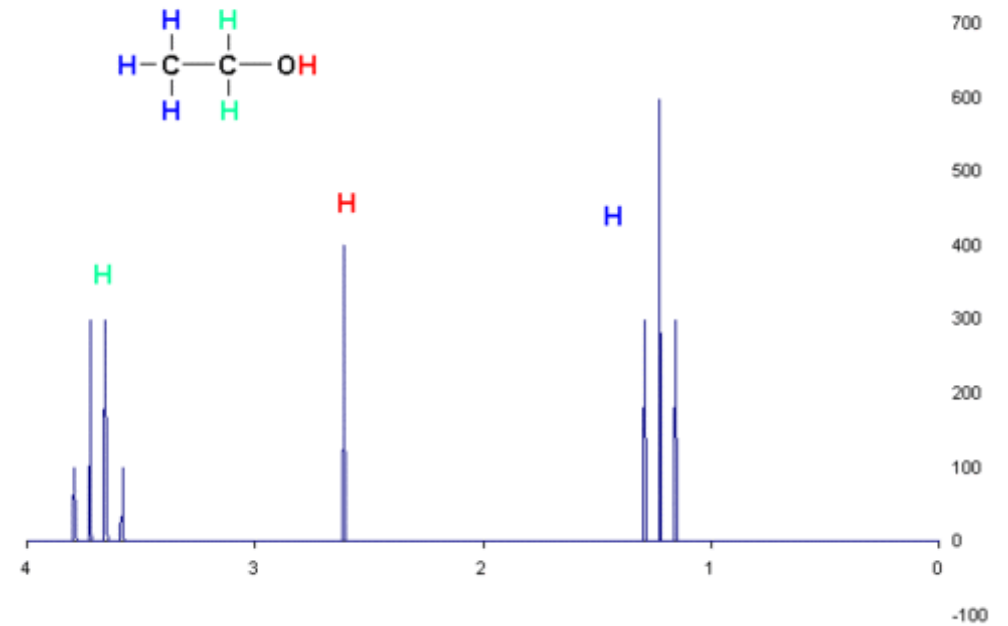


Identificazione delle molecole

$^1\text{H-NMR}$
tabella degli
spostamenti
chimici



Ethanol

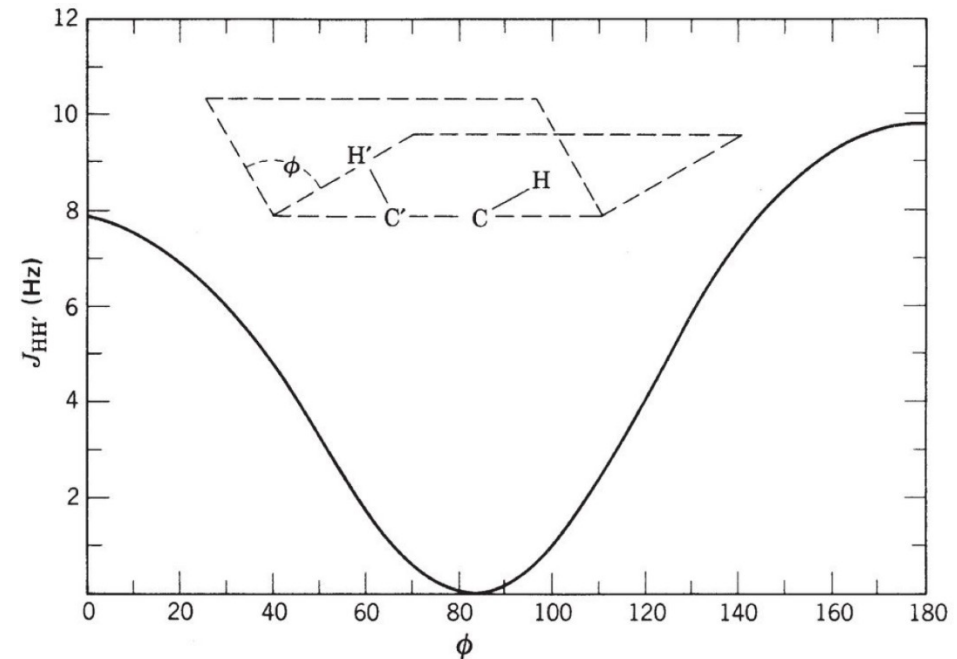


J-coupling e conformazione

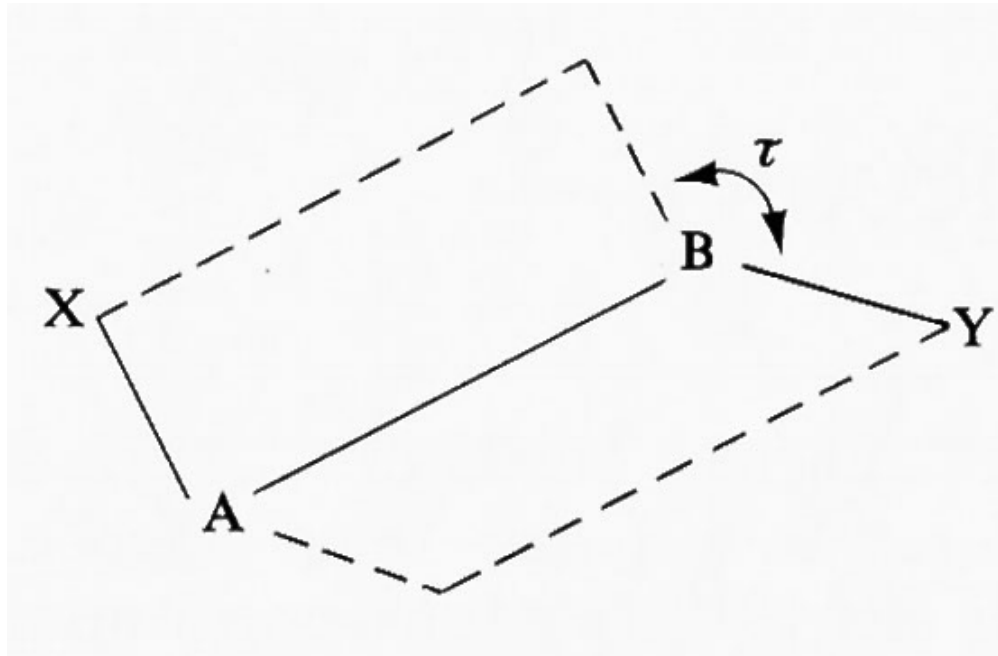
$$J_{HH'}(\phi) = A + B \cos(\phi) + C \cos(2\phi)$$

Questa equazione è nota come *equazione di Karplus* e stabilisce una relazione tra la costante di accoppiamento tra protoni legati ad atomi vicini ($J_{HH'}$) e il loro angolo diedro ϕ

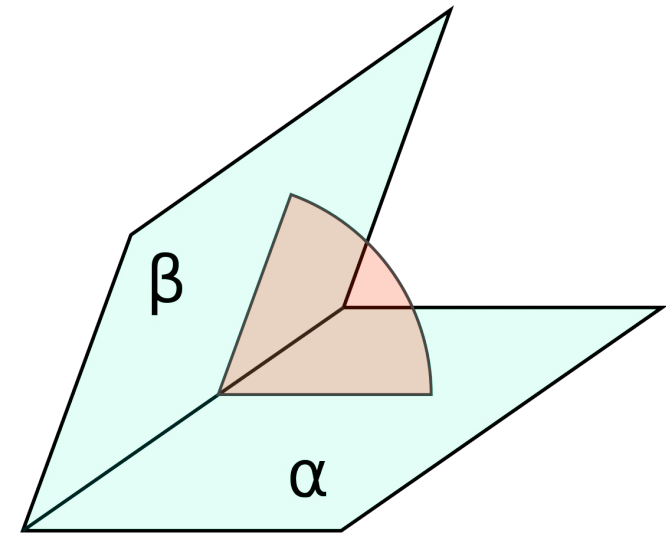
Questa relazione è molto importante perché fornisce informazioni sulla conformazione della molecola.



Angolo diedro tra legami

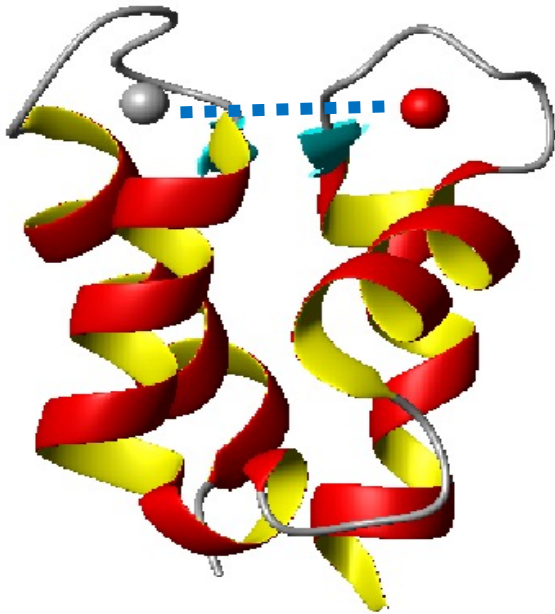


Angolo diedro (o di torsione) tra gli atomi X-A-B-Y



Effetto Nucleare Overhauser

A differenza del J-coupling, l'**effetto nucleare Overhauser (NOE)** noto anche come **accoppiamento dipolare**, si trasmette nello spazio (non attraverso gli elettroni di legame). E' un **trasferimento di magnetizzazione tra nuclei spazialmente vicini**.



Il NOE è quindi importantissimo per la determinazione del folding (struttura terziaria)

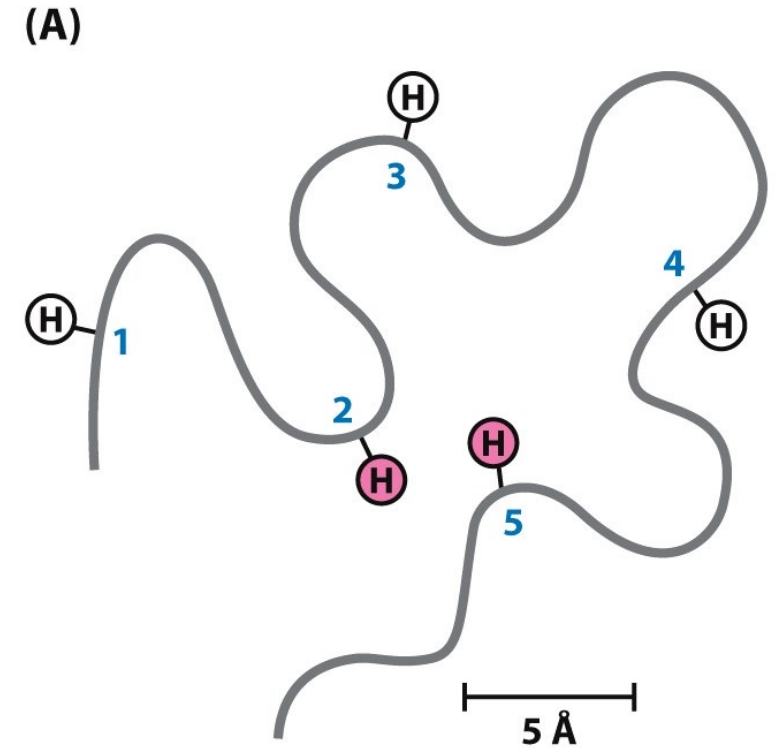


Figure 3.46
Biochemistry, Seventh Edition
© 2012 W. H. Freeman and Company

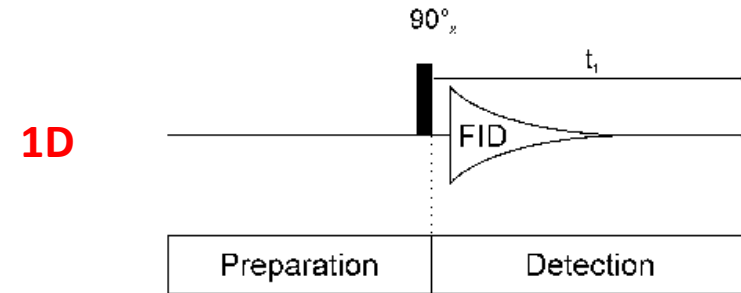
Sequenze di Impulsi in NMR

L'esperimento NMR può essere di una certa complessità.

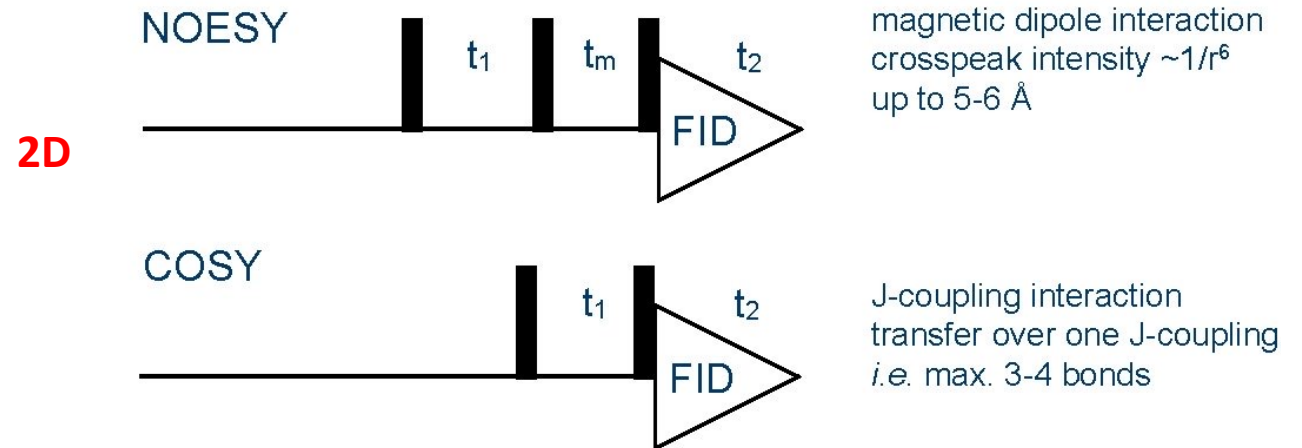
Selezionando:

- Il numero di impulsi (dimensione)
- La durata di ogni impulso (angolo)
- Il tempo tra un impulso e il successivo (Evoluzione o Mixing)

Si possono generare spettri diversi con proprietà diverse.



Homonuclear ¹H NMR

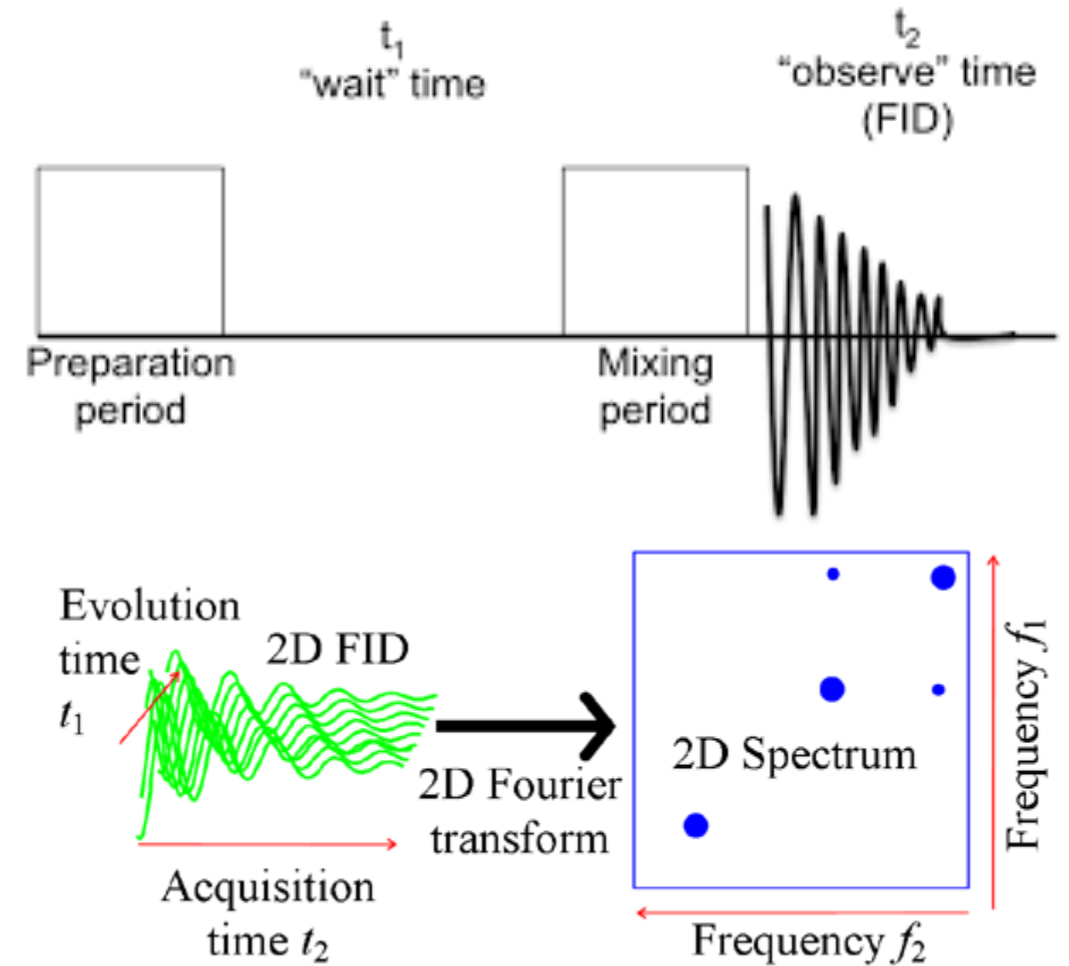


Tecniche bidimensionali in NMR

Le tecniche multidimensionali permettono l'acquisizione in tempi rapidi degli accoppiamenti tra nuclei (sia di natura scalare che dipolare).

Lo schema generale è il seguente:

- 1) Preparazione (primo impulso)
- 2) Evoluzione (il sistema si rilassa per un tempo t_1)
- 3) Mixing (secondo impulso)
- 4) Rilevamento (per un tempo t_2)
- 5) Una volta riequilibrato il sistema, ripeto la sequenza per un tempo $t_1 = t_1 + \Delta t$ (in genere si parte da $t_1 = 0$ fino ad un valore massimo predefinito. In tal modo esploro 'osservo' l'evoluzione del mio sistema di nuclei al variare di t_1 .



Tecniche bidimensionali

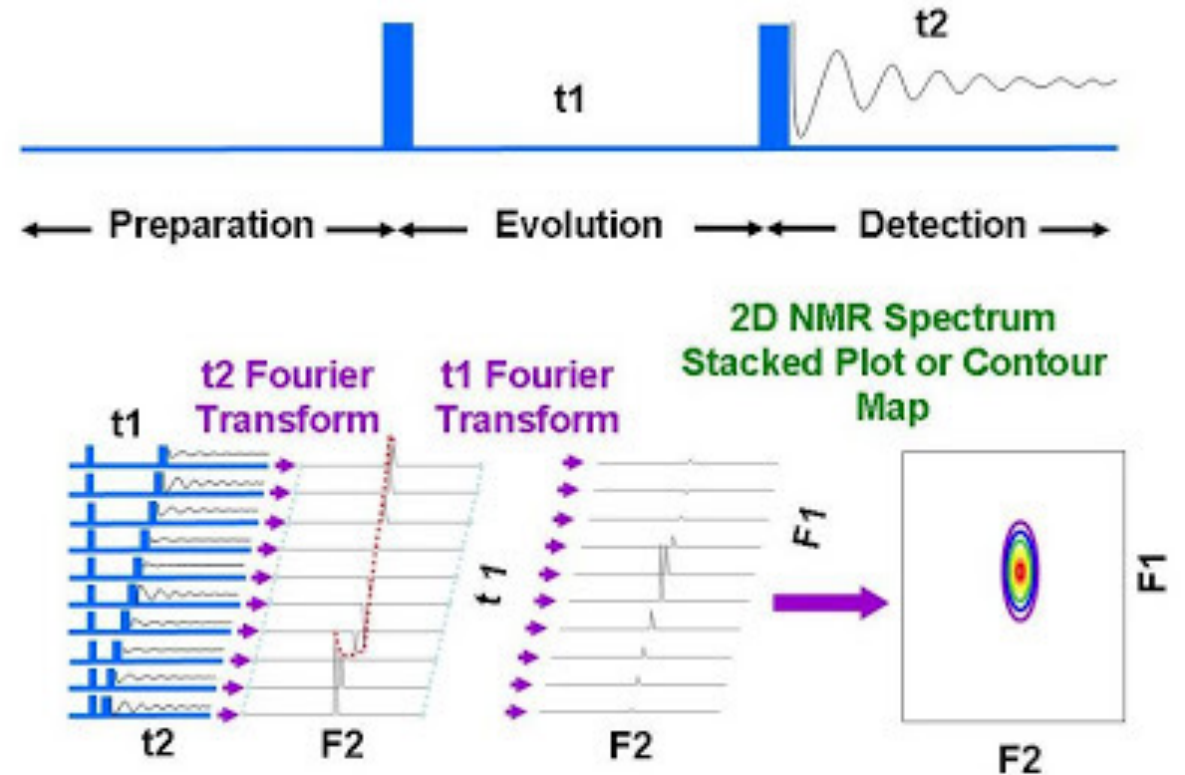
Il FID viene registrato solo durante t2

L'insieme dei FID registrati a tempi t1 diversi viene 'trasformato' (secondo Fourier) in modo da ottenere una mappa bidimensionale dove in ascisse e ordinate ho le frequenze (non più i tempi).

