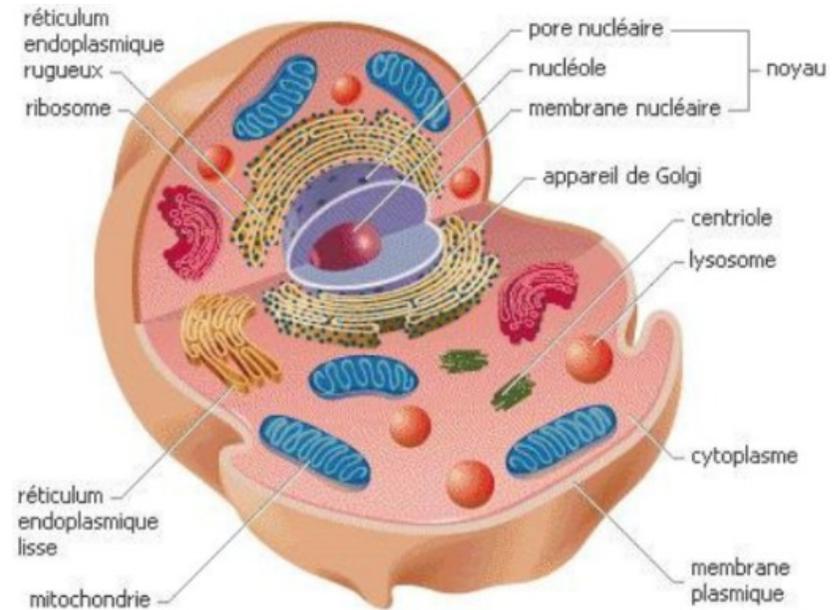
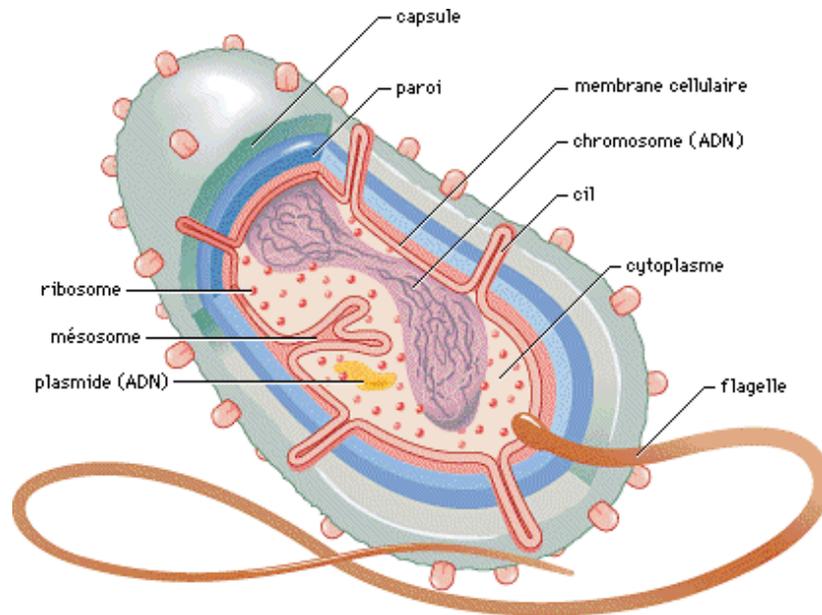


# INTERACTIONS NON COVALENTES ET RECONNAISSANCE MOLECULAIRE AU SEIN DU VIVANT

Tâp HA DUONG  
BioCIS – HM1 – 1804  
[tap.ha-duong@universite-paris-saclay.fr](mailto:tap.ha-duong@universite-paris-saclay.fr)

# Introduction : Bases moléculaires au sein du vivant

- La cellule est l'unité de base de tous les êtres vivants.
- Contrairement aux procaryotes, les eucaryotes sont compartimentés en organites spécialisés où se déroulent des processus métaboliques spécifiques.



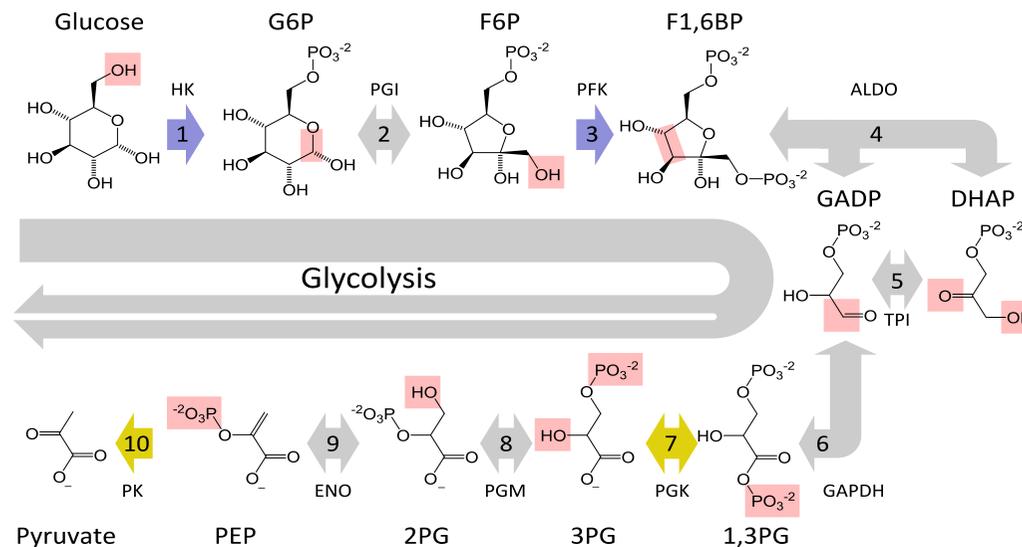
# Introduction : Bases moléculaires au sein du vivant

---

- Les activités principales d'une cellule sont :
  - se nourrir
  - produire de l'énergie
  - transformer des molécules
  - communiquer
  - se reproduire
  - se réparer
  - mourir
- Ces activités sont généralement réalisées par des « usines cellulaires » (plus ou moins localisées dans la cellule) qu'on appelle **processus métaboliques** ou **voies métaboliques**.

# Introduction : Bases moléculaires au sein du vivant

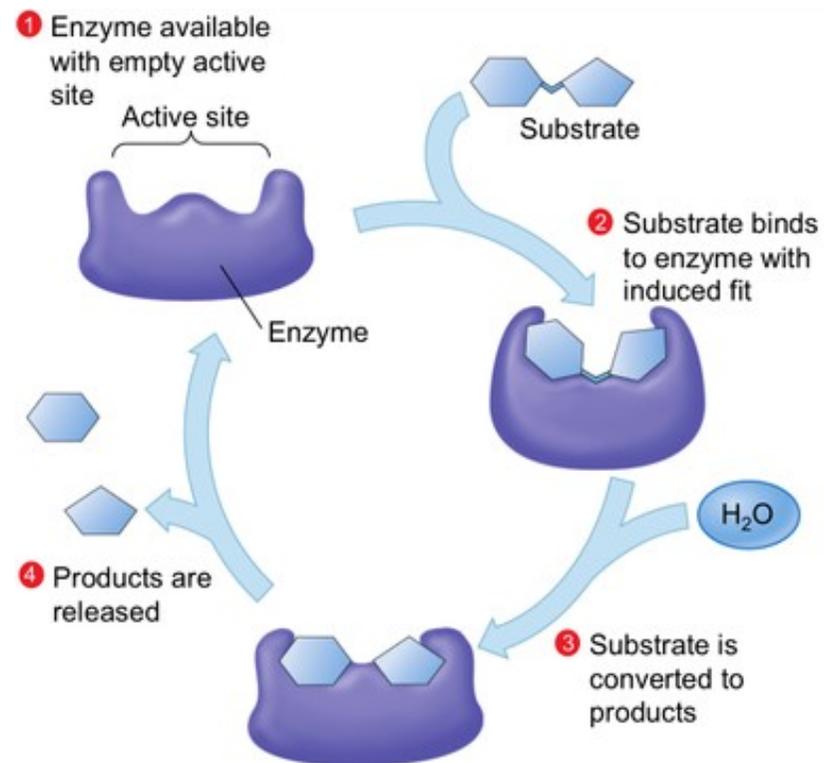
- Processus ou voie métabolique = ensemble de **transformations physiques ou chimiques** que subissent les substances introduites ou formées dans la cellule...
- Exemple : la glycolyse = voie métabolique de dégradation du glucose en pyruvate pour produire de l'ATP.



# Introduction : Bases moléculaires au sein du vivant

---

- Base moléculaire d'une catalyse enzymatique :



# Introduction : Bases moléculaires au sein du vivant

---

- Processus ou voie métabolique = ensemble de **transformations physiques ou chimiques** que subissent les substances introduites ou formées dans la cellule...
- ... et qui sont effectuées par des **complexes de biomolécules** (constitués principalement de protéines) qu'on peut appeler « machines moléculaires ».

Voie métabolique = ensemble des actions de complexes biomoléculaires



« Usine cellulaire » = ensemble des actions de « machines moléculaires »

# Introduction : Bases moléculaires au sein du vivant

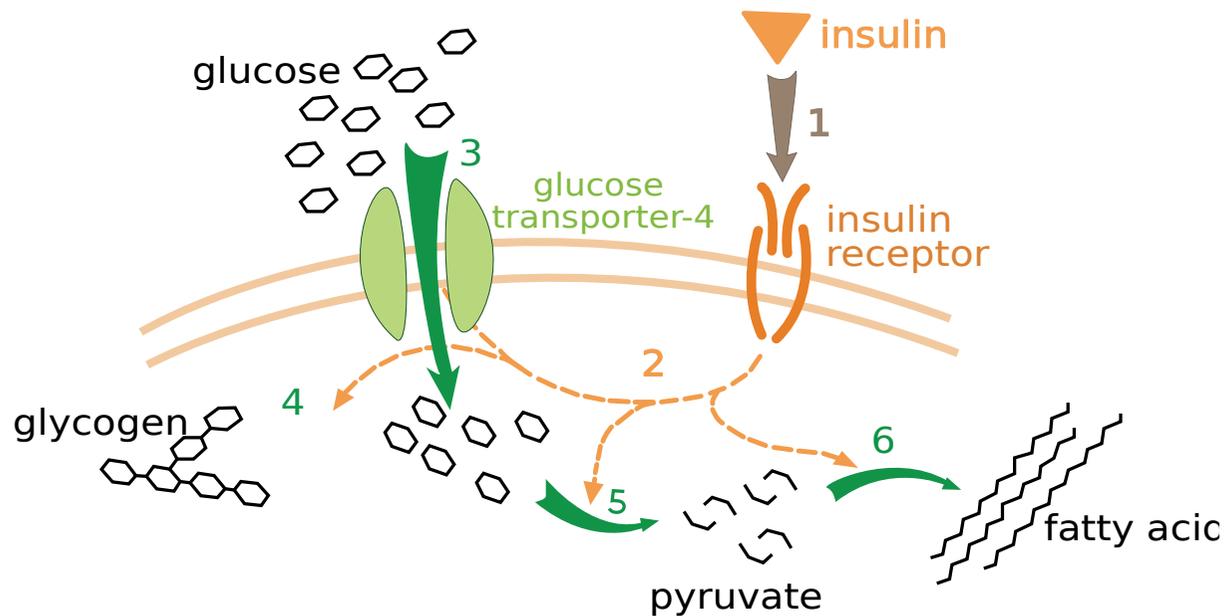
---

- Les activités principales d'une cellule sont :
  - se nourrir
  - produire de l'énergie
  - transformer des molécules
  - communiquer
  - se reproduire
  - se réparer
  - mourir
- Les voies métaboliques sont coordonnées entre elles grâce à un « réseau de communication cellulaire » appelé voie de signalisation.

# Introduction : Bases moléculaires au sein du vivant

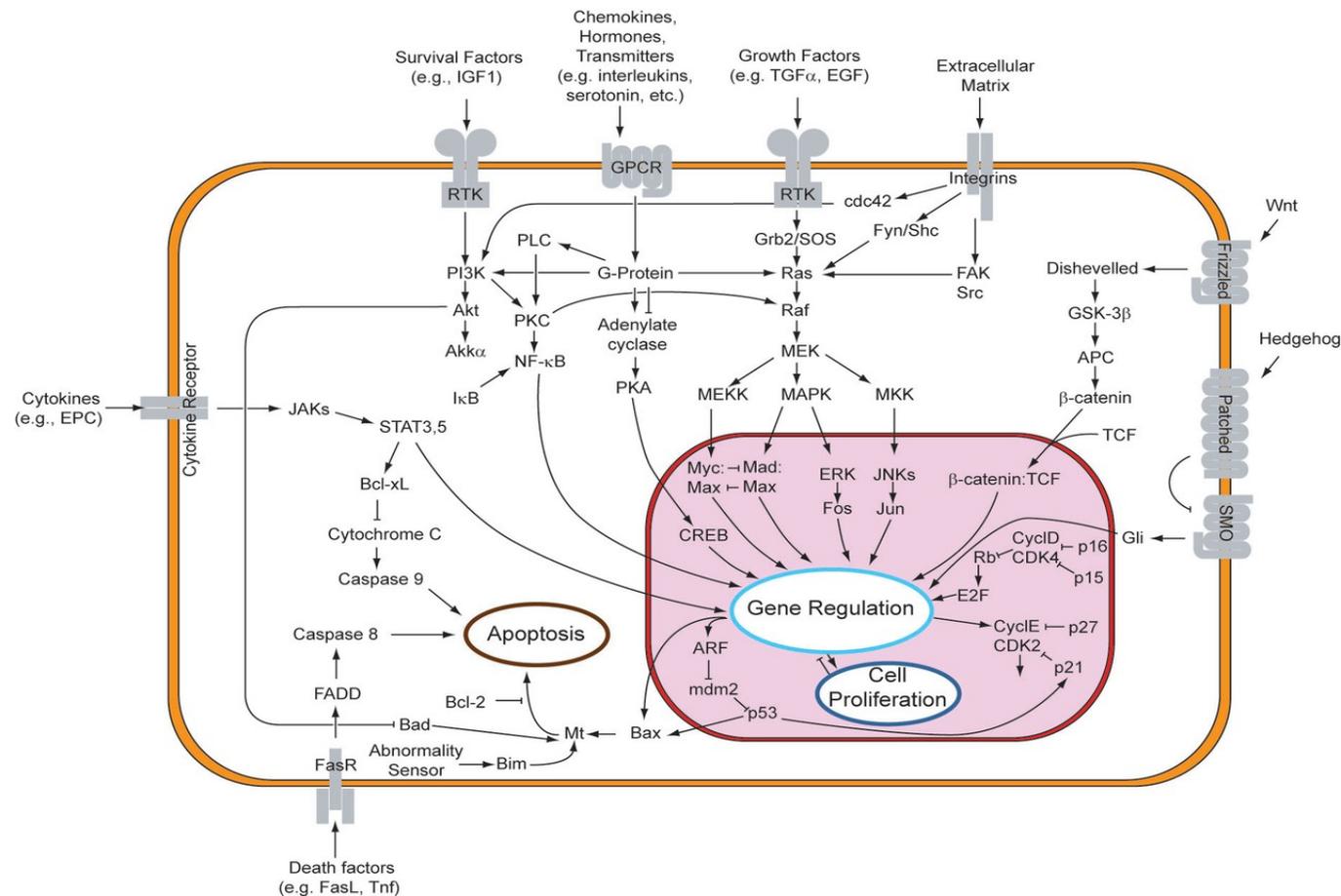
---

- Les **voies de signalisation** permettent de d'activer, de désactiver et de réguler les voies métaboliques.
- Exemple : la glycolyse est régulée par la voie de signalisation de l'insuline.



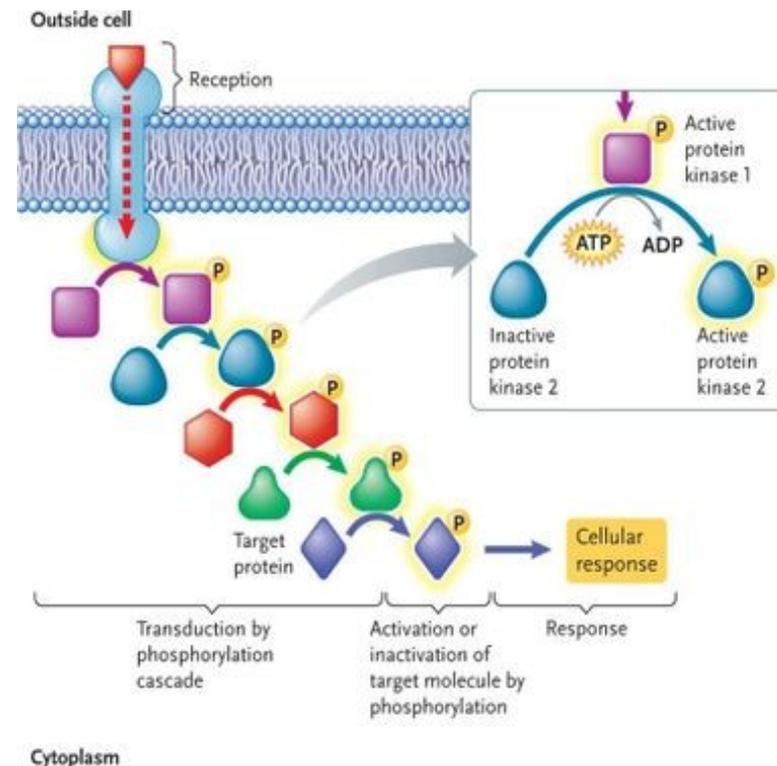
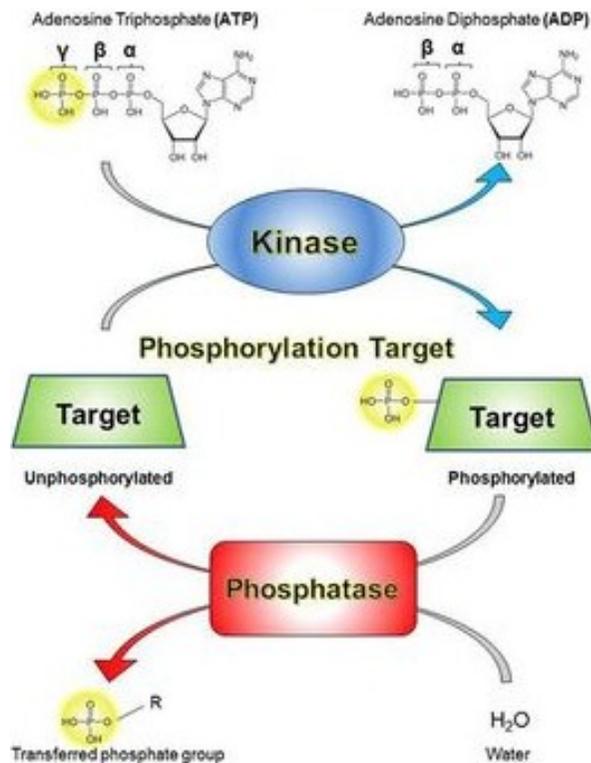
# Introduction : Bases moléculaires au sein du vivant

- Les **voies de signalisation** sont constituées d'une cascade d'activation/désactivation de protéines (très souvent, des kinases).



# Introduction : Bases moléculaires au sein du vivant

- Base moléculaire d'une signalisation : association non-covalente avec une autre protéine et modification chimique induisant une transition de structures entre un *état inactif* vers un *état actif* (*switch moléculaire*).



# Introduction : Bases moléculaires au sein du vivant

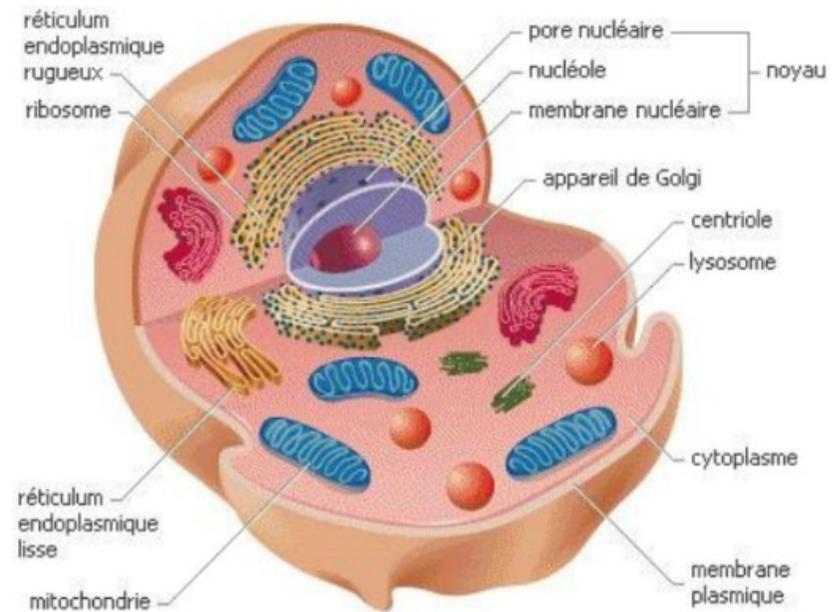
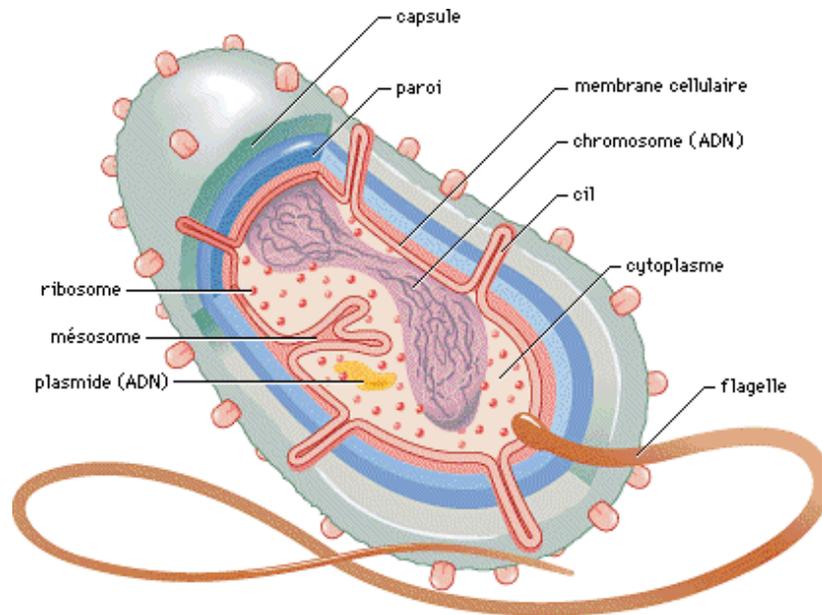
---

- Les activités principales d'une cellule sont :
  - se nourrir
  - produire de l'énergie
  - transformer des molécules
  - communiquer
  - se reproduire
  - se réparer
  - mourir
- Ces activités sont réalisées par des **voies métaboliques** et régulées par des **voies de signalisation**.
- Ces processus cellulaires sont effectués par des associations et **assemblages de biomolécules**, principalement des protéines.

# Introduction : Bases moléculaires au sein du vivant

---

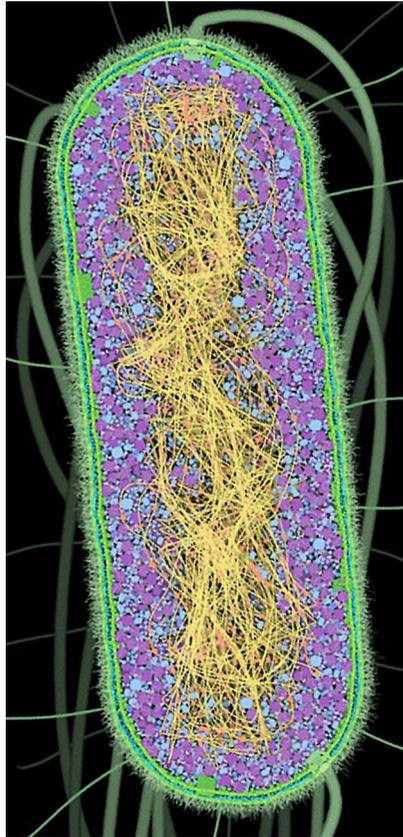
- Question : Peut-on voir ces assemblages de biomolécules à l'œuvre ?
- Réponse : Encore très difficilement au niveau cellulaire.



# Introduction : Bases moléculaires au sein du vivant

---

- Sauf pour la science artistique ou l'art scientifique...