A seconda dell'esperimento posso avere sia accoppiamenti omonucleari (es: $^1\text{H} - ^1\text{H}$) o eteronucleari ($^1\text{H} - ^{13}\text{C}$, $^1\text{H} - ^{15}\text{N}$)

Nelle mappe bidimensionali sono importanti i picchi fuori diagonale (indicano le interazioni tra nuclei)

Two Dimensional NMR Spectroscopy

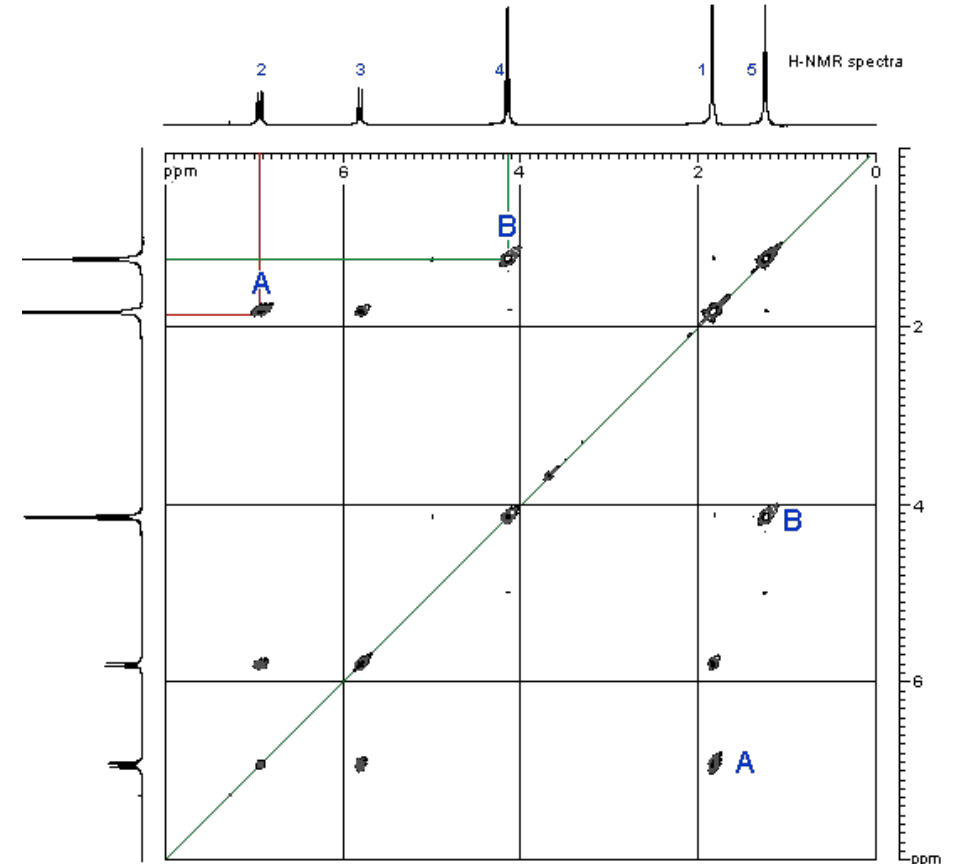
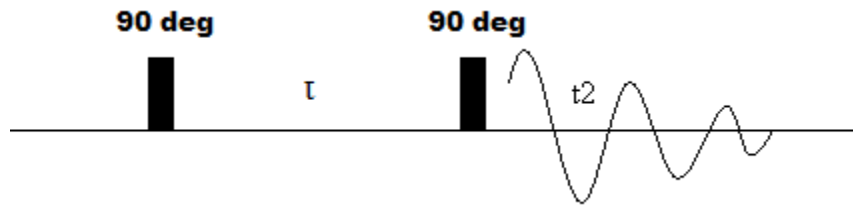


COSY

Nel COSY ho trasmissione di magnetizzazione tra nuclei uguali (^1H ma anche ^{13}C) e legati ad atomi vicini.

La magnetizzazione si trasmette lungo i legami (J-coupling).

Dalla posizione dei picchi fuori della diagonale posso capire quali nuclei sono accoppiati e quindi vicini.

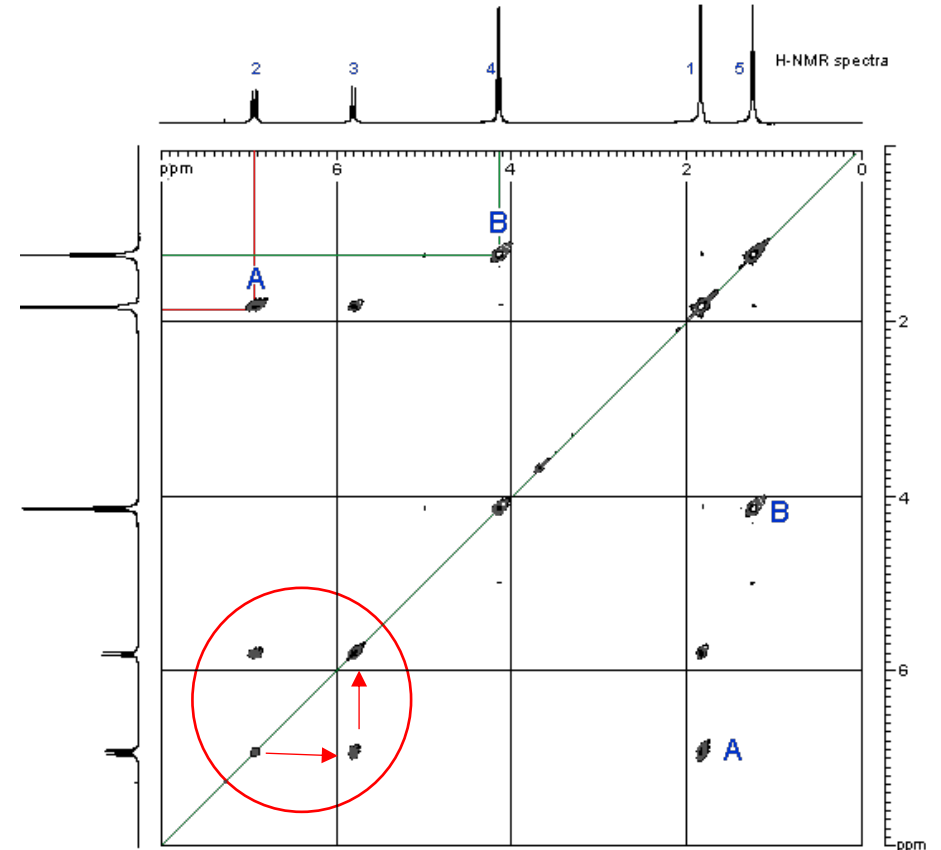


COSY

Sui due assi riporto gli spettri

I picchi lungo la diagonale corrispondono a quelli dello spettro 1-D, i **picchi fuori della diagonale indicano correlazione tra nuclei vicini**.

In questo modo è possibile stabilire quali nuclei sono vicini ed assegnare i picchi agli atomi.

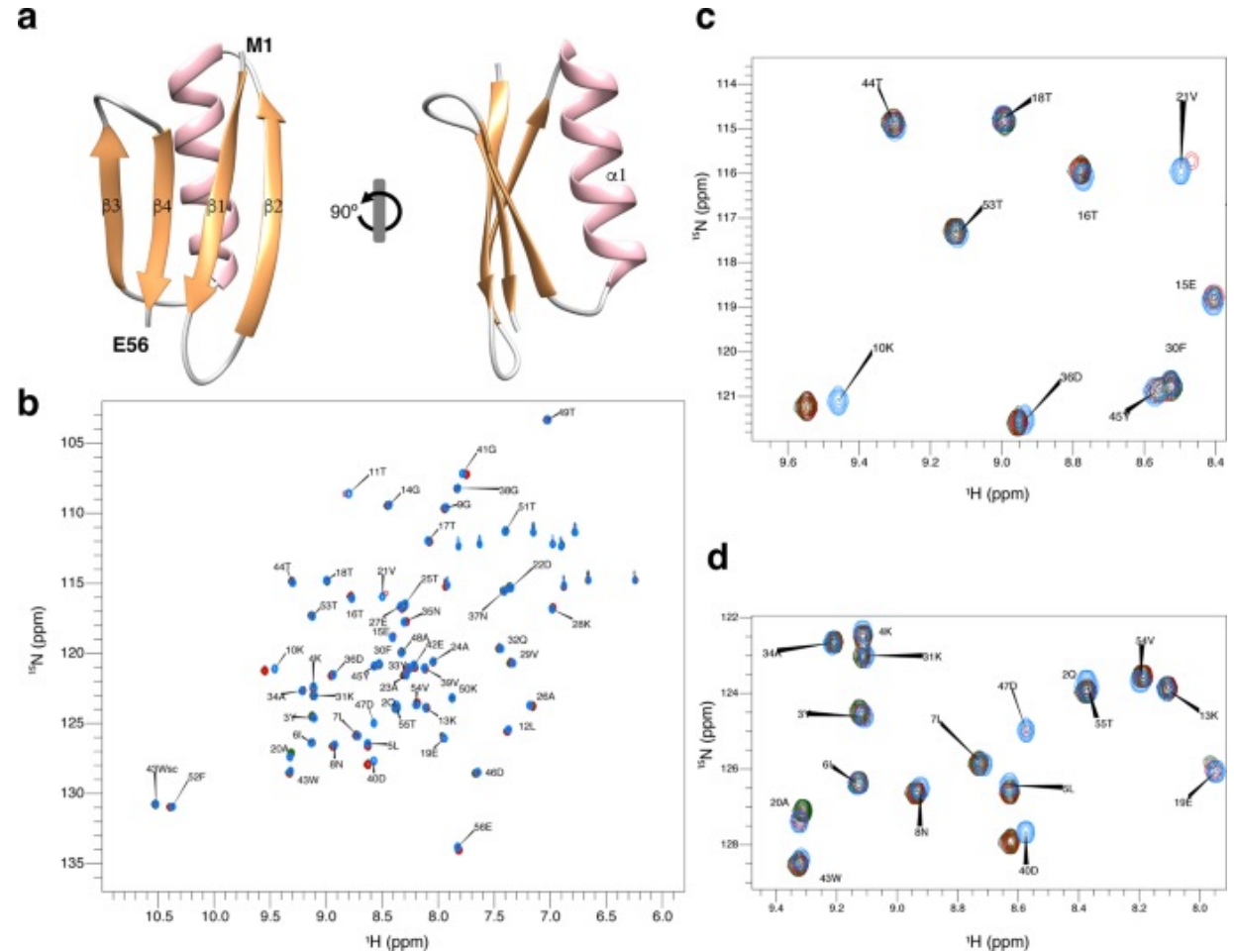


Spettri HSQC

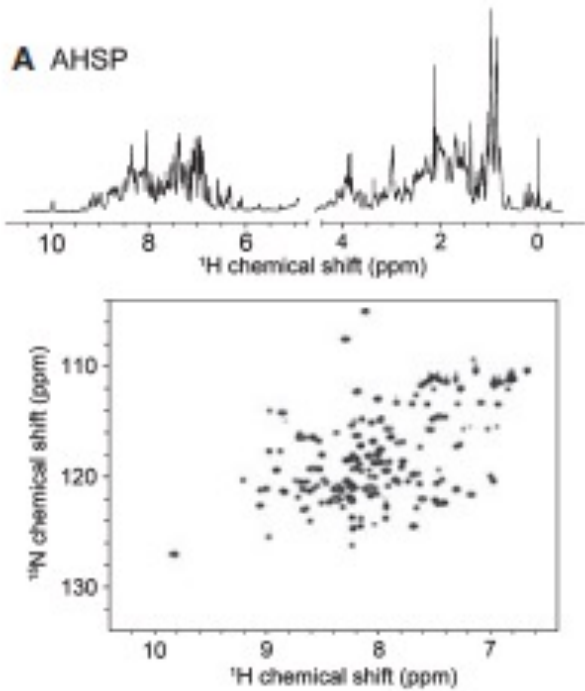
Nella tecnica bidimensionale (multidimensionale) **Heteronuclear Single-Quantum Coherence (HSQC)** si correla lo spettro ^1H con quello di un eteroatomo (^{13}C o ^{15}N)

Lo spettro bidimensionale ^1H - ^{15}N HSQC permette di assegnare (identificare) i protoni ammidici di una proteina.

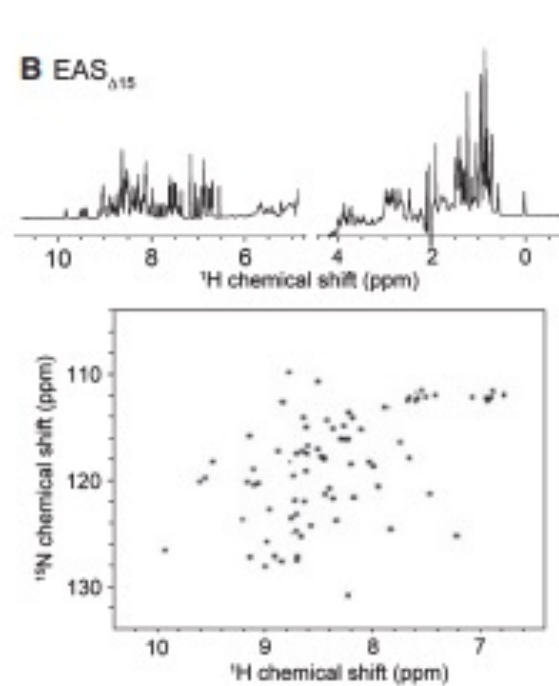
E' necessaria la marcatura della proteina con ^{15}N



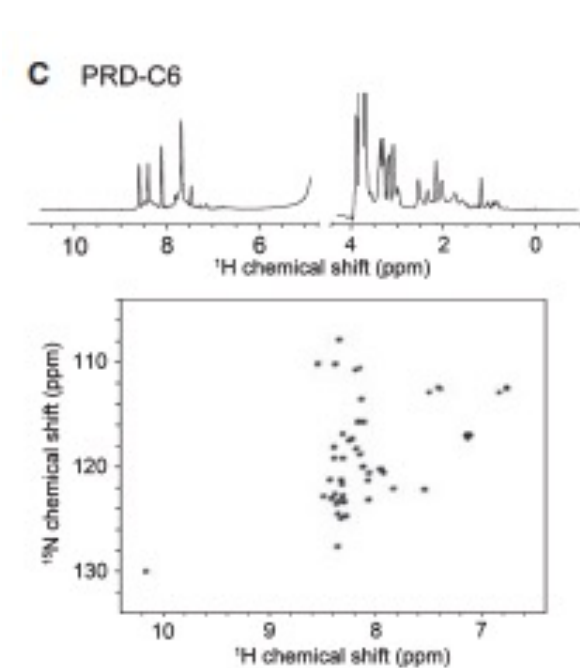
HSQC e struttura secondaria



α -elica



foglietto β



Proteina non struttura

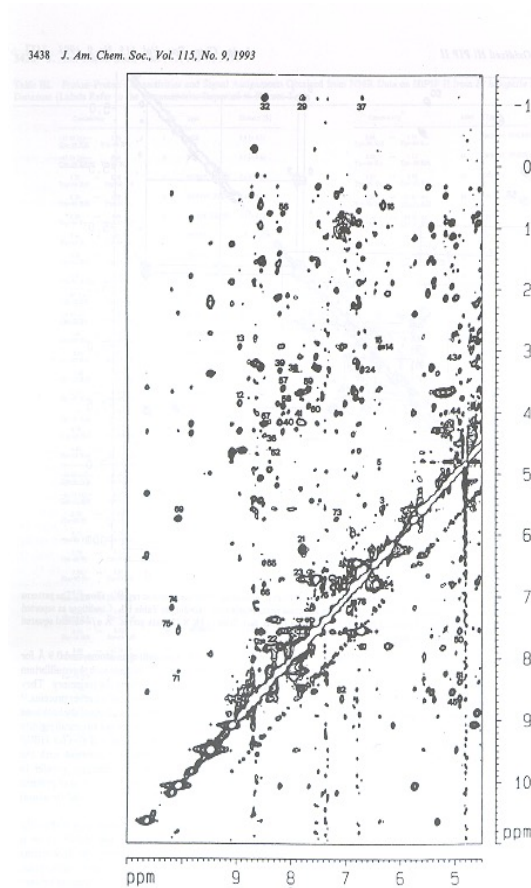
Lo spettro ^1H - ^{15}N -HSQC è sensibile alla struttura secondaria

NOESY

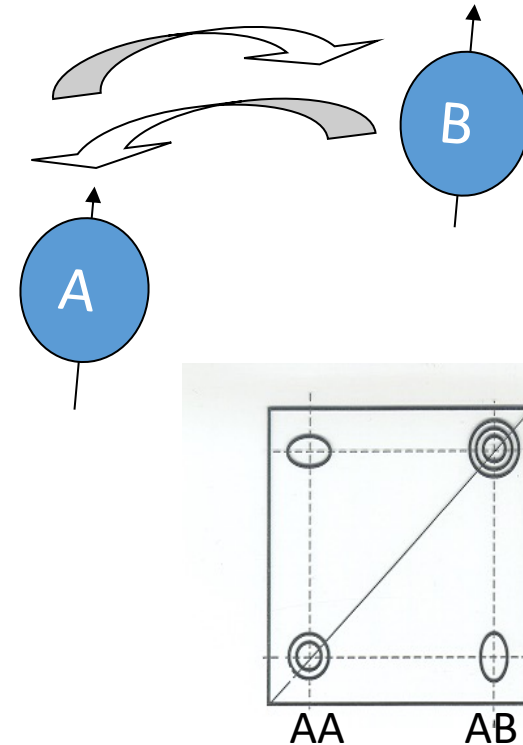
L'accoppiamento tramite NOE viene generalmente valutato tramite mappe bidimensionali.

Per capire quale atomo è correlato spazialmente con altri atomi si ricorre all'esperimento **NOESY**.

Le mappe possono essere molto complesse.



NOESY
75 aminoacidi



NOESY

Anche nel NOESY, riporto i due spettri sui due assi.

In questo caso la trasmissione della magnetizzazione è nello spazio e non attraverso i legami.

Considero la trasmissione di magnetizzazione tra nuclei dello stesso tipo.

I picchi al di fuori della diagonale mi indicano quale nucleo e correlato con quale e quindi quali nuclei sono spazialmente vicini.