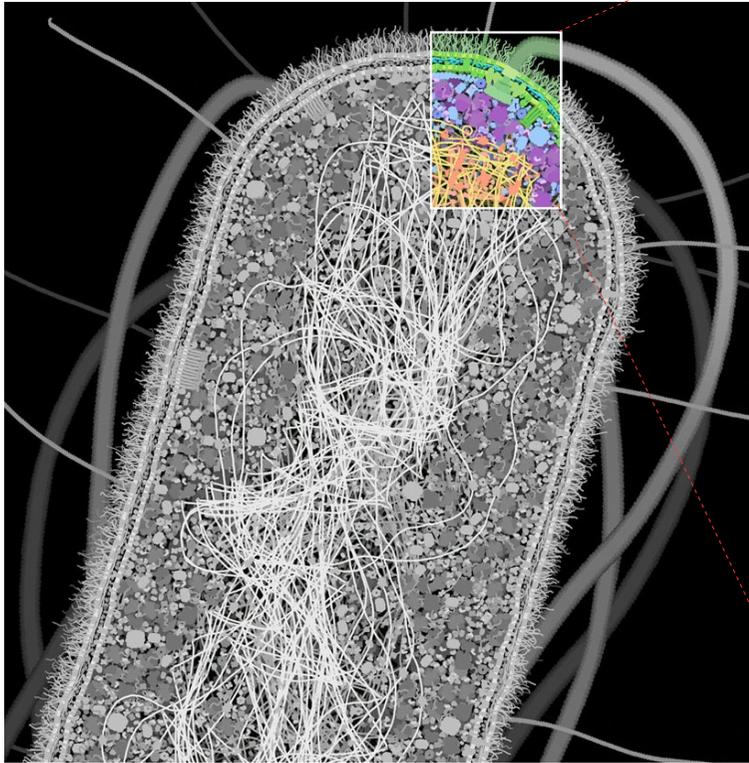


- David Goodsell : <https://ccsb.scripps.edu/goodsell/art/>

# Introduction : Bases moléculaires au sein du vivant

---

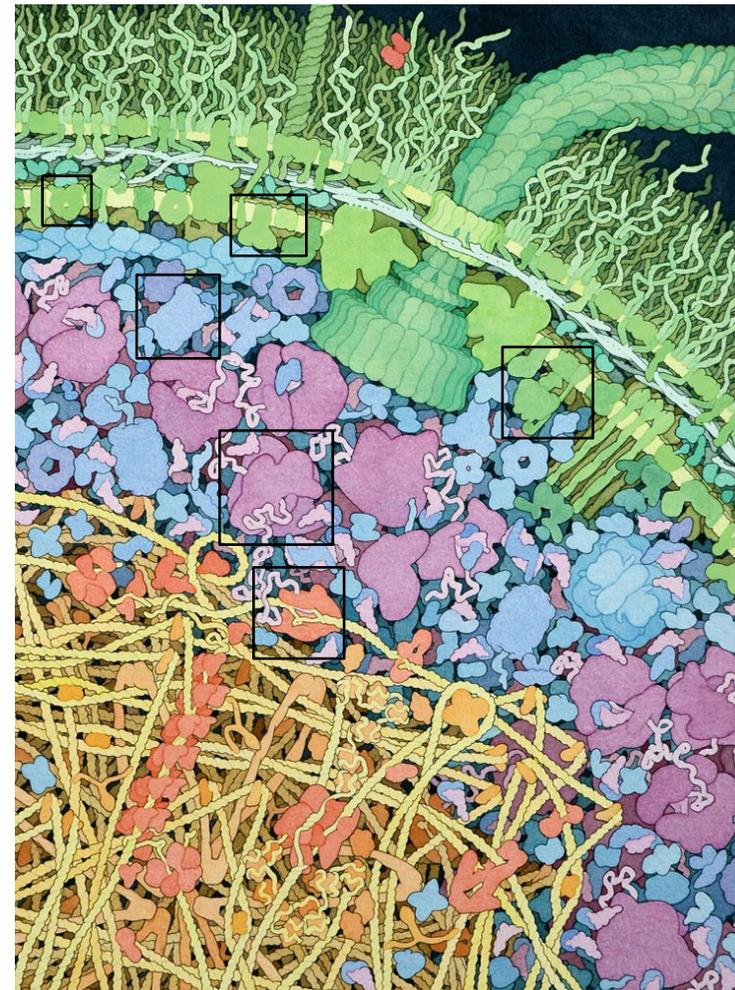
- <https://ccsb.scripps.edu/goodsell/machinery-of-life-reducedillustrations/>



# Introduction : Bases moléculaires au sein du vivant

---

- Goodsell D.S., *Biochem. Mol. Biol. Educ.* 2009, 37: 325-332



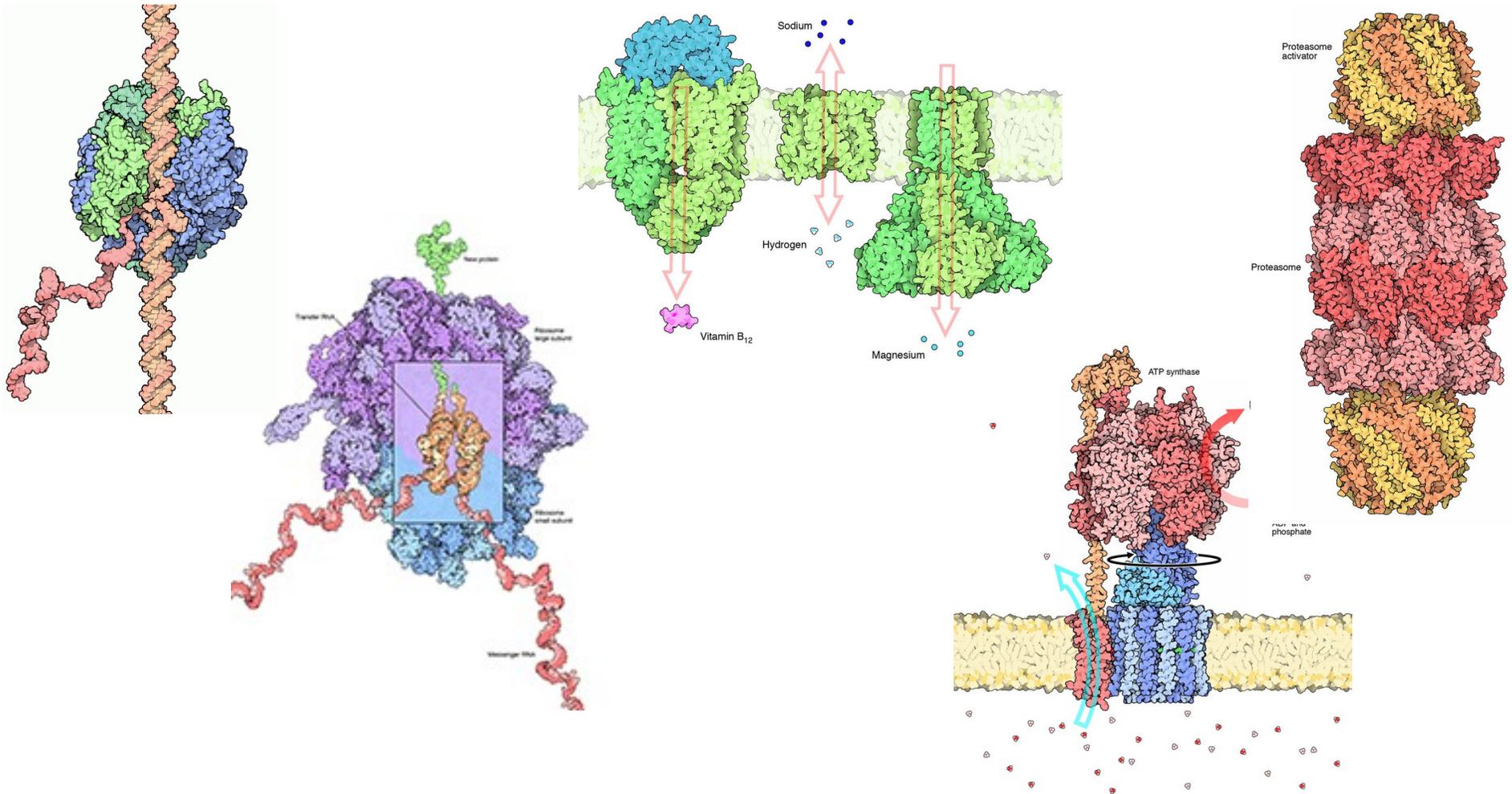
# Introduction : Bases moléculaires au sein du vivant

---

- *Extracellular*: **1**, Enterotoxin.
- *Outer membrane*: **2**, lipopolysaccharide; **3**, lipoprotein; **4**, porin; **5**, OmpA; **6**, fimbrial usher; **7**, pilus; **8**, iron transport protein FhnA.
- *Periplasm*: **9**, peptidoglycan; **10**, periplasmic binding proteins; **11**, beta-lactamase; **12**, superoxide dismutase; **13**, heat shock protein/chaperone DegP; **14**, proline isomerase FkpA.
- *Inner membrane*: **15**, magnesium transporter MglE; **16**, vitamin B12 transporter BtuCD-F; **17**, shape-determining proteins MreCD and penicillin-binding protein PBP2; **18**, mechanosensory channel MscL; **19**, molybdenum transporter ModBC-A; **20**, drug efflux pump AcrAB and TolC; **21**, magnesium transporter CorA; **22**, sodium/proton antiporter NhaA; **23**, nitrate reductase NarGHI; **24**, succinate dehydrogenase; **25**, ATP synthase; **26**, ubiquinol oxidase; **27**, aspartate receptor; **28**, signaling proteins CheAY; **29**, secretory channel SecAB; **30**, NADH dehydrogenase; **31**, zinc transporter YjiP; **32**, calcium pump.
- *Flagellar motor*: **33**, flagellum; **34**, flagellar hook; **35**, rotor; **36**, motor.
- *Cytoplasm*: **37**, cytoskeletal protein MreB; **38**, ribosome; **39**, transfer RNA; **40**, elongation factor Tu; **41**, elongation factor Ts; **42**, elongation factor G; **43**, initiation factors; **44**, aminoacyl-tRNA synthetase; **45**, chaperone GroEL; **46**, proteasome HslVU; **47**, glycolytic enzymes; **48**, tricarboxylic acid cycle enzymes; **49**, catalase; **50**, Iron superoxide dismutase; **51**, alkyl hydroperoxide reductase; **52**, phosphoenolpyruvate: sugar phosphotransferase system; **53**, nucleoside diphosphate kinase; **54**, glycerol kinase; **55**, acyl carrier protein system; **56**, aspartate carbamoyltransferase; **57**, aspartate aminotransferase; **58**, glutamine synthetase.
- *Nucleoid*: **59**, DNA; **60**, RNA polymerase; **61**, messenger RNA; **62**, catabolite activator protein; **63**, lac repressor; **64**, topoisomerase; **65**, HU; **66**, H-NS; **67**, IHF; **68**, Fis; **69**, Lrp; **70**, condensin MukBEF; **71**, RecA; **72**, RecBCD; **73**, DNA methyltransferase Hha1; **74**, DNA glycosylase MutM; **75**, DNA polymerase; **76**, single strand binding protein.

# Introduction : Bases moléculaires au sein du vivant

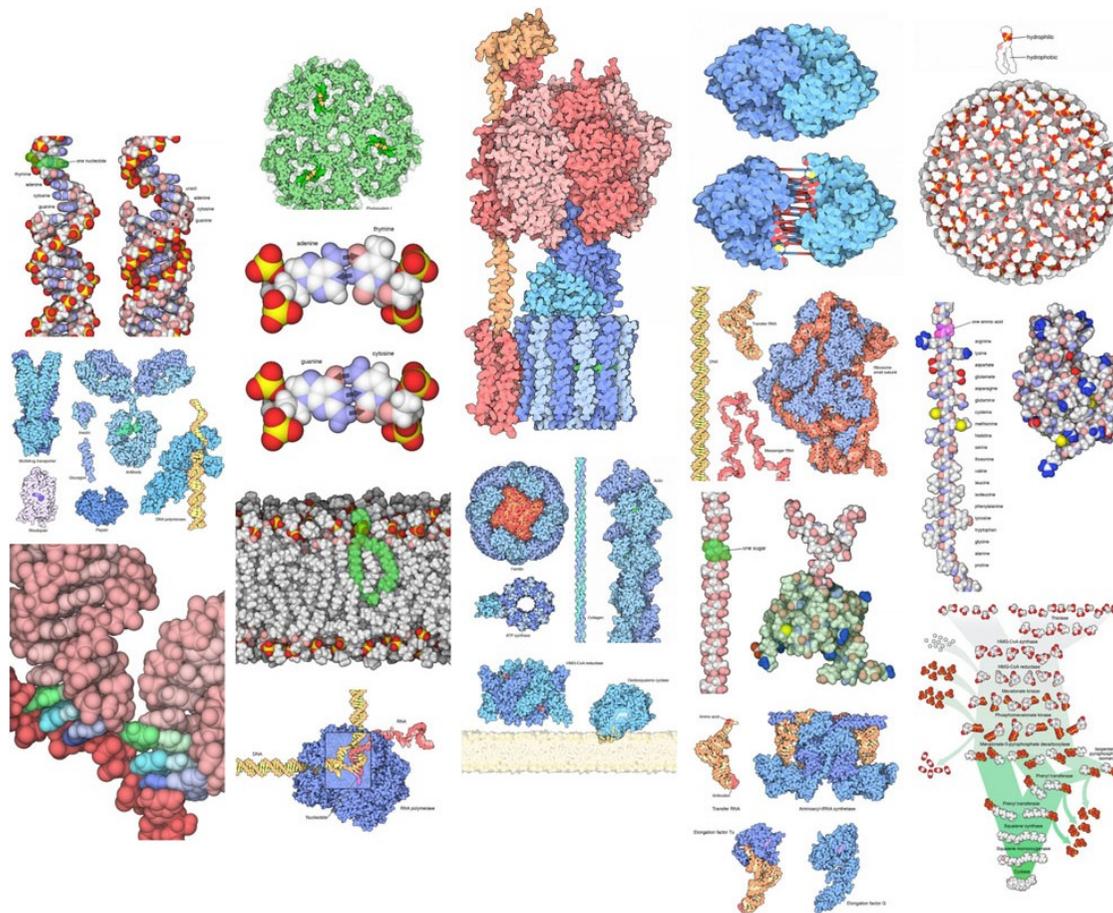
- Si on zoome encore un peu plus, on peut alors distinguer les détails des « machines moléculaires » qui peuvent être étudiées séparément.



# Introduction : Bases moléculaires au sein du vivant

---

- Les processus métaboliques et de signalisation cellulaire sont effectués par une multitude d'associations/dissociations de biomolécules que l'on peut observer expérimentalement.



# Introduction : Bases moléculaires au sein du vivant

---

- Question : Comment ces biomolécules s'associent entre elles pour former ces machines moléculaires ou pour transporter un signal ?
- Réponse : Par **reconnaissance moléculaire** = processus d'**association sélective** par complémentarité entre deux entités moléculaires.
- La reconnaissance moléculaire est l'objet d'étude de la **chimie supramoléculaire** (au delà de la molécule).
- Les forces physiques qui gouvernent les reconnaissances moléculaires sont les **interactions non-covalentes**.

# Plan du cours

---

## I. Aspect structurale de la reconnaissance moléculaire

- I.1. Complémentarités de forme
- I.2. Structures des biomolécules
- I.3. Changements conformationnels

## II. Aspect énergétique de la reconnaissance moléculaire

- II.1. Interactions non-covalentes
- II.2. Importance du solvant
- II.3. Énergie libre d'association

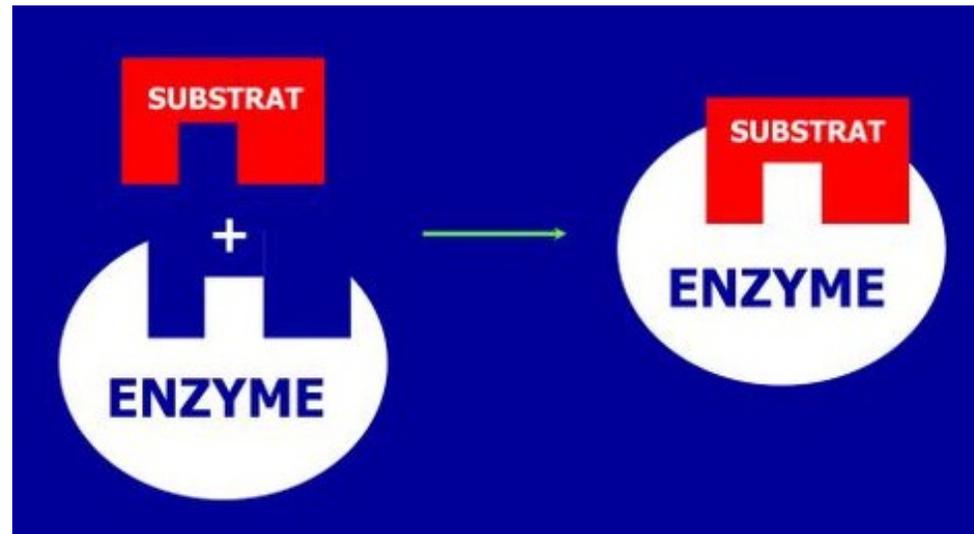
## III. Prédictions des assemblages biomoléculaires

- III.1. Docking protéine – petite molécules
- III.2. Docking protéine – macromolécule
- III.3. Docking sur protéines flexibles

## I.1. Complémentarité de forme

---

- Emil Fisher (1894) : Un substrat reconnaît spécifiquement une enzyme s'il possède une **géométrie complémentaire** au site actif permettant au substrat de s'emboîter exactement dans l'enzyme (modèle de clé-serrure).

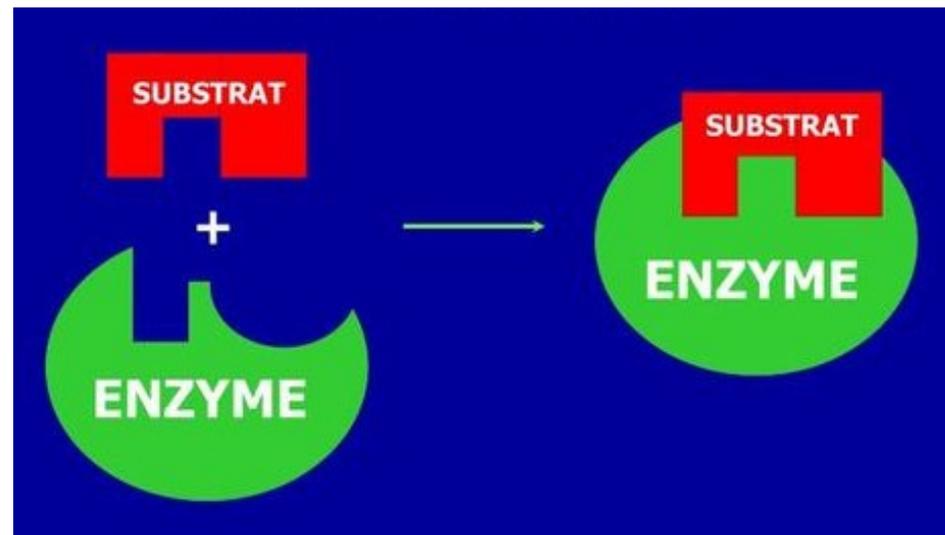


- Forme  $\Leftrightarrow$  Géométrie  $\Leftrightarrow$  Conformation  $\Leftrightarrow$  Structure

## I.1. Complémentarité de forme

---

- Daniel Koshland (1958) : Le site actif d'une enzyme peut se déformer pour s'adapter à la forme du substrat (modèle de l'ajustement induit).

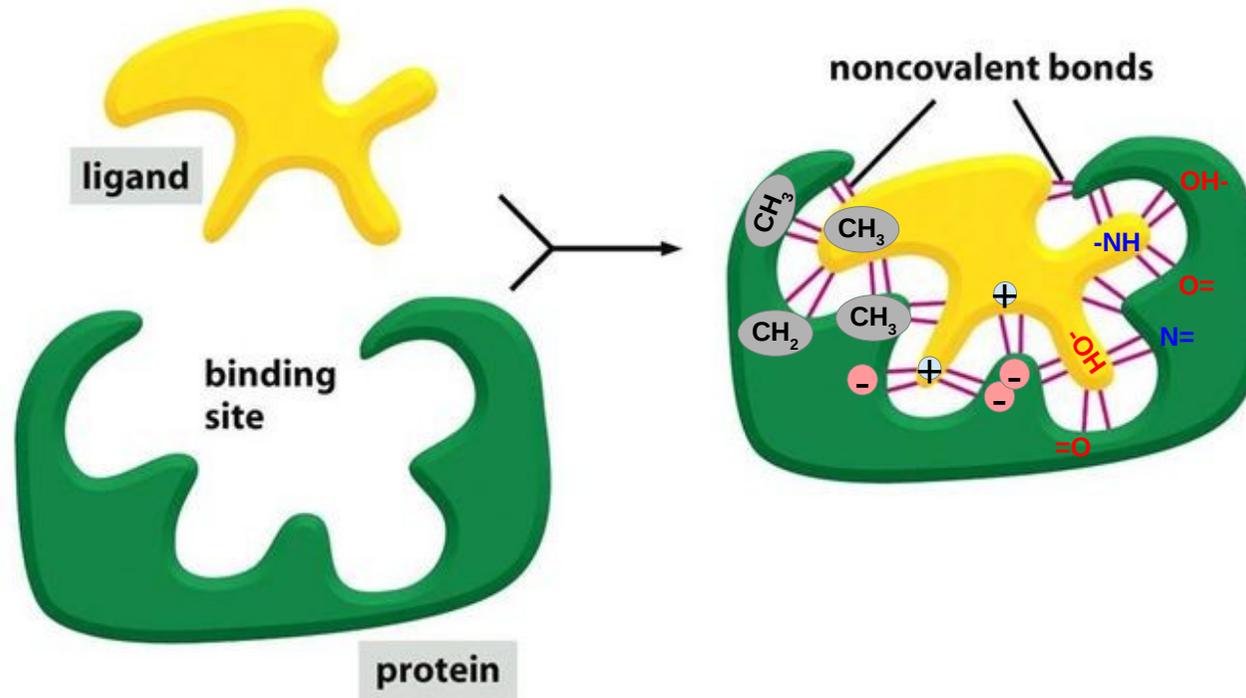


- Déformation  $\Leftrightarrow$  Changement conformationnel

## I.1. Complémentarité de forme

---

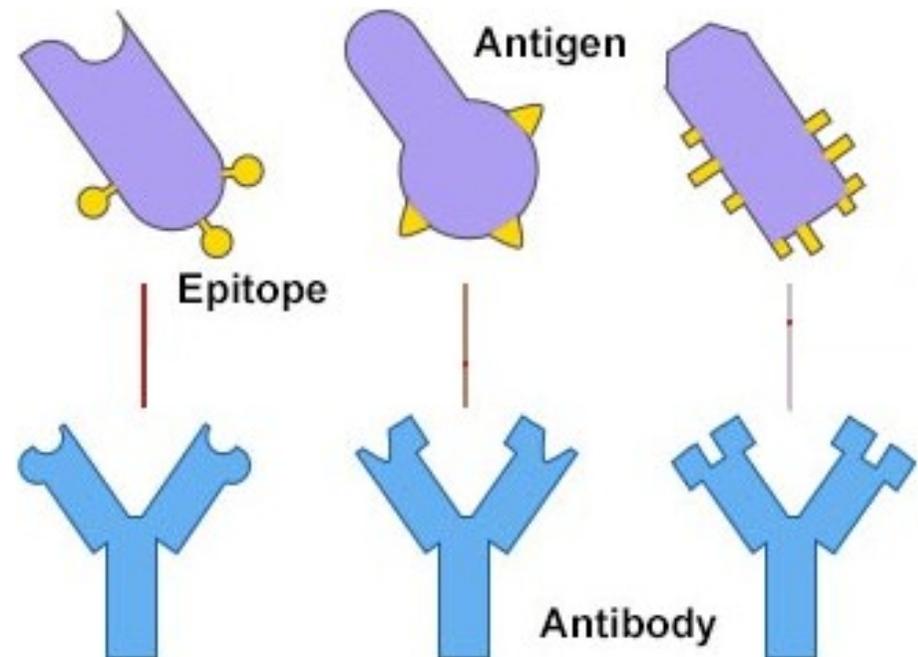
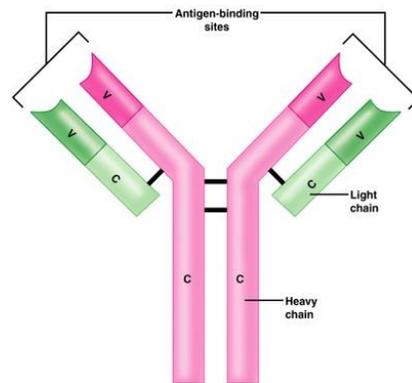
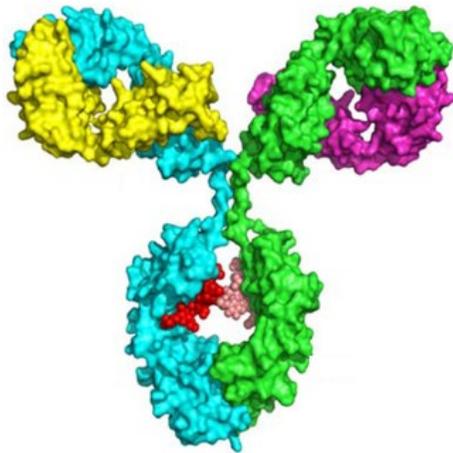
- Aspect négligé des modèles de complémentarité géométrique : La complémentarité physico-chimique (aspect énergétique de la reconnaissance moléculaire).