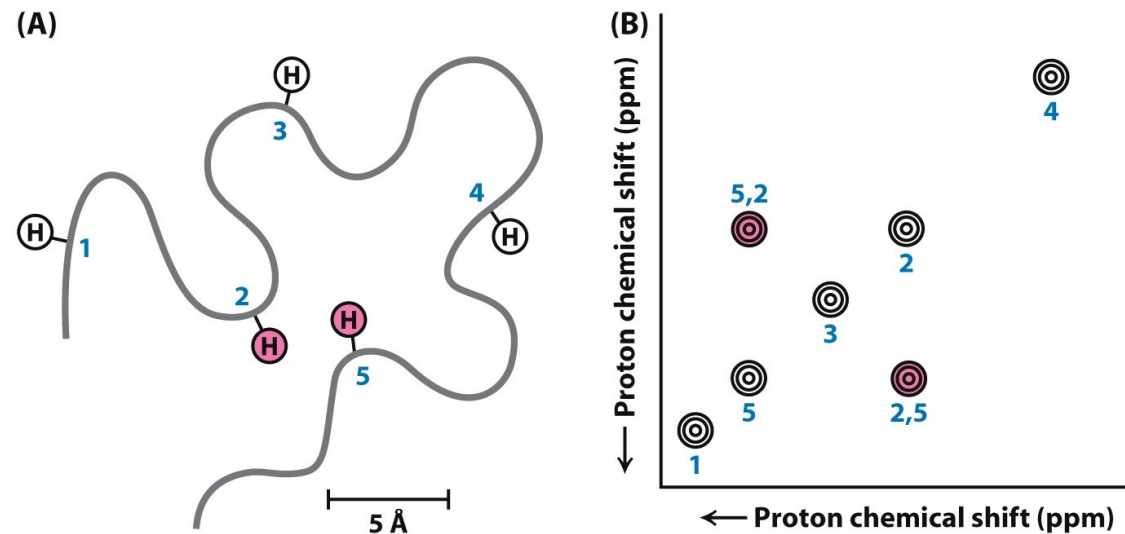
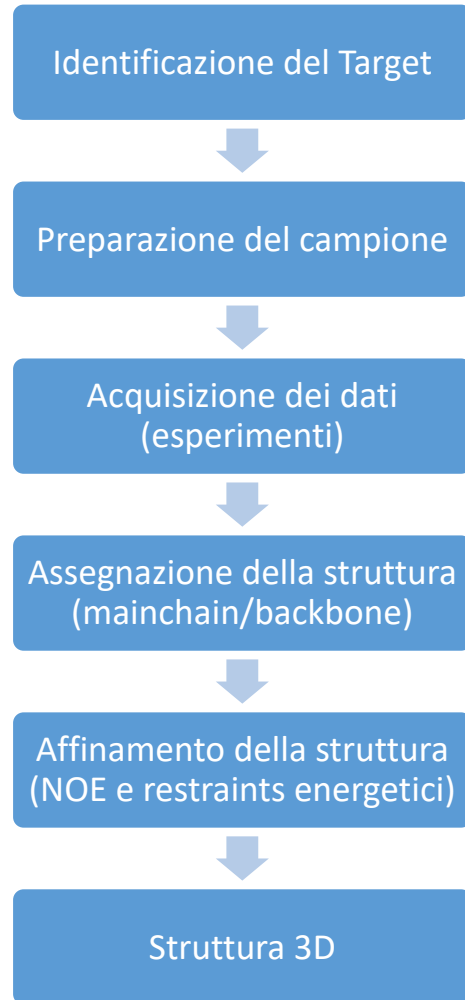


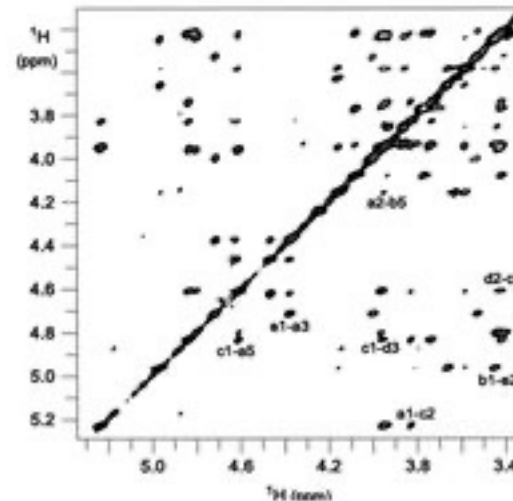
Figure 3.46
Biochemistry, Seventh Edition
© 2012 W. H. Freeman and Company

NMR per la determinazione strutturale



Esistono una molteplicità di tecniche che contribuiscono alla determinazione strutturale di proteine in soluzione.

La determinazione strutturale con NMR rimane comunque un problema complesso.



Determinazione strutturale di una proteina

1. Il primo step consiste nell'assegnazione delle diverse risonanze, ai diversi aminoacidi. In questa fase è fondamentale l'uso di esperimenti multidimensionali, come HSQC.
2. Si determinano le correlazioni NOE tra i diversi protoni. Questo è possibile in seguito all'assegnazione delle diverse risonanze. **I NOE costituiscono i principali restraints strutturali.**
3. Viene generata una struttura che deve rispettare i criteri strutturali noti ed essere coerente con i dati NMR osservati. Questa struttura viene modellata (automaticamente) cercando di migliorare (***minimi quadrati***) l'accordo tra i restraints (NOE) osservati e quelli calcolati a partire dal modello. Viene utilizzata la ***Dinamica Molecolare*** con l'aggiunta di restraints (NOE) sperimentali.
4. Si ripete fino a generare un certo numero di strutture tra loro coerenti.

$$F(x, y, z) = Energy(x, y, z) + Restraints(x, y, z)$$

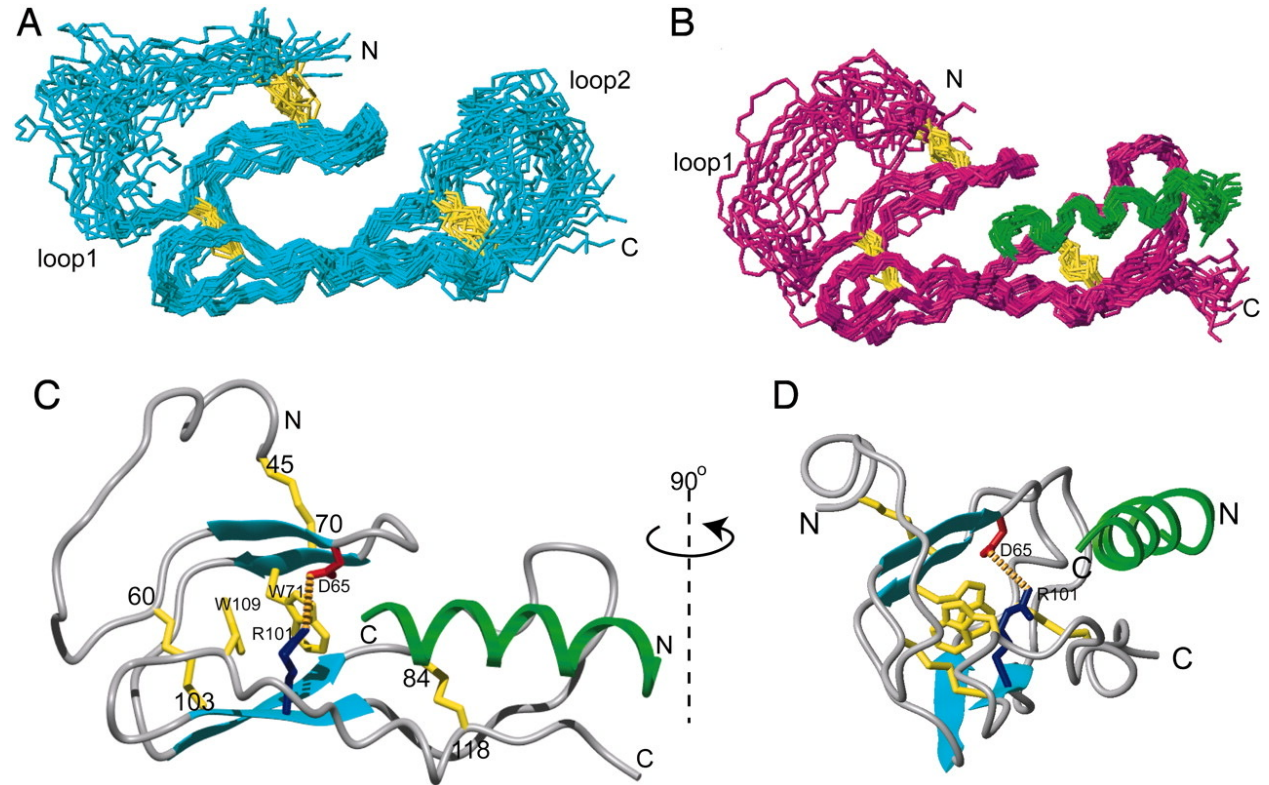
$F(x, y, z)$ viene minimizzata

La struttura NMR

Non si ottiene un modello unico ma un insieme di modelli.

Le parti più flessibili della struttura avranno pochi restraints strutturali (NOE) e quindi avranno maggiore variabilità nei modelli generati.

I modelli devono soddisfare i requisiti stereochimici noti per le macromolecole (distanze ed angoli di legame, plot di Ramachandran, collisioni tra atomi).



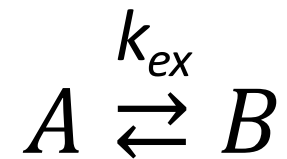
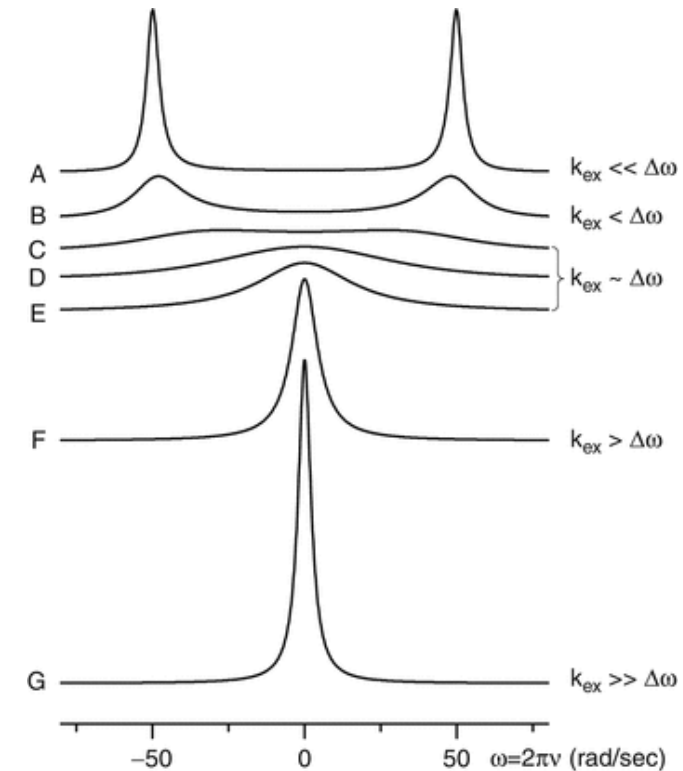
NMR e dinamica

L'NMR è probabilmente la tecnica strumentale più potente per lo studio della dinamica di una macromolecola, ovvero della sua evoluzione temporale in un determinato contesto biologico /chimico /fisico

Il movimento di un nucleo influenzerà la forma e posizione del suo assorbimento nello spettro NMR.

L'interconversione di A e B, con due protoni con chemical shift differenti mostrerà dei picchi distinti o meno, in relazione alla velocità di interconversione.

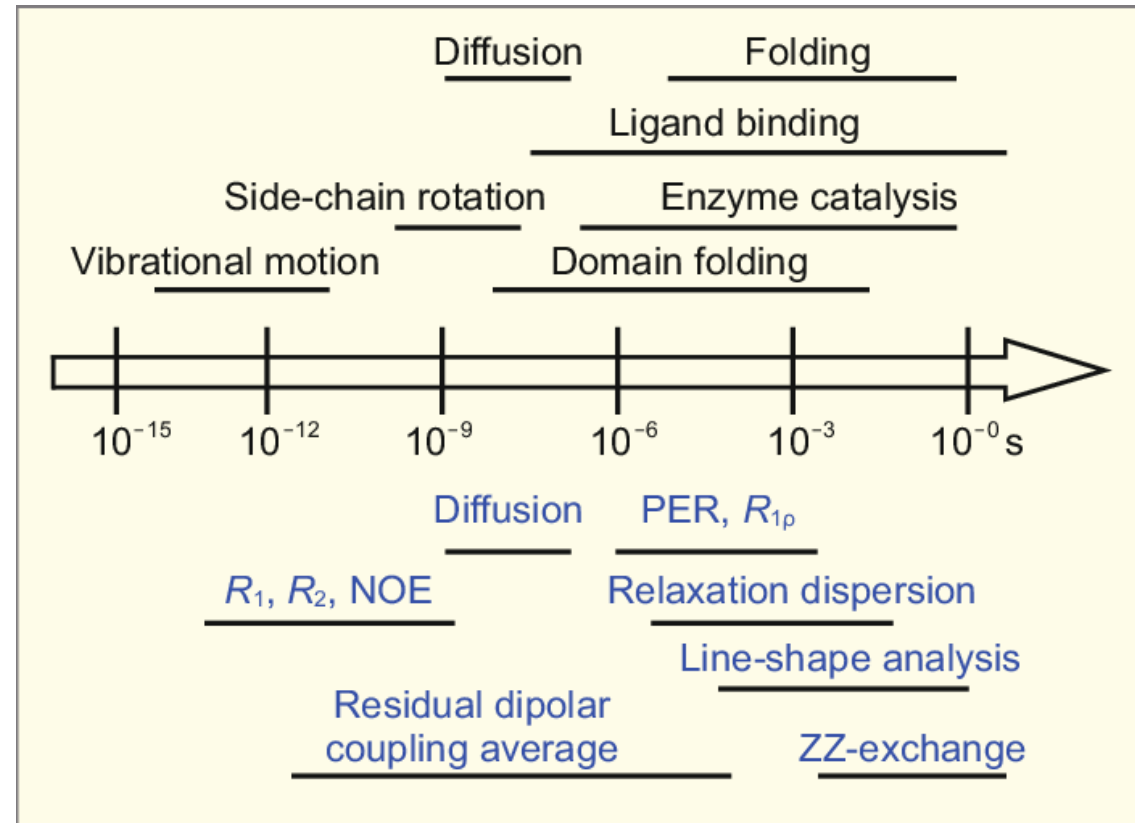
- Se il processo è lento in confronto alle frequenze di assorbimento, allora osservo due picchi distinti
- Se il processo è rapido, osservo un unico picco (medio)



Scala temporale in NMR

L'NMR è una tecnica estremamente versatile, scegliendo il metodo più opportuno è possibile studiare fenomeni che avvengono su scale temporali che vanno dai ps (rotazioni di catene laterali) fino ai ms e oltre (cinetica enzimatica, folding).

L'NMR è probabilmente lo strumento più versatile per l'analisi della dinamica di una (macro)molecola, l'unico limite è che presuppone l'assegnazione del picco spettrale di interesse



NMR e drug discovery

L'NMR trova una larghissima applicazione anche nello studio dell'interazione tra proteine e ligandi (es: proteina-farmaco).

L'interazione tra target (proteina) e ligando (piccola molecola) può essere seguito analizzando:

- Gli effetti dell'interazione sullo spettro NMR del **target** (macromolecola)
- Gli effetti sullo spettro del **ligando** (es: inibitore).

Se si è interessati al fenomeno del binding (costante di dissociazione: K_d), ***seguire il ligando è più semplice***
Ci sono molti approcci distinti, tra i più usati:

- **WaterLogsy,**
- **STD**

Chemical shift Perturbation

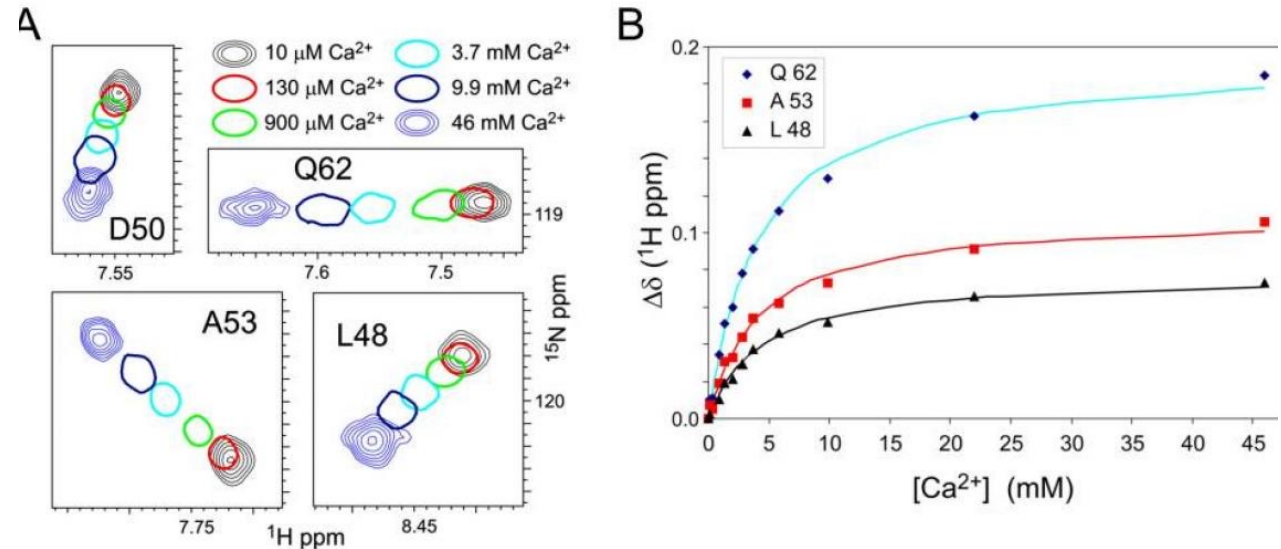
L'interazione tra proteina e ligando determina una variazione del chemical shift degli atomi (H, N, C) coinvolti.

La variazione di chemical shift si osserva bene nelle mappe bidimensionali ^1H - ^{15}N (HSQC)

In questo modo è possibile sia titolare il ligando, per ottenere una K_d , che avere informazioni sui residui coinvolti.

E' un metodo molto attraente, ma ha due requisiti:

- I chemical shift devono essere assegnati
- Richiede un esperimento di tipo bidimensionale (con marcatura etc...)



STD

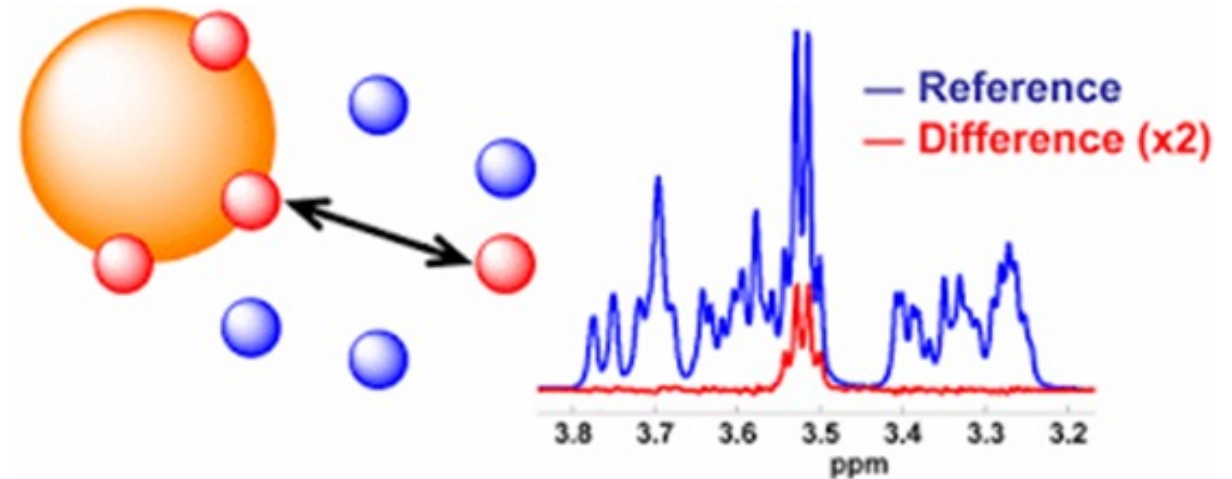
Il metodo STD (*Saturation from target to ligand*) in cui si analizza il trasferimento di magnetizzazione dalla proteina al ligando, che può avvenire solo se esiste interazione tra i due (proteina e ligando sono vicini).