# I.1. Complémentarité de forme

---

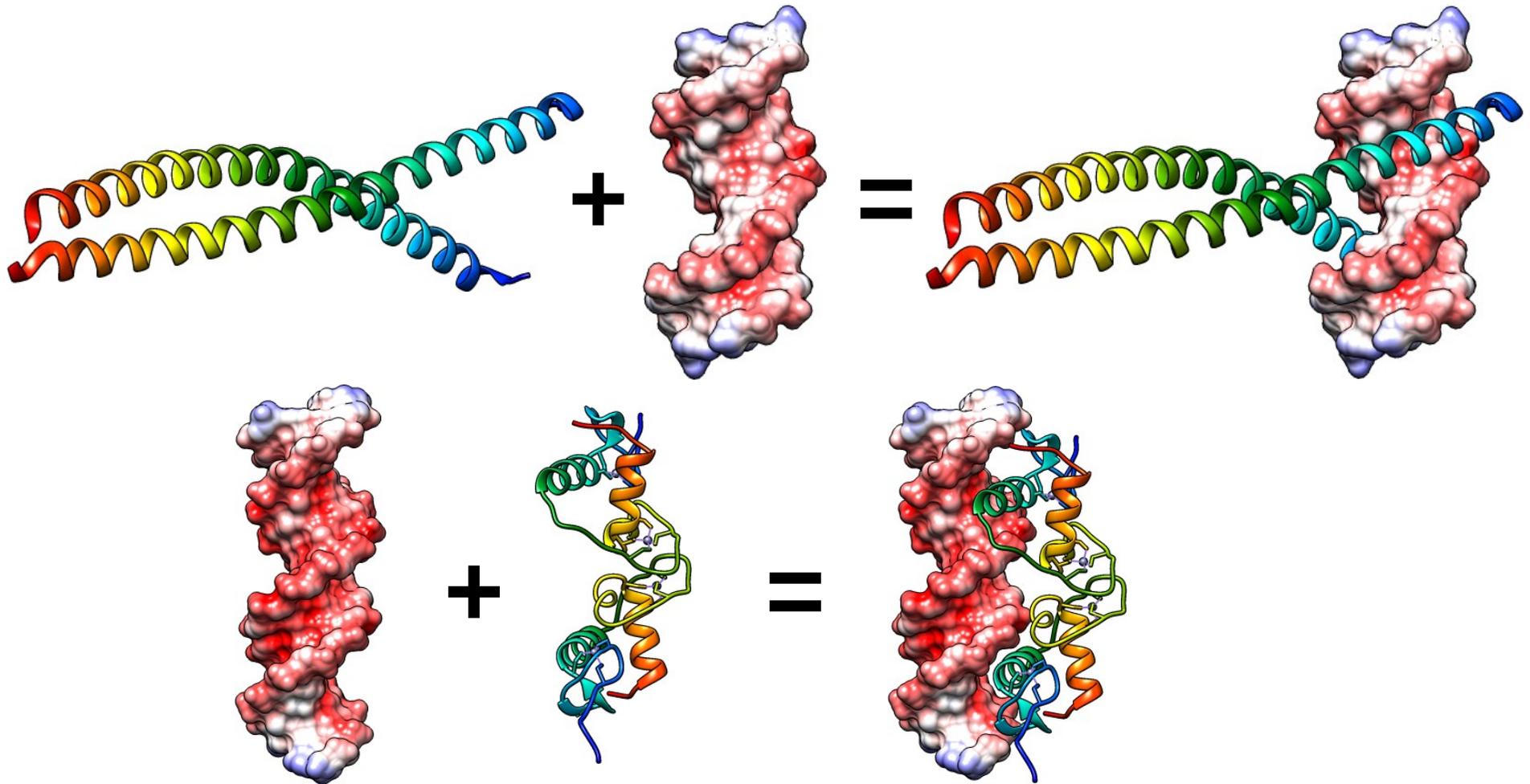
- Les mêmes principes s'appliquent à la reconnaissance moléculaire entre anticorps et antigènes, ...



## I.1. Complémentarité de forme

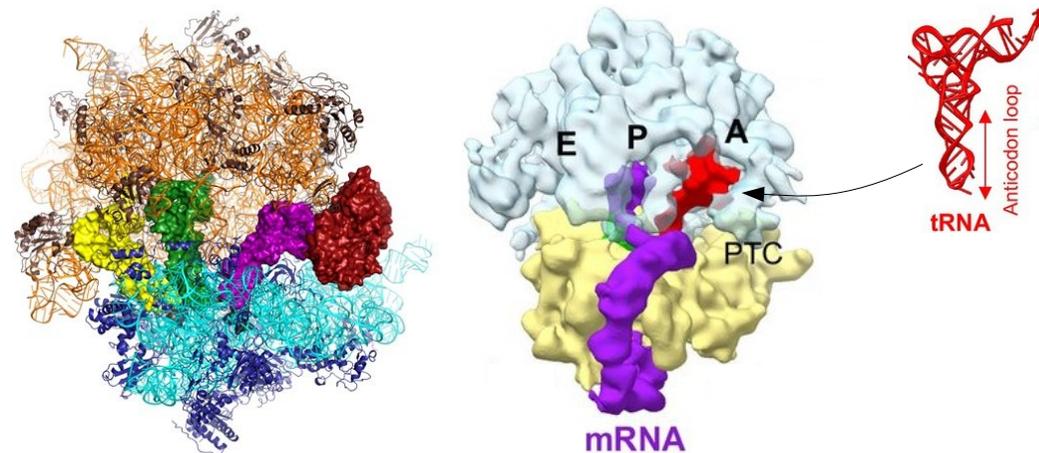
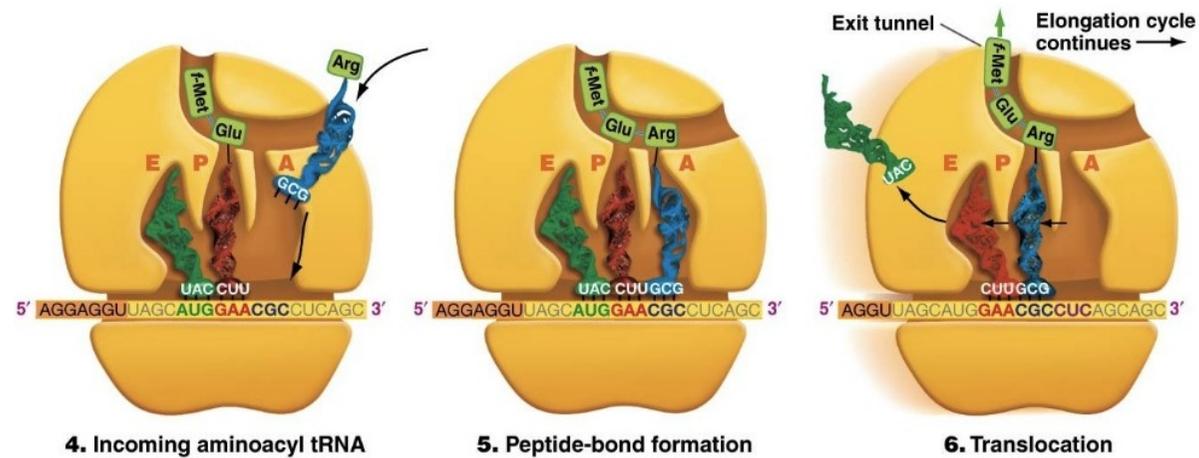
---

- ... entre ADN et protéines, ...



# I.1. Complémentarité de forme

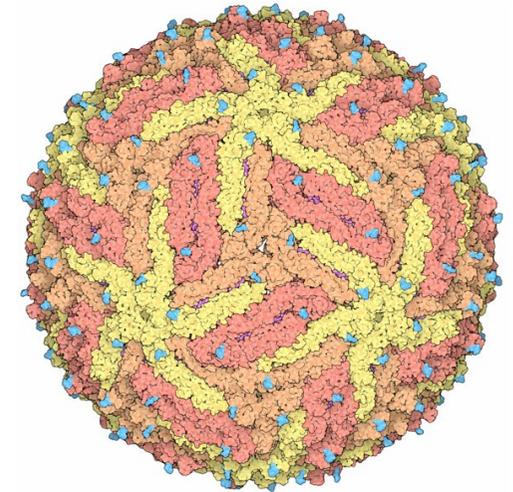
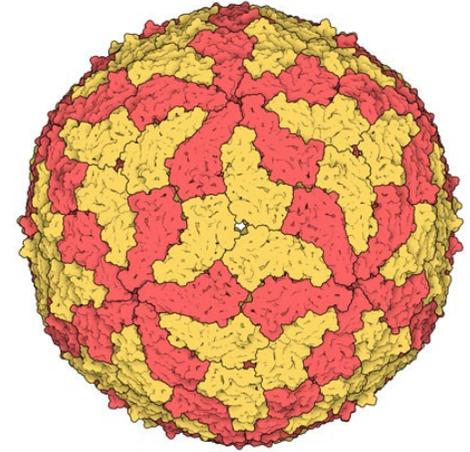
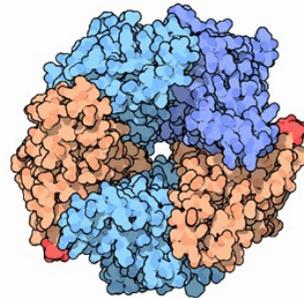
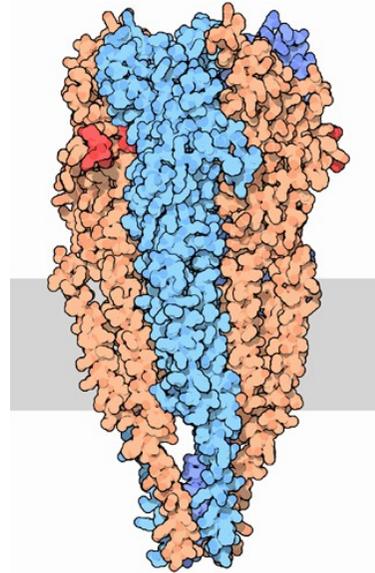
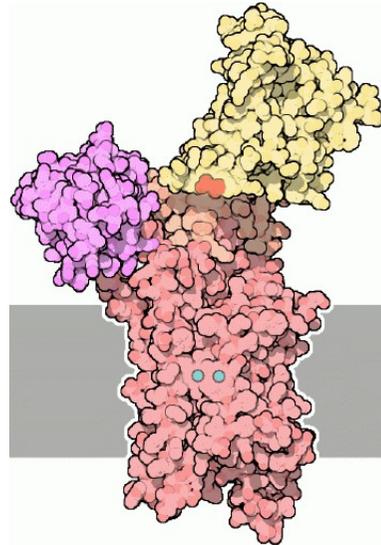
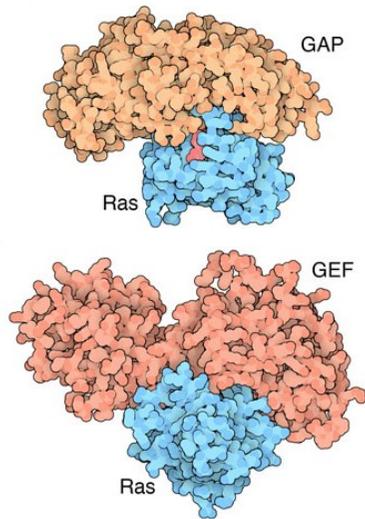
- ... entre RNA et ribosome, ...



# I.1. Complémentarité de forme

---

- ... et entre protéines.



# Plan du cours

---

## I. Aspect structurale de la reconnaissance moléculaire

- I.1. Complémentarités de forme
- I.2. Structures des biomolécules
- I.3. Changements conformationnels

## II. Aspect énergétique de la reconnaissance moléculaire

- II.1. Interactions non-covalentes
- II.2. Importance du solvant
- II.3. Énergie libre d'association

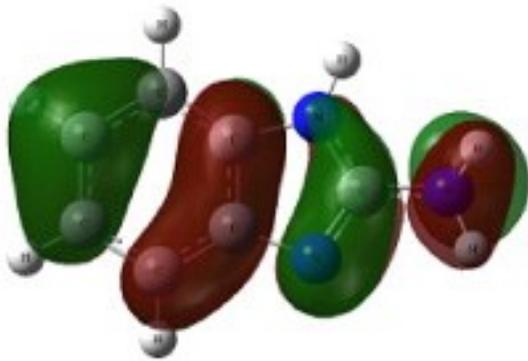
## III. Prédictions des assemblages biomoléculaires

- III.1. Docking protéine – petite molécules
- III.2. Docking protéine – macromolécule
- III.3. Docking sur protéines flexibles

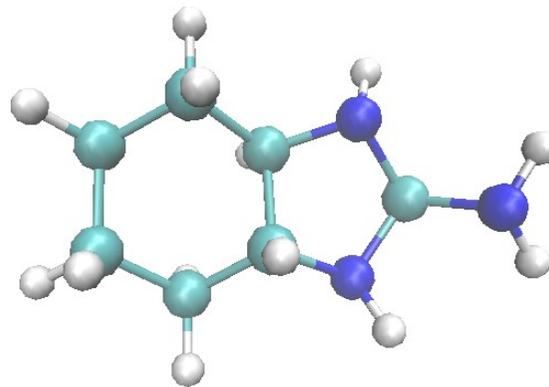
## I.2. Structures des biomolécules

---

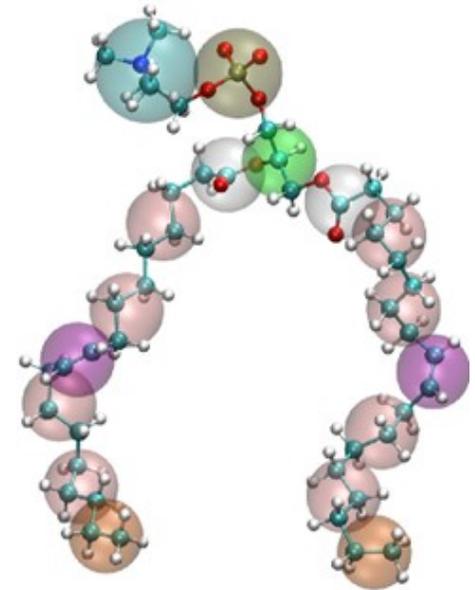
- Différentes descriptions des molécules à différentes résolutions :



**quantique**



**atomique**

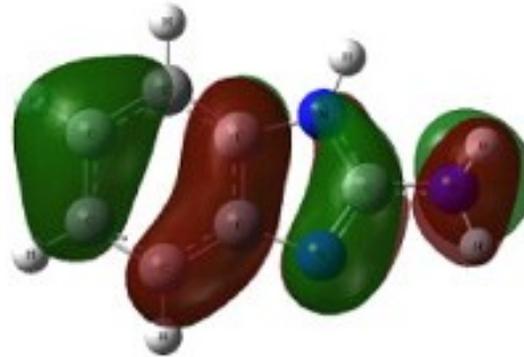


**gros-grain**

## I.2. Structures des biomolécules

---

- Pour les modèles quantiques, la structure d'une molécule = position spatiale des noyaux + densité des nuages électrons (orbitales).



- La densité électronique est déterminée en résolvant l'équation de Schrödinger (généralement pour chaque position des noyaux).

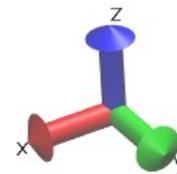
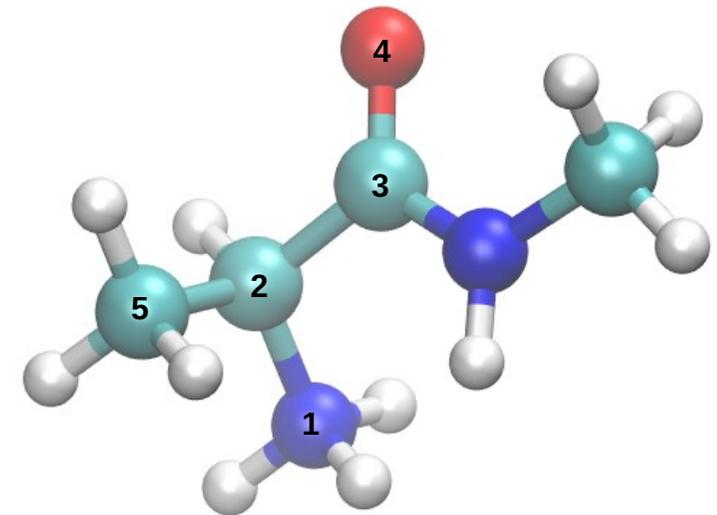
$$\frac{-\hbar^2}{2m} \nabla^2 \Psi(r) + V(r) \Psi(r) = E \Psi(r)$$

## I.2. Structures des biomolécules

---

- Pour les modèles atomiques (ou gros-grain), la structure d'une molécule = position spatiale (coordonnées cartésiennes) des atomes (ou grains).

ATOM	1	N	ALA	A	47	-5.140	57.756	31.835
ATOM	2	CA	ALA	A	47	-4.869	57.318	33.199
ATOM	3	C	ALA	A	47	-5.900	57.881	34.166
ATOM	4	O	ALA	A	47	-6.382	57.176	35.054
ATOM	5	CB	ALA	A	47	-3.459	57.741	33.617
ATOM	6	HA	ALA	A	47	-4.926	56.230	33.229
ATOM	7	HB1	ALA	A	47	-3.267	57.409	34.637
ATOM	8	HB2	ALA	A	47	-2.730	57.289	32.945
ATOM	9	HB3	ALA	A	47	-3.375	58.827	33.566
ATOM	10	H1	ALA	A	47	-6.072	57.472	31.569
ATOM	11	H2	ALA	A	47	-4.469	57.337	31.207
ATOM	12	H3	ALA	A	47	-5.067	58.762	31.784
ATOM	13	N	ALA	A	48	-6.231	59.157	33.990
ATOM	14	CA	ALA	A	48	-7.206	59.826	34.843
ATOM	15	H	ALA	A	48	-5.794	59.678	33.243
ATOM	16	HA1	ALA	A	48	-7.322	60.861	34.521
ATOM	17	HA2	ALA	A	48	-6.860	59.804	35.876
ATOM	18	HA3	ALA	A	48	-8.165	59.313	34.770



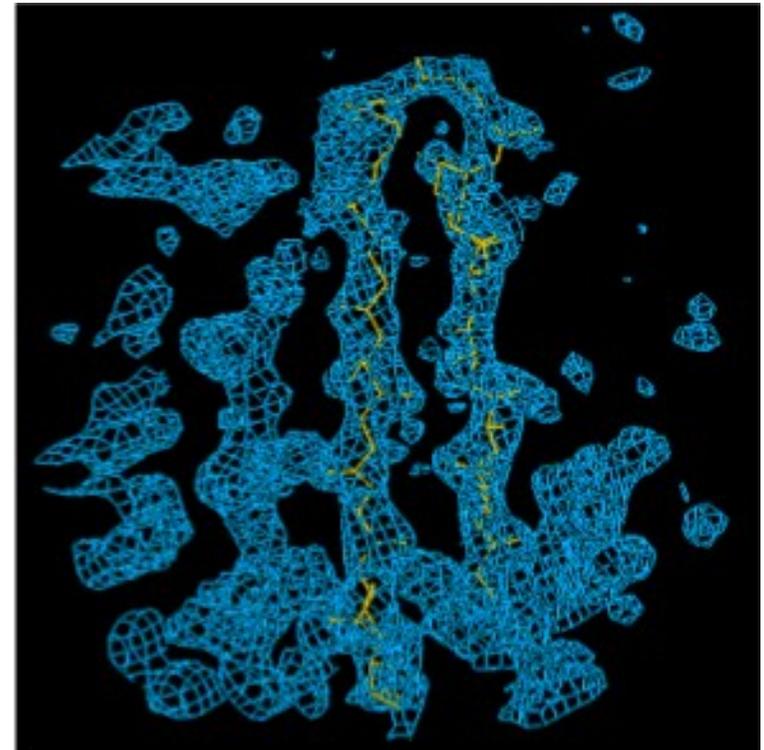
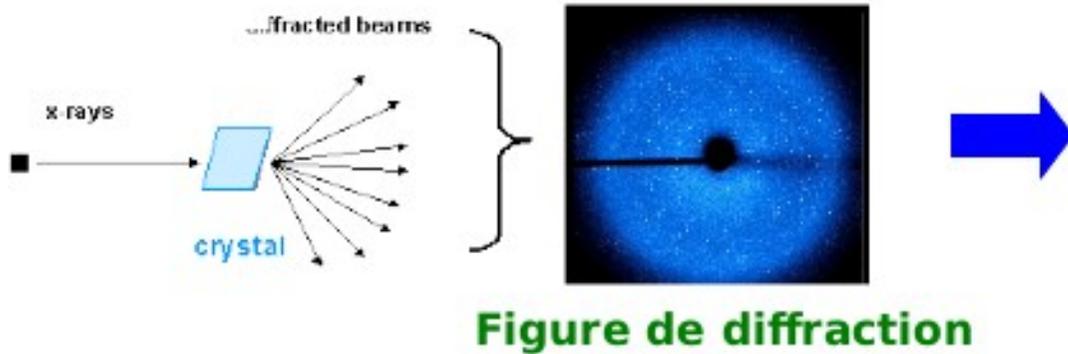
## I.2. Structures des biomolécules

---

- Méthodes de détermination expérimentales :
  - Cristallographie aux rayons X (radiocristallographie)
  - Résonance Magnétique Nucléaire (RMN)
  - Cryo-microscopie électronique (cryo-EM)

# Cristallographie aux rayons X

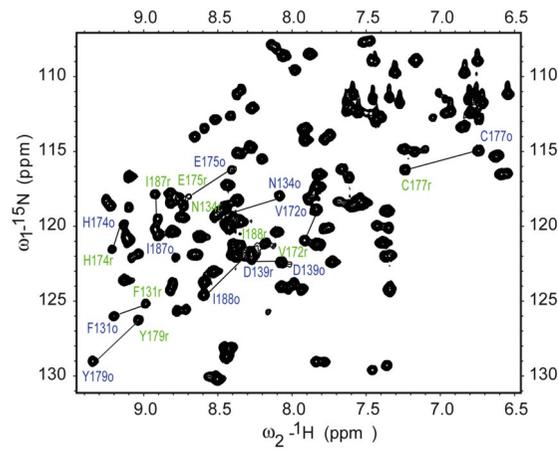
---



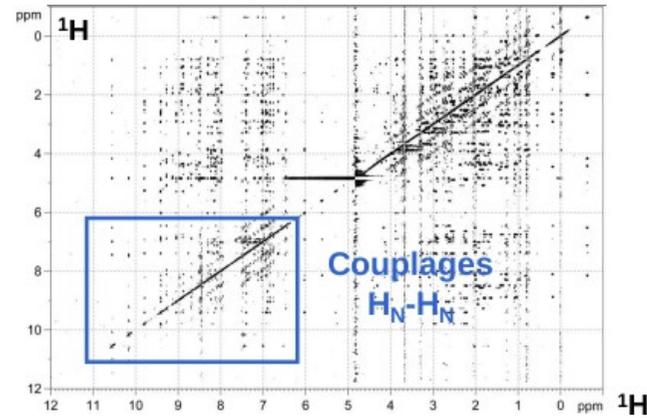
**Carte de densité électronique**

- La molécule doit cristalliser.
- Les régions très flexibles ne sont pas bien résolue.

# Résonance Magnétique Nucléaire

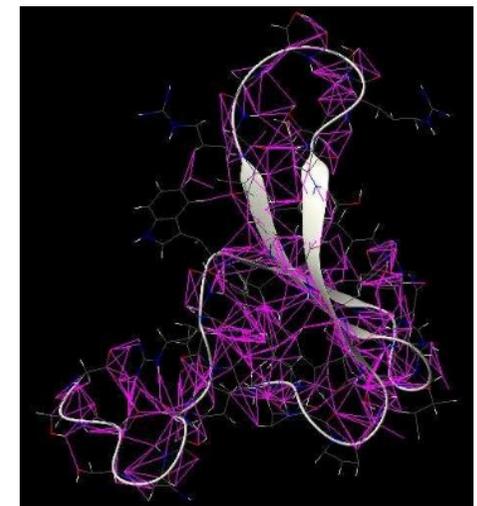
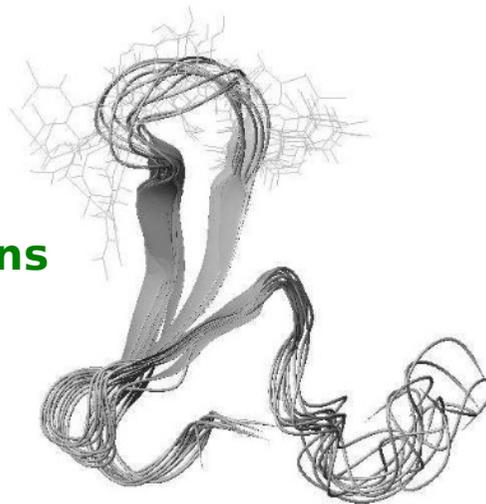


Spectre HSQC > Attribution



Spectre NOESY > Corrélations

Ensemble de conformations

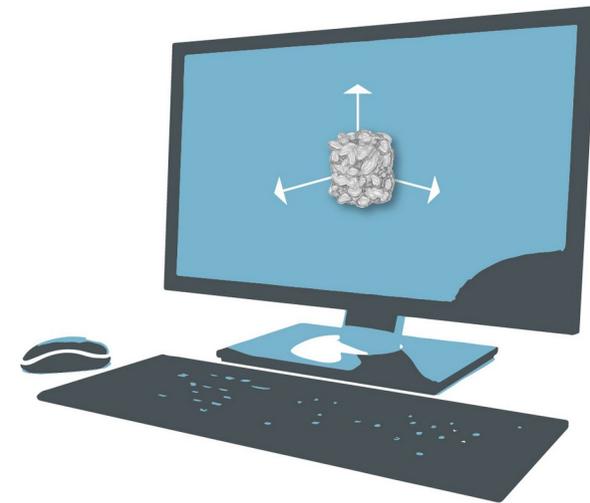
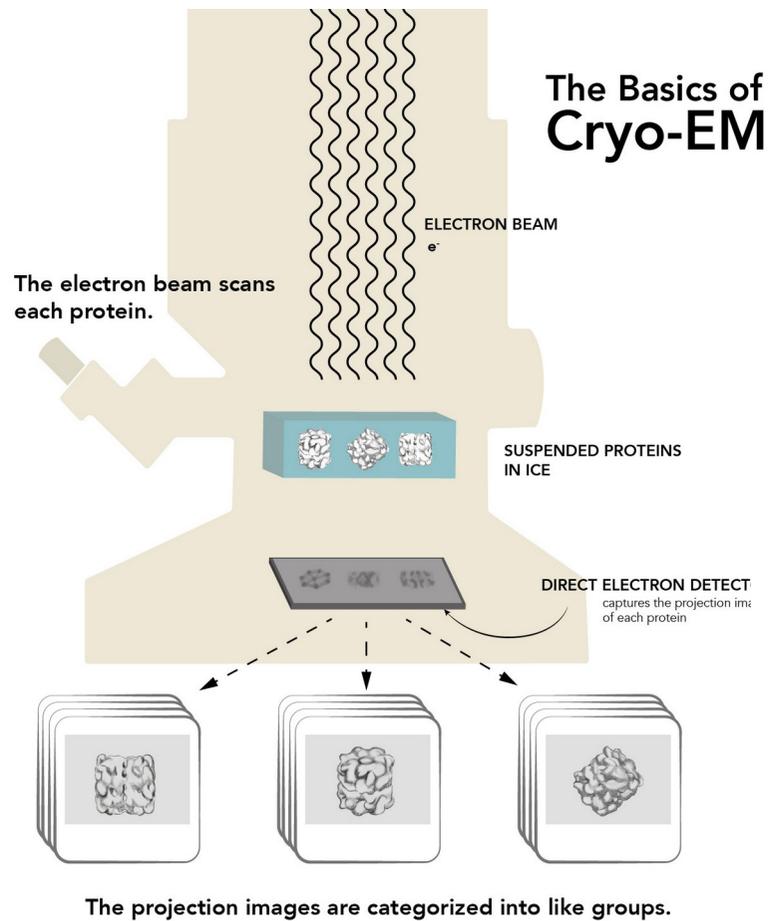


Ensemble de contraintes  $\text{H}_\text{N}-\text{H}_\text{N}$

- La molécule doit être soluble

# Cryo-microscopie électronique

---



Then they are **COMBINED** to create a hi-res 3D model.