Con il sistema proteina+ligando vengono misurati due spettri:

1. Uno spettro che non satura i protoni della proteina
2. Un secondo spettro che satura i protoni della proteina
3. Viene calcolata la differenza tra i due spettri

I picchi positivi nello spettro differenza sono attribuibili all'interazione con il ligando.

L'esperimento è ripetuto per concentrazioni crescenti di ligando.



Con l'STD è possibile misurare anche K_d dell'ordine del mM

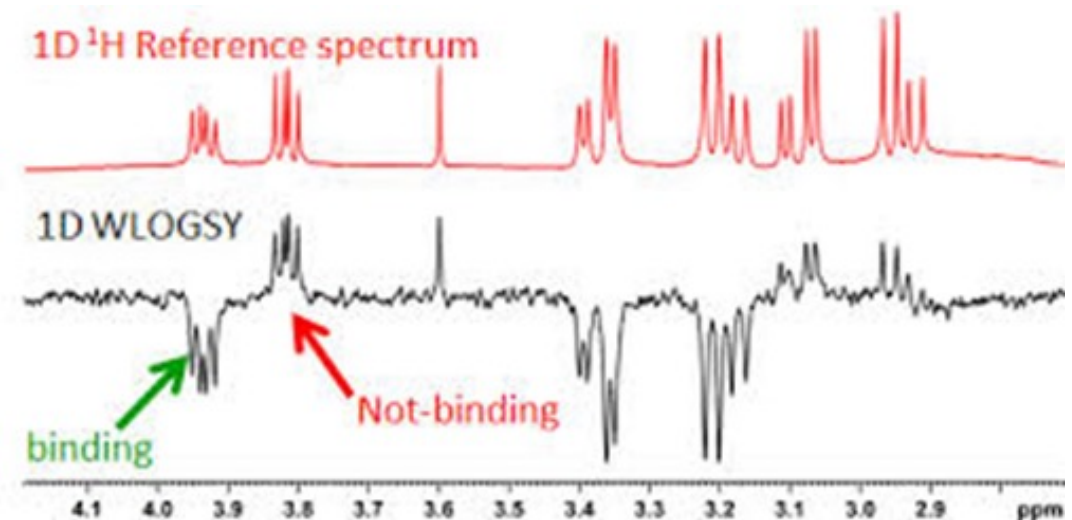
(Saturare: eccitare contemporaneamente tutti i nuclei interessati)

WaterLogsy

Il metodo WaterLogsy è basato sul diverso **trasferimento di magnetizzazione tra ligando e H₂O a seconda che il ligando interagisca con la proteina o meno.**

Il trasferimento di polarizzazione dall'acqua al ligando avviene diversamente se il ligando è legato o meno.

La diversa solvatazione del ligando non-legato/legato determina dei **picchi negativi** nello spettro WaterLOGSY



Vantaggi e svantaggi dell'NMR (per strutture 3D)

Vantaggi:

- L'NMR viene fatto in soluzione (non richiede cristalli)
- L'NMR può fornire informazione sulla dinamica delle proteine

Svantaggi:

- La complessità degli spettri limita le dimensioni dei sistemi studiabili (100 kDa?)
- La proteina è in una soluzione molto concentrata (non esattamente condizioni fisiologiche)
- Può essere richiesta una discreta mole di lavoro in laboratorio (marcatura con isotopi)
- Il modello finale non è univoco (ma può essere un vantaggio)



L'NMR per la determinazione strutturale richiede strumenti molto potenti e costosi, non dissimili dalle moderne beamlines presso i sincrotroni.

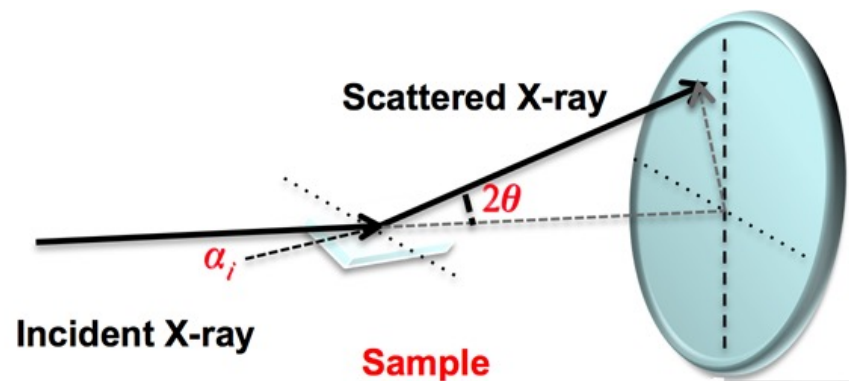
Metodi basati sullo Scattering (e CryoEM)

Metodi Basati sullo 'Scattering'

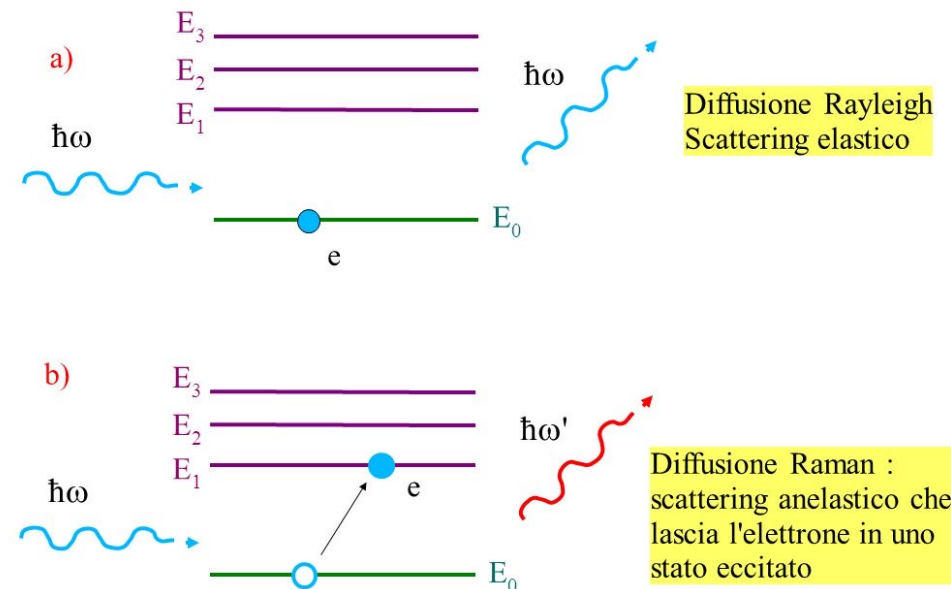
A differenza delle spettroscopie, dove la radiazione elettromagnetica viene assorbita dalla materia ed emessa a lunghezza d'onda più lunga (fenomeno dello scattering non-elastico), nei cosiddetti metodi basati sullo scattering (diffusione, in italiano) la radiazione elettromagnetica interagisce con la materia senza cambiare lunghezza d'onda ovvero senza perdere energia (scattering elastico).

Esistono diversi metodi utilizzati in biologia strutturale basati sullo scattering (elastico) della radiazione elettromagnetica:

- Dynamic Light Scattering (DLS)
- Small Angle X-ray Scattering (SAXS)
- **Metodi basati sulla diffrazione**



Interazione fotone – atomo



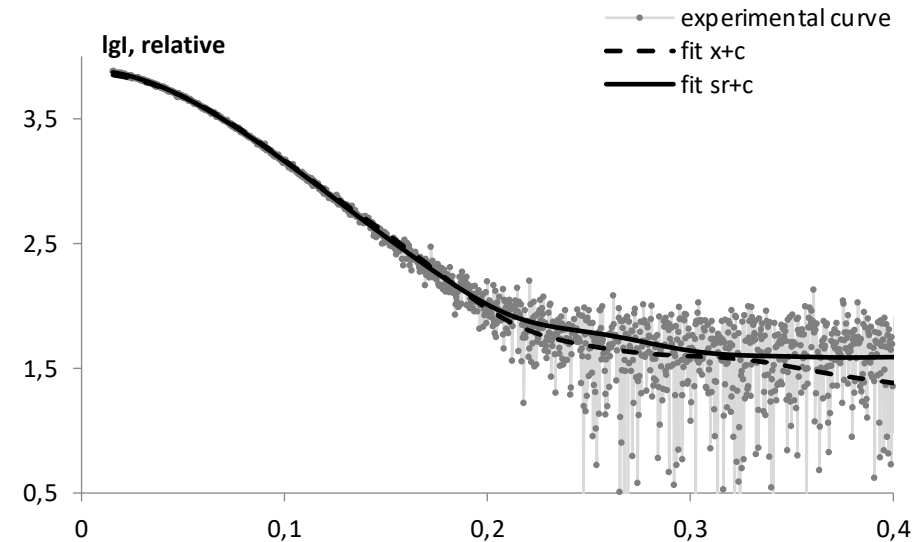
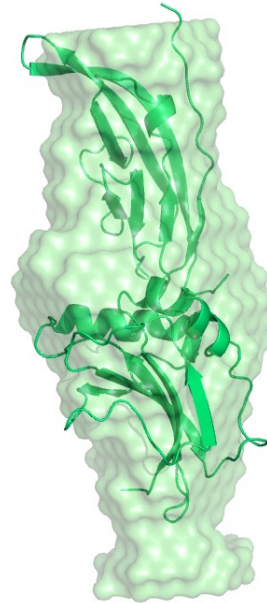
Small Angle X-ray Scattering

Se un campione omogeneo, in soluzione, viene irradiato con raggi-X, la radiazione stessa verrà deflessa con diversa intensità a vari angoli, in relazione alle dimensioni, forma e natura chimica delle 'molecole' in soluzione (**Small Angle X-ray Scattering, SAXS**).

Il SAXS, permette la ricostruzione della forma (shape) della macromolecola.

Quindi il SAXS è utilissimo per avere indicazioni strutturali a bassa risoluzione ma in condizioni quasi fisiologiche.

Esiste una variante che usa i **neutroni** invece che i raggi-X (**SANS**)



La curva delle intensità in funzione dell'angolo è la curva SAXS, e può essere utilizzata per estrarre parametri utili a ricostruire la forma della molecola in soluzione.

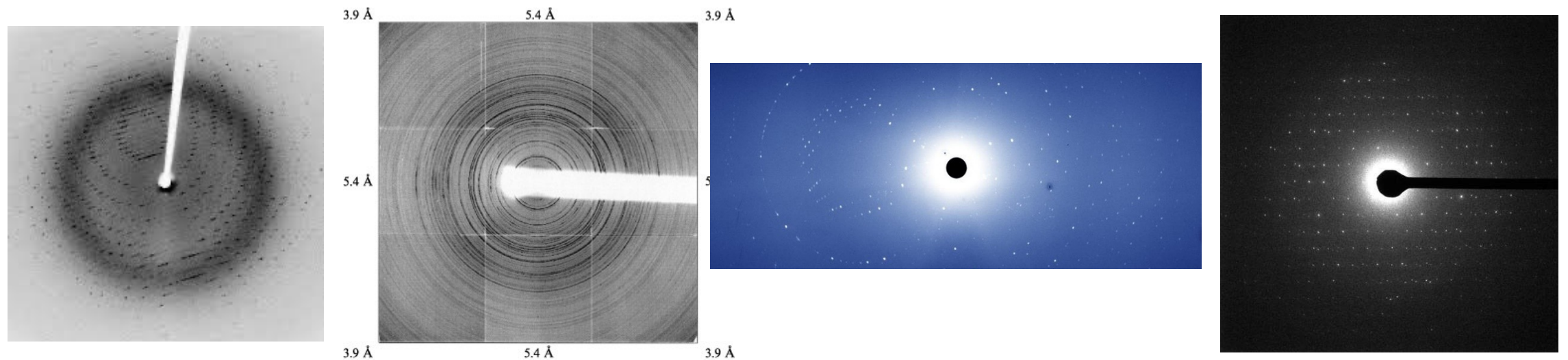
Lo scattering è analizzato solo ad angoli di 'deflessione' piccoli.

Diffrazione

Nella diffrazione ho sempre un'interazione senza perdita di energia tra radiazione elettromagnetica e materia (scattering elastico). Se la materia è ordinata in un reticolo cristallino si verifica il fenomeno della diffrazione, ovvero interferenza costruttiva e distruttiva delle onde 'diffuse'.

La diffrazione può utilizzare *radiazioni di natura diversa* (raggi-X, neutroni, elettroni)

La diffrazione può investigare materiali diversi (cristalli singoli, polveri cristalline, cristalli bidimensionali, fibre)

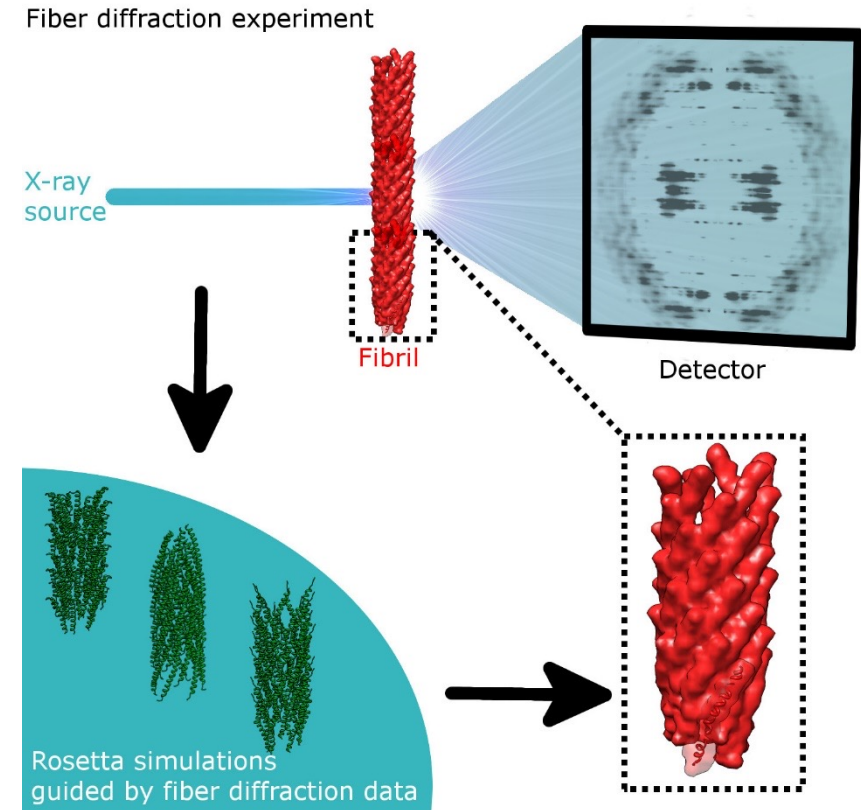
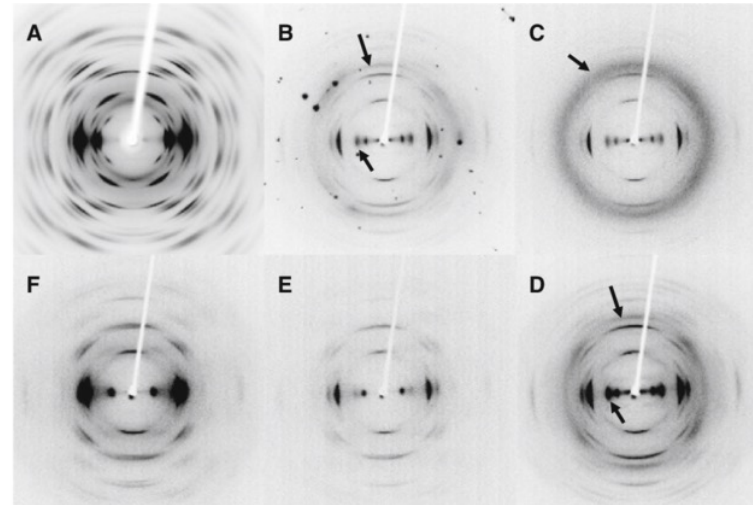


Diffrazione da fibre

Le fibre, come il DNA, alcuni polimeri e le fibrille non possiedono ordine in 3 dimensioni, tuttavia sono ordinati lungo 1 dimensione. Per questo motivo le fibre danno luogo a figure di diffrazione da cui è possibile estrarre alcune informazioni strutturali:

- **Simmetria di ciò che si ripete** (es: Elicoidale)
- **Spaziatura tra le unità che si ripetono** (passo dell'elica)
- **'Larghezza' di ciò che si ripete**

La diffrazione di fibre trae giovamento dall'uso della radiazione di sincrotrone (intensità e collimazione)

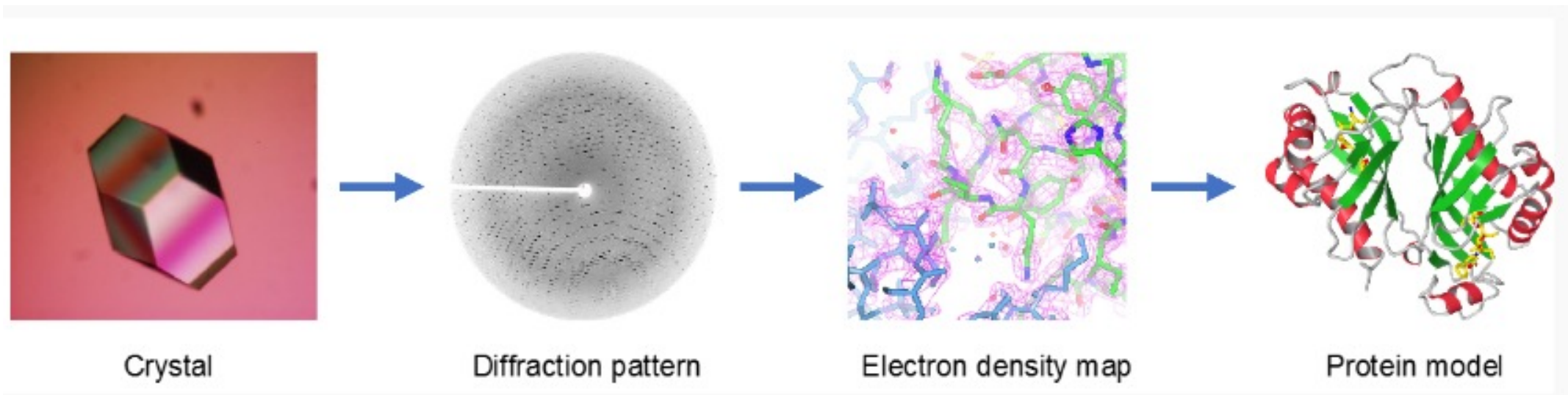


Diffrazione da cristallo singolo

La tecnica della diffrazione di raggi-X da cristallo singolo è una tecnica potentissima e consolidata, in grado di fornire modelli strutturali ad alta risoluzione di molecole di qualsiasi peso molecolare.

Questa tecnica ha fondamentalmente un solo limite: **richiede un cristallo che diffranga**.