- Résolutions atomiques autour de 3-4 Å.

## I.2. Structures des biomolécules

---

- Où trouver la structure tridimensionnelle d'une biomolécule ?
  - Petites molécules : Cambridge Structural Database (CSD)
  - Acides nucléiques : Nucleic Acid Database (NDB)
  - Protéines : Protein Data Bank (PDB)
  - Glucides : Carbohydrate Structure Data Bank (CSDB)

# Cambridge Structural Database (CSD)

---

- Gérée par le Cambridge Crystallographic Data Centre (CCDC) : [www.ccdc.cam.ac.uk](http://www.ccdc.cam.ac.uk)
- En 2020, la base contient ~1 075 000 structures de molécules (avec moins de 500 atomes) résolues par cristallographie.

## The Cambridge Structural Database (CSD)

---



The Cambridge Structural Database (CSD) is a highly curated and comprehensive resource.

Established in 1965, the CSD is the world's repository for small-molecule organic and metal-organic crystal structures. Containing over one million structures from x-ray and neutron diffraction analyses, this unique database of accurate 3D structures has become an essential resource to scientists around the world.

Each one of the entries added to the database undergoes **automated checking, and manual curation** by one of our expert in-house scientific editors. Every structure is enriched with chemical representations, as well as bibliographic, chemical and physical property information, adding further value to the raw structural data. These editorial processes are vital for enabling scientists to interpret structures in a chemically meaningful way. Ensuring the CSD is of the highest quality allows scientists around the globe to generate rich insights into the properties and behavior of molecular structures and crystalline materials. This new knowledge has been applied across academia and industry in pursuit of new drugs, novel materials and a greater understanding of chemical and crystallographic phenomena.

# Nucleic Acid Database (CSD)

- Gérée par Rutgers the State University of New Jersey : <http://ndbserver.rutgers.edu>
- En 2020, la base contient ~11 000 structures d'ADN ou ARN.



A Portal for Three-dimensional Structural Information about Nucleic Acids  
As of 26-Aug-2020 number of released structures: **10915**

Search DNA

Search RNA

Advanced Search

Enter an NDB ID or PDB ID



Search for released structures

## Welcome to the NDB

The NDB contains information about experimentally-determined nucleic acids and complex assemblies. Use the NDB to perform searches based on annotations relating to sequence, structure and function, and to download, analyze, and learn about nucleic acids.

### Search Structures

#### Search DNA

Search DNA and its complexes

#### Search RNA

Search for RNA structures in the NDB archive or in the Non-Redundant list

#### Advanced Search

Search for structures based on structural features, chemical features, binding modes, citation and experimental information



### Featured Tools

**RNA 3D Motif Atlas**, a representative collection of RNA 3D internal and hairpin loop motifs

**Non-redundant Lists** of RNA-containing 3D structures

**RNA Base Triple Atlas**, a collection of motifs consisting of two RNA basepairs

**WebFR3D**, a webserver for symbolic and geometric searching of RNA 3D structures

**R3D Align**, an application for detailed nucleotide to nucleotide alignments of RNA 3D structures



# Protein Data Bank (PDB)

- Gérée par **RUTGERS** | **UC San Diego** **SDSC** | **UCSF** : [www.rcsb.org](http://www.rcsb.org)
- En 2020, la base contient plus de ~170 000 structures de protéines.

The screenshot shows the RCSB PDB website homepage. At the top, the logo for RCSB PDB Protein Data Bank is displayed, along with the text "168358 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education". A search bar with the placeholder "Enter search term(s)" and a magnifying glass icon is present. Below the search bar, there are links for "Advanced Search" and "Browse Annotations". The header also features logos for PDB-101, Worldwide PDB, EMDatabank, Nucleic Acid Database, and Worldwide Protein Data Bank Foundation. A navigation menu on the left includes "Welcome", "Deposit", "Search", "Visualize", "Analyze", "Download", and "Learn". The main content area is divided into two sections: "A Structural View of Biology" and "September Molecule of the Month".

**A Structural View of Biology**

This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

As a member of the wwPDB, the RCSB PDB curates and annotates PDB data.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.

**COVID-19 CORONAVIRUS Resources**

**September Molecule of the Month**

SARS-CoV-2 RNA-dependent RNA Polymerase

# Carbohydrate Structure Data Bank (CSDB)

---

- Gérée par la Zelinsky Institute of Organic Chemistry (Russie) : <http://csdb.glycoscience.ru/database/index.html>
- En 2020, la base contient plus de ~25 000 structures de glucides.

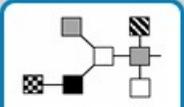


This is CSDB version 1 merged from Bacterial (BCSDB) and Plant&Fungal (PFCSD) databases.  
CSDB contains manually curated natural carbohydrate structures, taxonomy, bibliography, NMR data etc.  
Coverage is close to complete up to: 2018 (bacteria and archaea), 2011 (fungi), 1997 (plants).

Dear scientists! Please cite CSDB properly: [How to cite](#)

---

### Database search

 Structures

 Composition

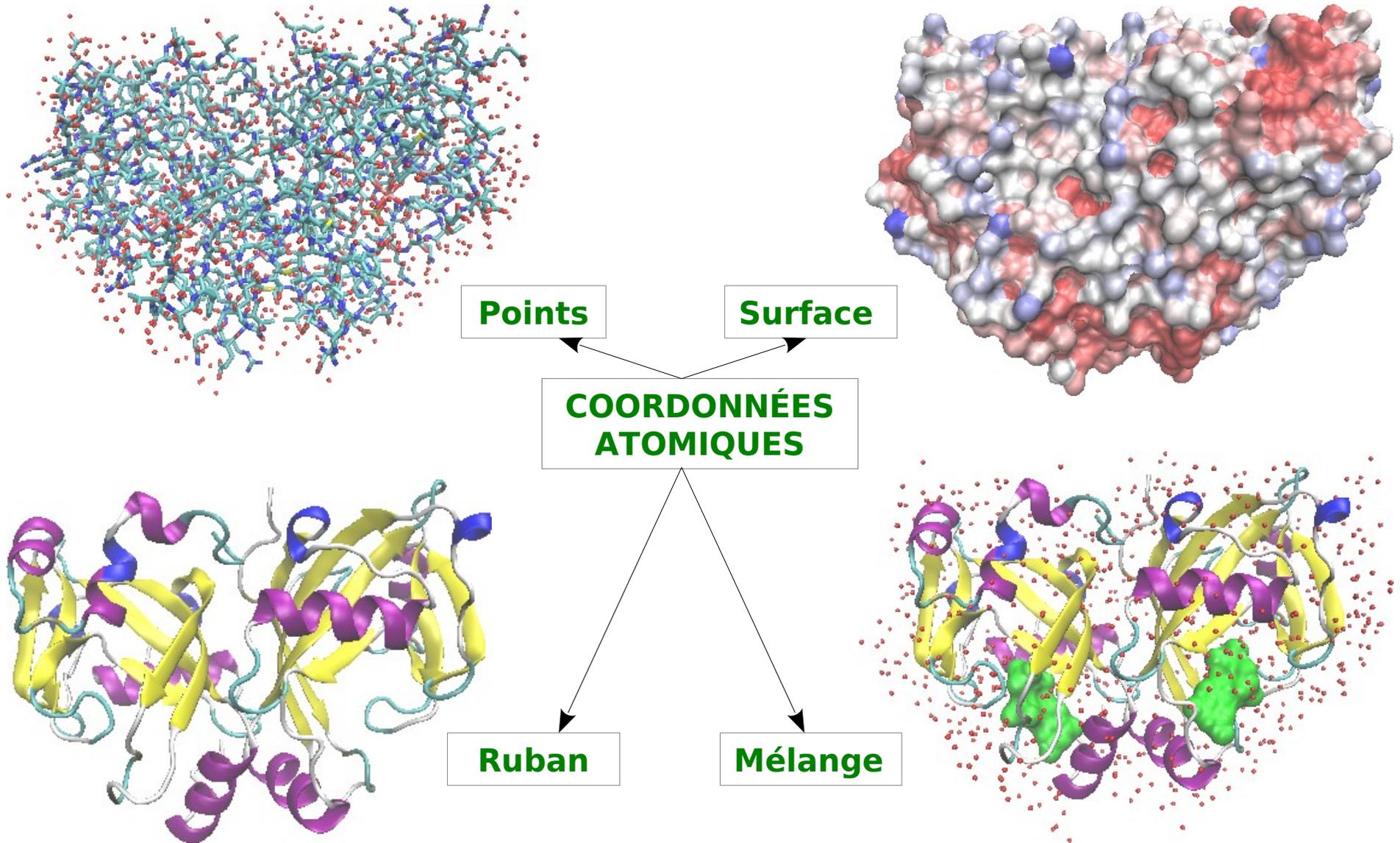
 Organism

 Publications

 NMR signals

# Logiciel de visualisation : Chimera, VMD, PyMol...

---



# Plan du cours

---

## I. Aspect structurale de la reconnaissance moléculaire

- I.1. Complémentarités de forme
- I.2. Structures des biomolécules
- I.3. Changements conformationnels

## II. Aspect énergétique de la reconnaissance moléculaire

- II.1. Interactions non-covalentes
- II.2. Importance du solvant
- II.3. Énergie libre d'association

## III. Prédictions des assemblages biomoléculaires

- III.1. Docking protéine – petite molécules
- III.2. Docking protéine – macromolécule
- III.3. Docking sur protéines flexibles

## I.3. Changements conformationnels

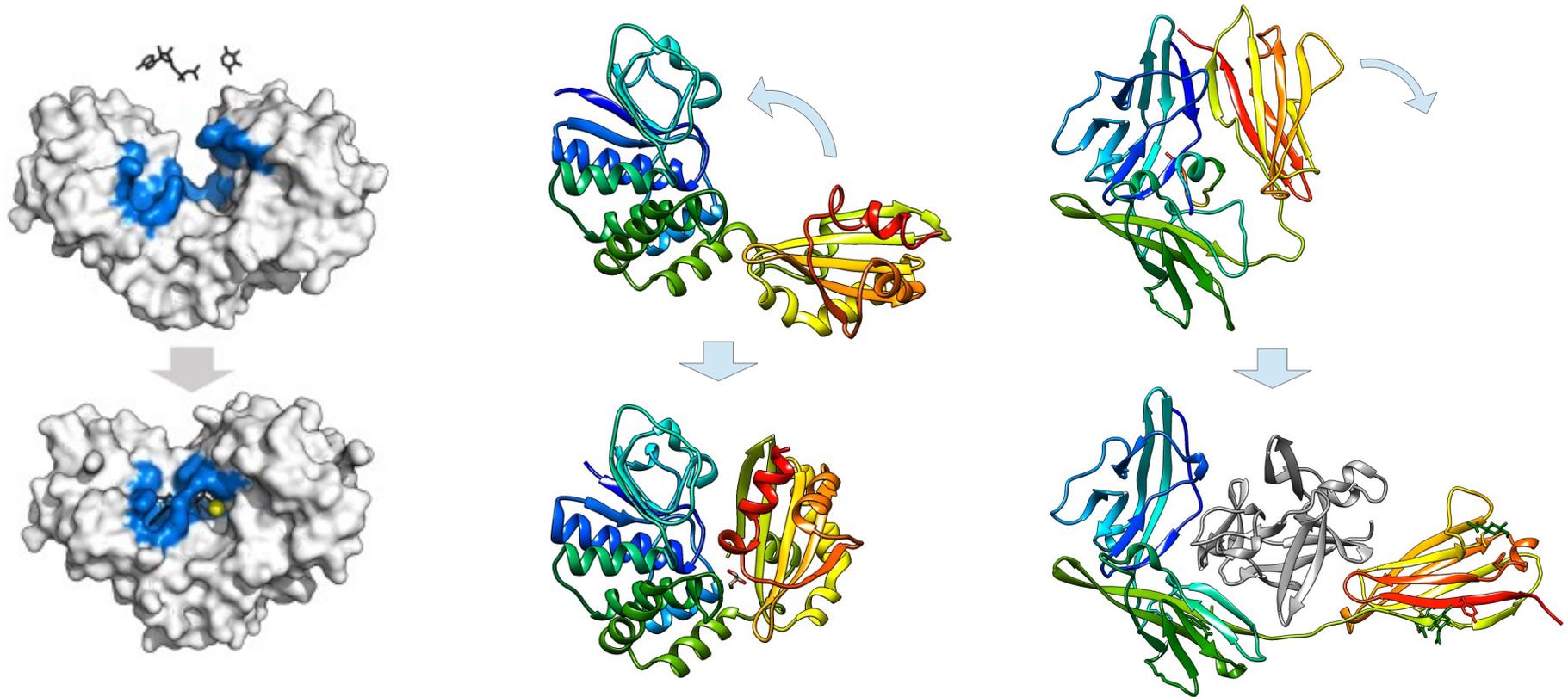
---

- Influence des changements conformationnels d'une biomolécule sur la reconnaissance moléculaire de ses ligands :
  - Ajustement induit
  - Effet allostérique

# Ajustement induit

---

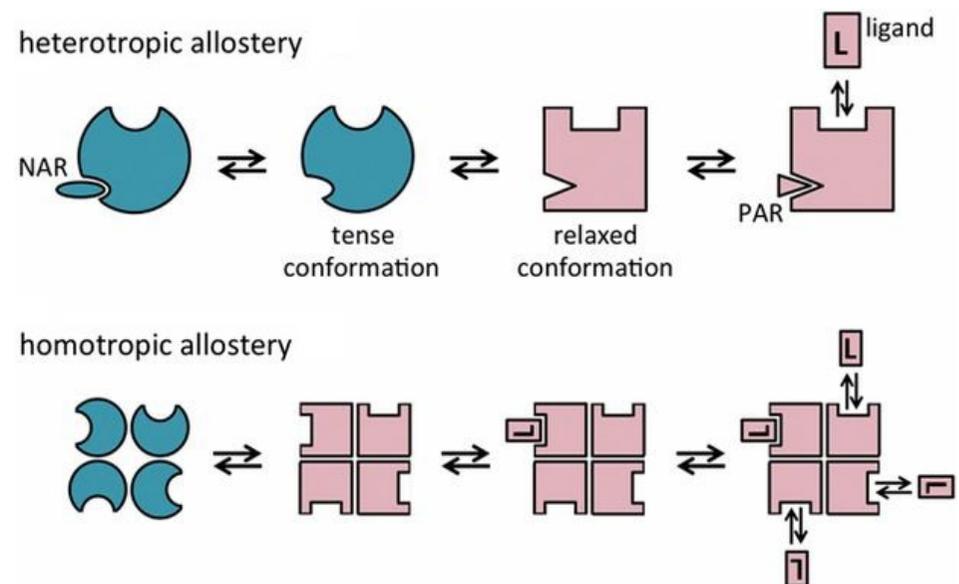
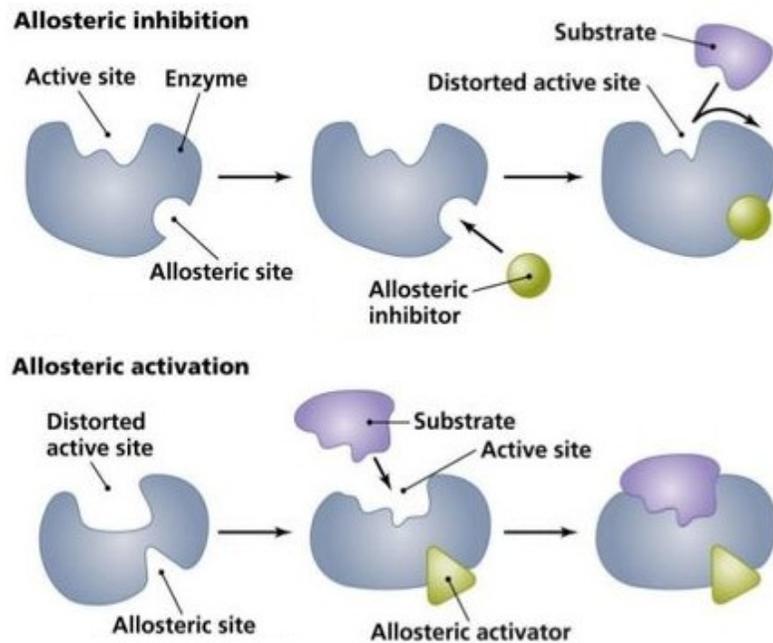
- Changement conformationnel de plus ou moins grande amplitude.



- La forme du site de liaison s'adapte à la présence du ligand pour maximiser la complémentarité de forme.

# Effets allostériques

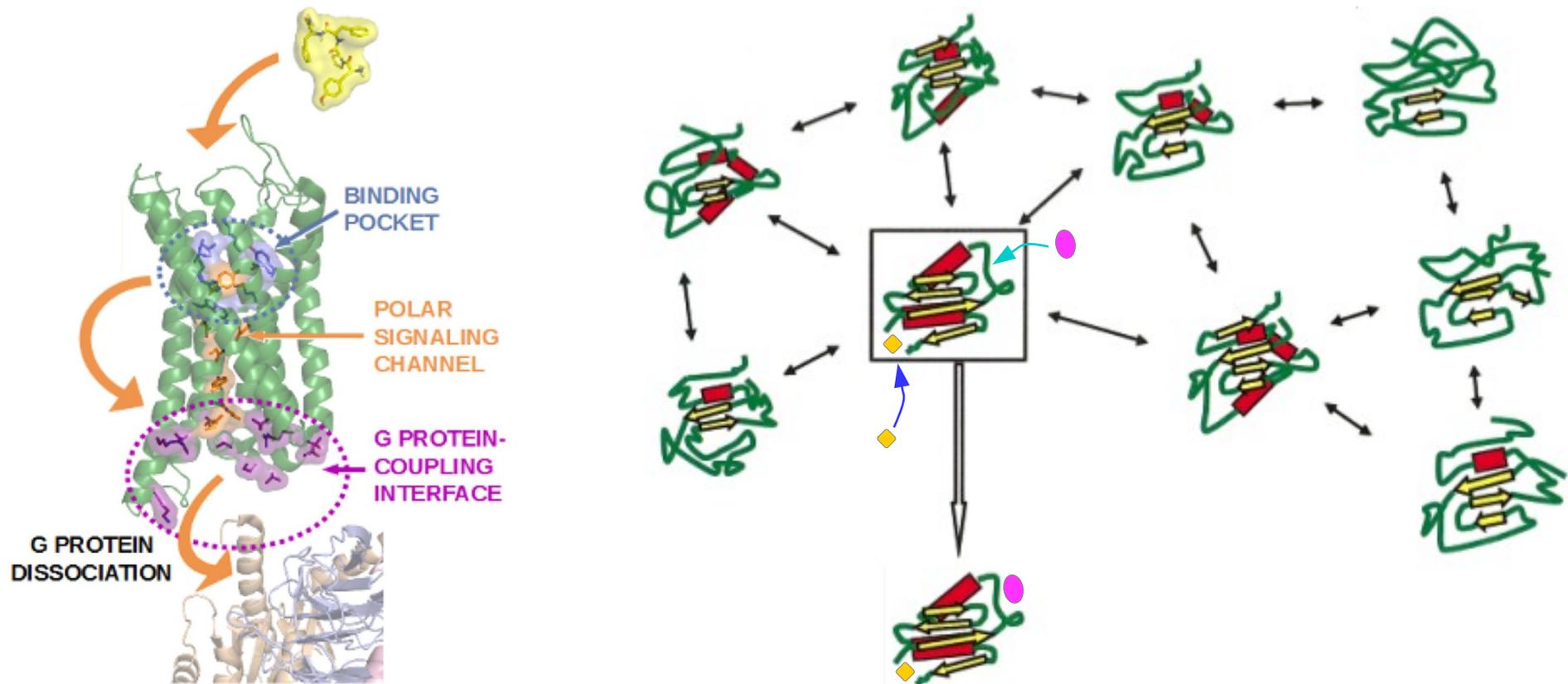
- J. Monod, J.P. Changeux, F. Jacobs, « Allosteric protein and cellular control systems », *J. Mol. Biol.* **1963**, 6 : 306-329.



- Un ligand allostérique change la forme du site de liaison orthostérique et donc la complémentarité géométrique avec le substrat.

# Mécanismes moléculaire de l'allostérie ?

- Changement conformationnel déclenché à distance (via une transmission de signal intramoléculaire) ou sélection conformationnelle :



- Dans les deux cas, la dynamique conformationnelle d'une biomolécule influence la reconnaissance de ses sites de liaison par ses ligands.

# Plan du cours

---

## I. Aspect structurale de la reconnaissance moléculaire

- I.1. Complémentarités de forme
- I.2. Structures des biomolécules
- I.3. Changements conformationnels

## II. Aspect énergétique de la reconnaissance moléculaire

- II.1. Interactions non-covalentes
- II.2. Importance du solvant
- II.3. Énergie libre d'association

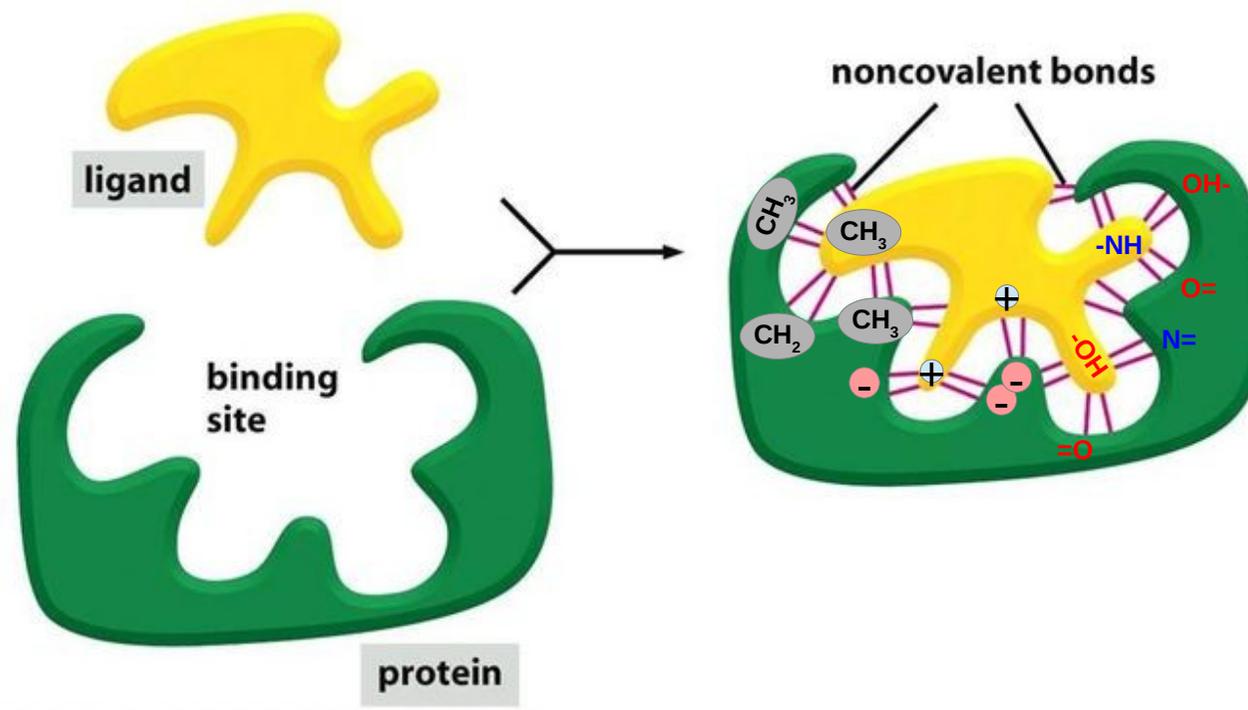
## III. Prédictions des assemblages biomoléculaires

- III.1. Docking protéine – petite molécules
- III.2. Docking protéine – macromolécule
- III.3. Docking sur protéines flexibles

## II. Aspect énergétique de la reconnaissance moléculaire

---

- Deux biomolécules vont s'associer spécifiquement si elles ont une complémentarité géométrique ET une complémentarité physico-chimique.