E' una tecnica largamente basata sulla radiazione di sincrotrone (ma non è indispensabile)



Diffrazione da elettroni

Gli elettroni sono quasi sempre concepiti come particelle, tuttavia in virtù della dualità onda/particella anche gli elettroni hanno una loro lunghezza d'onda (dell'ordine di 0.05 \AA), assolutamente compatibile con i fenomeni di diffrazione da parte di cristalli

- Nel 1927 viene dimostrato da Clint Davisson e Lester Germer che gli elettroni danno luogo al fenomeno della diffrazione
- Negli anni 50, più o meno nel periodo in cui Kendrew e Perutz lavorano sulla diffrazione di raggi-X, si fanno i primi esperimenti di diffrazione di elettroni su cristalli di proteine
- Anni 60-70 Aaron Klug lavora alla diffrazione di elettroni sui Virus
- Nel 1975 R. Henderson pubblica la struttura a 7 Angstrom di risoluzione della Batteriorodopsina (cristalli 2D)
- Nel 1990 R. Henderson pubblica la struttura della Batteriorodopsina a risoluzione atomica
- Introduzione della tecnica della micro Electron Diffraction (micro-ED)

La diffrazione di elettroni in biologia strutturale

Gli elettroni interagiscono fortemente con la materia, molto più dei raggi-X. Questo è sicuramente un vantaggio perché è richiesta meno 'materia' per avere un pattern di diffrazione accettabile, ma anche uno svantaggio perché il campione risulta danneggiato più rapidamente in seguito all'interazione con gli elettroni (richiede quindi trattamenti o condizioni particolari).

A differenza della diffrazione di raggi-X che usa cristalli 'tridimensionali', per molto tempo la diffrazione di elettroni ha indagato cristalli bidimensionali, ovvero cresciuti in 2 sole direzioni, questo perché l'assorbimento degli elettroni da parte della materia è molto elevato.

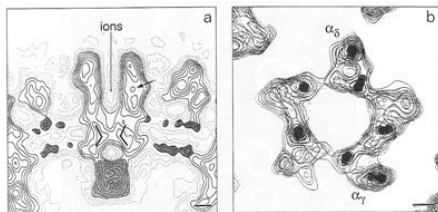
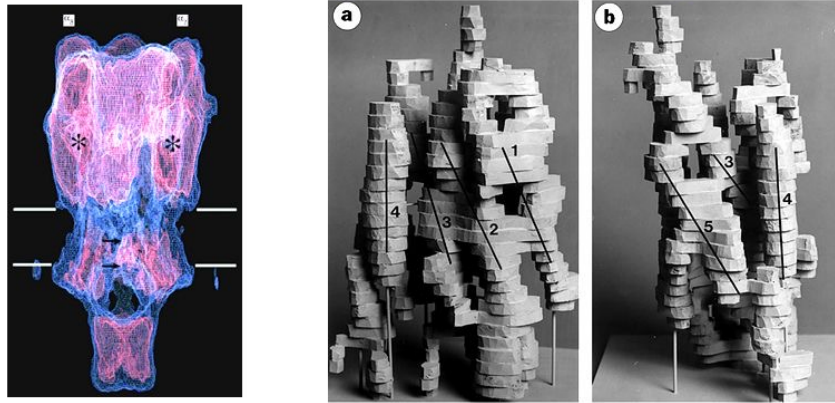
Crescere cristalli bidimensionali è forse più complesso che crescere cristalli tridimensionali

Un vantaggio della diffrazione di elettroni è che, a differenza della diffrazione dei raggi-X, non si pone 'il problema della fase', potendo ricavare le fasi direttamente dall'analisi delle immagini.

Ad oggi esiste una nuova tecnica, nota come micro Electron Diffraction (**micro-ED**), che non opera più su cristalli bidimensionali, ma su microcristalli, non adeguati a studi con raggi-X. La tecnica sembra molto promettente.

Diffrazione di elettroni: esempi

Electron diffraction from 2D crystals: Nicotinic Acetylcholine Receptor

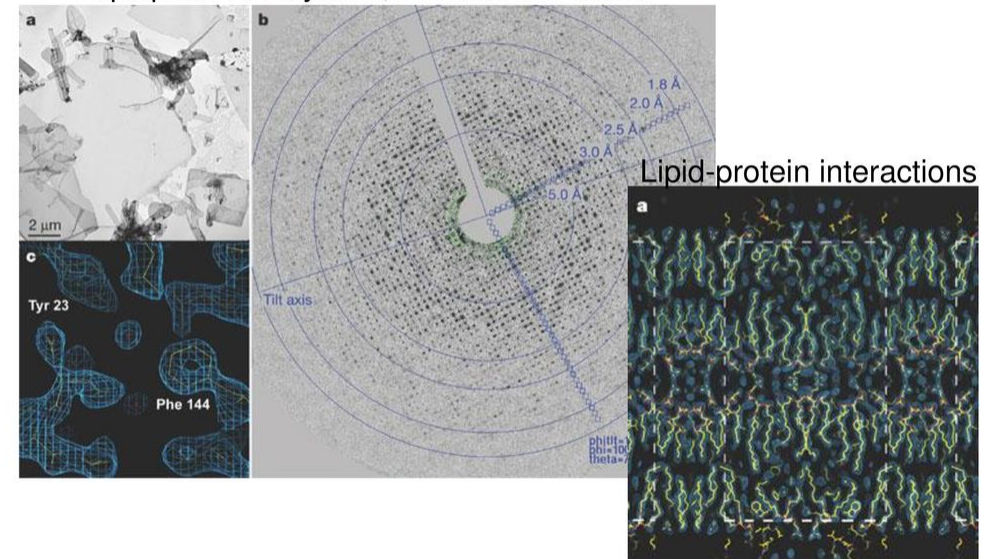


G-protein coupled receptors

Unwin et al, 2005

2D Crystals

Aquaporin 2D crystals, electron diffraction



test

Gonen et al. 2005. Nature 438:633-8.

CryoEM lecture Biophysical Chemistry Fall 2006

7

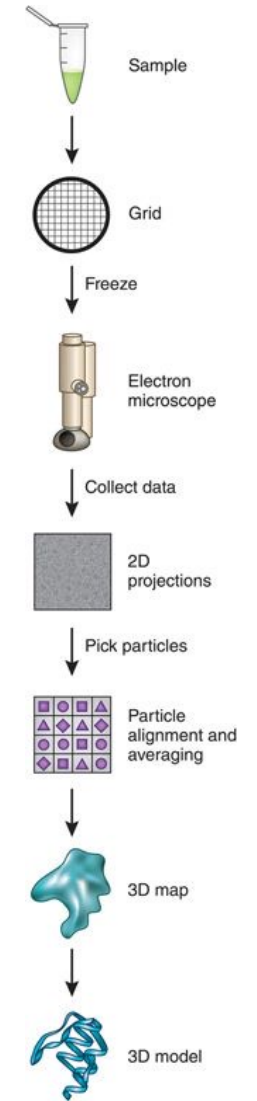
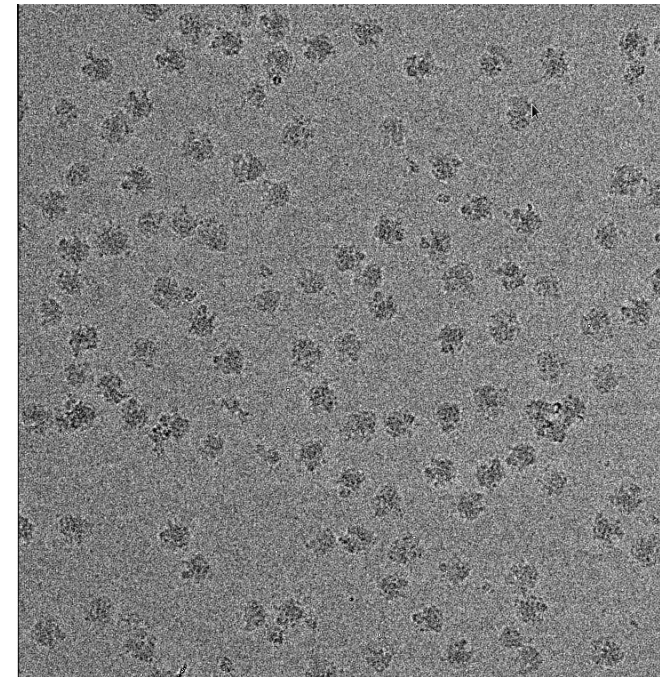
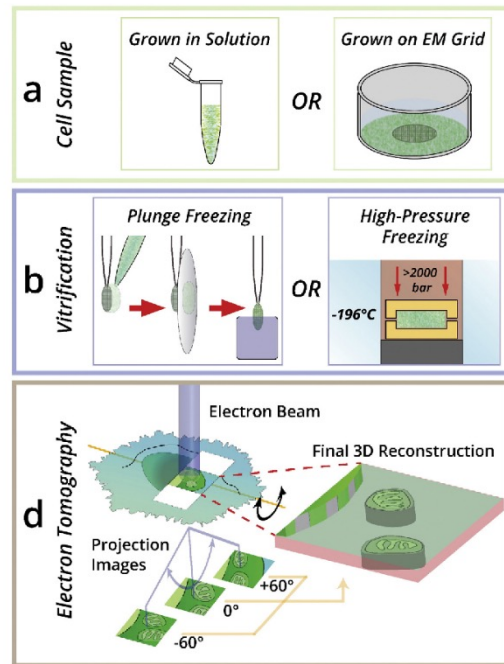
17 Jan 2006

24

Nel complesso la diffrazione di elettroni, fino alla fine degli anni 90, è stata una tecnica di nicchia, utile per esempio nello studio strutturale di proteine di membrana nel loro doppio strato lipidico. La risoluzione della struttura finale era sempre medio-bassa (per via delle problematiche sperimentali).

Single Particle Cryo-EM

I progressi tecnici e metodologici nella microscopia e nella diffrazione di elettroni, hanno portato a risultati eccezionali. Allo stato attuale la Cryo-Electron Microscopy (**Cryo-EM**) su molecola singola è in grado di fornire strutture molecolari a risoluzione atomica di molecole di dimensioni notevoli.



Cryo-EM per lo studio dei sistemi biologici

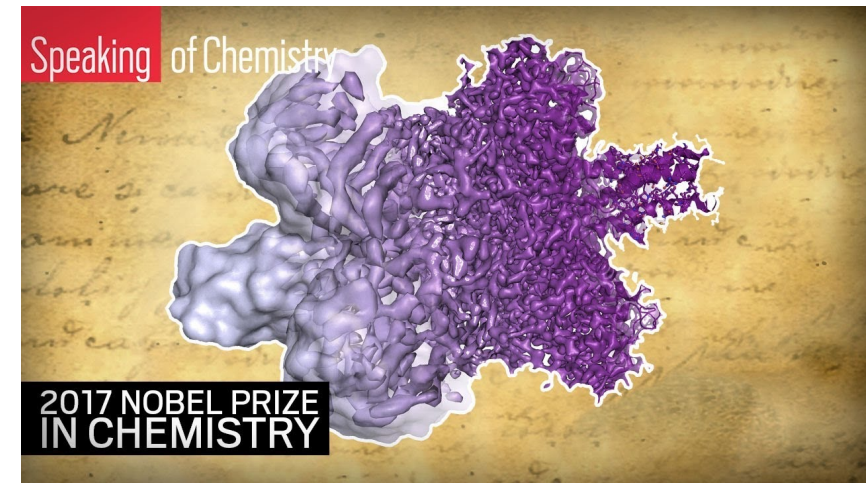
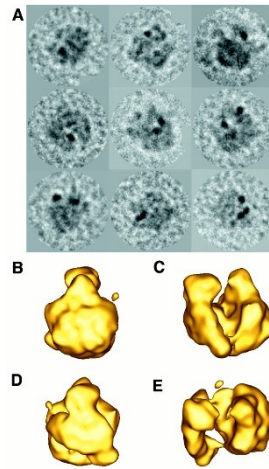
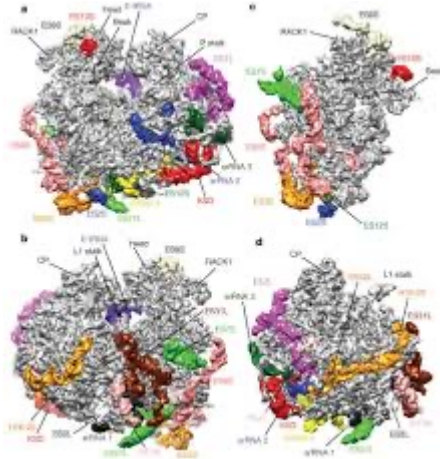
La Cryo-EM si sta dimostrando una tecnica potentissima.

Rispetto alla cristallografia di proteine mostra i seguenti vantaggi:

- Non ha bisogno di cristalli
- Può individuare stati conformazionali tra loro diversi

Gli svantaggi:

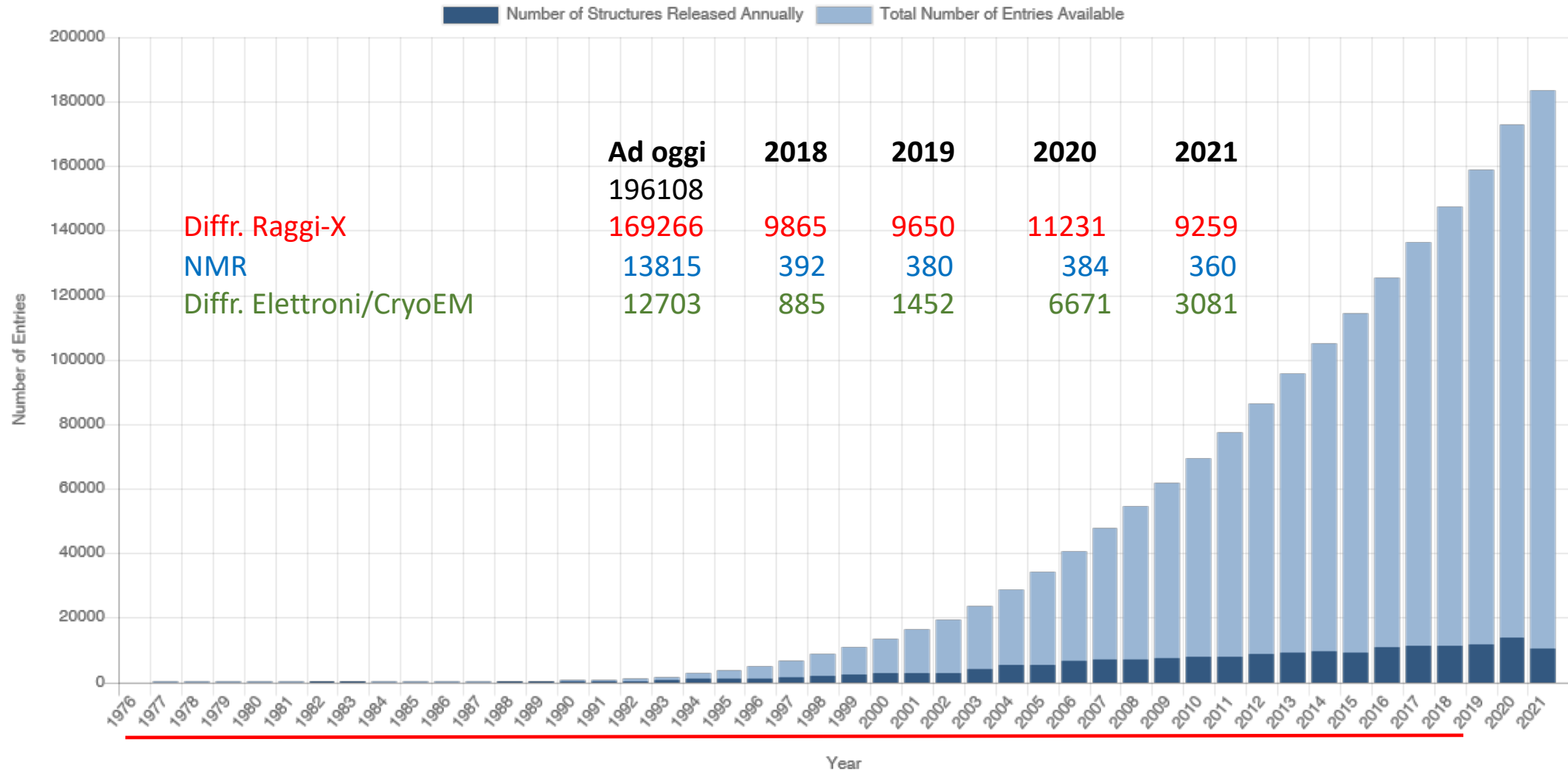
- La risoluzione media ottenibile è inferiore a quella ottenibile dai raggi-X (raramente si scende sotto i 2 Å di risoluzione, più frequentemente $> 3 \text{ \AA}$).
- La molecola studiata deve avere un peso molecolare di almeno 50 kDa



Vantaggi e Svantaggi delle diverse tecniche

Tecnica	Vantaggi	Svantaggi
Cristallografia di raggi-X:	<ul style="list-style-type: none">- Non ha limiti di peso molecolare- Tecnica consolidata- Risoluzione Atomica	<ul style="list-style-type: none">- Ha bisogno di cristalli!- Struttura influenzata dal cristallo- E' una tecnica statica
NMR:	<ul style="list-style-type: none">- E' in soluzione- Fornisce informazioni sulla dinamica- E' una tecnica molto flessibile	<ul style="list-style-type: none">- Peso molecolare \leq 50 kDa
Cryo-EM:	<ul style="list-style-type: none">- Non sono richiesti cristalli- Funziona su sistemi molto grandi	<ul style="list-style-type: none">- Risoluzione media- Danno da Radiazione- Peso molecolare \geq 50 kDa

Confronto tra tecniche



Metodi In-silico

Metodi *in-silico*

L'evoluzione dei computer, e la conoscenza sempre più profonda dei principi fisici o empirici che regolano la struttura di una macromolecola, ha portato all'evoluzione di metodi computazionali in grado di fornire modelli strutturali sufficientemente attendibili.

A seconda delle necessità avremo

- Metodi fisici
- Modelling (sequenza/similitudine)
- Docking

Metodi Fisici (Dinamica Molecolare)

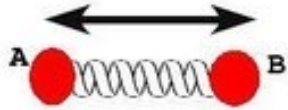
Dinamica Molecolare: generalità

Nella Dinamica Molecolare un sistema molecolare viene 'modellato' come masse connesse tra loro da molle e aventi una certa carica.

Le masse obbediscono alle leggi della meccanica classica (leggi di Newton e di Coulomb) ma anche alle leggi della meccanica statistica

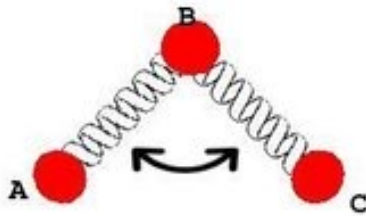
La Dinamica Molecolare è utilizzata per descrivere l'evoluzione temporale di un sistema di atomi, ma non per predire la struttura molecolare ab-initio

Dinamica Molecolare - modello



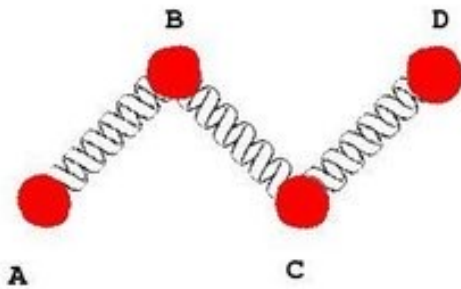
1) bonded interactions between 2 atoms

Interaction between A and B "spring in between 2 particles"



2) bonded interactions 2 bonds (angle interactions)

Interaction between AB bond and BC bond via another "spring"



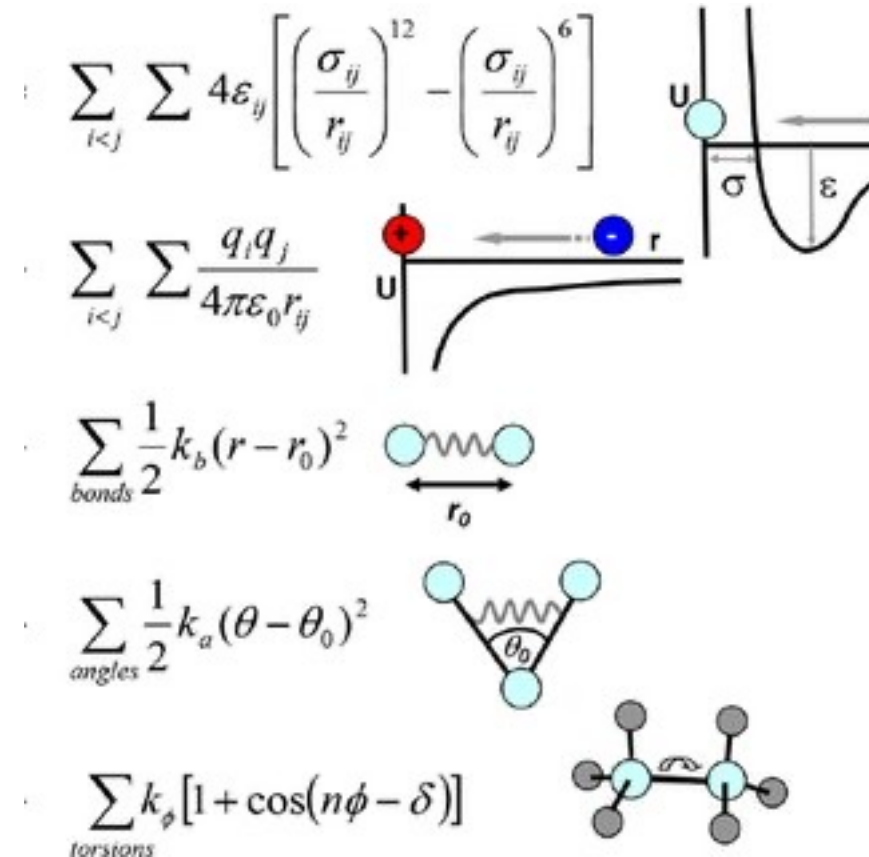
3) Dihedral interactions

These interactions are between two planes ABC and BCD in 3D space

Dinamica Molecolare

La molecola è vista come un sistema in cui gli atomi sono descritti come 'palline' legate da molle agli altri atomi.