- Cette complémentarité physico-chimique se traduit par des **interactions favorables non-covalentes**.

# Plan du cours

---

## I. Aspect structurale de la reconnaissance moléculaire

- I.1. Complémentarités de forme
- I.2. Structures des biomolécules
- I.3. Changements conformationnels

## II. Aspect énergétique de la reconnaissance moléculaire

- II.1. Interactions non-covalentes
- II.2. Importance du solvant
- II.3. Énergie libre d'association

## III. Prédictions des assemblages biomoléculaires

- III.1. Docking protéine – petite molécules
- III.2. Docking protéine – macromolécule
- III.3. Docking sur protéines flexibles

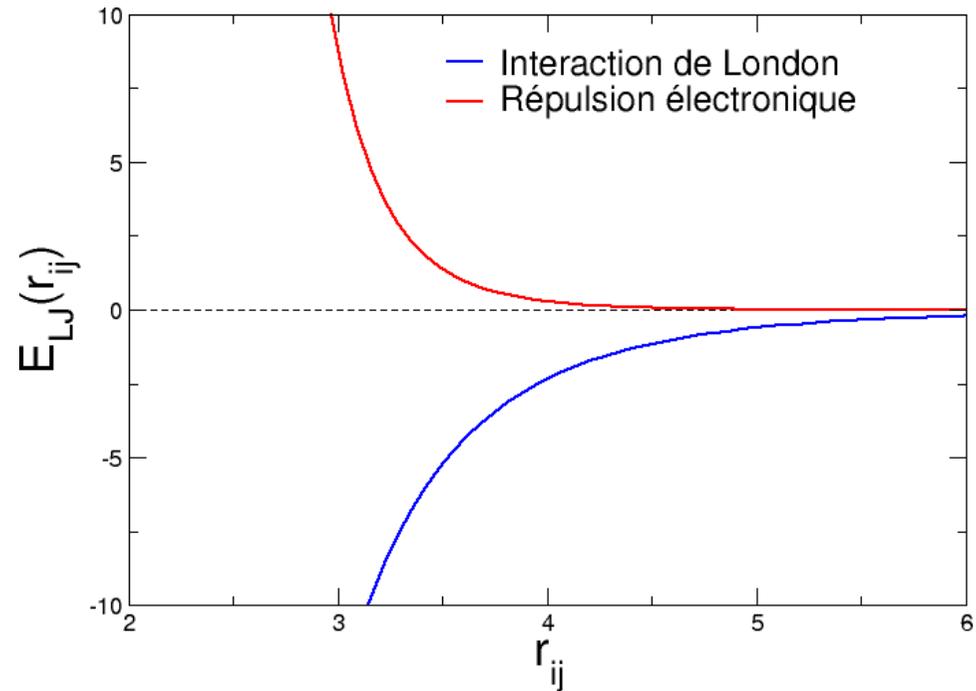
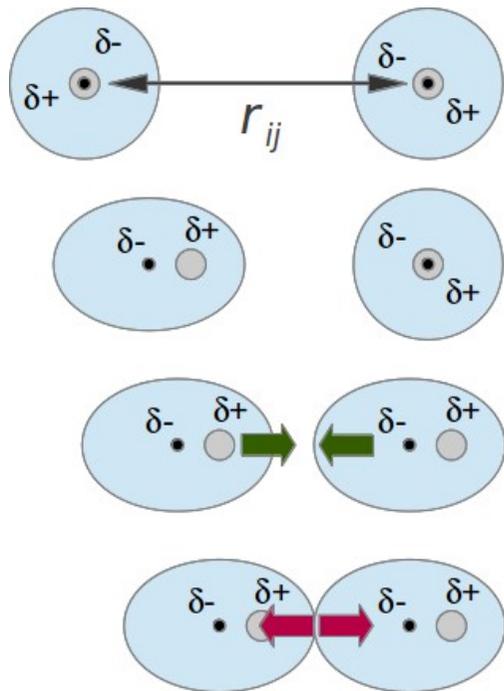
## II.1. Interactions non-covalentes

---

- Les interactions non-covalentes n'impliquent pas la mise en commun d'électrons.
- Elles sont plus faibles que les liaisons covalentes ( $\sim 100$  kcal/mol).
- On peut distinguer 4 types d'interactions non-covalentes :
  - Interactions de van-der-Waals ( $\sim 1$  kcal/mol)
  - Interactions  $\pi$ - $\pi$  et  $\pi$ -cation ( $\sim 2$  kcal/mol)\*
  - Liaisons hydrogène ( $\sim 3$  kcal/mol)\*
  - Interactions ioniques ( $\sim 4$  kcal/mol)\*

# Interactions de van der Waals - London

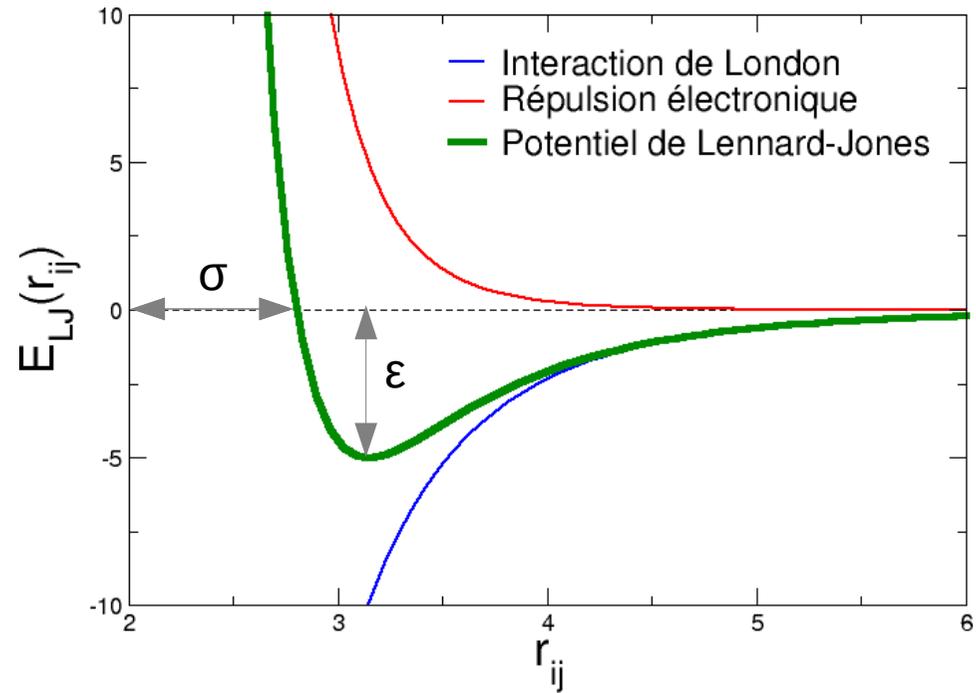
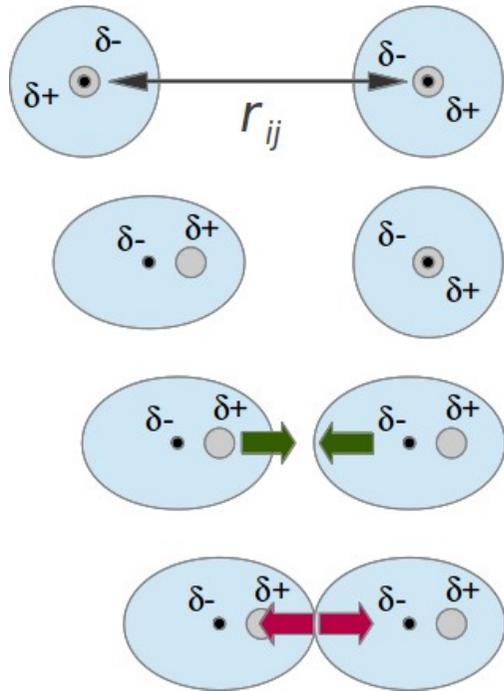
- Au niveau atomique, les forces de dispersion de London sont dues à la déformabilité des nuages électroniques :



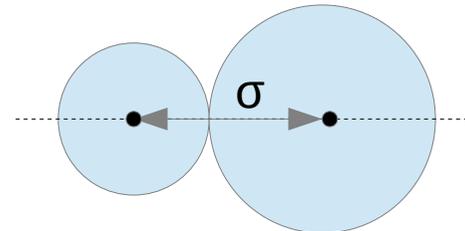
- Mais à courte distance, les nuages électroniques se repoussent et les atomes subissent une importante force répulsive.

# Potentiel de Lennard-Jones

- Il en résulte une énergie potentielle dite de Lennard-Jones :



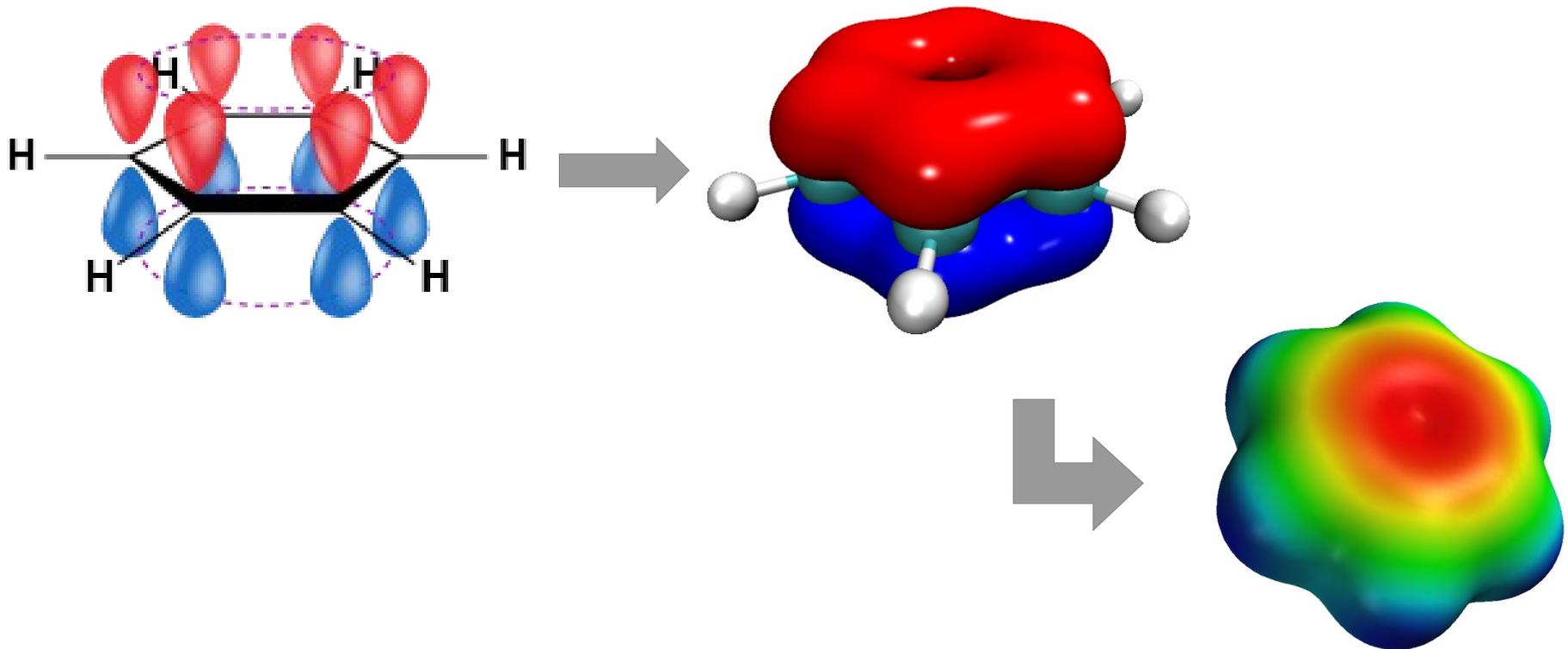
$$E_{LJ} = \sum_{i,j} 4\epsilon \left[ \left( \frac{\sigma}{r_{ij}} \right)^{12} - \left( \frac{\sigma}{r_{ij}} \right)^6 \right]$$



# Interactions $\pi$ - $\pi$ et $\pi$ -cation

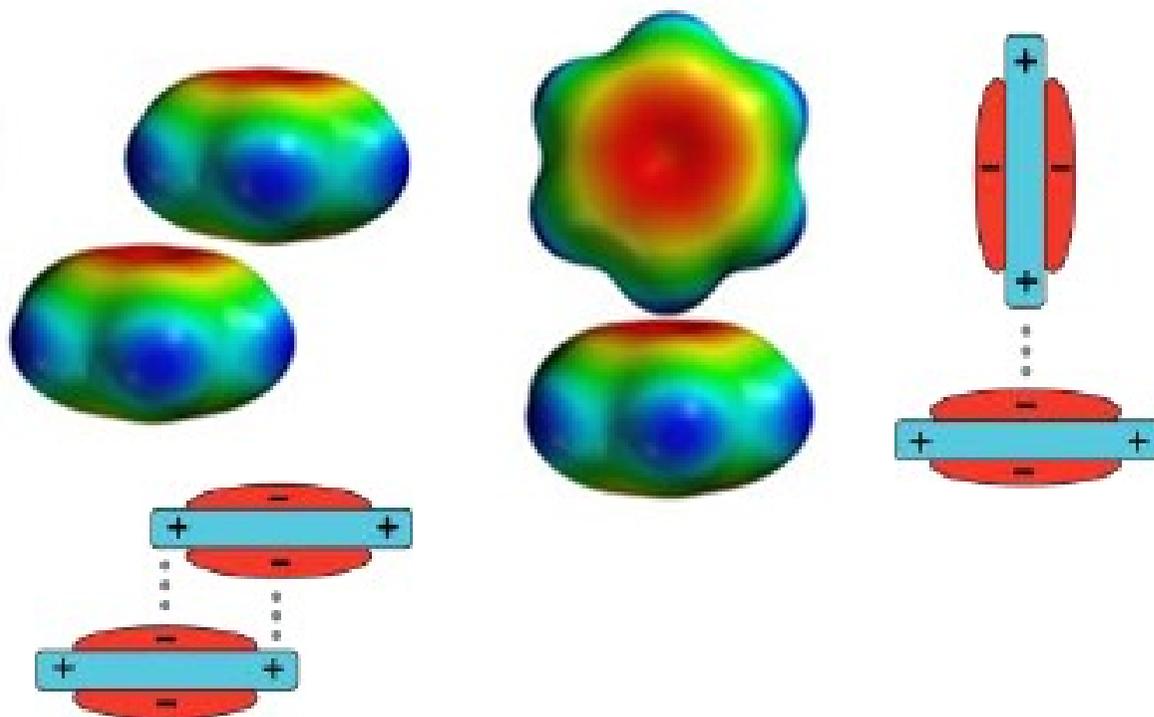
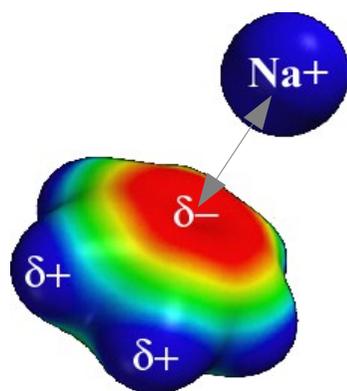
---

- Ce sont des interactions qui impliquent des systèmes conjugués  $\pi$  constitués d'orbitales p et d'électrons délocalisés.
- Exemple : le benzène



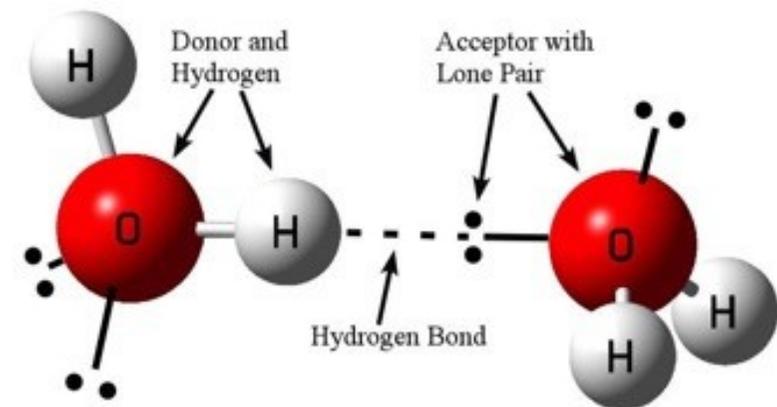
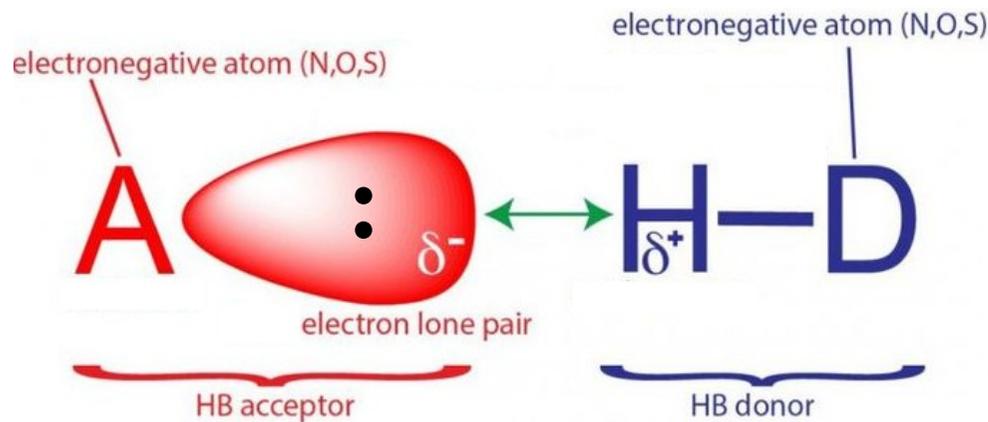
# Interactions $\pi$ - $\pi$ et $\pi$ -cation

- Le potentiel électrostatique des cycles conjugués  $\pi$  peut interagir favorablement avec un cation ou avec un autre cycle aromatique.



# Liaisons hydrogène

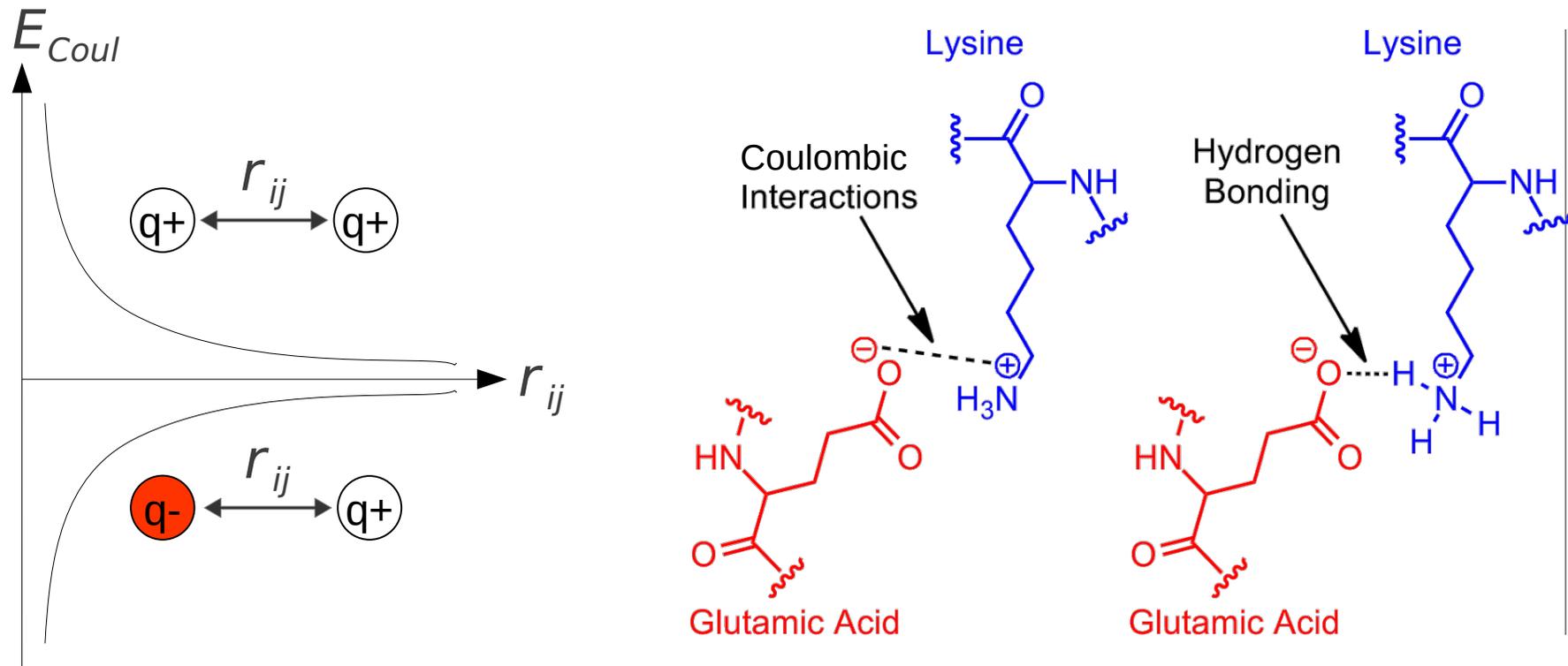
- C'est une interaction entre un hydrogène porté par hétéro-atome électro-négatif et un autre hétéro-atome possédant un doublet d'électrons non liant.



# Interactions ioniques

- Ce sont des interactions entre charges électriques.

- Loi de Coulomb :  $E_{coul} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$



# Plan du cours

---

## I. Aspect structurale de la reconnaissance moléculaire

- I.1. Complémentarités de forme
- I.2. Structures des biomolécules
- I.3. Changements conformationnels

## II. Aspect énergétique de la reconnaissance moléculaire

- II.1. Interactions non-covalentes
- II.2. Importance du solvant
- II.3. Énergie libre d'association

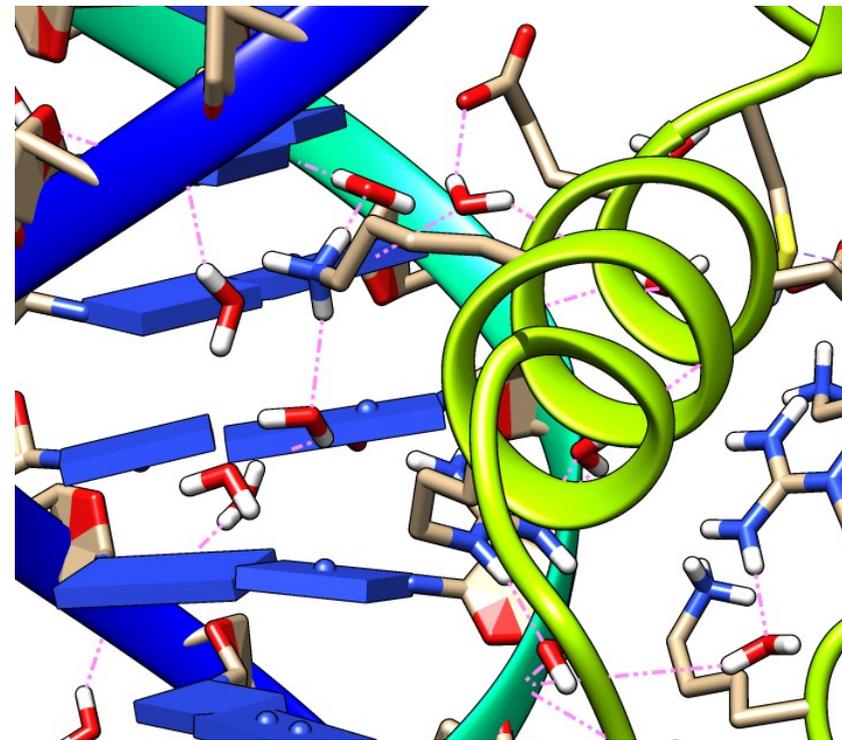
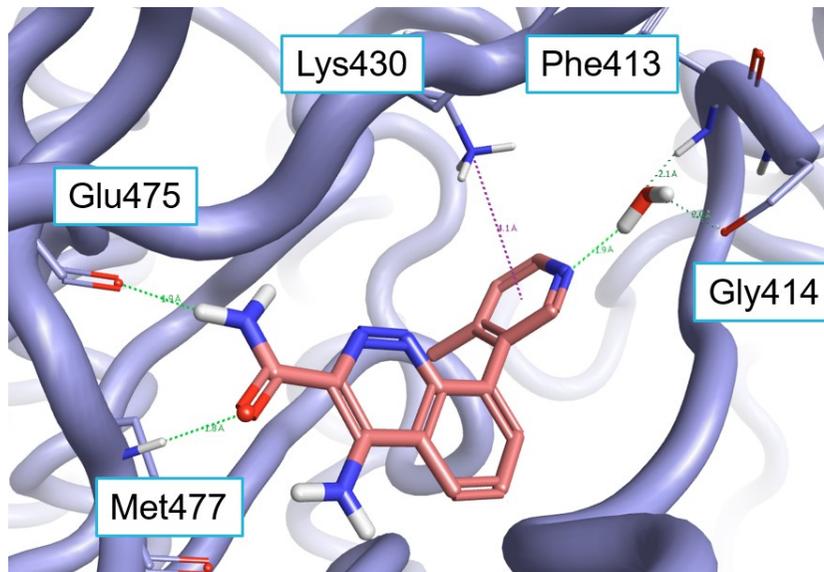
## III. Prédictions des assemblages biomoléculaires

- III.1. Docking protéine – petite molécules
- III.2. Docking protéine – macromolécule
- III.3. Docking sur protéines flexibles

## II.2. Importance du solvant

---

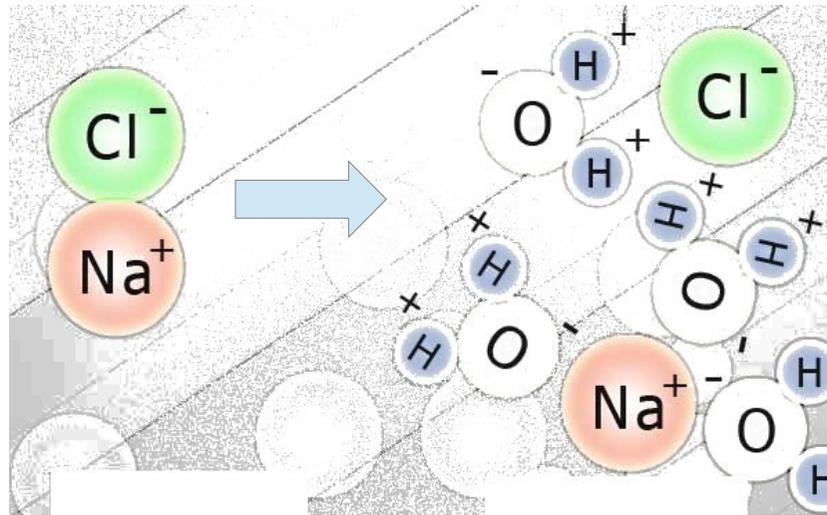
- Les molécules d'eau peuvent servir de pont entre molécules en interaction.



## II.2. Importance du solvant

---

- Les molécules d'eau polaires interagissent favorablement avec les groupes chargés. Elles affaiblissent leurs interactions de Coulomb :

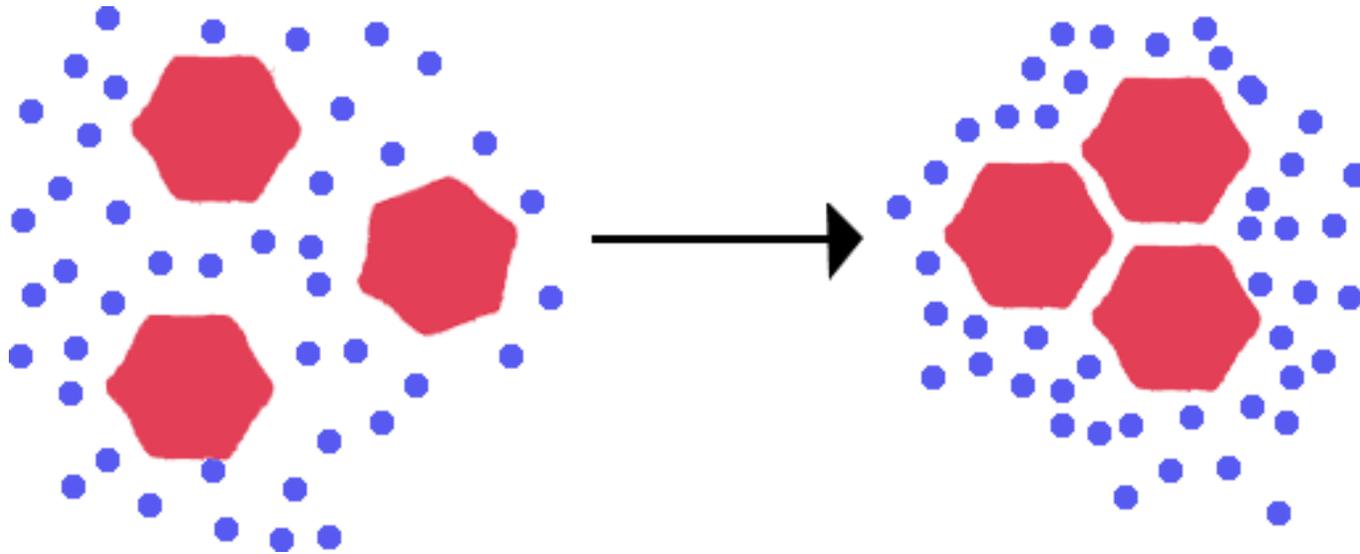


$$E_{coul} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \quad \longrightarrow \quad E_{coul} = \frac{q_i q_j}{4\pi\epsilon_0 \epsilon_r r_{ij}} \quad (\epsilon_r \approx 80)$$

## II.2. Importance du solvant

---

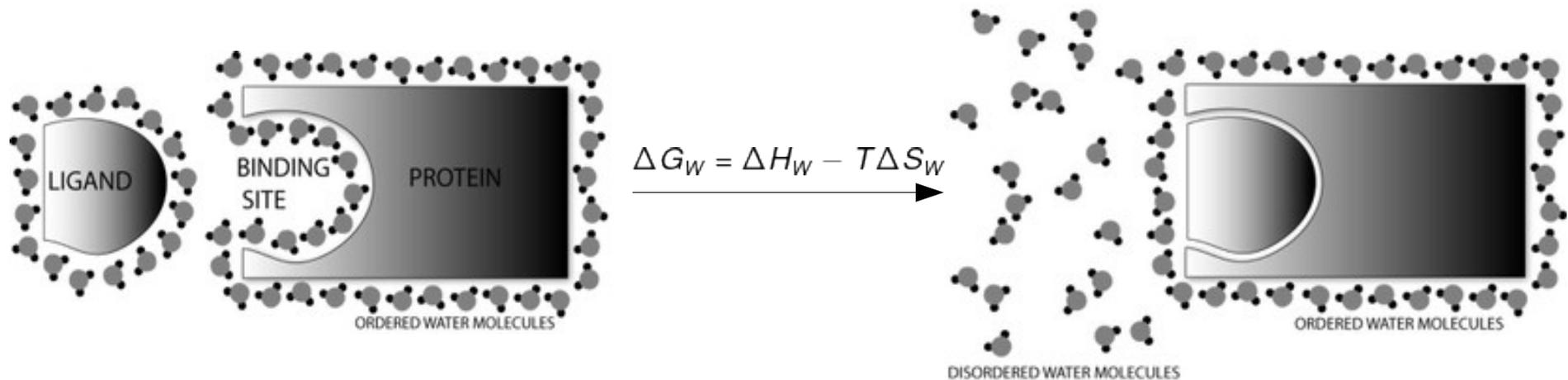
- Le solvant aqueux induit des interactions hydrophobes entre groupes chimiques apolaires.



## II.2. Importance du solvant

---

- Les interactions hydrophobes sont des forces entropiques.



- Les molécules d'eau désordonnées augmentent l'entropie  $\Rightarrow -T\Delta S_W < 0$
- Si solutés polaires  $\Rightarrow \Delta H_W > 0 \Rightarrow \Delta G_W \approx 0$  (pas d'effet hydrophobe)
- Si solutés apolaires  $\Rightarrow \Delta H_W \approx 0 \Rightarrow \Delta G_W < 0$  (attraction hydrophobe)

## Pour résumer

---

- Les interactions non-covalents n'impliquent pas la mise en commun d'électrons.
- Elles sont plus faibles que les liaisons covalentes ( $\sim 100$  kcal/mol).
- On peut distinguer 5 types d'interactions non-covalentes (directes):
  - Interactions de van-der-Waals ( $\sim 1$  kcal/mol)
  - Interactions  $\pi$ - $\pi$  et  $\pi$ -cation ( $\sim 2$  kcal/mol)\*
  - Liaisons hydrogène ( $\sim 3$  kcal/mol)\*
  - Interactions ioniques ( $\sim 4$  kcal/mol)\*
  - Interactions hydrophobes ( $\sim 1$  kcal/mol par  $-\text{CH}_3$  ou  $-\text{CH}_2$ )\*\*

\* Les interactions électrostatiques sont affaiblies dans l'eau.

\*\* Sont générées par le solvant aqueux.

# Plan du cours

---

## I. Aspect structurale de la reconnaissance moléculaire

- I.1. Complémentarités de forme
- I.2. Structures des biomolécules
- I.3. Changements conformationnels

## II. Aspect énergétique de la reconnaissance moléculaire

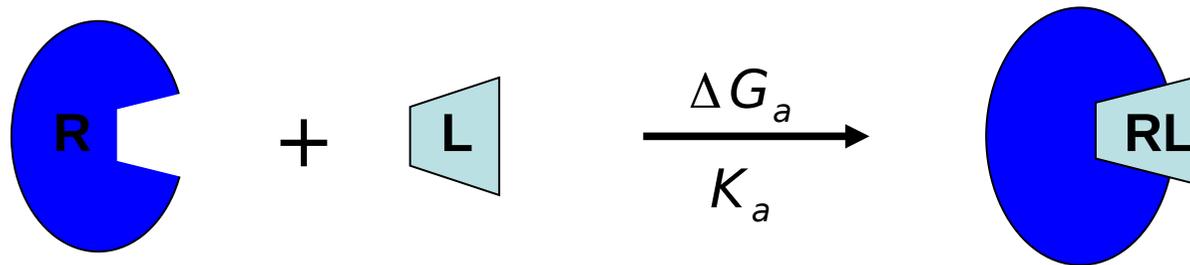
- II.1. Interactions non-covalentes
- II.2. Importance du solvant
- II.3. Énergie libre d'association

## III. Prédictions des assemblages biomoléculaires

- III.1. Docking protéine – petite molécules
- III.2. Docking protéine – macromolécule
- III.3. Docking sur protéines flexibles

## II.3. Énergie libre d'association

- Les interactions non-covalentes sont difficiles à quantifier individuellement.
- Seule l'énergie libre d'association entre deux molécules peut vraiment être mesurée expérimentalement.

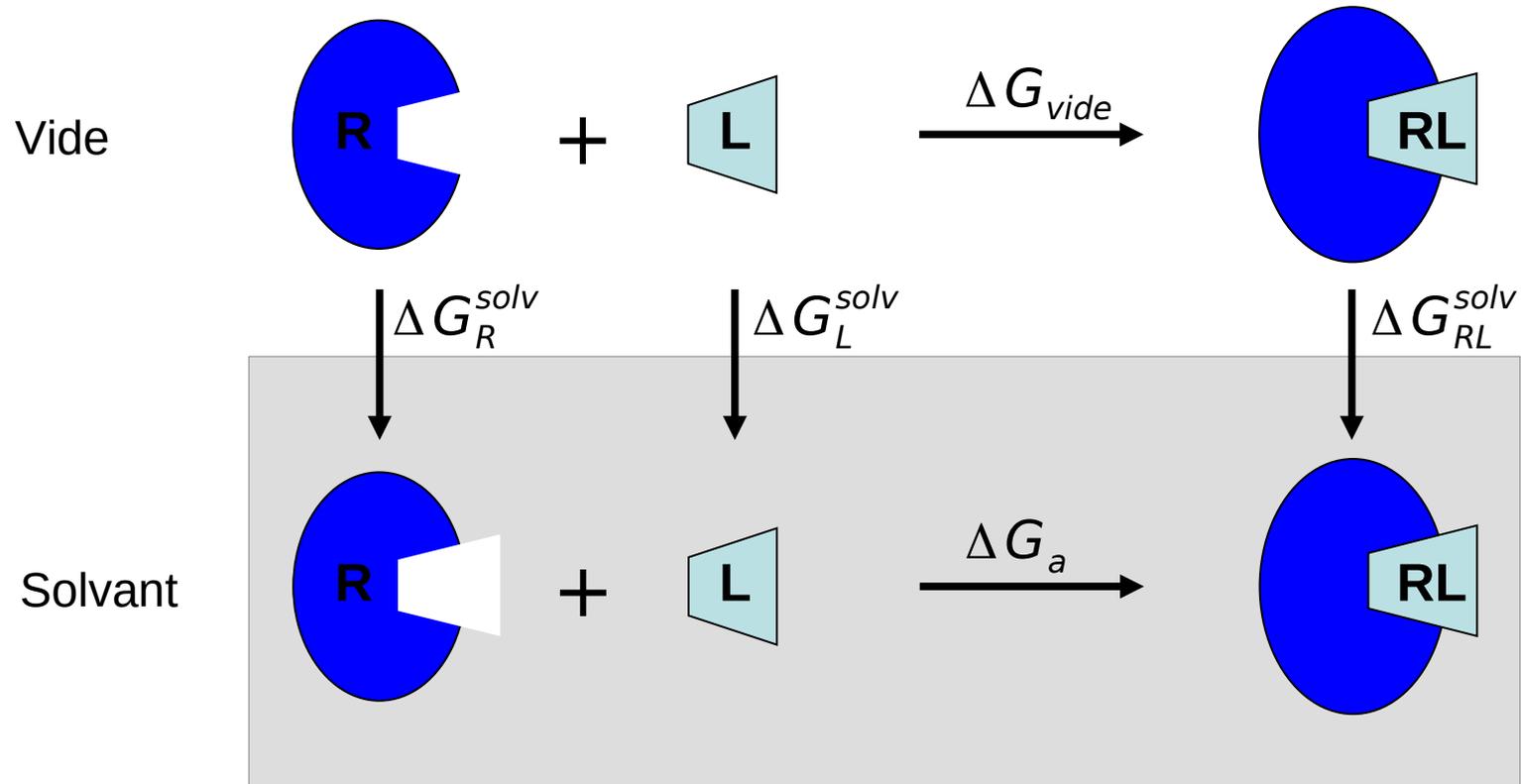


$$K_a = \frac{[\text{recepteur} - \text{ligand}]}{[\text{recepteur}][\text{ligand}]} = 1/K_d$$

$$\Delta G_a = -RT \ln K_a$$

$K_d$ (mol/l)	$10^{-9}$	$10^{-6}$	$10^{-3}$	1	$10^3$	$10^6$	$10^9$
$\Delta G_a$ (kcal/mol)	-12,3	-8,2	-4,1	0	4,1	8,2	12,3
Association	Favorable				Défavorable		

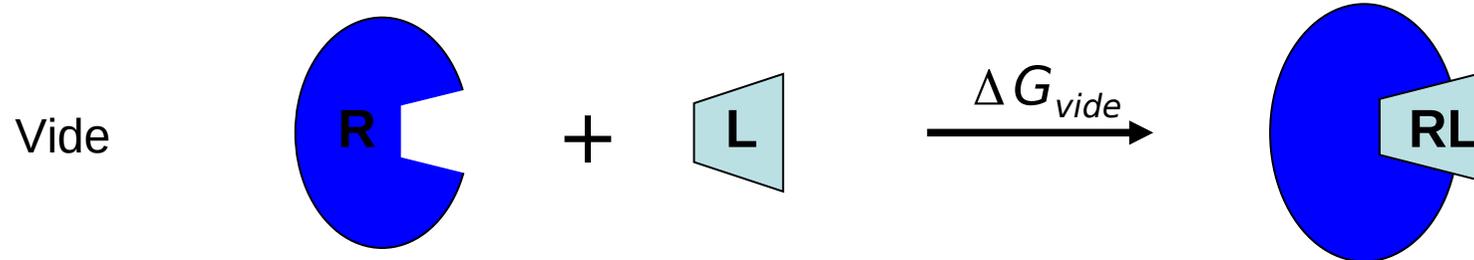
# Relation entre interactions non-covalentes et énergie libre ?



$$\Delta G_a = \Delta G_{vide} + \Delta G^{solv} = \Delta G_{vide} + (\Delta G_{RL}^{solv} - \Delta G_R^{solv} - \Delta G_L^{solv})$$

# Relation entre interactions non-covalentes et énergie libre ?

---



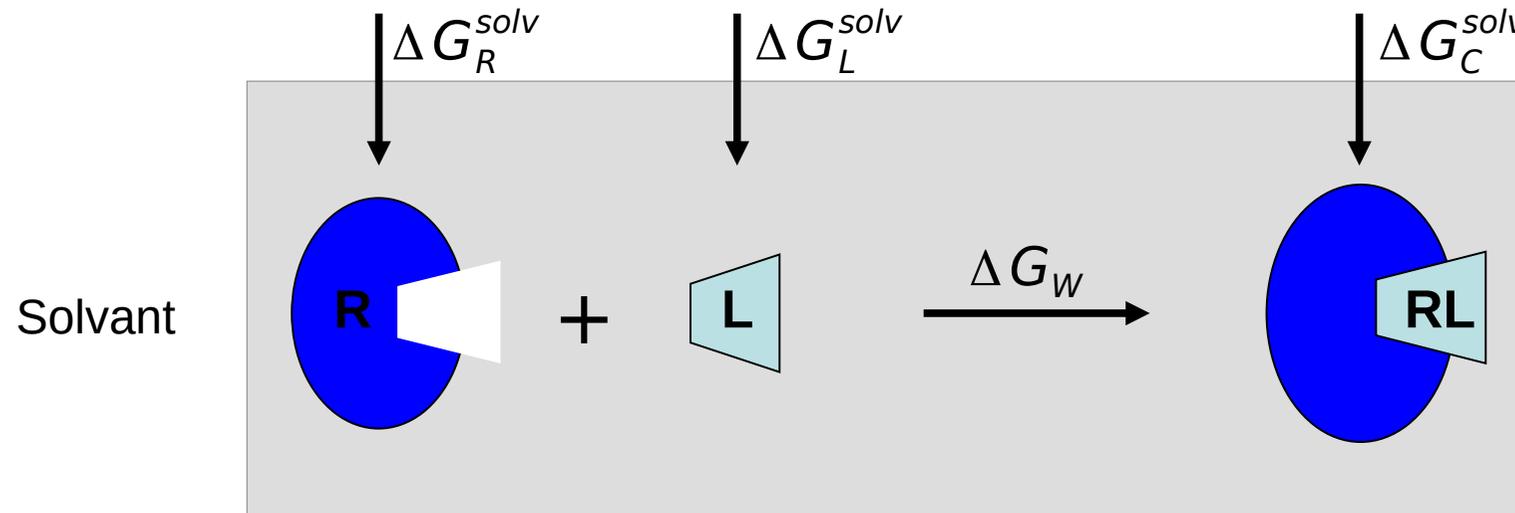
$$\Delta G_{vide} = \Delta H_{vide} - T \Delta S_{vide}$$

$$|\Delta H_{vide}| \gg |\Delta S_{vide}|$$

$$\Delta H_{vide} = \Delta \langle E^{non-covalent} \rangle = \langle E_{RL}^{vdW} \rangle + \langle E_{RL}^{pi/cation} \rangle + \langle E_{RL}^{liaisonH} \rangle + \langle E_{RL}^{ionique} \rangle$$

# Relation entre interactions non-covalentes et énergie libre ?

---

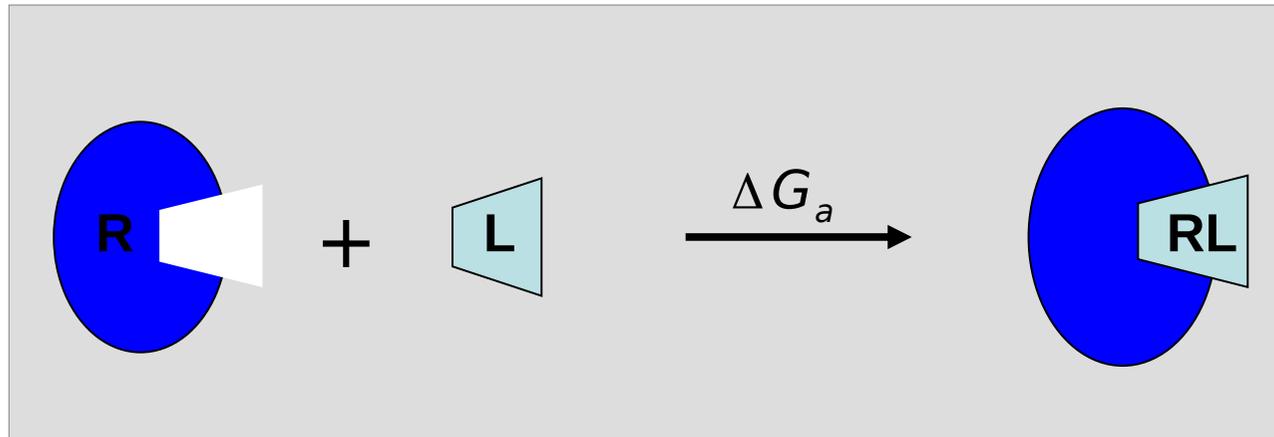


$$\Delta G^{solv} = \Delta G_{RL}^{solv} - \Delta G_R^{solv} - \Delta G_L^{solv} = \Delta G_W$$

$$\Delta G_W = \Delta H_W - T \Delta S_W$$

$$-T \Delta S_W \approx \langle E_{RL}^{hydrophobic} \rangle$$

# Relation entre interactions non-covalentes et énergie libre ?



$$\Delta G_a = \langle E_{RL}^{vdW} \rangle + \langle E_{RL}^{piliation} \rangle + \langle E_{RL}^{liaisonH} \rangle + \langle E_{RL}^{ionique} \rangle + \langle E_{RL}^{hydrophobic} \rangle + \Delta H_w$$

$\Delta H_w$  affaiblit les interactions électrostatiques

$$\Delta G_a = \langle E_{RL}^{vdW} \rangle + \langle E_{RL}^{piliation*} \rangle + \langle E_{RL}^{liaisonH*} \rangle + \langle E_{RL}^{ionique*} \rangle + \langle E_{RL}^{hydrophobic} \rangle$$

# Plan du cours

---

## I. Aspect structurale de la reconnaissance moléculaire

- I.1. Complémentarités de forme
- I.2. Structures des biomolécules
- I.3. Changements conformationnels

## II. Aspect énergétique de la reconnaissance moléculaire

- II.1. Interactions non-covalentes
- II.2. Importance du solvant
- II.3. Énergie libre d'association

## III. Prédictions des assemblages biomoléculaires

- III.1. Docking protéine – petite molécules
- III.2. Docking protéine – macromolécule
- III.3. Docking sur protéines flexibles

### III. Prédiction des assemblages biomoléculaires

---



- Technique principale : amarrage ou *docking* moléculaire
- L'objectif du *docking* est double :
  1. Prédire le **mode de liaison** = **position** + **orientation** + **conformation** d'un ligand avec un récepteur.
  2. Prédire l'**affinité** ou **énergie libre d'association** (relative) du ligand pour le récepteur.

### III. Prédiction des assemblages biomoléculaires

---

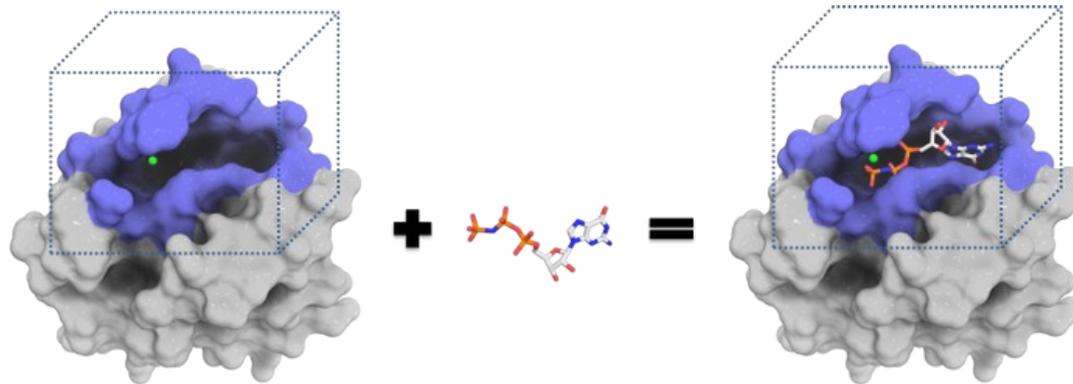


- Technique principale : amarrage ou *docking* moléculaire
- Tout logiciel de *docking* comporte deux composantes principales :
  1. Un **algorithme de recherche** des modes de liaison = position + orientation + conformation du ligand sur le récepteur.
  2. Un **score énergétique** pour déterminer les meilleurs ou plus probables modes de liaison.