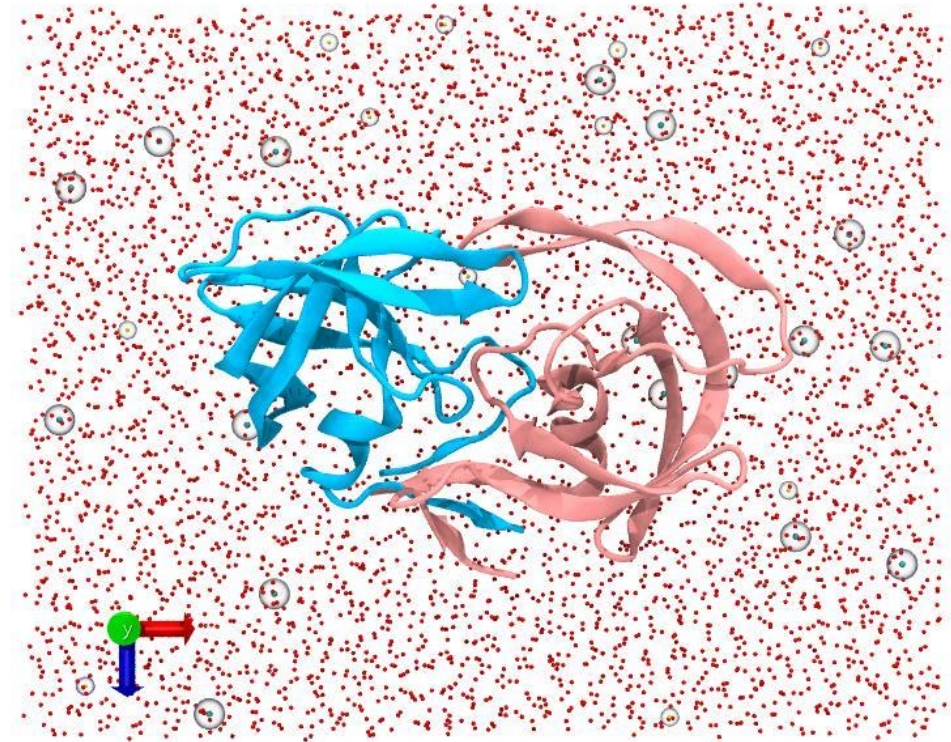
Inoltre sono considerate le interazioni tra atomi non legati (interazioni elettrostatiche. Van der Waals, angoli torsionali) secondo un punto di vista 'classico' (non quantistico).



Dinamica Molecolare: Solvatazione

La macromolecola viene 'immersa' in una 'scatola' di volume opportuna 'riempita' di molecole di acqua (o di altro solvente) e ioni opportuni.

In questo modo 'simulo' la proteina nel suo ambiente 'fisiologico'.



Dinamica Molecolare: Potenziale

Viene descritta una funzione potenziale, contenente i contributi di tutti i legami, angoli, angoli di torsione, termini di Van der Waals, termini elettrostatici...

Tutti gli atomi (proteina e solvente), i legami e le interazioni sono descritti da questa funzione.

In questo modo abbiamo una rappresentazione 'fisica' della nostra macromolecola immersa in acqua.

The Potential Energy Function

$$U(\vec{R}) = \underbrace{\sum_{\text{bonds}} k_i^{\text{bond}} (r_i - r_0)^2}_{U_{\text{bond}}} + \underbrace{\sum_{\text{angles}} k_i^{\text{angle}} (\theta_i - \theta_0)^2}_{U_{\text{angle}}} + \underbrace{\sum_{\text{dihedrals}} k_i^{\text{dihe}} [1 + \cos(n_i \phi_i + \delta_i)]}_{U_{\text{dihedral}}} + \underbrace{\sum_i \sum_{j \neq i} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \sum_i \sum_{j \neq i} \frac{q_i q_j}{\epsilon r_{ij}}}_{U_{\text{nonbond}}}$$

U_{bond} = oscillations about the equilibrium bond length

U_{angle} = oscillations of 3 atoms about an equilibrium bond angle

U_{dihedral} = torsional rotation of 4 atoms about a central bond

U_{nonbond} = non-bonded energy terms (electrostatics and Lenard-Jones)

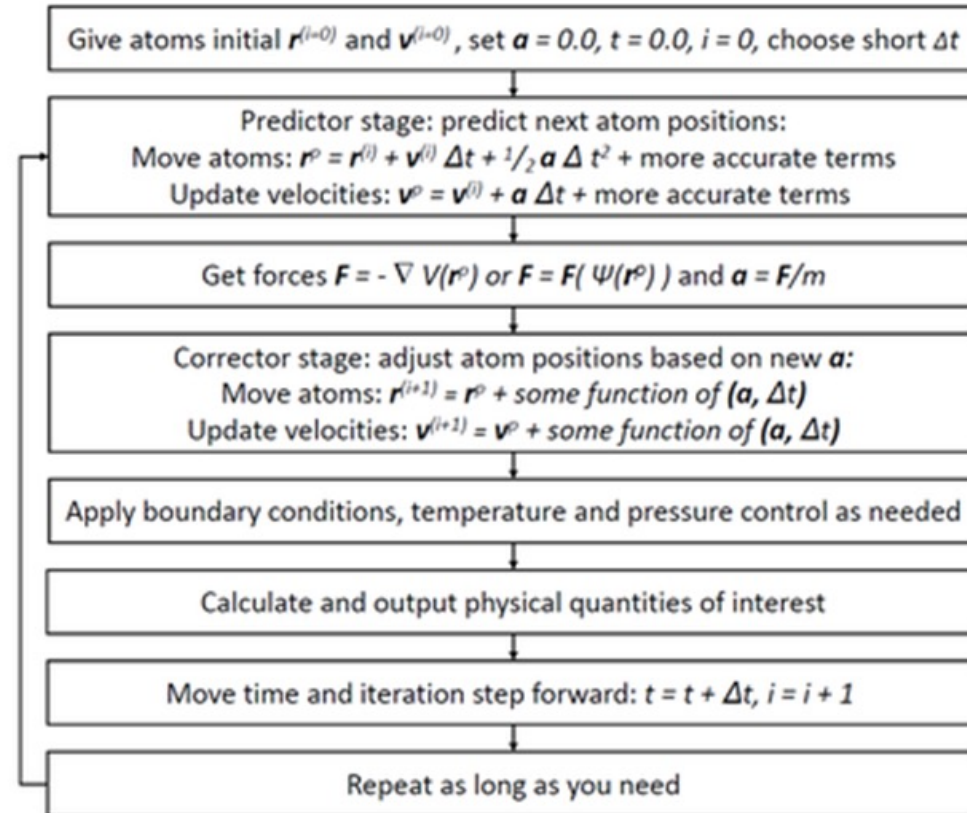
Dinamica Molecolare: Algoritmo

1) Al tempo $t=0$, agli atomi viene associata un'energia (velocità) iniziale dipendente dall'energia totale del sistema. Questa energia è definita dalle posizioni degli atomi e quindi dal potenziale U calcolato e dalla temperatura e pressione del sistema.

2) Minimizzando il potenziale, al tempo $t=t_1$, gli atomi vengono spostati in nuove posizioni.

3) A partire dalle nuove posizioni e applicando le opportune condizioni di temperatura e pressione, viene calcolato il nuovo potenziale per il tempo $t_2 = t_1 + \Delta t$ e così via fino ad un tempo t_{\max}

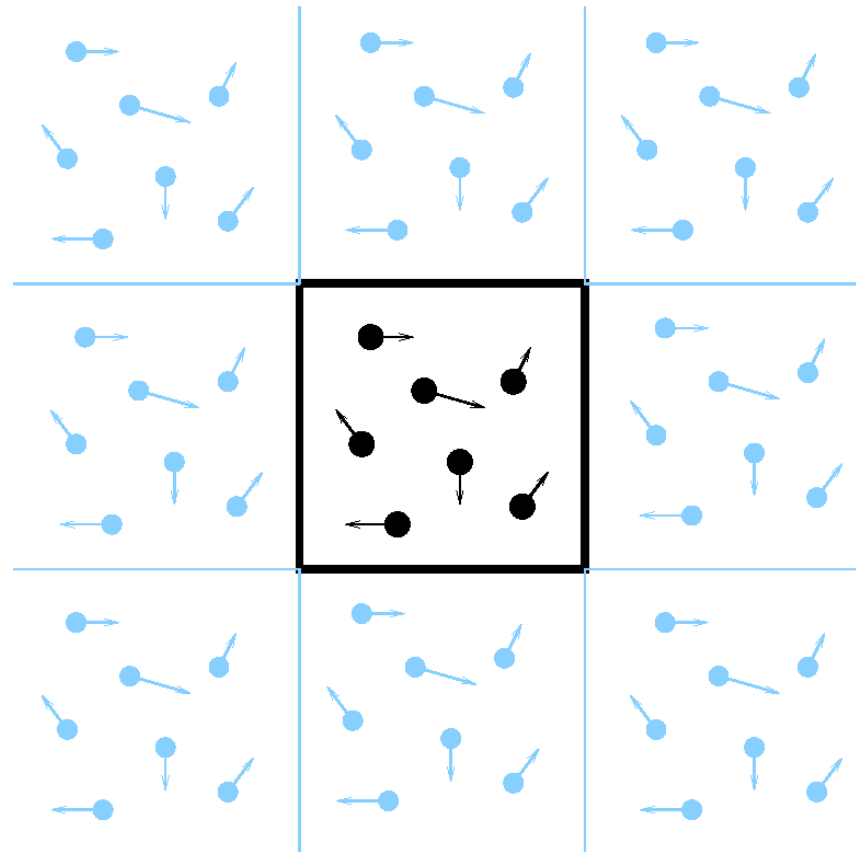
Simplified schematic of the molecular dynamics algorithm



Dinamica Molecolare

Simulo il comportamento di un insieme di atomi, per un certo periodo di tempo, tenendo conto delle interazioni esistenti tra atomi (legami, cariche elettriche...).

Gli atomi saranno soggetti a forze, che ne spostano la posizione.



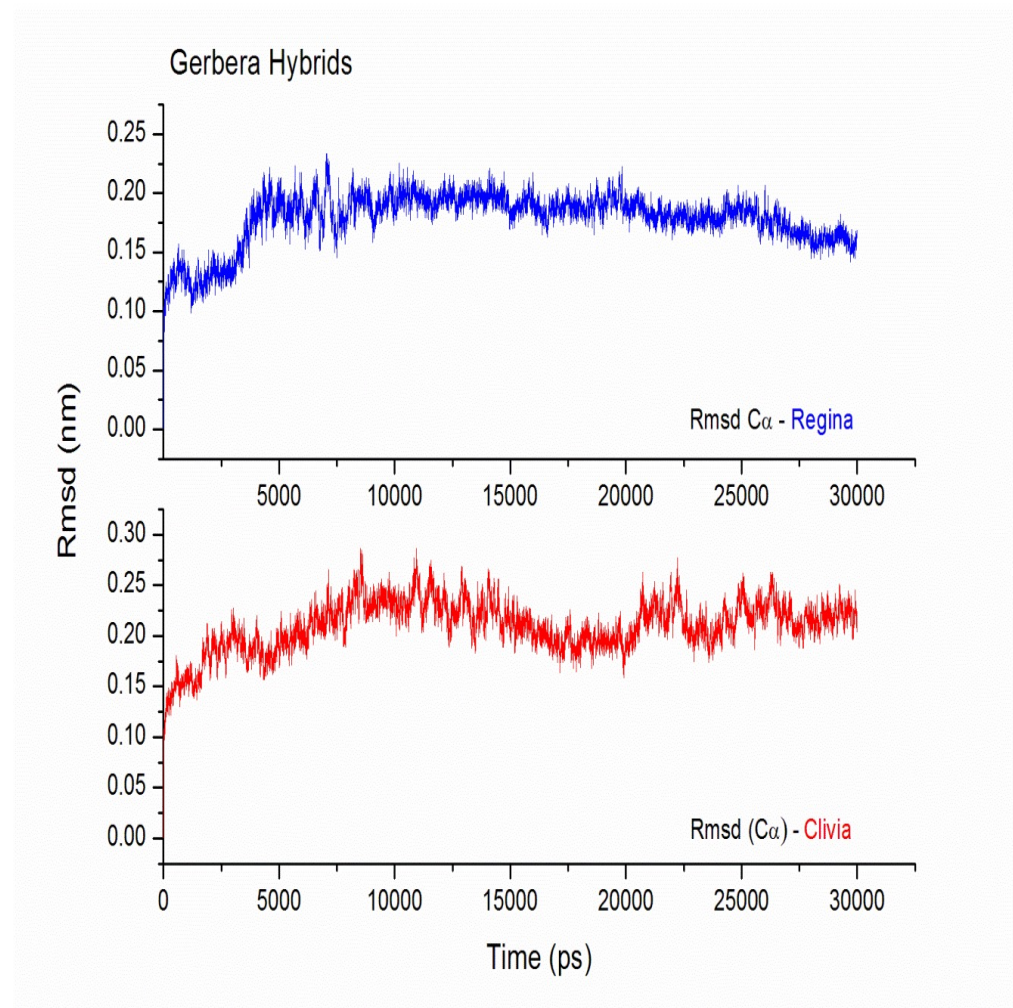
Dinamica Molecolare: Risultati

Con la dinamica molecolare posso simulare l'**evoluzione temporale** di una (macro)molecola in soluzione.

Per le macromolecole si arriva anche a simulare una durata temporale dell'ordine di ms.

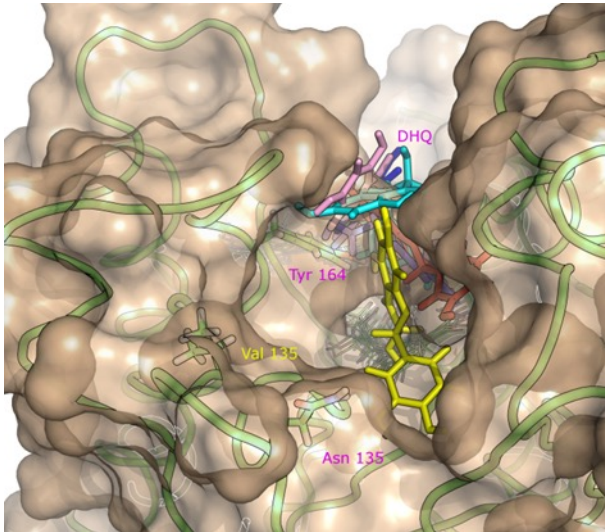
Tramite la dinamica molecolare è possibile avere informazioni sulla flessibilità di una data molecola

Spesso viene riportato lo scarto quadratico medio tra la struttura molecolare al tempo t e la struttura molecolare al tempo iniziale t_0

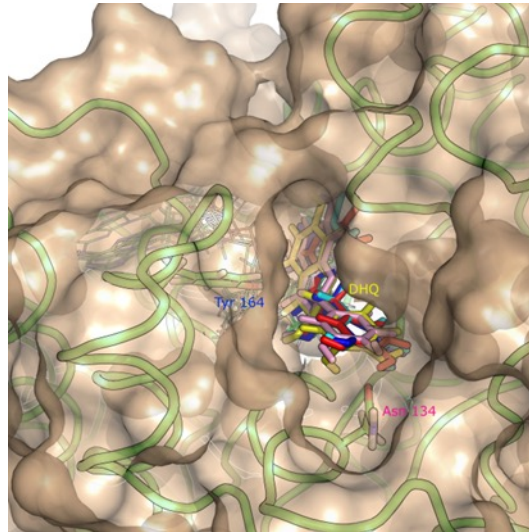


Dinamica Molecolare: risultati

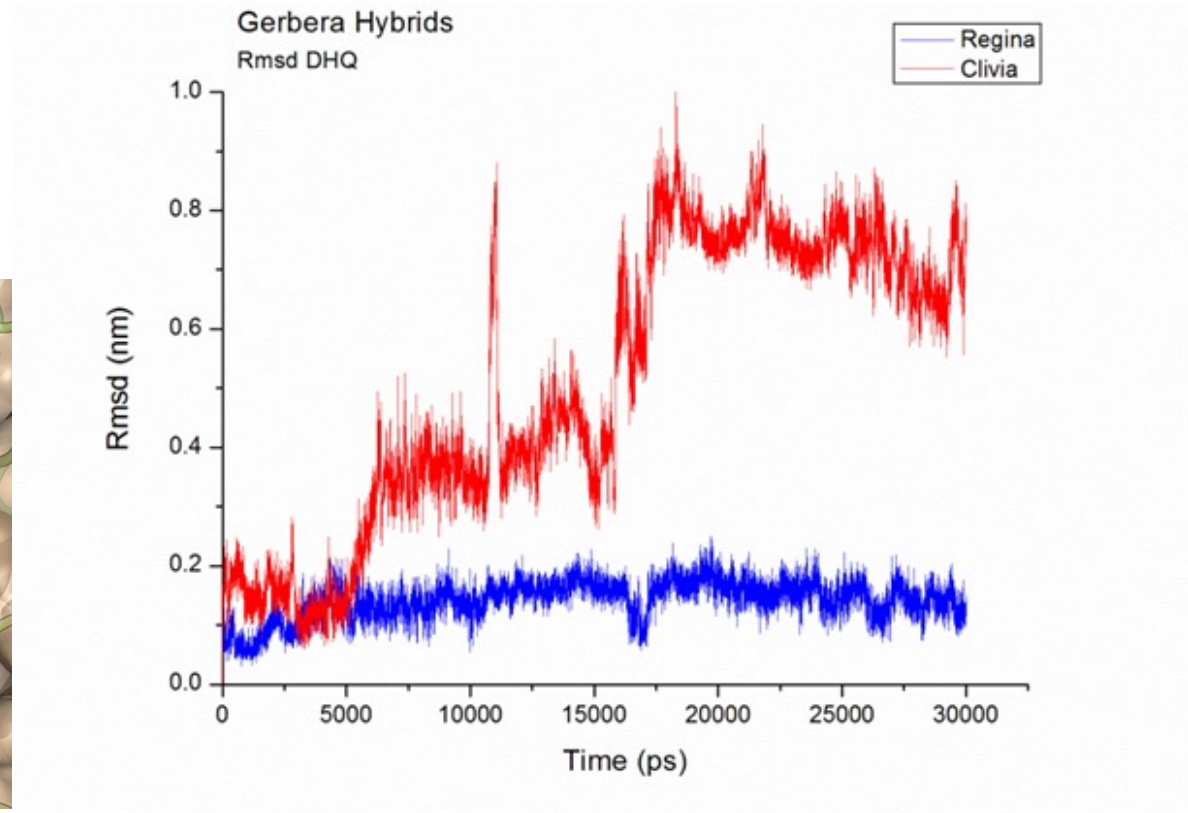
Si può simulare il comportamento nel tempo di un ligando interagente con la macromolecola, quindi valutare la stabilità dell'interazione