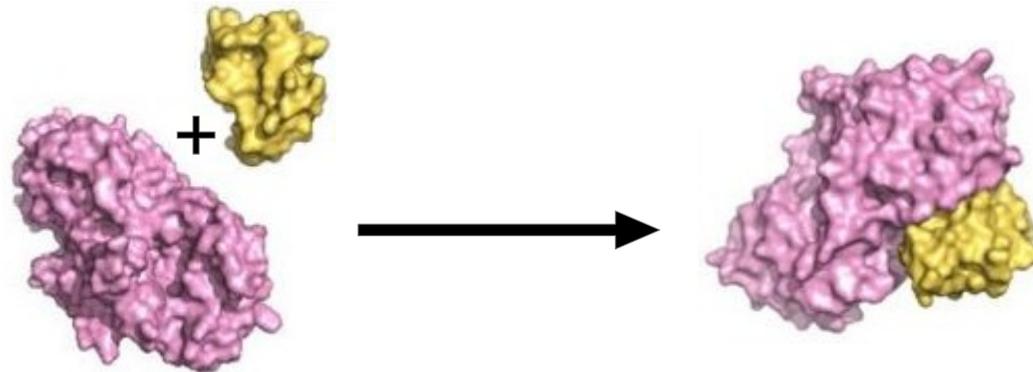
### III. Prédiction des assemblages biomoléculaires

---

- *Docking* protéine – petite molécule



- *Docking* protéine – macromolécule



# Plan du cours

---

## I. Aspect structurale de la reconnaissance moléculaire

- I.1. Complémentarités de forme
- I.2. Structures des biomolécules
- I.3. Changements conformationnels

## II. Aspect énergétique de la reconnaissance moléculaire

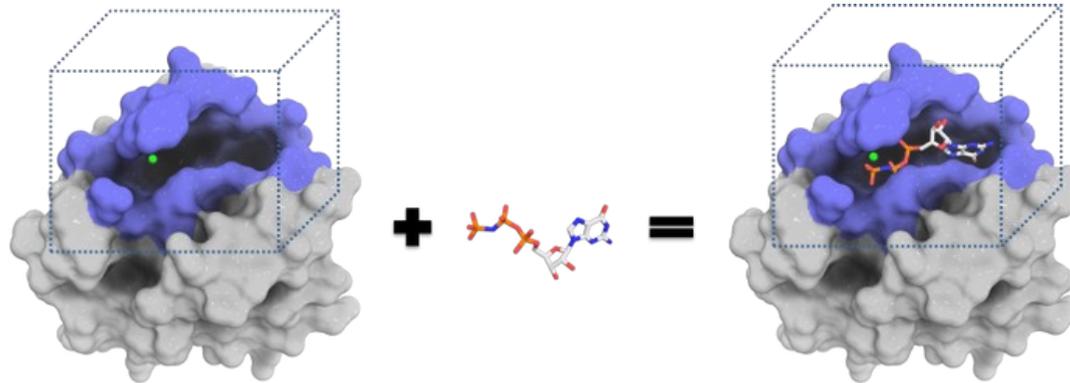
- II.1. Interactions non-covalentes
- II.2. Importance du solvant
- II.3. Énergie libre d'association

## III. Prédictions des assemblages biomoléculaires

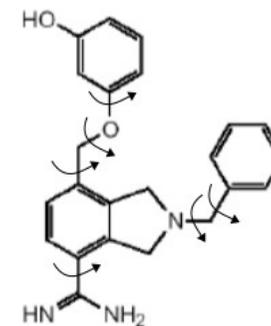
- III.1. Docking protéine – petite molécules
- III.2. Docking protéine – macromolécule
- III.3. Docking sur protéines flexibles

## III.1. Docking protéine – petite molécule

---



- Le site de liaison est généralement connu (site actif d'une enzyme).
- Les principaux degrés de liberté du problème sont l'orientation et la conformation (angles de torsions) du ligand.



- On cherche principalement à optimiser la complémentarité chimique du ligand avec le site de liaison.

## III.1. Docking protéine – petite molécule

---

- Le score énergétique est généralement une fonction approchée de l'énergie libre d'association :

$$\Delta G_a = w_i \delta G_{ionique} + w_l \delta G_{liaisonH} + w_p \delta G_{pi/cation} \\ + w_h \delta G_{hydrophobe} + w_v \delta G_{vdW} + w_d \delta G_{déformation}$$

$$\delta G_{ionique} \approx -4 \text{ kcal/mol par liaison}$$

$$\delta G_{liaisonH} \approx -3 \text{ kcal/mol par liaison}$$

$$\delta G_{pi/cation} \approx -2 \text{ kcal/mol par cycle aromatique/cation}$$

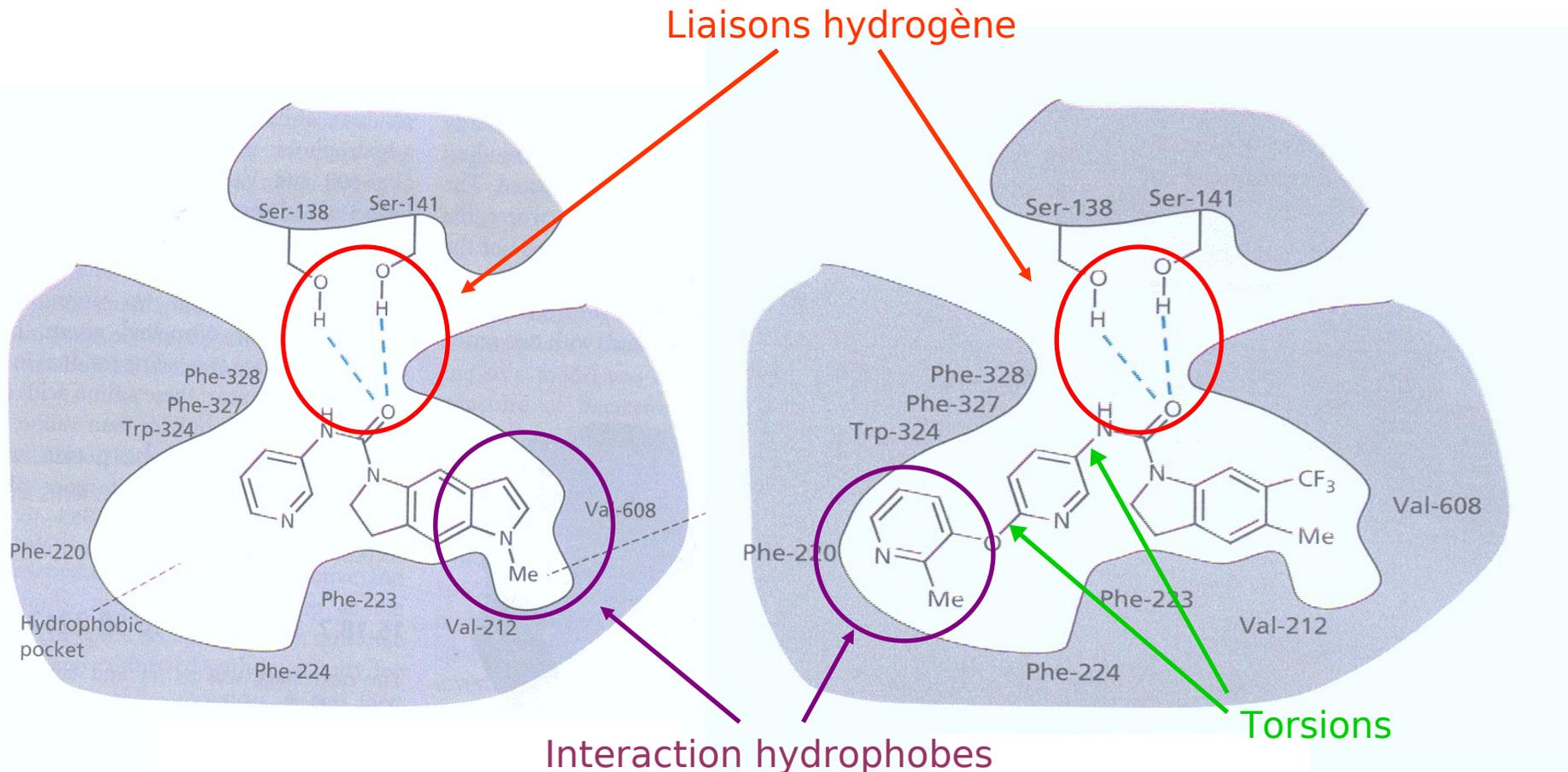
$$\delta G_{hydrophobe} \approx -1 \text{ kcal/mol par groupe méthyle/méthylène}$$

$$\delta G_{vdW} \approx -1 \text{ kcal/mol par atome en contact}$$

$$\delta G_{déformation} \approx +0,5 \text{ kcal/mol/torsion}$$

## III.1. Docking protéine – petite molécule

- Le calcul du score consiste alors à compter le nombre de liaisons hydrogène, le nombre de torsions déformés, le nombre de groupes méthyle/méthylène en contact...

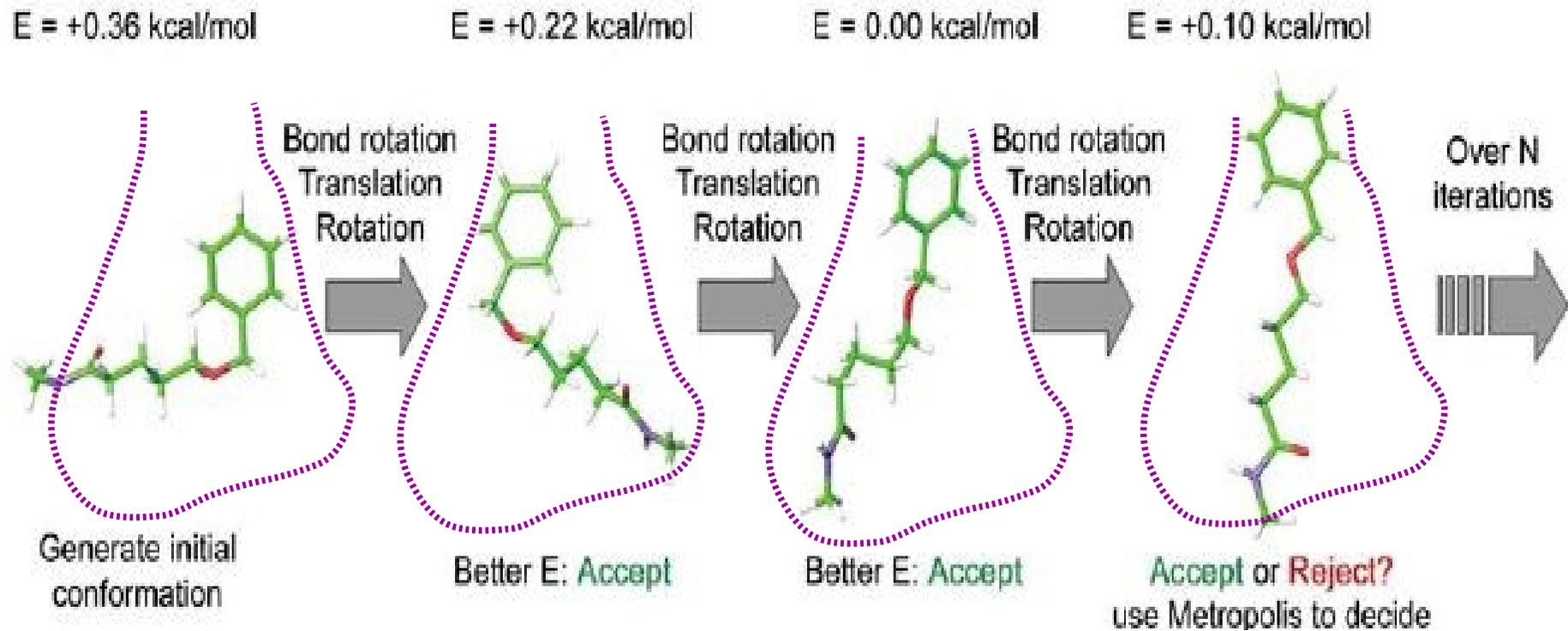


## III.1. Docking protéine – petite molécule

---

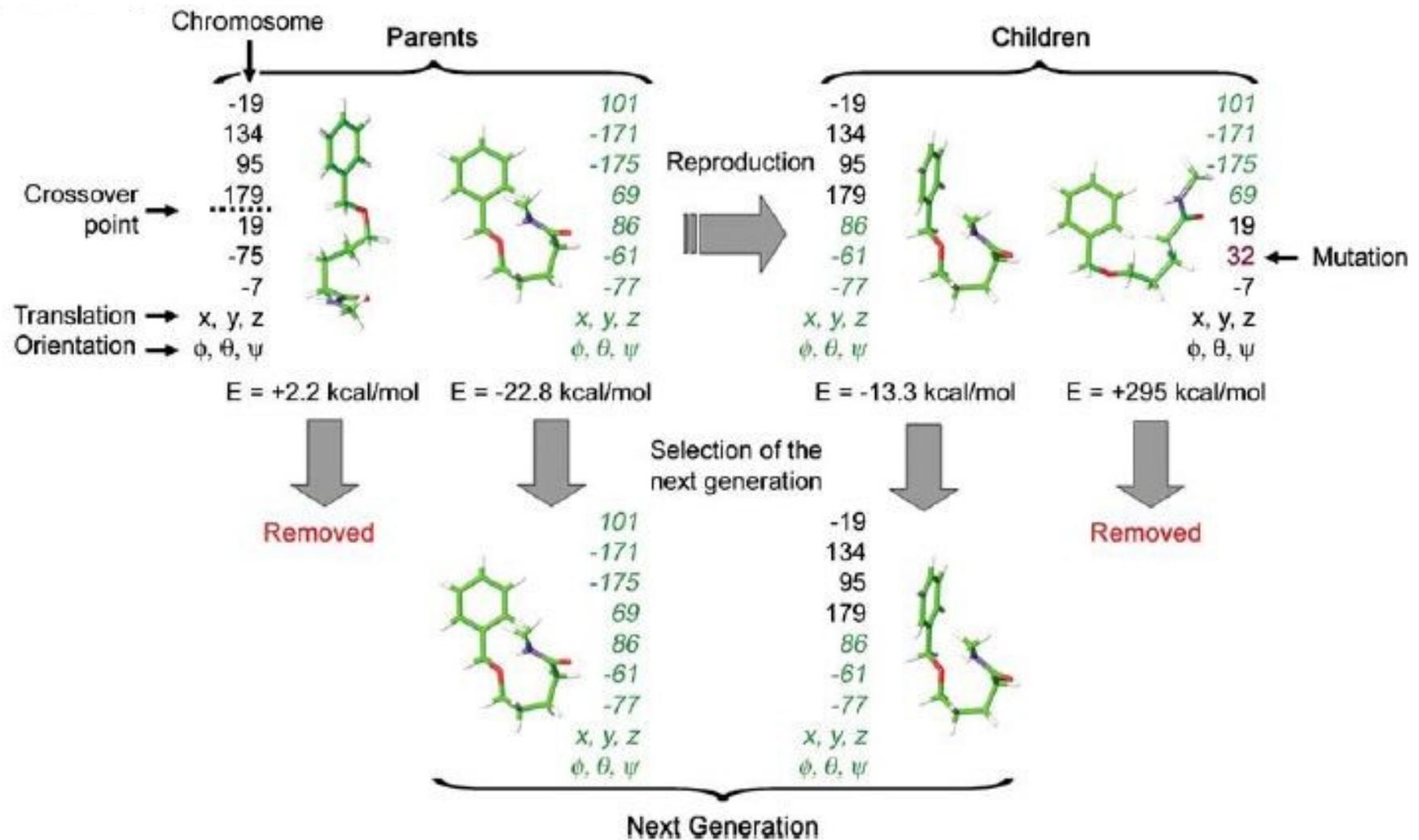
- Approximation : Le **ligand est flexible** et la protéine reste rigide. Mais plus le ligand est flexible, plus la recherche des modes de liaison est longue.
- Hypothèse plus réaliste : **Le ligand et la protéine sont flexibles**. Mais exploration des modes de liaison très compliquée et très longue.
- Algorithme de recherche des modes de liaison :
  - Monte Carlo
  - Algorithme génétique

# Monte Carlo



- La méthode permet de générer des conformations très différentes (à condition de tomber sur une structure de basse énergie).

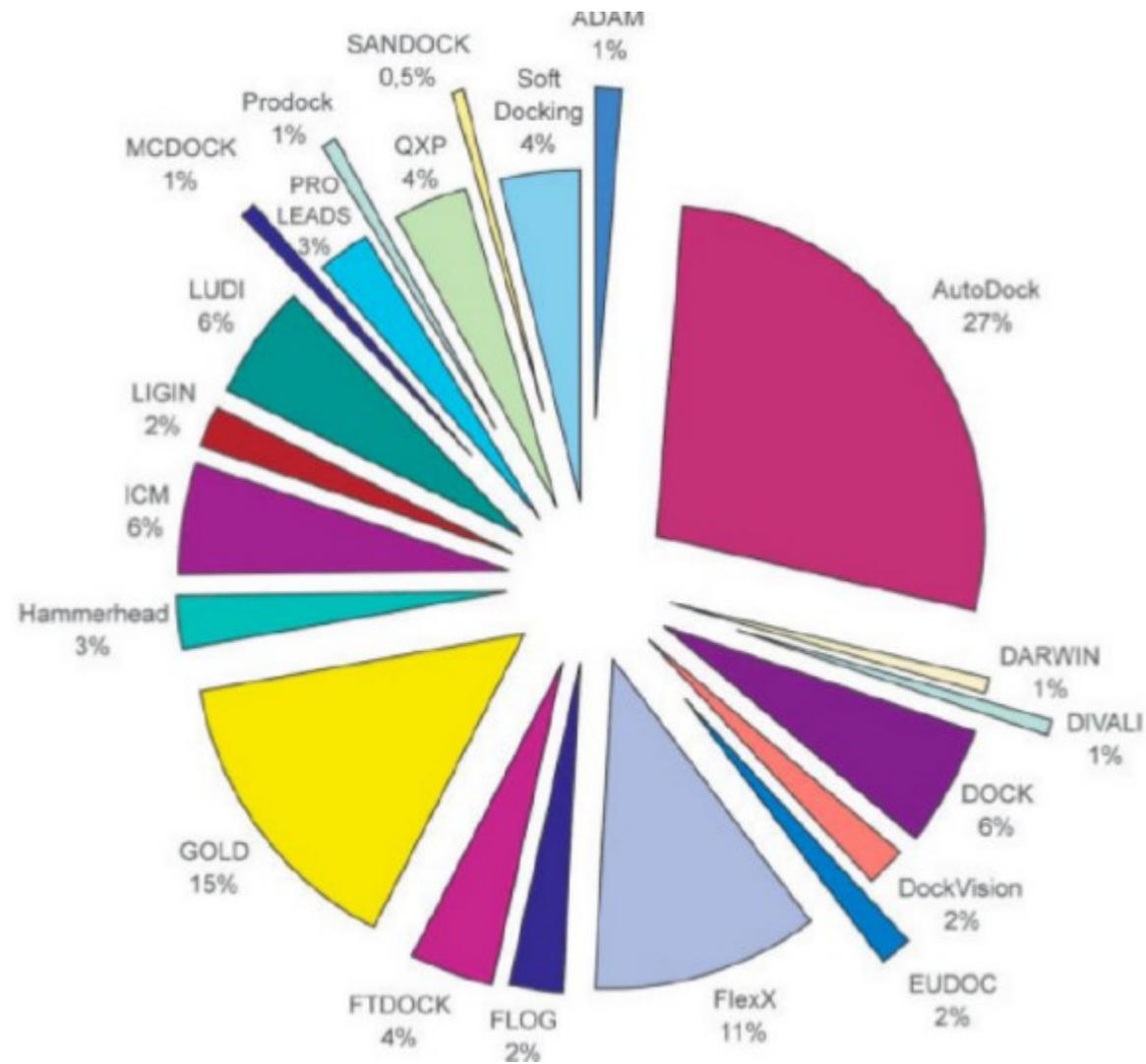
# Algorithme génétique



- Le résultat de l'algorithme est une population de conformères probables.

## III.1. Docking protéine – petite molécule

---



# Plan du cours

---

## I. Aspect structurale de la reconnaissance moléculaire

- I.1. Complémentarités de forme
- I.2. Structures des biomolécules
- I.3. Changements conformationnels

## II. Aspect énergétique de la reconnaissance moléculaire

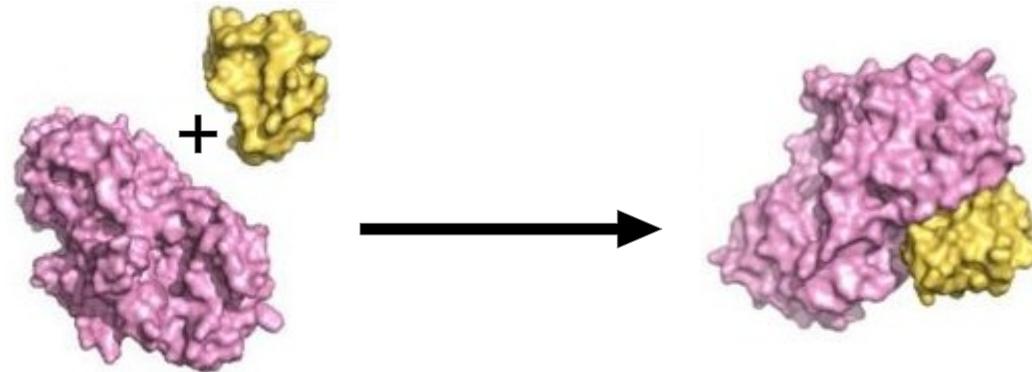
- II.1. Interactions non-covalentes
- II.2. Importance du solvant
- II.3. Énergie libre d'association

## III. Prédictions des assemblages biomoléculaires

- III.1. Docking protéine – petite molécules
- III.2. Docking protéine – macromolécule
- III.3. Docking sur protéines flexibles

## III.2. Docking protéine – macromolécule

---

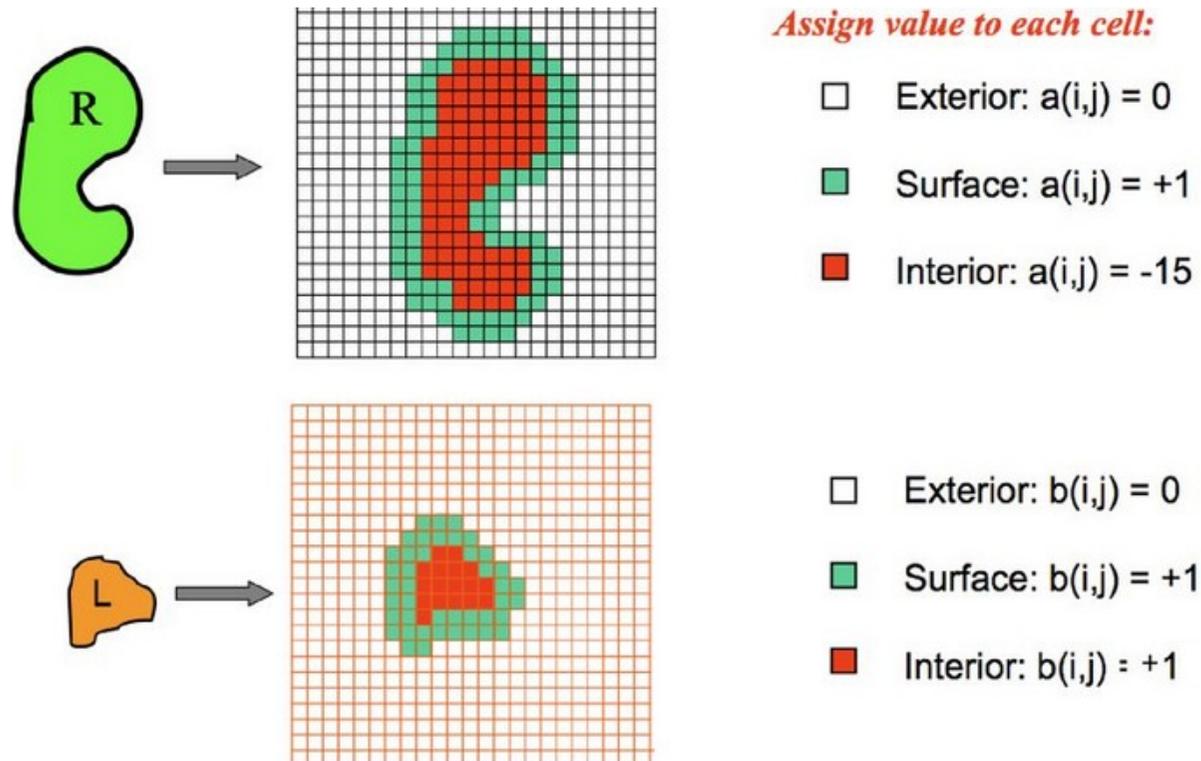


- Le site de liaison et l'interface d'interaction ne sont pas connus.
- Les principaux degrés de liberté du problème sont la translation et l'orientation globales du ligand par rapport au récepteur.
- On cherche principalement à optimiser la **complémentarité de forme** entre le ligand et le récepteur.