"Clivia"

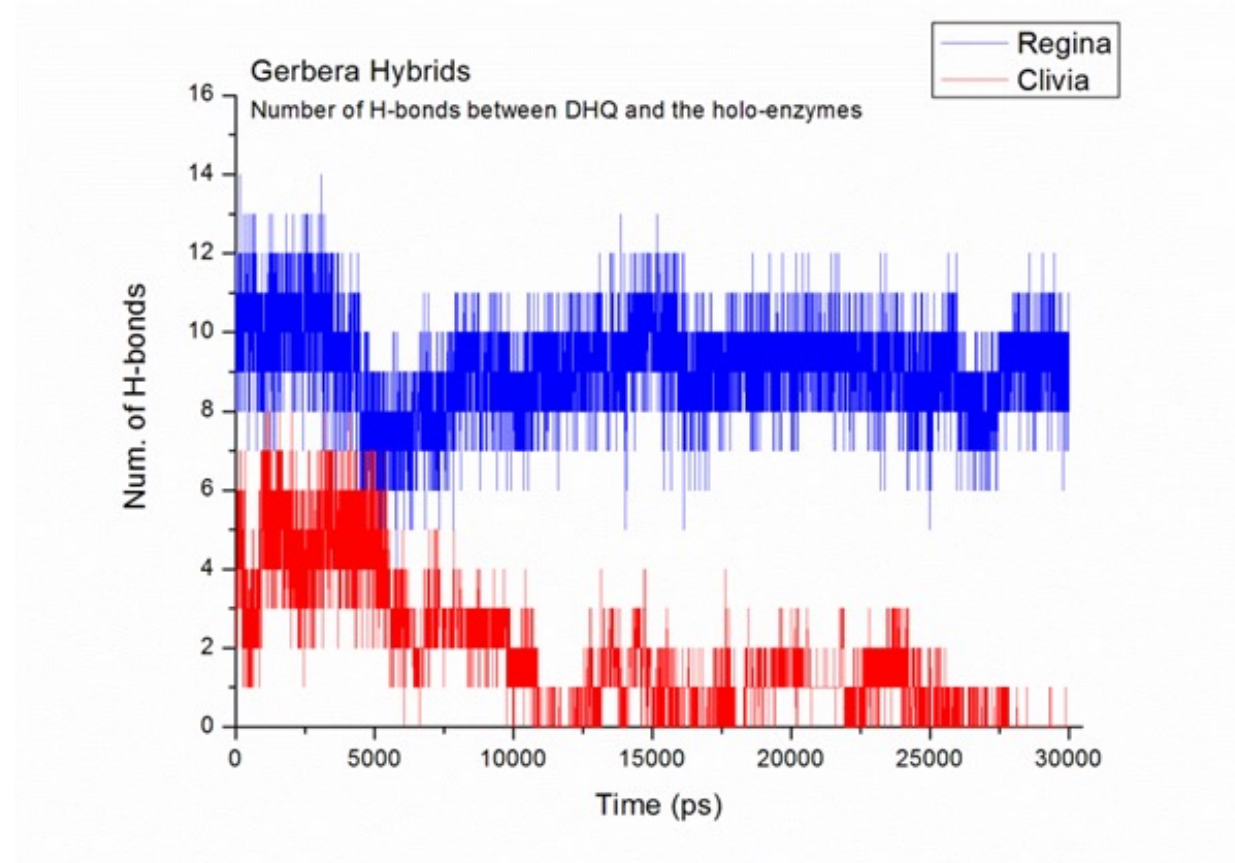


"Regina"



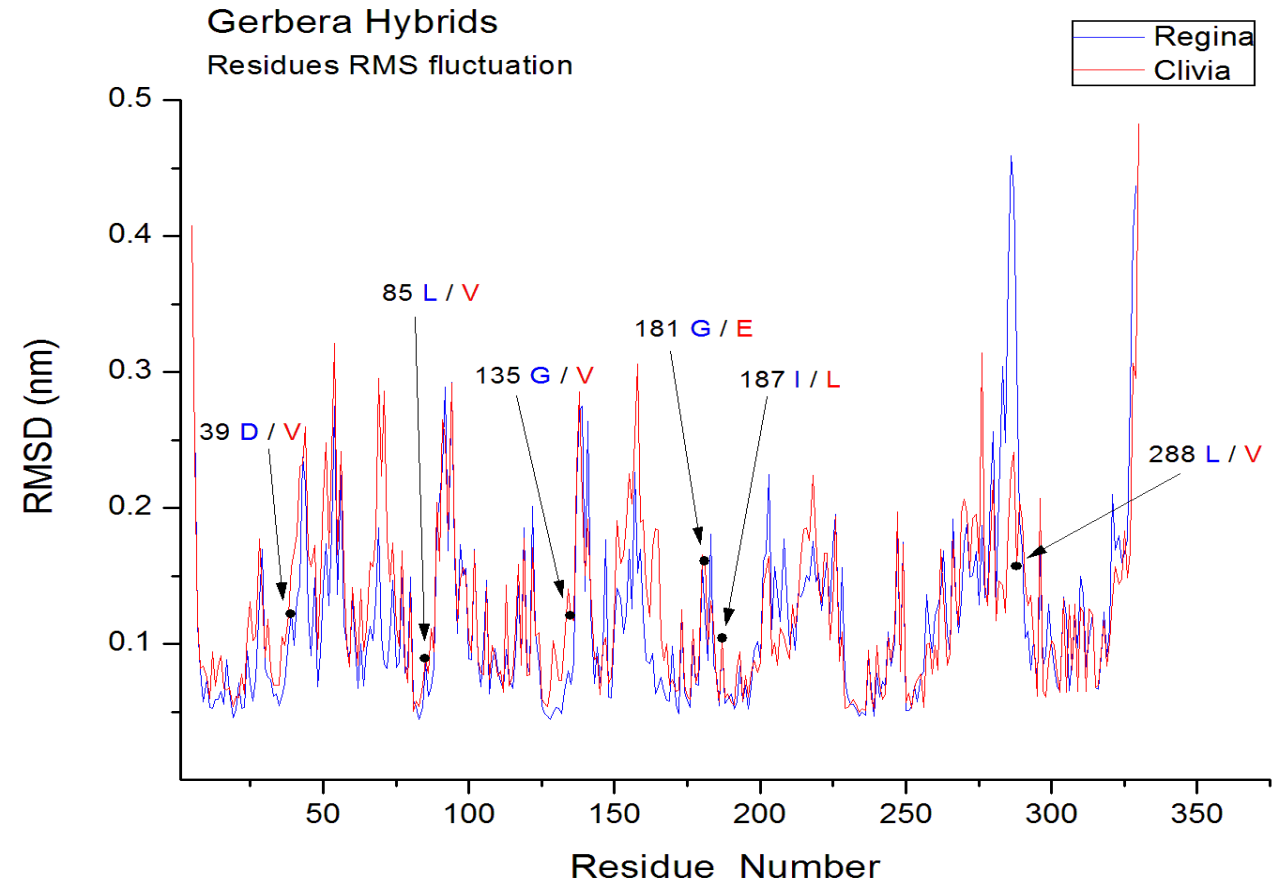
Dinamica Molecolare: Risultati

La stabilità di un'interazione proteina ligando può anche essere valutata analizzando l'evoluzione temporale dei legami idrogeno (o un'altra interazione specifica)



Dinamica Molecolare: flessibilità e dinamica

E' possibile valutare la flessibilità delle varie parti della macromolecola, studiando quanto 'si muovono' nel corso del tempo.



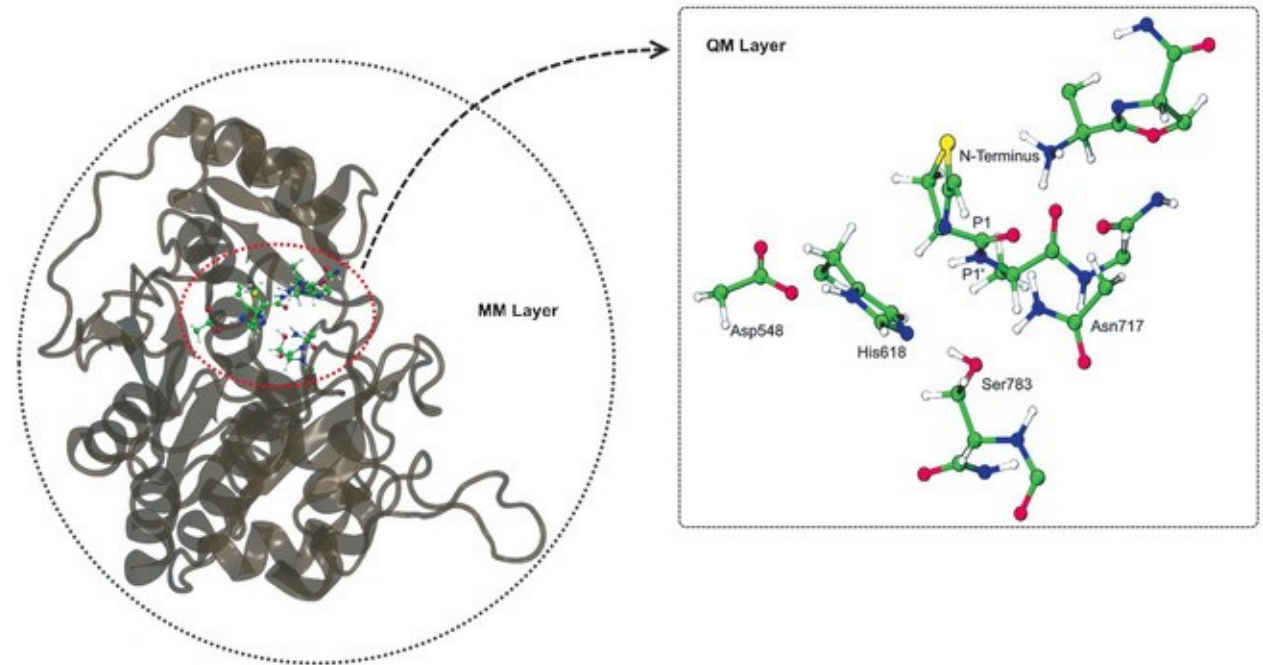
MM/QM

Una variante è la MM-QM, in cui una parte del sistema è trattato quantisticamente, mentre la parte rimanente è trattata in modo classico (come visto per la dinamica molecolare)

MM/QM trova la sua massima applicabilità in sistemi in cui la trattazione classica è insoddisfacente:

- Interazioni con metalli di transizione
- Descrizione di reazioni chimiche

MM/QM è computazionalmente molto più impegnativa della dinamica molecolare 'classica'



Modeling Molecolare

Modelling

Lo scopo del modelling è quello di generare per via puramente computazionale la struttura tridimensionale di una proteina, quando non ho a disposizione un modello sperimentale della mia macromolecola.

A differenza della dinamica molecolare, che ha comunque scopi diversi, **il modelling delle proteine non si basa su modelli 'fisici' della macromolecola** ma principalmente su **relazioni empiriche** dettate dalla similitudine tra la proteina di cui voglio costruire il modello (**target**) e le proteine di struttura nota (**templates**) che utilizzo per costruire il modello del target

Ci sono vari approcci, tra i più utilizzati:

- **Homology Modelling** (Modelling Comparativo)
- **Modelling by Threading**

Homology Modelling (Comparative Modelling)

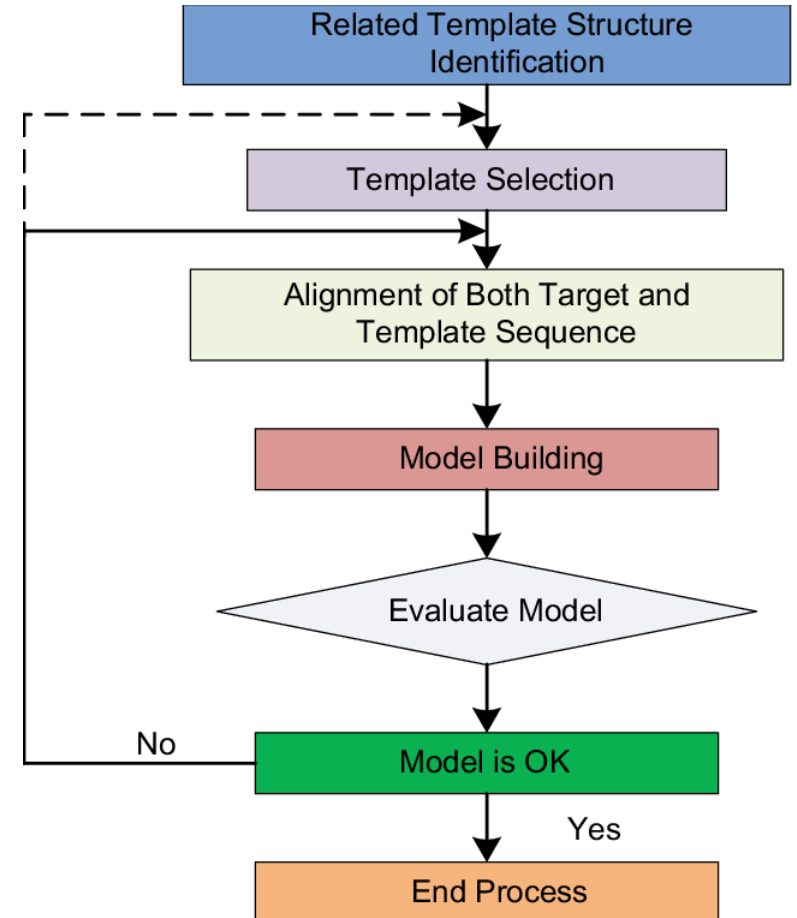
Si basa su di un principio generalmente vero:

Proteine con sequenza amminoacidica simile, hanno struttura tridimensionale simile

A partire dalla similitudine con la struttura primaria, viene generata una struttura basata una o più strutture sperimentalmente già determinate.

La struttura così generata, ma contenente gli aminoacidi della proteina che stiamo studiano (target), viene quindi sottoposta ad una minimizzazione di Energia, o una funzione pseudo-Energetica.

I vari modelli così generati vengono validati e vengono classificati sulla base della loro qualità.



Homology Modelling: limiti e problemi

1) La qualità dei modelli prodotti è strettamente legata all'identità di sequenza tra la proteina target e le proteine modello.

In genere si prende come limite di applicabilità il 20% di identità, ma **la qualità del modello è in generale tanto più bassa quanto più bassa è l'omologia di sequenza.**

2) Le zone 'flessibili' delle proteine (loops), per cui non esiste un'informazione strutturale sperimentale, sono generalmente di bassa qualità, ovvero prodotte molto male.

Homology Modelling è in genere piuttosto rapido e funziona bene per elevate omologie, ad esempio mutanti di una proteina la cui struttura è già nota sperimentalmente.

Modelling by Threading

Alla base del Modelling by Threading c'è il seguente principio: **Il numero di architetture (fold) assunte dalle proteine non è infinito.** L'architettura molecolare della proteina target sarà una tra quelle già note.

Come l'homology modelling è basato sul modelling di una proteina target a partire da un certo numero di templati scelti per similitudine di sequenza. I templati sono rappresentativi dei folds esistenti.

- A differenza dell'homology modelling non è richiesta l'esistenza di una proteina 'omologa' di struttura nota.
- La sequenza del target viene comparata con una library di proteine di fold (architettura molecolare) noto.
- Il fold della proteina viene assegnato sulla base della **similitudine** di sequenza.
- La struttura viene quindi raffinata e valutata.

Homology e Threading

Il modelling by Threading si applica a casi 'difficili', dove non esiste una proteina di struttura nota, omologa alla proteina target.

L'evoluzione dei metodi di modelling ha attenuato le differenze tra questi due approcci.

I metodi moderni di modelling fanno ampio uso di tutte le informazioni derivanti dagli allineamenti di sequenza, le relazioni evolutive tra sequenze e le strutture macromolecolari note sperimentalmente

Modelling: Server

Esistono molti server che producono predizioni di strutture molecolari a partire dalla sequenza aminoacidica, tra i più noti e utilizzati (e gratuiti):

Swiss-Model: <https://swissmodel.expasy.org/>

Phyre2: <http://www.sbg.bio.ic.ac.uk/phyre2>

RaptorX: <http://raptorx.uchicago.edu/>

I-Tasser: <https://zhanggroup.org/I-TASSER/>

Rosetta: <https://www.rosettacommons.org/software/servers>

Modeller: <https://modbase.compbio.ucsf.edu/modweb/>

Alcuni di questi server richiedono una *key*, gratuita

AlphaFold2

Article | [Open Access](#) | [Published: 15 July 2021](#)

Highly accurate protein structure prediction with AlphaFold

[John Jumper](#) , [Richard Evans](#), [...] [Demis Hassabis](#) 

[Nature](#) **596**, 583–589 (2021) | [Cite this article](#)

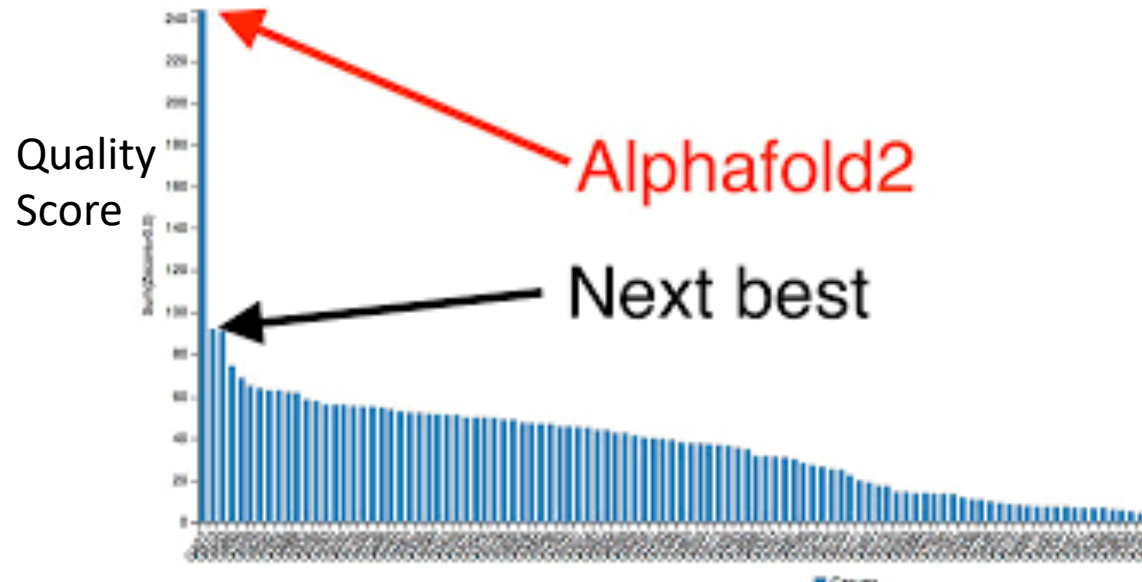
410k Accesses | **69** Citations | **2803** Altmetric | [Metrics](#)

AlphaFold2 è un programma prodotto da DeepMind (Google).

E' basato sull'utilizzo di metodi di Intelligenza Artificiale applicato al problema del riconoscimento del folding.

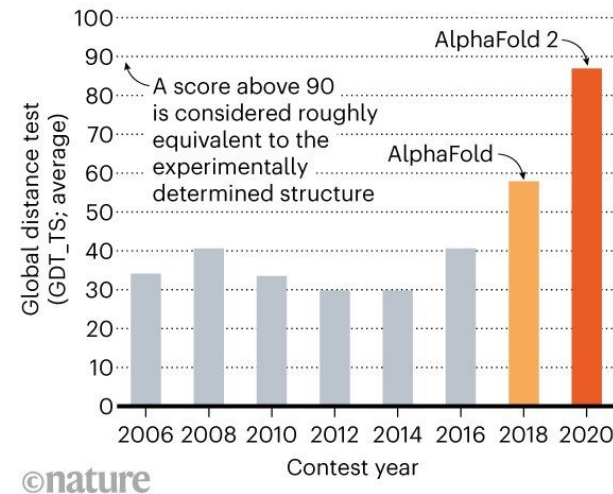
AlphaFold2

Nel Casp14 ('gara' di predizione di strutture incognite) ha surclassato ogni altro software ottenendo risultati impensabili solo tre anni fa



STRUCTURE SOLVER

DeepMind's AlphaFold 2 algorithm significantly outperformed other teams at the CASP14 protein-folding contest — and its previous version's performance at the last CASP.



AlphaFold2: Deep Learning

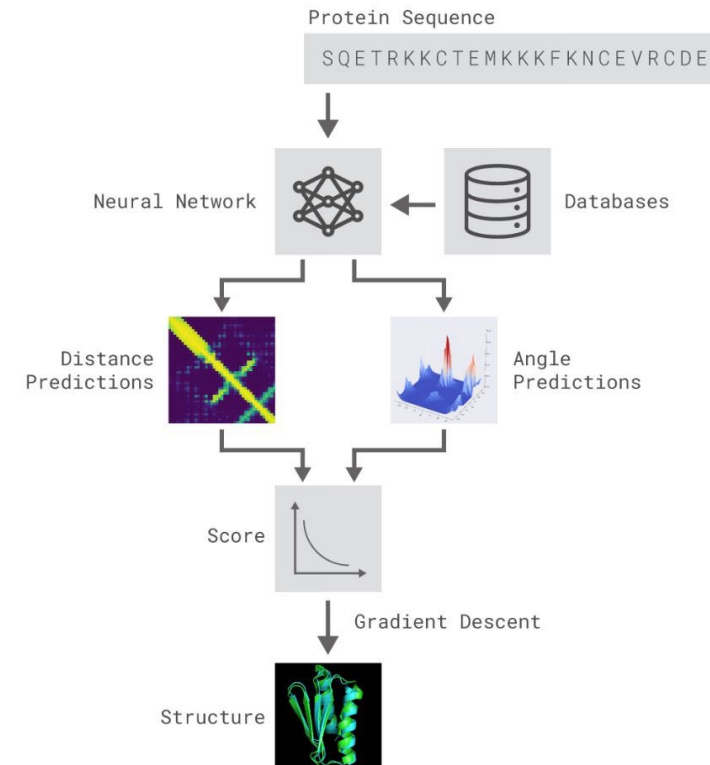
AlphaFold2, basato è su algoritmi di intelligenza artificiale per l'analisi di Big Data

- Relazione tra sequenza e struttura secondaria, accessibilità al solvente, angoli di torsione...
- Relazione tra sequenza e contatti interatomici all'interno di una struttura molecolare

Sono generati dei databases contenenti le relazioni tra sequenze e dati utilizzati.

<https://alphafold.ebi.ac.uk/>

<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>



AlphaFold2 – All proteins database

<https://www.alphafold.ebi.ac.uk/>

AlphaFold
Protein Structure Database

Developed by DeepMind and EMBL-EBI

Search for protein, gene, UniProt accession or organism BETA Search

Examples: [Free fatty acid receptor 2](#) [At1g58602](#) [Q5VSL9](#) [E. coli](#) Help: [AlphaFold DB search help](#)

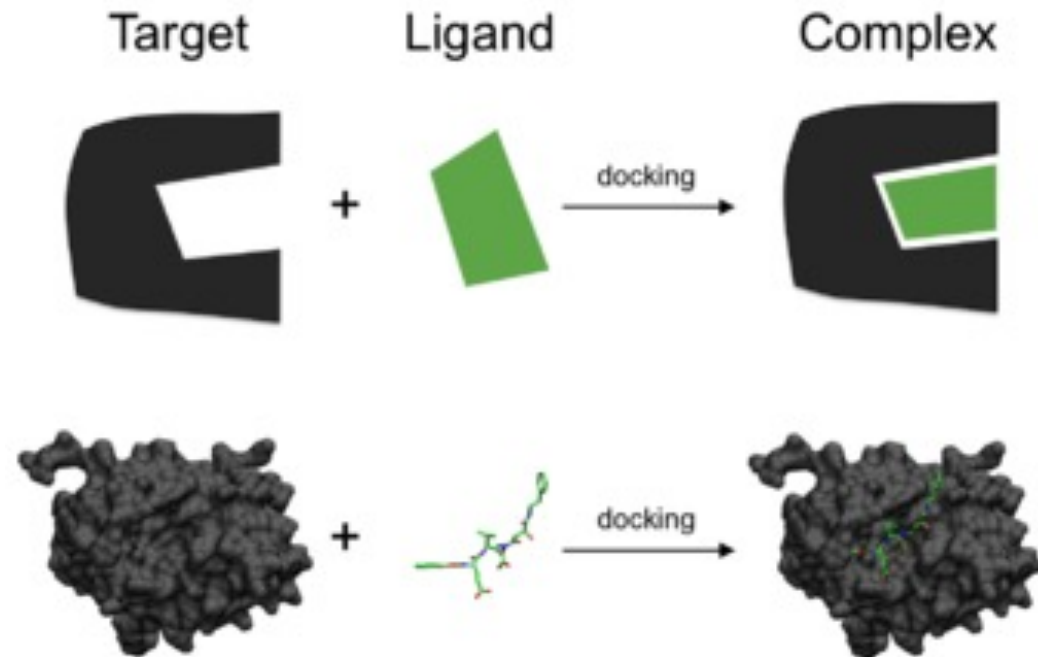
Feedback on structure: [Contact DeepMind](#)

Ligand Docking

E' un metodo computazionale per **identificare molecole (inibitori) potenzialmente in grado di interagire con il mio target (proteina)** o più semplicemente per **ipotizzare dove il sito di interazione tra target e inibitore (ligando)**.

Il ligando può essere sia una macromolecola (es: un anticorpo) che una piccola molecola (farmaco potenziale)

Il protocollo di docking può essere automatizzato in un vero e proprio virtual screening

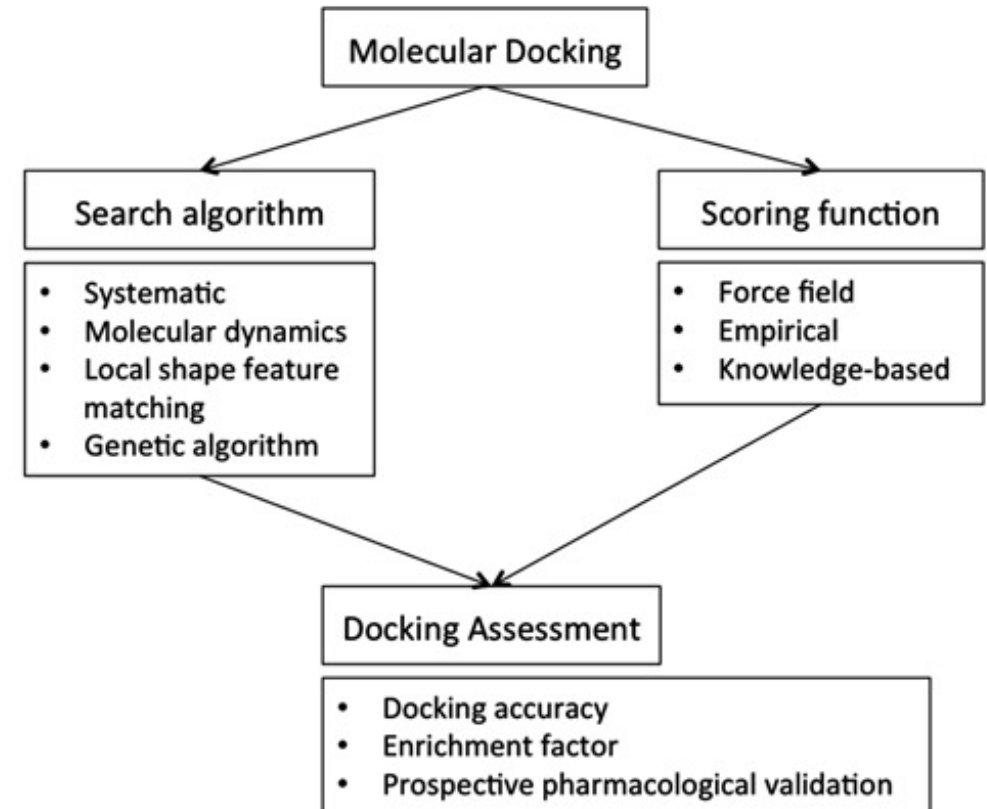


Docking: protocollo

A partire dalla struttura nota (sperimentalmente o per modelling) del mio target , iterativamente posiziono l'inibitore in punti diversi del target, valutando la 'compatibilità' (plausibilità dell'interazione) per mezzo di una scoring function.

Al termine del procedimento ho una 'classifica' (score) dei 'modi di interazione più probabili.

In questo modo posso comparare diverse molecole ma anche valutare potenziali siti di interazione (sul target) con gli inibitori e decidere quali sono le soluzioni più promettenti



Docking: Software e Servers

Anche per il docking esistono molti server in grado di fare predizioni (più o meno attendibili) sull'efficacia e le modalità di interazione tra proteina e ligando.

Server - Proteina-Proteina

- HDock: <http://hdock.phys.hust.edu.cn/> (proteina-proteina)
- ZDOCK: <https://zdock.umassmed.edu/> (proteina-proteina)

Server - Proteina- piccole molecole

Swiss-Dock: <http://www.swissdock.ch/>

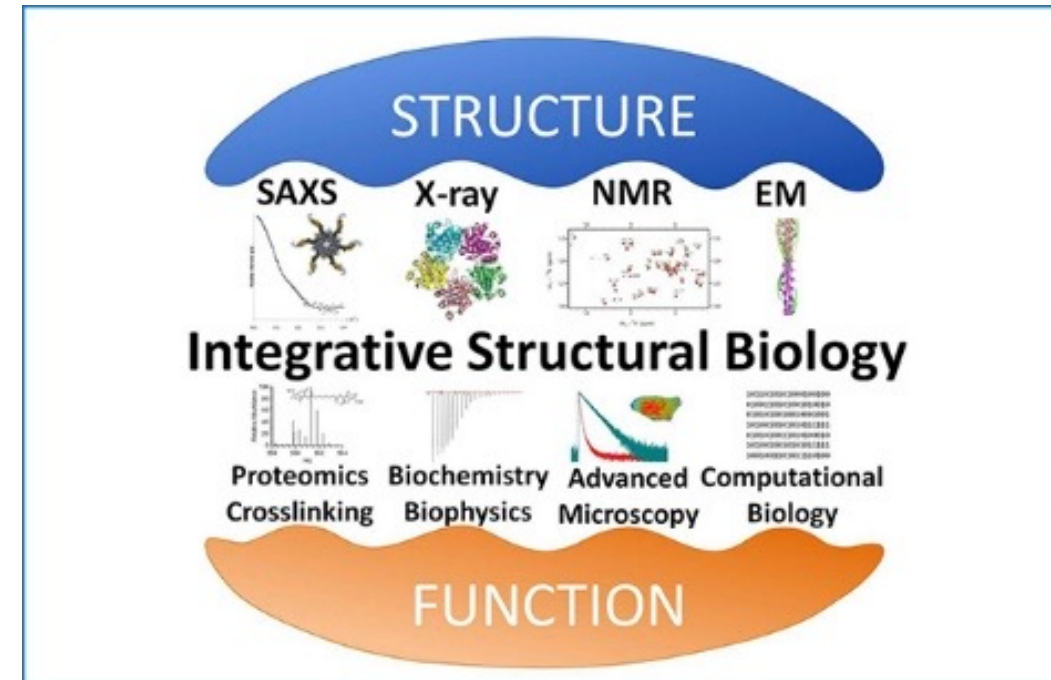
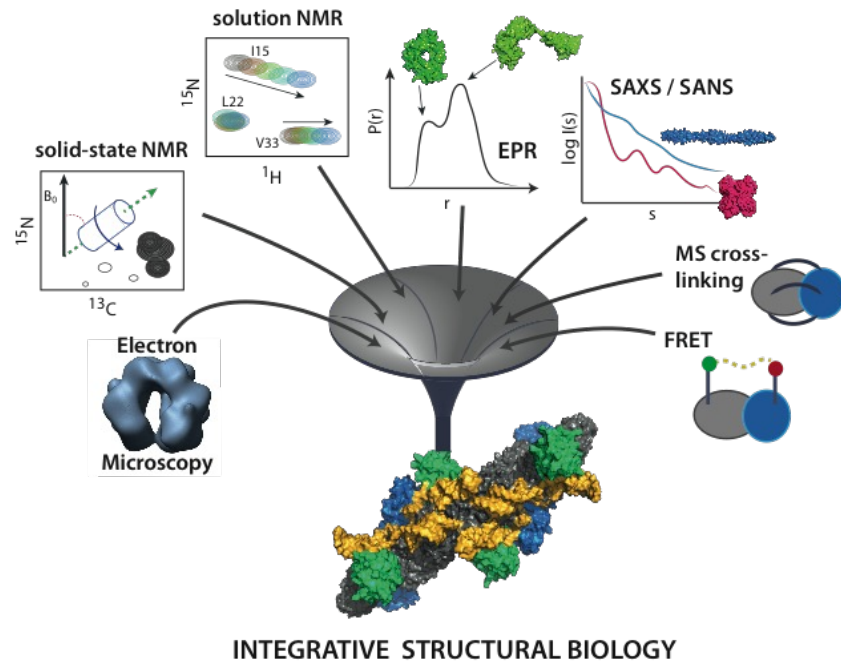
Software – Proteina – piccole molecole

- Autodock: <https://autodock.scripps.edu/>
- Autodock – Vina: <https://vina.scripps.edu/>

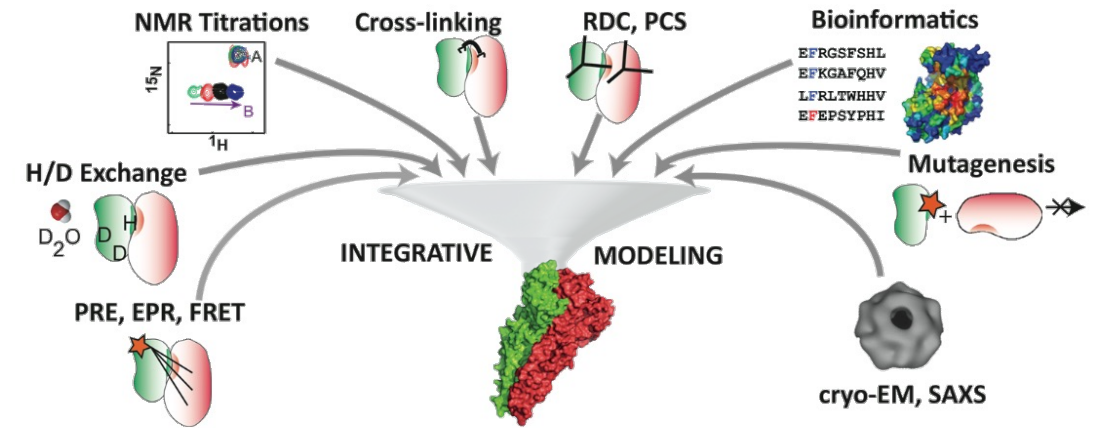
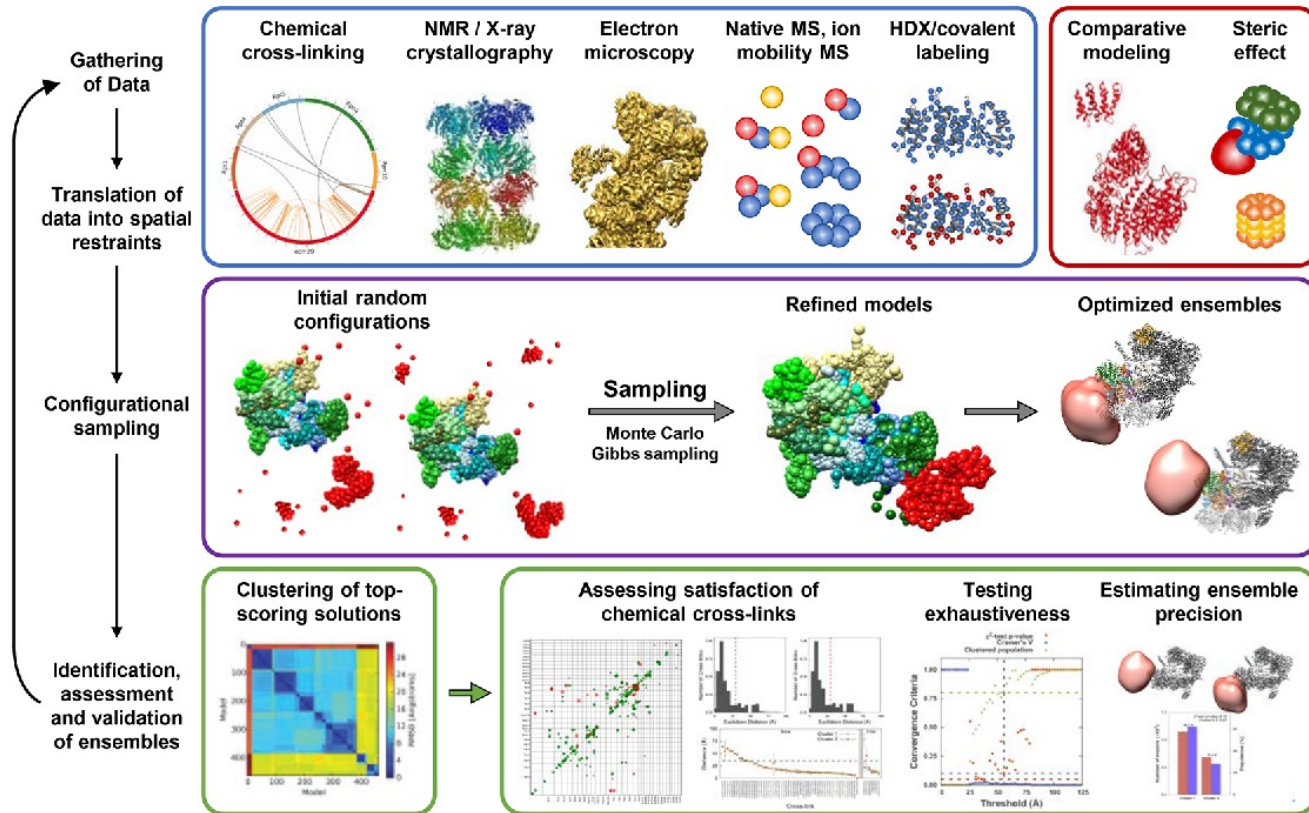
https://en.wikipedia.org/wiki/List_of_protein-ligand_docking_software

Integrative Structural Biology

Come abbiamo visto il biologo strutturale ha **diversi strumenti** a disposizione, basati su **principi fisici differenti** e in grado di fornire **informazioni strutturali diverse**. A volte la loro combinazione permette di ottenere un'informazione più approfondita e completa, si parla di Biologia Strutturale Integrata (*Integrative Structural Biology*)



Combinare le tecniche



La combinazione di tecniche, individualmente non in grado di fornire un modello sperimentale 3D, può portare a ipotizzare/definire un modello tridimensionale ragionevole.

HADDOCK

HADDOCK è un software di Docking che fa pieno uso di informazioni sperimentali, note da altri esperimenti, al fine di migliorare la qualità e l'attendibilità del docking.