## III.2. Docking protéine – macromolécule

---

- Comment quantifier la complémentarité de forme (score) ?
- Approche sur grille : On discrétise la forme des biomolécules sur des grilles 3D :

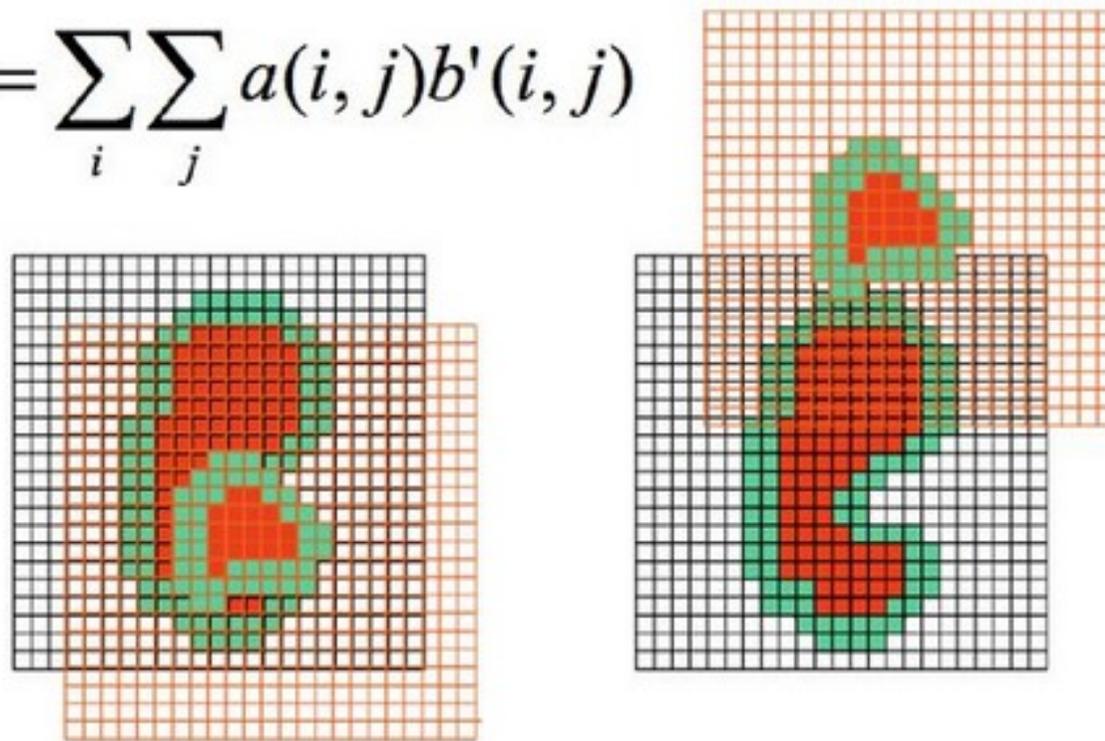


## III.2. Docking protéine – macromolécule

---

- Pour chaque superposition de grilles, on calcul le score :

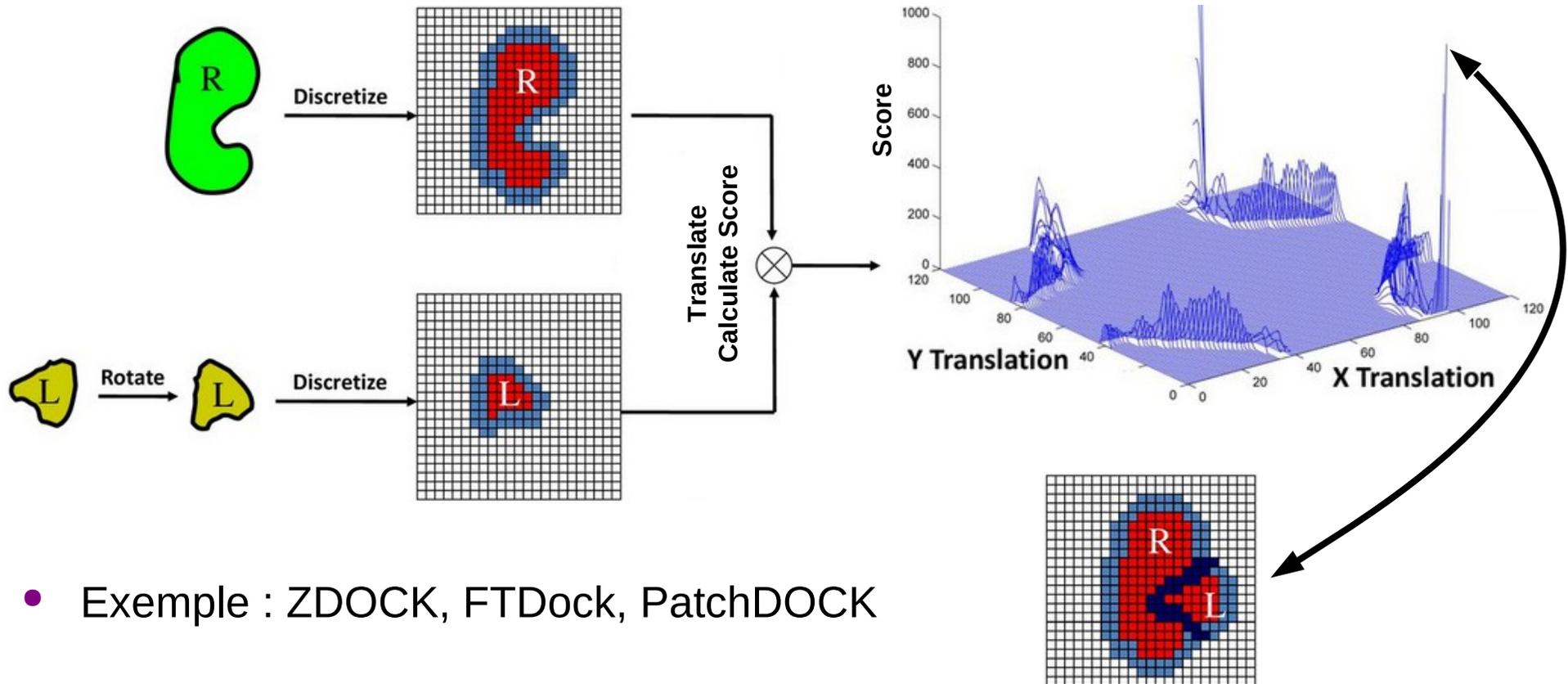
$$Score = \sum_i \sum_j a(i, j) b'(i, j)$$



- Plus le score est grand, plus les surfaces des biomolécules sont complémentaire

## III.2. Docking protéine – macromolécule

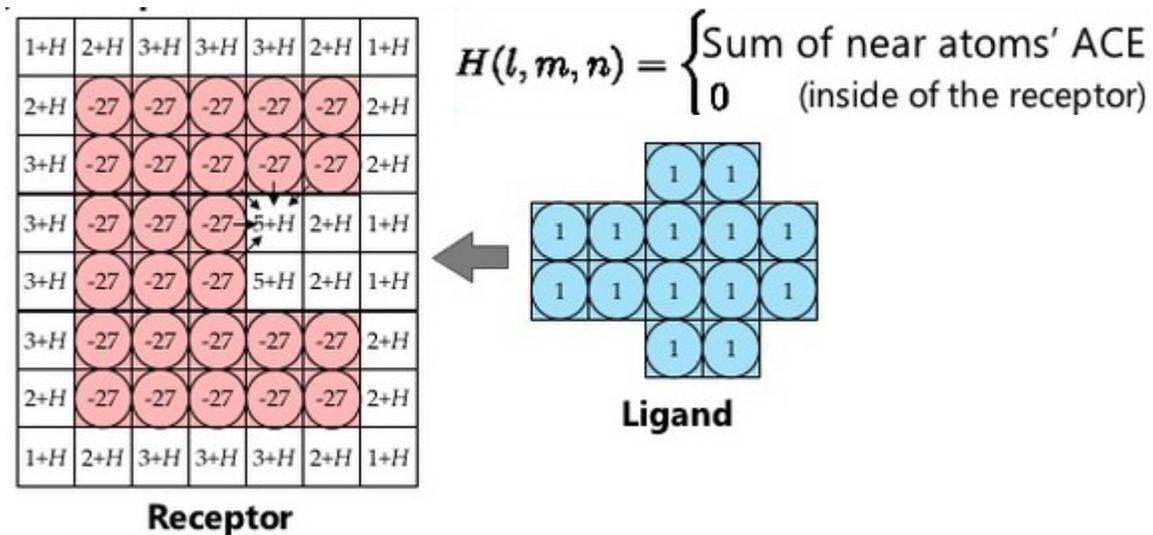
- L'algorithme de recherche consiste à appliquer des translations (en 3D) de la grille du ligand pour maximiser le score.



- Exemple : ZDOCK, FTDock, PatchDOCK

## III.2. Docking protéine – macromolécule

- Remarque : Pour rendre compte d'une certaine complémentarité chimique, on peut nuancer la couleur de chaque point de la surface du récepteur en leur attribuant une valeur qui dépend de la nature chimique des acides aminés de surface (chargée, polaire ou apolaire).

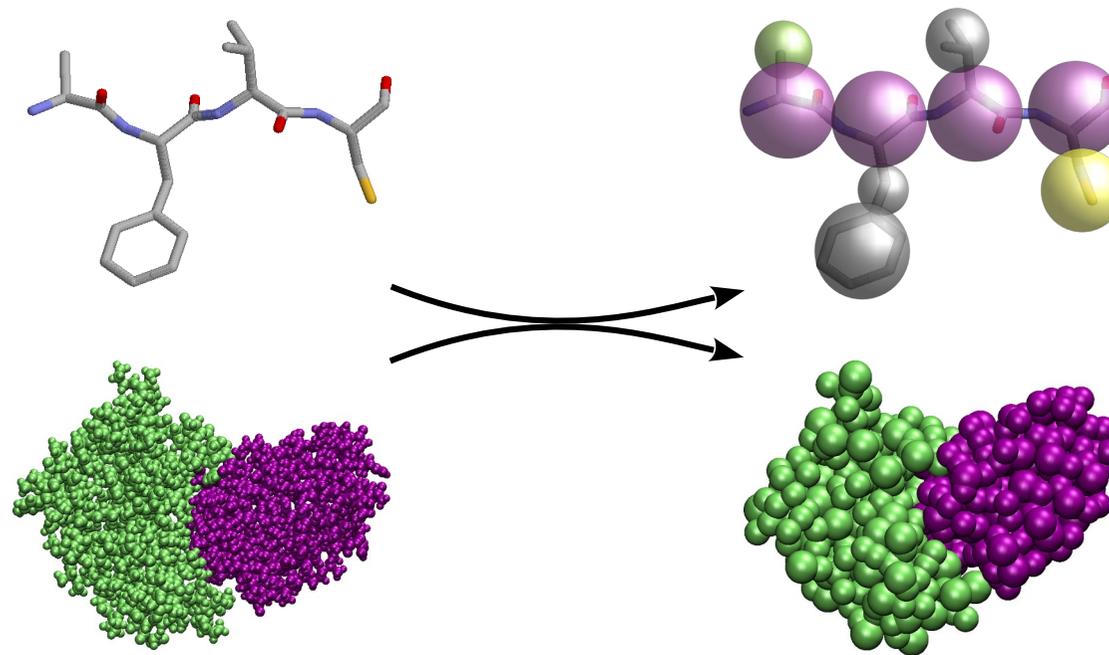


- ACE = Atomic Contact Energy is a desolvation free energy measure, defined as the free energy of replacing a protein atom / water contact, by a protein atom / protein atom contact.

## III.2. Docking protéine – macromolécule

---

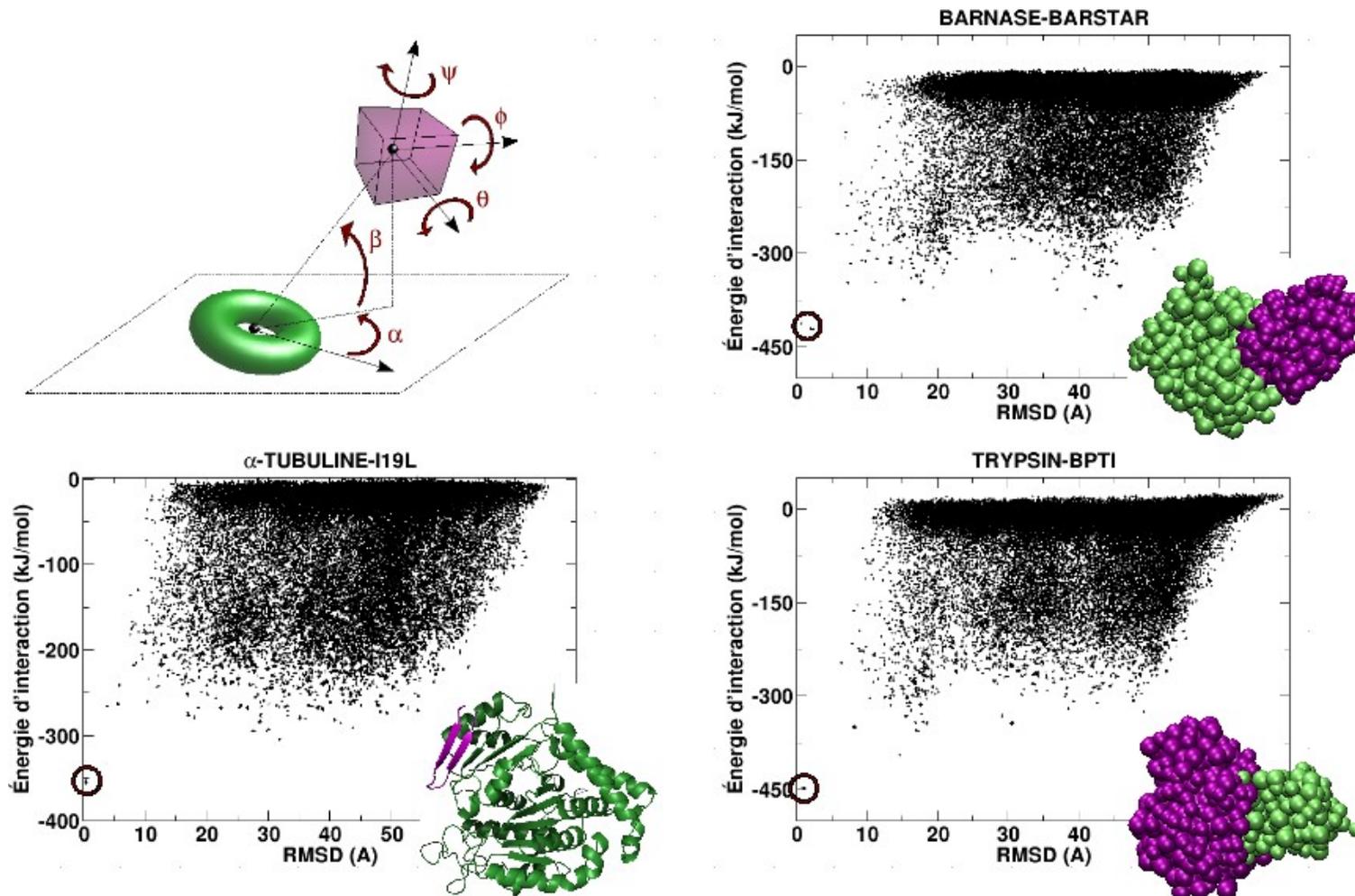
- Approche alternative : Les modèles gros grains.



- Les gros grains permettent de rendre compte de la forme mais aussi des propriétés physico-chimiques des biomolécules.

## III.2. Docking protéine – macromolécule

- La réduction du nombre de particules pour modéliser les biomolécules permet de faire des minimisations plus efficacement.



# Plan du cours

---

## I. Aspect structurale de la reconnaissance moléculaire

- I.1. Complémentarités de forme
- I.2. Structures des biomolécules
- I.3. Changements conformationnels

## II. Aspect énergétique de la reconnaissance moléculaire

- II.1. Interactions non-covalentes
- II.2. Importance du solvant
- II.3. Énergie libre d'association

## III. Prédictions des assemblages biomoléculaires

- III.1. Docking protéine – petite molécules
- III.2. Docking protéine – macromolécule
- III.3. Docking sur protéines flexibles

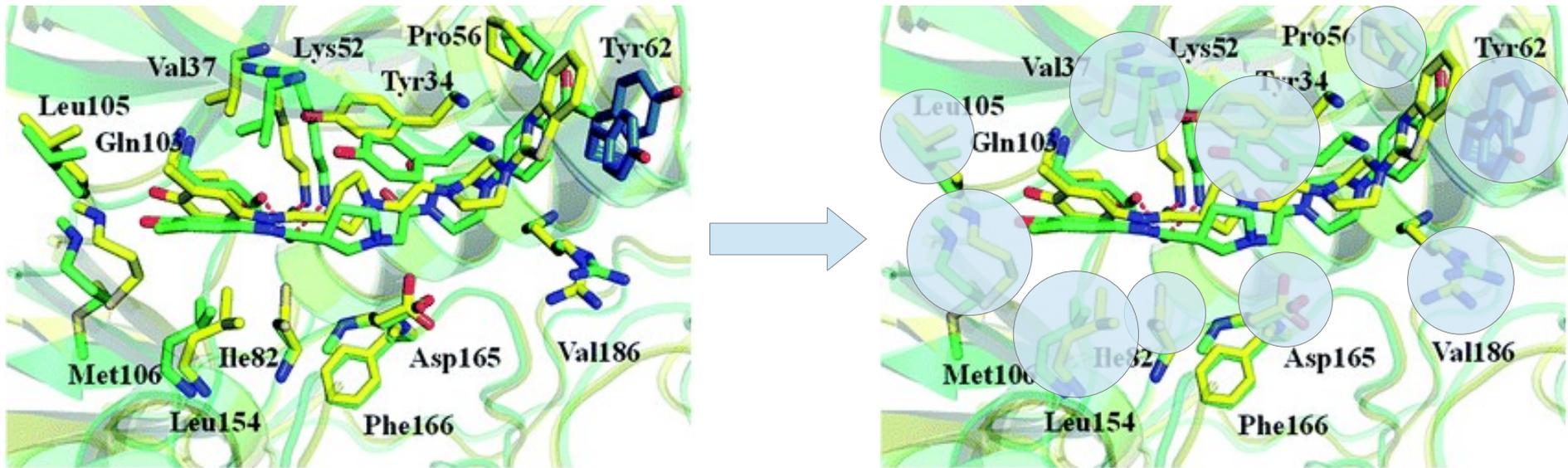
### III.3. Docking sur protéines flexibles

---

- Comment faire pour prendre en compte la flexibilité des biomolécules pendant un docking ?
  - Flexibilité des chaînes latérales des protéines
  - Flexibilité du squelette des protéines
  - Mouvement de charnière

# Flexibilité des chaînes latérales

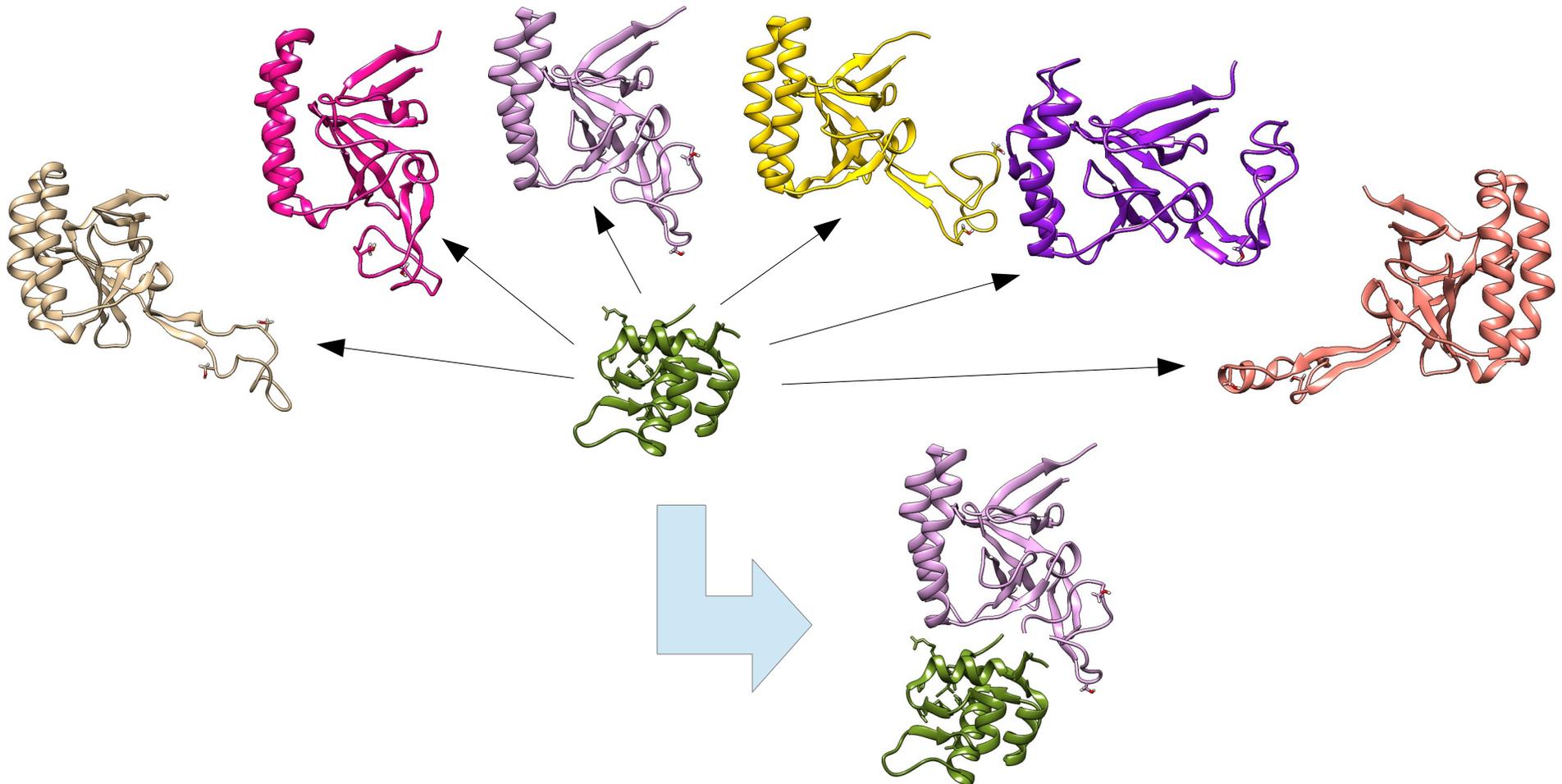
- Soft docking :



# Flexibilité du squelette

---

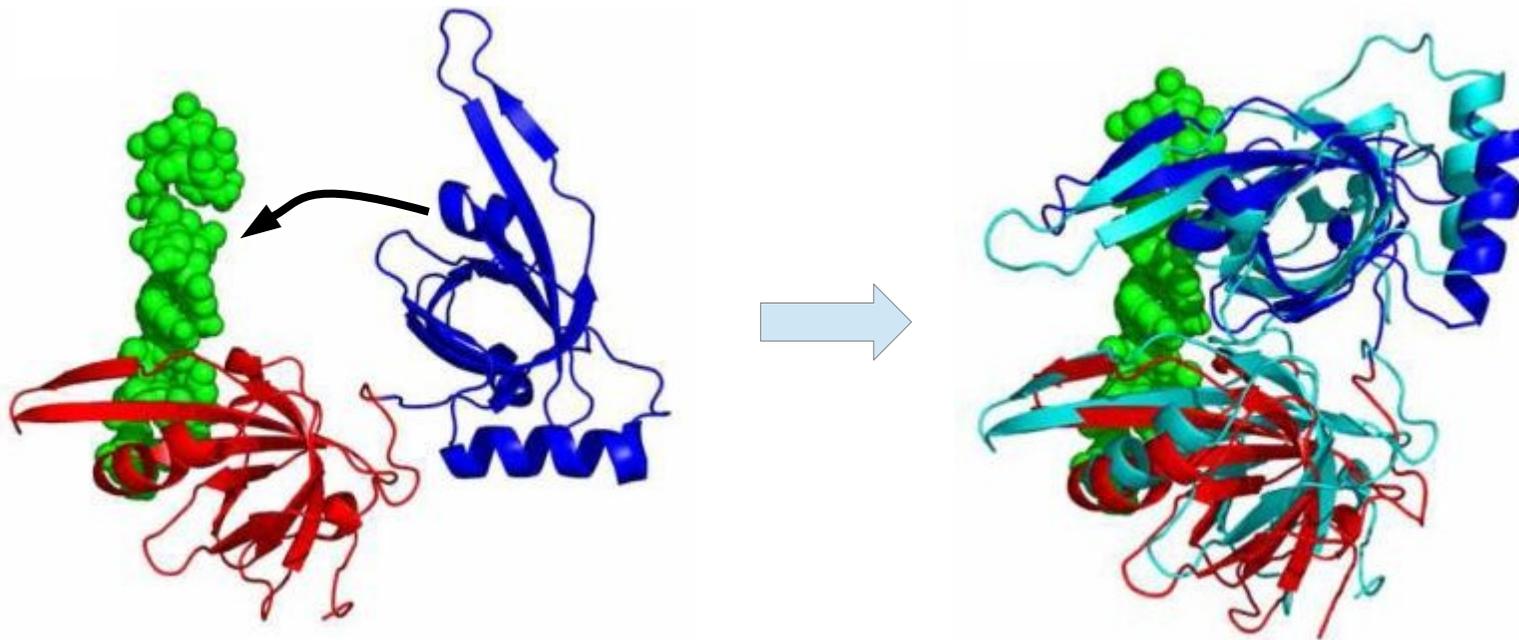
- Ensemble docking :



# Mouvement de charnière

---

- Docking par domaine ou fragment :



# Conclusions

---

- Les processus cellulaires (métabolique et signalisation) sont effectués par **assemblage de biomolécules** (principalement des protéines).
- Ces assemblages se forment par **reconnaissance moléculaire** = processus d'association **spécifique** (dépend de la nature des biomolécules) **et sélective** (produit une forme prépondérante du complexe).
- Les deux facteurs qui déterminent la reconnaissance moléculaire sont la **complémentarité structurale** et la **complémentarité chimique**.
- Importance de la **caractérisation structurale** (éventuellement dynamique) des biomolécules.
- L'affinité se traduit par des **interactions non-covalentes** entre les partenaires.
- L'eau est un solvant polaire qui affaiblit les **interactions électrostatiques** et crée des **interactions hydrophobes**.

# Références

---

